Purification and Characterization of a *Clostridium perfringens* 120-Kilodalton Collagenase and Nucleotide Sequence of the Corresponding Gene

OSAMU MATSUSHITA, KUNIKO YOSHIHARA, SEI-ICHI KATAYAMA, JUNZABURO MINAMI, AND AKINOBU OKABE*

Department of Microbiology, Kagawa Medical School, Kagawa 761-07, Japan

Received 16 February 1993/Accepted 18 October 1993

*Corresponding author. Mailing address: Department of Microbiology, Kagawa Medical School, 1750-1 Miki-cho, Kita-gun, Kagawa 761-07, Japan. Phone: +81 (878) 98-5111, ext. 2512. Fax: +81 (878) 98-7109. Electronic mail address: a64235@sakura.kudpc.kyoto-u.ap.jp.

*Clostridium perfringens* type C NCIB 10662 produced various gelatinolytic enzymes with molecular masses ranging from ~120 to ~80 kDa. A 120-kDa gelatinolytic enzyme was present in the largest quantity in the culture supernatant, and this enzyme was purified to homogeneity on the basis of sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The purified enzyme was identified as the major collagenase of the organism, and it cleaved typical collagenase substrates such as azocoll, a synthetic substrate (4-phenylazobenzoyl-carbonyl-Pro-Leu-Gly-Pro-n-Arg [Pz peptide]), and a type I collagen fibril. In addition, a gene (*colA*) encoding a 120-kDa collagenase was cloned in *Escherichia coli*. Nested deletions were used to define the coding region of *colA*, and this region was sequenced; from the nucleotide sequence, this gene encodes a protein of 1,104 amino acids (*M*₀ 125,966). Furthermore, from the N-terminal amino acid sequence of the purified enzyme, which was found in this reading frame, the molecular mass of the mature enzyme was calculated to be 116,339 Da. Analysis of the primary structure of the gene product showed that the enzyme was produced with a stretch of 86 amino acids containing a putative signal sequence. Within this stretch was found PLGP, the amino acid sequence constituting the Pz peptide. This sequence may be implicated in self-processing of the collagenase. A consensus zinc-binding sequence (HEXXH) suggested for vertebrate Zn collagenases is present in this bacterial collagenase. *Vibrio alginolyticus* collagenase and *Achromobacter lyticus* protease I showed significant homology with the 120-kDa collagenase of *C. perfringens*, suggesting that these three enzymes are evolutionarily related.

Certain bacteria produce collagenases, proteases which cleave native collagen in its triple helical conformation (17, 18, 21, 24, 34). Bacterial collagenases possess broad substrate specificities and degrade both native and denatured collagens (28), while vertebrate collagenases preferentially cleave the native form at a specific site (3). Much of our knowledge of bacterial collagenases has come from studies of the enzymes produced by *Clostridium histolyticum*. Seven different polypeptides with collagenolytic activity have been demonstrated in a culture of this organism, and these are divided into two groups on the basis of substrate specificity (6-8, 25, 35). Some other clostridia can also produce collagenase, but little is known regarding their biochemical properties.

*Clostridium perfringens* is one of the collagenase-producing clostridia (36). This organism is widely distributed in nature, particularly in soil and water contaminated with feces. It also lives in the intestinal tracts of humans and animals and sometimes causes infection, representing the most common pathogen of clostridial histotoxic infection (12). The collagenase produced by *C. perfringens* appears to be involved in tissue necrosis, along with other toxins such as phospholipase C and thiol-activated hemolysin (12, 14, 29). It also seems to play an important role in carcass putrefaction by the saprophytic life style of the organism, since collagen represents the most abundant protein in an animal body (12). A collagenase with a molecular weight of 80,000 was purified from *C. perfringens* several years ago (14), but it has not yet been fully characterized despite its importance as a virulence factor and as an enzyme involved in geochemical cycles.

In order to characterize the biological properties and the molecular structure of the *C. perfringens* collagenase, we have chosen to study the enzyme from *C. perfringens* type C NCIB 10662. While this strain produces collagenase in a moderate quantity, it expresses extremely low levels of alpha-toxin (16), which is often a contaminant in collagenase preparations. Therefore, this strain was judged to be useful for purification of the enzyme. This paper describes our successful purification and characterization of the 120-kDa collagenase as well as the nucleotide sequence of the corresponding gene (*colA*). Furthermore, since certain other bacterial proteases exhibit hydrolytic activity against denatured collagen, we suspected that the amino acid sequence of the collagenase might be similar to those of other bacterial proteases. We have therefore investigated similarities between these enzymes at the protein sequence level.

**MATERIALS AND METHODS**

**Bacterial strains, plasmids, and bacteriophage.** *C. perfringens* type C NCIB 10662 (16) was used throughout this study. *Escherichia coli* LE392 (22) was used to grow bacteriophage λ2001 (15). Recombinant λ phage were grown on *E. coli* P2392 (a P2 lysogen of LE392). A plasmid vector, pUC18 (41), and the host strain *E. coli* DH5a (2, 11) were used for all recombinant plasmid experiments.

**Media and culture conditions.** *C. perfringens* was precultured in cooked meat medium (Nissui, Tokyo, Japan) and was grown in GAM broth (Nissui) unless otherwise stated.
NZCYM medium (Gibco, Paisley, United Kingdom) (22) was used to grow λ2001 phage. Recombinant phage were plated on NZCYM agar plates containing 22 g of NZCYM broth powder (Gibco), 2 g of maltose, and 15 g of agar per liter. Transformed E. coli cells were selected on LB plates containing 20 g of LB broth base (Gibco) and 15 g of agar per liter, supplemented with 100 µg of ampicillin (Sigma Chemical Co., St. Louis, Mo.) per ml, 0.2 mM isopropyl-β-d-thiogalactopyranoside (IPTG) (Sigma), and 0.004% 5-bromo-4-chloro-3-indolyl-β-d-galacto-pyranoside (X-Gal) (Sigma).

Assay for collagenase. Protein concentrations were determined by using the Pierce bicinchoninic acid protein assay reagent (32) with bovine serum albumin as a standard. Collagenase activity was assayed with azocoll by the method of Kameyama and Akama (13) with slight modifications. The standard assay conditions were as follows. Azocoll (Sigma) powder was washed once with and resuspended in 0.2 M boric acid-borate buffer (pH 7.2)-0.15 M NaCl to give a concentration of 5 mg of azocoll per ml. Four hundred microliters of this assay mixture was taken in a 1.5-ml microcentrifuge tube, and 5 to 10 µl of enzyme solution was added. The tube then was placed horizontally and shaken at 37°C for 1 h. The supernatant was obtained by centrifugation at 15,000 × g for 15 min. The A_{420} of the supernatant was read. One unit of enzyme activity is defined as the amount of enzyme which causes an increase of 1.0 A_{420} unit per min. In some experiments collagenase activity was assayed by using 4-phenylazobenzyloxy-carbonyl-Pro-Leu-Gly-Pro-α-Arg (Pz peptide) as described by Wünsch and Heidrich (40).

Zymography. Gelatin zymography was carried out essentially by the method of Wilson et al. (37). GAM broth was inoculated with a 5% volume of C. perfringens preculture, and the cells were grown at 37°C without shaking. Portions of the culture were withdrawn after 2, 4, 6, and 8 h of incubation and chilled on ice, and the cells were removed by centrifugation.

The samples were mixed with nonreducing loading dye and applied to a 10% sodium dodecyl sulfate (SDS)-polyacrylamide gel, which had been copolymerized with gelatin (Wako Pure Chemicals, Osaka, Japan). After running, SDS was removed by washing the gel with 2.5% Triton X-100-100 mM NaCl, and then the gel was incubated overnight at 37°C in 50 mM Tris-HCl (pH 7.5)-10 mM CaCl₂ and stained with Coomassie brilliant blue R-250.

The activity of purified collagenase against collagen fibril was examined by the method of Birkedal-Hansen and Tayler (4), using acid-soluble type I collagen from calf skin (Sigma). Four micrograms of the purified collagenase was loaded on a conventional SDS-10% polyacrylamide gel under nonreducing conditions. After removal of SDS, the gel was briefly rinsed with phosphate-buffered saline (pH 7.4) containing 1 mM CaCl₂ and 1 mM ZnCl₂, and then it was placed on a collagen fibril film and incubated overnight at 37°C. The film was then stained with Coomassie brilliant blue R-250 in order to visualize an unstained band on a blue background.

Purification of collagenase. C. perfringens was grown in cooked meat medium and then in GAM broth at 37°C for 12 h. The final preculture was diluted 10-fold into 10 liters of GAM broth and incubated overnight at 25°C. Cells were removed by centrifugation at 15,000 × g for 30 min at 4°C, and the culture supernatant was concentrated to approximately 700 ml by ultrafiltration using a Pellicon apparatus and a polysulfone membrane (PTGC [5 ft²]; Millipore Co., Bedford, Mass.). Ammonium sulfate was then added to 60% saturation, and the precipitate was collected by centrifugation at 20,000 × g for 60 min at 4°C. The pellet was dissolved in 50 ml of 50 mM Tris-HCl buffer (pH 7.5)-5 mM CaCl₂, and the fraction was dialyzed against 1 liter of the same buffer with two changes.

The material was then applied to a DEAE-Sepharose A-25 column (2.2 by 50.0 cm). Proteins were eluted with a linear gradient from 0 to 500 mM NaCl in the starting buffer (over 4 bed volumes). Collagenase activity eluted in a single peak at approximately 100 mM NaCl.

The pooled collagenase fractions were concentrated by ultrafiltration with a dialysis tubing to approximately 10 ml. After dialysis against 50 mM Tris-HCl (pH 7.5)-5 mM CaCl₂, a 2-ml aliquot of the sample was applied to an ion-exchange fast protein liquid chromatography column (MonoQ; bed volume, 1 ml; Pharmacia LKB Biotechnology, Uppsala, Sweden) which was preequilibrated with 100 mM Tris-HCl (pH 7.0)-5 mM CaCl₂. Proteins were eluted by a 20-ml linear gradient from 0 to 0.6 M NaCl in 100 mM Tris-HCl (pH 7.0)-5 mM CaCl₂. Collagenase activity eluted at 210 mM NaCl. The pooled collagenase fractions were concentrated by ultrafiltration with a Centricon 30 (Amicon, Danvers, Mass.) to approximately 500 µl. A 100-µl aliquot of the sample was applied to a size exclusion column (Superose 12; bed volume, 25 ml; Pharmacia). The column was run in 100 mM Tris-HCl (pH 7.0)-5 mM CaCl₂, and collagenase activity eluted at 0.44 bed volume.

Determination of molecular weight and N-terminal amino acid sequence. Estimations of the molecular weight of the collagenase by gel filtration chromatography were made by using the Superose 12 column with detection at 280 nm. Approximately 30 µg of purified enzyme in 50 µl of 100 mM Tris-HCl (pH 7.0)-5 mM CaCl₂ was run in the system, using the same buffer at a flow rate of 0.5 ml/min. The following molecular weight standards were used at concentrations of 2 to 5 mg/ml; cytocrome c (Mₙ, 12,400), ovalbumin (Mₙ, 43,000), bovine albumin (Mₙ, 67,000), and bovine gamma globulin (Mₙ, 150,000).

Estimation of the molecular weight of the purified enzyme by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed by the method of Laemmli (19). A mixture of approximately 0.5 µg of purified enzyme, 62.5 mM Tris-HCl (pH 6.8), 1% SDS, 1% 2-mercaptoethanol, 0.001% bromophenol blue, and 10% glycerol was heated to 100°C for 3 min and then electrophoresed on a 10% polyacrylamide gel. The gel was stained with 0.2% Coomassie brilliant blue R by the conventional method. Pharmacia high-molecular-weight standards (Mₙ, 55,000, 76,000, 116,000, 170,000, and 212,000) were used as standards.

The N-terminal amino acid sequence of the purified enzyme was determined on a protein sequencer (model 476A; Applied Biosystems, Foster City, Calif.).

Effects of pH, temperature, and metal ions on enzyme activity. The effect of pH was examined by using azocoll under the standard assay conditions, except for the buffer. The buffers used were 0.2 M sodium phosphate (pH 4.5 to 8.5) or 0.2 M sodium borate (pH 6.8 to 9) containing 0.15 M NaCl. The effect of temperature was also examined under the standard assay conditions by varying the temperature (20 to 57°C). For examination of metal ion specificity, divalent metal ions present in the enzyme solution and the reaction mixture were removed prior to the assay. To the purified enzyme in 100 mM Tris-HCl (pH 7.0)-5 mM CaCl₂ was added EDTA at a final concentration of 10 mM (5 mM excess). Removal of the metal-chelator complex with Sephadex G-10 was carried out by the spin-column method of Salvesen and Nagase (30). Divalent metal ions which might be present in the Pz peptide assay mixture were removed by Na⁺-charged chelating Sepharose (Pharmacia) column chromatography. CaCl₂, MgCl₂, and ZnCl₂ were added to the metal ion-free assay mixture at
various concentrations. The reaction was initiated by adding the metal-free enzyme solution to the mixture.

**DNA manipulations.** Restriction endonucleases were purchased from Takara Shuzo (Kyoto, Japan), Toyobo (Osaka, Japan), and New England Biolabs (Beverly, Mass.). Klenow enzyme and the DNA ligation kit were products of Takara Shuzo. All recombinant DNA procedures were carried out as described by Maniatis et al. (22).

**Construction of genomic library of C. perfringens.** C. perfringens cells were harvested from a 1-liter culture by centrifugation at 8,000 × g for 20 min at 4°C. The cell pellet was resuspended in 60 ml of 100 mM Tris-HCl (pH 8.0)–10 mM EDTA. A 1/10 volume of a 2-mg/ml lysozyme solution was added; this was followed by incubation at 37°C for 5 min. A 1/10 volume of 20% SDS was then added to the suspension, and cells were lysed by further incubation at 37°C for 5 min. An equal volume of ice-cold phenol saturated with 100 mM Tris-HCl (pH 8.0) was added to the sample, and the mixture was gently shaken at 4°C for 30 min. The sample was extracted once with phenol-chloroform at 37°C and once with chloroform. The aqueous phase was recovered and treated with RNase A (10 μg/ml; Sigma) for 30 min at 37°C, and then the protein was removed by phenol-chloroform extractions as described above. The DNA was precipitated with 2 volumes of ethanol and dissolved in 10 mM Tris-HCl (pH 8.0)–1 mM EDTA.

Nine hundred micrograms of the above-described DNA was digested with 60 U of MboI at 37°C for 2 min, and then the DNA was precipitated with ethanol and DNA fragments of 9 to 20 kb were isolated by sucrose density gradient centrifugation (22). The sticky ends of both the size-fractionated insert DNA and Xhol-digested λ2001 were partially filled in (to prevent self-ligation) by treating each with the Klenow fragment in the presence of dATP and dGTP and of dCTP and dTTP, respectively. The DNA were then ligated by T4 DNA ligase and packaged in vitro.

**Preparation of anticolonagense antiserum and cloning of the colA gene.** Collagenase was purified from the culture supernatant of C. perfringens type A NCTC 8237 as described by Kameyama and Akama (14). Rabbits were immunized with this purified collagenase by a single subcutaneous injection in complete adjuvant followed by a booster injection in incomplete adjuvant over 42 days. The resulting antiserum was treated with the E. coli extract of the ProtBlot kit (Promega, Cleveland, Ohio) to reduce the background of immunoscreening.

Plaques of recombinant phage were grown on E. coli P2392 host cells and lifted on nitrocellulose membranes (BA85; Schleicher & Schuell, Inc., Keene, N.H.). The filters were probed with the anticolonagense rabbit antiserum by using the ProtBlot kit (Promega) according to the supplier’s instructions.

**Nucleotide sequencing.** DNA was sequenced by the dyeoxy chain termination method (31), using modified T7 DNA polymerase (Sequenase; United States Biochemicals, Cleveland, Ohio). Deoxyadenosine 5'-α-[35S]triorthophosphate (>1,000 Ci/mmol) was purchased from New England Nuclear (Wilmington, Del.). Plasmid DNAs were prepared as described by Birnboim and Doly (5), purified by CsCl density gradient centrifugation with ethidium bromide (33), denatured, and neutralized for use as sequencing templates. Synthetic oligonucleotides for the sequencing primers were prepared on a DNA synthesizer (model 381; Applied Biosystems).

**Southern hybridization.** A nonradioactive DNA-labeling and detection kit (Genius; Boehringer Mannheim Biochemicals, Indianapolis, Ind.) was used for hybridization experiments. To prepare a DNA probe, pKY3135 plasmid DNA was completely digested with XbaI and Smal, and the 1.61-kb insert (see Fig. 3) was isolated by agarose gel electrophoresis and electroelution. The DNA fragment was purified by using the GeneClean kit (Bio 101 Inc., La Jolla, Calif.), and approximately 1 μg was labeled with digoxigenin-11-dUTP by random priming as described by the supplier (Boehringer Mannheim).

**RESULTS**

Gelatinolytic activities in the culture supernatant. Zymography revealed the presence of multiple gelatinases of different sizes in the culture supernatant of C. perfringens (Fig. 1). In the mid-logarithmic phase (2 h), gelatinolytic activity was detected as one major band corresponding to a molecular mass of ~120 kDa, with three very faint bands at ~110, ~102, and ~80 kDa. While the intensities of these three minor bands increased with incubation time, the 120-kDa band remained the major band. However, during stationary phase (8 h), the intensities of all four bands decreased.

**Purification and characterization of the 120-kDa enzyme.** The major 120-kDa enzyme was purified to homogeneity from the C. perfringens culture supernatant (Fig. 2A). The molecular mass of the collagenase was estimated to be 120 kDa by SDS-PAGE. The same value was obtained from analysis by gel filtration on Superose 12, indicating that the 120-kDa collagenase is in a monomeric form. From 10 liters of the culture supernatant, 1.6 mg of purified enzyme was obtained. To examine the activity of this enzyme on a type I collagen fibril, zymography using this substrate was carried out. As shown in Fig. 2B, the hydrolytic activity was visible as a single band, and the location of this band coincided with that of the 120-kDa enzyme after staining of the polyacrylamide gel. From these results, we conclude that the type I collagen fibril was cleaved only by the 120-kDa enzyme in the absence of other proteases and that the purified enzyme was indeed a collagenase.

The specific activity of the purified enzyme was 12.3 U/mg of protein when assayed for azocoll hydrolytic activity under the standard assay conditions and 2.25 μmol/min/mg of protein.
when assayed for Pz peptide. The effect of pH on azocoll hydrolytic activity was assayed in the borate and phosphate buffers. Enzyme activity was detectable over a broad range, from pH 4.5 to 9. The maximal enzyme activity in the borate buffer was attained at pH 7.2 and that in the phosphate buffer was at pH 7.0, the former being 1.7-fold higher than the latter. The optimal temperature for azocoll hydrolytic activity was 42°C, and at higher temperatures the activity fell, probably because of thermal instability of the enzyme. Like the collagenase of C. histolyticum, the 120-kDa collagenase appears to require zinc for hydrolytic activity, since the addition of 1 mM 1,10-phenanthroline to the assay mixture completely abolished the activity against azocoll, even in the presence of excess CaCl₂ (10 mM). In order to confirm this, divalent metal ions present in the enzyme and substrate solutions were removed, and metal ions such as Ca²⁺, Mg²⁺, and Zn²⁺ were added to the reaction mixture. No appreciable activity was observed when divalent metal ions were not added. The addition of Zn²⁺ (30 μM) resulted in reactivation of the enzyme activity to the highest level. Ca²⁺ also reactivated it to a similar extent, but a much higher Ca²⁺ concentration (≥1 mM) was required in order for activity to reach the maximal level, which corresponded to 90% of the maximal level attained by addition of Zn²⁺. The maximal reactivation attained by addition of Mg²⁺ was only 30% of that attained by addition of Zn²⁺. High concentrations (≥100 μM) of Zn²⁺ inhibited the enzyme activity, which is consistent with the finding reported for zinc-dependent metalloproteases (30). These results clearly indicate that the 120-kDa collagenase is a zinc protease. The N-terminal amino acid sequence of the purified enzyme was determined to be Ala-Arg-Asn-Asn-Lys-Ile-Tyr-Thr-Phe-Asp-Glu-Leu-Asn-Arg-Met-Asn-Tyr-Ser-Asp-Leu-Val-Glu-Leu-Ile-Lys-Thr-Ile by Edman degradation. **Cloning of the colA gene.** A genomic library of C. perfringens DNA constructed in λ2001 was immunologically screened for a clone expressing a collagenase gene. The antiserum raised against collagenase purified from a C. perfringens type A strain was shown to cross-react with the collagenase from the type C strain by Western blotting (immunoblotting) (data not shown), and was used for the screening. Approximately 4,000 recombinant phage plaques were examined, and approximately 60 plaques were reactive against the antiserum. Four of these immunologically positive plaques were randomly selected and purified four times by using E. coli P2392 host cells, which inhibits the growth of nonrecombinant phage. Lysates of the purified recombinant phages were examined for collagenolytic activity by using azocoll and Pz peptide. Three recombinants showed strong activities on both of the substrates, and one showed weak activities. λcol3, one of the former clones, was used for further analyses. Restriction analysis showed that the λcol3 recombinant phage contained a 12.9-kb MboI insert. The insert was excised from the recombinant with XbaI and ligated into the XbaI site of pUC18. Two different groups of recombinant plasmids were obtained, one carrying the entire 12.9-kb insert (Fig. 3, A) SDS-PAGE. Lane 1, purified collagenase (0.5 μg); lane 2, molecular mass markers. Numbers on the right are molecular masses (in kilodaltons) of the markers. (B) Zymography. Four micrograms of the purified collagenase was loaded; this was followed by zymography using soluble collagen (type I) as the substrate. **FIG. 1.** Zymogram of gelatinolytic activities in culture supernatants of C. perfringens. Portions of the C. perfringens culture were removed after 2 (lane 1), 4 (lane 2), 6 (lane 3), and 8 (lane 4) h. Fifteen microliters of each culture supernatant was applied to an SDS-polyacrylamide gel copolymerized with gelatin. Numbers on the left are molecular masses (in kilodaltons) of the markers. **FIG. 2.** Analysis of the purified collagenase by SDS-PAGE and zymography. (A) SDS-PAGE. Lane 1, purified collagenase (0.5 μg); lane 2, molecular mass markers. Numbers on the right are molecular masses (in kilodaltons) of the markers. (B) Zymography. Four micrograms of the purified collagenase was loaded; this was followed by zymography using soluble collagen (type I) as the substrate. **FIG. 3.** Partial restriction map of C. perfringens DNA fragments encoding collagenase and subclones used for nucleotide sequencing. A partial restriction map of the pKY313 insert is shown in the middle. The collagenase activities of two subclones (pKY3134 and pKY3132) are indicated by (+) and (−). Various deletions were introduced from the left end of the pKY3133 insert, resulting in loss of the enzymatic activity when the deletion size exceeds approximately 2.5 kb (arrow). Three subclones were used to construct sequencing templates. Deletions of various sizes were introduced from the end of each insert indicated by the arrowheads.
FIG. 4. DNA sequence of the C. perfringens collagenase gene (colA). An inverted repeat close to the termination codon is indicated by arrows. Some of the restriction sites are also shown. A deduced amino acid sequence which is the same as that of a synthetic peptide substrate for collagenases is indicated by a black background. The N-terminal amino acid sequence of the collagenase is shaded.

pKY313) and the other carrying a 2.7-kb XbaI fragment (Fig. 3, pKY361). The latter fragment was placed between the XbaI site of polynucleotide sequence of the vector and the XbaI site within the 12.9-kb insert by partial physical mapping. Two subclones carrying the leftmost 7.6-kb XbaI (within the vector polynucleotide site)-PstI fragment and the 4.9-kb XbaI (within the 12.9-kb insert)-PstI fragment were constructed from pKY313 and were designated pKY313 and pKY3134, respectively. While E. coli transformed with pKY313 showed hydrolytic activity against azocoll, that transformed with pKY3134 showed no activity. Nested deletions were introduced from the left end of pKY3132 to localize the left end of the colA gene. When the deletion exceeded approximately 2.5 kb from the left end of the insert, azocoll-hydrolyzing activity was lost, indicating that the left end of the colA gene is located 2.5 kb away from the left end of pKY3132. The right end of the colA gene was not determined but could be placed at about 0.2 kb to the left of the HindIII site on the basis of a maximum nucleotide length of the colA gene encoding the 120-kDa polypeptide.

Nucleotide sequencing of the colA gene. Three plasmids, pKY361, pKY3135, and pKY3136, and their derivatives produced by nested deletions of the three plasmid (Fig. 3) were sequenced by using the strategy outlined in Fig. 3. The nucleotide sequences of some regions were determined by using synthetic oligonucleotides in order to close gaps in the sequence or to confirm ambiguous nucleotides. The sequences thus obtained were aligned by their overlaps to form a single contig (Fig. 4). The open reading frame encoding the colA gene extends from 0.43 kb to the left of the XbaI site to 0.46 kb to the left of the HindIII site, as expected. This open reading frame encodes a protein with a molecular mass of 125,966 Da. The N-terminal amino acid sequence of the purified 120-kDa collagenase is the same as that predicted from the nucleotide sequence starting at nucleotide 629, and the open reading frame continues to nucleotide 3685, which is a TAA termination codon (Fig. 4). The molecular mass of the protein predicted from this reading frame is 116,339 Da, which is in good agreement with that determined by SDS-PAGE (120 kDa) for the purified collagenase of C. perfringens. The ATG initiation codon at nucleotide 371 is preceded by a consensus ribosome binding sequence site (AGGAGG) with a 9-bp gap, which is complementary to the 3' end of C. perfringens 16S rRNA (10). Since no other ATG or GTG initiation codons are present in this region, it seems certain that the collagenase is initiated at the ATG at nucleotide 371.

Southern hybridization using a colA gene probe. In order to determine whether the colA gene is duplicated, a Southern hybridization was carried out with a 1.61-kb XbaI-BglII fragment as a probe (Fig. 3). Chromosomal DNA from C. perfringens was digested with BglII, HindIII, or PvuII and separated on a 0.8% agarose gel. For the BglII and PvuII digests, hybridized bands were 1.8 and 3.0 kb long, respectively (Fig. 5), which
coincided with a 1.766-bp BglII fragment (nucleotides 642 to 2407) and a 3.043-bp PvuII fragment (nucleotides 486 to 3528). A 5.4-kb fragment hybridized in the HincII digest, in good agreement with the predicted size of the HincII fragment estimated from the restriction map. In all cases only one band was detected for a given digestion, suggesting that the 120-kDa collagenase is encoded by a single colA gene.

DISCUSSION

Zymography using gelatin revealed the presence of a major 120-kDa gelatinolytic enzyme and other minor ones in the culture supernatant of the C. perfringens type C strain. During the initial phases of this work, it became clear that a gelatinolytic enzyme fraction partially purified by the method described for purification of C. perfringens collagenase (14) contained predominantly enzymes with molecular masses ranging from 110 to 80 kDa and only a trace amount of the 120-kDa enzyme. We therefore examined both the culture conditions and the purification procedure with the goal of successfully obtaining sufficient quantities of the 120-kDa enzyme. We identified two modifications which were critical to isolating the 120-kDa enzyme. First, we found that large amounts of the 120-kDa enzyme were present in the culture supernatants even after prolonged incubation if the cells were grown at 25 rather than 37°C. In addition, for certain Zn proteases, Ca²⁺ has been shown to inhibit their autocatalysis (23), and the C. histolyticum collagenase has been successfully purified by the addition of Ca²⁺ to the buffers used (28). We therefore carried out all the purification steps in the presence of 5 mM CaCl₂. The combination of these improvements allowed purification of the 120-kDa enzyme, which cleaved type I collagen filament and was identified as a collagenase. The discrepancy between the molecular mass of 120 kDa that we obtained and the 80-kDa value reported by other groups (14, 39) for the collagenase of a C. perfringens type A strain does not seem to arise from a difference in the type of strain. In our hands, zymography indicated the presence of both 120- and 80-kDa gelatinolytic enzymes in the culture supernatants of both the type A and type C strains (data not shown). One possible explanation is that the 120-kDa enzyme may be proteolyzed to give the 80-kDa enzyme.

Inspection of the N-terminal amino acid sequence starting with a potential translational initiation codon (ATG) reveals a typical signal peptide, which consists of a positively charged stretch (MKKNLRGETTLKLLVQR), a hydrophobic core (WSAFITLAAFLF), and a polar region (NSSFKFVAA). However, Ala-87, the N-terminal amino acid of the mature enzyme, is separated by 46 amino acid residues from the putative signal peptide. Within these amino acid residues occurs a unique amino acid sequence (PVGIP), the same sequence as that of a synthetic peptide substrate for bacterial collagenases. The C. histolyticum collagenase is known to cleave a peptide bond between leucine and glycine residues in the amino acid sequence indicated above, which is probably the case for the 120-kDa C. perfringens collagenase. This may suggest that the 120-kDa enzyme can do self-processing, triggering subsequent removal of the following octapeptide by another protease.

While vertebrate collagenases cleave native type I collagen only at a specific site and other proteases are necessary for complete degradation, bacterial collagenases can degrade collagen in the absence of other proteases. The latter enzymes seem to be related to protease, and one might expect sequence and structural similarities between bacterial collagenases and proteases. We therefore investigated this possibility by searching for proteins with amino acid sequences similar to that of the C. perfringens collagenase, using the BLAST algorithm. This used an ungapped alignment algorithm to compare a query with all protein sequences deposited in the data bases. This search identified two proteins, Vibrio alginolyticus collagenase (34) and Achromobacter lyticus protease I (26), which show sequence similarity to the C. perfringens collagenase. The former showed higher similarity (Poisson P value, 2.1 × 10⁻¹⁵) than the latter (Poisson P value, 1.8 × 10⁻⁵). To test their statistical significance, a Monte-Carlo test was performed by using the Lipman-Pearson alignment algorithm with the Dayhoff similarity scoring (27). In this analysis, the best alignment of the two original sequences was compared with a control obtained from the best alignments between 100 different randomized sequences. The alignment between the V. alginolyticus and the C. perfringens collagenases showed a score that was 14.8 standard deviations higher than the control, and the alignment between A. lyticus protease I and the C. perfringens collagenase was 10.3 standard deviations higher than the control. It should be noted that the V. alginolyticus collagenase and A. lyticus protease I showed the highest similarity (19.3 standard deviations) among the three enzymes.

The alignment of the three homologous enzymes is shown in Fig. 6. Note that regions of sequence similarity are not highlighted but are scattered over the whole sequences. A consensus zinc-binding sequence (His-Glu-Xaa-Xaa-His) regarded as the catalytic center of eucaryotic zinc collagenases (20) is also conserved in both the C. perfringens collagenase.
premature proteins, The asterisks indicate the shaded.

FIG. 6. Amino acid sequence alignment of C. perfringens collagenase (C), V. alginolyticus collagenase (V) (34), and A. lyticus protease I (A) (26). Identical amino acids are indicated by a black background. The asterisks indicate the consensus sequence for zinc collagenases, HEXXH. The catalytic triad of the Achromobacter serine protease is shaded. The numbering is according to the amino acid sequence of the premature proteins, and the mature proteins for C, V, and A start at amino acid positions 87, 77, and 206, respectively.

C. PERFRINGENS COLLAGENASE GENE

(positions 502 to 506) and the V. alginolyticus collagenase (positions 477 to 481). On the other hand, this putative zinc-binding sequence is absent in A. lyticus protease I, which is known to be a serine protease. The catalytic triad (His-262, Asp-318, and Ser-399) of the serine protease is present in A. lyticus protease I but not in the other collagenases. These results may imply that the three enzymes, which differ in function and origin, are evolutionarily related. They might have evolved from a common ancestor protein by using different strategies so that each could increase its catalytic efficiency on the specific substrate in each bacterium. The collagenase of V. alginolyticus, a gram-negative bacterium, is more homologous to protease I of A. lyticus, also a gram-negative bacterium, than to the collagenase of C. perfringens, a gram-positive bacterium. Furthermore, members of the genus Clostridium are considered to be among the most ancient forms of life known (9, 38). The diversity of these enzymes is correlated with the phylogenetic relatedness of the three bacteria. Accumulation of more sequence data for bacterial collagenases and proteases would enable a more precise analysis of their homology and provide valuable insight into the molecular evolution of these proteins.

ACKNOWLEDGMENTS

We thank Jon D. Stewart (Department of Chemistry, The Pennsylvania State University, University Park) for invaluable discussions and assistance in preparing the manuscript. We also thank Akihisa Takanizawa and Mitsuo Takagi (Kanonji Institute, The Research Foundation for Microbial Diseases of Osaka University) for preparing synthetic oligonucleotides and concentrating protein samples. We are indebted to Yoshihiro Shionoya (Applied Biosystems Japan, Inc.) for the extensive protein sequencing.

This work was partly supported by a grant in aid for Scientific Research from the Ministry of Education and Culture, Japan.

REFERENCES