The Plant Cell, Vol. 4, 831–838, July 1992 © 1992 American Society of Plant Physiologists

A Tobacco DNA Binding Protein That Interacts with a Light-Responsive Box II Element

Olga Perisic [a] and Eric Lam [a,b,1]

[a] AgBiotech Center, Waksman Institute, Rutgers State University, Piscataway, New Jersey 08855
[b] Department of Biochemistry and Microbiology, Waksman Institute, Box 759, Rutgers State University, Piscataway, New Jersey 08855

Ribulose-1,5-bisphosphate carboxylase/oxygenase plays a key role in photosynthetic carbon fixation in higher plants. The small subunit of this chloroplast enzyme (rbcS), encoded by a family of nuclear genes, is regulated at the transcriptional level by light. Promoter analyses have previously identified the box II sequence as a *cis* element critical for the light-regulated expression of *rbcS* genes. Nuclear factor GT-1 binds specifically to this element and is one of the plant nuclear factors that has been detected and studied in great detail. Here we describe the cloning and characterization of a tobacco cDNA encoding a protein, designated B2F (Box II Factor), with similar binding specificity and mobility in gel retardation assays as nuclear GT-1. Steady state levels of mRNA encoding B2F do not appear to be regulated by light; this is consistent with the previous observation that nuclear GT-1 activity is present in extracts from both light-grown and dark-adapted plants. Sequence comparison with another plant *trans*-acting factor, GT-2, which binds to a GT-like element in the rice phytochrome promoter, shows striking homology in three putative α-helices that may be involved in DNA binding.

INTRODUCTION

Photocontrol of gene expression via phytochrome, blue light, and UV-B photoreceptors plays a critical role in plant growth and development. The expression of a broad array of genes, many of which encode photosynthetic enzymes, is positively or negatively regulated by light, and in many cases, the regulation occurs largely at the level of initiation of transcription (Tobin and Silverthorne, 1985; Fluhr and Chua, 1986; Silverthorne and Tobin, 1987; Schulze-Lefert et al., 1989). The identification of protein–DNA interactions involved in transcriptional regulation of light-responsive genes is an essential step in the elucidation of the light-induced signal transduction pathway (Gilmartin et al., 1990).

The expression of pea *rbcS-3A*, one of the most highly transcribed members of a gene family encoding the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (rbcS), is induced at least 20-fold by light, and it has served as a model system for study of light-regulated transcription (Fluhr et al., 1986). Sequence analyses of promoter regions from this and other known *rbcS* genes have identified a number of conserved *cis*-regulatory elements, whose function has been extensively studied by promoter mutation analyses in transgenic plants (Fluhr et al., 1986; Kuhlemeier et al., 1987, 1988). 5′-Deletion of pea *rbcS-3A* that removes the sequences upstream of position −166 is sufficient for wild-type levels of light-responsive expression in mature leaves of transgenic tobacco (Kuhlemeier et al., 1988). The same −166 *rbcS-3A* promoter can confer light regulation to the adjacent (and divergent) −150 nopaline synthase (*nos*) promoter within pMON200 (Gilmartin and Chua, 1990). This −166 deletion contains two DNA elements, box II (5′-GTGTGGTTAATATG-3′; −151 to −138) and box III (5′-ATCATTTTCACT-3′; −125 to −114), both well conserved among all pea *rbcS* genes examined. Furthermore, a 58-bp element spanning the −169 to −112 region (containing box II and box III) retains both light-regulated and organ-specific transcriptional activity when fused to a −50 *rbcS* promoter (Cuozzo-Davis et al., 1990). Both of these elements are bound in vitro by a nuclear factor, GT-1 (Green et al., 1987, 1988). Deletion to position −149, which removes two bases from the 5′ end of box II, abolishes the activity of this promoter in vivo and the binding of GT-1 in vitro (Kuhlemeier et al., 1987; Green et al., 1988). In addition, substitution mutation in either box II or box III within the −170 *rbcS-3A* promoter results in the complete loss of transcriptional activity in vivo and of binding to GT-1 in vitro (Kuhlemeier et al., 1988). These data strongly correlate promoter activity in vivo and GT-1 binding to box II.

A critical role for GT-1 binding sites in light-responsive expression of *rbcS* genes is further demonstrated by the ability of four direct repeats of the box II element to confer light responsiveness to a −90 truncation of the cauliflower mosaic virus 35S promoter, which contains two binding sites for the tobacco nuclear factor ASF-1 (Lam et al., 1989; Lam and Chua, 1990). However, the same tetramer is inactive when fused to

[1] To whom correspondence should be addressed.

a −46 truncated cauliflower mosaic virus 35S promoter lacking the as-1 sites or to a −50 *rbcS-3A* promoter (E. Lam and N.-H. Chua, unpublished data; Cuozzo-Davis et al., 1990). These data indicate that GT-1 sites, although necessary for light-regulated expression, are functionally dependent on additional elements.

To study the molecular nature of the productive interaction between GT-1 factor and box II elements that leads to light-responsive transcription, it is necessary to isolate and characterize the properties of nuclear proteins that constitute GT-1 activity. Toward that goal, we have now isolated a tobacco cDNA clone encoding a protein that binds specifically to the box II element.

RESULTS

Isolation of a cDNA Encoding a GT Element Binding Protein

We used a tetramer of the box II element from the pea *rbcS-3A* promoter to screen a primary λgt11 cDNA expression library prepared from mature leaves of greenhouse-grown tobacco plants. Approximately 2 × 10^6 phage plaques were screened, of which a single plaque was identified as expressing a protein that interacts specifically with a wild-type box II, but not with a mutant box II tetramer, as shown in Figure 1. The DNA binding activity of the protein was not isopropyl β-thiogalactoside inducible, indicating that the recombinant protein might not be a translational fusion with β-D-galactosidase encoded by the λgt11 vector. Indeed, we found that the longest open reading frame (ORF) of the cDNA insert is in the opposite orientation from the *lacZ* gene of the λgt11 vector, possibly being transcribed from the distant phage *Pr'* promoter. The cDNA insert was subcloned as a NotI fragment into a plasmid expression vector. (NotI/EcoRI adapters were used for the construction of the cDNA library.)

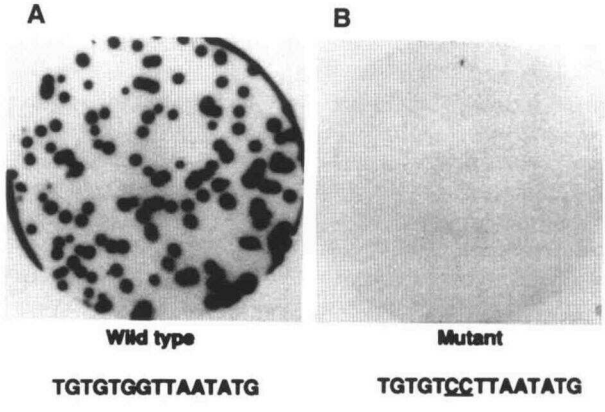

Figure 1. In Situ Detection of Box II–Specific Binding Activity.

(A) Purified recombinant λB2F phage incubated with tetramers of the box II element.

(B) Purified recombinant λB2F phage incubated with its mutant derivative containing a 2-bp mutation in the box II element. This box II mutant is known to inactivate nuclear GT-1 factor binding and function.

Sequences of the wild-type and mutant box II elements are given below, with the mutated bases underlined in **(B)**.

Sequence analyses of the subclone revealed two overlapping ORFs in a 1505-bp insert. Subsequent analyses of additional subclones revealed frequent and extensive rearrangement of the cDNA during subcloning, prompting us to sequence the λ DNA from the recombinant phage directly, without any subcloning step. Surprisingly, we found one additional base pair in the recombinant λ DNA compared to the sequence of the subclone (G/C pair at position 367). That difference gave rise to one long ORF of 392 amino acids (starting with the first amino acid) with a calculated molecular mass of 44,768 D, as shown in Figure 2. The second ATG, encoding Met-30, might be used as an initiator codon, because it is preceded by a good match to a bacterial ribosomal binding site (GAGG; Shine and Dalgarno, 1974).

The coding capacity of the ORF was tested by gel blot analyses of protein partially purified from lysogen extracts of the recombinant phage. As predicted from the sequence information, the lysogen extract contains a protein with a mass of approximately 43 kD that binds to the box II element specifically, as shown in Figure 3. Because this is not a fusion protein under the control of the isopropyl β-thiogalactoside-inducible promoter, the amount of the protein in the λ lysogen extract is low. This protein was not observed in lysogen extracts prepared from a control λ phage that contains an unrelated cDNA insert (data not shown). An N-terminal deletion of the protein that removed the first 122 amino acids was overexpressed in bacteria and tested for DNA binding activity. Using the same gel blot analysis, this truncated protein showed no significant binding activity to box II tetramers (data not shown), indicating that the N-terminal sequence is necessary for DNA binding of the recombinant protein. We designated this cDNA clone B2F (Box II Factor).

Recombinant B2F and Nuclear GT-1 Factor Have Very Similar Sequence Requirements for Interaction with Box II

Nuclear extracts from tobacco leaves and recombinant protein purified from lysogen extracts gave rise to a complex with similar mobility in gel retardation assays using box II tetramers as a DNA probe, as shown in Figure 4A. The relatedness of recombinant B2F and nuclear GT-1 was further tested by examining the specificity of binding to wild-type box II and to 2-bp mutations across box II. Mutations in the central six residues GGTTAA of box II decreased the binding of recombinant B2F dramatically (Figure 4B). The same core of six residues was found earlier to be critical for in vitro binding of the nuclear GT-1 factor (Green et al., 1988). These results suggest that B2F is likely a component of the nuclear GT-1 activity that interacts with box II.

```
1    CCAAAGTCCAAGCAAGTATTACCTATTTCTATCTTACATGTCTGATAAACTTCCAACATC   60
1     Q  S  P  S  K  Y  Y  L  F  L  S  Y  M  S  D  K  L  P  T  S    20

61   TATCAACTTGTACGAGGAGCAAGAATCAATGGATCAACACAGTAACCACATGATCATCGA  120
21    I  N  L  Y  E  E  Q  E  S  M  D  Q  H  S  N  H  M  I  I  E    40

121  AGTGGCCACCCCTAATGCTCATCTTCAACAACCCCAGCAAATTCTCCTTCCCGGAATCAG  180
41    V  A  T  P  N  A  H  L  Q  Q  P  Q  Q  I  L  L  P  G  I  S    60

181  TGGTGGTGACACCACCAGCAGCGGTGGCGAGGACAACAACAACAACGTAAAATTAGCACC  240
61    G  G  D  T  T  S  S  G  G  E  D  N  N  N  N  V  K  L  A  P    80

241  CAAGAAGCGAGCAGAAACATGGGTTCAAGAAGAGACACGAGCGCTCATTAGCCTCCGCAG  300
81    K  K  R  A  E  T  W  V  Q  E  E  T  R  A  L  I  S  L  R  R   100

301  AGAACTCGACTCACTCTTCAACACTTCAAAATCGAACAAGCACTTGTGGGACCAGATCTC  360
101   E  L  D  S  L  F  N  T  S  K  S  N  K  H  L  W  D  Q  I  S   120

361  CTTGAAGATGAGGGAAAAGGGGTTTGATAGGTCCCCTACCATGTGTACTGATAAATGGAG  420
121   L  K  M  R  E  K  G  F  D  R  S  P  T  M  C  T  D  K  W  R   140

421  GAACTTGTTAAAGGAGTTCAAAAAGGCTAAACACAATCAAGAACCAAATGGGTCTGCTAA  480
141   N  L  L  K  E  F  K  K  A  K  H  N  Q  E  P  N  G  S  A  K   160

481  GATGTCTTATCATAAGGAGATTGAGGAGATTCTCAAGGCCAGAAACAAAAATTATAAGAA  540
161   M  S  Y  H  K  E  I  E  E  I  L  K  A  R  N  K  N  Y  K  N   180

541  CCCTACGCTTAAAGTCGATACCTTTATGCAGTTTTCTCAGAAGGGTCTTGACGATACCAG  600
181   P  T  L  K  V  D  T  F  M  Q  F  S  Q  K  G  L  D  D  T  S   200

601  TATAACATTTGGACCTGTCGAAGAGAATGGGAGGCCAACTCTTAACTTAGAACGCCAGCT  660
201   I  T  F  G  P  V  E  E  N  G  R  P  T  L  N  L  E  R  Q  L   220

661  GGATCATGATGGGCATCCTCTTGCTATAACAGCTGCAGATGCGGTCACTGCAAGTGGATC  720
221   D  H  D  G  H  P  L  A  I  T  A  A  D  A  V  T  A  S  G  S   240

721  ACCTTGGAATTGGAGAGAGACGCCTGGAAATGGTGAGCAAAGTAACTCAGCTGAAGGTAG  780
241   P  W  N  W  R  E  T  P  G  N  G  E  Q  S  N  S  A  E  G  R   260

781  GGTCATTTCAGTCAAGTGGGGTGATTACACAAAAAGAATTGGTATCGATGGGACTGCAGA  840
261   V  I  S  V  K  W  G  D  Y  T  K  R  I  G  I  D  G  T  A  D   280

841  TGCCATCAAGCAAGCCATTAAATCTGCTTTTAGGTTGAGGACAGAGCGTGCATTTTGGTT  900
281   A  I  K  Q  A  I  K  S  A  F  R  L  R  T  E  R  A  F  W  L   300

901  AGAAGATGAACAGAATATTGTTCGAGCACTTGACAGGGATATGCCACTAGGGAGCTACAG  960
301   E  D  E  Q  N  I  V  R  A  L  D  R  D  M  P  L  G  S  Y  S   320

961  CTTGCATGTTGATGAAGGTCTGACAATAAAAGTTTGCATGTACGAGGAGGCAGATCACTC 1020
321   L  H  V  D  E  G  L  T  I  K  V  C  M  Y  E  E  A  D  H  S   340

1021 CTCAGTGCACACCGAGGATAAAACTTTCTACACTGCGAATGATTTTCGTGACTTCTTGTC 1080
341   S  V  H  T  E  D  K  T  F  Y  T  A  N  D  F  R  D  F  L  S   360

1081 CCACAGAGGCTGGACATGTTTAAGAGAGTACAATGGGTATCGGCATGTAGATATGTTGGA 1140
361   H  R  G  W  T  C  L  R  E  Y  N  G  Y  R  H  V  D  M  L  D   380

1141 TGAGCTTTGCCCTGGTGCAGTCTACCGAGGTGTCAATTGACGATAATAGGGGATACCCTG 1200
381   E  L  C  P  G  A  V  Y  R  G  V  N  *                       392

1201 TCCATCCACGGCTCGTTGGGGTGACTGCATTTGCCATCTTTGTAAGTAACATGATGTCGG 1260

1261 TGCCATCAGAGTGTGTTGATAATGCTGTTCTGGAAGTGTAAATGGTGCAGTGGTAACCTG 1320

1321 ATATAGTGATATAGTACTGAACAATTTAAATGTACAAGTCAATGAACATGAAACCTTGTT 1380

1381 CATATTAGTGTTCCGGATGGTGAAATCTTGTTCTGGAATTCATTTTCATTACATCTCCCA 1440

1441 TTGGCCTATGCCCGGAGACTAGAGTTATCTAGTATTTGTGCTGATGGGACAACAAAAAAA 1500

1501 AAAAAA                                                        1506
```

Figure 2. Nucleotide and Deduced Amino Acid Sequence of the Recombinant B2F Clone.

The nucleotide sequence was determined by direct sequencing of single-stranded DNA produced by an asymmetric polymerase chain reaction of purified λB2F DNA and by double-stranded sequencing of a B2F insert subcloned in the pBluescript II SK– vector. The subclone was obtained by digestion of λB2F DNA with the NotI enzyme, which gives the entire cDNA insert, followed by ligation into the NotI site of the pBluescript vector. The sequences of the λ DNA insert and of the subclone were identical except at position 367 (marked with a box): the subclone had a single base pair deletion (G/C missing), which interrupted the ORF. The first amino acid of the longest ORF is arbitrarily numbered 1. The four predicted helices are underlined (Garnier et al., 1978). All the sequence analyses were carried out using GCG software for the VAX (Devereux et al., 1984).

B2F mRNA Is Constitutively Expressed in Leaves and Roots, in Light and Dark

The genomic DNA gel blot pattern suggests that in *Nicotiana tabacum* there are at least two genes for B2F, as shown in Figure 5A. This is in line with the polyploid origin of *N. tabacum* (Gray et al., 1974). It is likely that each of the progenitors, *N. sylvestris* and *N. tomentosiformis*, contributed at least one copy of the gene to the *N. tabacum* nuclear genome. B2F mRNA was detected by RNA gel blots in all the tissues examined: roots, stems, and young and mature leaves from tobacco plants (Figure 5B). In addition, similar expression of B2F mRNA was observed in leaves from light-grown or dark-adapted plants, as well as from plants transferred from dark back to light (Figure 5B). The length of the transcript is approximately 1.65 kb. Because the B2F cDNA clone (1506 bp) has a poly(A)$^+$ tail, it is likely that only about 100 to 200 bases are missing from the 5′ end of the full-length clone.

A Putative Three-Helical Region of B2F Has Homology to GT-2

A computer search in the Swiss-Prot protein sequence data bank (release 18) revealed that B2F has no structural motifs characteristic of known transcription factors and no striking homology to any protein. Secondary structure prediction for B2F reveals several regions with the potential to form α-helical

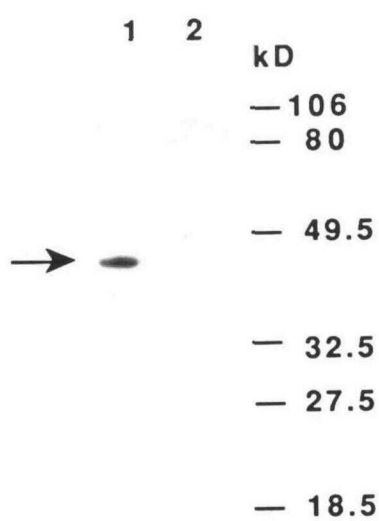

Figure 3. Gel Blot Analysis of λB2F Lysogen Extract.

The recombinant protein partially purified from λB2F lysogen extracts (20 μg/lane) was separated on a 10% polyacrylamide gel by SDS-PAGE. The proteins were transferred to an Immobilon-P membrane and incubated with binding buffer (EB buffer with 0.25% blocking reagent and 5 μg/mL polyd[I-C]) and 5 ng/mL of a ^{32}P–end-labeled tetramer of the wild-type (lane 1) or mutant (GG→CC, lane 2) box II element. The arrow marks the position of the 43-kD protein encoded by the B2F cDNA.

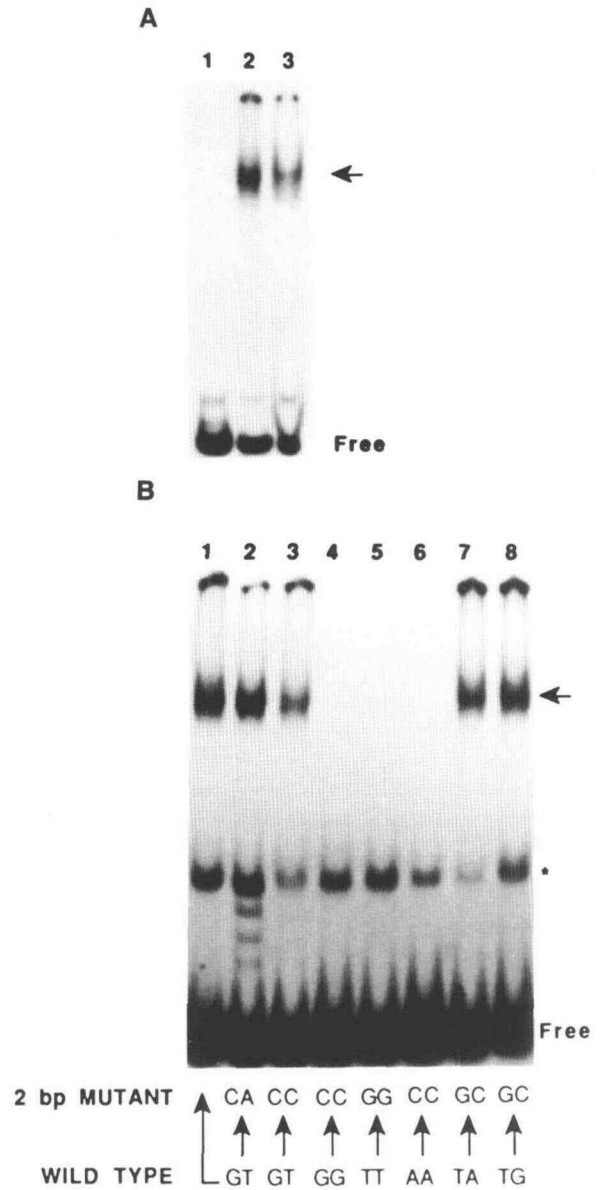

Figure 4. Binding Specificity of λB2F-Encoded Protein.

(A) Gel retardation assays with B2F recombinant protein and tobacco leaf nuclear extract. ^{32}P-end-labeled oligonucleotides containing a tetramer of the box II element were incubated with buffer (lane 1), tobacco nuclear extract (4 μg, lane 2), or recombinant B2F partially purified from lysogen extracts (0.5 μg, lane 3).

(B) Crude λB2F lysogen extract (3 μg) was incubated with a tetramer of wild-type box II (lane 1) or a tetramer containing a 2-bp substitution in each copy of box II (lanes 2 to 8).

Autoradiograms of the gels are shown. The arrows indicate the positions of the specific complexes, and the asterisk in **(B)** marks the position of the nonspecific complex.

structures, which are separated by regions likely to form turns or loops (Garnier et al., 1978). There are four putative helices between amino acids 75 and 174 (residues 75 to 103, 115 to 124, 138 to 150, and 159 to 174). The N-terminal deletion of 122 amino acids removes two of these helices, which might be the reason for the observed lack of DNA binding activity in the truncated protein.

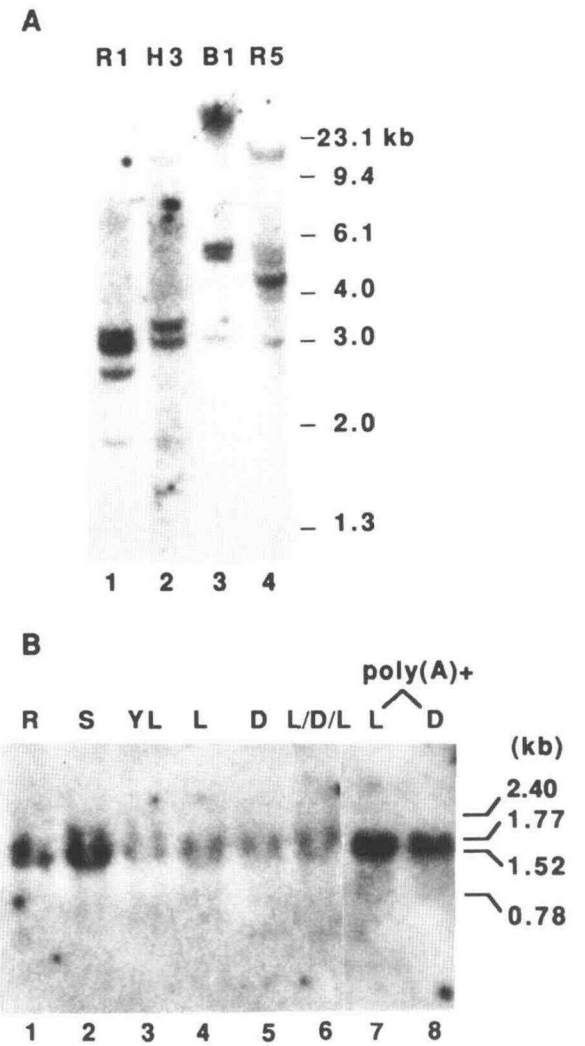

Figure 5. Genomic DNA and RNA Gel Blot Analyses of the B2F Gene.

The blots were probed with a 1.4-kb fragment labeled by random priming. The positions of the molecular length markers are indicated at right.

(A) Tobacco DNA (20 μg) digested with EcoRI (R1, lane 1), HindIII (H3, lane 2), BamHI (B1, lane 3), or EcoRV (R5, lane 4).

(B) Total cellular RNA (30 μg) from tobacco roots (R, lane 1); stems (S, lane 2); young leaves (YL, lane 3); mature leaves from plants grown in the light (L, lane 4), dark adapted for 2 days (D, lane 5), and shifted back to light (L/D/L, lane 6); poly(A)$^+$ RNA from mature leaves of a light-grown plant (1 μg, L, lane 7) or from a plant dark adapted for 2 days (0.5 μg, D, lane 8). Equal loadings of RNA were confirmed by ethidium bromide staining of the gel before the transfer.

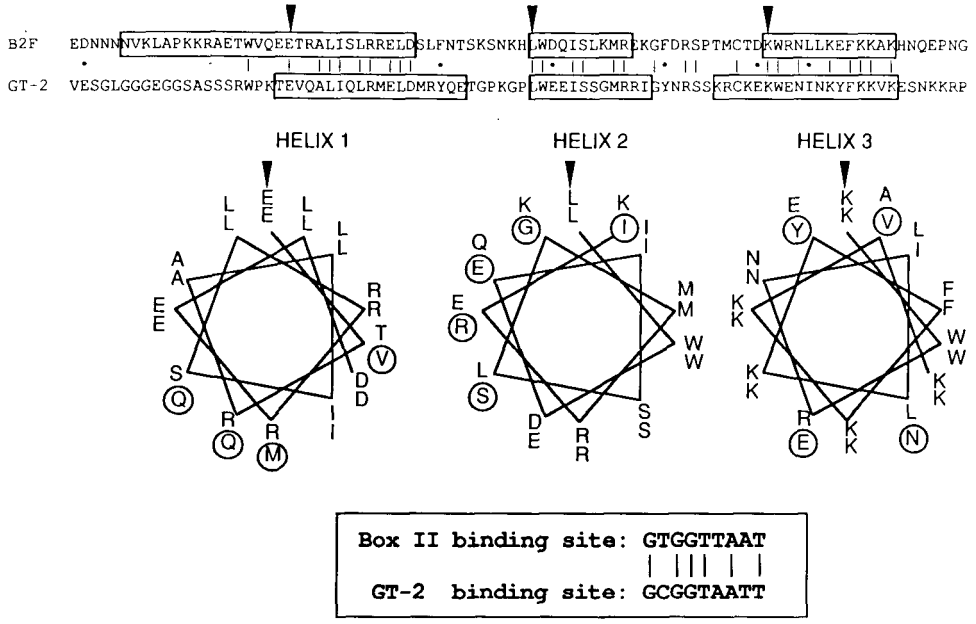

Figure 6. Sequence Comparison of B2F and GT-2.

The conserved region of B2F and GT-2 (residues 70 to 157 in both proteins) is displayed (Dehesh et al., 1990). The putative α-helices are boxed. Amino acid identity is marked with a vertical bar, and a conservative change is indicated with a dot. The amino acid residues of the three helices are also shown in a helical wheel representation, with B2F residues on the top and GT-2 on the bottom. Arrowheads indicate the first residue in each helix displayed on a wheel. The amino acid residues in GT-2 that are different from B2F are circled. Note that most of the different residues in helices 1 and 2 are located on one side of the helix. The binding site sequences for both proteins are also shown.

By visual inspection, we found that a region of B2F limited to 60 amino acids has significant homology to another recently isolated plant *trans*-acting factor, GT-2 (Dehesh et al., 1990) (44% identity, 55% similarity over a 60-amino acid region). GT-2 is an 81-kD protein encoded by a 3.8-kb mRNA, and its expression level shows a decline in white light (Dehesh et al., 1990; Quail, 1991). GT-2 protein binds specifically to a GCGGTAATT sequence in the rice phytochrome promoter, but much less to a related, GT-1–specific motif (GTGGTTAAT) present both in phytochrome and *rbcS* promoters (Dehesh et al., 1990). Similarly, based on competition experiments, B2F protein binds much less efficiently to the GT-2 element than to the box II element (O. Perisic and E. Lam, unpublished data). Conserved residues between the two proteins are found to be restricted to a region that, based on secondary structure prediction, can form three α-helices, as shown in Figure 6, in both B2F (first three helices underlined in Figure 2) and in GT-2. Within this region, the majority of the amino acids that differ between the two proteins, especially in helices 1 and 2, are clustered on one face of the putative helices. It is possible that some of these amino acids may be involved in selective binding of the proteins to related GT motifs. Unlike GT-2, which contains an acidic domain (Dehesh et al., 1990), B2F lacks any region that has the characteristics of known activation domains of transcription factors. This suggests that the two factors may function by different modes of action. The limited homology between B2F and GT-2 supports the previous suggestion that related *trans*-acting factors that bind to similar GT motifs may be involved in the differential regulation of light-responsive gene expression (Dehesh et al., 1990).

DISCUSSION

Mutational and gain-of-function experiments have shown previously that box II acts as a positive element in light-dependent transcriptional activation of cognate and heterologous promoters in transgenic tobacco (Gilmartin et al., 1990). Box II is a binding site for nuclear factor GT-1, and a tight correlation between GT-1 binding to box II in vitro and the function of box II as a positive element in vivo has been demonstrated (Green et al., 1987, 1988). It is therefore of great interest to identify individual proteins that constitute GT-1 nuclear activity and then study their mechanism of action. Toward that goal, we have isolated a clone encoding a protein, B2F, that binds to a box II element with the same sequence requirement as the nuclear GT-1 factor. This will enable us to determine its relationship to nuclear GT-1 and its potential involvement in light-responsive transcription.

Consistent with the previous observation that similar levels of GT-1 binding activity are present in nuclear extracts from light-grown and dark-adapted plants, steady state levels of B2F mRNA do not appear to be modulated by light. There are

various models for GT-1-mediated light activation of transcription. It was proposed that GT-1, while constitutively bound to box II in light and dark, might be activated in light by covalent modification (e.g., phosphorylation) or interaction with another activating factor (Green et al., 1987; Lam and Chua, 1990). There are many examples of such mechanisms. For example, the yeast heat shock factor (HSF) is constitutively bound to DNA, but its ability to stimulate transcription coincides with its phosphorylation during heat shock (Sorger and Pelham, 1988). A plant nuclear protein AT-1 binds DNA only in the nonphosphorylated form and completely loses its binding activity upon phosphorylation by a plant nuclear kinase (Datta and Cashmore, 1989). The opposite is true for a bean nuclear factor silencer binding factor (SBF-1), which loses its DNA binding activity after treatment with alkaline phosphatase (Harrison et al., 1991). Alternatively, GT-1 activity could be negatively regulated in the dark. Indeed, *rbcS* and several other normally light-regulated genes are constitutively expressed in etiolated *det1* (Chory and Peto, 1990) and *cop1* seedlings (Deng et al., 1991) and in mature, dark-adapted *cop1* plants (Deng et al., 1991). *det1* (de-etiolated) and *cop1* (constitutively photomorphogenic) are mutants in Arabidopsis that show in dark-grown plants many of the morphological and molecular characteristics of light-grown plants (Chory and Peto, 1990; Chory, 1991; Deng et al., 1991). In addition to deregulated expression in the dark, *det1* plants also show a tissue aberrant accumulation of *rbcS* and *cab* mRNA in roots, where they are normally not detectable in wild-type Arabidopsis (Chory and Peto, 1990).

These data are consistent with the presence of negative regulators that can suppress transcription from light-responsive promoters in the dark and in appropriate tissues (e.g., roots). We have shown that B2F mRNA is constitutively expressed not only in dark and light, but also in roots, stems, and young leaves, where the expression of *rbcS* genes is lower or undetectable (Fluhr et al., 1986; Kuhlemeier et al., 1988). The negative factors may interact, directly or indirectly, with GT-1 to prevent its function and *rbcS* expression. In *det1* and *cop1* mutants, the absence of these regulators could thus allow GT-1 to function constitutively. At present, we cannot rule out the possibility that related genes encoding GT-1 might be differentially expressed in various tissues and light/dark regimes. Alternatively, the GT-1 factor may be synthesized or activated only in appropriate tissues and under certain light conditions. The isolation of the B2F clone may enable us to generate the tools necessary to elucidate the molecular mechanisms of the switch that controls light-responsive gene transcription.

METHODS

cDNA Cloning and Sequencing

Poly(A)$^+$ RNA isolated from mature, green leaves of greenhouse-grown *Nicotiana tabacum* cv SR1 plants was used for the construction of an oligo(dT)-primed cDNA library in the λgt11 vector, using the Librarian XI kit (Invitrogen, San Diego, CA). Upon completion of double-stranded cDNA synthesis, NotI/EcoRI adapters were added, and cDNAs were ligated to λgt11 that was cut with EcoRI. The primary library of 2×10^6 phages was screened in Y1090 cells using ^{32}P–end-labeled 84-bp tetramers of the box II binding site (Green et al., 1988) as a probe, essentially as described in Ausubel et al. (1987). Filters were blocked for 1 hr in EB buffer (20 mM Hepes, pH 7.0, 40 mM KCl, 0.5 mM EDTA, 10% glycerol, 1 mM DTT) supplemented with 5% nonfat dry milk, washed twice for 10 min in EB buffer, and then stacked and incubated for 2 hr in EB buffer with labeled DNA probe (2×10^6 counts per minute per mL) and 5 μg/mL denatured salmon sperm DNA. Filters were washed twice for 10 min with EB buffer.

The insert of a positive clone was subcloned as a NotI fragment into the pBluescript II SK– vector and sequenced by the dideoxy sequencing method using a Sequenase kit (U.S. Biochemical Corp.). For direct sequencing, without a subcloning step, of λ DNA containing the B2F insert, the single-stranded insert DNA was first produced by an asymmetric polymerase chain reaction using a 50:1 ratio of two primers (either λ or insert-specific primers) and 10 ng of purified λB2F DNA. After the polymerase chain reaction, free nucleotides and primers were removed using a GeneClean II kit (Bio-101, La Jolla, CA), and the DNA was sequenced as described above.

An internal BglII site (at position 354) of a B2F subclone and the BamHI site in the pBluescript II SK– vector were used to generate a 5′ deletion fragment that was cloned into the BamHI site of the pET3b expression vector (Studier et al., 1990). The construct was transformed into BL21(DE3) bacteria, which were then used for overproduction of the truncated B2F protein (more than 20% of the total bacterial protein).

Gel Blot Analysis

λB2F recombinant lysogens were generated in Y1089 cells, and lysogen extracts were prepared as described by Ausubel et al. (1987). To enrich B2F protein, the crude lysogen extracts were partially purified by heparin-agarose and DNA-affinity chromatography. Crude extract (90 mg) obtained from a 200-mL culture was dialyzed against 0.1 M NaCl in buffer A (20 mM Hepes, pH 7.5, 1 mM EDTA, 10% glycerol, 1 mM DTT) and applied to a 20-mL heparin-agarose column equilibrated in the same buffer. The column was washed with the same buffer and eluted with a gradient of NaCl from 0.1 to 1.5 M in buffer A. Fractions containing B2F (as determined by gel retardation assays using end-labeled tetramers of box II) were pooled and applied to a specific box II DNA–Sepharose 4B column that was made by coupling the concatamers of the box II element to CNBr-activated Sepharose 4B. The column was washed and eluted with a 0.1 to 1.5 M gradient of NaCl in buffer A. Proteins were separated by electrophoresis on 10% SDS-polyacrylamide gels. For gel blot analysis, gels were incubated twice for 15 min in 20% glycerol, 50 mM Tris, pH 7.5, and once in the transfer buffer (Miskimins et al., 1985). The proteins were transferred to a Immobilon-P membrane (Millipore) by electroblotting, and the membrane was blocked in EB buffer supplemented with 2% blocking reagent (Boehringer Mannheim) for 1 hr at 4°C. After overnight incubation with binding buffer (EB buffer with 0.25% blocking reagent, 5 μg/mL polyd(I-C), and 5 ng/mL of ^{32}P–end-labeled tetramer probes), membranes were washed three times for 5 min in EB buffer before autoradiography.

Gel Retardation Assays

Reactions (10 μL), containing lysogen extract or tobacco nuclear extract, were incubated in EB buffer in the presence of 4 μg of polyd(I-C)

and 0.2 to 0.5 ng of ^{32}P-end-labeled DNA probe for 30 min at room temperature. The reactions were analyzed by electrophoresis on 3.2% polyacrylamide gels in 0.25 × Tris-borate-EDTA buffer.

Genomic DNA Gel Blot Analysis

DNA was prepared as described by Junghans and Metzlaff (1990). After electrophoresis on a 0.7% agarose gel, samples were transferred to a Zeta-Probe GT membrane (Bio-Rad) and hybridized at 42°C in 50% formamide, 0.25 M NaCl, 7% SDS, and 0.12 M Na_2HPO_4, pH 7.2. The blot was washed first at low stringency and then in a solution containing 0.1 × SSC (1 × SSC is 0.15 M NaCl, 0.015 M sodium citrate), 0.1% SDS at 65°C.

RNA Analysis

Total cellular RNA from various plant tissues was prepared as described previously (Nagy et al., 1987), except that 2 mM aurin tricarboxylic acid and 14 mM β-mercaptoethanol were used in the lysis buffer, and following LiCl precipitation RNA was resuspended in water, extracted once with phenol, and precipitated with 0.3 M sodium acetate and ethanol. Poly(A)$^+$ RNA was prepared from 0.25 g of tissue using a Micro-FastTrack kit (Invitrogen). Samples were separated on a 1% agarose-formaldehyde gel, transferred to Duralose-UV membrane (Stratagene), and hybridized with 25 ng of labeled probe in QuickHyb solution (Stratagene). The blot was washed first with 2 × SSC, 0.1% SDS at room temperature and then in 0.2 × SSC, 0.1% SDS at 60°C for 45 min.

ACKNOWLEDGMENTS

We thank Peter Day, Michael Lawton, Danny Reinberg, and Xiang-jun Liu for critical reading of the manuscript. This work was supported by a grant from the U.S. Department of Agriculture Competitive Grants Program. We also thank Dr. Nam-Hai Chua (Rockefeller University, New York) for the box II wild-type and mutant tetramers used for gel retardation assays.

Received March 11, 1992; accepted May 26, 1992.

REFERENCES

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds) (1987). Current Protocols in Molecular Biology (New York: Wiley).

Chory, J. (1991). Light signals in leaf and chloroplast development: Photoreceptors and downstream responses in search of a transduction pathway. New Biologist **3,** 538–548.

Chory, J., and Peto, C.A. (1990). Mutations in the *DET1* gene affect cell-type-specific expression of light-regulated genes and chloroplast development in *Arabidopsis.* Proc. Natl. Acad. Sci. USA **87,** 8776–8780.

Cuozzo-Davis, M., Yong, M.-H., Gilmartin, P.M., Goyvaerts, E., Kuhlemeier, C., Sarokin, L., and Chua, N.-H. (1990). Minimal sequence requirement for the regulated expression of *rbcS-3A* from *Pisum sativum* in transgenic tobacco plants. Photochem. Photobiol. **52,** 43–50.

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.

Dehesh, K., Bruce, W.B., and Quail, P.H. (1990). A *trans*-acting factor that binds to a GT-motif in a phytochrome gene promoter. Science **250,** 1397–1399.

Deng, X.-W., Caspar, T., and Quail, P.H. (1991). *Cop* 1: A regulatory locus involved in light-controlled development and gene expression in *Arabidopsis.* Genes Dev. **5,** 1172–1182.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucl. Acids Res. **12,** 387–395.

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., Kuhlemeier, C., Nagy, F., and Chua, N.-H. (1986). Organ-specific and light-induced expression of plant genes. Science **232,** 1106–1112.

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.

Garnier, J., Osguthorpe, D.J., and Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120,** 97–120.

Gilmartin, P.M., and Chua, N.-H. (1990). Localization of a phytochrome-responsive element within the upstream region of pea *rbcS-3A.* Mol. Cell. Biol. **10,** 5565–5568.

Gilmartin, P.M., Sarokin, L., Memelink, J., and Chua, N.-H. (1990). Molecular light switches for plant genes. Plant Cell **2,** 369–378.

Gray, J.C., Kung, S.D., Wildman, S.G., and Sheen, S.J. (1974). Origin of *Nicotiana tabacum* L. detected by polypeptide composition of Fraction I protein. Nature **252,** 226–227.

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.

Harrison, M.J., Lawton, M.A., Lamb, C.J., and Dixon, R.A. (1991). Characterization of a nuclear protein that binds to three elements within the silencer region of a bean chalcone synthase gene promoter. Proc. Natl. Acad. Sci. USA **88,** 2515–2519.

Junghans, H., and Metzlaff, M. (1990). A simple and rapid method for the preparation of total plant DNA. BioTechniques **8,** 176.

Kuhlemeier, C., Fluhr, R., Green, P.J., and Chua, N.-H. (1987). 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., Cuozzo, M., Green, P.J., Goyvaerts, E., Ward, K., and Chua, N.-H. (1988). 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.

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

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

Miskimins, W.K., Roberts, M.P., McClelland, A., and Ruddle, F.H. (1985). Use of a protein-blotting procedure and a specific DNA probe to identify nuclear proteins that recognize the promoter region of the transferrin receptor gene. Proc. Natl. Acad. Sci. USA **82,** 6741–6744.

Nagy, F., Boutry, M., Hsu, M.-Y., Wong, M., and Chua, N.-H. (1987). The 5′-proximal region of the wheat Cab-1 gene contains a 268-bp enhancer-like sequence for phytochrome response. EMBO J. **6,** 2537–2542.

Quail, P.H. (1991). Phytochrome: A light-activated molecular switch that regulates plant gene expression. Annu. Rev. Genet. **25,** 389–409.

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

Shine, J., and Dalgarno, L. (1974). The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71,** 1342–1346.

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

Sorger, P.K., and Pelham, H.R.B. (1988). Yeast heat shock factor is an essential DNA-binding protein that exhibits temperature-dependent phosphorylation. Cell **54,** 855–864.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990). Use of T7 polymerase to direct expression of cloned genes. Methods Enzymol. **185,** 60–89.

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