Glutelins and prolaminos are the major storage proteins of rice (Oryza sativa) seeds. These proteins have been extensively studied recently at the molecular level. To date, more than 10 rice prolamin cDNA clones (Kim and Okita, 1988; Masumura et al., 1990), but far fewer genomic clones (Kim and Okita, 1988; Feng et al., 1990), have been described. In this report we describe another rice prolamin genomic nucleotide sequence with its 5'-flanking region containing putative regulatory sequences.

The RP6 prolamin gene contains a single open reading frame encoding 156 amino acids, 672 bp of 5'-flanking region, and 1058 bp of 3'-flanking region; no introns are present (Table I). The transcribed region of the RP6 nucleotide sequence exactly matches the rice prolamin XRM7 cDNA, except that the dinucleotide GC replaces CG at position 49, causing a predictable amino acid change from Arg to Ala in the 19-amino acid signal peptide of RP6 (Masumura et al., 1990).

Downstream from the stop codon, TAG, there are two poly(A) signals and a hairpin loop that may be formed right after the first poly(A) signal at position 508. In comparison with the previously described rice prolamin gene LProl 4a (Kim and Okita, 1988), 22 and 133 nucleotides in the 5' flanking regions have been deleted at positions -96 and -244, respectively, from the translation starting site of RP6. In other regions, they exhibit 70% homology.

In addition to the CAAT and TATA boxes, the upstream highly conserved motif, TAAAGTGA (Forde et al., 1985), which is in consensus with other cereal prolamin genes such as wheat α-gliadin (Rafalski et al., 1984) and barley B1 hordein (Forde et al., 1985). The upstream highly conserved motif, TAAAGTGA in RP6, also shares some homology with its two direct-repeat sequences at -542 and -386. Colot et al. (1989) reported that this motif was essential for endosperm-specific expression of cereal prolamin genes in transformed tobacco.

The downstream conserved motif, TGAGTCAT, appears to be the Jun/GCN4-binding site important for gene expression in a variety of eukaryotes (Curran and Franza, 1988). Kim and Wu (1990) suggested that this motif may play an important role in rice glutelin gene tissue-specific expression. There are 10 consecutive repeats (beginning at -633) of the trinucleotide, TTA, in addition to two 13-nucleotide (GTAA-

### Table I. Characteristics of a rice prolamin RP6 gene

| Organism: | Oryza sativa L. cv Tainung 67. |
| Function: | Seed storage protein. |
| Techniques: | λ EMBL genomic library screened with pS18 cDNA (Shyur et al., 1990); restriction fragment subcloning into pBluescript II (Stratagene); plasmid sequencing; deletion subcloning and deoxy sequencing of both strands. |
| Features of Gene Structure: | Putative TATA box (TATAATA) at -90. Putative CAAT box (CCAAAT) at -168. Putative "-300 element" (Forde et al., 1985) at -289 to -270. Direct repeats at -542 to -530 and at -386 to -374. Ten consecutive direct repeats of TTA at -634. Stem-loop structure at 508 between two poly(A) signals at 503 and 539. (G+C) Content: 36.4% (45.9% in coding region). |
| Structural Features of Protein: | Open reading frame, 156 amino acids; calculated M, 17,785. |
| Antibodies: | Polyclonal rabbit (Shyur et al., 1992). |
| Subcellular Location: | Protein body I. |

AGTTCTACA) direct repeats that may act as an enhancer-like element controlling the expression of soybean conglycinin gene (Chen et al., 1988).

These findings suggest that the above-mentioned loci in RP6 are likely to be the cis-acting elements in the regulation of tissue specificity, developmental expression, and activity level of rice prolamin genes.

Received September 8, 1992; accepted October 2, 1992.

Correspondence author: fax 886-2-782-7954.

**LITERATURE CITED**

Chen ZL, Pan NS, Beachy RN (1988) A DNA sequence element that confers seed-specific enhancement to a constitutive promoter. EMBO J 7: 297–302