ORIGINAL ARTICLE

Hideki Kato · Keita Sutoh · Takao Minamikawa

Identification, cDNA cloning and possible roles of seed-specific rice asparaginyl endopeptidase, REP-2

Abstract We previously showed that two major cysteine endopeptidases, REP-1 and REP-2, were present in germinated rice (*Oryza sativa* L.) seeds, and that REP-1 was the enzyme that digests seed storage proteins. The present study shows that REP-2 is an asparaginyl endopeptidase that acts as an activator of REP-1, and we separated it into two forms, REP-2α (39 kDa) and REP-2β (40 kDa), using ion-exchange chromatography and gel filtration chromatography. Although analysis of the amino terminals revealed that 10 amino acids of both forms were identical, their isoelectric points were different. SDS–PAGE/immunoblot analysis using an antiserum raised against legumain, an asparaginyl endopeptidase from jack bean, indicated that both forms were present in maturing and germinating rice seeds, and that their amounts transiently decreased in dry seeds. Northern blot analysis indicated that REP-2 mRNA was expressed in both maturing and germinating seeds. In germinating seeds, the mRNA was detected in aleurone layers but not in shoot and root tissues. Incubation of the de-embryonated seeds in 10^{-6} M gibberellic acid induced the production of large amounts of REP-1, whereas REP-2β levels declined rapidly. Southern blot analysis showed that there is one gene for REP-2 in the genome, indicating that both REP-2 enzymes are generated from a single gene. The structure of the gene was similar to that of β-VPE and γ-VPE isolated from *Arabidopsis thaliana*.

Keywords Asparaginyl endopeptidase · Cysteine endopeptidase · Development · Germination · *Oryza sativa*

H. Kato · K. Sutoh (✉) · T. Minamikawa
Department of Biological Sciences,
Tokyo Metropolitan University,
Minami-ohsawa 1-1,
Hachioji-shi,
192-0397 Tokyo, Japan
E-mail: s-keita@comp.metro-u.ac.jp
Fax: +81-426-772559

Abbreviations ABA: abscisic acid · Asn: asparagine · DAF: days after flowering · DAI: days after imbibition · GA$_3$: gibberellic acid · VPE: vacuolar processing enzyme

Introduction

Seed storage proteins are mainly present in the endosperm or cotyledon as sources of amino acids for supporting early seedling growth. In germinating seeds, the storage proteins are digested by proteinases belonging to a papain family that exists widely in monocot and dicot seeds. There are two well-known plant papain types of endopeptidase, SH-EP and EP-B, which are specifically expressed in germinating seeds and growing seedlings. SH-EP and EP-B have been isolated from germinating *Vigna mungo* and barley seedlings, respectively, and their primary structures and enzymatic characterizations have been studied extensively (Mitsuhashi et al. 1986; Akasofu et al. 1989; Koehler and Ho 1990a, 1990b). Papain-family proteinases including both enzymes have common features such as the ERFNIN motif comprising amino acid residues dispersed in the propeptide region (Karrer et al. 1993), a catalytic triad constituting a catalytic site (Kamphuis et al. 1985), and a long pro-sequence involved in the activation and correct folding of the enzyme (Vernet et al. 1991). SH-EP and EP-B are initially transcribed as inactive precursors with long prosequences. Both precursors were detected by SDS–PAGE/immunoblot analysis and shown to be activated via several intermediates by multiple cleavages. In the case of activation of *V. mungo* SH-EP, the inactive SH-EP precursor was initially processed by another endopeptidase, VmPE-1, that exhibits strict substrate specificity to the carboxy-terminal asparagine (Asn) residues of the peptide (Okamoto and Minamikawa 1995).

Previously, we detected two major cysteine endopeptidases, REP-1 and REP-2, in germinating rice seeds and showed that REP-1, a member of the papain family

of proteinases, is mainly responsible for the degradation of the major seed storage protein, glutelin (Kato and Minamikawa 1996). Since REP-1 contains a long pro-sequence predicted by cDNA, we have speculated that its activation is mediated by another endopeptidase such as REP-2, which co-exists with REP-1 in germinated seeds and young seedlings, as in the case of the activation of SH-EP.

Recently, it was shown that two types of cysteine endopeptidase, papain-type and legumain-like proteinases, are involved in the successive degradation of storage globulin in vetch seeds (Becker et al. 1995; Schlereth et al. 2001). It can be speculated that REP-2 cooperates with REP-1 in the degradation of seed storage proteins.

In the present study, we attempted to purify and characterize REP-2 in order to clarify the role of REP-2 in germinating rice seeds. We partially purified REP-2 and determined that it is an asparaginyl endopeptidase. Although asparaginyl endopeptidase activity has been reported within germinating (Sutoh et al. 1999) and maturing (Bottari et al. 1996) wheat seeds, no asparaginyl endopeptidases have been isolated from monocot seeds. The present study is therefore the first to identify and characterize a monocot asparaginyl endopeptidase.

Materials and methods

Plant materials

The rice (*Oryza sativa* L. cv. Koshihikari) seeds used for purification were purchased from the Shibata-shi Agricultural Cooperative Association (Japan). The rice seeds (cv. Nihonbare) used for the analysis of mRNAs and genes were a gift from Dr. Seiichiro Kiyota (National Institute of Agrobiological Resources, Tsukuba, Japan). All rice seeds were germinated in continuous darkness at 27 °C.

Measurement of endopeptidase activity of REP-2

Endopeptidase activity was assayed using a synthetic substrate, Z-Ala-Ala-Asn-MCA (benzyloxycarbonyl-alanyl-alanyl-asparaginyl-7-(4-methyl)-coumarylamide), which was produced according to Kembhavi et al. (1993). The reaction mixture consisted of 5 µl of 10 mM Z-Ala-Ala-Asn-MCA (dissolved in dimethyl sulfoxide), 1980 µl of 50 mM sodium acetate (pH 6.0) containing 10 mM 2-mercaptoethanol, and 15 µl of enzyme solution. The mixture without the enzyme was pre-incubated for 5 min at 30 °C, and after the addition of the enzyme solution the mixture was incubated for a further 10 min. The enzymatic reaction was stopped by heating in a boiling-water bath for 5 min, and the solution was subsequently cooled in water at room temperature. The amount of free AMC (7-amino-4-methyl-coumarin) released during the incubation compared with that of zero-time incubation was assessed according to Okamoto et al. (1999). One unit of enzyme was defined as the amount of endopeptidase releasing 1 µmol AMC under the above-mentioned conditions.

Purification of REP-2

All the following procedures were conducted at 4 °C. Two kilograms of day-9 germinated seeds were homogenized with 4 l of 50 mM sodium acetate (pH 6.0), containing 10 mM 2-mercaptoethanol. The homogenate was centrifuged at 6,500 g for 30 min, and solid ammonium sulfate was added to the clarified solution to a

saturation of 40–75%. The precipitate was dissolved in 50 mM sodium acetate (pH 6.0) containing 10 mM 2-mercaptoethanol and 1.0 M ammonium sulfate, and was fully dialyzed against the same buffer. The solution was loaded on a column (16 mm × 400 mm) of butyl-Cellulofine (Seikagaku Kogyo, Tokyo, Japan). The column was washed with the same buffer, and adsorbed proteins were eluted from the column with a linear gradient (180 ml/180 ml) of 1.0 to 0 M ammonium sulfate in the buffer. The fractions of REP-2 were combined, and proteins were precipitated by adding solid ammonium sulfate to 80% saturation. The precipitate was dissolved in 5 ml of 20 mM sodium acetate (pH 6.0) containing 5 mM 2-mercaptoethanol, and the solution was thoroughly dialyzed against the same buffer. The solution was clarified by centrifugation at 11,000 g for 15 min and the supernatant was loaded on a column (10 mm × 100 mm) of QA-52 cellulose (Whatman). The column was washed thoroughly with the buffer, and adsorbed proteins were eluted from the column with a linear gradient (105 ml/105 ml) of 0 to 0.3 M KCl. Active fractions were concentrated by adding solid ammonium sulfate to 80% saturation. The precipitate was dissolved in 1 ml of 20 mM sodium acetate (pH 6.0) containing 5 mM 2-mercaptoethanol and 0.3 M KCl. Proteins were separated on a Sephacryl S-200 column (16 mm × 900 mm; Amersham Pharmacia Biotech), and the final enzyme fraction was obtained.

Determination of amino acid sequences of the amino terminal of REP-2

After the gel filtration chromatography, two forms of REP-2—REP-2α (39 kDa) and REP-2β (40 kDa)—were separated into fractions 16 and 15, respectively. Both fractions were lyophilized following desalting on PD-10 columns (Amersham Pharmacia Biotech), and the powders were dissolved in a small volume of SDS–PAGE sample buffer. The solutions were divided into two aliquots and analyzed by SDS–PAGE on a gel (Laemmli 1970). Proteins were transferred to a polyvinylidenedifluoride (PVDF) membrane (Millipore), after which the membrane was cut in half. One half of the membrane was stained with the antiserum against legumain (an asparaginyl endopeptidase from jack bean) with an ECL kit (Amersham Pharmacia Biotech) to act as a reference for the REP-2 bands. The other half was stained with Coomassie blue, and the bands of REP-2α and REP-2β were cut from the membrane for analysis of the amino-terminal amino acid sequences of the two REP-2 enzymes using an automated sequence analyzer (model 477A; Applied Biosystems).

Two-dimensional PAGE analysis of REP-2α and REP-2β expression

Aleurone layers and endosperm from 100 germinated seeds were collected, and other tissues including shoots and roots were discarded. Seeds were homogenized in 10 ml of 50 mM sodium acetate (pH 6.0) containing 10 mM 2-mercaptoethanol. The homogenate was centrifuged twice at 11,000 g for 15 min. The supernatant was passed through an 80-µm nylon mesh and added to 30 ml of cold acetone. Solutions were maintained at −20 °C overnight. The precipitated proteins were collected by centrifugation and dissolved in a small volume of distilled water; the solution was subsequently diluted to 1 ml with distilled water. Twenty microliters of each solution prepared from maturing and germinating seeds was analyzed by two-dimensional PAGE (Manabe et al. 1979).

Isolation of the REP-2 cDNA and gene

Total RNA was prepared from the aleurone layers and endosperm after removing shoots and roots from day-4 seedlings as described in Kato and Minamikawa (1996). Poly(A)$^+$ RNA was then

purified using oligo d(T)-cellulose (Amersham Pharmacia Biotech). The internal cDNA fragment of REP-2 cDNA was amplified by RT–PCR using four primers: one primer deduced from the amino-terminal amino acid sequence of REP-2 and three primers designed from motifs conserved among legume asparaginyl endopeptidases (Becker et al. 1995; Kinoshita et al. 1995a, 1995b; Okamoto and Minamikawa 1999). Single-stranded cDNA was synthesized using the oligo d(T) primer with a first-strand cDNA synthesis kit (Takara Shuzo, Shiga, Japan), and this was used as the template for the subsequent PCR steps. The first PCR was performed with two external primers, GGN GTN GGN ACN (A/C)GN TGG GC and GT(A/G) CAI TAC (G/C)T(T/C) AT(G/A) CCI (C/T)T, the primary product from which was used as the template for the second PCR. The fragment was amplified between the following internal primers: TT(C/T) ATG TA(C/T) GA(C/T) GA(C/T) AT(A/C/T) and NGC (A/G)TG (C/T)TT (C/T)TT (C/T)TT NA(A/G). The final PCR product was cloned into a TA-vector (Invitrogen), and the nucleotide sequence of interest was confirmed using an ABI310 automated sequence analyzer (Applied Biosystems). The amplified fragment was used as a probe for subsequent plaque lifting.

Double-stranded cDNAs were constructed with a Time Saver cDNA synthesis kit (Amersham Pharmacia Biotech) according to the manufacturer's instructions from day-4 poly(A)$^+$ RNA. Synthesized cDNAs were cloned into the λZAP II cloning vector (Stratagene) and packaged with Gigapack III gold (Stratagene). Plaques were lifted according to the method of Sambrook et al. (1989). The pBluescript phagemid having the REP-2 cDNA insert was excised in vivo according to the manufacturer's instructions. Serial deletions of pRAN151 (a REP-2 cDNA) from both orientations were obtained with a Kilo-sequence deletion kit (Takara Shuzo, Shiga, Japan), and nucleotide sequences of deletions were determined as described above.

Rice genomic DNA was isolated from day-9 seedlings under continuous darkness using the cetyltrimethylammonium bromide method (Rogers and Bendich 1985). The REP-2 gene was isolated from the genomic DNA by PCR using the following six primers: REP-2Fw, ATGGCGGCGCGGTGGTGCTTCG; REP-2Rv, TCGTATATACAGGGCATC-TAT; REP-2pfhFw, CGGAAGG GAGGGCTAAAGGAG; REP-2pfhRv, ATGATCAG-AG TAGAAGATAAA; REP-2pflFw, AAGCTTTACCTATACCAA GGT; and REP-2pflRv, CTGACCAGAAGGTCTAACAGC. The PCR products were cloned into pGEM-T easy vector (Promega). The nucleotide sequence of the REP-2 gene was determined as described above.

Northern and genomic Southern blot hybridization

For Northern blot analysis, total RNA was prepared from various tissues as described above. The RNA (10 µg) was loaded on a 1.4% agarose (w/v) gel, then transferred onto a Hybond-N membrane (Amersham Pharmacia Biotech) and immobilized by exposure to ultraviolet light. The membrane was prehybridized for 2 h in a hybridization solution (Shintani et al. 1997) at 42 °C. Hybridization was performed in the solution for 12 h at 42 °C, then the membrane was washed twice with 2×SSC containing 0.1% SDS (w/v) and once with 0.2×SSC containing 0.1% SDS (w/v) for 30 min at 42 °C.

For Southern blot analysis, genomic DNA was prepared from rice seedlings as described above. Nuclear DNA (10 µg) was digested by restriction enzymes, and electrophoresed on a 0.7% agarose (w/v) gel. After electrophoresis, the DNA was transferred onto a nylon membrane, and fixed under ultraviolet light. Hybridization was conducted with REP-2 cDNA (pRAN151) as a probe (Sambrook et al. 1989).

Treatment of de-embryonated seeds with gibberellic acid (GA$_3$) and abscisic acid (ABA)

One-hundred de-embryonated rice seeds were sterilized with 1% sodium hypochlorite (v/v) containing 0.05% Tween 20 (v/v) for 1 h, then rinsed thoroughly with sterile water and incubated in 10 ml of water at 25 °C for the duration of the 5-day test period. One-hundred microliters of 10^{-3} M GA$_3$ and ABA was added to the water to give a final concentration of 10^{-6} M.
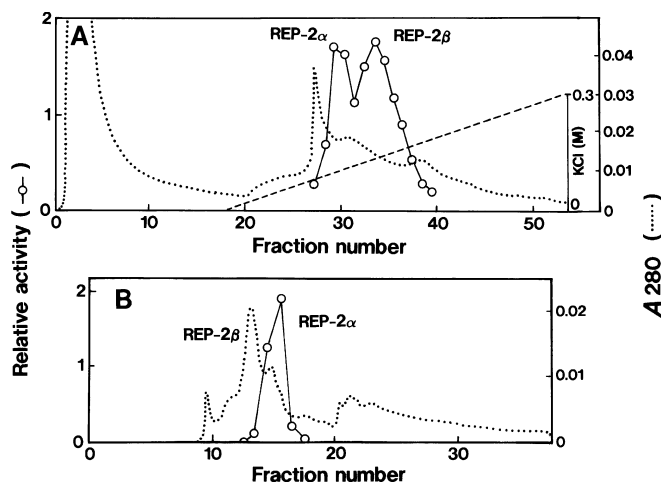
Results

Identification and characterization of REP-2

The REP-2 fraction was separated from the REP-1 fraction using hydrophobic chromatography as described by Kato and Minamikawa (1996). These fractions were used to examine the substrate specificities of REP-1 and REP-2 (Table 1). REP-1 has no exopeptidase activity and a strict substrate specificity toward carbobenzoxy-L-phenylalanyl-L-arginine-4-methyl-coumaryl-7-amide (Z-Phe-Arg-MCA), which is the most efficient substrate for papain-family endopeptidases (Davy et al. 1998; Okamoto et al. 1999). However, the REP-2 fraction exhibits exopeptidase activity, and efficiently digested Z-Ala-Ala-Asn-MCA but not Z-Phe-Arg-MCA (Table 1). As a result, during further purification of REP-2 we used Z-Ala-Ala-Asn-MCA as a substrate. We found that REP-2 was stable at around pH 6.0, but that it immediately lost its activity under other pH conditions, and that REP-2 is also sensitive to temperatures higher than 45 °C (data not shown). Anion-exchange chromatography was used to separate REP-2 into two fractions, named REP-2α (39 kDa) and REP-2β (40 kDa) (Fig. 1A). Both fractions were combined and concentrated, then loaded on a Sephacryl S-200 column. In this chromatographic analysis, REP-2 was distributed into fractions 15 and 16, (REP-2β and REP-2α, respectively; Figs. 1B and 2A). Finally, REP-2 was partially purified 303-fold (Table 2). Analysis of the amino-terminal sequences of REP-2α and REP-2β revealed 10 identical amino acids (Fig. 2B). Both forms of REP-2 were visualized with the antiserum against legumain (Fig. 2A). From these results, we concluded that REP-2 is an asparaginyl endopeptidase.

Table 1 Substrate specificities of rice (*Oryzasativa*) REP-2 and REP-1. REP-2 and REP-1 fractions were obtained from the first (butyl-Cellulofine)and second (QA-52) purification steps, respectively (Kato and Minamikawa 1996). 100% activity was defined as the amount of 7-amino-4-methyl-coumarin (AMC) released when Z-Ala-Ala-Asn-MCA and Z-Phe-Arg-MCAwere used for REP-2 and REP-1, respectively. *Z* Benzocarboxy, *MCA* 4-methyl-coumaryl-7-amide

Substrate	% Activity	
	REP-2	REP-1
	(Butyl-Cellulofine)	(QA-52)
Z-Ala-Ala-Asn-MCA	100	45.0
Z-Phe-Arg-MCA	0	100
Z-Arg-Arg-MCA	4.2	0
Arg-MCA	110	0
Ala-MCA	3.9	0
Leu-MCA	55.5	0
Phe-MCA	32.1	0

Fig. 1A, B Elution profiles of rice (*Oryza sativa*) REP-2 on column chromatography. The endopeptidase fraction obtained from ammonium sulfate fractionation (40–75% saturation) was loaded on a butyl-Cellulofine column for chromatography, resulting in the separation of REP-1 and REP-2 (Kato and Minamikawa 1996). **A** The REP-2 fraction from hydrophobic chromatography was loaded on a column of QA-52 cellulose. The column was fully washed after protein adsorption. Adsorbed proteins were eluted from the column by increasing the KCl concentration in the buffer. Active fractions (REP-2α and REP-2β) were combined and subjected to subsequent gel filtration chromatography. **B** Two active fractions, REP-2α and REP-2β, were re-separated by gel filtration chromatography

Fig. 2A, B Gel-electrophoretic separation and N-terminal amino acid sequences of rice REP-2α and REP-2β. **A** Fractions 15 and 16 obtained from gel filtration (Fig. 1B) were analyzed by SDS–PAGE followed by silver staining or immunoblotting. REP-2α and REP-2β differed in their molecular masses as estimated on the gel (at 39 kDa and 40 kDa, respectively). Both REP-2α and REP-2β cross-reacted with the antiserum against legumain on the membrane. **B** The bands of both REP-2 proteins blotted on a PVDF membrane were cut from the membrane, and the amino-terminal amino acid sequences were analyzed with an automated sequence analyzer ABI477A. *X* Unidentified residues

Molecular characterization of REP-2

Using degenerate primers for the amino-terminal amino acid sequence of REP-2 and primers for the consensus sequence for plant asparaginyl endopeptidases, we used RT–PCR to amplify a partial 0.48-kbp cDNA. The deduced amino acid sequence was highly homologous to known legume asparaginyl endopeptidases such as jack bean legumain (75%), vetch bean proteinase B (76%), and *V. mungo* VmPE-1 (72%) (Takeda et al. 1994; Becker et al. 1995; Okamoto and Minamikawa 1995). Therefore, we concluded that the amplified fragment was the partial cDNA for REP-2 mRNA. Prior to construction of a cDNA library, we preliminarily conducted Northern blot analysis using the amplified fragment as a probe. The highest level of REP-2 mRNA was found in day-4 seedlings. We prepared poly(A)$^+$ RNA from day-4 seedlings and constructed a cDNA library containing 50,000 independent clones. Subsequently, 13 clones were selected by plaque lifting, and they were grouped into a single clone represented by pRAN151 (DDBJ Nucleotide Sequence Database accession no. AB081464) on restriction-enzyme mapping.

The deduced amino acid sequence included 10 amino acids which were determined from the sequencing of REP-2 proteins, and revealed that REP-2 was initially transcribed as a 55-kDa precursor (Fig. 3). Hydropathy plot analysis predicted that REP-2 contains a signal sequence which is cleaved from the carboxy terminal of Ala22 followed by a small propeptide which is also removed to form the mature REP-2 (von Heijne 1983). A potential Asn-linked glycosylation site was found in the sequence at Asn153. REP-2 was highly homologous with legume asparaginyl endopeptidases, and the phylogenetic tree revealed that REP-2 is more similar to jack

Table 2 Partial purification of REP-2 from rice seedlings. The asparaginyl endopeptidase activity was measured using Z-Ala-Ala-Asn-MCAas a substrate. The values given for QA-52 and Sephacryl S-200 represent REP-2α plus REP-2β

Purification step	Total protein (mg)	Total activity (U×10^{-3})	Specific activity (U/mg)	Purification (*n*-fold)
Crude extract	2,700	9.1	3.3	1
Ammonium sulfate	1,200	8.9	7.4	2.2
butyl-Cellulofine	13	0.6	48	14
QA-52 cellulose	0.75	0.13	168	51
SephacrylS-200	0.1	0.1	1,000	303

```
1    ATGGCGGCGCGGTGGTGCTTCGCGTTGCTCTTGGCGCTGTCGGCGGCCGCGGCGGGGGCC
     M  A  A  R  W  C  F  A  L  L  L  A  L  S  A  A  A  A  G  A

61   GGAGCCAAGCGGACGTGGGAGCCGGTGATTCGGATGCCGGGGGAGGTGGTGGAGGAGGAG
     G  A  K  R  T  W  E  P  V  I  R  M  P  G  E  V  V  E  E  E
            ▲

121  GTGGCCACCGTGCCGCGCGGCAGCGAGGGCACCGAGGAGGAGGAGAAGGACGGCGTCGGG
     V  A  T  V  P  R  G  S  E  G  T  E  E  E  E  K  D  G̲ ̲V̲ ̲G̲

181  ACGAGGTGGGCGGTGCTCGTCGCGGGCTCCTCCGGCTACGGTAACTACAGGCACCAGGCT
     T̲  R̲  W̲  A̲  V̲  L̲  V̲  A̲  G̲  S  S  G  Y  G  N  Y  R  H  Q  A

241  GATGTATGCCATGCATCACCAGATATTACGGAAGGGAGGGCTAAAGGAGAGGAACATTGTC
     D  V  C  H  A  Y  Q  I  L  R  K  G  G  L  K  E  E  N  I  V

301  GTGTTTATGTATGACGACATTGCCAATAATATACTTAACCCGAGACCAGGGGTTATTGTC
     V  F  M  Y  D  D  I  A  N  N  I  L  N  P  R  P  G  V  I  V
                                                          →

361  AATCATCCACAGGGTGAAGATGTTTATGCAGGAGTTCCGAAGGATTACACTGGAGACGAA
     N  H  P  Q  G  E  D  V  Y  A  G  V  P  K  D  Y  T  G  D  E

421  GTCACTGCAAAAAACTTTTACGCGGTTCTCTTGGGCAATAAAACTGCAGTTACTGGGGGT
     V  T  A  K  N  F  Y  A  V  L  L  G  N  K  T  A  V  T  G  G

481  AGTAGGAAGGTCATAGATAGCAAACCAAATGACCACATATTTATCTTCTACTCTGATCAT
     S  R  K  V  I  D  S  K  P  N  D  H  I  F  I  F  Y  S  D  H
                                       ●

541  GGGGGTCCTGGAGTTCTTGGTATGCCCAATCTGCCATATCTTTATGCTGCTGACTTCATG
     G  G  P  G  V  L  G  M  P  N  L  P  Y  L  Y  A  A  D  F  M

601  AAGGTGTTACAAGAAAAACATGCCTCCAATACTTACGCTAAAATGGTTATATATGTAGAA
     K  V  L  Q  E  K  H  A  S  N  T  Y  A  K  M  V  I  Y  V  E

661  GCGTGTGAAAGTGGCAGTATTTTCGAAGGTTTGATGCCAGAGGACCTTAATATTTATGTC
     A  C  E  S  G  S  I  F  E  G  L  M  P  E  D  L  N  I  Y  V

721  ACAACAGCGTCTAATGCAGAAGAGAGTAGTTGGGGGACGTACTGCCCAGGAATGGAACCA
     T  T  A  S  N  A  E  E  S  S  W  G  T  Y  C  P  G  M  E  P

781  TCGCCTCCTTCCGAGTACATCACCTGCTTAGGTGACCTATACAGCGTTTCTTGGATGGAA
     S  P  P  S  E  Y  I  T  C  L  G  D  L  Y  S  V  S  W  M  E
                                   ←

841  GACAGTGAGACTCACAATCTGAAGGAGGAATCAATTAAGAAGCAGTACGAAGTGGTTAAG
     D  S  E  T  H  N  L  K  E  E  S  I  K  K  Q  Y  E  V  V  K

901  AAGCGTACCTCTGACATGAACAGTTACGGTGCTGGCTCTCATGTTATGGAGTACGGGGAC
     K  R  T  S  D  M  N  S  Y  G  A  G  S  H  V  M  E  Y  G  D

961  AGGACATTCAAGGATGACAAGCTTTACCTATACCAAGGTTTCGACCCTGCAAATGCCGAA
     R  T  F  K  D  D  K  L  Y  L  Y  Q  G  F  D  P  A  N  A  E

1021 GTTAAGAACAAGCTATCATGGGAAGGCCCAAAGGCTGCAGTCAACCAAAGAGATGCAGAT
     V  K  N  K  L  S  W  E  G  P  K  A  A  V  N  Q  R  D  A  D

1081 CTTCTTTTCCTCTGGAGAAGGTATGAGCTGTTACATGACAAGTCCGAAGAGAAGCTGAAA
     L  L  F  L  W  R  R  Y  E  L  L  H  D  K  S  E  E  K  L  K

1141 GCTCTGAGGGAGATCAGTGACACTGTTATGCATAGGAAGCTTCTTGACAGCAGTGTTGAT
     A  L  R  E  I  S  D  T  V  M  H  R  K  L  L  D  S  S  V  D

1201 CTTGTCGGAAAACTTCTCTTTGGATTTGGAAATGGGCCTTCTGTGCTTCAAGCTGTTAGA
     L  V  G  K  L  L  F  G  F  G  N  G  P  S  V  L  Q  A  V  R

1261 CCTTCTGGTCAGCCACTCGTTGATGATTGGGATTGTTTGAAGAGAATGGTACGGATCTTT
     P  S  G  Q  P  L  V  D  D  W  D  C  L  K  R  M  V  R  I  F

1321 GAGTCGCATTGCGGACCGTCACTCAGTACGGTATGAAGCACATGAGGGCGTTTGCTAAT
     E  S  H  C  G  P  L  T  Q  Y  G  M  K  H  M  R  A  F  A  N

1381 ATTTGCAACAACGGTATCTCTGGCGCTTCAATGAAGGAAGCAAGCATCGCTACTTGCAGC
     I  C  N  N  G  I  S  G  A  S  M  K  E  A  S  I  A  T  C  S

1441 AGTCACAACTCGGGTAGATGGAGCTCGCTGGTGCAGGGGTACAGTGCTTGATCTGCCGCA
     S  H  N  S  G  R  W  S  S  L  V  Q  G  Y  S  A  *

1501 ACACGTGATATAGCCTCAAAACTTATGCCTTGTACTCGTACTTCCTCAAAACAATACCTT

1561 CAGGACTGCAAATAACCGTCCAAGGATGTACCGCTGTATGCTACTGTACCTCTATGTGTG

1621 CAATATAGATGCCCTGTATATACGAGTATATGAATGAATCGTGTGCTTTTACAAAAAAAA

1681 AA  1682
```

Fig. 3 Nucleotide sequence of pRAN151 and deduced amino acid sequence of REP-2. Numbering starts from the adenine of the initiation codon. The *filled circle* indicates a potential site of glycosylation. The deduced amino acid sequence is described with a one-letter code. The *underlining* shows amino acids corresponding to the sequence of REP-2α. The *filled triangle* shows the putative cleavage site of a signal peptide. The stop codon is indicated by an *asterisk*. The *arrows* show the conserved region among plant asparaginyl endopeptidases, which was designed for the degenerated primer
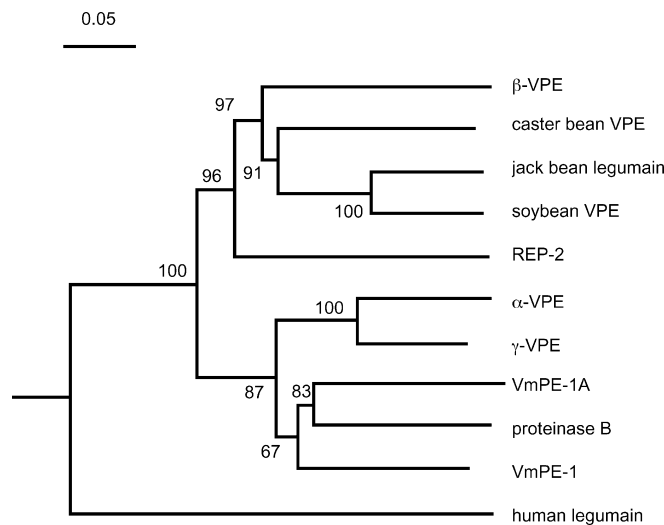
Fig. 4 A phylogenetic tree of plant asparaginyl endopeptidases. The phylogenetic tree was constructed by the neighbor-joining method (Saitou and Nei 1987). Numbers at branches indicate bootstrap values (Felsenstein 1985). The following amino acid sequences are used: α-VPE and β-VPE (Kinoshita et al. 1995a), γ-VPE (Kinoshita et al. 1995b), castor bean VPE (Hara-Nishimura et al. 1993), jack bean legumain (Takeda et al. 1994), soybean legumain (Shimada et al. 1994), VmPE-1 and VmPE-1A (Okamoto et al. 1999), proteinase B (Becker et al. 1995), and human legumain (Chen et al. 1997, used as the outgroup)

bean legumain than VmPE-1 (Fig. 4; Senyuk et al. 1998; Okamoto and Minamikawa 1999).

Expression of REP-2

Two-dimensional PAGE analysis using legumain antiserum revealed that REP-2α was not detectable in seeds 5 days after flowering (DAF). The level of REP-2β in seeds started to increase from 25 DAF and reached a maximum at 35 DAF. The amount of REP-2β transiently decreased in the dry (quiescent) seeds, increased again after imbibition, and reached a maximum level around 2–4 days after imbibition began (DAI), whereas the amount of REP-2α changed only slightly during this period (Fig. 5).
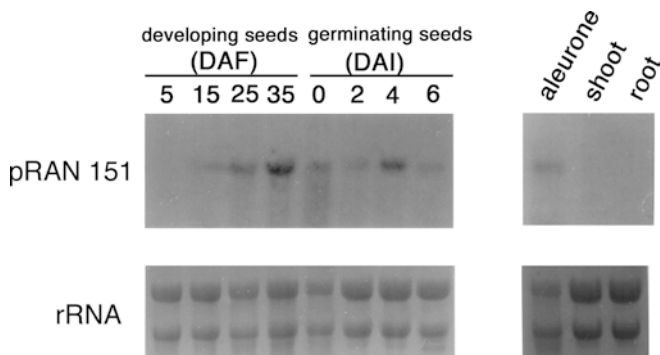
Northern blot analysis showed that REP-2 mRNA was not detectable in seeds 5 DAF but increased from 15 DAF, and its profile was similar to that of both REP-2 proteins throughout the maturing and germinating seed stages (Fig. 6). We prepared total RNA from the shoots and the embryos with roots from 4-DAI seeds. REP-2 mRNA was not detected in these tissues, indicating that in germinating seeds, REP-2 mRNA is expressed specifically in the aleurone layer.

Effects of plant hormones on REP-2 expression

The expression of REP-1 was controlled by GA_3 and ABA (Shintani et al. 1997). We were interested in
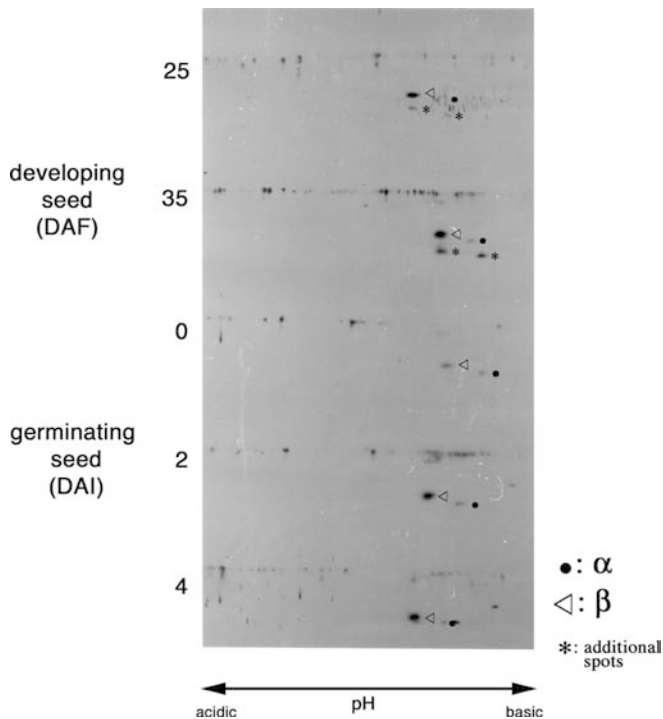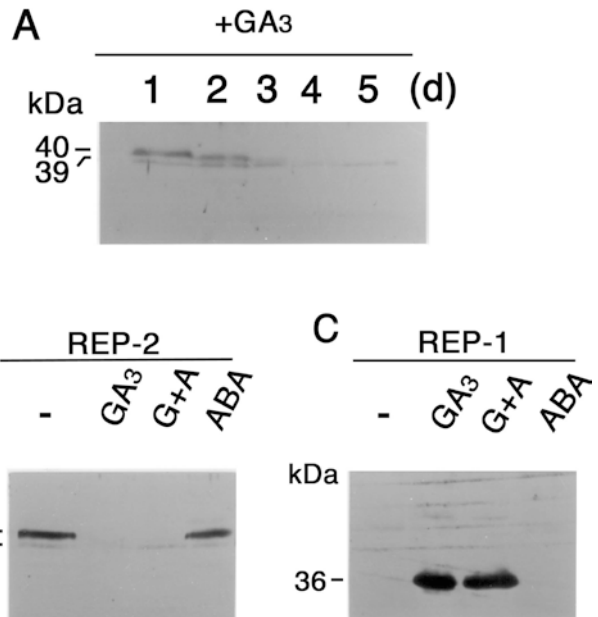
Fig. 5 Two-dimensional PAGE analysis of the levels of REP-2 in maturing and germinating rice seeds. The levels of REP-2α and REP-2β in seeds at the indicated stages were analyzed by two-dimensional PAGE/immunoblot. Ampholines (Amersham Pharmacia Biotech) of pH 3.5–10 and pH 5.0–8.0 were mixed at a ratio of 2:1. The first nondenaturing PAGE and the second SDS–PAGE were performed on 6% and 12.5% gels (w/v), respectively. *Filled circles* and *open triangles* indicate REP-2α and REP-2β, respectively. *Asterisks* show additional spots. *DAF* Days after flowering, *DAI* days after imbibition began

Fig. 7A–C The effects of ABA and GA₃ on the levels of REP-2. **A** De-embryonated seeds were incubated in the presence of 1 mM GA₃ for 5 days. The bands of both REP-2 proteins were visualized with the antiserum against legumain. **B, C** De-embryonated seeds were incubated in the absence of plant hormone (−), in the presence of 10^{-6} M GA₃ (*GA₃*) or 10^{-6} M ABA (*ABA*), and in the presence of both plant hormones (*G + A*) for 5 days. The bands of REP-2 and REP-1 were visualized using antisera against legumain and REP-1, respectively
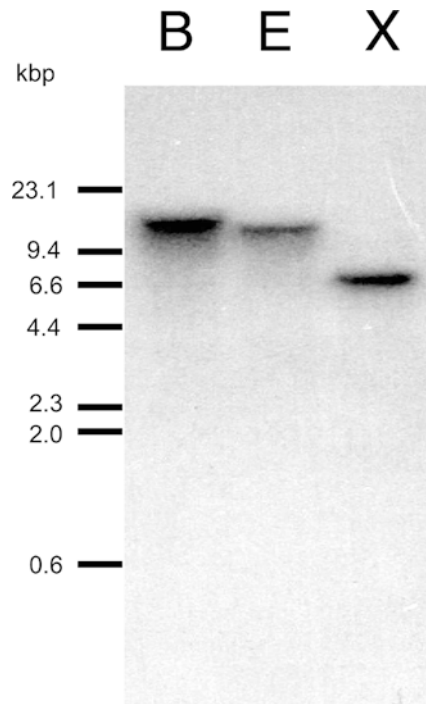

Fig. 6 Expression of REP-2 mRNA in maturing and germinating rice seeds. Total RNA was prepared from maturing, quiescent (0 DAI), and germinating seeds or each part of 4-DAI seedlings. The RNA (10 µg) was electrophoresed on a 1.4% agarose (w/v) gel. After electrophoresis, the gel was stained with methylene blue (rRNA) or analyzed by RNA blotting with pRAN151 cDNA insert as a probe

Fig. 8 Genomic Southern blot of the REP-2 gene. Rice genomic DNA (10 µg) was digested completely by 40 units of *Bam*HI (*B*), *Eco*RI (*E*), and *Xba*I (*X*), and electrophoresed on a 0.7% agarose (w/v) gel. The DNA was transferred onto a nylon membrane and hybridized with a ^{32}P-labeled pRAN151 cDNA insert

whether REP-2 expression was controlled by these plant hormones. We treated the de-embryonated seeds with GA₃, ABA, or only distilled water (mimic) for 5 days, and measured the protein levels. In the de-embryonated seeds, REP-2β started decreasing as soon as GA₃

```
   1 ATGGCGGCGCGGTGGTGCTTCGCGTTGCTCTTGGCGCTGTCGGCGGCCGCGGCGGGGGCCGGAGCCAAGCGGACGTGGGAGCCGGTGATTCGGATGCCGG 100
       M  A  A  R  W  C  F  A  L  L  L  A  L  S  A  A  A  A  G  A  G  A  K  R  T  W  E  P  V  I  R  M  P  G
 101 GGGAGGTGGTGGAGGAGGAGGTGGCCACCGTGCCCGCGCGCAGCGAGGGCACCGAGGAGGAGGAGGAGGACGGCGTCGGGACGAGGTGGGCGGTGCTCGT 200
       E  V  V  E  E  E  V  A  T  V  P  R  G  S  E  G  T  E  E  E  E  K  D  G  V  G  T  R  W  A  V  L  V
 201 CGCGGGCTCCTCCGGCTACGGTAACTACAGGCACCAGGTAAGCGCACGGGTCCAGTACTCCAGTGTCGTCGGCTTCGCCGTCGCTTAGTTCGTGCGCGTC 300
       A  G  S  S  G  Y  G  N  Y  R  H  Q
 301 TGGAGCTCGATTTGGGGGGCTTGGATGATGATGCAGCGTGGAACTGCTGTGAGGCTGTGTAGCTATGGGCAGCCGTTTCTGAAAGTCGTTCAGATCATTC 400
 401 GTAACGAGATCTTTAGTATTCATGAATTTAAAGATTTATATGTTCAATTTTGTCGTGCCATGATACTTCAATGCTCCAAAATTTTCCGGTTTTTAGCTTT 500
 501 TGAAACCGAAAAAGGTTTGCCTTCAGTACTGAAATTAACCGAGCAATTTTCCCGTCAAAAAAAAGAATTAACTTGAGCAATAAAGTTAACGATTTGGGCTT 600
 601 GATGAGAAGTGAACTACTTGAAAGCAGCCCCAGGATCTTTTTTTTTGGGAATGCGCAAGGAGCACAATCATTGTATTAAGAAGAAATAAAGTTTATGTATA 700
 701 AGGCGGGAACAAAAGCAAAACAAGATGGGGATGGGATACAGGAGAGGATCAAGAGTACTAAAGCCTAATCTCCAAATCATTTGGCGACCCCTCCCTATC 800
 801 CGCATGGATTACTACGCTAGGAATGGCTGCCCTAGCCAGCCTTCAAAGATCAGCTTCGCCTGTCATGGCCCGAGCGATGTCCACCCCAAGATCTGCTGGT 900
 901 TTGATTGAAAACTCTATTGTTTCTCTCCTTCCAAATTGTCCATGCTACTCAAGGTTGCGATGGTGTCAAACCCCTGACGATGCGGCCTTGATCTATGTAC 1000
1001 GTTTGATCATGTGAGTGTGTGCGGACTGCAGAGTGCTAGTAGTTCAATCTCAGCATGATTTGCACACATCGCTATCCAAGTTGTACTTAGTCTTTTTGTC 1100
1101 AGGAAGTTCAACAGACTGATGCCTAGACTCATTTCAAGGAAGAAAATGGCGTGAGCAGCAATATAATTCTGGATAAGCAAACGAATTTGAAGTTTTCAAT 1200
1201 TTTTTCTCGGGTTCTAAGTTCTAGTAATATCCATGTTTAAGCAGATTGAGAAAAAAAATGGCACAACATATGATGAGTTTAAAATGGAGGGATCAATTT 1300
1301 AATACGAAAGAAATATCAGCTGAAATATTTAGTGGCTAATACTTCAGTGATTAATTTTTACTGAAATATATGTAGCAGAAAGAGACAGGCTCATAGCAAG 1400
1401 TTTGTATTTATATTGTTTCGTTTCAAATTTAGAAAGGAATTGAGACCCAACAAATGGTTTATTTCTCCAAGGACTGTTTGCTTGAGCGTTTGATGTATGG 1500
1501 GCTCCTTTGGTTGTGTTTGCACTTTGCAGGCTGATTGTATGCCATGCATACATTACGGAAGGGAGGGCTAAAGGAGGAGAACATTGTCGTGTTTAT 1600
                                                     A  D  V  C  H  A  Y  Q  I  L  R  K  G  G  L  K  E  E  N  I  V  V  F  M
1601 GTATGACGACATTGCCAATAATATACTTAACCCGAGACCAGGGGTTATTGTCAATCATCCACAGGGTGAAGATGTTTATGCAGGAGTTCCGAAGGTATTA 1700
       Y  D  D  I  A  N  N  I  L  N  P  R  P  G  V  I  V  N  H  P  Q  G  E  D  V  Y  A  G  V  P  K
1701 CACGATGCCGTACTGCCTCACTGTTATTCGTTCCTCGGGCTGGAGAGAGGACTGTTAGGGTCTTATGTCCAATTATTTACCCTTTTGCACTTCTTTATTA 1800
       H  D  A  V  L  P  H  C  Y  S  F  L  G  L  E  R  G  L  L  G  L  M  S  N  Y  L  P  F  C  T  S  L  L
1801 GCTTCAGGTGGCTTCAGATTTGCGTGTACATGGCACAAGTGCATAACATATATATTGACCTTTTTCTCTAAGTTTGATGAAGATAACACACTCTATAATT 1900
       A  S  G  G  F  R  F  A  C  T  W  H  K  C  I  T  Y  I  L  T  F  F  L  *
1901 CCAGGATTACACTGGAGACGAAGTCACTGCAAAAAACTTTTACGCGGTTCTCTTGGGCAATAAAACTGCAGTTACTGGGGGTAGTAGGAAGGTCATAGAT 2000
           D  Y  T  G  D  E  V  T  A  K  N  F  Y  A  V  L  L  G  N  K  T  A  V  T  G  G  S  R  K  V  I  D
2001 AGCAAACCAAATGACCACATATTTATCTTCTACTCTGATCATGGGGGTCCTGGAGTTCTTGGTATGTGATAGTTGCATATGCTGCTTTGCTTTTTGAGTT 2100
       S  K  P  N  D  H  I  F  I  F  Y  S  D  H  G  G  P  G  V  L  G
2101 AAATTTTGTCCATCATGCATGTTGACATGTCCAATTCTCAGGTATGCCCAATCTGCCATATCTTTATGCTGCTGACTTCATGAAGGTGTTACAAGAAAAA 2200
                                                       M  P  N  L  P  Y  L  Y  A  A  D  F  M  K  V  L  Q  E  K
2201 CATGCCTCCAATACTTACGCTAAAATGGTAAGATTTTTCTTGGTTACTTATTTGTGGAAGAGCCTCCTTGGATGGAATAAGCAATATATTTTTGGGAGGT 2300
       H  A  S  N  T  Y  A  K  M                                                                     V
2301 TATATATGTAGAAGCGTGTGAAAGTGGCAGTATTTTCGAAGGTTTGATGCCAGAGGACCTTAATATTTATGTCACAACAGCGTCTAATGCAGAAGAGAGT 2400
       I  Y  V  E  A  C  E  S  G  S  I  F  E  G  L  M  P  E  D  L  N  I  Y  V  T  T  A  S  N  A  E  E  S
2401 AGTTGGGGGACGTACTGCCCCAGGAATGGAACCATCGCCTCCTTCCAGGTACATCACCTGCTTAGGTGACCTATACAGCGTTTCTTGGATGGAAGACAGGT 2500
       S  W  G  T  Y  C  P  G  M  E  P  S  P  P  S  E  Y  I  T  C  L  G  D  L  Y  S  V  S  W  M  E  D  S
2501 AGCTTCACATATCTTTACAAATGTAAAGAATTACATATTGCACACAATGCCGTTCGTGGTTTTTTCGTACATTCTTGTAGAACTTTCGGTGGCAAATGTT 2600
2601 TTAATAATCAAATCTGAAGATAATACTATAAATGGTAAAAACTATTTAATTTATTCGCATCATGCAGATGTTATGTGGTCATCTATAGAAGGATTATTTT 2700
2701 TTTTTCTTTTATGTTGAGAGACTGACCTGTACGCACTCCTTCCTCGTGCACGGGACACAAGGTCCAAATTCTTTTCCAAAGCATGTTCATCATTGTACAT 2800
2801 CATTTTAAATGATAATTGGATTAAGTGTGTATTTATATGCATGGGACCGATTTAAATGTATGGGACAGGATATTAAGTGTGTATTTCTTCTGTTATCATTG 2900
2901 AACATGTGCATATCTTGTGTTAAGTGTGTATTTATGTGCATTGGAAATTTCCATACATATCAGCTTTAATGTTCCTTATTTCTATGCAGTTTGCAAAGTA 3000
3001 TTTACTCTATAATATGTAAAATTTATCTCGATGTTAATTGGCAGTGAGACTCACAATCTGAAGGAGGAATCAATTAAGAAGCAGTACGAAGTGGTAAGTC 3100
                                                                     E  T  H  N  L  K  E  E  S  I  K  K  Q  Y  E  V
3101 TTGATTTGCTCTGTACCTTTCCAATTTTCATTGTTAAACTCCCAGGTAGTTATAGCATTGTTGACAGGTTAAGAAGCGTACCTCTGACATGAACAGCTAC 3200
                                                                         V  K  K  R  T  S  D  M  N  S  Y
3201 GGTGCTGGCTCTCATGTTATGGAGTACGGGGACAGGACATTCAAGGATGACAAGCTTTACCTATACCAAGGTTTCGACCCTGCAAATGCCGAAGTTAAGA 3300
       G  A  G  S  H  V  M  E  Y  G  D  R  T  F  K  D  D  K  L  Y  L  Y  Q  G  F  D  P  A  N  A  E  V  K  N
3301 ACAAGCTATCATGGGAAGGCCCAAAGGCTGCAGTCAACCAAAGAGATGCAGATCTTCTTTTCCTCTGGAGAAGGGTATTTCCTAGTCTATTCTTTTGCAG 3400
       K  L  S  W  E  G  P  K  A  A  V  N  Q  R  D  A  D  L  L  F  L  W  R  R
3401 ACTTTTAGTTTGTTGTTTGTTATGTTCCTTCCAAAAATACTTACCAGTACTAGCTATTATCTAATCCCCTTTCTATACAACCATATGGTTGGAGTGCACT 3500
3501 TGCATTTGTTTTTGATGTCACAAATTATTCCAGCCTACAGAGATTTTTCTAAACTGAAGTTCATTATGACGTGTTGAAATTACTTGGTAGTGTACATCAGC 3600
3601 TAAATATCTTTGTTTCTAGTAAACTGTGCTGCTGTCTGCTGCTCGCAAGTTTCTTTTCATTGGAGTATTATCTTAGAGAGATTTTTTTTTTATAAAAGTAT 3700
3701 CTCCTGTTGCCATTTGGAGAACAATTGCACTAGACTATTAGCGAACCAAACTTAAATGCAGAATCTGTGAATGGGCAAGCATTTGTTTCAGTCATGTGCT 3800
3801 TCCATCCTATGGTTGCTAATTTTATTTCTTTGTTTGGTTTCAGTATGAGCTGTTACATGACAAGTCCGAAGAGAAGCTGAAAGCTCTGAGGGAGATCAGT 3900
                                                           Y  E  L  L  H  D  K  S  E  E  K  L  K  A  L  R  E  I  S
3901 GACACTGTTATGCATAGGAAGCTTCTTGACAGCAGTGTTGATCTTGTCGGGAAAACTTCTCTTTGGATTTGGAAATGGGCCTTCTGTGCTTCAAGCTGTTA 4000
       D  T  V  M  H  R  K  L  L  D  S  S  V  D  L  V  G  K  L  L  F  G  N  G  P  S  V  L  Q  A  V  R
4001 GACCTTCTGGTCAGCCACTCGTTGATGATTGGGATTGTTTGAAGAGAATGGTAAAGTTGGAAACTTCTTTATACTCGGCTTACTGAGAACTGACATCTGC 4100
       P  S  G  Q  P  L  V  D  D  W  D  C  L  K  R  M
4101 TATTACTCCACTGCGCTCCAATAGTTCCATATTGTCCTGAAATCTCCTAGTTTCCATTGTTTCTGTTTGTCTGACTTTAATATGAGTTTCTTCAGGTACG 4200
                                                                                                       V  R
4201 GATCTTTGAGTCGCATTGCGGACCGCTCACTCAGTACGGTATGAAGCACATGAGGGCGTTTGCTAATATTTGCAACAACGGTATCTCTGGCGCTTCAATG 4300
       I  F  E  S  H  C  G  P  L  T  Q  Y  G  M  K  H  M  R  A  F  A  N  I  C  N  N  G  I  S  G  A  S  M
4301 AAGGAAGCAAGCATCGCTACTTGCAGCAGTCACAACTCGGGTAGATGGAGCTCGCTGGTGCAGGGGTACAGTGCTTGATCTGCCGCAACACGTGATATAG 4400
       K  E  A  S  I  A  T  C  S  S  H  N  S  G  R  W  S  S  L  V  Q  G  Y  S  A  *
4401 CCTCAAAACTTATGCCTTGTACTCGTACTTCCTCAAAACAATACCTTCAGGACTGCAAATAACCGTCCAAGGATGTACCGCTGTATGCTACTGTACCTCT 4500
4501 ATGTGTGCAATATAGATGCCCTGTATATACGA 4532
```

Fig. 9 Nucleotide sequence of the rice REP-2 gene. The nucleotide numbered "*1*" indicates the adenine of the putative initial methionine. *Arrows* indicate the primers for amplification of the REP-2 gene. The *asterisk* shows a termination codon. The amino acid sequence of REP-2 appears under the nucleotide sequence

treatment began and continued decreasing for 5 days, whereas REP-2α was affected only slightly by the GA₃ treatment (Fig. 7A). When water was used, both REP-2 proteins remained in de-embryonated seeds after 5 days of treatment. The level of REP-2 was maintained in the seeds treated with water or ABA but not with GA_3 (Fig. 7B). However, when day-5 GA_3-treated de-embryonated seeds were treated with ABA, no increase in REP-2 was observed (data not shown). In contrast,

Fig. 10 Introns and exons of REP-2 and related *Arabidopsis thaliana* asparaginyl endopeptidases. *Open squares* and *solid lines* indicate exons and introns, respectively. The following gene structures are used: α-VPE and β-VPE (Kinoshita et al. 1995a), and γ-VPE (Kinoshita et al. 1995b)

1 kb
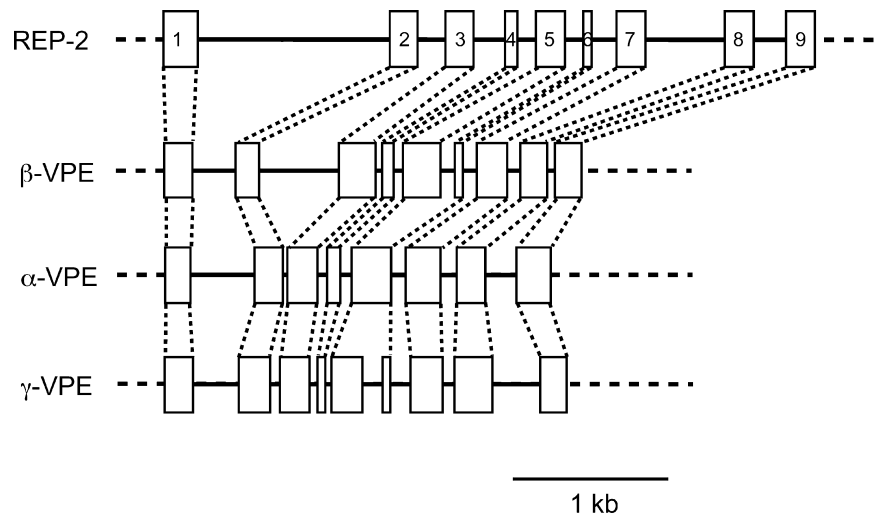
REP-1 could be detected in de-embryonated seeds treated with GA_3 but not with water or ABA. ABA treatment slightly supressed the accumulation of REP-1 (Fig. 7C; Shintani et al. 1997).

Structure of the REP-2 gene

Genomic Southern blot analysis using a full-length REP-2 cDNA as a probe revealed only one main band in each lane when the rice genome was completely digested by *Eco*RI, *Bam*HI, and *Xba*I (Fig. 8), indicating that the REP-2 gene exists as a single gene in the genome.

To isolate genomic clones for REP-2, six oligonucleotide primers were designed in pRAN151 (REP-2 cDNA, Figs. 3, 9). The primers were used to amplify a 5.0-kb fragment, and sequencing it revealed that the REP-2 gene (*Rep2*: DDBJ Nucleotide Sequence Database accession no. AB081465) consists of eight introns and nine exons (Fig. 10). The gene organization of REP-2 is similar to that of legumain-type cysteine endopeptidase genes of the *Arabidopsis* vacuolar processing enzyme β-VPE (Kinoshita et al. 1995a). Additionally, the phylogenic tree indicated that REP-2 is more similar to β-VPE than to α-VPE and γ-VPE (Fig. 4).

Discussion

Characterization of REP-2

We previously showed that the papain-family cysteine endopeptidase, REP-1, exists in germinating rice seeds, and is responsible for the degradation of a major seed storage protein, glutelin. REP-1 has a long prosequence following its signal peptide, as do SH-EP and EP-B, well-studied plant papain-family proteinases, and their prosequences are cleaved during maturation (Mitsuhashi

and Minamikawa 1989; Koehler and Ho 1990a; Kato and Minamikawa 1996). Okamoto and Minamikawa (1995) purified a processing enzyme, VmPE-1, from *V. mungo* seeds which processed the 43-kDa SH-EP precursor to a 36-kDa precursor, and they determined it to be an asparaginyl endopeptidase. In general, asparaginyl endopeptidases have strict substrate specificity toward a synthetic substrate, Z-Ala-Ala-Asn-MCA, and an examination of substrate specificity showed that REP-2 also efficiently digested it (Table 1). The exopeptidase activity of the REP-2 fraction from hydrophobic chromatography was attributed to exopeptidase coexistence. Furthermore, the antiserum against legumain, detected REP-2 on the nitrocellulose membrane (Fig. 2A), indicating that REP-2 is an asparaginyl endopeptidase that may act in the maturation of REP-1.

We showed that REP-2α and REP-2β have different isoelectric points, although they share an identical sequence on their amino terminal in the mature enzyme. The relationship between these forms is an interesting subject, and one possibility is that REP-2 is transcribed from one gene, with the differences in isoelectric point and molecular mass being derived from a difference in carboxy-terminal processing and/or posttranscriptional modification such as glycosylation. This speculation is supported by the results of Southern blot analysis using pRAN151 cDNA as a probe (Fig. 8). We also attempted to select a cDNA for REP-2 with the information obtained from the amino-terminal sequence of REP-2. This isolated only one cDNA for REP-2, pRAN151 (Fig. 3), and a putative amino acid sequence of REP-2 was determined. The deduced amino acid sequence of REP-2 indicated that there is a potential N-linked glycosylation site at Asn153. Hara-Nishimura et al. (1995) reported that asparaginyl endopeptidases are cleaved, by themselves, on the carboxy-terminal side of Asn residues exposed at the molecular surface, to form the mature enzyme. There are three Asn residues in the hydrophilic carboxy-terminal region of REP-2:

Asn338, Asn355, and Asn411 (Fig. 3). When the autocleavage occurs within the carboxy-terminal Asn338 or Asn355 residue without other posttranscriptional modifications, the putative molecular masses of REP-2 are 39 and 40 kDa, respectively. The difference in the molecular masses of these putative transcribed proteins apparently corresponds to that between REP-2α and REP-2β, as determined by SDS–PAGE analyses.

Senyuk et al. (1998) proposed the separation of asparaginyl endopeptidases into two groups based upon their amino terminal: within-VPE type and another type, where Asp and Asn are the amino-terminal residues, respectively. The amino-terminal amino acid of mature REP-2 was identified as Gly58 following Asp57, coinciding with the result from the phylogenetic analysis (Fig. 4), indicating that REP-2 is a VPE-type legumain.

Expression of REP-2 in maturing and germinating seeds

REP-2 mRNA exists from 15 DAF through to 6 DAI; similarly, both types of REP-2 are expressed in seeds from the middle stage of maturation to germination, suggesting that both forms function at that stage. This result is consistent with that obtained by other groups: Fischer et al. (2000) and Schlereth et al. (2000) reported that the vetch (*Vicia sativa*) seed VPE-type of asparaginyl endopeptidase and its mRNA also exist in cotyledons and axes of maturing and early germinating seeds. In the developing seeds of this study, we observed the gradual disappearance of additional spots that migrate faster than both REP-2 proteins on the SDS–PAGE gel. We could not identify these spots as other legumains, but the profile of their disappearance (as in Fig. 7) coincided with that in the vetch seed as observed by Fischer et al. (2000). The possibility that the additional spots are identical to the other legumains should not be discounted. The differences in the migration and isoelectric points found between the two additional spots are similar to those found between REP-2α and REP-2β. If these spots are not derivatives of REP-2, it is plausible that they are proteins that are subjected to modifications similar to REP-2.

The relation between the two forms of REP-2 is an interesting subject; one hypothesis is that the 39-kDa REP-2α is a digestion product of the 40-kDa REP-2β. When de-embryonated seeds were incubated in the presence of GA$_3$, the level of REP-2β declined rapidly, whereas REP-2α remained at pretreatment levels (Fig. 7A). The reduction in REP-2β may be due to digestion by other GA$_3$-inducible proteinases such as REP-1. If REP-2α is the digestion product of REP-2β, REP-2α is likely to be digested more efficiently as the level of REP-2β decreases. As judged by the Southern blot, both REP-2 proteins are generated from a single gene and their natures are essentially the same, so the difference in the sensitivity to GA$_3$ may reflect differences in posttranslational modification.

Possible role of REP-2

Asparaginyl endopeptidase was first identified as a processing enzyme for the precursor of concanavalin A in jack bean seed (Abe et al. 1993). Hara-Nishimura et al. (1991) showed that the asparaginyl endopeptidase processed several proprotein precursors to the mature form. Rice glutelin, accounting for 80% of the total amount of storage proteins, is known to be transcribed as a long precursor and to process the carboxy-terminal Asn residue like other storage proteins (Takaiwa et al. 1986; Hara-Nishimura et al. 1993). Therefore, detectable asparaginyl endopeptidases in maturing seeds, including both forms of REP-2, may function cooperatively in the processing of glutelin.

The expression profiles of REP-2 in various tissues indicate that REP-2 functions in aleurone layers. REP-1 expresses specifically in aleurone layers of germinating seeds and has a long prosequence like other papain-family proteinases. The prosequence of SH-EP is cleaved to form mature SH-EP by an asparaginyl endopeptidase, VmPE-1 (Okamoto et al. 1999). Previously, Hara-Nishimura et al. (1995) reported that asparaginyl endopeptidases cleaved the carboxy-terminal side of Asn residues that are exposed at the surface of potential substrates. The REP-1 precursor has four Asn residues in the prosequence following the signal peptide (Kato and Minamikawa 1996), all of which are positioned in a hydrophilic region, suggesting that proREP-1 is cleaved by REP-2.

Asparaginyl endopeptidases from kidney and vetch beans were reported to act as a trigger for storage-protein digestion in germinating seeds (Senyuk et al. 1998). We showed that the two REP-2 enzymes exhibit a notable difference in sensitivity to GA$_3$. Senyuk et al. (1998) reported that there are two types of asparaginyl endopeptidase in germinating kidney bean cotyledons, and suggested a difference in the roles of these enzymes in germinating seeds. The expression profiles of these enzymes are strictly different, and they speculated that the form expressed early might act indirectly on the digestion of storage proteins via the processing of other endopeptidases, whereas that expressed later acts directly in the limited digestion of storage proteins. The level of REP-2β was diminished by the addition of GA$_3$, whereas that of REP-2α was unaffected (Fig. 7A). Hence, REP-2β might function in the processing of REP-1 rather than in the limited digestion of storage proteins, and REP-2α might function mainly in the processing of REP-1 in germinating seeds.

References

Abe Y, Shirane K, Yokosawa H, Matsushita H, Mitta M, Kato I, Ishii S (1993) Asparaginyl endopeptidase of jack bean seeds. Purification, characterization, and high utility in protein sequence analysis. J Biol Chem 268:3525–3529

Akasofu H, Yamauchi D, Mitsuhashi W, Minamikawa T (1989) Nucleotide sequence of cDNA for sulfhydryl-endopeptidase (SH-EP) from cotyledons of germinating *Vigna mungo* seeds. Nucleic Acids Res 17:6733

Becker C, Shutov AD, Nong VH, Senyuk VI, Jung R, Horstmann C, Fischer J, Nielsen NC, Müntz K (1995) Purification, cDNA cloning and characterization of proteinase B, an asparagine-specific endopeptidase from germinating vetch (*Vicia sativa* L.) seeds. Eur J Biochem 228:456–462

Bottari A, Capocchi A, Galleschi L, Jopova A, Saviozzi F (1996) Asparaginyl endopeptidase during maturation and germination of durum wheat. Physiol Plant 97:475–480

Chen JM, Dando PM, Rawlings ND, Brown MA, Young NE, Stevens RA, Hewitt E, Watts C, Barrett AJ (1997) Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase. J Biol Chem 272:8090–8098

Davy A, Svendsen I, Sørensen SO, Sørensen MB, Rouster J, Meldal M, Simpson DJ, Cameron-Mills V (1998) Substrate specificity of barley cysteine endopeptidases EP-A and EP-B. Plant Physiol 117:255–261

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Fischer J, Becker C, Hillmer S, Horstmann C, Neubohn B, Schlereth A, Senyuk V, Shutov A, Müntz K (2000) The families of papain- and legumain-like cysteine proteinases from embryonic axes and cotyledons of *Vicia* seed: developmental patterns, intercellular localization and functions in globulin proteolysis. Plant Mol Biol 43:83–101

Hara-Nishimura I, Inoue K, Nishimura M (1991) A unique vacuolar processing enzyme responsible for conversion of several proprotein precursors into the mature forms. FEBS Lett 294:89–93

Hara-Nishimura I, Takeuchi Y, Nishimura M (1993) Molecular characterization of a vacuolar processing enzyme related to a putative cysteine proteinase of *Schistosoma mansoni*. Plant Cell 5:1651–1659

Hara-Nishimura I, Shimada T, Hiraiwa N, Nishimura M (1995) Vacuolar processing enzyme responsible for maturation of seed proteins. J Plant Physiol 145:632–640

Kamphuis IG, Drenth J, Baker EN (1985) Thiol proteases. Comparative studies based on the high-resolution structures of papain and actinidin, and on amino acid sequence information for cathepsins B and H, and stem bromelain. J Mol Biol 182:317–329

Karrer KM, Peiffer SL, DiTomas ME (1993) Two distinct gene subfamilies within the family of cysteine protease genes. Proc Natl Acad Sci USA 90:3063–3067

Kato H, Minamikawa T (1996) Identification and characterization of a rice cysteine endopeptidase that digests glutelin. Eur J Biochem 239:310–316

Kembhavi AA, Buttle DJ, Knight CG, Barrett AJ (1993) The two cysteine endopeptidases of legume seeds: purification and characterizaion by use of specific fluorometric assays. Arch Biochem Biophys 303:208–213

Kinoshita T, Nishimura M, Hara-Nishimura I (1995a) Homologues of a vacuolar processing enzyme that are expressed in different organs in *Arabidopsis thaliana*. Plant Mol Biol 29:81–89

Kinoshita T, Nishimura M, Hara-Nishimura I (1995b) The sequence and expression of the γ-VPE gene, one member of a family of three genes for vacuolar processing enzymes in *Arabidopsis thaliana*. Plant Cell Physiol 36:1555–1562

Koehler SM, Ho T-HD (1990a) Hormonal regulation, processing, and secretion of cysteine endopeptidase in barley aleurone layers. Plant Cell 2:769–783

Koehler SM, Ho T-HD (1990b) A major gibberellic acid-induced barley aleurone cysteine proteinases in barley aleurone layers. Plant Physiol 94:251–258

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685

Manabe T, Tachi K, Kojima K, Okuyama (1979) Two-dimensional electrophoresis of plasma proteins without denaturing agents. J Biochem 85:649–659

Mitsuhashi W, Minamikawa T (1989) Synthesis and posttranslational activation of sulfhydryl endopeptidase in cotyledons of germinating *Vigna mungo* seeds. Plant Physiol 89:274–279

Mitsuhashi W, Koshiba T, Minamikawa T (1986) Separation and characterization of two endopeptidases from cotyledons of germinating *Vigna mungo* seeds. Plant Physiol 80:628–634

Okamoto T, Minamikawa T (1995) Purification of a processing enzyme (VmPE-1) that is involved in post-translational processing of a plant cysteine endopeptidase (SH-EP). Eur J Biochem 231:300–305

Okamoto T, Minamikawa T (1999) Molecular cloning and characterization of *Vigna mungo* processing enzyme 1 (VmPE-1), an asparaginyl endopeptidase possibly involved in post-translational processing of a vacuolar cysteine endopeptidase (SH-EP). Plant Mol Biol 39:63–73

Okamoto T, Yuki A, Mitsuhashi N, Minamikawa T (1999) Asparaginyl endopeptidase (VmPE-1) and autocatalytic processing synergistically activate the vacuolar cysteine proteinase (SH-EP). Eur J Biochem 264: 223–232

Rogers S, Bendich A (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol 5:69–76

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Schlereth A, Becker C, Horstmann C, Tiedemann J, Müntz K (2000) Comparison of globulin mobilization and cysteine proteinases in embryonic axes and cotyledons during germination and seedling growth of vetch (*Vicia sativa* L.) J Exp Bot 51:1423–1433

Schlereth A, Standhardt D, Mock H-P, Müntz K (2001) Stored cysteine proteinases start globulin mobilization in protein bodies of embryonic axes and cotyledons during vetch (*Vicia sativa* L.) seed germination. Planta 212:718–727

Senyuk V, Rotari V, Becker C, Zakharov A, Horstmann C, Müntz K, Vaintraub I (1998) Does an asparaginyl-specific cysteine endopeptidase trigger phaseolin degradation in cotyledons of kidney bean seedlings? Eur J Biochem 258:546–558

Shimada T, Hiraiwa N, Nishimura M, Hara-Nishimura I (1994) Vacuolar processing enzyme of soybean that converts proproteins to the corresponding mature forms. Plant Cell Physiol 35:713–718

Shintani A, Kato H, Minamikawa T (1997) Hormonal regulation of expression of two cysteine endopeptidase genes in rice seedlings. Plant Cell Physiol 38:1242–1248

Sutoh K, Kato H, Minamikawa T (1999) Identification and possible roles of three types of endopeptidase from germinated wheat seeds. J Biochem 126:700–707

Takaiwa F, Kikuchi S, Oono K (1986) The structure of rice storage protein glutelin precursor deduced from cDNA. FEBS Lett 206:33–35

Takeda O, Miura Y, Mitta M, Matsushita H, Kato I, Abe Y, Yokosawa H, Ishii S (1994) Isolation and analysis of cDNA encoding a precursor of *Canavalia ensiformis* asparaginyl endopeptidase (legumain). J Biochem 116:541–546

Vernet T, Khouri HE, Laflamme P, Tessier DC, Musil R, Gour-Salin BJ, Storer AC, Thomas DY (1991) Processing of the papain precursor. Purification of the zymogen and characterization of its mechanism of processing. J Biol Chem 266:21451–21457

von Heijne G (1983) Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem 133:17–21