Purification, Characterization, and Gene Analysis of a Chitosanase (ChoA) from *Matsuebacter chitosanotabidus* 3001

JAE KWEON PARK, KUMIKO SHIMONO, NOBUHISA OCHIAI, KAZUTAKA SHIGERU, MASAKO KURITA, YUKARI OHTA, KATSUNORI TANAKA, HIDEYUKI MATSUDA, AND MAKOTO KAWAMUKAI*

Department of Life Science and Biotechnology, Faculty of Life and Environmental Science, Shimane University, 1060 Nishikawatsu, Matsue, Shimane 690-8504, Japan

Received 7 June 1999/Accepted 18 August 1999

The extracellular chitosanase (34,000 M_r) produced by a novel gram-negative bacterium *Matsuebacter* chitosanotabidus 3001 was purified. The optimal pH of this chitosanase was 4.0, and the optimal temperature was between 30 and 40°C. The purified chitosanase was most active on 90% deacetylated colloidal chitosan and glycol chitosan, both of which were hydrolyzed in an endosplitting manner, but this did not hydrolyze chitin, cellulose, or their derivatives. Among potential inhibitors, the purified chitosanase was only inhibited by Ag⁺. Internal amino acid sequences of the purified chitosanase were obtained. A PCR fragment corresponding to one of these amino acid sequences was then used to screen a genomic library for the entire *choA* gene encoding chitosanase. Sequencing of the *choA* gene revealed an open reading frame encoding a 391-amino-acid protein. The N-terminal amino acid sequence had an excretion signal, but the sequence did not show any significant homology to other proteins, including known chitosanases. The 80-amino-acid excretion signal of ChoA fused to green fluorescent protein was functional in *Escherichia coli*. Taken together, these results suggest that we have identified a novel, previously unreported chitosanase.

Chitosanase (EC 3.2.1.132) catalyzes the hydrolysis of glycosidic bonds of chitosan, which is a totally or partially deacetylated derivative of chitin. Chitin is a polymer of N-acetylglucosamine. It was proposed that chitosanases are enzymes that hydrolyze GlcN-GlcN, GlcN-GlcNAc, and GlcNAc-GlcN bonds but not GlcNAc-GlcNAc bonds (5, 13). However, the difference between chitosanase and other enzymes, such as chitinases, lysozymes, and N-acetyl-B-D-glucosaminidases is sometimes obscure, especially in the case of enzymes that display activity toward multiple substrates. It is thought that the ability to hydrolyze a 100% deacetylated chitin is an important criterion for classifying an enzyme as a chitosanase. Chitosanases are produced by many organisms, including actinomycetes (2, 17), fungi (4, 27), plants (6, 18), and bacteria (7, 16, 22, 29, 30). Bacterial chitosanases have received special attention because they are important for the maintenance of the ecological balance and have been used to determine the mechanism of chitosan hydrolysis at both biochemical and molecular levels. Compared to the numerous reports on the primary structure and function of chitinases, information on chitosanases is still relatively limited.

So far, chitosanase genes have been isolated from *Bacillus circulans* MH-K1 (1), *Nocardioides* sp. strain N106 (12), *Streptomyces* sp. strain N174 (13), and *Fusarium solani* (26). Among these, the former three chitosanases show structural similarity, but the one from *F. solani* does not. The chitosanase from *Streptomyces* sp. is the only example of a chitosanase whose three-dimensional structure has been determined (11, 23). We still did not know how many different types of chitosanase exist in nature. Some chitosanases have hydrolytic activity on substrates other than chitosan, such as chitin (28) and cellulose (21). A more detailed knowledge of the structure of a number

of chitosanases will be required for understanding their enzymatic differences and common structural elements.

We isolated *Matsuebacter chitosanotabidus* 3001 as a bacterium that produces chitosanase and classified it as a new genus and species belonging to the β -subclass of *Proteobacteria* (20). In this study, we report the properties of purified chitosanase from *M. chitosanotabidus* 3001 and describe the complete nucleotide sequence of the chitosanase (*choA*) gene.

MATERIALS AND METHODS

Materials. Various degrees of deacetylated chitosan (70 to 100%) were purchased from Funakoshi Co., Ltd. (Tokyo, Japan). Chitin was obtained from San-in Kensetsu Co., Ltd. Oligomers of glucosamine were purchased from Seikagaku Kogyo Co., Ltd. (Tokyo, Japan). Glycol chitosan was purchased from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). Carboxymethylchitosan and carboxymethylchitin were obtained from Nissui Co., Ltd. (Tokyo, Japan). λ ZapII and the packaging kit were purchased from Stratagene. Ampicillin, 5-bromo-4-chloro-3-indolyl- β -p-galactoside (X-Gal) and isopropyl- β -p-thiogalactopyranoside (IPTG) were purchased from Takara Biomedicals. Kyoto, Japan). Restriction endonuclease and bacterial alkaline phosphatase were obtained from Toyobo Co., Ltd. (Osaka, Japan), or Takara Biomedicals. The cassettes and cassette primers, 5'-d(CGTACATATGCGTAATACGACTCACTATAGGGAG A)-3' (C1) and 5'-d(CGTTAGAACGCGTAATACGACTCACTATAGGGAG A)-3' (C2), were purchased from Takara Biomedicals. All other chemicals were reagent grade or molecular biological grade.

Screening and culture conditions of *M. chitosanotabidus*. *M. chitosanotabidus* 3001 was identified as a novel bacterium on the basis of its 16S rRNA sequence, morphology, and physiological properties as described previously (20).

For growth in liquid culture, *M. chitosanotabidus* 3001 was grown in modified base medium containing 0.4% (wt/vol) colloidal chitosan. For chitosanase production, a single colony of *M. chitosanotabidus* 3001 was inoculated in 50 ml of colloidal chitosan liquid medium and grown for 4 days at 30°C with shaking (200 rpm). This medium consisted of 0.4% (wt/vol) colloidal chitosan, 0.5% MgSO₄, 0.3% KH₂PO₄, 0.7% K₂HPO₄, 0.25% yeast extracts, and 0.25% polypeptone at pH 7.0.

^{*} Corresponding author. Mailing address: Department of Life Science and Biotechnology, Faculty of Life and Environmental Science, Shimane University, 1060 Nishikawatsu, Matsue, Shimane 690-8504, Japan. Phone and fax: 81-852-32-6587. E-mail: kawamuka@life .shimane-u.ac.jp.

Concentration and purification of chitosanase. Bacterial cells were removed from culture broth by centrifugation at 12,000 rpm for 15 min in a Kubota KR-20000T rotor, and proteins in supernatant fluids were concentrated with ammonium sulfate (70% saturation). After incubation at 4°C for 1 h, the precipitates were collected by centrifugation at 12,000 rpm for 20 min. The precipitates were dissolved in an appropriate volume of 50 mM Tris-HCl buffer at pH 8.0 and dialyzed against 20 mM Tris-HCl buffer (pH 8.0) at 4°C overlight. The dialysates were centrifuged to remove the insoluble materials and were used as crude chitosanolytic enzyme fraction. Crude chitosanolytic enzyme was purified by isoelectric chromatography on a 110-ml column (LKB-Produkter) with am-



FIG. 1. Isoelectric-focusing chromatography of chitosanase produced by *M. chitosanotabidus* 3001. Open circles (\bigcirc) indicate the absorbance at 280 nm, and solid circles (\bullet) indicate the activity of the enzyme. Fractions 31 and 32 were pooled and used as a source of purified enzyme.

pholine (Sigma) as the carrier ampholite. Each fraction was collected, and the chitosanase activity was measured. Active fractions (numbers 31 and 32) were collected and used as a source of purified enzyme (Fig. 1). Proteins eluting from the column were detected by measuring the absorbance at 280 nm.

Determination of amino acid sequence and immunoblotting. To determine internal amino acid sequences, purified chitosanase from *M. chitosanotabidus* 3001 was digested with an appropriate concentration of trypsin, and the resulting peptides were separated by sodium dodecyl sulfate (SDS)–L2.5% polyacrylamide gel electrophoresis (PAGE). Chitosanase purified as mature extracellular protein from *E. coli* expressing the chitosanase gene (*choA*) was also used for determining of the N-terminal amino acid sequence. The enzyme partially digested with trypsin was blotted to polyvinylidene diffuoride (PVDF) membrane (Immobilon-PSQ; pore size, 0.45- μ m; IPVH 304FO; Millipore) by using the electroblotting system according to the manufacturer's instructions (Nippon Millipore Co.). Amino acid sequencing was performed by using a Shimadzu protein sequencer (PSQ-10). Rabbit antiserum raised against the chitosanase (anti-Cho) was prepared by Takara Biomedicals (Kyoto, Japan).

In Western blot analysis, cell extracts were subjected to SDS-PAGE (12.5%), and the protein bands were electrophoretically transferred onto a nitrocellulose membrane. For immunolabeling of ChoA, nitrocellulose membranes were incubated at room temperature with shaking in TBS-M buffer (20 mM Tris-HCI [pH 7.6], 0.137 M NaCl, 0.25% Tween 20, 5% dried milk) for at least 1 h. Afterwards, the membranes were rinsed several times with Tris-buffered saline (TBS-T,) buffer and then incubated for 1 h with rabbit antiserum. After several rinses in TBS buffer, the membranes were incubated with the horseradish peroxidase-conjugated second antibody, and the membrane-bound immunocomplexes were detected on X-ray film by the enhanced chemiluminescence method, which is based on measuring light emission from the oxidation of acridinium esters (ECL-Plus System; Amersham).

Construction and expression of ChoA-GFP fusions. The DNA fragment corresponding to the N-terminal region of ChoA was amplified by PCR by using the primer TAGGATCCTATGCAACTTCCTGGACCTGAT (to create a *Bam*HI site in front of the ATG codon) as the sense primer and TTGAATTCTTGGG GCCGGCCATGCC (to create an *Eco*RI site at the codon for the 48-amino-acid segment of ChoA) or TTGAATTCGGAATCACGCCCGCCGC (to create an *Eco*RI at the codon for the 88-amino-acid segment of ChoA) as the antisense primer. PCR was used to amplify the GFP (green fluorescent protein) gene from pGP110 by using the primer ACGAATTCTAGTAAAGGAGAAGAA (to create an *Eco*RI site before the ATG codon) and the primer ATCCCGGGGAAGC TTATTTGTATAGTTCATC (to create a *Hin*dIII site at the 3' end) (15). The two DNA fragments coding for ChoA and GFP were ligated into pBluescript SK(+) to yield pB48ChG (the N-terminal 48-amino-acid region was fused to GFP) and pB88ChG (the N-terminal 88-amino-acid region was fused to GFP).

E. coli JM109 harboring pB48ChG or pB88ChG was grown on Luria-Bertani medium to an optical density at 600 nm of 0.2 to 0.5. The cells were then further incubated for 2.5 h with 1 mM IPTG. Cells and supernatant were separated by centrifugation. Protein in the supernatant was concentrated by ultrafiltration through a Centricon 10 (Amicon) apparatus. *E. coli* cells were disrupted by sonication, and undisrupted cells were removed by centrifugation. Proteins (5 μ g) were loaded onto each lane and electrophoresed. Western blot analysis was done by using an anti-GFP antibody (Clontech). Western blot analysis was carried out as described above.

Chitosanase assay. Chitosanase activity was assayed by using colloidal chitosan as a substrate. The reaction mixture consisted of 0.5 ml of 0.5% colloidal chitosan, 1 ml of McIlvaine buffer (0.1 M citrate plus 0.2 M Na₂HPO₄) at pH 7.0, and 0.5 ml of the enzyme solution, and the mixtures were incubated at 30°C for 10 min. Reactions were stopped by boiling for 3 min, the reaction mixtures were centrifuged, and the supernatants were retained. The amount of reducing sugars produced was determined at A_{420} by the modified Schales method (8). One unit of chitosanase activity was taken as the amount of enzyme that produced 1 μ mol of reducing sugars (expressed as glucosamine equivalents) per min.

DNA manipulations and amplification (PCR) reactions. Oligonucleotide probes were synthesized by Takara Biomedicals. Standard procedures for restriction endonuclease digestion, agarose gel electrophoresis, purification of DNA from agarose gels, DNA ligations, and other cloning-related techniques were done as described by Sambrook et al. (24).

PCR amplification of the chitosanase gene was performed by using a DNA Thermal Cycler (Perkin-Elmer/Cetus). The two primers, designated S1 and S2 (5'-TCCGACAAGAAC(T)AAG(A)CG(T)CGCG-3' and 5'-GTACTT(C)GTC CTGCTG(T)CGTGT-3'), were used to amplify a piece of the choA gene from M. chitosanotabidus 3001. The mixtures contained 0.1 nM concentrations of each primer, 2.5 mM concentrations of each deoxynucleoside triphosphate, 100 ng of template DNA, 0.2 U of Ex-Taq DNA polymerase (Takara Biomedicals), and a 10× concentration of reaction buffer. Amplification was allowed to proceed through 30 cycles, where each cycle consisted of denaturation (94°C, 1 min), primer annealing (45°C, 2 min), and polymerization (72°C, 3 min). For amplification of a piece of the choA gene, the cassettes and cassette primer method (9, 14) were used to amplify a neighboring DNA of the known DNA region. Chromosomal DNA of M. chitosanotabidus was digested with BamHI and ligated with double-stranded DNA cassettes possessing BamHI sites. Then, cassette-ligated DNA was amplified by PCR by using the known sequence primers (S2, R1) and cassette primers (C1, C2).

Construction of genomic libraries and hybridization conditions. In order to construct a genome library, the total chromosomal DNA from *M. chitosanota-bidus* 3001 was isolated by the method described by Sambrook et al. (24). Genomic DNA was completely digested with *SacI*, and the fractions containing 2- to 6-kbp DNA fragments were obtained by density gradient centrifugation. The fragments were cloned into the *SacI* site of *SacI*-digested λ ZapII. The PCR-amplified 0.3-kbp DNA fragment containing a portion of the chitosanase gene of *M. chitosanotabidus* 3001 was used to probe the genomic library by plaque hybridization. Southern transfer of DNA onto a nylon membrane was performed according to the method of ECL Systems (Amersham, United King-dom).

Competent *E. coli* XL1-Blue MRF' cells were transfected with the packaged phage. Approximately 1,800 plaques were obtained, and these were screened by plaque hybridization with the PCR-amplified 0.3-kbp DNA fragment as a probe. A positive clone that contained the *choA* gene was recovered as a plasmid by in vivo excision.

DNA sequencing and analysis. Nested deletions of pChoU1 were generated by using the Kilo-Sequence Deletion Kit (Takara). DNA sequencing was performed on both double-stranded templates obtained after the subcloning of appropriate segments in the pBluescript II SK(-) and KS(-) vectors by using the dideoxy DNA sequencing system of the Ampli CycleR Sequence kit (Perkin-Elmer/Cetus) according to the procedure supplied by the manufacturer. DNA fragments were analyzed on an ABI Prism 377 DNA automated sequencer according to the instructions of the manufacturer. Nucleotide and amino acid sequence comparisons were made with the help of the GenBank database. The DNA sequence of the *choA* gene was deposited in the DDBI/EMBL/GenBank data base with the accession AB010493.

RESULTS

Purification of the chitosanase. *M. chitosanotabidus* 3001, isolated from soil collected in Matsue, Japan, was selected for its ability to form clear zones on plates of agar medium containing 1% colloidal chitosan (20). Apparently, this bacterium was an active chitosanase producer and was chosen for further studies of this enzyme. Chitosanase was purified from the supernatant fluids of broth cultures of *M. chitosanotabidus* 3001 as described in Materials and Methods. The steps used to purify the chitosanase are summarized in Table 1. The specific activity of the final preparation was 250 U/mg. The homogeneity of the purified enzyme was verified by SDS-PAGE (Fig. 2A, lane 2). The molecular weight of the purified chitosanase was estimated to be 34,000. The isoelectric point (pI) of the enzyme was determined to be 9.6 by isoelectric focusing.

To determine the pH dependence of chitosanase activity, McIlvaine buffer (pH 2.0 to 8.0), Tris-HCl buffer (pH 9.0 to

TABLE 1. Purification of chitosanase from*M. chitosanotabidus* 3001

Step	Protein (mg)	Total activity (U)	Sp act (U/mg)	Yield (%)	Purifi- cation (fold)	
Culture filtrate Ammonium sulfate precipitation	14.5 5.6	761.8 495.8	52.5 88.5	100 65.1	1 1.7	
Isoelectric-focusing chromatography	3.1	371.0	250.2	48.7	4.8	

10.0), and 0.05 M Na₂HPO₄ plus 0.1 M NaOH buffer (pH 10.0 to 12.0) were used to create a range of pH values in the reaction mixtures. The purified chitosanase showed optimal activity at pH 4.0. The stability of the chitosanase was determined in the same buffers used in the experiments described above. After 30 min of incubation at 37°C, the residual chitosanase activity was determined after readjustment of the pH to 4.0 with 0.5 M sodium acetate buffer. Chitosanase activity was found to be stable over a wide range of pH values, ranging from 4.0 to 9.0 (data not shown).

The optimum temperature of the purified enzyme under standard assay conditions was 30 to 40°C in the presence of chitosan. To determine the temperature stability of the purified enzyme, the residual activity was measured after incubation of the enzyme at various temperatures for 1 h at pH 4.0 in the absence of substrate. Ca. 92, 95, and 95% of the initial chitosanase activity remained after incubation at 20, 30, and 40°C for 1 h, respectively. At least 35% of the initial chitosanase activity remained when reaction mixtures were kept at under 50°C (data not shown).

Effects of metal ions on chitosanase activity. The effect of several metal ions on the activity of purified chitosanase was tested. Chitosanolytic activity was assayed after metal ions, including Ag⁺, Ba²⁺, Co²⁺, Hg²⁺, Ca²⁺, Cu²⁺, Mg²⁺, Fe²⁺, Mn²⁺, and Zn²⁺, were each added to the reaction mixtures at final concentrations of 1 mM. The enzymatic activity was assayed after preincubation at 30°C for 30 min. Chitosanolytic



FIG. 2. SDS-PAGE (A) and immunoblotting (B) of the purified enzyme from *M. chitosanotabidus.* (A) The homogeneity of the purified enzyme was examined by using SDS-PAGE and staining with Coomassie brilliant blue R-250 (horizontal arrow). (B) Results of immunoblots of the purified enzyme and trypsin-digested proteolytic products. The positions of molecular size standards are indicated to the left of panel A. A 19-kDa protein is the digested product which was used to determine the amino acid sequence. Lanes 1, crude proteins (1.0 mg) from culture supernatant; 2, purified chitosanase (30 ng); 3, immunoblot of the purified chitosanase (50 ng) from *M. chitosanase*.

TABLE 2. Effect of substrate specificity on the chitosanase activity^a

Substrate	Relative activity (%)
Colloidal chitosan ^b	100
Glycol chitosan ^c	113
Flaked chitosan ^d	35
Carboxymethylchitosan	6
Colloidal chitin	0
Ethyleneglycol chitin	0
Carboxymethyl chitin	0
Cellulose	0

^{*a*} Relative activity determined with colloidal chitosan is indicated as the standard. Chitosan and chitin-related compounds were used as the substrates of purified enzyme.

80% deacetylated chitosan was used.

 c 90% deacetylated chitosan was used.

^d 100% deacetylated chitosan was used.

activity was not affected by Ba^{2+} , Co^{2+} , Hg^{2+} , Ca^{2+} , Cu^{2+} , Mg^{2+} , Fe^{2+} , Mn^{2+} , or Zn^{2+} , but it was almost completely inhibited by Ag^+ (data not shown).

Specific activities of the purified enzyme with different substrates. Various derivatives of chitin and chitosan were used to determine the specificity of chitosanolytic activity (Table 2). Glycol chitosan and carboxymethylchitosan showed 113 and 6% susceptibilities to the enzyme, respectively. Chitin derivatives and cellulose were not degraded by the purified enzyme. The chitosanolytic enzyme, therefore, showed high specificity for chitosan and its derivatives. We also examined the ability of the enzyme to degrade chitosan that had been subjected to various degrees of deacetylation. Compared to the 90% deacetylated chitosan, 100% deacetylated preparations were poor substrates, as were preparations containing less than 70% deacetylation (Table 2).

Reaction pattern of the chitosanase with chitosan or its oligosaccharides. The products of hydrolysis of chitosan and GlcN oligosaccharides by the purified enzyme were analyzed by thin-layer chromatography. Table 3 shows the products of hydrolysis by chitosanase. When 80% deacetylated chitosan was employed, the amounts of tetramer, pentamer, and hexamer of glucosamine increased during the first 24 h of incubation, but their concentrations decreased after 30 h in parallel with the increase in the concentration of the dimer and trimer. The dimer and trimer of glucosamine were the final products of chitosan degradation. The monomer was never detected with the purified enzyme. The purified chitosanase could not hydrolyze the dimer and the trimer of glucosamine, but the enzyme could cleave the tetramer, pentamer, and hexamer into a dimer and a trimer. These results show that the

TABLE 3. Production of chitosan oligomers in the degradation of chitosan and glucosamine oligomers with chitosanase

Substrate	Product(s) ^a
Chitosan ^b	(GlcN)~(GlcN)
Chitosan ^c	(GlcN) ₂ ~(GlcN)
(GlcN) ₂	(GlcN)2
(GlcN) ₃	(GlcN)3
(GlcN) ₄	(GlcN) ₂
(GlcN) ₅	(GlcN) ₂ , (GlcN) ₃
(GlcN) ₆	(GlcN) ₂ , (GlcN) ₃

^{*a*} GlcN indicates residues of glucosamine, and the numbers outside the parentheses indicate the degree of polymerization.

^b Crude protein from M. chitosanotabidus 3001 was used.

^c Purified chitosanase from M. chitosanotabidus 3001 was used.



FIG. 3. Various plasmid constructions derived from the *choA* gene. pChoU1 is the chitosanase-positive clone that contained a 2.5-kbp *SacI* fragment. The names of the clones carrying the various deletion derivatives from pChoU1 are shown on the right. The direction and coding regions of the *choA* gene are indicated by the large arrow. Small arrows indicate the directions of the *lacZ* promoter. To determine the chitosanase activity, *E. coli* cells carrying pChoU1 or its deletion derivatives were transferred to the colloidal chitosan plates containing 0.4% colloidal chitosan and 50 mg of ampicillin per ml. Chitosanase activities were detected by the ability to form clear halos around the colonies on colloidal chitosan plates: +, visible halo; -, no halo. pB88ChG and pB48ChG carry the fusion genes that encode the N-terminal 88-amino-acid region and the 48-amino-acid region of ChoA fused to the GFP protein, respectively. pBG carries only the GFP gene on the pBluescript SK(+). Restriction sites are designated as follows: Sa, *SacI*; SI, *SalI*; B, *Bam*HI.

tetramer is the minimum molecular size necessary for digestion by chitosanase. These results also suggest that chitosanase from M. chitosanotabidus 3001 is an endosplitting-type chitosanase.

Immunoblotting and determination of amino acid sequence. We next determined the internal amino acid sequence of the purified enzyme. To determine the internal amino acid sequence, purified chitosanase was digested with the appropriate concentration of trypsin and then separated by SDS–12.5% PAGE. Purified enzyme and partially degraded chitosanase were blotted onto PVDF membrane and then detected by Western blotting by using an antibody raised against chitosanase (Fig. 2B). The amino acid sequence of the amino terminus of a 19-kDa peptide (Fig. 2B, lane 4), which was obtained after treatment with trypsin, was determined to be SDKNKRA ALTKIXGALQSAFDTQQDKY.

Cloning of the chitosanase gene from M. chitosanotabidus 3001. Based on the 27-amino-acid sequence described above, a mixture of two oligonucleotide primers, 5'-TCCGACAAG AAC(T)AAG(A)CG(T)CGCG-3' (S1) and 5'-GTACTT(C) GTCCTGCTG(T)CGTGT-3' (S2), was prepared, and a 80-bp DNA fragment was obtained by PCR amplification and then cloned into the vector pBluescript II SK(-). The sequence of this DNA fragment matched the amino-terminal amino acid sequence of the 19-kDa peptide. Then, the cassette-cassette PCR reaction (see Materials and Methods) was done by using M. chitosanotabidus 3001 genomic DNA to extend the region of the cloned fragment. Two oligonucleotide primers, S2 and R1 (5'-AGATCTTGGTCAGCGCCG-3'), were used to amplify DNA fragments corresponding to the coding region of the choA gene sequence. The cassette-cassette primers (C1 and C2) and the S2 and R1 primers were used to amplify the flanking fragment. A single major band of 300-bp was amplified by PCR by using the two primers, C2 and R1. Then, this 300-bp was used as a hybridization probe for genomic library screening

Southern hybridization was performed with total *M. chito-sanotabidus* 3001 DNA digested with several restriction enzymes by using the PCR products as probes. This resulted in the appearance of 1.6-, 2.5-, 4.0-, and 10-kb bands after *SalI*,

SacI, PstI, and EcoRI digestion, respectively (data not shown). We chose the SacI enzyme to construct the genomic library. The \sim 2.5-kbp SacI DNA fragment was isolated from an agarose gel by electroelution and cloned into λ ZapII.

The library was screened by plaque hybridization with the 300-bp PCR-amplified fragment as a probe. A few positive plaques carried the 2.5-kb fragment in alternative orientations. The DNA in a positive plaque of λ ZapII was changed into a plasmid by the in vivo excision method, and the plasmid obtained was designated pChoU1. (For a restriction map of the cloned fragment in pChoU1, see Fig. 3). When *E. coli* cells harboring pChoU1 were plated on a medium containing colloidal chitosan as a sole carbon source for chitosanase production, a clear degradation zone was observed, confirming that pChoU1 carries the chitosanase gene.

Deletion analysis and restriction mapping. To determine the coding region of the chitosanase gene, we performed deletion analysis. Clones containing various deletion derivatives from both directions were obtained (Fig. 3). The number associated with the name of the plasmid indicates the deletion site on the nucleotide sequence shown in Fig. 4. These deletion derivatives were transformed into *E. coli* DH10B and were tested for their ability to form clear zones on the colloidal chitosan plate (Fig. 5).

Deletion derivatives containing a 396-bp deletion from upstream to downstream (pChoU396) and a 733-bp deletion from downstream to upstream (pChoR1821) were not defective in their ability to form clear zones. However, deletion derivatives containing a 541-bp deletion (pChoU541) from upstream to downstream and a 1,072-bp deletion from downstream to upstream (pChoR1482) did not show any chitosanase activity on the colloidal chitosan plate (Fig. 5). Therefore, these results indicate that the chitosanase gene is located in the 1,425-bp region between positions 396 and 1,821.

Sequence analysis of the chitosanase gene (*choA*). The nucleotide sequence of the cloned *SacI* fragment (2,557 bp) was determined for both strands (Fig. 4). The presumed coding sequence started with initiation codon ATG at position 529, which was preceded by a putative ribosome-binding site, GGA GAGA, and extended to the stop codon, TGA, located at

CCA CAC GAT CAG CCC GAA CTC CTC CAG CTC GTT CGT GAT GTT GGC CAC CGC CTC CTG CGG GAT GAA ATC CAG CAC AAC GCG GAT GTC TT GGG GTA GGT TGT GGC CCC GGG GGC GTG TT GGC GGC GGC GC GGT GGA ACC GGG GGC GTT GTA CGG GT GGT GGC TG GGG GGC CTC CTG GGC GGG GGC GGC GGT GGA ACC GGC GGC GGG GGG GGC GGC GGC GGC	5'	GAG	CTC	GAA	GGC	GAT	CGG	CGG	CTG	GCC	GCG	CGA	CTT	GTC	CAC	CTT	CTG	CCG	CCG	54		
$ \begin{array}{c} \text{CGG} \text{ GCT} TMA CGA CAC CAT GTC CCC CAT GTC TTA CGC GCA CTG GCC GTG CTG CTG 210 GCG GCG GGC CGC GT CTA CAT GCG GGC GTTA CGC GCC GCG GCC GCG 214 TGG TGC GGG GGC GTC CAT CCT GGG GC TTA AGC GCA GGC GCC GCG GCC CTC 314 TGG TGC GGG GGC GTC CAA CCG GGC GTC TAA AGC GCA GGC GCG GCG GCC CTC CTA 315 TAT TGC TGC GGG GGC GTC GA CCG GGC GTC CAG TGTA AGC GCA GGC GCG GCG GCC CTC CTA 316 GCG CCC CAC GT GGC TAC GGC CTC GGC GCC CCC GGG GCG GCA GCC CTC CTA 316 TGC GCG GCC GCC TCC GCC TCC GCG GCC CCC GGG GCG GC$		CGA	CAC	GAT	CAG	CCC	GAA	CTC	CTC	CAG	CTC	GTT	CGT	GAT	GTT	GGC	CAC	CGT	CTG	108		
$ \begin{array}{c} \operatorname{GCG} \operatorname{GAT} \operatorname{GAA} \operatorname{ATC} \operatorname{CAG} \operatorname{CAC} \operatorname{AAC} \operatorname{GCG} \operatorname{GCG} \operatorname{GTT} \operatorname{GTA} \operatorname{CCG} \operatorname{GCC} \operatorname{CAT} \operatorname{GCG} \operatorname{GCT} \operatorname{TTG} \operatorname{GCG} \operatorname{GCG} \operatorname{GGG} \operatorname$		CGG	GCT	TAG	CGA	CAC	CAT	GTC	CGC	GAT	GTC	CTT	GCG	CGA	GGC	GGC	GCC	CTG	CTG	162		
GCC TCC TCC GCC GCC ATT CCT TC AT GCG GCT GC TCT TA AC GCC GCC GCC GCC GCC 324 TGG TGC GGG GCC GCG GTT CCA TCG GGC GCC TAT ACC GCC GCC GCG GCC CCC 332 ATT TCA TCA TA CAT CC CC GCT GC CG GGG GCT CAG GCG GGA CTT CCT 372 CAT TCA GCC CCC GCG GGC GCG GCG GCG GCG CTC GCA GGC TCT CTT 372 CAT TCA GCC CCC GCT TC GCG GCG GCG CCC GCG GGC CTC GCA GGC TAC GAC 486 CCA CCT GAT CTG CC GCT TC GCG GCG GCG CCG CGG GCG CTG CCA GGC ACC TTG GCG GCT 594 S TCC ACC GC GCG GCG GCG GCG GCG GCG CCA AGG ACG TAG GCG TAC GTG CAG 468 CCA CCT GAT CTG CC GCC GCC GCG GCG CCC AAG ACG GCC AGC AG		GCG	GAT	GAA	ATC	CAG	CAC	AAC	GCG	GCG	GTT	GTA	CGG	GGC	ACT	GGA	CTG	TTG	GGT	216		
$ \begin{array}{c} \begin{tabular}{lllllllllllllllllllllllllllllllllll$		GGC	TÇC	TCG	GCC	AAT	CGT	TCT	CAT	GCG	GCT	GCT	GGG	TTC	GTT	GCG	CCC	GGG	ATG	270		
The The Graph and the field of the transformation of tr		GCG	GGC	GGG	CCC	CCG	GTG	TCA	GGG	CAG	CGT	TAT	ACG	CGT	GCC	GCG	CGG	ACC	CCG	324		
ATT TCA ATA CAT CCA CTG GTT TTA ATA ATT CAC GCC ACC GCG GA CTC CTC CTAC GAG 432 CAT TCA GG GCG CGC ACG GC TTG GA GCA FTC GCG CGC CTG CAC GAG ACTT CT 466 CCC CCA CT GGC TAC GGC CTG GCG GCG CGC GG GG CTG CGG CTG CGG GG GG GAC AAC 540 CGA CCT GAT CTG CGC TCG CGC GGC GGC GG GG GG CTG GGG GG		TGG	TGC	GGG	GCG	GGT	CGA	CCG	GGC	GCT	CAG	TGT	AAT	CCC	CGC	CCC	GGG	CCT	CAC	378		
$ \begin{array}{c} ch \ trc \ cc \ cc \ cc \ cc \ cc \ cc \ c$		ATT	TCA	ATA	CAT	CCA	CTG	GTT	TTA	AAT	ATT	GAC	GGC	AGC	GCG	GAA	CCT	CTC	CTA	432		
$ \begin{array}{c} ccc cch cor acc the cor the cor cor cor cor cor cor cor cor the chart to the form of the cor cor form of the cor cor cor cor cor cor cor cor cor cor$		CAT	TCG	GCG	CCG	GCG	AGC	GGC	TTG	GAA	GCA	GTC	GCG	CGG	CTC	GCA	GGC	TAC	GAC	486		
1 Constant of the first of the		CCC	CCA	CGT	GGC	TAC	GGC	CAC	ACG	ACT	TCA	GAG	AGG	CAA	CCC	ATG	CAA	CTT	CCT	540		
ca ccr car cr car ccr car cr cac cc cac ca	1															м	0	т.	P			
5 R. P. D. L. R. R. F. A. R. R. A. A. L. P. L. L. A. A. G. A. C. C. C. C. C. C. C. A. C.	-	CGA	сст	GAT	CTG	CGT	CGC	TTC	GCG	CGC	CGC	GCG	GCG	CTG	CCG	CTG	CTC	600	GOT	594		
TTC FAC FTG GCG GCG TTG GGC GCG GCG GCG GCG GCG G	5	R	p	D	т.	R	R	-1- F	Δ	R	R	Δ	Δ	т.	P	т.	τ.	Ă	Δ	371		
23 S. T. L. Q. X. X. L. G. X. L. Z. S. C. Z. X. S. S. C. Z. X. S. S. J. L. Q. X. G. K. C. K. C. AC G. TAC G. C. TAC G. C. TAC G. TAC G. C. TAC G. TAC G. C. TAC G. TA	2	TCC	ACC	CTG	GCC	GCC	GCG	TTC	GGC	808	GCC	TCG	CCG	GCC	CTG	606	aca	GGC	AAC	648		
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	23	200	<u>т</u>	т.	2000	2000	200	- TC TT	ä	Δ	a a	5	D	2	T.	200	200	G	N	010		
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	25	COT	GTC	200		ATC	202	200	200	110	AGC	000	600	100	700	000	TAC	CTTC	202	702		
and the form of t	4.1	201	510	GCC	000	- M	GCC 7	660	CCC D	- AAG	AGC	000	GCC	AGC C	AGC C	GCC 2	TAC	316	CAG	/02		
$ \begin{array}{cccc} part Get 1 for the Cite fact Cite fact Cite fact Cite Cite Cite Cite Cite Cite fact Cite fact for $	41			THE C	CILC.	m 2.0	100	010	100	010	200		200		010					750		
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	50	GAT	GGC	TGG	UTC T	TAC	ACC	CAC	ACC	CAC	ACG	GCC	ACC	GGC	GAG	CCG	CTG	GIC	ACC	/ 20		
be are and be defined by an analysis of the form of t	29				200	222	200	and a	202	and a	-T-	_A_		G		5.5	Trans.		T and			
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		GCC	ACC	AAG	GCC		GUU	600	606	660	GTG	ATT	006	GTC	GGC	GAC	TCG	CGC	GTC TT	810		
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	11	_A_	T	K.	-A/		<u>A</u>	<u>A</u>	<u>A</u>	G	<u>_v</u>	1	P	<u>V</u>	G	0	<u> </u>	<u></u>				
95 $\frac{\mathbf{y}}{\mathbf{x}}$ $\frac{\mathbf{x}}{\mathbf{x}}$ $\frac{\mathbf{y}}{\mathbf{x}}$ $\frac{\mathbf{y}}{\mathbf{x}}$ $\frac{\mathbf{x}}{\mathbf{x}}$ $\frac{\mathbf{y}}{\mathbf{x}}$ $\frac{\mathbf{y}}{\mathbf{x}}$ $\frac{\mathbf{x}}{\mathbf{x}}$ $\frac{\mathbf{x}}{\mathbf{x}}}$ $\frac{\mathbf{x}}{\mathbf{x}}$ $\frac{\mathbf{x}}{$		TAT	GGC	GCG	GTC	<u>117C</u>	GAC	AAG	GGC	_CGC	AAG	CTG	ACG	GTC	AAT	CAA	TGG	CAG	GCC	864		
GTG CTG AGC AGC GCC TAC CCC GAG AAC GGC ACG ACC AAC TAC CAG GAG GTC918GGC CCC TGG GGC GTA TT CC GAC TAT TGC GAA TTC GAC TAC GAG GCG GC AA GGG ATC TCC GAC 72TAC CGC GGC GGC GTA TT TG GGT CGG GTC GGC GGC GGA GGG GAC TTC CCC 1026119YNY918YQYTYQ17YQYTYQ1YQQTTQQ1YQQTTQQPPYGC1080TXRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR <td <="" colspan="2" td=""><td>95</td><td>Y</td><td><u> </u></td><td><u>A</u></td><td><u>v</u></td><td>F</td><td>D</td><td>K</td><td>G</td><td>R</td><td>K</td><td>L_</td><td>т</td><td>v</td><td>N</td><td>Q</td><td>W</td><td>Q</td><td>A</td><td></td></td>	<td>95</td> <td>Y</td> <td><u> </u></td> <td><u>A</u></td> <td><u>v</u></td> <td>F</td> <td>D</td> <td>K</td> <td>G</td> <td>R</td> <td>K</td> <td>L_</td> <td>т</td> <td>v</td> <td>N</td> <td>Q</td> <td>W</td> <td>Q</td> <td>A</td> <td></td>		95	Y	<u> </u>	<u>A</u>	<u>v</u>	F	D	K	G	R	K	L_	т	v	N	Q	W	Q	A	
113 V L S M D A Y P E N G T T N Y Q E V GGC CCC TGG CCC TAT CCC GAA GTC GAC TAC GAC GTG GCG ACA GGA ATC CA GG TTC CCC TGG GCC ACC TTC GGC CAC CTC GCC GC GTG CCA GGG ATC TCC CAC TTC CGC GCC GAC ACC TTC GGC CTC GTC GCC TAC CAC GTG GCG GCA ATC TC CCC 1026 GAC TAC TTC AAC AAG GCC TTC GGC CTC GTG GCC TTG GCA AGA ACC AAC GGA ACC TTC CAC TTC AGA AAG GCC TTC GGC GC GTG CAG GTC ACC GCG GC GCA ACC 1131 S A D M L A W G V Q V T G V T A A G N TTC CAG GCG GAC ATG CTG GCC TGG GCC GTG GAC GTC AAG AGC CGC GC GAC AC 1132 AAC GCC GAC ATG CTG GCC TGG GCC GTG GAT CCG TAT CCG TCC AAG AGC CGC GC GAC AC 1135 N A D M L A W G V Q V T G V T A G N TTC CAG GCG GAC GAC ACG GCG CTG GAT CCG TAT CCG TCC AAG AGC CGC GCC ACC AAG AAC AAG CGG GCG CTG ACC AGA ATC TG GCG CGC GCC GC GC AA TCG GCC TTC 1242 AAG AAC AAG CGG GCG CTG ACC AAG ATC TG GCG CGC CTG CAA AT CG GCC TTC AAG CTG GTC CCG GTG CTG AAC AAG ATC TC CAG GCA CTC ACC GCC TTC AAG CTG GTC CCG GTG CTG AAC AAG TAC GCG CAC TTC GAC CAG GAC 1350 AAG CTG GTC CCG GTG CTG AAC AAG TAC GCG CAC GTC GAC CAG GAC ATC 1404 175 R Y N L V F V L N A L K G I G F T A F D CGC TAC AAC CTG GTG GTG GTG AAG CAC GTG AAA CAG GGC TCG AAC TAC GG CAC TC 1404 175 R Y N L V G L A F Q V Q V N T G S I AGC CG GAA CTC GCC TTC CTC CG GTG AAA CAG GAC CTC GG CTG 1404 175 R Y N L V G L A F Y Q V Q V N T G S I 1660 176 GAC ATC TAC AAC CTG GTG GCT GCA CAT CCG GAC GAC AAC TAC CGG CAG CTC 1458 176 GAC ATC TAC AAG AAG GAC CCG ACC GAC AAC TAC CGG CGG CAG CTC 1458 177 A F Y N L V G L A F S S V K S A G N C C G S L 170 AAG CTG GCC TGC TGC TGC TGC GAG GAA AAC ATC GG CTG ACC TTC 174 176 CTC AGC CTG GCC TGC TCC CTG GGA AAC ACG GCC CGC GCG CGG CGC CTG GC TG 176 177 Y N N L V G F A T Y L T D Q Y I R W L 176 CT CAC ACC CG GCC TGC GCC GCG GCG GCC CGC GCG GCG		GTG	CTG	AGC	ATG	GAC	GCC	TAC	ccc	GAG	AAC	GGC	ACG	ACC	AAC	TAC	CAG	GAG	GTC	918		
GC CCC TGG CGC TAT TGC GAA GTC GAC TAC GAG GCG CCA GGC ACTC TCC GAC 972 TAC CGC GGC GAC ACC TTC GGT CGC GTC GGC GTC ACC ACG GTG GGC GAC TTC CCC 1026 TAC CGC GGC GAC ACC TTC GGT CGC GTC GGC GTC ACC ACG GTG GGC GAC TTC CCC 1026 GAC TAC TTC AAG AAG GCC TTC GGC CTG GGC GTC GTC GTC GGC AAG AGC AAC GGC ACC 1080 TAC GGC GAC ATG CTG GCC TGG GGC GTG CAG GTC ACC GGC GTG ACG GCG GGC AAC 1134 TAC GGC GAC ATG CTG GCC TGG GGC GTG CAG GTC CG TAT CCG TCC AAG GGC GGC GAC AC 1138 TTC CAG GCG GAC CAA GC GCC GTG GAC GTC GTAT CCG TCC GAC AG AGC CGC TC GAC 1188 CAG AC AAG CGC GAC CAC AGG GCC GTG ACC AAG ATC TGC GGC GGC GGC AAC 1128 AAG AAC AAG CGC GGG GGC GTG ACC AAG ATC TGC GGC GGC GCC AATC GAC CAG GAC 1290 AAG AAC AAG CGC GGG GCG CTG ACC AAG ATC TGC GGC GCG CAC ATC GAC CAG GAC 1290 AAG CAC AAG CGC GGG GCG CTG GAC CAG ATG AGC CAT TAC GGC CAC ATC GAC CAG GAC 1290 AAG CAC ACG CGG GAC AAG TAC GTG ATG AGC CAT TAC GGC CAC ATC GAC CAG GAC 1290 CACALANCANCANCANCANCANCANCANCANCANCANCANCANCA	113	v	Г	S	M	D	A	Y	Р	Е	N	G	т	т	N	Y	Q	Е	v			
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$		GGC	ccc	TGG	CGC	TAT	TGC	GAA	GTC	GAC	TAC	GAG	GCG	GCG	CAA	GGG	ATC	TCC	GAC	972		
TAC CGC GGC GAC ACC TTC GGT CCG GTC GGC GTC ACC ACG GTC GGC GAC TTC CCC 102 GAC TAC TTC AAG AAG GCC TTC GGC CCT AC GTC GGC CAC AG GAC ACC GGC ACC TTC GGC GAC ATT AG AAG GCC TTC GG CCT GG GC CTAC GTC GGC CAC GGC ACG GGC ACC 1134 AAC GCC GAC ATG CTG GCC TGG GGC GTG GAC GTC ACC GGC GTG ACG GCG GAC ACC TTC CAG GCG GAC CATG CTG GCC TGG GGC GTG CAG GTC CAC AG AGC GCC GCC GCC ACC TTC CAG GCG GAC CACG GCG CTG GAT CCG TAT CCG TCC AAG AGC CGC GCC GCC CTC AAG AAC AAG CGC GCG GCG GCG CTG ACC AAG ATC TGC GCC CAC ATC GAC CAC GAC ATC AAC ACC CAC CAC AAG TAC GTG ATC AGC CAT TAC CGC CAC ATC GAC CAC GAC ACC AAG AAC AAG CGC GCG GCG GCG TTC ACC AAG ATC TGC GCC CAC ATC GAC CAC GAC ATC AAC ACC CAC CAC AAG TAC GTG ATC AGC CAT TAC CGC CAC ATC GAC CAC GAC ATC AAC ACC CAC CAG CAC AAG TAC GTG CTT CAC GGC ATT CAC CCC CC CTC CAC ATC AAC CTG GTC CCG GTG GTG GAC GCG CTC CAG GGC ATC GAC CAC GAC CAC GAC ATC AAC CTG GTC CCG GTG GG CTG AAC GCG CTC TC CAG GTC AAC CG GC TCC AAC AAG CTG GTC CCG GTG GG CTT CCC GG GTC AAG GCC ATC GAC CAC GAC CAC TC CAC GAC ATC AAG CTG GTC CCG GTG GG CTT CCC GG GTC AAG GCC ATC GAC CAG GTC AAC CG GC CTC 140 AAG CTG GTC CCG GTG GCC TTC CCC GG CTC CCAG GTC AAC CCG GC CAC ATC AAG CTG GTC CGG GTG GC TTC CCC GC TTC CAG GTC AAC CCG GC CAC ATC CTC GCC TC GCC ACC AAC CTG GTC GCC TTC CCC GC TTC CGG GAC CAC TAC CGC GC CAC CTC 140 AAG CTC GCAG ACC TGC TTC GCC ACC TAC CTG AGG GAC CAC TAC CGC GG CAC CTC AAG CCG CAG ACC TGC TTC GCC ACC TAC CTG AGG GAC CAC TAC CGC GGC ACC CTC AAG CCG CAG ACC TCC TCC CCC GCC AACA CCG CGC CAC CTG CCC GCC GCC GCC CCC CCC CCC CCC CC	131	G	₽	W	R	Y	С	É	v	D	Y	Е	Α	А	Q	G	I	s	D			
149 Y R G D T F G P V G V T T V G D F F P GGC TAC TTC ANG AAG GCC TTC GCG CCC TTC GCC CCC TAC GC GGC GGC AG AGC AGC GGC GGC AAC 1134 185 N A D G V T C G V T A G N T A G N T A G N T A G N T A G N T A A G V T T C G C C V T C G C <td< td=""><td></td><td>TAC</td><td>CGC</td><td>GGC</td><td>GAC</td><td>ACC</td><td>TTC</td><td>GGT</td><td>CCG</td><td>GTC</td><td>GGC</td><td>GTC</td><td>ACC</td><td>ACG</td><td>GTG</td><td>GGC</td><td>GAC</td><td>TTC</td><td>CCC</td><td>1026</td></td<>		TAC	CGC	GGC	GAC	ACC	TTC	GGT	CCG	GTC	GGC	GTC	ACC	ACG	GTG	GGC	GAC	TTC	CCC	1026		
GAC TAC TTC AAG AAG GCC TTC GCG CCC TAC GTG GTG GGC AAG AGC AAC GCG ACC 1070 AAC GCC GAC ATG CTG GCC TGG GGC GTG CAG GTC ACC GGC GTG ACG GCG ACC 1134 185 N A D M L A W G V Q V T G V T A G N TTC CAG GCG GAC GAC ACG GCG CTG GAT CCG TAC CG CC AAG AGC CCC TCC GAC 1138 203 F Q A D D T A L D P Y P S K S R $\frac{1}{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{\sqrt{$	149	Y	R	G	D	т	F	G	Р	v	G	v	т	т	v	G	D	F	P			
167 D Y F K K A F F A P Y V L G K S N A T AC GC GAC AC GG GG CTG GC TG GG GT GAG GT CAG GC GT AC GG GG GC AC 1134 185 N A D M L A W G V Q V T G V T A G N TTC CAG GG GAC GAC ACG GG CTG GAT CG TA CG TC CAA GAC GC <u>GC TC CAC</u> AC AC AC <u>AC CG GC GC GC GC GT GAT CG TT CG TC AA GAC CG <u>CC CC CAC</u> AC <u>AC AAC CC GC GC GC GC GC AC AA GAT CG GC GC GC GT AAA AC CAC CG GC GC GC GT AC AA GAT CG GC TT 1242 203 F Q A D D T A L D P Y P S K S R <u>S. R. C. C.</u></u></u>		GAC	TAC	TTC	AAG	AAG	GCC	TTC	GCG	CCC	TAC	GTG	CTG	GGC	AAG	AGC	AAC	GCG	ACC	1080		
AACGCCGACATCCTGGCCTGGCCGCGTCACCGCCGCCGCCGCCACCGCCACC	167	D	Y	F	ĸ	ĸ	А	F	А	₽	Y	v	L	G	ĸ	S	N	Α	т			
185NADMLAWGVQVTGVTAGNTTCCAGGGGACACCGGCTGGACCGTCCGACACCGCCGGACACCGCTTALDPYPSKSRSSSSSSSSSSSSSSSSSS <td< td=""><td></td><td>AAC</td><td>GCC</td><td>GAC</td><td>ATG</td><td>CTG</td><td>GCC</td><td>TGG</td><td>GGC</td><td>GTG</td><td>CAG</td><td>GTC</td><td>ACC</td><td>GGC</td><td>GTG</td><td>ACG</td><td>GCG</td><td>GGC</td><td>AAC</td><td>1134</td></td<>		AAC	GCC	GAC	ATG	CTG	GCC	TGG	GGC	GTG	CAG	GTC	ACC	GGC	GTG	ACG	GCG	GGC	AAC	1134		
TTC CAG GCG GAC AGC AGG GCG CTG GAT CCG TAT CCG TCC AAG AGC CCG TCC GAC TTC CAG GCG GAC GAC AGG GCG CTG GAT CCG TAT CCG TCC AAG AGC CCG TC GAC AAG AAC AAG CGC GCG GCG CTG ACC AAG ATC TGC GGC GCG CTG CAA TCG GCC TTC AAG AAC AAG CGC GCG GCG CTG ACC AAG ATC TGC GGC GCG CTG CAA TCG GCC TTC AAG CAG CAG CAG GAC AAG TAC GTG ATG AGC CAT TAC GCG CAC ATC GAC CAG GAC T290 AAG CTG GTC CCG GTG CTG AAC GCG CTC AAG GGC ATC GGC TTC ACC GCC TTC GAC TAC GCG CAC CTG GTG GAC AGG GCC TTG CAG GGC ATC GAC CAG GAC TCG CTAC AAC CTG GTG GAC GCG CTG GAC GTG CAG GGC ATC GAC CAG GAC TCG TAC AAC CTG GTG GGC GTG GAA GGC CTT CCAG GTG CAG CTG CAC GCC TTC GAC TCG TAC AAC CTG GTG GCC TTC CTC GG GTG AAG GGC AAC ACG GGC TCC ATG GGC AGC ATC TCG GCC TTC CTC TCG GTG AAG TCG GGC GC AAC TGC GGC AGC TTC AAG CTC CAG CCTG GCC TTC CCC TGC GTG AAG TCG GGC CAC TTC GGC GGC AGC TTC AGG GCC GAG ACC TGC TTC GCC CTC CTG GTG AAG CAG GGC AAC TGC GGC AGC CTC AAG CTC CAG CCTG GCC TTC CCC TAC GTG AGG GGC CAG TAC ATC CCG TGG CTG 1512 311 S A E T C F F A T Y L T T D Q Y I R W L AAG TCC TC GAC CTG GGC GAC GAT CCG GAC AAC TGC GGC GGC AGC ATG CGG 329 K S S S L G D D P T M G S V S V V N Q GTC ATC AAC GCG AGC TAT CCC GGC GAC GAC GTG GGC GAC GTG AAT CAG GTC ATC AAC GCG AGC TAT CCC GGC GAC GAG AGC CGG GCG GCG GCG GCG	185	N	Α	D	м	T.	Α	W	G	v	0	v	T	G	v	T	Α	G	N			
110 CAG GCG GAC ACG GAC ACG GCG C1G GAT CCG TAT CCG TAT CCG TCG ACG CCG C1G CAT CCG CCG C1G CAT CCG C1G CAT C1242 203 \overrightarrow{Mac} AAC AAG CGC GCG GCG CTG ACC AAG ATC TGC GCG CCG CTG CAA TCG GCC TTC 1242 221 $KKKKKKKKKK.$		mmo	C . C		C 2 C	<u></u>	200		and	сл. т.		- 	000	таа		200			- N -	1100		
203 F Q A D D T A L D P Y P S K S R S R S Q AAG AAG AAG CGC GGC GTG ACC AAG ATC TGC GGC GC GC GC GC GC GC GC TTC ACC GAC TTC 221 K.M.K.K.R.M.K.A.K.L.T.K.K.L.T.K.L.T.C.G.G.A.K.L.G.G.S.K.K.F. G <u>A</u> ACG CAG CAG GAC AAG TAC GTG ATG AGC CAT TAC GCG CAC TC GAC CAG GAC 1290 239 M.C.T.Q.Q.Q.L.K.Y V M S H Y A H I D Q D AAG CTG GTC CCG GTG CTG AAC GCG CTC AAG GGC ATC GGC TTC ACC GCC TTC GAC 1350 257 K L V P V L N A L K G I G F T A F D CGC TAC AAC CTG GTG GGC GTG GCC TTC CAG GTG CAG GTG AAC ACG GGC TCG ATC 1404 275 R Y N L V G L A F Q V Q V N T G S I GGC ACC ATC CG GTG TTC TC TC CG GTG AAG TCG GCC GC AAC TGC GCC AGC CTC 1458 293 G S I S A F S S V K S A G N C G G S L AGC GCC GAG ACC TGC TTC GCC GCC ACC TAC GG GAC CAG TAC ATC CGC TGG CTG 1512 311 S A E T C F A T Y L T D Q Y I R W L AAG TCC TC CAG CTG GGC GAC GAT CGG GAC AAC TGC TGC GGC AAC TGG GG 1566 239 K S S S L G D D P D N C W R A S M A CTG GAC ATC TAC AAG AAG GAC CGG ACG ATG GGC AGC GTC AGC GTG GC AAC TGC GGC GCC ACC TG GC GAC ATC GG CTG 310 S Y I N A S Y P G N S G K C C P T S G I AAG TGC TC CC CGC GCC GAT CCG GCA AGC GGC GGC AAC TGC GGC GGC GCG CTG 347 L D I Y K K D D T M G S V S V V N Q GTC ATC AAC GCG AGC TAT CCC GCC ACC GAG GCG GCG GCG CGC GCG GCG		TIC	CAG	GCG	GAC	GAC	ACG	GCG	CIG	GAT	CCG	TAT	CCG	TUU	AAG	AGC	CGC	TUU	GAC	1100		
AAG AAC AAG CGC GGG GGG GTG ACC AAG ATC TGC GGC GCG CTG CAA TCG GCC TTC1242221K.M.K.R.R.A.A.G.A.G.TAC GTG ATC AAG ATC TGC GGC GCG CTG CAA TCG GCC TTCAAG CGC CAG CAAG TACGTG ATG AGC CAA TAC GGC CTC CAG GAC ATC GAC CAG GAC CAG GAC1290239D.M.L.C.C.M.G.L.K.X.XVMSHYAHIDQDAAG CTG GTC CCG GTG CTG AAC GCG CTC AAG GGC ATC GGC TTC CAC GCC TTC GACGCC TTC CAC GCC TTC GAC1350257KLVPVLNALKGFTAFDcGC TAC AAC CTG GTG GGC TTC CTC GG GTG CAG GTG AAC TAC GG GCC GCC ATC1404275RYNLVGLAFQVNTGSIgGC ACC ATC TGG GCC TTC TCC TCG GTG GTG AAG TAC GG GC AAC TGC GGC AAC TGC GGC ACC TTCTCG GC ATC TCG GCC TTC TCC TCG GTG AAG TAC GGC ACC TAC CTG ACG GAC CAC TAC CTG CGC GCG ACC ATG GCG1458293GSISAFSSVKSAGNCGSLAGC CCC GAG ACC TGC TGC GCC GCC CAC CTA CTG ACG GAC CAC TAC CTG GCC GCG GCG ACC TGTTTDQYIRWL11SAFTTLTDQYIRML121SAFTTLTDQYIRML132K<	203	F	Q	A	D	D	т	A	L	D	P	Y	P	s	ĸ	s	R	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	~R			
221 K. N. K. R. A. L. T. K. L. C. G. A. L. O. S. A. F. GAC ACG CAG CAG GAC AAG TAC GTG ATG AGC CAT TAC GCG CAC ATC GAC CAG GAC 1290 D. C. C. C. G. C. G. C. G. A. G. G. C. C. ATC GAC CAG GAC CAG GAC AAG CTG GTC CCG GTG CTG AAC GCG CTC AAG GGC ATC GGC TTC ACC GCC TC GAC 1350 C. C. G. C. G. G. C. G. A. G. G. C. C. A. G. G. C. C. G. C. C. C. G. C.										110	3 6 6	mco	000	000	OT C	033	maa	000	mma			
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $		AAG	AAC	AAG	CGC	GCG	GCG	CTG	ACC	AAG	ATC	TGC	GGC	GCG	CIG	CAA	TCG	ほしし	110	1242		
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	221	AAG K	AAC N	AAG K	CGC R	A	GCG A	CTG L	ACC T	K	T		GGC	A	CTG T.	O	S	A	F IIC	1242		
239 D.T.T.Q.M.Q.M.Q.M.X.X.X V M S H Y A H I D Q D AAG CTG GTC CCG GTG CTG AAC GCG CTC AAG GCC ATC GGC TTC ACC GCC TTC GAC 1350 257 K L V P P V L N N A L K G I G F T A F D CGC TAC AAC CTG GTG GGC CTG GCC TTC CAG GTG CAG GTG AAC ACG GGC TCG ATC 1404 275 R Y N L V G L A F Q V Q V N T G S I GGC AGC ATC TCG GCC TTC TCC CTG GTG AAG CGC GGC GAC ACC GGC GC ACC TC 1458 293 G S I S A F S S V K S A G N C G S S L AGC GCC GAG ACC TGC TTC GCC ACC TAC CTG AGG GAC CAG TAC ATC CGC TG CTG 1512 311 S A E T C F A T Y L T D Q Y I R W L AAG TCC TCC AGC CTG GGC GAC GAT CCG GAC AAC TGC GGC AGC ATG GCG 1566 329 K S S S L G D D P D N C W R A S M A CTG GAC ATC TAC AAG AAG GAC CG CG GAT GGC GAG GGC GAC GTG GTG GTG 1566 329 K S S S L G D D P T M G S V S V V N Q GTC ATC AAC GCG AGC TAT CCC GGC AAC GCG GGG AAA TGC CG GG GAG ATG GCG 1566 347 L D I Y K K D P T M G S V S V V N Q GTC ATC AAC GCG AGC TAT CCC GGC CAC GCG GGG AAA TGC CG CG CG CG CG CT 1728 383 K W S K N M S W Q * 392 ACG CG CAC GCC CTG GAC CTC CCC CCC CGC CGA GCC CGC GCG CGC CGC 1778 AAG TGG TCG AAG AAC ATG TCG CG CG CTC CGA CGC CGC CGC GGG CGC CGC 1782 ACG CG CAC GCC CCT GAT CCC CGC CCC CGA CGC CCC GGG CGC CGC CGC	221	AAG Kan	AAC Nor	AAG					ACC	AAG	MTC			- Andrew			-100 			1242		
AAG CTG GTC CCG GTG CTG AAC GCG CTC AAG GGC ATC GCC TTC ACC GCC TTC GAC 1350 257 K L V P V L N A L K G I G F T A F D CGC TAC AAC CTG GTG GGC CTG GCC TGC CAG GTG AAC ACG GGC TC ATC 1404 275 R Y N L V G L A F Q V Q V N T G S I GGC AGC ATC TCG GCC TTC TCC TCG GTG AAG TCG GCC GGC AAC TGC GGC ACC TC 1458 293 G S I S A F S S V K S A G N C G S L AGC GCC GAG ACC TGC TTC GCC ACC TAC CTG ACG GGC ACC CAG TAC ATC CGC TGG CTG 1512 311 S A E T C F A T Y L T D Q Y I R W L AAG TCC TCC AGC CTG GGC GAC GAT CCG GGC AAC TGC GGC GGC GAC ATT CG 1562 294 K S S S L G D D P D N C W R A S M A CTG GAC ATC TAC AAG AAG GAC CCG ACG ATG GGC AGC GTC AGC GTG GAAT CAG 1620 347 L D I Y K K D P T M G S V S V V N Q GTC ATC AAC GCG AGC TAT CCC GGC CAC GATG GGC AAC CCG ACG CTG GG AAT CAG AAG TGG TCG AAG AAC ATG TCG TGG CGC ACC GATG GGC AAC CCG ACG CGC GCG CC GCG CCG CCG	221	GAC	AAC	AAG ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	CAG	GAC	AAG	TAC	ACC GTG	AAG	AGC	CAT	TAC	GCG	CAC	ATC	GAC	CAG	GAC	1242		
257 K L V P V L N A L K G I G F T A F D 275 R X N L V G G A F Q V Q V N T G S I 275 R Y N L V G G S T A F Q V Q V N T G S I 1404 275 R Y L A F Q V Q V N T G S I A F C	221 239	GAC D	AAC ACG	AAG ~~~~ CAG		GCG GAC	AAG	TAC	ACC GTG V	AAG ATG M	AGC	CAT H	TAC Y	GCG GCG A	CAC H	ATC I	GAC D	CAG Q	GAC D	1242		
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	221 239	AAG GAC D AAG	AAC ACG ACG CTG	AAG CAG CAG GTC		GAC GAC GTG	AAG	TAC AAC	GTG V GCG	AAG ATG M CTC	AGC S AAG	CAT GGC	TAC Y ATC	GCG GCG A GGC	CAC H TTC	ATC ACC	GAC GAC D GCC	CAG Q TTC	GAC D GAC	1242 1290 1350		
275 R Y N L V G L A F Q V Q N T G S I GGC AGC ATC TCG GCC TTC TCC TCG GTG AAG TCG GCC AGC CAC TGC GGC AGC CTC 1458 293 G S I S A F S V K S A G N C G S I AGC GCC GAG ACC TGC TTC GCC ACC TAC CTG ACG GAC GAC CAG TAC ATC CGC GG GCA GAC ATG CCG GAC AAC TGC TGG CGG GAC ATG GCG 1512 311 S A E T C F A T Y L T D Q Y I R W L AAG TCC CAGC CTG AGC ATC GGC GAC GAC GAC GAC GAC ACG TGG GAC AAC TGC TGG CGG GAC ATG ATC CAG I D N N A S M A CTG GAC ATC TAC AG AGA AG GAC CTG TGC CGC ACG AGC ATG CCG GAC ATG CCG GCG AGC GAC CGA GAC ATG CCA CGC AGC AGC GCG AGC GCC CGA CGC AGC GCG CGA GAC ATG CCA CGA CAG ATG CCC GGC AGC GCA GAC CGA CGC GCG AGC CCC AGC GCG CGC GCG CGC GCG CGC GCG CGC GCG CGC GCG CGC CGC GCG CGC CGC GCG CGC CGC GCG CGC GCG CGC CGC GCG CGC CGC GCG CGC CGC GCG CGC CGC GGC CGC CGG GGC CGC GGG CGC CGG GGC CGG GGC CCC CGG GGC CGC CGC G	221 239 257	GAC GAC AAG K	AAC ACG CTG L	AAG CAG CAG GTC V	CGC CAG CAG CCG P	GCG GAC GAC GTG V	AAG CTG	TAC AAC	GTG V GCG A	AAG ATG M CTC L	AGC S AAG K	CAT H GGC	TAC Y ATC I	GCG GCG A GGC GGC	CAC H TTC	ATC I ACC	GAC GAC D GCC A	CAG Q TTC F	GAC GAC D GAC D	1242 1290 1350		
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	221 239 257	AAG GAC D AAG K CGC	AAC ACG CTG L TAC	AAG CAG GTC V AAC	CGC CAG CAG CCG P CTG	GCG GAC GAC GTG V GTG	AAG AAG CTG L GGC	TAC TAC AAC N CTG	GTG V GCG A GCC	AAG ATG M CTC L TTC	AGC S AAG K CAG	CAT H GGC GTG	TAC TAC Y ATC I CAG	GCG GCG A GGC GGC GTG	CAC H TTC F AAC	ATC ATC I ACC T ACG	GAC GAC D GCC A GCC	CAG CAG Q TTC F TCG	GAC GAC D GAC D ATC	1242 1290 1350 1404		
293 G S I S A F S S V K S A G N C G S L AGC GCC GA ACC TG TT G C TT D Q Y I R W L AAG TCC TC F A T Y L T D Q Y I R W L AAG TCC TC F A T Y L T D Q Y I R W L L D Q Y I N K L D D Q Q I R M A S M A S M A S M A S M A S M A S N M A S N N S S V N N S G X N <t< td=""><td>221 239 257 275</td><td>AAG GAC AAG K CGC R</td><td>AAC ACG CTG L TAC Y</td><td>AAG CAG GTC V AAC N</td><td>CGC CAG CAG CCG P CTG L</td><td>GAC GAC GTG V GTG V</td><td>AAG AAG CTG GGC G</td><td>TAC TAC AAC N CTG</td><td>GTG GTG CCG A GCC A</td><td>AAG ATG M CTC L TTC F</td><td>AGC S AAG K CAG O</td><td>CAT H GGC GTG V</td><td>TAC Y ATC CAG</td><td>GCG GCG A GGC GTG V</td><td>CAC H TTC F AAC N</td><td>ATC ATC ACC T ACC</td><td>GAC GAC D GCC A GCC GCC GC</td><td>CAG Q TTC F TCG S</td><td>GAC GAC D GAC D ATC T</td><td>1242 1290 1350 1404</td></t<>	221 239 257 275	AAG GAC AAG K CGC R	AAC ACG CTG L TAC Y	AAG CAG GTC V AAC N	CGC CAG CAG CCG P CTG L	GAC GAC GTG V GTG V	AAG AAG CTG GGC G	TAC TAC AAC N CTG	GTG GTG CCG A GCC A	AAG ATG M CTC L TTC F	AGC S AAG K CAG O	CAT H GGC GTG V	TAC Y ATC CAG	GCG GCG A GGC GTG V	CAC H TTC F AAC N	ATC ATC ACC T ACC	GAC GAC D GCC A GCC GCC GC	CAG Q TTC F TCG S	GAC GAC D GAC D ATC T	1242 1290 1350 1404		
AGC GCC GAG ACC TGC TTC GCC ACC TAC CTG ACG GAC CAG TAC ATC CGC TGG CTG1512311SAETCFATYLTDQYIRWLAAG TCC TCA GC CTG GGC GAC GAT CCG GAC AAC TGC TGG CGC GG GAG CATG GCG1566329KSSSLGDDPDNCWRASMACTG GAC ATC TAC AAG AAG GAC CCG ACG ATG GGC AGC GTC AGC GTG GTG AAT CAG1620347LDIYKKDPTMGSVVNQGTC ATC AAC GCG AGA GTT CCC GGC AAC AGC GGG AAA TGC CCG ACG TCA GGG ATC1674AASYPGNSSVVNQGTC ATC AAC GCG AGA CTT CCC GGC AGA GGG GAC GCG GCG GCG CGC GCG CGG CGG CGC CGG 1728383KWSKNNSWQ*392ACG CGG CAC GCC GCG CGC CGC CGC CGC GCG CGC GGC G	221 239 257 275	AAG GAC D AAG K CGC R GGC	AAC ACG CTG L TAC Y AGC	AAG CAG GTC V AAC N ATC	CGC CAG CAG CCG P CTG L TCG	GCG GAC GTG V GTG V GCC	AAG AAG CTG GGC G TTC	TAC TAC AAC N CTG L TCC	GTG GTG V GCG A GCC A TCG	AAG ATG M CTC L TTC F GTG	ATC AGC S AAG K CAG Q AAG	CAT H GGC GTG V TCG	TAC Y ATC CAG GCC	GCG GCG A GGC G GTG GTG V GGC	CAC H TTC F AAC N AAC	ATC I ACC T ACG T TGC	GAC GAC D GCC A GCC GCC GCC GCC	CAG Q TTC F TCG AGC	GAC D GAC D ATC I CTC	1242 1290 1350 1404 1458		
311 S A E T C F A T Y L T D Q Y I R W L AAG TCC TC G G GA CCG GG GAC AAC TG GG GG GAC ATC TA AAG AAG GAC CCG ACG GG GAC AAC AAG GAC CCG AAC AAG GG AAC TA AA AAC TA AAC TA CA D D P D N C W AAC GG AAC GG GAC AAC GG GAC AAC GG AAC CCG GAC AAC GCG AAC GCG TCC GAC GG AAC GCG AAC TCC GCC TCC AAC GCG AAC ACG GG AAC ACG GG AAC	221 239 257 275 293	AAG GAC DAAG KCGC GGC GGC GGC	AAC ACG CTG L TAC Y AGC S	AAG CAG GTC V AAC N ATC T	CGC CAG CCG P CTG CTG TCG S	GAC GAC GTG V GTG V GCC A	AAG CTG GGC GGC G TTC F	TAC TAC AAC N CTG L TCC S	GTG GTG V GCG A GCC A TCG S	AAG ATG M CTC L TTC F GTG V	ATC AGC S AAG K CAG Q AAG K	CAT H GGC GTG V TCG S	GGC TAC Y ATC I CAG GCC A	GCG A GCG GGC GGC GTG GGC GGC GGC	CAC H TTC F AAC N AAC	ATC I ACC T ACG T TGC C	GAC GAC D GCC A GCC GCC GCC GCC GCC	CAG Q TTC F TCG S AGC S	GAC D GAC D ATC I CTC L	1242 1290 1350 1404 1458		
AAG TCC TCC AGC CTG GGC GAC GAT CCG GAC AAC TGC TGG CGC GCG AGC ATG GGG 1566 329 K S S L G D D P D N C W R A S M A 347 L D I Y K K D P T M G S V S V N Q GTC ATC AAC GCG AGC TAT CCC GAC CG ACG AGG GGA GAGC GTC AGC GCG CGA CG CGG CGG CGG CGG CGG CGG	221 239 257 275 293	AAG GAC D AAG K CGC R GGC GGC AGC	AAC ACG CTG L TAC Y AGC S GCC	AAG CAG GTC V AAC N ATC I GAG	CGC CAG CCG CCG CCG CCG CTG CTG CTG CTG S ACC	GAC GAC GTG V GTG V GTG C A TGC	AAG CTG GGC G TTC F TTC	TAC TAC AAC N CTG L TCC S GCC	GTG GTG V GCG A GCC A TCG S ACC	AAG ATG M CTC L TTC F GTG V TAC	ATC AGC S AAG K CAG Q AAG K CTG	CAT H GGC GTG V TCG ACG	TAC TAC Y ATC CAG GCC A GAC	GCG GCG A GGC GTG GTG GGC GGC GGC GGC	CAC H TTC F AAC N AAC N TAC	ATC I ACC T ACC T TGC C ATC	GAC GAC D GCC A GCC GC GC GC GC GC CGC	CAG Q TTC F TCG S AGC S TGG	GAC GAC D GAC D ATC I CTC L CTG	1242 1290 1350 1404 1458 1512		
329 K S S S L G D D D N C W R A S M A 329 K S S S L G D D D D W R A S M A 347 L D I Y K K D P T M G S V S V V N Q GTC ATC AAC GCG GAC TAT CCC GCC AAC AGG GGA AAC GCG GAC ACG GGG AAA TG TG GAC AGG AAC AGG CGC CCG	221 239 257 275 293 311	AAG GAC DAAG CGC CGC CGC CGC CGC CGC CGC CGC CGC	AAC ACG CTG L TAC Y AGC S GCC A	AAG CAG GTC V AAC N ATC I GAG E	CGC CAG CCG CCG CCG CCG CTG CTG CTG CTG ACC	GCC GAC D GTG V GTG V GTG C A TGC C	AAG CTG CTG GGC G TTC F TTC F	CTG TAC AAC N CTG L TCC S GCC A	GTG V GCG A GCC A TCG S ACC T	AAG ATG M CTC L TTC F GTG V TAC Y	ATC AGC S AAG K CAG Q AAG K CTG L	CAT H GGC GTG V TCG ACG T	TAC Y ATC CAG GCC A GAC D	GCG A GCC GCC GCC GCC GCC GCC GCC CAG	CAC H TTC F AAC N AAC N TAC	ATC I ACC T ACG T TGC C ATC T	GAC D GCC A GCC GCC GCC GCC CCC R	CAG Q TTC F TCG S AGC S TGG W	GAC GAC D ATC I CTC L CTG L	1242 1290 1350 1404 1458 1512		
CTG GAC ATC TAC AAG AAG GAC CCG ACG ATG GGC AGC GTC GAC GTG GTG GTG AAT CAG 1620 347 L D I Y K K D P T M GSC AGC GTC AGC GTG ATA TCAG 1620 347 L D I Y K K D P T M GSC ACG CGT GTG GTG AAT CAG 1620 367 L D I Y K K D P T M GS V V N Q 365 V I N A S Y P G N S K C P T S G I 1728 AAG TGG CGA AAG AAC ATG TCG TGC CGC CGA CGA GCA CGA CGA CGC GCG CGC GCG CGC CGC	221 239 257 275 293 311	AAG GAC AAG GGC GGC GGC GGC AGC AAG	AAC ACG CTG L TAC Y AGC S GCC A TCC	AAG CAG GTC V AAC N ATC I GAG E TCC	CGC CAG CAG CCG P CTG TCG ACC T CG ACC	GCG GAC GTG GTG GTG GTG GCC A CCG CTG	AAG CTG GGC TTC F TTC GGC	TAC TAC AAC N CTG L TCC S GCC A GAC	GTG GTG V GCG A GCC A TCG S ACC T GAT	AAG ATG M CTC L TTC F GTG V TAC Y CCG	ATC AGC S AAG K CAG Q AAG K CTG L GAC	CAT H GGC GTG TCG ACG T AAC	TAC TAC Y ATC CAG GCC GCC A GAC D TGC	GCG GCG A GGC GGC GGC GGC GGC CAG CAG TGG	CAC H TTC F AAC N AAC N TAC Y CGC	ATC ATC ACC T ACC T TGC C ATC I GCG	GAC GAC D GCC A GCC GCC GCC GCC CGC R AGC	CAG Q TTC F TCG S AGC S TGG W ATG	GAC GAC D GAC D ATC I CTC L CTG L GCG	1242 1290 1350 1404 1458 1512		
347 L D I Y K K D P T M G S V S V N Q GTC ATC AAC GCG GAC TAT CC GGC AAC ACG GCG AAC ATG TCG TGG TGG TGG CGC GCC G	221 239 257 275 293 311 329	AAG GAC AAG CGC R GGC GGC AGC S AAG K	AAC ACG CTG L TAC Y AGC S GCC A TCC S	AAG CAG GTC V AAC N ATC I GAG E TCC S	CGC CAG CCG CCG CCG CCG CCG CCG CCG CCG	GAC GAC GTG GTG V GTG CC A CC CTG L	AAG AAG CTG GGC G TTC F TTC F GGC	TAC TAC AAC N CTG L TCC S GCC A GAC	GTG GTG V GCG A GCC A TCG S ACC T GAT	AAG ATG M CTC L TTC F GTG V TAC Y CCG	ATC AGC S AAG K CAG Q AAG K CTG L GAC D	CAT H GGC GTG TCG ACG T AAC	TAC TAC Y ATC CAG GCC A GAC D TGC C	GCG GCG A GGC GGC GGC GGC GGC CAG CAG TGG	CAC H TTC F AAC N AAC N TAC Y CGC R	ATC ATC I ACC T ACG T TGC C ATC I GCG	GAC GAC D GCC A GCC GC GC GC CGC R AGC S	CAG Q TTC F TCG S AGC S TGG W ATG	GAC GAC D GAC D ATC L CTC CTG GCG A	1242 1290 1350 1404 1458 1512 1566		
GTC ATC AAC GCG AGC TAT CCC GGC AAC AGC GG AAA TGC CCG ACG TCA GG ATC 1674 365 V I N A S Y P G N S G K C P T S G I AAG TGG TCG AAG AAC ATG TCG TGG CAG TGG GCA GCG ACG GCG CGC CGC GCG CTG 1728 383 K W S K N M S W Q * 392 ACG CGG CAC GCC CCT GAT CCC TCC CAC CGA TGC ATT CC CAC CGG GGG CGC CGC 1782 TCG TGC GCG GTG CCC CGC CTC CGC TCT CGG CTC GGA CCT GGC CCC CGC CGC 11782 CCC TCC CCC CGC CAT GAA GAC CTT CGC TCC CGA CCT GGC CCC CGG CGC CCG 1940 CTC CGT CCC CGC CAT GAA GAC CTT CGC TCC CGA CCT GGC CCC CGG CGC CGC 1944 GCT GCC GTC CGC CGC GCC GCC CGC GCC GGC GG	221 239 257 275 293 311 329	AAG GAC AAG CGC C GGC GGC AGC AAG AAG CTG	AAC ACG CTG TAC Y AGC S CCC A CCC S C CCC S C CCC S C CCC S C CCC S C CCC S C CCC S C	AAG CAG GTC GTC V AAC N ATC I GAG E TCC S ATC	CGC CAG CCG CCG CCG CTG CTG ACC TCG ACC TAC	GAC GAC GTG GTG GTG GTG GTG CTG CTG CTG L AAG	AAG AAG CTG GGC GGC TTC F GGC GGC AAG	TAC TAC AAC CTG CTG CTG GCC GAC GAC	GTG GTG V GCG A GCC A GCC A TCG S ACC T GAT D CCG	AAG ATG M CTC L TTC F GTG V TAC Y CCG P ACG	ATC AGC S AAG K CAG Q AAG K CTG L GAC D ATG	CAT H GGC GTG TCG ACG TCG ACG T AAC N GGC	TAC Y ATC CAG GCC GCC A GAC TGC C AGC	GCG GCG GCG GGC GGC GGC GGC CAG GGC CAG GGC GGC	CAC H TTC F AAC N AAC N TAC Y CGC R	ATC I ACC T ACC T C ATC C ATC GCG A C C ATC	GAC GAC D GCC A GCC GCC GCC C GCC C C C C C C C C	CAG CAG TTC F TCG AGC S TGG ATG ATG	GAC GAC D GAC D ATC I CTC L GCG CAG	1242 1290 1350 1404 1458 1512 1566		
365 V I N A S Y F G N S G K C P T S G I AAG TGG TG AAG TGG TG TG G N S G K C P T S G I AAG TGG TGG TG CC TCC TCC TCC TGC TGC CGC CCC GC TRC TG GG CC GG CC TCC TGC CGC CCC GC TCC GG CCT GC TCC TCC GG CCT GC TCC TCC TCC GG CCT GC GC <td>221 239 257 275 293 311 329 347</td> <td>AAG GAC AAG CGC CC CGC CC CGC CC CGC AGC A</td> <td>AAC ACG CTG L TAC Y AGC S GCC A TCC S GAC</td> <td>AAG CAG GTC V AAC N ATC I GAG E TCC S ATC</td> <td>CGC CAG CCG CCG CCG CTG CTG ACC TCG ACC TAC</td> <td>GAC GAC GTG V GTG V GTG C C C C C C C C C C C C C C C C C C</td> <td>AAG AAG CTG GGC GGC TTC F TTC GGC GGC AAG</td> <td>TAC TAC AAC CTG CTG CTG GCC GAC GAC GAC</td> <td>GTG GTG V GCG A GCC A GCC A TCG S ACC T GAT CCG P</td> <td>AAG ATG M CTC L TTC F GTG V TAC Y CCG P ACG</td> <td>ATC AGC S AAG CAG Q AAG K CTG L GAC D ATG M</td> <td>CAT H GGC GTG GTG TCG ACG T AAC GGC GGC</td> <td>TAC Y ATC CAG GCC A GAC D TGC C AGC</td> <td>GCG A GCG GCG GCG GCG GCG GCG GCG GCG GTC V GTC</td> <td>CAC H TTC F AAC N AAC N TAC Y CGC R AGC</td> <td>ATC I ACC T ACG T GCG ATC GCG ATC GCG A GTG</td> <td>GAC GAC D GCC A GGC GGC GGC CGC R GGC CGC R GGC GGC CGC R GGC CGC C</td> <td>CAG Q TTC F TCG AGC S TGG W ATG AAT</td> <td>GAC GAC D GAC D ATC I CTC L GCG CTG CTG CTG CTG CTG CTG CTG CTG CTG</td> <td>1242 1290 1350 1404 1458 1512 1566 1620</td>	221 239 257 275 293 311 329 347	AAG GAC AAG CGC CC CGC CC CGC CC CGC AGC A	AAC ACG CTG L TAC Y AGC S GCC A TCC S GAC	AAG CAG GTC V AAC N ATC I GAG E TCC S ATC	CGC CAG CCG CCG CCG CTG CTG ACC TCG ACC TAC	GAC GAC GTG V GTG V GTG C C C C C C C C C C C C C C C C C C	AAG AAG CTG GGC GGC TTC F TTC GGC GGC AAG	TAC TAC AAC CTG CTG CTG GCC GAC GAC GAC	GTG GTG V GCG A GCC A GCC A TCG S ACC T GAT CCG P	AAG ATG M CTC L TTC F GTG V TAC Y CCG P ACG	ATC AGC S AAG CAG Q AAG K CTG L GAC D ATG M	CAT H GGC GTG GTG TCG ACG T AAC GGC GGC	TAC Y ATC CAG GCC A GAC D TGC C AGC	GCG A GCG GCG GCG GCG GCG GCG GCG GCG GTC V GTC	CAC H TTC F AAC N AAC N TAC Y CGC R AGC	ATC I ACC T ACG T GCG ATC GCG ATC GCG A GTG	GAC GAC D GCC A GGC GGC GGC CGC R GGC CGC R GGC GGC CGC R GGC CGC C	CAG Q TTC F TCG AGC S TGG W ATG AAT	GAC GAC D GAC D ATC I CTC L GCG CTG CTG CTG CTG CTG CTG CTG CTG CTG	1242 1290 1350 1404 1458 1512 1566 1620		
ANG TGG TCG ANG ANC ATG TCG TGG CAG TGA GCC GCG ACG GCG CGC GCG CGC TCG 1728 383 K W S K N M S W Q * 392 ACG CGG CAC GCC CCT GAT CCC TCC CAC CGA TGC CAT CGA CCC GCG GGG CCC CGT 1782 TCG TGC GCG GCG CCC GAT CCC TCC CGC CGC CGA CGA CCC CGA CCC GCG GGG CCC CGT 1782 TCG TGC GCG GTG CCC CGC CTC CGG CTT GGG ATT TCT CCC TGA CCC GGT CAC CGA 1836 CGC CCT CCC CGC CAT GAA GAC CTT CGC TCC CGA CCT GGC TCG CCA CGG CCG CGG 1890 CTC CGT CCC GCC CAT GAA GAC CTT CGC TCC CGA CCT GGC TCG CCA CGG CCG 1844 GCT GCC GTC CGC CGC GCC GGC CGC CGC GGC G	221 239 257 275 293 311 329 347	AAG GAC AAG CGC CC CGC CGC CGC CGC CGC C	AAC ACG CTG L TAC Y AGC S CCC A TCC S CCC A TCC S CCC	AAG CAG GTC V AAC N ATC GAG E TCC S ATC S ATC	CGC CAG CCG P CTG L TCG ACC TAG CCG TCG ACC TAG CCG	GAC GAC GTG GTG GTG GTG GTG CC CC L AAK AC	AAG AAG CTG GGC GGC TTC F GGC AAG AAK	TAC TAC AAC N CTG L TCC S GCC A GAC D GAC C C C	GTG GTG V GCG A GCC A GCC A TCG S A CC T GAT CCG GAT CCG GCC	AAG ATG M CTC L TTC F GTG V CCG P ACG ACG	ATC AGC S AAG K CAG Q AAG K CTG L GAC D ATG M AGC	CAT H GGC GTG TCG ACG AAN GGC GCG	TAC Y ATC I CAG GCC AC C C AGC AGC SAA	GCG A GCG GCG GCG GCG GCG GCG GCG GCG GC	CAC H TTC F AAC N AAC N CGC R AGC SCC	ATC I ACC T ACG T C ACG C ATC GCG ACG GCG V C C	GAC GAC GCC GCC GCC GCC GCC CCC AC GCC CCC C	CAG Q TTC F TCG S AGC S TGG W ATG AAT N	GAC GAC D GAC D ATC CTC CTC CTG GCG A CAG QC	1242 1290 1350 1404 1458 1512 1566 1620		
383 K W S 105 105 KM KM KM 105 105 105 005 000 104 000 000 000 000 000 000 000 000	221 239 257 275 293 311 329 347 265	AAG GAC AAG CGC CGC CGC CGC CGC CC CC CC CC CC CC	AAC ACG CTG L TAC Y AGC S GCC A TCC S GAC D ATC	AAG CAG GTC V AAC N ATC I GAG E TCC S ATC I AAC	CGC CAG CCG P CTG CTG CTG ACC TCG ACC TAC CCG CTG CTG CTG CTG CTG CTG CTG CTG CCG CTG CT	GAC GAC GTG V GTG CC A GTG CC CTG L AAG K AGC	AAG CTG CTG GGC GGC TTC F GGC AAG TAT	AAC TAC TAC TAC TAC TAC TAC CTG CTG CTG CTG CCC GCC A GAC CCC GAC CCC CCC	GTG GTG V GCG A GCC A GCC A TCG A CCG GCC C GCC GCC GCC GCC GCC GCC G	AAG ATG M CTC L TTC F GTG V TAC Y CCG P ACG TAC N	AGC S AAG K CAG Q AAG K CTG L GAC D ATG AGC	CAT H GGC GTG TCG ACG AAC GGC GGG GGG	TAC Y ATC I CAG GCC A GCC A GCC A GCC A GCC A GC S AA	GCG GCG GGC GGC GGC GGC GGC GGC GGC GGC	CAC H TTC F AAC N AAC N CGC R AGC CGC CG CCG	ATC I ACC T ACG T C ACG C ATC G C G C G C G C G C G C G C G C G C G	GAC GAC GCC GCC GCC GCC GCC CGC CGC CGC	CAG Q TTC F TCG S AGC S G G G G G G G G	GAC GAC D GAC C C C C C C C C C C C C C C C C C C	1242 1290 1350 1404 1458 1512 1566 1620 1674		
ACG CGC GCA CGC CT GAT CCC TCC CAC CGA TGC CAT CGA CCC GCG GGG CGC CGT 1782 TCG TGC GCG GTG CCC CGC TTC CGA CTC CGA TGC CAT CGA CCC GCG GGG CGC CGT 1782 CGC CCT CCC CGC CGC CGC CTC CGG CTT GTG ATT TCT CCC CCT GAC CCC GGT CAC CGA 1836 CGC CCT CCC CGC CAT GAA GAC CTT CGC TCC CGA CCC TGG CCC GGC CGG 1890 CTC CGT CCT GAC GAT GAC CGC CGC CGC GGC GAT GGC CGC GGG CGC CGG CGC 1944 GCT GCC GTC CGC CGC GGC GCT GGC CGC GCC GGC GCT GGC CGC GCG CGC 1944 GCT GCC GTC CGC CGC GGC GCT GGC CGC GGC G	221 239 257 275 293 311 329 347 365	AAG GAC AAG CR GGC GGC AGC AGC CTG GC CTG GTC AC	AAC ACG CTG L TAC SGCC A CTG GAC CTG CTG CTG CTG CTG CTG CTG CTG CTG CT	AAG CAG GTC V AAC N ATC I GAG E TCC S ATC I AAC N TCC N	CGC CAG CCG CCG CTG CTG CTG ACC TCG ACC TAC CCG ACC	GAC GAC GTG V GTG CC CTG L AAG AGC AAG AGC	AAG CTG GGC TTC F GGC GGC GGC AAG K TAY	TAC TAC TAC TAC TAC TAC TAC TAC TAC TAC	GTG GTG V GCG A GCC A CCG A CCG GAT CCG GGC GGC GCC	AAG ATG ATG CTC L TTC F GTG V TAC Y CCG P ACG T AAC	ATC AGC S AAG CAG Q AAG K CTG L GAC D ATG AGC S CAG S CAG S CTG CAG S CTG C CTG C CTG C CTG C CTG C CTG C C C C	CAT GGC GTG GTG CAT GGC GTG ACG GGG GGG GGG GGG GGG GGG GGG	TAC Y ATC CAG GCC AC CAG CAC CAC CAC AGC AAA CC	GCG GCG GCG GGC GGC GGC GGC GGC GGC GGC	CAC H TTC F AAC N AAC N TAC CGC R CGC CC CC CC	ATC I ACC T ACC T C ATC C ATC GCG ATC I GCG ATC I GCG ACC C C C C C C C C C C C C C C C C	GAC GAC GCC GCC GCC GCC GCC ACC SCC ACC SCC CCC ACC SCC CCC ACC CCC C	CAG CAG TTC F CG S CAG TCG S CG W ATG AAT N GGG GGG	GAC GAC D GAC D CTC L CTC L GCG CAG Q ATC I CTC	1242 1290 1350 1404 1458 1512 1566 1620 1674		
RCG CGC CGC GCC CGC CGC CGC CGC CGC CGC	221 239 257 275 293 311 329 347 365	AAG GAC D AAG K CGC R GGC AGC AAG CTG GTC V AAG	AAC ACG ACG CTG L TAC Y AGC S GCC A TCC S GAC D ATC I TGG GAC U TGG	AAG CAG GTC V AAC I GAG E TCC S ATC I AAC N TCC	CGC CAG CAG CCG P CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG	GAC GAC GTG GTG GTG GTG GTG CTG CTG CTG L AAG AAG AAG AAG	GCG AAG CTG GGC G G TTC F C TC F C G G C G G C T TC F C T T C F C T T C T C T G C C G C G C G C G C G C	TAC TAC TAC TAC TAC TCC SCC A CTG CCC CCC P CCC CCC CCC CCC CCC	ACC GTG GTG V GCG A GCC A GCC A TCG GAT D CCG G G G G G G G G G G G G G G	AAG AAG ATG CTC L TTC F GTG V CCG P ACG ACG AAC N CAC	ATC AGC S AAG CAG Q AAG K CAG Q AAG CAG CAG D AAG CAG CAG CAG CAG CAG CAG CAG CAG CAG	CAT H GGC GTG GTG TCG S ACG T CAC N GGC GGG GCC	TAC Y ATC CAG GCC A GAC C C AGC S AAA K GCG	GCG GCG GCG GGC GGC GGC GGC GGC GGC GGC	CAC H TTC F AAC N AAC N CGC R AGC S CCG GCG	ATC I ACC T ACG T TCC C ATC I GCG A GTG V ACG T CCC	GAC D GCC A GCC G GCC G GCC G GCC C GC C G	CAG Q TTC F TCG S AGC S G G G G G G G C G C G C G C C AG C AG	GAC GAC D GAC D ATC L CTC L GCG CTG CTG CTG CTG CTG	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728		
CGT CCC CCC CGC CGC CGC CGC CGC CGC CGC	221 239 257 275 293 311 329 347 365 383	AAG GAC DAAG GGC GGC GGC CTG GGC CTG CTG CTG CTG CT	AAC ACG ACG CTG CTG TAC Y AGC S CCG S CCC A TCC S S CCC D ATC I TGG W CCG CCG CCG CCG CCG CCG CCG CCG CCG	AAG CAG GTC V AAC N ATC I GAG E TCC S ATC I AAC N TCG S CAC	CGC CAG CAG CCG P CTG L TCG S ACC T AGC S TAC Y GCG A AAG K	GCG GAC GTG GTG GTG C GTG C C C C C C C C C C	GCG AAG CTG GGC G G G G G CTC F C TTC F C GGC G G AAG K TAT Y ATG M M	TAC TAC N CTG L CTG CTG CTG CTG CCC CCC CCC CCC CCC CCC	ACC GTG GTG V GCG A GCC A TCG S ACC TCG GAT D CCG G G TGG G G TGG G G TGG G G T G G C G T G C G T G C G C	AAG ATG ATG M CTC L TTC F GTG GTG V CCG P ACG T AAC N CAG Q Q CAC	AGC AGC S AAG CAG Q AAG K CAG Q AAG CTG AAG S TGA * CAG S CAG CCA	CAT H GGC G G G G G G G G G G G G G G G G G	TAC Y ATC CAG Q CCA GCC C AGC C C AGC S AAAA K GCG	GCG GCG GCG GCG GGC GGC GGC GGC GGC GGC	CAC H TTC F AAC N AAC N CGC R AGC S CCG P GCG	ATC I ACC T ACG T T GCG A GCG A GTG V ACG T C CC C C C C C C C C C C C C C C C	GAC D GCC A GCC G GCC G GCC G GCC C R A GCC S GTG V TCA S CGC	CAG Q TTCC F TCG S AGC S TGG W AATG M AAT N GGG G CG	GAC GAC D ATC CTC CTC CTG CTG CTG CTG CTG CTG CTG	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728		
CGC CGT CCC GAC GGC GGC GCC GGC GGC GGC GGC GGC	221 239 257 275 293 311 329 347 365 383	AAG GAC GAC CGC R GGC GGC CTC GTC V AAG K CTC GTC V AAG K CTC CTC C C C C C C C C C C C C C C C	AAC ACG ACG CTG CTG CTG CTG CTG S CCC A ACC S CCC S GAC C C C C C C C C C C C C C C C C C C	AAG CAG GTC V AAC N ATC I GAG E CAC S ATC I CAC S CAC S CAC	CGC CAG CAG CCG P CTG CTG S ACC TCG S ACC TAC S CCG CGC A AGC CCG CCG CCG CCG CCG CCG C	GCG GAC GTG GTG V GTG V GTG C C C C C C C C C C C C C C C C C C	GCG GCG AAG CTG GGC CTG F TTC F GGC C F GGC G AAG K TAT Y ATG M AG	TAC TAC TAC TAC TAC TAC TCG SCC CA SAC D CCC TCG SCC CCC TCG SCC CCC	ACC GTG GTG V GCG A GCC A GCC A CCG S ACC T GAT D CCG G GCC G GCC T GGC T C CCG C C C C C	AAG ATG ATG M CTC L TTC F GTG V TAC Y CCG P ACG T AAC N CAG Q CAC	ATC AGC S AAG K CAG Q AAG CTG AAG CTG ATG AGC S TGA S CGA CGA	GCCAT H GGC GTG V TCG GGC GGC GGC GGC GGC GGC GGC GGC GGC	TAC Y ATC CAG Q GCC A GAC C AGC S AAAA K GCG CAT	GCG GCG GCG GCG GC GCC GCC GCC GCC GCC	CAC CAC H TTC F AAC N AAC N TAC Y CGC R AGC CCG CCG CCC	ATC I ACC T ACG T C C ATC C C ATC GCG GCG GCG GCG GCG GCG	GAC D GCC A GCC G GCC G GCC G G G G G G G G G	CAG Q TTC F TCG S CAG S TGG W ATG M ATG M ATG G G G CGC CGC CGC CGC	GAC GAC D GAC C C C C C C C C C C C C C C C C C C	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782		
GCT CCT CGC CGC CGC CGC CGC CGC CGC CGC	221 239 257 275 293 311 329 347 365 383	AAG GAC DAAG CGC GGC GGC CCR GGC CCR GGC CCR GGC CCR CTG CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCR CCC CCC CCR CCCC	AAC ACG CTG CTG L TAC Y AGC S CTG S CCTG CTG S CCTG CTG CTG CTG CTG CTG CTG CTG CTG CT	AAG CAG GTC V AAC I GAG E TCC S ATC I AAC N TCG S CAC GCG	CGC CAG CAG CCG P CTG CCG CCG P CTG CCG S A CCG S CCG CCG CCG CCG CCG CCG CCG CCG C	GCG GAC GTG GTG V GTG V GCC CTG CTG CTG CTG S AAC N CCT CCC CCC	GCG GCG AAG CTG GGC CTG F TTC F GGC G AAG K TAT Y ATG M GAT CGC	TAC TAC TAC TAC TAC TAC TCG SCCC TCG SCCC TCG SCCC TCG SCCC TCG	GCC GTG V GCG A GCC A GCC A CCG S ACC T GAT D CCG G GCC G GCC G CCG C CCG C CCG C CCG C C CCG C	AAG ATG ATG M CTC L TTC F GTG V TAC T CCG P ACG T AACG N CAG Q CAC CTT	ATC AGC S AAG K CAG Q AAG CTG CTG AAG S TGA * CGA GTG	CAT H GGC G GTG S ACG GGG GGG GGG GGG GGG GGC 392 TGC ATT	TAC Y ATC CAG GCC C A GCC C A GCC C A AAA K GCG C AT TCT	GCG GCG A GGC G GGC G GGC G G CAG G CAG V TGC C ACG CGA	CAC CAC H TTC F AAC N AAC N TAC Y CGC R AGC S CCG P GCG CCC TGA	ATC I ACC T ACC T T C C ATC C ATC C ATC C C ATC C C ATC C C C	GAC D GCC A GCC G GCC G GCC G GCC G GCC G GCC C GCC C GCC C GCC C GCC G G C	CAG Q TTC F TCG S CAG S TGG W ATG M AAT N GGG GCG CAC CAC CAC	GAC D GAC D ATC I CTC CTG CTG CTG CTG CTG CTG CCGT CCGT	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836		
GCT GCC GTT GCC CGC GGT GGC CGC GGT GGC GGC	221 239 257 275 293 311 329 347 365 383	AAG GAC DAAG GGC R GGC R GGC R GGC R CTG GTC C AAG CTG GTC C CGC C CGCC C CGC C CGCC	AAC ACG CTG TAC S GCC A CTG S GCC A TCC S GAC C S GAC C C G GAC C C G GCC C C C C C C C C	AAG CAG GTC V AAC I GAG E TCC S ATC I AAC S CAC GCG CCC	CGC CAG CCG P CTG CCG CCG CCG CCG CCG CCG CCG CCG CCG	GCG GAC GTG GTG V GTG V GTG C C C GTG C C C C C C C C C C C C C	AAG AAG CTG GGC TTC F TTC GGC G AAG K TAT Y GAT CGC GAA	TAC TAC TAC TAC TAC TCC S CCC C C C C C C C C C C C C C C	ACC GTG GTG V GCG A GCC A TCG S ACC T TCG G G G G G C C G G C G C G C G C	AAG ATG ATG TG CTC F GTG GTG F GTG V TAC F ACG T AAC C ACG C ACG C CAC CTT CGCC	ATC AGC S AAG K CAG Q AAG K CTG CTG ATG AAG S TGA * CGA GTG TCC	CAT H GGC G G G G G G G G G G G G G G G G G	TAC Y ATC I CAG Q GCC A GAC C A GCC C A GCC C A GCC C C A GCC C C A C C C C	GCG GCG GGC GGC GGC GGC GGC GGC GGC CAG GGC CCC GGC CCC GGC CCC GGC	CAC CAC H TTC F AAC N AAC N TAC V CGC R AGC S CCG GCG CCC TGA	ATCC I ACCC T ACCC T TCCC C ATCC C ATCC C ATCC C C ATCC C C C	GAC D GCC A GCC G GCC G GCC G GCC G GCC C GCC S C GCC S C C C C	CAG Q TTC F G S AGC S TGG W ATG G G G G G G C G C C C C C C C C C C C	GAC GAC D ATC CTC CTC CTC CTC CCC CCC CCC CCC CCC	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836 1890		
CGT CCA GTT GCC CAC GGA CCC GCT GCT GGT GCA GAT CAA CCA GGT CGC GCT GGA 2052 GCG CGT CGG TCC CAA GCG GGC GGC GGT CGT CGT CGA GTC CAC CGG CGC GGC CGG GAG CGG 2106 CCG CCT ATC GCG TGC TGT CCG GCG GCA AGA CCA GTC GCG CGC GGC CGG GAG CGG 2106 CGC TGC CCG CGT TCA GCG AAT GGG GCG CGG GCA AGC GCC AGG GCA CGC TCA ACC 2214 TCT CGT CGC CGG TGA AAG CCG GCG GCA ATA ACC GCG TCA AGG CGA CGC TCA ACC 2268 GGC GCG GGG TCT CCG CGC GCG GCG ATA ACC GCG CGC GCG CGC TGT CCA CGA CCA 2322 TCG CCG ACA AGA CCG TGG ACT ATT TCC ATC GCG GCC GCC ACA CCG GCG GCG 2375 ATC GCA GCA TCC GCA TCT TCG AGA CCG GCA AGC GCC GCG TCG ACC GCG GCG 2430 GGA AGG ACG CGG GCG GCG AGA CCG GCA AGT ACC TCT CGC ACC TGG GCT ACG CCA 2484 ACT ACT TCA GCC CGC AGA AGA CCT CGA TGG TCG CCT GGG TGC TCA CGC C2358 GCA GAC CTC GCC GC 3'	221 239 257 275 293 311 329 347 365 383	AAG GAC CGC CGC CGC CGC CGC CTCC CGC CGC CGC C	AAC ACG CTG TACC Y AGC S GCC A TCC S GAC D ATC I TGG CGG TGC CCT CGT	AAG CAG GTC V AACC I GAG E C S ATC S ATC S ATC S CAC GCG CCC CCT	CGC CAG CCAG CCG P CTG CTG CTG CTG CTG CTG CCG A AAG CCG CGC CGC CGC CGC CGC CGC	GTG GTG GTG V GTG V GCC A CTG C CTG C CTG L AAG C CTG C CTG C C CTG C C C C C C C C C	AAG AAG CTG GGC G TTC F GGC G AAG K TAT Y ATG GAT CGC GAA GAC	TAC TAC TAC TAC TAC TAC TAC TAC TCC S CCC A C CCC CCC TCC S CCC CCC CCC CCC CCC CCC	ACC GTG GTG GCG A GCC A C C A C C G A T C G A T C G A T C G G C T G G C G C G C G C G C G C G	AAG ATG ATG M CTC F GTG GTG V TAC Y CCG P ACG P ACG T AAC N CAG Q CAC CTT CGC CGC CGC	ATC AGC S AAG K CAG Q AAG K CTG GAC CTG ATG AAG S TGA S GGC GGC GGC	CAT H GGC G G G G G G G G G G G G G G G G G	TAC Y ATC I CAG Q GCC A GAC D C C A GCC C A GCC C A GCC C C A C C C C	GCG GCG GGC GGC GGC GGC CAG Q CAG Q CAG CCC C C ACG CCC CGC CGC CGC	CAC CAC H TTC F AAC N AAC N TAC R AAC R AGC S CCG GCT TCG GCT	ATC I ACC T ACC T T C C ATC T C C ATC I G C G A T C C C ATC C C ATC C C ATC C C ATC C C ATC C C ATC C C C	GAC D GCC A GGC G G G C G C G C G C G C G C G	CAG Q TTC F F CCG S AGC S TGG W ATG G G G G G G C G C C C C C C C C C C C	GAC GAC D GAC D CTC L CTC L GCG A GCG A CTC C C GCG C C GCG C C C G C C C C C C	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836 1890 1944		
GCG CGT CGG TCC CAA GCG GGC GCG GCG GGC CGG GAC CGG 2106 CCG CCT ATC GCG TGC TGT CCG GGG GGA GAG ACG TCG CGC GGC CGG GAA GCG 2160 CGC TGC CCG CGT TCA GCG AAT GGG GCG CGG GCA AGC GCT ATT TCC AGG TGG ACT 2144 TCT CGT CGC TGG ACA AGG CGG GGA ATT ACC GCG TCG AGG GCA CGC TCG AGG 2268 GGC GCG GGG TGT TCC CGC CGG GCG ATT ACC GCG CGC GCG CGC GCG CGC CGC GCG 2322 TCG CCG ACA AGA CGC TGG ACT ATT TCC AGG CGG CGC GCC GCC GCC GCG CGC GCG CGC 2322 TCG CCG ACA AGA CGC TGG ACT ATT TCC AGC GGG GCG CGC CGC ACA CCG GCG GCG CGC 2375 ATC GCA GCA TCC GCA TCT TCG AGA CCG ACC GCC GCC GCC ACA CCG GCG G	221 239 257 275 293 311 329 347 365 383	AAG GAC D AAG CGC CGC CCG CCG CCG CCG CCG CCC CCC	AAC ACG CTG L L TAC Y AGC S CCT S CCT TGC CCT CCT CCT CCT CCT	AAG CAG GTC V AAC I GAG E TCC S ATC I CAC S CAC C CCC CCC CCC CCC	CGC CAG CCGG P CTG CCG S CCG T TCG S CCG T TCG S CCG C CGC GTG CGC CGC CGC CGC CGC	GCG GAC GTG V GCC A GCC C C C C C C C C C C C C C C C	AGC AAG CTG GGC G G TTC F C G G C TTC F C G G C T T C F C T T C T C T T C T T C T T C T C	TAC N CTG CTG CTG CTG CTG CCG CCC CCC CCC CCC	GTG GTG CC GCG A GCC A TCG S CCG A TCG GGT TCG GGC TCG GGC CCGC CCGC	AAG ATG ATG CTC L TTC F GTG V TAC CCG P ACG T CCG C CCG CCT CGC CGC CGC	ATC AGC S AAG K CAG Q AAG K CTG AAG GAC D ATG M CCG S TGA * CGA GTG TCC GGC CGC	GCCAT H GGC G GTG G GTG S ACG T CGC S ACG C GGG G GCC 3922 TGC ATT CGCA GAT CGCA	TAC Y ATC CAG Q GCC A GAC C A GCC C A GCC C C C C C C C	GCG GCG GCG GCG GCA GCA GCA GCA GCA GTC CA CA CA CCC CCC CCC CCC CCC CCC CCCC CCCC	CAC CAC H TTC F AAC N AAC N AAC N CGC R AGC S CCG CCC TCG GCT TCG GCT TCG GCT	ATCC I ACC T ACC T T C C C T T C C C I T C C C I T C C C I T C C C C	GAC GAC D GCC A GGC G GGC G GGC CGC CGC CGC CGC C	CAG Q TTC F TCG S AGC S TGG W ATG M AAT N GGG G CGC CCG CCG CCG CCG CCG	GAC GAC D ATC L CTC L CTC C C C C C C C C C C C C C	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836 1890 1944		
CCG CCT ATC GCG TGC TCT CCG GCG GCA AGA CCG TCG CGC AGG GCA CGC TCA AGC 2160 CGC CGC CGC CGT CA GCA GCA GCA GCG GCA AGC GCA AGC GCA CGC TCA AGC 2214 TCT CGT CGC TGG ACA AGG CCG GCG ATT ACC GCG TCG AGG GCA CGC TCG GCG GCA 2264 GGC GCG TGG TCT CCC CGC CGC GCG TGC ACC GCG CGC TGT TCA CGA CCA 2322 TCG CG ACA AGA CGC TGG ACT ATT TCC AGC GCG CGC GCC GCG TGT TCA CGA CCA 2322 TCG CG ACA AGA CGC TGG ACT ATT TCC ATC GCG GCC GCC CGC ACA CCG GCG GCG 2375 ATC GCA AGA TCC GCA TCT TCG AGA CCG ACC ACC GCG CGC TGC ACG GCG GCG CCC 2430 GGA AGG ACG CGG GCG AGA CCG GCA AGT ACC TCT CGC ACC TGG GCT ACG CCA 2484 ACT TCA GCC CGC CGC AGA CCT CGA TGG TCG CCT GGG TGC TGG CCT ACA CGC 2538 GCA GAC CTC GCC GC 3' 2552	221 239 257 275 293 311 329 347 365 383	AAG GAC Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q	AAC ACG CTG L L TAC Y AGC S GCC C A TCC S GAC I TGG W CGG CCT CCT GCC CCA	AAG CAG GTC V AAC N ATC I GAG E CC S ATC I CAC S CAC S CAC C CC TCG GCG CCCT GTC GTC	CGC CAG CCG P CTG CTG CTG CTG S ACC TTCG S ACC Y GCG A AAG K GCC GTG CGC CGC CGC CGC	GTG GTG V GTG V GCC C C C C C C C C C C C C C C C C	AAG GCC CTG GCC GCC GCC F F GGCC F F GGC K TAT Y ATG M GAT CGCC GAA GCT GCA	TAC TAC TAC TAC TAC TCC SCC CTC SCC CCC TCC SCC CTC CTC SCC CTC CT	ACC GTG GTG C G G G G G G G G G G G G G G	AAG ATG M CTC F GTG GTG GTG V CCG P ACG P ACG T CCC CGA CAC CTC CGC CGC CGC CGC	ATC AGC S AAG CAG Q AAG CTG CTG CTG CCTG CCTG CTG ATG AGC S TGA * CGA GTC CGC CGC CGC CGC	GCCAT H GGC GGG GTG V TCG S GGG GGG GGG GGG GGC 392 TGCC ATT CGA GGC GCC ATT CGCAT CGCAT CGCAT CGCAT CGCAT CGCAT CGCAT CGCAT CAT CAT CAT CAT CAT CAT CAT CAT CAT	TAC TAC Y ATC CAG GCC A GCC A GCC C A GCC C C C C C	GCG GCG GCG GCG GCC GCC GCC CCAG GCC CCAG GCC CCA CCA	CAC CAC H TTC F AAC N AAC N TAC C C C C C C C C C C C C C C C N C N C	ATCC I ACCG T ACCG T TCC C C T TCC C C T TCC C C T TCC C C T TCC C C C T TCC C C C C C C C C C C C C C C C C C C C	GAC D GCC G G G G G G G G G G G G G G G C G C C C C C C C R A S C G C C R C R C G C C C C C C C C C C C	CAG CAG Q TTC F TCG S AGC S TGG W ATG G G G G CGC CAC CCC CCC CCC CCC CCC GCT GCT	GAC GAC D CTC CTC CTC CTC CTC CTC CTC CCAG CCC CCC CCC CCC CCC CCC CCC CCC C	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836 1890 1944 1998 2052		
CGC TGC CCG CGT TCA GCG AAT GGG GCG CGG GCA AGC GCT ATT TCG AGG TGG ACT 2214 TCT CGT CGC TGG ACA AGG CCG GCG AGT ACC GCG AGG CGA CGC TCG GCG ACT 2214 TCT CGT CGC TGG ACA AGG CCG GCG GCG ATT ACC GCG AGG CGA GCG CCG TCG GCG CCG 2268 GGC GCG TGG TCT CCG CGC CGC TGC GCG GCG	221 239 257 275 293 311 329 347 365 383	AAG GAC D AAG CGC R CGC CGC CGC CGC CTC CCC CTC CCC CTC CCC CTC CCC CTC CCC CTC CCC CTC CCC CTC CCC CTC CCC CCC CCC CCC C CCC C C C C C C C C	AAC ACG CTG L TAC Y AGC S GAC V AGC S GAC C C TGC C C C T G C C G C C T G C C T G C T G C T G C T G C T G C T G C T G C T G C T G C T G C T G C T G C T G C T G C T G C T G C T G C T G C T G C C G C C C G C C C G C C C G C	AAG CAG GTCC V AAC N AACC S CAC GAG E TCC S ATC I AAC N TCG S CAC GCG CCC CCT GTT CCG GTT CCG CCC	CGC CAG CCG CCG CCG CCG CCG CCG CCG CCG	GCG GAC GTG GTG V GTG V C C G GTG V C C G G C C G C C G C C C C C C C C	AAG AAG CTG GGC G GGC F F GGC G AAG K TAT Y ATG GAA GAC GGA GCT GGA	TACC TACC AAAC CTG CTG CTG CTG CTG CTG CCC P CCC CTC CGAC CCC CCC CCC CCC CCC CCC CCC CCC C	ACC GTG GTG C G G G A G C C G A T C G G T C G G G T G G G C G G G G G G	AAG ATG M CTC L TTC F GTG GTG V CCG P ACG T AAC N CAG Q CAC CTT CGC CGC CGC CGC CGC	ATC AGC S AAG C CAG C CAG C CAG C CAG C CGA C CGA C CGA C CGA C CGA C CGC CG	CAT H GGC GTG GTG V TCG GGC GGC GGC GGC GGC GGC GGC GGC GGC	TAC Y ATC CAG GCC AC GCC C C C C C C C C C C C C	GCG GCG GCG GCC GCC GCC GCC GCC GCC CCA CCA	CAC CAC H TTC F AAC N AAC N AAC C R AGC S CCG F GCG CCC TGA TCG GCT CCA CGC CCA CGC	ATCC I ACCG T ACCG T C C C C C C C C C C C C C C C C C	GAC GAC D GCC A GGC G GGC G GGC G GGC C GC C	CAG CAG Q TTCC F TCG S AGC S AGC S TGG W ATG M AAT N GGG G GCG CCC CCG CCG CCG CCG CCG CCG	F F GAC D GAC D ATC L CTG L CTG CTG CTG CTG CCG CCG CCG CCG	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836 1890 1944 1998 2052 2106		
TCT CGT CGC TGG ACA AGG CCG GCG ATT ACC GCG TCG AGG CGA CGC TCG GCG GCA 2268 GGC GCG TGG TCT CCG CGC GCG GGC GGC GCG CGC GCG CGC TGT TCA CGA CCA 2322 TCG CCG ACA AGA CGC TGG ACT ATT TCC ATC CGG GCG CGC GCC TGT TCA CGA CCA 2322 TCG CCG ACA AGA CGC TGG ACT ATT TCC ATC CGC GCG CCC ACA CCG GCG GCG C2375 ATC GCA GCA TCC GCA TCT TCG AGA CCG ACC GCC GCC ACA CCG GCG CGC GCG CCC 2430 GGA AGG ACG CGG GCG GCG ACA CCG GCA AGT ACC TCT CGC ACC TGG GCT ACG CCA 2484 ACT ACT TCA GCC CGC AGA CCT CGA TGG TCG CCT GGG TGC TGG CCT ACA CGC 2538 GCA GAC CTC GCC CG 3' 2552	221 239 257 275 293 311 329 347 365 383	AAG GAC D AAG GC CC CC CC CC CC CC CC CC CC CC CC CC	AACG ACG TAC TAC S CTG S CTG S CCT S CCT S CCT CCT CCT CCT CCT CCT C	AAG CAG GTCC V AAC N ATC I GAG E TCC S ATCC S CAC CCC CCC CCC CCC CCC CCC CCC CCC	CGC CAG CCG CCG P CTG CCG CCG CCG CCG CCG CCG CCG CCG CCG	GCG GAC GTG GTG V GTG V GCC CTG CTG CTG CTG CTG AAG K AGC S AAC N CCT CCC CCC CCC CCC CCC CCC CCC CCC	AAG AAG CTG GGC TTC F GGC G AAG K TAT TATG AAG GAA GAC GCG GCG GCG GCG	TAC TAC X AAC N CTG CTG CC C C C C C C C C C C C C C C	ACC GTG GTG V GCG A GCC A GCC A GCC A CCG T G G T CCG G G G G G G G G G G G G	AAG ATG M CTC L TTC F GTG V TAC Y CCG P ACG T AACG T AACG CAG CCC CCA GCT CCGC CGCA	ATC AGC S AAG K CAG Q AAG K CTG L GAC CTG ATG AAG S TGA * CGAG GGC CGC CGC AAG	CAT H G G G G G G G G G G G G G G G G G G	TAC Y ATC CAG Q CCA GCC A GCC A GCC C A GCC C A GCC C A GCC C A GCC C C A GCC C C C	GCG GCG GCG GCG GCG GCC GCC GCC CAG CCC CCC	CAC CAC H TTC F AAC N TAC Y CGC R AGC S CCG C CCA C C C C C C C C C C C C C C C	ATC I ACC T ACC T C ATC C ATC C ATC C ATC C C ATC C C ATC C C ATC C C ATC C C ATC C C ATC C C ATC C C ATC C C C	GAC GAC D GCC G GCC G G G G G G G G G C G C G	CAG Q TTCC F TCG S AGC S TGG W ATG G G G G C G C C CCG C CCG C CCG C CCG C CCG C CCG C CCG C CCG C CCG C CCAG S C AG S S AG S S AG S S C AG S S C AG S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S S C S	GAC D GAC D CTC CTG CTG CTG CTG CCG CCG CCG CCG CCG	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836 1890 1944 1998 2052 2106 2160		
GGC GCG TGG TCT CCG CGC CGG TGC GGG TGC GCG CGC TGT TCA CCA CCA 2322 TCG CCG ACA AGA CGC TGG ACT ATT TCC ATC GCG GCC GCC GCC TGT TCA CCA CCG 2375 ATC GCA GCA TCC GGA TGG ACT ATT TCC ATC GCG GCC GCC ACA CCG GCG GCG CG 2375 GGA AGG ACG CGA CTC TTCG AGA CCG ACC GCC GCG TGC ACG GCG GCG CG 2430 GGA AGG ACG CGG GCG GCG ACA CCG GCA AGT ACC TCT CGC ACC TGG GCT ACG CCA 2484 ACT TCA GCC CGC AGA CCT CGA TGG TCG CCT GGG TGC TGG CCT ACA CGC 2538 GCA GAC CTC GCC GC 3' 2552	221 239 257 275 293 311 329 347 365 383	AAG GAC Q Q AAG G G C C C C C C C C C C C C C C C C	AAC ACG CTG TAC Y AGC S CCC Y AGC S CCA CCT CCT CCT CCT CCT CCT CCT CCT TGC	AAG CAG GTC V AAC N AAC I GAG E TCC S AAC I AAC N CAC C CCC GCC CCC GTC GTC GTC GTC CCC CCC	CGC CAG CCG CCG CCG CCG CCG CCG CCG CCG	GCC GAC U GTG GTG C GTG V GTG C C C C C C C C C C C C C C C C C C	AAG AAG CTG GCTG GCT TTC F TTC GGC C GGC AAG GGC K TAT Y ATG GAT CGC GCC GCC GCC GCC GCC CCC C C C C C	TAC TAC X A N CTG L C T C C C C C C C C C C C C C C C C C	ACC GTG GTG V GCG A GCC A GCC A CC G C T C G G T C C G G T G G G C T G G C C G C T G G C C A C C C A C C C A C C C C C C C	AAG ATG M CTC L TTC F GTG CCC G T TAC CCC Q T AAC N CAC CTT CCC CCT CCC CCC CCC CCC CCC CCC	ATC AGC S AAG K CAG Q AAG C C G AAG C C G AAG C C G C G C C G C C C C	CAT H GGC GTG V TCG S ACG TCS ACG C GGG GGC GGC GGC GCC GCC GCC GCC G	TAC Y ATC CAG Q CAC CAG CAC CAG CC CAG CC CAC CCCT CGC CGC CGC CGC CGC CGC CGC CGC C	GCG GCG GCG GCG GCA GCA GCA GCA GCA CCA C	CAC CAC H TTC F AAC N TAC C C C C C C C C C C C C C C C C C C	ATCC I ACCG T C C ATCC C ATCC C ATCC C C C C C C C	GAC GAC D GCC A GGC GGC GGC GGC CGC CGC CGC CGC C	CAG CAG Q TTCC F TCG S S TGG G G G G G G G G G G G G G G C C C C	F GAC D GAC D CTG L CTG CTG CTG CTG CTG CTG CCG CCG	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1728 1728 1728 1836 1890 1944 1998 2052 2160 2214		
TCG CCG ACA AGA CGC TGG ACT ATT TCC ATC GCG GCC GCC ACA CCG GCG CGG GCG 2375 ATC GCA GCA TCC GCA TCT TCG AGA CCG ACC GCC GCC ACA CCG CGC GCG CC 2430 GGA AGG ACG CGG GCG GCG ACA CCG GCA AGT ACC TCT CGC ACC TGG GCT ACG CCA 2484 ACT ACT TCA GCC CGC AGA CCT GGA CCT CGA TGG TCG CCT GGG TGC CCT ACA CGC 2538 GCA GAC CTC GCC CG 3' 2552	221 239 257 275 293 311 329 347 365 383	AAG GAC GAC C D AAG K C G C C R C R C R C R C R C R C R C R C	AACC ACT CTG L TAC Y AGC S GCC A TCC S GAC D ATC CGG CCT GAC CCT GAC CCT GCC CCT CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT GCC CCT CCT	AAG CAG GTC V AAC N C AAC S C C C C C C C C C C C C C C C C	CCGC CAG CCG CCG CCG CCG CCG CCG CCG CCG	GAC GAC GTG GTG V GTG V C C C C C C C C C C C C C C C C C C	AAG CTG GCCTG GCTG TTC F TTC F TTC F CGC GAA GAT CGC GAA GAC GCC GCC GCC GCC GCC GCC GCC	TAC TAC X AAC N CTG L CTG GCC D GAC D CCC P TCG GAC CCC CCC CCC CCC CCC CCC CCC CCC C	ACC GTG GTG V GCG A GCC A GCC A CCG TCG GAT TCG GGC TCG GGC TCG GGC TCG GGC TCG GGC CCG CCG	AAG ATG M CTC L TTC F GTG V TAC CCG GT CCC CCC CCC CCC CCC CCC CCC CCC	AGC S AAG K CAG Q AAG K CTG CTG CTG AAG S TGA AGC S GTG TCC CGA AGC CGG AAG ACC	CAT H GGC G G G G G G G G G G G G G G G G G	TAC Y ATC I CAG Q CA GAC C C C C C C C C C C C C C C C C	GCG GCG GCG GCG GCC GCC GCC GCC CCA CCC CCA CCA	CAC CAC H TTC F AAC N TAC R AGC CCG CCC CCC CCC CCC CCC CCC CCC CCC	ATC I ACC T ACG T C C ATC C ATC C C ATC C C C A T C C C C	GAC GAC D GCC G G G G G G G G G G G G C G C	CAG CAG Q TTCC F TCG S AGC S TGG W AAT N AAT N GGG GCG CCC GCC GCC GCC GCC GCC GCC G	FICE F F F C C C C C C C C C C C C C	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836 1890 1944 1998 2052 2106 2160 2214 2268		
ATC GCA GCA TCT CG ACA TCG ACA CCG ACC GCC GCG TCG ACG GCG GCG 2430 GGA AGG ACG CGG GCG GCG ACA CCG GCA AGT ACC TCT CGC ACC TGG GCT ACG CCA ACT ACT TCA GCC CGC AGA AGA CCT CGA TGG TCG CCT GGG TGC TGG CCT ACA CGC GCA GAC CTC GCC CG 3'	221 239 257 275 293 311 329 347 365 383	AAG GAC GAC GAC C C C C C C C C C C C C	ACG ACG CTG L TACC S CCTG S CCT S CCT CCT CCT CCT CCT CCT CCT CCT	AAG CAG GTC V AAC N ATC I GAG E C S ATC I C S C ATC S C ATC S C C C C C C T C G T C S C C T C S C C T C C S T C V A C N C V A C V A C V A C V A C V A C V A C V A C V A C V A C V A C V A C V A C V A A C V A C C V A A C S T C V A A C S T C V A A C S C C V A A C S C C V A A C S C C V A A C S A C S C C V A A C S C C S C S C C S C S C S C S C C S C S C S C C S C S C S C S C S C C S C S C C S C C S S C C C S C S C C S C C S C C C C C S C	CCGC CAG P CTG CCG P CTG S CTG S CTC C CGC C CGC CGC CGC CGC CGC CGC CGC	GAC GAC V GTG C C C C C C C C C C C C C C C C C C	GCG AAG CTG GGC G G G G G G G G G G G G G G G	TAC TAC N CTG C C C C C C C C C C C C C C C C C C	ACC GTG GTG V GCG A GCC A GCC TCG GAT CCG C GGC CCG CCTT CCC CCGC CCG	AAG ATG M CTC L TTC F GTG V TACC F P ACG T AAC N CAG Q CAC CCT CGC CGC CGC CGC CGC CGC CGC CGC ATT GCG	ATC AGC S AAG K CAG Q AAG K CTG AAG K CTG AAG C CGA CGC CGC CGC CGC CGC CGC CGC	GCCAT GGC GGG GGG GGG GGC GGC GGC GG	TAC Y ATC CAG GCC A GCC C A GCC C C C C C C C C	GCG GCG GCG GTG GCA GCA GCA GCA GCA CAG CCA CCC CCA ACC CCA CCA	CAC CAC H TTC F AAC N AAC N TAC S CCG C CCG C CCG C CCG C CCC CCC AGG CCC CCC	ATC I ACC T ACC T T GCC A T C C C T T C C C T T C C C T T C C C T T C C C T T C C C T T C C C T T C C C T T C C C C T T C C C T T C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C T T C C C T T C C C T T C C C T T C C C T T C C C T T C C C T T C C C T T C C C T T C C C T T C C C T T C C C C T T C C C C T T C C C C T C	GAC D GCC A GGC G G G G G G G G G G G G C G C	CAG CAG Q TTC F TCG S AGC S AGC S AGC G G G G G G G G G G G C CAC CCC CCC	F GAC D GAC D GAC C C C C C C C C C C C C C	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836 1890 1944 1948 2052 2106 2160 2214 2268 2322		
GGA AGG ACG CGG GCG GCG ACA CCG GCA AGT ACC TCT CGC ACC TGG GCT ACG CCA ACT ACT TCA GCC CGC AGC AGA CCT CGA TGG TCG CCT GGG TGC TGG CCT ACA CGC GCA GAC CTC GCC CG 3'	221 239 257 275 293 311 329 347 365 383	AAG GAC GAC C C C C C C C C C C C C C C	ACG ACG L TACC L TACC S C C C C C C C C C C C C C C C C C	AAC CAG GTC GTC N AAC S CAC GGTC S CAC GCG CCCT GTC CCC GTT CCG GTT CCC CCC C	CCGC CAG Q CCGG P CCGG L TCG S ACC TCG S ACC S CT ACC S CT C C C G C G C G C G C C G C C G C C G C C G C C G G C C G G C C G G C C G G S C C G G S S A C C G G S S A C C G G S S C C G G S S A C C G S S A C C G S S A C C G S S A C C G S S A C C G S S A C C G S S A C C G S S A C C G S S A C C G S S A C C S S S A C C S S S A C C S S S A C C S S S A C C S S S A C C S S S S	GCC GAC GTG GTG C GTG C C C C C C C C C C C C	AAG AAG CTG GGC TTC F GGC TTC F GGC AAG GAT GAT GAT GCG GCA GCG GCA GCC GCA GCC GCC GCA CCC CCC	TAC TAC XAAC N CTG GCC C C C C C C C C C C C C C C C C	ACC GTG GTG V GCG A GCC A TCG GAT TCG GAT TCG GAT TCC GCG GCG GCG GCG GCG GCG GCG GCG GC	AAG ATG M CTC F GTG TTC C F GTG V V CCG P AACG T CAC C T C CAC C C C C C C C C C C C C	ATC AGC S AAG CAG Q AAG CTG CTG CG AAG AGC S TGA * CGA GGG GGG CGC GGGT CGG AGA CGG AGA CGG ACC	CAT GGC GTG GTG S ACG GGC GCG GCG GCC GCC GCC GCC GCC GCC	ACCC CAC	GCG GCG GCG GCG GCG GCC GCC GCC GCC CAG GCC CCC GCC CCC C	CAC CAC H TTC F F AAC N AAC N AAC N CGC C C C C C C C C C C C C C C C C	ATC I ACC T C C T C C T C C C T C C C T C C C T C C C T C C C T C C C T C C C T C C C T C C C C T C	GAC GAC D GGC G G G G G G G G G G G G G G G C G C	CAG CAG Q TTC F F TCG S AGC S AGC S AGC S CGC GCG GCG GCG CCCC GCC GCG CCCC GCG CCCC GCG CCCC CCG CCCC CCCC CCCC GCG CCCC GCG CCCC CCCC CCC CCCC CCCC CCCC CCCC CCCC CCCC	F F GAC D GAC D ATC L CTG CTG CTG CTG CTG CTG CTG CTG	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1788 21836 1890 1944 1994 2052 21060 2214 2268 2325		
ACT ACT TCA GCC CGC AGC AGA CCT CGA TGG TCG CCT GGG TGC TGG CCT ACA CGC 2538 GCA GAC CTC GCC CG 3' 2552	221 239 257 275 293 311 329 347 365 383	AAG GAC AAG GGC R GGC CC R GGC CC CC CC CC CC CC CC CC CC CC CC CC	ACG ACG CTGC L CTG S CCTG S CCC A TCC S CCC T TCC S CCC T TCC CCT CCT CCC CCC	AAG CAG GCAG V AACC I GAG TCC S ATCC I CAC N TCC S CACC CCC CCC CCC CCC CCC CCC CCC CC	CGC CAG CCG P CTG S CTG S CTG S CTG S CTG S CTG C CGC CGC CGC CGC CGC CGC CGC CGC C	GCG GAC GTG GTG C GTG C C C C C C C C C C C C	AAG AAG GGC G G G T T C T G G G C T T T C T F C F G G C C T C T C T C T C T C C G G C T T C C F C G G C C T T C C F C G G C C T T C C C G G C C T T C C C G G C T T C C F C G G C C T T C C F C F C G G C C T T C C F C G G C C F C F C G G C C F C F C	TAC TAC X X AAC N CTG C C C C C C C C C C C C C C C C C C	ACC GTG GTG V GCG A GCC A TCG S ACC T GAT CCG G G G CCG G G G CCG G G G CCG G G CCG G G CCG G G CCG G G CCG A A CCG G CCG A CCG A CCG A CCG A CCG A CCG A CCG A CCG A CCG A CCG A CCG A CCG C C C C	AAG ATG M CTC F GTG V CCG P ACG T CCG CAG CAG CAG CAG CAG CAG CGC CGC CGC	ATC AGC S AAG K CAG Q AAG K CTG D ATG AGC TGA CGG GGT CGGA AGC CGG CGGA AGC CGG ACC CGG AACC CGG CGG	ICCAT GGC GGG GGG GGG GGC GGC GGC GGC GGC GGC	GGC TAC Y ATC I GAC D GGC C A GGC C A G C C C C C C C C C C C	GCG GCG GGCG GGCG GGCG CAG GGC CAG CCA CCC CCA CCC CCA CCC CCA CCC CCA CCC CCA CCCC CCA CCCC CCA CCCC CCCC CCCC CCCC CCCC CCCC CCCCC CCCC	CAC CAC H TTC F AAC N AAC N CGC C CGC CCA CGC CCCA CGC CCCA CGC CCCA CGC CCCA CGC CCCA CGC CCCA CGC CCCA CGC CAC CA	ATC I ACC T T C C T T C C C T T C C C T T C C T T C C T T C C C T T C C C T T C C C T T C C C T T C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C C C C T T C	GAC D GCC G GCC G G GCC G G G G C G C G C	CAG CAG Q TTCC F TCG S AGC S S TGG G G G G G G G G G G C G C CCG C CCG C CCG C CGG G CGG G CGG G CGG G CGG G CGG G CGG G CGG G CGG C CAG S S C CAG S S C S S C S S S S S S S S S S S S S	F GAC D GAC D C L C C L C C C A C C C C C C C C C C C C C	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836 1890 1944 1998 2052 2106 2160 2214 2268 2325 2375 2430		
GCA GAC CTC GCC CG 3' 2552	221 239 257 275 293 311 329 347 365 383	AAG GAC AAG GGC G GGC C C G C C C G C C C G C	AACG CTG CTG CTG CTG CTG CTG CTG CCG CCG C	AAG CAG GCAG GCAG GCA CACA CACA CACA CA	CGC CAG CCAG PG CCG PG CTG S ACC TCG S ACC TCG S ACC TCG S CCG CGC CGC CGC CGC CGC CGC CGC CGC	GGAC GTG GTG V GTG C C C C C C C C C C C C C C C C C C	AAG GGC GGC CTG GGC TTC F GGC G AAG GAT CGC GAA GAA GGAC GCA GCC GCC GCC GCC GC	TAC TAC TAC SC CTG C C C C C C C C C C C C C C C C C	ACC GTG GTG V GCG A GCC A TCG GCC TCG G GCC TCG G GCC G G TCG G GCC G G C TCG G G C C C G G C C C G C C A A C C C G C G	AAG ATG M CTC F GTG GTG TTC C C C C C T TACC P ACG T C C C C C C C C C C C C C C C C C C	ATC AGC S AAG CAG Q AAG CTG CTG CTG CGA AGC CGA AGC CGGCC CGGC	GCCAT H GGC G G G G G G G G C C C A T C G G G G G G G G G G G G G G G G G G	ATC I CAG Q GCC C AGC C AGC C C AGC C C C C C C C C	GCG GCG GCG GCG GCG GCG CAG CCAG CCAG C	CCAC H TTC F AAC N AAC N TAC Y CGC R AGC S CCG F GCG CCC TGA TCC GGT CCCA CGC ACG ACC ACCA ACC	ATC I ACC T ACG T T GCG ATC C ATC C ATC C ATC C C C C C C C C C	GAC D GCC A GGC G GGC G GGC C GCC C C C C C C C C C C C C C C C C C C C	CAG CAG TTC F TCG S AGC S S AGC S S AGC M AAT N GGG G GCG CCAC CCAC CCAC CCAC CCAC CC	F GAC D GAC D GAC C C C C C C C C C C C C C	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836 1890 1944 1798 2052 2106 2214 22375 2430 2484		
	221 239 257 275 293 311 329 347 365 383	AAG GAC AAG GGAC CGC R GGC CGC CGC CGC CGC CGC CGC CGC	AACG ACG CTG CTG CTG CTG CTG CTG CCT CCT	AAG CAG GCAG V AACC I GAG E TCC S CAC GCG CCC CAC GCG CCC CGC CCC CGC CCC CC	CAG CAG CCAG P CTG CTG CTG S ACC T TCG S ACC T T AGC S TAC S CGC G CGC CGC CGC CGC CGC CGC CGC CG	GGAC GAC GTG GTG V GCC C C TG C C C C TG C C C C C C C C C	GCG GCC GCC GCC GCC F TTC F GGC G G AAG GAC GCC GCA GCC GCA GCC GCA GCC GCC	TAC TAC N CTG GCT G GAC D CCC P TCG GAC CCC CTC GAC CCC CTC GAC CCC CTC CCC CTC CCC CTC CCC CTC CCC CTC CCC CTC CCC CTC C CCC C C C C C C C C C C C C C C C C	ACC GTG GTG GCG A GCC A CCG G CC T CCG G CCT CCG CCT CCC CCT CCC CCT CCC CCC	AAG M ATG M CTC L F GTG GTG V TAC Y CCG P ACG O TAC Y TAC Y CAG Q CAC CGT CGC CGC CGC CGC CGC CGC CCG CCG CC	ATC S AAG C S AAG C C C C	GCCAT GGC GGG GGG GGG GGC GGC GGC GG	GGC TAC Y ATC I GCC Q GCC A GCC C C C C C C C C C C C C C C C	GCG GCG GCG GCG GCG GCG CAG QCAG CAG CCC CCC CCC CCCC C	CCAC H TTC F AAC N TAC TAC C C C C C C C C C C C C C C C C	ATC ATC T ACC T ACC T T G C C ATC C T T G C C ATC C C ATC C C ATC C C ATC C C C ATC C C C C C C C C C C C C C	GAC D GCC A GGC G G GC G G G C G C G C G C G	CAG Q TTC F TCG S TGG S TGG W M ATG G G G G G G G G G G G G G C G C C C C	FIGERAL CONTRACTOR OF CONTRACT	1242 1290 1350 1404 1458 1512 1566 1620 1674 1728 1782 1836 1890 1944 1998 2052 2106 2166 2160 214 2268 2375 2430 24430		

FIG. 4. Nucleotide sequence and deduced amino acid sequence of the choA gene from M. chitosanotabidus 3001. The putative Shine-Dalgarno sequence is indicated by a double underline. The one-letter amino acid code is aligned below the nucleotide sequence. The signal peptide sequence is underlined with a dashed line, and a signal sequence cleavage site is represented by a vertical arrow between the alanine residues at positions 80 and 81. Boxed amino acids were determined by N-terminal amino acid sequencing of the purified chitosanase from E. coli expressing the choA gene. A wavy line indicates the amino acid sequence obtained from the purified chitosanase. Two arrows indicate the nucleotide sequence of the primers that were used for PCR. The single bold arrow indicates the primer used to amplify the PCR fragment by using the cassette-cassette primer PCR amplification method.

position 1,702. The open reading frame of 1,173 bp would code for a 391-amino-acid polypeptide. The deduced N-terminal 80 amino acids showed the typical features of signal peptides, such as the presence of hydrophobic amino acids. A comparison of the signal sequence region of three chitosanases from M. chitosanotabidus, Nocardioides sp. strain N106, and Streptomyces sp. strain N174 is shown in Fig. 6. Although the signal sequence regions are similar among these chitosanases, other regions did not show any significant homologies.

N-terminal cleavage site of ChoA. We found that the signal peptide was cleaved between alanine residues 80 and 81 by determining the amino-terminal sequence of purified mature chitosanase from M. chitosanotabidus. Cleavage at amino acid position 80 would result in a protein of 311 amino acids with a



FIG. 5. Detection of chitosanase activity on the chitosanase medium plate. *M. chitosanotabidus* 3001 (A) and *E. coli* carrying the plasmid pChoU1 and its deletion derivatives (B) were spotted onto the colloidal chitosan plate to assay for chitosanase expression. The plates were incubated at 30 and 37°C for 3 days, respectively. Chitosanase activity was identified as a clear zone around the colony on the colloidal chitosan plate. *E. coli* carrying the plasmid pChoU1 and its deletion derivatives are indicated as follows: 1, pChoU1; 2, pChoU268; 3, pChou396; 4, pChoU541; 5, pBluescript II SK(-); 6, pChoR2552; 7, pChoR1821; and 8, pChoR1482.

calculated molecular weight of 34,000. This is consistent with the size of the 34-kDa chitosanase estimated by SDS-PAGE. The 15 N-terminal amino acids of the 34-kDa protein produced by *E. coli* carrying the cloned *choA* gene was determined to be AAAAGVIPVGDSRVY. This sequence is identical to the one predicted from the *choA* DNA sequence and lies between Ala-81 and Tyr-95. The 80-amino-acid sequence is considered to be the signal sequence responsible for the export of ChoA into the medium both in *M. chitosanotabidus* and *E. coli*. We determined the enzymatic activity in a culture of *E. coli* expressing *choA*. Whereas the extracellular chitosanase activity was 3.4 U/mg, the cell-associated chitosanase activity was only 0.13 U/mg, indicating that most of the enzyme had been released.

To verify that the N-terminal signal sequence of ChoA is functional in E. coli, the N-terminal sequence of ChoA was fused to GFP. GFP fusions bearing the 48-amino-acid (pB48ChG) or the 88-amino-acid (pB88ChG) region of ChoA were constructed and inserted into the expression plasmid (Fig. 3). E. coli cells harboring these plasmids were grown, and extracellular and intracellular proteins were detected by Western blot analysis with the GFP antibody. While both of these proteins were detected intracellularly (Fig. 7, lanes 2, 4, and 6), the extracellular protein contained the GFP fusion protein with the 88amino-acid region (lane 5) but not the GFP fusion protein with the 44-amino-acid region or GFP alone (lanes 1 and 3). The cleaved ChoA protein was only seen when the 88-amino-acid-GFP fusion was employed, and a fraction of the extracellular proteins still possessed the signal sequence (lane 5). These results suggest that the N-terminal signal sequence was functional in E. coli and that this region needs to be longer than 48 amino acids to be functional.

DISCUSSION

Table 4 summarizes all of the microbial chitosanases that have been identified to date. The data are still relatively limited compared to similar data for the many chitinase genes that have been analyzed thus far (19). The molecular weights of chitosanase are relatively low (26,000 to 50,000) compared to those of chitinases (31,000 to 115,000) (19). In the present study the 34-kDa chitosanase produced by M. chitosanotabidus 3001 was purified and characterized, and its corresponding gene was cloned. Although the enzymatic properties of M. chitosanotabidus 3001 chitosanase are similar to those of other chitosanases, the optimal temperature and pH of our chitosanase were the lowest among known chitosanases (Table 4). The chitosanase from *M. chitosanotabidus* 3001 is highly active on glycol chitosan, but it does not hydrolyze carboxymethylchitosan, colloidal chitin, or carboxymethylchitin. Most known chitosanases do not hydrolyze 100% chitin or cellulose. However, a chitosanase from Myxobacter AL-1 (7, 21) has the unique feature of being able to hydrolyze carboxymethyl cellulose, and a chitosanase from Acinetobacter sp. strain CHB101 hydrolyzes glycol chitin (28). Recently, a part of the chitosanase sequence from Myxobacter AL-1 was reported to have structural similarity to an endoglucanase from Bacillus subtilis (21). Even though there are widespread similarities in the activities of chitosanases, there seem to be many different structural types of chitosanases in nature.

Purified chitosanase from *M. chitosanotabidus* 3001 had highest activities with 90% deacetylated chitosan and efficiently hydrolyzed 70 to 100% deacetylated chitosan. The fact that 90% deacetylated chitosan is more susceptible to hydrolysis than 100% deacetylated chitosan may suggest that *N*-acetylglucosamine residues in the chitosan are important in the recognition

ChitoM 1 MQLPRPDLREFARRALPLLAASTLAAAFGAASPALAAGNRV 42 ChitoN 1 MRLKHPT-ARLALALLVAVPRSVAAAGTVHAAPAPAGATRL 41 ChitoS 1 MHSQHRT-ARIALAVULTAIPASLATAGVGYASTQASTAVK 40

FIG. 6. Comparison of the N-terminal amino acid sequences of three chitosanases. N-terminal amino acid sequences of the chitosanases from *M. chitosanotabidus* 3001 (ChitoM), *Nocardioides* sp. strain N106 (12) (ChitoN), and *Streptomyces* sp. strain N174 (13) (ChitoS) are shown. Identical amino acids are boxed.

TABLE 4. Comparison of properties of purified chitosanases produced by microorganisms

		No. of	Optimum activity							
Microorganism	Mol wt ^a	amino acids ^b	pH	Temp	Inhibitor(s)	Substrate specificity (% deacetylation)	Source or reference			
Matsuebacter chitosanotabidus 3001	34,000	311	4.0	30-40	Ag ⁺	Chitosan (90), glycol chitosan	This study			
Myxobacter AL-1	31,000		5.0, 6.8	70	Hg ²⁺ , Ag ⁺ , Cd ²⁺ , Zn ²⁺	CMC^{f} , chitosan	7, 21			
Acinetobacter sp. strain CHB101							28			
Chitosanase I	37,000		5–9	65	ND^{g}	Chitosan (70–90)				
Chitosanase II	30,000		5–9	50	ND	Glycol chitosan				
Streptomyces sp. N174	29,500	238	5.5	65	ND	Chitosan (40–99)	2, 13			
Fusarium solani f. sp. phaseoli	36,000	285	5.6	40	No inhibitor ^c	Chitosan (70–100), glycol chitosan, CMC	26			
Bacillus megaterium P1 chitosanase A	43,000		4.5-6.5	50	ND	Chitosan (81), CMC	22			
Amycolatopsis sp. strain CsO-2	27,000		5.3	55	Hg^{2+} , pCMB ^d , NBS ^e	Chitosan (100)	16			
Penicillium islandicum	30,000		4.5-6.0	45	Hg^{2+}, Cu^{2+}	Chitosan (70)	4			
Enterobacter sp. strain G-1	50,000		7.0	50	NĎ	Chitosan (80)	30			
Bacillus circulans MH-K1	30,000	259	6.5	50	Cu ²⁺ , Hg ²⁺ , Ni ²⁺ , Zn ²⁺ , pCMB	Chitosan (80)	29			
Nocardioides sp.	29,500	237			NĎ	Chitosan (100)	12			

^a Determined by SDS-PAGE analysis with purified enzyme.

^b The processed amino acid number as deduced from the gene sequence.

^c Tested chemicals were Ca²⁺, Mg²⁺, Co²⁺, Ni²⁺, Fe²⁺, Hg²⁺, and Cu²⁺. ^d pCMB, *para*-chloromercuribenzoate.

^e NBS, N-bromosuccinimide.

f Carboxymethyl cellulose.

^g ND, not determined.

mechanism of the substrate by the enzyme. The ability to degrade 100% deacetylated chitosan is an important characteristic used to categorize enzymes as chitosanases. However, enzymes in nature have different specificities depending on the different degrees of deacetylation of chitosan, as indicated in Table 4.

The chitosanase activity of M. chitosanotabidus 3001 was not inhibited by Ba^{2+} , Co^{2+} , Hg^{2+} , Ca^{2+} , Cu^{2+} , Mg^{2+} , Fe^{2+} , Mn^{2+} , and Zn^{2+} , but it was completely inhibited by 1 mM Ag⁺. Chitosanase from *Myxobacter* AL-1 is an example of another enzyme inhibited by Ag^+ (7, 21). However, this chitosanase differs from that of *M. chitosanotabidus* 3001 in several ways (Table 4).

The choA gene encoding a 34-kDa chitosanase was cloned into E. coli, and the complete nucleotide sequence of the gene was determined. This is the fifth chitosanase whose primary structure has been determined (Table 4). The deduced amino acid sequence of the choA gene is identical to the amino acid sequence of the purified chitosanase (Fig. 4). The choA gene encodes a polypeptide of 391 amino acid residues which includes an 80-residue signal sequence. This 80-residue signal sequence is functional in E. coli (Fig. 7). Fusion of the signal sequence of ChoA with GFP suggested that the signal for excretion required more than 48 amino acids but fewer than 88 amino acids (Fig. 7). Surprisingly, however, the GFP fusion protein with the 88-amino-acid region, as well as the cleaved protein, was found in both the intracellular and extracellular fractions. We speculate that the signal peptidase of E. coli recognizes the cleavage site of the fusion protein but that the fusion protein may not necessarily have to be cleaved for excretion.

Computer analyses of the deduced amino acid sequence did not reveal a strong identity with any other protein sequences of glycolytic enzymes, including the chitosanases, chitinases, and N-acetylglucosaminidases included in protein databases. Three bacterial chitosanases from B. circulans MH-K1, Streptomyces sp. strain N174, and Nocardioides sp. have a common primary structure (3). However, ChoA from M. chitosanotabidus, as well as a fungal chitosanase, is different. Thus far, at least three

different types of chitosanases have been identified on the basis of their primary structure.

There is an 80-amino-acid signal polypeptide at the N terminus of ChoA of M. chitosanotabidus 3001 (Fig. 6). The Nterminal 15-amino-acid sequence determined from the 34-kDa mature chitosanase produced in E. coli was identical to the N-terminal amino acid sequence of chitosanase from M. chitosanotabidus 3001, indicating that the mature chitosanase is cleaved at the same site in E. coli. This property will be useful for the mass production of chitosanase in E. coli.

Boucher et al. (3) suggested that two invariant carboxylic amino acid residues of Glu22 and Asp40 conserved in Nterminal segments of chitosanase from Streptomyces sp. strain N174 are essential for its catalytic activity. We have looked for the corresponding amino acid combination in ChoA of M. chitosanotabidus 3001. There are two candidate residues: residues Glu121 and Asp139 or residues Glu137 and Asp152. However, since the regions surrounding these amino acids are not conserved in other chitosanases, it is difficult to predict



FIG. 7. Western blot analysis of the chitosanase-GFP fusion proteins. Extracellular proteins (lanes 1, 3, and 5) and intracellular proteins (lanes 2, 4, and 6) from E. coli JM109 harboring pBG (lanes 1 and 2), pB48ChG (lanes 3 and 4), or pB88ChG (lanes 5 and 6) were loaded onto the gel. Western blot analysis was done by using an anti-GFP antibody. The extracellular ChoA-GFP protein was only seen in lane 5.

whether they may constitute the catalytic center or not. Further analysis will be required to define the active region of M. chitosanotabidus 3001 chitosanase.

ACKNOWLEDGMENTS

This work was supported by a grant-in-aid from the Ministry of Education, Science, and Culture of Japan and by San-in Kensetsu Kogyo Co., Ltd. J. K. Park was supported by the commemorative foundation of Yoneyama International Rotary and Iwatanai Naozi. S. Imamura and F. Ishii participated in some of this work.

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. Boucher, I., A. Dupuy, P. Vidal, W. A. Neugebauer, and R. Brzezinski. 1992. Purification and characterization of a chitosanase from Streptomyces N174. Appl. Microbiol. Biotechnol. 38:188-193.
- Boucher, I., T. Fukamizo, Y. Honda, G. E. Willick, W. A. Neugebauer, and R. 3. Brzezinski. 1995. Site-direct mutagenesis of evolutionary conserved carboxylic amino acids in the chitosanase from Streptomyces sp. N174 reveals two residues essential for catalysis. J. Biol. Chem. 270:31077-31082.
- 4. Fenton, D. M., and D. E. Eveleigh. 1981. Purification and mode of action of a chitosanase from Penicillium islandicum. J. Gene. Microbiol. 126:151-165.
- 5. Fukamizo, T., T. Ohkawa, Y. Ikeda, and S. Goto. 1994. Specificity of chitosanase from Bacillus pumilus. Biochim. Biophys. Acta 1205:183-188.
- 6. Grenier, J., N. Benhamou, and A. Asselin. 1991. Colloidal gold-complexed chitosanase: a new probe for ultrastructural localization of chitosan in fungi. J. Gen. Microbiol. 137:2007-2015.
- 7. Hedges, A., and R. S. Wolfe. 1974. Extracellular enzyme from Myxobacter AL-1 that exhibits both β-1,4-glucanase and chitosanase activities. J. Bacteriol. 120:844-853.
- 8. Imoto, T., and K. Yagishita. 1971. A simple activity measurement of lysozyme. Agric. Biol. Chem. 35:1154-1156.
- 9. Isegawa, Y., J. Sheng, Y. Sokawa, K. Yamanishi, O. Nakagomi, and S. Ueda. 1992. Selective amplification of cDNA sequence from total RNA by cassetteligation mediated polymerase chain reaction (PCR): application to sequencing 6.5-kb genome segment of hantavirus strain B-1. Mol. Cell. Probes 6:467-475
- 10. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680-685.
- 11. Marcotte, E. M., A. F. Monzingo, S. R. Ernst, R. Brzezinski, and J. D. Robertus. 1996. X-ray structure of an anti-fungal chitosanase from Streptomyces N174. Nat. Struct. Biol. 3:155-162.
- 12. Masson, J. Y., I. Boucher, W. A. Neugebauer, D. Ramotar, and R. Brzezinski. 1995. A new chitosanase gene from a Nocardioides sp. is a third member of glycosyl hydrolase family 46. Microbiol. 141:2629-2635.
- 13. Masson, J. Y., F. Denis, and R. Brzezinski. 1994. Primary sequence of the chitosanase from Streptomyces sp. strain N174 and comparison with other endoglycosidases. Gene 140:103-107.
- 14. Matsuo, Y., M. Kurita, J. K. Park, K. Tanaka, T. Nakagawa, M. Kawamukai,

and H. Matsuda. 1999. Purification, characterization and gene analysis of N-acetylglucosaminidase from Enterobacter sp. G-1. Biosci. Biotechnol. Biochem. 63:1261-1268

- 15. Nabeshima, K., S. Saitoh, and M. Yanagida. 1997. Use of green fluorescent protein for intracellular protein localization in living fission yeast cells. Methods Enzymol. 283:459-471.
- 16. Okajima, S., A. Ando, H. Shinoyama, and T. Fujii. 1994. Purification and characterization of an extracellular chitosanase produced by Amycolatopsis sp. CsO-2. J. Ferment. Bioeng. 77:617-620.
- 17. Okajima, S., T. Kinouchi, Y. Mikami, and A. Ando. 1995. Purification and some properties of a chitosanase of Nocardioides sp. J. Gen. Appl. Microbiol. 41.351-357
- 18. Osswald, W. F., J. P. Shapiro, H. Doostdar, R. E. McDonald, R. P. Niedz, C. J. Narin, C. J. Hearn, and R. T. Mayer. 1994. Identification and characterization of acidic hydrolases with chitinase and chitosanase activities from sweet orange callus tissue. Plant Cell Physiol. 35:811-820.
- 19. Park, J. K., K. Morita, I. Fukumoto, Y. Yamasaki, T. Nakagawa, M. Kawamukai, and H. Matsuda. 1997. Purification and characterization of the chitinase (ChiA) from Enterobacter sp. G-1. Biosci. Biotechnol. Biochem. 61:684-689.
- 20. Park, J. K., Y. Matsuo, K. Shimono, K. Tanaka, T. Nakagawa, A. Yokota, H. Matsuda, and M. Kawamukai, Matsuebacter chitosanotabidus gen, nov., sp. nov., an aerobic chitosanase producing gram-negative bacterium belonging to the beta subclass of Proteobacteria. Submitted for publication.
- 21. Pedraza-Reyes, M., and F. Gutierrez-Corona. 1997. The bifunctional enzyme chitosanase-cellulase produced by the gram-negative microorganism Myxobacter sp. AL-1 is highly similar to Bacillus subtilis endoglucanases. Arch. Microbiol. 168:321-327.
- Pelletier, A., and J. Sygusch. 1990. Purification and characterization of three chitosanase activities from Bacillus megaterium P1. Appl. Environ. Microbiol. 56:844-848
- 23 Robertus, J. D., A. F. Monzingo, E. M. Marcotte, and P. J. Hart. 1998. Structural analysis shows five glycohydrolase families diverged from a common ancestor. J. Exp. Zool. 282:127-132.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a 24 laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- 26. Shimosaka, M., M. Kumehara, X. Y. Zhang, M. Nogawa, and M. Okazaki. 1996. Cloning and characterization of a chitosanase gene from the plant pathogenic fungus Fusarium solani. J. Ferment. Bioeng. 82:426-431.
- 27 Shimosaka, M., M. Nogawa, Y. Ohno, and M. Okazaki. 1993. Chitosanase from the plant pathogenic fungus, Fusarium solani f. sp. phaseoli: purification and some properties. Biosci. Biotechnol. Biochem. 57:231-235.
- Shimosaka, M., M. Nogawa, X. Y. Wang, M. Kumehara, and M. Okazaki. 28. 1995. Production of two chitosanases from a chitosan-assimilating bacterium, Acinetobacter sp. strain CHB101. Appl. Environ. Microbiol. 61:438-442
- 29. Yabuki, M., A. Uchiyama, K. Suzuki, A. Ando, and T. Fujii. 1988. Purification and properties of chitosanase from Bacillus circulans MH-K1. J. Gen. Appl. Microbiol. 34:255-270.
- Yamasaki, Y., I. Hayashi, Y. Ohta, T. Nakagawa, M. Kawamukai, and H. 30. Matsuda. 1993. Purification and mode of action of chitosanolytic enzyme from Enterobacter sp. G-1. Biosci. Biotechnol. Biochem. 57:444-449.