The EMBO Journal vol.9 no.6 pp.1717 - 1726, 1990

Mutation of either G box or I box sequences profoundly affects expression from the *Arabidopsis rbcS-1A* promoter

Robert G.K.Donald and Anthony R.Cashmore

Plant Science Institute, Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA

Communicated by M.van Montagu

A deletion analysis of the *Arabidopsis thaliana rbcS-1A* promoter defined a 196 bp region (−320 to −125) sufficient to confer light-regulated expression on a heterologous *Arabidopsis* alcohol dehydrogenase (*Adh*) reporter gene in transgenic *Nicotiana tabacum* (tobacco) leaves. This region, which contains DNA sequences I, G and GT boxes, with homology to other ribulose-1,5-bisphosphate carboxylase small subunit (RBCS) gene promoter sequences, directed expression independent of orientation and relative position in the *Adh* promoter. Site-specific mutagenesis of these conserved sequences and subsequent expression analysis in transgenic tobacco showed that both G box and I box mutations in the context of the full (−1700 to +21) *rbcS-1A* promoter substantially reduced the expression of *Adh* and β-glucuronidase (GUS) reporter genes. The G box has previously been shown to specifically bind *in vitro* a factor isolated from nuclear extracts of tomato and *Arabidopsis*. This factor (GBF) is distinct from the factor GT-1 which binds to adjacent GT boxes in the pea *rbcS-3A* promoter. Multiple mutations in putative *Arabidopsis rbcS-1A* promoter GT boxes had no pronounced affect on expression, possibly due to a redundancy of these sites. Experiments in which *rbcS-1A* promoter fragments were fused to truncated 35S CaMV (cauliflower mosaic virus) promoter−GUS reporter constructs showed that *cis*-acting CaMV promoter elements could partially restore expression to G-box-mutated *rbcS-1A* sequences.

Key words: Adh reporter/chimeric genes/GBF/plant transformation/RBCS

Introduction

The structure and organization of light-regulated nuclear genes encoding the ribulose-1,5-bisphosphate carboxylase small subunit (RBCS) has been determined in a variety of plants including duckweed (Wimpee *et al.*, 1983), soybean (Berry-Lowe *et al.*, 1982), petunia (Dunsmuir *et al.*, 1983; Dean *et al.*, 1985a, 1987), pea (Cashmore, 1983; Coruzzi *et al.*, 1983, 1984; Timko *et al.*, 1985; Fluhr *et al.*, 1986), tomato (Pichersky *et al.*, 1986; Sugita *et al.*, 1987) and *Arabidopsis* (Timko *et al.*, 1987; Krebbers *et al.*, 1988). Individual members of these gene families often show a range of expression levels, and differences in developmental and tissue-specific expression (Dean *et al.*, 1985b, 1987; Sugita *et al.*, 1987; Fluhr *et al.*, 1986). Since regulation of RBCS expression is determined, at least in part, at the level of transcription (reviewed by Tobin and Silverthorne, 1985; Kuhlemeir *et al.*, 1987), much effort has been directed at characterizing promoter sequences mediating this expression.

To identify evolutionarily conserved sequences that may be important for RBCS gene expression, comparisons have been made in 5′ flanking regions of available RBCS gene sequences (Kuhlemeir *et al.*, 1987; Giuliano *et al.*, 1988; Manzara and Gruissem, 1988). In addition to the 'TATA' and 'CAAT' box sequences common to most eukaryotic promoters, at least three classes of conserved sequences have been identified. One class, present in multiple copies in all RBCS genes sequenced so far, has been functionally characterized in the pea *rbcS-3A* promoter. This sequence, which we will refer to as the GT box, has the broad sequence consensus (G/T-A/T-GTG-Pu-A/T-AA-A/T-Pu-A/T) and is bound specifically by a nuclear factor GT-1 (Green *et al.*, 1987, 1988; Kuhlemeir *et al.*, 1988). None of these sequences if mutated separately in the context of the full promoter has a profound affect on gene expression (Kuhlemeier *et al.*, 1987, 1988).

A second class of conserved DNA sequence, the G box (C/A-CACGTGGC) is bound specifically by a factor (GBF), distinct from GT-1, isolated from *Arabidopsis* and tomato nuclear extracts (Giuliano *et al.*, 1988). This sequence is also present in light-regulated chalcone synthase promoters (Schulze-Lefert *et al.*, 1989; Staiger *et al.*, 1989) and the alcohol dehydrogenase (*Adh*) promoter of *Arabidopsis* (Chang and Meyerowitz, 1986). Whereas mutation of this sequence has been shown to affect transient expression of chalcone synthase in plant protoplasts (Schulze-Lefert *et al.*, 1989), a requirement for this sequence has not previously been demonstrated for the expression of RBCS genes in transgenic plants.

A third class of conserved DNA sequence, previously referred to as the I box (Giuliano *et al.*, 1988) or sequence 2 (Manzara and Gruissem, 1988) has the consensus 5′-GATAAG. This sequence, which is found in most but not all RBCS genes, is also commonly found 5′ proximal to the 'TATA' box of most chlorophyl a/b binding protein (CAB) genes (Castresana *et al.*, 1987; Gidoni *et al.*, 1989). We have demonstrated that a common factor (GA-1) binds to these sequences in both RBCS and CAB gene promoters (U.Schindler and A.Cashmore, submitted). This factor may be related to factor ASF-2 which binds to a GATA motif in the 35S CaMV (cauliflower mosaic virus) promoter and a petunia CAB promoter (Lam and Chua, 1989). As in the case of the G box sequence, previous RBCS gene expression studies with transgenic plants have failed to demonstrate a requirement for these I box sequences, although mutation of what are likely to be functionally related sequences in CAB genes does affect expression (Gidoni *et al.*, 1989).

Here we describe a deletion analysis of the *Arabidopsis rbcS-1A* promoter which defines a 196 bp region, containing GT, G and I boxes, capable of conferring light-regulated and tissue-specific expression on a heterologous *Arabidopsis Adh* reporter gene. We report experiments which show that

site-specific mutations in either G or I, but not GT boxes, in the context of the full *rbcS-1A* promoter, drastically affect the *rbcS-1A* promoter-dependent expression of *Adh* and GUS reporter genes.

Results

Mutational analysis of the rbcS-1A promoter using Adh as a reporter identifies sequences essential for mediating rbcS-1A−Adh expression in transgenic tobacco leaves

Initial *Agrobacterium*-mediated gene transfer experiments in transgenic tobacco in which sequences from the 5′ flanking region of the *Arabidopsis rbcS-1A* gene were fused to the coding region of the bacterial reporter gene neomycin phosphotransferase II (nptII), indicated that the sequences necessary for photo-regulated and tissue-specific expression of the *rbcS-1A* gene reside within 1.7 kb of the promoter (Timko *et al.*, 1987). The *rbcS-1A* gene is unlinked from three other RBCS genes, which reside in tandem within an 8 kb stretch of chromosome and which constitute a separate subfamily based on their linkage and amino acid similarities (Krebbers *et al.*, 1988).

To investigate the sequences mediating *rbcS-1A* expression in greater detail we have developed the *Arabidopsis Adh* gene as a convenient heterologous reporter to monitor *rbcS-1A* expression in the large numbers of transgenic plants typically generated in a mutational analysis (R.Donald, unpublished results). ADH enzyme activity was not detectable in crude extracts prepared from mature leaves of untransformed tobacco. However, high levels of *Arabidopsis* ADH enzyme activity and light-regulated mRNA levels were detected in leaves of transgenic plants carrying chimeric *rbcS-1A−Adh* transcriptional or promoter fusions. Transgenic *Arabidopsis* ADH activity could be readily distinguished from endogenous tobacco ADH activities present in roots by its distinct mobility in non-denaturing native protein gels (see later).

For the experiments presented here, three cloning vectors (pAdh-T, pAdh-P and pAdh-F) were constructed to facilitate the subcloning of *rbcS-1A* promoter fragments for expression analysis in transgenic tobacco (Figure 1A). These vectors, derived from the plant binary transformation vector pEnd4K (Klee *et al.*, 1985), carry the entire coding region of the *Arabidopsis Adh* (Chang and Meyerowitz, 1986) gene and variable portions of non-coding 5′ *Adh* DNA. Plasmid pAdh-T was used to subclone derivatives of the 1.7 kb *rbcS-1A* promoter fragment Δ2006 bearing deletions or site-directed mutations (see Materials and methods, Figure 2A). The resulting *rbcS-1A−Adh* constructs (e.g. rbcS1A-T, Figure 1B) are transcriptional gene fusions in which the chimeric transcript includes 21 bp of *rbcS-1A* untranslated leader, and 12 bp of the *Adh* untranslated leader mRNA in addition to *Adh* coding DNA. Plasmid pAdh-P was used to subclone *rbcS-1A* promoter fragments lacking *rbcS-1A* transcription initiation signals. In the resulting *rbcS-1A−Adh* promoter fusions, sequences 124 bp 5′ proximal to the transcription start including 'CAAT' and 'TATA' boxes (Chang and Meyerowitz, 1986), are contributed by the *Adh* promoter. Deletions tested in the context of *rbcS-1A−Adh* promoter fusions were subcloned by inserting deleted *rbcS-1A* sequences spanning the −1.7 kb *Eco*RI site and the −60 bp *Hin*dIII site into pAdh-P. The expression of these deletion constructs was compared in transgenic plants with

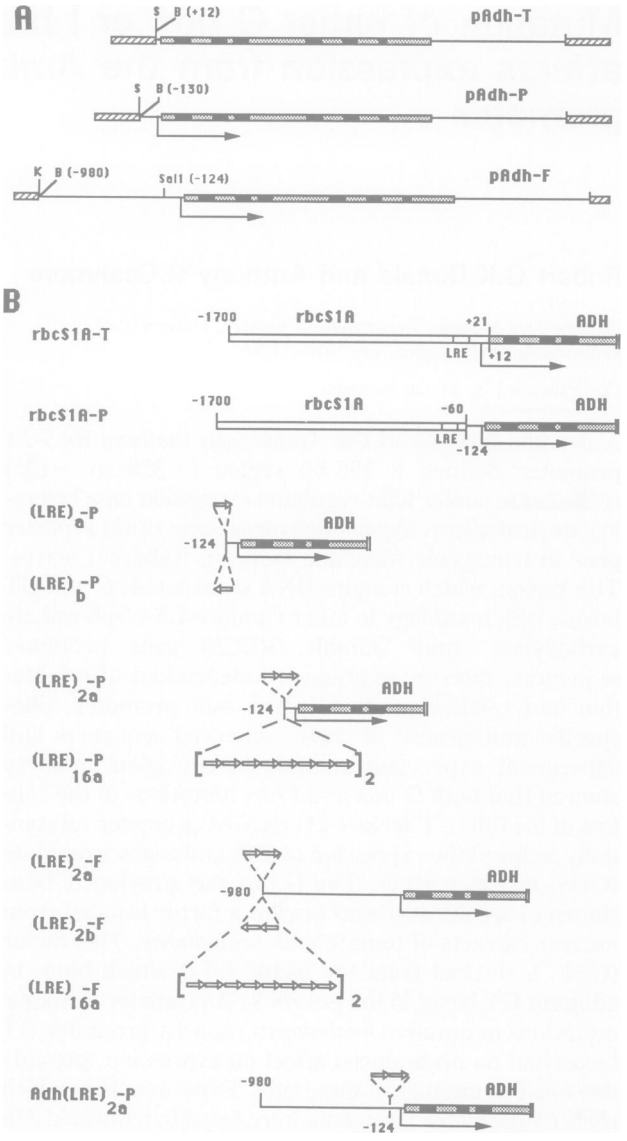

Fig. 1. Structure of *Adh* expression vectors and some *rbcS-1A−Adh* fusions derived from them. (**A**) Plasmid vectors used for the construction of *rbcS-1A−Adh* promoter and transcriptional fusions. Plasmids pAdh-T, pAdh-P and pAdh-F were derived from the 3.7 kb genomic *Sac*I fragment containing the *Arabidopsis Adh* gene (Chang and Meyerowitz, 1986) and the plant transformation vector pEnd4K (Klee *et al.*, 1985). Grey and black boxes represent *Adh* exon and intron sequences, respectively. Hatched boxes denote pEnd4K sequences. S, K and B denote vector pEnd4k *Sal*I, *Kpn*I and *Bam*HI sites. Numbers indicate positions of these sites and of the genomic *Sal*I site relative to the *Adh* transcriptional start. (**B**) Promoter structure of selected *rbcS-1A−Adh* constructs referred to in the text and in Figures 2B, 4 and 5. Open boxes denote *rbcS-1A* sequences, open arrows show the orientation and copy number of the *rbcS-1A* LRE sequence located between −320 and −125 in the intact *rbcS-1A* promoter. Numbers indicate endpoints of genomic *rbcS-1A* and *Adh* sequences in these constructs.

the control construct rbcS1A-P (Figure 1B) which contains undeleted *rbcS-1A* sequences upstream of −60 bp (Figure 2A) fused to the truncated *Adh* promoter of plasmid pAdh-P. Plasmid pAdh-F, containing 980 bp of the *Adh* promoter, was used to test for the ability of *rbcS-1A* promoter fragments to confer leaf-specific expression on the full *Adh* gene in transgenic plants. While transgenic tobacco plants

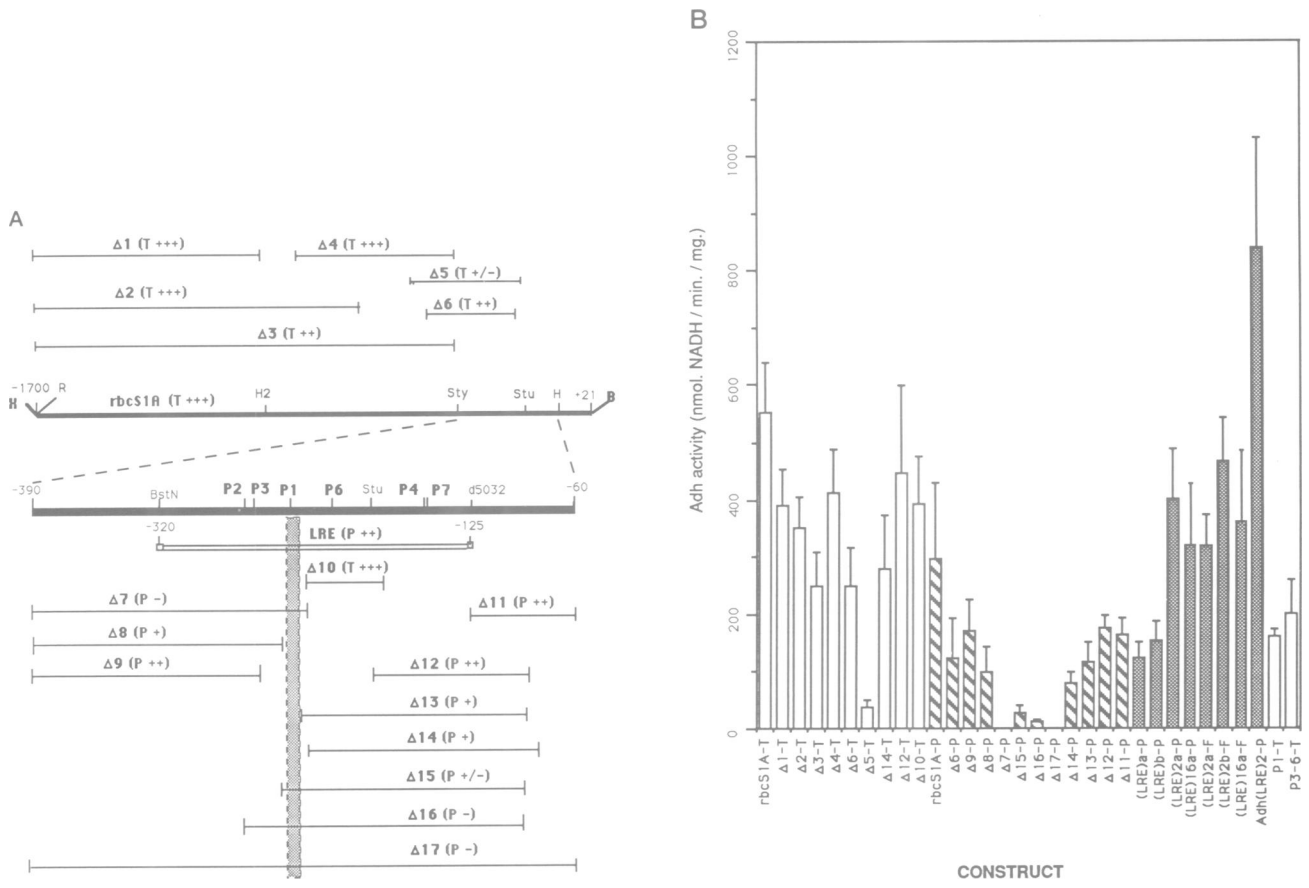

Fig. 2. Mutational analysis of the *Arabidopsis rbcS-1A* fusions: ADH reporter. (**A**) Deletions (denoted Δ) within the 1.7 kb *rbcS-1A* promoter. The relative ADH activity of deletion constructs expressed in leaves of trangenic T_0 plants, is presented qualitatively as +++ high, ++ moderate, + low, +/− very low and − no expression. The context of the deletion in the construct with respect to either a transcriptional (T) or promoter (P) *rbcS-1A−Adh* fusion background is indicated. X and B show cleavage sites for *Xho*I and *Bam*HI located in the vector sequences of plasmid pRΔ2006 (Materials and methods). Genomic sites shown are *Eco*RI (R), *Hin*dII (H2), *Sty*I (Sty), *Stu*I (Stu), *Hin*dIII (H) and *Bst*NI (BstN). Also shown are the location of sequences within the LRE element (see text) targeted by site-directed mutagenesis (I, G and GT boxes; = P1−P7, see also Figure 3). 'd5032' marks the 3′ endpoint of *rbcS-1A* sequences contained within plasmid pRΔ5032 from which LRE sequences were subcloned (Materials and methods). The stippled box emphasizes a region defined by deletions Δ5, Δ6, Δ7, Δ8, Δ10, Δ13, Δ14 and Δ15, containing the G box, which is apparently indispensable for *rbcS-1A* expression (see text; Figures 2B and 3). (**B**) Quantitative mean ADH activity levels in leaves of T_0 transgenic plants expressing *rbcS-1A−Adh* fusions bearing *rbcS-1A* and mutated *rbcS-1A* promoter sequences. Open bars correspond to deletions, and site-specific double I box (P3−P6) and G box (P1) mutations analyzed in the context of transcriptional *rbcS-1A−Adh* fusions, hatched bars to deletions measured in the context of *rbcS-1A−Adh* promoter fusions. Grey bars show the expression levels of constructs in which the LRE sequence is fused to the −124 promoter-truncated *Adh* gene or positioned at −124 or −980 of the full *Adh* gene (see text, Figure 1B). Error bars indicate the standard error of the mean of each T_0 population.

carrying pAdh-F expressed transgenic *Arabidopsis* ADH activity in a strictly root- and seed-specific manner similar to the ADH activities expressed in wild-type *Arabidopsis* or tobacco, transgenic tobacco transformed with pAdh-P showed no detectable transgenic *Arabidopsis* ADH activity in either roots or seed (R.Donald, unpublished results).

Figure 2 shows the results of the *rbcS-1A* mutational analysis in transgenic T_0 tobacco plants using the *Adh* reporter gene. Effects of deletions on ADH reporter activity in either the context of the *rbcS-1A−Adh* transcriptional fusions (labeled 'T', open bars, Figure 2B), or promoter fusions (labeled 'P', hatched bars, Figure 2B) are indicated. For each construct analyzed in which significant ADH activity was detected, data from between six and 15 transgenic plants expressing ADH at detectable levels are presented. For constructions showing little or no detectable ADH reporter activity (Δ7-P, Δ-16P and Δ17-P), at least 10 plants were analyzed.

Figure 3 shows the *rbcS-1A* DNA sequences (Timko *et al.*, 1987) within the region −320 to −125 bp relative to the transcription start, emphasizing the G box, I boxes and GT boxes conserved in other RBCS genes (Giuliano *et al.*, 1988; Manzara and Gruissem, 1988) and indicating the precise position of all the deletion endpoints. This region is termed the light regulatory element (LRE) since, as will be subsequently described, it could confer the ability to mediate light-regulated expression on both the truncated *Adh* promoter (pAdh-P) and the full *Adh* promoter (pAdh-F) (Figures 1, 2 and 4). We have shown that factors GBF and GA-1, isolated from tobacco or *Arabidopsis* nuclear extracts in the case of GBF or from tobacco, pea and tomato nuclear extracts in the case of GA-1, bind specifically to LRE G box and I boxes, respectively (Giuliano *et al.*, 1988; U.Schindler and A.Cashmore, submitted). Although we have not identified specific GT-1 binding sites within the LRE fragment, the fragment does compete strongly for binding with defined GT-1 oligonucleotides of an activity isolated from tomato, tobacco or pea nuclear extracts which

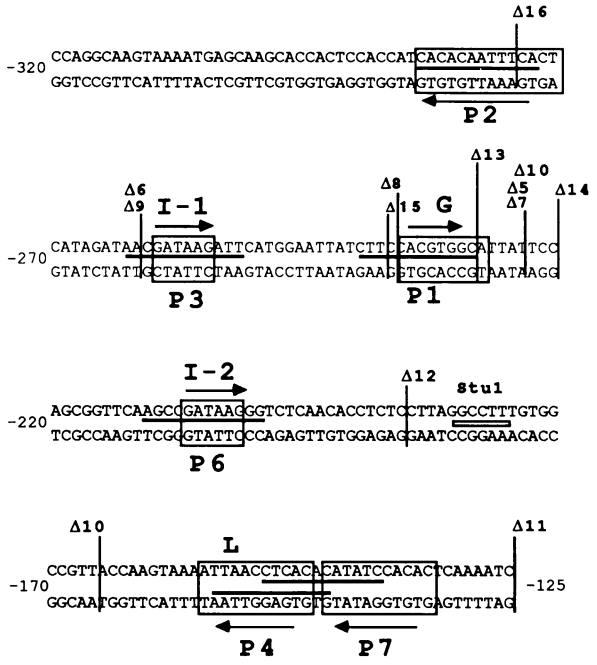

Fig. 3. DNA sequence of the *rbcS-1A* LRE (Timko *et al.*, 1987) showing endpoints of *Bal*31 deletions (see text, Figure 2A) and regions of homology with other RBCS genes that were targeted by site-specific mutagenesis (P1–P7). Boxes show sequences with homology to the I box, G box or GT boxes. Arrows indicate the relative orientation of the conserved sequences. Coding strand sequences underlined and labeled P1, P2, P3, P4, P6 and P7 were substituted with *Eco*RV/*Bgl*II restriction enzyme recognition sequences as follows: P1, P2, P3, P7–5′-GATATCAGATCT; P4, P6–5′-AGATCTGATATC. Mutations P1, P3 and P6 alter, respectively, G box and I box sequences. We note that although the P3, P6 substitutions alter the conserved 'GATAAG' I box motif, a 'GATA' motif is present in the introduced *Eco*RV site. This introduced GATA sequence clearly does not serve as a functional I box (see Results). Sequences substituted by P2, P4 and P7 show sequence similarity to GT-1 boxes III*, II and II*, respectively, of the pea *rbcS-1A* gene (Green *et al.*, 1987). The sequence targeted by substitution P4 overlaps a sequence showing homology to the 'L' box (consensus 5′-AAATTAACCAA), which is conserved in RBCS upstream sequences of both tomato and tobacco (Giuliano *et al.*, 1988).

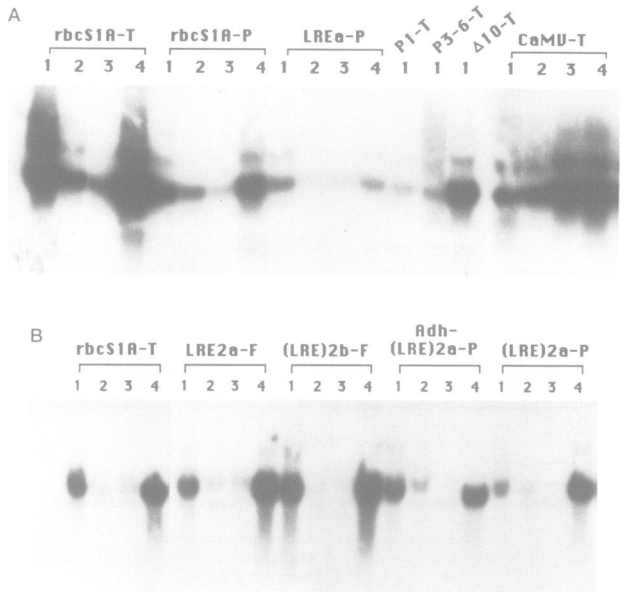

Fig. 4. Northern blot analysis of *rbcS-1A–Adh* fusion expression in transgenic tobacco leaves. Amounts of 20 μg of total RNA are loaded in each lane. (**A**) *Adh* transcript levels in leaves of transgenic plants expressing the indicated constructs. Leaves were sampled from mature greenhouse-grown plants as follows: (1) no treatment, (2) after 2 days of continuous darkness, (3) after 4 days of continuous darkness, (4) after re-exposure of 4 day dark-treated plants to 16 h of continuous light. For light/dark expression analysis of constructs rbcS1A-T, rbcS1A-P, $(LRE)_a$-P and CaMV-T, RNA was extracted from leaves of representative T_1 plants. These plants were progeny of T_0 transformants showing median ADH activity levels. In the case of constructs P1-T, P3-6-T and Δ10-T, only 'light-grown' leaves (1) were analyzed. For each of these constructs, RNA was extracted from leaves of five independent T_0 transformants and pooled. (**B**) *Adh* transcript levels in leaves of light/dark-treated representative T_1 plants expressing dimerized LRE-Adh constructs: rbcS1A-T control; $(LRE)_{2a}$-F; $(LRE)_{2b}$-F; Adh$(LRE)_{2a}$-P; $(LRE)_{2a}$-P.

resembles factor GT-1 in specificity (U.Schindler and A.Cashmore, submitted). Putative GT-boxes within the LRE fragment are designated in Figure 3 on the basis of their homology to sites recognized by factor GT-1 in the pea *rbcS-3A* promoter (Green *et al.*, 1987, 1989).

rbcS-1A promoter deletions 5′ to the *Sty*I site (−390 bp from transcriptional start) had no profound affect on ADH reporter activity. Although deletion of the entire −390 bp proximal region (Figure 2, Δ3-T) resulted in a 50% reduction in ADH activity, the effect could not be attributed to any specific 5′ DNA segment since deletions spanning this 1.3 kb upstream region failed to show a substantial decline in ADH activity (Δ1-T, Δ2-T and Δ4-T). In contrast, deletion Δ5-T which extends further 3′ into a region containing conserved RBCS sequences, drastically reduced activity to near background levels. Deletion Δ6-T, which also penetrates this region, though not as deeply, showed levels of activity comparable to the −390 bp Δ3-T deletion. Since the 5′ endpoints of deletions Δ5-T and Δ6-T reside in a region of the *rbcS-1A* promoter apparently dispensable for leaf expression (encompassed by deletion Δ4-T), it seemed likely that the substantial differences in ADH activity level observed for constructs Δ5-T and Δ6-T could be attributed to differences in the 3′ endpoints of these two deletions. Figure 3 shows that while deletion Δ6-T removes one of the putative GT boxes, deletion Δ5-T, which extends 38 bp further, in addition removes one of the two I boxes and the G box.

To more fully delineate the DNA sequences required for *rbcS-1A* expression and to confirm the results described above, further deletions were constructed. The strategy used was to make deletions extending into the G box, I box and adjacent GT boxes from both 5′ and 3′ directions. These deletions were tested in the context of *rbcS-1A–Adh* promoter fusion construct rbcS1A-P (Figure 1B), in which *rbcS-1A* sequences upstream of −60 were fused to the *Adh* promoter truncated at −124 by subcloning into pAdh-P (see earlier). It was expected that the use of *rbcS-1A* promoter fusions would facilitate the interpretation of the affects of the 3′ *rbcS-1A* promoter deletions by providing heterologous ubiquitous *Adh* transcription elements potentially deleted from the *rbcS-1A* promoter. Although the *rbcS-1A* promoter lacks a perfect CAAT box (Krebbers *et al.*, 1988), it does have sequences at −120 bp which match the RBCS general consensus (5′-AATCCAA[C]-3′ Manzara and Gruissem, 1988) and which are removed by deletions Δ11 through Δ17 (Figure 2). However, the effects of deletions Δ12 and Δ14 measured in the context of the transcriptional *rbcS-1A–Adh*

fusions (Δ12-T and Δ14-T, Figure 2) now show that this region is dispensable in the context of the full *rbcS-1A* promoter.

Deletions Δ9-P, Δ8-P and Δ7-P extended 3′ from the −390 *Sty*I site and deleted successively sequences containing a putative GT box, an I box and the G box (Figures 2 and 3). Deletions Δ9-P and Δ7-P share, respectively, the same 3′ endpoints as the deletions Δ6-T and Δ5-T mentioned previously. When compared with the control construct rbcS1A-P, deletion Δ9-P showed no significant reduction in ADH activity. In comparison, Δ8-P showed a modest but significant drop in activity and Δ7-P a total loss of activity. Similar results were obtained with the set of deletions Δ12-P, Δ13-P, Δ14-P, Δ15-P and Δ16-P. Deletion Δ12-P, which deleted sequences that included two putative GT boxes, showed no significant loss of ADH activity when compared with constructs rbcS1A-P or Δ11-P. Deletions Δ14-P and Δ13-P, which also removed an I box, showed a 25−50% loss of ADH activity. Deletions Δ15-P and Δ16-P, which extended beyond the G box, reduced activity to close to background levels. The removal of the entire −390 to −60 region with deletion Δ17-P completely eliminated expression.

Deletion Δ10, tested in the context of a transcriptional *rbcS-1A*−*Adh* fusion (Δ10-T) and hence comparable with rbcS1A-T, Δ6-T, Δ5-T and Δ14-T, had no effect on ADH activity (Figures 2 and 3). When considered together with the effects of other deletions in this −390 to −60 region, measured in the context of *rbcS-1A*−*Adh* promoter fusions, these results suggest that unique sequences indispensable for *rbcS-1A*−*Adh* expression reside within a narrow region flanked by, but not spanned by, deletions Δ6-T and Δ10-T. This 38 bp region (Figure 3) contains an I box, the G box, but no obvious GT boxes. However, it is possible that sequences such as the GT box which are present in multiple copies and not completely removed by any one of the described deletions (Figure 2A), might also be essential for *rbcS-1A* promoter expression.

An rbcS-1A element (LRE) confers light-regulated and leaf-specific expression independent of orientation and position in the Adh gene promoter

To determine whether an *rbcS-1A* sequence containing I, G and GT boxes was capable of mediating light-regulated and leaf-specific *Adh* reporter gene expression, the fragment encompassing the region −320 to −125 was fused to the truncated *Adh* promoter in either orientation by subcloning into vector pAdh-P (Figure 1A). Transgenic plants carrying the resulting constructs $(LRE)_a$-P and $(LRE)_b$-P expressed levels of leaf ADH enzyme comparable with constructs rbcS1A-P (Figures 1B and 2B). Furthermore, light-regulated expression of *Arabidopsis Adh* mRNA in the leaves was detected for plants carrying either $(LRE)_a$-P or $(LRE)_b$-P. Figure 4A shows an mRNA hybridization analysis of *Arabidopsis Adh* transcripts expressed in light/dark-adapted leaves of individual representative transgenic plants carrying $(LRE)_a$-P, rbcS1A-P, rbcS1A-T and CaMV-T constructs. The CaMV-T control construct consisted of a 450 bp CaMV promoter fragment subcloned into vector pAdh-T to create a transcriptional fusion in which the chimeric *Adh* transcript included 7 bp of CaMV untranslated leader. In contrast with the *rbcS-1A*−*Adh* constructs shown, this construct expressed *Adh* mRNA in the dark. We can conclude that *rbcS-1A* sequences sufficient for conferring light-regulated and leaf-specific expression on the truncated *Adh* gene (pAdh-P) reside within the −320 to −125 bp region of the *rbcS-1A* promoter which we designate the *rbcS-1A* LRE (light-regulatory element).

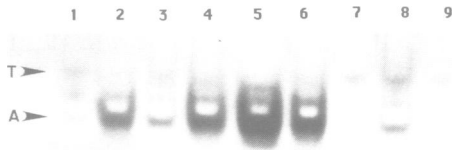

Fig. 5. Native gel activity staining of ADH extracted from roots of individual transgenic T_1 plants expressing (LRE)−Adh constructs. Crude extract protein (30−40 μg) was loaded in each lane. **Lane 1**, Adh-F control; **lanes 2** and **3**, $(LRE)_{2a}$-F; **lanes 4** and **5**, $(LRE)_{2b}$-F; **lane 6**, Adh$(LRE)_{2a}$-P; **lane 7**, SR-1 untransformed control; **lanes 8** and **9**, $(LRE)_{2a}$-P. Relative mobilities of tobacco (T) and *Arabidopsis* (A) ADH isozymes are indicated. Subunit mixing of *Arabidopsis* and endogenous tobacco enzymes accounts for activities of intermediate mobilities.

To determine if the LRE-dependent expression of the *Adh* reporter could be enhanced by higher LRE copy number and to determine the effect on *Adh* expression of relative LRE position within the *Adh* promoter, the LRE was polymerized 2-fold and 16-fold in direct repeats and subcloned into pAdh-P and pAdh-F to form constructs $(LRE)_{2a}$-P, $(LRE)_{16a}$-P, $(LRE)_{2a}$-F, $(LRE)_{2b}$-F and $(LRE)_{16a}$-F (Figure 1B). To obtain a construct in which LRE sequences were inserted within the full *Adh* promoter at −124, the upstream portion of *Adh* promoter (−125 to −980) was added to construct $(LRE)_{2a}$-P to form construct Adh$(LRE)_{2a}$-P (Figure 1B). *Adh* reporter expression analysis of T_0 transgenic plants carrying these constructs is shown in Figure 2B (grey bars). Constructs $(LRE)_{2a}$-P, $(LRE)_{16a}$-P, $(LRE)_{2a}$-F, $(LRE)_{2b}$-F and $(LRE)_{16a}$-F all showed a 2- to 3-fold higher level of ADH activity than single copy constructs $(LRE)_a$-P and $(LRE)_b$-P. Construct Adh$(LRE)_{2a}$-P showed the highest leaf ADH activity of all the *rbcS-1A*−*Adh* constructs analyzed, including transcriptional fusion rbcS1A-T. The light-regulated expression of the transgenic constructs containing the dimerized LRE [$(LRE)_{2a}$-P, $(LRE)_{2a}$-F, $(LRE)_{2b}$-F and Adh$(LRE)_{2a}$-P] was examined in leaves of representative transgenic plants (Figure 4B). All of these constructs showed strong light-regulated *Arabidopsis Adh* mRNA expression comparable to construct rbcS1A-T.

The rbcS-1A LRE enhances root-specific Adh expression

Examination of *Arabidopsis* ADH activity in roots of transgenic plants expressing LRE−*Adh* constructs showed that the LRE enhanced root-specific expression of the truncated *Adh* and full *Adh* gene constructs. Root ADH activity in representative T_1 plants expressing constructs Adh-F (full *Adh* gene), $(LRE)_{2a}$-F, $(LRE)_{2b}$-F, Adh$(LRE)_{2a}$-P and $(LRE)_{2a}$-P is shown in Figure 5. Constructs $(LRE)_{2a}$-F, $(LRE)_{2b}$-F and Adh$(LRE)_{2a}$-P, which contain the LRE fused to or inserted within the 980 bp *Adh* promoter, expressed *Arabidopsis* ADH at substantially higher levels in transgenic roots than the *Adh* gene control (Adh-F) (Figure 5, lanes 1−6). Lower levels of *Arabidopsis*

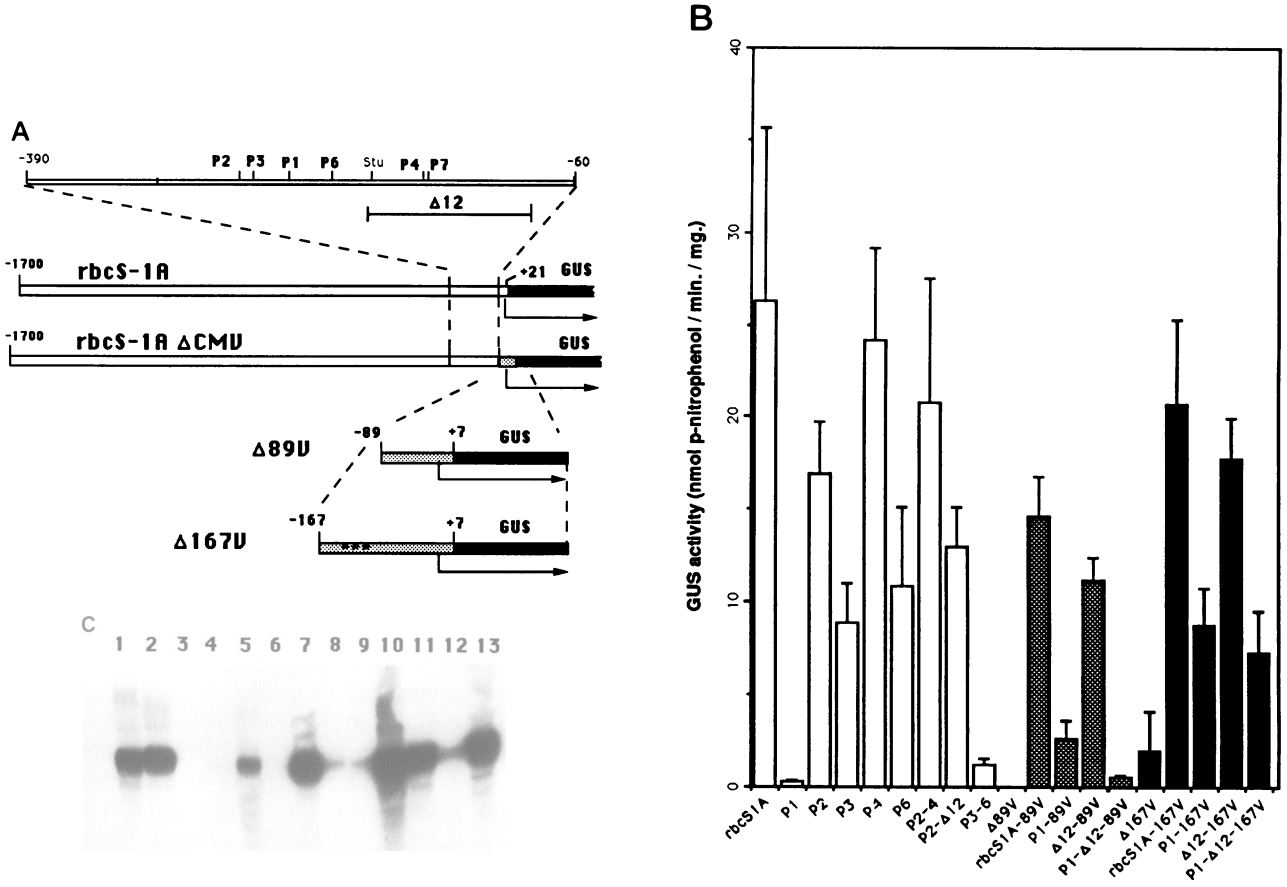

Fig. 6. Mutational analysis of the *Arabidopsis rbcS-1A* promoter in transgenic T_0 plants: GUS reporter. (**A**) Structure of rbcS1A−Gus and rbcS1A-ΔCaMV-GUS constructs derived from GUS expression vector pBI101 (see Materials and methods). In the construction of rbcS1A-ΔCaMV-GUS, *rbcS-1A* promoter sequences downstream of −60 bp (relative transcription start) are replaced with CaMV promoter sequences between −89 and +7 (89V) and −167 and +7 (167V). In construct Δ167V, sequences homologous to the SV40 core enhancer are indicated by an asterisk (Ow *et al.*, 1987). (**B**) Quantitative expression analysis of transgenic T_0 plants carrying rbcS1A−Gus, rbc1A-Δ89V-GUS and rbcS1A-Δ167V-GUS constructs represented respectively by open, grey and black bars and carrying the mutations indicated (see text). GUS activity was determined by the colorimetric assay (Jefferson *et al.*, 1987). (**C**) Northern blot analysis of GUS transcript levels in leaves of T_0 plants expressing some of the GUS constructs shown in (B). Lanes 1 and 2 contain 20 μg of total RNA extracted from two T_1 plants obtained as offspring of independent T_0 transformants expressing control construct rbcS1A−GUS. In each of lanes 3−13 is shown 20 μg RNA pooled from five T plants carrying each of the following constructs: lane 3, P1; lane 4, P3−6; lane 5, P2-Δ12; lane 6, φ-89V; lane 7, Δ12-89V; lane 8, P1-Δ12-89V; lane 9, φ-167V; lane 10, Δ12-167V; lane 11, P1-Δ12-167V; lane 12, P1-89V; lane 13, P1-167V.

ADH root activity were observed for construct $(LRE)_{2a}$-P (Figure 5, lanes 8−9) in which the light-regulatory element is fused to the promoter-truncated *Adh* gene. The promoter-truncated *Adh* gene construct Adh-P yields no activity in either seed or in root tissue (R.Donald, unpublished results).

Levels of root ADH activity comparable with $(LRE)_{2a}$-P were also detected in transgenic plants expressing rbcS1A-T and rbcS1A-P constructs (not shown). Because the *rbcS-1A* promoter is incapable of directing root expression on bacterial reporter genes encoding chloramphenicol acetyl transferase (CAT) and β-glucuronidase (GUS) (R.Donald, unpublished results), the results suggest a general cooperativity between *Arabidopsis Adh* reporter sequences and *rbcS-1A* promoter sequences.

Site-specific mutation of either the G box or the I boxes within the rbcS-1A promoter reduces expression of both the Adh reporter and the bacterial GUS reporter

To test more specifically the affect of I and G box mutations on *rbcS-1A−Adh* expression in leaf tissue, we used site-directed mutagenesis (Kunkel, 1985) to introduce 12 bp substitutions into both I boxes (mutations P3 and P6, Figure 3) and the G box (mutation P1, Figure 3) in the context of the full 1.7 kb *rbcS-1A* promoter. Compared with the rbcS1A-T control, the G box substitution (construct P1-T) reduced ADH enzyme activity by 60−70%, the double I box mutation (construct P3-6-T) by 50−60% (Figure 2B). At the RNA level the affects of G box and double I box mutations on *Adh* expression were much more apparent as seen in the Northern analysis shown in Figure 4. RNA pooled from plants expressing constructs P1-T (G box mutant) and P3-6 (double I box mutations) showed, respectively, 40- and 10-fold lower levels of *Adh* mRNA compared with control construct rbcS1A-T, and 20- and 5-fold lower levels than a second control construct, Δ10-T, which removes a single I box and additional sequences immediately distal to the G box (Figure 2A). When considered together with the results of the deletion analysis, these data strongly implicate the G box and adjacent I boxes as important elements mediating *rbcS-1A−Adh* expression.

To confirm and extend the results of the *rbcS-1A* promoter site-directed mutagenesis obtained with the *Adh* reporter, site-directed mutations in the *rbcS-1A* promoter were also

analyzed with the GUS reporter gene. Site-directed mutations in the same 1.7 kb *rbcS-1A* promoter fragment (Figure 2A) as analyzed with the *Adh* reporter were introduced into plant binary GUS expression vector pBI101 (Jefferson *et al.*, 1987) and their affects tested in leaves of transgenic T_0 plants (Figure 6). The affects of G box and I box mutations were even more dramatic than those observed with the *Adh* reporter. The G box mutation (construct P1) reduced GUS activity and GUS mRNA to near-undetectable levels (by >99%) as compared with the intact parental promoter (construct rbcS1A). Mutations in either I box reduced GUS activity ~50% (constructs P3 and P6). Double 1 box mutations resulted in >90% reduction in levels of GUS enzyme and GUS mRNA (construct P3−6). Consistent with the results of *rbcS-1A* promoter deletions analyzed with the *Adh* reporter, deletion or site-specific mutation of GT boxes (Δ12, P2, P4 and P7), either individually or in combination, did not have a profound effect on GUS activity (constructs P2, P4, P2-4 and P2-Δ12), although mutation P2-Δ12 which eliminates all three of the GT boxes immediately adjacent to the *rbcS-1A* G box did show a small but statistically significant decrease in GUS enzyme and mRNA levels (20−50% for both GUS enzyme and mRNA).

35S CaMV promoter elements can restore expression to a G-box-mutated rbcS-1A promoter

The plant binary transformation vector pBI101 (Jefferson *et al.*, 1987) was also used to construct two rbcS1A-ΔCaMV-GUS fusion constructs in which sequences mediating transcription initiation are provided by truncated 35S CaMV promoter fragments (Figure 6A). The purpose of the rbcS1A-ΔCaMV-GUS constructions was to test the affect of G box mutations in the context of other known *cis*-acting elements.

The first 35S CaMV promoter fragment Δ89V spanning −89 to +7 contains a CCAAT-like box at an expected position (−85) and two other CCAAT-like boxes closer to the start of transcription. The −85 CCAAT box is indispensable for the expression of a minimal −148 CaMV promoter in carrot protoplasts (Ow *et al.*, 1987). Although the −89 promoter core is incapable of independently directing reporter gene expression, sequences residing in it (−90 to −46) are important for mediating expression directed by upstream CaMV sequences in transgenic tobacco (Fang *et al.*, 1989). A protein factor (ASF-1) has been identified which binds to two TGAC motifs in the region −83 to −63 (Katagiri *et al.*, 1989; Lam *et al.*, 1989; Prat *et al.*, 1989).

The second CaMV promoter fragment Δ167V which spans −167 to +7 contains three sequences homologous to the SV40 'core' enhancer, GTGG(A/T)(A/T)(A/Y)G (Weiher *et al.*, 1983), which are located between −148 and −89. The 5′ deletions of the 35SCaMV promoter fused to the firefly luciferase reporter and analyzed in carrot protoplasts showed that, relative to a −148 minimal promoter, a −108 deletion removing two of these elements resulted in a dramatic (5-fold) drop in reporter expression (Ow *et al.*, 1987). A second 35S CaMV promoter binding protein (ASF-2) has also been identified which binds to a −100 GATA motif which is distinct from these SV40 core element related sequences (Lam and Chua, 1989).

RbcS-1A G box mutations were tested in the context of these 35S CaMV truncated promoters by fusing *rbcS-1A* sequences upstream of −60 to both Δ89V-GUS and Δ167V-GUS plasmids (Figure 6A). Analysis of leaves of transgenic plants carrying the Δ89V-GUS and Δ167V-GUS control constructs showed that only for the Δ167V-GUS construct could significant GUS activity be detected with either the colorimetric assay or with Northern RNA analysis (see constructs Δ89V and Δ167V, Figure 6B and C). With the more sensitive fluorogenic assay or with longer autoradiogram exposures a low but above-background GUS activity could be detected for the Δ89V-GUS construct (data not shown). Fusion of *rbcS-1A* sequences (−60 to −1700) to either Δ89V-GUS or Δ167V-GUS resulted in GUS enzyme activity and GUS RNA levels comparable to that of the *rbcS-1A*-GUS construct (compare constructs rbcS1A-89V and rbcS1A-167V with construct rbcS1A, Figure 6B and C).

A comparison of GUS expression data from transgenic plants carrying *rbcS-1A* G box mutations in the context of the intact rbcS1A−GUS and rbcS1A-ΔCaMV-GUS fusions showed that elements within the truncated CaMV promoter can act co-operatively with *rbcS-1A* promoter sequences to restore reporter gene activity in the absence of a functional G box. The effect of the G box mutation on the expression of the rbcS1A-ΔCaMV-GUS constructs was less severe than its affect on the expression of the rbcS1A−GUS construct. The GUS activity observed in G box mutant construct P1-Δ89V-GUS was 15−20% the level observed for the non-mutated parental rbcS1A-Δ89V-GUS construct (compare construct P1-89V with construct rbcS1A-89V, Figure 6B). In the rbcS1A-Δ167V-GUS construct background, the GUS activity of the G box mutant derivative was ~50% that of the non-mutated control (compare construct P1-167V with construct rbcS1A-167V).

Similar results were obtained when the G box mutation was tested in the context of *rbcS-1A* deletion Δ12 fused to either the −89 or −167 truncated CaMV promoter (Figures 2A and 6A−C). The purpose of these constructs was to examine the effect of the G box mutation introduced into the *rbcS-1A* promoter deleted of closely linked GT sites (Figure 3). Although the Δ12 deletion reduced slightly the level of rbcS1A-ΔCaMV-GUS expression (compare construct Δ12-89V with construct rbcS1A-89V and construct Δ12-167V with construct rbcS1A-167V), the G box mutation exerted the same relative effect on Δ12-rbcS1A-ΔCaMV-GUS activity as on the rbcS1A-ΔCaMV-GUS constructs described above (compare construct P1-Δ12-89V with construct Δ12-89V and construct P1-Δ12-167V with construct Δ12-167V). The results indicate that the effect of the G box mutation on GUS reporter activity is independent of the adjacent GT sites removed by deletion Δ12. Furthermore, these results confirm that the loss of GUS activity observed for *rbcS-1A* constructs modified by the 12 bp G box substitution (mutation P1) is reproducible when the substitution is introduced independently into a distinct *rbcS-1A* template.

Discussion

Conserved G box and I box sequences are important for expression from the rbcS-1A promoter

To investigate the functional significance of evolutionarily conserved sequence elements within the promoters of RBCS genes (Giuliano *et al.*, 1988; Manzara *et al.*, 1988) we performed a mutational analysis of the *Arabidopsis* 1.7 kb

rbcS-1A promoter in transgenic tobacco plants using *Adh* and GUS reporter genes. Previous studies from the laboratory of Chua have established a role for the factor GT-1 in the expression of the pea *rbcS-3A* promoter (Kuhlemeier *et al.*, 1987, 1988; Green *et al.*, 1988). These studies showed a functional redundancy among the six GT-1 binding sites (Green *et al.*, 1988). A functional requirement for GT-1 binding sites on *rbcS-1A* expression could be demonstrated after reducing the *rbcS-1A* promoter to 170 bp of flanking DNA (Kuhlemeier *et al.*, 1987, 1988). Our results, which fail to show a functional requirement for GT-boxes in the context of the full *Arabidopsis rbcS-1A* 1.7 kb promoter are not inconsistent with these results.

In our studies, we have clearly established a requirement for conserved G box and I box sequences for expression from the *Arabidopsis rbcS-1A* promoter. Since the *rbcS-1A* G box is bound by nuclear factor GBF (Giuliano *et al.*, 1988) and the I boxes are bound by factor GA-1 (U.Schindler and A.Cashmore, submitted), we have implicated the functional significance of second and third DNA binding proteins regulating RBCS gene expression in transgenic plants. The precise role that any of these proteins plays in photo-regulated gene expression remains to be determined.

G-box-like promoter elements and GBF-like factors are not restricted to plant RBCS genes

The role that GBF plays in mediating promoter expression is apparently not limited to plant RBCS genes. Similar studies in this laboratory with the tobacco *Cab-E* promoter have shown that mutation of a G box sequence significantly reduces expression of a GUS reporter gene (P.Bringmann and A.Cashmore in preparation). In *in vivo* dimethyl sulfate footprinting experiments with cell cultures, G-box-like sequences at −200 in the *Arabidopsis Adh* promoter (Ferl and Laughner, 1989) and at −169 in the *Petroselinum crispum* (parsley) chalcone synthase (CHS) promoter (Schulze-Lefert *et al.*, 1989) were bound by factors. In the case of the CHS promoter, site-directed mutagenesis of the G-box-related seqence (CCACGTGGC), resulted in complete loss of expression in transient expression assays using the GUS reporter. More recent experiments in this laboratory have established the existence of a yeast 'GBF-like' factor which can mediate the G-box-dependent expression of a bacterial β-galactosidase reporter gene fused to a truncated yeast *iso*-1-cytochrome *c* (CYC1) promoter (Donald *et al.*, 1990).

Sequences in the Adh gene and the CaMV 35S promoter can partially complement G box mutations

Although *rbcS-1A* G box and double I box mutations had profound affects on the expression of both reporter genes in transgenic tobacco leaves, the affects were much more drastic for the GUS reporter than for the *Adh* reporter gene. Although additional G and I boxes are not obviously present within the *Adh* reporter sequences, observed differences might be explained by the presence of sequences within the *Adh* reporter capable of providing a low level compensatory enhancing activity. *Adh* sequences showing enhancing activity have been identified for the maize *Adh1-S* gene in electroporation-mediated gene transfer experiments with maize cells. In these experiments, it was demonstrated that intron sequences of the maize *Adh1-S* gene were not only essential for maize *Adh1-S* gene expression, but could stimulate chimeric gene expression when located near the 5′ end of the mRNA (Callis *et al.*, 1987). A possible advantage that the *Adh* reporter affords compared with the bacterial GUS reporter is that the residual *Adh* reporter activity driven by mutated *rbcS-1A* G box and I box promoter sequences may constitute a means of determining in future studies the extent of involvement of these sequences in the light-regulated and tissue-specific expression directed by the *rbcS-1A* gene promoter. We have also shown here that the severity of the effect of the G box mutation is influenced by context not only with respect to sequences 3′ to the promoter, but also by the presence of other *cis*-acting promoter elements. Compared with the inactive G-box-mutated rbcS1A−GUS construct, substitution of *rbcS-1A* sequences distal to −60 by a −89 CaMV core promoter resulted in a partial (15−20%) restoration of GUS activity relative to a non-mutated control; substitution by a −167 CaMV promoter fragment, which contains additional CaMV sequences encompassing known enhancer elements, restored substantial GUS activity (50−60%).

Arabidopsis Adh and rbcS-1A promoter elements act synergistically to enhance both leaf and root expression

Expression analysis of *rbcS-1A−Adh* mixed promoter constructions showed that *rbcS-1A* sequences within the 196 bp LRE, which include a G box, and I and GT boxes, were capable of mediating light-regulated *Adh* mRNA expression in leaves of transgenic plants independent of their orientation and relative position within the *Adh* promoter. Surprisingly, an apparent co-operativity between *Adh* and *rbcS-1A* LRE sequences in enhancing both leaf- and root-specific ADH reporter activity was observed. The addition of an upstream *Adh* promoter fragment (−980 to −125) to construct $(LRE)_{2a}$-P to create construct Adh$(LRE)_{2a}$-P, enhanced significantly the leaf expression conferred by the dimerized LRE elements on the truncated *Adh* promoter (Figures 1B and 2B). The LRE-dependent expression of construct $(LRE)_{2a}$-P could not be enhanced by higher LRE copy number. In root tissue, the LRE element was capable of significantly enhancing the root expression of either the full *Adh* gene construct or the truncated *Adh* gene construct. It is possible that general enhancing elements, either in the LRE or in the *Adh* promoter, are responsible for these effects.

Materials and methods

Nucleic acid manipulations

DNA manipulations were performed essentially as described (Maniatis *et al.*, 1982). Restriction enzymes were purchased from Promega; DNA modifying enzymes from Boehringer or Pharmacia.

Construction of Adh expression vectors

A 3.7 kb genomic *Arabidopsis Sac*I fragment which contains the *Adh* gene including 1 kb of promoter was excised from recombinant lambda phage At3101 (Chang and Meyerowitz, 1986) and subcloned into the *Hin*dII site of pUC19 in either orientation to give plasmids pAdh11 and pAdh12. In plasmid pAdh11, the pUC19 polylinker *Bam*HI site was oriented 5′ to the *Adh* promoter. Plasmid pAdh-F (Figure 1) was constructed by excision of the 3.7 kb *Adh* insert of pAdh11 with *Bam*HI and *Sph*I and ligation to *Agrobacterium* binary vector pEnd4k (Klee *et al.*, 1985) linearized by digestion of its polylinker restriction sites with *Bam*HI and *Sal*I. The respective *Sph*I and *Sal*I ends of the *Adh* fragment and pEnd4K vector were made ligatable by Klenow polymerase/exonuclease. Plasmid pAdh-P was

constructed after first modifying the −124 *Adh SalI* site in plasmid pAdh11 with the insertion of a *Bam*HI linker, followed by excision of the 2.8 kb promoter-truncated *Adh* gene fragment with *Bam*HI and *Sph*I, and subsequent ligation with vector pEnd4K cut with *Bam*HI and *Kpn*I. In this case, the vector polylinker *Kpn*I and *Adh* fragment *Sph*I sites were prepared for blunt-end ligation with Klenow polymerase/exonuclease. Similarly, plasmid pAdh-T was prepared by subcloning a 2.7 kb *Bam*HI−*Sph*I fragment containing sequences distal to +12 bp of the *Adh* gene from plasmid pAΔ1503 into pEnd4K polylinker *Bam*HI and *Kpn*I sites. Plasmid pAΔ1503 was derived from pAdh11 by 5′ *Bal*31 deletion from the unique −124 *Adh* promoter *Sal*I site. The polylinker *Bam*HI site of pAdh11 was linked to the deletion endpoint by recutting linearized *Bal*31-digested pAdh11 with *Bam*HI and subsequent screening of ligated plasmids for *Bam*HI site regeneration. Double-stranded DNA sequencing (Kraft *et al.*, 1988) was used to determine the precise position of the *Bal*31 deletion endpoint in pAΔ1503 and other pAdh11 deletion derivatives.

Preparation of rbcS-1A−GUS constructs

Plant transformation vector pBI101 (Jefferson *et al.*, 1987) was used to create *rbcS-1A*−GUS fusions. A *Xho*I−*Bam*HI fragment spanning −1700 to +21 of the *rbcS-1A* promoter and contained in plasmid pRΔ2006 (see below) was subcloned into the polylinker *Sal*I and *Bam*HI sites. Mutated derivatives of this fragment were introduced into pBI101 as *Eco*RI−*Bam*HI fragments after excision from replicative form M13mp18 containing the *Xho*I−*Bam*HI Δ2006 fragment inserted in the M13mp18 *Sal*I and *Bam*HI sites. rbcS1A-ΔCaMV-GUS constructs were made in two steps. First CaMV fragment Δ89 (−89 to +7) and Δ167 (−167 to +7) were cloned into pBI101 cut with *Xba*I and *Bam*HI as *Eco*RV−*Bam*HI and *Ava*II−*Bam*HI fragments, respectively. The 35S CaMV fragments were obtained from CaMV sequences of plasmid pCaMVneo (Fromm *et al.*, 1986). Next, *rbcS-1A* fragments prepared by *Eco*RV−*Hin*dIII digestion (spanning −1700 to −60) were introduced by blunt-end cloning/ligation using the pBI101 *Sal*I site immediately proximal to the ΔCaMV sequences. The orientation and integrity of these subcloned fragments was verified by restriction enzyme analysis.

Mutagenesis of rbcS-1A promoter sequences

Plasmid pRΔ2006, which contains as insert sequences spanning −1700 to +21 of the *rbcS-1A* promoter was the starting point for all the *rbcS-1A* promoter manipulations described. It was derived from plasmid pATS-3 (Timko *et al.*, 1987) by *Bal*31 deletion from a unique *rbcS-1A*-promoter-distal *Sph*I site. pATS-3 contains a 2 kb *Arabidopsis* genomic insert including 1.7 kb of the *rbcS-1A* promoter and 420 bp of the coding region. In a strategy analogous to the construction of plasmid pAΔ1503 (described above), the 3′ deletion endpoints of linearized *Bal*31 digested pATS3 fragments were fused to the polylinker *Bam*HI site by recutting with *Bam*HI, intra-molecular ligation and screening for plasmids containing regenerated *Bam*HI sites and the desired deletion endpoint. For the construction of *rbcS-1A−Adh* transcriptional fusions, the desired deletion endpoints of 5′ *Bal*31 *Adh* gene deletions and 3′ *Bal*31 *rbcS-1A* promoter deletions reside in the untranslated leader mRNA of either gene. Both plasmids pAΔ1503 and pRΔ2006 contained the appropriate *rbcS-1A* promoter fragments and *Adh* gene fragments necessary for constructing such *rbcS-1A−Adh* transcriptional fusions.

Unique *Hin*dII, *Sty*I and *Stu*I sites were used to create *Bal*31 deletions within plasmid pRΔ2006. In the case of deletions initiated from pRΔ2006 *Hin*dII and *Sty*I sites, *Bgl*II linkers were ligated to *Bal*31 digested ends before intramolecular ligation and *E.coli* transformation. This made possible the construction of deletions Δ4, Δ2, Δ7, Δ8 and Δ9 by allowing combination of deletion endpoints with *rbcS-1A* fragments prepared by restriction digestion. Deleted plasmids were analyzed by restriction enzyme mapping and DNA sequence analysis using double-stranded DNA as template and oligonucleotide primers complementary to *rbcS-1A* and polylinker sequences. *RbcS-1A* sequences encompassing the LRE (−320 to −125) were isolated by subcloning a *Bst*NI−*Bam*HI fragment from plasmid pRΔ5032 into the pUC19 *Hin*dII site in either orientation to form plasmids pIGL1 and pIGL2. Plasmid pRΔ5032, like pRΔ2006 is derived from pATS-3 (Timko *et al.*, 1987) by *Bal*31 deletion from the promoter-distal *Sph*I site and has a deletion endpoint at −125 bp from the start of transcription. To facilitate the polymerization of the LRE, a *Bgl*II linker was inserted into the *Pst*I site of pIGL2. After repeated cycles of ligation of the *Bam*HI/*Bgl*II-cut LRE insert to *Bam*HI-cut parental vector and selection of recominant plasmids containing LRE sequences duplicated in direct repeats, the LRE was polymerized between 2 and 16 times.

Site directed mutagenesis was performed by the method of Kunkel (1985) with oligonucleotides synthesized by an Applied Biosystems model 380B. Site-directed mutations were introduced as 12 bp substitutions into the 1.7 kb *rbcS-1A* promoter fragment Δ2006 subcloned into M13mp18. The 12 bp substitutions included the restriction enzyme recognition sequences for *Bgl*II and *Eco*RV to facilitate the screening and identification of mutations after mutagenesis and subsequent subcloning manipulations (see Figure 3). Mutated *rbcS-1A* sequences were excised after sequence verification from replicative form M13mp18 phage and subcloned directly into *Adh* and GUS plant expression vectors (see above).

Tobacco transformations and expression analysis of transgenic plants

Agrobacterium-mediated transformation of tobacco leaf explants was performed essentially as described (Ueda *et al.*, 1989) with modifications for the use of binary transformation vectors. Binary vector pEnd4K and pBI101 constructs were introduced into *Agrobacterium* strain LBA4404 by direct freeze−thaw transformation of *E.coli* plasmid minipreparations and integrity of plasmids checked immediately before plant transformation by a modified plasmid quick-screen method as described by An *et al.* (1988). As claimed by these authors, we have found that for binary vectors, the direct freeze−thaw transformation reduces considerably the incidence of detectable plasmid rearrangements often observed after transfer from *E.coli* by conjugation. To minimize the possible contribution of undetectable plasmid mutation to the subsequent plant expression analysis, pooled Agrobacteria containing plasmids derived from at least three independent *E.coli* transformants (selected directly from cloning/ligation reactions) were used to innoculate tobacco strain SR-1 leaf explants. During tissue culture shoot regeneration, Carbenicillin (500 μg/ml) was used for *Agrobacterium* counter selection.

Transgenic T_0 tobacco plants were tested for leaf reporter gene expression after 1−2 weeks of greenhouse growth, 3−4 weeks after initial transfer of rooted tissue culture plantlets to soil. For leaf RNA analysis, T_0 or T_1 plants were grown to maturity (40−60 cm tall) in the greenhouse. Crude extracts were prepared from leaf tissue for ADH or GUS assays, by grinding small leaf samples (1 cm^2) placed in 1.5 ml Eppendorf tubes in extraction buffer with a conically tipped plastic rod attached to a motor-driven tissue homogenizer. For ADH assays, the extraction buffer consisted 0.2 M TBE (0.2 M Tris−HCl, pH 8.8, Borate 10 mM, EDTA, 10 mM), 20% sucrose, 5 mM β-mercaptoethanol, and 5 mM PMSF. After a 15 min microfuge spin, the supernatant was assayed for ADH by measuring the rate of ethanol-dependent NAD^+ reduction at 340 nm (Dolferus and Jacobs, 1984) in a Beckman DU-70 spectrophotometer. Protein was determined using the Bradford reagent (Bradford, 1976). GUS assays were performed as described (Jefferson *et al.*, 1987). Representative T_0 transgenic plants selected after leaf expression analysis were retained to obtain seed from self-pollinated flowers.

For native gel analysis, up to 50 μg of total crude extract protein was loaded onto a 7% acrylamide gel (1.5 mm thick) containing 0.2 M TBE and electrophoresed in 0.2 M TBE running buffer (pH 8.8) for 12 h at 80 V and at 5°C. The gel was stained for ADH activity in 100 mM Tris−HCl buffer, pH 8.8 containing 0.1 mM NAD^+, 0.1% ethanol, 0.1 mM nitro-blue-tetrazolium, and 0.1 mM phenazine methosulfate. Root samples were homogenized by grinding in liquid nitrogen and ADH extraction buffer.

For Northern blot RNA hybridization analysis, 1−5 g of leaf tissue was ground in liquid nitrogen with a pestle and mortar and RNA extracted with a guanidinium thiocyanate buffer followed by phenol/chloroform extraction, and ethanol/lithium chloride precipitations as described (Castresana *et al.*, 1988). Equal amounts of total RNA (20 μg) was resolved on 1% MOPS/formaldehyde agarose gel (Fourney *et al.*, 1988), and transferred to a nylon membrane (Genetran45^{TM}) as recommended by the supplier. DNA probes corresponding to *Adh* and GUS coding regions were labeled by the random priming method (Feinberg and Volgelstein, 1983). Hybridizations were performed as described (Maniatis *et al.*, 1982) and RNA signals quantitated by excising labeled bands from the nylon filters and counting ^{32}P emissions in liquid scintillant.

Acknowledgements

The authors wish to thank Dr Eliot Meyerowitz for the recombinant lambda phage At3101, Dr Peter Bringmann for ΔCaMV−GUS plasmids, advice with site-directed mutagenesis and helpful discussions. For technical assistance in the greenhouse we thank Kirsten Savinese and for critical reading of the manuscript we thank Drs M.Ahmad, U.Schindler, A.Batschauer, A.Sagar and W.Terzaghi. This work was supported by a National Institutes of Health Grant GM-38409 (to A.R.C.) and a post-doctoral National Science Foundation fellowship (to R.G.K.D.).

References

An,G., Ebert,P.R., Mitra,A. and Ha,S.B. (1988) In Glevin,S.B. and Schilperoot,R.A. (eds), *Plant Molecular Biology Manual*. Kluwer, Dordrecht, Vol. A3, pp. 7–10.

Berry-Lowe,S.L., McKnight,T.D., Shah,D.M. and Meagher,R.B. (1982) *Mol. Appl. Genet.*, **1**, 483–498.

Bradford,M.M. (1976) *Anal. Biochem.*, **72**, 248–254.

Callis,J., Fromm,M. and Walbot,V. (1987) *Genes Dev.*, **1**, 1183–1200.

Cashmore,A.R. (1983) In Kosuge,T., Meredith,C.P. and Hollaender,A. (eds), *Genetic Engineering of Plants*, Plenum Press, New York/London pp. 29–38.

Castresana,C., Staneloni,R., Malik,V.S. and Cashmore,A.R. (1987) *Plant Mol. Biol.*, **10**, 117–126.

Castresana,C., Garcia-Luque,I., Alonso,E., Malik,V.S. and Cashmore,A.R. (1988) *EMBO J.*, **7**, 1929–1936.

Chang,C. and Meyerowitz,E.M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1408–1412.

Coruzzi,G., Broglie,R., Cashmore,A. and Chua,N.-H. (1983) *J. Biol. Chem.*, **258**, 1399–1402.

Coruzzi,G., Broglie,R., Edwards,C. and Chua,N.-H. (1984) *EMBO J.*, **3**, 1671–1679.

Dean,D., van den Elzen,P., Tamaki,S., Duunsmuir,P. and Bedbrook,J. (1985a) *Proc. Natl. Acad. Sci. USA*, **82**, 4964–4968.

Dean,D., van den Elzen,P., Tamaki,S., Duunsmuir,P. and Bedbrook,J. (1985b) *EMBO J.*, **4**, 3055–3061.

Dean,C., Favreau,M., Dunsmuir,P. and Bedbrook,J. (1987) *Nucleic Acids Res.*, **15**, 4655–4668.

Dolferus,R. and Jacobs,M.J. (1984) *Biochem. Genet.*, **22**, 817–838.

Donald,R.G.K., Schindler,U., Batschauer,A. and Cashmore,A.R. (1990) *EMBO J.*, **9**, 1727–1735.

Dunsmuir,P., Smith,S. and Bedbrook,J. (1983) *Nucleic Acids Res.*, **11**, 4177–4183.

Fang,R.-X., Nagy,F., Sivasubramaniam,S. and Chua,N.-H. (1989) *Plant Cell*, **1**, 141–150.

Feinberg,A.P. and Volgelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.

Ferl,R.J. and Laughner,B.H. (1989) *Plant Mol. Biol.*, **12**, 357–366.

Fluhr,R., Kuhlemeier,C., Nagy,F. and Chua,N.-H. (1986) *Science*, **232**, 1106–1172.

Fourney,R.M., Miyakashi,J., Day,R.F. and Paterson,M.C. (1988) *Focus*, **10**, 5–7.

Fromm,M.E., Taylor,L.P. and Walbot,V. (1986) *Nature*, **319**, 791–793.

Gidoni,D., Brosio,P., Bond-Nutter,D., Bedbrook,J. and Dunsmuir,P. (1989) *Mol. Gen. Genet.*, **215**, 337–344.

Giuliano,G., Pichersky,E., Malik,V.S, Timko,M.P., Scolnic,P.A. and Cashmore,A.R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 7089–7093.

Green,P.J., Kay,S.A. and Chua,N.-H. (1987) *EMBO J.*, **6**, 2543–2549.

Green,P.J,. Young,M.-H., Cuozzo,M.Y., Murakami,K., Silverstein,P. and Chua,N.-H. (1989) *EMBO J.*, **7**, 4035–4044.

Jefferson,R.A., Kavanagh,T.A. and Bevan,M. (1987) *EMBO J.*, **6**, 3901–3907.

Katagiri,F., Lam,E. and Chua,N.-H. (1989) *Nature*, **340**, 727–729.

Klee,H.J., Yanofsky,M.F. and Nester,E.W. (1985) *Biotechnology*, **3**, 637–642.

Kraft,R., Tardiff,J., Krauter,K.S. and Leinwand,L.A. (1988) *Biotechniques*, **6**, 544–547.

Krebbers,E., Seurinck,J., Herdies,L., Cashmore,A.R. and Timko,M.P. (1988) *Plant Mol. Biol.*, **11**, 745–759.

Kuhlemeier,C., Fluhr,R., Green,P. and Chua,N.-H. (1987) *Genes Dev.*, **1**, 247–255.

Kuhlemeier,C., Cuozzo,M., Green,P.J., Goyvaerts,E., Ward,K. and Chua,N.-H. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4662–4666.

Kunkel,T.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 488–492.

Lam,E. and Chua,N.-H. (1989) *Plant Cell*, **1**, 1147–1156.

Lam,E., Benfey,P.N., Gilmartin,P.M., Fang,R.-X. and Chua,N.-H. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 7890–7894.

Manzara,T. and Gruissem,W. (1988) *Phytosynth. Res.*, **16**, 117–130.

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

Ow,D.W., Jacobs,J.D. and Howell,S.H. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4870–4874.

Pichersky,E., Bernatzky,R., Tankersley,S.D. and Cashmore,A.R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3880–3884.

Prat,S., Willmitzer,L. and Sanchez-Serrano,J.J. (1989) *Mol. Gen. Genet.*, **217**, 209–214.

Schulze-Lefert,P., Dangl,J.L., Becker-André,M., Hahlbrock,K. and Schulz,W. (1989) *EMBO J.*, **8**, 651–656.

Staiger,D., Kaulen,H. and Schell,J. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6930–6934.

Sugita,M,. Manzara,T., Pichersky,E., Cashmore,A. and Gruissem,W. (1987) *Mol. Gen. Genet.*, **209**, 247–256.

Timko,M.P., Kausch,A.P., Hand,J.M., Cashmore,A.R., Herrera-Estrella,L., Van den Broeck,G. and Van Montagu,M. (1985) In Steinbeck,K.E., Bonitz,S., Arntzen,C.J. and Bogorad,L. (eds), *Molecular Biology of the Photosynthetic Apparatus*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 381–396.

Timko,M.P., Herdies,L., de Almeida, El., Cashmore,A.R., Leemans,J. and Krebbers,E. (1987) In Phillips,H., Shoemaker,S., Ottenbrite,R.H. and Middlekauf,R.D. (eds), *Impact of Chemistry of Biotechnology*. ACS Books, Washington, DC, pp. 279–295.

Tobin,E.M. and Silverthorne,J. (1985) *Annu. Rev. Plant Physiol.*, **36**, 569–593.

Ueda,T., Pichersky,E., Malik,V. and Cashmore,A.R. (1989) *Plant Cell*, **1**, 217–227.

Weiher,H., Koenig,M. and Gruss,P. (1983) *Science*, **219**, 626–631.

Wimpee,C.F., Stiekma,W.J. and Tobin,E.M. (1983) In Goldberg,R.B. (ed.), *Plant Molecular Biology, UCLA Symposium on Molecular and Cellular Biology*. Alan R. Liss, New York, Vol. 1–2, pp. 391–401.

Received on July 28, 1989; revised on January 22, 1990