Mol Gen Genet (1990) 224:289–293

MGG

Short communications

Interaction of an immature seed-specific *trans*-acting factor with the 5′ upstream region of a rice glutelin gene

Fumio Takaiwa and Kiyoharu Oono

Department of Cell Biology, National Institute of Agrobiological Resources, Tsukuba, Ibaraki 305, Japan

Received May 21, 1990

Summary. The 5′ flanking region of a glutelin gene was analyzed for interactions with nuclear proteins from immature rice seed. The specific region between positions −272 and −99 was shown to interact with nuclear proteins from immature seeds, but not with those of leaves and roots. Methylation interference experiments revealed that one factor interacted with a specific sequence element between positions −130 and −120 relative to the transcriptional start site. The sequence specificity of this DNA-protein interaction was confirmed by competition experiments using synthetic oligonucleotides. By using a synthetic oligonucleotide as a probe it was also shown that the binding activity was closely correlated with the mRNA levels of this gene during seed maturation.

Key words: DNA-binding protein – Glutelin gene – Rice – Transcription

Introduction

It has been shown that transcription by RNA polymerase II is regulated through DNA-protein interactions between nuclear factors (*trans*-acting factors) and specific *cis*-regulatory DNA sequences (Maniatis et al. 1987; Ptashne 1988; Schleif 1988). Binding of specific *trans*-acting factors to specific *cis*-acting elements is required for the transcriptional regulation associated with tissue- and stage-specific expression. Consequently, identification of the sequence-specific binding proteins and their recognition sequences is necessary for understanding the mechanism of transcriptional regulation in eukaryotic gene expression.

Glutelin is the major seed storage protein of rice and accounts for about 80% of total endosperm protein. Glutelin genes are expressed only in the endosperm during seed development and are regulated at the transcriptional level. Therefore, these genes are good candidates for studying transcriptional regulation of the tissue- and stage-specific expression of plant genes. In order to understand the molecular mechanism of regulation of glutelin genes, we have characterized five glutelin genes and detected a common 13 bp sequence motif between positions −80 and −60 in the 5′ flanking region of all members of glutelin gene families characterized to date (Takaiwa et al. 1987a; Okita et al. 1989; F. Takaiwa et al., submitted). Furthermore, we have shown that the 441 bp 5′ flanking sequence is sufficient to direct correct endosperm-specific expression in transgenic tobacco seeds and that there is a positive regulatory element between positions −441 and −237 (F. Takaiwa et al., in press). In this paper, we report specific *trans*-acting factors in maturing seed, which bind to the region between positions −130 and −120 of the glutelin upstream sequence. The interaction between the glutelin upstream region and the binding factors was highly sequence specific and this binding activity was closely correlated with the glutelin mRNA level.

Offprint requests to: F. Takaiwa

Materials and methods

DNA fragment isolation. The 5′ upstream region of a glutelin gene was prepared by digestion with *Bam*HI of plasmid pHIR13-7 (F. Takaiwa et al., in press) which carries DNA derived from a type II glutelin gene (Takaiwa et al. 1987a). This was further digested with *Hin*fI and *Dde*I and end-labeled with $\alpha[^{32}P]$dCTP using Klenow enzyme. These subfragments were electrophoresed in a native 6% acrylamide gel, and recovered from the gel by incubation overnight in 0.5 M ammonium acetate, 10 mM $MgCl_2$, 1 mM EDTA and 0.1% SDS.

Preparation of crude nuclear extracts. Maturing rice seeds (*Oryza sativa* L. cv. Mangetsumochi) were harvested 10–18 days after flowering, frozen in liquid nitrogen and stored at −80° C until use. The preparation of nuclear extracts was carrid out essentially by the

method of Jensen et al. (1988). The frozen seeds were ground to a fine powder in liquid nitrogen using a coffee mill, and resuspended at approximately 0.25 g of tissue/ml in nuclear extraction buffer A (10 mM NaCl, 10 mM MES-NaOH pH 6.0, 5 mM EDTA, 0.15 mM spermine, 0.5 mM spermidine, 2 mM dithiothreitol, 0.2 mM phenylmethylsulfonyl fluoride (PMSF), 0.4 mM salicylhydroxamic acid, 10 µg/ml butylated hydroxytoluene, 0.6% Triton X-100, 0.25 M sucrose). The suspension was stirred and homogenized with a Polytron for 30 s. The homogenate was filtered through two layers of Miracloth and centrifuged for 10 min at $2000 \times g$. The crude nuclear pellet was washed once in buffer A and resuspended in buffer B (6 g 5 × buffer A, 45 g Percoll). The suspension was centrifuged for 5 min at $4000 \times g$ and the white floating nuclei were collected and diluted with buffer A. The nuclei were pelleted by centrifugation at $2000 \times g$ for 10 min and washed once in buffer A. Nuclear proteins were prepared from this crude nuclear pellet by resuspension and sonication in 0.5–1 ml of extraction buffer (20 mM HEPES-KOH pH 7.9, 25% glycerol, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT). The mixture was incubated for 1–2 h on ice, and nuclear debris was removed by centrifugation for 15 min at $15000 \times g$. Extracts containing nuclear proteins were dialyzed for 2 h against 100–200 volumes of 20 mM HEPES-KOH pH 7.9, 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT. Aliquots were frozen and stored at −80° C.

Gel retardation assays. Binding reactions were carried out at 25° C for 30 min in a total volume of 15 or 30 µl containing 10 mM HEPES-KOH pH 7.9, 10 mM $CaCl_2$, 1 mM DTT, 1 µg poly dI-dC, 10 or 20 µl of crude nuclear extract and labeled DNA fragment. Competition reactions were conducted with various competitors added in the molar ratios described in the text. Following incubations, reactions were loaded on 4% or 8% polyacrylamide gels (30:1, acrylamide:bisacrylamide) in 6.7 mM Tris-HCl pH 7.9, 3.3 mM sodium acetate, 1 mM EDTA (Strauss and Varshavsky 1984), which had been pre-run at 12 V/cm for 90 min with buffer recirculation. Electrophoresis was performed under the same conditions. Gels were blotted on Whatman 3MM paper, dried and exposed to X-ray film.

Preparation of oligonucleotides. Oligonucleotides that span the presumptive binding site; 33mer W (−142 to −110), TTATTCATCCACCTTTCGTGTACCACACTTCAT, 33mer M, TTATTCATCCACTTTTTATGTACCACACTTCAT and 20mer W (−137 to −118), CATCCACCTTTCGTGTACCA and a control oligonucleotide C (−216 to −192) ATGTCATATTGCAAAAGAAAGAGAG; and their respective complementary sequences were chemically synthesized and annealed as described by Sambrook et al. (1989).

Methylation interference assay. The *Hin*fI fragment extending from −272 to −99 was repaired with Klenow enzyme and cloned into the *Sma*I site of pUC118. Labeled probes were prepared by digesting with *Hind*III and *Eco*RI to excise the insert and end-labeling with $\alpha[^{32}P]$dCTP or $\alpha[^{32}P]$dATP using Klenow enzyme. In the case of labeling with $\alpha[^{32}P]$dATP, the labeled probe was further digested with *Bam*HI to remove the end-labeled *Hin*dIII site. Probes were treated with dimethyl sulfate as for the Maxam-Gilbert G reaction (Maxam and Gilbert 1980) and precipitated twice with ethanol in the presence of carrier DNA. Binding reactions with the partially methylated probes were done as described above, except that the binding reaction volumes were tenfold greater. Free and bound DNAs were purified from the gel by electroelution. After elution, samples were extracted with chloroform-phenol (1:1, v/v) and the DNA precipitated with ethanol. These DNAs were treated with piperidine as in the Maxam and Gilbert (1980) reaction, and separated in denaturing 7% sequencing gels.

RNA isolation and Northern blotting analysis. Total RNA was prepared from immature and mature seeds at various stages of development as described previously (Takaiwa et al. 1987b), denatured with glyoxal at 50° C for 1 h and resolved by electrophoresis in 1.2% agarose gels (Thomas 1980). The gel was blotted onto nitrocellulose filter, which was hybridized with the ^{32}P-labeled pREE99 cDNA insert (Takaiwa et al. 1987b) in 50% formamide, 5 × SSC, 1 × Denhardt's solution, 10% dextran sulfate and 250 µg/ml denatured salmon sperm DNA at 48° C for 24 h. The filters were washed four times in 2 × SSC and 0.1% SDS at room temperature and twice in 0.1 × SSC and 0.1% SDS at 60° C.

Results and discussion

Binding of nuclear proteins to the type II glutelin gene promoter

Gel retardation assays were carried out in order to identify sequences capable of binding to specific nuclear proteins within the type II glutelin promoter. As shown in Fig. 1, the insert of pHIR 13-7, spanning positions −1329 to +1, was completely digested with *Hin*fI and *Dde*I. These fragments were labeled with ^{32}P by filling-in with Klenow enzyme and then were incubated with nuclear extract prepared from immature rice seeds. The protein-DNA interactions were analyzed in native polyacrylamide gels. When these end-labeled fragments were incubated with the nuclear extract from immature seeds, mobility of fragment 4 was retarded, indicating that it bound to nuclear protein. As shown in Fig. 1, three retarded bands were detected. On the other hand, when incubated with nuclear extracts from leaves, it was shown that fragment 2 was retarded but fragment 4 did not bind to the nuclear proteins. The retardation of fragments 2 and 4 could not be observed on incubation with nuclear extract from roots. This evidence suggests that factors extracted from immature rice seed nuclei form specific complexes with fragment 4 only. The specificity of this binding was further confirmed by the following

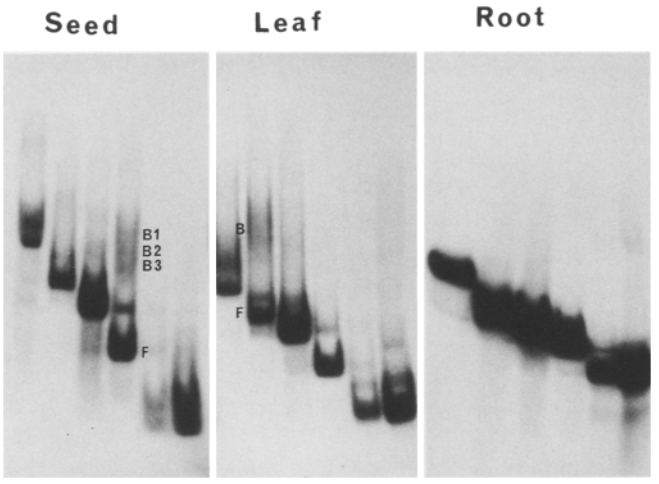

Fig. 1. Nuclear protein binding to the 5′ flanking region of the type II glutelin gene. The 5′ flanking region (positions −1336 to +1) was digested with *Hin*fI and *Dde*I and the resultant fragments were end-labeled with ^{32}P using Klenow enzyme. Fragments were separated in a 6% acrylamide gel, and then eluted from the gel and used in the binding assays. Fragment sizes are as follows: 1, 389 bp; 2, 314 bp; 3, 270 bP; 4, 174 bp; 5, 103 bp; 6, 94 bp. Nuclear extracts were prepared from immature seeds, young levels grown under light and root tissues. F, free DNA not bound to protein; B, DNA bound to nuclear protein factor.

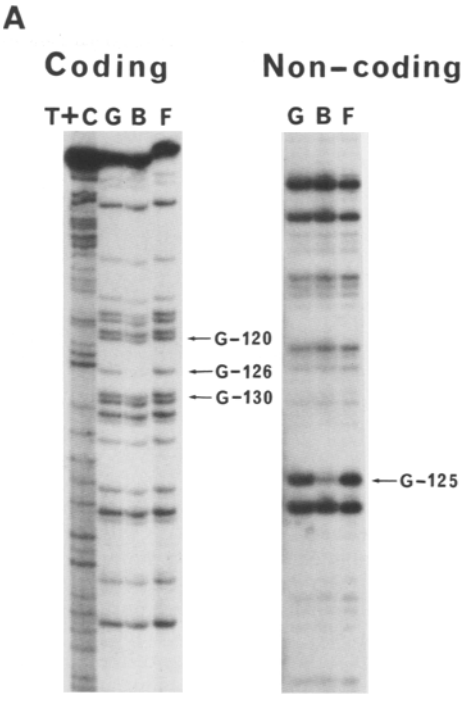

Fig. 2A and B. Methylation interference experiments with the coding and non-coding strands in the *Hin*fI fragment 4 of the type II gene. **A** The ^{32}P-labeled probe was treated with dimethyl sulfate then incubated with the crude nuclear extract from immature seeds. After electrophoresis in 4% native gels, the free and retarded bands were extracted from the gels, and then cleaved with 1 M piperidine and analyzed in 7% acrylamide gels. Positions (relative to the transcription initiation site) of the guanine residues resistant to chemical cleavage are indicated by *arrows*. F, chemical cleavage patterns of the free probe; B, the cleavage patterns of protein-bound probe. **B** A schematic representation of the sequence showing the G residues (*dotted*) involved in protein binding, as deduced from the methylation interference experiments

experiments. First, competition experiments demonstrated that the addition of an excess of unlabeled homologous fragment inhibited this interaction (data not shown). Second, binding ability was completely abolished by treating the extracts with heat (68° C for 10 min) and proteinase K, whereas RNase A was ineffective (data not shown), indicating that proteins were involved in the interaction.

Localization of binding sites

To identify the precise location of the binding site, methylation interference experiments were performed. Fragment 4 was end-labeled, partially methylated, and reacted with nuclear protein extracts isolated from maturing rice seeds. The resulting DNA-protein complexes and free DNA band were isolated, cleaved at methylated guanin sites by piperidine, and fractionated by gel electrophoresis. As shown in Fig. 2 it was found that methylation of guanin residues at position −125 of the non-coding strand and positions −120, −126 and −130 of the coding strand interfered with binding. These results show that the DNA sequence involved in binding nuclear proteins is located between positions −130 and −120 and represents the minimal region required for binding of a *trans*-acting factor.

Next, the binding specificity was investigated by competition with synthetic wild-type DNA sequences and mutated derivatives. The oligonucleotide of 33 residues complementary to positions −142 and −110 was chemically synthesized and annealed. Addition of a 10- or 100-fold molar excess of this unlabeled oligonucleotide to the binding reaction containing end-labeled fragment 4 resulted in inhibition of the band shift (Fig. 3C). In constrast, even a 100-fold molar excess of the upstream fragment (−272 to −175) did not compete for binding (Fig. 3B). Then, we examined whether the 33mer oligonucleotide from positions −142 to −110 contained sufficient sequence information for binding nuclear factors. When this 33mer oligonucleotide was incubated with nuclear extract isolated from immature rice seeds and subjected to gel retardation assay, its mobility was retarded as shown in Fig. 4. Three retarded bands (B1, B2 and B3) could be clearly detected. The specificity of interaction was further confirmed by competition experiments. Addition of excess amounts of the unlabeled 33mer oligonucleotide equally prevented the formation of these slowly migrating complexes (Fig. 4, 33W). As a control, the 25mer synthetic oligonucleotide spanning the region between positions −216 and −192 was synthesized; it contains the direct repeat sequence characteristic of the type II glutelin gene promoter (Tak-

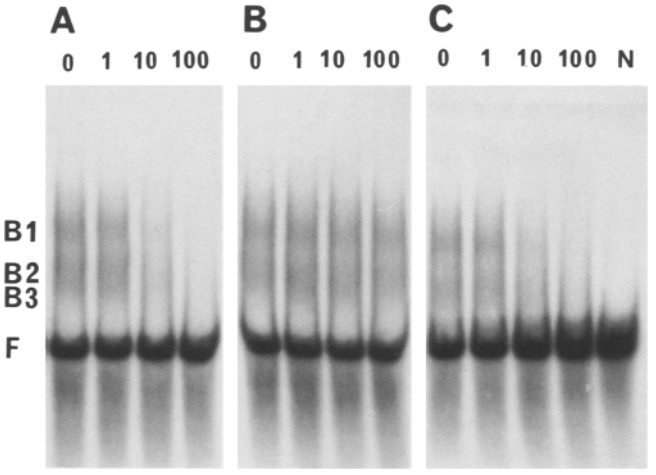

Fig. 3A–C. Competition analysis for DNA binding factors. The ^{32}P-labeled fragment 4 was used in binding assays with unlabeled DNA fragments. **A** DNA fragment between positions −174 and −99. **B** DNA fragment between positions −272 and −175. **C** Synthetic double-stranded oligonucleotide between positions −142 and −110. N, no nuclear extract added; O, no competitor added; 1, 10 and 100, 1-, 10- and 100-fold molar excess of the unlabeled fragments was added to the reaction mixture as indicated. F, free DNA; B, DNA-protein complex

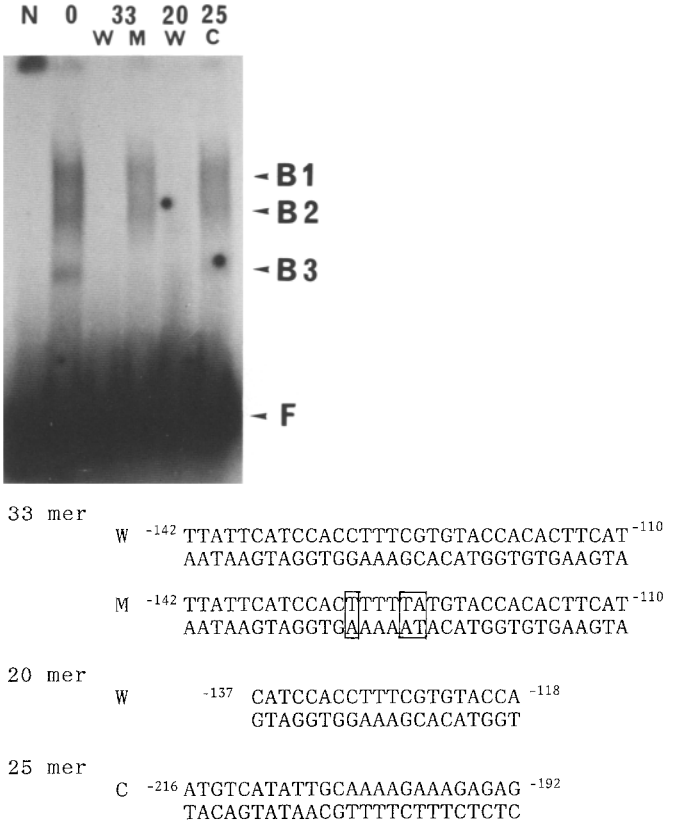

Fig. 4. Top. Competition analysis of the 33mer wild-type synthetic oligonucleotide with the other synthetic oligonucleotides. Gel retardation analysis was carried out with the end-labeled 33mer synthetic oligonucleotide as probe. A 100-fold molar excess of the unlabeled fragments of the same wild-type oligonucleotide (33 W), the 33mer mutated oligonucleotide (33 M), the 20mer oligonucleotide corresponding to part of the 33mer wild-type oligonucleotide (20 W), or the 25mer control oligonucleotide (25 C) were added as competitors. N, no nuclear extract added; O, no unlabeled oligonucleotide competitor added; F, free DNA not bound to protein; B, DNA bound to nuclear protein factor. Bottom. The sequences of synthetic double-stranded oligonucleotides

aiwa et al. 1987b). Addition of this unlabeled oligonucleotide in excess did not interfere with the interaction between the factor and the labeled 33mer oligonucleotide (Fig. 4, 25C). These observations support the conclusion that binding of the factor depends on the sequence of the upstream region from positions −142 to −110.

The sequence specificity of the binding was further ascertained by using the mutated 33mer oligonucleotide in a gel shift assay. This mutated oligonucleotide had three base substitutions (positions −120, −125 and −126) within the region characterized by the methylation interference experiment. The mutated oligonucleotide had no effect on gel shift, even though a 100-fold molar excess was added as competitor (Fig. 4, 33M). On the other hand, the smaller 20mer oligonucleotide spanning the region between positions −137 and −118 competed well with the wild-type 33mer oligonucleotide (Fig. 4, 20W). These results indicate that the sequence CTTTCGTGTA between positions −130 and −120 is responsible for the binding. However, this sequence may not be sufficient for factor binding for the following reasons. First, there is a very similar sequence (CTTTCGTGTC) between positions −252 and −243, although the binding of *trans*-acting factors could not be detected at any sites other than between positions −130 and −120 by the methylation interference experiments. Second, the competition experiments showed that the *Hin*fI-*Fnu*4H fragment between positions −272 and −175 did not affect complex formation (Fig. 3B). These results suggest that some additional sequences flanking this core site may also be required for binding of *trans*-acting factors.

The core site sequence (CTTTCGTGTAC) required for binding of the *trans*-acting factor is quite different from those characterized for other seed storage protein genes (Jofuku et al. 1987; Maier et al. 1987; Allen et al. 1989; Bustos et al. 1989; Jordano et al. 1989). In addition, the immature seed-specific *trans*-acting factor does not interact with either the positive regulatory element indicated by transformation experiments (F. Takaiwa et al., in press), or the common motif (AACAAACTCTAT from positions −78 to −65) found in all members of the glutelin genes. We do not yet have any evidence about the molecular mechanism involved in the interaction between the DNA regulatory sequences characterized to date and this *trans*-acting factor. A similar situation exists for the 60 kDa *trans*-acting factor characterized in the lectin gene (Jofuku et al. 1987).

Change of DNA binding activity during seed maturation

It is interesting to speculate as to whether the *trans*-acting factor binding to the region between positions

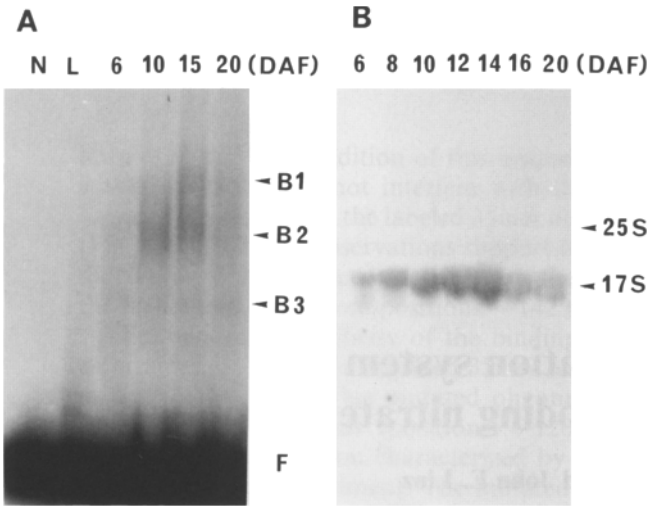

Fig. 5. A Developmental regulation of DNA-binding activity during seed maturation. The 33mer double-stranded synthetic oligonucleotide was end-labeled with ^{32}P and used as a probe for binding reactions. Nuclear proteins were isolated from maturing seeds at 6, 10, 15 and 20 days after flowering (DAF). N, the control reaction without addition of protein; L, the reaction containing the leaf nuclear proteins; B, DNA bound to factor(s); F, free DNA not bound to protein. **B** Northern blot analysis of the type II glutelin mRNA during seed maturation. Total RNA (20 μg) isolated from developing endosperm was electrophoretically fractionated in a 1.2% agarose gel. Hybridization was carried out under high stringency conditions. Rice 25 S (3377 bases) and 17 S (1812 bases) ribosomal RNAs were used as size markers

−130 and −120 regulates type II glutelin gene expression during seed maturation. The relative levels of DNA binding activities were examined during seed maturation. Equal amounts of nuclear extracts from four developmental stages were reacted with the labeled 33mer oligonucleotide fragment and the reaction was assayed by gel retardation. Very low levels of binding activity could be detected in extracts prepared from seeds 6 days after flowering. This binding activity increased and reached maximum levels at the 15th day, and then declined (Fig. 5). The binding activity increased in parallel with the amount of type II glutelin mRNA during seed development, indicating that the mRNA level of the type II glutelin gene is closely correlated with the binding activity in nuclear extracts. In contrast, nuclear extracts from leaves did not contain any binding activity for this oligonucleotide fragment. Therefore, the observed correlation between the DNA-binding activity and the mRNA level supports the notion that the binding factor may be involved in regulating transcription of type II glutelin genes. Similar correlations between gene expression and binding activity have been observed in the soybean lectin gene (Jofuku et al. 1987), soybean β-conglycinin gene (Allen et al. 1989) and tomato fruit-ripening genes (Cordes et al. 1989). However, based on the experiments done so far it cannot be ruled out that changes in the level of the binding activity are due to differences in the quality of the nuclear extract preparations from different developmental stages and not to modulation of the concentration or activity of the factor. Further investigation is required in order to examine whether qualitative or quantitative differences may be involved in the changes in level of binding activity during seed development.

Acknowledgements. We thank Dr. Kazuo Shinozaki and Dr. Robert J. Henry for critical reading of this manuscript. This work was supported in part by a project grant from the Ministry of Agriculture, Forestry and Fisheries and by Grant-in-Aid of Scientific Research of Priority Areas from the Ministry of Education, Science and Culture of Japan.

References

Allen RD, Bernier F, Lessard PA, Beachy RN (1989) Nuclear factors interact with a soybean β-conglycinin enhancer. Plant Cell 1:623–631

Bustos MM, Guiltinan MJ, Jordano J, Begum D, Kalkan FA, Hall TC (1989) Regulation of β-glucuronidase expression in transgenic tobacco plants by an A/T-rich, cis-acting sequence found upstream of a french bean β-phaseolin gene. Plant Cell 1:839–853

Cordes S, Deikman J, Margossian LJ, Fischer R (1989) Interaction of a developmentally regulated DNA-binding factor with sites flanking two different fruit-ripening genes from tomato. Plant Cell 1:1025–1034

Jensen EO, Marcker KA, Schell J, Bruijn FJ (1988) Interaction of a nodule specific *trans*-acting factor with distinct DNA elements in the soybean leghaemoglobin lbc3 upstream region. EMBO J 7:1265–1271

Jofuku KD, Okamura JK, Goldberg RB (1987) Interaction of an embryo DNA binding protein with a soybean lectin gene upstream region. Nature 328:734–737

Jordano J, Almoguera C, Thomas TL (1989) A sunflower helianthinin gene upstream sequence ensemble contains an enhancer and sites of nuclear protein interaction. Plant Cell 1:855–866

Maier UG, Brown JWS, Toloczyki C, Feix G (1987) Binding of a nuclear factor to a consensus sequence in the 5′ flanking region of zein genes from maize. EMBO J 6:17–22

Maniatis T, Goodbourn S, Fischer JA (1987) Regulation of inducible and tissue-specific gene expression. Science 236:1237–1245

Maxam A, Gilbert W (1980) Sequencing end-labeled DNA with base specific chemical cleavage. Methods Enzymol 65:499–560

Okita TW, Hwang TS, Hnilo J, Kim WT, Aryan AP, Larson R, Krishnan HB (1989) Structure and expression of the rice glutelin multigene family. J Biol Chem 264:12563–12581

Ptashne M (1988) How eukaryotic transcriptional activators work. Nature 335:683–689

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Schleif R (1988) DNA binding by proteins. Science 241:1182–1187

Strauss F, Varshavsky A (1984) A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell 37:889–901

Takaiwa F, Ebinuma H, Kikuchi S, Oono K (1987a) Nucleotide sequence of a rice glutelin gene. FEBS Lett 221:43–47

Takaiwa F, Kikuchi S, Oono K (1987b) A rice glutelin gene family – A major type of glutelin mRNAs can be divided into two classes. Mol Gen Genet 208:15–22

Thomas P (1980) Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc Natl Acad Sci USA 77:5201–5205

Communicated by J. Schell