Molecular Cloning and Characterization of a Cysteine-rich 16.6-kDa Prolamin in Rice Seeds

Norihiro MITSUKAWA, Ryoichi KONISHI, Masato UCHIKI, Takehiro MASUMURA, and Kunsuke TANAKA

Laboratory of Genetic Engineering, Faculty of Agriculture Kyoto Prefectural University, Shimogamo, Kyoto 606-8522, Japan

Received February 15, 1999; Accepted July 26, 1999

An alcohol-soluble storage protein, a 16.6-kDa prolamin found in rice seeds, was purified from both the total protein body and purified type I protein body fractions. The partial amino acid sequences of three tryptic peptides generated from the purified polypeptide were analyzed. A part of the 16.6-kDa prolamin cDNA was amplified from developing seed mRNA by the reverse transcribed polymerase chain reaction using an oligo (dT) primer and a primer which was synthesized based on the partial amino acid sequence. The amplified product was used to isolate the full-length cDNA clone (\( \lambdaRP16 \)) from a developing seed cDNA library. The cDNA has an open reading frame encoding a hydrophobic polypeptide of 149 amino acids. The polypeptide was rich in glutamine (20.0%), cysteine (10.0%), and methionine (6.9%). The cysteine content was higher than those of most other rice storage proteins. Messenger RNA of the 16.6-kDa prolamin was detected in seeds, but not in other aerial tissues.

Key words: seed maturation; rice prolamin; protein body; storage protein; sulfur-rich protein

Rice storage protein is an important protein source for people who eat rice as their staple food. The physiological role of storage proteins is to provide nutrients such as nitrogen and sulfur for germination. The rice starchy endosperm contains more total protein than the embryo and the aleurone layer. Up to 95% of the endosperm protein is deposited in proteinaceous particles called protein bodies (PBs). Two types of PBs, type I protein bodies (PB-Is) and type II protein bodies (PB-IIs), have been isolated, and it has been shown that these are mainly prolamins and glutelins, respectively. PB-I proteins account for 18–35% of all endosperm protein. Most of the PB-I proteins are indigestible for people because the heat-denatured PB-Is passes through the human digestive tract as discrete particles. As a prerequisite for analyzing the protein composition of PB-Is and try to find the cause of their low digestibility.

PB-I is enriched by a microheterogeneous set of alcohol-soluble proteins, prolamins, which differ from each other in their biochemical characteristics. This microheterogeneity has impeded isolation of these proteins as single homogeneous polypeptides, but nevertheless the amino acid sequences of some PB-I proteins have been reported. PB-I proteins separate into a number of bands with apparent molecular sizes of 10–17 kDa by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), but the designations for the individual proteins differ among researchers. Ogawa et al. defined the separated bands as “16-kDa prolamins, 13a prolamins, 13b prolamins, and 10-kDa prolamins” in order of size on SDS-PAGE. In this paper, we will use these designations, and also use “13-kDa prolamins” defined by Tanaka et al. when it is hard to distinguish between 13a or 13b prolamins. The 16-kDa polypeptides can be separated into three, and the 13-kDa polypeptides into 4 or 5 spots by two-dimensional electrophoresis.

It has been shown that the 10-kDa and some 16-kDa prolamins are rich in sulfur-containing amino acids. Yamagata et al. found that S-methionine was incorporated by the 10-kDa and some 16-kDa prolamins during seed development. Hibino et al. reported that the amounts of cysteine and methionine in the 10-kDa and one 16-kDa prolamins are higher than those in the 13-kDa prolamins. A thiol-specific fluorescent reagent can react with the 10-kDa prolamin; some 16-kDa and 13a prolamins can be labeled with it, but not 13b prolamins. The 10-kDa prolamin encoded by \( \lambdaRP10 \) cDNA is rich in sulfur-containing amino acids with a composition of 10% cysteine and 20% methionine. Other rice prolamin cDNAs, \( \lambdaRM7 \), pS18, and pProl 17, encode proteins with sulfur-containing amino acids. The 13-kDa prolamins encoded by \( \lambdaRM7 \) and pS18 cDNAs contain about 3% cysteine but almost no methionine. pProl 17 cDNA encodes a prolamin containing about 5% cysteine and 3% methionine, but there has been no report of purification or amino acid sequencing of any corresponding protein. The possibility that

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Database under the accession number D88210.

* To whom correspondence should be addressed. Phone and Fax: +81-75-703-5675; E-mail: k_tanaka@kpu.ac.jp

** Present address: Forestry Research Institute, Oji Paper Co., Ltd., Noibono 24-9, Kameyama, Mei 519-0212, Japan

Abbreviations: ABD-F, 4-(aminosulfonyl)-7-fluoro-2,1,3-benoxadiazole; CBB, Coomassie Brilliant Blue; DAF, days after flowering; FPLC, Fast Protein Liquid Chromatography; PBs, protein bodies; 2-ME, 2-mercaptoethanol; PB-Is, type I protein bodies; PB-IIs, type II protein bodies; PCR, polymerase chain reaction; SDS-PAGE, SDS-polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid; TBP, tri-n-butyl phosphate
this gene product is a 16-kDa prolamin cannot be excluded, but to date there have been no unambiguous reports of cDNA isolation and purification of any sulfur-rich 16-kDa prolamin. Therefore, to illuminate its biochemical properties, it is necessary to purify a sulfur-rich 16-kDa prolamin and identify the corresponding cDNA.

In this study, we report the purification of one of the 16-kDa prolamins, 16.6-kDa prolamin, and the isolation of a corresponding cDNA. The 16.6-kDa prolamin is shown to be a novel protein that is rich in sulfur-containing amino acids, 10% cysteine and 7% methionine. We found that the 16.6-kDa prolamin is similar in sequence to proteins encoded by previously isolated rice prolamin cDNAs including the 10-kDa prolamin cDNA, which suggests that all isolated rice prolamin cDNAs have evolved from a single ancestral gene. Moreover, the rice 16.6-kDa prolamin sequence is highly similar to oat avenins and other cereal prolamins, especially with regard to conserved cysteine residues.

Materials and Methods

Extraction and fluorescent cysteine labeling of 16-kDa prolamins from the PB-I fraction. The PB-I fraction was prepared from rice seeds (Oryza sativa, L. Japonica cv. Nipponbare) by the method of Ogawa et al. with some modifications. Polypeptides were extracted from the PB-I with UMS buffer (62.5 mm Tris-HCl, pH 7.0, 4 M urea, 2% (v/v) 2-mercaptoethanol [2-ME], and 2% (w/v) SDS) or U buffer (62.5 mm Tris-HCl, pH 7.0, and 4 M urea). After shaking, suspensions were centrifuged at 15,000 × g for 10 min. The residue of extraction with U buffer was re-extracted with UMS buffer (62.5 mm Tris-HCl, pH 7.0, 4 M urea, and 2% (v/v) 2-ME). Equal volumes of ethanol containing 2% (v/v) 2-ME were added to the extractants. Ethanol soluble fractions were run on SDS-PAGE, after resuspending in R buffer (62.5 mm Tris-HCl, pH 7.0 with 4 M urea, and 10 mm tri-n-butyl phosphate [TBP], and 2% (w/v) SDS).

Fluorescent cysteine labeling of protein fractions by reaction with 4-(aminosulfonyl)-7-fluoro-2,1,3-benzoxadiazole (ABD-F, Dojin) was done by the method previously described by Kirby. Reactions were done in 50 mm Tris-HCl, pH 7.0, 4 M urea, 2% (w/v) SDS, 15% (v/v) DMSO, 8 mm ABD-F in the presence or absence of 2 mm TBP. The reaction mixtures were separated by SDS-PAGE with no additional reductant. Fluorescently labeled polypeptides in the gels were detected under UV illumination (365 nm) and photographed through a red filter using Polaroid film (No.667).

Purification and amino acid sequencing of the 16.6-kDa prolamin. Rice seeds harvested at the middle-developing stage (on average, 14 days after flowering [DAF]) were homogenized in 50 mm Tris-HCl, pH 7.5, 50 mm potassium acetate, 5 mm magnesium acetate, 5 mm 2-ME, and 0.5 M sucrose. The total PB fraction was precipitated by centrifugation at 2,000 g, and defatted with acetone. Prolamin polypeptides were extracted from the defatted total PB fraction with 20 mm Tris-HCl, pH 8.0, 55% (w/v) l-propanol, and 0.2% (v/v) 2-ME. The prolamin fraction was precipitated by evaporation of l-propanol in vacuo and washed with distilled water. The 16-kDa polypeptides were prepared from the prolamin fraction by preparative SDS-PAGE. After electrophoresis, polypeptides were stained with 0.3 μM CuCl2. The pieces of 16-kDa polypeptide band were excised from the gels; the protein was eluted in 20 volumes of 10 mm Tris-HCl, pH 7.0, 0.1% (w/v) SDS, and 0.2% (v/v) 2-ME with shaking. The eluted polypeptides precipitated during desalting by ultrafiltration using Centriprep-10 (Amicon).

S-pyridylethylated of cysteine residues in the 16-kDa polypeptides by 4-vinylpyridine was done as described previously. The alkylated polypeptide was further fractionated by the Fast Protein Liquid Chromatography (FPLC) system using a preRPC HR5/10 column (Pharmacia). The polypeptides were eluted with a linear acetonitrile gradient from 0% to 80% in 0.1% trifluoroacetic acid (TFA) at a flow rate of 0.3 ml/min.

The protein in the major peak was digested with trypsin (Promega) as described. Tryptic peptides were isolated by the FPLC system using a pepRPC HR5/5 column (Pharmacia). The peptides were eluted with a linear acetonitrile gradient from 0% to 100% in 0.1% TFA at a flow rate of 0.5 ml/min. Two peptide peaks (T1 and T2) were further purified by re-chromatography.

The 16.6-kDa prolamin was also purified from the PB-I fraction. Proteins soluble in 4 M urea were removed from the PB-I fraction by extraction with U buffer. The residue was extracted with UM buffer. Equal volumes of ethanol containing 2% (v/v) 2-ME were added to the extractants. The ethanol soluble fraction was concentrated by ultrafiltration. The purified polypeptide was digested with trypsin as described above. Tryptic peptides were separated by the Tricine SDS-PAGE system using 10% gels. The separated peptides were blotted to a PVDF membrane (Bio-Rad), and one piece of the peptide (T3) was excised from the blotted membrane.

The peptides (T1, T2, and T3) were sequenced using a model 477A protein sequencer (Applied Biosystems). The amounts of proteins were estimated by the Bradford method using bovine γ-globulin as the standard.

Amplification and screening of a 16.6-kDa prolamin cDNA. The oligonucleotide primer 5′-CAICAITGIGICAICA1ATIG3′ was synthesized according to the amino acid sequence of a tryptic peptide from 16.6-kDa prolamin, Gln-Gln-Cys-Cys-Gln-Gln-Met-Arg, as a sense primer. Using a combination of the sense primer and an oligo (dT)18 primer, a 16.6-kDa prolamin cDNA fragment was isolated by the polymerase chain reaction (PCR). The amplified fragment (P16) was blunt-ended and cloned into the Smal I site of the plasmid vector pBluescript KS− (Stratagene).

To obtain the full-length rice 16.6-kDa prolamin cDNA clone, a Agt11 cDNA library prepared from developing rice seeds was screened with the [α-32P] dCTP P16 fragment as probe. Phage plaques were blotted onto Hybond-N nylon membranes (Amersham), the membranes were hybridized with the probe for 16 h at
42°C in 50% formamide, 6 × SSC, 0.1% SDS, 5 × Denhardt’s solution, and 200 μg/ml denatured salmon sperm DNA. Then, the membranes were washed twice with 0.3 × SSC, 0.1% SDS at 62°C and autoradiographed. Inserts of positive plaques were sub-cloned into the EcoRI I site of pBluescript KS⁺.

**DNA sequencing analysis.** Nucleotide sequence analysis of the cloned DNA was done using a Taq Dye Primer Cycle Sequencing kit with a 373A DNA sequencer (Applied Biosystems). Nucleotide and amino acid sequence data were analyzed using the GENETYX (SDC), as well as a multiple sequence alignment program, CLUSTAL W,⁵ and a phylogenetic tree drawing program, TREEVIEW,⁶ and a signal peptide prediction program, SignalP Ver.1.0.⁷

**Northern and Southern blot analyses.** Total RNA was prepared by the SDS-phenol method.⁴ For Southern blotting, genomic DNA was prepared from rice germ as described,⁵ then digested with restriction enzymes (BamHI, EcoRI, EcoRV, and Hind III). Total RNAs and digested DNAs were separated on a 1.2% (w/v) agarose gel containing formaldehyde and a 0.8% agarose gel, then blotted onto Hybond-N membranes (Amersham). The membranes were hybridized with 32P-labeled ARP16 cDNA in 50% formamide, 6 × SSPE, 5 × Denhardt’s solution, 0.5% SDS, and 100 μg/ml salmon sperm DNA at 42°C, and washed as described in figure legends.

**Results**

**Heterogeneity of the 16-kDa polypeptides in the rice PB-I fraction**

To confirm the heterogeneity of the 16-kDa polypeptides in their solubilities under different reducible conditions, we extracted polypeptides from the PB-I fraction with non-alcoholic solutions containing 4 M urea in the absence (U buffer) or presence (UM buffer) of 2% 2-ME. The PB-I fraction was purified from the total PB fraction of rice seeds by centrifugation and pepsin digestion. SDS-PAGE analysis of proteins from the PB-I fraction in the reducing condition (R buffer) showed that the fraction contained two 16-kDa polypeptide groups having molecular masses of 16.3 and 16.6 kDa as well as 13-kDa polypeptides (Fig. 1A, lane 1). These 16-kDa polypeptide groups were also soluble in 50% ethanol containing 2 M urea.

The 16.3-kDa polypeptides were extracted with the U buffer containing 4 M urea. The 16.3-kDa polypeptides further separated into two bands having apparent molecular masses of 16.4 and 16.1 kDa when these polypeptides were resolved by SDS-PAGE in the nonreducing conditions (Fig. 1B, lane 1). After reduction with

---

**Fig. 1.** Extraction and Characterization of 16-kDa Polypeptides from the Rice PB-I Fraction.

(A) Sequential extraction of polypeptides from the rice PB-I fraction. Polypeptides were extracted from the PB-I fraction with the U buffer (62.5 mM Tris-HCl, pH 7.0, 4 M urea, 2% (v/v) 2-ME and 2% (w/v) SDS) (lane 1) and the UM buffer (62.5 mM Tris-HCl, pH 7.0, 4 M urea and 2% (v/v) 2-ME) (lane 2). 16.6-kDa polypeptides (lane 3) were extracted with the UM buffer (62.5 mM Tris-HCl, pH 7.0, 4 M urea, and 2% (v/v) 2-ME) from the residue of extraction with U buffer. The extracted polypeptides were reduced with TBP, and separated by SDS-PAGE using a 14% (w/v) gel. After separation, a gel was stained by Coomassie Brilliant Blue (CBB) R-250. Open and closed arrowheads indicate a 16.6-kDa and a 16.3-kDa polypeptide, respectively. (B) Fluorescent cysteine labeling of 16-kDa polypeptides with ABD-F. The 16-kDa polypeptides were extracted with the U buffer (lanes 1, 2, 3, and 5) or the UM buffer (lanes 4 and 6). The extracted polypeptides were reacted with 2 mM TBP (lane 2) or, 8 mM ABD-F and 2 mM TBP (lanes 3, 4, 5, and 6). After separation by SDS-PAGE using a 14% (w/v) gel, the gel was photographed under UV light (right) and stained by CBB R-250 (left). Open arrowheads indicate the 16.6-kDa prolamin used for internal amino acid sequencing.
TBP these polypeptides migrated to the same position to form a single protein band (Fig. 1B, lane 2). These polypeptides were strongly labeled with ABD-F, which reacts specifically to thiols (Fig. 1B, lane 3 and 5).

After extraction with U buffer containing 4 M urea, the residual proteins were separated with UM buffer containing 4 M urea and 2% 2-ME. SDS-PAGE analysis and ABD-F labeling of the extracted proteins showed that 16.6-kDa polypeptides can be extracted only in the presence of reductant such as 2-ME (Fig. 1A, lane 3), and can be labeled with ABD-F weaker than 16.3-kDa polypeptides (Fig. 1B, lane 6). The cause of weak labeling may be insufficient removal of 2-ME which competes with ABD-F in the labeling reaction.

**Purification and partial amino acid sequencing of the 16.6-kDa prolamin**

To get amino acid sequences for cDNA screening, the 16.6-kDa prolamin was purified from the total PB fraction of developing seeds. The total PB fraction was prepared from homogenates of developing seeds by centrifugation, and then used for extraction of the prolamin fraction with 55% 1-propanol in the presence of 2-ME. About 30 mg of prolamin fraction was extracted from 100 g of rice seeds with the alcoholic solution. SDS-PAGE analysis showed that the prolamin fraction consisted of three polypeptide groups with apparent molecular masses of 16.6 kDa, 14.3-15.0 kDa, and 11.5 kDa (Fig. 2A). The 16.6-kDa polypeptide band was excised from preparative SDS-PAGE gels. After preparative SDS-PAGE, 2.90 mg of the polypeptides was obtained. For further purification of the 16.6-kDa prolamin, 0.50 mg of the polypeptide was S-pyridylethylated, and then separated by reverse-phase column chromatography. The major peak and two minor shoulder peaks were eluted from the column (Fig. 2B). The SDS-PAGE analysis showed that the major peak was a single polypeptide (Fig. 2A). The chromatography yielded 0.30 mg of the 16.6-kDa prolamin, which is about 60% of the alcohol-soluble 16-kDa polypeptides.

Tryptic fragments of the 16.6-kDa prolamin separated by reverse-phase column chromatography (Fig. 2C) were sequenced, because the NH₂-terminal amino acid of the prolamin was blocked. Amino acid sequences of two tryptic peptides showed that the sequence of T1 and T2 peptides are Tyr-Cys-Ser-Thr-Pro-Cys-Lys and Glu-Gln-Cys-Cys-Gln-Gln-Met-Arg.

To confirm localization of the 16.6-kDa prolamin in PB-I, a 16-kDa polypeptide was purified again from the PB-I fraction prepared by pepsin digestion of total PBs. The protein soluble in the U buffer containing 4 M urea was removed from the PB-I fraction. The 16.6-kDa polypeptide was extracted with the UM buffer containing 4 M urea and 2% 2-ME from the residual PB-I fraction. The large-scale extraction from 200 g of seeds yielded 7.5 mg of the 16.6-kDa polypeptide. The purified polypeptide was digested with trypsin. The tryptic peptides of the polypeptide were resolved by Tricine SDS-PAGE, then blotted onto a PVDF membrane. The NH₂-terminal sequence of the 5.7-kDa tryptic peptide (T3) was Leu-Met-Ala-Gln-Gln-Xaa-Xaa-Cys-Xaa-Ala-Ile-Xaa-Xaa-Met.
Isolation and characterization of the 16.6-kDa prolamin cDNA clone

The 16.6-kDa prolamin cDNA clones were isolated from a cDNA library prepared from developing rice seeds. Oligonucleotide primers were synthesized based on the tryptic peptide sequence for PCR amplification (see Materials and Methods). The amplified DNA fragment was used as a probe to screen the cDNA library. Out of 40,000 recombinant plaques, 21 positive clones were isolated. The cDNA insert of the longest clone, λRP16, was subcloned into the plasmid vector pBluescript KS+ and sequenced. The cDNA sequence encodes a protein of 149 amino acids with a molecular mass of 16,674 Da. The computer program Signalp Ver. 1.0 predicted the location of the signal peptide sequence to be from Met1 to Ala19. The molecular mass of the mature protein was 14,785 Da. The internal sequences of the purified 16.6-kDa prolamin were identified at Gln72, Arg80 (T2), Leu81-Met84 (T3), and Tyr134-Lys140 (T1) (Fig. 3).

A striking feature of the mature protein was the unusually high content of the sulfur-containing amino acids, Met (7%) and Cys (10%). The protein was also rich in Gln (20%), Pro (7%), and Ala (7%) similar to other cereal prolamins. Hydrophilicity analysis showed that the 16.6-kDa prolamin was slightly more hydrophilic than the 13-kDa prolamin encoded by the λRM7 cDNA (data not shown).

Figure 3 shows there is a high degree of similarity between the 16.6-kDa prolamin and rice prolams (36.4-56.4% amino acid identities) and an oat prolamin, avenin encoded by the pAv100 (44.6% amino acid identities). Comparison of the amino acid sequence of the 16.6-kDa prolamin and that of the avenin indicated that the eight cysteine residues were conserved. The alignment revealed that the signal sequence of the 16.6-kDa prolamin shared a high degree of similarity with those of rice prolams, oat avenin, and maize 10-kDa and 15-kDa zeins. In contrast, the mature sequences of the rice 16.6-kDa prolamin and zeins showed lower similarity than the signal sequences.

A phylogenetic tree of prolamins is shown in Figure 4. The 10-kDa prolamin was placed closer to the 16.6-kDa prolamin than to 13-kDa prolamins in the tree. Comparison of amino acid sequences of the 10-kDa prolamin and the 16.6-kDa prolamin reveals higher similarity (36.4%) than between those of the 10-kDa prolamin and rice 13-kDa prolamins (25.8-34.3%).

Northern blot analysis

Messenger RNA of the 16.6-kDa prolamin was detected in developing and matured seeds (Fig. 5). No

---

Fig. 3. Comparison of Deduced Amino Acid Sequences of the Rice 16.6-kDa Prolamin and Cereal Prolamins.

Alignment of deduced amino acid sequences from the rice 16.6-kDa prolamin, λRP16 (this study), oat avenin, pAv100 (M38721), rice 13-kDa prolamin, pProl17 (M23745), λRM720 (X14392) and pProl14 (M23744), rice 10-kDa prolamin, λRP10 (X15231), maize 10-kDa zein, 10k1-10 (X07355) and maize 15-kDa zein, cZ15A3 (M12147). GenBank accession numbers of the sequences are indicated in parentheses. Gray boxes indicate regions in which the residues are identical to the sequence of the 16.6-kDa prolamin, and solid boxes indicate conserved cysteine residues. Dashes were introduced to optimize alignment. The arrowhead shows the expected signal peptide recognition sites. The internal amino acid sequences found from tryptic peptides (T1, T2, and T3) of purified 16.6-kDa prolamin are double underlined.
Fig. 4. Unrooted Phylogenetic Tree Inferred from Amino Acid Sequences of Prolamins.

The phylogenetic tree was constructed with the deduced amino acid sequences from the following sequences; rice 16.6-kDa prolam in, \( \lambda \)RP16 (this study), oat avenin, pAv10 (M38721), rice 13-kDa prolamin s, pProl 17 (M23745), \( \lambda \)RMT7 (X14392) and pProl 14 (M23744), rice 10-kDa prolam in, \( \lambda \)RP10 (X15231), maize 10-kDa zein, 10kz-1 (X07535) and maize 15-kDa zein, cZ1A3 (M12147). The amino acid sequences were aligned using the multiple alignment program Clustal W. Based on this alignment, a phylogenetic analysis was done by the neighbor-joining method and a phylogenetic tree was drawn using the software TREEVIEW. The scale bar indicates a divergence of 0.1 amino acid substitution per site.

Fig. 5. Northern Blot Analysis of Rice Tissues Probed with 16.6-kDa Prolamin cDNA.

Total RNA was isolated from developing seeds 8, 14, 22, and 28 DAF, matured seeds, germs, etiolated leaves, and callus. The RNA gel blots containing 5 μg of these samples per lane were hybridized with \(^{32}P\)-labeled insert of the \( \lambda \)RP16 cDNA clone and washed with 3× SSC and 0.1% SDS at 42°C.

hybridizing band was seen in leaves, etiolated leaves, stems, germs, or callus. The mRNA band (about 900 nt) was seen in early stage seeds (8 DAF) at a low level. The amounts of the mRNA increased until 22 DAF, and remained at high levels until the mature stage (Fig. 5). Expression of the gene was correlated with the accumulation of 16-kDa prolam in protein.

Fig. 6. Southern Blot Analysis of the 16.6-kDa Prolamin in Rice Genomic DNA.

Samples of 2 μg of rice germ DNA were digested with EcoRI, EcoRV, Hind III and Xba I, and separated on a 0.8% agarose gel. The DNA gel blot containing digested DNAs was hybridized with \(^{32}P\)-labeled insert of the \( \lambda \)RP16 cDNA clone and washed with 0.1× SSC and 0.1% SDS at 62°C.

**Genomic Southern blot analysis**

The \(^{32}P\)-labeled 622-bp fragment of the \( \lambda \)RP16 cDNA probe hybridized to an EcoRV DNA fragment of about 17 kbp, while two or three bands were found after DNA digestion with EcoRI, Hind III, and Xba I (Fig. 6). Using primers with end-sequences of the \( \lambda \)RP16 cDNA, fragments of about 600 bp were amplified from rice genomic DNA by PCR (data not shown). These results indicate that the prolam in gene has no introns. The gene copy number of the 16.6-kDa prolam in is two or three per haploid genome because the cDNA sequence has no EcoRI, Hind III, or Xba I sites.

**Discussion**

We purified the 16.6-kDa prolam in from rice PB-Is and isolated a cDNA clone encoding this protein. The 16.6-kDa prolam in is a sulfur-rich protein that contains 10% cysteine and 7% methionine, and must account-at least in part—for the strong \(^{35}S\)-methionine labeling of the 16-kDa prolamin s observed by Yamagata et al. In contrast, the 16-kDa prolam in reported by Hibino et al. contains 1.3% cysteine and 3.8% methionine and therefore is clearly distinct from the 16.6-kDa prolam in amino acid compositions reported here. The 16-kDa prolam in was purified by preparative SDS-PAGE and isoelectric focusing, and its amino acid composition was analyzed. Our extraction conditions included a reductant, enabling extraction of the 16.6-kDa prolam in into an alcoholic solution. We found that the 16.6-kDa prolam in is soluble only in the presence of a reductant such as 2-ME. The amino acid composition of the 16-
kDa prolamin reported by Hibino et al.\textsuperscript{15} is close to that of the NaCl-soluble protein Rf7.\textsuperscript{30} The NH\textsubscript{2}-terminal sequence of Rf7 is similar to that of the allergenic protein RA5.\textsuperscript{30} We also isolated a 16-kDa polypeptide by preparative SDS-PAGE from a fraction containing PBs and endoplasmic reticulum. The NH\textsubscript{2}-terminal amino acid sequence of this protein is similar to the NH\textsubscript{2}-termini of proteins Rf7 and RA5 (Mitsukawa et al. unpublishe results). We demonstrated that at least three polypeptides can be distinguished among the 16-kDa proteins in PB-Is on the basis of solubility and migration on SDS-PAGE. In this study we purified and characterized this novel 16-kDa prolamin, but there remains no report on the purification of the third 16-kDa prolamin.

The 16.6-kDa prolamin is the first protein purified from PB-Is whose internal sequences were analyzed. By isolating PB-Is before carrying out the alcoholic extraction step, we were able to clarify the localization of the protein in vivo. The 16.6-kDa prolamin was purified from the total PBs fraction, but in order to confirm its presence specifically in PB-Is it was purified again, this time from PB-Is prepared by pepsin digestion of total PBs. The amino acid sequence of the re-isolated protein was almost identical to the amino acid sequence deduced from the 16.6-kDa prolamin cDNA. The prolamins for which amino acid sequences have previously been reported were purified directly from developing seeds,\textsuperscript{39} from a mixture of aleurone layer and starchy endosperm,\textsuperscript{10} or from starchy endosperm.\textsuperscript{8} It was necessary to establish the PB-I localization of these previously reported prolamins retrospectively by direct extraction from PB-Is.

Our sequence analysis of the 16.6-kDa prolamin cDNA clone clarifies the phylogenetic relationship among previously cloned rice prolamins and other cereal prolamins. The 16.6-kDa prolamin shows similarity (35.2–44.6\% amino acid identities) to oat avenins.\textsuperscript{29} Like oat avenins, rice 16.6-kDa prolamin has 3 non-repetitive sequence domains that are conserved among cereal sulfur-rich proteins.\textsuperscript{39} The previously isolated rice prolamins show only low sequence similarity with other cereal prolamins and so were thought to have evolved from a separate ancestral gene.\textsuperscript{16} The 16.6-kDa prolamin fills in the evolutionary gaps between the rice prolamins and the other cereal prolamins, as well as between the 10-kDa prolamin and the other rice prolamins.

The low solubility of 16.6-kDa prolamin may be attributed to intermolecular disulfide bonds that in effect polymerize the individual polypeptides into much larger aggregates. All eight cysteine residues in the avenin that are conserved in 16.6-kDa prolamin form intramolecular disulfide bonds,\textsuperscript{39} and so the avenin is soluble in alcholic solutions lacking reductants. 16.6-kDa prolamin has four cysteine residues in addition to the eight conserved residues. These four cysteine residues may be involved in inter-molecular disulfide bonds because the amount of 16.6-kDa prolamin extracted from PB-Is increases in proportion to the concentration of added reductant.\textsuperscript{16} Amounts of 16-kDa and 10-kDa prolamins in seeds of a rice mutant line (CM1675)\textsuperscript{32} were decreased. Electron microscopy analysis indicated that PB-Is of the mutant lack the lamellar structure that is found in PB-Is of wild type. One of the 16-kDa prolamins acting together with 10-kDa prolamin may form that structure.\textsuperscript{13} However, PB-Is of the mutant and wild type have no great difference in susceptibility to pepsin digestion, suggesting that a partial decrease in sulfur-rich prolamin alone is insufficient to improve PB-I digestibility.\textsuperscript{13} High hydrophobicity is another potential cause of PB-I low digestibility. As a supplemental or alternative route to improving PB-I digestibility, we will try to make transgenic rice expressing chimeric genes encoding prolamin fused with hydrophilic proteins such as bovine casein.

Acknowledgments
We are grateful to Dr. Daisuke Shibata (Kazusa DNA Research Institute) for his stimulating discussion and suggestions. We are indebted to Dr. Robert F. Whittier (Amersham Pharmacia Biotech) for his critical reading of the manuscript. The work was supported by grants from the Ministry of Education, Science, and Culture (Japan), and the “Research for the Future” Program of the Japan Society for the Promotion of Science.

References


