Molecular Cloning of Two *Haemaphysalis Longicornis* Cathepsin L-like Cysteine Proteinase Genes

Albert MULENGA, Chihiro SUGIMOTO, Geoffrey INGRAM, Kazuhiko OHASHI and Misao ONUMA

Laboratory of Infectious Diseases, Department of Disease Control, Graduate School of Veterinary Medicine, Hokkaido University, Sapporo 060–0818, Japan

(Received 5 November 1998/Accepted 6 January 1999)

**ABSTRACT.** Immunological protection of mammalian hosts against tick infestation has been proposed as the most sustainable alternative tick control method to the current use of acaricides which has several limitations. The success of this method is dependent on the identification of key molecules for use as tick vaccine antigens. Proteolytic enzymes are involved in a wide range of cellular processes in eukaryotes such as development regulation and nutrition, thus they can be considered as good target antigens for a tick vaccine. In the present study we used primers designed based on the consensus amino acid motifs flanking the conserved active sites C$^{25}$ and N$^{175}$ present in all papain-like cysteine proteinases to amplify by polymerase chain reaction, sequence and characterize two *Haemaphysalis longicornis* tick cysteine proteinase genes. Based on the nucleotide and deduced amino acid sequences, both genes were identified as members of the cysteine proteinase gene family by presence in their sequences of consensus motifs flanking the conserved active sites C$^{25}$, H$^{150}$ and N$^{175}$ that are present in all papain-like cysteine proteinases. Both genes are about 1.2 kb in size and show high sequence homology predominantly to invertebrate cathepsin L-like cysteine proteinases.—**KEY WORDS:** cathepsin L-like cysteine proteinase, *Haemaphysalis longicornis*, tick vaccine.

Ticks are obligate ectoparasites and vectors of arboviruses, rickettsiae, spirochetes and parasitic protozoa of humans and domestic animals [19]. Ticks cause large economic losses directly as pests and indirectly by transmitting diseases and pose a major threat to the livestock industry of the world. At present, ticks are predominantly controlled by use of chemical acaricides, but due to the emergence of tick resistance to acaricides, contamination of the food chain and environment and the increasing community awareness to the health ill-effects of chemicals, it has become necessary to develop alternative tick control methods [26, 35]. Evidence that has been collected from studies based on crude vaccine antigens have shown that host vaccination against tick infestation is a sustainable alternative tick control method [32]. However, for this method to succeed, identification and production of key tick molecules in sufficient amounts for vaccination is needed [11].

The present study is part of our ongoing project aimed at identifying tick molecules with potential for use as targets for the development of alternative tick control methods. Proteolytic enzymes may represent interesting target molecules for an anti-tick vaccine or drug design because of their involvement in the mediation of a wide range of cellular processes in living organisms such as protein metabolism and processing of precursor proteins [8, 27, 36]. While available data is limited there is evidence to show that proteolytic enzymes are involved in the mediation of key physiological functions such as tick embryo development [13, 14, 21] and digestion of host red blood cells and proteins [1, 24]. Despite the potential to target proteolytic enzymes as candidate antigens for anti-tick drugs or vaccines, several aspects on their involvement in tick physiology remain unknown.

While presence of cathepsin L-like cysteine proteinases enzymes has been demonstrated in ticks by use of substrate hydrolysis and inhibitor sensitivity analysis [13, 14, 24], the present study is the first report on molecular characterization of cysteine proteinase genes from a hard tick.

**MATERIALS AND METHODS**

**Purification of tick mRNA:** Adult female *H. longicornis* ticks were obtained from a colony of ticks maintained on Japanese white rabbits. Ticks partially fed for 4 days were detached from rabbits by traction using a pair of forceps and left at room temperature for 1 hr to shed of remnants of rabbit tissue and hair, after which they were pulverized in liquid nitrogen. Total RNA was extracted using the TRIZOL reagent (GIBCO BRL, U.S.A.) according to the manufacturers instructions.

**Cloning of the full length *Haemaphysalis longicornis* tick cysteine proteinase genes A and B (HLCG-A and -B) by RACE:** A possible tick specific cysteine proteinase oligonucleotide primer (PF-1, see Fig. 1) was designed from consensus amino acid sequences flanking the active site, C$^{25}$ (residue position is based on the papain amino acid sequence [25]) [10, 27, 29, 30]. In order to design this primer, the consensus amino acid sequence was reverse translated using the *H. longicornis* codon usage frequency table that was generated from the deduced nucleotide and amino acid sequences of two *H. longicornis* serine proteinases genes (DDJB accession numbers, ABO20543 and ABO20544, [Mulenga et al., unpublished]) using the
About 5 µg of total RNA was used in a standard 3' RACE RT-PCR protocol to generate the PCR template first strand cDNA according to the manufacturer’s instructions (GIBCO BRL, U.S.A). One µl aliquot of the RT-PCR mixture was used in the first PCR with PF-1 plus RACE universal amplification primer (AUAP, GIBCO BRL, U.S.A.). Following cloning and DNA sequencing of the 3' end cysteine proteinase gene fragments, gene specific primers (GSPA-1, -2 and -3 for gene A, GSPB-1, -2, -3 for gene B) shown in Fig. 1, were designed and used in a 5' RACE protocol to amplify the 5' end fragments of both HLCG-A and -B (2nd PCR). Following sequencing of the 5' ends, specific GSPA-4/GSPB-4 were designed and used in a 3' RACE protocol to amplify full length HLCG-A and -B (3rd PCR). All PCR was carried out in a buffer containing 1.25 U of Taq polymerase (Pharmacia, Sweden), 45 mM Tris-HCl (pH 8.8), 11 mM ammonium sulfate, 4.5 mM MgCl₂, 6.7 mM 2-mecaptoethanol, 4.4 mM EDTA, 113 mg/ml BSA and 1 mM dNTPs. The amplified fragments were ligated into pGEMT vector (Promega, U.S.A.) and transformed into Escherichia coli (DH5α). The nucleotide sequences of the cloned PCR products were determined as described below.

**DNA sequencing and analysis:** Using the dye terminator system (Applied Biosystems, U.S.A.) and the automated sequencer, 310 Genetic analyser (Applied Biosystems, U.S.A.), the nucleotide sequences of the cloned PCR fragments were determined using the vector specific T7 and SP6 promoter primers as well as gene specific primers where necessary. The template for sequencing was generated by purification of the recombinant plasmid DNA using the QIAprep plasmid DNA isolation spin columns (Qiagen, U.S.A.) according to the manufacturer’s instructions. DNA sequence analysis was done using the GENETYX-MAC software package in combination with the GenBank and SWISS-PROT data bases for comparison with other known gene or protein sequences.

**Northern blotting analysis:** Thirty µg of total RNA extracted from partially fed adult ticks was electrophoresed on a 1% formaldehyde agarose gel in a formamide running buffer as described elsewhere [31]. The electrophoresed RNA was transferred to Hybond N+ (Amersham, U.K.) by capillary transfer method as described elsewhere [31] and UV-crosslinked for 3 min using a UV transilluminator (UVP Inc. U.S.A.). The cloned full length cDNAs, HLCG-A and HLCG-B used as probes were radiolabeled with α³²P using the Multiprime Labelling kit (Amersham, U.K.) according to the manufacturer’s instructions and hybridization was carried out overnight at 42°C in a buffer containing 50% formamide, 5X SSPE, 1X Denhardt’s salts, 0.1% SDS and 20 µg/ml salmon sperm DNA. The membranes were washed to a final stringency of 0.1X SSC plus 0.1% SDS for 90 min at 65°C for 15 min and subsequently exposed to an X-ray film at -80°C for two weeks.

**Southern blotting analysis:** Tick genomic DNA extracted according to methods described elsewhere [31], digested by restriction enzymes, Hind III, EcoRI, XhoI (New England Biolab Inc. U.S.A.) overnight at 37°C was electrophoresed on a 0.8% agarose gel and blotted on to Hybond N+ membranes by capillary transfer method [31]. Following the transfer, DNA was crosslinked by baking the dried membranes at 120°C for 30 min. The hybridization probes were prepared as described above and hybridization carried out overnight at 65°C in a buffer containing 0.5 M sodium phosphate pH 7.2, 1 mM EDTA and 7% SDS [7]. The membranes were washed to a final stringency of 0.1X SSC plus 0.1% SDS for 90 min at 65°C and subsequently exposed to an X-ray film as described above.
RESULTS

PCR amplification, cloning and analysis of tick cysteine gene fragments: A specific 700 bp 3’ end cysteine gene fragment was amplified with PF-1 in the first PCR based on a 3’RACE protocol. DNA sequencing and structural analysis of 10 randomly selected clones of the 700 bp product showed that they were all cathepsin L-like cysteine proteinases. Nine of these clones were identical and referred to in the text as H. longicornis cysteine proteinase gene A (HLCG-A) while the remaining clone, referred to as HLCG-B was distinctly different from HLCG-A. The cloned gene fragments were identified as members of the cysteine proteinase gene family by presence of consensus motifs highly conserved among all known papain-like cysteine proteinases in their deduced amino acid sequences [25, 27, 29, 30].

Cloning, DNA sequence and analysis of full length HLCG-A and HLCG-B: The cloning strategy used to amplify and clone full length HLCG-A and HLCG-B is described under Materials and Methods and summarised in Fig. 1. HLCG-A has an open reading frame (ORF), extending from position 135 in frame to position 1068 encoding a 312 amino acid residue polypeptide with 33.7 kDa predicted molecular mass. The ORF of the second molecule, HLCG-B, extends in frame from position 44 to 1042 and encodes a 332 amino acid residue polypeptide with 37 kDa predicted molecular mass. The two genes showed homologies of 68.1% and 33.7% homology at nucleotide and amino acid levels respectively. Conserved motifs of cysteine proteinases were identified in the deduced HLCG-A and -B (Fig. 3, boxes).

HLCG-A and -B sequence homology to other known cysteine proteinases on the data base: In order to compare the similarity of HLCG-A and -B gene predicted amino acid sequences to the already known cysteine proteinases on the data base, sequences of both genes were used to search the SWISS-PROT data base. Both genes, HLCG-A and -B showed homology to invertebrate cathepsin L-like cysteine proteinase genes (Fig. 2). Homologies to known proteinase genes listed in Fig. 2 were between 33% to 42.9%.

Northern blotting analysis: Based on the sizes of the PCR products for both genes (result not shown) as well as the observed mRNA sizes on the Northern blot analyses of tick RNA (Fig. 3a and b), the sizes of full length HLCG-A and -B is about 1.2 kbp. As shown in Fig. 3a, the mRNA band for HLCG-A was stronger than that of HLCG-B (Fig. 3b).

Southern blotting analysis: On the Southern blot analyses of tick genomic DNA digested with indicated restriction enzymes, the hybridization patterns showed that HLCG-A (Fig. 4a) is a single copy. As shown in Fig. 4b, the HLCG-B probe hybridized to a single band for HindIII-digested DNA and 2 bands for EcoRI-digested DNA. This result indicated that there were 2 or more copies of HLCG-B gene, because there was no EcoRI site in HLCG-B.

DISCUSSION

Cysteine proteinases have been identified from a wide range of organisms including viruses, bacteria, protozoa, yeast, plants, helminths, insects, soft ticks, and mammals [2, 4–6, 13, 14, 33]. The current study describes cloning and characterization of two cathepsin L-like cysteine proteinase genes from a hard tick, H. longicornis. Several workers have successfully used degenerate PCR primers synthesized from the cysteine proteinase consensus motifs flanking the active sites C25 and N175 to amplify the cysteine gene fragments [17, 29, 30]. However in our study, we did not succeed to amplify the tick cysteine proteinase genes using degenerate primers, because of high background amplification, due to low temperatures required for degenerate primers to anneal. We were able to overcome this, by designing possible gene specific tick cysteine proteinase primers which could anneal at higher temperatures, with minimal background amplification. The primer design strategy and primers used in this study will be useful to clone homologues from other tick spp. Domain structure analysis showed that the active sites involving C25, H150 and N175 present in all papain-like cysteine proteinases [25] were conserved in both HLCG-A and HLCG-B amino acid sequences. In addition, the amino acid residues surrounding C25, H150 and N175 were equally highly conserved, although the overall homologies of the proteinases in this study to other known proteinases on the data base were in the range of between 33–43%.

Cysteine proteinases have been shown to play important roles in the pathogenesis of several parasitic protozoa and helminth infections [8, 9, 22, 28]. For instance, cysteine proteinases are involved in the degradation of hemoglobin by Plasmodium falciparum [27], adult schistosome worms [20], Rhodnius prolixus [3], the growth and development of Trypanosoma cruzi [15, 18, 23] and excystation of Giardia, a step necessary for infection to occur [37]. While it cannot be taken for granted that observations on cysteine proteinases in other parasites will be consistent in ticks, it is logical to assume that, the functions of cysteine proteinase enzymes in ticks are also involved in many biological activities.

There are no studies dealing with cysteine proteinases as candidates for either vaccine or drug design against hematophagous arthropods. However, studies on other classes of proteinases have provided indirect evidence to show that inhibition of proteolytic enzyme activity either by host vaccination or use of specific inhibitors can interfere with the biological capacity of hematophagous parasites. For example, vaccination of cattle against Hypoderma lineatum (cattle grub) with a purified serine protease stimulated a strong cellular and humoral immunity leading to 95% protection of cattle against cattle grub infestation [5]. Addition of serine protease inhibitors such as soybean trypsin to a blood meal inhibited stable fly egg production up to 71% [34] and killed or inhibited egg development of buffalo flies [12]. In another study, addition of pepstatin an
Fig. 2. Sequence similarity of HLCG-A and HLCG-B to already known cystein proteinase genes on the data base. Deduced amino acid sequences of HLCG-A and HLCG-B were used for comparison with already known proteinases on the SWISS-PROT data base. HLCG-A = *H. longicornis* cysteine proteinase gene A (DDJB accession number = ABO 20492), HLCG-B = *H. longicornis* cysteine proteinase gene B (ABO 20491, amino acid homology to HLCG-A, 33.7%), SZCLA1 (JC 5442, 36.5%) and SZCLA2 (D165533, 35.7%) = *Sitophilus zeamais* cathepsin L-like cysteine proteinase, DCLA3 (AF 012089-1, 35.2%) = *Drosophila melanogaster* cysteine proteinase-1 protein, DROCP (D31970-1, 42.9%) = *D. melanogaster* cathepsin L-like cysteine proteinase, RATCPL (Y00697-1, 37.5%) = *Rattus norvegicus* prepro-cathepsin L-like protein. Asterisks (*) = homologous regions. Box = motifs conserved in all papain-like cysteine proteinases. Boldfaced amino acid residues (C 25, H 150 and N 175, residue position based on the amino acid sequence of papain [25]) = conserved active sites.
inhibitor of acid proteinases to a blood meal completely inhibited molting and oviposition in *Rhodnius prolixus* [16]

Studies to determine the biological functions, the expression organ(s) in the tick and the vaccine effect of HLCG-A and -B against tick infestation are underway. It is anticipated that data from these studies will provide evidence with respect to the suitability of targeting cysteine proteinases as candidates for a tick vaccine.

ACKNOWLEDGEMENT. This work was funded by grants from the Ministry of Education, Science, Sports and Culture, Japan.

REFERENCES


