JOURNAL OF BACTERIOLOGY, June 1986, p. 937–944
0021-9193/86/060937-08$02.00/0

Vol. 166, No. 3

Specific Excretion of *Serratia marcescens* Protease through the Outer Membrane of *Escherichia coli*

NOBORU YANAGIDA,* TAKESHI UOZUMI, AND TERUHIKO BEPPU

Department of Agricultural Chemistry, The University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Received 26 November 1985/Accepted 13 March 1986

A DNA fragment of *Serratia marcescens* directing an extracellular serine protease (M_r, 41,000) was cloned in *Escherichia coli*. The cloned fragment caused specific excretion of the protease into the extracellular medium through the outer membrane of *E. coli* host cells in parallel with their growth. No excretion of the periplasmic enzymes of host cells occurred. The cloned fragment contained a single open reading frame of 3,135 base pairs coding a protein of 1,045 amino acids (M_r 112,000). Comparison of the 5′ nucleotide sequence with the N-terminal amino acid sequence of the protease indicated the presence of a typical signal sequence. The C-terminal amino acid of the enzyme was found at position 408, as deduced from the nucleotide sequence. Artificial frameshift mutations introduced into the coding sequence for the assumed distal polypeptide after the C terminus of the protease caused complete loss of the enzyme production. It was concluded that the *Serratia* serine protease is produced as a 112-kilodalton proenzyme and that its N-terminal signal peptide and a large C-terminal part are processed to cause excretion of the mature protease through the outer membrane of *E. coli* cells.

Escherichia coli, a gram-negative bacterium, has two distinct membrane systems, the outer membrane and the inner cytoplasmic membrane. Several enzymes of *E. coli* such as alkaline phosphatase and β-lactamase are localized in the space between these two membranes, which is called the periplasm. The N-terminal signal coding sequence found in the periplasmic protein genes is involved in their secretion through the cytoplasmic membrane. However, the outer membrane functions as a barrier to prevent further secretion into the extracellular medium. The signal sequences of these periplasmic proteins have been used for the construction of vectors which can facilitate the secretion of cloned heterologous proteins into the periplasm but not into the extracellular medium (2, 6, 10, 17).

However, several observations have suggested that direct excretion of proteins through the outer membrane is possible in some cases, even in *E. coli*. Hemolysin is excreted directly into media from *E. coli* cells (25), and a cloned penicillinase of an alkalophilic *Bacillus* sp. has been shown to be excreted from host *E. coli* cells at the late-stationary phase (11, 12). Also, a lysis gene has been identified which facilitates extracellular excretion of colicins from host cells (4).

Serratia marcescens, a gram-negative bacterium belonging to the family *Enterobacteriaceae*, is known to produce large amounts of extracellular metalloprotease into the surrounding medium (18). During the course of cloning the protease, we obtained an *E. coli* clone which produced a hitherto unknown serine protease of *Serratia* origin. In this study we examined the cloning of this *Serratia* serine protease and its excretion through the outer membrane of *E. coli*. The nucleotide sequence of the cloned DNA and the N- and C-terminal amino acid sequences of the excreted protease suggested processing of the long C-terminal peptide of a large proenzyme, as well as processing of the N-terminal signal peptide during the excretion of the mature protease through the outer membrane.

* Corresponding author.

MATERIALS AND METHODS

Bacterial strains and plasmids. *S. marcescens* IFO-3046 was used as a DNA donor; and *E. coli* C600 (*hsdR hsdM leu thr thi supE*), *E. coli* KO-60 (K strain wild type, this laboratory), and *E. coli* JM105 [Δ(*lac pro*) *thi rpsL endA sbcB15 hsdR4* F′ *traD36 proAB lacI*^{q} *lacZ*Δ*M15*] (purchased from Amersham Co. Ltd., England) were used as hosts. pBR322 was used as a cloning vector. pMC1403 was used to detect promoter activity of the cloned fragment (3).

Media. *S. marcescens* and *E. coli* were grown in Penassay broth (Difco Laboratories, Detroit, Mich.) or in L broth aerobically at 30 and 37°C, respectively. *E. coli* transform-

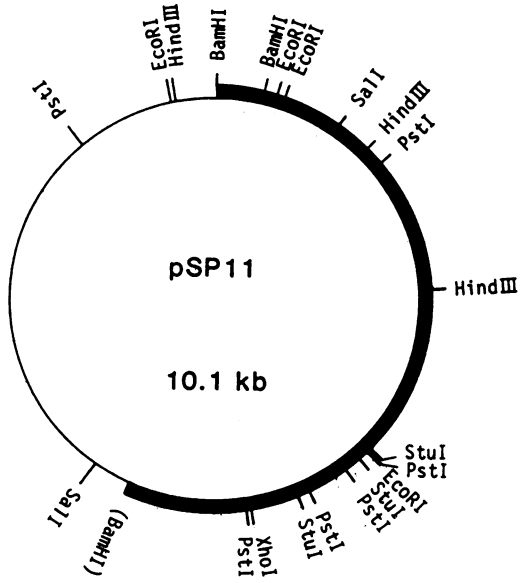

FIG. 1. Restriction map of pSP11. The thick line represents the 5.8-kb chromosomal DNA insert from *S. marcescens* IFO-3046, and the thin line represents pBR322.

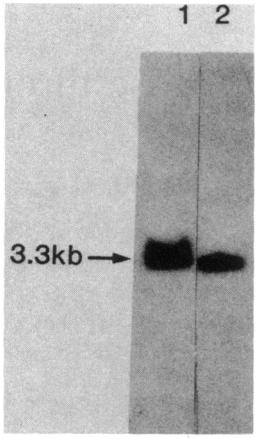

FIG. 2. Southern blot DNA-DNA analysis. Hybridization analysis of the Southern transfer of the DNAs from the electrophoresed 1.0% agarose gel. The ^{32}P-labeled 3.3-kb *Eco*RI fragment was used as a probe. Lane 1, *S. marcescens* IF0-3046 chromosomal DNA digested with *Eco*RI; lane 2, pSP11 digested with *Eco*RI.

ants were plated on Penassay broth containing 50 μg of ampicillin (Sigma Chemical Co., St. Louis, Mo.) per ml, 1% skim milk, and 1.2% agar; and halo-forming colonies were detected after 12 h of cultivation at 37°C.

Recombinant DNA techniques. Chromosomal DNA of *S. marcescens* IFO-3046 was purified by the method of Marmur (14). Preparation of a large amount of plasmid was performed by the method of Tanaka et al. (26). For small-scale plasmid isolation, the method of Holmes and Quigley (8) was employed. *Serratia* chromosomal DNA was partially digested with *Sau*3AI, and 3- to 20-kilobase (kb) fragments were collected from electrophoresed agarose gel. The fragments were ligated to the linear pBR322 DNA cleaved with *Bam*HI and dephosphorylated with bacterial alkaline phosphatase. Transformation of *E. coli* with the ligated mixture was performed by the method of Norgard et al. (19). Various restriction endonucleases and other enzymes were purchased from Takara Shuzo Co. Ltd., Kyoto, Japan.

Southern blot analysis. Transfer of DNA fragments from agarose gel to nitrocellulose paper was performed as described by Southern (24). Hybridization conditions and labeling of DNA for hybridization probes by nick translation were as described by Rigby et al. (21).

DNA sequencing. Specific restriction fragments of the cloned DNA were ligated into the appropriate M13 vector mp10 or mp11 (15) and sequenced by the chain-terminating dideoxy method (23). dGTP was replaced by dITP in some reaction mixtures to resolve compression of bands in the gels due to stable secondary structures in unusually GC-rich regions (1).

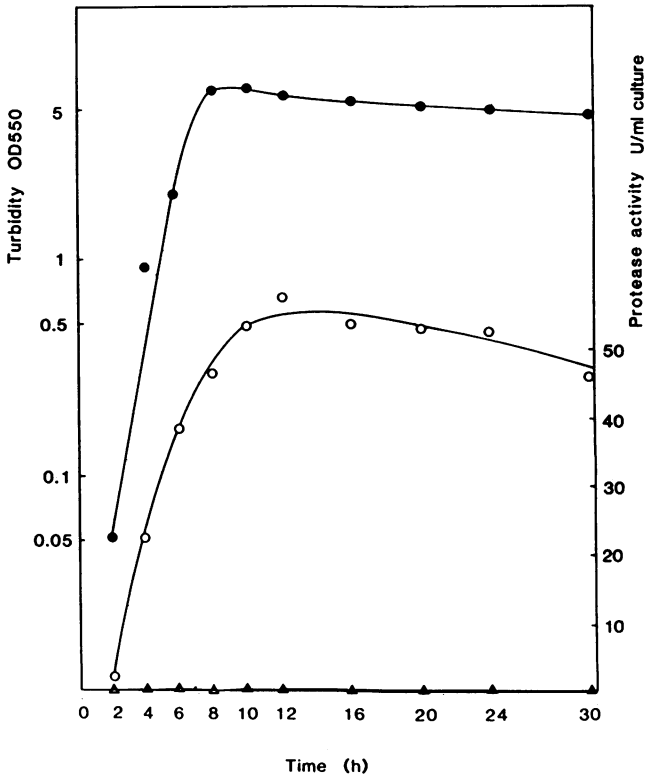

FIG. 3. Bacterial growth and protease production by *E. coli* KO-60 carrying pSP11. *E. coli* KO-60(pSP11) was inoculated into L broth containing 50 μg of ampicillin per ml and cultured aerobically at 37°C. Bacterial growth (OD_{550}, optical density at 550 nm; [A_{550}; ●]; and protease activity (units per milliliter of culture, extracellular [○]; periplasmic [△] were determined.

Assay of enzyme activities. Protease activity was assayed by the modified method of Hagihara et al. (7). A total of 700 μl of prewarmed 0.6% casein in 50 mM phosphate buffer (pH 7.5)–100 μl of enzyme solution incubated for 30 min at 37°C. The reaction was stopped by adding 700 μl of stop solution containing 0.11 M trichloroacetic acid, 0.22 M acetic acid, and 0.33 M sodium acetate. After the reaction mixture was allowed to stand for 30 min at room temperature, it was centrifuged, and the A_{280} of the supernatant was measured. One unit of the protease activity was defined as the amount causing an increase in absorbance of 0.1 during the 30-min incubation. To measure the effect of inhibitors, the enzyme was preincubated with the inhibitors for 30 min at 37°C.

TABLE 1. Localization of the protease and other enzymes in *E. coli* KO-60(pSP11) and *E. coli* (pBR322)[a]

Enzymes[b]	Plasmids carried	Extracellular activity (U)	Periplasmic activity (U)	Cytoplasmic activity (U)
Protease	pSP11	39.1 (76)[c]	2.02 (4)	10.1 (20)
	pBR322	0 (0)	0 (0)	9.9 (100)
β-Lactamase	pSP11	40 (14)	224 (80)	16 (6)
	pBR322	35 (4)	852 (93)	28 (3)
β-Galactosidase	pSP11	0.013 (4)	0.026 (8)	0.273 (88)
	pBR322	0.001 (0.3)	0.007 (2)	0.265 (97)

[a] Both strains were grown aerobically in L broth for 8 h at 37°C.
[b] Enzymatic activities of protease, β-Lactamase, and β-Galactosidase were measured as described in the text.
[c] Values in parentheses indicate percent distribution of the enzymes in each strain.

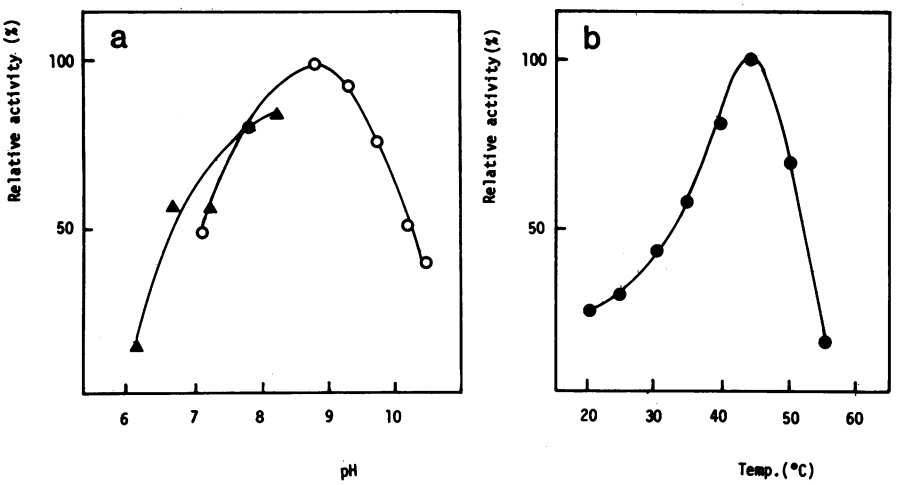

FIG. 4. Optimum pH (a) and optimum temperature (b) of the cloned protease activity. (a) Effect of pH was examined in 50 mM Na_2HPO_4–NaH_2PO_4 buffer (pH 6.1 to 8.2 [▲]) and 50 mM NH_4OH–NH_4Cl buffer (pH 7.1 to 10.5 [○]) at 37°C. (b) Effect of temperature was examined in 50 mM Na_2HPO_4–NaH_2PO_4 buffer (pH 7.5).

Activities of β-lactamase and β-galactosidase were measured by the method of macroiodometry (22) and by the method of Miller (16), respectively.

Fractionation of extracellular, periplasmic, and cytoplasmic enzymes in *E. coli* cells. Fractionation was performed by the method of Cornelis et al. (5). *E. coli* cells were harvested at the early-stationary phase and washed twice with 10 mM Tris hydrochloride buffer (pH 7.5) containing 25% sucrose. The washed cells were suspended in the same buffer containing 25% sucrose and 1 mM EDTA, and the suspension was shaken for 10 min at room temperature. After centrifugation at 7,000 × *g* for 10 min, the cells were quickly and vigorously suspended in ice-cold water. The suspension was further shaken for 10 min at 4°C and centrifuged at 9,000 × *g* for 10 min. The precipitated cells were suspended in 10 mM Tris hydrochloride buffer (pH 7.5) and disrupted by sonication. The extracellular enzyme fraction was the sum of the activities in the culture supernatant, the two washes, and the supernatant of the EDTA treatment. The periplasmic fraction was the activity found in the supernatant after treatment with cold water. The cytoplasmic fraction was the activity found in the supernatant after sonication.

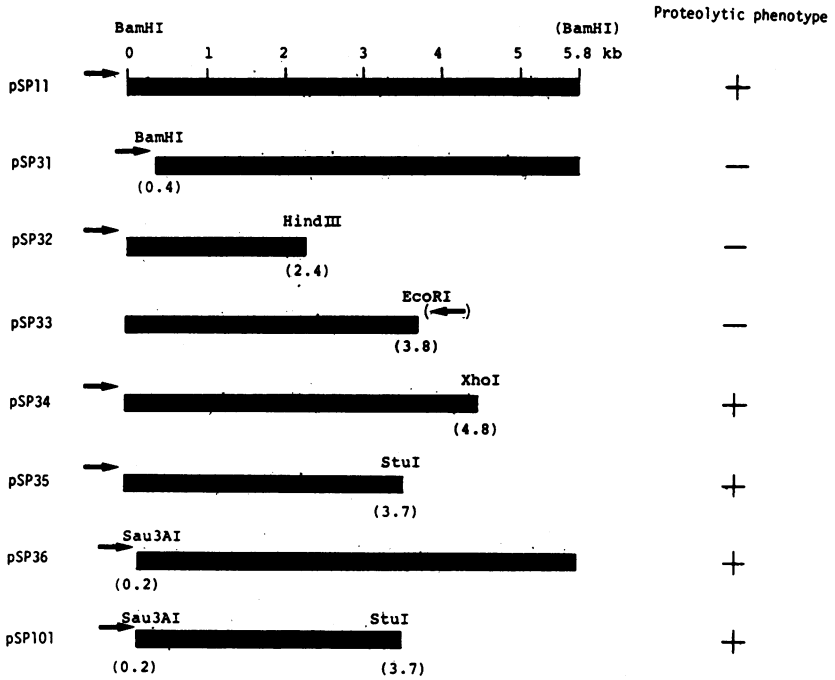

FIG. 5. Subcloning of the DNA coding the *Serratia* serine protease. Arrows indicate the direction of transcription from the tetracycline promoter on the vector. The arrow in parentheses for pSP33 indicates the opposite orientation of the insert, although the tetracycline promoter on the vector was deleted in this construction. The proteolytic phenotype was determined by halo formation on 1% skim milk-agar medium. Numbers in parentheses are in kilobases.

```
                                                                BamHI/Sau3AI                -240
                                                                   GATCC TTACGCCCCG TCCCGATGGT CGGTCATTCC GGCGCAACGG

-220                -200                -180                -160                -140                -120
GCGAAACGTA CATTCGCCTC CTCTCCAGAC GTCCTGCGCC ATAGCCGCGG CGCGATGGCA TCCCTTTGTC TGTCCGCGCC CATTTGCATC TCTCAGCATG TCTGTTCGTG

          -100                -80                 -60                 -40                 -20
CGGCCCGCGC CGGGTCCGCA TGTGTTATCC GTCGTTTCAC GATCATAGGA GGTTTCAGCG CGGTTAAAAC GTAAGTTTTC TTCTGGGTAT AAATAAACAT CGAGCGTAAA

1
ATG ATA CTT AAT AAA AGA TTG AAG TTA GCG TAT TGC GTT TTT CTG GGT TGT TAT GGC TTA TCC ATT CAT TCT TCT CTT GCC GCT TAT CAG
met ile leu asn lys arg leu lys leu ala tyr cys val phe leu gly cys tyr gly leu ser ile his ser ser leu ala ala tyr gln
 1                                  10                                  20                                  30

GAT CCC GGT CGA TTG GGC GCG CCT GAC AGT TGG AAA ACC GCG GAG TTT AAT CGC CAA TGG GGG CTT GAA GCT ATT TCC GCC GAA TTC GCC
asp pro gly arg leu gly ala pro asp ser trp lys thr ala glu phe asn arg gln trp gly leu glu ala ile ser ala glu phe ala
                                   40                                  50                                  60

TAT GCC AGA GGC TAT ACC GGA AAA GGT ATA ACC ATC GGC GTT ATC GAT AAC GCT ATT CTT TCC CAT TCT GAA TTC TCC GGT AAA CTG ACG
tyr ala arg gly tyr thr gly lys gly ile thr ile gly val ile asp asn ala ile leu ser his ser glu phe ser gly lys leu thr
                                   70                                  80                                  90

CGC CTG GAT AAC GGC AGT TAT AAT TTC TCG TAT GAT AAA CAA GAT AAT ATG TCT TTC GGC GAT CAC GGC ACG CAC GTA GCC GGT ATC GCT
arg leu asp asn gly ser tyr asn phe ser tyr asp lys gln asp asn met ser phe gly asp his gly thr his val ala gly ile ala
                                  100                                 110                                 120

GCG GCT AAA AGA GAT GGC GCA GGC ATG CAC GGC GTC GCT TTC GAC GCG GAT ATT ATC GGC ACC AAA TTG AAT GAT TAC GGT AAT CGC AAC
ala ala lys arg asp gly ala gly met his gly val ala phe asp ala asp ile ile gly thr lys leu asn asp tyr gly asn arg asn
                                  130                                 140                                 150

GGC CGT GAA GAG CTG ATT CAG AGC GCG GCT CGC GTC ATC AAT AAC AGC TGG GGG ATC GCG CCG GAT ATC CGG CGA GAC GCC AAA GGC GAT
gly arg glu glu leu ile gln ser ala ala arg val ile asn asn ser trp gly ile ala pro asp ile arg arg asp ala lys gly asp
                                  160                                 170                                 180

ATC ATC TGG TTG CCG AAC GGC AGG CCG GAC TAC GTG GCA TTC GTA AAA AGC GAA GTG ATC GCC GAG ATG ATG CGC AGC AAA TCC AGC GTG
ile ile trp leu pro asn gly arg pro asp tyr val ala phe val lys ser glu val ile ala glu met met arg ser lys ser ser val
                                  190                                 200                                 210

GAA TGG GGC AGC GAA CAA CCG GTG CCC ACC GGC GGG CAC AGC GCC ATG TCG ACG CTG CTG CGG GCG GCC AGG CAC GGC AAG CTG ATC GTC
glu trp gly ser glu gln pro val pro thr gly gly his ser ala met ser thr leu leu arg ala ala arg his gly lys leu ile val
                                  220                                 230                                 240

TTC TCG GCG GGC AAT TAC AAC AAT TAC AAT ATT CCG GAA GCG CAA AAG TCA CTG CCT TAC GCT TTC CCG GAC GTA CTG AAT AAT TAC CTG
phe ser ala gly asn tyr asn asn tyr asn ile pro glu ala gln lys ser leu pro tyr ala phe pro asp val leu asn asn tyr leu
                                  250                                 260                                 270

ATT GTG ACC AAC CTG AGT GAC GAA AAT CAG TTA AGC GTT TCC TCG ACC AGC TGC GGG CAA ACG GCC AGC TAT TGT GTT TCC GCG CCG GGT
ile val thr asn leu ser asp glu asn gln leu ser val ser ser thr ser cys gly gln thr ala ser tyr cys val ser ala pro gly
                                  280                                 290                                 300

TCT GAT ATT TAC AGC ACC GTC GGC CGG CTG GAG TCC AAT ACC GGC GGC GCC GTG AAT CGC GAA GCT TAT AAT AAG GGC GAG CTG TCG CTT
ser asp ile tyr ser thr val gly arg leu glu ser asn thr gly gly ala val asn arg glu ala tyr asn lys gly glu leu ser leu
                                  310                                 320                                 330

AAT CCT GGT TAC GGC AAC AAA TCC GGC ACC TCA ATG GCG GCG CCG CAT GTG ACC GGC GTC GCC GCC GTA TTG ATG CAG CGC TTC CCA TAT
asn pro gly tyr gly asn lys ser gly thr ser met ala ala pro his val thr gly val ala ala val leu met gln arg phe pro tyr
                                  340                                 350                                 360

ATG AGC GCC GAT CAA ATA TCT GCA GTC ATT AAA ACC ACC GCC ACC GAT TTA GGC GTT GCC GGC ATT GAT AAT TTA TTC GGC TGG GGG CGC
met ser ala asp gln ile ser ala val ile lys thr thr ala thr asp leu gly val ala gly ile asp asn leu phe gly trp gly arg
                                  370                                 380                                 390

GTC AAT TTA CGC GAT GCG ATC AAC GGC CCG AAA ATG TTT ATT ACC AAA GAG GAT ATC CCG CAG GAA TAT TAT GTG CCG GGT TCC TAC AGT
val asn leu arg asp ala ile asn gly pro lys met phe ile thr lys glu asp ile pro gln glu tyr tyr val pro gly ser tyr ser
                                  400                                 410                                 420

GAG AAA CAG TTT GTG GTG AAT ATC CCC GGC CTG GGA AAT ATC GTC GAA CCC GGC ACG CCG GTT GAG CGG CGC TGC ACG TCG AGC GAA TGC
glu lys gln phe val val asn ile pro gly lue gly asn ile val glu pro gly thr pro val glu arg arg cys thr ser ser glu cys
                                  430                                 440                                 450

AGT TTC GAT TCG TGG AGT AAC GAC ATC AGC GGG CAC GGC GGC TTG ACC AAA ACG GGC GCC GGT ACG CTG GCG CTG TTG GGT AAT AAC ACC
ser phe asp ser trp ser asn asp ile ser gly his gly gly leu thr lys thr gly ala gly thr leu ala leu leu gly asn asn thr
                                  460                                 470                                 480

TAT CGC GGT GAT ACC TGG GTG AAA CAG GGC GTG TTG GCG ATC GAC GGT TCG GTG GCA TCC AAC GTC TAT ATC GAA AAT AGC GGC ACA TTG
tyr arg gly asp thr trp val lys gln gly val leu ala ile asp gly ser val ala ser asn val tyr ile glu asn ser gly thr leu
                                  490                                 500                                 510

TCC GGC GAG GGC ACC GTA GGC GCC TTC AGA GCG GCT CGG AGC GGC AGC GTC GCG CCG GGC AAC GGC ATC GGC ACG CTG CAT GTG TTG CAC
ser gly glu gly thr val gly ala phe arg ala ala arg ser gly ser val ala pro gly asn gly ile gly thr leu his val leu his
                                  520                                 530                                 540
```

```
GAT GCC ATT TTT GAT CGC GGT TCG CAG TAT AAC GTG GAA GTG GCG GAC AAT GGC CGC AGC GAT AAG ATC GCC GCG CGG CGC GCT TTC CTC
asp ala ile phe asp arg gly ser gln tyr asn val glu val ala asp asn gly arg ser asp lys ile ala ala arg arg ala phe leu
                                  550                                 560                                 570

AAC GGC GGC AGT GTG AAT GTC AGC CTG GAA CGC AGC CAA AAC CTG CTG TCG CAG AAT GAG GCG CAG AGT CTG CTG GGC AAC AAG TAC ACC
asn gly gly ser val asn val ser leu glu arg ser gln asn leu leu ser gln asn glu ala gln ser leu leu gly asn lys tyr thr
                                  580                                 590                                 600

ATC CTG ACC ACG ACG GAC GGC GTT ACC GGT AGA TTC GAA AAC GCC AAC CCA TCG TAT CCG TTT GTT AAA GTC GCG CTG GAT TAT CGG GGC
ile leu thr thr thr asp gly val thr gly arg phe glu asn ala asn pro ser tys pro phe val lys val ala leu asp tyr arg gly
                                  610                                 620                                 630

AAT GAC GTC GGC CTT GGC ATC ACG CGC ACC GAC GCC AGC TTT GAC AGC CTG GCC AGC ACT GAG AAC GAG AAA GCG GTA GCT CGT GCG GTG
asn asp val gly leu gly ile thr arg thr asp ala ser phe asp ser leu ala ser thr glu asn glu lys ala val ala arg ala val
                                  640                                 650                                 660

GAG ACG CTC AAC GCG ACG GAA CCG GTC ACG GAA ACG GCC AAA CGC AGC GTG GCG ATC CCG GCG GCC GAA GAG GCC AAC CTG CTG CAA AGC
glu thr leu asn ala thr glu pro val thr glu thr ala lys arg ser val ala ile pro ala ala glu glu ala asn leu leu gln ser
                                  670                                 680                                 690

GAT GGG GGC GAG GCG CAA GCC GTG AAC GAA GAG GCG AGC ATC GTG GCG GGC CAT CCG ATC TAT GAA AGC TTC CTC GGC TTC ACC TCG GCC
asp gly gly glu ala gln ala val asn glu glu ala ser ile val ala gly his pro ile tyr glu ser phe leu gly phe thr ser ala
                                  700                                 710                                 720

AGA GAA TTG CAA CAG GCG ACC CGT CAA CTG TCC GGC CAG ATC CAC GCG GAT ATG GCT TCC GCC CAG ATC AAC GAA AGC CGT TAC CTG CGC
arg glu leu gln gln ala thr arg gln leu ser gly gln ile his ala asp met ala ser ala gln ile asn glu ser arg tyr lue arg
                                  730                                 740                                 750

GAT ACC GCC ACC GAG CGC TTG CGC CAG GCG GAA GGC CGC CGC ACC GCT ACC GAC ATT AAA GCG GAT GAC AAC GGC GCC TGG GCG AAA CTG
asp thr ala thr glu arg leu arg gln ala glu gly arg arg thr ala thr asp ile lys ala asp asp asn gly ala trp ala lys leu
                                  760                                 770                                 780

CTG GGT AGC TGG GGG CAT GCT TCC GGC AAC GAC AAC GCC ACC GGT TAC CAG ACC TCC ACC TAT GGC GTG CTG TTA GGT CTG GAC AGC GAA
leu gly ser trp gly his ala ser gly asn asp asn ala thr gly tyr gln thr ser thr tyr gly val leu leu gly leu asp ser glu
                                  790                                 800                                 810

CTG TTT GGC GAC GGC CGG CTT GGC ATG ATG ACC GGG TAT ACC CGC ACT TCG CTG GAT GGA GGT TAT CAG TCA GAT GCT CAC AGC GAC AAC
leu phe gly asp gly arg leu gly met met thr gly tyr thr arg thr ser leu asp gly gly tyr gln ser asp ala his ser asp asn
                                  820                                 830                                 840

TAC CAT CTG GGG CTG TAC GGC GAC AAA CGC TTC GGC GCG TTG GCG CTG CGA GCG GGC GGC ACC TAT ACC TGG CAT CGC ATC GAC ACC TCG
tyr his leu gly leu tyr gly asp lys arg phe gly ala leu ala leu arg ala gly gly thr tyr thr trp his arg ile asp thr ser
                                  850                                 860                                 870

CGT TCG GTG AAC TAC GGC GCG CAG TCG GAT CGC GAG AAG GCC AAG TAT AAC GCG CGC ACC GGT CAG CTG TTC ATC GAA AGC GGC TAC GAT
arg ser val asn tyr gly ala gln ser asp arg glu lys ala lys tyr asn ala arg thr gly gln leu phe ile glu ser gly tyr asp
                                  880                                 890                                 900

TGG ACG AGC GAT GCG GTC AAC CTT GAG CCG TTC GCC AAC CTG GCG TAT ACC CAT TAC CGT AAC GAG GAG ATC AAC GAG CAA GGC GGG GCA
trp thr ser asp ala val asn leu glu pro phe ala asn leu ala tyr thr his tyr arg asn glu glu ile asn glu gln gly gly ala
                                  910                                 920                                 930

GCG GCG CTG CGC GGC GAC AAA CAA AGT CAG TCC GCC ACC GCC TCG ACG TTG GGT CTG CGC GCC GAC ACC GAG TGG CAA ACC GAC AGC GTG
ala ala leu arg gly asp lys gln ser gln ser ala thr ala ser thr leu gly leu arg ala asp thr glu trp gln thr asp ser val
                                  940                                 950                                 960

GCG ATC GCG CTG CGC GGC GAG CTG GGT TGG CAG CAT CAG TAC GGC AAG CTG GAG CGT AAA ACG CAG CTG ATG TTC AAA CGC ACT GAT GCG
ala ile ala leu arg gly glu leu gly trp gln his gln tyr gly lys leu glu arg lys thr gln leu met phe lys arg thr asp ala
                                  970                                 980                                 990

GCG TTC GAC GTG AAC AGC GTG CCT GTT TCT CGC GAT GGC GCG ATT CTG AAA GCG GGC GTC GAT GTA TCG ATT AAC AAA AAC GCC GTC CTG
ala phe asp val asn ser val pro val ser arg asp gly ala ile leu lys ala gly val asp val ser ile asn lys asn ala val leu
                                 1000                                1010                                1020

TCC CTT GGC TAC GGC GGG CAG CTG TCG TCC AAC CAC CAG GAC AAC AGC GTC AAC GCC GGT CTG ACC TGG CGC TTC TGA TCGCGGT
ser leu gly tyr gly gly gln leu ser ser asn his gln asp asn ser val asn ala gly leu thr trp arg phe TRM
                                 1030                                1040

TGCATGCTTT GCTTTATTCA GCCCTCATCG AGACGATGGG GGCTTTTTTT TCGCGGCCAG CCGGCCTTCT GCGGCATGAG CGATACTTTC TCCGCGACCT ATTCCCCCGG

CCGGTGTTCG CCGCCGGGTT AATCAATCAG AGGAAAACAA ACGCTGCTGC CAATAGCGCA ATTGCGTCAG CGTGATCGGC GCAGCGGCGG CAATGACCGC GCGATTTTCC

                                    StuI
AACGCAATCT CCCGCGCACG TTGCAACGGC TCGGCGCTGA GGCCT
```

FIG. 6. Nucleotide sequence of the *Serratia* serine protease gene. The large box (left) indicates the sequence corresponding to the mature protease deduced from the N- and C-terminal amino acid sequences. Asp^{76}, His^{112}, and Ser^{341}, which predict the active site of this serine protease, are boxed. The position of the possible ribosome-binding sequence is underlined. A typical ρ-independent transcription termination region is indicated by the arrows facing each other.

TABLE 2. Purification of the *Serratia* serine protease

	Total activity (U)	Total protein (mg)	Sp act (U/μg of protein)	Yield (%)
Culture supernatant	337,500	500	0.675	100
Ammonium sulfate precipitate	300,000	297	1.010	89
After DEAE column	220,000	9.8	22.4	65
After G100 column	170,000	6.6	25.6	50

Purification of the cloned *Serratia* protease. Five liters of the stationary-phase culture of *E. coli* C600(pSP11) was centrifuged at 9,000 × *g* for 15 min, and the supernatant was treated with solid ammonium sulfate to give 70% saturation. After standing for 24 h at 4°C, the precipitates were collected by centrifugation and dissolved in 50 ml of 10 mM Tris hydrochloride (pH 7.0). The solution was dialyzed three times against 30 liters of the same buffer, and the dialysate was applied to a column (50 by 100 mm) of DEAE-cellulose DE 52. The enzyme was eluted with 400 ml of a linear gradient of NaCl (0 to 400 mM) in the same buffer. The eluted enzyme fraction was concentrated with a Diaflo membrane (PM30; Amicon Corp., Lexington, Mass.), and 30 to 3 ml and was subjected to gel filtration on a Sephadex G 100 column (30 by 900 mm). Purification steps were monitored by sodium dodecyl sulfate- (SDS)-polyacrylamide gel electrophoresis (13). The molecular weight of the protease was estimated by gel filtration by using a high-performance liquid chromatograph (Shimazu LC-4A; Shimazu Co., Ltd., Kyoto, Japan) equipped with a Shodex PROTEIN WS-803 column (Showa Denko Co. Ltd., Tokyo, Japan) and SDS-polyacrylamide gel electrophoresis.

N- and C-terminal amino acid sequencing. The N-terminal amino acid sequence of the purified *Serratia* protease excreted by *E. coli* C600(pSP11) was determined with a gas-phase amino acid sequencer (system 890D; Beckman Instruments, Inc., Fullerton, Calif.) reverse-phase high-performance liquid chromatograph equipped with a Senshu Pak SEQ-4 column (Senshu Scientific Co. Ltd., Tokyo, Japan) for analysis of the phenylthiohydantoin amino acids.

The C-terminal amino acid sequence of the protease was determined by using the acid carboxypeptidase of *Aspergillus saitoi* (9) and an amino acid analyzer (automatic amino acid analyzer 835-30; Hitachi Co., Ltd., Tokyo, Japan) by the method of Takeuchi and Ichishima (M. Takeuchi and E. Ichishima, Agric. Biol. Chem., in press).

Artificial frameshift mutation by linker insertion. pSP11 DNA (10.1 kb) was partially digested with *Hin*cII or *Pvu*II, and linear DNAs of 10.1 kb were collected by agarose gel electrophoresis. An 8-mer *Bgl*II linker $\frac{CAGATCTG}{GTCTAGAC}$ (obtained from Nippon Zeon Co. Ltd.) was ligated to these linear DNAs and successively digested with *Bgl*II. After removal of the small fragments, the full-length linear plasmids were religated and transformed into *E. coli*. The insertion site in each plasmid was identified by determining loss of the restriction site after complete digestion with *Hin*cII or *Pvu*II.

RESULTS

Cloning of *S. marcescens* DNA coding an extracellular protease. The partial *Sau*3AI fragments of chromosomal DNA of *S. marcescens* IFO-3046 were inserted into the *Bam*HI site of pBR322, and the ligated DNA was transformed into *E. coli* C600. Among 25,000 transformants, 2 formed a distinct turbid zone around their colonies on the plate containing 1% skim milk. Colonies of the parental *Serratia* strain formed a large clear zone because of the extracellular metalloprotease (18). The turbid halo of the transformants suggested extracellular production of a different protease which caused precipitation of casein by limited proteolysis. A recombinant plasmid containing a 5.8-kb segment of insert DNA was recovered from both transformants and designated pSP11. Restriction analysis resulted in the map of pSP11 shown in Fig. 1. Southern blot DNA-DNA hybridization of the *Eco*RI-digested chromosomal DNA of *S. marcescens* with ^{32}P-labeled 3.3-kb *Eco*RI fragments within the insert used as probes, showed a positive band identical to that of the cloned fragment (Fig. 2).

Excretion of the cloned *Serratia* protease from *E. coli* cells. When *E. coli* KO-60 harboring pSP11 was cultured in liquid L broth aerobically at 37°C, the cells produced the protease in the medium in parallel with their growth (Fig. 3). Almost all the protease activity was found in the extracellular fraction of the cells at the early-stationary phase, while most of the β-lactamase and β-galactosidase activities were localized in the periplasm and the cytoplasm, respectively (Table 1). No accumulation of the protease in the periplasm was observed even at the beginning of the excretion in the early growth phase (Fig. 3). The maximum amount of the protease accumulated in the medium was estimated to be 4 μg/ml.

Enzymatic properties of the cloned *Serratia* protease. The protease produced by *E. coli* C600(pSP11) in the medium was purified to give an almost single band on SDS-polyacrylamide gel electrophoresis (Table 2). The purified protease was completely inhibited by 1 mM phenylmethylsulfonyl fluoride (PMSF) but not by 10 mM EDTA. In contrast, the major extracellular protease of *S. marcescens* IFO-3046 was inhibited by EDTA but not by PMSF. When the *S. marcescens* strain was cultured in the presence of 1 mM EDTA, a trace amount of the PMSF-sensitive protease was detected in the medium. These observations indicate that the cloned protease is a hitherto unknown serine protease, a small amount of which is produced extracellularly by *S. marcescens* with a large amount of the extracellular metalloprotease.

The optimum pH of the serine protease was between 8.0 and 9.0 (Fig. 4a), and its optimum temperature was 45°C (Fig. 4b). The molecular weight of the enzyme was estimated to be approximately 43,000 from gel filtration and about 58,000 from SDS-polyacrylamide gel electrophoresis.

Subcloning and sequencing of the DNA coding the *Serratia* serine protease. Subcloning experiments were performed to identify the minimum region of the insert DNA necessary for directing the protease. The results (Fig. 5) indicate that the 3.7-kb region from the *Bam*HI site at the end of the insert to the *Stu*I site nearest from the *Bam*HI site was essential.

The total base sequence of this region was determined and was found to contain a single open reading frame of 3,135 base pairs (bp) directing a polypeptide composed of 1,045 amino acid residues (Fig. 6).

The plasmid pSP33 carrying the essential coding region of the protease in the opposite direction failed to produce the enzyme (Fig. 5). When the 360-bp *Bam*HI fragment containing the 250-bp sequence farthest upstream from the Shine-Dalgarno sequence of the gene was cloned into the promoter probe vector pMC1403 (3), two recombinant plasmids carrying the insert in the opposite direction were obtained. Neither of these recombinants showed any expression of the β-galactosidase gene on the vector. These results indicate that the *Serratia* protease gene is expressed by readthrough from the promoter of the tetracycline resistance gene on

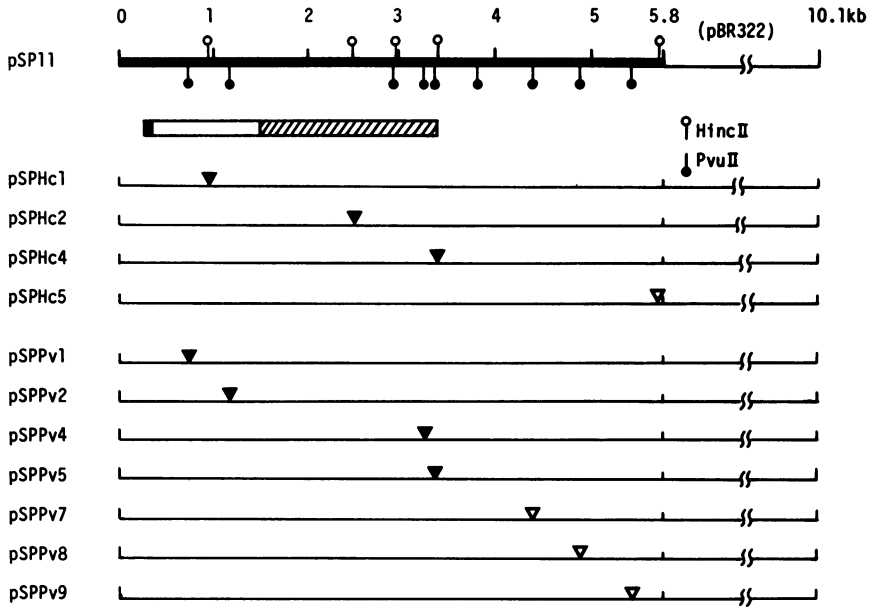

FIG. 7. Effect of artificial frameshift mutations in the cloned *Serratia* DNA fragment. The open reading frame is shown under the restriction map of pSP11. The putative signal peptide, mature enzyme, and distal large polypeptide of the C-terminal side are represented by ■, □, and ▨, respectively. Linker insertion sites causing loss of the protease production and those without effect are indicated by ▼ and ▽, respectively.

pBR322 and that the *Serratia* promoter on the fragment does not function in *E. coli*.

N- and C-terminal amino acid sequences of the *Serratia* serine protease. The purified *Serratia* serine protease excreted by *E. coli* C600(pSP11) was analyzed for its N- and C-terminal amino acid sequences. The N-terminal sequence of Ala-Tyr-Gln-Asp-Pro-Gly-Arg-Leu- was found, which corresponds to the sequence after Ala 28 predicted from the base sequence. The N-terminal sequence from Met 1 to Ala 27 is similar to a typical signal sequence containing proximal basic amino acids following a sequence rich in hydrophobic amino acids.

Two amino acids at the C terminus, Glu and Asp, were determined to exist in the purified extracellular protease. This dipeptide sequence was found only at positions 407 and 408 in the predicted sequence. The polypeptide from Ala 28 to Asp 408 gives a molecular weight of 40,905 which coincides well with that of the protease measured by gel filtration.

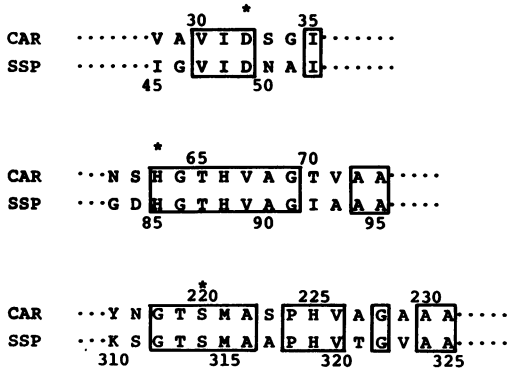

FIG. 8. Partial homology of the amino acid sequences between cloned *Serratia* serine protease (SSP) and subtilisin Carlsberg (CAR). Putative essential catalytic residues are marked with asterisks. The numbers above and below the sequence indicate the amino acid numbers of each mature protein. Homologous amino acid sequences are boxed.

Artificial frameshift mutations in the cloned protease gene. The results described above indicate the possibility that a proenzyme of 112 kb is synthesized in the cells and that processing of not only the N-terminal signal peptide but also the distal large polypeptide on the C-terminal side is involved in excretion of the mature enzyme. To confirm the possible role of this distal processing, artificial frameshift mutations were introduced at various *Hin*cII and *Pvu*II sites in the inserted DNA of pSP11. Not only the mutations in the coding region for the mature protease but also those in the sequence for the distal processed part caused total loss of the protease activity in both the extracellular medium and periplasm of host cells (Fig. 7). On the other hand, no inactivation was observed with mutations outside of the assumed open reading frame of the proenzyme. These results indicate that the distal processed part of the proenzyme is essential for extracellular production of the protease.

DISCUSSION

Protein excretion through the outer membrane of *E. coli* cells was recently reported with a cloned penicillinase of an alkalophilic *Bacillus* strain (11, 12). Excretion of the cloned penicillinase occurred at the late-stationary phase, accompanying excretion of the periplasmic enzymes of the host cells. Some structural damage to the outer membrane was induced by the cloned *Bacillus* DNA or the vector that was used. In contrast, excretion of the cloned *Serratia* serine protease proceeded in parallel with growth of the host without accompanying significant leakage of the host periplasmic enzymes. It seems highly probable that the specific excretion of the protease is directed by characteristic sequences within the protease gene.

The identified open reading frame in the cloned *Serratia* DNA is composed of 3,135 bp coding 1,045 amino acids, which corresponds to a protein of M_r 112,000. Comparison

of the sequence with the N- and C-terminal amino acid sequences of the extracellularly produced protease indicates the processing of a proenzyme of 112 kilodaltons not only at the N-terminal signal sequence but also at the middle of the polypeptide. In the predicted amino acid sequence for the mature enzyme, three possible essential residues for catalytic function, Asp^{49}, His^{85}, and Ser^{314}, can be assigned according to the close homology of their neighboring sequences with those of subtilisin, an extracellular serine protease of *Bacillus subtilis* (Fig. 8) (20). Although we failed to detect the assumed 112-kilodalton proenzyme in the cytoplasm or periplasm of the host cells, loss of the protease production because of frameshift mutations within the distal coding sequence suggests its essential role for extracellular production of the enzyme (Fig. 7). It seems reasonable to assume that the typical N-terminal signal sequence of the putative proenzyme and its processing are involved in secretion through the inner cytoplasmic membrane, while processing of the large C-terminal part of the proenzyme is involved in excretion through the outer membrane. Studies involving site-directed mutation and immunological detection of the protease to identify the functional roles of these sequences in the excretion of the protease are now in progress.

ACKNOWLEDGMENTS

We are grateful to Eiji Ichishima and Michio Takeuchi for helpful discussions and suggestions for the determination of the C-terminal amino acid sequence.

LITERATURE CITED

1. **Barnes, W. M., M. Bevan, and P. H. Son.** 1983. Kilosequencing: creation of an ordered nest of asymmetric deletions across a large target sequence carried on phage M13. Methods Enzymol. **101**:98–122.
2. **Bassford, P. J., T. J. Silhavy, and J. R. Beckwith.** 1979. Use of gene fusion to study secretion of maltose-binding protein into *Escherichia coli* periplasm. J. Bacteriol. **139**:19–31.
3. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1980. In vitro gene fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. **143**:971–980.
4. **Cole, S. T., B. Saint-Joanis, and A. P. Pugsley.** 1985. Molecular characterisation of the colicin E2 operon and identification of its products. Mol. Gen. Genet. **198**:465–472.
5. **Cornelis, P., C. Digneffe, and K. Willemot.** 1982. Cloning and expression of a *Bacillus coagulans* amylase gene in *Escherichia coli*. Mol. Gen. Genet. **186**:507–511.
6. **Emr, S. D., M. N. Hall, and T. J. Silhavy.** 1980. A mechanism of protein localization: the signal hypothesis and bacteria. J. Cell Biol. **86**:701–711.
7. **Hagihara, B., H. Matsubara, M. Nakai, and K. Okunuki.** 1958. Crystalline bacterial proteinase. I. Preparation of crystalline proteinase of *Bac. subtillus*. J. Biochem. **45**:185–194.
8. **Holmes, D. S., and M. Quigley.** 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114**: 193–197.
9. **Ichishima, E., and T. Arai.** 1973. Specificity and mode of action of acid carboxypeptidase from *Aspergillus saitoi*. Biochim. Biophyts. Acta **293**:444–450.
10. **Ito, K., and J. R. Beckwith.** 1981. Role of the mature protein sequence of maltose-binding protein in its secretion across the *E. coli* cytoplasmic membrane. **25**:143–150.
11. **Kato, C., T. Kudo, and K. Horikoshi.** 1984. Gene expression and production of *Bacillus* no. 170 penicillinase in *Escherichia coli* and *Bacillus subtilis*. Agric. Biol. Chem. **48**:397–401.
12. **Kudo, T., C. Kato, and K. Horikoshi.** 1983. Excretion of the penicillinase of an alkalophilic *Bacillus* sp. through the *Escherichia coli* outer membrane. J. Bacteriol. **156**:949–951.
13. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
14. **Marmur, J.** 1961. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J. Mol. Biol. **3**:208–218.
15. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
16. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Miyake, T., T. Oka, T. Nishizawa, F. Misoka, T. Fuwa, K. Yoda, M. Yamasaki, and G. Tamura.** 1985. Secretion of human interferon-α induced by using secretion vectors containing a promoter and signal sequence of alkaline phosphatase gene of *Escherichia coli*. J. Biochem. **97**:1429–1436.
18. **Miyata, K., K. Maejima, K. Tomoda, and M. Isono.** 1970. *Serratia* protease Part 1. Purification and general properties of the enzyme. Agric. Biol. Chem. **34**:310–318.
19. **Norgard, M. V., K. Keen, and J. J. Monahan.** 1978. Factors affecting the transformation of *Escherichia coli* strain 1776 by pBR322 plasmid DNA. Gene **3**:279–292.
20. **Ottensen, M., and I. Svendsen.** 1970. The subtilisins. Methods Enzymol. **19**:199–215.
21. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
22. **Ross, G. W., and C. H. O'Callaghan.** 1975. β-Lactamase assays. Methods Enzymol. **43**:69–85.
23. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
24. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
25. **Springer, W., and W. Goebel.** 1980. Synthesis and secretion of hemolysin by *Escherichia coli*. J. Bacteriol. **144**:53–59.
26. **Tanaka, T., and B. Weisblum.** 1975. Construction of a colicin E1-R factor composite plasmid in vitro: means for amplification of deoxyribonucleic acid. J. Bacteriol. **121**:354–362.