J. Korean Agric. Chem. Soc.
한국농화학회지 32(1), 64—72(1989)

Molecular and Cellular Studies of Seed Storage Proteins from Rice and Wheat

Woo-Taek Kim

Institute of Biological Chemistry, Washington State University,
Pullman, WA99164-6340, U.S.A.

Abstract

Near full length cDNA clones encoding the rice seed storage protein, prolamine, were isolated and divided into two homology classes based on cross-hybridization and DNA sequencing analysis. These cDNA clones contain a single open reading frame encoding a putative rice prolamine precursor(M.W.=17,200) possessing a typical 14 amino acid signal peptide. Clones of these two homology classes diverge mainly by insertions/deletions of short nucleotide stretches and point mutations. The deduced primary structures of both types of prolamine polypeptides are devoid of any major tandem repetitive sequences, a feature prevalent in other cereal prolamines. No significant homology was detected between the rice prolamine and other cereal prolamines, indicating that the rice gene evolved from a different ancestor that gave rise to other cereal prolamine genes. Developing wheat and rice endosperms were examined using ultrathin sections prepared from tissues harvested at various days after flowering. By immunocytochemical localization techniques, wheat prolamines are localized within vesicles from Golgi apparatus and in homogeneous regions of protein bodies. The involvement of the goli apparatus in the packaging of wheat prolamines into protein bodies indicates a pathway which differs from the mode of other cereal prolamines and resembles the mechanism employed for the storage of rice glutelin and legume globulins.

Introduction

Prolamines, typified by their solubility in alcohol solutions, are the major seed storage proteins in most of the cereals. These proteins accumulate during endosperm development and serve as a source of a nitrogen, carbon, and generally sulphur for the young developing seedling. The rice prolamines have mol sizes of about 12~17kD and, as seen for other cereal prolamines, contain a high mole percentage of glutamine residues and low levels of lysine, histidine, cysteine, and methionine.[1,2] They are initially synthesized around 10 DAF[3,4] and are deposited in protein bodies formed by direct dilation of the rough ER lumen.[5] SDS-PAGE analysis of *in vitro* translation products purified by immunoprecipitation using a rice prolamine antibody revealed the synthesis of a 16kD precursor form presumably containing a signal peptide.[6,7]

Recently we showed that the rice prolamines are immunologically distinct from other cereal prolamines.[8] DNA sequence analysis of a single near full length prolamine cDNA clone revealed

Received Jan. 15, 1989
A Paper Presented at the Annual Meeting of the Korean Agricultural Chemical Society, Kyung Sang University, November 19, 1988.

that the derived primary sequence of the rice prolamine did not exhibit significant homology to prolamines from the major cereals.[9,10] In this study, we report that the rice prolamines are encoded by at least two gene classes whose respective mRNA transcripts differ by about 3 ~4 fold in abundance levels during the early stages of endosperm development. Southern blot analysis of rice DNA and sequence data of recombinant DNA clones suggest that the rice prolamines are encoded by complex family of genes.

Isolation and Classification of Prolamine cDNA Clones

By antibody screening of a cDNA lambda gt 11 library, several putative prolamine cDNA clones were obtained. Restriction enzyme digests of DNAs isolated from these clones indicated that all, except pProl 7, contained inserts of about 400~500bp, approximately 50% of the mRNA transcript size as estimated by Northern blotting(data not shown). To obtain full length inserts representative of the prolamine genes, a second cDNA library was constructed using the methods of Heideker and Messing.[11] From analysis of about 500 cDNA clones with a radiolabeled prolamine cDNA insert, 9 positive polamine clones were obtained. Subsequent analysis by restriction enzyme digestion and agarose gel electrophoresis revealed that all of the clones contained inserts of at least 700bp in length. Two clones, pProl 14 and pProl 17, were 870 bp in length and were estimated to contain almost all of the sequence information of the prolamine transcript.

These near full length cDNA clones were further characterized by physical restriction endonuclease mapping(Fig. 1). Single restriction sites for FokI, KpnI, SphI, PstI, XbaI, and EcoRV were evident at the same relative position for both pProl 14. In contrast, pProl 17 shared only the unique PstI site and lacked sites for KpnI, SphI and XbaI. Identical analysis of the remaining recombinant DNA clones indicated that they could be classified either with pProl 7 or pProl 17 by these criteria.

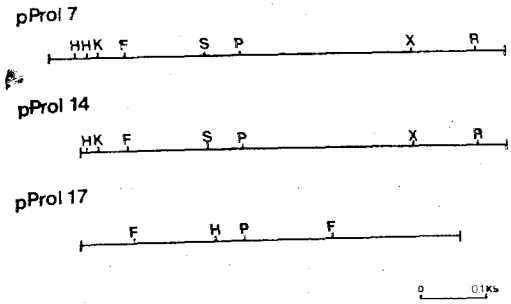

Fig. 1. Restriction enzyme map analysis of three near full length prolamine cDNA clones: F, FokI; H, HaeII; K, KpnI; P, PstI; S, SphI; R, EcoRV; X, XbaI

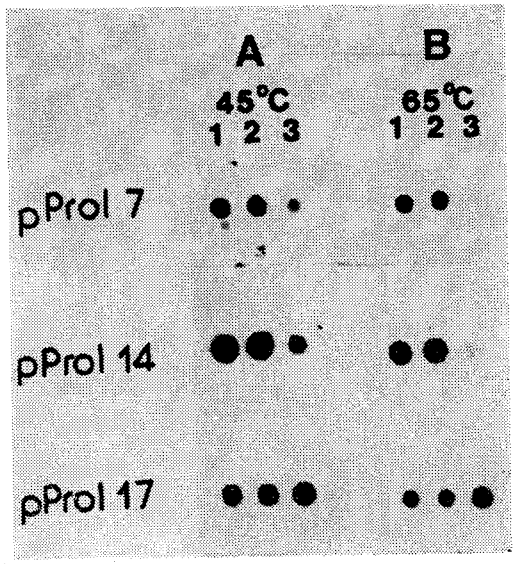

Fig. 2. Dot blot hybridization analysis of the Prolamine cDNA sequences. 250ng of plasmid DNAs containing prolamine cDNA inserts was immobilized onto Zeta-Probe membrane filters and hybridized to ^{32}P-labeled prolamine cDNA inserts according to the manufacturer's recommendations. Filters were washed at 45°C(panel A) or 65°C(panel B) with 0.2×SSPE and 0.1% SDS: (1)pProl 7, (2) pProl 14, (3) pProl 17. The cDNA probes utilized are indicated at the left

The extent of homology among cDNA clones

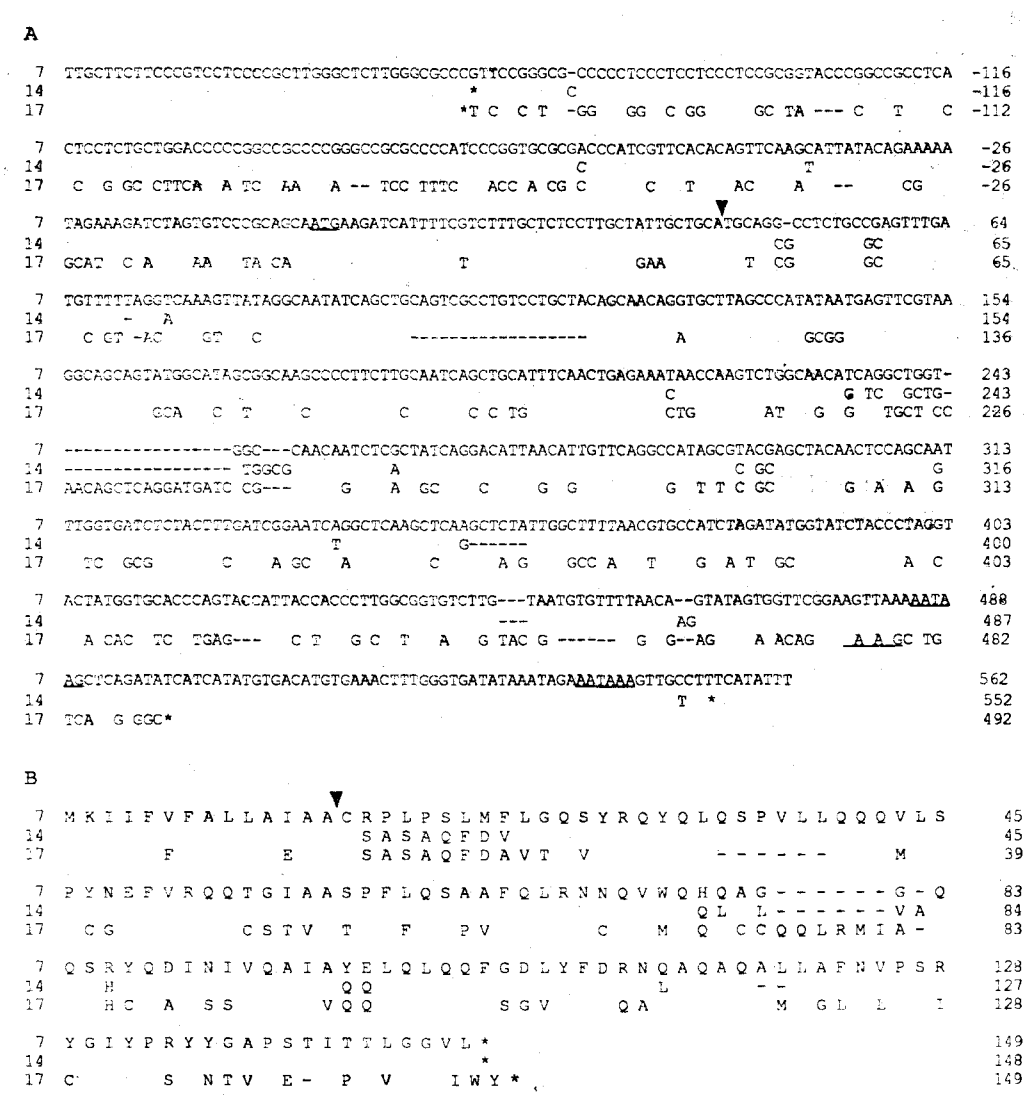

Fig. 3. Nucleotide(panel A) and deduced amino acid sequences(panel B) of prolamine cDNA clones: (7) pProl 7, (14) pProl 14, (17) pProl 17. Only the differences in the nucleotide and amino acid sequences among the cDNA clones are shown. Numbers on the right are in bp(panel A) or amino acid residues(panel B) relative to the translational start site. The putative translational *initiation and polyadenylation signals are underlined. Dotted lines indicate gaps to align the* nucleotide and amino acid sequences. The 5′-and 3′-ends of the cDNA rclones are indicated by asterisks. The arrowhead represents the putative cleavage site of the signal peptide

was examined by the dot blot hybridization technique. Purified, randpm-primed cDNA inserts were hybridized at 45°C to Zeta-Probe blotting membrane filters containing 250ng of each plasmid and two types of cross-hybridization pattern were obtained. Figure 2 represents the dot blot cross-hybridizations of the three near full length cDNA clones. When membrane filters were washed at 45°C after hybridization, pProl 7 and pProl 14 inserts cross-hybridized extensiyely to each other, but these two inserts hybridized only weakly to pProl 17(Panel A, lane 1 and 2). When the insert of pProl 17 was used as a probe, strong hybridization signals were

obtained to its own DNA but cross-hybridization to pProl 7 and pProl 14 occurred to a lesser degree(lane 3). When these filters were washed at 65°C, these differences in crosshybridization patterns were even more evident(Fig. 2, Panel B). Overall, these results together with those obtained by physical mapping of restriction enzyme sites indicate that the prolamines are encoded by at least two classes of genes.

Structure and Primary Sequence of Rice Prolamine cDNA

DNA sequence analysis of pProl 14 and pProl 17 revealed that these two prolamine transcripts contain a single open reading frame encoding 148 and 149 amino acids, respectively(Fig. 3). Both pProl 14 and pProl 17 possess a relatively long 5-untranslated region of about 160 and 150 bases in length, respectively, followed by coding regions of 444~447bp. The nucleotide sequence of the 5-untranslated region of pProl 14 is identical to the analogous region of pProl 7 except for two point mutations and a single nucleotide insertion.[9] In contrast, the 5 end of pProl 17 is highly divergent from comparable sequences of pProl 7 or pProl 14. Despite the heterogeneity displayed by the 5' untranslated segments of pProl 14 and pProl 17, both regions are highly enriched for G-C nucleotides and contain no potential AUG translational initiation codons. The 170 nucleotide 3-untranslated region of pProl 14 contain two putative polyadenylation signals (AATAAG and AATAAA), whereas, only a single polyadenylation signal(AATAAG) is present within the 40 base long 3-untranslated region of pProl 17.

The encoded proteins of pProl 14 and pProl 17 are about 17.2kD and this deduced mol size is in good agreement with the value obtained by SDS-PAGE of the putative prolamine precursor.[6] The composition of the first 14 amino acid residues of both proteins is typical of a signal peptide in displaying a basic amino acid, lysine, at residue 2 followed by a core of hydrophobic amino acids(Fig. 3B). These 14 amino acid residues are highly conserved in pProl 14 and pProl 17 except for Val/Phe and Ile/Glu exchanges at residues 6 and 12, respectively. Although the N-terminus of the rice prolamine is blocked to Edman degradation,[12] the cleavage site of the signal peptide is likely to occur between alanine(residue 14) and cysteine(residue 15) since it best conforms to von Heijne's rule for determining the signal peptide cleavage site.[13] Consistent with the results from dot blot hybridization analysis, the coding sequence of pProl 14 is 95% homologous to pProl 7, whereas, pProl 17 shares only 75% nucleotide sequence homology with pProl 7 or pProl 14. The predicted derived primary sequences of pProl 14 and pProl 17 share only about 63% homology. Short stretches (6~15 nucleotides) of insertions/deletions and numerous point mutations are evident between pProl 14 and pProl 17, most of which result in amino acid replacements. In spite of the close DNA homology between pProl 7 and pProl 14,

Table 1. Amino acid composition(%) of deduced mature prolamine polypeptides

	pProl 14	pProl 17
Ala	10.4	8.1
Val	6.7	8.1
Leu	11.9	6.7
Ile	4.5	5.2
Pro	4.5	5.2
Phe	5.2	5.2
Trp	0.7	0.7
Met	0.0	3.0
Lys	0.0	0.0
His	0.7	0.7
Gly	5.2	4.4
Thr	2.2	3.7
Cys	0.7	5.9
Tyr	7.5	4.4
Asx	7.5	3.7
Glx	20.1	23.7
Ser	7.5	8.2
Arg	4.5	3.0

eight out of the nine amino acid residues at the N-terminus of the deduced mature proteins are dissimilar between these two polypeptides. This heterologous N-terminus is due to a single nucleotide(C) insertion at +48 bp and a nuclentide (T) deletion at the +72 bp of pProl 14.

The deduced mature proteins of pProl 14 and pProl 17 are about 15.3kD in size. The amino acid composition of the mature polypeptide of pProl 14 showed a high mole percentage of glutamine and hydrophobic amino acids and relatively low levels of charged residues such as lysine and histidine(Table 1). The proline content of rice prolamine(5%), however, is lower than the 10~30mole percentage evident in other cereal prolamines.[14] The relatively higher proline content of prolamines from maize, wheat, barley and rye[14] is due in part to the presence of a tandemly repeated conserved peptide enrich in this amino acid. Therefore, the lower content of proline in the rice prolamine is expect since these proteins lack repetitive peptides prevalent in other cereal prolamines. No methionine and lysine residues are found in this rice polypeptide. This amino acid composition is in good agreement with the known rice prolamine composition,[2,2] supporting the identification of this clone. The deduced mature protein of pProl 17 also contains a high mole percentage of glutamine as well as hydrophobic amino acids. However, in contrast to pProl 7 or pProl 14, 4methionine and 8 cysteine residuces were found in this polypeptide(Table 1). The presence of cysteine residuces suggests that the pProl 17 polypeptide is likely to be soluble in alcohol solutions only in the presence of reducing agents.

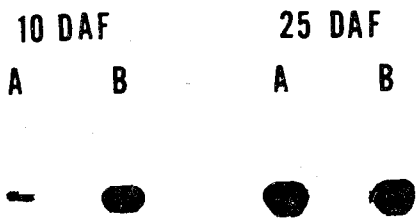

Fig. 4. Northern blot analysis of rice seed RNA. Five μg of total RNA isolated from 10 and 25 DAF seeds were resolved on a 1.2% agarose-formaldehyde gel alongside identically treated EcoRI/ScaI cut pUC 19 fragments as molecular size standards. The gel was blotted onto membrane filter and the blot hybridized to ^{32}P-labeled cDNA inserts of pProl 14(lane A) and pProl 17 (lane B), respectivelly

Relative Abundance of Prolamine mRNA Transcripts

The relative abundance levels for the two different types of prolamine mRNA transcripts were examined by Northern blot analysis(Fig. 4). Total RNA was isolated from rice seeds at 10 and 25 DAF and immobilized onto membrane

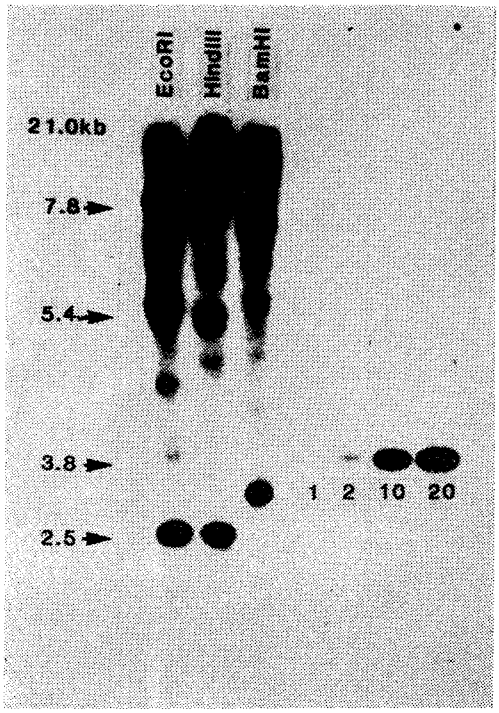

Fig. 5. Southern blot analysis of rice genomic DNA. Five μg or rice leaf genomic DNA was digested with EcoRI, HindIII, and BamHI and resolves on a 0.4% agarose gel along with prolamine cDNA gene copy number standards. The genomic blot was hybridized with ^{32}P-labeled pProl 14 cDNA insert. Bands labeled 1, 2, 10 and 20 indicate the number of gene copies of the prolamine insert sequence per haploid rice genome

filters. These filters were hybridized with ^{32}P-labeled pProl 14(lane A) and pProl 17(lane B) cDNA inserts, respectively. At 10 DAF of rice seed development, mRNA transcripts for pProl 17 are present at 3~4 fold higher levels than pProl 14 transcripts, whereas, the relative abundance of these two mRNA transcripts is similar at 25 DAF. Overall, the expression levels of both prolamine mRNA transcripts are higher at 25 DAF than at 10 DAF.

Organization of the Rice Prolamine Genes

The organization and copy number of the prolamine genes were determined by using the genomic Southern blotting technique(Fig. 5). Rice leaf DNA was digested with EcoRI, HindIII, and probed with the ^{32}P-labeled pProl 14 insert. Amounts of pProl 14 equivalent to 1, 2, 10, and 20 copies per haploid genome were run on gel lanes adjacent to the rice DNA fragments. With all three enzymes, complex patterns of rice DNA restriction fragments containing the prolamine gene sequenes were obtained. The EcoRI and HindIII digestions of rice DNA gave about 10~20 copies of a 2.5kb fragment when its autoradiographic signals are compared to those displayed by the internal standards. Several additional bands of 5kb or larger size were also detected, all containing multiple copies of the rolapmine sequence. Overall, the results from genomic Southern blot analysis is consistent with the notion that the number of genes for prolamines is extremely large(~80~100 copies per haploid genome).

The Cellular Pathways of Protein Body Formation

The storage proteins in plants are packaged into discreete membrane delimited organelles called protein bodies. All storage proteins are initially synthesized with a transient signal peptide which mediates binding of the nascent chain to the ER. Vectorial transport and proteolytic processing of the signal sequence results in internalization of the mature setorage protein within the rough ER lumen. Within this compartment, the storage protein, depending on the type, i.e. globulin or prolamine, can undergo two fates with regard to their eventual packaging into protein bodies. The two different cellular processes in protein body formation are illustrated in rice. Ultrastructural analysis of rice endosperm reveals at least two different types of protein bodies distinquishable both by morphology and staining. One type of protein body assumes a spherical shape and is electron luscent while a second type displays an irregular shape structure and stains very densely(Fig. 6). Since rice endosperm accumulates glutelins(which are actually globulins) and prolamines, the questions raised were whether these proteins are packaged separately into distinct organelles or were intermixed within the same, and what were the cellular processes thatm ediate the formation of morphological distinct protein bodies in rice endosperm.

Using immunocytochemistry where antigen-antibody complexes are visualized by gold particles conjugated with protein A, Krishnan et al.[5] demonstrated that the rice glutelins and prolamines were packaged into separate organelles. Protein contents contained within the electron-luscent spherical protein bodies reacted with the prolamine antibodies while electron-dense, irregular shaped protein bodies were gold particle labeled when glutelin antibodies were employed(Fig. 6). Very little crosslabeling was evident indicating that these organelles preferentially contain only one type of storage protein. The surrounding membrane of the spherical protein bodies were studded with ribosome-like particles and in many instances a direct connection to the rough ER were observed. These observation indicate that rice prolamine is packaged via the process first described for the maize prolamine, zein, where these proteins aggregate directly within the rough ER lumen.[15]

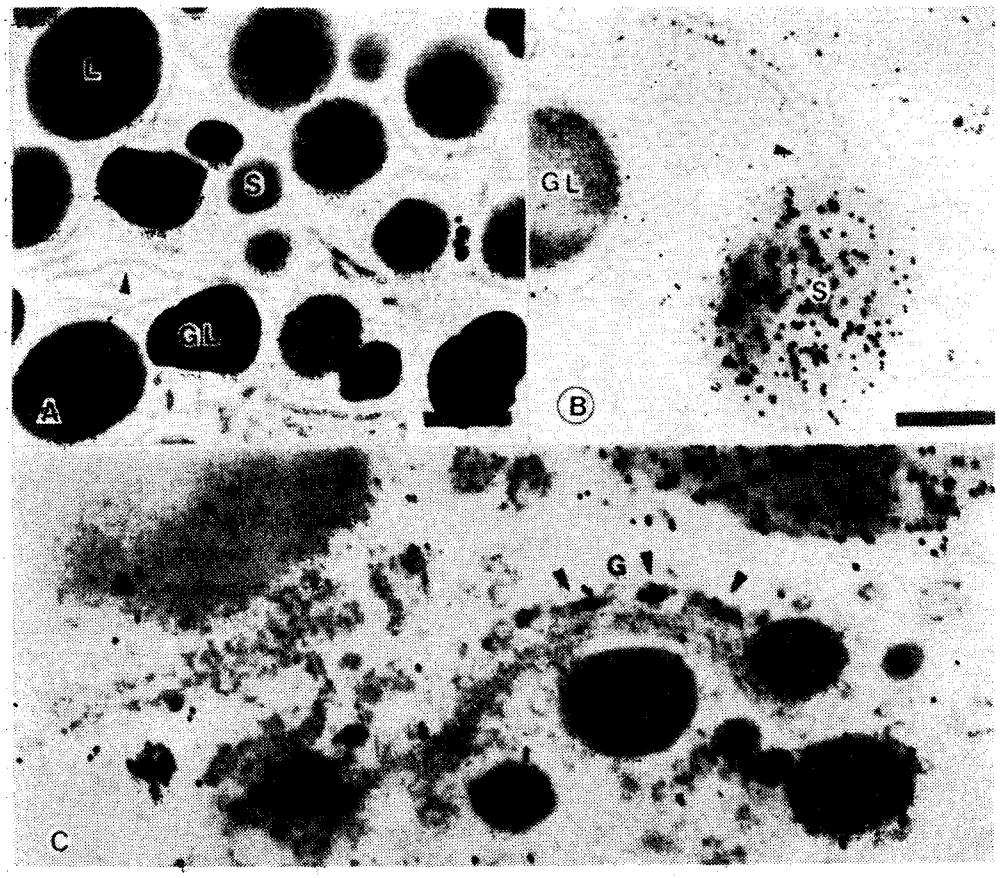

Fig. 6. Protern body formation in rice endosperm. Panel A is a low magification view of 14 day old endosperm and shows the presence of three types of protein bodies. Extensive rough ER is routinely observed(arrowhead). L and S, large and small spherical prolamine protein bodies, respectively; GL, irregular shaped glutelin protein body. Panel B shows the anti-prolamine mediated immunogold labeling of spherical protein body(note the direct connection to rough ER). Panel C depicts gold particle labeling of Golgi-derived vesicles after exposure to glutelin antibodies(Bar=1μm)

In contrast, several lines of evidence indicate that the rice glutein is packaged by a different mechanism. The irregular-shaped protein bodies which are highly enriched for glutelin are surrounded by smooth membranes devoid of ribosomes. At higher magnification, small electron-dense particles, labeled with protein A conjugate, were commonly found in the vicinity of the Golgi complex strongly suggesting a role of this organelle in sorting of glutelins to protein bodies(Fig. 6). Yamagata and Tanaka[16] have suggested that the irregular shaped protein bodies were formed by the discharge of protein from the dictysomal vesicles. A vacuolar site of glutelin deposition was observed in subaleurone cells(the outer 2~3 layer of endosperm cells adjadent to the aleurone).

Vacuoles, however, were not evident in the bulky parenchymous cells of the endosperm, indicating an alternative path way of for glutelin protein body formation and maturation. Staining of endosperm sections with phosphotungstic acid, considered a selective stain of polysaccharides of the plasma membrane, revealed an elaborate infrastructure of the irregular shaped protein bodies. Distinct protein inclusions of various

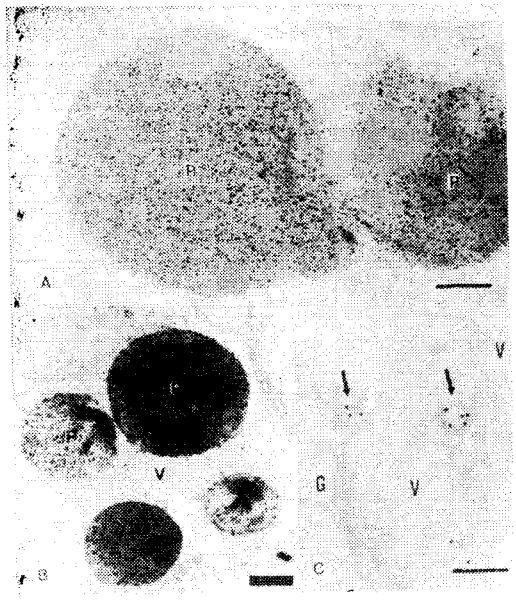

Fig. 7. Protein body formation in wheat endosperm. Panel A shows the immunogold labeling patterns of protein body after incubation with gliadin antibody. The electron-opaque inclusions are unreactive to the antibody(Bar=2μm). Panel B shows four gold particle labeled protein inclusions contained within a vacuolelike body(Bar =1μm). Panel C demonstrates gold labeling of Golgiderived vesicles(Bar=0.5μm).

sizes separated by electron-dense material were evident. The distinct structural features of this organelle suggested to these workers that irregular shaped protein bodies enlarge by fusion with Golgi-derived vesicles or with smaller protein bodies.

The prolamines of rice, maize and sorghum are packaged into distended rough ER.[5,15,17] The packaging of the wheat prolamines has been under investigation by several laboratories but a consensus view of protein-body formation has yet to be been agreed upon. For this reason, protein body formation of wheat endosperm was reinvestigated by Kim et al.[18] who employed immunocytochemical approach of wheat endosperm sections. As early as 9 days after flowering, wheat endosperm cells contained abundant starch granules and numerous protein bodies up to 20mm in diameter. Consistent with Bechtel's earlier study,[19] these workers could not demonstrate any connections between the rough ER and protein bodies, although numerous membranous structures were associated with the larger protein bodies. The larger protein bodies, contained one or more spherical electron-opaque inclusions which contained material unreactive to gliadin specific antibody(Fig. 7A) or antibodies which cross-reacted with the wheat glutenins. The Golgi apparatus was readily demonstrable in endosperm tissue as early as 5 days after flowering. This organelle consisted of 2~3 loosely stacked cisternae layers. In 9 day old developing seed, electrondense vesicles were associated with the Golgi apparatus and in many instances, these vesicles were directly connected to the distal ends of the Golgi cistenae. The contents of these electron-dense vesicles were specifically labeled with gold particles when the tissue section was preincubated with gliadin antibodies(Fig. 7). Overall, these results support the hypothesis that the Golgi apparatus play a specific role in the packaging of wheat prolamines into protein bodies. Although it is difficult to totally eliminate the rough ER as the site of protein deposition, it is not the major mechanism by which discrete, condensed protein bodies are formed in wheat.

References

1. Mandac, B.E. and Juliano, B.O.: Phytochem., 17 : 611(1978)
2. Padhye, V.W. and Salunkhe, D.K.: Cereal Chem., 57 : 389(1979)
3. Luthe, D.S.: Plant Sci. Lett., 32 : 147(1983)
4. Yamagata, H., Sugimoto, T., Tanaka, K. and Kasai, Z.: Plant Physoil., 70 : 1094 (1982)
5. Krishnan, H.B., Franceschi, V.R. and Okita, T.W.: Planta 169 : 471(1986)
6. Krishnan, H.B. and Okita, T.W.: Plant Physiol., 81 : 748(1986)
7. Yamagata, H., Tamura, K., Tanaka, K. and Kasai, Z.: Plant Cell Physiol., 27 : 1419

(1986)
8. Okita, T.W., Ksishnan, H.B. and Kim, W. T.: Plant Sci., 57 : 103(1988)
9. Kim, W.T. and Okita, T.W.: FEBS Lett., 231 : 308(1988)
10. Kim, W.T. and Okita, T.W.: Plant Physiol., 88 : 649(1988)
11. Heidecker, G. and Messing, J.: Nucleic Acid Ses., 11 : 4891(1983)
12. Bietz, J.A.: Biochem. Genet., 20 : 1039(1982)
13. von Heijne, G.: Nucleic Acid Res., 14 : 4683 (1986)
14. Kreis, M., Shewry, P.R., Forde, B.G., Ford,e J. and Miflin, B.J.: Oxford Surveys of Plant Mol. and Cell Biol, 2 : 253(1985)
15. Larkins, B.A. and Hurkman, W.T.: Plant Physiol., 62 : 356(1978)
16. Yamagata, H. anp Tanaka, K.: Plant Cell Physiol., 27 : 135(1986)
17. Taylor, J.R.N., Schussler, L. and N.v.d.W.: S. Afr. J. Bot., 51 : 35(1985)
18. Kim, W.T., Franceschi, V.R., Krishnan, H.B. and Okita, T.W.: Planta 176 : 173 (1988)
19. Bechtel, D.B. and Barnett, B.D.: Cereal Chem., 63 : 232(1986)