Heterogeneity in cDNA clones encoding rice glutelin

Co-Shine Wang, Kamini Shastri, Lisa Wen, Jenq-Kuen Huang, Burachai Sonthayanon,
S. Muthukrishnan and Gerald. R. Reeck

Department of Biochemistry, Kansas State University, Manhattan, KS 66506, USA

Clones encoding the rice storage protein glutelin were selected from a cDNA library of immature rice seeds. Sequence analysis and hybridization studies on these clones provide insight into the nature of heterogeneity in glutelin genes. Based particularly on major differences in the 3'-noncoding regions, it appears that glutelin genes fall into two sub-families.

Glutelin; cDNA clone; Nucleotide sequence; Amino acid sequence; (Rice)

1. INTRODUCTION

Glutelin, the major storage protein in rice [1], is a heterogeneous group of apparently related proteins [2]. Glutelin is synthesized as a 57 kDa polypeptide which is processed to yield 37 and 22 kDa polypeptides [3]. Recently, a sequence was reported for a cDNA clone that apparently encoded a complete glutelin precursor [4]. However, no detailed information is yet available about heterogeneity in the glutelin precursor.

In this report, we unambiguously identify a cDNA clone that encodes rice glutelin. Sequence analysis and hybridization data on this and other cloned cDNAs indicate that rice glutelin genes fall into two sub-families.

2. MATERIALS AND METHODS

2.1. Plant material

Rice (*Oryza sativa* L. cv. Newbonnet) was field-grown at Kansas State University. Seeds were generously provided by Robert Dilday of the

Correspondence address: G.R. Reeck, Department of Biochemistry, Kansas State University, Manhattan, KS 66506, USA

USDA Rice Research Center (Stuttgart, AR). Rice kernels were harvested 10–17 days after flowering and stored at −70°C.

2.2. RNA preparation and construction of cDNA library

Total RNA was prepared from dehulled rice seeds by a guanidine isothiocyanate/CsCl procedure [5]. Poly(A)$^+$ RNA was isolated by oligo(dT)-cellulose chromatography [6] and used to construct a cDNA library according to Gubler and Hoffman [7]. Transformants which were tetracycline-resistant and ampicillin-sensitive were lysed onto nitrocellulose and screened using ^{32}P-labelled cDNA synthesized from a poly(A)$^+$ RNA template by avian myoblastosis virus reverse transcriptase (Life Science Inc.).

2.3. Hybrid-selected mRNA translation and immunoprecipitation

The procedure was that of Miller et al. [8]. Plasmid DNA was digested with *Bam*HI and denatured before spotting onto diazophenyl thioether filters (Schleicher and Schuell). mRNA released from filters was translated in a rabbit reticulocyte lysate (Promega Biotec) and followed by immunoprecipitation with antiserum specific

for rice glutelin. Antisera were raised in rabbits against either the 37 or 22 kDa polypeptides which had been excised from a two-dimensional gel [9].

2.4. DNA analysis

Plasmid DNA was purified by a cleared lysate procedure followed by two rounds of equilibrium centrifugation in a CsCl-ethidium bromide gradient [10]. DNA sequence analysis of the insert DNA was performed on both strands by the dideoxynucleotide chain termination method after subcloning into M13mp19 [11]. Synthetic primers were synthesized to extend the sequence analysis along the entire insert.

3. RESULTS

3.1. Hybrid-selected mRNA translation and immunoprecipitation

A cloned plasmid, pCSW-197, that was selected on the basis of its strong hybridization with a mixed cDNA probe prepared from dehulled seed mRNA, was found to select an mRNA that encoded a polypeptide of about 57 kDa (fig.1). This polypeptide was immunoprecipitated by antibodies against 37 and 22 kDa glutelin polypeptides (lane 7).

Fig.1. Hybrid-selected mRNA translation and immunoprecipitation. Polypeptides translated in a rabbit reticulocyte lysate containing [^{35}S]Met and Cys were displayed on the 15% polyacrylamide gel. Samples were the polypeptides translated: 1, without adding RNA; 2, with added RNA 'selected' by pBR322; 3, with added mRNA selected by pCSW-197; 4,5,8, with added total RNA from dehulled seeds. Samples for lanes 6 and 7 were immunoprecipitates from samples for lanes 2 and 3, respectively. The immunoprecipitated material migrates slightly faster than the glutelin band due to the presence of a large excess of the heavy chain of immunoglobulin in the same region of the gel.

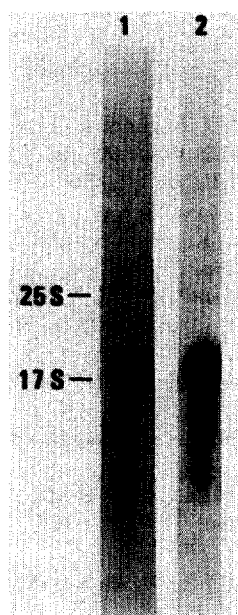Fig.3. Hybridization of glutelin cDNA clones with total RNA from dehulled rice seeds harvested 10–17 days after flowering. Total RNA (5 μg per lane) was subjected to electrophoresis in an agarose gel containing 2.2 M formaldehyde and blotted onto nitrocellulose, which was then hybridized with one of the ^{32}P-labelled inserts of pCSW-197 (lane 1) and pCSW-321 (lane 2). The position of rice rRNAs is indicated.

3.2. cDNA sequence analysis

pCSW-197 and pCSW-321, a clone to which pCSW-197 hybridized, were subjected to nucleotide sequence analysis. The insert in pCSW-197 was 229 base pairs not including poly(A)$^+$ and G-C tails. Its longest open reading frame encoded 38 amino acid residues. The sequence of pCSW-197 is exactly that of a portion of the rice cDNA clone pREE61 of Takaiwa et al. [4] except for a C for T substitution at position 1419 in pREE61.

pCSW-321 had an insert of 913 nucleotides not

including poly(A)$^+$ and G-C tails as shown in fig.2. Its longest open reading frame encodes 226 amino acid residues and the inferred sequence stops at a position corresponding exactly to the C-terminal residue inferred from the nucleotide sequence of pREE61 reported by Takaiwa et al. [4]. Within their coding regions, pCSW-321 and pREE61 differ at 6% of the aligned positions, producing 16 amino acid substitutions. In the 3'-noncoding region the two clones differ at 34% of the aligned positions. In addition, the 3'-noncoding region of pCSW-321 is 114 nucleotides longer than the corresponding region of pREE61.

3.3. Northern hybridization

Poly(A)$^+$ RNA was isolated from immature rice kernels for Northern analysis in which radiolabelled inserts from pCSW-197 and pCSW-321 were probes. As shown in fig.3, the inserts of both cDNA clones hybridized to RNA with a length of about 2000 nucleotides.

4. DISCUSSION

By hybrid-selected translation and immunoprecipitation, we have unambiguously identified a cDNA clone (pCSW-197) that encodes rice glutelin. This clone appears to be a portion of a larger cDNA clone (pREE61) of Takaiwa et al. [4]. Our identification of pCSW-197 allowed us to use it to select other glutelin cDNA clones. Our analysis of several such clones leads to the view that there are two sub-families of glutelin clones. This is particularly clear when examining the 3'-noncoding regions. These regions are much less similar than are the coding regions (see section 3). We have been able to distinguish between the subfamilies by hybridization with a probe from the 3'-noncoding region of pCSW-197 (not shown).

Particularly interesting are differences in the sequenced clones' polyadenylation signals. The glutelin-encoding insert of pCSW-197 (or,

Fig.2. Nucleotide sequences and inferred amino acid sequences for rice glutelin cDNA clones. The top row is the nucleotide sequence of pCSW-321. The second row is the nucleotide sequence of pREE61 [4]. Only the positions where this clone differs from pCSW-321 are indicated. The third row is the predicted amino acid sequence encoded by pCSW-321. The fourth row shows amino acid residues encoded by pREE61 where they differ from those encoded by pCSW-321. The box indicates the overlapping sequence of AATAAA. The solid line shows the nucleotide sequence of pCSW-321 cDNA that is missing in pREE61.

equivalently, of pREE61) has two regions with potential polyadenylation signals (AATAAATA-AA and GATAAA) at 27 and 89 nucleotides downstream from the TAG stop codon. On the other hand, pCSW-321 has only one region that appears to contain polyadenylation signals. This signal is composed of four overlapping sequences of AATAAA, and thus is quite different from the putative polyadenylation signals in pCSW-197. This suggests that alternative mechanisms of polyadenylation, or its control, might exist in the expression of genes encoding rice glutelin polypeptides.

ACKNOWLEDGEMENTS

REFERENCES

[1] Houston, D.F. and Mohammad, A. (1970) Cereal Chem. 47, 5–12.

[2] Yamagata, H., Sugimoto, T., Tanaka, K. and Kasai, Z. (1982) Plant Physiol. 70, 1094–1100.

[3] Krishnan, H.B. and Okita, T.W. (1986) Plant Physiol. 81, 748–753.

[4] Takaiwa, F., Kikuchi, S. and Oono, K. (1986) FEBS Lett. 206, 33–35.

[5] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.

[6] Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408–1412.

[7] Gubler, U. and Hoffman, B.J. (1983) Gene 25, 263–269.

[8] Miller, J.S., Paterson, B.M., Ricciardi, R.P., Cohen, L. and Roberts, B.E. (1983) Methods Enzymol. 101, 650–674.

[9] Pan, S. (1983) MS Thesis, Kansas State University.

[10] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

[11] Messing, J. (1983) Methods Enzymol. 101, 20–78.