Two Genes Encoding GF14 (14–3–3) Proteins in Zea mays¹

Structure, Expression, and Potential Regulation by the G-Box-Binding Complex

Nick C. de Vetten² and Robert J. Ferl*

Department of Horticultural Sciences, Program in Plant Molecular and Cellular Biology, 1243 Fifield Hall, University of Florida, Gainesville, Florida 32611

Two maize (Zea mays) genes, designated GRF1 and GRF2, have been isolated and characterized. The proteins encoded by these genes, called GF14 proteins, participate in protein/DNA complexes and show more than 60% identity with a highly conserved, widely distributed protein family, collectively referred to as 14-3-3 proteins. Members of the 14-3-3 protein family have been reported to activate Tyr and Trp hydroxylases, modulate protein kinase C activity, and activate ADP-ribosyltransferase. The mRNAs of the GRF genes are encoded by six exons interrupted by five introns. The transcriptional units of the GRF genes were found to be very similar, with complete conservation of the intron positions. In addition, the length and nucleotide sequences of the two genes' introns were highly conserved. The 5' flanking sequences of the two GRF genes were compared and regions of homology and divergence identified. This comparison revealed the presence of a conserved G-box element in the 5' flanking region of both genes. Electrophoretic mobility shift assays of maize protein extract with the GRF G-box indicates that GBF binds to this G-box site in the 5' upstream region of GRF. Antibody supershifts indicate that GF14 protein is associated with the G-box-binding complex that interacts with the GRF upstream region.

A focal point in the field of eukaryotic gene regulation is the understanding of the mechanisms by which transcription is controlled. The prevailing view is that this regulation is mediated, in part, by an interplay between distinct DNA sequence elements found in the promoter region of a gene and sequence-specific DNA-binding proteins. Recently, in plant systems there has been a major effort to identify DNAbinding proteins specifically interacting with their cognate promoter sites and to elucidate how the binding of such proteins results in increased or decreased transcription of the associated gene (Katagiri and Chua, 1992; Brunelle and Chua, 1993). Studies of plant DNA-binding proteins are most advanced for the GBF and the TGA1a proteins (reviewed by Katagiri and Chua, 1992; de Vetten and Ferl, 1994). The G-box is a conserved *cis*-acting promoter element important for induction of a number of environmentally inducible genes (Schulze-Lefert et al., 1989; Donald and Cashmore, 1990; Guiltinan et al., 1990). Several GBFs have been isolated from a number of plant species using the Gbox as probe (Guiltinan et al., 1990; Oeda et al., 1991; Weisshaar et al., 1991; Schindler et al., 1992a). They all belong to the basic Leu zipper class of DNA-binding proteins, with a basic amino acid domain abutting a Leu repeat (Johnson and McKnight, 1989). Considerable conservation (90–95% identity) among the isolated GBFs was found exclusively within the basic domain, the region that is thought to come in direct contact with DNA. Recent studies have shown that the N-terminal Pro-rich domain of the *Arabidopsis* GBF-1 can activate transcription in vivo (Schindler et al., 1992b).

Although transcriptional control is executed by DNAbinding proteins, it is apparent that transcription is also regulated by additional regulatory molecules that are associated with the DNA-binding protein. The formation of these multicomponent complexes apparently allows the cell to respond to the complexity of the temporally regulated processes that must occur for growth and development. These complexes may include, in addition to the DNA-binding protein, the general transcription factors, and RNA polymerase II, transcription activators, co-activators, adaptors, and various accessory proteins (Martin, 1991).

In an attempt to identify proteins that are part of the Gbox-binding complex, we have partially purified GBF activity and generated monoclonal antibodies to the DNA-binding complex (de Vetten et al., 1992; Lu et al., 1992). Using this approach, we identified a protein designated G-box Factor 14-3-3 homolog (GF14) as being associated with the DNAprotein complex. Evidence for this was provided by EMSAs showing that anti-GF14 monoclonal antibodies are able to shift the G-box-protein complex to a slower mobility and by co-immunoprecipitation demonstrating the association of GF14 with a 46-kD GBF (de Vetten et al., 1992).

GF14 has extensive homology with a class of protein kinase-dependent regulatory proteins (Aitken et al., 1992). This class of proteins, known as 14–3–3 proteins, has recently received much attention by investigators studying a broad

¹ This work is contribution No. R-03546 from the Florida Agricultural Experiment Station and was supported by National Institutes of Health grant RO1 GM 40061 and U.S. Department of Agriculture-National Research Initiative grant 93-37304-9608 to R.J.F.

² Present address: Department of Genetics, Free University, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands.

^{*} Corresponding author; fax 1-904-392-4072.

Abbreviations: EMSA, electrophoretic mobility shift assay; GBF, G-box-binding factor; PKC, protein kinase C; TGA1a, tobacco transcription activator.

range of biological systems and has been implicated in many diverse physiological functions. The 14-3-3 proteins are a family of acidic proteins that have molecular masses of about 30 kD and are abundantly present in mammalian brain tissue. After the original identification (Moore and Perez, 1967), Ichimura et al. (1987) showed that purified 14-3-3 proteins are able to activate Tyr and Trp hydroxylase in the presence of type II calmodulin-dependent kinase. Independently, Aitken et al. (1991) isolated polypeptides with homology (approximately 70% identical residues) to the 14-3-3 proteins that were found to be potent inhibitors of phospholipid, Ca²⁺-dependent PKC. In contrast to this, the 14-3-3 protein that was found to activate Tyr and Trp hydroxylases stimulates PKC activity (Isobe et al., 1992). This raises the possibility that different members of the 14-3-3 family have distinct cellular activities. A member of the 14-3-3 family was also identified as being a stimulator of Ca2+-dependent exocytosis in permeabilized adrenal chromaffin cells (Morgan and Burgoyne, 1992). It has been suggested that this stimulation of exocytosis is potentiated by PKC activation and, therefore, the 14-3-3 protein may influence the PKCmediated control of Ca2+-dependent exocytosis (Morgan and Burgoyne, 1992).

Although much work remains to elucidate the exact physiological role of the 14–3–3 proteins, it seems likely that these proteins have a key role as regulators in the signal transduction pathways. This regulatory function of the 14–3–3 protein is further substantiated by gene disruption and overexpression studies in yeast (Van Heusden et al., 1992).

In addition to maize (Zea mays) and Arabidopsis GF14, cDNAs for 14-3-3 protein homologs have been isolated from Oenothera, spinach, barley, and rice (Brandt et al., 1992; Hirsch et al., 1992; Kidou et al., 1993). Studies of mammalian systems demonstrated that the 14-3-3 protein exists in multiple molecular forms, as many as seven or eight isoforms, which probably arise from different genes (Isobe et al., 1991; Toker et al., 1992). Here we report the isolation of two maize GF14 cDNA clones using the anti-GF14 monoclonal antibodies. To study their regulation we isolated genomic clones encoding the corresponding members of the GF14 gene family. The 5' flanking sequences of these genes have been compared and putative regulatory elements identified, one of which is a G-box element. EMSA demonstrates that partially purified maize GBF activity or maize GBF produced in Escherichia coli binds to the G-box sequence in the 5' upstream region of the GF14 genes.

MATERIALS AND METHODS

Plant Material

Mature maize (*Zea mays* line XL80) plants were grown under greenhouse conditions. Most experiments were conducted with maize suspension culture cell line P3377 derived from immature embryo, which originated from Dr. J. Widholm (University of Illinois; Duncan et al., 1985). Cell suspensions were maintained on a commercial mixture of Murashige and Skoog salts (Gibco–BRL) with 2,4-D at a level of 2 mg/L.

Isolation of cDNA Clones

The isolation of the GF14 cDNA clones using the anti-GF14 monoclonal antibody cell supernatant was as described previously (de Vetten et al., 1992).

Isolation of Genomic Clones

A maize genomic library was purchased from Clontech (Palo Alto, CA). Approximately 600,000 plaques were plated on *Escherichia coli* K803 and screened at high stringency with ³²P-labeled pGF14–12 cDNA. The hybridization was performed at 65°C for 18 h in 0.5 M sodium phosphate (pH 7.2), 1 mM EDTA, 7% SDS, and 1% BSA. The filters were washed for 1 h with three changes of 40 mM sodium phosphate (pH 7.2), 24 mM NaCl, 0.1% SDS, and 1 mM EDTA at 65°C and exposed to Kodak XAR-5 film with intensifying screen. Hybridizing recombinant phage were rescreened at lower plaque density to plaque purity. Phage particles were isolated, and DNA was purified according to the method of Sambrook et al. (1989).

Subcloning and Sequencing

cDNA inserts and genomic DNA fragments were isolated from λ -DNA by agarose gel electrophoresis. The purified fragments were subcloned into pUC18. The nucleotide sequence of both strands of the plasmid DNA was determined by automated dideoxy chain termination on an ABI 373 (Applied Biosystems, Foster City, CA) according to the manufacturer's protocols. Nucleotide sequence analysis was performed using SeqEd (Applied Biosystems). Nucleotide and amino acid sequence alignments were carried out using GeneWorks (IntelliGenetics, Mountain View, CA).

Protein Expression in E. coli

The expression vector pET15b (Novagen, Madison, WI) was used to express both of the GF14 cDNA clores in *E. coli* cells. The pET15b carries an N-terminal leader sequence of a six consecutive His residue that can be cleaved by thrombin (see "Results"). Oligonucleotide-directed mutagenesis was used to create an *NdeI* site at the initiating ATG of the pGF14–6 and pGF14–12 coding sequence and a *Bam*HI site 3' of the stop codon. The plasmids pET-GF14–6 and pET-GF14–12 were constructed by subcloning the 800-bp *NdeI-Bam*HI fragment of pGF14–6 and pGF14–12 into pET15b.

GF14 proteins were expressed in the bacterial strain BL21 (DE3). Cells transformed with the expression plasmids were grown at 37°C in Luria broth with 100 μ g/mL ampicillin. The cells were grown to an A_{600} of 0.5. Isopropylthio- β -galactoside was then added to the culture to a final concentration of 1 mm, and cells were grown for an additional 3 h.

Induced cells were harvested by centrifugation and resuspended in binding buffer (20 mM Tris-HCl [pH 7.9], 0.5 M NaCl) containing 5 mM imidazole. Cells were frozen and then thawed on ice. Lysozyme was added to a final concentration of 0.2 mg/mL, and the cell suspension was incubated on ice for 20 min. The lysate was centrifuged at 39,000g^c for 20 min to remove debris. Soluble protein was loaded on an iminodiacetic acid-agarose column equilibrated with 50 mM NiSO₄ and extensively washed with binding buffer. The column was then washed with binding buffer containing 60 mM imidazole, and proteins were eluted with binding buffer containing 1 M imidazole. The presence of the 30-kD GF14 protein in the eluate fraction was confirmed by SDS-PAGE and immunoblotting. The GF14 fusion protein was dialyzed against 20 mM Tris-HCl (pH 8.4), 150 mM KCl, and 2.5 mM CaCl₂ and digested overnight with 2 units of thrombin at room temperature to remove the His leader sequence (see "Results"). The GF14 recombinant protein after cleavage was passed for a second time over the NiSO₄-charged iminodiacetic acid-agarose column to remove the His leader sequence and metal affinity proteins. The GF14, which was not retained on the second metal affinity column, was purified to at least 95% homogeneity.

Immunoblotting Analysis

Total maize protein was extracted as described previously (de Vetten et al., 1992). Proteins were separated on a 12.5% SDS-polyacrylamide gel and transferred to nitrocellulose filters (Schleicher & Schuell) according to the method of Towbin et al. (1979). Filters were blocked overnight in 10% nonfat dry milk, washed, and incubated with a 1:2000 dilution of anti-GF14 monoclonal antibody ascites for 1 h. After three 10-min washes with 50 mм Tris-HCl, pH 7.5, 150 mм sodium chloride, 0.1% Tween 20, the filters were incubated with a 1:2000 dilution of sheep anti-mouse horseradish peroxidaseconjugated IgG (Amersham). After washing the filters with 50 mm Tris-HCl, pH 7.5, 150 mm sodium chloride, 0.1% Tween 20, immunodetection was performed with an enhanced chemoluminescence western blotting detection reagent (Amersham). Autoradiograms were scanned in a computing densitometer (Molecular Dynamics, Mountain View, CA).

EMSA

EMSAs were performed essentially as described by McKendree et al. (1990). Maize whole-cell extract was enriched for GBF activity by heparin-agarose (Bio-Rad) and fast protein liquid chromatography Mono-Q (Pharmacia LKB Biotechnology Inc.) as described by DeLisle and Ferl (1990). E. coli-expressed maize GBF (N.C. de Vetten and R.J. Ferl, unpublished data) was isolated and purified as described above. The standard assay mixture consisted of 5 µL of enriched maize extract or 4 ng of purified recombinant GBF, 2 µg of tRNA, 1 µg of poly(dI-dC), 50 ng of competitor DNA, and 0.5 ng of end-labeled G-box oligonucleotide. The samples were incubated for 10 min on ice before electrophoresis on a 5% nondenaturing polyacrylamide gel. In some experiments, 2 µL of anti-GF14 antibody cell supernatant were added to the maize GBF-enriched protein extract and incubated for 1 h before addition of end-labeled probe and electrophoresis.

RESULTS

Isolation and Characterization of GF14 Genes

Previously, we described the isolation of a partial cDNA clone designated pGF14-12, representing a transcript encod-

ing a GF14 protein associated with the G-box-binding complex (de Vetten et al., 1992). Five other clones were isolated with anti-GF14 monoclonal antiserum, four of which are clones overlapping pGF14–12. One of these cDNAs extended the 5' region of the previously reported pGF14–12 sequence by about 100 bp. The fifth clone is closely related to pGF14– 12 and was termed pGF14–6.

The complete nucleotide and deduced amino acid sequences of pGF14-6 and pGF14-12 were recovered. Their nucleotide sequences (1136 and 1175 bp, respectively) are 94% identical. Each cDNA contains one large open reading frame comprising 261 amino acid residues, which is preceded by a typical Kozak initiation sequence (AAGATGG) and is terminated by a TAA stop codon (Kozak, 1984). The deduced amino acid sequences of the two cDNAs are 98% identical, with only four amino acid substitutions of 261 amino acid residues. The calculated molecular sizes of the GF14-6 and GF14-12 proteins are 29,666 and 29,640 D, respectively, which is similar to the size of the immunoreactive protein detected by immunoblotting (de Vetten et al., 1992). Both the pGF14-6 and pGF14-12 polypeptides have a putative protein kinase A and PKC recognition site and a motif resembling a hydrophobic heptad repeat of Leu and Ile (de Vetten et al., 1992). The polypeptides are very acidic with an overall negative charge of 16 at neutral pH, which is concentrated predominantly at the N and C termini of the protein. The calculated isoelectric point of the polypeptides of 4.6 is similar to that reported for GF14 homologous proteins (Nielsen, 1991; Brandt et al., 1992; Zupan et al., 1992).

Sequence Comparisons of the 14-3-3 Protein Homologs

A computer search for related protein sequences revealed a definite similarity between maize GF14 and a class of proteins isolated predominantly from mammalian brain termed 14-3-3 proteins (Moore and Perez, 1967; Ichimura et al., 1988). Figure 1 shows an amino acid alignment of the two maize GF14 proteins and a selected number of 14-3-3 homologs isolated from plants, mammals, and yeast. Overall identity among the proteins is 47% and is concentrated in highly conserved motifs. Some of these conserved motifs have been implicated in a possible mechanism of PKC regulation and are indicated by stippling in Figure 1. The first motif contains the conserved sequence GARR, reminiscent of part of the pseudosubstrate domain of PKC (Aitken et al., 1992). This pseudosubstrate domain may account for the PKC inhibitory activity of the 14-3-3 proteins. A second motif is similar to a domain of annexins, a family of Ca²⁺and lipid-binding proteins (Aitken et al., 1992). This site may be a potential binding site for the regulatory domain of PKC, thereby inhibiting its activity. A third motif shows similarity with the EF-hand calcium-binding motif in domain III of calmodulin (Babu et al., 1985). Recently, Lu et al. (1994b) demonstrated that the C-terminal domain of the Arabidopsis GF14, containing this potential EF-hand motif, binds calcium. Since the motifs mentioned above are highly conserved among different species, they could play a similar important role in mammals, higher plants, and yeast.

A gene genealogical tree of full-length protein sequences of the GF14/14-3-3 gene family isolated to date is shown in

Maize GF14-6 Maize GF14-12 Arabidopsis GF14 Yeast Bovine 14-3-3 Sheep KCIP PKC pseudosub	MASAELSREENVIMAKLAEQAERVEEMVEEMVEFMENVAKTVDSEEITVEEMNLLSVAYKNVIG MASAELSREENVIMAKLAEQAERVEEMVEFMENVAKTVDSEEITVEEMNLLSVAYKNVIG MASGREEIVIMAKLAEQAERVEEMVEFMENVSAAVDGIEITVEEMNLLSVAYKNVIG MSTSRED-VYLAKLAEQAERVEEMVENMKTVASSGQEISVEEMNLLSVAYKNVIG MDKSQIVQKAKLAEQAERVDIMAAMKAVTEQGHELSNEEHDLLSVAYKNVVG MDDREDIVYQAKLAEQAERYDEMVESMKNVAGMDVEITVEEHNLLSVAYKNVVG RKG	60 60 57 54 53 54
Maize GF14-6 Maize GF14-12 Arabidopsis GF14 Yeast Bovine 14-3-3 Sheep KCIP PKC pseudosub	ARRASWRI ISSIEQKEEGRG-NEDRVTLIKDYRGKIEHELTKILDGILKILETHIVPSST ARRASWRI ISSIEQKEEGRG-NEDRVTLIKDYRGKIEHELTKILDGILKILESHIVPSST ARRASWRI ISSIEQKEEGRG-NDDHVTATREYRSKIEHELSGILDGILKILDSRUIPAAA ARRASWRI VSSIEQKEESKEKSEHQVELICSYRSKIEHELTKI SDDILSVLDSRUIPAAA ARRASWRI VSSIEQKEESKEKSEHQVELICSYRSKIEHELTKI SDDILSVLDSHUIPSAT ARRASWRI VSSIEQKEENKGG-EDKLKMIREYROMVEHELKILCCDILDVLDKHUIPAAN ARRASWRI ISSIEQKEENKGG-EDKLKMIREYROMVEHELKILCCDILDVLDKHUIPAAN ALROK	119 119 116 114 110 113
Maize GF14-6 Maize GF14-12 Arabidopsis GF14 Yeast Bovine 14-3-3 Sheep KCIP Annexin V	APESKVFYLKMKGDYYRYTAE FKTGAERKDAAENTMVAYKAADDTALAELAFTHPIRLGL APESKVFYLKMKGDYYRYTAE FKTGAERKDAAENTMVAYKAADDTALAELAFTHPIRLGL SGDSKVFYLKMKGDYHRYTAE FKTGDERKDAAEHTLAAYKSADDTANAELAFTHPIRLGL TGESKVFYTKMKGDYHRYTAE FKTGDERKDAAEHTLAAYKSADDTANAELAFTHPIRLGL QPESKVFYLKMKGDYHRYTAE FSGGDAREKATNASLEAYKTASETATTDLPFTHPIRLGL QPESKVFYLKMKGDYFRYTSEVASGDNKQTTVSNSQQAYDEAFETSKKEMQPTHPIRLGL TGE <u>SKVFYLKMKGDYHRYTAEFATC</u> NDRKEAAENSLV <u>AY</u> KAASCTAMTELP <u>PTHPIRLGT</u> GDYKKALLLL.CGEDD	179 179 176 174 170 173
Maize GF14-6 Maize GF14-12 Arabidopsis GF14 Yeast Bovine 14-3-3 Sheep KCIP Calmodulin III	ALNFSVFYYEILNSPDFACSLAK DAFDFAISELDTLSEESYKDSTLIMQLLADNLTLWTS ALNFSVFYYEILNSPDFACSLAK DAFDFAISELDTLSEESYKDSTLIMQLLADNLTLWTS ALNFSVFYYEILNSPDFACKLAK DAFDFAIAELDTLSEESYKDSTLIMQLLKDNLTLWTS ALNFSVFYYEILNSPDFACKLAK DAFDFAIAELDTLSEESYKDSTLIMQLLKDNLTLWTS ALNFSVFYYEILNSPERACSLAK FAFDFAIAELDTLSEESYKDSTLIMQLLKDNLTLWTS ALNFSVFYYEILNSPDFACKLAK DAFDFAIAELDTLSEESYKDSTLIMQLLKDNLTLWTS ALNFSVFYYEILNSPDFACKLAK DAFDFAIAELDTLKEESYKDSTLIMQLLKDNLTLWTS ALNFSVFYYEILNSPDFACKLAK DAFDFAIAELDTLKEESYKDSTLIMQLLKDNLTLWTS ALNFSVFYYEILNSPDFACKLAK DAFDFAIAELDTLKEESYKDSTLIMQLLKDNLTLWTS ALNFSVFYYEILNSPDFACKLAK DAFDFAIAELDTLKEESYKDSTLIMQLLKDNLTLWTS ALNFSVFYYEILNSPDFACKLAK DAFDFAIAELDTLKEESYKDSTLIMQLLKDNLTLWTS	239 239 236 234 230 233
Maize GF14-6 Maize GF14-12 Arabidopsis GF14 Yeast Bovine 14-3-3 Sheep KCIP	DISEDPAEEIREA-PKRDSSEGQ DISEDPAEEIREA-PKHDLSEGQ DMQDDAADEIKEAAAPKPTEEQQ DMSESGQAEDQQQQQHQQQQPPAAAEVKHQSK ENQGDEGDAGEGEN- DMQGDGEEQNKEALQDVEDENQ	261 261 259 267 244 255

Figure 1. Comparison of the amino acid sequences among maize GF14–6 and GF14–12, *Arabidopsis* GF14, sheep protein kinase inhibitor protein (KCIP), bovine 14–3–3 β chain, and yeast BMH1. Completely conserved residues are indicated by boxes. Gaps are introduced to maximize matching. The sequence around the PKC "pseudosubstrate" (pseudosub) and the area that is similar to the annexin V (Aitken et al., 1992) and domain III of calmodulin (Babu et al., 1985) are stippled. The sources of the sequences are: sheep KCIP, Toker et al. (1992); maize GF14–6 and GF14–12, this work; *A. thaliana* GF14, Lu et al. (1992); yeast BMH1, Van Heusden et al. (1992); bovine 14–3–3 β chain, Isobe et al. (1991).

Figure 2. There is an unambiguous and separate grouping of the mammalian and plant GF14/14-3-3 proteins. Like the mammalian 14-3-3 proteins, the plant GF14 protein isoforms are 70 to 98% identical. The only exception is the sheep 14-3-3 ϵ -isoform, which is more related to the plant isoforms than to the mammalian 14-3-3 proteins. However, the confidence in the branch points at this level is low by this method (Ferl et al., 1994), as indicated by the overlapping error bars in the secondary bifurcations. Pairwise comparison of the plant and insect/mammalian isoforms presented in Figure 2 shows that they are between 50 and 60% identical (data not shown). The yeast 14-3-3 protein isoform is more related to the plant isoforms (62-69% identical residues) than to the insect/mammalian isoforms (51-59% identical residues). There appears to be no clear separation between the dicot and monocot isoforms, since the the tomato GF14 and Arabidopsis GF14 isoforms are grouped with the monocot proteins.

Expression of GF14 Isoforms in E. coli

The authenticity of both cDNAs was verified by expression experiments in *E. coli*. To express the polypeptides encoded

by the GF14 cDNAs in E. coli, the cDNAs were fused to the 5' region of the β -galactosidase gene in pET15b such that hybrid polypeptides are formed of GF14 with a leader sequence of six consecutive His residues separated by a thrombin cleavage site (Fig. 3A). The isopropylthio- β galactoside-induced fusion polypeptides migrated at 32 kD on SDS-PAGE and were recognized by anti-GF14 antiserum (data not shown). The fusion proteins were purified by nickel affinity chromatography, cleaved by thrombin, and passed over the nickel affinity column for a second time to remove the histine leader sequence and contaminating bacterial proteins. Overall purity of the GF14 polypeptides was estimated to be more than 95%. Figure 3B shows a Coomassie bluestained gel (lanes 1-3) and an immunoblot (lanes 4-6) of crude protein extract from maize P3377 cell suspension (lanes 1 and 4), purified GF14-6 (lanes 2 and 5), and GF14-12 (lanes 3 and 6). Purified GF14-6 and GF14-12 polypeptides are indistinguishable on the basis of their mobility on an SDS-PAGE gel, both having a molecular size of 30 kD. This is close to the calculated 29.6 kD of the deduced amino acid sequences. The immunoblot analysis showed that the size of the purified expressed GF14 polypeptides is identical with



Figure 2. Gene genealogy illustrating evolutionary relationships among the GF14/14-3-3 proteins. The full-length sequences were used for the calculations. The numbers and sE values indicate the validity of the branch points derived from the UPGMA analysis in GeneWorks. The sources of the sequences are: sheep ϵ_i Toker et al. (1992); maize GF14-6 and GF14-12, this work; A. thaliana GF14 phi, GF14 psi, and GF14 upsilon, Ferl et al. (1994); A. thaliana GF14 omega, Lu et al. (1992); Oenothera PHP-O, Hirsch et al. (1992); yeast BMH1, Van Heusden et al. (1992), bovine 14-3-3 β , Isobe et al. (1991); tomato GF14, B. Laughner and R.J. Ferl, unpublished data; rice 14-3-3, Kidou et al. (1993); barley 14-3-3, Brandt et al. (1992); bovine 14-3-3 η , Ichimura et al. (1988); human PLA₂, Zupan et al. (1992); human T-cell, Nielsen (1991); human HME1, Prasad et al. (1992); human stratifin and 1054, Leffers et al. (1993); rat 14-3-3 ζ, Rosenfeld et al. (1991); rat 14-3-3 α and γ , Watanabe et al. (1993); and Drosophila 14-3-3, Swanson and Ganguly (1992). The human 14-3-3 HS1 (Leffers et al., 1993), rat 14-3-3 η (Watanabe et al., 1991), and sheep ζ (Toker et al., 1992) are not included in the analysis because they are 100% identical with the human T-cell, bovine 14-3-3 η , and rat 14-3-3 ζ , respectively.



Figure 3. Expression of GF14 proteins. A, Partial restriction map of pET-GF14 plasmids used for expression of the cDNAs in *E. coli*. See "Materials and Methods" for details of construction. The nucleic acid and amino acid sequences at the fusion junction of the His leader sequence and the modified pGF14–6 and pGF14–12 inserts are shown. The nucleotides corresponding to the His leader are shown in lowercase letters, whereas those of pGF14 are shown in uppercase. B, Expression of the GF14–6 and GF14–12 polypeptides. Coomassie blue-stained gel (lanes 1–3) and immunoblot analysis (lanes 4–6) of total protein extracted from maize cell suspension (lanes 1 and 4) and purified *E. coli*-expressed GF14–6 (lanes 2 and 5) and GF14–12 (lanes 3 and 6). Lanes 1 and 4 contain 100 μ g of protein, lanes 2 and 3 contain 4 μ g of protein, and lanes 5 and 6 contain 0.2 μ g of protein. C and D, Quantification of GF14 protein extracted from cell or leaf or root tissue. D, Purified GF14 protein concentrations plotted against the relative signal intensities on immunoblot are indicated by the closed circles and are connected with a straight line. Relative immunoblot signal intensity of 50 μ g of total protein extracted from cell and leaf and root tissue is compared with those obtained from the purified GF14 protein.

that of polypeptides extracted from the maize suspension cells, suggesting that there is little, if any, major posttranslation modification of GF14 polypeptides in maize cells. Immunoblot analysis also detected an additional protein of approximately 26 kD in both the maize cell extract and the extracts of *E. coli*-expressed GF14 (lanes 4–6). This is likely a proteolytic cleavage product of the 30-kD polypeptide and appears to be processed identically in maize cells and *E. coli*. Since the 26-kD polypeptide is retained on the nickel affinity column, the cleavage point is probably located 4 kD from the C terminus.

The cellular concentration of the GF14 proteins in maize tissue was estimated using both the monoclonal antiserum to GF14 and purified GF14 protein. A standard curve was prepared by running different amounts of purified GF14–12 protein on an SDS-PAGE gel, blotting to nitrocellulose, and probing with anti-GF14 antiserum (Fig. 3C). The signal obtained by multiple exposures of chemiluminescence detection was measured for each quantity of GF14 protein by a scanning densitometer. These values were plotted against nanograms of GF14 protein electrophoresed, resulting in a linear relationship (Fig. 3D). Above 250 ng of GF14 protein, the values did not increase linearly and are not represented in the standard curve. Subsequently, from 25 to 100 μ g of crude protein extract from maize suspension cells or leaf or root tissue were subjected to blot analysis (Fig. 3C). The values were linear between 25 and 50 μ g of protein but outside the curve in case of 100 μ g of protein. Values obtained for 50 μ g of crude protein are represented in the standard curve and extrapolate to relative GF14 protein concentrations in maize suspension cells and leaf and root tissue of 0.33%, 0, and 0.24%, respectively. These values represent the values of both GF14–6 and GF14–12 protein, since the anti-GF14 monoclonal antibodies cannot distinguish between the isoforms.

GRF Gene Structure and Organization

To characterize the organization and structure of the GF14 genes, a library of maize genomic DNA fragments cloned in λ EMBL3 was screened at high stringency with pGF14-12. Twelve recombinant phage were isolated and purified. Southern blot analysis using pGF14-12 as probe suggested the presence of two classes of genomic clones, designated *GRF1* and *GRF2* (Fig. 4). *GRF2* was identified as being homologous to pGF14-12 after preliminary sequence analysis, and the *GRF2* clones all were truncated, since none of them hybridized with a 3' probe of pGF14-12.

The sequence of the complete GRF1 gene is shown in Figure 5. Comparison of the nucleotide sequence of the cloned genomic regions with the cDNA clone pGF14-6 indicated that GRF1 encoded the transcript in six exons inter-



Figure 4. Restriction map of *GRF1* and *GRF2*. Restriction map of the insert of each genomic clone is shown at the top. Below the maps are the fragments of each gene that were subcloned and sequenced. The exons are filled blocks, and the 5' and 3' untranslated regions are open blocks. The lines connecting the filled blocks represent the introns. R, Right; L, left; S, *Sall*; B, *Bam*HI; E, *EcoRI*; K, *KpnI*; Hc, *HinclI*; Sc, *Sacl*; P, *PstI*; Sp, *SpeI*; SI, *SfiI*; Hd, *HindIII*.

rupted by five introns (Fig. 4). The exons of *GRF1* were almost identical with the sequence of pGF14-6 with 13 silent nucleotide substitutions, probably due to allelic differences in DNA sources between the cDNA and genomic libraries. As mentioned before, the transcriptional unit of *GRF2* was not completely represented in the isolated genomic clones, which were found to contain the first four exons and three introns (Fig. 4). The introns of *GRF1* and *GRF2* were conserved in length and nucleotide sequence, with the exception of intron IV, which did not reveal any homology (Table I). Both the *GRF1* and *GRF2* genes displayed intron junction sequences similar to the consensus for plants reported by Brown (1986). The conservation of nucleotide sequence and intron number and position indicates that these genes likely arose by gene duplication of a single primordial gene.

GBF-Binding Site within the 5' Upstream Region of the GRF Genes

Comparison of the nucleotide sequences of the 5' flanking regions of *GRF1* and *GRF2* revealed long stretches of extensive similarity out to approximately -650 (Fig. 6). Optimal alignment of these sequences required the introduction of several short gaps indicated by dots in both sequences.

To study the regulation of the GRF genes, we first investigated the possibility that these genes may contain binding sites for GBF and associated proteins. Interestingly, about 450 bp upstream of the transcription start site a potential class I/II GBF-binding site (5'-TCACGTGG-3') is present (Williams et al., 1992). To determine whether this sequence was involved in binding GBF, EMSA were performed with heparin-agarose and fast protein liquid chromatography Mono-Q-enriched maize GBF activity and E. coli-produced maize GBF (Fig. 7B). Figure 7A shows the DNA probes and competitor DNA that were used in these analyses, previously described by McKendree and Ferl (1992). EMSA of GBFenriched maize extract, using the 5' end-labeled GF14 G-box oligo as probe, demonstrated two major bound complexes (Fig. 7B, lane 2) that were stable to competition by greater than 200-fold excess of poly(dI-dC). Further experiments showed that the homologous GF14 G-box (lane 3) and the dyad G-box (lane 4) specifically competed for the bound complex. However, an equivalent amount of mutant G-box competitor does not effectively compete for GBF binding (lanes 5 and 6). Some competition is observed for the GM1 mutation, a mutation that destroys one side of the G-box dyad. However, the GM2 mutation, which destroys the ACGT core of the G-box, is unable to compete. These binding competition data clearly indicate that GBF specifically interacts with the GF14 G-box. Results obtained using E. coliproduced maize GBF are consistent with those using the cellular GBF activity (Fig. 7C, lanes 1-6). One G-box-GBF complex is observed in the absence of competitor DNA (lane 2); however, it was completely abolished by adding unlabeled GF14 G-box (lane 3) or dyad G-box (lane 4) but not in the presence of GM1 or GM2 (lanes 5 and 6).

Because the sequences flanking the CACGTG G-box core motif have been found to be very important for GBF binding (Williams et al., 1992) and since these are different between the GF14 G-box and the dyad G-box, we tested the relative

ggtaccgacaaagagatagataaaagactgcatgatacaattgtgagcctttaaatcatagatacatttaaagag	-1416
a castatgttagagacgaactctttataaagaacacttctcttgtatttttctattaacaccctagaggcottta	-1341
a gas a totttottotgacta a agttt a gtt c g c a t c c a t c a a t t t g a g t a t c a a t t a a a t a t a a a t a g a g a g	-1266
atgttttataatttcatctcgtatctttactaaacttcaagagaaaatcccttcaatgccatccaatcttatatc	-1191
astgootocastgcastcatasttatagcttcccttcastgccatcgasttaasttttastcoctgcastgcc	-1116
attaccgttagtttggagtaacgcogttagtccaatttaaagtctatctctcccaataccactaaatctatactga	-1041
tocottcaatgtcagtagaaattggttcgatotatttgggtttaaaaattocagaatcaaaaatatggatcaga	-966
tttttaanattccacgatttttaaaaaottcagtccattgaagaaattgacattgagatattgactttaaattg	-891
and an congration of the second test of	-816
	-741
	-666
aggetacaaaacaeeoacatgeeogattteaaateegetgetateaeteegatgacagtgeege	-000
gegeeteatgteacetgaegetogtgeaggagegggeetettoggegacaegeegacaegeettgte	-391
getecocacgecocgecetgtgtgtetegtgagtegtgacgtecagaggtecagatecatecegeeg	-516
cgcccgcgcccgccagcgccagttcttacccgcgcccgacgcacgaaaggacg <u>tcacgtgg</u> ccagcgaggcacgc	-441
gggcccacgggtcggtaactcaggcagctttgcggtccgtgcgagatgaatcatttccctttcggcoccgcgccg	-366
gtcagtocagogogtccaggctgtctcgoggcgcggggcgccggagacttacgaccagggcgcacaggg	-291
occaocatcacccggcaggcaggtaggcctgacaggtggagcaggctgtcccagtcoccacaagtcgtgagccgt	-216
gacggcototgaattogtogatggotacgcototocotgtooggtgggocatacgac <u>cacat</u> gcacogogogga	-141
$\tt ctagggctgccatcgtggcgccgacacgcgtcacc\underline{atat}cattggcccgcgggactcggcgacaggccttttgta$	-66
gaccgcagccggcaccatetettgccgcacccccattcgaacgccggagetogacaaggacgacCGACGACGCA	10
CGCCGGAGgttacctctcqccccccctctcctccgtcoctcggagtctcctgctgtttgatttcgggtcgctgct	85
gtttogatcccctttocttoctgacgatttgcttatttttgcgttggatttgtaacctaaacatctogtttccgc	160
accontitiogggccocontetactoggtcocttiagategatagggggacacatecccc	235
	310
gitt gitte gate gate gate gate gate gate gate and	385
grand at the sector and the sector of the sector and the sector at the s	460
	5 2 5
	535
tgetgagaattttgeattaeeeetggaaateeeeggaattgaetagetag	010
caataaagtgttgtoccatatattgatgtatatttgcagATCCGGCGTACTGCAAATTTTGAAACTGAACTGTAA	685
GATGGCATCAGCAGAGCTTTCCCGTGAGGAAAATGTGTACATGGCGAAGCTCGCTGAGCAGGCGGAGAGGTACGA M A S A B L S R B B N V Y M A K`L A B Q A B R Y B	760
GENGRTGETTENGTTCRTGENGANGETNGCCANGACTETTENCTCGENGENGCTCACTETGENGENGCGANACCT	835
	(50)
L S V A Y K N V I G A R R A S W R I I S S I B Q K	(75)
GGAGGAGGGCCGAGGCAATGAGGACCGAGTAACACTCATCAAGGATTATCGTGGCAAGATTGAGACTGAGCTGAC B B G R G N B D R V T L I K D Y R G K I B T E L T	985 (100)
CAAGATCTGTGATGGCATCCTCAAGCTGCTTGAGACCCATCTTGTGCCGTCTTCCACTGCCCCCGAGTCCAAGGT	1060
CTTCTATCTCAAGATGAAGGGTGATTACTACAGgtaatatgattgtatggatoccagtaagactttctcatatgc	1135
FYLKMKGDYYR	(136) 1210
	(146)
GAGAAAGGACGCCGCTGAGAACACGATGGTGGCATACAAGGCTGCCCAGgtcaacaatttttgtttcattctta R K D λ λ B N T M V λ Y K λ λ Q	1285
atcagtgtgcaattttgttcatgaagcaatatgcttgtagtttttttt	1360
ganaggggtttgtggtgcattcagGACATTGCTCTGGCTGAGCTTGCTCCAACTCACCCTATTAGGCTTGGACT	1435
D I A L A B L A P T H P I R L G L GGCACTTAACTTCTCAGTGTTCTACTATGAGATTCTGAACTCACCTGATCGCGCCTGCAGTCTTGCAAAGCAGgt	(179) 1510
ALNPSVPYYEILNSPDRACSLAKQ	(203)
tt ut ct ct tt tt cat a t a t a t a t a t a	1660
	1735
gyana cycal cycal a cyc	1010
tttggttcatctctagactcaatgctacaaccctatactttggtgagactttcaattattttttctggtlatga	1810
actatgaagcoctccagatattagatattagtatgcoctgtgtttttgcatgcatgttagttttacccatgtca	1885
aatatttctctctttgaccaccaattttgtaaaaattaatactccctcc	1960
${\tt tt} {\tt maaaaataaactagcgggcgacaaatattcgagaatggaggtagtaaattatataggctaacaacataagttt}$	2035
at gtttttggattcattatgagaaagactttcatactatgcggctcacatgttgtgacactacatgtcttctttt	2110
${\tt tg} a {\tt a} {\tt a} {\tt a} {\tt tg} {\tt ta} {\tt a} {\tt a}$	2185
tanttytttttgttgtcacacctgctgtcataaatgtgtactttttatacaaggattggtttgagcacaatttg	2260
tgatatggacaggaaacctaatattatggttsttcactattctgtcattgacttatatacaacactggaaatttc	2335
ttggaatcggtcttttgggagacagcatcttttaatggcagtcaactgttgtgctgagtgcgtgc	2410
atoctattgggtgtacctaaccttatgttattatgtgactttggttgactgctactttgtattaagttatgcogt	2485
topacagttatgcaaatecttttagttcaatgtcaaagtcccgttaggagaagaaaatatcacagtatattctgg	2560
	2635
	2710
attggtattergtatterggggggatatggattergggtttttttgatgattgatgattttttgatg	2785
	2960
tgaaagaagaagaatgaacatcttaattgtttagagtttagtgttattgagagttagtgattagtgattagtgatga	2000
A F D B A I S E L D T L S	(216)
GAGGAGTCCTACAAGGACAGCACTTGATCATGCAGCTCCTTCGTGATAACTTGACCTCTAGGACCTCTGACATC B B S Y K D S T L I N Q L L R D N L T L W T S D I	3010 (241)
$\label{eq:construct} \begin{split} \texttt{TCGgtgagacccatgtocctaacactaagttttottogcactcttgaattggatcttoctgaaatcagcotcocccccc} \\ \texttt{c} \end{split}$	3085
octgancaattgtatttctggtgcttggctgcagGAGGACCCTGCTGAAGAGATTAGGGAGGCGCCCAAGCGGGA	3160
E D P A E E I R E A P X R D	(256) 3235
SSBGQ.	(261)
CTNPGCANTGTGATCTGGAACCTGGTTTCTTGGGTGCGACGCTTGAAGCCGTGTTATGACTGTCGTCCGTC	3310 3385
CONCRETE: CONCRETE: CONCRETE: CONTREMENT OF CONTREMENTS CONTREMENTS OF CONTREMENTS OF CONCRETE: CONTREMENTS OF	3460
OULOUTOTOTIOGIIOGAACULTIAACUIIGULIGATIIACUIIICALUIGULGULG	25.25
IUNDIGGULUTINICALGGGEGGEGGEGEGEGEGEGEGEGEGEGEGEGEGEGEGE	3610
Ly ay Ly yy y y ca Ly ag ca Li Lo Lo Ly Lo Coatol ca al ya Ly La ay Li Ci Cy Cy Cy Co Ca Ag a Cy	2010
acticagaateegtaaatetetgaagtetaaaegetetatetaaaeettiggattegattaeaattgtagatata	2895
gaaatotgtagaagggtgcaattcagaatactttgccatgaaatgaatatoottottgtgagactotgtattaag	3/60
actamacagercacattgatagagattergtgatteetteetacaegatgagageetttgatatätgtgeata	2050
TT CONTENT OF ACABACCANTIC	2028

ttccatctttctagaaggaattc

affinity of GBF for the GF14 G-box. To assess this, we compared the binding of E. coli-produced maize GBF to that of the GF14 G-box sequence and the dyad G-box (Fig. 7C, lanes 7-10). It is apparent that both sequences are recognized by GBF; however, the formation of the G-box-GBF complex with the dyad G-box occurs at a 5- to 10-fold higher concentration compared to that with the GF14 G-box probe (cf. lanes 8 and 10). Thus, GBF can bind to the GF14 G-box, although the binding is significantly weaker than the binding to the dyad G-box.

To test whether GF14 is itself part of the GBF-protein complex that binds to the GF14 G-box, we performed EMSA in the presence of anti-GF14 monoclonal antibodies (Fig. 7B, lanes 7-11). Incubation of the G-box-protein complex with anti-GF14 antiserum, either monoclonal antibody 2A3 (lane 9) or monoclonal antibody 2A6 (lane 10), shifted some of the formed complexes to a slower mobility, whereas a control monoclonal antibody had no effect (lane 8). Antibodies do not have DNA-binding activity (lane 11). It should be noted that in this instance less DNA-protein complex was shifted than was reported by de Vetten et al. (1992). We attribute this to the age of the monoclonal antibody supernatant used in these experiments. However, it is clear that GF14 is part of a G-box-binding complex that is able to bind to the upstream region of its own gene.

DISCUSSION

In this report we describe the cloning and characterization of the genes for two GF14 isoforms from maize. GF14 proteins are members of the 14-3-3 protein family, a class of proteins about which information is rapidly emerging and that are thought to function as regulators of signal transduction (Aitken et al., 1992). One outstanding feature of this protein family is the extraordinarily high sequence conservation, more than 60% identity over a broad range of biological species, including plants, invertebrates, and yeast.

Our original probes for isolating the maize GF14 isoform cDNAs were monoclonal antibodies selected specifically for their ability to recognize proteins that are part of GBF complexes (de Vetten et al., 1992). Two cDNAs were isolated encoding proteins of 29.6 kD that are different in their amino acid sequence by only four amino acids (Fig. 1). The deduced amino acid sequences are homologous to the plant, invertebrate, and yeast 14-3-3 protein isoforms. The maize GF14 isoforms are between 75 and 83% identical with other plant GF14/14-3-3 isoforms, between 56 and 71% identical with the animal 14-3-3 proteins, and 65% identical with the yeast isoform. The homology is concentrated predominantly in

Figure 5. Complete nucleotide sequence of the GRF1 gene including introns and 5' and 3' flanking regions. The nucleotide at position +1 corresponds to the first nucleotide of the cDNA. The nucleotides upstream from the +1 position are indicated by negative numbers. The coding sequence of the gene is shown in uppercase letters. The remainder of the sequence is shown in lowercase letters. The derived amino acid sequence is presented below the DNA sequence and the number is given in parentheses. The putative TATA and CAAT boxes and polyadenylation signal are underlined.

fable I. Splic	e junction	sequences, intron length, and ir	tron identity in m	aize GF14 genes.	
Gene	Intron Number	Donor Length	Acceptor	Identity (%)	
GRF1 GRF2	١	GCGGAGgttacc638ttg	cagATCCGG	68	
GRF1 GRF2	II	CTACAGgtaata87gat CTACAGgtatta89gat	tagATACCT tagATACCT	85	
GRF1 GRF2	111	GCCCAGgtcaac126att GCCCAGgtcacc126att AAGCAGgtttgt1388tg	cagGACATT cagGACATT cagGCTTTT	87	
GRF1 GRF2	IV	CCGATCgtctac . >870			
GRF2	v	ATCTCGgtgaga 106 ctg	cagGAGGAC		
Consensus		CAGgta/c cag	A/GT/AC		

conserved domains, some of which have been implicated in functional aspects of the protein (Fig. 1). It is very likely that more than these two isolated GF14 isoforms are present in maize cells. The fact that only two GF14 isoforms were isolated from the maize expression library is probably the result of the specificity of the monoclonal antibodies. Reverse-phase HPLC of purified 14-3-3 protein prepared from mammalian brain tissue revealed seven to eight distinct polypeptides (Ichimura et al., 1988; Toker et al., 1992), likely transcribed from separate genes (Isobe et al., 1991). From Arabidopsis thaliana, we have isolated five different GF14 isoform cDNAs, all recognizing a different set of restriction fragments on a Southern blot of Arabidopsis genomic DNA (Ferl et al., 1994). The presence of more than two GF14 isoforms in maize is substantiated by the work of Keith et al. (1993), who identified by partial sequencing of random cDNA clones a GF14 distinctly different from the ones reported here.

Structure of the GRF Genes Encoding GF14 Proteins

The maize genes encoding the GF14 proteins are termed GRF1 and GRF2 for General Regulatory Factor. They consist of six exons interrupted by five introns (Fig. 4). The exon sequences of the two GRF genes are very similar (94%); they exhibit identical intron position and similar intron lengths. GRF2 was not completely represented by the cloned genomic region and is truncated at the 3' end of the transcription unit. The overall conservation of the two isolated GRF genes indicates that they likely arose through a relatively recent duplication of a single ancestral gene. The chromosomal locations of the GRF genes were determined by restriction fraction length polymorphism analysis in the laboratory of T. Helentjaris (Arizona State University, Tucson, AZ) and were mapped to chromosomes 2S and 10L. A considerable amount of evidence suggests, in fact, that these entire chromosome segments originate from a duplication event (Helentjaris et al., 1988).

GF14-6	. ARACTRACGATAATATCATTCAAGAGATTAAAATTTCAAGTTATTGAAGGGA AGC	-814
GF14-12	Tácáccásccólascáccáclolitetétetetetetetetetetetetetetetetetetet	-757
GF14-6	TATAAATTCTGATGACATTGAAGAAATTAATAATATAAAATTTGGATGTCATTGAAGGAATTTTCTTTT	-749
GF14-12	tchttridetdestideedtheesticheitshisddechattriestst	-709
GF14-6	AACTGAGGCTAACAAAACACCCCACATGCCCGATTTCAAATCCATCTGCTATCACTCCATCCGA	-686
GF14-12	CÓTGŤCÁĊGGGTTGĊĊĊĂTŤĊĂ.AĊĊĊĠĂ.ŤĊĊĂGĂŤGGĂĊAŤĠ.ŤĂŤAGAGA	-659
GF14-6	GT.GACAGTCGTTGCCGCGCGCCTCATGTCACCTGACGCTCGTGCAGGAGCGGGCCCCGCGCGCCT	-622
GF14-12	ctgcgcAgtcgttgAtgcgcgcCtcAcgtcAcgtcAtgCtcgcgcgggAgcgggCCCCCCGCGACCT	-594
GF14-6	CTTCGGCGACACGCCGACACGCCTTGTCGCCGCCGCCGCCGCCGCCGCCGGCGACGACGC	-557
GF14-12	CCTCGGCAACAGGCCTTGTCGCTCCCCATGCCCC.CCGCCC.GCCGCCCTGTGTGT	-540
GF14-6	CGTGACGTCCAGAGGTCCAGATCCATCCCGCCCGCGCCCGCGCCGCCAGCGCCAGTTCTT	-492
GF14-12	CCTGACGCCCGTTCAGGTCCAGATCGATGCCG	-508
GF14-6	ACCCGCGCCCCGACGCACGAAAGGACGTCACGTGGCCAGGCGACGCCCCGCGCCCCACGGGTCGGTA	-427
GF14-12	. CCCGCACCGACGACGAAGGGACG <u>TCACGTGG</u> CCAGCGAGACACGCGGGCCCACGGATTGGTA	-444
GF14-6	ACTCAGGCAGCTTTGCGGTCCGTGCGAGATGAATCATTTCCCTTTCGGCCCCGCGCCGGTC	-366
GF14-12	ACTCAGACAGCTTTGCGGCCCGTGCGAGATGAACCCTTTCGGCCCTGTGCGCCCGACC	-385
GF14-6		-320
GF14-12		- 271
GF14-0		-255
GE14=6	CTACCC CTACAGETGEAGCAGETGTCCACTCCCCACAAGTCGTGAGCCGTG	-218
GF14-17	CTACCCCGGCAGGTCACACTCACACTCGCAGCAGGCTGTCCCGGTCCCCACACGTCGTG	~197
GF14-6	ACGGCCTCTGAATTCGTCGATGGCTACGCCTCTCCCTGTCCGGTGGGCCATA.CGAC	-162
GF14-12	ACGGCCT.TGAATTTGTCGATGCTCCGCCTCCCCCCCCCC	-133
GF14-6	CACATGCACCGCGCGCGCACTAGGGCTGCCATCGTGGCGCGACACGCGTCACCATATCATTGGCC	-97
GF14-12	CACATGCACCGCGCAGTATTAGGGCTGCCACGGTGGGCCGCCACGCGTGGATATCTCATTGGAG	-68
GF14-6	cqcgqqActcqqcqAcAqGccttttgctAqAccqcAgccCGqCACCAtctttgcCGCACCCCCC.A	-33
GF14-12	TGCAGGACTCGGAGGCAGCCCATTTCTAG,GCAGCGCCGTCTTCGCGCACCCCCCAA	-10
GF14-6	TTCGAACGCCGGAGCTCGACAAGGACGACCGACGACGCCGCGAGG +14	
GF14-12	TTCGAACGACGGAGGGAGCACGACGCACGCAGGGAGG +21	

Figure 6. Comparing of the 5' flanking sequences of *GRF1* and *GRF2*. Numbering is from the start of the cDNA. Gaps have been introduced to maximize matching. Conserved nucleotides are boxed. A conserved G-box promoter element 5'-TCACGTGG-3' near -450 is indicated in bold type and is underlined.

1602



The GBF-Binding Complex May Regulate GF14 Expression

Inspection of the 5' flanking region of the GF14 genes revealed the presence of a well-defined G-box element 450 bp upstream of the start of the cDNA (Fig. 6). This sequence specifically binds to GBF from maize-enriched protein extract and GBF expressed in E. coli (Fig. 7). Furthermore, GF14 is itself associated with the GBF-protein complex that binds to the GF14 upstream region in a fashion similar to that presented previously (de Vetten et al., 1992). Physical association of GF14 and GBF was supported by partial co-fractionation of GBF activity and GF14, immunoprecipitation experiments, and localization studies demonstrating that both proteins are present in the nucleus (de Vetten et al., 1992). Binding of GBF and associated GF14 to a G-box element from within the upstream region of GF14 may suggest that GF14 activity in vivo is, in part, autoregulated at the transcriptional level. In plants, examples of autoregulation of genes encoding transcription factors are limited. Fromm et al. (1991) demonstrated that TGA1a can bind to a sequence in the upstream region of a TGA1a-related gene. However, it is unknown how these binding properties correspond to the regulation of expression of the TGA1a-related gene family in vivo. In mammalian systems, autoregulation of expression of genes encoding transcription factors is fairly well documented (Serfling, 1989). In case of the transcriptional regulation of the jun gene family, it was shown that c-Jun is positively autoregulated (Angel et al., 1988), whereas Jun-B can inhibit the activation by c-Jun (Chui et al., 1989). In our study we showed that GBF and associated GF14 binds to the upstream region of GF14. How these in vitro experiments correspond to regulation of GF14 gene expression in vivo is as yet unknown.

Figure 7. Maize GBF-enriched protein extract or E. coli-expressed GBF binds to the GF14 G-box promoter element. A, Competitor DNA oligonucleotides used in the EMSA. The dyad G-box oligonucleotide was derived from the Arabidopsis Adh 5' flanking sequence (McKendree et al., 1990). The two site-specific mutations of the dyad G-box oligonucleotide, GM1 and GM2, were previously described by McKendree and Ferl (1992). Only the top strand of the sequence is shown. B, Partially purified maize GBF activity interacts specifically with the GF14 G-box and is recognized by anti-GF14 antiserum. The probe was labeled GF14 G-box (see A). Binding reactions were carried out either in the absence (lane 2) or in the presence of 50 ng of the following oligonucleotides: GF14 G-box (lane 3), dyad G-box (lane 4), GM2 (lane 5), or GM1 (lane 6). Lanes 8 through 10 show standard binding reactions incubated with control monoclonal antibody (lane 8), anti-GF14 monoclonal antibody 2A3 (lane 9), or anti-GF14 monoclonal antibody 2A6 (lane 10) cell supernatant. Lane 11 is anti-GF14 monoclonal antibody cell supernatant without the addition of GBF-enriched protein extract. A, B, and F denote antibody-protein-DNA complex, protein-DNA complex, and free DNA, respectively. C, E. coli-expressed GBF interacts specifically with the GF14 G-box. Binding reactions contained 4 ng of purified GBF, 1 µg of poly(dI-dC) (lane 2), 50 ng of GF14 G-box (lane 3), 50 ng of dyad G-box (lane 4), 50 ng of GM2 (lane 5), or 50 ng of GM1 (lane 6). Lanes 7 though 10 show E. coliexpressed GBF binding to either GF14 G-box as probe (lane 8) or dyad G-box as probe (lane 10).

The Range of Possible Roles for the GF14 Polypeptides

The highly conserved structural features of the GF14/14-3-3 protein family suggests an important and very likely a similar role in the diverse biological systems studied thus far. The 14-3-3 proteins were first identified as activators of Tyr and Trp hydroxylases, rate-limiting enzymes in the biosynthesis of monoamines (Ichimura et al., 1987, 1988). Accumulation of 14-3-3 protein was estimated to be more than 1% of total brain proteins (Boston et al., 1982), and such abundance is apparently in keeping with the important nature of neurotransmitter production in the brain. However, discoveries of additional functions for these proteins and their presence at lower levels in nonbrain tissues expanded consideration of their roles. Aitken et al. (1991) found that 14–3–3 proteins isolated from sheep brain are highly effective inhibitors of Ca2+-/phospholipid-dependent PKC. Ichimura et al. (1991) showed that the tissue distribution of 14-3-3 almost parallels the known distribution of PKC. In situ hybridization revealed that one of the 14-3-3 isoforms was expressed in both monoamine-synthetic neurons and neurons that do not synthesize monoamines (Watanabe et al., 1991). Furthermore, Isobe et al. (1992) demonstrated that some 14-3-3 isoforms that were previously shown to activate Tyr and Trp hydroxylases are also able to activate PKC. The 14-3-3 isoforms have also been implicated in the stimulation of Ca²⁺-dependent exocytosis (Morgan and Burgoyne, 1992). Again, it is thought that the ability of the 14-3-3 protein to activate exocytosis is potentiated by PKC activation. Thus, an increasing amount of evidence implicates 14-3-3 proteins in the mediation of calcium- and protein kinase-regulated events (Aitken et al., 1992; Schlaepfer et al., 1992; Roth et al., 1993), considerably broadening the scope of potential roles of GF14s.

Our estimates of GF14 abundance are (0.24–0.33%) somewhat lower than the abundance of 14-3-3 proteins in the brain. Nonetheless, plant GF14/14-3-3 proteins have a relatively wide distribution and demonstrate properties of the animal 14-3-3 proteins. They have been shown to modulate PKC activity isolated from mammalian brain (Hirsch et al., 1992; Lu et al., 1994a), and plant proteins with properties of PKC have been described (Elliott and Kokke, 1987). Therefore, plant GF14/14-3-3 homologs may act as regulators of these protein kinases. Phosphorylation also affects DNA binding and transcriptional activation properties of several nuclear proteins (Klimczak et al., 1992). Therefore, it is likely that the participation of the GF14 isoforms in the G-boxbinding complex regulates phosphorylation of proteins associated with this complex. This, in turn, may regulate promoter activity of certain genes, including GRFs.

ACKNOWLEDGMENTS

We thank Beth Laughner and Ernie Almira for discussion and technical assistance. Access to automated DNA sequence analysis and hybridoma core facilities was provided through the Interdisciplinary Center for Biotechnology Research at the University of Florida.

Received April 14, 1994; accepted July 12, 1994. Copyright Clearance Center: 0032-0889/94/106/1593/12.

LITERATURE CITED

- Aitken A, Collinge DB, Van Heusden BPH, Isobe T, Roseboom PH, Rosenfeld G, Soll J (1992) 14–3–3 proteins: a highly conserved widespread family of eukaryotic proteins. Trends Biochem Sci 17: 498–501
- Aitken A, Ellis CA, Sellers LA, Toker A (1991) Kinase and neurotransmitters. Nature 344: 594
- Angel P, Hattori K, Smeal T, Karin M (1988) The Jun protooncogene is positively autoregulated by its product, Jun/Ap-1. Cell 55: 875–885
- Babu YS, Sack JS, Greenhough TJ, Bugg CE, Means AR, Cook WJ (1985) Three-dimensional structure of calmodulin. Nature 315: 37-40
- Boston PF, Jackson P, Kynoch AM, Thompson RJ (1982) Purification, properties, and immunohistochemical localization of human brain 14-3-3 protein. J Neurochem 38: 1466-1474
- **Bradford MM** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem **72:** 248–254
- Brandt J, Thordal-Christensen H, Vad K, Gregersen PL, Collinge DB (1992) A pathogen-induced gene of barley encodes a protein showing high similarity to a protein kinase regulator. Plant J 2: 815-820
- **Brown JWS** (1986) A catalogue of splice junctions and putative branch point sequences from plant introns. Nucleic Acids Res 14: 9549–9559
- Brunelle AN, Chua N-H (1993) Transcription regulatory proteins in higher plants. Curr Opin Genet Dev 3: 254-258
- Chui R, Angel P, Karin M (1989) Jun-B differs in its biological properties from, and is a negative regulator of, c-Jun. Cell 59: 979–986
- DeLisle AJ, Ferl RJ (1990) Characterization of the Arabidopsis Adh G-box binding factor. Plant Cell 2: 547–557
- De Vetten NC, Ferl RJ (1994) Transcriptional regulation of environmentally inducible genes in plants by an evolutionary conserved family of G-box binding factors. Int J Biochem 26: 1055–1068
- De Vetten NC, Lu G, Ferl RJ (1992) A maize protein associated with the G-box binding complex has homology to brain regulatory proteins. Plant Cell 4: 1295–1307
- **Donald RGK, Cashmore AR** (1990) Mutation of either G box or I box sequences profoundly affects expression from the *Arabidopsis rbcS*-1A promoter. EMBO J **9**: 1717–1726
- Duncan DR, Williams ME, Zehr BE, Widholm JM (1985) The production of callus capable of plant regeneration from immature embryo of numerous Zea mays genotypes. Planta 165: 322-332
- Elliott D, Kokke YS (1987) Partial purification and properties of a protein kinase C type enzyme from plants. Phytochemistry 26: 2929–2935
- Ferl RJ, Lu G, Bowen BW (1994) Evolutionary implications of the family of 14-3-3 brain protein homologs in *Arabidopsis thaliana*. Genetica 92: 129-138
- Fromm H, Katagiri F, 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 MJ, Marcotte WM, Quatrano RS (1990) A leucine zipper protein that recognizes an abscisic acid response element. Science 250: 267–271
- Helentjaris T, Weber D, Wright S (1988) Identification of the genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. Genetics 118: 353-363
- Hirsch S, Aitken A, Bertsch U, Soll J (1992) A plant homologue to mammalian brain 14–3–3 protein and protein kinase C inhibitor. FEBS Lett 296: 222–224
- Ichimura T, Isobe T, Okuyama T, Takahashi N, Araki K, Kuwano R, Takahashi Y (1988) Molecular cloning of cDNA coding for brain-specific 14-3-3 protein, a protein-kinase-dependent activator of tyrosine and tryptophan hydroxylases. Proc Natl Acad Sci USA 85: 7084-7088
- Ichimura T, Isobe T, Okuyama T, Yamauch T, Fujisawa H (1987) Brain 14-3-3 protein is an activator protein that activates trypto-

phan 5-monooxygenase and tyrosine 3-monooxygenase in the presence of Ca^{2+} , calmodulin-dependent protein kinase II. FEBS Lett **219**: 79-82

- Ichimura T, Sugano H, Kuwano R, Sunaya T, Okuyama T, Isobe T (1991) Widespread distribution of the 14–3–3 protein in vertebrate brains and bovine tissues: correlation with the distributions of calcium-dependent protein kinases. J Neurochem 56: 1449–1451
- Isobe T, Hiyane Y, Ichimura T, Okuyama T, Takahashi N, Nakajo S, Nakaya K (1992) Activation of protein kinase C by the 14-3-3 proteins homologous with Exo1 protein that stimulates calcium-dependent exocytosis. FEBS Lett 308: 121–124
- Isobe T, Ichimura T, Sunaya T, Okuyama T, Takahashi N, Kuwano R, Takahashi Y (1991) Distinct forms of the proteinkinase-dependent activator of tyrosine and tryptophan hydroxylases. J Mol Biol 217: 125-132
- Johnson PF, McKnight SL (1989) Eukaryotic transcriptional regulation regulatory proteins. Annu Rev Biochem 58: 799–839
- Katagiri F, Chua N-H (1992) Plant transcription factors: present knowledge and future challenges. Trends Biochem Sci 8: 22-27
- Keith CS, Hoang DO, Barrett BM, Feigelman B, Nelson MC, Thai H, Baysdorfer C (1993) Partial sequence analysis of 130 randomly selected maize cDNA clones. Plant Physiol 101: 329–332
- Kidou S-I, Umeda M, Kato A, Uchimiya H (1993) Isolation and characterization of a rice cDNA similar to the bovine brain-specific 14–3–3 protein gene. Plant Mol Biol 21: 191–194
- Klimczak LJ, Schindler U, Cashmore AR (1992) DNA binding activity of the *Arabidopsis* G-box binding factor GBF1 is stimulated by phosphorylation by casein kinase II from broccoli. Plant Cell 4: 87–98
- **Kozak M** (1984) A compilation and analysis of sequences upstream from the translation start site in eukaryotic mRNAs. Nucleic Acids Res 12: 857–872
- Leffers H, Madsen P, Rasmussen HH, Honore B, Andersen AH, Walbum E, Vandekerckhove J, Celis JE (1993) Molecular cloning and expression of the transformation sensitive epithelial marker stratifin: a member of a protein family that has been involved in the protein kinase C signaling pathway. J Mol Biol 231: 982–998
- Lu G, DeLisle AJ, de Vetten NC, Ferl RJ (1992) Brain proteins in plants: an Arabidopsis homolog to neurotransmitter pathway activators is part of a DNA binding complex. Proc Natl Acad Sci USA 89: 11490–11494
- Lu G, de Vetten NC, Sehnke PC, Isobe T, Ichimura T, Fu H, van Heusden GPH, Ferl RJ (1994a) A single *Arabidopsis* GF14 isoform possesses biochemical characteristics of diverse 14–3–3 homologs. Plant Mol Biol 25: 659–667
- Lu G, Sehnke P, Ferl RJ (1994b) Phosphorylation and calcium binding properties of an *Arabidopsis* GF14 brain protein homolog. Plant Cell 6: 501-510
- Martin KJ (1991) The interactions of transcription factors and their adaptors, coactivators and accessory proteins. Bioessays 13: 499-503
- McKendree WL, Ferl RJ (1992) Functional elements of the Arabidopsis Adh promoter include the G-box. Plant Mol Biol 19: 859-862
- McKendree WL, Paul A-L, DeLisle AJ, Ferl RJ (1990) In vivo and in vitro characterization of protein interactions with the dyad Gbox of the Arabidopsis Adh gene. Plant Cell 2: 207–214
- **Moore BW, Perez VJ** (1967) Physiological and biochemical aspects of nervous integration. *In* FD Carlson, ed, Symposium on Physiological and Biochemical Aspects of Nervous Integration. Marine Biological Laboratory, Woods Hole, MA, pp 343–359
- Morgan A, Burgoyne RD (1992) Exo1 and Exo2 proteins stimulate calcium-dependent exocytosis in permeabilized adrenal chromaffin cells. Nature 355: 833-836
- Nielsen PJ (1991) Primary structure of a human protein kinase regulator protein. Biochim Biophys Acta 1088: 425–428
- Oeda K, Salinas J, 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

- Prasad GL, Valverius EM, McDuffie E, Cooper HL (1992) Complementary DNA cloning of a novel epithelial cell marker protein, that may be down-regulated in neoplastic mammary cells. Cell Growth Differ 3: 507–513
- Rosenfeld GC, Sanborn B, Loose-Mitchell D (1991) C.oning, distribution and expression of rat 14–3–3 protein: proceedings of the FASEB conference, Atlanta (abstract No. 2692). FASEF J A834
- Roth D, Morgan A, Burgoyne RD (1993) Identification of a key domain in annexin and 14–3–3 proteins that stimulate calciumdependent exocytosis in permeabilized adrenal chromaffin cells. FEBS Lett 320: 207–210
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Schindler U, Menkens AE, Beckmann H, Ecker JR, Cashmore AR (1992a) Heterodimerization between light-regulated and ubiquitously expressed Arabidopsis thaliana GBF bZIP proteins. EMBO J 4: 1261–1273
- Schindler U, Terzaghi W, Beckmann H, Kadesch T, Cashmore AR (1992b) DNA binding site preference and transcription activation properties of the *Arabidopsis* transcription factor GBF1. EMBO J 11: 1275–1289
- Schlaepfer DD, Jones J, Haigler H (1992) Inhibition of protein kinase C by annexin V. Biochemistry 31: 1886–1891
- Schulze-Lefert P, Dangl JL, Becker-Andre M, Hahlbrock K, 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
- Serfling E (1989) Autoregulation—a common property of eucaryotic transcription factors? Trends Genet 5: 131–133
- Swanson KD, Ganguly R (1992) Characterization of a Drosophila melanogster gene similar to the mammalian genes encoding the tryosine/tryptophan hydroxylase activator and protein kinase C inhibitor proteins. Gene 113: 183-190
- Toker A, Sellers LA, Amess B, Patel Y, Harris A, Aitken A (1992) Multiple isoforms of a protein kinase C inhibitor (KCIP/14-3-3) from sheep brain. Eur J Biochem 206: 453-461
- Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA 76: 4350–4354
- Van Heusden GPH, Wenzel TJ, Lagendijk EL, De Steensma Y, van den Berg JA (1992) Characterization of the yeast BMH1 gene encoding a putative protein homologous to mammalian protein kinase II activators and protein kinase C inhibitors. FEBS Lett 302: 145–150
- Watanabe M, Isobe T, Ichimura T, Kuwano R, Takahashi Y, Kondo H (1993) Molecular cloning of rat cDNAs for β and γ subtypes of 14–3–3 protein and developmental changes in expression of their mRNAs in the nervous system. Mol Brain Res 17: 135--146
- Watanabe M, Isobe T, Okuyama T, Ichimura T, Kuwano R, Takahashi Y, Kondo H (1991) Molecular cloning of cDNA to rat 14-3-3 η chain polypeptide and the neuronal expression of the mRNA in the central nervous system. Mol Brain Res 10: 151-158
- Weisshaar B, Armstrong GA, Block A, Da Costa e Silva O, uHahlbrock 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
- Williams ME, Foster R, Chua N-H (1992) Sequences flanking the hexameric G-box core CACGTG affect the specificity of protein binding. Plant Cell 4: 485–496
- Zupan LA, Steffens DL, Berry CA, Landt M, Gross RW (1992) Cloning and expression of a human 14-3-3 protein mediating phospholipolysis. J Biol Chem 267: 8707-8710

1604