Homodimeric and Heterodimeric Leucine Zipper Proteins and Nuclear Factors from Parsley Recognize Diverse Promoter Elements with ACGT Cores

Gregory A. Armstrong, Bernd Weisshaar, and Klaus Hahlbrock
Max-Planck-Institut für Züchtungsforschung, Abteilung Biochemie, Carl-von-Linné Weg 10, D-5000 Köln 30, Germany

Four short nucleotide sequences (boxes I to IV) contribute to the light responsiveness of the parsley chalcone synthase promoter. The sequence-related boxes II and III resemble several plant, viral, and bacterial promoter elements that share ACGT core sequences and are associated with diversely regulated genes. We have analyzed the binding characteristics and protein–protein interactions of factors from nuclear extracts and of three putative leucine zipper (bZIP) transcription factors potentially involved in the regulation of this promoter. These common plant regulatory factors (CPRFs) bind specifically to boxes II and III as well as other ACGT-containing promoter elements (hex1, Em1a, and as-1), though with markedly different affinities. Intact bZIP domains are crucial for CPRF binding to DNA. Distinct ensembles of nuclear factors bind to boxes II and III, despite their sequence similarity. The parsley CPRFs bind to DNA as dimers, selectively form heterodimeric DNA binding complexes, and interact with nuclear proteins.

INTRODUCTION

Many complex biological processes in plants rely on light as an environmental trigger. Light-mediated transcriptional regulation of nuclear gene expression in plants results from the interplay of diverse families of nuclear factors with distinct classes of binding sites (reviewed in Gilmartin et al., 1990; Schindler and Cashmore, 1990; Weising and Kahl, 1991). Particular combinations of these factors confer specificity upon regulated plant promoters, a situation analogous to that more completely described in yeast and mammalian systems (reviewed in Johnson and McKnight, 1989; Mitchell and Tjian, 1989). Many eukaryotic transcription factors bind to DNA as dimers and can selectively heterodimerize, contributing to regulation by modifying binding site specificities, altering dimer stability, and permitting a new set of specific protein–protein interactions to occur at the promoter (reviewed in Lamb and McKnight, 1991).

We have focused our attention on the single-copy chalcone synthase (CHS) gene from parsley (Herrmann et al., 1988) as a model for the transcriptional regulation of plant nuclear genes by light. CHS catalyzes the first biochemical reaction unique to flavonoid biosynthesis (reviewed in Hahlbrock and Scheel, 1989). UV-absorbing flavonoids accumulate in parsley cell suspensions (Kreuzaler and Hahlbrock, 1973) and in epidermal cells of intact plants (Schmelzer et al., 1988) in response to irradiation with UV-containing white light. The massive light-induced elevation in the transcription of the CHS gene precedes sequential increases in the levels of CHS mRNA and enzyme in plants (Schmelzer et al., 1988) and cell cultures (Chappell and Hahlbrock, 1984), contributing ultimately to flavonoid accumulation.

The availability of a parsley protoplast system that faithfully mimics cellular responses to external stimuli such as light and fungal elicitor (Dangl et al., 1987) provided the means to functionally dissect the CHS promoter. In vivo genomic footprinting pinpointed four short nucleotide sequences, termed boxes I to IV, that displayed altered reactivity to dimethylsulfate upon light treatment of parsley cells (Schulze-Lefert et al., 1989a, 1989b). Boxes II (5′-CCACGTGGCC-3′) and III (5′-GTACGTGG-3′) are sequence related and are organized together with boxes I and IV, respectively, into two synergistically acting light-responsive units. Boxes I and II (unit 1) are both necessary, but neither is sufficient for light responsiveness within the context of a CHS minimal promoter containing unit 1. A 52-bp region corresponding to unit 1 suffices to confer light-mediated regulation upon a heterologous TATA box (Weisshaar et al., 1991a).

Box II closely resembles both the G-box, a sequence element first defined in light-regulated plant ribulose 1,5-bisphosphate carboxylase (rbcS) promoters (Giuliano et al., 1988), and the related dyad G-box (Ferl and Laughner, 1989). The palindromic nature of box II and the G-box suggests that dimeric nuclear factors bind to these sites in vivo. A detailed functional analysis of box II identified an asymmetric heptanucleotide core

1 Current address: Institute for Plant Sciences, Department of Plant Genetics, ETH-Zentrum, Universitätstr. 2, CH-8092 Zürich, Switzerland.
2 To whom correspondence should be addressed.
The Plant Cell

(5'-ACGTTGGC-3') crucial for activity of the CHS minimal light-responsive promoter, however (Block et al., 1990). The possibility that box II-like sequences are not exclusively involved in photoregulation arose from the observation of similar sequences in promoters from plant genes subject to regulation by environmental and developmental signals other than light (Schulze-Lefert et al., 1989a). The functional core of box II not only coincides exactly with a portion of the G-box, but also with various other plant promoter elements (Weisshaar et al., 1991a, 1991b). These include the EmIa and motif I elements involved in regulation by abscisic acid in the wheat Em (Marcotte et al., 1989; Guiltinan et al., 1990) and rice abscisic acid--responsive (rab) gene promoters (Mundy et al., 1990; Oeda et al., 1991), respectively, and sequences overlapping the hex1 motif from the wheat histone H3 promoter (Tabata et al., 1989), a putative regulatory element suggested to play a role in cell cycle--dependent transcription (Tabata et al., 1991). Functional significance has been attached to the G-box from the Arabidopsis rbcS-1A promoter (Donald and Cashmore, 1990) and the EmIa element (Guiltinan et al., 1990). Box II also shares at least an ACGT core with plant pathogen promoter elements (Weisshaar et al., 1991a), such as the as-1 element from the cauliflower mosaic virus (CaMV) 35S promoter (Lam et al., 1989) and the bacterial octopine synthase (ocs) enhancer sequence (Bouchez et al., 1989). The family of ACGT-containing promoter elements can be divided into a box II/G-box--like subfamily (e.g., box II, box III, G-box and dyad G-box, motif I, EmIa) and a TGACG-like subfamily (e.g., as-1, ocs, hex1), although the hex1 site in fact bridges this classification (Guiltinan et al., 1990; Weisling and Kahl, 1991; Weisshaar et al., 1991a, 1991b).

To analyze the components of the signal transduction chain linking light perception to gene expression, we have pursued the identification of proteins that interact with sequences critical for light regulation of the parsley CHS promoter. This approach resulted in the recent cloning of cDNAs encoding three putative transcription factors that interact with boxes II and III in vitro. These putative transcription factors were designated common plant regulatory factors (CPRFs) because of their interaction with members of the ACGT-containing family of promoter elements. Nuclear extracts from many plant species also contain factors able to recognize the box II/G-box--like subfamily of ACGT elements (Giuliano et al., 1988; Staiger et al., 1989, 1991; DeLisle and Ferl, 1990; Donald et al., 1990; Guiltinan et al., 1990; McKendree et al., 1990; Oeda et al., 1991). Individual CPRFs may, however, participate in specific regulatory processes. Of the three parsley factors cloned, CPRF-1 appears to be a good candidate for involvement in the light-induced transcription of the CHS gene. CPRF-1 mRNA accumulates strongly in response to irradiation with a more rapid time course than that observed for CHS mRNA (Weisshaar et al., 1991a).

CPRF-1, CPRF-2, and CPRF-3 contain putative basic domain/leucine zipper (bZIP) DNA binding motifs. bZIP motifs (Vinson et al., 1989) are characterized by a region of basic amino acids that contact the DNA (Talanian et al., 1990) immediately N-terminal to a heptad leucine repeat (Landschulz et al., 1988), or more generally a 4-3 hydrophobic repeat, required for dimerization (O'Shea et al., 1989). Dimeric bZIP transcription factors, typified by yeast GCN4 and the mammalian ATF-CREB, Jun-Fos, and C/EBP families, recognize binding sites with dyad symmetry (reviewed in Busch and Sassone-Corsi, 1990; Lamb and McKnight, 1991). cDNAs encoding other plant bZIP proteins that recognize ACGT-containing promoter elements have been isolated in several laboratories (Hartings et al., 1989; Katagiri et al., 1989; Tabata et al., 1989, 1991; Guiltinan et al., 1990; Schmidt et al., 1990; Singh et al., 1990; Lohmer et al., 1991; Oeda et al., 1991). The bZIP domains of these proteins and of CPRF-1, CPRF-2, and CPRF-3 are highly conserved, in particular the basic regions. In addition, tobacco TGACG-sequence-specific binding protein 1a (TGA1a) (Katagiri et al., 1989) and wheat histone DNA binding protein-1b (HBP-1b) display 51% overall sequence identity (Tabata et al., 1991), including significant similarity outside of their respective bZIP domains.

We have identified specific and distinct binding activities for boxes II and III in parsley nuclear extracts. To explore the correspondence between nuclear proteins that interact with boxes II and III and the pool of putative transcription factors recognizing these sequences, we have studied the binding characteristics of CPRF-1, CPRF-2, and CPRF-3, and of nuclear extracts to an array of ACGT-containing promoter elements. Our data indicate the potential of the CPRFs for complex and selective interactions with one another and with native nuclear proteins.

RESULTS

The CPRF bZIP Domains Suffice for Binding to Box II

We have previously characterized three cDNAs encoding putative bZIP proteins (CPRF-1, CPRF-2, and CPRF-3) that displayed specific interactions with a multimerized box II probe in a filter binding assay. To determine whether the bZIP domains alone would qualitatively suffice for DNA binding, we prepared several expression constructs. Each directed the synthesis of different portions of the respective CPRF by coupled in vitro transcription/translation, as shown in Figure 1A. Further diversity in the proteins studied was generated by 3' deletions of the transcription template. Several of the chosen restriction sites lie within the heptad leucine repeats of the respective bZIP domains (Figure 1B). CPRF derivatives were assayed for binding to a box II monomer by gel retardation, as shown in Figure 2. The amounts of CPRFs were determined by analysis of 35S-methionine-labeled translation products taking into account the number of labeled amino acids per protein, although this proved to be difficult for CPRF-3 derivatives, each of which possesses only one methionine.
Three different truncated derivatives of CPRF-1 bound to box II, including CPRF-1b (Figure 1B; see also Weisshaar et al., 1991a), indicating that regions of the polypeptide located N-terminal to the bZIP domain are not essential for DNA binding. In CPRF-1c, a C-terminal deletion removing the last residue of the heptad leucine repeat (Leu335; Figure 1B) resulted in the formation of two DNA–protein complexes (Figure 2). This effect is likely a consequence of incomplete digestion of the transcription template, as an additional XbaI site occurs in the cDNA between the XbaI and the SalI sites shown for CPRF-1 in Figure 1A. In CPRF-1d, deletion of the last two residues of the heptad leucine repeat (Leu328 and Leu335) also did not abolish DNA binding. For C-terminal deletions of CPRF-2 (Figure 1A), complex formation was most evident with the three longest protein derivatives (CPRF-2a, CPRF-2b, CPRF-2c; Figure 2). Further deletions proximal to (CPRF-2d) and just inside the bZIP domain (CPRF-2e), in the latter case removing the last residue of the heptad leucine repeat (Leu253; Figure 1B), severely reduced but did not eliminate binding. A further deletion that completely removed the bZIP domain (CPRF-2f) abolished DNA binding. N-terminal deletions of CPRF-2 (CPRF-2g, CPRF-2h) did not eliminate box II binding, although the complexes appeared more diffuse than for CPRF-2a, CPRF-2b, and CPRF-2c. Two different derivatives of CPRF-3 (Figure 1A) were tested for binding to box II. CPRF-3b, containing an intact bZIP domain, generated a single complex (Figure 2),
Figure 2. The CPRF bZIP Domains Suffice for Binding to Box II.
mRNA indicates the in vitro synthesized CPRF transcript added to the translation extract. An unspecific (U) DNA–protein complex often observed in the no mRNA controls and the free (F) probe are shown. One or more additional complexes were present in lanes 2 to 4, 6 to 10, 12 to 13, and 15 upon translation of CPRF mRNA.

whereas CPRF-3c, lacking the last three residues in the heptad hydrophobic repeat (Ile^{245}, Val^{252}, Ile^{259}; Figure 1B), did not bind to DNA.

Characteristics of Nuclear CPRFs and CPRF-1, CPRF-2, and CPRF-3 Binding to Diverse ACGT-Containing Promoter Elements

We tested the ability of parsley nuclear proteins to bind specifically to box II using a monomer probe in a gel retardation assay, as shown in Figure 3. Nuclear extracts from light-treated and dark-grown parsley cells yielded indistinguishable and complicated patterns of DNA–protein complexes. The addition of proteinase K (lane 4 in Figure 3) abolished complex formation. All labeled complexes were titrated away by increasing amounts of the unlabeled box II competitor (up to 125-fold), though most complexes persisted in the presence of either an unspecific sequence (lanes 8 to 11 in Figure 3) or an oligonucleotide containing a mutated version of box II (lanes 12 to 15 in Figure 3). Nuclear proteins from light-treated or dark-grown parsley cells formed at least five distinct and specific complexes, denoted N1 to N5, with the box II probe. A prominent unspecific complex, migrating immediately below N4 (compare lanes 12 to 15 in Figure 3), was not obvious in other experiments. We consider the proteins forming complexes N1 to N5 to be nuclear CPRFs, according to our definition.

To further characterize nuclear extract binding activities, we prepared a series of oligonucleotide probes, shown in Figure 4A, corresponding to plant (box III, hex1, and Em1a) and viral (as-1) promoter elements related to box II by at least an ACGT motif. As controls, we used mutated versions of boxes II and III (boxes IIm and IIIm), in which the core ACGT motif had been altered at all four positions. Figure 4B shows the pattern of complexes obtained with nuclear extracts. Although boxes II and III are sequence related, substantially different patterns of complexes were observed with these probes (compare lanes 2 and 4 in Figure 4B). Mutations of the ACGT core in boxes IIm and IIIm eliminated almost all binding (compare lanes 3 with 2 and 5 with 4). In the case of box IIIm, residual binding may be an unintentional consequence of the nature of the mutation (compare portions of boxes III 5'-TGTACGtgG-3' and IIIm 5'-TGTACGgaG-3' in Figure 4A; boldface letters indicate residues at the position of the ACGT core, and uppercase letters indicate conserved residues). Nuclear extracts also contain a strong as-1 binding activity, reflected by the presence of a dominant complex not prominent in the binding patterns for boxes II or III (lane 6 in Figure 4B). The hex1 and Em1a probes, which contain the heptanucleotide functional core of box II (Figure 4A), yielded similar binding patterns to those observed with box II, although a major new complex was also observed with the hex1 probe (lane 7 in Figure 4B).

To compare the binding specificities of CPRF-1, CPRF-2, and CPRF-3 with the nuclear CPRFs for an identical array of ACGT-containing probes, gel retardation assays were performed with

Figure 3. Proteins from Parsley Nuclear Extracts Bind Specifically to Box II.

Nuclear extracts from dark-grown (lane 2) or light-treated (lanes 3 to 16) parsley cells were incubated with a box II probe. Competitor (Comp) DNAs used and the molar excesses of competitors to the probe are as shown (see Figure 4A for boxes II and IIm sequences). The unspecific competitor was prepared by hybridization of the following complementary oligonucleotides: 5'-CAACTGATGATGAGCAATGACAGATATCAGATAAGTCGCTT-3' and 5'-GTGAGCTTATCTGAGTACGTTCTTGTACTATCCATAG-3'. At the left margin, the free (F) probe and specific box II–nuclear protein complexes (N1 to N5) that were stable to competition with the unspecific oligonucleotide (lanes 9 to 11) and box IIm (lanes 13 to 15) are indicated.
in vitro translated polypeptides (Figures 4C to 4E). CPRF-2a, CPRF-1b, and CPRF-3b (Figure 1A), the largest respective CPRF derivatives tested, were used for this analysis. As seen in Figures 4C to 4E, the CPRFs all bound strongly to the closely related box II, hex1, and Em1a probes (lanes 2, 7, and 8), moderately to box III (lane 4), and not at all to boxes Ilm and IlIlm (lanes 3 and 5). Although a direct comparison between the CPRFs may not be appropriate due to differences in the sizes of the truncated polypeptides studied, we note that CPRF-2a interacted more strongly with the as-1 probe, relative to other probes, than did CPRF-1b or CPRF-3b (compare lanes 6 and 4 for Figures 4D, 4C, and 4E). CPRF-2a, but neither CPRF-1b

Figure 4. Binding of Parsley Nuclear Factors and CPRFs to Promoter Elements Containing ACGT Cores.

(A) Oligonucleotide probes and competitors used in this study (upper strands only). Nucleotides in uppercase letters indicate the extents of the promoter elements. The central ACGT is denoted by boldface letters. Dots above the sequences of boxes II and III, positions -243 to -223 and -175 to -152 in the parsley CHS promoter, respectively, indicate residues defining the in vivo footprints. The large box encloses residues of functional importance in box II and the analogous positions in the other sequences. The small boxes enclose nucleotides that were altered in the box II and box III mutant (m) series. Oligonucleotides for boxes II, Ilm, III, IlIlm, and the as-1 element have been previously described (Weisshaar et al., 1991a). The additional mutant oligonucleotides for boxes II and III described here are numbered sequentially 5' to 3' across the binding site beginning with m2. The as-1 sequence is derived from the CaMV 35S promoter (positions -83 to -62; Lam et al., 1989), hex1 from the wheat histone H3 promoter in the reverse orientation (positions -160 to -180; Tabata et al., 1989), and Em1a from the wheat Em promoter sequence (positions 1088 to 1108, EMBL accession number X52103; Guiltinan et al., 1990).

(B) Factors in parsley nuclear extracts bind to diverse ACGT-containing promoter elements. Mix (lane 1) signifies the use of a mixture of all seven probes from lanes 2 to 8 in the absence of nuclear extract as a control. Nuclear extracts from light-treated parsley cells, free probes, and the DNA-protein complexes are labeled as given in Figure 3.

(C) to (E) CPRF-1, CPRF-2, and CPRF-3 bind to diverse ACGT-containing promoter elements. CPRF mRNAs added to the translation extract are specified at the top of each panel, and the unspecific complex (U) is labeled. The free probes are not shown. The positions of the respective specific complexes (see Figure 2) are indicated in the left margin. A novel complex observed with CPRF-2a and the as-1 probe is labeled at the left as 2a*.
nor CPRF-3b, also formed a novel complex designated 2a*, with the as-1 probe (Figure 4D; see Discussion).

To demonstrate the binding specificity of nuclear CPRFs for boxes II and III, competition series were performed, as shown in Figures 5A and 5B, using ACGT-containing oligonucleotides for as-1, hex1, Em1a, box II, box III, and a series of mutants in the latter two sequences (Figure 4A). Competition for box II binding was extremely efficient with box II, Em1a, and hex1 (lanes 11, 5, and 4 in Figure 5A), efficient with box III (lane 6), and weak with as-1 (lane 3). Of the box II mutant sequences, only box Ilm did not compete for binding at all (lane 15), although single or double point mutations in flanking positions (lanes 13, 14, and 16) clearly reduced the ability of these sequences to compete. Competition for box II binding was efficient with boxes III and Ilm2 (lanes 6 and 7), weak with box Ilm3 (lane 9), and absent with box Ilm (lane 8). The competition pattern for complex N4 with box II differed from those of N1 to N3 and N5 in that binding was relatively strong in the presence of hex1 or box Ilm4 competitors (lanes 4 and 14 in Figure 5A). Thus, nuclear CPRF–box II complex formation in the presence of ACGT-containing competitors mirrors the binding pattern observed in Figure 4B.

The competition profiles for box III binding by nuclear extracts (Figure 5B) differed dramatically from the results obtained with box II (Figure 5A). Although the patterns of box III complexes observed with extracts from dark-grown and light-treated cells were again indistinguishable, both efficient and poor competitors for box II binding competed similarly for box III binding. Box II, hex1, and Em1a competed weakly for formation of the box III complexes N6 and N9 (lanes 11, 4, and 5 in Figure 5B), whereas box II, hex1, Em1, and as-1 (lane 3) competed somewhat for N7 formation, and as-1 and Em1a competed for formation of N8. On the other hand, none of these competitors was as efficient as either box III itself or box Ilm3 (lanes 6 and 9). Mutations in boxes Illm2 and Illm reduced the competition for formation of some DNA–protein complexes compared to box III, although the nature of the mutation in box Illm (see above) may have permitted some unexpected competition. Competition profiles obtained for binding to boxes II or III using nuclear extracts isolated from dark-grown parsley cells were indistinguishable from the results shown in Figures 5A and 5B (data not shown).

![Figure 5](image-url)

**Figure 5.** Diverse ACGT-Containing Promoter Elements Compete for the Binding of Nuclear Extracts and the CPRFs to Boxes II and III.

(A) and (B) Box II and box III nuclear binding competitions, respectively. The indicated competitors (Comp), listed in Figure 4A, were used at a 125-fold molar excess to the labeled probe. Nuclear extracts from either dark-grown (lane 1) or light-treated (lanes 2 to 19) parsley cells were mixed with the indicated probes. Box II–protein complexes are labeled as given in Figure 3. Complexes formed with box III are designated as N6 to N9.

(C) to (E) Competition series for CPRF binding to box II. Competitors (Comp) were used at a 125-fold molar excess to the labeled probe. CPRF mRNAs translated are specified at the top of each panel. DNA–protein complexes are labeled as given in Figures 4C to 4E.
The binding site requirements of the CPRFs were further explored by performing an identical competition series (Figures 5C to 5E). In all three cases, competition was very efficient with box II, hex1, and Emf1 (lanes 11, 4, and 5 in Figures 5C to 5E), strong with boxes III and Ilm2 (lanes 6 and 7), poor with box Ilm3 (lane 9), and not detectable with box Ilm (lane 8). The differences in as-1 binding abilities (Figures 4C to 4E) were also evident in the competition analysis. CPRF-2a binding to box II was more efficiently competed by as-1 than was the case for CPRF-1b or CPRF-3b binding (lane 3 in Figures 5C to 5E). When comparing the ability of box II mutant derivatives to compete for CPRF binding to box II, further subtle differences emerged. CPRF-3b displayed a broader binding site requirement than did either CPRF-1b or CPRF-2a, both of which were similar to one another. With either CPRF-1b or CPRF-2a, boxes Ilm2, Ilm3, Ilm4, Ilm, Ilm5, and Ilm6 competed less efficiently for box II binding than did box II itself (lanes 12 to 17). For CPRF-3b, box Ilm7 also exhibited reduced competition (lane 18). As observed with the nuclear CPRFs, box Ilm did not compete for binding (lane 15). The effectiveness of other single and double point mutants of box II in competitions decreased with the proximity of the mutated nucleotides to the ACGT core. We assume that the binding specificities of truncated and intact CPRF-1, CPRF-2, and CPRF-3 are similar, as we observed no differences in competition profiles between different truncated derivatives tested. Binding experiments using boxes Ilm, III, Ilm, Ilm2, and Ilm3 as probes reflected the competition results obtained with these sequences (Figures 4C to 4E and data not shown).

CPRF-1, CPRF-2, and CPRF-3 Bind to DNA as Dimers, Selectively Form Heterodimers, and Interact with Nuclear CPRFs

A prediction based on the fact that CPRF-1, CPRF-2, and CPRF-3 contain bZIP domains is that these proteins bind to DNA as dimers. To test this hypothesis, mRNAs for various combinations of CPRF derivatives were cotranslated in vitro and the binding properties of the mixtures of products were studied by gel retardation assays. Templates for protein synthesis were selected such that dimers formed by two polypeptides of different sizes should yield a DNA–protein complex of intermediate mobility to that of homodimers of either of the two polypeptides.

Cotranslation of CPRF-1b and CPRF-1d led to the formation of a major new box II complex (CPRF-1b1d) with an intermediate mobility, as shown in lane 15 of Figure 6. Similarly, cotranslation of CPRF-2a and CPRF-2g also yielded a prominent intermediate complex (CPRF-2a2g in lane 19). Although not directly tested, we assume that CPRF-3 also binds to DNA as a homodimer. Cotranslation of CPRF-1b and CPRF-3b resulted in the appearance of a major complex of intermediate mobility (CPRF-1b3b in lane 7 of Figure 6), in addition to the homodimeric complexes. Similarly, cotranslation of CPRF-2a and CPRF-3b led to formation of a minor intermediate complex (CPRF-2a3b in lane 11). Although as much CPRF-2a homodimeric complex was evident in the CPRF-2a and CPRF-3b cotranslations as in the CPRF-2a and CPRF-1b cotranslations (compare lanes 11 and 3 of Figure 6), no new complex was apparent for cotranslated CPRF-2a and CPRF-1b. This suggests that these proteins do not readily form box II binding heterodimers. These data also indicate that CPRF-1, CPRF-2, and CPRF-3 are capable of selectively forming DNA binding heterodimers with each other. Similar results were obtained by mixing the in vitro translated proteins prior to the binding reactions, although dimer formation was in no case as evident as with cotranslation.

To determine whether the box II binding specificities of the CPRF-1, CPRF-2, and CPRF-3 heterodimers differed from those of the homodimers, we performed a competition series with the same array of ACGT-containing promoter elements and mutant derivatives shown in Figure 5. Figure 7A illustrates the simultaneous competition patterns observed for CPRF-1b, CPRF-3b, and the CPRF-1b3b heterodimeric complex. The latter displayed a competition profile similar to that of its homodimeric parents. Similarly, the competition pattern for CPRF-2a3b binding to box II did not deviate substantially from that of either individual CPRF homodimer (Figure 7B). These results indicate that both the heterodimeric and homodimeric CPRF complexes formed recognize approximately the same sites on the DNA probe.

To delve further into the relationships between box II binding by nuclear CPRFs and CPRF-1, CPRF-2, and CPRF-3, we tested whether mixed DNA binding complexes could be generated. We combined nuclear extracts from light-treated parsley
Figure 7. Diverse ACGT-Containing Promoter Elements Compete for Binding of CPRF Heterodimers to Box II.

(A) Competition of CPRF-1b3b heterodimer binding.
(B) Competition of CPRF-2a3b heterodimer binding.

DNA-protein complexes, competitors (Comp), and mRNAs translated are labeled as given in Figures 4 to 6. Competitors were used at a 125-fold molar excess to the labeled probe.

cells with CPRF-1b, CPRF-2a, or CPRF-3b proteins, before initiation of the binding reaction and analysis by gel retardation. The results are shown in Figure 8. CPRF-1b and CPRF-3b formed new complexes with native nuclear proteins (designated N:1b and N:3b in lanes 3 and 7 of Figure 8, respectively). For CPRF-2a, formation of a new complex absent from the extracts alone was not evident, although the similarity in the mobilities of the CPRF-2a complex and the slowest migrating major nuclear extract complex complicated this determination. The apparent weakness of nuclear extract binding to box II (see Figures 3 and 8) is most likely a consequence of the mixing conditions.

DISCUSSION

Features of CPRF-DNA Binding and Selective Dimerization

A complex pattern of protein-DNA interactions results from binding of parsley nuclear extracts to the set of box II/G-box-like probes used in this study. These sequences also bind multiple factors and/or form multiple complexes with nuclear extracts from various other plant species (Giuliano et al., 1988; Mikami et al., 1989; Staiger et al., 1989, 1991; DeLisle and Ferl, 1990; Donald et al., 1990; Guiltinan et al., 1990; McKendree et al., 1990; Oeda et al., 1991). In parsley nuclear extracts, the patterns of DNA-protein complexes formed are similar regardless of whether box II, Em1a, or hex1 is used as the probe. Therefore, we have avoided new designations for DNA binding activities of nuclear CPRFs that would indicate a special association between box II and complexes N1 to N5. We conclude from these data that the detection of different subgroups from a pool of nuclear CPRFs that are able to interact with box II/G-box-like promoter elements depends on the particular probe used.

Box II was physically defined as a light-dependent in vivo footprint (Schulze-Lefert et al., 1989a). Surprisingly, the nuclear proteins present in extracts from either dark-grown or light-treated parsley cells gave rise to identical patterns of box II-protein complexes. Other researchers have also reported no (Staiger et al., 1989) or subtle (Giuliano et al., 1988; Guiltinan et al., 1990) differences in the complexes formed between boxII/G-box-like sequences and nuclear proteins isolated from plants or cells maintained in various environmental backgrounds. One explanation in parsley could be the sequestering or inactivation of box II binding factors in the dark in vivo, possibly through differential modification (reviewed by Berk, 1989), which may be lost upon isolation of nuclear extracts. Alternatively, the complexity of interactions between box II and parsley nuclear CPRFs may obscure subtle light/dark differences relevant to regulation of the CHS promoter.

Figure 8. CPRF-1 and CPRF-3 Interact Selectively with Box II Binding Components of Parsley Nuclear Extracts.

Nuclear extracts from light-treated parsley cells, DNA-protein complexes, and CPRF mRNAs translated are labeled as given in Figure 4, except that novel complexes formed (lanes 3 and 7) after mixing nuclear extracts with in vitro translated CPRFs are indicated at the right as N:1b and N:3b, respectively.
Unexpectedly, boxes II and III gave rise to different patterns of complexes with nuclear proteins in both binding and competition experiments. We do not favor the hypothesis that certain minor components of box II binding activity correspond to prominent box III complexes, given the divergence in the competition patterns for binding to these two probes. Box III was originally classified as a box II-like element (Schulze-Lefert et al., 1989a), but actually resembles even more closely the motif I sequence conserved among rice rab and cotton late embryo abundant (lea) genes (Mundy et al., 1990). Box III can, however, functionally replace box II in an orientation-dependent fashion within the context of the minimal light-responsive CHS promoter (Block et al., 1990). These results indicate that distinct sets of nuclear proteins interact with boxes II and III, despite their sequence similarity, and that in this case the geometry rather than the identity of the factors bound to these promoter elements may be significant for function. The observation that unit 2 (boxes III and IV) constitutes a light-responsive promoter region analogous to unit 1 (boxes I and II) (Schulze-Lefert et al., 1989b) also supports this hypothesis as boxes I and IV are not sequence related.

Truncated CPRF-1, CPRF-2, and CPRF-3 polypeptides lacking all amino acids N-terminal to the respective bZIP domains continued to bind to DNA. The heptad leucine repeat of CPRF-1 contains 6 leucine residues, the last 2 of which are qualitatively dispensable for dimerization and DNA binding. In CPRF-3, deletion of nonleucine residues in the last three heptad positions of the hydrophobic repeat abolished DNA binding, presumably by preventing dimerization of the polypeptide. This suggests that the deleted hydrophobic residues at the repeat positions may participate in dimerization, consistent with studies on the requirements for dimerization of the Fos and Jun proteins (Ryseck et al., 1990; Schueremann et al., 1991). For CPRF-2, C-terminal deletions close to or just inside the leucine repeat drastically reduced but did not abolish DNA binding. We suggest that residues located C-terminal to the leucine repeat may help to stabilize CPRF-2 dimer formation and hence DNA binding, as reported for other bZIP transcription factors (Schueremann et al., 1991). On the other hand, the CPRF-2 bZIP domain was demonstrated, by deletion of the entire region in the CPRF-2f polypeptide, to be absolutely required for binding to DNA. Two studies to date have provided data on the protein requirements for DNA binding of other cloned plant bZIP proteins. Tabata et al. (1991) demonstrated that truncated HBP-1a and HPB-1b polypeptides containing intact bZIP domains continued to bind to DNA, whereas Singh et al. (1990) have shown that a C-terminal truncation of ocs element binding factor-1 destroyed DNA binding when a portion of the heptad leucine repeat was deleted. As predicted on the basis of these and other studies of bZIP proteins (reviewed in Johnson and McKnight, 1989; Busch and Sassone-Corsi, 1990), we observed that the bZIP domains of the parsley CPRF-1, CPRF-2, and CPRF-3 are crucial for DNA binding and dimerization.

Direct evidence for dimerization of CPRF-1, CPRF-2, and CPRF-3 was obtained by cotranslation of different polypeptide derivatives. Although CPRF homodimers displayed similar interaction patterns with the ACGT-containing promoter elements tested and demonstrated particularly strong binding to box II/G-box-like sequences, subtle differences emerged in both binding and competition experiments. The novel low mobility 2a* complex formed between CPRF-2a and the as-1 probe (Figure 4D) probably represents the interaction of two CPRF-2a dimers with both of the tandem TGACG sites present in the as-1 sequence. Similarly, the tobacco activation sequence factor-1 formed two complexes with the as-1 element, depending upon the occupation of both TGACG sites (Lam et al., 1989; Katagiri et al., 1989). The fact that neither CPRF-1b nor CPRF-3b produced novel complexes with the as-1 probe may arise from differences in the protein–protein interactions possible with each of the three truncated CPRFs tested.

Cotranslation and analysis of different pairwise combinations of CPRF-1, CPRF-2, and CPRF-3 lead to the conclusion that these proteins selectively form DNA binding heterodimers, a demonstration not yet reported for other plant bZIP proteins. Differences in the amounts of box II complex formed between homodimers and heterodimers could be the consequences of altered affinities for target sites or of relative dimer stabilities (Lamb and McKnight, 1991). CPRF-1b3b and CPRF-2a3b heterodimers bind to box II with patterns similar to those of the parent homodimers, suggesting the latter possibility. Heterodimerization between members of box II/G-box binding and TGACG binding subfamilies of CPRFs, leading to novel intermediate binding activities, cannot be ruled out, however. The formation of CPRF homodimers and selective heterodimers, as well as the subtle differences in their binding to ACGT elements, parallel the properties of mammalian factor families (Lamb and McKnight, 1991). These qualities of CPRFs may be representative of other families of plant transcription factors. The individual homodimers or heterodimers of CPRF-1, CPRF-2, and CPRF-3 have definite, though related, binding-site preferences and retain, at least in principle, the potential to interact with promoter elements from diversely regulated genes. The reported binding patterns of most other cloned plant bZIP proteins differ from the spectrum described here for CPRF-1, CPRF-2, and CPRF-3. Most closely related in terms of binding specificities are the Em binding protein-1 (EmBP-1) (Guitian et al., 1990) and the tobacco transcription factor TAF-1 (Oeda et al., 1991). EmBP-1 demonstrates the highest affinity for the Em1a sequence among the binding sites tested, but can also interact with the hex1 site, whereas TAF-1 binds to motif I, but interacts equally well with hex1 and even more efficiently with the G-box. At least CPRF-2 defines a novel subgroup of putative bZIP transcription factors with overlapping but distinct affinities for box II/G-box-like (box II, Em1a, hex1, and box III) and TGACG-like sequences (as-1).

Of the amino acid sequences deduced to date from cDNAs encoding plant bZIP ACGT-binding proteins (CPRFs by our definition), only TAF-1 (Oeda et al., 1991) exhibits significant similarity outside of the bZIP domain with parsley CPRF-1, CPRF-2, or CPRF-3. The partial sequences available for CPRF-1 and TAF-1 are 43% identical outside the bZIP region and 94 and 58% identical within the basic region and leucine.
repeat, respectively. We therefore make two specific predictions and offer one proposal addressing the relationship between CPRF-1 and TAF-1. The near absolute identity of the TAF-1 and CPRF-1 basic regions indicates that TAF-1 should, using our experimental conditions, interact weakly with the TGACG subfamily of promoter elements, including as-1. More importantly, the strong conservation of the CPRF-1 and TAF-1 leucine repeats suggests that both proteins, assuming that TAF-1 behaves like normal bZIP proteins, will form heterodimers with a common partner(s) from the pool of available CPRFs. Patches of conserved solvent-exposed residues within leucine repeats have been postulated to represent surfaces for protein–protein contacts (Lamb and McKnight, 1991). We further propose, based on the similarities in overall protein sequence and in binding site preferences, that CPRF-1 and TAF-1 may participate in related regulatory pathways. TAF-1 functions as a transcriptional activator in transgenic tobacco plants, although it does not seem to play a direct role in abscisic acid–regulated gene expression (Oeda et al., 1991). CPRF-1 may be involved in the light responsiveness of the CHS promoter as the light-dependent transient accumulation of CPRF-1 and CHS mRNAs occurs sequentially in parsley cells (Weisshaar et al., 1991a). This brief comparison suggests specific experiments to test the above proposal. As previously noted (Weisshaar et al., 1991a), the leucine repeats of the maize Opaque-2 protein (Hartlings et al., 1989; Schmidt et al., 1990) and CPRF-2 are also highly conserved (64% identity compared to about 25% between combinations of CPRF-1, CPRF-2, and CPRF-3). By analogy to CPRF-1 and TAF-1, we hypothesize that Opaque-2 and CPRF-2 may heterodimerize with a common partner(s).

Relationships between CPRF-1, CPRF-2, and CPRF-3 and Plant Nuclear CPRFs

Additional experiments will be necessary to determine the contributions of CPRF-1, CPRF-2, and CPRF-3 to the spectrum of box II–nuclear CPRF complexes. Both binding and competition experiments demonstrate that all of these CPRFs bind to box II specifically in vitro and thus have the potential to interact in vivo. The characteristics of box II binding by nuclear proteins (complexes N1 to N5) and by the CPRF-1, CPRF-2, and CPRF-3 homodimers reveals some differences in the competition patterns, however. The functional core of box II also does not correspond exactly to those residues important for binding of CPRF homodimers. We have demonstrated, however, that CPRF-1 and CPRF-3 form heterodimeric box II binding complexes with native nuclear CPRFs present in extracts from light-treated parsley cells. Formation of a single dominant heteromeric complex in both cases suggests that CPRF-1 and CPRF-3 interact preferentially with specific CPRFs. Whether or not this preferential interaction represents dimerization of the native full-length CPRF with its truncated counterpart remains to be established. One intriguing possibility is that light-induced de novo synthesis of CPRF-1 could alter the cellular equilibrium of CPRF heterodimers able to interact specifically with box II. The asymmetry of the box II functional core centered around the ACGT (Block et al., 1990) supports the hypothesis that a heterodimeric factor(s) binds to box II in vivo in the context of the minimal light-responsive CHS promoter. We therefore suggest that nonpromiscuous dimer formation among nuclear CPRFs, including CPRF-1, CPRF-2, and CPRF-3, produces a pool of factors capable of interacting with box II/G-box–like and other ACGT-containing promoter elements in vivo.

The results described here provide the starting point to decipher the complex relationships between CPRF-1, CPRF-2, CPRF-3, and nuclear proteins contributing to the light regulation of the CHS promoter. Emerging analogies between plant bZIP proteins may strengthen our understanding of the contributions of protein–protein interactions to selective dimerization and transcriptional regulation.

METHODS

Maintenance of Parsley Cell Suspensions and Preparation of Nuclear Extracts

The maintenance and irradiation of parsley (Petroselinum crispum) cell suspensions have been previously described (Dangl et al., 1987). Parsley nuclear extracts were prepared by a modification (P. Schweizer, personal communication) of a method used for the isolation of wheat nuclei (Schweizer et al., 1989), using 50 g of light-treated (4 hr of continuous irradiation) or dark-grown parsley cells. Cells were ground to a fine powder under liquid N2 and suspended in 2.2 volumes (v/v) of nuclear isolation buffer B (70% glycerol [v/v], 20 mM Tris–Cl, pH 7.8, 5 mM MgCl2, 5 mM KCl, 250 mM sucrose). The homogenate was strained through nylon meshes of 60 and 20 μm, respectively. The filtrate was centrifuged for 1 hr at 3500g at 4°C, and the supernatant was decanted. The pellet was resuspended in buffer B (1.5 volumes [v/v]) and centrifuged as above for 40 min, and the supernatant was again removed. The pellet was then resuspended in 10 mL of high salt buffer (20 mM Tris–Cl, pH 7.8, 5 mM MgCl2, 0.5 mM NaCl), incubated on ice for 15 min with slow stirring, and centrifuged for 20 min at 20,000g at 4°C. The final supernatant was dialyzed against 2 L of dialysis buffer (20% glycerol [v/v], 0.1 mM EDTA, 50 mM KCl, 25 mM Hepes/KOH, pH 7.8) at 4°C for 2 hr. All buffers contained freshly added 0.1% (v/v) 2-mercaptoethanol, plus 0.1 μM E-64 (Sigma) and 200 μM phenylmethylsulfonyl fluoride as protease inhibitors. Dialyzed extracts were aliquoted and stored at −70°C until use.

Oligonucleotide Probe and Competitor Preparation

Double-stranded oligonucleotides with engineered restriction site overhangs for use as either probes or competitors were generated by annealing of gel-purified complementary single-stranded oligonucleotides (Figure 4A), prepared with an Applied Biosystems ( Weiterstadt, Germany) DNA synthesizer. Probe DNAs (100 ng) were labeled with the Klenow fragment of DNA polymerase I by filling in the 5' overhangs in the presence of the appropriate [32P]-labeled d-deoxyribonucleoside 5' triphosphates to a specific activity of ∼1 × 108 cpm/μg DNA. Competitor DNAs were prepared in similar fashion using unlabeled nucleotides.
CPRF Protein Expression Plasmids

The construction of pCPRF-lb (originally pCPRF-IlbZIP), used for expression of the CPRF-lb protein, has been previously described (Weisshaar et al., 1991a). CPRF-lc and CPRF-ld proteins were generated by restriction of pCPRF-lb at XbaI and PstI sites in the cDNA insert, respectively, followed by in vitro transcription/translation. pCPRF-1b contains four additional C-terminal amino acids (LOAC) resulting from the restriction of pCPRF-lb 3' to the cDNA insert at a HindIII site.

pCPRF-2a was constructed by cloning a NotI fragment derived from the CPRF-2 cDNA into the NotI site of the expression plasmid pBSEX. A modified pBS+ vector that contains EcoRI and NotI cloning sites in-frame with a start codon located 3' to a T7 RNA polymerase promoter (O. da Costa e Silva, personal communication). pCPRF-2a was restricted at a Sall site 3' to the cDNA insert prior to in vitro transcription. The protein derivatives CPRF-2b, CPRF-2c, CPRF-2d, and CPRF-2e were synthesized from pCPRF-2a templates restricted at NheI, Accl, EcoRV, HinClI, and XbaI sites in the cDNA, respectively. pCPRF-2g was constructed using the polymerase chain reaction (PCR) technique. An internal primer that introduced an EcoRI site immediately 5' to the bZIP domain and an endogenous PstI site 3' to the stop codon were used to produce a fragment that was cloned into EcoRI/PstI-restricted pBSEX. Before in vitro transcription, pCPRF-2g was restricted at a HindIII site 3' to the cDNA insert. pCPRF-2h was constructed in an analogous fashion except that an endogenous NheI site between the bZIP domain and the stop codon was selected as the 3' end of the PCR fragment. This EcoRI-NheI fragment was then cloned into EcoRI/BamHI-restricted pBSEX. CPRF-2h contains seven new C-terminal amino acids (RVDLQAC) introduced by restriction of pCPRF-2h at a HindIII site 3' to the cDNA insert before in vitro transcription.

pCPRF-3b was produced by using a similar PCR approach to introduce an EcoRI site 5' to the bZIP domain and by taking advantage of an endogenous BamHl site in the cDNA 3' to the stop codon. This fragment was ligated into EcoRI/BamHI-restricted pBSEX. pCPRF-3b was restricted at a HindIII site 3' to the cDNA insert prior to in vitro transcription. A further protein derivative, CPRF-3c, was synthesized from a pCPRF-3b template restricted at a PstI site in the cDNA insert.

All CPRF-1 and CPRF-3 protein derivatives used in this study, as well as CPRF-2g and CPRF-2h, contain six new N-terminal amino acids (MAAAEKF) encoded by the vector DNA. CPRF-2a through CPRF-2f protein derivatives contain four new N-terminal amino acids (MAAA). Protein expression plasmids were characterized in all cases by DNA sequencing across the vector/insert junctions, and, in the case of PCR products, the inserts were sequenced in their entirety.

Coupled In Vitro Transcription/Translation

In vitro transcribed CPRF mRNAs were prepared as described in Weisshaar et al. (1991a). To generate unlabeled in vitro translated proteins, similar amounts of the RNAs, as determined by agarose gel electrophoresis and ethidium bromide staining, were added to each 25-μL translation reaction using a nuclease-treated rabbit reticulocyte lysate extract (Boehringer Mannheim). Cotranslation experiments were performed by adding equal amounts of both selected mRNAs to the translation reaction. The 12.5-μL translation reactions containing 10 μCi of 35S-methionine were carried out to assess protein synthesis. Unincorporated 35S-methionine was removed by dialysis against 200 mL of nuclear extract dialysis buffer at 4°C for 2 hr using type V5 0.025-μm pore size filter discs (Millipore, Bedford, MA). Synthesis of proteins was compared by analysis on 15% SDS-polyacrylamide gels after normalizing the amounts of translation extract loaded to reflect the differing numbers of methionines in the respective polypeptides. Protein gels were subsequently fixed, treated with Enlightening (Du Pont), dried, and subjected to fluorography.

Gel Retardation Experiments

Gel retardation assays using parsley nuclear extracts from cells and/or in vitro translated proteins were performed as follows, unless otherwise noted. Binding reactions were initiated by the addition of 5 μL (4 μg of protein) of nuclear extract to a mixture of 4× binding buffer (Hollandsworth and Laties, 1989; Weisshaar et al., 1991a), 400 ng poly(dI- dC), 0.1 ng of probe DNA, and, in some cases, specific competitor in a 20-μL total volume. The mixture was incubated at room temperature for 10 min. Incubation of nuclear extracts with 2 μL of proteinase K (5 mg/mL) was performed during the binding reaction. Unlabeled competitor DNAs were used at a 125-fold molar excess over the labeled probe. In the binding reactions using in vitro translated proteins, equal amounts of in vitro synthesized proteins were added to each binding reaction. In the cotranslation experiments, the protein input was doubled. Mixing of nuclear extracts and in vitro translated proteins was performed for 1 hr at 37°C prior to initiation of the binding reaction. DNA–protein complexes were separated from unbound probe on 5% native polyacrylamide gels run in 0.5 x TBE buffer (Sambrook et al., 1989) at 15 V/cm at 4°C for 2 hr and 15 min or for 3 hr when resolving dimeric complexes. Following electrophoresis, gels were dried and subjected to autoradiography at −70°C with an intensifying screen.

ACKNOWLEDGMENTS

We thank Oswaldo da Costa e Silva, Michael Feldbrügge, and Catharina Maulbecker for critical reading of this manuscript, and Wolfgang Smallenbach and Sabine Schäfer for synthesis of oligonucleotides. This work was in part supported by a National Science Foundation–North Atlantic Treaty Organization Postdoctoral Fellowship (G.A.A.), by a National Science Foundation Fellowship for Long-Term and Medium-Term Research at Foreign Centers of Excellence (G.A.A.), and by a grant from Fonds der Chemischen Industrie (K.H.).

Received December 26, 1991; accepted March 20, 1992.

REFERENCES


Homodimeric and heterodimeric leucine zipper proteins and nuclear factors from parsley recognize diverse promoter elements with ACGT cores.

G A Armstrong, B Weisshaar and K Hahlbrock

*Plant Cell* 1992;4;525-537

DOI 10.1105/tpc.4.5.525

This information is current as of November 4, 2017