Cysteine Endopeptidase from *Vigna mungo*: Gene Structure and Expression

Daisuke Yamauchi, Harue Akasofu and Takao Minamikawa

*Department of Biology, Tokyo Metropolitan University, Minami-ohsawa, Hachioji, Tokyo, 192-03 Japan*

A cysteine endopeptidase (SH-EP) is involved in the degradation of storage globulin in cotyledons of germinating seeds of *Vigna mungo*. Using SH-EP cDNA as a probe, we isolated two overlapping genomic clones for SH-EP from a *V. mungo* genomic library. The results of Southern blot analysis of the nuclear DNA suggested that there is a single gene for SH-EP. The site of initiation of transcription was located 31 bp upstream from the initiation codon. Protein immunoblotting showed that, in addition to the presence of the protein in cotyledons of germinating seeds, levels of a protein that reacted with an antiserum against SH-EP increased in pods of maturing fruits of *V. mungo*, with very low levels being present in other organs of the seedlings. The change in levels of SH-EP mRNA in cotyledons of germinating seeds followed a time course similar to that of α-amylase mRNA. Levels of both enzymes, as well as levels of their mRNAs, were also examined in cotyledons of seedlings of three other legumes by protein immunoblotting and RNA blot hybridization, respectively.

**Key words:** Amino acid sequence — α-Amylase — Cysteine endopeptidase — Gene structure — Seed germination — *Vigna mungo*.

When seeds of *Vigna mungo* are germinated, seed storage proteins in cotyledons are mobilized as amino acids for the growing embryonic axes. One of the major hydrolytic enzymes involved in this mobilization is a cysteine endopeptidase, designated SH-EP, that acts to degrade storage globulin in combination with a serine endopeptidase (Mitsuhashi et al. 1986). Protein immunoblot analysis with antiserum raised against SH-EP showed that this enzyme is synthesized in the cotyledons after the onset of imbibition and the level of the enzyme increases until day 4 and decreases thereafter, when germination proceeds at 27°C in the dark (Mitsuhashi and Minamikawa 1989). Analysis of the products of translation in vitro and the results of in vitro processing experiments with enzyme extracts suggested that SH-EP of 33 kDa is synthesized on membrane-bound polysomes as a large, inactive 45-kDa precursor, which is cotranslationally processed to a 43-kDa intermediate through cleavage of a signal peptide, and that this intermediate is processed further to the 33-kDa mature enzyme (Mitsuhashi and Minamikawa 1989). We isolated a cDNA clone for SH-EP from a cDNA expression library of cotyledons of seedlings of *V. mungo* and determined the sequence of its insert (Akasofu et al. 1989). Using the cDNA clone as a probe, we then isolated two overlapping genomic clones for SH-EP from a *V. mungo* genomic library. The nucleotide sequence of the gene for SH-EP has been reported elsewhere (Akasofu et al. 1990). In this report, we describe the genomic organization and site of initiation of transcription of the gene.

In a recent paper (Tanaka et al. 1991), we characterized one of the major endopeptidases, designated EP-C1, in pods of maturing fruits of *Phaseolus vulgaris*. EP-C1, a 34-kDa polypeptide, is thought to participate in the mobilization of proteins in pods during the development of seeds (Endo et al. 1987). This endopeptidase is immunologically homologous to SH-EP, and the deduced amino acid sequence showed a high degree of homology (94%) to that of SH-EP (Tanaka et al. 1991). In the present study, we examined the occurrence of the protein that reacts with the antiserum against SH-EP in pods of maturing fruits of *V. mungo*, and we observed increases in levels of the protein in senescing pods. The protein was detectable at very low levels in organs of the seedlings other than the cotyledons.

SH-EP and α-amylase have similar developmental patterns of expression in cotyledons during germination and
similar patterns of distribution in organs of *V. mungo* (Minamikawa et al. 1992). We demonstrate here that levels of mRNAs for SH-EP and a-amylase show similar temporal changes in cotyledons during germination. In addition, using protein immunoblotting and RNA blt hybridization, we examined the expression of cysteine endopeptidase and a-amylase in cotyledons of seedlings of three other legumes.

**Materials and Methods**

**Plant materials**—Germination of seeds and incubation of cotyledons detached from dry seeds of *Vigna mungo* were conducted in the dark at 27°C (Mitsuhashi and Minamikawa 1989). Seeds of other legumes were germinated under the same conditions. Green leaves were collected from day-6 plants grown in a phytotron maintained at 27°C with 14 h of light and 10 h of darkness. For experiments with pods, plants were grown at the experimental farm of Tokyo Metropolitan University from May to August and, at appropriate developmental stages, maturing pods were harvested and stored at −20°C.

**Preparation of nuclear DNA from *V. mungo* and construction of a genomic library**—Nuclear DNA was isolated from *V. mungo* and a genomic library consisting of 4 × 10⁶ independent recombinant phages was constructed as described previously (Yamauchi and Minamikawa 1990a). The amplified library was screened with [³²P]-labeled cDNA for SH-EP (Akaosfu et al. 1989). Subclones were constructed in pUC118 ( Vieira and Messing 1987) and mp18/19 (Yanisch-Perron et al. 1985) for nucleotide sequencing.

**Preparation of RNA and RNA blot analysis**—Total RNA was prepared as described previously (Mitsuhashi and Minamikawa 1989). The poly(A)⁺-RNA was obtained by fractionation of total RNA on an oligodeoxynucleotide-column (Aviv and Leder 1972). RNA blot analysis was carried out as described elsewhere (Yamauchi et al. 1988) under the same hybridization conditions as those used for Southern blot analysis (Yamauchi et al. 1989).

**S1 nuclease mapping and primer extension**—The S1 nuclease mapping and primer extension were performed as described by Sakurai et al. (1988) using the poly(A)⁺-RNA from day-2 seeds.

**Southern blot analysis**—Southern blot analysis of DNA from *V. mungo* was carried out as described previously (Yamauchi et al. 1989) with [³²P]-labeled cDNA for SH-EP (Akaosfu et al. 1989).

**Assays of enzymatic activity and protein immunoblot analysis**—Endopeptidase was assayed with azoalbumin as substrate essentially as described previously (Endo et al. 1987). The reaction mixture, consisting of 0.3 ml of a solution of enzyme that contained 5 mM 2-ME and 0.1 ml of a 1% solution of azoalbumin (w/v) in 0.1 M sodium acetate buffer (pH 5.4), was incubated in a plastic microtube for 60 min at 30°C. The reaction was terminated by the addition of 0.4 ml of cold 10% trichloroacetic acid. The mixture was left at least for 30 min at 0 to 4°C, then centrifuged at 17,000 × g, and the absorbance at 366 nm of the supernatant solution was recorded against the zero-time control in a spectrometer-data processor (model 150-20; Hitachi, Tokyo). One unit of activity is defined as the amount of enzyme that gives an increase in absorbance of 1.0 unit under the conditions of the assay. The preparation of extracts from cotyledons of *V. mungo* and the protein immunoblot analysis were performed as described elsewhere (Minamikawa et al. 1992). Roots, axes and green leaves were powedered in cold acetone, and the dried powder was homogenized with 0.2 M sodium acetate buffer (pH 5.4) containing 5 mM 2-ME (Endo et al. 1987).

**Results**

**Genomic organization of the gene for SH-EP**—Two overlapping genomic clones (designated pEP1 and pEP3) for SH-EP from a *Vigna mungo* genomic library in λDASH were isolated with the SH-EP cDNA probe (pEP36; Akaosfu et al. 1989) (Fig. 1a). The nucleotide sequence of the gene, together with flanking regions, has been reported elsewhere (Akaosfu et al. 1990). S1 nuclease mapping and primer extension analysis indicated that the site of initiation of transcription is located at the cytosine nucleotide 31 bp upstream from the initiation codon (Fig. 2). The results revealed that the gene is composed of 4 exons and 3 introns.

The genomic DNA was digested with restriction enzymes and probed with pEP36 cDNA (Fig. 1b). Under the given hybridization conditions, we obtained one band of high intensity (2.5 kb) and two of low intensity (1.7 and 4.1 kb) in the case of *Hind* III, and one band of high intensity (2.0 kb) and two of low intensity (3.7 and 4.1 kb) in the case of *EcoR* I. Two small fragments (0.2 kb with *Hind* III and 0.5 kb with *EcoR* I), which were anticipated from the restriction map of pEP1 (Fig. 1a), were not detected. The results suggest that there is a single gene for SH-EP in *V. mungo*. The 4.1-kb fragment (*Hind* III) and the 3.7-kb fragment (*EcoR* I) may be due to the presence of a non-functional gene or a gene for a cysteine endopeptidase other than SH-EP.

**Changes in levels of SH-EP in pods during maturation of fruits**—SH-EP, which is very active in cotyledons of germinating seeds of *V. mungo* (Mitsuhashi et al. 1986), exhibits a high degree of homology in terms of amino acid sequence with EP-C1, an endopeptidase expressed in pods of maturing *Phaseolus vulgaris* fruits (Akaosfu et al. 1989, Tanaka et al. 1991). Therefore, we examined the endopeptidase activity in extracts of pods of maturing fruits of
V. mungo (Fig. 3a). The results indicated that the enzymatic activity, expressed per pod or per g fresh weight of pod, increased during the maturation of fruits, with the highest activity being found in extracts from pods at 20 DAF, despite the progress of senescence. Protein immunoblotting showed that levels of the protein that reacted with the antiserum against SH-EP and had the same molecular mass (33 kDa) as the enzyme found in cotyledons also increased in pods as the maturation of fruits proceeded (Fig. 3a). Although the total endopeptidase activity in extracts from pods was not necessarily attributable exclusively to levels of SH-EP, change in levels of SH-EP corresponded closely to changes in the endopeptidase activity during the maturation of fruits.

Occurrence of SH-EP in various organs of seedlings—Immunoblotting failed to reveal any SH-EP protein in extracts from dry and day-1 cotyledons, but the protein became detectable after day 2 with a maximum level at day 4. These temporal changes in the level of SH-EP are very similar to the temporal changes in its activity (Mitsuhashi and Minamikawa 1989). In the present study, we examined the occurrence of SH-EP in various organs of seedlings. The results showed that, as compared to that in cotyledons, endopeptidase activity was very low in axes and roots of dark-grown seedlings and was even lower in young green leaves (Fig. 3b). Only bands of the 33-kDa polypeptide were detected in the analysis of extracts of cotyledons, roots and green leaves by immunoblotting, but a band corresponding to a 34-kDa polypeptide and a faint band corresponding to a 32-kDa were observed in the analysis of extracts of axes. However, from the present observations we cannot conclude that the 34-kDa and 32-kDa polypeptides detected in extracts of axes were due to the product of translation of SH-EP mRNA, even though these proteins were immunologically homologous to SH-EP.

Levels of SH-EP mRNA in cotyledons of germinating seeds—The RNA blot hybridization with pEP36 cDNA as probe indicated that SH-EP mRNA was barely detectable in cotyledons of dry seeds, and it was first detectable in day-1 cotyledons (Fig. 4). The level of mRNA, expressed per seed or per μg RNA, increased as germination proceeded; it reached a maximum on day 4 and decreased thereafter. When cotyledons detached from dry seeds were allowed to imbibe and were incubated under the same conditions as germinated seeds, only a small increase in the level of SH-EP mRNA was apparent. The results coincided with the observation that the endopeptidase activity and amount of SH-EP remained very low in detached and incubated cotyledons (Mitsuhashi et al. 1986), suggesting that the enzymatic activity is regulated at the mRNA level.

The developmental pattern of levels of SH-EP in cotyledons of germinating seeds is similar to that of α-amylase.

Fig. 1 Restriction maps of the gene for SH-EP and Southern blot analysis of nuclear DNA from Vigna mungo. (a) Restriction maps of genomic clones for the gene for SH-EP (λP1 and λP3) and the genomic subclone (pgEP1). The arrow indicates the direction of transcription. The black bars denote exons in the gene for SH-EP. The thick lines on the map of pgEP1 indicate exons 1 to 4. The restriction enzyme sites indicated are: B, BamH I; E, EcoR I; H, Hind III; S, Sac I; X, Xba I and Xh, Xho I. (b) Southern blot analysis of the gene for SH-EP. Nuclear DNA (16 μg) from axes of seedlings of V. mungo, digested with Hind III (H) or EcoR I (E), was fractionated on a 0.7% agarose gel, blotted onto a nylon filter, and hybridized with [32P]-labeled λEP36 cDNA (Akasofu et al. 1989).
The results of S1 mapping analysis are shown in lanes 1 to 3. The 5'-end-labeled probe (positions -48 to +112) was allowed to hybridize at 37°C with the poly(A)$^+$-RNA from day-2 seedlings of Vigna mungo and digested with S1 nuclease at concentrations of 250, 500, and 1,000 units ml$^{-1}$. The results of primer extension analysis are shown in lane 4. The 5'-end-labeled primer (+65 to +112) was allowed to hybridize at 37°C with poly(A)$^+$-RNA from day-2 seedlings for 16 h and extended by reverse transcriptase. Products of degradation of the probe for S1 mapping analysis are shown in lane 5 (G reaction) and lane 6 (A$>$C reaction) (Maxam and Gilbert 1980). The arrow indicates the site of initiation of transcription.

(Minamikawa et al. 1992). SH-EP and a-amylase are also distributed similarly in various organs. We examined changes in levels of a-amylase mRNA in cotyledons of germinating seeds using the λAMY11 cDNA clone, which contains the complete coding sequence for a-amylase from cotyledons of germinating seeds of V. mungo (Yamauchi and Minamikawa 1990b). a-Amylase mRNA was not detected in dry seeds but its level, expressed per μg RNA or per seed, increased sharply upon imbibition, reached a peak on day 3 and fell thereafter (Fig. 4). These temporal changes were similar to those in levels of SH-EP mRNA, but the changes in levels of a-amylase mRNA preceded the latter by one day. Both SH-EP mRNA and a-amylase mRNA remained at very low levels in embryonic axes of germinating seeds, as compared to the levels in cotyledons, and SH-EP mRNA and a-amylase mRNA became detectable only on day 5 and day 3 post-imbibition, respectively. The low rates of synthesis of SH-EP and a-amylase in the axes were also demonstrated by protein immunoblotting (data not shown).

Cysteine endopeptidase and a-amylase in cotyledons of other legumes—As observed in cotyledons of V. mungo (Mitsuhashi et al. 1986), activities of endopeptidase and a-amylase increased in cotyledons of V. radiata, V. unguiculata and Canavalia gladiata as germination proceeded. Extracts of cotyledons of germinating seeds of these legumes were examined for the presence of cysteine endopeptidase and a-amylase by protein immunoblotting (Fig. 5a). With the antiserum against SH-EP, a polypeptide of the same molecular mass as the 33-kDa SH-EP of V. mungo was observed in extracts from V. radiata. A 34-kDa polypeptide from V. unguiculata and a 30-kDa polypeptide from C. gladiata, both of which were different from SH-EP in terms of molecular mass, were also detect-
Cysteine endopeptidase from Vigna mungo

Fig. 4 Changes in levels of mRNAs for SH-EP and α-amylase during germination of seeds of Vigna mungo. Relative amounts of mRNAs from cotyledons were estimated by scanning the autoradiographs with a densitometer and are expressed per μg of RNA (a) and per seed (b). —, SH-EP mRNA in cotyledons of germinating seeds; ——, SH-EP mRNA in detached cotyledons; ---, α-amylase mRNA in cotyledons of germinating seeds.

Discussion

A conspicuous increase in the level of SH-EP occurs in cotyledons of germinating seeds of Vigna mungo and the enzyme is responsible for the degradation of seed storage globulin, acting cooperatively with serine endopeptidase (Mitsuhashi et al. 1986, Mitsuhashi and Minamikawa 1989). The results of RNA blot hybridization indicated that the synthesis of SH-EP in cotyledons during germination is regulated at the mRNA level (Fig. 4). SH-EP was present only at low levels in axes, roots and green leaves as compared to levels in cotyledons (Fig. 3b). SH-EP exhibits a high degree of homology to EP-C1, an endopeptidase in pods of maturing fruits of Phaseolus vulgaris (cf. Fig. 6). Therefore, we examined the occurrence of SH-EP in pods of maturing fruits of V. mungo. We found a pronounced increase in levels of SH-EP in senescing pods (Fig. 3a). These observations suggested that the gene for SH-EP is expressed mainly in cotyledons of germinating seeds and senescing pods, in both of which storage proteins are actively mobilized. From the results of genomic DNA hybridization (Fig. 1b), it appears that there is only a single gene that encodes SH-EP in V. mungo, in contrast to the case of rice and barley, in which there are multiple genes for the cysteine endopeptidases that constitute rather large gene families (Watanabe et al. 1991, Koehler and Ho 1990). However, the possibility that genes for other cysteine endopeptidases are present in the genome of V. mungo cannot be excluded at present.

Temporal change in levels of SH-EP in cotyle-
Table 1 Homologies between SH-EP and other cysteine proteases in terms of amino acid sequence

<table>
<thead>
<tr>
<th>Protease</th>
<th>Signal peptide region (%)</th>
<th>Propeptide region (%)</th>
<th>Mature enzyme region (%)</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP-C1</td>
<td>90</td>
<td>90</td>
<td>96</td>
<td>94</td>
</tr>
<tr>
<td>EP-B</td>
<td>37</td>
<td>45</td>
<td>60</td>
<td>54</td>
</tr>
<tr>
<td>Alerain</td>
<td>36</td>
<td>24</td>
<td>40</td>
<td>34</td>
</tr>
<tr>
<td>Oryzain\ a</td>
<td>28</td>
<td>27</td>
<td>56</td>
<td>37</td>
</tr>
<tr>
<td>Oryzain\ y</td>
<td>29</td>
<td>24</td>
<td>39</td>
<td>34</td>
</tr>
<tr>
<td>Actinidin</td>
<td>25</td>
<td>30</td>
<td>50</td>
<td>42</td>
</tr>
<tr>
<td>Papain</td>
<td>37</td>
<td>32</td>
<td>42</td>
<td>39</td>
</tr>
<tr>
<td>Cathepsin H</td>
<td>30</td>
<td>23</td>
<td>41</td>
<td>35</td>
</tr>
<tr>
<td>Cathepsin L</td>
<td>36</td>
<td>22</td>
<td>45</td>
<td>38</td>
</tr>
</tbody>
</table>

Homologies were calculated from the sequence alignments in Figure 6. For references, see the legend to Figure 6.

dons of germinating seeds of \textit{V. mungo} (Mitsuhashi and Minamikawa 1989) and the distribution of the enzyme in various organs (Fig. 3) were similar to those of \textit{a}-amylose (Minamikawa et al. 1992), and temporal changes in levels of mRNAs for SH-EP and \textit{a}-amylose in cotyledons during germination were also found to be very similar (Fig. 4). These results suggest the presence of common factors in the control of the expression of both genes during the life cycle of \textit{V. mungo}.

Protein immunoblotting showed that there are cysteine endopeptidases that differ from one another in terms of molecular mass even among \textit{Vigna} species (Fig. 5a), and this variation in endopeptidases among species seem to be larger than that in \textit{a}-amyloses. Results of RNA blot hybridization analysis suggested differences in terms of nucleotide sequences between the cysteine endopeptidases from different species of legume.

In Figure 6, the amino acid sequence of SH-EP, as deduced from the nucleotide sequence of the cDNA, is compared with those of other cysteine proteases reported to date. SH-EP exhibits from 34\% to 54\% homology to the other proteases, which include cathepsins H and L from rat, and an exceptionally high degree of homology (94\%) to EP-C1 is apparent (Table 1). Cys-152, which corresponds to the residue that is directly involved in the catalytic sites of papain and actinidin (Kamphuis et al. 1985), and sequences in the vicinity of this residue are well conserved among the enzymes cited. When the sequences of the mature enzymes are compared, SH-EP exhibits still higher homology (39\% to 60\%) to the other enzymes, with 96\% homology to EP-C1. Thus, the signal peptide and the propeptide regions of SH-EP exhibit relatively low homology to those of the other enzymes. The propeptide region of SH-EP exhibits higher homology (45\%) to that of EP-B than to those of the other proteins (22\% to 32\%), with the exception of EP-C1. Since both SH-EP (Mitsuhashi and Minamikawa 1989) and EP-B (Koehler and Ho 1990) are known to be synthesized via multiple post-translational processing events, we can postulate that the propeptide regions of both enzymes include similar or common amino acid sequences that are responsible for their processing and intracellular transport.

James and Sielecki (1986) studied the mechanism of activation of porcine pepsinogen, the zymogen of an aspartic proteinase, and observed that, in the zymogen molecule, the positively charged proenzyme region is bound to the pepisin region through ionic interactions that result in the masking of catalytic and substrate-binding sites. Therefore, we examined the relative levels of charged amino acid residues in the propeptide and mature enzyme regions of...
SH-EP and other cysteine endopeptidases (Table 2). The propeptide regions of these enzymes obviously contain a higher proportion of charged amino-acid residues than the mature enzyme regions. In particular, propeptide regions of SH-EP, EP-C1 and EP-B include relatively high proportions of positively charged residues. Thus, it seems likely that the propeptide regions of these three enzymes share common functions, such as the masking of active sites in the region that corresponds to the mature enzyme.

SH-EP and EP-C1 have a carboxyterminal sequence of Lys-Asp-Glu-Leu (KDEL) (Fig. 6). In mammalian and yeast cells, a KDEL or HDEL sequence at the carboxyterminus functions as a signal for the retention of proteins in the ER and, probably, for the return of proteins from other compartments to the ER (Warren 1987, Pelham 1989). In higher plants, the carboxyterminal sequence KDEL has been found in the ER-associated auxin-binding protein of maize (Inohata et al. 1989) and in other proteins (Chrispeels 1991). Herman et al. (1990) modified the gene of the vacuole protein phytohemagglutinin-L so that the amino acid sequence would end in KDEL, and they engineered the expression of the modified gene in transgenic tobacco under the control of a seed-specific promoter. They found that a substantial portion of phytohemagglutinin-KDEL was associated with the ER and the nuclear envelope. The roles of the carboxyterminal KDEL sequence in SH-EP and EP-C1, both of which should be transferred to and accumulated in vacuoles, remain to be characterized. Putatively glycosylated Asn residues are found at three sites in the carboxyterminal sequences of SH-EP and EP-C1. These carboxyterminal regions of SH-EP and EP-C1 may also function in targeting the proenzymes to specific intracellular compartments since the glycosylated carboxyterminal propeptide of barley lectin is known to be a determinant in the sorting of vacuolar proteins (Bednarek and Raikhel 1991).

This work was supported in part by grants from the Ministry of Education, Science and Culture, Japan.

Table 2 Number of charged amino acid residues, as a percentage of the total, in cysteine propeptidases

<table>
<thead>
<tr>
<th>Protease</th>
<th>Propeptide region</th>
<th>Mature enzyme region</th>
</tr>
</thead>
<tbody>
<tr>
<td>SH-EP</td>
<td>+ 26 12</td>
<td>+ 9 14</td>
</tr>
<tr>
<td>EP-C1</td>
<td>+ 26 11</td>
<td>+ 9 14</td>
</tr>
<tr>
<td>EP-B</td>
<td>+ 25 14</td>
<td>+ 9 12</td>
</tr>
<tr>
<td>Alearain</td>
<td>+ 17 10</td>
<td>+ 8 11</td>
</tr>
<tr>
<td>Oryzain a</td>
<td>+ 19 21</td>
<td>+ 10 14</td>
</tr>
<tr>
<td>Oryzain γ</td>
<td>+ 20 11</td>
<td>+ 8 11</td>
</tr>
<tr>
<td>Actinidin</td>
<td>+ 17 15</td>
<td>+ 6 10</td>
</tr>
<tr>
<td>Papain</td>
<td>+ 11 18</td>
<td>+ 11 7</td>
</tr>
<tr>
<td>Cathepsin H</td>
<td>+ 21 7</td>
<td>+ 9 8</td>
</tr>
<tr>
<td>Cathepsin L</td>
<td>+ 20 12</td>
<td>+ 11 12</td>
</tr>
</tbody>
</table>

+, Positively charged residues; -, negatively charged residues. For references, see the legend to Figure 6.

References


Cysteine endopeptidase from Vigna mungo

797


(Received March 13, 1992; Accepted June 28, 1992)