The Major Secreted Cathepsin L1 Protease of the Liver Fluke, *Fasciola hepatica*

**A LEU$^{12}$ TO PRO$^{12}$ REPLACEMENT IN THE NONCONSERVED C-TERMINAL REGION OF THE PROSEGMENT PREVENTS COMPLETE ENZYME AUTOACTIVATION AND ALLOWS DEFINITION OF THE MOLECULAR EVENTS IN PROSEGMENT REMOVAL**

Received for publication, December 15, 2006, and in revised form, March 2, 2007 Published, JBC Papers in Press, April 2, 2007, DOI 10.1074/jbc.M611501200

Colin M. Stack$^{1,1}$, Sheila Donnelly$^{1,2}$, Jonathan Lowther$^3$, WeiBo Xu$^4$, Peter R. Collins$^{1,5}$, Linda S. Brinen$^3$, and John P. Dalton$^4$

From the $^1$Institute for the Biotechnology of Infectious Diseases, University of Technology Sydney, Building 4, Harris Street, Ultimo, Sydney, New South Wales 2007, Australia, $^2$School of Biotechnology, Dublin City University, Dublin 9, Republic of Ireland, and the $^3$Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94143-2140

A protease secreted by the parasitic helminth *Fasciola hepatica*, a 37-kDa procathepsin L1 (FheproCL1), autocatalytically processes and activates to its mature enzyme (FheCL1) over a wide pH range of 7.3 to 4.0, although activation is more rapid at low pH. Maturation initiates with cleavages of a small proportion of molecules within the central region of the prosegment, possibly by intramolecular events. However, activation to fully mature enzymes is achieved by a precise intermolecular cleavage at a Leu$^{12}$-Ser$^{11}$↓His$^{10}$ sequence within the nonconserved C-terminal region of the prosegment. The importance of this cleavage site in enzyme activation was demonstrated using an active site variant FheproCL1Gly$^{26}$ (Cys$^{26}$ to Gly$^{26}$) and a double variant FheproCL1Pro$^{12}$/Gly$^{26}$ (Leu$^{12}$ to Pro$^{12}$), and although both of these variants cannot autocatalytically process, the former is susceptible to trans-processing at a Leu$^{12}$-Ser$^{11}$↓His$^{10}$ sequence by pre-activated FheCL1, but the latter is not. Another *F. hepatica* secreted protease FheCL2, which, unlike FheCL1, can readily accept proline in the S2 subsite of its active site, can trans-process the double variant FheproCL1Pro$^{12}$/Gly$^{26}$ by cleavage at the Pro$^{12}$-Ser$^{11}$↓His$^{10}$ sequence. Furthermore, the autoactivation of a variant enzyme with a single replacement, FheproCL1Pro$^{12}$, was very slow but was increased 40-fold in the presence of FheCL2. These studies provide a molecular insight into the regulation of FheproCL1 autocatalysis.

*Fasciola hepatica* is a parasitic helminth of livestock and humans (1). Virulence is associated with the secretion of a variety of proteins, the most predominant of which are cathepsin L proteases. The cathepsin L proteases are synthesized by epithelial cells that line the parasite gut and are stored as inactivezymogens in specialized secretory vesicles prior to extrusion into the parasite gut lumen (2, 3). Here the proteases mix with blood derived from the host and catabolize proteins to peptides and amino acids that are absorbed and used as nutrient by the parasite (3). The cathepsin L proteases are also secreted into the tissues of the host where they modulate the reaction of the host to the parasite by suppressing the elaboration of protective Th1-driven immune responses via a mechanism that, albeit unknown, is dependent on enzyme activity (4, 5). Their importance in these virulence-related functions has made cathepsin L proteases prime targets for vaccines against the disease they cause; we have shown that vaccine-induced antibodies can disrupt their function and protect cattle and sheep hosts against infection (6–8).

The *F. hepatica* cathepsin Ls are orthologous to the mammalian cathepsin Ls, and like these and other members of the papain superfamily, theirzymogens possess an N-terminal extension or prosegment. Prosegments function as regulators of activity by binding to the substrate cleft in a reverse substrate mode (9–11), and they act as intermolecular chaperones that are crucial to protein folding (12–14) and intracellular trafficking (15–17). Cleavage between the prosegment region and the catalytic domain is necessary to generate the active mature enzyme and usually occurs in acidic environments, whether that be intracellular (9, 18–21) or extracellularly (10, 22–25). Various mammalian cathepsins, including cathepsins L (9, 10, 18, 26–29), S (10, 30–32), K (10, 33, 34), and B (9, 24–27), autoactivate in vitro under acidic and reducing conditions. The *F. hepatica* cathepsin Lzymogen autoactivates within the slightly acidic pH of the parasite gut lumen and, similarly to the mammalian enzymes, autoactivates in vitro at low pH (2).

The presence of a functional active site cysteine (Cys$^{25}$, papain numbering) is essential for enzyme activity, and site-directed mutagenesis studies with human (9, 15) and *F. hepatica* cathepsin L (2) have shown that autoactivation is prevented.
if this residue is substituted with another amino acid. Because enzyme activation proceeds quicker than proteolytic processing, it was suggested that enzyme maturation occurs in two steps as follows: the activation of the proenzyme followed by enzymatic cleavages within the prosegments to give rise to a mature enzyme (15). The initial “trigger” that initiates activation still remains elusive (15) but may be the disruption of a salt bridge involving Asp \(^{-36}\) (papain numbering) of the prosegment, contained within a conserved GXNXXFD motif, as the pH decreases (35). Subsequent removal of the prosegment involves possible intramolecular and/or intermolecular cleavages that take place within the central portion of the prosegment and then at the C-terminal region of the prosegment close to the N terminus of the mature enzyme (9, 10, 35).

The C-terminal parts of cathepsin prosegments (~20 residues) do not make significant contact with the mature enzyme, and removal of these C-terminal residues from free prosegments has no effect on their binding kinetics to their cognate enzymes (9, 15). The high B-factors observed in this portion of the proenzyme structure are believed to facilitate proteolytic cleavage at this site (9, 10), and several studies have shown that this region is particularly susceptible to trans-cleavage with the same or different protease (18, 25, 28, 32, 33). As a consequence of this lack of structural stability, within the papain superfamily, the C-terminal part of prosegments show little or no conservation (9).

We have investigated the molecular events involved in the activation of the \(F.\ hepatica\) procathepsin L1 (FheproCL1).\(^5\) Autocatalytic processing of FheproCL1 to a fully active mature enzyme (FheCL1) occurs most quickly at low pH, which is relevant to the digestive function of the enzymes in the slightly acid environment of the parasite gut, but can also occur at physiological pH. Our studies suggest that the initial events in thezymogen activation process involve the generation of a small proportion of catalytically active molecules by cleavages, possibly intramolecular, in the central portion of the prosegment. Further processing occurs by a precise cleavage at the \(\text{Leu}^{12}\)-Ser\(^{-11}\) ↓ His\(^{-10}\) sequence within the C terminus of the prosegment that gives rise to a fully active mature enzyme with several prosegment amino acids still attached. We show that an active site variant, FheproCL1Gly\(^{26}\), that cannot autoactivate is susceptible to \(\text{trans}\)-processing with pre-activated wild-type FheCL1, but the double variant, FheproCL1Gly\(^{26}/\text{Pro}^{-12}\), is resistant. However, a second \(F.\ hepatica\) protease, FheCL2, that, unlike FheCL1, can readily accept proline in the S2 subsite of its active site can \(\text{trans}\)-process the FheproCL1Gly\(^{26}/\text{Pro}^{-12}\) by cleaving within the new \(\text{Pro}^{-12}\)-Ser\(^{-11}\) ↓ His\(^{-10}\) sequence. The single cathepsin L mutant, FheproCL1Pro\(^{-12}\), exhibits a very low rate of activation that is increased 40-fold following addition of FheCL2. These studies reveal the critical importance of the \(\text{Leu}^{12}\)-Ser\(^{-11}\) ↓ His\(^{-10}\) sequence in regulating the activation of the major secreted protease of the helmint pathogen \(F.\ hepatica\).

---

\(^5\) The abbreviations used are: FheproCL1, \(F.\ hepatica\) procathepsin L1; Z, benzylxoxycarbonyl; NHMec, 4-methylcoumarinyl-7-amide; PBS, phosphate-buffered saline.

### EXPERIMENTAL PROCEDURES

**Transformation and Expression of \(F.\ hepatica\) Procathepsin L and Variants in \(Pichia\ pastoris\)**—The \(F.\ hepatica\) procathepsin L (FheproCL1) was amplified by PCR from the \(pAAH5\ Saccharomyces\ cerevisiae\) expression vector into which the full-length cDNA had been cloned previously (36–38). Primers were used (see primers A and D below) to incorporate an SnaBI restriction site at the 5′ end of the gene and an AvrII restriction site and His\(_{\text{tag}}\) sequence at the 3′ end. The 980-bp fragment was inserted into the AvrII/SnaBI site of \(P.\ pastoris\) expression vector pPIC9K (Invitrogen).

From the pPIC9K-FheproCL1 DNA template an active site mutant was prepared by replacing Cys\(^{26}\) with a glycine using the PCR-based site-directed mutagenesis method of gene splicing by overlap extension (SOEing) (39). The primers used were as follows: primer A, 5′-GCGGCTACGTATCGAATGATGA-TTGTGGCAT-3′; primer B, 5′-GAATGCACCAACCGGAGCCAC-3′; primer C, 5′-GTGGCTCCGGTTGGCCATTC-3′; and primer D, 5′-GCGCTAGTCAGTGGTGTGGTGGTGTTGGGCC-CCC′. The underlined nucleotides indicate the replacements introduced. Each reaction used one flanking primer that hybridized at one end of the target sequence (primer A or D) and one overlapping internal primer that hybridized at the site of the mutation and contained the mismatched base (primer B or C). In the first round of amplification, two sections of the cDNA were amplified using primer combinations A + B and C + D. These two PCR products, with an overlap of 21 bp at one end of each fragment, were then combined in a second PCR to amplify the entire FheproCL1Gly\(^{26}\) cDNA. Primers for this reaction were the two outside primers used in each of the two first round reactions (primers A and D). High Fidelity Advantage Taq polymerase (Clontech) was used in all PCRs (25 cycles at 94 °C for 30 s, 55 °C for 1 min, and 72 °C for 2 min).

The FheproCL1 cDNA and the variant FheproCL1Gly\(^{26}\) were inserted into the AvrII/SnaBI site of expression vector pPIC9K. The plasmids were linearized by digestion with SalI and introduced to \(P.\ pastoris\) GS115 cells by spheroplasting (40). Both FheproCL1 and FheproCL1Gly\(^{26}\) plasmid inserts were sequenced to verify the sequence and substitutions.

**Screening of Transformed Yeast and Identification of Variants**—Yeast transformants were replica-plated onto yeast/peptide/dextrose medium plates and grown overnight. Transformants were lifted onto nitrocellulose filters (0.45 μm, Sartorius) by laying the filters on top of the colonies, and those secreting enzyme were identified by colony blotting as described (41). Yeast transformants were cultured in 50 ml of buffered glycerol-complex medium; protein expression was induced with 1% methanol in buffered methanol-complex medium, and samples of culture supernatants were analyzed by 12% SDS-PAGE to verification of protein secretion (2).

During the screening of yeast clones for the Cys\(^{26}\) to Gly\(^{26}\) variant, another procathepsin L variant was identified by its noticeably slower migration in SDS-polyacrylamide gels. The DNA encoding this proenzyme was sequenced and shown to have, in addition to the Cys\(^{26}\) to Gly\(^{26}\) replacement, a proline substituted for a leucine residue at position –12 in the proseg-
ment (cathepsin L1 numbering), and this clone was termed FheproCL1Pro$^{12}$/Gly$^{26}$. Because this clone became of interest for this study, a variant containing an unchanged active site Cys$^{26}$ but with the single Leu$^{12}$ to Pro$^{12}$ substitution was constructed using a similar procedure as described above and termed FheproCL1Pro$^{12}$. Purification of Recombinant Proteases—Upscale production and isolation of recombinant proteases were performed essentially as described by Collins et al. (2) with the following important modifications: yeast transformants were cultured in 1 liter of buffered glycerol-complex medium, buffered to pH 8.0, in 5-liter baffled flasks at 30 °C until an $A_{650}$ of 2–6 was reached. Cells were harvested by centrifugation at 10000 × g for 5 min, and protein expression was induced by resuspending the cells in 200 ml of buffered methanol-complex medium, buffered to pH 8.0, containing 1% methanol (38). The cultures were grown at 30 °C with shaking at 225 rpm for 12 h. Because the pH of the medium drops slowly over this period, it was readjusted with 0.1 M NaOH every 3 h to pH 8.0. Recombinant proteins were purified from yeast medium by affinity chromatography using nickel-nitrilotriacetic acid-agarose (Qiagen) as described by Collins et al. (2). However, bound proteins were eluted using 10 ml of 50 mM sodium phosphate buffer, pH 7.0, containing 300 mM NaCl and 250 mM imidazole, and collected into tubes containing 5 ml of 100 mM sodium phosphate, pH 7.8, before being dialyzed into PBS, pH 7.3. Samples were stored at −80 °C for long term or at 4 °C for up to a week during experiments. Samples of purified recombinants were analyzed by SDS-PAGE followed by Coomassie Blue R staining.

Autocatalytic Processing and Activation and trans-Processing of Recombinant Proteases—Auto-processing and activation of purified recombinant enzymes (100 μg in 50 μl) was performed by adding these to activation buffer (0.1 M sodium acetate buffer, pH 4.5, sodium phosphate, pH 6.0, or PBS, pH 7.3, containing 1 mM dithiothreitol) at 37 °C, and samples were removed at various time points and transferred into tubes containing 1 μl of 1 mM E-64 to halt the reaction. Proteolytic cleavage of the prosegment was visualized by 12% SDS-PAGE.

To correlate the proteolytic cleavage events that occur during enzyme activation with the emergence of functionally active mature recombinant FheCL1, the proenzyme was autoactivated at pH 4.5, and aliquots were snap-frozen in liquid nitrogen at time 0 and at various time points thereafter. The following day the samples were thawed on ice and (a) analyzed by reducing SDS-PAGE, and (b) assayed for proteolytic activity for 10 min against the small peptide substrate Z-Phe-Arg-NHMec (see below) and against the protein substrate chicken ovalbumin (ratio of 20:1 protein to enzyme). Other aliquots were acidified by adding 20 μl of 5% trifluoroacetic acid to the frozen sample and then allowed to thaw. These were then loaded onto a biobrene-treated, preycled filter and subjected to five cycles of N-terminal (Edman) sequencing on an Applied Biosystems 494 Procise protein sequencing system. Other aliquots taken from the various time points in the autocatalytic reactions were diluted 50:50 with acetonitrile, water, 1% formic acid solution. The diluted samples were analyzed at the Australian Proteome Analysis Facility (Sydney, Australia) by electrospray ionization mass spectrometry using a Q-TOF Ultima (Water/Micromass, Manchester, UK).

The trans-processing of FheproCL1Gly$^{26}$ and FheproCL1Pro$^{12}$/Gly$^{26}$ was carried out by mixing 50 μg of purified mutant recombinant with 5.0 μg of the recombinant FheCL1 or FheCL2 that had been activated by the endogenous activation method described above. The mixture was then incubated for 3 h at 37 °C, and samples were removed at various time points for analysis by 12% SDS-PAGE. The trans-processing of FheproCL1Pro$^{12}$ by FheCL2 was performed by mixing 100 μg of purified mutant recombinant with 5 μg of the activated enzyme in a volume of 500 μl of 0.1 mM sodium phosphate, pH 6.0. Aliquots of 10 μl were taken over a period of 1 h and analyzed by 12% SDS-PAGE. Samples taken at time 0 were also assessed for enzyme activity over the subsequent 10 min as described below.

In experiments involving trans-processing by FheCL2, proteins were transferred to polyvinylidene difluoride membrane using a semi-dry transfer cell (Bio-Rad) at 15 V for 20 min. The membrane was washed with distilled H2O and stained with 0.025% Coomassie Brilliant Blue R-250 in 40% methanol (37, 38). Protein bands were subjected to N-terminal sequencing at Genosphere Biotechnologies (Paris, France) or the Australian Proteomic Analysis Facility (Sydney, Australia).

Fluorogenic Peptide Assays—Cathepsin L activity was determined by a fluorometric assay employing the substrate Z-Phe-Arg-NHMec (37, 38). The activity of the samples was calculated from a standard curve of NHMec ranging from 0 to 10 μM and presented as nanomoles of NHMec min$^{-1}$ ml$^{-1}$. Protein concentration was determined by the method of Bradford using bovine serum albumin as standard.

The progress of the autocatalytic activation of wild-type FheproCL1 and variant FheproCL1Pro$^{12}$ was observed by incubating the zymogen in the presence of the substrate Z-Phe-Arg-NHMec and measuring fluorescence over time. Processing was initiated by adding diluted zymogen to the assay mixture containing 50 mM acetate buffer, 2 mM substrate, and 2 mM dithiothreitol. Progress curves were fitted to Equation 1 as described previously by Menard et al. (15) to analyze the rate of activation of human procathepsin L to mature enzyme,

$$[P] = v_F t - [(v_F - v_E)/k_{obs})](1 - e^{-k_{obs}t})$$

(Eq. 1)

where $v_F$ and $v_E$ represent the initial rate of product release for proenzyme and mature enzyme, respectively, and $k_{obs}$ is a first-order rate constant.

Prosegment Sequences Comparisons—The prosegment sequence of the F. hepatica FheproCL1 was aligned with other sequences using ClustalX 1.81. Protein sequences included Carica papaya, human cathepsin L, human cathepsin S, and human cathepsin K.6 Sequences were numbered according to papain numbering, and prosegment residues were recorded as a
Autoprocessing and Activation of F. hepatica Cathepsin L1

Homology-based Molecular Modeling—A model structure of FheproCL1Gly26 was built using the ModWeb interface to the Modeller program for comparative protein structure modeling (42–44). The 1.8 Å structure of human procathepsin L (Protein Data Bank code 1CS8) was used as a three-dimensional template of related fold. Generated models were visualized and compared with Coot (45) and with PyMOL (56).

RESULTS

Autocatalytic Activation of FheproCL1—We reported previously the expression of FheproCL1 in the yeast P. pastoris, and its isolation by affinity chromatography (2). In that study we found that FheproCL1 in the yeast medium was present as a 37-kDa component and several breakdown products, including a major intermediate form of 30 kDa. We have since discovered that this breakdown occurred during the 24-h fermentation process as the pH of the medium decreased to as low as pH 6.0. By readjusting the pH of the medium to pH 8.0 every 3 h, and by reducing the induction time to 12 h, a single protein product was purified that migrated to 37 kDa (see Fig. 1A, 2nd lane). Mass spectrometry of the purified preparation identified a species of 38,884.22 Da, and N-terminal sequencing showed that the product represents the FheproCL1 zymogen with five amino acids of the yeast α-factor signal peptide still attached (see Table 1). These data agree with the predicted molecular mass for FheproCL1 with these additional five amino acids of 36,843.04 Da.

To investigate the processing events involved in the activation of FheproCL1 to its mature enzyme, the 37-kDa zymogen, which was stored at 4 °C in PBS, pH 7.3, was diluted 10-fold in 0.1M sodium acetate, pH 4.5, at 37 °C. Aliquots of the reaction

| Cleavage site no. | N-terminal sequence | Mass spectrometry | Predicted molecular mass (kDa) | Amount of first residue (picomoles)*** |
|-------------------|--------------------|--------------------|-------------------------------|---------------------------------------|
|                   | Obtained           | Data               | Predicted molecular mass      |                                       |
|                   |                    | (kDa)              | (kDa)*                        |                                       |
| 1                 | EAEAY...           | 36.884.22          | 36.843.04                     | 24 37 47.2 51.1 88.1^ 108.7^ 123^    |
| 2                 | AYVSN...           | ND**               | 36.513.73                     | 7.6 9.0 15.6 31.7 70.8 88.6 70.8     |
| 3                 | NDDL...           | 36.002.15         | 36.094.26                     | 2.6 3.4 8.5 17.2 40.8 46.2 68.3      |
| 4                 | HGVPV...           | 26.351.10         | 26.352.00                     | - 3.7 18.6 27.5 44.9 65.1 67.7      |
| 5                 | DILS...           | 26.679.97         | 26.753.87                     | - 11.1 38.3 49.0 71.9                |
| 6                 | EANN...           | 25.810.00         | 25.771.77                     | - - - - 88.1^ 108.7^ 123^           |
| 7                 | RMY...            | -                  | -                             | - - - - 17.3 19.9                    |
| 8                 | L.R...            | -                  | -                             | - - - - 10.9 14.7                    |
Autoprocessing and Activation of F. hepatica Cathepsin L1

The first major processing event to occur within the prosegment is at the Leu$^{12}$-Ser$^{11}$ → His$^{10}$ sequence (cleavage 4) in the nonconserved region of the prosegment and appears first at 5 min and increasingly throughout the subsequent 180-min reaction (Fig. 2A and Table 1). The molecular mass of the fragment produced was determined to be 26.351 kDa by mass spectrometry. Subsequent to, and therefore likely dependent on, cleavage 4, a slightly higher molecular mass fragment of 26.679 kDa was observed that resulted from a cleavage at Ser$^{15}$ → Asp$^{14}$, four amino acids to the N terminus (cleavage 5). These fragments would co-migrate in the SDS-PAGE and thus both represent the 26-kDa fragment observed (Fig. 2A). Following these cleavages, the procathepsin L was further processed at Tyr$^{6}$ → Glu$^{5}$ to finally give rise to a mature protein with five amino acids of the prosegment still attached; the final product had a molecular mass of 25.81 kDa determined by mass spectrometry (see Table 1).

The data show that processing first occurs at the Leu$^{12}$-Ser$^{11}$ → His$^{10}$ sequence and that this is followed by further trimming toward the juncture between the prosegment and mature enzyme (Table 1). This fully activated FheCL1 cleaved fluorogenic and protein substrates over a wide pH range, pH 3.5 to 8.5, with an optimum at pH 6.2 ($pK_a = 3.87 ± 0.07$ and $pK_a = 8.14 ± 0.08$) (data not shown).

To determine the relative activity of the FheCL1 during the autocatalytic processing event, and to correlate activity with the cleavage fragments produced, parallel samples to those shown in Fig. 2A were analyzed for proteolytic activity against the Z-Phe-Arg-NMhcr at pH 6.2, the optimal pH for enzyme activity, for 10 min (Fig. 2B). It is important to note that the proFheCL1 enzyme continues to activate during these assay conditions (see below), and hence, true calibration curves could not be established. As shown in Fig. 2B, the first time point does not begin at 0. Nevertheless, it can be observed that after an initial slow start between 0 and 5 min, the enzyme activity increases in the first 60 min and slows between time points 120 and 180 min, when processing is complete. The total enzyme activity at time point 180 was 40 times greater than at the zero time point. When aliquots of the same reaction were mixed with chicken ovalbumin (which migrates as two bands at 26 kDa), this band reduces slightly further in size as the experiment is monitored to 180 min (Fig. 2A).

The sequence of processing events in autocatalytic activation of FheproCL1 at pH 4.5 was followed by removing aliquots and halting the reaction by snap-freezing in liquid nitrogen. The samples were thawed and analyzed visually by 12% SDS-PAGE (Fig. 2A) for peptidolytic activity against chicken ovalbumin ($\alpha$-factor signal peptide in samples taken at time 0 and within the first 5 min of activation (cleavages 1, 2, and 3) suggesting that the enzyme activates very rapidly during the thawing period (these $\alpha$-factor residues are exposed on the recombinant proteins and are highly susceptible to cleavage) (Table 1).
cleavage points that occurred within these peptides at 120 and 180 min (cleavages 7 and 8).

Surprisingly, mass spectrometry of the reaction samples resolved only one small molecularly sized species of 9.359 kDa and only in the samples taken at 180-min samples. Analysis of the prosequence suggests that this predominant peptide represents a fragment beginning at cleavages 3 and an ending at cleavage 5 (predicted mass 9,357.40 Da); hence, the data indicate that at 180 min the prosequence remains in largely intact form (Table 1).

Reaction Progress Curves—The activation of FheproCL1 was examined in a continuous fluorescence assay by mixing the enzyme with substrate Z-Phe-Arg-NHMec and monitoring the release of -NHMec over time (Fig. 3A). These progress curves reveal that the autoactivation process increases with time and that the rate of activation is faster with increasing zymogen concentrations. When the first-order rate constants for autocatalytic processing, $k_{obs}$, were calculated, by fitting the progress curve data to Equation 1, as described by Menard et al. (15) and plotted as a function of enzyme concentration, it was observed that the rate of enzyme activation increases linearly with zymogen concentration (Fig. 3B). A similar linear relationship was observed for human cathepsin L, but as pointed out by Menard et al. (15), the equation may not exactly reflect the reaction rates as it does not take into account factors such as inhibition of activated enzyme by released propeptide (which remains intact throughout the 180-min of our reactions). The data show that the rate of autocatalytic activation is dependent on enzyme concentration and is an indication that activation occurs by a bimolecular process, i.e. intermolecular cleavage and activation of zymogen by activated mature enzyme.

Our experiments show that FheproCL1 is consistent with other papain-like cysteine proteases, which also autocatalytically process and activate at low pH (9, 10, 35). However, we found that FheproCL1 autocatalytically activated even in PBS at pH 7.3 when incubated at 37 °C, although the rate of activation was much slower than at pH 4.5 (Fig. 3C). Examination of the autocatalytic process at pH 7.3 by 12% SDS-PAGE and mass spectrometry shows that cleavage at the Leu$^{−12}$, Ser$^{−11}$ His$^{−10}$ sequence was observed in the first 60 min and processing to the complete mature enzyme at Tyr$^{−5}$ Glu$^{−4}$ occurred between 60 and 180 min (Fig. 1B).

We examined the autoactivation process at pH 7.3 more closely by stopping the reaction with the addition of E-64 immediately and at 2, 5, and 10 min after 10-fold dilution with activation buffer (PBS, pH 7.3, containing dithiothreitol and EDTA) at room temperature. As shown in Fig. 1C, FheproCL1 samples stored in PBS, pH 7.3, were homogeneous (time 0 min). However, samples taken immediately after addition of the reaction medium (time 0) contained very minor bands that migrated with and above the molecular size marker of 30 kDa. Bands in this region were also present at the 2-min time point but had disappeared by 5 min. At 10 min, the minor bands had disappeared, although the 26.351-kDa fragment increased in intensity (Fig. 1C). The precise sizes of the >30 band proteins were not determined because they were not detected in spectra of samples analyzed by mass spectrometry.

Expression of Recombinant Variants FheproCL1Gly$^{26}$ and FheproCL1Pro$^{−12}$/Gly$^{26}$—We prepared an active site variant of procathepsin L by replacing Cys$^{26}$ with a Gly, FheproCL1Gly$^{26}$. This variant migrated as a single band at ~37 kDa and had an identical N-terminal amino acid sequence to the wild-type FheproCL1 (Table 1). While screening the FheproCLGly$^{26}$ yeast colonies, we discovered a yeast clone that secreted a protein that migrated slightly slower on 12% SDS-PAGE than the recombinant FheproCL1Gly$^{26}$ variant (Fig. 4A). Amino acid sequence analysis revealed that this protein had a similar N-terminal sequence to the 37-kDa FheproCL1 (including the five amino acids of the yeast α-factor signal peptide); however, sequencing of the DNA revealed that a single nucleotide change
Autoprocessing and Activation of F. hepatica Cathepsin L1

FIGURE 4. trans-Processing of variant recombinant proteins FheproCL1Gly26 and FheproCL1Pro12/Gly26 by FheCL1. A, purified variants FheproCL1Gly26 and FheproCL1Pro12/Gly26 were incubated at pH 4.5 (0.1 mM sodium acetate) for 120 min. Aliquots of the reaction mixtures were removed at time 0, 60, and 120 min and analyzed by 10% SDS-PAGE. Neither variant is capable of autoactivation. B and C, 50 μg of purified recombinant FheproCL1Gly26 and FheproCL1Pro12/Gly26 were incubated with fully activated mature FheCL1 (5 μg) for up to 240 min in 0.1 mM sodium acetate, pH 4.5 (total reaction volume 100 μl). Aliquots (5 μl) were removed from the mixtures at time 0, 10, 30, 60, 90, 120, 180, and 240 min and analyzed by 10% SDS-PAGE. D, recombinant FheproCL1Gly26 and FheproCL1Pro12/Gly26 (shown in lanes marked G and PG, respectively) were incubated with fully activated mature FheCL1 in 0.1 mM sodium acetate at pH 3.0, 4.0 and 5.0 for 240 min. Lanes marked act show the fully activated mature FheCL1 incubated alone in control reactions.

trans-Processing of Variants FheproCL1Gly26 and FheproCL1Pro12/Gly26 by Active FheCL2—We noticed that the amino acid residue immediately N-terminal to the cleavage within the Leu12-Ser11 ↓ His10 sequence places a Leu residue in the P2 position, and hence, this would occupy the S2 subsite of the protease active site. The specificity of papain-like protease for peptide bonds is determined by the P2 residue (10, 11), and previous studies in our laboratory using a range of fluorogenic peptides demonstrated that leucine is particularly favored in this position by FheCL1 (37, 38). In contrast, we have found that FheCL1 does not efficiently cleave peptides with Pro in the P2 position. However, a second cathepsin L enzyme characterized by us, termed FheCL2, can cleave substrates containing Pro in the P2 position (37, 38). We therefore tested the ability of this FheCL2 to trans-process the recombinant variant FheproCL1Pro12/Gly26. The results, shown in Fig. 5, demonstrated that although activated FheCL1 could not trans-process the FheproCL1Pro12/Gly26 variant to a mature enzyme, under the same conditions FheCL2 cleaved the variant to produce the mature enzyme of 26 kDa, which was determined to be 26,352 Da by mass spectrometry and thus confirmed that the cleavage was at the Pro15-Ser14 ↓ His10 site.

Expression and Processing of FheproCL1Pro12 Variant—To study the effect of the Leu12 to Pro12 substitution alone, a new variant, FheproCL1Pro12, containing the Pro12 substitution but retaining the active site Cys26 was constructed. A comparison of the activation rates of wild-type FheproCL1 and variant FheproCL1Pro12 monitored at pH 4.5 by continuous fluorogenic assay showed that in comparison with the wild-type
zymogen, the rate of activation of the variant is ~15- to 20-fold slower (Fig. 6A). Autocatalytic activation of the variant FheproCL1Pro\textsuperscript{12} occurred faster at pH 4.5 compared with pH 6.0 (Fig. 6B).

We then assessed whether the FheproCL1Pro\textsuperscript{12} could be trans-activated by mixing it with a 20-fold less quantity of pre-activated FheCL2 at pH 6.0 (close to pH optimum for this enzyme). Analysis by 12% SDS-PAGE shows that over a 60-min period no significant autoactivation of FheproCL1Pro\textsuperscript{12} is observed (Fig. 7A). However, following mixing with FheCL2, the FheproCL1Pro\textsuperscript{12} band at 37 kDa gradually diminishes, and two lower bands appear above the 25-kDa marker that were identical to the sizes observed for autoactivation of FheproCL1 (Fig. 7B). Activation curves of samples removed at time 5 min show that the FheproCL1Pro\textsuperscript{12} alone exhibits low level activity, whereas FheCL2 alone is slightly higher (Fig. 7C). The reaction rate of the mixture of the FheproCL1Pro\textsuperscript{12} and FheCL2 is, however, ~40-fold higher than either of the enzymes alone, indicating that FheCL2 is capable of trans-activating the FheproCL1Pro\textsuperscript{12}. Addition of pre-activated FheCL1 to the FheproCL1Pro\textsuperscript{12} variant did not cause processing or activation (data not shown).

**Homology-based Molecular Modeling of FheproCL1Gly\textsuperscript{26}**—The modeling of FheproCL1Gly\textsuperscript{26} revealed similar structural features to that of human procathepsin L (Fig. 8A). There are minimal interactions between the C-terminal 20 residues of the prosegment and the mature domains of the enzyme, suggesting that the lack of structural stability in this region is consistent in the parasite enzyme. The potential trigger step of disruption of a salt bridge suggested by Vernet et al. (35, 46) within the prosegment in the conserved GXNXFD is maintained as putative interactions between Asp\textsuperscript{36} (papain numbering) and Lys\textsuperscript{77} (this Lys is an Arg 21p in the structure of human procathepsin L) exist at distances of 2.5 and 3.0 Å.

**DISCUSSION**

*F. hepatica* secretes abundant amounts of FheCL1 protease into the tissues of its host where the enzyme functions in tissue degradation and immunosuppression (3, 5, 8). The secreted form is fully processed and activated, and studies in our laboratory suggest that these events occur within the slightly acidic, pH 5.0, environment of the parasite gut lumen following release of the stored zymogen from vesicles of the gut epithelial cells (3, 8). In this study we show that recombinant FheproCL1 can autokatallytically activate in the pH range 7.3 to 4.5; the rate of activation increases with decreasing pH. By extension, we can suggest that autoactivation of the native zymogen to the mature active enzyme can be accomplished in the parasite gut. The broad pH range at which the mature enzyme is active, pH 3.5 to 8.5, not only allows it to perform its role in the digestion of blood and host tissue in the acidic gut but also accommodates its extracorporeal function in the degradation and penetration of host organs and suppression of immune effector cell activity.

The nature of the initial events that transform the inactive zymogen of papain-like cysteine proteases to active enzymes...
Autoprocessing and Activation of F. hepatica Cathepsin L1

still remains a mystery. It has been suggested that this may happen as a result of intramolecular cleavages that remove the N-terminal part of the prosegment and opens up the active site of a small proportion of molecules that then activate otherzymogens by intermolecular cleavages (15, 32, 33). Several studies have indicated an important function for a conserved GXXNFXD motif within the prosegment in these intramolecular activation events (9, 10, 14, 29, 31, 35) (see Fig. 8). Using random mutagenesis studies, Vernet et al. (35) showed that changes within the GXXNFXD motif of papain that disturb the charge state of Asp<sup>−36</sup> prevented proper processing of theprotease in the yeast S. cerevisiae. This motif is situated at a bend in the prosegment after α-helices 1 and 2, which make up the N-terminal domain, and before α-helix 3, which turns the prosegment into the active site cleft (see position of GXXNFXD motif in Fig. 2A). Elucidation of the human procathepsin L three-dimensional structure revealed that Asp<sup>−36</sup> forms a salt bridge with Arg<sup>−77</sup> that is important in maintaining the globular fold of the N-terminal domain of the prosegment (9). Its contribution to the structural integrity of the prosegment adds support to the proposal of Vernet et al. (35) that this residue is involved in the pH-dependent triggering of processing in papain-like proteases; therefore, as the pH decreases, protonation of the negatively charged Asp<sup>−36</sup> causes a conformation change in the motif making it susceptible to cleavage. However, under conditions of neutral pH, the motif lies a distance from the active site (9), and therefore, these pH-induced conformational changes would not only need to cause weakening of the prosegment structure but also bring the GXXNFXD motif in close proximity to the active site Cys<sup>26</sup> to facilitate its cleavage. This is unlikely, as circular dichroism studies have shown that freeprosegments do not undergo significant structural changes during processing (9, 32, 33).

The precise cleavage site within the GXXNFXD motif was not identified by Vernet et al. (35) but was postulated to be Phe-Y <sup>—36</sup> and D motif (albeit the Gly residue is replaced by Ala in cathepsin K) (33) (see Fig. 8). It should also be noted that because B-like proteases do not contain GXXNFXD motifs, and hence, processing events differ with these enzymes and involve cleavages in the C-terminal region of the protease only (25–27).

Nevertheless, in a previous study in our laboratory, we identified a 30-kDa intermediate secreted by P. pastoris transformed with FheproCL1, and N-terminal sequencing showed that it resulted from cleavage of the zymogen after Gly<sup>42</sup> within the GXXNFXD motif (2). Although we postulated that this cleavage may be a product of an intramolecular cleavage, supporting the proposal of Vernet et al. (35) described above, we cannot exclude the possibility that it resulted from random proteolytic degradation of the FheproCL1 as we have since discovered that under the fermentation conditions used the pH of the yeast media drops to pH 6.0. In this study, by maintaining the fermentation pH at 8.0, breakdown of the FheproCL1zymogen was prevented and a single protein of 37 kDa obtained. By using this homogeneous FheproCL1zymogen, our analysis at pH 7.3 revealed that minor intermediates of 30 kDa and above appear in the first moments of FheproCL1 autoprocessing. Because these intermediates were so minor, we could not identify cleavage sites by mass spectrometry. However, it is probable that at least one of these fragments corresponds to the 30-kDa intermediate described by us previously (2). To investigate the role of the GXXNFXD motif in zymogen autoactivation more directly, a mutagenesis approach is being undertaken in our laboratory.

Within 5 min of the activation process at pH 7.3, the >30-kDa components disappear and a lower fragment increases in intensity, which was identified by mass spectrometry as having a molecular mass of 26.351 kDa. This cleavage product was also observed in reactions performed at pH 4.5, where it occurred more rapidly. Our time course studies show that activation of

![Image](68x563 to 299x615)

![Image](68x652 to 299x704)
Autoprocessing and Activation of F. hepatica Cathepsin L1

A.

B.

FheCL1  TYTLGNCFTDRTFEEFKAKYTVEMS-RASDILSHGVYBAN-NRavpdikwer
Papain  SYMLGNCFTDRTFEEFKAKYTVEMS-RASDILSHGVYBAN-NRavpdikwer
Hm CL   SFMTAGNCFTDRTFEEFKAKYTVEMS-RASDILSHGVYBAN-NRavpdikwer
Hm CS   SYDLGNCFTDRTFEEFKAKYTVEMS-RASDILSHGVYBAN-NRavpdikwer
Hm CK   SYELGNCFTDRTFEEFKAKYTVEMS-RASDILSHGVYBAN-NRavpdikwer

FIGURE 8. The prosegment region of FheproCL1 and other papain-like cysteine proteases. A, three-dimensional model of FheproCL1Gly26 is shown to illustrate the relationship of the prosegment (yellow) with the mature enzyme domain (blue) and to highlight the position of GNSXFD motif (purple) and the nonconserved C-terminal region of the prosegment (orange). The position of Leu-12 is indicated and highlighted in red. B, alignment of C-terminal region of the FheproCL1 prosegment region with that of papain, human cathepsin L (Hm CL), human cathepsin S (Hm CS), and human cathepsin K (Hm CK) highlighting identified cleavage sites. Amino acids that are conserved in all sequences are indicated by an asterisk and were used to align the sequences. Less conserved sequences are indicated by single or double dots. Prosegment residues are in capital letters, and those of the N terminus of the mature enzyme are indicated by lowercase letters. Arrowheads indicate cleavage sites identified in the prosegment of FheCL1 (Ref. 2 and this study), papain (35), human cathepsin L (15), human cathepsin S (32), and human cathepsin K (33).

the enzyme correlates with the production of this fragment (cleavage 4 in Table 1) and generates an enzyme capable of not only cleaving small peptide substrates like Z-Phe-Arg-NMHEc but also large proteins such as chicken ovalbumin. Accordingly, these activated enzyme molecules are available to activate other FheproCL1 molecules in a bimolecular process (see below). Thus the enzyme became more active in the course of our 180-min reactions. Enzyme samples taken at time 180 min were ~80 times more active than enzyme taken at time 0; however, a true calibration of activity at each time point could not be obtained as the enzyme activates under the conditions of the assay system.

The 26.351-kDa fragment resulted from a cleavage within a Leu-12-Ser-11↓His-10 sequence situation in the C-terminal region of the propeptide. This cleavage site places Leu-12 in the P2 position, which is a hydrophobic residue highly favored by this enzyme (37, 38). Substitution with a proline placed an unfavorable residue in the P2 position and, hence, prevented the ability of FheCL1 to trans-process FheproCL1Pro-12/Gly-26 at the Pro-12-Ser-11↓His-10 site. This substitution also caused slowing of the processing and activation of the variant FheproCL1Pro-12 by 15–20-fold compared with the wild-typezymogen. However, it did not prevent trans-processing of both of these variants at this site by pre-activated FheCL2, an enzyme that we have shown to accept proline in the P2 position (37, 38). Furthermore, addition of pre-activated FheCL2 to FheproCL1Pro-12 increased its activation rate 40-fold. Collectively, the data pinpoint cleavage of the Leu-12-Ser-11↓His-10 peptide bond as a critical step in the processing and activation of FheproCL1.

Continuous fluorescence assays were used to monitor the autocatalytic activation of thezymogen FheproCL1 and showed that by increasing enzyme concentration the rate of activation increased correspondingly. Plots of the first-order rate constants, $k_{obs}$, as a function of enzyme concentration supported the idea that the activation occurs by a bimolecular process, i.e. as thezymogen is processed and activated, the newly activated molecules cleave otherzymogens. A similar bimolecular process of activation has been described for the human cathepsin L (9). Additional support for a bimolecular process is provided by the results showing that the rate of activation of the FheproCL1Pro-12 variant (a) reaches a similar rate to the wild type if the concentration of enzyme is increased (Fig. 6A) and (b) by addition of pre-activated FheCL2 (Fig. 7C). Furthermore, progress curves carried out at different conditions of pH show that low pH increases the rate of the autocatalytic activation of FheproCL1, a characteristic also observed for other papain-like cysteine proteases (9, 10, 35). Because autoactivation is faster at pH 4.5 than at the pH optimum for activity of the mature FheCL1 enzyme (pH 6.2), lowering of the pH must speed up this process by aiding release of the prosegment rather than activating the enzyme itself. Circular dichroism studies on the human cathepsin Lzymogen show that the secondary structure of the proseg-
ment is unperturbed by lowering the pH; however, the association of the prosegment with the mature domain of the enzyme is weakened, and this may speed up the autoactivation process (9, 32, 33).

The C-terminal portion of the prosegment of papain-like proteases is the region that runs from the active site to the N terminus of the mature protein (9) (see Fig. 2A). A comparison of this region from various papain-like proteases, as shown in Fig. 8, demonstrates the lack of conservation in the properties and number of amino acids that span it and reflects its weak contact with other parts of the molecule and absence of structural constraints (9). The loose structure appears to make it susceptible to proteolysis, both cis and trans, and is a critical feature for both processing and activation of plant, invertebrate, and mammalian cathepsins, including cathepsins L, S, and K (10, 15, 32, 33, 35). For example, within the mammalian cell lysosome, cathepsin L may be activated by auto-catalytic processing or by trans-processing by other enzymes with quite distinct specificities, such as cathepsin D (48). When the C-terminal cleavage sites that have been mapped for various papain-like cysteine proteases are compared (Fig. 8) a common hydrolytic site is not observed, and hence, this region seems to represent a general “protease-receptive” region. However, Menard et al. (15) pointed out that although accessibility to this region is a predominant factor determining processing specificity of cysteine proteasezymogens, on the other hand, the accessible region must contain residues compatible with the enzyme substrate specificity. Thus, cleavage sites within the C-terminal region tend to be after Ser for cathepsin S (32), within one or two amino acid positions from Pro for cathepsin K (33), and contain a hydrophobic amino acid in the P2 position for cathepsin L (15). In experiments with propaya proteinase IV, a protease unable to autocalytically process, the replacement of a Leu residue with a Gly at the prosegment-mature junction created an enzyme with autocalytic ability because the new site overcame the restricted specificity of enzyme for a P1 Gly (49). These observations are consistent with our findings showing that a Leu residue is located at the P2 position of the autocalytic cleavage site Leu-12, Ser-11, and support the contention by Menard et al. (15) that despite the apparent lack of conservation within the C terminus of the pro-segment of papain-like proteases some constraints to proteolytic cleavage in this region do apply.

Cathepsin L proteases are secreted by many parasitic helminths and are associated with the virulence and pathogenesis of infection (50, 51). Thus they are foremost on the list of targets of infection (50, 51). Activation of procathepsin L could begin immediately by activation of a few molecules, possibly by intramolecular cleavages, but would speed up as the pH of the gut decreases, aided by putative protein pumps in the parasite gut wall and protons generated by peptide bond hydrolysis (3). The reaction would further quicken as more active enzyme would be generated by cleavages within the Leu-12, Ser-11, and His-10 sequence of the C-terminal region of the prosegment, and as the concentration of active mature enzyme increases. Final removal of the remaining prosegment residues may be performed by the FheCL1 molecules themselves, as observed in this study, together with di- and amino-exopeptidases that are also secreted by the gut epithelial cells (3, 53). The enzyme secreted by the parasite when it empties its gut contents (every 3 h) would therefore be in a fully processed and activated form, similar to that found in parasite in vitro culture medium (37, 38). Therefore, the complete processing and activation of F. hepatica can occur by a unimolecular mechanism.

Whereas this proposal excludes the need for other proteases in FheproCL1 activation, it is still possible that proteases such as FheCL2, asparaginyl endopeptidase, and the aforementioned exopeptidases that are also secreted by F. hepatica cooperate within the milieu of the parasite gut to make the more rapidly activating and efficient digestive mix. Finally, we have previously hypothesized that helminth asparaginyl endopeptidase was involved in the trans-activation of several other parasite zymogens, including FheproCL1 (3), which have conserved Asn residues at the juncture between prosegment and mature enzyme. In vitro experiments by Sajid et al. (54) with schistosome cathepsin B have demonstrated that this is possible. In this study we show that FheproCL1 can autoactivate, and therefore, the function of the asparaginyl endopeptidase might seem redundant, particularly given that asparaginyl endopeptidases are expressed in very low amounts relative to the FheproCL1 (3, 55). Accordingly, we have modified our hypothesis and suggest that the function of asparaginyl endopeptidases may be to initiate or “kick-start” the activation process by trans-processing and activating the first molecules in the activation cascade of FheCL1 and other parasite zymogens. This hypothesis would obviate the need for rapid intramolecular events to take place in the initial stages of the autocalytic activation of the FheproCL1 in the parasite tissues.

REFERENCES

1. Andrews, S. J. (1999) in Fasciolosis (Dalton, J. P., ed) pp. 1–29, CABI Publishers, Wallingford, Oxon, UK
2. Collins, P. R., Stack, C. N., O’Neill, S. M., Doyle, S., Ryan, T., Brennan, G. P., Mousley, A., Stewart, M., Maule, A. G., Dalton, J. P., and Donnelly, S. (2004) J. Biol. Chem. 279, 17038–17046
3. Dalton, J. P., Skelly, P., and Halton, D. W. (2004) Can. J. Zool. 82, 211–232
4. Brady, M. T., O’Neill, S. M., Dalton, J. P., and Mills, K. H. G. (1999) Infect. Immun. 67, 5372–5378
5. O’Neill, S. M., Mills, K. H., and Dalton, J. P. (2001) Parasite Immunol. (Oxf.) 23, 541–547
6. Dalton, J. P., McGonigle, S., Rolph, T. P., and Andrews, S. J. (1996) Infect. Immun. 64, 5066–5074
7. Mulcahy, G., O’Connor, F., McGonigle, S., Dowd, A., Clery, D. G., Andrews, S. J., and Dalton, J. P. (1998) Vaccine 16, 932–939
8. Dalton, J. P., O’Neill, S. M., Stack, C., Collins, P., Walsh, A., Sekiya, M., Doyle, S., Mulcahy, G., Hmole, E., and Moire, N., Brennan, G., Mousley, A., Kreshchenko, N., Maule, A., and Donnelly, S. (2003) Int. J. Parasitol. 33, 621–640
9. Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., Mort, J. S., and Cygler, M. (1996) EMBO J. 15, 5492–5503
