The Plant Cell, Vol. 4, 1309–1319, October 1992

TGA1 and G-Box Binding Factors: Two Distinct Classes of Arabidopsis Leucine Zipper Proteins Compete for the G-Box-Like Element TGACGTGG

Ulrike Schindler, Holger Beckmann, and Anthony R. Cashmore[1]

Department of Biology, Plant Science Institute, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6018

Regulatory elements containing the sequence ACGT are found in several plant promoters and are recognized by various basic/leucine zipper (bZIP) proteins. The Arabidopsis G-box binding factor 1 (GBF1), initially identified by its ability to bind to the palindromic G-box (CCACGTGG), also interacts with the TGACGT motif if this hexamer sequence is followed by either the dinucleotide GG—as found in the Hex motif of the wheat *histone 3* promoter—or GT. Here we describe the isolation of an Arabidopsis bZIP protein, denoted TGA1, that also recognizes ACGT-containing sequences. However, TGA1 differs from members of the GBF family in the spectrum of base pair permutations flanking the ACGT sequence that are required for DNA binding. TGA1 primarily requires a TGACG motif and preferentially binds to those pentamers that are followed by a T residue. We show that although both TGA1 and GBF1 bind to the Hex motif (TGACGTGG), this binding can be distinguished on the basis of their specific DNA-protein contacts. Furthermore, TGA1 also differs from members of the GBF family in that it apparently does not form heterodimers with any member of this family.

INTRODUCTION

Transcriptional regulation of gene expression is mediated by the concerted action of sequence-specific transcription factors that interact with regulatory elements residing in the promoter regions of the corresponding gene. One group of transcription factors is defined by a basic/leucine zipper (bZIP) motif (Landschulz et al., 1988; McKnight, 1991). This bipartite DNA binding structure consists of a region enriched in basic amino acids (basic region) adjacent to a leucine zipper that is characterized by several leucine residues regularly spaced at seven–amino acid intervals (Vinson et al., 1989). Whereas the basic region directly contacts the DNA, the leucine zipper mediates homodimerization and heterodimerization of protein monomers through a parallel interaction of the hydrophobic dimerization interfaces of two α-helices, resulting in a coiled-coil structure (O'Shea et al., 1989, 1991; Hu et al., 1990; Rasmussen et al., 1991).

The Arabidopsis bZIP family of G-box binding factors (GBF1, GBF2, and GBF3) interact with the palindromic G-box motif (CCACGTGG) found in several plant promoters (Schindler et al., 1992a). For ease of reference, we have numbered the individual base pairs encompassing the G-box from −4 to +4 (numbering from 5′ to 3′). We have demonstrated that the DNA binding specificity of GBF1 is strongly influenced by the nature of the nucleotides (positions −4, −3, +3, and +4) flanking the ACGT core sequence. For example, sequences that carry a TG at positions −4 and −3 and a T or G residue at position +4 are as efficiently bound by GBF1 as the palindromic G-box.

Interestingly, these two G-box–like elements (TGACGTGG and TGACGTGT) encompass the sequence TGACGT found in the cauliflower mosaic virus (CaMV) 35S promoter (the as-1 element), the enhancers of the nopaline and octopine synthase genes (the nos and ocs elements, respectively), and the wheat *histone 3* promoter (the hexamer or Hex element) (Bouchez et al., 1989; Katagiri et al., 1989; Tabata et al., 1989, 1991; Singh et al., 1990). Only the TGACGT motif found in the wheat *histone 3* promoter fulfills the binding site requirements for GBF1, because in this specific context the TGACGT motif is followed by two G residues (positions +3 and +4) (Schindler et al., 1992b).

Several plant bZIP proteins have been shown to interact with TGACG-related motifs and/or the G-box (Katagiri et al., 1989; Tabata et al., 1989, 1991; Guiltinan et al., 1990; Singh et al., 1990; Oeda et al., 1991; Weisshaar et al., 1991; Schmidt et al., 1992; Ueda et al., 1992). Because all these bZIP proteins bind to DNA motifs carrying the ACGT core sequence, they constitute a broad class of ACGT binding proteins (Tabata et al., 1991; Weisshaar et al., 1991; Armstrong et al., 1992). However, proteins belonging to this group that bind to the as-1 site have DNA binding site requirements distinct from those proteins interacting with the G-box (Tabata et al., 1991). As well as defining individual classes of bZIP proteins according to their DNA binding specificity, such proteins may also be classified according to their heterodimerization characteristics. These criteria have been used by Cao et al. (1991) in describing the properties of the mammalian C/EBP family of bZIP proteins.

To further explore the question of whether Arabidopsis encodes distinct classes of bZIP proteins with overlapping binding

[1] To whom correspondence should be addressed.

specificities, we isolated a cDNA encoding a member of the bZIP class of proteins that interacted with the as-1 and the Hex motif (as found in the wheat *histone 3* promoter), but not with the G-box. Both the deduced amino acid sequence and the expression pattern are very similar to the tobacco protein TGA1a; consequently, we have designated this Arabidopsis protein TGA1. Although TGA1, like GBF1, binds to the Hex sequence, the protein–DNA contacts for these proteins are quite distinct. Using the random binding site selection assay, we further demonstrated that, in contrast to GBF1, TGA1 did not require a specific base pair combination following the TGACG sequence. Furthermore, TGA1 did not productively heterodimerize with members of the GBF family. The results presented here clearly establish criteria (DNA contacts and dimerization properties) that help discriminate between proteins belonging to the class of TGACG binding proteins and those belonging to the GBF family.

RESULTS

Isolation of an Arabidopsis cDNA Encoding TGA1

To isolate Arabidopsis cDNAs encoding proteins that interacted with TGACG-related motifs, we screened an Arabidopsis cDNA library under low-stringency hybridization conditions using a DNA fragment derived from the tobacco TGA1a sequence (Katagiri et al., 1989). Six positively hybridizing clones were isolated. All of these cDNAs were shown to be derived from the same mRNA species. The DNA and the deduced amino acid sequences of the longest of the six cDNAs are shown in Figure 1. The protein encoded by the Arabidopsis cDNA is 63% identical to the tobacco protein TGA1a, 60% identical to the tobacco protein PG13 (Fromm et al., 1991), and 37% identical to the wheat protein HBP-1b (Tabata et al., 1991); the protein differs by three of 30 amino acids from the partial sequence recently reported (Kawata et al., 1992) for a bZIP protein from Arabidopsis (Landsberg).

RNA gel blot analysis demonstrated that the mRNA corresponding to the Arabidopsis cDNA is elevated in roots and dark-grown leaf tissue (data not shown). Similar results have been observed for the tobacco TGA1a mRNA (Katagiri et al., 1989). Because of the similarity in the expression patterns and the strong amino acid sequence identity between the Arabidopsis protein and the tobacco protein TGA1a, we have designated the Arabidopsis protein TGA1.

TGA1 Interacts with TGACGT-Containing Sequences but Not with the G-Box Motif

To explore the DNA binding specificity of TGA1, we performed competitive DNA binding studies employing in vitro–generated TGA1 and various synthetic oligonucleotides, as shown in Figure 2A. The results in Figure 2B demonstrate that TGA1

```
   1  CCTTCCAAACACTATTGGGGAAATGGCTTCTTCTCTTTTTTCCGGTGTAGGTTTAAGG
  59 TTTTAGATTTGAAGGCGGATGGTGAGGAGTTTGTGTTGATGAAATGGGTGATTTGAAGTT
 119 AAATACCTTTAAGTTCTTTTAGCCTTAAGTTCTTTTAGCCTGTCATTACAATATATGTTT
 179 CAGTTGTGATCTTTTGGTGCTCTCCAAGGTCTTAGGGAATCTCCGTGTCCCCTCTGGTTT
 239 CTTTCTTTGTATAAAGTCTCTGGAATTTAAAGATGATTTTGTCCTGTTTTCTTCGAAGAA
 299 TTTGGCGAAGAAACATCAAAAGTTCACTATTTGCTTTACGTGTTATGCTTGTTTGACTTT
 359 GTGTGATCTAACTTTTGGTTGAAAACTAATGCTCATCTTTTGTTCTTTGGAATGTCTTGA
 419 TCTGCTTGATTTTGAATGCTTTGTCACATTGTTTAAACAATTTGTTCTTTGTTTTCAGTT
 479 GAGGAAAACATGAATTCGACATCGACACATTTTGTGCCACCGAGAAGAGTTGGTATATAC
              M  N  S  T  S  T  H  F  V  P  P  R  R  V  G  I  Y   17

 539 GAACCTGTCCATCAATTCGGTATGTGGGGGGAGAGTTTCAAAAGCAATATTAGCAATGGG
     E  P  V  H  Q  F  G  M  W  G  E  S  F  K  S  N  I  S  N  G   37

 599 ACTATGAACACACCAAACCACATAATAATACCGAATAATCAGAAACTAGACAACAACGTG
     T  M  N  T  P  N  H  I  I  I  P  N  N  Q  K  L  D  N  N  V   57

 659 TCAGAGGATACTTCCCATGGAACAGCAGGAACTCCTCACATGTTCGATCAAGAAGCTTCA
     S  E  D  T  S  H  G  T  A  G  T  P  H  M  F  D  Q  E  A  S   77

 719 ACGTCTAGACATCCCGATAAGATACAAAGACGGCTTGCTCAAAACCGCGAGGCTGCTAGG
     T  S  R  H  P  D  K  I  Q  R  R  L  A  Q  N  R  E  A  A  R   97

 779 AAAAGTCGCTTGCGCAAGAAGGCTTATGTTCAGCAACTGGAAACAAGCAGGTTGAAGCTA
     K  S  R  L  R  K  K  A  Y  V  Q  Q  L  E  T  S  R  L  K  L  117

 839 ATTCAATTAGAGCAAGAACTCGATCGTGCTAGACAACAGGGATTCTATGTAGGAAACGGA
     I  Q  L  E  Q  E  L  D  R  A  R  Q  Q  G  F  Y  V  G  N  G  137

 899 ATAGATACTAATTCTCTCGGTTTTTCGGAAACCATGAATCCAGGGATTGCTGCATTTGAA
     I  D  T  N  S  L  G  F  S  E  T  M  N  P  G  I  A  A  F  E  157

 959 ATGGAATATGGACATTGGGTTGAAGAACAGAACAGACAGATATGTGAACTAAGAACAGTT
     M  E  Y  G  H  W  V  E  E  Q  N  R  Q  I  C  E  L  R  T  V  177

1019 TTACACGGACACATTAACGATATCGAGCTTCGTTCGCTAGTCGAAAACGCCATGAAACAT
     L  H  G  H  I  N  D  I  E  L  R  S  L  V  E  N  A  M  K  H  197

1079 TACTTTGAGCTTTTCCGGATGAAATCGTCTGCTGCCAAAGCCGATGTCTTCTTCGTCATG
     Y  F  E  L  F  R  M  K  S  S  A  A  K  A  D  V  F  F  V  M  217

1139 TCAGGGATGTGGAGAACTTCAGCAGAACGATTCTTCTTATGGATTGGCGGATTTCGACCC
     S  G  M  W  R  T  S  A  E  R  F  F  L  W  I  G  G  F  R  P  237

1199 TCCGATCTTCTCAAGGTTCTTTTGCCACATTTTGATGTCTTGACGGATCAACAACTTCTA
     S  D  L  L  K  V  L  L  P  H  F  D  V  L  T  D  Q  Q  L  L  257

1159 GATGTATGCAATCTAAAACAATCGTGTCAGCAAGCAGAAGACGCGTTGACTCAAGGTATG
     D  V  C  N  L  K  Q  S  C  Q  Q  A  E  D  A  L  T  Q  G  M  277

1219 GAGAAGCTGCAACACACCTTGCGGACCGTTGCAGCGGGACAACTCGGTGAAGGAAGTTAC
     E  K  L  Q  H  T  L  R  T  V  A  A  G  Q  L  G  E  G  S  Y  297

1379 ATTCCTCAGGTGAATTCTGCTATGGATAGATTAGAAGCTTTGGTCAGTTTCGTAAATCAG
     I  P  Q  V  N  S  A  M  D  R  L  E  A  L  V  S  F  V  N  Q  317

1439 GCTGATCACTTGAGACATGAAACATTGCAACAAATGTATCGGATATTGACAACGCGACAA
     A  D  H  L  R  H  E  T  L  Q  Q  M  Y  R  I  L  T  T  R  Q  337

1499 GCGGCTCGAGGATTATTAGCTCTTGGTGAGTATTTTCAACGGCTTAGAGCCTTGAGCTCA
     A  A  R  G  L  L  A  L  G  E  Y  F  Q  R  L  R  A  L  S  S  357

1559 AGTTGGGCAACTCGACATCGTGAACCAACGTAGGTTTGAGTTATTTTGTAACAACCAAAT
     S  W  A  T  R  H  R  E  P  T  *                              367

1619 GAAGAAAATGGAAAGACCTCAAAAATGAAGAATGAGTGCATCTGAAAACAGAGGACTACT
1679 CTGAATAAATAGAGGGGTTGCTGCTGATATTTATTTTTACTCTGCGGCGGAATTAGAAAA
1739 TTTGAAAAACATCATGTATTGATAAGTTGTAAATATCAGAAAAAGGTGGGGGTGCAAAAA
1799 TTTGTACTTTTTAGCTTTTGAAAGAGGCAAGTTT
```

Figure 1. DNA Sequence of the cDNA Encoding TGA1 and Predicted Amino Acid Sequence.

Amino acids representing the basic region are boxed; the regularly spaced leucine and glycine residues within the leucine zipper domain are circled. Numbers on the left and right correspond to the base pair and amino acid positions, respectively. Sequence data has been submitted to EMBL, GenBank, and DDBJ as accession number X68053.

interacted strongly with the Hex oligonucleotide (lane 3) derived from the wheat *histone 3* promoter. Efficient competition was observed with the homologous DNA sequences (lanes 10 to 12). Similarly, strong competition was also obtained when an unlabeled oligonucleotide carrying the as-1 sequence of the CaMV 35S promoter was included in the binding reactions

(lanes 4 to 6). In contrast, little or no competition was observed when the G-1A oligonucleotide carrying the G-box of the Arabidopsis ribulose bisphosphate carboxylase small subunit 1A (*rbcS-1A*) promoter was included (lanes 7 to 9). Hence, TGA1 interacted specifically with TGACG-related sequences. However, TGA1 differed from members of the GBF family of bZIP proteins in that it did not exhibit significant binding specificity to the G-box.

To further characterize the binding site requirements of TGA1, we employed two mutant "TGACGTGG" oligonucleotides

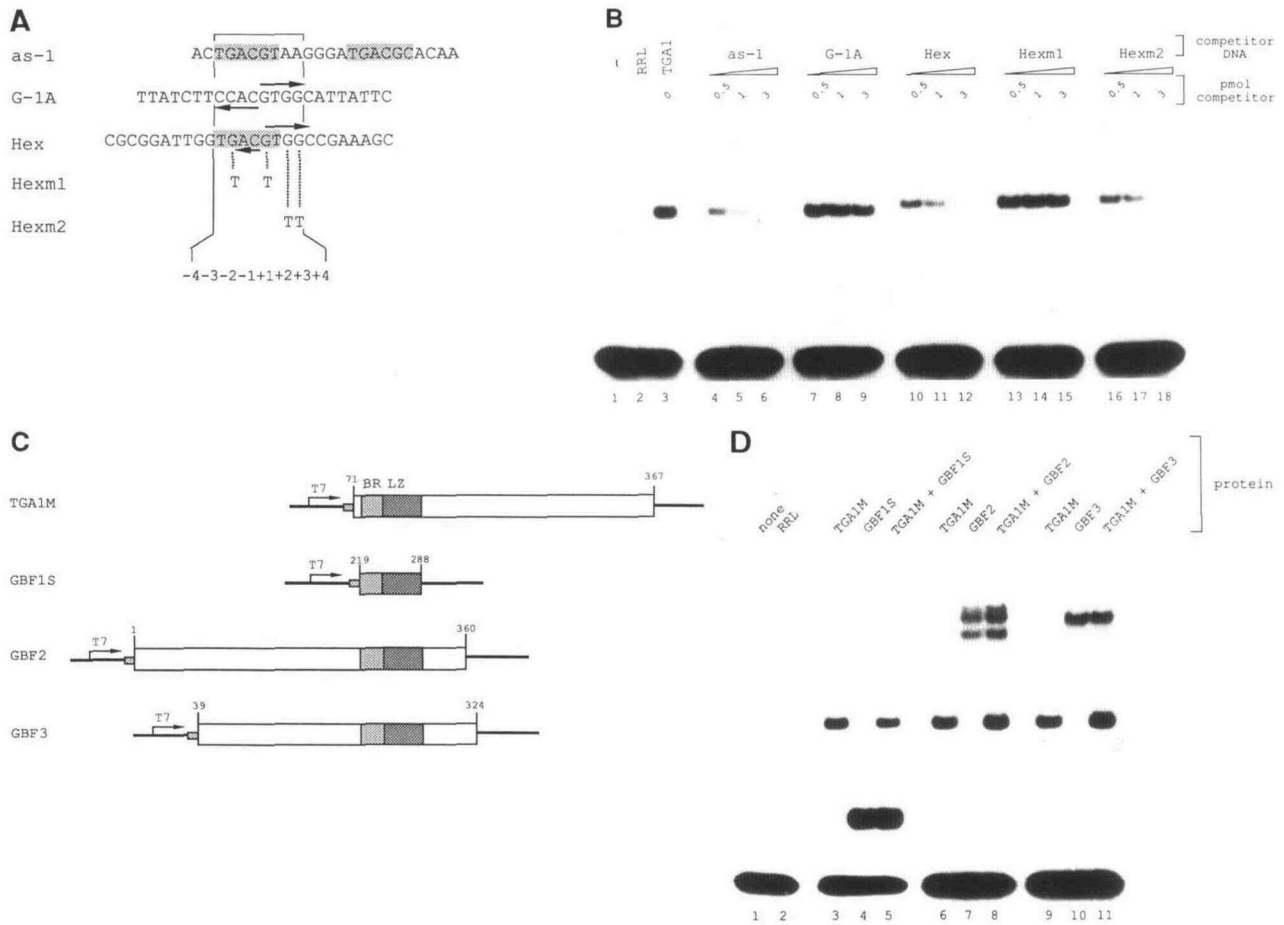

Figure 2. DNA Binding and Heterodimerization Properties of TGA1.

(A) Sequences of oligonucleotides employed in competitive DNA binding studies. as-1 is derived from the CaMV 35S promoter, G-1A carries the G-box of the Arabidopsis *rbcS-1A* promoter, and Hex contains the wheat *histone 3* hexamer. Hexm1 and Hexm2 are mutant derivatives of Hex. The sequences are aligned according to the ACGT core, and individual positions are numbered from −4 to +4. TGACG-related motifs are highlighted; the arrows mark the palindromic nature of the G-box and G-box–like motif.

(B) TGA1 interacts with TGACG-related motifs but not with the G-box. In vitro–generated TGA1 was incubated with the radiolabeled Hex oligonucleotide. Free and protein-complexed DNA fragments were separated on a 5% nondenaturing polyacrylamide gel and visualized by autoradiography. Lane 1, free DNA probe; lane 2, unprogrammed rabbit reticulocyte lysate (RRL); lanes 3 to 18, in vitro–generated TGA1 (1 μL of translation product); lane 3, no competitor DNA was added; lanes 4 to 18, various amounts of competitor DNA that are indicated above each lane were included in the binding reactions.

(C) Schematic presentation of the templates used for in vitro generation of full-length or truncated proteins employed in heterodimerization assays. The basic region (BR) and leucine zipper (LZ) domains are shaded. Numbers refer to amino acids and designate the start and end point of each in vitro translation product. The location of the T7 promoter is indicated.

(D) The Hex oligonucleotide is not recognized by heterodimeric proteins formed between TGA1 and either GBF1, GBF2, or GBF3. In vitro–generated Arabidopsis full-length and/or truncated derivatives of TGA1, GBF1, GBF2, and GBF3 were incubated with the radiolabeled Hex oligonucleotide under the same reaction conditions as described in **(B)**. Lane 1, free DNA probe; lane 2, unprogrammed rabbit reticulocyte lysate (RRL); lanes 3, 6, and 9, TGA1M; lane 4, GBF1S; lane 5, TGA1M and GBF1S (both were incubated for 30 min prior to the addition of the radiolabeled DNA); lane 7, GBF2; lane 8, same as in lane 5, except that GBF1S was substituted by GBF2; lane 10, GBF3; lane 11, same as lane 5, except that GBF1S was substituted by GBF3.

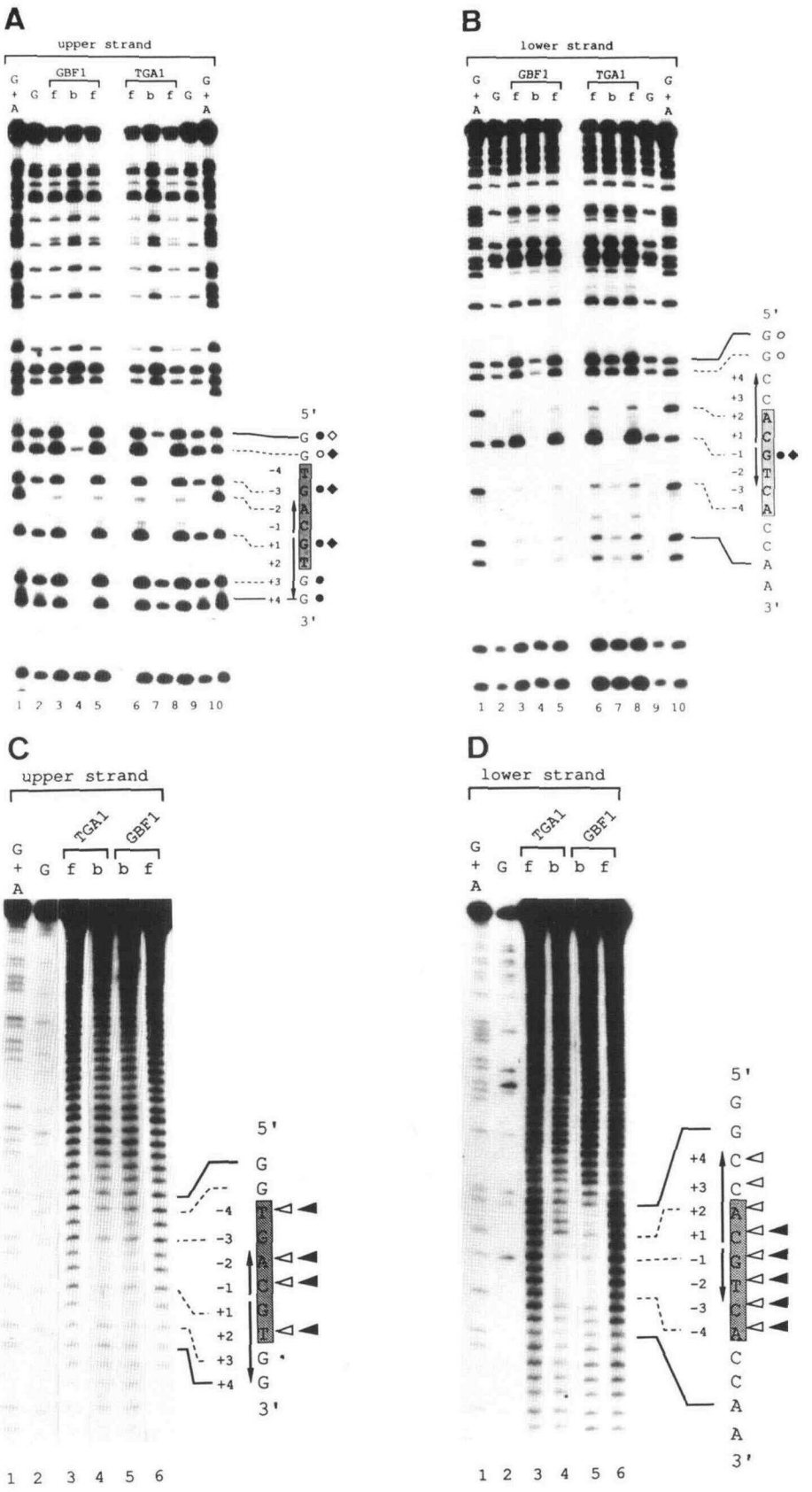

Figure 3. Binding of TGA1 and GBF1 to the Hex Oligonucleotide Can Be Discriminated on the Basis of Protein–DNA Contacts.

(Hexm1 and Hexm2; Figure 2A) in our competitive DNA binding studies. The first mutant oligonucleotide (TtACtTGG) lacks both the TGACGT motif and the GTGG sequence; this latter sequence is characteristic of the G-box. As shown in Figure 2B (lanes 13 to 15), TGA1 did not interact with this sequence. Similar results concerning the Hexm1 sequence have been observed with the Arabidopsis proteins GBF1 (Schindler et al., 1992b), GBF2, and GBF3 (U. Schindler, A. Menkens, and A. R. Cashmore, unpublished results). In contrast to the results observed with Hexm1, the second mutant oligonucleotide Hexm2 (TGACGTtt) was bound as efficiently by TGA1 as the wild-type Hex sequence (lanes 16 to 18). These results differed from those observed with the G-box–specific proteins GBF1, GBF2, and GBF3, which do not recognize this Hexm2 sequence (Schindler et al., 1992b). In summary, these data indicated that binding of TGA1 to the Hex sequence was affected by mutations within the TGACGT sequence (Hexm1). However, TGA1 binding did not require the two G residues following the TGACGT sequence (positions +3 and +4). Hence, TGA1 has different DNA binding site requirements than members of the GBF family.

TGA1 Does Not Productively Heterodimerize with Members of the GBF Family

Heterodimerization between different polypeptides can by used as one criteria to determine whether proteins belong to the same or distinct classes of bZIP proteins (Cao et al., 1991). Given that GBF1, GBF2, and GBF3 heterodimerize promiscuously, they belong to the same bZIP family or class (Schindler et al., 1992a). To determine whether TGA1 is a member of the same class, we investigated whether TGA1 would heterodimerize with members of the Arabidopsis GBF family. We used three different DNA sequences (Hex, as-1, and G-box), arguing that heterodimers, if formed, would be likely to recognize at least one of these DNA motifs. Different portions of four cDNAs—corresponding to full-length or truncated versions of GBF1, GBF2, GBF3, and TGA1 (Figure 2C)—were transcribed and translated in vitro, and the translation products were employed in DNA binding studies.

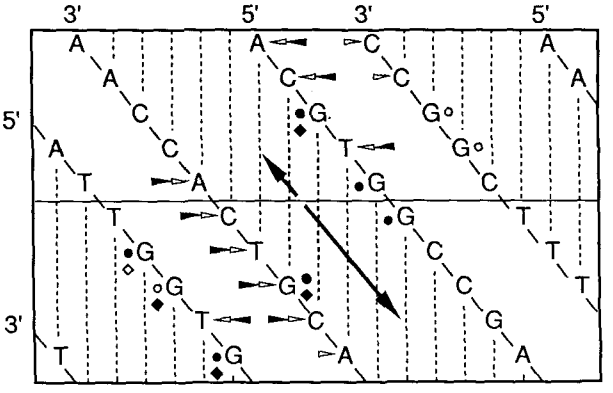

Figure 4. Summary of the Methylation Interference and Missing Nucleoside Assays Data.

The DNA double helix is displayed in a planar representation (Siebenlist and Gilbert, 1980). The dotted vertical lines represent the plane of the base pairs, the diagonal lines indicate the phosphate backbone, and the horizontal line shows the axis of the DNA. The sequence of the Hex oligonucleotide is indicated. Arrows mark the location of the nonpalindromic G-box–like sequence. N7-methylguanine residues that completely or partially inhibit TGA1 (closed and open diamonds, respectively) or GBF1 (closed and open circles, respectively) binding are shown in the major groove of the DNA double helix. Nucleosides that are essential for TGA1 or GBF1 binding are marked by closed and open arrowheads, respectively.

Figure 3. (continued).

(A) Methylation interference on the upper strand of the Hex oligonucleotide.

(B) Methylation interference on the lower strand of the Hex oligonucleotide.

The Hex oligonucleotide was radiolabeled in separate reactions on either end, partially methylated, and incubated with either in vitro–generated TGA1 or GBF1 (10 μL of in vitro–generated proteins). Free (f) and protein-complexed (b) DNA fragments were separated on 5% low ionic strength polyacrylamide gels, eluted, and cleaved with piperidine. The cleavage products were analyzed on 15% sequencing gels. Lanes 1, 2, 9, and 10 contain the Maxam-Gilbert sequencing reactions (Maxam and Gilbert, 1980); lanes 3, 5, 6, and 8, free fractions; lane 4, GBF1-complexed fractions; lane 7, TGA1-complexed fractions. The DNA sequence of the recognition sites are given; individual nucleotides are designated −4 to +4. The TGACGT motif is highlighted and the G-box–like sequence is indicated by arrows. Methylguanine residues that completely (closed circles) or partially (open circles) impair GBF1 binding are indicated. Closed and open diamonds designate guanine residues, which when methylated completely or partially inhibit TGA1 binding, respectively.

(C) Missing nucleoside analysis on the upper strand of the Hex oligonucleotide.

(D) Missing nucleoside analysis on the lower strand of the Hex oligonucleotide.

The Hex oligonucleotide was radiolabeled in separate reactions on either end and treated with iron and EDTA to partially remove individual nucleosides. The DNA was then incubated with either TGA1 or GBF1 (10 μL of in vitro–generated proteins); free (f) and protein-complexed (b) DNA fragments were separated, eluted, and analyzed on 15% sequencing gels. Lanes 1 and 2 contain the Maxam-Gilbert sequencing reactions (Maxam and Gilbert, 1980); lanes 3 and 6, free fractions; lanes 4 and 5, GBF1- and TGA1-complexed fractions, respectively. The DNA sequences of the protein binding sites are given; individual nucleotide positions are designated −4 to +4. Open and filled arrowheads indicate the nucleosides that are required for binding of TGA1 and GBF1, respectively. The TGACGT motif is highlighted, and the G-box–like sequence is indicated by arrows.

As shown in Figure 2D, the protein TGA1M, carrying the bZIP domain but missing parts of the C terminus of TGA1, was still capable of binding DNA (lane 3). Similarly, the truncated version of GBF1 bearing only the bZIP region (GBF1S) also bound efficiently to the Hex oligonucleotide (lane 4). When both proteins were synthesized separately and then incubated together prior to the addition of DNA, again the two corresponding protein–DNA complexes were observed (lane 5). Significantly, no additional protein–DNA complexes, which would have been indicative of the formation of heterodimers, were obtained. Similarly, no heterodimers were seen when the mRNAs corresponding to the two proteins were cotranslated and the products assayed for DNA binding (data not shown). The same results were also obtained when GBF2 and GBF3 were analyzed. Although both proteins efficiently recognized the Hex oligonucleotide (lanes 7 and 10), no heterodimeric complexes were observed when either of the two proteins was incubated with TGA1 prior to the addition of DNA (lanes 8 and 11) or when the mRNAs were cotranslated (data not shown). Also, no heterodimeric complexes were obtained when either of the cotranslation products involving TGA1 was assayed using a radiolabeled as-1 sequence or the G-box (data not shown).

These results indicated that no heterodimeric complexes involving Arabidopsis TGA1 and GBF1, GBF2, or GBF3 were formed that were capable of binding either the Hex motif, the G-box, or the as-1 sequence.

TGA1 Requires a Smaller Recognition Sequence than Does GBF1

TGA1 and members of the GBF family exhibited distinct DNA binding properties; however, they share the ability to bind the Hex oligonucleotide. Thus, we compared the protein–DNA contacts mediated by TGA1 and GBF1 in more detail using the Hex oligonucleotide. We performed methylation interference experiments to show whether methylation of the same or different G residues interfered with binding of both proteins. The data, illustrated in Figures 3A and 3B and summarized in Figure 4, revealed that binding of TGA1 to the Hex oligonucleotide was inhibited when the G residues at positions −3 and +1 (upper strand) and −1 (lower strand) were methylated (Figures 3A and 3B, lane 7). Similarly, methylation of the same G residues interfered with GBF1 binding (lane 4). However, in contrast to TGA1, GBF1 binding was also impaired when G residues +3 and +4 (upper strand) were methylated. These data are in agreement with our results obtained with the Hexm2 mutant oligonucleotide (Figure 2B) and our previous data which suggested that not only the hexamer sequence (TGACGT) but the entire imperfect palindrome (TGACGTGG) of the Hex oligonucleotide is required for GBF1 binding to this sequence (Schindler et al., 1992b). Binding of both GBF1 and TGA1 was impaired when the two 5′ flanking G residues in the upper strand (positions −5 and −6) were methylated. In the case of GBF1, methylguanine residues at positions +5 and +6 in the lower strand also partially inhibited protein binding. Similar results to those obtained for GBF1 were also observed for GBF2 and GBF3 (data not shown). These data document that binding to the Hex sequence of proteins belonging to the GBF family requires distinct and additional contacts than that required by TGA1.

Methylation interference experiments determine whether a methyl group at the N7 position of a G residue interferes with protein binding. However, the assay does not distinguish whether the impaired binding is due to steric hindrance or whether important interactions between the protein and corresponding G residues are disrupted. For example, the methylation interference studies could be interpreted as implicating an essential role for the G residues at positions −5 and −6 (upper strand) in the Hex oligonucleotide for the binding of TGA1.

Therefore, we delineated the contacts of both TGA1 and GBF1 to the Hex oligonucleotide more precisely using the "missing nucleoside assay" (Dixon et al., 1991). Hydroxyl radical treatment that removes individual nucleosides was performed prior to protein binding. Protein-complexed and unbound DNA fragments were separated and analyzed on denaturing polyacrylamide gels. In this assay, nucleosides required for DNA–protein interactions appear as missing bands on the autoradiograph. The results of this analysis are illustrated in Figures 3C and 3D and summarized in Figure 4. On the upper strand, the nucleosides at positions −4 (T), −2 (A), −1 (C), and +2 (T) were seen to be required for the binding of both proteins and no substantial differences were observed (Figure 3C, lanes 4 and 5). In contrast, the nucleoside requirements for both proteins on the lower strand were significantly different (Figure 3D, lanes 4 and 5). Whereas binding of both proteins required the nucleosides at positions −4 to +1, GBF1 binding also required the nucleosides at positions +2 to +4 (lane 5). In support of our competitive DNA binding assays (Figure 2B), these data demonstrated that TGA1 binding required only the sequence TGACGT, whereas GBF1 binding demanded the entire nonpalindromic G-box–like sequence TGACGTGG.

The Pentameric TGACG Motif Is Required for High-Affinity DNA Binding of TGA1

So far our results indicated either that (1) in contrast to GBF1, TGA1 required only a 5-bp recognition sequence, or (2) if there was a preference of TGA1 for additional nucleotides, this preference was yet to be demonstrated.

To distinguish these possibilities and, more generally, to determine if TGA1 exhibited any preference for certain base pair combinations at positions +2 to +4, we used the random binding site selection assay that is based on the selection of specific DNA binding sites from a pool of randomized oligonucleotides. Thirteen random base pairs were inserted into the center of a synthetic oligonucleotide pool, as shown in Figure 5. After three rounds of selection involving binding to TGA1, the bound oligonucleotides were cloned and individually subjected to DNA binding analysis (data not shown). Oligonucleotides that were

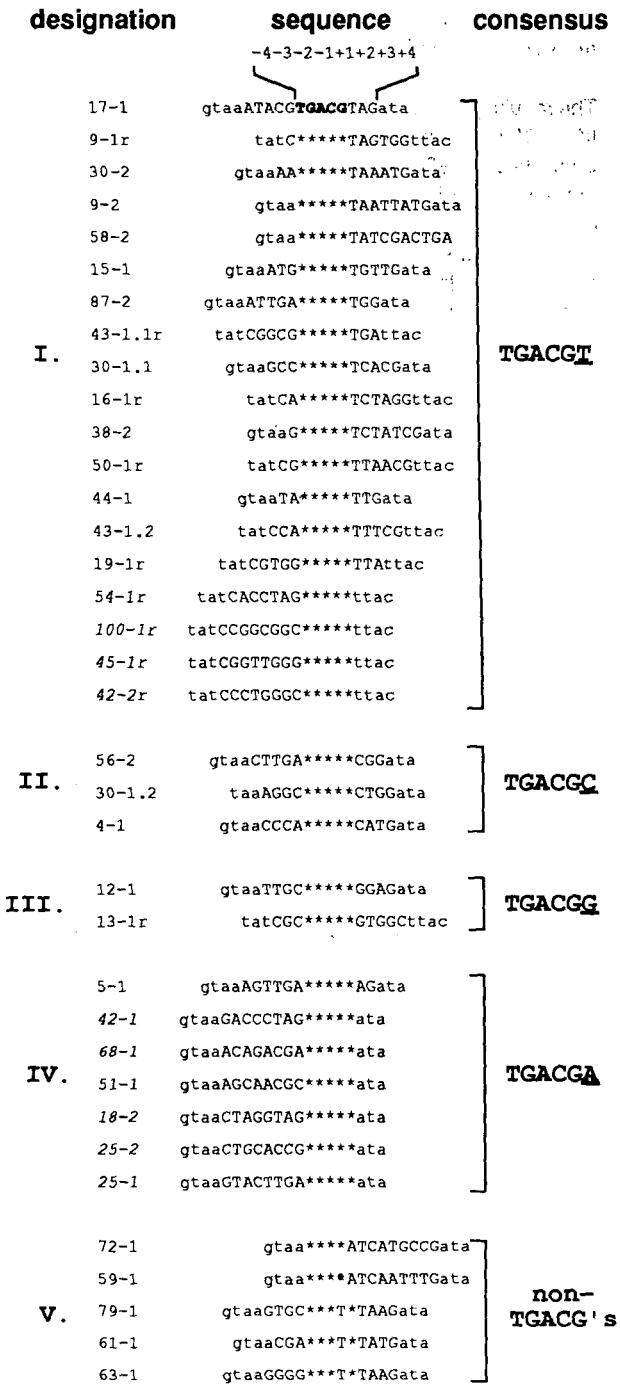

Figure 5. Preferential Binding of TGA1 to TGACG Motifs That Are Followed by a T Residue.

DNA binding sites were identified using the random binding site selection assay. The binding sites were selected from a pool of oligonucleotides carrying 13 random base pairs in the center of the sequence (5′-CGCGACGTCGGAAGACAAGCTTGTAAN$_{13}$ATAGGATCCCTCACCTCAGACAGAC-3′). The selected oligonucleotides were digested with BamHI and HindIII, ligated into pBluescript SK+, and subjected to DNA sequence analysis. The identified binding sites were grouped according to the nature of the nucleotide following the TGACG motif. Only the center portions of the oligonucleotides are shown. Nucleotides given in lowercase letters had not been randomized in the original pool. Asterisks mark the nucleotides that are identical to the sequence TGACG. Italicized designations represent those cases where nonrandom base pairs are part of the sequence extending from −4 to +4. Roman numerals indicate the five distinct groups.

bound with high affinity, and a few of the ones that were recognized with lower affinity, were then sequenced.

The identified DNA binding sites are illustrated in Figure 5. The sequences are arranged (groups I to IV) according to the nature of the nucleotides occupying position +2. All oligonucleotides that were bound with low affinity are included in group V. Significantly, all high-affinity binding sites contained an intact TGACG motif (groups I to IV), whereas all low-affinity binding sites carried one base pair substitution within the TGACG motif. Furthermore, TGA1 preferentially bound to those TGACG motifs that were followed by a T residue (position +2, group I, Figure 5); this pentamer is also present in the as-1, nos, ocs, and Hex motifs. It was noted that oligonucleotides carrying the base pair combinations ATA (majority of group IV) and TTA (parts of group I) at positions +2 to +4 are likely to be overrepresented in our compilation because these base pair combinations are derived from the nonrandomized region of the original oligonucleotide pool. These data also demonstrated that in contrast to GBF1, TGA1 binding did not require a specific base pair at position +2, and there was no indication of a preference for any particular nucleotides at positions +3 or +4, at least in the case of the TGACGT class that represents the largest number of bound oligonucleotides. This finding supports the results obtained in the missing nucleoside assays (Figures 3C and 3D), indicating that binding of TGA1 to the Hex oligonucleotide required only the sequence TGACG (−4 to +1), whereas GBF1 binding required the additional nucleotides TGG at positions +2 to +4.

DISCUSSION

TGA1 and Members of the GBF Family Differ in Their DNA Binding Properties

In this study we have described the isolation of an Arabidopsis cDNA, TGA1, encoding a bZIP protein that shows extensive sequence homology to the tobacco protein TGA1a (Katagiri et al., 1989). TGA1 bound to the Hex element of the wheat *histone 3* promoter as well as to the as-1 motif of the CaMV 35S promoter; however, in contrast to members of the GBF family, it did not bind the G-box.

Both TGA1 (this study) and GBF1 (Schindler et al., 1992b) recognize the Hex sequence. Using methylation interference studies and missing nucleoside analysis, we showed that TGA1

binding required only the TGACGT sequence of the Hex oligonucleotide. However, GBF1 binding required more extensive contacts spanning the entire G-box–like Hex motif (TGACGTGG). These results are consistent with our previous data showing that GBF1 binding to TGACGT-containing sequences was only observed when the hexamer was followed by GG or GT (positions +3 and +4; Schindler et al., 1992b). Furthermore, these observations explain why GBF1 does not recognize the as-1 element (TGACGTaa and TGACGCac) of the CaMV 35S promoter, whereas in contrast, both the Hex motif and the as-1 element are substrates for TGA1 binding. The differences in the interactions observed for TGA1 and the GBF proteins might simply reflect the differences in the basic domains of the two proteins. Alternatively, other differences between the two proteins may also affect their binding properties.

TGA1 was shown to be capable of interacting with an extensive array of sequences containing the pentamer TGACG (positions −4 to +1). Although TGA1 binding was favored when the pentamer was followed by a T residue (position +2), sequences carrying a C, G, or A residue at this position were also identified, albeit at a lower frequency. This suggests that certain combinations of sequences derived from the pentamer sequence followed by a C, G, or A residue are presumably bound with a somewhat lower affinity than those sequences containing the TGACGT motif. These data demonstrated that both TGA1 and GBF1 belong to the broad group of ACGT binding proteins (Weisshaar et al., 1991; Armstrong et al., 1992). Whereas in neither case is the tetranucleotide ACGT (positions −2 to +2) sufficient for DNA binding, the two proteins are quite distinct with respect to their requirements for additional nucleotides. TGA1 binding, in contrast to GBF1, demands the presence of the dinucleotide TG at positions −4 and −3.

Productive Heterodimerization Does Not Occur between the Two Classes of Arabidopsis bZIP Proteins

We previously demonstrated that individual members of the Arabidopsis GBF family promiscuously heterodimerize and that these heterodimeric complexes show DNA binding specificities similar to, but not necessarily identical to, the original homodimers (Schindler et al., 1992a). The results presented in this report show that TGA1 and members of the GBF family do not form heterodimeric complexes that recognize the G-box, the as-1 element, or the Hex motif. These results suggest that (1) GBF/TGA1 dimeric complexes do not bind DNA, (2) these complexes gain a new DNA binding specificity compared with the parental heterodimeric complexes, or (3) heterodimeric complexes between individual members of these protein classes are not formed per se. This inability to dimerize would presumably reflect an incompatibility of the leucine zippers that are believed to dictate dimerization specificity (Kouzarides and Ziff, 1989). Consistent with this last interpretation is the observation that cross-linking data obtained with the wheat proteins HBP-1a and HBP-1b showed that the two proteins do not heterodimerize in solution (Tabata et al., 1991).

Arabidopsis bZIP Proteins Can Be Divided into at Least Two Distinct Classes

The results presented here, together with the previously identified DNA binding site preferences and heterodimerization properties of GBF1, enable us to classify the Arabidopsis bZIP proteins into at least two classes: the GBF family and a second class exemplified by TGA1. This latter protein differs from members of the GBF family both in its DNA binding characteristics and also in its heterodimerization properties.

Additional criteria that distinguish TGA1 from members of the GBF family include differences in overall structure and the nature of the activation domains of these proteins. The DNA binding domain of TGA1 is located at the N terminus, whereas the C-terminal domain is enriched in glutamine and acidic amino acids; in the case of tobacco TGA1a, this C-terminal domain has been implicated in transcriptional activation (Katagiri et al., 1990; Yamazaki et al., 1990). In contrast, members of the GBF family are characterized by a bZIP motif at the C terminus and an N-terminal proline-rich region, which in the case of GBF1 activates transcription when fused to a heterologous DNA binding domain (Schindler et al., 1992a, 1992b).

A similar classification of bZIP proteins may be applied to other plant species. For example, the tobacco protein TAF1 (Oeda et al., 1991) exhibits DNA binding characteristics similar to members of the Arabidopsis GBF family, whereas TGA1a (Katagiri et al., 1989) belongs to the TGACG binding class (Lam et al., 1990). The wheat proteins HBP-1a and the Em binding protein EmBP-1 also have DNA binding properties that are similar to the Arabidopsis GBFs; HBP-1b on the other hand behaves like a member of the TGACG binding class (Tabata et al., 1989, 1991; Guiltinan et al., 1990). In parsley, three common plant regulatory factors (CPRF-1, CPRF-2, and CPRF-3) belonging to the bZIP class of proteins have been identified that strongly interact with the G-box (box II) of the parsley chalcone synthase promoter (Weisshaar et al., 1991). In contrast to CPRF-2, CPRF-1 and CPRF-3 exhibit very little affinity for the as-1 element (Weisshaar et al., 1991; Armstrong et al., 1992). Furthermore, CPRF-1 and CPRF-3 heterodimerize efficiently, whereas very little or no heterodimerization is observed between CPRF-2 and CPRF-3 or CPRF-2 and CPRF-1, respectively (Armstrong et al., 1992). Based on our criteria and by analogy to the situation observed in Arabidopsis, CPRF-1 and CPRF-3 fall into one class of bZIP proteins (the GBF family or the G-box/box II binding proteins). So far, no parsley proteins with binding and heterodimerization properties similar to TGA1 have been isolated. However, it appears that CPRF-2 bridges the G-box and the TGACG binding class (Armstrong et al., 1992). This suggestion is supported by the overall structure of CPRF-2 because the DNA binding domain is located in the center of the protein (Weisshaar et al., 1991).

An example of a plant bZIP protein that is distinct from the G-box and TGACG binding proteins is provided by the recently characterized bZIP protein PosF21 (Aeschbacher et al., 1991). PosF21 exhibits little sequence similarity to either one of the GBFs or to TGA1. Precedence for the existence of multiple bZIP

families is well established in other organisms (for a review, see Ziff, 1990).

Our classification of plant bZIP proteins into distinct families does not imply that members of a particular family can functionally substitute for each other. For example, the observed sequence differences within the activation domains between members of one family might reflect the possibility that these proteins may be involved in different signal transduction pathways, as observed for several transcription factor families in mammalian cells (for a review, see Karin et al., 1990; Ziff, 1990; He and Rosenfeld, 1991; Schöler, 1991). This argument is supported by the presence of the G-box in various plant promoters that are regulated by different environmental stimuli and/or in different tissues (Schulze-Lefert et al., 1989; DeLisle and Ferl, 1990; Donald and Cashmore, 1990; Guiltinan et al., 1990; Skriver et al., 1991).

METHODS

Screening of the *Arabidopsis thaliana* cDNA Library

The DNA probe used to isolate the cDNA encoding Arabidopsis TGA1 was generated as follows. Using total RNA isolated from tobacco leaf tissue as template and oligo(dT) as primer, first-strand cDNA synthesis was performed. This cDNA pool was used as template for polymerase chain reactions (PCRs). A DNA fragment encoding the basic/leucine zipper region of tobacco TGA1a was amplified using two gene-specific primers (5′-primer, CCGGgatatcGTAAACCCGTCGAGAAGGTACTTAGACGT; 3′-primer, TAGCggatccAGAGTAACTTAGCTGGCTAGCATCTAC; small letters indicate nucleotide changes that had been inserted to create sites for restriction endonucleases). The following conditions were used for the amplification reaction: 1 min at 94°C, 1 min at 55°C, and 1 min at 72°C, 30 cycles. The PCR products were digested with StuI, and the smaller DNA fragment encoding the basic region was gel purified and radiolabeled using random hexamers (Feinberg and Vogelstein, 1984). The radiolabeled probe was used for screening an Arabidopsis (ecotype Columbia) cDNA library (Schindler et al., 1992a). The filters were prehybridized in 30% formamide, 5 × Denhardt's (1 × Denhardt's is 0.02% each Ficoll, polyvinylpyrolidone, BSA), 5 × SSPE (1 × SSPE is 0.15 M NaCl, 10 mM sodium phosphate, 1 mM EDTA, pH 7.4), 10% dextran sulfate, 1% SDS, 10 μg/mL salmon sperm DNA for 14 hr at 42°C. The hybridization was performed for 24 hr under the same conditions after adding the radiolabeled probe (5 × 10^6 cpm/mL). The filters were washed twice for 30 min in 30% formamide, 0.5% SDS, 5 × SSPE at 42°C, and once for 30 min in 3 × SSPE at 42°C.

DNA Sequence Analysis

Double-stranded DNA was used for sequence analysis employing the dideoxy chain termination reaction (Sanger et al., 1977) and suitable subclones of the cDNA insert or gene-specific internal primers.

Plasmids

All plasmids were constructed by using standard techniques (Sambrook et al., 1989). The generation of the proteins GBF1S, GBF2, and GBF3 was described previously (Schindler et al., 1992a). In vivo excision, yielding the recombinant plasmid pTGA1, was performed as recommended by the manufacturer (Stratagene). Templates used for in vitro transcription and translation reactions were generated as described by Schindler et al. (1992a). N- and C-terminal end points of the translation products are specified in the text.

Preparation of Radiolabeled Probes

All oligonucleotides were cloned into the BamHI and BglII sites of pBgl (Donald et al., 1990). The oligonucleotides were excised from the vector, radiolabeled by filling in the 5′ overhangs with α-^{32}P-dATP and the Klenow fragment of DNA polymerase I, and gel purified. For methylation interference and missing nucleoside experiments, the Hex oligonucleotide was radiolabeled at the BamHI or BglII site and released with KpnI or SacI, respectively.

In Vitro Transcription and Translation, Mobility Shift Assays, and Methylation Interference Experiments

The assays were performed as described previously (Schindler and Cashmore, 1990; Schindler et al., 1992a). Competitive binding assays were performed in the presence of specific competitor DNAs as indicated in the legend to Figure 2. Formation of heterodimers was investigated by incubating two different in vitro translation products for 30 min at room temperature prior to the addition of the radiolabeled DNA binding site.

Missing Nucleoside Analysis

Experiments were carried out essentially as described by Dixon et al. (1991). Briefly, DNA (5 × 10^5 cpm) was treated in a final volume of 20 μL containing 0.25 mM Fe(II), 10 mM EDTA, 0.015% H_2O_2, and 10 mM ascorbic acid. After 2 min at room temperature, the reaction was terminated and the DNA was precipitated. Band shift assays employing the modified DNA were carried out as described for methylation interference experiments (Schindler and Cashmore, 1990).

Random Binding Site Selection Assay

Random binding site selection assays were carried out essentially as described by Schindler et al. (1992b) with the following modifications. In vitro–generated TGA1 protein was used and partially purified as follows: four in vitro translation reactions were loaded onto a 400-μL Q-Sepharose column; fractions containing TGA1 binding activity were dialyzed against buffer D (10 mM Tris-HCl, pH 7.5, 40 mM NaCl, 1 mM EDTA, 10% glycerol, 1 mM DTT) and used in the random binding site selection assays employing synthetic oligonucleotides with 13 random base pairs inserted in the center. After three rounds of selection, subsequent steps were performed as described by Schindler et al. (1992b). For each round of selection, the oligonucleotides were radiolabeled using 10 PCR cycles under the following conditions: 1 min at 94°C, 1 min at 55°C, and 40 sec at 72°C. Reactions were carried out in a 20 μL volume according to the manufacturer's specifications (Perkin-Elmer Cetus), except that dCTP was replaced by 10 μL (3000 Ci/mmol) of α-^{32}P-dCTP.

ACKNOWLEDGMENTS

U.S. and H.B. made equal contributions to this manuscript. We thank Margaret Ahmad for synthesizing the oligonucleotides and critical reading of the manuscript. We are also grateful to Anne Menkens for useful discussions and to Joseph R. Ecker and John Chan for helpful comments concerning the manuscript. This work was supported by National Institutes of Health Grant No. GM38409 to A.R.C.

Received May 14, 1992; accepted August 10, 1992.

REFERENCES

Aeschbacher, R.A., Schrott, M., Protrykus, I., and Saul, M.W. (1991). Isolation and molecular characterization of *PosF21*, an *Arabidopsis thaliana* gene which shows characteristics of a b-Zip class transcription factor. Plant J. **1,** 303–316.

Armstrong, G.A., Weisshaar, B., and Hahlbrock, K. (1992). Homodimeric and heterodimeric leucine zipper proteins and nuclear factors from parsley recognize diverse promoter elements with ACGT cores. Plant Cell **4,** 525–537.

Bouchez, D., Tokuhisa, J.G., Llewellyn, D.J., Dennis, E.S., and Ellis, J.G. (1989). The ocs-element is a component of the promoters of several T-DNA and plant viral genes. EMBO J. **8,** 4197–4204.

Cao, Z., Umek, R.M., and McKnight, S.L. (1991). Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. Genes Dev. **5,** 1538–1552.

DeLisle, A.J., and Ferl, R.J. (1990). Characterization of the *Arabidopsis Adh* G-box binding factor. Plant Cell **2,** 547–557.

Dixon, W.J., Hayes, J.J., Levin, J.R., Weidner, M.F., Dombroski, B.A., and Tullius, T.D. (1991). Hydroxyl radical footprinting. Methods Enzymol. **208,** 380–413.

Donald, R.G.K., and Cashmore, A.R. (1990). Mutation of either G box or I box sequences profoundly affects expression from the *Arabidopsis rbcS-1a* promoter. EMBO J. **9,** 1717–1726.

Donald, R.G.K., Schindler, U., Batschauer, A., and Cashmore, A.R. (1990). The plant G-box promoter sequence activates transcription in *Saccharomyces cerevisiae* and is bound *in vitro* by a yeast activity similar to GBF, the plant G-box binding factor. EMBO J. **9,** 1727–1735.

Feinberg, A., and Vogelstein, B. (1984). A technique for radiolabelling DNA-restriction endonuclease fragments to high specific activity. Anal. Biochem. **137,** 266–267.

Fromm, H., Katagiri, F., and Chua, N.-H. (1991). The tobacco transcription activator TGA1a binds to a sequence in the 5′ upstream region of a gene encoding a TGA1a-related protein. Mol. Gen. Genet. **229,** 181–188.

Guiltinan, M.J., Marcotte, W.R., and Quatrano, R.S. (1990). A plant leucine zipper protein that recognizes an abscisic acid response element. Science **250,** 267–271.

He, X., and Rosenfeld, M.G. (1991). Mechanisms of complex transcriptional regulation: Implications for brain development. Neuron **7,** 183–196.

Hu, J.C., O'Shea, E.K., Kim, P.S., and Sauer, R.T. (1990). Sequence requirements of coiled-coils: Analysis with λ repressor-GCN4 leucine zipper fusions. Science **250,** 1400–1403.

Karin, M., Castrillo, J.-L., and Theill, L.E. (1990). Growth hormone gene regulation: A paradigm for cell-type–specific gene activation. Trends Genet. **6,** 92–96.

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

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

Kawata, T., Imada, T., Shiraishi, H., Okada, K., Shimura, Y., and Iwabuchi, M. (1992). A cDNA clone encoding HBP-1b homologue in *Arabidopsis thaliana*. Nucl. Acids Res. **20,** 1141.

Kouzarides, T., and Ziff, E. (1989). Leucine zippers of *fos*, *jun* and GCN4 dictate dimerization specificity and thereby control DNA binding. Nature **340,** 568–571.

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

Landschulz, W.H., Johnson, P.F., and McKnight, S.L. (1988). The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. Science **240,** 1759–1764.

Maxam, A., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65,** 499–560.

McKnight, S.L. (1991). Molecular zippers in gene regulation. Sci. Am. **264,** 54–64.

Oeda, K., Salinas, J., and Chua, N.-H. (1991). A tobacco bZip transcription activator (TAF-1) binds to a G-box–like motif conserved in plant genes. EMBO J. **10,** 1793–1802.

O'Shea, E.K., Rutkowski, R., and Kim, P.S. (1989). Evidence that the leucine zipper is a coiled coil. Science **243,** 538–542.

O'Shea, E.K., Klemm, J.D., Kim, P.S., and Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science **254,** 539–544.

Rasmussen, R., Benvegnu, D., O'Shea, E., Kim, P.S., and Alber, T. (1991). X-ray scattering indicates that the leucine zipper is a coiled coil. Proc. Natl. Acad. Sci. USA **88,** 561–564.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74,** 5463–5467.

Schindler, U., and Cashmore, A.R. (1990). Photoregulated gene expression may involve ubiquitous DNA binding proteins. EMBO J. **9,** 3415–3427.

Schindler, U., Menkens, A.E., Beckmann, H., Ecker, J.R., and Cashmore, A.R. (1992a). Heterodimerization between light-regulated and ubiquitously expressed *Arabidopsis* GBF bZIP proteins. EMBO J. **11,** 1261–1273.

Schindler, U., Terzaghi, W., Beckmann, H., Kadesch, T., and Cashmore, A.R. (1992b). DNA binding site preferences and transcriptional activation properties of the *Arabidopsis* transcription factor GBF1. EMBO J. **11,** 1274–1289.

Schmidt, R.J., Ketudat, M., Aukerman, M.J., and Hoschek, G. (1992). Opaque-2 is a transcriptional activator that recognizes a specific target site in 22-kD zein genes. Plant Cell **4,** 689–700.

Schöler, H.R. (1991). Octamania: The POU factors in murine development. Trends Genet. **7,** 323–328.

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

Siebenlist, U., and Gilbert, W. (1980). Contacts between *Escherichia coli* RNA polymerase and an early promoter of phage T7. Proc. Natl. Acad. Sci. USA **77,** 122–126.

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

Skriver, K., Olsen, F.L., Rogers, J.C., and Mundy, J. (1991). *cis*-Acting DNA elements responsive to gibberellin and its antagonist abscisic acid. Proc. Natl. Acad. Sci. USA **88,** 7266–7270.

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

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

Ueda, T., Waverczak, W., Ward, K., Sher, N., Ketudat, M., Schmidt, R.J., and Messing, J. (1992). Mutations of the 22- and 27-kD zein promoters affect transactivation by the Opaque-2 protein. Plant Cell **4,** 701–709.

Vinson, C.R., Sigler, P.B., and McKnight, S.L. (1989). Scissors-grip model for DNA recognition by a family of leucine zipper proteins. Science **246,** 911–916.

Weisshaar, B., Armstrong, G.A., Block, A., da Costa e Silva, O., and Hahlbrock, K. (1991). Light-inducible and constitutively expressed DNA-binding proteins recognizing a plant promoter element with functional relevance in light responsiveness. EMBO J. **10,** 1777–1786.

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

Ziff, E.B. (1990). Transcription factors: A new family gathers at the cAMP response site. Trends Genet. **6,** 69–72.