JOURNAL OF BACTERIOLOGY, Feb. 1992, p. 736–742
0021-9193/92/030736-07$02.00/0

Vol. 174, No. 3

Molecular Cloning and Sequencing of the Gene for a Halophilic Alkaline Serine Protease (Halolysin) from an Unidentified Halophilic Archaea Strain (172P1) and Expression of the Gene in *Haloferax volcanii*

MASAHIRO KAMEKURA,[1]* YUKIO SENO,[1] MELISSA L. HOLMES,[2] AND MICHAEL L. DYALL-SMITH[2]

Noda Institute for Scientific Research, 399 Noda, Noda-shi, Chiba-ken, 278 Japan,[1] and Department of Microbiology, University of Melbourne, Parkville, Victoria 3052, Australia[2]

Received 5 July 1991/Accepted 21 November 1991

The gene of a halophilic alkaline serine protease, halolysin, from an unidentified halophilic archaea (archaebacterium) was cloned and its nucleotide sequence was determined. The deduced amino acid sequence showed that halolysin consists of 411 amino acids, with a molecular weight of 41,963. The highest homology was found with thermitase from *Thermoactinomyces vulgaris*. Halolysin has a long C-terminal extension of approximately 120 amino acids which has not been found in other extracellular subtilisin type serine proteases. The gene, *hly*, was expressed in another halophilic archaea, *Haloferax volcanii*, in a medium containing 18% salts by using a plasmid shuttle vector which has a novobiocin resistance determinant as a selectable marker.

Halobacteria, extremely halophilic red-pigmented bacteria, were first studied extensively because they attacked salted fish during long sea voyages and caused pinkeye (5). Their extraordinary ability to grow in hypertonic solution (above 300 g of NaCl per liter) and their potential ability to hydrolyze proteins (7) are the main reasons for the reddening of salted fish. Though a few articles have dealt with the purification and characterization of proteases from *Halobacterium salinarium* (14, 18) and *H. halobium* (12), detailed mechanisms to explain the halophilicity of these enzymes have not been elucidated because of their extreme instability even in the presence of 3 to 4 M NaCl (12). In a previous article, Kamekura and Seno reported the purification and characterization of a halophilic alkaline serine protease (F-II) of a halophilic archaebacterium (strain 172P1) and also determined the sequence of the first 35 N-terminal amino acids (13). F-II is thus the first serine protease from archaebacteria (alternatively called archaea, according to a new proposal by Woese et al. [29]) whose amino acid sequence has been partially determined. Many articles have been published on the amino acid sequences of alkaline serine proteases from eubacteria and eukaryotes (bacteria and eucarya, respectively, according to Woese et al. [29]). Thus, from the phylogenetic point of view, it is very interesting to compare the complete amino acid sequence of F-II with those of other serine proteases.

In addition to its evolutionary significance, the protease F-II is extraordinary in that it is extremely halophilic, with an optimal NaCl concentration of 4 to 4.5 M for activity, and also thermophilic, having maximal activity at 70°C. Fortunately, it is quite stable when kept at 5°C in the presence of 4 M NaCl; thus, F-II is a good candidate for further investigation of the mechanism of halophilicity of an enzyme at the molecular level. Sequencing of the gene coding for F-II is a prerequisite for this purpose.

In this article, we propose to designate the protease F-II halolysin. The molecular cloning and sequencing of the gene for the halolysin (*hly*) are reported, as well as its expression in another halophilic archaea, *Haloferax volcanii*, with a plasmid vector developed by Holmes and Dyall-Smith (9). Selection for plasmid maintenance in *Haloferax* cells was done by including a novobiocin resistance gene in the vector (10).

MATERIALS AND METHODS

Microbial strains. Detailed taxonomical study of the halophilic archaea strain 172P1 used in this study is now in progress. The strain was deposited with the Fermentation Research Institute, MITI (Tsukuba, Japan), as FERM-P 10747. *Haloferax volcanii* WFD11, a derivative of *H. volcanii* NCMB 2012 (DS2) cured of the endogenous plasmid pHV2, was obtained from W. F. Doolittle (Dalhousie University, Halifax, Canada) and used as a host for the plasmid vector containing the halolysin gene. Competent cells of *Escherichia coli* JM109 and HB101 were obtained from Takara Shuzo Co., Ltd, Japan. *E. coli* JM110 (31) was used to produce unmethylated plasmid DNA.

Isolation of DNA. Strain 172P1 was cultivated in a 25% (wt/vol) artificial seawater medium containing 0.5% yeast extract as described by Rodriguez-Valera et al. (20). Chromosomal DNA was prepared by lysing the cells in 10 mM Tris-hydrochloride (pH 7.6)–50 mM EDTA–2% *N*-lauroylsarcosine, followed by digestion with RNase A and proteinase K before phenol extraction and isopropanol precipitation.

Preparation of a probe by PCR. The protease halolysin was purified, and a powdered phenylmethane sulfonyl-halolysin (PMS-halolysin) was obtained as described in a previous article (13). The PMS-halolysin was digested completely with *N*-tosyl-L-phenylalanine chloromethyl ketone (TPCK)-trypsin, and resultant oligopeptides were separated by high-pressure liquid chromatography (HPLC) with a column of VYDAC 218TP54. The amino acid sequences of two well-separated oligopeptides were determined by a gas-phase

* Corresponding author.

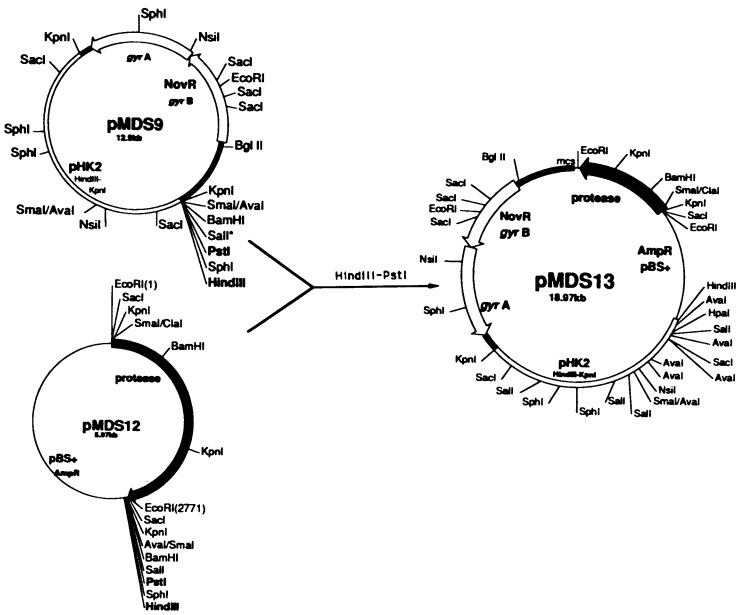

FIG. 1. Construction of the *hly*-containing plasmid vector pMDS13.

amino acid sequencer (model 477A; Applied Biosystems Inc.).

For the isolation of the gene for halolysin, a probe was prepared by polymerase chain reaction (PCR). The sense and antisense primers were synthesized by the phosphoramidite method (16). The sense primer had the base sequence 5'-GCC-ACC-CCG-AAC-GAC-CCC-CAG-TAC-GGC-CAG-CAG-TAC-GCG-CCC-CAG-CAG-GTC-AAC-3' (amino acid sequence NH_2-Ala-Thr-Pro-Asn-Asp-Pro-Gln-Tyr-Gly-Gln-Gln-Tyr-Ala-Pro-Gln-Gln-Val-Asn). The antisense primer had the base sequence 3'-CGG-CCG-TTG-CCC-ATG-CCG-TTG-5' (amino acid sequence Ala-Gly-Asn-Gly-Tyr-Gly-Asn). PCR was performed in 50 μl of 10 mM Tris-hydrochloride (pH 8.3)–50 mM KCl–1.5 mM $MgCl_2$–0.01% (wt/vol) gelatin–200 μM each deoxynucleotide triphosphate; 20 μM primer, 0.1 μg of chromosomal DNA, and 2 U of AmpliTaq polymerase were added. After 24 cycles of amplification (1 min at 94°C, 2 min at 37°C, and 3 min at 72°C), a product of the expected size (510 bp) was detected by 1.5% agarose gel electrophoresis. The fragment was eluted from the gel and purified by phenol-chloroform extraction. To ensure blunt ends, the fragment was treated with T4 DNA polymerase, then ligated to *Sma*I-cut plasmid pUC119, and transformed into *E. coli* JM109. White colonies on TY agar plates containing 50 μg of ampicillin per ml, isopropyl-β-D-thiogalactopyranoside (IPTG), and 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) were picked, recombinant plasmid DNAs were extracted and digested with *Eco*RI and *Pst*I, and the released fragments of 580 bp were isolated from the agarose gel by electroelution. The fragment was labeled with digoxigenin-dUTP by using a Nonradioactive DNA labeling and detection kit from Boehringer Mannheim and used as a probe in Southern blot and colony hybridization experiments (1).

Cloning and sequencing of halolysin gene. The chromosomal DNA from strain 172P1 was digested with *Eco*RI and *Pst*I. After electrophoresis in a 0.7% agarose gel, DNA fragments of approximately 6 kb were eluted from the gel. Plasmid pBR322 was also digested with *Eco*RI and *Pst*I, and a fragment of 3.6 kb was purified by electrophoresis. The two DNA species were ligated and used to transform *E. coli* HB101, MC1061, or JM109 to tetracycline resistance and ampicillin sensitivity. Colonies were transferred to nitrocellulose filters and hybridized with the probe prepared above. Isolation and purification of the transforming plasmid and subsequent DNA extraction were carried out by standard methods (1). The nucleotide sequence of the halolysin gene and its flanking regions was determined by the dideoxy sequencing method (21), using subclones constructed from plasmids pUFII15 and pUFII3 (see Results) by the kilosequencing method (8, 23, 31). Single-stranded DNA was prepared as described by Vierira and Messing (26), and sequencing reactions were performed with dye-labeled M13 universal primer (−21; Applied Biosystems Inc.) and run in a DNA sequencer (ABI 373A).

Construction of a plasmid vector and transformation of *H. volcanni*. The halolysin gene was inserted into the halobacterial plasmid vector pMDS9 (10) by the strategy outlined in Fig. 1. The resulting construct, pMDS13, was then grown in *E. coli* JM110 to avoid a restriction barrier in *H. volcanii* (10).

H. volcanii WFD11 was transformed with plasmid pMDS13 by the method described previously (9, 10). Transformants were plated on 1.5% agar plates of transformation medium containing 18% seawater, 15% sucrose, 0.8% Bacto skim milk (13), and 0.3 μg of novobiocin per ml and incubated at 37°C. Transformants expressing the halolysin gene *hly* were detected as colonies surrounded by an area of clearing in the milk agar (i.e., a clear halo).

Assay of protease activity of halolysin. Cells of the *H. volcanii* transformant cultivated as described in Results were harvested by centrifugation, and the protease activity of the supernatant was assayed with azocasein as a substrate as described previously (13) except that the reaction mixture was incubated for 1 h. The concentration of protease was calculated from a calibration curve constructed by using a purified halolysin preparation from strain 172P1.

Purification of halolysin (F-II) and F-III from a transform-

ant. The supernatant was concentrated with an ultrafiltration module (SEP-1013) with a molecular weight cutoff of 3,000 (Asahi Kasei Industry, Ltd.). The concentrate was centrifuged to get rid of insoluble material and applied to a column (3 by 9.5 cm) of phenyl-Sepharose CL-4B (Pharmacia, Uppsala, Sweden) equilibrated with 25% NaCl–10 mM Tris-hydrochloride (pH 7.6), washed with the same buffer, and eluted with 18% NaCl–10 mM Tris-hydrochloride (pH 7.6). The eluate was concentrated with a Diaflo membrane (PM-10). One hundred-microliter aliquots of the concentrate were loaded onto an HPLC column of phenyl-Superose HR5/5 (0.5 by 5 cm; Pharmacia) which had been equilibrated with 10 mM Tris-acetate, 10 mM calcium acetate, and 1.83 M sodium acetate, pH 7.4. F-II and F-III were eluted as separate peaks by a continuous flow of the buffer at a rate of 0.5 ml/min. Gel filtration HPLC was performed with a column (TSK G3000 SWXL; 0.75 by 30 cm; Tosoh Corp.). The solvent was 10 mM Tris-acetate–10 mM calcium acetate–500 mM sodium acetate (pH 7.4), and the flow rate was set at 0.5 ml/min.

For the differentiation of halolysin (F-II) from F-III (see Results), a 10-μl portion of protease solution was separated by nondenaturing polyacrylamide gel electrophoresis (PAGE) at 5°C in a 16% acrylamide slab gel, in parallel with 172P1 halolysin. After completing the run, the gel was flooded with a reaction mixture containing 25% NaCl, 10 mM Tris-hydrochloride (pH 7.6), and 50 μg of a synthetic substrate, succinyl-Ala-Ala-Pro-Phe-4-methyl-coumaryl-7-amide, per ml and then exposed to light (365-nm wavelength). Protease activity was detected as a fluorescent band of amino methyl coumarin (AMC), and a thin needle was inserted to mark the center of the band, after which it was stained with Coomassie brilliant blue.

The two purified active components F-II and F-III were applied to protein sequencer ABI 473A to determine their N-terminal amino acid sequences.

Materials. Restriction endonucleases were from Takara Shuzo Co., Ltd. (Kyoto, Japan) or Boehringer Mannheim Biochemicals, Indianapolis, Ind. Proteinase K, RNase A, novobiocin sodium salt, and azocasein were obtained from Sigma Chemical Co. A protease substrate, succinyl-Ala-Ala-Pro-Phe-MCA, was purchased from Protein Institute, Inc., Osaka, Japan. Lambda DNA cleaved with *Sty*I and φX174 DNA digested with *Hin*cII were used as DNA size markers and were purchased from Nippon Gene Co. Ltd.

Nucleotide sequence accession number. The nucleotide sequence for *hly* reported in this article has been assigned DDBJ, EMBL, and GenBank accession number D90432.

RESULTS

Cloning of the halolysin gene. The amino acid sequences of two tryptic oligopeptides of 172P1 halolysin were determined as described in Materials and Methods. A sequence (AAAGN) in one oligopeptide was noticeable, because this sequence has been found in all subtilisins (from residues 151 through 155 of BPN'[28]) and thermitase (17). To isolate the gene coding for the halolysin, which we designated *hly*, a probe was prepared by PCR with a sense primer deduced from the N-terminal 18-amino-acid sequence determined previously (13) and an antisense primer deduced from the sequence of amino acids corresponding to residues 153 to 159 of subtilisin BPN' (28). The base sequences of the two primers were determined from the codon usage frequencies of the genes of bacteriorhodopsin (3) and ribosomal A protein (11) of *H. halobium*. Southern hybridization of

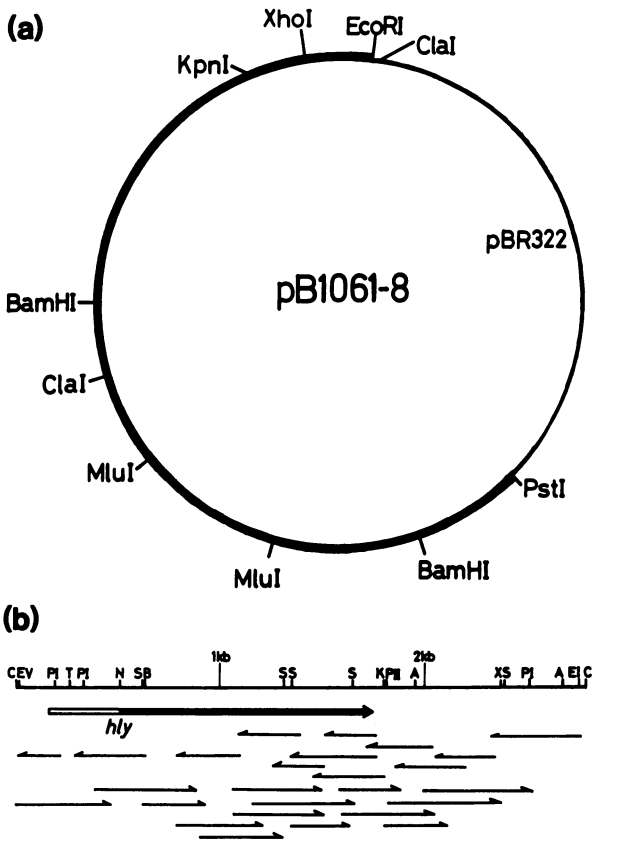

FIG. 2. (a) Restriction map of pB1061-8. (b) Strategy for determination of the sequence of *hly* by the kilosequencing method. Restriction sites: C, *Cla*I; EV, *Eco*RV; PI, *Pvu*I; T, *Tth*III; N, *Nru*I; S, *Sal*I; B, *Bam*HI; K, *Kpn*I; PII, *Pvu*II; A, *Ava*I; X, *Xho*I; EI, *Eco*RI.

chromosomal DNA from 172P1 digested with various restriction enzymes, singly or in combination, with the labeled probe suggested that *hly* resides in an approximately 6-kb DNA fragment produced by digestion with *Eco*RI and *Pst*I. Digestion with *Bam*HI gave two bands of 3.6 and 4 kb on Southern hybridization, suggesting the existence of a *Bam*HI site in *hly*. The 6-kb DNA was ligated to pBR322 which had been digested with *Eco*RI and *Pst*I and transformed into competent cells of *E. coli*, and colony hybridization was performed as described in Materials and Methods.

A colony giving significant staining was identified from among the transformants. Plasmid (pB1061-8) DNA was purified from this transformant, cut with *Eco*RI and *Pst*I, and electrophoresed in a 0.7% agarose gel, and the two DNA fragments (observed after ethidium bromide staining) were transferred to a nitrocellulose filter. The plasmid contained a 5.8-kb insert which hybridized with the probe, and a detailed restriction map is shown in Fig. 2a. To reduce the size of the insert, pB1061-8 was cleaved with *Cla*I, and the isolated 2.7-kb fragment was made blunt-ended and inserted into plasmid pUC119 at the *Sma*I site. Two plasmids with the insert in opposite directions, pUFII15 and pUFII3, were selected.

DNA sequencing. Plasmids pUFII15 and pUFII3 were sequenced to localize the structural gene for the halolysin and its 5'- and 3'-flanking regions. The strategy for DNA sequencing by the kilosequencing method is depicted in Fig.

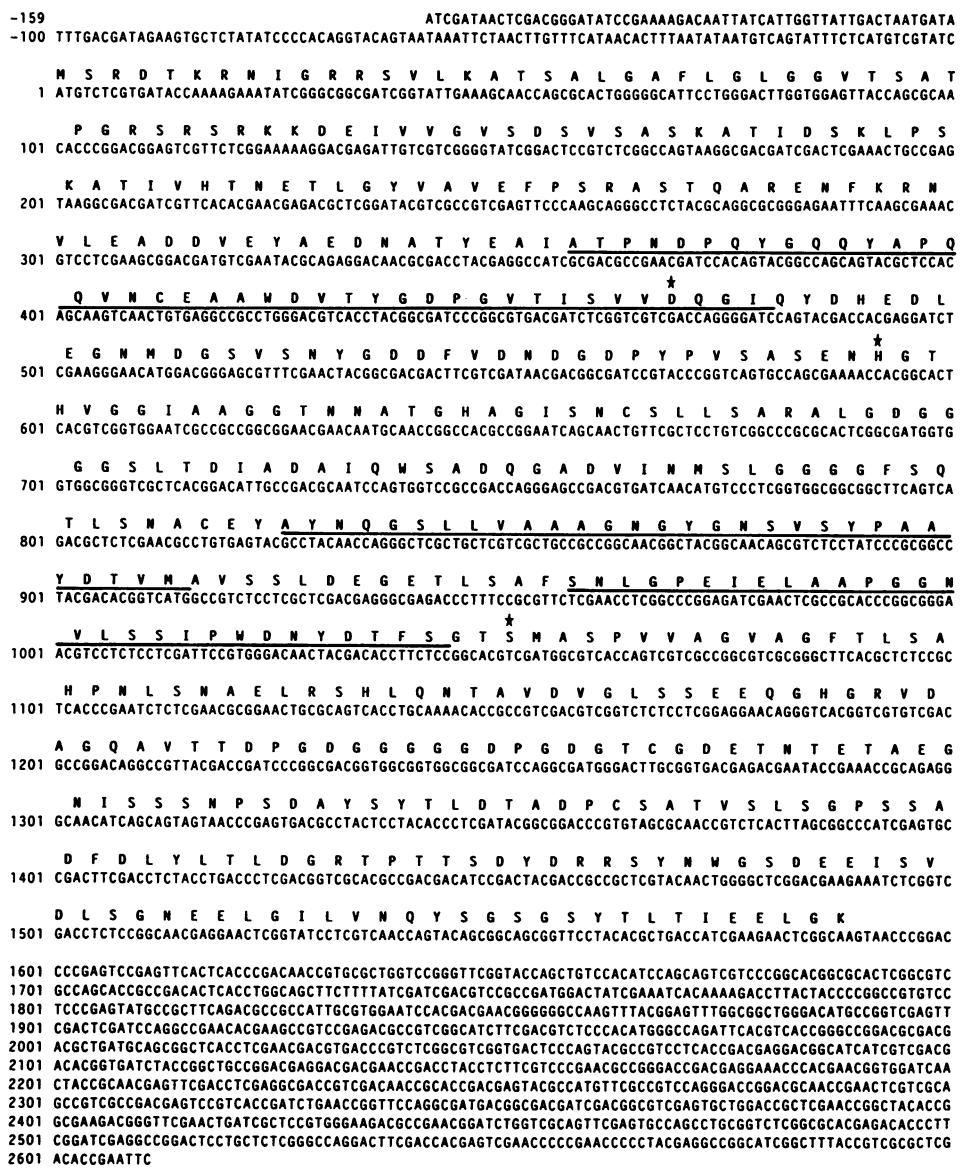

FIG. 3. Nucleotide sequence of *hly* and the deduced amino acid sequence of halolysin. Although the full length of the *Cla*I-*Cla*I fragment was sequenced, the 3′-terminal region of the fragment, *Eco*EI-*Cla*I of pBR322, has been omitted from the figure. Underlined amino acid sequences were determined with an ABI 477A protein sequencer. Stars indicate essential amino acids constituting the active site.

2b. The determined nucleotide sequence and the deduced amino acid sequence are shown in Fig. 3. The sequence included portions (underlined) corresponding to the N-terminal region determined previously (13) and in this study and two oligopeptides sequenced in this study.

A computer search for restriction sites showed that there was a *Cla*I site at position 1737. This site is preceded by TTCTTTT, which might interrupt cleavage by the enzyme.

Structural analysis of *hly*. According to the nucleotide sequence, there is a long open reading frame encoding 119 amino acids preceding the coding region for the halolysin, which is 411 amino acids in length, with a molecular weight of 41,963. These data suggest that halolysin is translated as a precursor polypeptide consisting of 530 amino acids and that the N-terminal 119 amino acid residues span the putative prepro region. Within the first 40 amino acids of the precursor, there is a hydrophobic region following Lys-16 which could function as a signal sequence for secretion from the cell.

Without transcriptional mapping of the 5′ end of the mRNA, it is not possible to identify the promoter for the *hly* gene, although within the sequence upstream of the start of translation (shown in Fig. 3) there are a number of A/T-rich motifs that may fulfill this function (25).

Expression of *hly* gene in *H. volcanii*. Transformation of *H. volcanii* WFD11 with plasmid pMDS13 (Fig. 1) gave two transformant colonies which produced clear halos on the skim milk agar plates. A transformant surrounded by a larger halo was transferred to a fresh 18% seawater–1.5% agar plate containing 0.8% skim milk and 0.3 μg of novobiocin per ml. The transformant was stable on the plate and kept producing clear haloes upon repeated subcultures after incubation for 18 to 24 h at 37°C.

Characterization of the protease produced by the trans-

formant. The transformant *H. volcanii*/pMDS13 was precultured in 18% seawater medium containing novobiocin for 2 days at 37°C, and a 20-ml portion was inoculated into 2 liters of the same medium in a 3-liter jar fermentor. Cultivation was performed at 37°C with aeration (2 liters per 2 liters of medium) and agitation (300 rpm) for 58 h, and the cells were removed by centrifugation. For comparison, strain 172P1 was grown in 2 liters of 18% seawater medium at 40°C for 58 h with aeration (2 liters per 2 liters of medium) and agitation (110 rpm), these being the optimal culture conditions. The amount of protease produced in the two supernatants was equivalent to 2.3 and 65.2 μg of halolysin per ml for 172P1 and *H. volcanii*/pMDS13, respectively.

The supernatant of *H. volcanii*/pMDS13 (5.23 liters total from three cultures) was concentrated to 68 ml with an ultrafiltration module, and 10-μl portions of the resultant 77-fold concentrates were separated by native PAGE. Protease activities were detected in the gel with the substrate succinyl-Ala-Ala-Pro-Phe-MCA, and total proteins were then stained with Coomassie brilliant blue. The concentrate from the transformant culture showed a major protein band with protease activity, which had the same relative mobility (0.60) as F-III, and a weaker band of halolysin (F-II), with a relative mobility of 0.84.

The two active components in 50 ml of the 77-fold concentrate were purified as described in Materials and Methods. The protease activities were separated into two peaks by hydrophobic chromatography with a phenyl-Superose HR5/5 column, an early fraction and a late fraction. The former showed the same behavior as halolysin on hydrophobic chromatography and nondenaturing PAGE. The latter comigrated with F-III from 172P1 on PAGE. Amino acid sequencing, however, showed that the two components were identical in sequences of at least 30 N-terminal amino acids, which were exactly the same as those of 172P1 halolysin. Gel permeation HPLC with a G3000 SWXL column clearly demonstrated that the molecular weights of the two components were the same as that of 172P1 halolysin. Furthermore, the specific activities of the two components were the same, within experimental error.

N-terminal amino acid sequence of 172P1 F-III. F-III was purified from a culture supernatant of strain 172P1. Nondenaturing PAGE of a concentrated supernatant showed that F-III is the single active component, and a halolysin band was not detected. The 172P1 F-III was purified by the same procedure used for the purification of F-II and F-III from the transformant.

The amino acid sequence of the N-terminus was shown to be exactly the same as that of halolysin (F-II).

DISCUSSION

Few proteases have been purified from archaea. The thermophilic archaea *Sulfolobus acidocaldarius* produces a thermostable acid protease, thermopsin, and the encoding gene was cloned and sequenced (6, 15). Cowan et al. (2) purified an extremely thermostable extracellular protease, archaelysin, from a strain of the sulfur-metabolizing thermophilic archaea *Desulfurococcus mucosus*. Archaelysin seems to be a serine protease because it was inhibited by chymostatin and phenylmethylsulfonyl fluoride, as was halolysin. More recently, another hyperthermophilic archaea, *Pyrococcus furiosus*, was shown to have a hyperthermoactive serine protease, pyrolysin (4). These two archaeal serine proteases are outstanding in their thermophilic characteristics, but their genes have not been cloned.

As yet, no serine protease from methanogens has been reported. The methanogens are another major branch of archaea, and this gap in knowledge needs urgent attention.

Concerning the halophilic proteases from halophilic archaea, a serine protease has been purified from *Halobacterium halobium* (12), and more recently, an extracellular serine protease, ESP4 of *Halobacterium* sp. strain TuA4, was also purified (22). Detailed characterization of these two proteases has not been reported yet. Halolysin is so far the first protease from halophilic archaea and the first serine protease from archaea whose gene has been sequenced.

Expression of the *hly* gene was first tried in an osmophilic yeast, *Zygosaccharomyces rouxii* (19), in media containing 18% NaCl; however, plasmid vectors containing *hly* were quite unstable, and thus no expressions was observed.

In a previous article (13), the authors showed that at least three active protease moieties, F-I, F-II, and F-III, were obtained from the culture filtrate of strain 172P1. F-III was the major component throughout the exponential and stationary phases of growth, as demonstrated by nondenaturing PAGE of concentrates of culture supernatants. During the following purification procedures, the amount of F-II finally exceeded that of F-III. Furthermore, a partially purified F-III preparation was found to convert to F-II during storage for a few weeks at 5°C in the presence of 25% NaCl. Thus, the authors of the previous article (13) speculated that F-III might be a "precursor" of F-II halolysin, though there was no direct evidence.

In this study, we showed that *H. volcanii*/pMDS13 produced F-III as the major component, thereby demonstrating that the genes for F-III and F-II reside in the 2.7-kb *Cla*I-*Cla*I fragment. N-terminal amino acid sequencing of F-II and F-III and gel permeation HPLC clearly demonstrated that the two components have the same N-terminal amino acid sequences and almost the same molecular weights. However, they behaved quite differently in nondenaturing PAGE and hydrophobic HPLC with a phenyl-Superose column. From these data, we speculate that F-III is converted to F-II by posttranslational modification; e.g., deamination of glutamine and asparagine to glutamic acid and aspartic acid, respectively, thereby making F-II a more acidic protein than F-III. If this speculation is correct, it follows that the amino acid sequence deduced from the base sequence as shown in Fig. 3 is that of F-III. The determined amino acid sequences (underlined in Fig. 3) of three regions of F-II (172P1 halolysin) coincided completely with the sequences deduced from the base sequences; no conversion of N to D or Q to E was detected in these regions. Thus, deamination might occur in other parts of the molecule, possibly at residues close to the C terminus. Other possible modifications are phosphorylation, sulfation, and attachment or detachment of carbohydrate chains. Proof of these speculations requires amino acid sequencing of whole molecules of F-II and F-III.

A characteristic of halolysin is its high molecular weight compared with other serine proteases. In the previous article (13), the molecular weight of halolysin was estimated at 42,000 to 44,000, and in this study it was calculated as 41,963 from the known sequence of 411 amino acids. Various subtilisins from *Bacillus* spp. and subtilisin-type serine proteases have molecular weights of approximately 28,000, with about 280 amino acid residues (17). In contrast, an extracellular protease produced by *Serratia marcescens* is an exceptionally large serine protease, comprising 381 amino acids, with a total molecular weight of 40,905 (30).

Alignment of five serine proteases from archaea, bacteria,

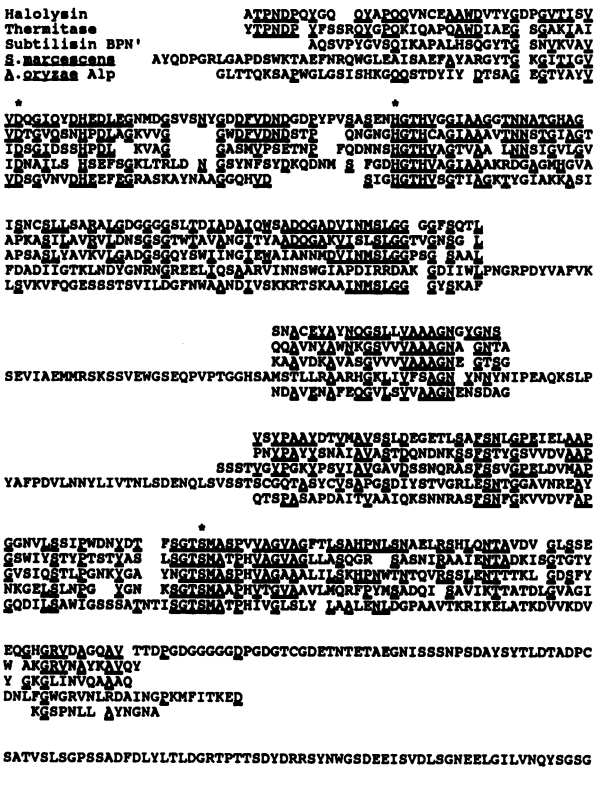

FIG. 4. Alignment of amino acid sequence of halolysin with those of other subtilisin-type serine proteases. References: *Thermoactinomyces vulgaris* thermitase (17), *Bacillus amyloliquefaciens* subtilisin BPN' (28), *Serratia marcescens* serine protease (30), and *Aspergillus oryzae* alkaline protease Alp (24). Stars indicate essential amino acids constituting the active site. Underlined residues are identical to the corresponding halolysin residues.

and eucarya (Fig. 4) revealed that halolysin has a long C-terminal extension or "tail" of approximately 120 amino acids which is not present in other extracellular subtilisin-type serine proteases. On the other hand, *S. marcescens* protease is characteristic in having two long extra sequences which are not found in the others. Despite these tails and the gaps observed in the alignments, very high homologies are observed in some regions, particularly around the essential amino acids constituting the active site, Asp-32, His-64, and Ser-221 (according to subtilisin BPN' numbering). Though the long gap and tail make it difficult to calculate the similarity of halolysin to other serine proteases, simple counting of the number of matched amino acids (in the alignment of Fig. 4) would give an idea of its similarity to other proteases; 130 for thermitase, 103 for subtilisin BPN', 70 for *S. marcescens* protease, and 79 for the alkaline serine protease of *A. oryzae*. Altogether, of the 298 amino acids of halolysin, at least 183 occur in one of the four other proteases.

Secretory proteins are generally synthesized with a signal sequence at the N terminus for translocation through a membrane. The signal sequence usually consists of 20 to 40 amino acids and contains a highly hydrophobic domain in the middle and positively charged residues close to the N terminus. Watson (27) stated that the general format of a signal seems to include a charged residue within the first five amino acids, followed by a core of at least nine hydrophobic residues, which should be sufficient to span a membrane, and that a helix-breaking residue (glycine or proline) or a large polar residue frequently occurs four to eight residues before the cleavage site. The prepro-region of the halolysin contains a putative signal sequence which starts at Lys-16, which is followed by 11 hydrophobic amino acids and Gly-Gly-Val, but it is not yet clear how cleavage to the F-III form occurs.

Now that the amino acid sequence of halolysin is known and expression of the cloned gene is possible, we aim to engineer mutations in *hly* to identify regions of the protein that are responsible for transport, processing, and halophilicity.

ACKNOWLEDGMENTS

We thank A. Matsuyama for his help in the preparation of synthetic oligodeoxynucleotides. We thank Y. Ogawa for the kind gift of plasmids used for the construction of a yeast plasmid vector.

REFERENCES

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1989. Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York.
2. **Cowan, D. A., K. A. Smolenski, R. M. Daniel, and H. W. Morgan.** 1987. An extremely thermostable extracellular proteinase from a strain of the archaebacterium *Desulfurococcus* growing at 88°C. Biochem. J. **247**:121–133.
3. **Dunn, R., J. McCoy, M. Simsek, A. Majumdar, S. H. Chang, U. L. RajBhandary, and H. G. Khorana.** 1981. The bacteriorhodopsin gene. Proc. Natl. Acad. Sci. USA **78**:6744–6748.
4. **Eggen, R., A. Geerling, J. Watts, and W. M. de Vos.** 1990. Characterization of pyrolysin, a hyperthermoactive serine protease from the archaebacterium *Pyrococcus furiosus*. FEMS Microbiol. Lett. **71**:17–20.
5. **Flannery, W. L.** 1956. Current status of knowledge of halophilic bacteria. Bacteriol. Rev. **20**:49–66.
6. **Fusek, M., X.-L. Lin, and J. Tang.** 1990. Enzymic properties of thermopsin. J. Biol. Chem. **265**:1496–1501.
7. **Gibbons, N. E.** 1957. The effect of salt concentration on the biochemical reactions of some halophilic bacteria. Can. J. Microbiol. **3**:249–255.
8. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351–359.
9. **Holmes, M. L., and M. L. Dyall-Smith.** 1990. A plasmid vector with a selectable marker for halophilic archaebacteria. J. Bacteriol. **172**:756–761.
10. **Holmes, M. L., S. D. Nuttall, and M. L. Dyall-Smith.** 1991. Construction and use of halobacterial shuttle vectors and further studies on *Haloferax* DNA gyrase. J. Bacteriol. **173**:3807–3813.
11. **Itoh, T., T. Kumazaki, M. Sugiyama, and E. Otaka.** 1988. Molecular cloning and sequence analysis of the ribosomal 'A' protein gene from the archaebacterium, *Halobacterium halobium*. Biochim. Biophys. Acta **949**:110–118.
12. **Izotova, L. S., A. Y. Strongin, L. N. Chekulaeva, V. E. Sterkin, V. I. Ostoslavskaya, L. A. Lyublinskaya, E. A. Timokhina, and V. M. Stepanov.** 1983. Purification and properties of serine protease from *Halobacterium halobium*. J. Bacteriol. **155**:826–830.
13. **Kamekura, M., and Y. Seno.** 1990. A halophilic extracellular protease from a halophilic archaebacterium strain 172P1. Biochem. Cell Biol. **68**:352–359.
14. **Kovaleva, N. A., A. A. Konnova, I. A. Tsaplina, and K. V. Gorbunov.** 1977. Biosynthesis of protease from *Halobacterium salinarium*. Appl. Biochem. Microbiol. **13**:501–508 (in Russian).
15. **Lin, X.-L., and J. Tang.** 1990. Purification, characterization, and gene cloning of thermopsin, a thermostable acid protease from *Sulfolobus acidocaldarius*. J. Biol. Chem. **265**:1490–1495.

16. **Matteucci, M. D., and M. H. Caruthers.** 1981. Synthesis of deoxyoligonucleotides on a polymer support. J. Am. Chem. Soc. **103**:3185–3191.
17. **Meloun, B., M. Baudys, V. Kostka, G. Hausdorf, C. Frommel, and W. E. Hohne.** 1985. Complete primary structure of thermitase from *Thermoactinomyces vulgaris* and its structural features related to the subtilisin-type proteinases. FEBS Lett. **183**:195–200.
18. **Norberg, P., and B. V. Hofsten.** 1969. Proteolytic enzymes from extremely halophilic bacteria. J. Gen. Microbiol. **55**:251–256.
19. **Ogawa, Y., H. Tatsumi, S. Murakami, Y. Ishida, K. Murakami, A. Masaki, H. Kawabe, H. Arimura, E. Nakano, H. Motai, and A. Toh-e.** 1990. Secretion of *Aspergillus oryzae* alkaline protease in an osmophilic yeast, *Zygosaccharomyces rouxii*. Agric. Biol. Chem. **54**:2521–2529.
20. **Rodriguez-Valera, F., F. Ruiz-Berraquero, and A. Ramos-Cormenzana.** 1981. Characteristics of the heterotrophic bacterial populations in hypersaline environments of different salt concentrations. Microb. Ecol. **7**:235–243.
21. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
22. **Schmitt, W., U. Rdest, and W. Goebel.** 1990. Efficient high-performance liquid chromatographic system for the purification of a halobacterial serine protease. J. Chromatogr. **521**:211–220.
23. **Southern, E. M.** 1975. Deletion of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
24. **Tatsumi, H., Y. Ogawa, S. Murakami, Y. Ishida, K. Murakami, A. Masaki, H. Kawabe, H. Arimura, E. Nakano, and H. Motai.** 1989. A full length cDNA for the alkaline protease from *Aspergillus oryzae*: structural analysis and expression in *Saccharomyces cerevisiae*. Mol. Gen. Genet. **219**:33–38.
25. **Thomm, M., G. Wich, J. W. Brown, G. Frey, B. A. Sherf, and G. S. Beckler.** 1989. An archaebacterial promoter sequence assigned by RNA polymerase binding experiments. Can. J. Microbiol. **35**:30–35.
26. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153**:3–11.
27. **Watson, M. E. E.** 1984. Compilation of published signal sequences. Nucleic Acids Res. **12**:5145–5164.
28. **Wells, J. A., E. Ferrari, D. J. Henner, D. A. Hestell, and E. Y. Chen.** 1983. Cloning, sequencing, and secretion of *Bacillus amyloliquefaciens* subtilisin in *Bacillus subtilis*. Nucleic Acids Res. **11**:7911–7925.
29. **Woese, C. R., O. Kandler, and M. L. Wheelis.** 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. USA **87**:4576–4579.
30. **Yanagida, N., T. Uozumi, and T. Beppu.** 1986. Specific excretion of *Serratia marcescens* protease through the outer membrane of *Escherichia coli*. J. Bacteriol. **166**:937–944.
31. **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.