Two Distinct Cystatin Species in Rice Seeds with Different Specificities against Cysteine Proteinases

MOLECULAR CLONING, EXPRESSION, AND BIOCHEMICAL STUDIES ON ORYZACYSTATIN-II*

Hiroto Kondo, Keiko Abe, Ikuko Nishimura, Hirohito Watanabe, Yasufumi Emori†, and Soichi Araij
From the Department of Agricultural Chemistry, The University of Tokyo, Yayoi-cho, Bunkyo-ku and the Department of Molecular Biology, The Tokyo Metropolitan Institute of Medical Science, Hinkomagome, Bunkyo-ku, Tokyo 113, Japan

Oryzacystatin (oryzacystatin-I) is a proteinaceous cysteine proteinase inhibitor (cystatin) in rice seeds and is the first well defined cystatin of plant origin. In this study we isolated cDNA clones for a new type of cystatin (oryzacystatin-II) in rice seeds by screening with the oryzacystatin-I cDNA probe. The newly isolated cDNA clone encodes 107 amino acid residues whose sequence is similar to that of oryzacystatin-I (approximately 55% of identity). These oryzacystatins have no disulfide bonds, and so could be classified as family-I cystatins; however, the amino acid sequences resemble those of family-II members more than family-I members. Oryzacystatin-I and -II are remarkably distinct in two respects: 1) their specificities against cysteine proteinases; and 2) the expression patterns of their mRNAs in the ripening stage of rice seeds. Oryzacystatin-I inhibits papain more effectively (\(K_i 3.0 \times 10^{-8} \text{ M}\)) than cathepsin H (\(K_i 0.79 \times 10^{-8} \text{ M}\)), whereas oryzacystatin-II inhibits cathepsin H (\(K_i 1.0 \times 10^{-8} \text{ M}\)) better than papain (\(K_i 0.83 \times 10^{-8} \text{ M}\)). The mRNA for oryzacystatin-I is expressed maximally at 2 weeks after flowering and is not detected in mature seeds, whereas the mRNA for oryzacystatin-II is constantly expressed throughout the maturation stages and is clearly detected in mature seeds. Western blotting analysis using antibody to oryzacystatin-II showed that, as is the case with oryzacystatin-I, oryzacystatin-II occurs in mature rice seeds. Thus, these two oryzacystatin species are believed to be involved in the regulation of proteolysis caused by different proteinases.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) J05506.

§ To whom correspondence should be addressed.

1 In this paper, the previously reported oryzacystatin is renamed oryzacystatin-I (type I oryzacystatin) and the newly isolated oryzacystatin is designated as oryzacystatin-II (type II oryzacystatin).

2 The abbreviations used are: OC-I, oryzacystatin-I; OC-II, oryzacystatin-II; 2,2-bis(hydroxymethyl)-2,2-nitrotritriethanol; BANA, N-benzoyl-DL-arginine-2-naphthylamide; Leu-NA, L-leucine-2-naphthylamide; PAGE, polyacrylamide gel electrophoresis.

of Abe et al. (25) which encoded a fusion protein containing 27 cDNA sequence and was designated "pNOC1-1." The expression open reading frame (aa. l-107, OC-II numbering) deduced from the residues which shows papain-inhibitory activity (25).

Construction of a cdNA Library—Total RNA was extracted by the phenol-SDS method (17) from rice seeds harvested at 2 weeks after flowering, and poly(A)+ RNA was purified by oligo(dT)-cellulose column chromatography (18). After the synthesis of single-stranded cdDNA by reverse transcriptase with total poly(A)+ RNA as a template, double-stranded cdDNA was synthesized essentially according to the method of Gubler and Hoffman (19). Flush ends of the cdDNA were generated with T4 DNA polymerase, and the double-stranded cdDNA was treated with EcoRI methylase. Phosphorylated EcoRI linker (12-mer) was ligated, digested with EcoRI, and inserted into the phage vector Agt10. The DNA was then packaged into bacteriophage particles using the packaging extract Gigapack-Gold (Stratagene) and grown on E. coli C600H11, with a yield of 5.0 x 10^8 plaques/µg of mRNA.

Isolation of cdNA Clones for Oryzacystatins—Recombinant plasmids were transferred onto Nylon filters (Hybond N, Amersham Corp.). The filters were fixed with ultraviolet light and prehybridized for 20 h in 50 mM Tris-HCl, pH 5.0, containing 1 M NaCl, 10 mM EDTA, 0.2% Ficoll 4000, 0.1% SDS, 200 ng/ml heat-denatured E. coli cdNA, and 200 µg/ml salmon sperm DNA at 65°C. After prehybridization, the filters were hybridized with 32P-labeled OC-I cdNA (3) at 65°C in the same solution used for prehybridization. The filters were finally washed in 3 x SSC, 0.1% SDS at 65°C and then exposed.

Results

Detection and Partial Purification of OC-II from Rice Seeds—Unhulled rice seeds (4 kg) were homogenized with 5 liters of 25 mM sodium phosphate buffer, pH 7.0, containing 0.15 M NaCl. The homogenate was centrifuged at 20000 x g for 30 min and the supernatant was immediately heated at 80°C for 10 min and centrifuged again to obtain a heat-stable extract. The supernatant was saturated with solid ammonium sulfate at 20-80%. The precipitate (20-80% saturation) was suspended in 400 ml of 50 mM sodium acetate buffer, pH 4.9 (buffer A) and dialyzed against buffer A. The dialysate was applied to a C5 Sephadex C-50 column (1.5 x 80 cm). Fractions (eluting at 0.2 M NaCl) were collected, pooled, and dialyzed against 25 mM sodium phosphate buffer, pH 7.5, and applied to a Butyl-Toyopearl 650 column (25 mM sodium phosphate buffer, pH 7.5, containing 30% ammonium sulfate (buffer C)). The column was washed with buffer C containing 0.1 M ammonium sulfate decreasing to 0%. Active fractions (eluting at 7% ammonium sulfate) were pooled and concentrated.

Assay of Inhibitory Activity and Determination of K Values—Papain and cathepsin H activities were measured by the method of Barrett (20). Papain and cathepsin H were equilibrated at a final concentration of 4 nM in 100 mM sodium phosphate buffer, pH 6.5, containing 3 mM dithiothreitol and 2 mM EDTA. The solutions were preincubated at 37°C for 10 min in the presence of various concentrations of expressed OC-I or OC-II, prior to addition of the appropriate substrate, BANA or Leu-NA, at a final concentration of 100 µM. Inhibitory activity was represented as the percent inhibition of the BANA-hydrolyzing or Leu-NA-hydrolyzing activity. K, values were obtained from Dixon plots (27).

RESULTS

Isolation and Identification of a Second Type of Oryzacystatin cdNA Clones—From 1.2 x 10^9 independent cdDNA clones of a rice seed cdNA library constructed previously (3), eight positive clones were isolated by hybridization using the 3'P-labeled OC-I cdNA fragment (3) as a probe. Seven of these positive clones had strong signals on x-ray films and were considered to be cdDNA clones encoding OC-I. One of the clones designated Amc1A, had a weak signal on an x-ray film, making it seem likely to encode a similar, but somewhat different, protein. This clone was then subjected to nucleotide-sequence analysis and proved to have a deduced amino acid sequence (31 residues) similar to that of OC-I, including the

Purification of OC-I and OC-II Expressed in E. coli—E. coli strain YA21 was used for expression. E. coli YA21 cells carrying recombinant plasmids (pOC26-5'-1, pNOC1-1) were cultured for 18 h in χ-broth. Synthesis of the encoded proteins was then induced with isopropyl-β-D-thiogalactopyranoside in M9 medium as described previously (25). After cultivation, the cells were harvested by centrifugation and disrupted by a French press, followed by centrifugation at 10000 x g. The supernatant was dialyzed against 0.6 M NaCl. The fractions containing OC-I or OC-II eluted at a concentration of 0.25 M NaCl and were collected. These fractions were then applied to a MonoQ column (fast protein liquid chromatography system) equilibrated with 20 mM bis-tris, pH 7.1, and eluted with a linear gradient to 0.5 M NaCl. The fractions eluted at a concentration of 0.12 M NaCl were collected.

Preparation of Antibodies to OC-II—Purified OC-II protein (200 µg) expressed in E. coli was mixed with an equal volume of Freund’s complete adjuvant and injected into a female rabbit (New Zealand White). After 2 weeks, the same amount of OC-II in Freund’s complete adjuvant was again injected (booster). Immunity was then confirmed by the Ouchterlony double-diffusion method. This antibody preparation had no cross-reaction with OC-I expressed in E. coli.

Downloaded from http://www.jbc.org/ by guest on October 22, 2017
Gln-Val-Val-Gly-Gly sequence that corresponds to the central conserved sequence among animal cystatins (28). Therefore, λOCP was considered to contain a partial cDNA fragment encoding a second type of oryzacystatin. By screening a newly constructed cDNA library of 2.1 × 10^6 independent plaques with the cDNA insert of λOCP as a probe, two positive clones were isolated. Judging from restriction mapping analysis, these clones were derived from a single mRNA species and overlapped with each other. One of these clones, designated λOCl, was subjected to further analysis to determine its sequence according to the strategy shown in Fig. 1. The complete nucleotide sequence of λOCl (652 base pairs) was determined and is shown, along with the deduced amino acid sequence, in Fig. 2. The nucleotide sequence contained a poly(A) tail of 18 base pairs, a poly(A) addition signal, and its conserved sequence among animal cystatins (28). Therefore, the putative initiation codon (ATG) was assigned based on the following reasoning. 1) The length of the expected mRNA (approximately 700 bases; Fig. 5) corresponds roughly to the total length of the cDNA insert. 2) Both the molecular weight (11,900) and the deduced amino acid sequence of the encoded protein resemble those of cystatins, especially OC-I (3). The molecular weight (12,000) of the immunoreactive protein in the OC-II-containing fraction of the rice seed extract coincides with the deduced value (11,900) as described below (Fig. 6).

**Comparison of the Nucleotide and Amino Acid Sequences of OC-I and OC-II**—Comparison of the nucleotide sequences of OC-I (3) and OC-II shows 59% identity. In particular, the nucleotide sequences of the coding regions are highly homologous, showing approximately 70% identity. However, little sequence homology is observed in the 5’ and 3’ noncoding regions. On the other hand, an overall comparison of the deduced amino acid sequences shows 55% identity (Fig. 3). In the NH2-terminal regions (residues 1–20), OC-II numbering, little sequence homology is observed and the residue numbers differ between OC-I and OC-II. In the rest of the region, however, high sequence homology (78%) is observed. It should also be noted that the commonly conserved sequence Gln-Val-Val-Ala-Gly, in the central part of the molecule of most cystatin superfamily members including OC-I, is also present in OC-II (Glu-Val-Val-Gly-Gly). OC-I numbering, respectively.

**Comparison of OC-II and Other Animal Cystatins**—Comparison of the amino acid sequence of OC-II with those of all cystatin-I and oryzacystatin-II. Identical amino acid residues are boxed. Altogether, OC-II resembles family II more closely than family I cystatins (approximately 22%).

**Distinct Inhibitory Activities of OC-I and OC-II**—OC-I and OC-II were expressed in E. coli and purified as described under "Experimental Procedures." The resultant supernatant gave main bands on SDS-PAGE (Fig. 4) corresponding to oryzacystatin-I and oryzacystatin-II. Identical amino acid residues are boxed.
Molecular Cloning and Expression of Oryzacystatin-II

Fig. 4. SDS-polyacrylamide gel electrophoresis of OC-I and OC-II expressed in E. coli. Samples were prepared as described under “Experimental Procedures.” Samples applied were as follows: crude extract supernatant (5 μg) of OC-I (lane 2), the final purified preparation (3.5 μg) of OC-I (lane 4), the final purified preparation (3.5 μg) of OC-II (lane 5). A homogenate of pUC18 cells (6 μg) was electrophoresed as a control (lane 1), and size markers (lane 6) were also included. Positions of marker proteins are shown on the right.

Fig. 5. RNA blot analysis. Ten μg each of total RNA from rice seeds harvested 1 week (lane 1), 2 weeks (lane 2), 3 weeks (lane 3), 4 weeks (lane 4) and 10 weeks, i.e. at maturation (lane 5), after flowering were electrophoresed, blotted onto a nitrocellulose membrane, and hybridized with 32P-labeled OC-I cDNA fragment (3) (A) or 32P-labeled OC-II cDNA fragment (B). The positions of 18 S and 28 S ribosomal RNA markers are shown at the center.

Fig. 6. Separation of oryzacystatin-II fraction on Butyl-Toyopearl 650. The sample treated as described under “Experimental Procedures” was loaded onto a column of Butyl-Toyopearl 650 (1.2 cm diameter; 20 ml bed volume) pre-equilibrated with 25 mM sodium phosphate buffer, pH 7.5, containing 30% ammonium sulfate. The column was washed with 40 ml of the buffer and the protein was eluted with a 360-ml linear gradient of 30 to 0% ammonium sulfate. Fractions of 3 ml were collected. Fraction A, eluted by 12% ammonium sulfate, was the oryzacystatin-I fraction showing papain-inhibitory activity. Fraction B, eluted by 7% ammonium sulfate, was the oryzacystatin-II fraction showing cathepsin-H-inhibitory activity.

TABLE I

Inhibitory activities of oryzacystatins

<table>
<thead>
<tr>
<th></th>
<th>K&lt;sub&gt;i&lt;/sub&gt; values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Oryzacystatin-I</td>
</tr>
<tr>
<td>Papain</td>
<td>3.02 × 10&lt;sup&gt;-8&lt;/sup&gt;</td>
</tr>
<tr>
<td>Cathepsin H&lt;sup&gt;+&lt;/sup&gt;</td>
<td>0.79 × 10&lt;sup&gt;-4&lt;/sup&gt;</td>
</tr>
<tr>
<td>Cathepsin H&lt;sup&gt;+&lt;/sup&gt;</td>
<td>1.29 × 10&lt;sup&gt;-4&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> BANA as substrate.
<sup>b</sup> Leu-NA as substrate.

fusion proteins of a 17-kDa band for OC-I containing the additional NH₂-terminal sequence (25) and a 12-kDa band for OC-II. The purified preparations of OC-I and OC-II gave single bands. OC-I and OC-II thus expressed and purified were evaluated for their inhibitory activities against papain and cathepsin H. OC-I exhibited papain-inhibitory activity as described previously (25). A crude extract containing OC-II and purified OC-II showed significant cathepsin H-inhibitory activities, whereas the crude extract from a control experiment containing pUC18 showed no inhibitory activity nor any protease activity that might affect the activity assay. OC-II, thus, has the characteristics of a cysteine proteinase inhibitor and was confirmed to be a cystatin. We next determined the K<sub>i</sub> values for OC-I and OC-II against papain and cathepsin H. The K<sub>i</sub> value for OC-I versus papain was 3.02 × 10<sup>-8</sup> M whereas the K<sub>i</sub> values versus cathepsin H were higher, 0.79 × 10<sup>-4</sup> M and 1.29 × 10<sup>-4</sup> M, when BANA and Leu-NA were used as substrates, respectively (Table I). On the other hand, the K<sub>i</sub> value for papain inhibition by OC-II was 0.83 × 10<sup>-6</sup> M and those for inhibition of the BANA and Leu-NA hydrolyzing activities of cathepsin H were 1.00 × 10<sup>-4</sup> M and 1.84 × 10<sup>-4</sup> M (Table I), respectively. Neither OC-I and OC-II showed any effective inhibition against cathepsin B or L (data not shown). Thus, OC-II has a different inhibitory action from OC-I; that is, OC-I is a papain-inhibitory type of cystatin while OC-II is a cathepsin H-inhibitory type.

Continuous Expression of the OC-II mRNA throughout Rice Seed Maturation—OC-II cDNA was used as a probe for Northern analysis to detect OC-II mRNA. As shown in Fig. 5, a single band of ~700 bases was detected in rice seeds throughout seed maturation (from 1 to 10 weeks after flowering). The expression pattern of the OC-II mRNA was different from that of OC-I. The OC-I mRNA was expressed distinctly at 2 weeks after flowering and not detected both at 1 week after flowering and in mature seeds (Fig. 5A). As opposed to this pattern, the OC-II mRNA was expressed distinctly at 1 week after flowering and expression continued up to complete maturation (Fig. 5B), although the amounts are very low.

Detection and Partial Purification of OC-II from Rice Seeds—We tried to detect and purify OC-II from rice seeds by isolating the cathepsin H-inhibitory fraction as described under “Experimental Procedures” to verify the existence of OC-II in the seeds. Butyl-Toyopearl 650 chromatography yielded two cystatin fractions (fractions A and B) showing strong papain- and cathepsin H-inhibitory activities, respectively (Fig. 6). Fig. 7 shows the results of SDS-PAGE of OC-II purified from mature rice seeds and Western blotting analysis using the antiserum to OC-II. Western blotting analysis (Fig. 7B) revealed that the fraction showing cathepsin H-inhibitory activity (fraction B) exhibits an immuno-
ern blotting analysis of oryzacystatin-II. A, profile of SDS-blotting analysis stained with antibody to OC-II. Samples prepared as described under “Experimental Procedures” were as follows: supernatant of crude extract (8 μg) of rice seeds (lanes 1 and 7), fractions (8 μg) after treatment with CM-Sephadex (lanes 2 and 8), solution (8 μg) after treatment with Sephades G-75 (lanes 3 and 9), the partially purified OC-II fraction (2 μg) after treatment with Butyl-Toyopearl 650, the same as fraction B in Fig. 6 (lanes 4 and 10). OC-II expressed in E. coli and size markers were included (lanes 5 and 6, respectively).

reactive band, and OC-II exists in rice seeds as a protein with a molecular weight of approximately 12,000, consistent with the estimate from the deduced amino acid sequence (Fig. 2). In Fig. 7A, major bands gather around M, 12,000 in lanes 1, 2, and 3, because in rice seeds there are many kinds of proteins whose molecular weights are approximately 12,000. The band due to OC-II itself is not so dense and may be almost the same as that in lane 4, judging from the Western blotting analysis (Fig. 7B).

DISCUSSION

In this study we isolated cDNA clones encoding a second type of oryzacystatin and named it oryzacystatin-II (OC-II). OC-II is composed of 107 amino acid residues and has a sequence similar to those of other cystatins, especially to that of OC-I. We established an expression system for OC-II in E. coli and used it to study the biochemical characteristics of OC-II.

Northern hybridization analysis showed that the OC-II mRNA is approximately 700 bases in length (Fig. 5), and we concluded that the OC-II cDNA represents a nearly full length copy. The length of the OC-II mRNA was somewhat longer than that of OC-I (Fig. 5). The expression patterns of mRNA in seeds during maturation differ between OC-I and OC-II. The OC-I mRNA is expressed predominantly 2 weeks after flowering. OC-II mRNA, on the other hand, is expressed throughout seed maturation, similar to the expression pattern of glutelin mRNAs (3). Analyses of gene structure and function will be necessary to elucidate the mechanism of cooperative expression of these proteins. OC-I exists in rice seeds and has been purified for investigation of its properties (1-3), but no evidence for the occurrence of OC-II has yet been provided from biochemical studies. Here we confirm that OC-II, with cathepsin H-inhibitory activity, really exists in rice seeds as an approximately 12-kDa protein (Figs. 6 and 7). Thus, it is presumed that the multiple cystatin system works in a rice seed as in animal tissues.

OC-II, as well as OC-I, have no disulfide bonds, and can be classified as family I members of the cystatin superfamily. With respect to amino acid sequence, however, OC-I and OC-II more closely resemble family II cystatins. Taking these points into consideration, we propose that, besides the conventional three cystatin families that comprise animal cystatins, a new category, such as phytocystatins, should be organized to include OC-I and OC-II. In any event, it is likely that phytocystatins diverged from a cognate ancestral cystatin before the divergence of the three animal cystatin families.

Several hypotheses concerning the active site of cystatins have been proposed. According to a scheme proposed for the interaction between cystatin and papain (28, 29), the NH$_2$-terminal region, including a Gly residue, a Glu-Val-Val-Ala-Gly sequence or its derivatives, and a Pro-Trp sequence are thought to play important roles in the interaction. Sequence comparison of OC-I and OC-II shows that the Glu-Val-Val-Ala-Gly and Pro-Trp sequences in OC-I are changed, respectively, to Glu-Val-Val-Ser-Gly and Ala-Trp sequences in OC-II (Figs. 2 and 3). According to the x-ray structural analysis of chicken cystatin (29), β-turn-like structures exist at both the Glu-Leu-Val-Ser-Gly and Pro-Trp sequences. It is likely that similar turn structures are conserved in OC-I and OC-II. However, somewhat different partial structures may be formed due to differences in these key sequences. In fact, OC-I and OC-II have mutually different specificities toward target cysteine proteinases; the former inhibits papain much more effectively than cathepsin H, whereas the reverse is true for the latter (Table I). This difference in specificity is considered to arise from the structural differences discussed above. Low sequence similarity in NH$_2$-terminal regions of OC-I and OC-II (Fig. 3) may also contribute to their different specificities toward target enzymes.

Knowledge of the structure and inhibitory action of this new oryzacystatin in rice seeds will define a major mechanism involving the proteolytic regulation of rice storage proteins. Known animal cystatins, for example kininogen (30), cystatin S (31), rat muscle cystatin (32), and human amniotic fluid cystatin (33), do not show significant differences in their specificities against papain and cathepsin H. On the other hand, rat cystatins α and β have different inhibitory activities against cathepsins B, H, and L. This may be important physiologically because there are several cystatins that specifically inhibit different target proteinases in animal organs and tissues. It is very interesting that OC-I and OC-II differentially between two cysteine proteinases, papain and cathepsin H. These two cystatins, possibly along with cysteine proteinases, might play different roles in controlling the accumulation of storage proteins synthesized in rice seeds during the period from flowering to ripening and their degradation during germination. In fact, we have found a cysteine proteinase in rice seeds that efficiently hydrolyzes glutelin and is almost completely inhibited in vitro by an equimolar concentration of OC-I (34), although the effect of OC-II remains to be investigated. Moreover, our recent work has revealed that two types of cysteine proteinases, like papain and cathepsin H occur in rice seeds. The reason two types of cystatins, OC-I and OC-II, occur in rice seeds is probably that they have their own target enzymes, as discussed above, or that their localizations are different. These problems will be explored using immunologic and genetic methods.

Acknowledgment—We thank Dr. Meera Rhee of the Department of Agricultural Chemistry, University of Tokyo, for helping with the purification and preparation of cathepsins B, H, and L.

REFERENCES
Two distinct cystatin species in rice seeds with different specificities against cysteine proteinases. Molecular cloning, expression, and biochemical studies on oryzacystatin-II.

H Kondo, K Abe, I Nishimura, H Watanabe, Y Emori and S Arai


Access the most updated version of this article at http://www.jbc.org/content/265/26/15832

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/265/26/15832.full.html#ref-list-1