A 45-bp proximal region containing AACA and GCN4 motif is sufficient to confer endosperm-specific expression of the rice storage protein glutelin gene, *GluA-3*

Toshihiro Yoshihara[a,b], Haruhiko Washida[a,c], Fumio Takaiwa[a,*]

[a]*Dept. of Cell Biol., Natl. Inst. of Agrobiological Resources, Tsukuba, Ibaraki 305, Japan*
[b]*Dept. of Biotechnology, Central Res. Inst. Electric Power Industry, 1646 Abiko, Chiba 270-11, Japan*
[c]*Faculty of Agriculture, Meiji University, Kawasaki, Kanagawa 214, Japan*

Received 20 November 1995; revised version received 20 February 1996

Abstract A 45-bp proximal region of the rice glutelin promoter (−104/−60) containing two putative *cis*-elements, the AACA motif and GCN4 motifs, was fused to a truncated CaMV 35S promoter (−90/+9; −90Δ35S)/GUS. The 45-bp fragment specifically enhanced the promoter activity in endosperm tissue of transformed tobacco. A substitution mutation of the GCN4 motif reduced the promoter activity, whereas mutation of the AACA motif increased the activity in the embryo as well as in the endosperm. These results suggest that the GCN4 motif generally enhances the promoter activity but that the combination of the two motifs confers the endosperm specificity. Furthermore, the function of the two motifs was dependent on the orientation and/or distance from a G-box element in −90Δ35S, suggesting that synergistic interaction between the factors that recognize those motifs and the G-box element is important for transcriptional regulation.

Key words: Rice glutelin gene; AACA motif; GCN4 motif; Endosperm-specific expression; Promoter analysis

1. Introduction

Glutelin is a major seed storage protein of rice (*Oryza sativa* L.), which constitutes a small multigene family of about 10 members per haploid genome. These genes are clearly classified into two subfamilies designated as *GluA* and *GluB* [1–5]. This genomic organization is similar to that of other seed storage protein genes. The expression of storage protein genes is strictly controlled during seed development and is regulated primarily at the transcriptional level [5–7]. Therefore, they provide a model system for studying the tissue specificity of plant genes [8–10]. To date, several sequences, such as the −300 element [11,12], GCN4 motif [13,14], RY element [15–17] and G-box motif [18], have been reported to be involved in seed-specific expression.

GluA-3 is one member of the *GluA* subfamily that has thus far been well analyzed since it is different from other glutelin genes in temporal expression [5,6,19]. These observations suggest that several putative *cis*-elements are responsible for tissue specificity and temporal regulation. Zhao et al. [20] have recently reported that the *cis*-element responsible for stage-specific expression can be separated from that for endosperm specificity. The latter element was located between positions −346 and −263. We have also shown that the sequence between −437 and −317 of *GluA-3* is sufficient to confer seed-specific expression [21]. However, the functional importance of the proximal region of the gene has not been thoroughly analyzed in spite of the high sequence homologies with other members of the *GluA* subfamilies. It is noteworthy that the AACA motif around −70 and the GCN4 motif around −100 are conserved in all members of glutelin genes. In this paper, we demonstrate clearly by gain-of-function experiments that the 45-bp proximal region of the *GluA-3* promoter (−104/−60), including the AACA motif and GCN4 motifs, is sufficient to confer endosperm-specific expression. This is the first report clearly showing the elements required for endosperm specificity in a cereal seed storage protein gene.

2. Materials and methods

2.1. Construction, plant transformation and cultivation

A 45-bp oligonucleotide and three types of mutagenized oligonucleotides of the *GluA-3* promoter (−104/−60) were synthesized with the *Sau*3AI recognition site at both the 5′-and 3′-end. Two putative *cis*-elements, the AACA and GCN4 motif, were substituted to create *Apa*I and *Sma*I recognition sites, respectively. These were blunted at their *Sau*3AI sites and cloned into the *Hin*cII site of the plasmid pUC18. After digestion with *Hin*dIII and *Bam*HI, fragments including the 45-bp oligonucleotides were excised from the plasmid and then subcloned into the upstream region of the CaMV 35S truncated promoter (−90/+9) in pLP19 binary vector [21] as shown in Fig. 1. The orientation and accurateness of the inserted oligonucleotides were determined by sequencing [22].

The plasmid pLP19 containing these synthesized oligonucleotides was introduced into *Agrobacterium tumefaciens* strain LBA4404 by triparental mating using pRK2013 as helper. Tobacco (*Nicotiana tabacum* cv. Petit Havana SR1) was used for the *Agrobacterium* infection. The infected tobacco leaf disks were cultured and regenerated as described previously [23]. The transformants were selected by kanamycin (100 ppm), and confirmed by Southern hybridization [24] and the PCR method (data not shown). After regeneration, all the transformants were transferred into nutrient soil and grown in a greenhouse under a 14-h light (about 500 μE/cm^2, 25°C) and 10-h dark (23°C) cycle.

2.2. GUS assays

One pod from each of 7–20 transformed tobacco plants was harvested every 4 or 5 days between 7 and 24 days after pollination (DAP). To compare the maximum activity in the seeds for each construction, the highest value during endosperm development was used. Pods were frozen instantly in liquid nitrogen and stored at −80°C until use. A leaf disc and a section of stem and root were also harvested and stored similarly. GUS assays were performed as described by Jefferson [25]. Total protein content was determined to normalize the GUS activities by the method of Bradford [26]. Data are shown for each transgenic line separately, and statistically confirmed the significance in the mean difference by Student's *t*-test. For histochemical staining of GUS activity, a hand-cut section of the seed was dipped into X-glucuronide solution overnight at 37°C.

*Corresponding author. Fax: (81) (298) 38 8397.

SSDI S0014-5793(96)00233-5

3. Results

It was first investigated whether a 45-bp fragment of the *GluA-3* gene is sufficient to confer seed-specific expression to a truncated CaMV 35S promoter (−90/+9: −90Δ35S)/GUS construct. Subsequently, the functional importance of two putative *cis*-elements in the 45-bp fragment, GCN4 and AACA motifs, was examined by the substitutional mutations of individual elements as shown in Fig. 2. The normal 45-bp construct (Nm) resulted in 2.6-fold induction of GUS activity as compared with that of the control level (−90Δ35S) in seeds, while it did not show any significant effects on the activities of non-seed tissues. It was notable that high GUS activity was observed in every tissue when the AACA motif was mutagenized (MAm). The relative level of GUS activity to the control (−90Δ35S) in seeds, leaves, stems and roots was 4.3, 4.7, 1.9 and 2.3, respectively. Furthermore, when either the GCN4 (MGm) or both motifs (MAGm) were mutagenized, little activation was observed in every tissue.

To clarify the location of the expression sites directed by the 45-bp fragment, histochemical analysis was carried out as shown in Fig. 3. The radicle of the embryo was stained in the −90Δ35S construct. Similar, albeit partially staining patterns in the radicle of the embryo were also observed in the constructs of MGm and MAGm. In contrast, it was noteworthy that the endosperm tissue is specifically stained in the Nm construct, whereas whole tissue in the seeds is highly stained in the MAm construct.

In order to examine whether the promoter activity of the 45-bp fragment is dependent on orientation, the normal and mutagenized 45-bp fragments (Nm, MAm, MGm and MAGm constructs) were changed in the reverse manner relative to −90Δ35S (Figs. 1 and 4). The effect of the reverse orientation on promoter activity was calculated as the relative GUS activity to that of the corresponding forward construct. The reverse Nm and MAm constructs (rNm and rMAm) seemed to decrease the relative GUS activity to about 0.7 and 0.5 in seed, respectively. In contrast, the reverse MGm construct (rMGm) increased the relative GUS activity to about 4.0 in seeds. It was notable that the rMGm construct also markedly enhanced the relative GUS activity in leaves, stems and roots to about 12.1, 4.7 and 2.2, respectively. In non-seed tissues, the reverse orientation of Nm had little effect on promoter activity in any tissues other than leaf, in which it resulted in 2.3-fold increase. Furthermore, there was little difference in activity of non-seed tissues between the rMAm and MAm constructs. The rMAGm construct showed slightly lower activities than the MAGm construct in every tissues excepting root.

We further examined whether GUS activity would be enhanced by the trimer of the 45-bp fragment. The trimer of the Nm fragment (Nt) demonstrated almost the same effect as the monomer on −90Δ35S promoter activity. This tendency was similar in every tissue. The relative GUS activity of the trimer to the monomer fragment in seeds, leaves, stems and roots was about 0.9, 1.0, 0.9 and 1.0, respectively (data not shown).

4. Discussion

Both the AACA and GCN4 motifs are conserved in all the members of glutelin genes around −70 and around −100, respectively, suggesting that these elements should be considered as candidates for the *cis*-regulatory element responsible for endosperm-specific expression. However, the previous 5′-

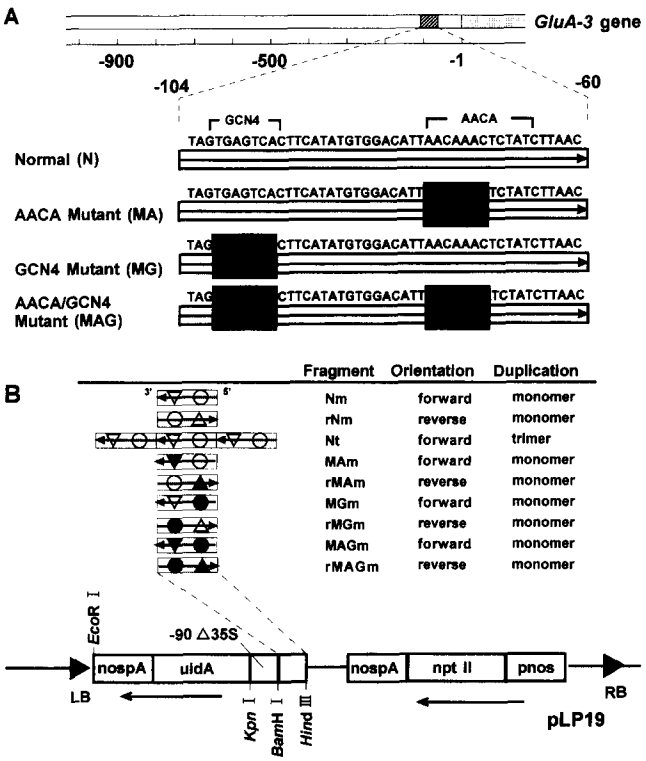

Fig. 1. Construction of *GluA-3*::−90Δ35S chimeric promoters. (A) Sequence of *GluA-3* synthetic oligonucleotides used for the chimeric promoters. (B) Orientation and duplication of the *GluA-3* synthetic oligonucleotides and structure of pLP19 binary vector. Open circles and triangles indicate the normal GCN4 and AACA motifs, and closed circles and triangles denote the mutated GCN4 and AACA motifs, respectively.

deletion analysis of *GluA-3* has only indicated the functional importance of the upstream sequence between −346 and −263 [20] or −437 and −317 [21]. In this study, it was tested whether a 45-bp fragment in the proximal region of the *GluA-3* gene (Nm: −104/−60) is sufficient to confer seed-specific expression when fused to a truncated CaMV 35S promoter (−90Δ35S). Subsequently, the functional importance of two putative *cis*-elements in the 45-bp fragment, GCN4 and AACA motif, was examined by the substitutional mutation of each element. Our results indicated that the small 45-bp fragment specifically enhances the activity of −90Δ35S in endosperm of transgenic tobacco. It is noteworthy that the mutation of the AACA motif (MAm) shows higher GUS activity than the Nm construct in every tissue, although the mutation of the GCN4 motif (MGm) and both motifs (MAGm) did not exert any significant effect on the GUS activity. These indicated that the GCN4 motif generally enhances the promoter activity but that the combination of the two motifs confers endosperm-specific gene expression.

The GCN4 motif consists of the consensus seven nucleotides (TGAPuTCA) forming a palindromic structure, which is recognized by a bZIP type *trans*-acting factor [27,28]. It has been reported that the GCN4 motif is implicated in some plant gene promoter activation in combination with other elements (i.e. endosperm motif [13,29]). De Pater et al. [14] demonstrated that a trimer of a 22-bp fragment in the proximal region of the pea lectin promoter containing the GCN4 motifs confers seed-specific gene expression and GUS staining in whole seed tissues. However, it was shown here that the GCN4 motif enhances the promoter activity in every tissue or the enhancement is due to mutation of a putative negative regulatory element AACA (the MAm construct). This discrepancy may be explained by the difference in the CaMV 35S promoter used: De Pater et al. [14] used the −46 to +8 region (−46Δ35S) of the CaMV 35S promoter to express the oligomerized GCN4 motif, whereas the −90 to +9 region (−90Δ35S) was used in this study. The region between −90 and −46 of the 35S promoter includes a G-box which recognizes the b-ZIP *trans*-acting factor, TGA1a [30,31]. Therefore, it was suggested that the combinatorial interaction between the G-box and GCN4 motifs directed the expression in every tissue. It is interesting to note that our 45-bp fragment, containing the GCN4 and AACA motifs, could not enhance the activity of the core −46Δ35S promoter (data not shown). Consequently, it is suggested that the synergistic interaction of GCN4, AACA and G-box may determine the distinctive seed specificity while the interaction between GCN4 and G-box may enhance the promoter activity in every tissue. The CATGCAT sequence, named the RY element, is necessary but sufficient for the embryo-specific expression of leguminous seed storage protein genes [15,16], such specificity is also determined by the RY element in conjunction with other less defined elements [17]. Salinas et al. [18] showed that the tetramer of G-box sequences confers seed-specific expression when

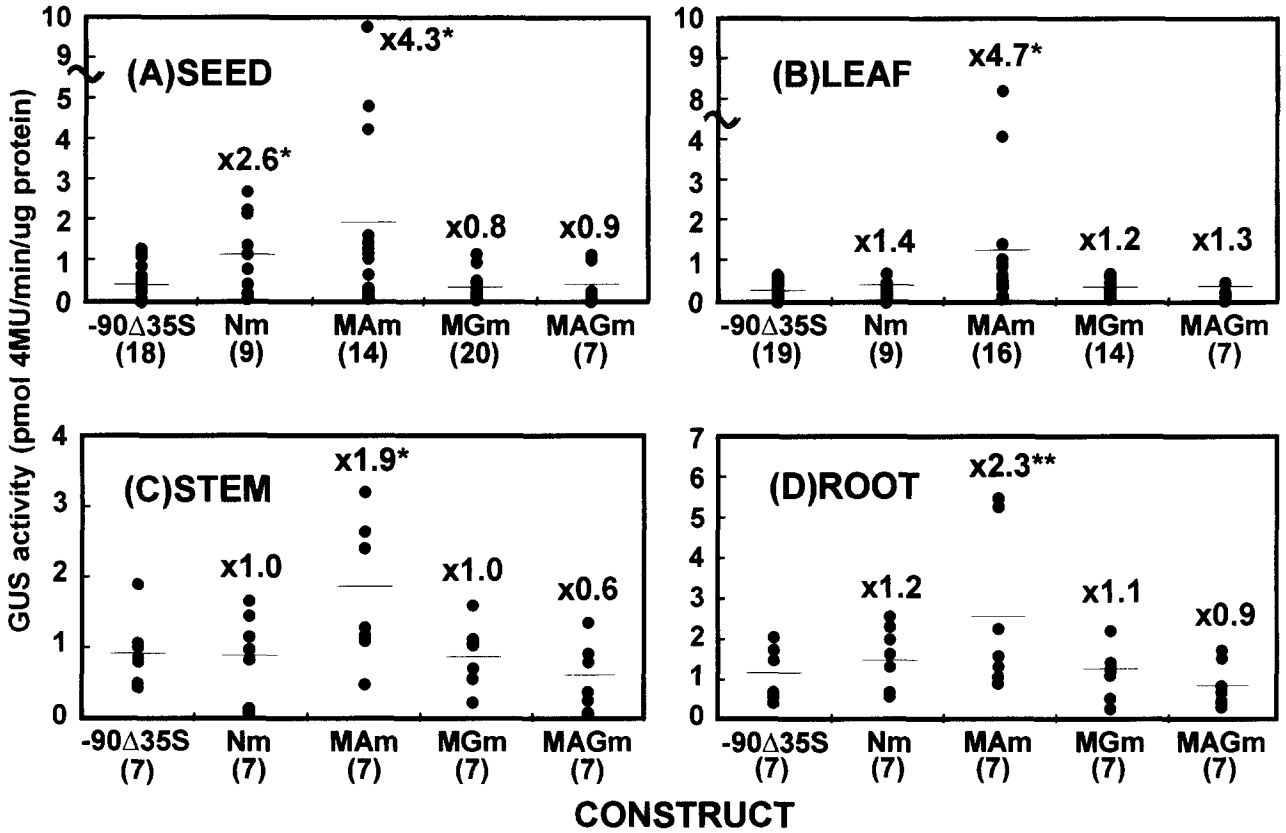

Fig. 2. Functional analysis of the normal and mutated 45-bp fragments. GUS activities of each transformant were examined in (A) seed, (B) leaf, (C) stem and (D) root, individually. Constructs used in the study are shown in Fig. 1. Each dot represents the GUS activity for each individual transgenic tobacco line. The number of plants tested is shown in parentheses. Bar indicates the average value of the GUS activity for individual construct. Relative level of GUS activity to the control (−90Δ35S) is also shown (e.g. ×2.6). Significant difference between the mean value of each construct and that of the control was examined by *t*-test and is indicated by a single ($p < 0.05$) or double ($p < 0.1$) asterisks beside the relative GUS activity.

fused to −90Δ35S. Kawagoe et al. [32] reported that synergism between the G-box and CACCTG elements is required for activation of embryo-specific expression of the bean β-*phaseolin* promoter.

It is shown here that endosperm specificity is clearly determined by a set of the GCN4 and AACA motifs when fused to −90Δ35S. However, little is known about the function of the AACA motif. More than one AACA motif is repeated in the 5′-flanking region of some glutelin genes, in addition to the AACA motif around −70 conserved in all the glutelin genes [4,33]. Homology search of the AACA motif indicates that the sequence of the AACA motif is very similar to that of a gibberellic acid response element. It was demonstrated by the gain-of-function experiment that an AACA motif located in the upstream region of *GluA-3* acts as a specific negative regulator suppressing the expression in tissues other than endosperm [21]. Furthermore, it was interesting to note that the negative effect of the AACA motif is converted to a positive one when the 45-bp fragment is fused to −90Δ35S in the reverse orientation with respect to the TATA box (rMGm), while the positive effect of the GCN4 motif is little affected besides seed by changing the orientation (rMAm). These results suggest that a set of the AACA and GCN4 motif does not function as one large orientation-dependent enhancer. It is also suggested that the effect of the AACA motif depends on the orientation of this motif and/or its position against the other elements. However, it could not simply explain why the same reversed AACA motif enhances the promoter activity only in the rMGm construct but not in the rNm construct. It is possible that the presence of the GCN4 motif in the 45-bp fragment affects the intensity of the interaction between the AACA motif and the other. On the other hand, the possibility cannot be ruled out that the changes in expression we observed here are not due to the oligonucleotides with which we have substituted the original sequence, because these oligonucleotides are a palindromic sequence and they could bind unknown factors. Extensive studies by more detailed base substitution mutations may be required to characterize these motifs.

In conclusion, it has clearly been demonstrated that one set of the AACA and GCN4 motif plays an important role in the endosperm-specific gene expression of the *GluA-3* gene, while the set alone did not result in high level gene expression. Consequently, the level of gene expression may depend on the combination with other sequences, such as endosperm motif and G-box motif in the upstream region. It is interesting to note that a combination of one set of the AACA and GCN4 motifs and an endosperm motif or G-box is found in the 5′-flanking region of *GluA-3*, whereas two sets of the AACA and GCN4 motifs are found within the 250 bp of

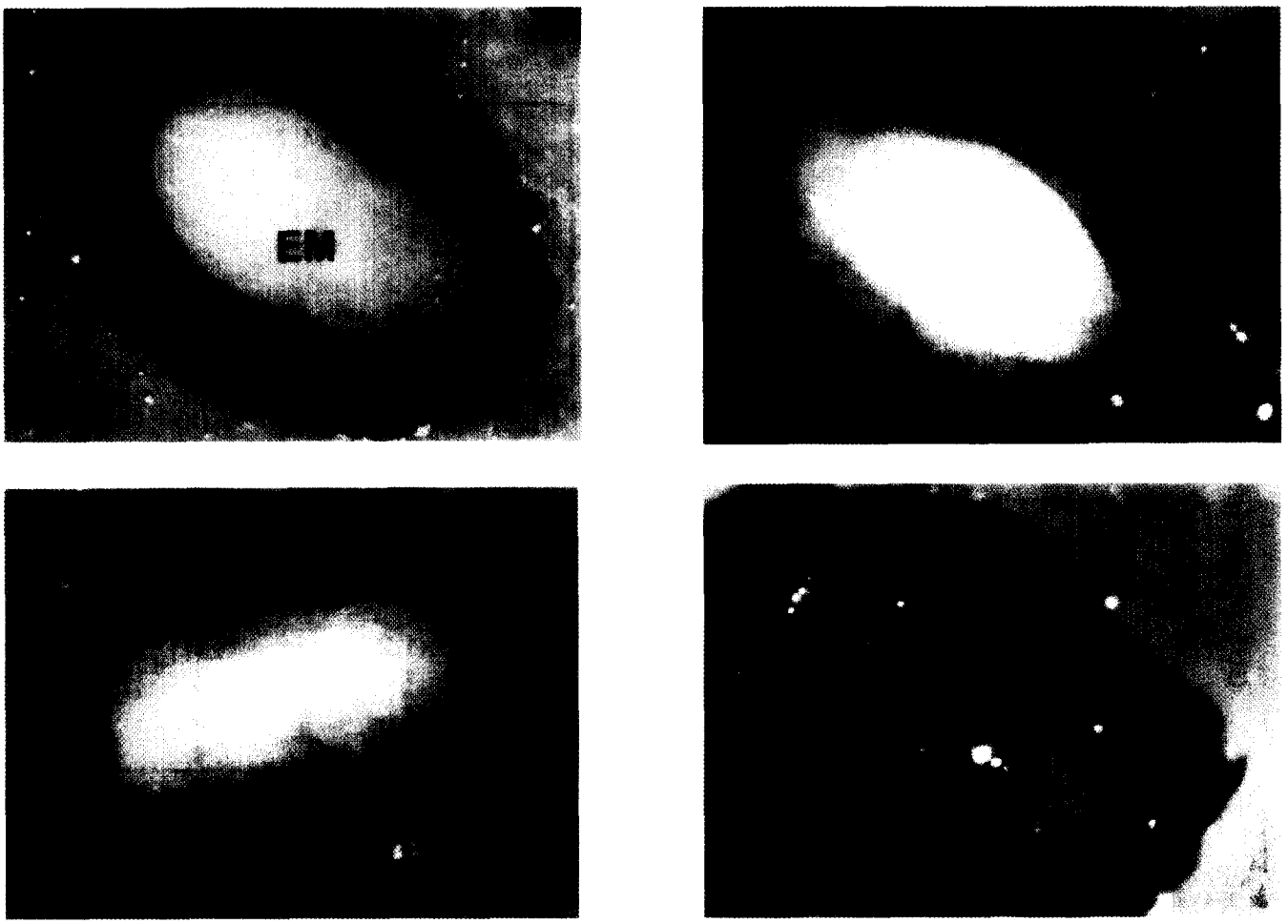

Fig. 3. Histochemical GUS staining of transgenic tobacco seeds carrying the −90Δ35S with various *GluA-3* truncated fragments. Mature seeds from transformed tobacco were sectioned and then incubated with X-Gluc. solution at 37°C, overnight. A, non-transformant; B, control (−90Δ35S); C, Nm construct; D, MAm construct; Em, embryo; En, endosperm. It should be noted that the staining site is clearly different for each construct (A, none; B, radicle of the embryo; C, endosperm; D, whole).

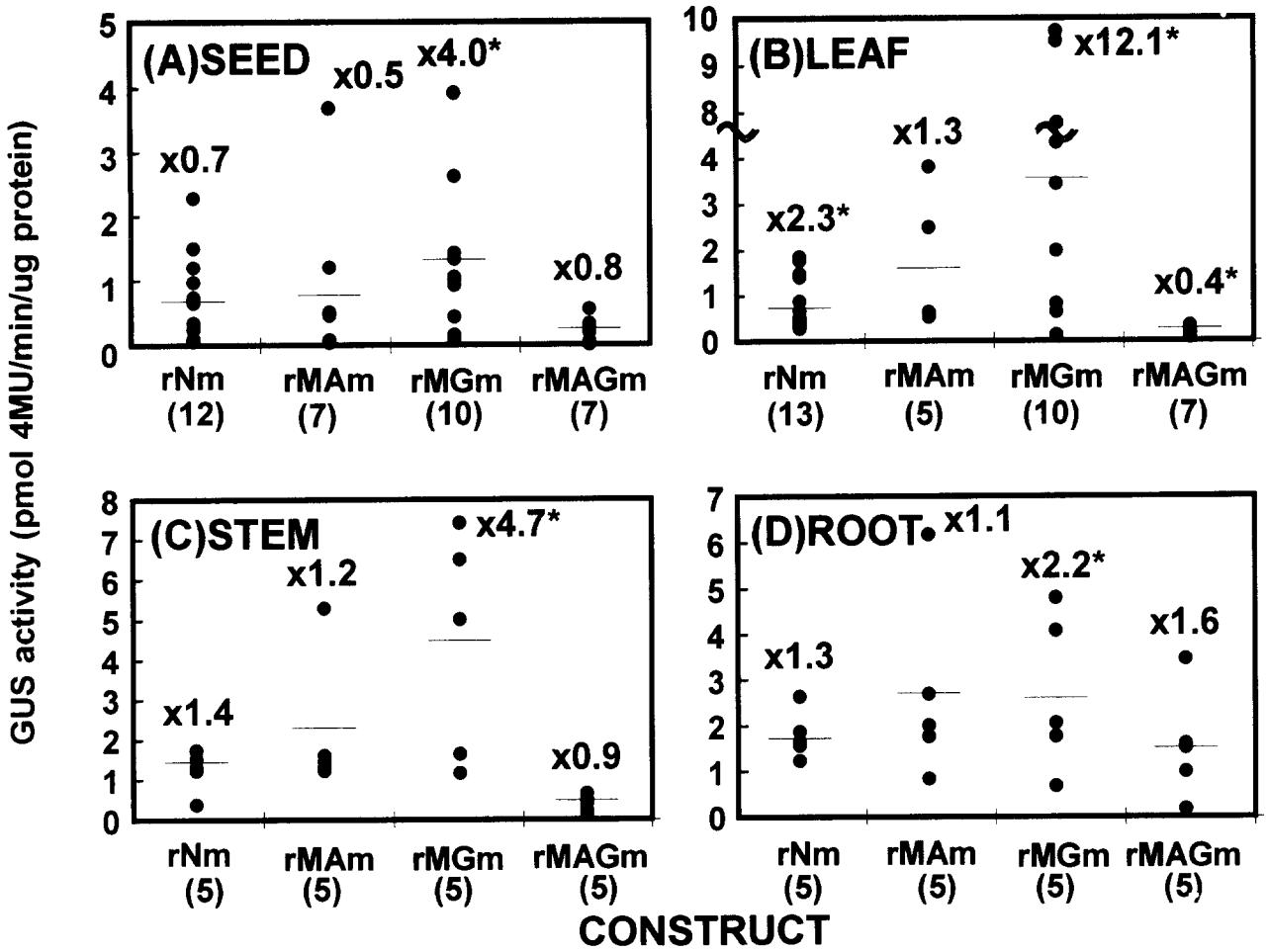

Fig. 4. Effect of the orientation of normal and mutated 45-bp fragments on GUS activity in (A) seed, (B) leaf, (C) stem and (D) root of each transformant. Constructs used in the study are shown in Fig. 1. Each dot represents the GUS activity for each individual transgenic line. The number of plants tested is shown in parentheses. Bar indicates the average value of the GUS activity for individual construct. Relative level of GUS activity to each corresponding forward orientation (shown in Fig. 2) is shown (e.g. ×4.0 = 4-fold). Significant difference between the mean value of each construct and that of the corresponding forward orientation was examined by *t*-test ($p < 0.05$) and is designated by an asterisk beside the relative GUS activity.

the 5′-flanking region of the *GluB* subfamily genes [4,33]. We have recently shown by 5′-deletion analysis of the *GluB-1* gene that both sets of the AACA and GCN4 motifs are required to confer endosperm-specific expression [33]. It is possible that the difference in motifs combined with one set of the AACA and GCN4 motif makes a difference in the gene expression level between the two glutelin gene families.

Acknowledgements: We wish to express our gratitude to Dr. B. de Lumen for critical reading of the manuscript and to Dr. J. de Bruijn (Michigan State University) for providing the binary vector pLP19. We also thank Ms. M. Utsuno, Ms. Y. Suzuki and Ms. H. Onodera for technical assistance. This research was supported in part by research grants from the Enhancement of Center-of-Excellence, the Special Coordination Funds for promoting the Science and Technology to Fumio Takaiwa.

References

[1] Takaiwa, F., Ebinuma, H., Kikuch, S. and Oono, K. (1987) FEBS Lett. 221, 43–47.
[2] Takaiwa, F., Kikuchi, S. and Oono, K. (1987) Mol. Gen. Genet. 208, 15–22.
[3] Takaiwa, F., Kikuchi, S. and Oono, K. (1989) Nucleic Acids Res. 17, 3289.
[4] Takaiwa, F., Oono, K., Wing, D. and Kato, A. (1991) Plant Mol. Biol. 17, 875–885.
[5] Okita, T.W., Hwang, Y.S., Hnilo, J., Kim, W.T., Aryan, A.P., Larson, R. and Krishnan, H.B. (1989) J. Biol. Chem. 265, 12573–12581.
[6] Takaiwa, F. and Oono, K. (1991) Jap. J. Genet. 66, 161–171.
[7] Kim, W.T., Li, X. and Okita, W.T. (1993) Plant Cell Physiol. 34, 595–603.
[8] Shotwell, M.A. and Larkins, B. (1989) in: The Biochemistry of Plants: A Comprehensive Treatise, vol. 15 (Marcus, A. ed.) pp. 296–345, Academic Press, Orlando, FL.
[9] Thomas, T.L. (1993) Plant Cell 5, 1401–1410.
[10] Müller, M., Muth, J.R., Gallusci, P., Knudsen, S., Maddaloni, M., Motto, M., Schmitz, D., Sørensen, M.B., Salamini, F., Von Wettstein, D. and Thompson, R.D. (1995) J. Plant Physiol. 145, 606–613.
[11] Quayle, T. and Feix, G. (1992) Mol. Gen. Genet. 231, 369–374.
[12] Ueda, T., Wang, Z., Pham, N., Messing, J. (1994) Mol. Cell. Biol. 14, 4350–4359.
[13] Müller, M. and Knudsen, S. (1993) Plant J. 4, 343–355.
[14] de Pater, S., Pham, K., Chua, N.-H., Memelink, J. and Kijne, J. (1993) Plant Cell 5, 877–886.
[15] Bäumlein, H., Nagy, I., Villarroel, R., Inze, D. and Wobus, U. (1992) Plant J. 2, 233–239.
[16] Lelievre, J.-M., Oliveira, L.O. and Nielsen, N.C. (1992) Plant Physiol. 98, 387–391.
[17] Fujiwara, T. and Beachy, R.N. (1994) Plant Mol. Biol., 24, 261–272.

[18] Salinas, J., Oeda, K. and Chua, N.-H. (1992) Plant Cell 4, 1485–1493.
[19] Leisy, D.J., Hnilo, J., Zhao, Y. and Okita, T.W. (1989) Plant Mol. Biol. 14, 41–50.
[20] Zhao, Y., Leisy, D.J. and Okita, T.W. (1994) Plant Mol. Biol. 25, 429–436.
[21] Yoshihara, T. and Takaiwa, F. (1996) Plant Cell Physiol. 37, 107–111.
[22] Szabodos, L., Ratet, P., Grunenberg, B., de Bruijin, F.J. (1990) Plant Cell 2, 973–986.
[23] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[24] Maniatis, T., Fritssh, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[25] Jefferson, R.A. (1987) Plant Mol. Biol. Rep. 5, 387–405.
[26] Bradfold, M.M. (1976) Anal. Biochem. 72, 248–254.
[27] Schmidt, R.J., Ketudat, M., Aukerman, M.J. and Hoschek, G. (1992) Plant Cell 4, 689–700.
[28] de Peter, S., Katagiri, F., Kijne, J. and Chua, N.-H. (1994) Plant J. 6, 133–140.
[29] Hammond-Kosack, M.C.U., Holdworth, M.J. and Bevan, M.W. (1993) EMBO J. 12, 545–554.
[30] Prat, S., Willmitzer, L. and Sánchez-Serrano, J.J. (1989) Mol. Gen. Genet. 217, 209–214.
[31] Benfey, P.N., Ren, L. and Cua, N.-H. (1989) EMBO J. 8, 2195–2202.
[32] Kawagoe, Y., Campell, B.R. and Murai, N. (1994) Plant J. 5, 885–890.
[33] Takaiwa, F., Yamanouchi, U., Yoshihara, T., Washida, H., Tanabe, F., Kato, A. and Yamada, K. (1996) Plant Mol. Biol. (in press).