Cloning and Sequencing of a Gene Encoding a Novel Extracellular Neutral Proteinase from *Streptomyces* sp. Strain C5 and Expression of the Gene in *Streptomyces lividans* 1326

JAY S. LAMPEL,^{1,2} JAYANT S. APHALE,¹ KEITH A. LAMPEL,³ AND WILLIAM R. STROHL^{1*}

Department of Microbiology, The Ohio State University, Columbus, Ohio 43210¹; Crop Genetics International, Hanover, Maryland 21076²; and Food and Drug Administration, HFF-235, Washington, D.C. 20204³

Received 18 October 1991/Accepted 19 February 1992

The gene encoding a novel milk protein-hydrolyzing proteinase was cloned on a 6.56-kb *Sst*I fragment from *Streptomyces* sp. strain C5 genomic DNA into *Streptomyces lividans* 1326 by using the plasmid vector pIJ702. The gene encoding the small neutral proteinase (*snpA*) was located within a 2.6-kb *Bam*HI-*Sst*I restriction fragment that was partially sequenced. The molecular mass of the deduced amino acid sequence of the mature protein was determined to be 15,740, which corresponds very closely with the relative molecular mass of the purified protein (15,500) determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The N-terminal amino acid sequence of the purified neutral proteinase was determined, and the DNA encoding this sequence was found to be located within the sequenced DNA. The deduced amino acid sequence contains a conserved zinc binding site, although secondary ligand binding and active sites typical of thermolysinlike metalloproteinases are absent. The combination of its small size, deduced amino acid sequence, and substrate and inhibition profile indicate that *snpA* encodes a novel neutral proteinase.

Streptomycetes are known to contain a wide variety of both serine and neutral proteinases with varied substrate specificities (10, 38, 42). Most of the investigations of streptomycete proteinases, however, have centered around the structures and cleavage mechanisms of the various proteinases of *Streptomyces griseus* that constitute pronase (38). Little is actually known about the regulation of proteinase expression in streptomycetes, largely because only a few genes encoding streptomycete proteinases have been cloned and sequenced. Henderson et al. (22) found that colonies of *Streptomyces lividans* 1326 did not produce significant clearing zones when grown on milk-agar (MA) plates, and Henderson et al. used this property to assist their cloning of two genes encoding serine proteinases (SGPA and SGPB) from *S. griseus* in *S. lividans*. Recently a gene from *Streptomyces cacaoi* encoding a neutral proteinase with a deduced M_r of 35,000 was isolated and sequenced (11). The deduced amino acid sequence of the *S. cacaoi* neutral proteinase indicated that it was significantly different from the thermolysinlike neutral proteinases that are common to *Bacillus* strains.

We found previously that *Streptomyces* sp. strain C5, a mutagenized strain isolated as an overproducer of anthracycline antibiotics, produces negligible extracellular proteinase activity against azocasein but strong extracellular proteolytic activity against the proteins in Carnation dry milk (20, 21). In the present study, we used the respective differences in proteolysis profiles of *Streptomyces* sp. strain C5 and *S. lividans* 1326 to clone the gene (*snpA*) that encodes a milk protein-degrading neutral proteinase (SnpA) from the former strain into the latter. We also purified the recombinant protease to electrophoretic homogeneity and found the enzyme to be a neutral proteinase. The characteristics of this proteinase, however, including its size, substrate specificities, sensitivity to inhibitors, and deduced amino acid sequence, are remarkably different from those of other known

neutral proteinases. Thus, we propose that *snpA* encodes a novel neutral proteinase that may comprise a new subclass of neutral proteinases.

MATERIALS AND METHODS

Bacterial strains, maintenance, and cultivation conditions.

The cultures used in this study are described in Table 1. These strains are maintained in the laboratory as described previously (12).

R2YE, R5, and YEME media, described by Hopwood et al. (25), were typically used for normal growth of the strains. Generalized milk protein-hydrolyzing activity was detected by screening for zones of hydrolysis around colonies growing on MA plates, which consisted of the following components (in grams per liter): morpholinopropanesulfonic acid buffer (pH 7.2), 1.49; nutrient broth (Difco, Detroit, Mich.), 1.0; KH_2PO_4 , 0.5; K_2HPO_4 , 0.5; fructose, 0.25; $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.735; and agar, 15. The pH was adjusted to 7.2 with 1 N sodium hydroxide before autoclaving. Powdered milk (Carnation Instant Dry Milk) was autoclaved separately as a 20% (wt/vol) solution in distilled H_2O and added to the medium after partial cooling to make a final concentration of 20 g/liter. Thiostrepton was typically added to the solid medium (MAT plates) at 50 $\mu\text{g/ml}$ for selective pressure when recombinant strains were grown.

Cloning procedures and genetic manipulations. Procedures for formation, transformation, and regeneration of streptomycete protoplasts were carried out as described by Hopwood et al. (25). Procedures used for the preparation of *Streptomyces* plasmid and chromosomal DNA also have been described (30). Restriction endonuclease mapping of the plasmids was carried out after single and double endonuclease cleavage by standard procedures.

To clone the gene conferring milk protein-hydrolyzing activity from *Streptomyces* sp. strain C5, DNA isolated from liquid cultures of *Streptomyces* sp. strain C5 grown in YEME medium was completely digested with *Sst*I and

* Corresponding author.

TABLE 1. Bacterial strains and plasmids used in this study^a

Strain or plasmid	Relevant characteristics	Source or reference
Bacteria		
<i>Streptomyces</i> sp. strain C5	Milk protein-hydrolyzing protease producer	FCRC (20, 21)
<i>S. galilaeus</i> 31133	Lacks milk protein-hydrolyzing proteolytic activity	ATCC
<i>S. peucetius</i> 29050	Lacks milk protein-hydrolyzing proteolytic activity	ATCC
<i>S. lividans</i> 1326	Minor milk protein-hydrolyzing proteolytic activity	D. A. Hopwood
<i>E. coli</i> DH5αF'	φ80dlacZΔ(lacZYA-argF)U169 <i>deoR recA1 endA1 hsdR17</i> (r _K ⁻ m _K ⁺) <i>supE44</i> λ ⁻ <i>thi-1 gyrA96 relA1</i>	BRL
Plasmids		
pIJ702	5.686 kbp, derivative of pIJ350, HC, Thio ^r Mel ⁺	D. A. Hopwood
pUC19	2.686 kbp, Amp ^r	J. N. Reeve
pGEM-7ZF(+)	3.00 kbp, Amp ^r	Promega
pKC505	18.7 kbp, derivative of SCP2*, <i>E. coli-Streptomyces</i> shuttle vector, LC, Apr ^r	
pANT21	12.25 kbp, pIJ702 with 6.56-kbp <i>Sst</i> I cloned DNA fragment from <i>Streptomyces</i> sp. strain C5, confers strong milk protein-hydrolyzing activity on <i>S. lividans</i> ^b	This work
pANT22	23.5 kbp, pKC505 with 4.8-kb <i>Bam</i> HI fragment from pANT21 (carrying part of pIJ702, which includes <i>melC1</i> , as well as 3.25 kb of <i>Streptomyces</i> sp. strain C5 DNA), no proteolytic activity conferred	This work
pANT23	26.15 kbp, pKC505 with 7.45-kbp <i>Bam</i> HI fragment from pANT21 (carrying part of pIJ702, including the <i>tsr</i> gene, as well as 3.31 kbp of <i>Streptomyces</i> sp. strain C5 DNA), confers milk protein-hydrolyzing activity on <i>S. lividans</i> ^b	This work
pANT25	9.25 kbp, pUC19 with 6.56-kbp <i>Sst</i> I fragment from pANT21	This work
pANT42	8.76 kbp, pIJ702 with 3.31-kbp <i>Bam</i> HI- <i>Sst</i> I fragment from pANT21, confers strong milk protein-hydrolyzing activity on <i>S. lividans</i> ^b	This work
pANT54	3.9 kbp, pUC19 with 1.2-kbp <i>Kpn</i> I- <i>Sph</i> I fragment from pANT42	This work
pANT55	6.9 kbp; pIJ702 with 1.2-kbp <i>Sst</i> I (from polylinker of pUC19)- <i>Sph</i> I fragment from pANT54, confers moderate but slow milk-protein hydrolyzing activity on <i>S. lividans</i> ^b	This work
pANT60	6.79 kbp; pIJ702 with 1.1-kbp <i>Sph</i> I- <i>Sst</i> I fragment from pANT42, confers strong milk protein-hydrolyzing activity on <i>S. lividans</i> ^b	This work
pANT62	3.79 kbp, pUC19 plus 1.1-kbp <i>Sph</i> I- <i>Sst</i> I fragment from pANT42	This work
pANT63	4.1 kbp, pGEM-7ZF+ plus 1.1-kbp <i>Sph</i> I- <i>Sst</i> I fragment from pANT42	This work

^a Abbreviations: ATCC, American Type Culture Collection; BRL, Bethesda Research Laboratories; FCRC, Frederick Cancer Research Center; LC, low-copy-number plasmid; HC, high-copy-number plasmid; Thio^r, thiostrepton resistance; Amp^r, ampicillin resistance; Apr^r, apramycin resistance; Mel⁺, production of melanin.

^b Also confers milk protein-hydrolyzing activity on *S. galilaeus*.

ligated with T4 DNA ligase into the *Sst*I site of the high-copy-number streptomycete vector pIJ702 (27) (Fig. 1). The ligated plasmid-plus-insert DNA was introduced by transformation into protoplasts of *S. lividans* 1326 by standard methods described previously (25). After 16 h of growth at 30°C, the protoplasts were challenged with 20 μg of thiostrepton per ml and regenerated as described previously (25). The transformants were allowed to sporulate and then were replica plated with sterile velvet pads to plates containing MA plus 50 μg of thiostrepton per ml (MAT plates). After 24 to 36 h of incubation, the plates were screened for colonies producing cleared zones of hydrolysis. Milk protein-hydrolyzing transformants were picked over to fresh MAT plates to verify the proteinase phenotype. Other subclones were constructed from the original plasmid (pANT21) conferring milk protein-hydrolyzing activity by similar procedures (Fig. 1, Table 1).

Small-scale *Escherichia coli* plasmid preparations suitable for restriction and ligation were performed as described by Birnboim and Doly (5). For large-scale preparations (500-ml cultures), plasmid DNA was further purified by CsCl density gradient centrifugation. Ethidium bromide was extracted from plasmid preparations with water-saturated butanol and then dialyzed against a buffer composed of 10 mM Tris-HCl and 1.0 mM EDTA (pH 8.0). Overnight cultures of *E. coli* were transformed by standard techniques as described by Maniatis et al. (34). The cells were incubated for 60 min at

37°C before they were plated to permit expression of antibiotic resistance determinants. When pKC505 (44) or derivatives of it were used, the plates were incubated at 31°C for expression of the apramycin resistance gene.

DNA sequencing and sequence analysis. Double-stranded DNA of plasmids pANT54, pANT60, pANT62, and pANT63 was sequenced by the dideoxynucleoside termination method of Sanger et al. (48) with a Sequenase kit from U.S. Biochemical Corp. (Cleveland, Ohio), [α-thio-³⁵S]dATP (>1,000 Ci/mmol; Amersham Corp., Arlington Heights, Ill.), and double-stranded templates. Universal and reverse primers were used to obtain the initial sequences within the inserts, and then specific primers for the sequences within the inserts were generated. Conditions for DNA sequencing are described in the brochure accompanying the Sequenase enzyme (U.S. Biochemical). Denaturing polyacrylamide gel electrophoresis (PAGE; 6% [wt/vol] polyacrylamide) at 60 W was used to separate the reaction products. The gels were exposed to X-ray film (Kodak Omat-AR) overnight. All of the DNA shown was sequenced in both directions.

The sequenced DNA was subjected to a FRAME program (4) analysis to determine the direction of transcription, reading frame, and size of the open reading frame. The DNA and deduced amino acid sequences were analyzed by using the sequence analysis software package of the Genetics Computer Group, University of Wisconsin, Madison, Wis. (16), and the microcomputer programs Clone Manager (Sci-

FIG. 1. (A) Cloning and subcloning of the neutral proteinase gene (*snpA*) from *Streptomyces* sp. strain C5 into *S. lividans* 1326. *SstI* DNA fragments from *Streptomyces* sp. strain C5 were ligated into pIJ702 to make pANT21, which was isolated from a milk-hydrolyzing colony of recombinant *S. lividans* 1326. Plasmid pANT42 was constructed by subcloning the 3.31-kb *BamHI-SstI* fragment from pANT21 into *BglII-SstI*-digested pIJ702 (the 236-bp *BglII-SstI* DNA fragment from pIJ702 was lost in the construction of pANT42). (B) Restriction map of the insert in pANT42, showing the location of the inserts in pANT55 [pIJ702 containing the 1.2-kb *SstI-SphI* fragment from pANT54; the *SstI* site denoted as (*SstI*)* is from the polylinker of pUC19] and pANT60 (pIJ702 containing the 1.1 kb *SphI-SstI* fragment) and the location of *snpA* within these sequences. The open part of the thick arrow (▬) represents the sequences encoding the potential pre-pro portion of the protein, whereas the blackened portion represents the location of the sequences encoding the mature protein.

entific & Educational Software, Inc., State Line, Pa.), Genepro (Riverside Scientific, Seattle, Wash.), and DNA-STAR (Madison, Wis.). Similarity of the deduced amino acid sequence from *snpA* with other amino acid sequences was determined with the TFASTA program (41).

Milk hydrolysis assay and protein determination. The milk hydrolysis activity of the proteinase encoded by *Streptomyces* sp. strain C5 *snpA* was determined in a spectrophotometric assay with 750 μ l of 50 mM *N*,2-hydroxyethylpiperazine-*N'*,2-ethanesulfonic acid (HEPES) buffer (pH 7.0) and 200 μ l of an autoclave-sterilized solution containing 20 mg of dry milk (Carnation) per ml in the same buffer. The substrate solution and buffer were added to a cuvette and prewarmed to 37°C for 2 min. To start the reaction, 50 μ l of the enzyme preparation was prewarmed separately at 37°C for 5 min and added to the cuvette containing the substrate-buffer solution to give a final volume of 1,000 μ l. The decrease in A_{580} was continuously monitored for 2 to 5 min at 37°C by using a dual-beam spectrophotometer (Bausch and Lomb Spectronic 2000) equipped with a temperature control cuvette holder attached to a Haake recirculating water bath. One unit of milk-hydrolyzing enzyme activity is defined as the activity required to decrease the A_{580} of the milk solution by 0.001 absorbance unit per min. Specific activity is defined as units per milligram of protein. Protein was determined by the method of Bradford (6) with a commercial protein assay reagent (Bio-Rad, Richmond, Calif.) and bovine serum albumin as the standard.

In preliminary experiments it was necessary to determine whether the activity was due to combined activities of multiple proteinases (including potential serine proteinases). Thus, the enzyme solution was pretreated with 3 mM diisopropylfluorophosphate (DIFP) for 60 min at the ambient temperature before the milk hydrolysis activity was assayed. Since the proteinase encoded by *snpA* is not sensitive to DIFP, this inhibitor was typically added to assays of the proteinase from crude culture supernatants.

Purification of the milk-hydrolyzing proteinase from *S. lividans*(pANT42). The neutral proteinase (SnpA) was purified from 10-liter fermentation cultures of *S. lividans* (pANT42) and from 500-ml shake flask cultures of *Streptomyces galilaeus*(pANT42). For shake-flask cultures, spores were scraped off of MAT solid medium and used to inoculate a 250-ml flask containing 50 ml of milk broth containing 15 μ g of thioestrepton per ml (MBT [MAT lacking agar]) and a coiled spring for mycelial dispersion (12). After incubation for 24 h at 30°C with rotary shaking (250 rpm), the entire culture (primary seed culture) was used to inoculate a 2-liter flask containing either 450 ml [for *S. galilaeus*(pANT42)] or 950 ml of MBT and a coiled spring. After incubation for 36 h, the entire contents of the *S. lividans*(pANT42) secondary seed culture were used to inoculate a 14-liter stirred-tank fermentor (New Brunswick Scientific) containing 9 liters of MBT (ca. 11% [vol/vol] inoculum). Five milliliters of undiluted antifoaming agent (Mazu DF-60P; Mazer Chemical Co., Gurnee, Ill.) was added to the medium before autoclaving. Fermentation conditions were as follows: agitation, 400 rpm; air flow, 1 volume of air per volume of medium per min; temperature, 30°C; pH, initially 7.2 and not controlled. Mycelia were harvested after 24 h of fermentation by cross-flow filtration with a Millipore Pelicon cassette system equipped with 0.45- μ m-pore-size membranes (2 ft² [ca. 1,858 cm²] of membrane area).

All protein purification steps were carried out at 4°C unless otherwise stated. The mycelium-free broth was concentrated by using a Millipore filtration unit equipped with a

membrane with a molecular weight cutoff of 10,000 (5 ft² [ca. 4,645 cm²] of membrane area). The retentate, which contained all of the milk-hydrolyzing proteinase activity, was used for further purification and stored at -20°C when not in use. Acetone was chilled to -70°C and added slowly, with gentle stirring, to the enzyme preparation on an ethanol-ice (9:1, wt/vol) bath to reach a final saturation of 55% (vol/vol). The resulting precipitate was recovered by centrifugation for 20 min at 16,000 $\times$ g. The supernatant from this step, which contained less than 1% of the milk-hydrolyzing activity, was discarded. Residual acetone was removed from the pellet under a vacuum, and then the enzyme was resuspended in a minimal volume of 50 mM HEPES buffer (pH 6.5), placed into dialysis tubing (average flat width of 35 mm; Sigma), and dialyzed for 24 h against a 1,000-fold volume of the same buffer. The dialyzed enzyme solution was applied via a flow adapter to a carboxymethyl Sepharose column (2.5 by 10 cm) equilibrated with the same buffer. The flow rate was adjusted to 150 ml/h, and 5-ml fractions were collected. Active fractions were pooled and lyophilized with a Savant vacuum centrifuge equipped with a refrigerated condensation trap. The dry powder from the pooled active fractions was resuspended in 1.5 ml of 50 mM HEPES (pH 7.0) and applied to a Sephadex G-75 column (1 by 90 cm) that had previously been equilibrated with the same buffer. The flow rate was adjusted to 6 ml/h, and 2-ml fractions were collected. Fractions containing milk-hydrolyzing activity were pooled, lyophilized as described above, and resuspended in 1 ml of 50 mM HEPES (pH 7.0). The size of the native form of the purified proteinase was analyzed by gel filtration on a TSK-G2000XL high-pressure liquid chromatography column with a running buffer composed of 50 mM HEPES (pH 7.0) plus 50 mM NaCl. The approximate size of the native protein was determined by linear regression of the purified proteinase in comparison with known size standards.

Electrophoresis procedures. Protein samples were analyzed by sodium dodecyl sulfate (SDS)-PAGE with a 12% polyacrylamide resolving gel and a 3% polyacrylamide stacking gel as described by Laemmli (29). The final concentration of SDS in both gels was 0.1%. Samples were prepared by mixing 5 to 10 μ g of protein in a sample buffer containing 125 mM Tris-HCl (pH 6.8), 5% (vol/vol) β -mercaptoethanol, 10% (wt/vol) glycerol, 0.1% (wt/vol) SDS, and a few crystals of bromophenol blue. A constant amperage of 20 mA was applied until the dye front had reached the bottom edge of the gel. The proteins in the SDS-PAGE gels were detected by using silver nitrate as described by Merrill et al. (35). The relative molecular mass (M_r) of the denatured proteinase was determined by linear regression in comparison to known size standards after separation by SDS-PAGE.

Substrate specificity of the recombinant proteinase. The purified milk-hydrolyzing enzyme was tested for its ability to cleave several synthetic amino acid *p*-nitroanilides: *N*-succinyl-alanyl-alanyl-prolyl-phenylalanyl *p*-nitroanilide, a substrate for chymotrypsin-like proteinases (14); *N*- α -benzoyl-L-arginine *p*-nitroanilide, a substrate for trypsinlike proteinases (18); benzyloxycarbonyl-glycyl-glycyl-leucyl *p*-nitroanilide, a substrate for subtilisinlike proteinases (33); and L-leucyl *p*-nitroanilide, a substrate for leucine aminopeptidases (1). Stock solutions of 5.0 mM were used for each synthetic substrate to yield a final concentration in the assay of 0.25 mM. *N*- α -Benzoyl-L-arginine *p*-nitroanilide and benzyloxycarbonyl-glycyl-glycyl-leucyl *p*-nitroanilide were dissolved in dimethylsulfoxide, *N*-succinyl-alanyl-alanyl-prolyl-phenylalanyl *p*-nitroanilide was dissolved in 100 mM HEPES buffer (pH 7.0), and L-leucyl *p*-nitroanilide was

dissolved in ethanol. The reactions were carried out in 100 mM HEPES (pH 7.0) in a final volume of 1.0 ml at 37°C. The A_{410} of the *p*-nitroaniline released by hydrolysis of all of the amino acid *p*-nitroanilides was measured continuously with a Bausch and Lomb Spectronic 2000 spectrophotometer equipped with a temperature control cuvette holder attached to a Haake circulating water bath. The molar extinction coefficient of *p*-nitroaniline in this buffer was $8,480 \text{ M}^{-1} \cdot \text{cm}^{-1}$ (57). Additional assays were carried out as above but at a pH of 8.0, and similar assays were carried out in the presence and absence of 10 mM CaCl_2 . Thermolysinlike neutral proteinase activity was measured by the hydrolysis of *N*-(3-[2-furyl]acryloyl)-Gly-Leu amide as described by Feder (19), and esterase activity was assayed by the hydrolysis of *N*- α -*p*-tosyl-L-arginine methyl ester, as described by Roberts and Elmore (45). The enzymatic hydrolysis of the specific collagenase substrates *N*-(3-[2-furyl]acryloyl)-leucyl-glycyl-prolyl-alanine and 4-phenylazobenzyloxycarbonyl-prolyl-leucyl-glycyl-prolyl-D-arginine were measured as described by Van Wart and Steinbrink (55) and Wunsch and Heidrich (59), respectively. The enzymatic hydrolysis of azocasein was assayed as described previously (20, 21), and the hydrolysis of azocoll was carried out as described by Rufo et al. (46).

Effect of inhibitors and influence of metal ions. The following proteinase inhibitors (47) were assessed for their ability to inhibit the milk-hydrolyzing activity of the purified enzyme: DIFP, EDTA, ethylene glycol-bis(β -aminoethyl ester)-*N,N,N',N'*-tetraacetic acid (EGTA), 1,10-phenanthroline, *N*-tosyl-L-phenylalanine chloromethyl ketone, *N*-*p*-tosyl-L-lysine chloromethyl ketone, *N*-(benzyloxycarbonyl)-L-leucyl-L-methionylglycine hydroxamic acid (36), aprotinin, bestatin, phosphoramidon (28), pepstatin, dithiothreitol (DTT), iodoacetate, and HgCl_2 . All inhibitors were tested in the concentration range of 1 to 10 mM against the milk-hydrolyzing activity of the purified proteinase. Additionally, some of the inhibitors were also assayed against the azocoll-hydrolyzing activity of the purified proteinase. For these assays, the enzyme was preincubated with the inhibitor for 1 h at ambient temperature before residual enzyme activity was measured. Controls included incubation of the enzyme in an equal volume of buffer for the same period and incubation of the enzyme in solvents in which the inhibitors were dissolved. Various cations were added to the purified enzyme to determine whether they could counteract the effect of certain inhibitors on enzyme activity, and the ability of these cations to restore activity to previously inactivated enzyme was also determined. Inactivated enzyme was incubated with each cation preparation for 15 min at ambient temperature before residual activity was measured.

Metal ion determination. The purified protein was dialyzed for 24 h against a 1,000-fold volume of double-distilled water with two changes of the water before metal ions were determined. The presence of heavy metal associated with the proteinase was determined by energy-dispersive X-ray fluorescence spectrometry with a SpecTrace 4050 spectrometer (Tracor X-ray, Inc., Mountainview, Calif.). Zinc, determined by energy-dispersive X-ray fluorescence spectrometry to be the major heavy metal in the pure proteinase, was then quantified in a preparation of extensively dialyzed, pure proteinase by atomic absorption spectrometry with a Perkin-Elmer (Norwalk, Conn.) model 403 atomic absorption spectrometer at 213.8 nm. The content of zinc per molecule of dialyzed proteinase was calculated by dividing the moles of zinc in the purified proteinase preparation by the moles of

proteinase, based on an M_r of 15,500 for the enzyme (see Results).

Effect of temperature and pH on enzyme activity. The purified enzyme was assayed at various temperatures by using a temperature control cuvette holder attached to a Haake recirculating water bath. The temperature inside the cuvette was measured with a temperature microprobe (model 42SC telethermometer; Yellow Springs Instrument Co., Yellow Springs, Ohio) to determine the actual reaction temperature. The thermostability of the enzyme was determined by incubating the enzyme at appropriate temperatures for periods of time, during which samples taken at time intervals for each temperature were assayed for residual enzyme activity. The effect of pH on enzyme activity was determined by running the standard milk hydrolysis assay in 100 mM HEPES buffer at various pH values between 6.5 and 8.5.

N-terminal sequencing of purified neutral proteinase. After electrophoresis, staining, and destaining, an SDS-PAGE gel containing the purified neutral proteinase was soaked in cold transfer buffer [10 mM 3-(cyclohexylamino)-1-propanesulfonic acid, 10% methanol (pH 11.0)] for 15 min. A polyvinylidene difluoride membrane (Immobilon P; Millipore, Inc., Boston, Mass.) was also rinsed briefly in 100% methanol, soaked in deionized water for 5 min, and stored in transfer buffer. The gel, sandwiched between the polyvinylidene difluoride membrane and several sheets of blotting paper (Whatman 3 mm), was assembled into a blotting apparatus (Trans-Blot; Bio-Rad). Protein was transferred at 4°C at a constant voltage of 90 V (200 mA) for 150 min with the polyvinylidene difluoride membrane toward the cathode. The polyvinylidene difluoride membrane was then rinsed in deionized water for 5 to 10 min, air dried, and stored at -20°C. The N-terminal amino acid sequence of the neutral proteinase (10 pmol) purified from the culture broth of *S. lividans*(pANT42) was determined by Edman degradation with an automated gas-phase protein sequencer (model 470A; Applied Biosystems, Inc., Foster City, Calif.) and an on-line phenylthiohydantoin analyzer (model 120A; Applied Biosystems).

Nucleotide sequence accession number. The DNA sequence data described in this paper have been deposited at EMBL and GenBank with the accession number M86551.

RESULTS

Cloning of a gene from *Streptomyces* sp. strain C5 encoding milk-hydrolyzing proteinase. Preliminary results showed that *S. galilaeus* ATCC 31133 and *Streptomyces peucetius* ATCC 29050 produced essentially no zones of hydrolysis, and *S. lividans* 1326 produced very weak zones of hydrolysis, on powdered milk. Moreover, the zones of milk hydrolysis from *S. lividans* cultures were only visible after several days of incubation on MA plates. All other streptomycetes tested, including *Streptomyces* sp. strain C5, *Streptomyces coelicolor* A3(2), *Streptomyces albus* G, and *S. griseus* IMRU 3499, produced large zones of inhibition on powdered milk within the first 36 to 48 h of incubation (data not shown). Although *Streptomyces* sp. strain C5 produced large zones of inhibition on powdered milk plates, this strain did not significantly hydrolyze azocasein (20, 21). All other strains tested, on the other hand, hydrolyzed azocasein strongly (21). Thus, we attempted to clone the *Streptomyces* sp. strain C5 milk-hydrolyzing proteinase that was incapable of azocasein hydrolysis.

From approximately 5,000 transformants of thiostrepton-

resistant *S. lividans* obtained after introduction of pIJ702 containing DNA fragments from *Streptomyces* sp. strain C5, only three colonies were observed to have large zones of hydrolysis on the initial MAT plates. Colonies on control MAT plates (i.e., transformants receiving only pIJ702) did not produce zones of hydrolysis that were any larger or produced any faster than those produced by untransformed *S. lividans* grown on MA plates. The milk-hydrolyzing transformants were picked over to fresh MAT plates to verify the proteinase phenotype; only one of the strains maintained the strong milk-hydrolyzing phenotype. This transformant contained a 12.25-kb plasmid (pIJ702 plus a 6.56-kb DNA insert) that was named pANT21 (Fig. 1).

Subcloning of pANT21. To localize the milk hydrolysis-conferring gene within the 6.56-kb *Sst*I fragment, pANT21 was digested with *Bam*HI and ligated into the *Bam*HI site of pKC505, an 18.7-kb, low-copy-number, *E. coli-Streptomyces* shuttle vector to make plasmids pANT22 (23.5 kb) and pANT23 (26.15 kb) in *E. coli* DH5 α (Table 1). Plasmids pANT22 and pANT23 were then each introduced into *S. lividans* protoplasts, and apramycin-resistant transformants were picked over to MA plates and screened for milk protein-hydrolyzing activity. *S. lividans*(pANT23) displayed milk-hydrolyzing activity on milk plates containing apramycin, whereas *S. lividans*(pANT22) did not (Table 1), indicating that the proteinase gene was located on the 3.31-kb *Bam*HI-*Sst*I DNA fragment. This fragment was ligated with *Bgl*III- and *Sst*I-digested pIJ702 to make pANT42 (Fig. 1, Table 1). Further subcloning of the ca. 1.1-kb *Sph*I-*Sst*I DNA fragment of pANT42 into *Sph*I- and *Sst*I-digested pIJ702 resulted in pANT60 (Fig. 1B; Table 1), which also conferred stable, strong milk-hydrolyzing activity on *S. lividans* 1326. The sequence of the *Sph*I site partially encodes a Met residue at the N terminus of the open reading frame (ORF) encoding the proteinase (Fig. 2). Similarly, the *Sph*I site in pIJ702 encodes the initiation codon ATG of the *melC1* gene (2). The *melC1* promoter lies just upstream of the *Sph*I site in pIJ702, and we have deduced from these data that, in pANT60, transcription of *snpA* is actually from the *melC1* promoter and not from the (uncharacterized) promoter of *snpA*. Preliminary fermentation data have indicated that proteinase production from *S. lividans*(pANT60) is very differently regulated than the protease produced in cultures of *S. lividans*(pANT42) (data not shown), supporting the notion that different promoters are being used in the two separate constructs.

Activity conferred by the upstream sequence. The 1.2-kb *Sst*I-*Sph*I fragment located immediately upstream of the *Streptomyces* sp. strain C5 DNA cloned in pANT60 was subcloned from pANT54 (Table 1) into pIJ702 to make pANT55. Transformation of *S. lividans* 1326 with pANT55 reproducibly resulted in the formation of a slow (visible only after 36 h of incubation) but distinguishable milk hydrolysis activity above background levels. Transformation of *S. galilaeus* with pANT55, on the other hand, resulted in no proteolytic activity whatsoever, even after 48 to 54 h of incubation (data not shown).

Analysis of nucleotide sequence and open reading frame. The ca. 1.1-kbp *Sph*I-*Sst*I DNA fragment conferring strong milk-hydrolyzing activity isolated from *Streptomyces* sp. strain C5 plus about 330 nucleotides upstream of the *Sph*I site were sequenced completely in both directions. The pertinent area sequenced is shown in Fig. 2. A FRAME computer analysis of the G+C content of the nucleotide positions within each codon (4) indicated the presence of an ORF of 687 nucleotides (*snpA*) with a high G+C content in

the third position in the sequence, reading left to right. Combining the FRAME analysis with an ORF analysis (including translation starts and stops), we determined that this ORF begins at the ATG (formyl-methionine) located 6 nucleotides upstream of the *Sph*I site. This deduced ORF has the capacity to encode a protein of 229 amino acids (Fig. 2). It is possible, however, that other initiation codons may be used. Two other ATGs are found at nucleotides 288 and 333, either of which may serve as an alternative translation start site; however, as discussed below, this is probably not the case. No other ATG or GTG translation start site is found in the region demarcated as the N-terminal region of the ORF by FRAME analysis. The codon usage for *snpA* (data not shown) was similar to that found in other previously described streptomycete genes (4). The first, second, and third positions of codons in typical streptomycete ORFs have G+C contents of ca. 63 to 75%, 46 to 54%, and >90%, respectively (4). Gene *snpA* (nucleotides 327 to 1013), which starts with ATG, had G+C contents for codon positions 1, 2, and 3 of 64, 59, and 94%, respectively. The overall G+C content of *snpA* was 72 mol%.

A good potential ribosome binding site (GAGGA) can be found six nucleotides upstream of *snpA*. This sequence has a binding energy to the 3' end of streptomycete 16S rRNA (ΔG of -11.6 kcal/mol, determined as described by Tinoco et al. [53]) that is somewhat stronger than that of the average streptomycete ribosome binding site (50). A pair of inverted repeats and a single direct repeat can be found in the sequences upstream of *snpA* in the region that would be expected to contain the promoter to this gene (Fig. 2). Structures such as these are quite common in streptomycete promoter regions and are thought to be involved as recognition and binding sites for transcriptional regulatory proteins (50).

A potential stable stem-loop structure for the mRNA was found in the sequence directly downstream of *snpA* (Fig. 2). The ΔG of formation for this stem-loop of -36.4 kcal (ca. -152 kJ/mol) (53) was similar to the free energy of known streptomycete terminators (15, 31, 43). Among other streptomycete stem-loop terminators that have been characterized by S1 mapping and other transcriptional analyses (15, 31, 43), none is followed by a run of uracils, which would be typical of *E. coli* rho-independent terminators.

Comparison of the deduced gene product with other proteins. The deduced amino acid sequence of the protein encoded by *snpA* (SnpA) was compared with other potential protein sequences by using the TFASTA program. No sequences that contained significant homology with SnpA outside of the conserved zinc ligand site were found in the Genbank or EMBL data base (dated June 1991). Similarly, the deduced amino acid sequence of SnpA was compared directly (one-to-one) with the amino acid sequences of *Bacillus thermoproteolyticus* thermolysin (54), carboxypeptidase A (7), *Pseudomonas aeruginosa* elastase (3), *S. cacaoi* metalloproteinase (11), two *Bacillus subtilis* metalloproteinases (49, 60), and proteinase B from *Erwinia chrysanthemi* (13); no significant homology was found between SnpA and any of these other metalloproteinases. This is not necessarily surprising, given the differences between SnpA and these other neutral proteinases with respect to size, substrate specificity, and sensitivity to inhibitors. Thus, it appears that SnpA is significantly different from all previously published neutral (metallo) proteinases.

A gene encoding a normally silent proteinase from *S. lividans* 1326 that has greater than 70% homology with the sequence of *snpA* was recently cloned (9, 32). The protease

FIG. 2. The nucleotide sequence of the 1,200-bp DNA fragment and deduced amino acid sequence of the *snpA* gene product. The numbers at the right indicate nucleotide positions. Approximately 260 nucleotides separate the *Dra*III site (nucleotide 1165) and the *Sst*I site (Fig. 1B). The ribosome binding site (RBS) is bracketed. The thick horizontal arrows above the sequences indicate a direct repeat, the thick horizontal arrows beneath the sequences represent inverted repeats, and the thick vertical arrows indicate the probable sites for cleavage of the pre and pro regions of the signal sequence from the mature protein. The deduced amino acids which are identical to those sequenced from the N-terminal end of the mature protein are noted, as is the deduced amino acid sequence of the conserved zinc binding site. Restriction endonuclease sites are noted by name and double underlining of the sequences recognized.

encoded by the *S. lividans* gene appears to be larger (M_r , 22,000) than and slightly different from SnpA (32).

Purification of SnpA from *S. lividans*(pANT42) and *S. galilaeus*(pANT42). The milk-hydrolyzing enzyme encoded by *Streptomyces* sp. strain C5 *snpA* was purified from a 24-h-old fermentation culture of *S. lividans*(pANT42). The spectrophotometric milk hydrolysis assay was used to quantify the activity of the neutral proteinase, since, at the time of

purification, the enzyme was not found to hydrolyze any of the other substrates tested. All of the milk-hydrolyzing activity was in the filtrate after separation of the mycelia from the culture broth by tangential flow filtration with a membrane with a 0.45- μ m cutoff. Two steps, ultrafiltration and acetone precipitation, were used to concentrate the proteinase. Essentially all of the milk-hydrolyzing proteinase activity from the 0.45- μ m-cutoff fraction was retained by a

FIG. 3. (A) SDS-PAGE of proteinase, encoded by *Streptomyces* sp. strain C5 *snpA*, purified from 10 liters of culture broth of *S. lividans* 1326(pANT42). This pure enzyme (right lane) was the result of the purification protocol described in Results. The standards (left lane) are (M_r s given within parentheses) bovine albumin (66,000), egg albumin (45,000), glyceraldehyde-3-phosphate dehydrogenase (subunit; 36,000), carbonic anhydrase (29,000), bovine trypsinogen (24,000), soybean trypsin inhibitor (20,100), and α -lactalbumin (14,200) alongside the enzyme preparation in the gel. The relative molecular mass of the purified denatured proteinase was estimated by linear regression calculation. (B) Protein standards as described above (left lane) and SnpA proteinase purified from a 24-h-old shake flask culture of *S. galilaeus*(pANT42) (right lane). The proteins in both panels were stained with silver nitrate.

membrane with a 10,000-MW cutoff. Although 5- and 7-fold losses in yield were obtained for the filtration and acetone precipitation steps, respectively, the enzyme was concentrated about 12-fold in these steps. Attempts to use alternative methods for concentration of the enzyme, including precipitation with either ammonium sulfate or ethanol, resulted in even lower yields. The milk-hydrolyzing enzyme was eluted in the flow-through volume of the carboxymethyl Sepharose ion-exchange step, which afforded a ca. 2.4-fold purification over the previous step. Moreover, this step also removed an uncharacterized brown pigment that partially interfered with other purification steps attempted during early stages of this project. The pooled and lyophilized active fractions from the carboxymethyl Sepharose column

were applied to a Sephadex G-75 column, from which the milk-hydrolyzing enzyme was eluted in a single, sharp peak that yielded a pure protein as demonstrated by SDS-PAGE analysis (Fig. 3). The final specific activity indicated that the pure protein was about fivefold purified over the culture broth, with only a 0.3% yield (data not shown).

By using a scaled-down procedure of the one described above, SnpA was also purified to electrophoretic homogeneity from shake flask cultures of *S. galilaeus* ATCC 31133(pANT42) (Fig. 3B). *S. galilaeus* ATCC 31133 was used as an alternative host for expression of the proteinase gene because it also did not produce a milk protein-degrading proteolytic activity (51).

Physical properties of SnpA. SnpA purified from cultures of *S. lividans*(pANT42) or *S. galilaeus*(pANT42) was a monomeric protein with an M_r of 15,500 as demonstrated by SDS-PAGE (Fig. 3) and an M_r of 16,000 as demonstrated by gel filtration with a TSK-G2000XL high-pressure liquid chromatography column (data not shown). The optimum pH for activity of the proteinase was 7.0; approximately half the amount of activity was still present at a pH of 8.0 (data not shown). The optimal temperature for maximal activity of the neutral proteinase was 55°C, with 50, 73, and 89% of optimal activity obtained at 24, 37, and 44°C, respectively (data not shown). Only ca. 88 and 52% of optimal activity were obtained after incubations at 60 and 65°C, respectively (data not shown). SnpA activity was completely lost after incubation at 80°C for 30 min (data not shown).

Effect of inhibitors and metals on proteinase activity. The milk-hydrolyzing activity of SnpA was not inhibited by DIFP, *N*-tosyl-L-phenylalanine chloromethyl ketone, *N*-*p*-tosyl-L-lysine chloromethyl ketone, aprotinin, bestatin, phosphoramidon, or disodium EDTA but was completely inhibited by 10 mM 1,10-phenanthroline (Table 2), indicating that this enzyme probably requires a metal for either stability or activity. After complete inactivation of the enzyme by 1,10-phenanthroline, the milk-hydrolyzing activity of the proteinase could be reactivated 100 or 120% by the addition of 10 mM Zn^{2+} or 10 mM Co^{2+} , respectively (Table 2), but activity was not restored by the addition of 10 mM Ca^{2+}

TABLE 2. Effects of inhibitors on proteinase encoded by the *Streptomyces* sp. strain C5 *snpA* gene

Inhibitor ^a (concn)	% inhibition of proteinase activity on substrate tested	
	Milk hydrolysis	Azocoll hydrolysis
PMSF (10 mM)	0	0
DIFP (3.5 mM)	0	0
Phosphoramidon (5 mM)	0	0
CBZ-PLG-hydroxymate (10 mM)	ND	0
Aprotinin (10 mM)	0	ND
Bestatin (10 mM)	0	ND
TPCK (10 mM)	0	ND
TLCK (10 mM)	0	ND
EDTA (10 mM)	0	ND
EGTA (10 mM)	0	ND
PHEN (1, 3, 5, 10 mM, respectively)	7, 44, 56, 100	100 (10 mM PHEN)
PHEN (10 mM) plus 10 mM Zn^{2+} ^b	0	ND
PHEN (10 mM) plus 10 mM Co^{2+} ^c	0	ND
DTT (5, 10, 20 mM, respectively)	15, 21, 28	ND
Iodoacetate (20 and 30 mM, respectively)	15, 20	ND
HgCl ₂ (1, 5, and 10 mM, respectively)	85, 100, 100	ND

^a Abbreviations: ND, not done; TPCK, *N*-tosyl-L-phenylalanine chloromethylketone; CBZ-PLG-hydroxymate, *N*-(benzyloxycarbonyl)-L-leucyl-L-methionyl-glycine hydroxamic acid; TLCK, *N*-*p*-tosyl-L-lysine chloromethylketone; PHEN, 1,10-phenanthroline.

^b Added as zinc acetate.

^c Added as cobalt acetate.

(data not shown). The specific inhibitor of collagenase, *N*-(benzyloxycarbonyl)-L-leucyl-L-methionylglycine hydroxamic acid, did not inhibit the azocoll-hydrolyzing activity of the purified proteinase (Table 2).

Determination of metal ligand. Purified SnpA was extensively dialyzed against 10 mM HEPES buffer (pH 7.0) or double-distilled water to determine whether its putative metal cofactor could be removed easily. Even after extensive dialysis against either buffer or water, the enzyme remained essentially 100% active, indicating that the presumed metal ligand was tightly bound to the enzyme. X-ray analysis with a Tracor X-ray dispersive analysis instrument showed that both undialyzed and extensively dialyzed preparations of the enzyme contained significant levels of zinc and trace levels of copper (data not shown), suggesting that zinc was likely to be the metal cofactor in this neutral proteinase. Atomic absorption spectroscopy analysis showed that 0.166 ± 0.004 (average $\pm$ standard deviation of three protein determinations) mg of the purified proteinase contained 0.401 ± 0.004 (average $\pm$ difference of actual values of two determinations) μ g of zinc, indicating that, under the conditions tested, the enzyme contains 0.58 mol of zinc per mol of extensively dialyzed, pure enzyme.

Biological properties of milk-hydrolyzing proteinase. SnpA demonstrated the milk-hydrolyzing activity described above, which was not inhibited by DIFP but was completely inhibited by 10 mM 1,10-phenanthroline (Table 2). The enzyme showed no activity with the synthetic substrates L-leucine *p*-nitroanilide, *N*- α -benzoyl-L-arginine *p*-nitroanilide, *N*-succinyl-alanyl-alanyl-prolyl-phenylalanyl *p*-nitroanilide, *N*- α -tosyl-L-arginine methyl ester, or *N*-(3-[2-furyl]acryloyl)-Gly-Leu-amide at pH 7.0 or 8.0. The purified enzyme did not hydrolyze azocasein, but its activity against azocoll fit standard Michaelis-Menton kinetics (data not shown); however, because azocoll is an insoluble substrate, kinetic constants are not given. Since azocoll is a standard substrate used by collagenases, the enzyme was tested for its ability to hydrolyze the specific collagenase substrates *N*-(3-[2-furyl]acryloyl)-leucyl-glycyl-prolyl-alanine (55) and 4-phenylazobenzoyloxycarbonyl-prolyl-leucyl-glycyl-prolyl-D-arginine (59). No activity for the purified neutral proteinase was found on these collagenase substrates (data not shown). The enzyme appeared during exponential growth. Its optimum pH for activity was 7.0, the optimum temperature was 55°C, and the optimum Zn^{2+} concentration was 20 mM. The enzyme was inhibited by 1,10-phenanthroline, DTT, iodoacetamide, and $HgCl_2$. It was not inhibited by PMSF, DIFP, *N*-tosyl-L-phenylalanine chloromethyl ketone, *N*-*p*-tosyl-L-lysine chloromethyl ketone, EDTA, EGTA, aprotinin, bestatin, or phosphoramidon.

N-terminal amino acid sequences of the purified enzyme subunits. The sequence of the first 20 amino acids from the N terminus of *Streptomyces* sp. strain C5 SnpA, purified from the culture broth of *S. lividans* 1326(pANT42), was determined. The N-terminal sequence (AAVTVVYNASGAPS FATQIA) of the purified proteinase was identical to a region of the deduced amino acid sequence of the cloned gene (Fig. 2).

DISCUSSION

The milk-hydrolyzing proteinase encoded by *Streptomyces* sp. strain C5 *snpA* was found to have an M_r of 15,500 (as determined by SDS-PAGE analyses) whether it was purified from the culture broth of *S. lividans*(pANT42) or *S. galilae-*

us(pANT42). Gel filtration also indicated an M_r of 16,000, indicating that the enzyme is a monomeric protein. Moreover, the sequence of the first 20 amino acids at the N terminus of the pure protein was determined and matched exactly an amino acid sequence deduced from the gene sequence of *Streptomyces* sp. strain C5 *snpA*. The deduced amino acid sequence of the mature protein, based on the position of the N-terminal amino acid sequence, includes 148 amino acids that would theoretically encode a protein of M_r 15,740. This corroborates the SDS-PAGE and gel filtration data and demonstrates that the protein purified and characterized in this work was encoded by the *snpA* gene from *Streptomyces* sp. strain C5. Since a gene that is very similar to *Streptomyces* sp. strain C5 *snpA* was also found in *S. lividans* 1326 (9, 32), the physical data substantiate the assertion that *Streptomyces* sp. strain C5 SnpA and the *S. lividans* homolog are different proteins.

Lichenstein et al. (32) purified the SnpA homolog from *S. lividans* 1326 and found it to be a protein of M_r 22,000. Similarly, Butler et al. (9) have also indicated that the *S. lividans* 1326 *snpA* homolog encodes potentially multiple proteins that are somewhat larger (i.e., M_r 20,000 to 24,000) than *Streptomyces* sp. strain C5 SnpA. This has posed the possibility that the 15,500- M_r protein we purified is a proteolytic degradation product of a larger initial protein encoded by *snpA*. In crude culture broths of *S. lividans*(pANT42) and *S. galilaeus*(pANT42) and in samples from early purification steps, however, the major band specifically associated with SnpA activity was always observed to be ca. M_r 15,500 to 16,000. Furthermore, we did not observe in any of our experiments a larger protein product of *snpA* that appears to be cleaved to the final size of ca. M_r 15,500. The amino acid sequence of the leader peptide region of *S. lividans prt* (32) is considerably different from the peptide leader sequence of the *Streptomyces* sp. strain C5 *snpA*. It is possible that the considerable differences in those leader peptide amino acid sequences allow differential cleavage of the leader peptides, thus resulting in the different sizes of the mature products. Finally, the similarity in sizes of the SnpAs from *S. galilaeus*(pANT42) and *S. lividans*(pANT42) also indicates that these two organisms recognize and cleave off the leader peptide from mature SnpA in an analogous manner.

The 81 amino acids upstream of the N terminus of the mature proteinase probably constitute a leader sequence, as has been found with most bacterial extracellular proteinases (11, 22, 37, 49, 56, 60). The N terminus of this region, however, is somewhat different from most bacterial leader sequences (58) in that it contains only a single positively charged amino acid (Arg) that is placed between the two Met residues. If the second Met residue were found to be the actual translation start site, there would be no positively charged residues in this region. This, and the location of the potential strong ribosome binding site, is why we assume at this time that the upstream Met is the probable N terminus of the protein. The leader peptide sequences of D,D-carboxypeptidase (17) and *Streptomyces* subtilisin inhibitor (40), both exported streptomycete proteins, also contain only a single positively charged amino acid in the N-terminal region of the leader sequence. Furthermore, the N-terminal regions of several other peptide leader sequences of exported streptomycete proteins (e.g., HAIM, SapA, Lep-10, LTI, Amy, AmlV, AmlG) contain only two positively charged amino acids (51), indicating that the oft-cited requirement of multiple positively charged amino acids in this region of the leader peptide sequence may not be so stringent in streptomycetes. The methods described by von Heijne

Gene	Residue	Residue	Residue
1. C5-SnpA	VTAHETGHVGL 82-93		
2. SL-SnpA	VTAHETGHVGL 160-171		
3. NpII	TTLHEFTHAPGV 125-136		
4. Npr	TTVHEAGHSLMG 199-210	* FREYFLTD 315-322	+
5. ProB	SFTHEIGHALGL 171-182	LNEGSFSD 304-311	DFKG-HYSAG 224-232
6. Serratia	TFTHEIGHALGL 173-184	LNEKSFS 305-312	DNGG-HYAAA 225-233
7. Elastase	VAAHEVSHGFTD 137-148	MNE-AFSD 162-168	D---VHYSSG 221-227
8. Thermolysin	VVGHELTAVTD 139-150	INE-AISD 164-170	DNGGVHINS 226-235
9. BS-Npr	VTAHEMTHGVTQ 140-151	LNE-SFSD 165-171	DYGGVHTNSG 223-242

FIG. 4. Comparison of the conserved zinc ligand and active site of several neutral proteinases. Lines: 1, *Streptomyces* sp. strain C5 SnpA (this work); 2, *S. lividans* 1326 homolog of SnpA (called SL-SnpA by Butler et al. [9] and Prt by Lichenstein et al. [32]); 3, *A. oryzae* neutral proteinase NpII (52); 4, *S. cacaoi* Npr (11); 5, *E. chrysanthemi* protease B (13); 6, *Serratia* neutral proteinase (37); 7, *P. aeruginosa* elastase (3); 8, *B. thermoproteolyticus* thermolysin (54); 9, *B. subtilis* neutral proteinase (60). *, zinc ligands; +, active site (where known); -, gap.

(58) and data on known leader peptide sequences in streptomycetes, recently reviewed by Brawner et al. (8), indicate that there is a potential cleavage site in the leader peptide sequence (Fig. 2). This site likely represents the cleavage between the "pre" region and the "pro" region, as found for a variety of other proteinases (11, 22, 56, 60). Using the predicted cleavage site (PASA↓TPT) shown in Fig. 2, the pre, or leader sequence, region of the upstream peptide would be 28 amino acids long, which is approximately the same size as several other streptomycete leader peptides (51).

A conserved sequence of 12 amino acids was found in the deduced amino acid sequence of SnpA for the primary zinc binding site of metalloproteinases (Fig. 4). The zinc binding site of SnpA shared between 3 and 8 amino acids with all of the other sequences shown; it shared 5 amino acids with the zinc-binding site of *S. cacaoi* metalloproteinase and 4 amino acids with thermolysin, the classic metalloproteinase from *B. thermoproteolyticus*. This conserved zinc-binding motif is characteristic of thermolysinlike neutral proteinases (26). Moreover, metalloproteinases of the thermolysin superfamily also have additional secondary ligand sites, including a sequence conserved with INEAISD, of which the glutamic acid residue is the zinc ligand (24), and DNGGVHINS, of which the histidine is an active site (24); the amino acid sequences shown are for thermolysin). SnpA does not contain amino acid sequences resembling these additional ligand-binding and active sites, indicating that it is considerably different from the thermolysinlike metalloproteinases. Our data also indicate that, at most, one zinc molecule is bound to each molecule of SnpA and that no calcium ions are bound to the proteinase, which is considerably different from the one zinc and four calcium molecules found for each molecule of thermolysinlike metalloproteinases (24). The sequence of proteinase NpII, an *Aspergillus oryzae* protein with an M_r of 19,000 that also contains a single zinc binding site, was recently described (Fig. 4); however, it lacks the second zinc ligand Glu residue and the active-site His residue (52). Nevertheless, the deduced amino acid sequence of this proteinase (NpII) shows little similarity with that of SnpA. Moreover, the substrate specificities of NpII are very different from those of SnpA (52), and NpII is markedly thermostable whereas SnpA loses activity rapidly at 80°C (52), indicating that NpII is very different from SnpA.

The zinc/proteinase molar ratio was approximately 0.6:1. Based on data obtained with other neutral proteinases (24) and the presence of a single zinc binding domain in the enzyme, we expected a value somewhat closer to a 1:1 ratio. The low value obtained may be due to loss of zinc during dialysis, especially since secondary zinc binding domains are

not present in the proteinase. Cobalt was found to substitute for zinc (and even yielded a slightly better activity), but calcium was not associated with the enzyme, nor did calcium stabilize the enzyme or increase the specific activity of the enzyme against milk proteins. Whether the zinc is involved in the catalytic mechanism is unknown at this time. The lack of secondary zinc binding sites, the binding of zinc to the proteinase, and the potential role of the zinc in catalytic activity are under further investigation.

Microbial endolytic proteinases fall into four major groups: the serine or alkaline proteinases, which have a serine residue at the active site and which are inhibited by hydroxyl-reactive organofluorides such as DIFP and phenylmethylsulfonyl fluoride (PMSF) (47); the sulfhydryl proteinases, which contain a cysteine at the active site and are markedly stimulated by the presence of thiol reagents such as DTT (39); the acid proteinases, which typically have an aspartic acid residue at the active site and are active under low-pH conditions (24); and the neutral proteinases, which are active at neutral pH, typically contain a metal ligand as part of the active site, and are typically inhibited by chelating agents such as EDTA, EGTA, and 1,10-phenanthroline (24, 47). The major characteristics of *Streptomyces* sp. strain C5 SnpA are summarized above. This small proteinase has optimal activity at pH 7.0, and it contains a zinc ligand that can be chelated from the enzyme by relatively high concentrations of 1,10-phenanthroline but not EDTA or EGTA. SnpA is not inhibited by DIFP or PMSF, indicating that it is not a serine proteinase, and DTT slightly inhibits rather than stimulates activity of the enzyme, indicating that it is not a sulfhydryl proteinase. Furthermore, high concentrations (20 and 30 mM) of iodoacetate have only a marginal inhibitory effect on the proteinase, further indicating it is not a sulfhydryl proteinase (47). Thus, SnpA has characteristics typical of neutral proteinases. The small size (ca. 15,500 Da) of the mature recombinant *Streptomyces* sp. strain C5 neutral proteinase expressed in *S. lividans*, however, is highly unusual, since nearly all neutral proteinases isolated from bacteria thus far have ranged in molecular mass from 28,000 to 57,000 Da. Moreover, *N*-(3-[2-furyl]acryloyl)-Gly-Leu amide, a substrate designed for measuring the activity of thermolysinlike neutral proteinases of *Bacillus* spp. (19), is not a substrate for SnpA. Additionally, phosphoramidon, a relatively specific inhibitor for thermolysinlike neutral proteinases (28), did not inhibit SnpA, even at very high concentrations (Table 2). Normally, phosphoramidon inhibits neutral proteinases at micromolar concentrations (28, 47), but it did not inhibit SnpA milk- or azocoll-hydrolyzing activity, even at a concentration of 5 mM. The same concentration of phosphoramidon inhibited *Clostridium* collagenase completely in a control experiment (data not shown).

Other than the primary zinc binding domain (HExxH) found in a variety of zinc-containing proteinases (26), additional conserved amino acid sequences normally associated with zinc binding and active sites of thermolysinlike neutral proteinases were not found in the deduced amino acid sequence of SnpA (Fig. 4). Calcium ions, which are normally associated with thermolysin and several of the proteinases like it (24), also were not found in association with purified SnpA, nor did Ca^{2+} stimulate activity or restore activity of 1,10-phenanthroline-treated SnpA, further indicating the difference between this proteinase and thermolysinlike proteinases. Likewise, SnpA has no similarity to carboxypeptidase A-like neutral proteinases (7). Finally, although azocoll was a good substrate for SnpA, the proteinase had no activity on two specific collagenase substrates, and the azocoll-hydrolyzing activity of SnpA was not inhibited by the specific collagenase inhibitor, *N*-(benzyloxycarbonyl)-L-leucyl-L-methionylglycine hydroxamic acid, indicating that it is not an unusual collagenase. Thus, SnpA appears to be a highly unusual neutral proteinase.

Since the cloned milk-hydrolyzing proteinase contained two cysteine residues (based on the deduced sequence of the mature enzyme), it appears that a disulfide bond-linked cysteine residue pair may be present in the native protein. The inhibition of the milk-hydrolyzing activity of the purified proteinase by DTT (Table 2) further suggests the presence in the native protein of disulfide bridges that are required for the active conformation. The presence of a disulfide linkage would enhance the stability of the enzyme. We have noted that pure preparations of this enzyme are stable at 4°C for the period of a year, indicating that this enzyme is very stable in solution. Similarly, the enzyme is equally stable and active when suspended in distilled water and buffer. Based on its amino acid sequence deduced from the gene, P_{rt}, a similar proteinase from *S. lividans*, also contains two cysteine residues (32). Recently, an unusual neutral proteinase of *B. subtilis* (Mpr) that also appeared to contain disulfide bridges was purified (46), and the small, highly thermostable, neutral proteinase (NpII) from *A. oryzae* has six Cys residues that are thought to form disulfide bridges to yield the remarkable thermostability of that enzyme (52).

Upstream from *S. lividans prt* (the *snpA*-like gene) is a divergently transcribed gene (9, 32) that encodes a protein with a relatively high sequence similarity with members of the LysR family of bacterial activator proteins (23). We partially sequenced this upstream region (e.g., the pANT55 insert) and found that 17 of the first 19 amino acid residues of the divergently reading ORF (Fig. 2) are identical to those of the *orf-1* gene of *S. lividans* 1326 (9, 32) and that several stretches of DNA sequences with a very high degree of identity to the *orf-1* gene of *S. lividans* 1326 (9, 32) were present. Since *S. lividans* contains an *snpA* homolog (i.e., *prt* [9, 32], although it is normally silent) and *S. galilaeus* apparently does not (as determined by Southern hybridization; data not shown), it seems probable that the transformation of *S. lividans* with pANT55, a plasmid presumably carrying the *Streptomyces* sp. strain C5 *orf-1* gene equivalent, activated the normally silent *prt* gene, which was visualized phenotypically as a slow, milk-hydrolyzing activity.

ACKNOWLEDGMENTS

We thank Michael J. Butler of Cangene Corp., Mississauga, Ontario, and Henri Lichenstein of Amgen Corp., Inc., Thousand Oaks, Calif., for sharing data with us before publication. We also thank Jane Tolley of the Ohio State University Biochemical Instru-

mentation Laboratory for determining the N-terminal amino acid sequence of the protein, Donald Ordaz and the Ohio State University fermentation facility for use of the fermentors and harvesting equipment, and Cynthia Miller and Robert Pfister for their assistance in metal determinations.

REFERENCES

1. Aretz, W., K. P. Koller, and G. Riess. 1989. Proteolytic enzymes from recombinant *Streptomyces lividans* TK24. FEMS Microbiol. Lett. 65:31-36.
2. Bernan, V., D. Filpula, W. Herber, M. Bibb, and E. Katz. 1985. The nucleotide sequence of the tyrosinase gene from *Streptomyces antibioticus* and characterization of the gene product. Gene 37:101-110.
3. Bever, R. A., and B. H. Iglewski. 1988. Molecular characterization and nucleotide sequence of the *Pseudomonas aeruginosa* elastase structural gene. J. Bacteriol. 170:4309-4314.
4. Bibb, M. J., P. R. Findlay, and M. W. Johnson. 1984. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. Gene 30:157-166.
5. Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1513-1523.
6. Bradford, M. M. 1976. A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248-254.
7. Bradshaw, R. A., L. H. Ericsson, K. A. Walsh, and H. Neurath. 1969. The amino acid sequence of bovine carboxypeptidase A. Proc. Natl. Acad. Sci. USA 63:1389-1394.
8. Brawner, M., G. Poste, M. Rosenberg, and J. Westpheling. *Streptomyces*: a new host for heterologous gene expression. In M. Rosenberg (ed.), Expression systems, current opinion in biotechnology, in press. Current Biology, Philadelphia, Pa.
9. Butler, M. J., C. C. Davey, P. Krygsman, E. Walczyk, and L. T. Malek. Cloning of genetic loci involved in endopeptidase activity in *S. lividans* 66: a novel neutral protease gene with an adjacent divergent putative regulatory gene. Can. J. Microbiol., in press.
10. Chandrasekaran, S., and S. C. Dhar. 1987. Multiple proteases from *Streptomyces moseratus*. Arch. Biochem. Biophys. 257:395-401.
11. Chang, P. C., T.-C. Kuo, A. Tsugita, and Y.-H. W. Lee. 1990. Extracellular metalloprotease gene of *Streptomyces cacaoi*: structure, nucleotide sequence and characterization of the cloned gene product. Gene 88:87-95.
12. Dekleva, M. L., J. A. Titus, and W. R. Strohl. 1985. Nutrient effects on anthracycline production by *Streptomyces peucetius* in a defined medium. Can. J. Microbiol. 31:287-294.
13. Delepelaire, P., and C. Wandersman. 1989. Protease secretion by *Erwinia chrysanthemi*. Proteases B and C are synthesized and secreted as zymogens without a signal peptide. J. Biol. Chem. 264:9083-9089.
14. Delmar, E. G., C. Largman, J. W. Brodrick, and M. C. Geokas. 1979. A sensitive new substrate for chymotrypsin. Anal. Biochem. 99:316-320.
15. Deng, Z., T. Kieser, and D. A. Hopwood. 1987. Activity of *Streptomyces* transcriptional terminator in *Escherichia coli*. Nucleic Acids Res. 15:2665-2675.
16. Devereux, J., P. Haeblerli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12:387-395.
17. Duez, C., C. Piron-Fraipont, B. Joris, J. Dusart, M. S. Urdea, J. A. Martial, J.-M. Frere, and J.-M. Ghuyssen. 1987. Primary structure of the *Streptomyces* R61 extracellular DD-peptidase. Eur. J. Biochem. 162:509-518.
18. Erlanger, B. F., N. Kowowski, and W. Cohen. 1961. The preparation and properties of two new chromogenic substrates of trypsin. Arch. Biochem. Biophys. 95:271-278.
19. Feder, J. 1968. A spectrophotometric assay for neutral protease. Biochem. Biophys. Res. Commun. 32:326-332.
20. Gibb, G. D., M. L. Dekleva, J. S. Lampel, and W. R. Strohl.

1987. Isolation and characterization of a *Streptomyces* C5 mutant strain hyperproductive of extracellular protease. *Biotechnol. Lett.* **9**:605–610.
21. Gibb, G. D., D. E. Ordaz, and W. R. Strohl. 1989. Overproduction of extracellular protease activity by *Streptomyces* C5-A13 in fed-batch fermentations. *Appl. Microbiol. Biotechnol.* **31**: 119–124.
22. Henderson, G., P. Krygsman, C. J. Liu, C. C. Davey, and L. T. Malek. 1987. Characterization and structure of genes for proteases A and B from *Streptomyces griseus*. *J. Bacteriol.* **169**:3778–3784.
23. Henikoff, S., G. W. Haughn, J. M. Calvo, and J. C. Wallace. 1988. A large family of bacterial activator proteins. *Proc. Natl. Acad. Sci. USA* **85**:6602–6606.
24. Hofmann, T. 1985. Metalloproteinases, p. 1–64. In P. M. Harrison (ed.), *Metalloproteins, part 2: metal proteins with non-redox roles*. Verlag Chemie GmbH, Weinheim, Germany.
25. Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schrepf. 1985. Genetic manipulation of *Streptomyces*, a laboratory manual. The John Innes Foundation, Norwich, United Kingdom.
26. Jongeneel, C. V., J. Bouvier, and A. Bairoch. 1989. A unique signature identifies a family of zinc-dependent metalloproteinases. *FEBS Lett.* **242**:211–214.
27. Katz, E., C. J. Thompson, and D. A. Hopwood. 1983. Cloning and expression of the tyrosinase gene from *Streptomyces antibioticus* in *Streptomyces lividans*. *J. Gen. Microbiol.* **129**:2703–2714.
28. Komiyama, T., H. Suda, T. Aoyagi, T. Takeuchi, and H. Umezawa. 1975. Studies on the inhibitory effect of phosphoramidon and its analogs on thermolysin. *Arch. Biochem. Biophys.* **171**:727–731.
29. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
30. Lampel, J. S., and W. R. Strohl. 1986. Transformation and transfection of the anthracycline-producing streptomycetes. *Appl. Environ. Microbiol.* **51**:126–131.
31. Li, Y., D. C. Dosch, W. R. Strohl, and H. G. Floss. 1990. Nucleotide sequence and transcriptional analysis of the nosiheptide-resistance gene from *Streptomyces actuosus*. *Gene* **91**: 9–17.
32. Lichenstein, H. S., L. A. Busse, G. A. Smith, L. O. Narhi, M. O. McGinley, M. F. Rohde, J. L. Katowitz, and M. M. Zukowski. 1992. Cloning and characterization of a gene encoding extracellular metalloprotease from *Streptomyces lividans*. *Gene* **111**: 125–130.
33. Lyublinskaya, L. A., S. V. Belyaev, A. Y. Strongin, L. E. Matyash, E. D. Levin, and V. M. Stepanov. 1974. A new chromogenic substrate for subtilisin. *Anal. Biochem.* **62**:371–376.
34. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. Merrill, C. R., D. Goldman, and M. L. Van Keuren. 1983. Silver staining methods for polyacrylamide gel electrophoresis. *Methods Enzymol.* **96**:230–239.
36. Moore, W. M., and C. A. Spilburg. 1986. Purification of human collagenases with a hydroxamic acid affinity column. *Biochemistry* **25**:5189–5195.
37. Nakahama, K., K. Yoshimura, R. Marumoto, M. Kikuchi, I. S. Lee, Y. Hase, and H. Matubara. 1986. Cloning and sequencing of *Serratia* protease gene. *Nucleic Acids Res.* **14**:5843–5855.
38. Narahashi, Y. 1970. Pronase. *Methods Enzymol.* **19**:651–664.
39. North, M. J. 1985. Cysteine proteinases of cellular slime molds. *Biochem. Soc. Trans.* **13**:288–290.
40. Obata, S., S. Taguchi, I. Kumagai, and K.-I. Miura. 1989. Molecular cloning and nucleotide sequence determination of gene encoding *Streptomyces* subtilisin inhibitor (SSI). *J. Biochem.* **105**:367–371.
41. Pearson, W. R., and D. J. Lipmann. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**:2444–2448.
42. Pokorny, M., L. J. Vitale, V. Turk, M. Renko, and J. Zuvanic. 1979. *Streptomyces rimosus* extracellular proteases. I. Characterization and evaluation of various crude preparations. *Eur. J. Appl. Microbiol. Biotechnol.* **8**:81–90.
43. Pulido, D., and A. Jiménez. 1987. Optimization of gene expression in *Streptomyces lividans* by a transcription terminator. *Nucleic Acids Res.* **15**:4227–4240.
44. Richardson, M. A., S. Kuhstoss, P. Solenberg, N. A. Schaus, and R. N. Rao. 1987. A new shuttle cosmid vector, pKC505, for streptomycetes: its use in the cloning of three different spiramycin-resistance genes from a *Streptomyces ambofaciens* library. *Gene* **61**:231–241.
45. Roberts, D. V., and D. T. Elmore. 1974. Kinetics and mechanism of catalysis by proteolytic enzymes. *Biochem. J.* **141**:545–554.
46. Rufo, G. A., Jr., B. J. Sullivan, A. Sloma, and J. Pero. 1990. Isolation and characterization of a novel extracellular metalloprotease from *Bacillus subtilis*. *J. Bacteriol.* **172**:1019–1023.
47. Salvesen, G., and H. Nagase. 1989. Inhibition of proteolytic enzymes, p. 83–104. In R. J. Beynon and J. S. Bond (ed.), *Proteolytic enzymes, a practical approach*. IRL Press, Oxford.
48. Sanger, R., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
49. Sloma, A., C. F. Rudolph, G. A. Rufo, Jr., B. J. Sullivan, K. A. Theriault, D. Ally, and J. Pero. 1990. Gene encoding a novel extracellular metalloprotease in *Bacillus subtilis*. *J. Bacteriol.* **172**:1024–1029.
50. Strohl, W. R. Compilation and analysis of DNA sequences associated with streptomycete promoters. *Nucleic Acids Res.*, in press.
51. Strohl, W. R. Unpublished data.
52. Tatsumi, H., S. Murakami, R. F. Tsuji, Y. Ishida, K. Murakami, A. Masaki, H. Kawabe, H. Arimura, E. Nakano, and H. Motal. 1991. Cloning and expression in yeast of a cDNA clone encoding *Aspergillus oryzae* neutral protease II, a unique metalloprotease. *Mol. Gen. Genet.* **228**:97–103.
53. Tinoco, I. Jr., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla. 1973. Improved estimation of secondary structure in ribonucleic acid. *Nature (London) New Biol.* **246**:40–41.
54. Titani, K., M. A. Hermodson, L. H. Ericsson, K. A. Walsh, and H. Neurath. 1972. Amino-acid sequence of thermolysin. *Nature (London) New Biol.* **238**:35–37.
55. Van Wart, H. E., and D. R. Steinbrink. 1981. A continuous spectrophotometric assay for *Clostridium histolyticum* collagenase. *Anal. Biochem.* **113**:356–365.
56. Vasantha, N., L. D. Thompson, C. Rhodes, C. Banner, J. Nagle, and D. Filpula. 1984. Genes for alkaline protease and neutral protease from *Bacillus amyloliquefaciens* contain a large open reading frame between the regions coding for signal sequence and mature protein. *J. Bacteriol.* **159**:811–819.
57. Vinci, V. A. 1988. Biochemical and genetic analysis of serine proteases of *Streptomyces* spp. Ph.D. thesis, The Ohio State University, Columbus.
58. von Heijne, G. 1986. A new method for predicting signal cleavage sites. *Nucleic Acids Res.* **14**:4683–4690.
59. Wünsch, E., and H.-G. Heidrich. 1963. Zur quantitativen Bestimmung der Kollagenase. *Z. Physiol. Chem.* **333**:149–151.
60. Yang, M. Y., E. Ferrari, and D. J. Henner. 1984. Cloning of the neutral protease gene of *Bacillus subtilis* and the use of the cloned gene to create an in vitro-derived deletion mutation. *J. Bacteriol.* **160**:15–21.