JOURNAL OF BACTERIOLOGY, June 1991, p. 3311–3317
0021-9193/91/113311-07$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 173, No. 11

Cloning and Nucleotide Sequence of the *Vibrio cholerae* Hemagglutinin/Protease (HA/Protease) Gene and Construction of an HA/Protease-Negative Strain

CLAUDIA C. HÄSE AND RICHARD A. FINKELSTEIN*

Department of Molecular Microbiology and Immunology, School of Medicine, University of Missouri—Columbia, Columbia, Missouri 65212

Received 13 December 1990/Accepted 28 March 1991

The structural gene *hap* for the extracellular hemagglutinin/protease (HA/protease) of *Vibrio cholerae* was cloned and sequenced. The cloned DNA fragment contained a 1,827-bp open reading frame potentially encoding a 609-amino-acid polypeptide. The deduced protein contains a putative signal sequence followed by a large propeptide. The extracellular HA/protease consists of 414 amino acids with a computed molecular weight of 46,700. In the absence of protease inhibitors, this is processed to the 32-kDa form which is usually isolated. The deduced amino acid sequence of the mature HA/protease showed 61.5% identity with the *Pseudomonas aeruginosa* elastase. The cloned *hap* gene was inactivated and introduced into the chromosome of *V. cholerae* by recombination to construct the HA/protease-negative strain HAP-1. The cloned fragment containing the *hap* gene was then shown to complement the mutant strain.

Vibrio cholerae O group 1 classical and El Tor biotype strains, as well as non-O1 serotypes, produce several hemagglutinins (9, 14, 19, 22) which may be involved in adherence of the vibrios to the human gut. In particular, the soluble hemagglutinin, which also has proteolytic activity (16) and is called hemagglutinin/protease (HA/protease), may play a role in the pathogenesis of cholera (for reviews see references 8, 10, and 11). This secreted HA/protease has been purified and characterized as a zinc-dependent metalloprotease (6) with the ability to cleave several physiologically important substrates, including mucin, fibronectin, and lactoferrin (15). It also nicks and thus activates the A subunit of the cholera enterotoxin and related enterotoxins (7). We have previously reported the similarities between HA/protease and the elastase of *Pseudomonas aeruginosa* (20). This paper describes the cloning of the structural gene for the HA/protease. We also report the DNA sequence and deduced amino acid sequence for the protease gene and describe the construction and complementation of an HA/protease-negative strain of *V. cholerae*. The deduced amino acid sequence is compared with that of *P. aeruginosa* elastase.

MATERIALS AND METHODS

Bacterial strains, plasmids, media, and culture conditions. The plasmids utilized are described in Table 1. *Escherichia coli* DH5α (Bethesda Research Laboratories, Gaithersburg, Md.) was used as the host for recombinant plasmids, and cells were grown at 37°C in Luria broth medium (24). *E. coli* JM101 (35) was the host strain for recombinant M13 bacteriophage (Bethesda Research Laboratories). *V. cholerae* O1 Ogawa serotype El Tor biotype strain 3083 (17) and HAP-1, an HA/protease-negative derivative of 3083 (this study), were grown at 37°C on meat extract agar medium containing 1.5% skim milk (Difco). All liquid cultures were grown in Luria broth at 37°C with vigorous aeration. Antibiotics (Sigma Chemical Co., St. Louis, Mo.) were added at the following concentrations (in micrograms per milliliter): ampicillin, 50; kanamycin, 100; chloramphenicol, 25; tetracycline, 15; and apramycin, 10. Isopropyl β-D-thiogalactopyranoside (IPTG, Sigma) was used at 0.1 mM; 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal, Sigma) was used at 40 μg/ml.

DNA preparation, manipulation, and analysis. *V. cholerae* chromosomal DNA was prepared as described previously (2). Plasmid DNA was extracted from *E. coli* DH5α cells by the alkaline lysis method of Birnboim and Doly (4). Standard techniques (24) were used in the construction of recombinant plasmids. Restriction enzymes were generally obtained from Promega (Madison, Wis.), and digestions were done with buffers provided by the suppliers under the recommended conditions. Restriction fragments were electrophoresed in horizontal 1% agarose gels in TAE buffer (24) and stained with ethidium bromide (0.5 μg/ml). For isolation of restriction fragments, the DNA was purified from agarose gels by using the GeneClean Glassmilk kit (Bio101 Inc., La Jolla, Calif.). Radiolabelled oligonucleotide probes were generated by kinase reactions with [γ-^{32}P]ATP (NEN Research Products, Boston, Mass.) and T4 polynucleotide kinase (Promega) as described previously (1a). Purified restriction fragments were radiolabelled with [α-^{32}P]dCTP (NEN) by using a random oligonucleotide extension kit (Pharmacia LKB, Piscataway, N.J.). Southern blot transfers were performed in 20× SSC (3 M NaCl, 0.3 M Na_3citrate, pH 7.0) as described previously (24). The nitrocellulose filters were baked at 80°C for 1 h, prehybridized (5× SSC, 5× Denhardt's solution, 50% formamide, 25 mM sodium phosphate) at 42°C for 2 h, and hybridized (5× SSC, 1× Denhardt's solution, 45% formamide, 25 mM sodium phosphate) at 42°C overnight. The filters were washed twice at room temperature (1× SSC, 1% sodium dodecyl sulfate [SDS]) and twice at 65°C (0.1× SSC, 1% SDS) and then exposed to XRP5 film (Kodak).

Genomic library construction and screening. Size-fractionated chromosomal libraries from *V. cholerae* 3083 were generated in the plasmid vector pGEM3Zf(+) or pACYC 184. Vibrio chromosomal DNA was digested with an appropriate restriction enzyme, and then several DNA fractions of

* Corresponding author.

TABLE 1. Plasmids used

Plasmid	Description[a]	Source
pACYC184	Cml^r Tet^r	G. McDonald
pGEM3Zf(+)	Amp^r *lacZ* ColE1	Promega
pTB42	Apr^r ColE1	M. Calcutt
M13mp18/19	*lacZ*	Bethesda Research Laboratories
pCH1	Amp^r, 700-bp *Sal*I fragment in pGEM3Zf(+)	This study
pCH2	Cml^r, 3.2-kb *Hin*dIII fragment in pACYC184	This study
pCH3	Amp^r Kan^r, 2.4-kb *Xho*I fragment in pCH1	This study
pCH5, pCH6	Amp^r, 3.2-kb *Hin*dIII fragment in pGEM3Zf(+) in both orientations	This study

[a] Abbreviations for drug resistance phenotypes: Cml, chloramphenicol; Tet, tetracycline; Amp, ampicillin; Apr, apramycin; Kan, kanamycin.

the desired size were isolated. The isolated DNA fractions were retested by Southern blot analysis, and positive fractions were ligated (T4 DNA ligase, Bethesda Research Laboratories) to restriction nuclease-cleaved and alkaline phosphatase (Boehringer Mannheim Biochemicals, Indianapolis, Ind.)-treated plasmid DNA. The ligated products were transformed into competent *E. coli* DH5α cells. Resultant drug-resistant colonies were inoculated in duplicate onto sterile nitrocellulose filters on antibiotic-containing media. After growth on the filters overnight the colonies were lysed, and the DNA was denatured and neutralized as described previously (1a). The filters were then treated as for the Southern blots.

Subcloning and DNA sequencing. Suitable restriction fragments were purified and subcloned into pGEM3Zf(+), M13mp18, or M13mp19. Exonuclease III deletions of pCH5 and pCH6 were prepared by using the Erase-a-base system (Promega). The subclones obtained are shown in Fig. 1. Double-stranded sequencing of plasmid DNA, isolated by the alkaline lysis method, was accomplished with universal or reverse oligonucleotide primers, which were synthesized by the University of Missouri DNA Core Facility. Single-stranded DNA from recombinant phages was purified by the method outlined by Amersham Corp. (1) and sequenced with the universal primer. Sequence analyses were carried out by using dideoxy-chain termination reactions (29) with Sequenase (United States Biochemical Corp., Cleveland, Ohio). The sequencing data were analyzed with the computer program of the Genetics Computer Group (Madison, Wis.).

Electroporation of *V. cholerae*. Purified plasmids were introduced into *V. cholerae* with a Gene Pulser apparatus (Bio-Rad Laboratories, Richmond, Calif.) by using a slight modification of a procedure of Stoebner and Payne (32). Fresh mid-log-phase cells were collected, washed several times in SEB (272 mM sucrose, 7 mM sodium phosphate buffer [pH 7.4], 1 mM $MgCl_2$) and resuspended in 1/10 of the original volume on ice. Plasmid DNA and cells were mixed in a cuvette and electroporated (25 μF, 2,500 V). The cells were then grown in Luria broth for 1 h at 37°C, and fractions were plated on antibiotic-containing media and incubated at 37°C overnight.

Western immunoblot analysis. Gel electrophoresis and Western blot analysis were performed as previously described (20). Samples were prepared from 35-ml Luria broth cultures in 200-ml plastic storage bottles (Corning Inc., Corning, N.Y.) inoculated with fresh colonies. After overnight growth with vigorous shaking, the bacteria were pelleted. Samples of the supernatants were concentrated fivefold in a SpeedVac concentrator (Savant), mixed with 2× sample buffer, and boiled for 5 min. Five microliters of each sample was loaded per lane on an SDS–12% polyacrylamide gel. After electrophoresis and transfer to nitrocellulose, the filter was incubated with a 1:1,000 dilution of polyclonal rabbit antiserum raised against purified HA/protease (16).

Hemagglutination. Microtiter quantitation of hemagglutination was performed as described previously (19). "Responder" and "nonresponder" chicken erythrocytes were obtained from individual White Leghorn chickens as previously reported (19); erythrocytes from some chickens (nonresponders) are not agglutinated by the HA/protease, whereas erythrocytes from other chickens (responders) are. The titer was defined as the reciprocal of the highest protein dilution which caused hemagglutination.

Nucleotide sequence accession number. The nucleotide sequence data reported in this paper have been registered (accession number M59466) with GenBank, the EMBL Data Library, and the DNA Data Bank of Japan.

RESULTS

Cloning of the structural gene (*hap*) for the *V. cholerae* HA/protease. The N-terminal amino acid sequence of the purified extracellular HA/protease from *V. cholerae* 3083 has previously been determined and was shown to have striking homology to the N-terminal amino acid sequence of the mature elastase from *P. aeruginosa* (20). The oligonucleotide 5'GGCCCCGGCGGCAACCAGAAG3' derived from the elastase DNA sequence (3) coding for a short N-terminal region which is highly conserved between the two proteins was synthesized. It hybridized strongly to *V. cholerae* 3083

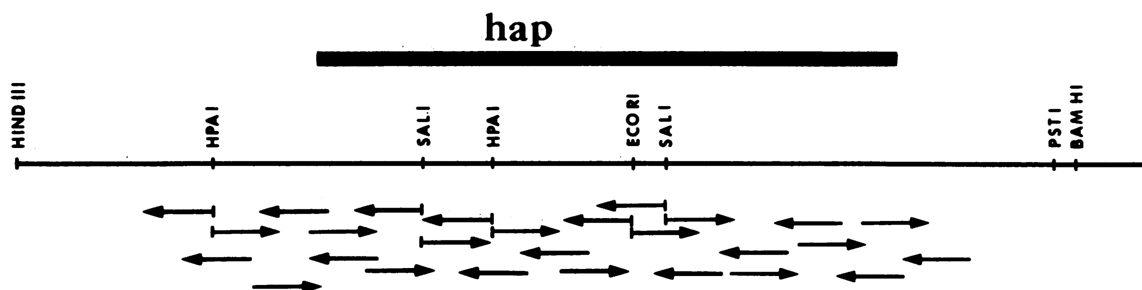

FIG. 1. Restriction map and sequencing strategy for the 3.2-kb DNA fragment containing the *hap* gene. Arrows indicate the direction of sequencing from the subclones. ↦, subclones containing a specific restriction fragment; →, subclones generated by exonuclease digestion.

```
   1  AATTTAAAGGCGCAAACCTATCAGTAAGATTCTATCTTTATGATTTTATTTTCTTTATTTCGG

  64  CTTTCTTATCGAGTATTCCTGCTTACTGAGTAGTTTGTGACACGTAGAGTTCACACTGCTCGT

 127  TAACATTTTTTTAACTCGAATTTGGTGTTTTTCATTGACTTGTACCACTTTGATTGAACGCGG

 190  TATTTCGCGTGCAATCGGCGTTGATGGGAGCGGTTCAACGCTTAATACATTCAACATCTCTAG

      SD  SD
 253  GATTGAGAAATGAAAATGATACAACGTCCTCTGAATTGGTTAGTTCTGGCCGGAGCGGCAACT
                MetLysMetIleGlnArgProLeuAsnTrpLeuValLeuAlaGlyAlaAlaThr      18

 316  GGCTTCCCTCTCTATGCGGCACAAATGGTCACGATTGATGATGCATCAATGGTTGAACAAGCG
      GlyPheProLeuTyrAlaAlaGlnMetValThrIleAspAspAlaSerMetValGluGlnAla     39

 370  TTGGCGCAGCAACAGTACAGTATGATGCCTGCCGCCAGCGGTTTTAAAGCCGTCAATACGGTA
      LeuAlaGlnGlnGlnTyrSerMetMetProAlaAlaSerGlyPheLysAlaValAsnThrVal     60

 442  CAGTTGCCGAATGGTAAGGTGAAAGTGCGTTACCAGCAGATGTACAACGGGGTTCCTGTCTAT
      GlnLeuProAsnGlyLysValLysValArgTyrGlnGlnMetTyrAsnGlyValProValTyr     81

 505  GGCACCGTTGTGGTGGCAACCGAGTCCAGTAAAGGGATTTCGCAAGTGTATGGTCAAATGGCT
      GlyThrValValValAlaThrGluSerSerLysGlyIleSerGlnValTyrGlyGlnMetAla    102

 568  CAGCAGTTGGAAGCCGATCTCCCAACCGTGACCCCTGACATTGAAAGCCAGCAGGCCATCGCT
      GlnGlnLeuGluAlaAspLeuProThrValThrProAspIleGluSerGlnGlnAlaIleAla    123

 631  TTAGCGGTTAGCCATTTTGGTGAACAACACGCTGGAGAATCGCTCCCGGTGGAAAACGAAAGT
      LeuAlaValSerHisPheGlyGluGlnHisAlaGlyGluSerLeuProValGluAsnGluSer    144

 694  GTGCAACTGATGGTACGTTTGGATGATAACCAACAGGCTCAGTTAGTGTACTTGGTCGACTTT
      ValGlnLeuMetValArgLeuAspAspAsnGlnGlnAlaGlnLeuValTyrLeuValAspPhe    165

 757  TTTGTCGCTTCAGAAACACCTTCGCGTCCGTTCTACTTTATCAGTGCGGAAACGGGAGAAGTG
      PheValAlaSerGluThrProSerArgProPheTyrPheIleSerAlaGluThrGlyGluVal    186

 820  CTAGACCAATGGGATGGCATTAACCACGCACAGGCAACAGGAACCGGCCCCGGCGGTAACCAA
      LeuAspGlnTrpAspGlyIleAsnHisAlaGlnAlaThrGlyThrGlyProGlyGlyAsnGln    207

 883  AAAACGGGACGTTATGAATACGGCAGTAACGGTTTACCCGGTTTCACGATTGATAAGACCGGA
      LysThrGlyArgTyrGluTyrGlySerAsnGlyLeuProGlyPheThrIleAspLysThrGly    228

 946  ACCACCTGTACTATGAATAACAGTGCGGTAAAAACCGTTAACCTCAATGGCGGCACCTCGGGT
      ThrThrCysThrMetAsnAsnSerAlaValLysThrValAsnLeuAsnGlyGlyThrSerGly    249

1009  AGCACGGCGTTCAGTTATGCTTGTAACAACAGCACTAACTACAACAGCGTGAAAACAGTGAAT
      SerThrAlaPheSerTyrAlaCysAsnAsnSerThrAsnTyrAsnSerValLysThrValAsn    270

1072  GGTGCTTACTCACCGCTTAACGACGCGCACTTCTTTGGAAAAGTGGTGTTTGATATGTATCAG
      GlyAlaTyrSerProLeuAsnAspAlaHisPhePheGlyLysValValPheAspMetTyrGln    291

1135  CAGTGGTTGAATACTTCGCCGCTGACTTTCCAATTAACCATGCGTGTGCACTACGGCAATAAC
      GlnTrpLeuAsnThrSerProLeuThrPheGlnLeuThrMetArgValHisTyrGlyAsnAsn    312

1198  TATGAAAATGCCTTCTGGGATGGACGCGCCATGACTTTTGGTGATGGCTATACCCGTTTCTAT
      TyrGluAsnAlaPheTrpAspGlyArgAlaMetThrPheGlyAspGlyTyrThrArgPheTyr    333

1261  CCTTTGGTGGATATCAACGTTAGTGCCCATGAGGTCAGCCACGGTTTTACTGAGCAGAATTCA
      ProLeuValAspIleAsnValSerAlaHisGluValSerHisGlyPheThrGluGlnAsnSer    354

1324  GGCCTCGTTTACCGAGATATGTCCGGTGGTATTAACGAAGCATTCTCGGATATCGCAGGGGAA
      GlyLeuValTyrArgAspMetSerGlyGlyIleAsnGluAlaPheSerAspIleAlaGlyGlu    375

1387  GCGGCAGAGTACTTTATGCGTGGCAATGTCGACTGGATTGTCGGCGCGGATATTTTTAAATCC
      AlaAlaGluTyrPheMetArgGlyAsnValAspTrpIleValGlyAlaAspIlePheLysSer    396

1450  TCCGGTGGTCTACGTTATTTCGATCAGCCGTCACGTGATGGCCGCTCGATAGATCATGCTTCA
      SerGlyGlyLeuArgTyrPheAspGlnProSerArgAspGlyArgSerIleAspHisAlaSer    417

1513  CAGTATTACAGCGGTATTGATGTTCACCATTCGAGTGGCGTGTTTAACCGCGCGTTTTACCTA
      GlnTyrTyrSerGlyIleAspValHisHisSerSerGlyValPheAsnArgAlaPheTyrLeu    438

1576  CTCGCCAATAAATCGGGTTGGAACGTACGTAAAGGTTTTGAAGTGTTTGCCGTGGCTAACCAG
      LeuAlaAsnLysSerGlyTrpAsnValArgLysGlyPheGluValPheAlaValAlaAsnGln    459

1639  TTGTACTGGACACCGAACAGCACGTTTGATCAAGGTGGCTGTGGGGTAGTGAAAGCGGCGCAG
      LeuTyrTrpThrProAsnSerThrPheAspGlnGlyGlyCysGlyValValLysAlaAlaGln    480

1702  GATCTCAACTACAACACCGCAGACGTTGTGGCAGCCTTTAATACCGTGGGTGTCAATGCTTCT
      AspLeuAsnTyrAsnThrAlaAspValValAlaAlaPheAsnThrValGlyValAsnAlaSer    501

1765  TGTGGCACCACGCCACCACCTGTCGGCAAAGTGCTTGAGAAAGGTAAACCGATCACAGGACTG
      CysGlyThrThrProProProValGlyLysValLeuGluLysGlyLysProIleThrGlyLeu    522

1828  AGCGGCTCACGTGGAGGAGAAGATTTCTATACCTTTACGGTGACCAATTCAGGCAGTGTTGTT
      SerGlySerArgGlyGlyGluAspPheTyrThrPheThrValThrAsnSerGlySerValVal    543

1891  GTGTCCATCAGTGGTGGAACGGGCGATGCGGATCTGTATGTCAAAGCGGGCAGCAAACCCACC
      ValSerIleSerGlyGlyThrGlyAspAlaAspLeuTyrValLysAlaGlySerLysProThr    564

1954  ACCTCTTCTTGGGATTGTCGTCCATACCGTTCAGGCAATGCCGAGCAGTGTTCCATCTCTGCG
      ThrSerSerTrpAspCysArgProTyrArgSerGlyAsnAlaGluGlnCysPheIleSerAla    585

2017  GTCGTGGGTACGACATACCATGTCATGTTACGCGGTTACAGTAACTATTCTGGTGTGACGTTA
      ValValGlyThrThrTyrHisValMetLeuArgGlyTyrSerAsnTyrSerGlyValThrLeu    606

2080  CGCTTGGACTAACTTCCTTGCCACCTACCTGCAACGCCTCAGCAAAGTCTGAGGGCGTTGTTT
      ArgLeuAsp*                                                         609

2143  TTGAAGGGCAGTTTCTAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA

2206  CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA
```

FIG. 2. Nucleotide sequence of the *hap* gene and adjacent DNA regions. The deduced polypeptide sequence is shown. Putative Shine-Dalgarno (SD) sequences are overscored. The heavy underline designates the previously determined N-terminal amino acid sequence of the extracellular HA/protease (20). Possible cleavage sites of the signal peptide are underlined. The asterisk indicates the stop codon.

chromosomal DNA in Southern blots and was used to identify suitable restriction fragments for cloning (data not shown). Size-fractionated genomic libraries were created in the plasmid vector pGEM3Zf(+) and screened with the radiolabelled oligonucleotide probe. A 700-bp *Sal*I restriction fragment that hybridized strongly with the oligonucleotide was isolated. It was predicted that this fragment was too small to contain the entire gene for the HA/protease; therefore, the fragment was radiolabelled and used to clone a 3.2-kb *Hin*dIII restriction fragment from a size-fractionated genomic library created in the plasmid vector pACYC184.

Nucleotide sequence of the *hap* gene. A detailed restriction map of the 3.2-kb *Hin*dIII fragment and the subclones generated for sequence analysis of the *hap* gene and flanking regions are shown in Fig. 1. The nucleotide sequence of the *hap* gene region, shown in Fig. 2, contains an open reading frame of 1,827 bp starting at nucleotide 262 and ending at nucleotide 2,089. This open reading frame could encode a 609-amino-acid polypeptide with a predicted molecular weight of 69,300. The N-terminal region of the predicted polypeptide contains a putative signal sequence (Met-1 or Met-3 to Ala-16–Ala or Ala-24–Ala, Fig. 2) presumably necessary for the polypeptide to cross the cytoplasmic membrane. The N-terminal amino acids of the purified extracellular HA/protease, previously determined by protein sequencing (20), were located within the open reading frame starting at Ala-196 and showed absolute agreement with the deduced amino acid sequence (Fig. 2), suggesting subsequent processing at the N terminus. This deduced HA/protease (starting at Ala-196) consists of 414 amino acids and is calculated to be 46.7 kDa, which is substantially larger than the 32 kDa of the purified extracellular HA/protease (16). We have previously reported (20) that the *V. cholerae* HA/protease and the elastase from *P. aeruginosa* are immunologically and functionally related. A comparison between the deduced amino acid sequences of the two mature proteins revealed 61.5% identity and 70% similarity (Fig. 3). A comparison of the full-size polypeptides revealed 63% similarity (51.5% identity).

Construction of an HA/protease-negative strain (HAP-1) of *V. cholerae*. To create a mutant strain deficient in HA/protease, the clone containing the N-terminal segment of the *hap* gene on a 700-bp *Sal*I fragment (pCH1) was mutagenized and introduced into *V. cholerae*. The kanamycin resistance (Kan^r) gene from Tn*5* (kindly provided by M. Bauer), residing on a 2.4-kb *Xho*I restriction fragment, was isolated and cloned into the unique *Hpa*I restriction site of pCH1 after the *Xho*I ends were filled in with Klenow enzyme and deoxyribonucleotides to create a blunt-end fragment. This Amp^r Kan^r plasmid (pCH3) was purified from *E. coli* and then electroporated into the wild-type *V. cholerae* strain 3083 with selection for Amp^r and Kan^r. A single colony was grown and electroporated with pTB42, an apramycin resistance (Apr^r) plasmid of the same incompatibility group as pCH3 (ColE1). The cells were grown overnight to segregate the plasmids and then plated on kanamycin- and apramycin-containing medium. Any apramycin-resistant colony should have lost pCH3 but can still be Kan^r if the kanamycin resistance gene is integrated into the chromosome by homologous recombination between the *hap* sequences on the plasmid and the chromosome. Resulting Kan^r Apr^r colonies were tested for the loss of Amp^r. The mutant strain, named HAP-1, should be isogenic to 3083 except for the lack of HA/protease production and the presence of the kanamycin

```
196  AlaGlnAlaThrGlyThrGlyProGlyGlyAsnGlnLysThrGlyArgTyrGluTyrGly
198  AlaGluAla   Gly   GlyProGlyGlyAsnGlnLysIleGlyLysTyrThrTyrGly

216  SerAsn   GlyLeuProGlyPheThrIleAspLysThrGlyThrThrCysThrMetAsn
216  SerAspTyrGly   Pro   LeuIleValAsnAspArg         CysGluMetAsp

235  AsnSerAlaValLysThrValAsnLeuAsnGlyGlyThrSer   Gly   SerThrAla
231  AspGlyAsnValIleThrValAspMetAsnSerSerThrAspAspSerLysThrThrPro

253  PheSerTyrAlaCysAsnAsnSerThrAsnTyrAsnSerValLysThrValAsnGlyAla
251  PheArgPheAlaCysProThrAsnThr   Tyr         LysGlnValAsnGlyAla

273  TyrSerProLeuAsnAspAlaHisPhePheGlyLysValValPheAspMetTyrGlnGln
267  TyrSerProLeuAsnAspAlaHisPhePheGlyGlyValValPheLysLeuTyrArgAsp

293  TrpLeuAsnThrSerProLeuThrPheGlnLeuThrMetArgValHisTyrGlyAsnAsn
287  TrpPheGlyThrSerProLeuThrHisLysLeuTyrMetLysValHisTyrGlyArgSer

313  TyrGluAsnAlaPheTrpAspGlyArgAlaMetThrPheGlyAspGlyTyrThrArgPhe
307  ValGluAsnAlaTyrTrpAspGlyThrAlaMetLeuPheGlyAspGlyAlaThrMetPhe

333  TyrProLeuValAspIleAsnValSerAlaHisGluValSerHisGlyPheThrGluGln
327  TyrProLeuValSerLeuAspValAlaAlaHisGluValSerHisGlyPheThrGluGln
                                   *  Δ        *
353  AsnSerGlyLeuValTyrArgAspMetSerGlyGlyIleAsnGluAlaPheSerAspIle
347  AsnSerGlyLeuIleTyrArgGlyGlnSerGlyGlyMetAsnGluAlaPheSerAspMet
                    Δ                          *
373  AlaGlyGluAlaAlaGluTyrPheMetArgGlyAsnValAspTrpIleValGlyAlaAsp
367  AlaGlyGluAlaAlaGluPheTyrMetArgGlyLysAsnAspPheLeuIleGlyTyrAsp

393  IlePheLysSerSerGlyGlyLeuArgTyrPheAspGlnProSerArgAspGlyArgSer
387  IleLysLysGlySerGlyAlaLeuArgTyrMetAspGlnProSerArgAspGlyArgSer

413  IleAspHisAlaSerGlnTyrTyrSerGlyIleAspValHisHisSerSerGlyValPhe
407  IleAspAsnAlaSerGlnTyrTyrAsnGlyIleAspValHisHisSerSerGlyValTyr
                                             Δ
433  AsnArgAlaPheTyrLeuLeuAlaAsnLysSerGlyTrpAsnValArgLysGlyPheGlu
427  AsnArgAlaPheTyrLeuLeuAlaAsnSerProGlyTrpAspThrArgLysAlaPheGlu

453  ValPheAlaValAlaAsnGlnLeuTyrTrpThrProAsnSerThrPheAspGlnGlyGly
447  ValPheValAspAlaAsnArgTyrTyrTrpThrAlaThrSerAsnTyrAsnSerGlyAla

473  CysGlyValValLysAlaAlaGlnAspLeuAsnTyrAsnThrAlaAspValValAlaAla
467  CysGlyValIleArgSerAlaGlnAsnArgAsnTyrSerAlaAlaAspValThrArgAla

493  PheAsnThrValGlyValAsnAlaSerCysGlyThrThrPro      HA/protease
487  PheSerThrValGlyValThr      CysProSerAlaLeu      Elastase
```

FIG. 3. Comparison of the deduced amino acid sequences of the isolated forms of the *V. cholerae* HA/protease and *P. aeruginosa* elastase. |, identity; :, conservation of side group size or charge; *, putative zinc binding domain; Δ, active site.

resistance gene, which now resides within the N-terminal region of the HA/protease gene (Fig. 4). Restriction fragments of chromosomal DNA from HAP-1 compared with those of DNA from 3083 showed the expected increase in size (2.4-kb Kan^r gene) in Southern blots with the cloned HA/protease fragment as a probe (Fig. 4). Additionally, only the mutant DNA hybridized to a Tn*5* probe (data not shown). The mutant strain lacked the HA/protease band in Western blots with specific anti-HA/protease serum (Fig. 5A) and had reduced extracellular proteolytic activity compared with the parent strain 3083 as demonstrated on milk-containing medium (Fig. 5B). Nevertheless, the mutant strain still produces some extracellular proteolytic activity. That the residual proteolytic activity expressed by HAP-1 is distinct from the HA/protease was demonstrated by the failure of anti-HA/protease serum to inhibit the activity of the secondary protease(s) in milk agar, unlike the results obtained with strain 3083 (data not shown). The mutant strain showed no soluble hemagglutination ability with responder chicken erythrocytes.

Complementation of HAP-1 with the cloned *hap* gene. The cloned *hap* gene was tested for complementation of the recombinant HA/protease-negative *V. cholerae* strain (HAP-1). HAP-1 was electroporated with pCH2, which contains the 3.2-kb *Hin*dIII fragment in pACYC184. As a negative control the vector without an insert was also introduced into the mutant strain. The resulting chloramphenicol-resistant vibrio cells showed marked differences in their milk-cleaving abilities (Fig. 5B). HAP-1 cells transformed with pCH2 regained the ability to produce large clearing zones on milk agar, whereas cells transformed with pACYC184 did not. Additionally, the HA/protease-specific band found in Western blots of culture supernatants of 3083, which is not present in HAP-1, was restored when HAP-1 was transformed with pCH2 (Fig. 5A). No reaction with anti-HA/protease-specific serum was observed with HAP-1 carrying pACYC184. The increased size of the milk-clearing zone produced by HAP-1(pCH2) as compared with that produced by 3083 (Fig. 5) is probably due to the higher copy number of the *hap* gene present in these cells. This effect is not observed in the Western blot because even though the same amount of culture supernatant was loaded the bacteria grown under selective pressure do not grow as fast and so there were fewer cells in the culture of HAP-1(pCH2). Culture supernatants of the mutant strain harboring pCH2 showed hemagglutinating reactions with responder but not with nonresponder chicken erythrocytes, which was not observable with HAP-1 carrying pACYC184 (data not shown).

DISCUSSION

The structural gene for the HA/protease of *V. cholerae* was cloned on a 3.2-kb *Hin*dIII restriction fragment. The nucleotide sequence revealed an open reading frame of 1,827 nucleotides which could code for a 609-amino-acid peptide with a computed molecular weight of 69,300. This is larger than the previously reported 32 kDa for the purified extracellular *V. cholerae* HA/protease (16). The N-terminal amino acid sequence of the purified extracellular HA/protease has previously been reported (20) and was located within the open reading frame, as shown in Fig. 2. The deduced mature form of the protein would then consist of 414 amino acids with a molecular size of 46.7 kDa, which is still much larger than the 32-kDa purified product. However, when the HA/protease was purified in the presence of protease inhibitors, a larger (approximately 45-kDa) form of the HA/protease was isolated and found to have the same N-terminal amino acid sequence as the 32-kDa form (5a). These observations suggest that the secreted HA/protease may undergo several stages of processing, including cleavage of the signal peptide, processing of the proprotein at Ala-196 to form the mature N terminus, and a further proteolytic processing of its C-terminal region to result in the isolated 32-kDa form. Whether these steps all represent maturation stages is not yet determined, nor is the cellular location of such processing events. Since no significant amount of HA/protease could be detected in the periplasmic or cytoplasmic fractions of *V. cholerae* (results not shown), it is likely that the protein is rapidly transported after synthesis. Black et al. (5) pointed out that other bacterial zinc-containing metalloproteases (from several *Bacillus* species, *P. aeruginosa*, and *Legionella pneumophila*) have also been shown to be synthesized as larger precursors and processed to the mature forms.

The first methionine codon in the open reading frame at nucleotide 262 is followed by a second ATG codon at nucleotide 268 (Fig. 2). Since two possible Shine-Dalgarno

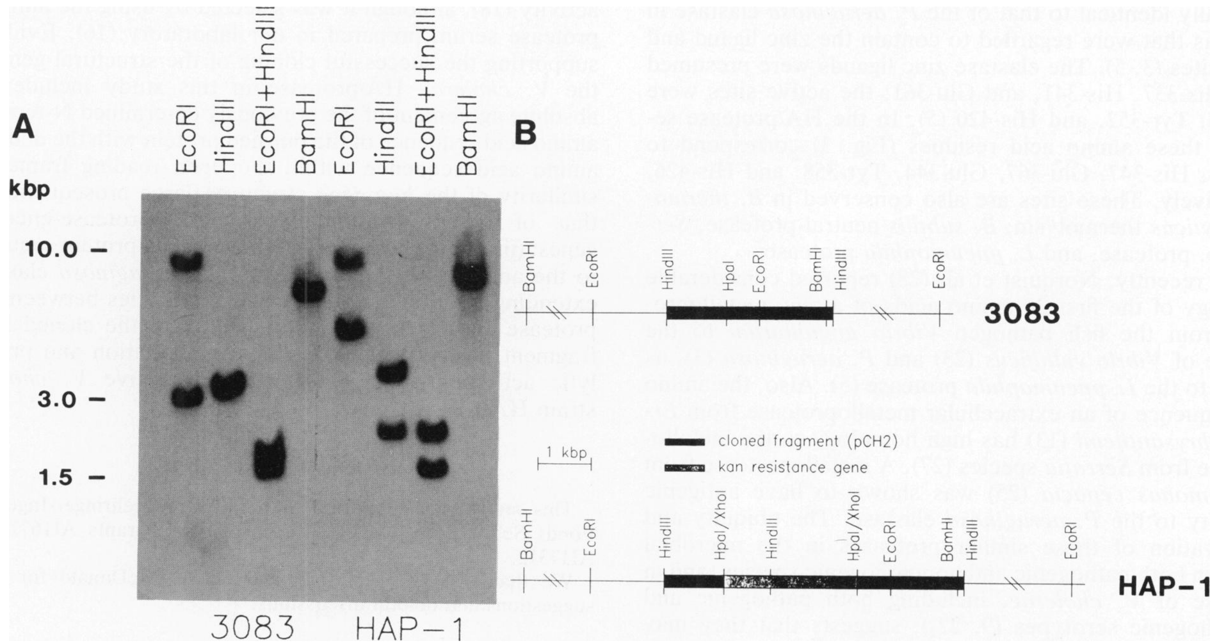

FIG. 4. Comparison of *hap* gene DNA regions of strain 3083 and the HA/protease-negative mutant HAP-1. (A) Southern blot of restriction enzyme-digested chromosomal DNA hybridized with the radiolabelled 3.2-kb *Hin*dIII DNA fragment containing the *hap* gene. (B) Restriction maps of the *hap* gene regions.

sequences (31) can be identified just upstream (Fig. 2), we predict that both start codons may be used. The length of the signal sequence is ambiguous because of the presence of two possible Ala-Ala cleavage sites (Fig. 2).

We have recently reported (20) that the *V. cholerae* HA/protease is structurally, functionally, and immunologically related to the elastase of *P. aeruginosa*. Elastase is also a zinc-dependent metalloprotease that has been shown to degrade or inactivate a variety of biologically important substances and is potentially involved in virulence of *P. aeruginosa* (12, 21, 26, 33, 34). The structural gene coding for elastase has been cloned (30) and sequenced (3). A comparison of the amino acid sequences of the HA/protease and elastase revealed high homology (70% similarity) between the two mature proteins, although the homology of the two prosequences was only 35%, and confirmed the previously observed relatedness of the two proteases. Recently other bacterial extracellular zinc-containing metalloproteases have been compared with the *P. aeruginosa* elastase. Bever and Iglewski (3) demonstrated amino acid sequence homology of the elastase to *Bacillus thermoproteolyticus* thermolysin, *Bacillus subtilis* neutral protease, and *Serratia* sp. protease. The sequence of the gene for an extracellular zinc metalloprotease from *L. pneumophila* revealed striking amino acid homology to the elastase (5). It is of special interest that the amino acid sequence of the HA/protease is

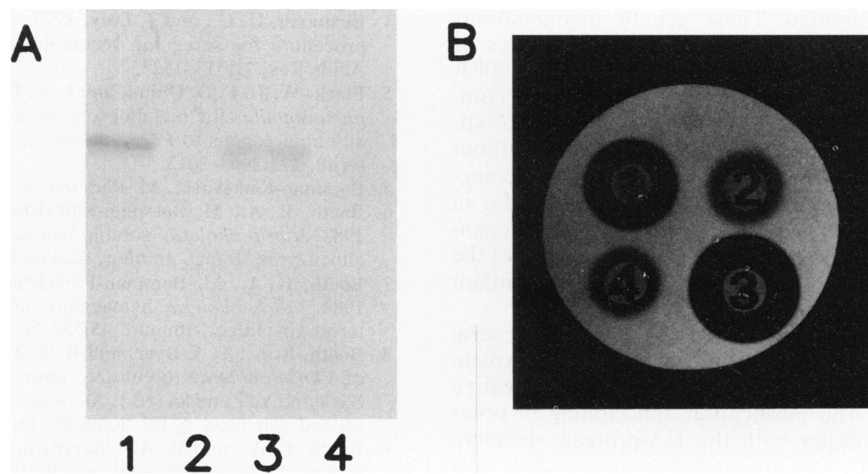

FIG. 5. Complementation of the HA/protease-negative mutant HAP-1 by pCH2. (A) Western blot of culture supernatants reacted with polyclonal anti-HA/protease serum. (B) Macrocolonies grown on milk-containing medium. 1, 3083; 2, HAP-1; 3, HAP-1(pCH2); 4, HAP-1(pACYC184).

practically identical to that of the *P. aeruginosa* elastase in the areas that were regarded to contain the zinc ligand and active sites (3, 5). The elastase zinc ligands were presumed to be His-337, His-341, and Glu-361; the active sites were Glu-338, Tyr-352, and His-420 (5). In the HA/protease sequence these amino acid residues (Fig. 3) correspond to His-343, His-347, Glu-367, Glu-344, Tyr-358, and His-426, respectively. These sites are also conserved in *B. thermoproteolyticus* thermolysin, *B. subtilis* neutral protease, *Serratia* sp. protease, and *L. pneumophila* protease.

Very recently, Norquist et al. (28) reported considerable homology of the first 20 amino acids of a zinc metalloprotease from the fish pathogen *Vibrio anguillarum* to the elastase of *Vibrio vulnificus* (23) and *P. aeruginosa* (3), as well as to the *L. pneumophila* protease (5). Also, the amino acid sequence of an extracellular metalloprotease from *Erwinia chrysanthemi* (13) has high homology to the metalloprotease from *Serratia* species (27). A metalloprotease from *Pseudomonas cepacia* (25) was shown to have antigenic similarity to the *P. aeruginosa* elastase. The ubiquity and conservation of these similar proteases in the microbial world, in both pathogenic and nonpathogenic species (and in the case of *V. cholerae*, including both pathogenic and nonpathogenic serotypes [9, 22]), suggests that they may provide some common survival advantage which is not necessarily essential for virulence but which, in the case of pathogenic species, may be associated with it.

E. coli carrying the *hap* gene on a plasmid did not show extracellular proteolytic activity on milk-containing plates and did not give specific bands in Western blots, indicating that the transcription or translation signals of the *hap* gene are not efficiently recognized by *E. coli*. Expression of the *hap* gene in *E. coli* could provide a selectable system for cloning *V. cholerae* regulatory, secretory, or processing elements.

In the present study the cloned *hap* gene was inactivated by insertion of the kanamycin resistance gene and introduced into the chromosome of *V. cholerae* by homologous recombination. This was accomplished by transforming plasmids into vibrio cells by electroporation, eliminating conjugational systems. After the original transformation with the mutant gene construct, the cells were "supertransformed" with an incompatible plasmid with a different resistance marker. By selecting for the appropriate phenotype, a desired HA/protease-negative derivative of *V. cholerae* 3083, named HAP-1, was isolated. These genetic manipulations had no discernible effect on the biosynthetic abilities of strain 3083 in that HAP-1 and HAP-1(pCH2) were, like 3083, capable of vigorous growth from small inocula in a completely chemically defined basal medium consisting of sucrose and inorganic salts (i.e., syncase medium without Casamino Acids [14a]). This recombinant HA/protease-negative mutant strain of *V. cholerae* will be very useful in examining the role, if any, of HA/protease in the pathogenesis of cholera. Virulence, adherence, and detachment of the mutant and parent strains will be compared in the infant rabbit model and in cultured cells (8).

Young and Broadbent (36) previously described several extracellular proteases in *V. cholerae* which could explain the residual proteolytic activity of the HA/protease-negative *V. cholerae* mutant. The purification and cloning of other proteases should be easier with the HA/protease-deficient strain.

The restriction map of the *hap* gene DNA region indicated no similarity to previously isolated DNA reportedly encoding a hemagglutinin of *V. cholerae* which had no protease activity (18), although it was selected by using the anti-HA/protease serum prepared in our laboratory (16). Evidence supporting the successful cloning of the structural gene for the *V. cholerae* HA/protease in this study includes: (i) absolute agreement of the previously determined N-terminal amino acid sequence of the purified protein with the deduced amino acid sequence within the open reading frame; (ii) similarity of the *hap* gene structure (large prosequence) to that of other zinc-containing metalloprotease-encoding genes; (iii) strong homology of the deduced protein sequence to the amino acid sequence of the *P. aeruginosa* elastase, extending the previously reported similarities between HA/protease and elastase; and (iv) ability of the cloned DNA fragment to complement the hemagglutination and proteolytic activities of the HA/protease-negative *V. cholerae* strain HAP-1.

ACKNOWLEDGMENTS

This study was supported in part by a Boehringer-Ingelheim Fonds Scholarship (to C.C.H.) and by NIH grants AI16776 and AI17312.

We thank M. A. McIntosh and G. A. McDonald for many suggestions and helpful discussions.

ADDENDUM

Since this paper was submitted, an HA/protease from *Vibrio mimicus* was shown to be immunologically cross-reactive with the *V. cholerae* HA/protease (11a). Additionally, the gene encoding a protein homologous to bacterial metalloproteases was cloned from *Listeria monocytogenes* (13a). The putative zinc-binding and active sites are also conserved in this protease.

REFERENCES

1. **Amersham Corp.** 1983. M13 cloning and sequencing handbook. Amersham Corp., Arlington Heights, Ill.

1a. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.

2. **Berman, M. L., L. W. Enquist, and T. J. Silhavy.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

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. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.

5. **Black, W. J., F. D. Quinn, and L. S. Tompkins.** 1990. *Legionella pneumophila* zinc metalloprotease is structurally and functionally homologous to *Pseudomonas aeruginosa* elastase. J. Bacteriol. **172:**2608–2613.

5a. **Boesman-Finkelstein, M.** Personal communication.

6. **Booth, B. A., M. Boesman-Finkelstein, and R. A. Finkelstein.** 1983. *Vibrio cholerae* soluble hemagglutinin/protease is a metalloenzyme. Infect. Immun. **42:**639–644.

7. **Booth, B. A., M. Boesman-Finkelstein, and R. A. Finkelstein.** 1984. *Vibrio cholerae* hemagglutinin/protease nicks cholera enterotoxin. Infect. Immun. **45:**558–560.

8. **Booth, B. A., T. J. Dyer, and R. A. Finkelstein.** 1990. Adhesion of *Vibrio cholerae* to cultured human cells, p. 19–35. *In* R. B. Sack and Y. Zinnaka (ed.), Advances in research on cholera and related diarrheas. KTK Scientific Publishers, Tokyo.

9. **Booth, B. A., and R. A. Finkelstein.** 1986. Presence of hemagglutinin/protease and other potential virulence factors in O1 and non-O1 *Vibrio cholerae*. J. Infect. Dis. **154:**183–186.

10. **Booth, B. A., C. V. Sciortino, and R. A. Finkelstein.** 1986. Adhesins of *Vibrio cholerae*, p. 169–182. *In* D. Mirelman (ed.),

Microbial lectins and agglutinins: properties and biological activity. John Wiley & Sons, Inc., New York.

11. **Booth, B. A., C. V. Sciortino, and R. A. Finkelstein.** 1987. On the trail of adherence mechanisms in *Vibrio cholerae*: many scents but not much sense, p. 220–241. *In* J. B. Robbins (ed.), Bacterial vaccines. Praeger Publishers, New York.

11a. **Chowdhury, M. A. R., S.-I. Miyoshi, and S. Shinoda.** 1990. Purification and characterization of a protease produced by *Vibrio mimicus*. Infect. Immun. **58**:4159–4162.

12. **Cryz, S. J.** 1984. *Pseudomonas aeruginosa* infections, p. 317–351. *In* R. Germanier (ed.), Bacterial vaccines. Academic Press, Inc., Orlando, Fla.

13. **Dahler, G. S., F. Barras, and N. T. Keen.** 1990. Cloning of genes encoding extracellular metalloproteases from *Erwinia chrysanthemi* EC16. J. Bacteriol. **172**:5803–5815.

13a. **Domann, E., M. Leimeister-Wächter, W. Goebel, and T. Chakraborty.** 1991. Molecular cloning, sequencing, and identification of a metalloprotease gene from *Listeria monocytogenes* that is species specific and physically linked to the listeriolysin gene. Infect. Immun. **59**:65–72.

14. **Finkelstein, R. A., M. Arita, J. D. Clements, and E. T. Nelson.** 1978. Isolation and purification of an adhesive factor ("cholera lectin") from *Vibrio cholerae*, p. 137–151. *In* Proceedings of the 13th Joint Conference on Cholera, U.S.-Japan Cooperative Medical Science Program. National Institutes of Health publication number 78-1590. National Institutes of Health, Bethesda, Md.

14a. **Finkelstein, R. A., P. Atthasampunna, M. Chulasamya, and P. Charunmethee.** 1966. Pathogenesis of experimental cholera: biologic activities of purified Procholeragen A. J. Immunol. **96**:440–449.

15. **Finkelstein, R. A., M. Boesman-Finkelstein, and P. Holt.** 1983. *Vibrio cholerae* hemagglutinin/lectin/protease hydrolyzes fibronectin and ovomucin: F. M. Burnet revisited. Proc. Natl. Acad. Sci. USA **80**:1092–1095.

16. **Finkelstein, R. A., and L. F. Hanne.** 1982. Purification and characterization of the soluble hemagglutinin (cholera lectin) produced by *Vibrio cholerae*. Infect. Immun. **36**:1199–1208.

17. **Finkelstein, R. A., P. Z. Sobocinski, P. Atthasampunna, and P. Charunmethee.** 1966. Pathogenesis of experimental cholera: identification of choleragen (procholeragen A) by disc immunoelectrophoresis and its differentiation from cholera mucinase. J. Immunol. **97**:25–33.

18. **Franzon, V. L., and P. A. Manning.** 1986. Molecular cloning and expression in *Escherichia coli* K-12 of the gene for a hemagglutinin from *Vibrio cholerae*. Infect. Immun. **52**:279–284.

19. **Hanne, L. F., and R. A. Finkelstein.** 1982. Characterization and distribution of the hemagglutinins produced by *Vibrio cholerae*. Infect. Immun. **36**:209–214.

20. **Häse, C. C., and R. A. Finkelstein.** 1990. Comparison of the *Vibrio cholerae* hemagglutinin/protease and the *Pseudomonas aeruginosa* elastase. Infect. Immun. **58**:4011–4015.

21. **Holder, I. A.** 1983. Experimental studies of the pathogenesis of infections due to *Pseudomonas aeruginosa*: effect of treatment with protease inhibitors. Rev. Infect. Dis. **5**(Suppl.):914–921.

22. **Honda, T., B. A. Booth, M. Boesman-Finkelstein, and R. A. Finkelstein.** 1987. Comparative study of *Vibrio cholerae* non-O1 protease and soluble hemagglutinin with those of *Vibrio cholerae* O1. Infect. Immun. **55**:451–454.

23. **Kothary, M. H., and A. S. Kreger.** 1987. Purification and characterization of an elastolytic protease of *Vibrio vulnificus*. J. Gen. Microbiol. **133**:1783–1791.

24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

25. **McKevitt, A. I., S. Bajaksouzian, J. D. Klinger, and D. E. Woods.** 1989. Purification and characterization of an extracellular protease from *Pseudomonas cepacia*. Infect. Immun. **57**:771–778.

26. **Morihara, K., and J. Y. Homma.** 1985. *Pseudomonas* proteases, p. 41–79. *In* I. A. Holder (ed.), Bacterial enzymes and virulence. CRC Press, Inc., Boca Raton, Fla.

27. **Nakahama, K., K. Yoshimura, R. Marumoto, M. Kikuchi, I. S. Lee, T. Hase, and H. Matsubara.** 1986. Cloning and sequencing of *Serratia* protease gene. Nucleic Acids Res. **14**:5843–5855.

28. **Norquist, A., B. Norrman, and H. Wolf-Watz.** 1990. Identification and characterization of a zinc metalloprotease associated with invasion by the fish pathogen *Vibrio anguillarum*. Infect. Immun. **58**:3731–3736.

29. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.

30. **Schad, P. A., R. A. Bever, T. I. Nicas, F. Leduc, L. F. Hanne, and B. H. Iglewski.** 1987. Cloning and characterization of elastase genes from *Pseudomonas aeruginosa*. J. Bacteriol. **169**:2691–2696.

31. **Shine, J., and L. Dalgarno.** 1975. Determinants of cistron specificity in bacterial ribosomes. Nature (London) **254**:34–38.

32. **Stoebner, J. A., and S. M. Payne.** 1988. Iron-regulated hemolysin production and utilization of heme and hemoglobin by *Vibrio cholerae*. Infect. Immun. **56**:2892–2895.

33. **Woods, D. E., and B. H. Iglewski.** 1983. Toxins of *Pseudomonas aeruginosa*: new perspectives. Rev. Infect. Dis. **5**(Suppl.):715–722.

34. **Wretlind, B., and O. R. Pavlovskis.** 1983. *Pseudomonas aeruginosa* elastase and its role in *Pseudomonas* infections. Rev. Infect. Dis. **5**(Suppl.):998–1004.

35. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.

36. **Young, D. B., and D. A. Broadbent.** 1982. Biochemical characterization of extracellular proteases from *Vibrio cholerae*. Infect. Immun. **37**:875–883.