The Plant Cell, Vol. 2, 369–378, May 1990 © 1990 American Society of Plant Physiologists

Molecular Light Switches for Plant Genes

Philip M. Gilmartin, Laura Sarokin, Johan Memelink,[1] and Nam-Hai Chua[2]
Laboratory of Plant Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, New York 10021-6399

REVIEW

PATTERNS OF LIGHT-RESPONSIVE GENE EXPRESSION

Light is essential for normal plant growth and development not only as a source of energy but also as a stimulus that regulates numerous developmental and metabolic processes. The plant's responses are varied and complex and dependent upon the quality and quantity of ambient light. The initial requirement for light is as a signal for germination in many plant species. After germination in complete darkness, seedlings have a morphology distinct from light-grown ones and do not express light-inducible genes. Upon illumination of these etiolated seedlings, modifications in the transcription of light-responsive genes occur and rapid light-induced morphological changes ensue. Adaptation of plants in continuous darkness for 2 to 3 days does not cause dramatic morphological changes but results in alterations of specific transcript levels (for recent reviews, see Ellis, 1986; Kendrick and Kronenberg, 1986; Cuozzo et al., 1987; Kuhlemeier et al., 1987b; Silverthorne and Tobin, 1987; Jenkins, 1988; Nagy et al., 1988).

The most extensively studied light-responsive genes are those encoding the small subunit of ribulose-1,5-bisphosphate carboxylase-oxygenase (*rbcS*) and the chlorophyll *a/b*-binding proteins (*cab*) (see Tobin and Silverthorne, 1985; Manzara and Gruissem, 1988; Dean et al., 1989c). In several plant species an increase in the transcript levels from these genes occurs in etiolated seedlings and dark-adapted plants in response to light. This increase is mediated by the photoreceptor phytochrome and is regulated at the transcriptional level (Gallagher and Ellis, 1982; Silverthorne and Tobin, 1984; Berry-Lowe and Meagher, 1985; Mosinger et al., 1985). Phytochrome is the best characterized of the three known photoreceptors. The other two, cryptochrome and the UV-B photoreceptor, mediate their effects in response to blue and UV light, respectively (see Kendrick and Kronenberg, 1986). The expression of many light-responsive genes is modulated by more than one wavelength of light (see Tobin and Silverthorne, 1984; Ellis, 1986; Kuhlemeier et al., 1987b).

[1] Current address: Department of Plant Molecular Biology, Leiden University, Clusius Laboratory, Lassenaarseweg 64, 2333 AL Leiden, The Netherlands.
[2] To whom correspondence should be addressed.

The observation that different light responses are mediated through distinct photoreceptors raises the question of whether genes that respond to more than one wavelength do so through distinct *cis*-acting elements or whether the signal transduction pathways converge to act upon the same regulatory sequence.

Analyses of the kinetics of light-responsive gene induction show that the rate of mRNA accumulation is variable among genes and can be dependent on the developmental state of the plant (Gallagher et al., 1985; Fluhr and Chua, 1986). This variation may be due to a requirement for distinct regulatory factors or because the genes have different thresholds for a specific regulator.

Several genes are down-regulated by light, specifically those encoding phytochrome (Lissemore and Quail, 1988; Kay et al., 1989), NADPH-protochlorophyllide reductase (Batschauer and Apel, 1984; Darrah et al., 1990), and asparagine synthetase (Tsai and Coruzzi, 1990). For each of these genes, the photoresponse is mediated by phytochrome. The ability of one photoreceptor to mediate opposite patterns of expression implies that there is a branch point in the signal transduction pathway leading to these different responses.

Studies of many light-regulated genes from different species demonstrate that DNA elements responsible for light-responsive expression are located within 5′ upstream sequences (see Kuhlemeier et al., 1987b; Silverthorne and Tobin, 1987; Jenkins, 1988; Benfey and Chua, 1989; Dean et al., 1989a; Stockhaus et al., 1989). However, there is evidence that other regions of the gene can mediate changes in transcript abundance in response to light. For example, in the case of a pea gene encoding ferredoxin, sequences within the transcribed region modulate mRNA levels by affecting transcript stability (Elliot et al., 1989a). In addition, nuclear run-on experiments with petunia *rbcS* show that both upstream and downstream sequences play a role in the transcriptional regulation of these genes (Dean et al., 1989b). In contrast, downstream sequences of pea *rbcS* do not affect steady-state transcript abundance (Kuhlemeier et al., 1988b). These differences may reflect subtle variations in the mechanisms that operate to regulate *rbcS* expression in different plant species.

Previous reviews have discussed the photoreceptors through which these responses are mediated, the expression conferred by upstream sequences from several light-responsive genes, and the conserved regions within these sequences. In this review we will focus on recent advances toward an understanding of the *trans*-acting factors and their DNA binding sites that are involved in the regulation of light-responsive genes. We will emphasize the regulatory elements within the upstream region of pea *rbcS-3A*; these elements will be discussed in relation to similar motifs present in the promoters of other genes in an attempt to define the critical components of the molecular light switch. We will limit our discussion to conserved elements that have been shown to interact with nuclear proteins in vitro or that affect transcription in vivo.

REGULATORY SEQUENCES OF PEA *rbcS-3A*

Pea *rbcS-3A* (Fluhr et al., 1986) has been used as a paradigm for the study of light-responsive gene expression (Kuhlemeier et al., 1987b). Figure 1 shows the upstream region of *rbcS-3A*. Analysis of this gene has uncovered three independent upstream regions (−410 to −330, −330 to −170, and −166 to −50; see Figure 1) that confer expression upon the −50 deleted *rbcS-3A* promoter (Davis et al., 1990) in response to white light. Such reiteration of genetic information may enable the gene to compensate for changing environmental conditions and respond to different developmental cues (Nagy et al., 1988). It is possible that the presence of multiple regulatory elements may enable *rbcS* expression to be maintained at various light fluences or if the abundance of specific *trans*-acting factors varies during different stages of development. In addition, distinct regulatory elements may mediate responses to the different wavelength components of white light. Most experiments aimed at identifying the specific elements that confer light responsiveness have focused on the phytochrome-responsive (P.M Gilmartin and N.-H. Chua, unpublished data) −166 promoter deletion (Kuhlemeier et al., 1987a). This region contains two physically separable light-responsive elements (LREs), one located between −166 and −50 and the other located downstream of −50 (Kuhlemeier et al., 1989). Additional LREs have also been identified between −410 and −170 of *rbcS-3A* (Kuhlemeier et al., 1988a). The multiplicity of LREs raises the question of whether different DNA-protein complexes are involved in light-responsive transcription.

Recent work on light-responsive genes has focused on the nuclear proteins that interact with specific elements within the upstream regions of these genes. Many of these elements are shared among different genes from a number of species (Manzara and Gruissem, 1988; Dean et al., 1989c). To facilitate our discussion of these recent findings, we will use pea *rbcS-3A* as a model. Figure 1 presents the sequence of pea *rbcS-3A* between −410 and −1 and shows the location of the factor binding sites defined within this promoter as well as those inferred by sequence homology to sites defined in other genes. The contribution of each of these different DNA-protein complexes to light-responsive gene expression is discussed in detail in the following sections.

GT-1 BOXES

Comparison of several different *rbcS* promoters reveals conserved sequence motifs including box II (−151 to −138) and box III (−125 to −114) within pea *rbcS-3A*. These two elements bind the nuclear factor GT-1. Four additional sequences located further upstream, termed boxes II* (−224 to −213), III* (−257 to −248), II** (−386 to −372), and III** (−360 to −347) also bind GT-1 (Green et al., 1987, 1988). We will refer to these elements as GT-1 boxes. Sequences homologous to the GT-1 boxes have also been noted within upstream regions of numerous other genes (Stockhaus et al., 1987; Manzara and Gruissem, 1988; Dean et al., 1989c; Elliot et al., 1989b; Kay et al., 1989). The sequences of defined GT-1 boxes show much variation when compared with the box II core sequence GGTTAA required for GT-1 binding in vitro (Green et al., 1988). These variations may indicate that GT-1 can bind to many related sequences. Alternatively, there may be a family of GT-1 proteins whose members do not differentiate between sites in vitro but do recognize specific sequences in vivo. Mutational analyses of the −166 deleted *rbcS-3A* promoter have uncovered differences in the ability of boxes II and III to activate transcription (P.M. Gilmartin and N.-H. Chua, unpublished results). A promoter containing paired box II elements, as opposed to one each of boxes II and III, retains wild-type levels of transcriptional activity. However, a promoter composed of paired box III elements confers 20-fold lower levels of transcription, and this reduction correlates with a decrease in affinity for GT-1 in vitro (P.M. Gilmartin and N.-H. Chua, unpublished data). The strength of GT-1 binding to an element may, in turn, influence the level of transcription. Therefore, different expression levels could be generated through the interaction of the same factor with related sequences.

There is an absolute requirement for the two GT-1 binding sites within a −170 deleted *rbcS-3A* promoter for transcriptional activity. Substitution of G^{-146} and G^{-145} within box II for C residues results not only in the loss of GT-1 binding in vitro (Green et al., 1988) but also in the loss of transcription in vivo (Kuhlemeier et al., 1988a). Further mutational analysis of boxes II and III within the −166 *rbcS-3A* promoter revealed that sequences outside of the in vitro defined box II and III core binding sequences are also required for transcriptional activity in vivo (L. Sarokin and N.-H. Chua, unpublished data). Indeed, a promoter deletion to −149 that removes 2 bp from the 5′ end of box II was previously shown to reduce GT-1 binding (Green et al., 1988) and to abolish transcriptional activity

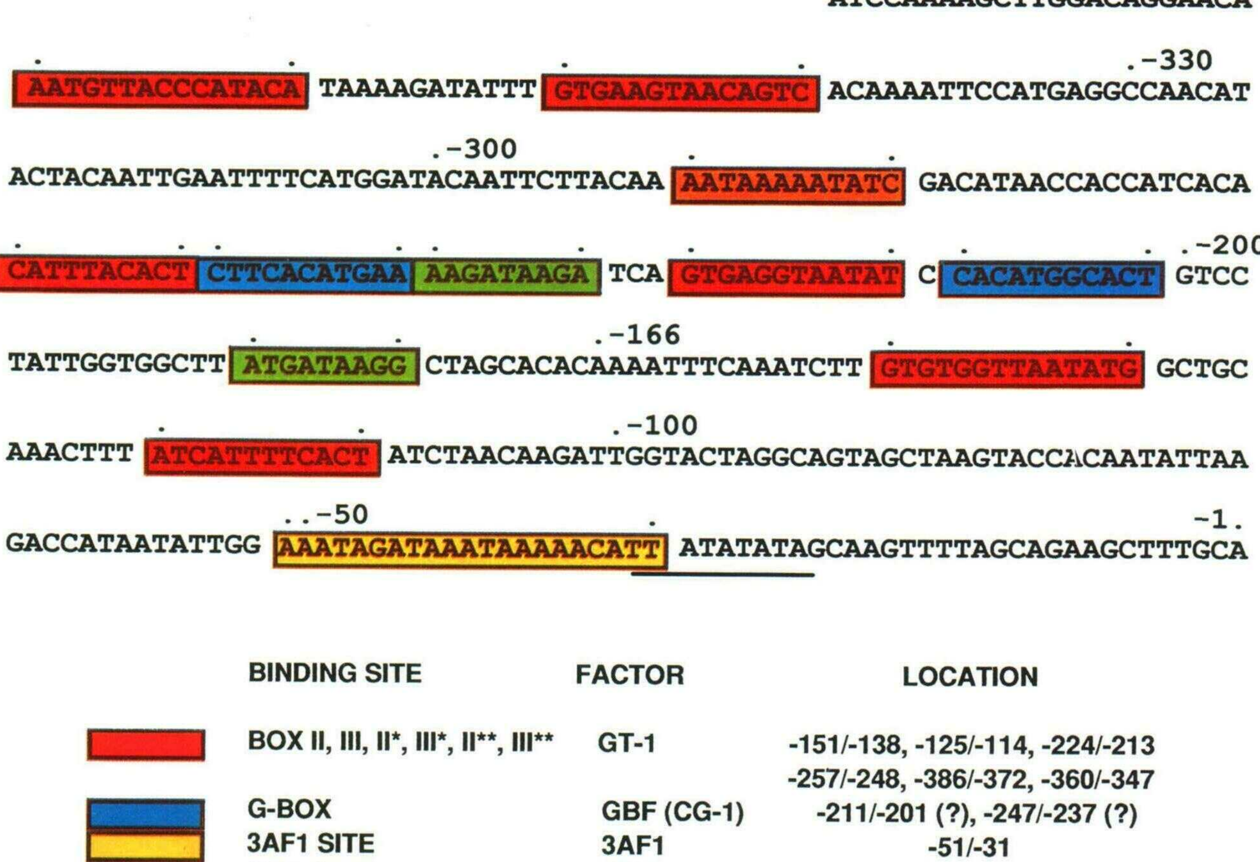

Figure 1. Nucleotide Sequence of Pea *rbcS-3A* between −410 and −1 Relative to the Transcription Start Site.

Elements defined as binding sites for nuclear proteins are indicated. The dots (·) denote the endpoints of these elements. The sequence locations indicated by a question mark (?) identify those elements that have not yet been shown to bind proteins within the context of *rbcS-3A* but show homology to binding sites defined in other light-responsive genes.

(Kuhlemeier et al., 1987a). These observations provide a correlation between GT-1 binding in vitro and *rbcS-3A* transcriptional activity in vivo.

There is a requirement for an interaction between the two GT-1 boxes within the −166 *rbcS-3A* promoter for transcriptional activity. Deletions of up to 8 bp from between these two sites do not significantly affect expression. By contrast, introduction of as few as 2 bp between these sites severely reduces promoter activity, as do insertions of up to 21 bp. However, neither expansion nor contraction of the distance between the GT-1 boxes affects binding of GT-1 in vitro. These results suggest that GT-1 can in some cases interact with its binding site in a nonproductive manner (Gilmartin and Chua, 1990). Consistent with this notion is the observation that a box II tetramer can bind GT-1 in vitro (Green et al., 1988) but cannot potentiate transcription when fused to either the −50 *rbcS-3A* or the −46 cauliflower mosaic virus (CaMV) 35S promoters (Davis et al., 1990). In striking contrast, this same tetramer can confer light-responsive transcriptional activity upon the −90 CaMV 35S promoter (Lam and Chua, 1990). Figure 2 summarizes these data. Control experiments demonstrated that a tetramer of box II elements containing the GG to CC mutation is not active when fused to −90 CaMV 35S (Lam and Chua, 1990). The mutant box II tetramer shows greatly reduced affinity for GT-1 in vitro (Green et al., 1988). These observations clearly implicate GT-1 binding in light-responsive transcription. Taken together, these data suggest that GT-1 binding is necessary, but not sufficient, for transcriptional activation.

The ability of a box II tetramer to confer light respon-

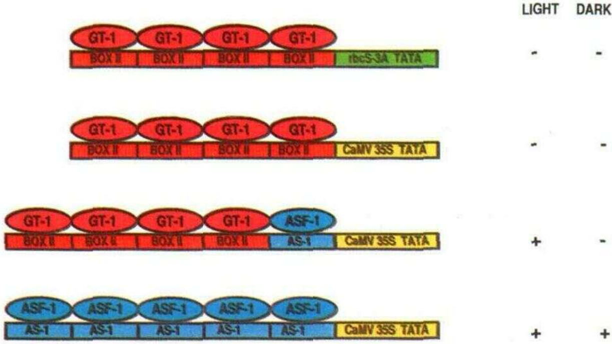

Figure 2. Chimeric Promoters Used To Define the Role of Specific Elements.

Truncated promoters used to assay *cis*-acting elements are the −50 deleted *rbcS-3A* (green) and the −46 to +15 CaMV 35S promoter (yellow) fused to the GUS coding region. GT-1 boxes (box II) and GT-1 are shown in red. The *as-1* elements and ASF-1 are shown in blue.

siveness only when fused to the −90 CaMV 35S promoter is most likely due to an interaction between GT-1 bound to the tetramer and additional factors bound to the CaMV −90 35S promoter. A likely candidate for such a factor is activation sequence factor 1 (ASF-1) that interacts with the activation sequence 1 (as-1) element located between −83 and −65 of the CaMV 35S promoter (see Figure 2 and Lam et al., 1989b). Fusion of an *as-1* tetramer to the −90 CaMV 35S promoter gives constitutive activity, demonstrating that box II is a critical component of the molecular light switch (Lam and Chua, 1990). These observations strongly suggest a requirement for an interaction between GT-1 and other proteins for the formation of a stable transcription initiation complex.

Insertion of GT-1 binding sites into the constitutive CaMV 35S promoter suggests an additional role for these sites in mediating light responsiveness (Kuhlemeier et al., 1987a). Within the *rbcS-3A* promoter, and when fused to the −90 CaMV 35S promoter, GT-1 boxes act as positive elements. However, when introduced into the CaMV 35S promoter, these same sequences act to silence transcription in the dark, suggesting that a DNA-protein complex at this site represses dark transcription. (Kuhlemeier et al., 1987a). This result does not exclude the possibility that a protein distinct from GT-1 may act as the repressor. The introduction of mutant GT-1 binding sites within this context would help to resolve this point.

THE G BOX

Gel retardation analyses with tomato *rbcS-3A* upstream sequences reveal the presence of a protein, termed GBF (G box factor), in both tomato and *Arabidopsis* nuclear extracts (Giuliano et al., 1988). Furthermore, upstream sequences from other tomato, pea, and *Arabidopsis rbcS* genes can compete for GBF binding. The GBF binding site was identified as a conserved sequence, the G box; a consensus, TCTTACACGTGGCAYY, was derived by sequence comparisons among several light-responsive promoters (Giuliano et al., 1988). A tobacco nuclear protein, CG-1, that interacts with the sequence CACGTG within the *Antirrhinum* chalcone synthase promoter (CHS) shows the same binding specificity as GBF (Staiger et al., 1989). This sequence is also found within the upstream regions of many other genes (Schulze-Lefert et al., 1989b).

Analysis of the upstream regions of different genes suggests a requirement for the G box for transcriptional activity. An internal deletion of tobacco *cabE* that removes a 134-bp fragment containing the G box results in a complete loss of light-responsive transcriptional activity (Castresana et al., 1988). However, in this deletion two GT-1 boxes are also removed. Similar observations on the critical role of the G box were made after mutation of this element within the *Arabidopsis rbcS-1A* promoter (see Datta and Cashmore, 1989). Deletion of the G box from the *pallida* gene of *Antirrhinum* after the imprecise excision of a transposable element severely reduces expression levels (Coen et al., 1986; Almeida et al., 1989). Within tobacco *cabE*, the G box is present in upstream sequences that confer a response to white light when fused to a heterologous promoter. Within both *Antirrhinum* and parsley chalcone synthase promoters, the CACGTG motif is located within regions that confer a response to UV light in transgenic tobacco (Kaulen et al., 1986) and in parsley protoplasts (Lipphardt et al., 1988; Schulze-Lefert et al., 1989a, 1989b), respectively. Within the parsley CHS promoter, mutation of the G box results in a loss of UV light-responsive expression (Schulze-Lefert et al., 1989a). However, it has not yet been demonstrated whether the G box acts as a regulatory component of these light-responsive elements or whether it just modulates the quantitative level of expression.

The strongest evidence of a regulatory role for the G box in UV light responsiveness comes from in vivo footprint analyses of the parsley CHS promoter. After illumination of parsley protoplasts with UV light, four UV light-inducible protected regions of the promoter are observed (Schulze-Lefert et al., 1989a, 1989b). Two of these regions correspond to the conserved CACGTG motifs (G box). Each of these sequences is adjacent to an additional protected region (Schulze-Lefert et al., 1989a, 1989b). The appearance of UV light-inducible, in vivo footprints suggests three possibilities: (1) the factor is synthesized in response to this stimulus, (2) it is modified to facilitate its binding, or (3) a second modified or labile factor facilitates its binding.

Two sequences with homology to the G box are present within the pea *rbcS-3A* 5′ region (see Figure 1). The TATA proximal element (−211 to −201) shows high homology

to the defined G box with a single base substitution in the core (CACATG), whereas the distal element (−247 to −237) contains the same CACATG core but is otherwise distinct. We note that this latter sequence (CTTCACATGA) shows homology to a sequence (CTTCACTTGATGTATC) identified as a UV light-inducible footprint adjacent to one of the parsley CHS G boxes (Schulze-Lefert et al., 1989a). The significance of this observation and whether GBF (CG-1) can interact with these regions in pea *rbcS-3A* are not known at present.

It is not yet clear whether the proteins responsible for protection of the CACGTG motif in vivo (Schulze-Lefert et al., 1989a, 1989b) are the same as the in vitro defined factors GBF (Giuliano et al., 1988) and CG-1 (Staiger et al., 1989). GBF is present in nuclear extracts from both light-grown and dark-adapted tomato plants, although the GBF-DNA complex formed with the extract from dark-adapted plants has a faster mobility compared with extract from light-grown plants (Giuliano et al., 1988). Studies of CG-1 activity in nuclear extracts from tobacco seedlings grown either with or without UV light demonstrate its presence in both (Staiger et al., 1989). These observations suggest that different proteins may have been identified in the in vivo and in vitro assays. Alternatively, they are the same protein and a modification during preparation of nuclear extracts from plants grown in the absence of UV light results in its ability to bind.

THE 3AF1 BINDING SITE

Analysis of the pea *rbcS-3A* promoter in Figure 1 identifies a region between nucleotides −51 and −31 that is protected from DNase I digestion by pea nuclear extract (E. Lam, Y. Kano-Murakami, P.M Gilmartin, B. Niner, and N.-H. Chua, manuscript submitted). Synthetic oligonucleotide probes containing four copies of the footprinted region, AAATAGATAAATAAAAACATT (−51 to −31), are shown by gel shift analyses to bind a protein designated 3AF1. A 3-bp mutation within each of the binding sites of the tetramer, AAATAGAGATCTAAAAACATT, severely reduces binding of 3AF1 in vitro.

Tetramers of the wild-type 3AF1 binding site ligated to the −90 CaMV 35S-β-glucuronidase (GUS) gene fusion confer transcriptional activity in all cell types in transgenic tobacco independently of light, suggesting that 3AF1 is present in all cell types. Furthermore, fusion of the 3AF1 tetramer to −46 CaMV 35S-GUS does not give any transcriptional activity, indicating that a tetramer of the 3AF1 binding site cannot function independently (E. Lam, Y. Kano-Murakami, P.M. Gilmartin, B. Niner, and N.-H. Chua, manuscript submitted). This observation is similar to that for the box II tetramer fused to the same CaMV 35S promoters (Lam and Chua, 1990). The role of the 3AF1 binding site within the *rbcS-3A* promoter is unclear, and it is not known whether 3AF1 binds to other genes. However, its contribution to *rbcS-3A* activity is most likely derived from an interaction with other proteins bound to adjacent sequence elements.

The recent cloning of a cDNA encoding a putative zinc-finger protein with the same binding specificity as 3AF1 reveals that this factor is encoded by a small multigene family and that various sized mRNA transcripts are present in different organs and in leaves of both light-grown and dark-adapted leaves (E. Lam, Y. Kano-Murakami, P.M. Gilmartin, B. Niner, and N.-H. Chua, manuscript submitted). The constitutive expression conferred by a tetramer of the 3AF1 binding site may, therefore, reflect transcriptional activation by distinct members of the family in different organs.

THE AT-1 BINDING SITE

The binding site for a factor termed AT-1 was defined by Bal31 deletion of labeled fragments of the pea *rbcS-3.6* promoter. These studies identify two overlapping motifs within the −574 to −433 region of this gene (Datta and Cashmore, 1989). A consensus sequence, AATATTTT-TATT, was derived by comparison of similar sequences within upstream regions of other *rbcS* and *cab* genes. More than one AT-1 consensus sequence is present within both tobacco *cabE* and pea *rbcS-3.6*. The presence of AT-1 boxes within the upstream regions of tobacco *cabE* and tomato *rbcS-3A* was confirmed by gel shift analyses and competition studies. It is likely that the similar motif within pea *rbcS-3A* (−288 to −276) (see Figure 1) interacts with the same protein, although AT-1 has not yet been shown to bind to this sequence. The relationship between AT-1 and 3AF1, which binds to a similar sequence, is not known. A promoter fragment containing the AT-1 box and oligonucleotides of the consensus binding site sequence form two DNA-protein complexes with nuclear extract (Datta and Cashmore, 1989). These results possibly suggest the presence of different proteins that recognize this sequence or the existence of different forms of AT-1 (Datta and Cashmore, 1989).

Analyses of the tomato *rbcS-3A* and tobacco *cabE* upstream sequences ascribe functions for fragments that include AT-1 boxes. A possible positive function for the AT-1 box was suggested by deletion analysis of the tomato *rbcS-3A* gene (Ueda et al., 1989). Deletion of only 90 bp from the 1100-bp promoter results in a dramatic reduction in transcriptional activity; this deletion included the 5′ end of the AT-1 box. Inversion of the same 90-bp fragment on the −984 truncated tomato *rbcS-3A* promoter does not restore transcriptional activity. The loss of activity is, therefore, most likely due to the disruption of the AT-1 box (Datta and Cashmore, 1989). By contrast, the AT-1 box of tobacco *cabE* is located within an AT-rich region with the properties of a negative element (Castresana et al., 1988). Additional mutational analysis of this region is required to

define a functional role for the AT-1 box and to delineate the sequences responsible for the observed negative regulation.

The most striking aspect of these studies is the demonstration that AT-1 binding is affected by its phosphorylation state (Datta and Cashmore, 1989). The addition of either Mg^{2+}-ATP or Mg^{2+}-GTP to the nuclear extract results in the loss of binding activity. This observation suggests that formation of an AT-1-DNA complex is dependent upon dephosphorylation of the factor in vitro. Protein phosphorylation is a mechanism that regulates the DNA-binding activity of serum response factor in mammalian cells, although in this case the in vitro binding activity of serum response factor is dependent upon phosphorylation of the factor (Prywes et al., 1988). There is at present no evidence that AT-1 is involved in the light response. Nevertheless, it would be interesting to determine whether dephosphorylation of this factor is light dependent.

GATA MOTIFS

Comparisons between the upstream sequence of three *cab* genes from *Nicotiana plumbaginifolia* showed a region of strong homology between nucleotides −150 and −100 relative to the translation start site. Further comparisons of this region to *cab* genes from other species uncovered a striking conservation of the sequence GATA within this region (Castresana et al., 1987).

Functional analysis of *N. plumbaginifolia cab-E* shows that sequences between −112 and +36 that contain three copies of the sequence GATA are required for light-responsive activity in the context of the full −1554-bp upstream region (Castresana et al., 1988). Replacement of the −112 to +36 region by the −145 truncated *nopaline synthase* promoter results in constitutive expression, suggesting either that an element is present between −112 and +36 that represses activity in the dark or that a factor that activates expression is only present in the light (Castresana et al., 1988). However, a specific role for the GATA motifs within this sequence has not been defined. In the case of petunia *cab22R*, the introduction of a 2-bp mutation into the TATA-proximal GATA motif results in a fivefold reduction in transcriptional activity. No further decrease was noted when mutations were introduced in all three GATA motifs (Gidoni et al., 1989). The combined data from the analysis of *cabE* and *cab22R* strongly suggest that the GATA elements play a role in high-level expression in the light.

A tobacco nuclear protein factor, termed activation sequence factor 2 (ASF-2), that binds to the −105 to −85 region of the CaMV 35S promoter was identified by DNase I footprinting (Lam and Chua, 1989). Gel shift competition analyses with a tetramer of this sequence, GTGGATT-GATGTGATATCTCC, demonstrate that extracts prepared from both light-grown and dark-adapted plants contain approximately equal amounts of ASF-2. Methylation interference studies defined 2 nucleotides within the binding site, G^{-93} and G^{-98}, as essential for ASF-2 binding (Lam and Chua, 1989). A tetramer of a mutant sequence containing substitutions of these 2 G residues, GTGGATT-CATGTAATATCTCC, is unable to interact with ASF-2. A similarity between the ASF-2 binding site (*as-2*) and the paired GATA motifs found within several *cab* genes was observed (Lam and Chua, 1989). Analyses with the petunia *cab22L* −75 to −59 region, GTAGATAGAGATATCAT (Gidoni et al., 1989), which contains tandem GATA motifs, show that it forms two complexes with tobacco nuclear extract (Lam and Chua, 1989); one of these is competed by the *as-2* tetramer. These data demonstrate that ASF-2 binds not only to the CaMV 35S promoter but also to the conserved tandem GATA motifs within the *cab22L* promoter, suggesting that the GATA elements within *cab* genes are most likely the same as those in *as-2*. The identity of the second complex has not yet been determined. An *as-2* tetramer fused to the −90 CaMV 35S promoter confers leaf-specific expression in transgenic tobacco that is insensitive to light (Lam and Chua, 1989). This observation suggests that the tandem GATA motifs mediate cell-type specificity but not light responsiveness.

Several GATA-like motifs are present within the pea *rbcS-3A* upstream region, but none are found as tandem repeats (see Figure 1). However, a conserved motif containing a single GATA element, ATGATAAGG, is present in virtually all *rbcS* and *cab* genes (Castresana et al., 1987; Manzara and Gruissem, 1988; Gidoni et al., 1989). This element was highlighted as a conserved region within petunia *rbcS* genes by Dean et al. (1985) and shows homology to the GATA motif proximal to the TATA box in many *cab* genes (Castresana et al., 1987; Gidoni et al., 1989). The reverse of this sequence, CCTTATCAT, was also noted within petunia *rbcS* genes (Dean et al., 1985) and independently identified by a computer search within the promoters of numerous phytochrome-responsive genes (Grob and Stuber, 1987). The sequence ATGA-TAAGG is located between −189 and −170 of the pea *rbcS-3A* upstream sequence (Figure 1). Giuliano et al. (1988) refer to this element in several other *rbcS* genes as the I box. Recent analyses with synthetic tetramers containing this motif have identified a factor termed GAF-1 (GA factor 1) in tobacco nuclear extracts that binds specifically to this sequence (J. Memelink, P.M. Gilmartin, and N.-H. Chua, unpublished data) (see Figure 1). GAF-1 is present in greater abundance in extracts prepared from light-grown as opposed to dark-adapted plants. Gel shift competition studies demonstrated that GAF-1 is distinct from ASF-2 and 3AF1, both of which bind to sequences containing GATA motifs (J. Memelink, P.M. Gilmartin, and N.-H. Chua, unpublished data). These observations point to a distinction between different GATA elements present within light-responsive promoters.

The ATGATAAGG element is located 5′ of the −166 *rbcS-3A* promoter (see Figure 1) and its functional role in *rbcS-3A* is currently being defined. Figure 1 shows that a sequence with homology to the ATGATAAGG element is located further upstream of *rbcS-3A*, but whether this sequence, AAGATAAGA (−236 to −228), can interact with GAF-1 is not known. The contribution of these elements to the light-responsive expression conferred by the −410 to −170 region of *rbcS-3A* is unclear at present because this region also contains binding sites for the factors GT-1 (Green et al., 1987), GBF (CG-1) (Giuliano et al., 1988; Staiger et al., 1989), and AT-1 (Datta and Cashmore, 1989), as shown in Figure 1. Further analyses of the expression conferred by chimeric promoters containing the ATGATAAGG motif should enable a functional definition of its role in *rbcS-3A*.

CONCLUDING REMARKS

Dissection of upstream sequences from several light-regulated genes has revealed the complexity of the *cis*-acting elements through which light-responsive transcription is mediated. Interpretation of the role played by individual elements is complicated by the distinct responses seen in different tissues at various stages of development (Gallagher et al., 1985; Fluhr and Chua, 1986). The choice of material used for these analyses is, therefore, critical. To define precisely the influence of a particular motif within an LRE, it is necessary to determine its contribution in terms of a specific photoreceptor. Mutation of a factor binding site and the analysis of the effect in vivo after illumination with white light may not uncover a function for the element. It is possible that distinct elements exert their effects in response to specific wavelengths of light, for example red, blue, or UV. The critical role of the element may only be observed after illumination with a defined wavelength of light.

The ability of several upstream sequences to confer regulated expression in heterologous plant species suggests that the mechanisms that mediate light-responsive transcription are conserved to some degree. However, not all mechanisms are conserved among all species, as suggested by the observation that wheat *rbcS* is inactive in transgenic tobacco plants (Keith and Chua, 1986). There is also a diversity of responses to light even for different members of a gene family. For example, tomato *rbcS-2* and *rbcS-3A* genes (Manzara and Gruissem, 1988) and maize *cab-1* (Sullivan et al., 1989) are expressed at a high level in the dark despite the presence of the same elements within their upstream sequences, as found in light-responsive *rbcS* and *cab* genes (Castresana et al., 1988; Manzara and Gruissem, 1988). The complexity of sequences that mediate light-responsive expression emphasizes the need for the dissection of upstream regions from different genes in order to define the critical components of the light-responsive elements.

Reiteration of sequence motifs in the upstream regions of light-responsive genes compounds the problem of defining the role of a specific element. After mutation of a sequence motif present in more than one copy, its function may not be uncovered because similar elements can compensate for the loss. The −410 *rbcS-3A* promoter serves as a clear example. Mutation of individual GT-1 boxes within this upstream region does not affect gene expression when assayed under high levels of white light (Kuhlemeier et al., 1987a). However, the use of different assay conditions (e.g., low light intensity) may uncover a function for such repeated elements. The presence of different sequences that perform similar functions may further complicate the analysis. For example, the two physically separable LREs within the −166 deleted *rbcS-3A* promoter share no conserved sequence elements (Kuhlemeier et al., 1989). This suggests that different elements may possess overlapping roles such that mutation of one may be compensated for by the other.

Differences among various *trans*-acting factor binding sites in the upstream regions of light-responsive genes may reflect distinct roles for these elements under diverse environmental conditions or at distinct developmental stages. Some elements play a regulatory role in mediating the light response, whereas others only contribute to the overall level of expression. The regulatory contribution of a specific element cannot be assessed easily by mutational analysis within the cognate promoter alone. The dramatic differences in expression pattern mediated by the *as-2* (GATA box) and GT-1 box tetramers fused to the −90 CaMV 35S promoter demonstrate this point (Lam and Chua, 1989, 1990). A positive role for both of these elements within an LRE was previously demonstrated by mutational analysis (Castresana et al., 1988; Kuhlemeier et al., 1988a; Fang et al., 1989; Gidoni et al., 1989); however, gain of function experiments have resolved the specific contribution of each of these sequences. These analyses clearly demonstrate that *as-2* (GATA box) mediates tissue specificity independently of light (Lam and Chua, 1989). By contrast, GT-1 boxes mediate both cell-type specificity and light responsiveness (Lam and Chua, 1990). Mutational analyses in the context of the native promoter did not permit this distinction.

Expression patterns conferred by the GT-1 box, *as-2* (GATA box), and 3AF1 site have been defined by fusion of tetramers of these elements to the truncated −46 and −90 CaMV 35S promoters (Lam and Chua, 1989, 1990; E. Lam, Y. Kano-Murakami, P.M. Gilmartin, B. Niner, and N.-H. Chua, manuscript submitted). These studies suggest that factors bound to all these sequences likely interact with ASF-1, the cognate factor for *as-1* (Lam et al., 1989b). It is, therefore, likely that factors bound to these various motifs within the upstream regions of light-responsive genes interact with each other to form a light-dependent

transcription initiation complex. Several of the proteins within this complex are likely to affect only the level of expression conferred by the LRE, and others will act as regulatory components. Plausible examples of the former are ASF-2 (GATA factor) (Lam and Chua, 1989) and 3AF1 (E. Lam, Y. Kano-Murakami, P.M. Gilmartin, B. Niner, and N.-H. Chua, manuscript submitted), whereas GT-1 (Green et al., 1987) may be an example of the latter. The category into which GBF (CG-1) (Giuliano et al., 1988; Staiger et al., 1989), AT-1 (Datta and Cashmore, 1989), and GAF-1 (J. Memelink, P.M. Gilmartin, and N.-H. Chua, unpublished data) fall is as yet unclear.

Increases in the transcription rates of pea *rbcS-3A* and other light-responsive genes after illumination may occur as a consequence of transcriptional activation in the light, transcriptional repression in the dark, or a combination of both. The mechanisms by which regulatory *trans*-acting factors could mediate these diverse responses are essentially similar. The factor may be synthesized de novo; alternatively, it may be present constitutively and the DNA binding domain could be modified to facilitate binding, or modification of its activation domain may affect its ability to interact with other components of the transcriptional machinery. The observation that binding of AT-1 depends upon its modification state (Datta and Cashmore, 1989) and the presence of both GT-1 (Green et al., 1987) and GBF (CG-1) (Giuliano et al., 1988; Staiger et al., 1989) in the light and dark may reflect such regulatory processes. The demonstration that the inhibition of calmodulin blocks *cab* expression in vivo suggests that the transcriptional response to light is mediated through protein modification by phosphorylation (Lam et al., 1989a). Moreover, the requirement of cytoplasmic protein synthesis for transcriptional activation of *rbcS* and *cab* in both pea and tobacco suggests that at least one component of the signal transduction chain is a labile factor (Lam et al., 1989c).

Isolation of the genes encoding proteins involved in light-responsive transcription will enable the characterization of the terminal components of the signal transduction chain. These proteins may then be used as biochemical substrates to identify the modification enzymes that modulate their transcriptional activity. The activities of these *trans*-acting proteins and their modified derivatives can be assessed in a recently developed plant in vitro transcription system (K. Yamazaki, F. Katagiri, H. Imaseki, and N.-H. Chua, manuscript submitted). The recent isolation of cDNA sequences encoding proteins with the same binding specificities of 3AF1 (E. Lam, Y. Kano-Murakami, P.M. Gilmartin, B. Niner, and N.-H. Chua, manuscript submitted) and GT-1 (P.M. Gilmartin, J. Memelink and N.-H. Chua, unpublished data) in conjunction with the use of transgenic plant technologies (for recent reviews see Weising et al., 1988; Willmitzer, 1988) may help to resolve the role of these proteins in the light-responsive expression of *rbcS-3A*.

Definition of the individual steps linking photoperception and transcriptional regulation will require combined approaches focusing on the photoreceptors that perceive the signal, the machinery that executes the transcriptional response, and the intermediary components of the transduction pathway. Recent advances in the identification of several *cis*-acting elements and *trans*-acting factors that interact to confer light responsiveness provide insight into a few of the components required for the modulation of specific transcript levels in response to the environmental stimulus of light.

ACKNOWLEDGMENTS

We would like to thank our colleagues in the laboratory for many helpful discussions and especially Drs. Steve Kay, Jörg Stockhaus, and Philip Benfey for their comments on the manuscript. P.M.G. is supported by a postdoctoral fellowship from the Winston Foundation, L.S. is supported by a National Institutes of Health postdoctoral fellowship, and J.M. was supported by a North Atlantic Treaty Organization postdoctoral fellowship.

Received March 16, 1990; revised March 27, 1990.

REFERENCES

Almeida, J., Carpenter, R., Robbins, T.P., Martin, C., and Coen, E. (1989). Genetic interactions underlying flower color patterns in *Antirrhinum majus.* Genes Dev. **3,** 1758–1767.

Batschauer, A., and Apel, K. (1984). An inverse control by phytochrome of the expression of two nuclear genes in barley. Eur. J. Biochem. **143,** 593–597.

Benfey, P.N., and Chua, N.-H. (1989). Regulated genes in transgenic plants. Science **244,** 174–181.

Berry-Lowe, S.L., and Meagher, R.B. (1985). Transcriptional regulation of a gene encoding the small subunit of ribulose-1,5-bisphosphate carboxylase in soybean tissue is linked to the phytochrome response. Mol. Cell. Biol. **5,** 1910–1917.

Castresana, C., Staneloni, R., Malik, V.S., and Cashmore, A.R. (1987). Molecular characterization of two clusters of genes encoding the Type I CAB polypeptides of PSII in *Nicotiana plumbaginifolia.* Plant Mol. Biol. **10,** 117–126.

Castresana, C., Garcia-Luque, I., Alonso, E., Malik, V.S., and Cashmore, A.R. (1988). Both positive and negative regulatory elements mediate expression of a photoregulated CAB gene from *Nicotiana plumbaginifolia.* EMBO J. **7,** 1929–1936.

Coen, E.S., Carpenter, R., and Martin, C. (1986). Transposable elements generate novel spatial patterns of gene expression in *Antirrhinum majus.* Cell **47,** 285–296.

Cuozzo, M., Kay, S.A., and Chua, N.-H. (1987). Regulatory circuits of light-responsive genes. In Temporal and Spatial Regulation of Plant Genes, D.P.S. Verma and R.B. Goldberg, eds (Vienna: Springer-Verlag), pp. 131–153.

Darrah, P., Kay, S., Teakle, G., and Griffiths, T. (1990). Cloning

and sequencing of protochlorophyllide reductase. Biochem. J. **265,** 789–798.

Datta, N., and Cashmore, A.R. (1989). Binding of a pea nuclear protein to promoters of certain photoregulated genes is modulated by phosphorylation. Plant Cell **1,** 1069–1077.

Davis, M.C., Yong, M.-H., Gilmartin, P.M., Goyvaerts, E., Kuhlemeier, C., Sarokin, L., and Chua, N.-H. (1990). Minimal sequence requirements for the regulated expression of *rbcS-3A* from *Pisum sativum* in transgenic tobacco plants. Photochem. Photobiol., in press.

Dean, C., van den Elzen, P., Tamagi, S., Dunsmuir, P., and Bedbrook, J. (1985). Differential expression of the eight genes of the petunia ribulose bisphosphate carboxylase small subunit multi-gene family. EMBO J. **4,** 3055–3061.

Dean, C., Favreau, M., Bedbrook, J., and Dunsmuir, P. (1989a). Sequences 5′ to translation start regulate expression of petunia *rbcS* genes. Plant Cell **1,** 209–215.

Dean, C., Favreau, M., Bond-Nutter, D., Bedbrook, J., and Dunsmuir, P. (1989b). Sequences downstream of translation start regulate quantitative expression of two petunia *rbcS* genes. Plant Cell **1,** 201–208.

Dean, C., Pichersky, E., and Dunsmuir, P. (1989c). Structure, evolution and regulation of *RbcS* genes in higher plants. Annu. Rev. Plant Physiol. **40,** 415–439.

Elliot, R.C., Dickey, L.F., White, M.J., and Thompson, W.F. (1989a). *cis*-Acting elements for light regulation of pea ferredoxin I gene expression are located within transcribed sequences. Plant Cell **1,** 691–698.

Elliot, R.C., Pedersen, T.J., Fristensky, B.A., White, M.J., Dickey, L.F., and Thompson, W.F. (1989b). Characterization of a single copy gene encoding ferredoxin I from pea. Plant Cell **1,** 681–690.

Ellis, R.J. (1986). Photoregulation of plant gene expression. Biosci. Rep. **6,** 127–135.

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.

Fluhr, R., and Chua, N.-H. (1986). Developmental regulation of two genes encoding ribulose-bisphosphate carboxylase small subunit in pea and transgenic petunia plants: Phytochrome response and blue-light induction. Proc. Natl. Acad. Sci. USA **83,** 2358–2362.

Fluhr, R., Moses, P., Morelli, G., Coruzzi, G., and Chua, N.-H. (1986). Expression dynamics of the pea *rbcS* multigene family and organ distribution of the transcripts. EMBO J. **5,** 2063–2071.

Gallagher, T.F., and Ellis, R.J. (1982). Light-stimulated transcription of genes for two chloroplast polypeptides in isolated pea leaf nuclei. EMBO J. **1,** 1493–1498.

Gallagher, T.F., Jenkins, G.I., and Ellis, R.J. (1985). Rapid modulation of transcription of nuclear genes encoding chloroplast proteins by light. FEBS Lett. **186,** 241–245.

Gidoni, D., Brosio, P., Bond-Nutter, D., Bedbrook, J., and Dunsmuir, P. (1989). Novel *cis*-acting elements in petunia *Cab* gene promoters. Mol. Gen. Genet. **215,** 337–344.

Gilmartin, P.M., and Chua, N.-H. (1990). Spacing between GT-1 binding sites within a light-responsive element is critical for transcriptional activity. Plant Cell **2,** 447–455.

Giuliano, G., Pichersky, E., Malik, V.S., Timko, M.P., Scolnik, P.A., and Cashmore, A.R. (1988). An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. Proc. Natl. Acad. Sci. USA **85,** 7089–7093.

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., Yong, M.-H., Cuozzo, M., Kano-Murakami, Y., Silverstein, P., and Chua, N.-H. (1988). Binding site requirements for pea nuclear protein factor GT-1 correlate with sequences required for light-dependent transcriptional activation of the *rbcS-3A* gene. EMBO J. **7,** 4035–4044.

Grob, U., and Stuber, K. (1987). Discrimination of phytochrome dependent light-inducible from non-light-inducible plant genes. Prediction of a common light-responsive element (LRE) in phytochrome dependent light-inducible genes. Nucl. Acids Res. **15,** 9957–9972.

Jenkins, G. (1988). Photoregulation of gene expression in plants. Photochem. Photobiol. **48,** 821–832.

Kaulen, H., Schell, J., and Kreuzaler, F. (1986). Light-induced expression of the chimeric chalcone synthase-NPT II gene in transgenic tobacco. EMBO J. **5,** 1–8.

Kay, S.A., Keith, B., Shinozaki, K., Chye, M.-L., and Chua, N.-H. (1989). The rice phytochrome gene: Structure, autoregulated expression, and binding of GT-1 to a conserved site in the 5′ upstream region. Plant Cell **1,** 351–360.

Keith, B., and Chua, N.-H. (1986). Monocot and dicot pre-mRNAs are processed with different efficiencies in transgenic tobacco. EMBO J. **5,** 2419–2425.

Kendrick, R.E., and Kronenberg, G.M.H., eds (1986). Photomorphogenesis in Plants. (Dordrecht, The Netherlands: Martinus Nijhoff).

Kuhlemeier, C., Fluhr, R., Green, P., and Chua, N.-H. (1987a). Sequences in the pea *rbcS-3A* gene have homology to constitutive mammalian enhancers but function as negative regulatory elements. Genes Dev. **1,** 247–255.

Kuhlemeier, C., Green, P.J., and Chua, N.-H. (1987b). Regulation of gene expression in higher plants. Annu. Rev. Plant Physiol. **38,** 221–257.

Kuhlemeier, C., Cuozzo, M., Green, P., Goyvaerts, E., Ward, K., and Chua, N.-H. (1988a). Localization and conditional redundancy of regulatory elements in *rbcS-3A*, a pea gene encoding the small subunit of ribulose-bisphosphate carboxylase. Proc. Natl. Acad. Sci. USA **85,** 4662–4666.

Kuhlemeier, C., Fluhr, R., and Chua, N.-H. (1988b). Upstream sequences determine the difference in transcript abundance of pea *rbcS* genes. Mol. Gen. Genet. **212,** 405–411.

Kuhlemeier, C., Strittmatter, G., Ward, K., and Chua, N.-H. (1989). The pea *rbcS-3A* promoter mediates light responsiveness but not organ specificity. Plant Cell **1,** 471–478.

Lam, E., and Chua, N.-H. (1989). ASF-2: A factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in *Cab* promoters. Plant Cell **1,** 1147–1156.

Lam, E., and Chua, N.-H. (1990). GT-1 binding sites confer light-responsive expression in transgenic tobacco. Science, in press.

Lam, E., Benedyk, M., and Chua, N.-H. (1989a). Characterization

of phytochrome-regulated gene expression in a photoautotrophic cell suspension: Possible role for calmodulin. Mol. Cell. Biol. **9,** 4819–4823.

Lam, E., Benfey, P.N., Gilmartin, P.M., Fang, R.-X., and Chua, N.-H. (1989b). 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., Green, P.J., Wong, M., and Chua, N.-H. (1989c). Phytochrome activation of two nuclear genes requires cytoplasmic protein synthese. EMBO J. **8,** 2777–2783.

Lipphardt, S., Brettschneider, R., Kreuzaler, F., Schell, J., and Dangl, J.L. (1988). UV-inducible transient expression in parsley protoplasts identifies regulatory *cis*-elements of a chimeric *Antirrhinum majus* chalcone synthase gene. EMBO J. **7,** 4027–4033.

Lissemore, J.L., and Quail, P.H. (1988). Rapid transcriptional regulation by phytochrome of the genes for phytochrome and chlorophyll *a/b*-binding protein in *Avena sativa*. Mol. Cell. Biol. **8,** 4840–4850.

Manzara, T., and Gruissem, W. (1988). Organization and expression of the genes encoding ribulose-1,5-bisphosphate carboxylase in higher plants. Photosynth. Res. **16,** 117–139.

Mosinger, E., Batschauer, A., Schafer, E., and Apel, K. (1985). Phytochrome control of in vitro transcription of specific genes in isolated nuclei of barley (*Hordeum vulgare*). Eur. J. Biochem. **147,** 137–142.

Nagy, F., Kay, S.A., and Chua, N.-H. (1988). Gene regulation by phytochrome. Trends Genet. **4,** 37–42.

Prywes, R., Dutta, A., Cromlish, J.A., and Roeder, R.G. (1988). Phosphorylation of serum response factor, a factor that binds to the serum response element of the *c-FOS* enhancer. Proc. Natl. Acad. Sci. USA **85,** 7206–7210.

Schulze-Lefert, P., Becker-André, M., Schulz, W., Hahlbrock, K., and Dangl, J.L. (1989a). Functional architecture of the light-responsive chalcone synthase promoter from parsley. Plant Cell **1,** 707–714.

Schulze-Lefert, P., Dangl, J.L., Becker-André, M., Hahlbrock, K., and Schulz, W. (1989b). Inducible in vivo DNA footprints define sequences necessary for UV light activation of the parsley chalcone synthase gene. EMBO J. **8,** 651–656.

Silverthorne, J., and Tobin, E. (1984). Demonstration of transcriptional regulation of specific genes by phytochrome action. Proc. Natl. Acad. Sci. USA **81,** 1112–1116.

Silverthorne, J., and Tobin, E. (1987). Phytochrome regulation of nuclear gene expression. Bioessays **7,** 18–23.

Staiger, D., Kaulen, H., and Schell, J. (1989). A CACGTG motif of the *Antirrhinum majus* chalcone synthase promoter is recognized by an evolutionarily conserved nuclear protein. Proc. Natl. Acad. Sci. USA **86,** 6930–6934.

Stockhaus, J., Eckes, P., Rocha-Sosa, M., Schell, J., and Willmitzer, L. (1987). Analysis of *cis*-acting sequences involved in the leaf-specific expression of a potato gene in transgenic plants. Proc. Natl. Acad. Sci. USA **84,** 7943–7947.

Stockhaus, J., Schell, J., and Willmitzer, L. (1989). Identification of enhancer elements in the upstream region of the nuclear photosynthetic gene *ST-LS1*. Plant Cell **1,** 805–813.

Sullivan, T.D., Christensen, A.H., and Quail, P.H. (1989). Isolation and characterization of a maize chlorophyll *a/b* binding protein gene that produces high levels of mRNA in the dark. Mol. Gen. Genet. **215,** 431–440.

Tobin, E.M., and Silverthorne, J. (1985). Light regulation of gene expression in higher plants. Annu. Rev. Plant Physiol. **36,** 569–593.

Tsai, F.-Y., and Coruzzi, G.M. (1990). Dark-induced and organ-specific expression of two asparagine synthetase genes in *Pisum sativum*. EMBO J. **9,** 323–332.

Ueda, T., Pichersky, E., Malik, V.S., and Cashmore, A.R. (1989). Level of expression of the tomato *rbcS-3A* gene is modulated by a far upstream promoter element in a developmentally regulated manner. Plant Cell **1,** 217–227.

Weising, K., Schell, J., and Kahl, G. (1988). Foreign genes in plants: Transfer, structure, expression, and applications. Annu. Rev. Genet. **22,** 421–477.

Willmitzer, L. (1988). The use of transgenic plants to study plant gene expression. Trends Genet. **4,** 13–18.