The Plant Journal (1992) **2**(6), 837–844

The promoter of the rice gene *GOS2* is active in various different monocot tissues and binds rice nuclear factor ASF-1

B. Sylvia de Pater[1,*], Frits van der Mark[2], Saskia Rueb[1], Fumiaki Katagiri[3], Nam-Hai Chua[3], Rob A. Schilperoort[1] and Lambert A.M. Hensgens[1]

[1]*MOLBAS Research Group, Institute of Molecular Plant Sciences, Leiden University, Wassenaarseweg 64, 2333 AL, The Netherlands,*

[2]*Center for Phytotechnology RUL/TNO, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands, and*

[3]*Laboratory of Plant Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, USA*

Summary

A single copy gene has been isolated, termed *GOS2*, from rice. Sequence comparison revealed highly similar genes in mammals and yeast, indicating that *GOS2* encodes an evolutionary conserved protein. *GOS2* mRNA was detected in all tissues examined. When the upstream region was translationally fused to the reporter gene *gusA* it was found to drive expression in a variety of rice tissues and in cell suspensions of other monocot species following introduction by particle bombardment. Therefore, the *GOS2* promoter is potentially useful for genetic engineering of monocots. A DNA-binding activity from rice, termed rice ASF-1, with similar binding specificity as the cloned tobacco transcription factor TGA-1a, was found to bind to a TGACG sequence motif in the *GOS2* promoter. Possible roles for rice ASF-1 in the transcriptional activation of the *GOS2* promoter are discussed.

Introduction

Many plant genes examined to date are regulated by various endogenous or environmental factors. Extensively studied examples are light-regulated and leaf-specific genes (Cuozzo *et al.*, 1988) and seed-specific genes (Goldberg *et al.*, 1989). In contrast, relatively little is known about plant genes that are ubiquitously expressed. Promoters that are active in many different plant tissues and that have been analysed in some detail are either from viral or bacterial origin: the 35S promoter from cauliflower mosaic virus (CaMV) (Benfey *et al.*, 1989) and promoters of the octopine synthase (*ocs*) (Kononowicz *et al.*, 1992) and nopaline synthase (*nos*) (An *et al.*, 1988) genes from the *Agrobacterium tumefaciens* T-DNA.

Sequence elements important for high-level expression of these promoters (Ellis *et al.*, 1987; Fang *et al.*, 1989; Mitra and An, 1989) contain the core motif TGACG. *Trans*-acting factors binding to this motif have been characterized (Fromm *et al.*, 1989; Lam *et al.*, 1989, 1990; Lohmer *et al.*, 1991; Prat *et al.*, 1989; Singh *et al.*, 1989; Tokuhisa *et al.*, 1990) and cloned (Hartings *et al.*, 1989; Katagiri *et al.*, 1989; Singh *et al.*, 1990; Tabata *et al.*, 1989, 1991). Sequence analysis of the cDNA clones showed that the proteins belong to the leucine zipper (bZIP) class of transcription factors (McKnight, 1991). The tobacco bZIP protein TGA-1a was shown to function as a transcriptional activator in human and plant *in vitro* transcription systems (Katagiri *et al.*, 1990; Yamazaki *et al.*, 1990).

The CaMV 35S promoter and the *nos* promoter are widely used to monitor plant transformation procedures and to express useful traits. Since the number of promoters known to be active in many different plant tissues is very limited, isolation of other promoters, in particular ones that are active in monocots, is desirable. Knowledge about *cis*-acting elements and *trans*-acting factors essential for the activity of such promoters increases their potential for sophisticated applications and in addition enhances our understanding of transcription processes in plants.

Here we describe the isolation of the rice gene *GOS2*, which is expressed in all tissues examined and we present an initial analysis of its promoter region.

Results

Isolation of cDNA clones corresponding to genes expressed in a wide variety of rice tissues

In a search for genes expressed in many different rice tissues, a cDNA library made on poly(A) RNA from 2-week-old rice seedlings (*Oryza sativa*, variety IR36) was differentially screened with cDNA probes made on shoot and root mRNA. Three cDNA clones (COS2, COS7 and COS8) that hybridized strongly with both probes were further characterized. The three clones did not cross-hybridize, nor was there any sequence homology.

Figure 1 shows that the steady state mRNA levels of the genes corresponding to these cDNA clones were similar in green shoots, etiolated shoots and roots of IR36 seedlings. Expression of these genes was also found in different parts of IR36 mature plants (Figure 2). In roots the expression

Received 5 December 1991; revised 9 April 1992.

*For correspondence at Center for Phytotechnology RUL/TNO, Nonnensteeg 3, 2311 VJ Leiden, The Netherlands (fax +31 71275088).

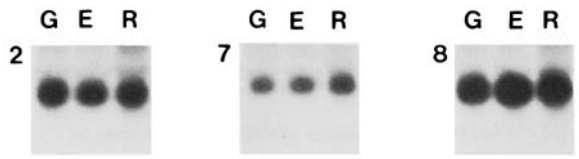

Figure 1. Expression patterns of COS2, COS7 and COS8 mRNAs in rice seedlings.
Gel blots containing total RNA isolated from green shoot tissue (G), etiolated shoot tissue (E) and roots (R) from 2-week-old IR36 seedlings were hybridized with ^{32}P-labelled COS2, COS7 or COS8 cDNA clones.

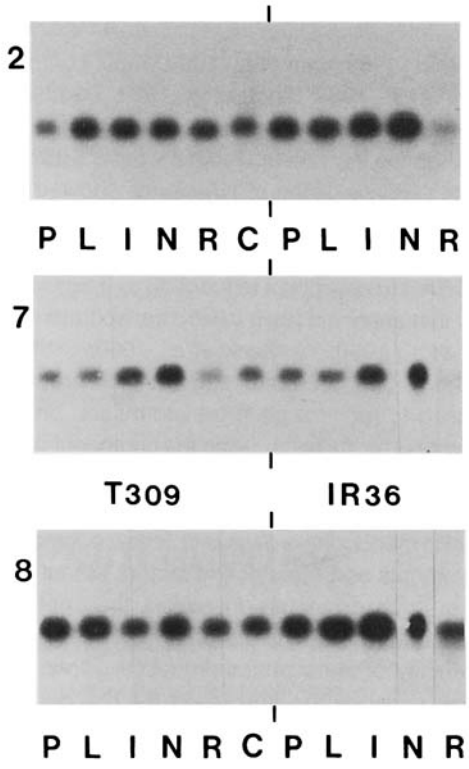

Figure 2. Expression patterns of COS2, COS7 and COS8 mRNAs in mature rice plants.
Gel blots containing total RNA isolated from panicles (P), leaves (L), internodes (I), nodes (N) and roots (R) from 3-month-old plants of the variety IR36 or T309 and RNA isolated from a T309 cell suspension (C) were hybridized with ^{32}P-labelled COS2, COS7 or COS8 cDNA clones.

of genes corresponding to COS2 and COS7 was somewhat lower. Similar mRNA levels were found in different plant parts and in a cell suspension of another variety, Japonica T309 (Figure 2). The COS2, COS7 and COS8 steady state mRNA levels were roughly two to four times lower than the steady state mRNA level of ribulose-1,5-bisphosphate carboxylase small subunit (*rbcS*) genes in green tissue (data not shown). The lengths of the mRNAs hybridizing to COS2, COS7 and COS8 were approximately 800 bases, 900 bases and 900 bases, respectively.

Subsequently, the copy numbers of the genes corresponding to COS2, COS7 and COS8 were estimated by Southern blot hybridization. DNA blots containing IR36

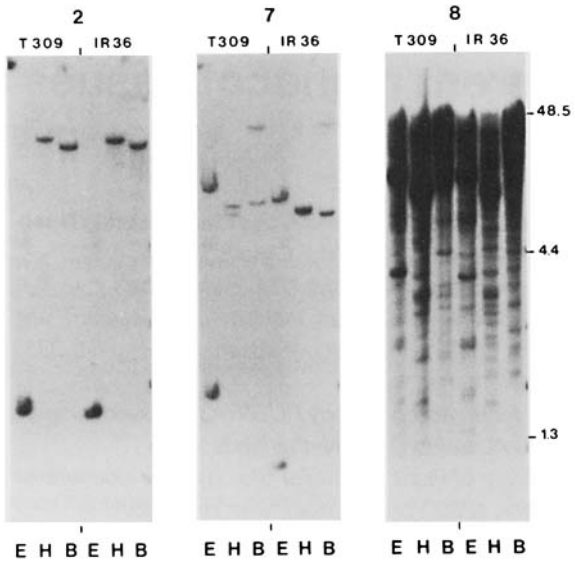

Figure 3. Copy number estimation of genes corresponding to COS2, COS7 and COS8 cDNAs by DNA blot hybridization.
Gel blots containing IR36 or T309 rice DNA digested with *Eco*RI (E), *Hin*dIII (H) or *Bam*HI (B) were hybridized with ^{32}P-labelled COS2, COS7 or COS8 cDNA clones. The sizes of marker fragments are indicated in kilobasepairs.

and T309 genomic DNA were hybridized with the three cDNA clones. The DNA was digested with *Eco*RI, *Bam*HI or *Hin*dIII, which have no recognition sites in the cDNA inserts. Figure 3 shows that hybridization with COS2 resulted in one band in each lane, which makes it likely that COS2 corresponds to a single copy gene. Two bands were visible in each lane after hybridization with COS7, and at least thirty bands after hybridization with COS8. Thus, COS7 mRNA is probably encoded by one or two genes, while COS8 mRNA corresponds to a large multigene family.

Isolation of the gene encoding COS2 mRNA

The high steady state COS2 mRNA level and the single copy character of the corresponding gene *GOS2* suggested that *GOS2* contains a strong promoter, which is potentially useful for genetic engineering purposes. Therefore, COS2 homologous sequences were isolated from an IR36 genomic library. Four overlapping clones with identical maps for four restriction enzymes were obtained, supporting the hypothesis that *GOS2* is a single copy gene. One of these clones was further analysed.

A restriction map of the insert is shown in Figure 4. The sequence of the region homologous to COS2 is shown in Figure 5. Since the cDNA clone is not a full-length clone, the nucleotide sequence of the 5′ end of the mRNA was determined by primer extension in the presence of dideoxynucleotides. The mRNA sequence and the corresponding genomic sequence are identical, which is

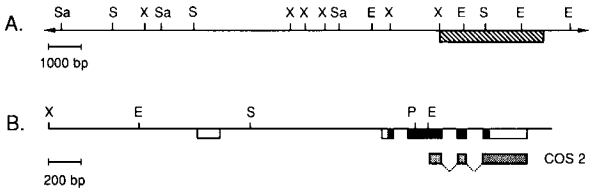

Figure 4. Structure of the rice gene *GOS2*.
(a) Restriction map of the insert of the genomic clone containing *GOS2*. Restriction sites are indicated for *Sal*I (Sa), *Sac*I (S), *Xba*I (X), *Eco*RI (E) and *Pst*I (P).
(b) The DNA fragment that was sequenced (indicated by a hatched bar in a) is shown in detail. Exons are indicated by bars: black bars represent the protein coding region. The stippled bars represent sequences present in the cDNA clone COS2.

additional evidence that *GOS2* is a single copy gene. The gene contains four introns of 998 bp, 94 bp, 90 bp and 105 bp. The largest intron is located in the leader sequence (Figure 4b). The *GOS2* mRNA contains an open reading frame of 345 nucleotides, located between the first AUG codon at position 2200 (in the genomic sequence of Figure 5) and an UGA codon at position 2827. The open reading frame potentially encodes a protein of 115 amino acids (molecular mass 12 732).

The encoded GOS2 protein shows 75% similarity (56% identity) to the deduced amino acid sequence of a mouse cDNA clone (RP4–17) and RP4–17 homologous sequences have also been found in human and bovine DNA (Schedl, personal communication). The GOS2 protein is also 52% identical to the protein encoded by the *sui1* gene of yeast, which is a suppressor locus of translational initiator mutants (Donahue, personal communication). This suggests that the GOS2 gene product is part of the translation machinery.

The 5′ ends of the mRNA were determined by primer extension and are located at positions 1024, 1025 and 1026 (Figure 5). At 33 bp upstream of the major start sites a TATA box is present.

The cDNA clone COS2 contains a poly(A) tail of 53 bp. A sequence AATAAt with homology to the poly(A) addition consensus (AATAAA) is present 27 bp upstream of the poly(A) tail. In the *GOS2* genomic sequence 16 bp downstream of the poly(A) tail the sequence TGTGTTgg is present with homology to the consensus sequence YGTGTTYY, which plays a role in mRNA 3′ end formation (McLauchlan *et al.*, 1985).

Transient expression of the gusA *reporter gene under control of the* GOS2 *promoter*

We tested whether the *GOS2* upstream region is able to drive the expression of a reporter gene. The construct used comprises the *Xba*I–*Eco*RI fragment, containing the *GOS2* upstream region, including the first and second intron, translationally fused to the coding region of the *Escherichia coli* β-glucuronidase reporter gene (*gusA*) (Figure 6). The 3′ end of the 0′ gene of the T_R-DNA of *Agrobacterium tumefaciens* functions as a poly(A) addition sequence. Introns were included in the construct since enhancement of gene expression by introns has been reported (Luehrsen and Walbot, 1991).

Ten-day-old seedlings of the Japonica variety T309 and cell suspensions of the rice varieties IR54 and Nipponbare were used for particle bombardment experiments. Expression of the construct was detected in leaf tissue (Figure 7a–c), root tissue (Figure 7d–f), coleoptile (Figure 7g) and cell suspension (Figure 7h and i). The *GOS2* promoter was also active in a cell suspension of the monocot barley (*Hordeum vulgare*, variety Igri, Figure 7j). In addition, expression was detected in cell suspensions of maize (*Zea mays*, variety Black Mexican Sweet) (van der Mark, unpublished results) and grass (*Lolium perenne*) (van der Maas, personal communication). Control experiments with a promoterless construct (pSK+-GUS-01) gave no detectable expression (data not shown). Thus, the 5′ DNA fragment used contains sequences that can drive expression of a reporter gene in different rice tissues. In addition, the *GOS2* promoter is active in other monocots.

Rice nuclear protein binds to the GOS2 *promoter*

Transcription is regulated through sequence-specific binding of transcription factors to *cis*-acting DNA elements within the promoter. Therefore, we investigated the binding of rice nuclear proteins to the *GOS2* promoter. The region upstream of the TATA box was subdivided into four DNA fragments, as shown in Figure 8a. Each fragment was incubated with 10 μg of nuclear protein isolated from rice leaves (Figure 8b). A distinct complex was detected with fragment II and some binding to I and IV could be observed, but this resulted in smears rather than distinct bands. No complex could be found with fragment III using the same amount of nuclear protein.

The rice GOS2 *binding activity has similar specificity as a cloned transcription factor (TGA-1a) from tobacco*

Inspection of the sequence of *GOS2* promoter fragment II, revealed that it contains a TGACG motif at positions 435–439 (Figure 5). This sequence is present as a core consensus in recognition sites for the cloned plant nuclear factors TGA-1a and TGA-1b (Katagiri *et al.*, 1989), OCSBF-1 (Singh *et al.*, 1990), HBP-1a and HBP-1b (Tabata *et al.*, 1991) and Opaque-2 (Lohmer *et al.*, 1991).

First, we investigated whether rice nuclear extract contains proteins with a similar binding specificity as TGA-1a. As a probe we used a tetramerized TGA-1a binding site termed 4A1. The binding site is derived from the CaMV 35S promoter, where it is located between positions −82

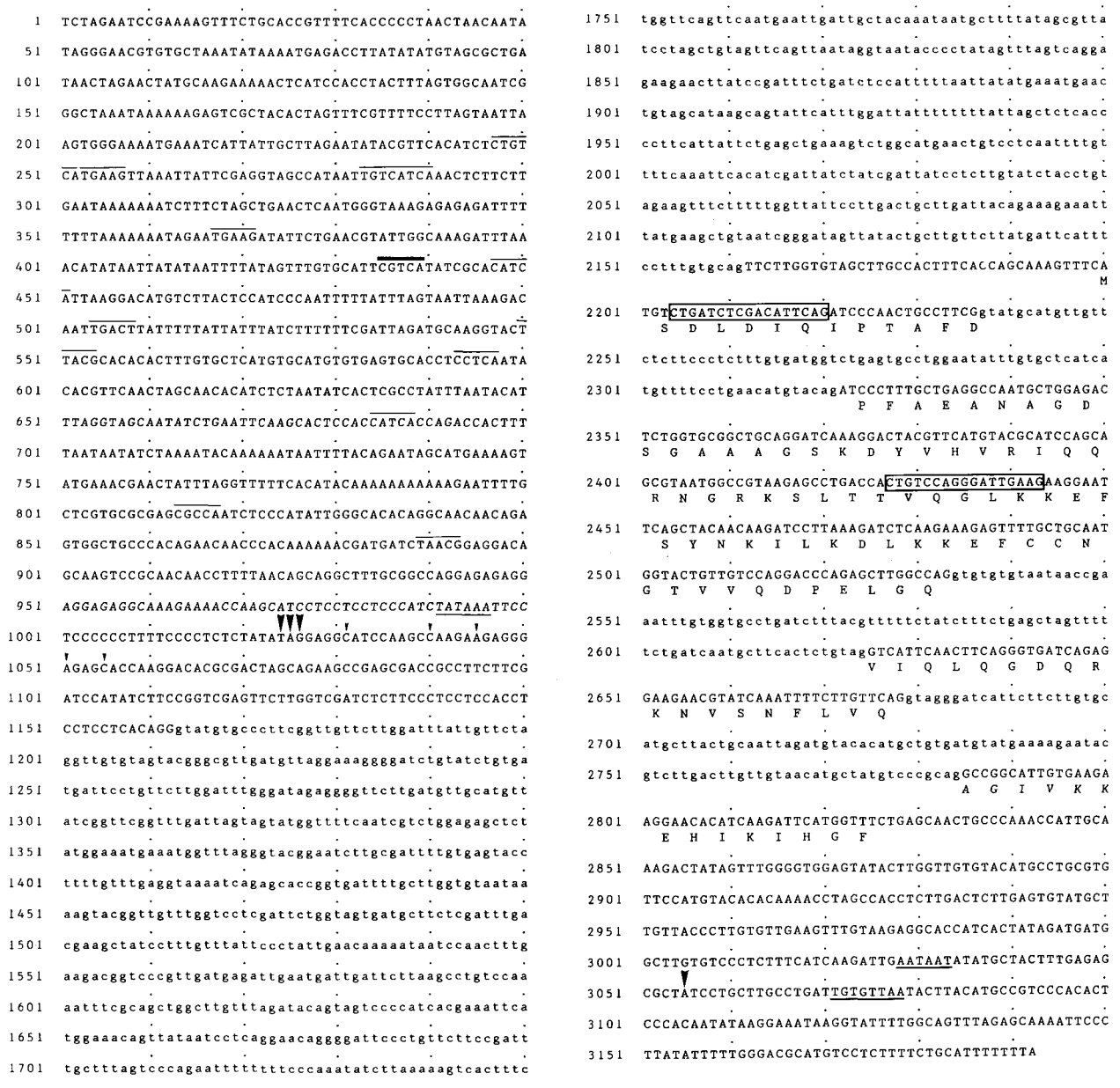

Figure 5. Nucleotide sequence of the rice gene *GOS2* and the amino acid sequence of the protein encoded by the largest open reading frame. The transcription start sites and poly(A) addition site are indicated with arrowheads. Intron sequences are shown in lower case letters. The putative TATA box and signals involved in 3′ processing of the mRNA are underlined. TGACG(-like) motifs are overlined. The sequences complementary to the 17-mer primers used for mRNA sequencing are boxed.

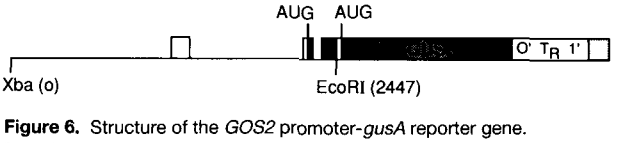

Figure 6. Structure of the *GOS2* promoter-*gusA* reporter gene. The *Xba*I–*Eco*RI fragment of *GOS2* containing 1024 bp of the upstream region, the first and second intron (exons are indicated by bars) was translationally fused to the *gusA* coding region and the T_R0'–T_R1' dual terminator of *Agrobacterium tumefaciens* containing the 3′ part of the T_R1' coding region (light-stippled box).

and −62 with respect to the transcription start site. Incubation of 4A1 with rice nuclear extract resulted in complexes (Figure 8b, lane 5) which are very similar to the complexes obtained with cloned TGA-1a purified from *E. coli* (Figure 8c, lane 15). Thus rice contains proteins that bind to 4A1. The binding activity will be referred to hereafter as rice ASF-1 in analogy to the tobacco 4A1 (as-1) binding activity ASF-1 (Lam *et al.*, 1989).

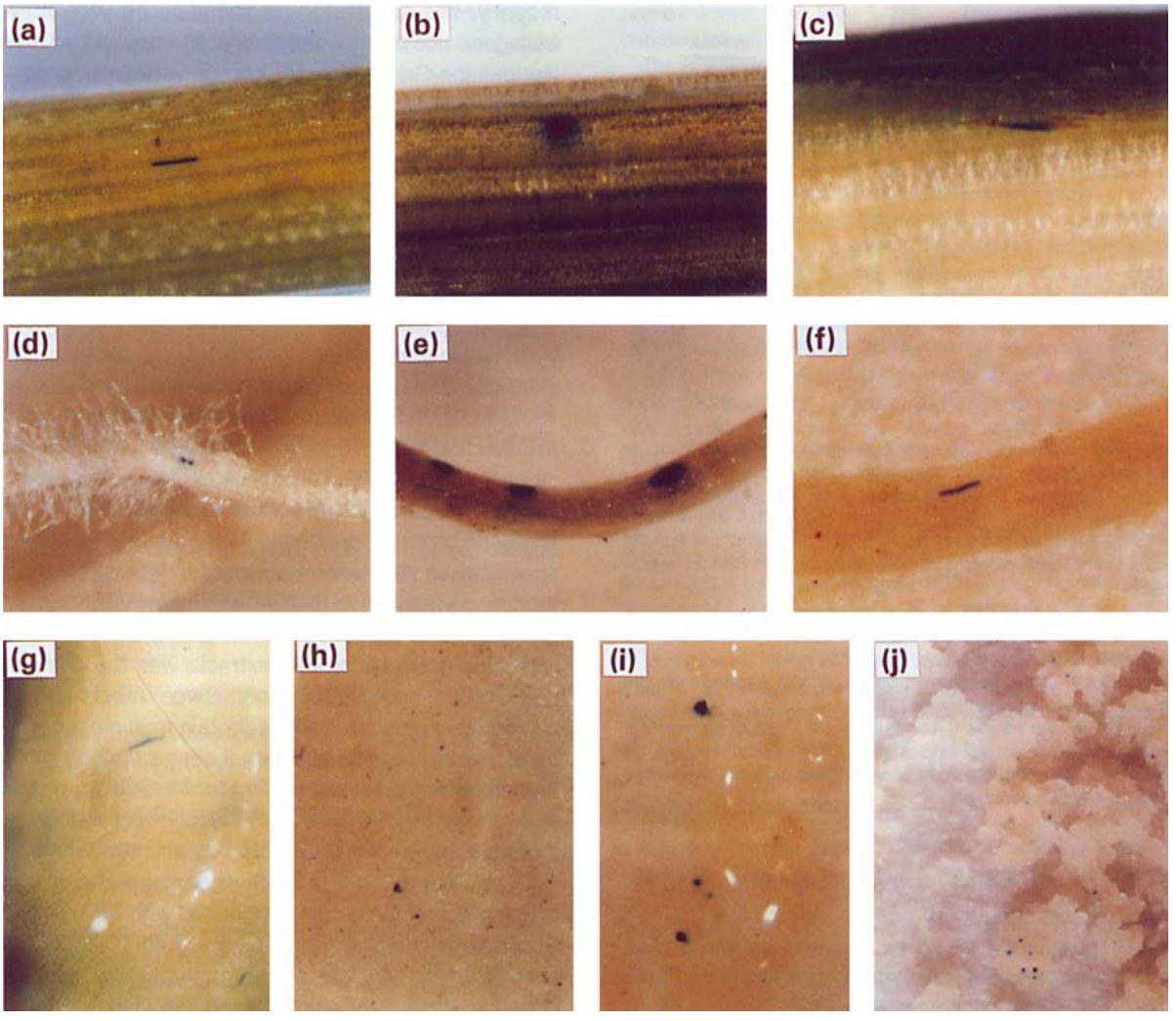

Figure 7. Histochemical analysis of *GOS2* promoter activity.
GUS activity in leaf (a,b,c), root (d,e,f) and coleoptile (g) of the rice variety T309, in cell suspension of the rice varieties IR54 (h) and Nipponbare (i) and in cell suspension of the barley variety Igri (j) 2 days after particle bombardment with a plasmid containing the *GOS2* promoter–*gusA* reporter fusion gene shown in Figure 6.

Subsequently, we determined whether the tobacco transcription factor TGA-1a binds to the *GOS2* promoter. Each *GOS2* promoter fragment was tested for binding to TGA-1a protein, which was purified to homogeneity from *E. coli*. A complex was observed in all cases, but the affinity of TGA-1a for the four fragments varied considerably (Figure 8c). The highest affinity was observed for fragment II, which has the TGACG sequence. It is likely that TGA-1a binds to this sequence, since it is homologous to TGA-1a binding sites in other promoters (Katagiri *et al.*, 1989). TGA-1a probably binds to TGACG-like motifs in the other *GOS2* subfragments I, III and IV. Twelve of these motifs, each containing one mismatch at different positions, are indicated in Figure 5.

Using the same amount of TGA-1a protein more 4A1 was shifted compared with fragment II (Figure 8c, lane 15), suggesting a higher affinity of TGA-1a for 4A1. However one must take into consideration that 4A1 contains eight TGACG motifs.

To show more directly that rice ASF-1 binds to the TGACG motif in fragment II, 4A1 and 4A3 were used as competitors. In the 4A3 tetramer the TGACG motifs contain two basepair mutations, that cause severe reduction of TGA-1a binding (de Pater, unpublished results). Otherwise the sequences of 4A1 and 4A3 are identical. Addition of a 500-fold molar excess of 4A1 to the reaction mixture competed for binding, whereas addition of the same amount of 4A3 had no effect (Figure 8d). Thus, it can be concluded that rice ASF-1 binds to the TGACG motif in fragment II of the *GOS2* promoter. In addition, it has similar specificity as TGA-1a for binding to 4A1 and 4A3. Although TGA-1a also forms complexes with fragments I, III and IV,

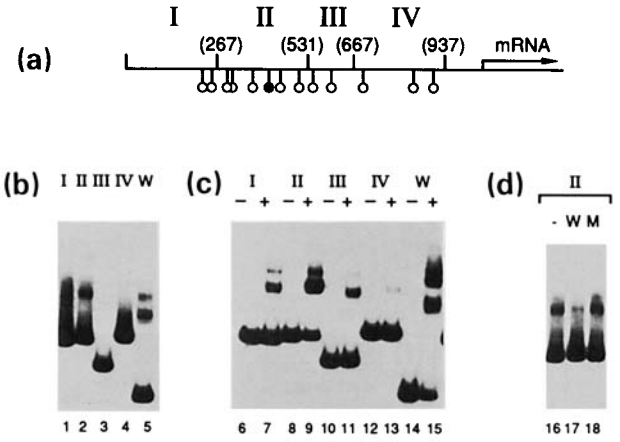

Figure 8. Rice contains a *GOS2* promoter-binding protein with similar binding specificity as the tobacco transcription factor TGA-1a.
(a) Promoter fragments of *GOS2* that were used for gel retardation (I, II, III, IV). The numbers indicate the end positions of the DNA fragments in relation to the sequence shown in Figure 5. The position of the TGACG motif (closed circle) and similar motifs with one mismatch (open circles) are indicated.
(b) Gel retardation of the *GOS2* promoter fragments shown in (a) and a tetramer of the TGA-1a binding site (4A1) from the CaMV 35S promoter (W) with rice nuclear extracts.
(c) Gel retardation of the fragments used in (b) after incubation in the absence (−) or presence (+) of purified TGA-1a.
(d) Gel retardation with fragment II and rice nuclear extract. Fragment II was incubated with rice nuclear extract without further additions (lane 16) or in the presence of a 500-fold molar excess of wild-type 4A1 competitor (W) or mutant 4A3 competitor (M).

no binding of rice ASF-1 to these fragments could be detected with the amount of nuclear extract used. Weak binding to fragment I could be masked by the smear (Figure 8, lane 1) and the affinity for fragments III and IV is probably too low to detect with crude nuclear extract. Thus, our interpretation of these observations is that the binding specificities of rice ASF-1 and TGA-1a for the *GOS2* promoter subfragments are similar, but detection of DNA-protein complexes using purified TGA-1a instead of crude nuclear protein extracts is more sensitive.

Discussion

A single copy gene of rice, termed *GOS2* was isolated. The steady state mRNA level in different plant parts is comparable with the steady state mRNA level of *rbcS* genes in green rice tissues. Since the *GOS2* mRNA is encoded by a single gene, we assumed that the promoter must be active, unless the mRNA is relatively stable. The promoter was able to drive the expression of the reporter gene *gusA* in all rice tissues tested. This observation is in agreement with the expression pattern of the endogenous *GOS2* gene. In addition, the *GOS2* promoter is active in the monocots barley, maize and grass and appears to be more active than the CaMV 35S promoter in monocots in transient assays (van der Mark, unpublished results). More recently, this construct was found to be very active in transgenic rice plants (Hensgens *et al.*, unpublished results). Besides the CaMV 35S promoter, several other promoters have been shown to drive expression of reporter genes in different monocot tissues (Dekeyser *et al.*, 1989; Klein *et al.*, 1989; McElroy *et al.*, 1990). Isolation of additional promoters active in monocots is not redundant, since improvement of plants by genetic engineering requires the expression of a multitude of genes.

We found nuclear rice protein(s), termed rice ASF-1, that binds to the *GOS2* promoter. Several promoters known to be active in many different plant tissues, contain TGACG elements that are important for expression (Ellis *et al.*, 1987; Fang *et al.*, 1989; Mitra and An, 1989) and that bind *trans*-acting factors (Fromm *et al.*, 1989; Lam *et al.*, 1989, 1990; Prat *et al.*, 1989; Tokuhisa *et al.*, 1990). This prompted us to test whether rice ASF-1 binds to a TGACG motif in the *GOS2* promoter. Indeed, rice ASF-1 has a similar binding specificity as the cloned transcription factor TGA-1a from tobacco, which binds to TGACG motifs in the CaMV 35S promoter. The only distinct complex that was observed with crude nuclear protein extracts was the complex of rice ASF-1 and fragment II. Since most eukaryotic promoters examined to date contain multiple recognition motifs for transcription factors, it seems unlikely that only one protein species would bind to the *GOS2* promoter *in vivo*. We assume that rice ASF-1 is relatively abundant and therefore can be easily detected by gel retardation.

The size of the 4A1-rice ASF-1 complex is comparable with the size of the complex obtained with TGA-1a (Figure 8, compare lane 5 with lane 15) and with TGACG-binding activity from pea (de Pater, unpublished results) as observed in gel retardation. Recently a wheat cDNA clone was isolated encoding a protein with 51% homology to TGA-1a (Tabata *et al.*, 1991). These observations indicate that this class of DNA binding proteins is conserved among different plant species including monocots and dicots. Additional evidence for their widespread occurrence throughout the plant kingdom is the presence of functionally important TGACG motifs in promoters of several genes (Fromm *et al.*, 1989; Lam *et al.*, 1989, 1990) that are active in a variety of plant species.

The CaMV 35S promoter was shown to have a modular organization. It consists of different *cis*-elements that each confer a specific expression pattern (Benfey *et al.*, 1990). A subfragment containing the TGA-1a binding site confers expression principally in root (Benfey *et al.*, 1989). It also interacts synergistically with another domain of the 35S promoter to activate expression in leaf (Benfey *et al.*, 1990). High expression in leaf is also conferred by multimers of the TGA-1a binding site (Lam and Chua, 1990). The *GOS2* promoter activity may result from one of the mechanisms described. The multiple TGACG-like motifs might act cooperatively to drive expression in different organs.

Alternatively, the perfect TGACG motif in fragment II might be the only functional ASF-1 binding site *in vivo*, and determine expression in root, whereas other unknown motifs may act in concert with the ASF-1 binding site to give expression in other organs.

Experimental procedures

Plant material

Rice plants (*O. sativa*, Indica variety IR36 or Japonica variety T309) were grown as described (de Pater *et al.*, 1990). Rice cell suspensions were grown on AA medium (Müller and Grafe, 1978) and barley cell suspension on the medium described by Lührs and Lörz (1988), in the dark, with weekly transfers to fresh medium.

RNA and DNA isolation and blot hybridization

RNA and DNA isolation (de Pater *et al.*, 1990) and blot hybridization (Memelink *et al.*, 1987) were done as described. Twenty micrograms of total RNA were applied in each lane of the RNA gel blots and 10 μg of DNA in each lane of the DNA gel blots.

Screening of cDNA and genomic libraries

A cDNA library from 2-week-old IR36 seedlings (de Pater *et al.*, 1990) was differentially screened as described (Memelink *et al.*, 1987). A rice genomic library (de Pater *et al.*, 1990) was screened for sequences hybridizing to the cDNA clone COS2.

Nucleic acid sequencing

The cDNA insert of COS2 and the *Xba*I fragment containing *GOS2* were subcloned in M13tg130/131 or Bluescript vector SKM13+ (Stratagene). Overlapping subclones and/or exonuclease III deletions were produced and sequenced by the dideoxy chain-termination method. The 5′ end of the mRNA was sequenced by the dideoxy chain-termination method essentially as described (Geliebter, 1987) using 20 μg of poly(A) RNA and the 17-mer primers shown in Figure 5.

Nucleotide sequence data were collected, assembled and analysed with a VAX computer fitted with the Genetics Computer Group Sequence Analysis Software Package (Devereux *et al.*, 1984).

Transient expression assays

The *GOS2* promoter was translationally fused to the *gusA* coding region by cloning the *Xba*I–*Eco*RI fragment of *GOS2* (positions 1–2447 in Figure 5) into a plasmid containing the *gusA* coding region and the *Agrobacterium tumefaciens* T_R0'–T_R1' dual terminator (pSK+-GUS-01) (van der Zaal *et al.*, 1991). Plasmid DNA was delivered to intact tissues or suspension cells by high velocity microprojectiles using a particle delivery apparatus (Klein *et al.*, 1988). For each microparticle preparation, 2 μg of gold powder (Aldrich, spherical, average diameter 1.5–3.0 μm) was resuspended in 20 μl of H_2O and mixed with 6 μl of plasmid DNA (1 μg μl^{-1}), 30 μl of a 16% polyethylene glycol 1500 solution and 3 μl of a 100 mM spermidine-trihydrochloride solution. Particles were allowed to settle, and 40 μl of the solution was removed. Two microlitres of the resulting particle suspension was used per bombardment. Following bombardment the tissues were incubated 20–40 h in the dark.

The *in situ* assays for GUS activity were performed as described (Wang *et al.*, 1988).

Gel retardation assays

TGA-1a, encoded by a tobacco cDNA clone (Katagiri *et al.*, 1989), was purified to homogeneity from *Escherichia coli* as described (Katagiri *et al.*, 1990). Nuclear protein extract was prepared from green tissue of 2-week-old seedlings essentially as described (Green *et al.*, 1989), except that the tissue was powdered in liquid N_2 prior to homogenizing in buffer and that the crude nuclear extract was adjusted to 55% saturation of ammonium sulphate to precipitate the proteins. Gel retardation experiments were performed essentially as described previously (Green *et al.*, 1987). *GOS2* promoter fragments I–IV were subcloned into Bluescript IIKS+. I is a 267 bp *Xba*I–*Taq*I fragment; II is a 264 bp *Taq*I fragment; III is a 136 bp *Taq*I–*Eco*RI fragment; IV is a 270 bp *Eco*RI–*Hae*III fragment. 4A1 is a tetramer of wild-type TGA-1a binding sites (as-1), containing a 21 bp region of the 35S promoter (position −82 to −62 with respect to the transcription start site) and 4A3 is a tetramer of mutant TGA-1a binding sites (as-1c) as described by Lam *et al.* (1989). Plasmids were digested with *Xba*I and *Hin*dIII (*GOS2* fragments) or with *Xho*I and *Hin*dIII (4A1) and 3′ end-labelled using Klenow enzyme and [α-^{32}P]dCTP. Labelled fragments were isolated from polyacrylamide gels. For each reaction 0.9 μg of purified TGA-1a or 10 μg of rice nuclear protein and 0.1 ng of probe were used.

Acknowledgements

We thank Thomas Donahue (Indiana University, Bloomington, USA), Heleen van der Maas (CPO, Wageningen, The Netherlands) and Andreas Schedl (German Cancer Research Center, Heidelberg, Germany) for information on their results prior to publication, Ellen van Os-Ruigrok and Liesbeth Bakker for technical assistance, Johan Memelink for many valuable discussions and comments on the manuscript and Peter Hock and the Biology Photographic Service for preparation of the figures. BSdP and LAMH were financially supported by the Rockefeller Foundation.

References

An, G., Costa, M.A., Mitra, A., Ha, S.-B. and Márton, L. (1988) Organ-specific and developmental regulation of the nopaline synthase promoter in transgenic tobacco plants. *Plant Physiol.* **88**, 547–552.

Benfey, P.N., Ren, L. and Chua, N.-H. (1989) The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. *EMBO J.* **8**, 2195–2202.

Benfey, P.N., Ren, L. and Chua, N.-H. (1990) Tissue-specific expression from CaMV 35S enhancer subdomains in early stages of plant development. *EMBO J.* **9**, 1677–1684.

Cuozzo, M., Kay, S.A. and Chua, N.-H. (1988) Regulatory circuits of light-responsive genes. In *Plant Gene Research* (Dennis, E.S., Hohn, B., Hohn, Th., King, P.J., Schell, J. and Verma, D.P.S., eds). Wien: Springer-Verlag, pp. 131–153.

Dekeyser, R., Claes, B., Marichal, M., Van Montagu, M. and Caplan, A. (1989) Evaluation of selectable markers for rice transformation. *Plant Physiol.* **90**, 217–223.

Devereux, J., Haeberli, P. and Smithies, O. (1984) A comprehensive set of sequence analysis programs of the VAX. *Nucl.*

Acids Res. **12**, 387–395.

Ellis, J.G., Llewellyn, D.J., Walker, J.C. Dennis, E.S. and Peacock, W.J. (1987) The *ocs* element: a 16 base pair palindrome essential for activity of the octopine synthase enhancer. *EMBO J.* **6**, 3203–3208.

Fang, R.-X., Nagy, F., Sivasubramaniam, S. and Chua, N.-H. (1989) Multiple *cis* regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. *Plant Cell*, **1**, 141–150.

Fromm, H., Katagiri, F. and Chua, N.-H. (1989) An octopine synthase enhancer element directs tissue-specific expression and binds ASF-1, a factor from tobacco nuclear extracts. *Plant Cell*, **1**, 977–984.

Geliebter, J. (1987) Dideoxynucleotide sequencing of RNA and uncloned cDNA. *Focus*, **9**, 5–9.

Goldberg, R.B., Barker, S.J. and Perez-Grau, L. (1989) Regulation of gene expression during plant embryogenesis. *Cell*, **56**, 149–160.

Green, P.J., Kay, S.A. and Chua, N.-H. (1987) Sequence-specific interactions of a pea nuclear factor with light-responsive elements upstream of the rbcS-3A gene. *EMBO J.* **6**, 2543–2549.

Green, P.J., Kay, S.A., Lam, E. and Chua, N.-H. (1989) *In vitro* DNA footprinting. In *Plant Molecular Biology Manual*, Suppl. II (Gelvin, S. and Schilperoort, R.A., eds). Dordrecht: Kluwer Academic, pp. 303–310.

Hartings, H., Maddaloni, M., Lazzaroni, N., Di Fonzo, N., Motto, M., Salamini, F. and Thompson, R. (1989) The *O2* gene which regulates zein deposition in maize endosperm encodes a protein with structural homologies to transcriptional activators. *EMBO J.* **8**, 2795– 2801.

Katagiri, F., Lam, E. and Chua, N.-H. (1989) Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. *Nature*, **340**, 727–730.

Katagiri, F., Yamazaki, K.-I., Horikoshi, M., Roeder, R.G. and Chua, N.-H. (1990) A plant DNA-binding protein increases the number of active preinitiation complexes in a human in vitro transcription system. *Genes Dev.* **4**, 1899–1909.

Klein, T.M., Gradziel, T., Fromm, M.E. and Sanford, J.C. (1988) Factors influencing gene delivery into *Zea mays* cells by high-velocity microprojectiles. *Bio/Technology*, **6**, 559–563.

Kononowicz, H., Wang, Y.E., Habeck, L.L. and Gelvin, S.B. (1992) Subdomains of the octopine synthase upstream activating element direct cell-specific expression in transgenic tobacco plants. *Plant Cell*, **4**, 17–27.

Lam, E. and Chua, N.-H. (1990) GT-1 binding site confers light responsive expression in transgenic tobacco. *Science*, **248**, 471–474.

Lam, E., Benfey, P.N., Gilmartin, P.M., Fang, R.-X. and Chua, N.-H. (1989) Site-specific mutations alter *in vitro* factor binding and change promoter expression pattern in transgenic plants. *Proc. Natl Acad. Sci. USA*, **86**, 7890–7894.

Lam, E., Katagiri, F. and Chua, N.-H. (1990) Plant nuclear factor ASF-1 binds to an essential region of the nopaline synthase promoter. *J. Biol. Chem.* **265**, 9909–9913.

Lohmer, S., Maddeloni, M., Motto, M., Di Fonzo, N., Hartings, H., Salamini, F. and Thompson, R.D. (1991) The maize regulatory locus *Opaque-2* encodes a DNA-binding protein which activates the transcription of the *b-32* gene. *EMBO J.* **10**, 617–624.

Luehrsen, K.R. and Walbot, V. (1991) Intron enhancement of gene expression and the splicing efficiency of introns in maize cells. *Mol. Gen. Genet.* **225**, 81–93.

Lührs, R. and Lörz, H. (1988) Initiation of morphogenic cell-suspension and protoplast cultures of barley (*Hordeum vulgare* L.). *Planta*, **175**, 71–81.

McElroy, D., Zhang, W., Cao, J. and Wu, R. (1990) Isolation of an efficient actin promoter for use in rice transformation. *Plant Cell*, **2**, 163–171.

McLauchlan, J., Gaffney, D., Whitton, J.L. and Clements, J.B. (1985) The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3′ termini. *Nucl. Acids Res.* **13**, 1347–1368.

McKnight, S.L. (1991) Molecular zippers in gene regulation. *Sci. Am.* **264**, 32–39.

Memelink, J., Hoge, J.H.C. and Schilperoort, R.A. (1987) Cytokinin stress changes the developmental regulation of several defence-related genes in tobacco. *EMBO J.* **6**, 3579–3583.

Mitra, A. and An, G. (1989) Three distinct regulatory elements comprise the upstream promoter region of the nopaline synthase gene. *Mol. Gen. Genet.* **215**, 294–299.

Müller, A.J. and Grafe, R. (1978) Isolation and characterization of cell lines of *Nicotiana tabacum* lacking nitrate reductase. *Mol. Gen. Genet.* **161**, 67–76.

de Pater, S., Hensgens, L.A.M. and Schilperoort, R.A. (1990) Structure and expression of a light-inducible shoot-specific rice gene. *Plant Mol. Biol.* **15**, 399–406.

Prat, S., Willmitzer, L. and Sánchez-Serrano, J.J. (1989) Nuclear proteins binding to a cauliflower mosaic virus 35S truncated promoter. *Mol. Gen. Genet.* **217**, 209–214.

Singh, K., Tokuhisa, J.G., Dennis, E.S. and Peacock, W.J. (1989) Saturation mutagenesis of the octopine synthase enhancer: Correlation of mutant phenotypes with binding of a nuclear protein factor. *Proc. Natl Acad. Sci. USA*, **86**, 3733–3737.

Singh, K., Dennis, E.S., Ellis, J.G., Llewellyn, D.J., Tokuhisa, J.G., Wahleithner, J.A. and Peacock, W.J. (1990) OCSBF-1, a maize ocs enhancer binding factor: isolation and expression during development. *Plant Cell*, **2**, 891–903.

Tabata, T., Takase, H., Takayama, S., Mikami, K., Nakatsuka, A., Kawata, T., Nakayama, T. and Iwabuchi, M. (1989) A protein that binds to a cis-acting element of wheat histone genes has a leucine zipper motif. *Science*, **245**, 965–967.

Tabata, T., Nakayama, T., Mikami, K. and Iwabuchi, M. (1991) HBP-1a and HBP-1b: leucine zipper-type transcription factors of wheat. *EMBO J.* **10**, 1459–1467.

Tokuhisa, J.G., Singh, K., Dennis, E.S. and Peacock, W.J. (1990) A DNA-binding protein factor recognizes two binding domains within the octopine synthase enhancer element. *Plant Cell*, **2**, 215–224.

Wang, Y.-C., Klein, T.M., Fromm, M., Cao, J., Sanford, J.C. and Wu, R. (1988) Transient expression of foreign genes in rice, wheat and soybean cells following particle bombardment. *Plant Mol. Biol.* **11**, 433–439.

Yamazaki, K.-I., Katagiri, F., Imaseki, H. and Chua, N.-H. (1990) TGA1a, a tobacco DNA-binding protein, increases the rate of initiation in a plant *in vitro* transcription system. *Proc. Natl Acad. Sci. USA*, **87**, 7035–7039.

van der Zaal, E.J., Droog, F.N.J., Boot, C.J.M., Hensgens, L.A.M., Hoge, J.H.C., Schilperoort, R.A. and Libbenga, K.R. (1991) Promoters of auxin-induced genes from tobacco can lead to auxin-inducible and root tip-specific expression. *Plant Mol. Biol.* **16**, 983–998.

EMBL Data Library accession number X51910 (*Oryza sativa, GOS2* gene).