JOURNAL OF BACTERIOLOGY, Aug. 1987, p. 3778–3784
0021-9193/87/083778-07$02.00/0

Vol. 169, No. 8

Characterization and Structure of Genes for Proteases A and B from *Streptomyces griseus*

GRAHAM HENDERSON, PHYLLIS KRYGSMAN, CI JUN LIU, CHERYL C. DAVEY, AND LAWRENCE T. MALEK*

Cangene Corporation, Mississauga, Ontario L4V 1T4, Canada

Received 17 February 1987/Accepted 15 May 1987

Protease A and protease B are extracellular proteins which are secreted by *Streptomyces griseus*. The genes encoding protease A (*sprA*) and protease B (*sprB*) were isolated from an *S. griseus* genomic library by using a synthetic oligonucleotide probe. Fragments containing *sprA* and *sprB* were characterized by hybridization and demonstration of proteolytic activity in *Streptomyces lividans*. Each DNA sequence contains a large open reading frame with the coding region of the mature protease situated at its carboxy terminus. The amino terminus of each reading frame appears to encode a 38-amino-acid signal peptide followed by a 76- or 78-amino-acid polypeptide, a propeptide, which is joined to the mature protease. Strong homology between the coding regions of the protease genes suggests that *sprA* and *sprB* originated by gene duplication.

Streptomyces griseus, an organism used for the commercial production of pronase, secretes many extracellular proteins (13). Protease A and protease B, two of the serine proteases secreted by *S. griseus*, have sequences which are 61% homologous on the basis of amino acid identity (8). These proteases also have similar tertiary structure, as determined by X-ray crystallography (6, 8, 11). Although the structures of proteases A and B have been extensively studied, the genes encoding the proteases have not been characterized.

This report describes the structure of two *S. griseus* genes which are responsible for the expression of protease A and protease B. The DNA sequences suggest that each protease is initially secreted as a precursor which is processed to remove an amino-terminal polypeptide (propeptide) from the mature protease. This propeptide may be important for the secretion of proteases in *Streptomyces* spp. We propose the genetic designations *sprA* and *sprB* for the unmapped genes for (serine) proteases A and B.

MATERIALS AND METHODS

Strains and plasmids. *S. griseus* IMRU3499 was obtained from the Waksman Institute of Microbiology, Rutgers University, Piscataway, N.J. *Streptomyces lividans* 66 (4) and plasmid pIJ702 (14) were from the John Innes Institute. *Escherichia coli* HB101 (ATCC 33694) was used for all transformations. Plasmids pUC8 (28) and pUC18 and pUC19 (21) were purchased from Bethesda Research Laboratories, Inc.

Media, growth, and transformation. Growth of *Streptomyces* mycelium for the isolation of DNA or the preparation of protoplasts was as described previously (10). Protoplasts of *S. lividans* were prepared by lysozyme treatment, transformed with plasmid DNA, and selected for resistance to thiostrepton, as described previously (10). Transformants were screened for proteolytic activity on LB plates (17) containing 30 μg of thiostrepton per ml and 1% skim milk. *E. coli* transformants were grown on YT medium (17) containing 50 μg of ampicillin per ml.

Materials. Oligonucleotides were synthesized by using an Applied Biosystems 380A DNA synthesizer. Columns, phosphoramidites, and reagents used for oligonucleotide synthesis were obtained from Applied Biosystems, Inc., through Technical Marketing Associates. Oligonucleotides were purified by polyacrylamide gel electrophoresis followed by DEAE-cellulose chromatography. Enzymes for digesting and modifying DNA were purchased from New England BioLabs, Inc., and used in accordance with the recommendations of the supplier. The radioisotopes [α-^{32}P]dATP (~3,000 Ci/mmol) and [γ-^{32}P]ATP (~3,000 Ci/mmol) were from Amersham. Thiostrepton was donated by E. R. Squibb & Sons.

Isolation of DNA. Chromosomal DNA was isolated from *S. griseus* as described previously (5), except that sodium dodecyl sarcosinate (final concentration, 0.5%) was substituted for sodium dodecyl sulfate. Plasmid DNA from transformed *S. lividans* was prepared by an alkaline lysis procedure (10). Plasmid DNA from *E. coli* was purified by a rapid boiling method (9). DNA fragments and vectors used for all constructions were separated by electrophoresis on low-melting-point agarose and purified from the molten agarose by phenol extraction and ethanol precipitation (15).

Construction of genomic library. Chromosomal DNA of *S. griseus* 3499 was digested to completion with *Bam*HI and fractionated by electrophoresis on a 0.8% low-melting-point agarose gel. DNA fragments ranging in size from 4 to 12 kilobase pairs (kbp) were isolated from the agarose gel. The plasmid vectors pUC18 and pUC19 were digested with *Bam*HI and treated with calf intestinal alkaline phosphatase (Boehringer Mannheim Biochemicals). The *S. griseus* *Bam*HI fragments (0.3 μg) and vectors (0.8 μg) were ligated in a final volume of 20 μl as described previously (15). Approximately 8,000 tranformants of *E. coli* HB101 were obtained from each ligation reaction.

Subcloning of protease gene fragments. A hybrid *Streptomyces-E. coli* vector was constructed by ligating pIJ702, which had been linearized with *Bam*HI, into the *Bam*HI site of pUC8. The unique *Bgl*II site of this vector was used for subcloning *Bam*HI and *Bgl*II fragments of the protease genes. Other fragments were adapted with *Bam*HI linkers to facilitate ligation into the *Bgl*II site. The hybrid vector, with pUC8 inserted at the *Bam*HI site of pIJ702, was incapable of

* Corresponding author.

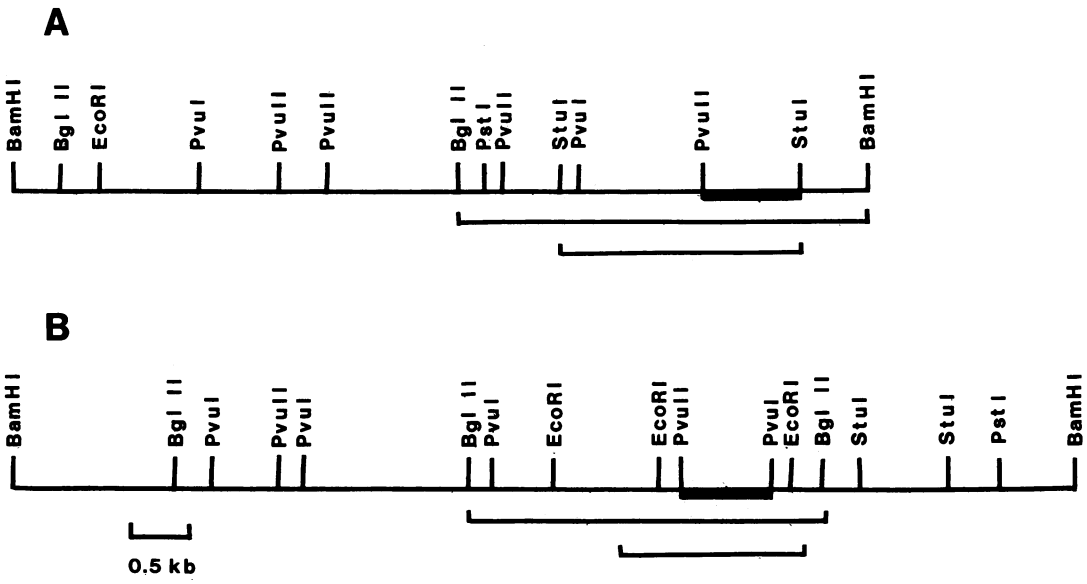

FIG. 1. Restriction endonuclease maps of the 6.8-kbp *Bam*HI fragment containing *sprA* (A) and the 8.4-kbp *Bam*HI fragment containing *sprB* (B). The thick lines indicate the minimum restriction fragments capable of hybridizing to the oligonucleotide probe. Subclones capable of directing secretion of proteolytic activity are represented below each restriction map.

replicating in *S. lividans*. However, the *E. coli* plasmid could be readily removed before transformation of *S. lividans* by digestion with *Bam*HI followed by recircularization with T4 DNA ligase.

Hybridization. A 20-mer oligonucleotide (5'-TTCCC$^{C}_{G}$A ACAACGACTACGG-3') was designed from an amino acid sequence (FPNNDYG) which is common to both proteases. For use as a hybridization probe, the oligonucleotide was end labeled by using T4 polynucleotide kinase (New England BioLabs) and [γ-^{32}P]ATP. Digested genomic or plasmid DNA was transferred to a Hybond-N nylon membrane (Amersham) by electroblotting and hybridized in the presence of formamide (50%) as described previously (10). The filters were hybridized with the labeled oligonucleotide probe at 37°C for 18 h and washed at 47°C. The *S. griseus* genomic library was screened by colony hybridization as described previously (29).

DNA sequencing. The sequences of *sprA* and *sprB* were determined by using a combination of the chemical cleavage (16) and dideoxy (22) sequencing methods. Restriction fragments were end labeled by using either polynucleotide kinase or the large fragment of DNA polymerase I (Amersham), with the appropriate radiolabeled nucleoside triphosphate. Labeled fragments were either digested with a second restriction endonuclease or strand separated, followed by electroelution from a polyacrylamide gel. Subclones were prepared in the M13 bacteriophages mp18 and mp19 (21), and the dideoxy sequencing reactions were run by using the −20 universal primer (New England BioLabs). In some areas of strong secondary structure, compressions and polymerase failure necessitated the use of either inosine (18) or 7-deazaguanosine (19) analogs in the dideoxy reactions to clarify the sequence. The sequences were compiled with the software of DNASTAR (7).

RESULTS

Screening for protease genes by hybridization. Proteases A and B are homologous proteins containing several segments of identical sequence (8). An oligonucleotide probe was designed from one of these consensus amino acid sequences by using the known codon bias for *Streptomyces* spp. (2, 3, 24). The utility of the oligonucleotide probe was demonstrated by hybridization to genomic DNA of *S. griseus*. Since the oligonucleotide was designed from an identical amino acid sequence of both proteases, two DNA fragments should have been detected by hybridization analysis. As anticipated, the probe hybridized equally to two fragments generated by either *Bam*HI (8.4 and 6.8 kbp) or *Bgl*II (11 and 2.8 kbp). The same fragments were detected in the genomic DNA of other *S. griseus* strains, but no such hybridization was observed with DNA from *S. lividans*.

By using the oligonucleotide probe, plasmids containing *sprA* and *sprB* were isolated from a genomic DNA library prepared from *S. griseus* DNA, which had been digested to completion with *Bam*HI. Of 15,000 *E. coli* transformants that were screened by colony blot hybridization, 12 were detected by the oligonucleotide probe and isolated for further characterization. These colonies contained two distinct classes of plasmids, based on restriction analysis. As expected based on the hybridization of genomic DNA, the plasmids contained either a 6.8- or a 8.4-kbp *Bam*HI fragment.

Characterization of protease genes. The DNA fragments isolated by hybridization screening were tested for expression of proteolytic activity. The 6.8- and 8.4-kbp *Bam*HI fragments were ligated into the *Bgl*II site of the vector pIJ702, to allow transformation of *S. lividans*, with selection for thiostrepton resistance (14). Transformants containing these constructions were then tested on a milk plate for secretion of proteases. A clear zone, representing the degradation of milk proteins, surrounded each tranformant that contained either *Bam*HI fragment. The clear zones were not found around *S. lividans* colonies which contained either pIJ702 only or no plasmid. Thus, the two *Bam*HI fragments each appear to contain a protease gene which is capable of effecting secretion in a different *Streptomyces* species.

The particular protease gene contained within each cloned

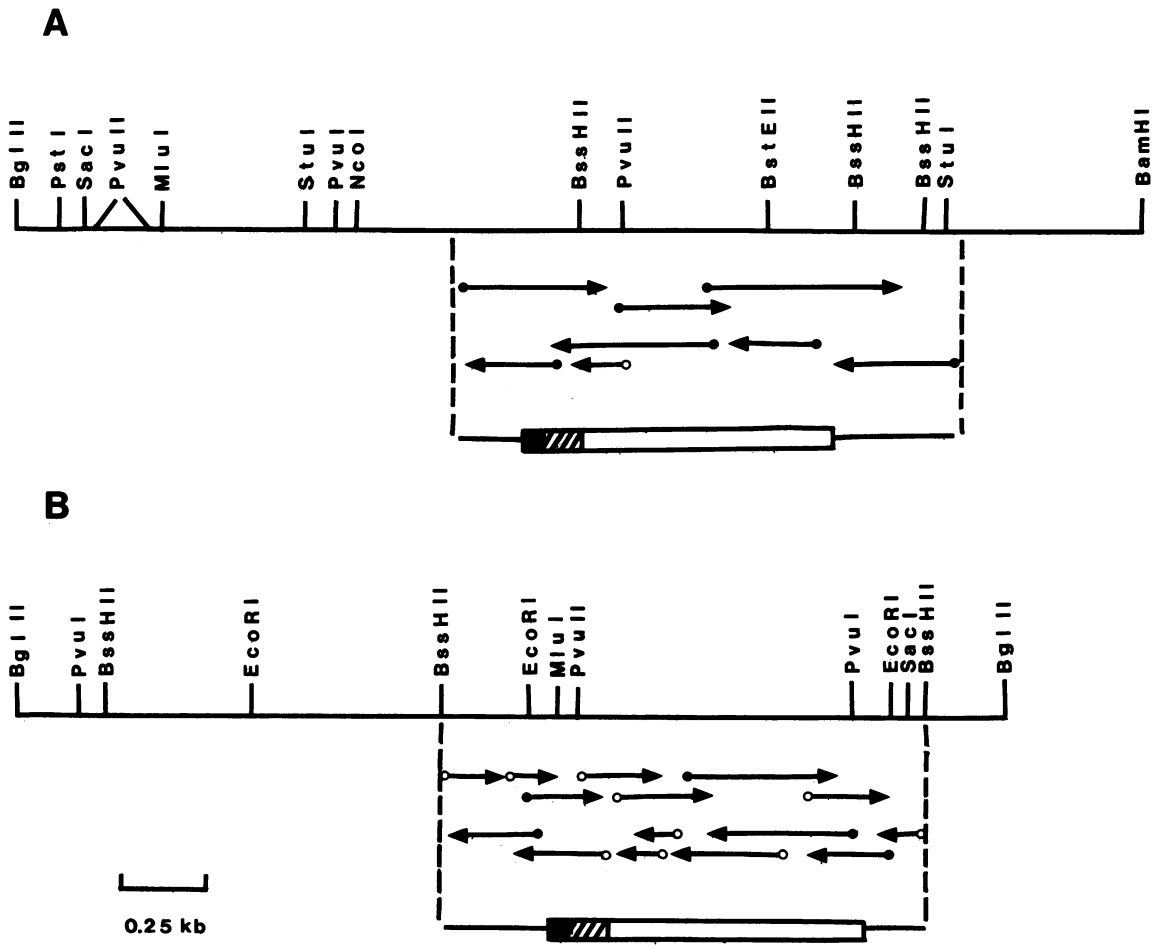

FIG. 2. Restriction endonuclease maps and sequencing strategies for the 3.2-kbp *Bgl*II-*Bam*HI fragment containing *sprA* (A) and the 2.8-kbp *Bgl*II fragment containing *sprB* (B). The arrows indicate the direction and length of the sequence determined by either the dideoxy (●) or chemical cleavage (○) methods. The organization of the structural genes, with signal peptides (■), propeptides (▨), and mature proteases (▭), is shown below each map.

*Bam*HI fragment could be readily determined by dideoxy sequencing of the plasmids by using the oligonucleotide probe as a primer. The 8.4-kbp *Bam*HI fragment was found to contain *sprB*, since a polypeptide deduced from the DNA sequence matched a unique segment of protease B (8). The 6.8-kbp *Bam*HI fragment, which could not be sequenced by this method, was assumed to contain *sprA*.

The protease genes were localized within each *Bam*HI fragment by determining which restriction fragments were capable of hybridizing to the oligonucleotide probe. Detailed restriction maps of the 6.8- and 8.4-kbp *Bam*HI fragments are shown in Fig. 1. Hybridization to the oligonucleotide probe was confined to a 0.9-kbp *Pvu*II-*Stu*I fragment of *sprA* and a 0.6-kbp *Pvu*II-*Pvu*I fragment of *sprB* (Fig. 1). Hybridization to the cloned *Bam*HI fragments and the 2.8-kbp *Bgl*II fragment of *sprB* was in excellent agreement with the hybridization to *Bam*HI and *Bgl*II fragments of genomic DNA. Thus, rearrangement of the *Bam*HI fragments containing the protease genes is unlikely.

The functional limits of the genes were determined by subcloning restriction fragments into pIJ702, transforming *S. lividans*, and testing for proteolytic activity. The 3.2-kbp *Bam*HI-*Bgl*II fragment of *sprA* and the 2.8-kbp *Bgl*II fragment of *sprB*, when subcloned into pIJ702 in either orientation, resulted in the secretion of a protease from *S. lividans*. The intact protease genes could be further delimited to a 1.9-kbp *Stu*I fragment for *sprA* and a 1.4-kbp *Bss*HII fragment for *sprB*. Each of these functionally active subclones (Fig. 1) contained the region which hybridized to the oligonucleotide probe.

DNA sequence of protease genes. The 3.2-kbp *Bam*HI-*Bgl*II fragment of *sprA* and the 2.8-kbp *Bgl*II fragment of *sprB* were subcloned into pUC18 to facilitate further structural characterization. The restriction maps of these subclones and the strategies used to sequence the 1.4-kbp *Sal*I fragment containing *sprA* and the 1.4-kbp *Bss*HII fragment containing *sprB* are shown in Fig. 2. The DNA sequences of these fragments are shown in Fig. 3. The predicted sequence of protease A differed from the published amino acid sequence (8) by the amidation of amino acid 133, whereas that of protease B was identical to the published sequence (8).

Each sequence contains a large open reading frame with the coding region for the mature protease situated at its 3′ end. For each gene, the sequence encoding the carboxy terminus of the protease is followed immediately by a translation stop codon. However, at the other end, the predicted amino acid sequences appear to extend beyond the amino termini of the mature proteases by an additional 116 amino acids for *sprA* and 114 amino acids for *sprB*. The putative GTG initiation codons at these positions are each

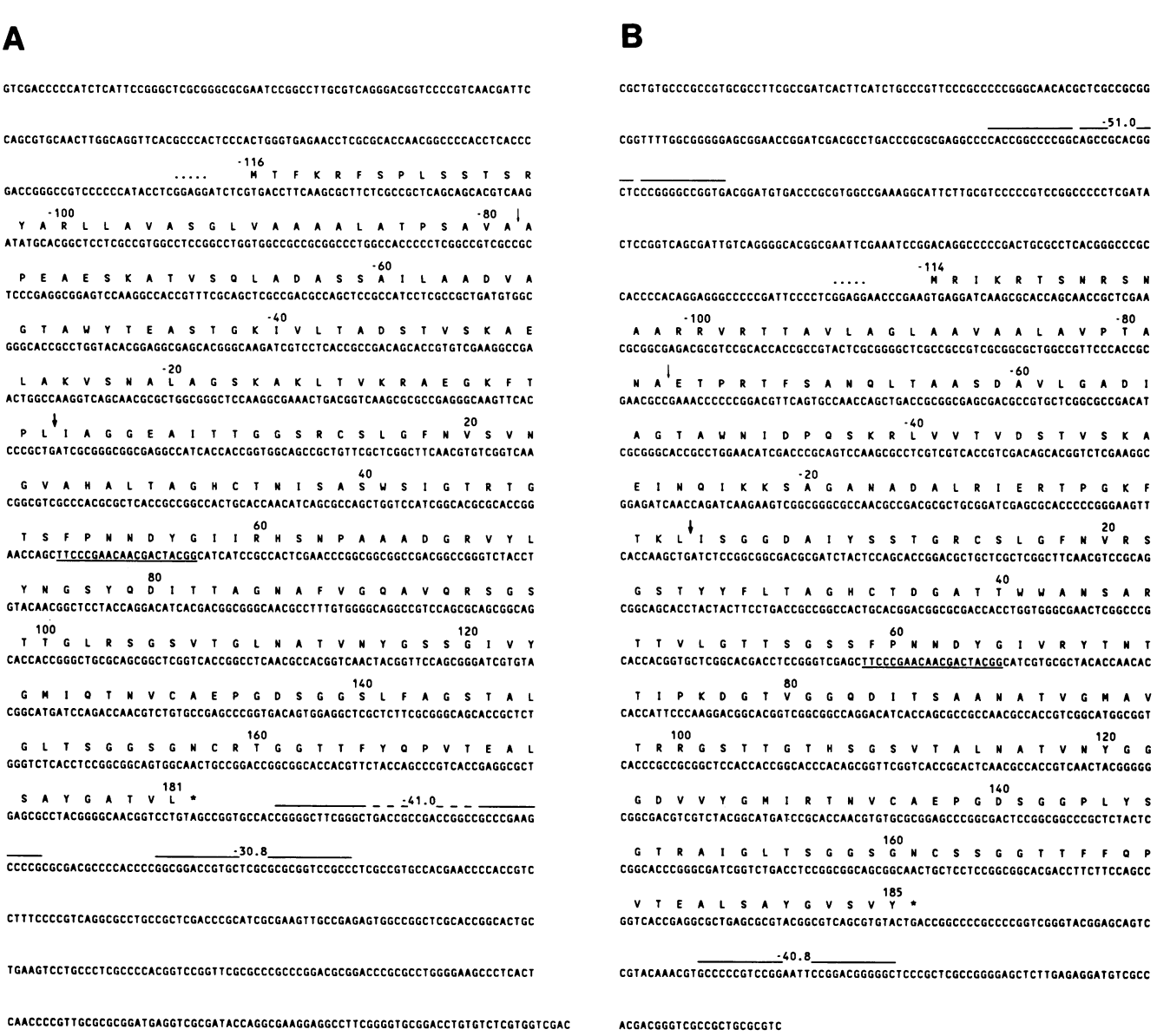

FIG. 3. DNA sequences of *sprA* (A) and *sprB* (B). The deduced amino acid sequences, which are numbered from the amino termini of the mature proteases, are shown above the DNA sequences. Stop codons are indicated by asterisks. The processing sites for the signal peptide (thin arrow) and the mature protease (thick arrow) are shown. Bases involved with the putative ribosome-binding sites are indicated by dots. Inverted repeat sequences are overlined. Indicated above these structures are the stabilities of the hypothetical hairpin loops (free-energy change in kilocalories [1 kcal = 4,184 J] per mole), calculated as described previously (25). DNA sequences which hybridized to the oligonucleotide probe are underlined.

preceded by a potential ribosome-binding site and followed by a sequence which could encode a signal peptide. A processing site for a signal peptidase can be predicted at 38 amino acids from the amino terminus of the putative precursor of each protease. The remaining sequence between the signal processing site and the mature amino terminus of each protease appears to represent a propeptide. The genomic organization of *sprA* and *sprB*, based on the interpretation of the DNA sequence data, is shown schematically in Fig. 2.

The translation stop codons at the end of each coding region are followed by inverted repeated sequences which could form stable hairpin loops. Two of these structures were found after the coding region of *sprA*. The positions of the inverted repeats and their free energies are indicated in Fig. 3. These structures, which are similar to those characterized in other *Streptomyces* genes (3, 24), are believed to be involved with termination of transcription. Interestingly, an inverted repeat was found 178 base pairs before the initiation codon of *sprB*. Preceding this structure was a potential coding region, with the characteristic *Streptomyces* codon usage, which was followed in frame by a TGA stop codon. If the inverted repeated sequence preceding the *sprB* coding region represents a transcription terminator, then the entire transcription unit of *sprB* would be defined. This corroborates the functional activity of the *Bss*HII subclone.

DISCUSSION

Codon usage and base content. The codon selection for *sprA* and *sprB* (Table 1), which is similar to that of other

TABLE 1. Codon usage in *sprA* and *sprB* genes

Amino acid	Codon	No. of codons in:		Amino acid	Codon	No. of codons in:	
		sprA	*sprB*			*sprA*	*sprB*
F	TTT	1	0	Y	TAT	1	0
	TTC	7	7		TAC	9	10
L	TTA	0	0	*[a]	TAA	0	0
	TTG	0	0		TAG	1	0
L	CTT	0	0	H	CAT	0	0
	CTC	11	8		CAC	3	2
	CTA	0	0	Q	CAA	0	0
	CTG	11	7		CAG	6	5
I	ATT	0	1	N	AAT	0	0
	ATC	11	12		AAC	13	17
	ATA	0	0	K	AAA	1	0
M	ATG	1	2		AAG	8	8
V	GTT	1	1	D	GAT	1	0
	GTC	13	16		GAC	6	12
	GTA	0	1	E	GAA	1	1
	GTG	8	6		GAG	7	4
S	TCT	0	0	C	TGT	2	0
	TCC	8	10		TGC	2	4
	TCA	1	0	*	TGA	0	1
	TCG	9	7	W	TGG	2	3
P	CCT	0	0	R	CGT	0	0
	CCC	4	5		CGC	7	13
	CCA	0	0		CGA	0	0
	CCG	4	4		CGG	3	3
T	ACT	0	0	S	AGT	2	1
	ACC	22	33		AGC	16	12
	ACA	0	0	R	AGA	1	1
	ACG	10	7		AGG	0	1
A	GCT	3	0	G	GGT	4	2
	GCC	28	17		GGC	27	30
	GCA	2	1		GGA	2	1
	GCG	14	20		GGG	4	4

[a] *, Stop codon.

sequenced *Streptomyces* genes (2, 3, 24), shows a clear bias for the use of C or G in position 3. The G+C content for each position with a codon is normal (54 to 60%) for positions 1 and 2 but anomalously high (92 to 96%) for position 3 (Table 2). In comparison, the base composition outside of the proposed open reading frames is uniformly distributed within each nucleotide triplet. The maintenance of a biased codon usage throughout each coding region and the disappearance of the triplet periodicity outside of the reading frames strengthens the assignment of initiation and termination codons for the proposed protease precursors.

Homology of *sprA* and *sprB*. The alignment of the amino acid sequences translated from the coding regions of the *sprA* and *sprB* genes (Fig. 4) indicates an overall homology of 54% based on amino acid identity. However, the sequence homology is not uniformly distributed throughout the coding regions (Table 3). The carboxy-terminal domains of the proteases (6) are 75% homologous, whereas the average

TABLE 2. G+C content of *sprA* and *sprB*

Position in codon	G+C content (%) in region:					
	Before coding[a]		Coding		After coding[b]	
	sprA	*sprB*	*sprA*	*sprB*	*sprA*	*sprB*
1	75	71	57	54	84	75
2	61	69	59	60	82	77
3	66	66	92	96	73	66
All	67	69	70	70	80	73

[a] 177 base pairs preceding the initiation codon.
[b] 132 base pairs following the termination codon.

A

```
sprA    MTFKRFSPLSSTSRYARLLAVASGLVAAAALATPSAVA
        M  KR S  S   R  R  AV  GL A AALA P A A
sprB    MRIKRTSNRSNAARRVRTTAVLAGLAAVAALAVPTANA
```

B

```
sprA    APEAESKATVSQLADASSAILAADVAGTAWYTEASTGKI
              A   QL  AS A L AD AGTAW
sprB    ETPRTFSAN--QLTAASDAVLGADIAGTAWNIDPQSKRL

sprA    VLTADSTVSKAELAKVSNALAGSKAK-LTVKRAEGKFTPL
        V T DSTVSKAE        AG  A  L   R  GKFT L
sprB    VVTVDSTVSKAEINQIKKS-AGANADALRIERTPGKFTKL
```

C

```
sprA    IAGGEAITTGGSRCSLGFNVSVNGVAHALTAGHCTNIS
        I GG AI     RCSLGFNV        LTAGHCT
sprB    ISGGDAIYSSTGRCSLGFNVRSGSTYYFLTAGHCTDGA

sprA    ASWS--------IGTRTGTSFPNNDYGIIRHSNPAAA-
          W          GT  G SFPNNDYGI R  N
sprB    TTWWANSARTTVLGTTSGSSFPNNDYGIVRYTNTTIPK

sprA    DGRVYLYNGSYQDITTAGNAFVGQAVQRSGSTTGLRSG
        DG V    G  QDIT A NA VG AV R GSTTG  SG
sprB    DGTV----GG-QDITSAANATVGMAVTRRGSTTGTHSG

sprA    SVTGLNATVNYGSSGIVYGMIQTNVCAEPGDSGGSLFA
        SVT LNATVNYG    VYGMI TNVCAEPGDSGG L
sprB    SVTALNATVNYGGGDVVYGMIRTNVCAEPGDSGGPLYS

sprA    GSTALGLTSGGSGNCRTGGTTFYQPVTEALSAYGATVL
        G  A GLTSGGSGNC  GGTTF QPVTEALSAYG  V
sprB    GTRAIGLTSGGSGNCSSGGTTFFQPVTEALSAYGVSVY
```

FIG. 4. Alignment of amino acid sequences deduced from *sprA* and *sprB*. The precursor proteases encoded by *sprA* and *sprB* are shown as indicated. The middle lines indicate the consensus sequences. Aligned separately are the signal peptides (A), propeptides (B), and mature proteases (C).

homology for the remainder of the proposed protease precursors is only 45%. Interestingly, the amino-terminal domains and the propeptides are similar in both the extent of homology and the distribution of consensus sequences. The unexpectedly high DNA sequence homology relative to that of the protein sequences was partially due to the 61% conservation in position 3 of each codon. The close homology between the *sprA* and *sprB* genes suggests that both originated by duplication of a common ancestral gene.

Function of the protease precursor. The putative signal peptides predicted by the sequences of *sprA* and *sprB* are similar to those from other gram-positive bacteria (1). Each has an extended, positively charged amino-terminus of 17 amino acids followed by a hydrophobic core of 16 amino acids and a beta-bend region with a predicted processing site. The function of the putative 38-amino-acid signal peptide was tested by replacing the *sprB* coding region, following the signal processing site, with a sequence encoding the mature amylase from *S. griseus*. In this construction, the *sprB* signal peptide was as effective as the natural amylase signal peptide in directing the secretion of correctly processed extracellular amylase (P. Krygsman, G. Henderson, L. Escote-Carlson, and L. Malek, manuscript in preparation).

TABLE 3. Homology of *sprA* and *sprB* coding regions

Coding region	Length (no. of codons)[a]	Protein homology (%)[b]	DNA homology (%)
Signal peptide	38	50	58
Propeptide	79	43	62
NT protease[c]	87	45	58
CT protease[d]	103	75	75
Total protease	190	61	67
Total	307	55	65

[a] Some values represent the aligned length with gaps included.
[b] Based on identical amino acids from alignment shown in Fig. 4.
[c] Amino termini of mature proteases (amino acids 1 to 87).
[d] Carboxy termini of mature proteases (amino acids 88 to 190).

The propeptide of the predicted protease precursor is probably not required for the initial membrane translocation stage of secretion, since this function was adequately performed by the signal peptide. However, the propeptide could be involved with facilitating secretion at a later stage, such as crossing the cell wall. The overall length and hydrophilic character of the propeptide could allow it to assume an extended conformation through a peptidoglycan matrix. However, amino acid homology suggests that the propeptide is not simply an inert spacer spanning the cell wall but probably has a definite, conserved structure. The propeptide may interact with the mature protease portion of the precursor, possibly inhibiting its activity until secretion is complete. Consistent with this hypothesis is the occurrence of the same consensus amino terminus, IXGG (Fig. 4C), in many eucaryotic serine proteases which have amino-terminal activation peptides (23, 27, 31).

The processing of each precursor protease is probably autocatalytic, since the amino acid sequence near the scissile bond should provide a good substrate (12, 20). Less likely would be the involvement of host proteases, which would have to be present in both *Streptomyces* species. The precursor proteases predicted from the structures of *sprA* and *sprB* are similar in general organization to those of the alkaline and neutral proteases of *Bacillus* spp., which have amino-terminal propeptides of 77 and 194 amino acids, respectively (26, 30). Also common is the proposed autocatalytic processing of the propeptides from the mature proteases. These similarities suggest that a precursor with an amino-terminal propeptide may be a general feature of proteases which are secreted by gram-positive bacteria.

ACKNOWLEDGMENTS

We are grateful to Robert T. Garvin and Eric James for their support of this work, helpful discussions, and critical review of the manuscript. We thank Gisela Soostmeyer for screening *Streptomyces* strains and preparing and testing transformants of *S. lividans*. We also acknowledge Donald Stewart and Burton Pogell for both helpful discussions and assays of proteolytic activity.

LITERATURE CITED

1. **Abrahmsen, L., T. Moks, B. Nilsson, U. Hellman, and M. Uhlen.** 1985. Analysis of signals for secretion in the staphylococcal protein A. EMBO J. 4:3901–3906.
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. **Bibb, M. J., M. J. Bibb, J. M. Ward, and S. N. Cohen.** 1985.

Nucleotide sequences encoding and promoting expression of three antibiotic resistance genes indigenous to *Streptomyces*. Mol. Gen. Genet. **199**:26–36.

4. **Bibb, M. J., J. L. Schottel, and S. N. Cohen.** 1980. A DNA cloning system for interspecies gene transfer in antibiotic-producing *Streptomyces*. Nature (London) **284**:526–531.
5. **Chater, K. F., D. A. Hopwood, T. Kieser, and C. J. Thomson.** 1982. Gene cloning in *Streptomyces*. Curr. Top. Microbiol. Immunol. **96**:69–95.
6. **Delbaere, L. T. J., W. L. B. Hutcheon, M. N. G. James, and W. E. Thiessen.** 1975. Tertiary structural differences between microbial serine proteases and pancreatic serine enzymes. Nature (London) **257**:758–763.
7. **Doggette, P. E., and F. R. Blattner.** 1986. Personal access of sequence databases on personal computers. Nucleic Acids Res. **14**:611–619.
8. **Fujinaga, M., L. T. J. Delbaere, G. D. Brayer, and M. N. G. James.** 1985. Refined structure of α-lytic protease at 1.7 A resolution; analysis of hydrogen bonding and solvent structure. J. Mol. Biol. **183**:479–502.
9. **Holmes, D. S., and M. Quigley.** 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114**:193–197.
10. **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. Schrempf.** 1985. Genetic manipulation of *Streptomyces*. A laboratory manual. The John Innes Foundation, Norwich, United Kingdom.
11. **James, M. N. G., A. R. Sielecki, G. D. Brayer, L. T. J. Delbaere, and C.-A. Bauer.** 1980. Structures of product and inhibitor complexes of *Streptomyces griseus* protease A at 1.8 A resolution. J. Mol. Biol. **144**:43–88.
12. **Johnson, P., and L. B. Smillie.** 1971. The disulfide bridge sequences of a serine protease of wide specificity from *Streptomyces griseus*. Can. J. Biochem. **49**:548–562.
13. **Jurasek, L., P. Johnson, R. W. Olafson, and L. B. Smillie.** 1971. An improved fractionation system for pronase on CM-Sephadex. Can. J. Biochem. **49**:1195–1201.
14. **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.
15. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Maxam, A., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74**:560–564.
17. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 433. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Mills, D. R., and F. R. Kramer.** 1979. Structure independent nucleotide sequence analysis. Proc. Natl. Acad. Sci. USA **76**:2232–2235.
19. **Mizusana, S., S. Nishimura, and F. Seela.** 1986. Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. Nucleic Acids Res. **14**:1319–1324.
20. **Narahashi, Y., and K. Yoda.** 1973. Studies on proteolytic enzymes (pronase) of *Streptomyces griseus* K-I. III. Purification and proteolytic specificity of alkaline proteinase C from pronase. J. Biochem. **73**:831–841.
21. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene **26**:101–106.
22. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
23. **Steffens, G. J., W. A. Gunzler, F. Oetting, E. Frankus, and L. Flohe.** 1982. The complete amino acid sequence of low molecular mass urokinase from human urine. Hoppe-Seyler's Z. Physiol. Chem. **363**:1043–1058.
24. **Thompson, C. J., and G. S. Gray.** 1983. Nucleotide sequence of a streptomycete aminoglycoside phosphotransferase gene and its relationship to phosphotransferases encoded by resistance plasmids. Proc. Natl. Acad. Sci. USA **80**:5190–5194.
25. **Tinoco, I., P. N. Borer, B. Dengler, M. Livine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) New Biol. **246**:40–41.
26. **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.
27. **Vehar, G. A., W. J. Kohr, W. F. Bennett, D. Pennica, C. A. Ward, R. N. Harkins, and D. Collen.** 1984. Characterization studies on human melanoma cell tissue plasminogen activator. Bio/Technology **2**:1051–1057.
28. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
29. **Wallace, R. B., M. J. Johnson, T. Hirose, T. Miyake, E. H. Kawashima, and K. Itakura.** 1981. The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequence of rabbit globin DNA. Nucleic Acids Res. **9**:879–894.
30. **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.
31. **Young, C. L., W. C. Barker, C. M. Tomaselli, and M. O. Dayhoff.** 1978. Serine proteases, p. 73–93. *In* M. O. Dayhoff (ed.), Atlas of protein sequence and structure, vol. 5, suppl. 3. National Biomedical Research Foundation, Washington, D.C.