Characterization of a zinc finger DNA-binding protein expressed specifically in *Petunia* petals and seedlings

Hiroshi Takatsuji¹, Masaki Mori²,
Philip N. Benfey, Ling Ren and Nam-Hai Chua³

Laboratory of Plant Molecular Biology, The Rockefeller University,
1230 York Avenue, New York, NY 10021, USA

¹Present address: National Institute of Agrobiological Resources,
Tsukuba Science City, Ibaraki, 305, Japan

²Present address: Kyushu National Agricultural Experimental Station,
Nishigoushi, Kikuchi-gun, Kumamoto, 861-11, Japan

³Corresponding author

Communicated by J.H. Weil

In *Petunia*, the expression of the 5-enolpyruvylshikimate-3-phosphate synthase gene (*EPSPS*) is tissue-specific and developmentally regulated. Nuclear extracts from *Petunia* petal contain a factor that interacts with the 5' upstream region of *EPSPS*. DNase I footprinting experiments revealed four strong binding sites (EP1–EP4) and several weaker sites that appear to bind the same factor. We have isolated a cDNA clone (*EPF1*) encoding a DNA-binding protein that has similar binding activity to that of the nuclear factor. The deduced amino acid sequence shows that the encoded protein, EPF1, contains two repeats of a Cys₂/His₂ zinc finger motif. EPF1 and the factor detected in nuclear extracts appear to differ in their molecular weight and Zn²⁺ requirements. Nevertheless, Northern blot analyses showed that the expression pattern of *EPF1* is remarkably similar to that of *EPSPS*. In addition, as determined by translational fusion of the *EPF1* upstream region to the β -glucuronidase reporter gene, the cell specific expression pattern of *EPF1* in flower and seedling is nearly identical to that of *EPSPS*. Taken together with the results of *cis*-element analyses, these observations suggest that *EPF1* may be one of the factors involved in the activation of *EPSPS*.

Key words: *EPSPS*/plant transcription factor/tissue-specific promoter

Introduction

5-Enolpyruvylshikimate-3-phosphate (EPSP) synthase catalyzes an essential step in the shikimate pathway leading to the biosynthesis of aromatic amino acids and chorismate-derived secondary metabolites such as lignins, flavonoids and anthocyanins. This enzyme is also the target for the broad spectrum herbicide, glyphosate (Steinrücken and Amerhein, 1980). The *Petunia* gene that codes for EPSP synthase (*EPSPS*) was found to be activated at two different developmental stages. In seedlings, expression is high in root and stem. In mature plants, the expression is high in petal, lower in floral tube and barely detectable in root, stem and leaf (Gasser *et al.*, 1988). We have shown that upstream sequences of the *Petunia EPSPS* confer petal-specific expression to a reporter gene in transgenic plants (Benfey

and Chua, 1989). A 5' deletion analysis indicated that upstream sequences between –1752 and –823 are sufficient to confer the tissue-specific expression pattern. Further deletion analysis showed that smaller 5' fragments that contained sequences from –1752 to –1270 and from –1234 to –823 still confer tissue specificity, although the expression level is significantly reduced (Benfey *et al.*, 1990). These results suggest that there is a functional reiteration of *cis* elements for tissue specificity. In addition to tissue specificity, the upstream sequences of *EPSPS* have been shown to confer developmental stage-specific regulation in petals. The expression level is very low in flower buds but increases dramatically during flower opening. A histochemical analysis of β -glucuronidase (GUS) reporter gene expression revealed that the *EPSPS* upstream fragment (–1752 to –823) is active in nearly all cell types of the limb of the flower but only in cells adjacent to and in the upper epidermis of the upper portion of the tube (Benfey and Chua, 1989).

Flower development is a complex process that involves differential gene regulation. Recently, two genes that appear to be responsible for homeotic changes in flower development, *deficiens* from *Antirrhinum majus* (Sommer *et al.*, 1990) and *agamous* from *Arabidopsis thaliana* (Yanofsky *et al.*, 1990), have been cloned. Both genes encode putative transcription factors that share high sequence homology in a region thought to be involved in DNA binding. The genetic and molecular characterization of these genes suggests that they act very early in flower development. On the other hand, very little work has been done on the later stages of corolla development, which are characterized by the cessation of cell division and rapid cell expansion resulting in the opening of the flower. Because *EPSPS* in *Petunia* is activated at the later stages of flower development, analysis of the promoter of this gene is expected to shed light on the molecular mechanisms underlying the developmental events at these stages.

In this report, we describe the characterization of a nuclear factor that binds to multiple sites in the *EPSPS* promoter region and the isolation of a cDNA clone encoding a DNA-binding protein (EPF1) with a similar binding specificity to that of the nuclear factor. The deduced amino acid sequence of EPF1 shows that it contains two repeats of the canonical Cys₂/His₂ type zinc finger motif previously demonstrated to be a DNA-binding domain (Evans and Hollenberg, 1988). This is the first demonstration of a Cys₂/His₂ type zinc finger protein from plants. However, differences in apparent molecular weight and zinc ion requirements suggest that *EPF1* is not the most abundant factor detected in petal nuclear extract. Nevertheless, the expression pattern of *EPF1* parallels that of *EPSPS*, suggesting that this factor is a tissue-specific and developmental stage-specific positive transcription factor of *EPSPS* which may act in concert with other factors to achieve transcriptional activation.

Results

Characterization of factors that bind to the *EPSPS* upstream region

To characterize *cis* elements and *trans*-acting factors that are responsible for flower- and stage-specific regulation of *EPSPS*, *in vitro* binding experiments were carried out using nuclear extracts from *Petunia* petals. The extracts were prepared from the petals of mature flowers, the tissue in which the expression of *EPSPS* is the highest. Four fragments from the upstream region were used as probes (−1752 to −1513, −1513 to −1270, −1234 to −1006, −1006 to −823). All four fragments bound factors found in the extract (data not shown). DNase I footprint analysis of these fragments showed four strong binding sites, EP1 (−1170 to −1150), EP2 (−1307 to −1292), EP3 (−1366 to −1345) and EP4 (−1425 to −1408) and several weaker binding sites (Figure 1a). Three of these sites (EP2, EP3, EP4) share some sequence homology, while EP1 is completely different in sequence and contains a perfect palindrome (Figure 2). Figure 1b shows that an excess amount of the EP1 tetramer in the footprint reaction mix abolished protection of all sites including the weaker binding

sites, whereas a tetramer of a mutant EP1 site (EP1m) in which two nucleotides were altered (Figure 2) did not affect the protection pattern. These observations suggest that all the sites are bound by the same factor. It is somewhat surprising that the same factor binds to such divergent sequences despite the apparently strict sequence specificity of the binding, as shown by the loss of protection with only a two base alteration.

To compare the relative affinity of the four strong binding sites, competition experiments with increasing amounts of the EP1 tetramer were carried out. The binding sites EP1, EP2 and EP3 were fully competed by a 50-fold molar excess of the competitor, while the EP4 site was fully competed by a 10- to 25-fold molar excess, indicating that the former three binding sites have a higher affinity for the factor than the latter (Figure 1b). Of the three stronger binding sites, we mainly used EP1 in the following experiments because of its interesting palindromic sequence.

Cloning of a cDNA encoding a sequence-specific DNA-binding protein

To isolate a cDNA encoding a protein that binds specifically to the EP1 site, we constructed a cDNA expression library

Fig. 1. (a) DNase I footprinting of the *EPSPS* 5' upstream region with nuclear extract from mature *Petunia* petal. Two fragments of the *EPSPS* 5' upstream region (−1513 to −1270 and −1234 to −1006) were labeled on either the top or bottom strand by Klenow fill-in with all four [32 P]NTPs. The end-labeled probes were incubated with either nuclear extract (1 μ g) from *Petunia* petal (+) or control buffer (−) and then treated with DNase I. Positions of guanines as determined by the method of Maxam and Gilbert (1980) are shown (G). The bracketed areas indicate the sequences protected from DNase I digestion by incubation with the nuclear extract. (b) Competition of DNase I footprints with *Petunia* petal nuclear extract by a synthetic tetramer of EP1 or mutant EP1. Tetramers of the 21 bp sequences of EP1 and mutant EP1 (shown in Figure 2) were synthesized. An increasing amount of the EP1 tetramer (10-, 25-, 50-, 100-fold excess of the probe) was included in the incubation mix with the nuclear extract before DNase I treatment. A 100-fold excess of the mutant EP1 tetramer was included in a separate incubation mix.

using poly(A) RNA of mature *Petunia* petals. From 1.5×10^6 plaques, we isolated five clones derived from the same gene, whose products show sequence-specific DNA-binding activity to the EP1 tetramer but not the EP1m tetramer. These cDNAs were not full length because their insert size (~ 0.5 kb) was shorter than the size of the corresponding mRNA as estimated by Northern blot analysis (~ 1.5 kb). This was probably due to digestion of the cDNA insert at an internal *EcoRI* site that resulted from the lack of adequate *in vitro* methylation during preparation of the cDNA library. Therefore, we constructed a second cDNA library without using *EcoRI* linkers. By screening the library using the partial cDNA as a probe, a larger clone (*EPFI*) was isolated. The size of the cDNA was 1.5 kb which is

close to the expected size. The 5' end of the cDNA was the same as the 5' end of the mRNA as mapped by primer extension. A genomic clone of the *EPF1* gene was subsequently isolated and sequenced. A TATA box-like sequence was found ~30 bases upstream of the 5' end of the cDNA. Therefore, we conclude that this clone is close to a full length cDNA.

To characterize the binding activity of the protein encoded by *EPF1*, we first attempted to express the protein in *Escherichia coli*. The expressed protein showed some binding activity; however, most of the protein was insoluble and we were unable to solubilize it. Moreover, the high non-specific binding activity in *E. coli* extracts hampered further binding analysis. To circumvent these problems, we expressed the *EPF1* cDNA in a baculovirus expression system. We found that the EPF1 protein obtained from the baculovirus system gave much better binding activity with little non-specific binding activity, although a substantial portion of the protein was still insoluble. Figure 3a shows that the EPF1 protein in the insect cell extract bound to the EP1 tetramer, whereas control cell extract gave no binding to the same probe. The EPF1 protein also bound to the EP1m tetramer but with much lower affinity. Excess amounts of the EP1 tetramer in the reaction mix abolished the binding, while the EP1m tetramer showed only weak competition, demonstrating the specificity of the binding (Figure 3b). In similar competition experiments, the formation of the EPF1-EP1 tetramer complex was weakly competed by an EP3 tetramer, while little or no competition was observed with an EP2 tetramer. By contrast, the EP2 and EP3 tetramers strongly competed with the binding of the nuclear factor to the EP1 sequence (data not shown). Addition of anti-EPF1 IgG raised against EPF1 protein expressed in *E. coli* shifted the DNA-protein complex nearly to the top of the gel (Figure 3c). This result indicates that the binding activity is certainly due to the protein product of the cloned *EPF1* cDNA. By contrast, the antibody did not recognize the DNA-protein complex obtained with petal nuclear extract (data not shown), suggesting that EPF1 is not present in significant amounts in the extract.

Sequence analysis of the full length cDNA of *EPF1* and the genomic clone containing *EPF1* revealed an open reading

Fig. 2. Locations and sequences of binding sites in the *EPSPS* 5' upstream region. Locations of four strong binding sites in the region between -1752 and -823 in the *EPSPS* 5' upstream region are shown as black boxes at the top and their sequences are shown below. Sequences of EP2 and EP4 are shown in the antisense orientation to emphasize sequence homology among EP2, EP3 and EP4. Altered nucleotides of the mutant EP1 (EP1m) are underlined. The palindromic structure in EP1 is indicated by arrows. Conserved sequences in EP2rev, EP3 and EP4rev are boxed.

Fig. 3. Gel shift assay with EPF1 protein expressed in a baculovirus system. (a) The [32 P]EP1 tetramer (lanes 1, 2 and 3) or the [32 P]EP1m tetramer (lanes 4, 5 and 6) were incubated with buffer (lanes 1 and 4), cell extract (4 μ g protein) from Sf9 infected with recombinant virus (RE) (lanes 2 and 5) or wild-type virus (WE) (lanes 3 and 6) and analyzed by gel shift. (b) Competition in the gel shift assay. Excess amounts of the EP1 tetramer (lanes 2 and 3) or the EP1m tetramer (lanes 4 and 5) [10-fold (lanes 2 and 4) or 100-fold (lanes 3 and 5)] were incubated with the extract (RE) and the [32 P]EP1 tetramer. (c) Recognition of the DNA-protein complex with anti-EPF1 IgG. Control buffer (—) (lane 1), anti-EPF1 (*E. coli*) IgG (lane 2) or preimmune IgG (lane 3) was included in the gel shift reaction with the extract (RE) and the [32 P]EP1 tetramer.

frame encoding 281 amino acids (Figure 4a). The deduced protein sequence contains two repeats of the canonical Cys₂/His₂ zinc finger motif (Cys-X₂-Cys-X₃-Phe-X₅-Leu-X₂-His-X₃-His). Figure 4b shows a sequence alignment of the zinc finger motifs in EPF1 and one of the motifs in the Kruppel protein (Rosenberg *et al.*, 1986). The two zinc fingers in EPF1 share extensive homology to each other, suggesting that the repeats were created by an internal duplication. In particular, six consecutive amino acids which form the right face of the finger structure are identical (Figure 4b). Recent crystallographic studies showed that the residues in these regions contact DNA (Pavletich and Pabo, 1991). These regions might be determinants of DNA-binding specificity, assuming that the two finger motifs have the same binding specificities. Unlike many other known zinc finger proteins, the two fingers in EPF1 are widely separated. An extremely serine/threonine-rich stretch is located near the N terminus of the protein and just upstream of this stretch there is a cluster of basic amino acids.

Zinc ion requirements for the DNA-binding activity of the EPF1 protein were demonstrated in two ways. Incubation of the protein with a high concentration of EDTA (5, 10 and 50 mM) abolished the binding activity (Figure 5a). Moreover, the zinc chelating agent 1,10-phenanthroline,

which inhibits the DNA-binding activity of *Xenopus* zinc finger transcription factor TFIIIA (Hanas *et al.*, 1983), completely abolished the DNA-binding activity of EPF1 (Figure 5a). By contrast, the DNA-binding activity in the petal nuclear extract was not inhibited by either treatment (Figure 5b). These results, together with the different mobilities of the DNA-protein complex in gel shift assays, suggest that there is more than one factor in the nuclear extract that can interact with the *EPSPS* upstream sequence and that EPF1 does not represent a major component among them. The inability of the anti-EPF1 IgG to recognize the nuclear factor is consistent with this view. In a control experiment, the binding activity of TGA1a, a DNA-binding protein cloned from tobacco (Katagiri *et al.*, 1989), was not inhibited by either treatment (Figure 5c).

To determine whether a domain that contains each of the finger motifs is sufficient for DNA binding, the N-terminal portion (1–142) and the C-terminal portion (143–281) of the EPF1 protein were expressed independently in *E. coli*. Gel shift experiments revealed that the C-terminal portion of the protein binds to the EP1 site but only very weakly to the mutant EP1 site, indicating that one finger motif is sufficient for specific DNA-binding activity (data not shown). The N-terminal portion of the protein binds to neither EP1

Fig. 4. (a) Nucleotide sequence and the deduced amino acid sequence of *EPF1*. The sequence of the full length cDNA and the 5' upstream sequence of the genomic DNA (nucleotides -2091 to -1) are shown. The sequences from 1 to 1120 were determined from cDNA clones and those from -2091 to +515 were from genomic DNA. The nucleotide number 1 represents the transcription start site. A TATA box-like sequence is underlined. The ORF which begins at nucleotide 91 and ends at 934 encodes a protein of 281 amino acids. The two zinc finger motifs are double underlined and the serine/threonine-rich sequence is shown in italic. Potential nuclear targeting sequences are shown by asterisks. The nucleotide and amino acid sequences of *EPF1* were deposited in the GenBank Library, accession number X60700 P.HYBRIDA,EPF1 DNA. (b) Sequence alignment of the zinc finger motifs. The upper two sequences show the two finger motifs in EPF1. EPF1-I represents the N-terminal motif and EPF1-II the C-terminal one. The lower one represents one of the zinc fingers in the *Kruppel* protein (Rosenberg *et al.*, 1986). The six consensus amino acids of the Cys₂/His₂ type zinc finger motif are boxed. A stretch of six consecutive amino acids which are identical in the two zinc finger motifs in EPF1 are indicated by asterisks.

nor any subfragments of the *EPSPS* 5' upstream region (data not shown). However, we cannot rule out the possibility that the N-terminal portion of the protein was improperly folded when expressed in *E. coli*.

Flower-specific and developmental stage-specific expression of *EPF1*

Northern blot analysis of RNA from various tissues of mature *Petunia* plants showed that the expression of *EPF1* is very high in the petal, low in floral tube and undetectable in leaf, stem and root (Figure 6a). This pattern of tissue specificity is very similar to that of *EPSPS*, except that the latter shows low but detectable expression in root and stem. The expression of *EPF1* in the petal is very low in floral buds (day 0) and increases dramatically during flower opening until the flower reaches a mature stage (day 3) (Figure 6b). This pattern of stage-specific expression parallels that of *EPSPS*. The same filter was sequentially hybridized to the cDNA of chalcone synthase from *Antirrhinum majus* (Sommer and Saedler, 1986) and cDNA of ATP synthase β -subunit from *Nicotinia plumbaginifolia* (Boutry and Chua, 1985). In contrast to *EPF1* and *EPSPS*, the chalcone synthase gene is activated at the very early stage of flower opening and then rapidly turned off, and the expression of the ATP synthase β -subunit gene is basically constitutive, with somewhat higher expression in the petal; its expression level increases only slowly during flower opening. The parallel expression of *EPF1* and *EPSPS* strongly suggests that *EPF1* is a positive transcription factor of *EPSPS*.

EPF1 belongs to a small gene family

Southern blot analysis of *Petunia* genomic DNA using *EPF1* cDNA as a probe, revealed one strong band and one or two

weaker bands in most of the restriction enzyme digests (Figure 7). This observation suggests that there are one or two related *EPF1* genes in the *Petunia* genome.

EPF1 5' upstream sequences confer a tissue-specific expression pattern

We have obtained a genomic clone which contains the entire *EPF1* coding sequence and its upstream sequence by screening a genomic DNA library constructed from *Petunia* DNA. The parallel mRNA expression pattern of *EPSPS* and *EPF1* prompted us to study the *EPF1* promoter activity in transgenic *Petunia* plants. To this end, we constructed a translational fusion gene by fusing 2 kb of the *EPF1* 5' upstream sequence including the first 27 bp of the coding sequence to the bacterial β -glucuronidase (GUS) coding sequence. This chimeric construct was introduced into *Petunia* by *Agrobacterium*-mediated transformation. A fluorometric determination of GUS activity in mature transgenic plants revealed that the GUS gene was highly expressed in petals, but only low level activity was detected in leaves and stems (Table I). Moreover, there was only very low expression in petals of floral buds but the expression increased ten times during flower opening (Table I). These results are consistent with those obtained by Northern analyses (Figure 6).

Unexpectedly, substantial activity was detected in roots (Table I). This observation is in contrast to the results of Northern analysis in which no expression of *EPF1* was

Fig. 5. Zinc ion requirement of *EPF1* for the DNA binding activity. The *EPF1* protein produced in a baculovirus system (a), *Petunia* nuclear factor (b) and tobacco factor TGA1a (c) were treated with control buffer (-), 5 mM, 10 mM or 50 mM EDTA, or 1 mM 1,10-phenanthroline for 20 min on ice. The treated proteins or extracts were subjected to gel shift analysis with [32 P]EP1 tetramer for *EPF1* and nuclear factor or [32 P]AS1 tetramer (the ASF-1 binding site from the 35S promoter) for TGA1a (Katagiri *et al.*, 1989).

Fig. 6. Northern blot analysis of *Petunia* RNA. (a) Total RNA (10 μ g) from various tissues. (b) Total RNA (10 μ g) from petal at various stages (0–3 days) of flower opening. The filter was sequentially probed with [32 P]-labeled *EPF1* cDNA from *Petunia*, *EPSPS* cDNA from *Petunia* (Gasser *et al.*, 1988), chalcone synthase cDNA from *A. majus* (Sommer and Saedler, 1986) and ATP synthase β -subunit cDNA from *N. plumbaginifolia* (Boutry and Chua, 1985).

Fig. 7. Southern blot analysis of *Petunia* genomic DNA probed with EPF1 cDNA. Genomic DNA of *Petunia* (10 µg each) was digested with the indicated restriction enzymes and run through a 0.8% agarose gel. The gel was blotted and the nitrocellulose filter was hybridized to a 5' fragment of EPF1 cDNA (nucleotides 1–424).

Table 1. β-Glucuronidase enzymatic activities in transgenic plants containing EPF1 promoter–GUS fusion constructs

Plant	Petal (days)				Tube	Leaf	Stem	Root
	0	1	2	3				
1	36	579	1503	6423	612	4	47	3240
2	621	2469	4593	13824	1630	828	879	5413
3	302	4752	4824	10140	2030	32	486	4512

Fluorometric analysis of GUS activity was performed as described by Benfey *et al.* (1989). The activities are given in pmol of 4-methylumbelliferone per mg protein per min. The results of three independent transgenic plants are shown.

detected in roots. One possibility is that *EPF1* sequences 3' of the fusion site might affect the expression of the gene by suppressing the transcription or destabilizing the transcript, resulting in reduced expression of *EPF1* in roots.

To see if the *EPF1* promoter conferred expression in specific cell types, histochemical staining of flower sections was carried out. In unopened flower, very little staining was observed in any cell type (Figure 8a). In mature flower, all cell types were stained in the limb of the petal (Figure 8b), only the upper epidermal cells were stained in the upper part of the floral tube (Figure 8c), and no staining was observed in the lower part of the tube (Figure 8d). These results indicate that the *EPF1* 5' upstream sequence confers a cell-type expression pattern remarkably similar to that of the *EPSPS* 5' upstream sequence in flower (Benfey and Chua, 1989).

The finding that *EPSPS* 5' upstream sequences confer activity in the root cortex, trichome and meristematic region in the stem of *Petunia* seedlings (Benfey *et al.*, 1990) led us to examine the expression of the *EPF1* 5' upstream region in seedlings. Seeds from the transgenic plants described

above were germinated and stained as a whole mount at seven days after germination. As shown in Figure 8e–g, the *EPF1* 5' upstream region gave expression in the root cortex (e), trichome (f) and meristematic region (g) in a pattern similar to that of the *EPSPS* 5' upstream region. In addition to these tissues, staining was observed also in root vascular tissue and root tip, and in leaf tip (data not shown).

The concomitant activation of *EPSPS* and *EPF1* in the same cell types suggests that *EPF1* is a positive transcription factor of *EPSPS*. To examine this possibility, multiple copies of the binding site EP1 were fused to the GUS coding sequence and the chimeric construct was introduced into *Petunia*. However, neither four nor 16 copies of the elements placed upstream of the basic TATA element of cauliflower mosaic virus 35S promoter (Benfey and Chua, 1989) conferred activity to the GUS reporter gene (data not shown). Taken together with previous *cis*-element analyses (Benfey *et al.*, 1990), these results suggest that one kind of *cis* element alone is not sufficient but a combination of this element with other kinds of *cis* elements is necessary for activation. It is possible that the abundant factor detected in nuclear extracts is one of the other factors that in combination with *EPF1* can activate transcription from the *EPSPS* promoter.

Discussion

Interaction of *EPSPS* promoter and factors found in nuclear extract

We have previously shown that sequences between –1752 and –823 of the *EPSPS* 5' upstream region are responsible for developmental stage-specific expression in *Petunia* petals (Benfey and Chua, 1989). *In vitro* binding analysis using nuclear extracts from *Petunia* petals revealed numerous binding sites spanning the entire region. These binding sites appear to interact with the same factor. Deletion analyses of the promoter in transgenic plants revealed a complex set of regulatory sequences. When the –1752 to –823 region was dissected into two pieces of ~500 bp (–1752 to –1270 and –1234 to –823), each of the smaller sequences conferred an expression pattern similar to that of the parental sequence but at a lower level. These results indicated that there is a functional repeat within the –1752 to –823 region. When the 500 bp fragments were cut down to even smaller fragments (–1752 to –1513, –1513 to –1270, –1234 to –1006, –1006 to –823), these sequences were inactive despite the presence of several copies of the factor binding sites within each of these fragments (Benfey *et al.*, 1990). Taken together, these observations suggest that there are numerous functionally redundant *cis* elements in the –1752 to –823 region but that a minimum number or a combination of functionally different *cis* elements is necessary for the activation of transcription. Unlike other plant promoter elements so far reported, these *cis* elements in the *EPSPS* 5' upstream region are located relatively far upstream from the transcription start site. Carey *et al.* (1990) have reported that for the yeast Gal4 binding site, the enhancer activity becomes weaker as the distance of the *cis* element from the transcription start site increases; however, the enhancer activity can be strengthened by increasing the number of the *cis* elements. These observations are consistent with the location and multiple copy requirement of the *cis* elements in the *EPSPS* 5' upstream region, although the significance is unknown.

Fig. 8. Histochemical staining of flower sections and seedlings of transgenic *Petunia* plants containing an *EPF1* 5' upstream sequence—GUS translational fusion construct. The histochemical staining was carried out as described by Benfey and Chua (1989). (a) Limb of an unopened petal (day 1). (b) Limb of a mature petal (day 3). (c) Upper part of a mature floral tube (day 3). (d) Lower part of a mature floral tube (day 3). (e) Seedling root. (f) Seedling trichome. (g) Meristematic region in seedling stems. UE, upper epidermis; LE, lower epidermis; C, root cortex; T, trichome; M, meristem.

A plant *Cys₂/His₂* type zinc finger DNA-binding protein

We have isolated a full length cDNA clone encoding a sequence specific DNA-binding protein by screening a

cDNA expression library constructed from *Petunia* petal RNA. The deduced amino acid sequence contains two repeats of a *Cys₂/His₂* zinc finger motif, which has been found in a number of transcription factors such as *Krüppel*

(Rosenberg *et al.*, 1986) and *hunchback* (Tautz *et al.*, 1987) (*Drosophila* transcription factors of the gap class), Sp1 (Kadonaga *et al.*, 1987) (a mammalian enhancer binding protein), PRD II-BF1 (Fan and Maniatis, 1990) (an enhancer-binding protein of human IFN- γ promoter) and MBP-1 (Baldwin *et al.*, 1990) (an enhancer-binding protein of the class I major histocompatibility complex and kappa immunoglobulin genes). Among these, the zinc finger motifs in EPF1 have the highest sequence similarity to those of *Krüppel*, a transcription factor that is responsible for the periodic expression of certain pair-rule genes in the early embryo of *Drosophila*. EPF1 is the first Cys₂/His₂ type zinc finger protein identified from plants.

The two zinc finger motifs in EPF1 are separated by 61 amino acids. This feature is unusual since zinc finger motifs are normally adjacent to one another with a well conserved H-C link sequence consisting of seven amino acids (Rosenberg *et al.*, 1986). Recently, zinc finger proteins with widely separated Cys₂/His₂ finger motifs devoid of the H-C link have been reported. These are Suvar(3)7 (Reuter *et al.*, 1990), teashirt (Fasano *et al.*, 1991) and TRS-1 (Pays and Murphy, 1987). Reuter *et al.* (1990) postulated that widely separated zinc fingers might bind large domains of DNA, resulting in the control of gene expression by opening or closing contiguous stretches of DNA. Suvar(3)7 and teashirt share common features besides widely separated finger motifs, in that the distance between the two histidines is five amino acids, whereas many of the Cys₂/His₂ class zinc finger proteins that are transcription factors have an inter-histidine distance of three amino acids. In EPF1, the inter-histidine distance is three amino acids, making it more similar to the latter group. An extremely serine/threonine-rich sequence is located near the N-terminus of the protein. Similar serine/threonine-rich regions are also present in the human transcription factor Sp1 and the sequences are reported to be heavily phosphorylated upon virus infection of tissue culture cells (Jackson *et al.*, 1990). A cluster of basic amino acids which may be a nuclear targeting signal is located just upstream of the serine/threonine-rich sequence. A close association of nuclear targeting sequence and phosphorylation sites has been reported for some DNA-binding proteins (Scheidtmann *et al.*, 1982). Many activator proteins have been demonstrated to have characteristic activation domains. Recently, a hydroxylated amino acid-rich region has also been shown to be an activation domain in the pituitary-specific transcription factor GHF-1 (Theill *et al.*, 1989). The serine/threonine-rich region in EPF1 is a candidate for an activation domain.

EPF1 may be a positive transcription factor of EPSPS

Northern blot analysis and promoter function analysis with the EPF1 promoter-GUS fusion construct revealed that the expression pattern of EPF1 is very similar to that of EPSPS in flowers at different developmental stages and also in young seedlings. This concomitant activation of both genes suggests that EPF1 is a positive transcription factor of EPSPS. If EPF1 is a simple activator protein, its binding sites are expected to confer an activity on a reporter gene. However, multiple copies of the binding site, EP1, were unable to activate the GUS reporter gene in transgenic plants. This observation, taken together with results from previous deletion analyses (Benfey *et al.*, 1990), suggest that a combination of the EPF1 binding sites with other *cis* elements is necessary for transcription activation. Among several

examples of such combinatorial regulation is that of the cauliflower mosaic virus 35S promoter. When the promoter is dissected into subdomains and selected subdomains are combined, they confer expression which is not detected from the isolated subdomains (Benfey *et al.*, 1989). This combinatorial regulation is likely to be mediated by *trans* factors which bind to these *cis* elements. Another example of combinatorial regulation has been reported for the regulation of *bronze*, a structural gene in the maize anthocyanin pathway (Goff *et al.*, 1990. Roth *et al.*, 1991). Two *trans*-acting factors, C1 and B, interact with the promoter sequence of *bronze* and the binding of both factors is required for the activation of *bronze*.

The regulation of the promoter of EPSPS seems to be quite complex. The promoter appears to consist of multiple kinds of *cis* elements and transcription regulation seems to involve more than one *trans* factor binding to the *cis* elements. We postulate that EPF1 represents one of several factors that are required to bind simultaneously to *cis* elements within the EPSPS promoter to activate transcription. To understand fully the activation from the promoter, characterization and isolation of other *cis* elements and *trans* factors will be necessary. Although we have only indirect evidence that EPF1 is involved in the transcriptional regulation of EPSPS, this factor is interesting because of its highly specific expression pattern, suggesting that it may play an important role in the activation of genes late in flower development.

Materials and methods

Preparation of nuclear extracts, DNase I footprinting and gel shift assays

Nuclear extracts from *Petunia* petal were prepared essentially as described by Green *et al.* (1987) except that the petals were frozen in liquid nitrogen and ground in a prechilled mortar on dry ice by a modification of a procedure used for nuclear run-on assays (Scheidtmann *et al.*, 1982). The use of frozen material is crucial for *Petunia* because its nuclei are so fragile that they are easily broken when unfrozen material is used. DNase I footprint experiments were carried out with an end-labeled probe under conditions similar to those described in Lam and Chua (1989). The positions of guanines were determined by the method of Maxam and Gilbert (1980). Gel shift assays were carried out in 0.7% agarose/3% polyacrylamide gels under conditions similar to those described in Mikami *et al.* (1987).

Construction of cDNA libraries and expression library screening

Total RNA was prepared from the petal of mature *Petunia* (cv. Mitchell diploid) flowers according to Fromm *et al.* (1985) and poly(A) RNA was selected by oligo(dT)-cellulose column chromatography (Aviv and Leder, 1972). Two cDNA libraries were constructed from the same RNA by different methods. The first was constructed in λ gt11 using *Eco*RI linkers to ligate the cDNA to the vector and *Eco*RI methylase for the protection of the internal *Eco*RI sites (Huyne *et al.*, 1985). This library was used for screening for DNA binding activity; however, the clones obtained were not full length because of imperfect protection of internal *Eco*RI sites. A second library was then constructed in λ ZAPII (Stratagene) using adaptors (Wu *et al.*, 1987) at the termini of cDNA instead of *Eco*RI linkers. A full length cDNA clone encoding EPF1 was isolated from this library. The cDNA library was screened for sequence-specific DNA-binding activity by the method of Singh *et al.* (1988). A tetramer of the EP1 sequence was end-labeled and used as a probe and a tetramer of mutant EP1 was used as a negative control probe. The sequence containing the full length EPF1 cDNA was excised *in vivo* from λ ZAPII in the plasmid pBluescript (Stratagene). Single-stranded recombinant DNA induced from these clones was sequenced by the dideoxy chain termination method (Rosenberg *et al.*, 1986).

Expression of EPF1 protein in a baculovirus expression system

The *Xba*I-SpeI fragment that contains EPF1 coding sequences was excised from the pBluescript SK(-) vector and inserted into the *Nhe*I site downstream of the polyhedron promoter in the pBlueBac transfer plasmid (Vialard *et al.*,

1990). Transfection and selection of recombinant virus was carried out according to Summers and Smith (1987). The purified recombinant virus was infected into Sf9 cells and total cell extract was prepared according to Paul *et al.* (1990).

Construction of genomic DNA library and isolation of the EPF1 clone

High molecular weight DNA was prepared from *Petunia* by lysis with sarcosyl and a CsCl gradient (Boutry and Chua, 1985). A genomic library was constructed from *Sau3A1* partially digested DNA (13–17 kb) in EMBL3 λ phage (Stratagene). The library was grown on *E. coli* strain ER1647 and screened using the entire *EPF1* cDNA.

Northern and Southern blot analyses

Total RNA was extracted as described by Kuhlemeier *et al.* (1988) except that 2% tri-isopropyl-naphthalene sulphonate was included in the extraction buffer. 10 μ g each of total RNA from different tissues of *Petunia* and from petals at different developmental stages was denatured in 50% formamide and 6% formaldehyde. The transcripts were separated on a 1% agarose gel containing 6% formamide. After transfer onto a nitrocellulose filter, the filter was hybridized with cDNAs labeled with 32 P by random priming. Restriction digest of *Petunia* genomic DNA was electrophoresed through a 0.8% agarose gel and blotted onto nitrocellulose. The filter was probed with a 32 P-labeled partial *EPF1* cDNA.

Construction of the EPF1 promoter – GUS translational fusion and Petunia transformation

The *XbaI*–*Clal* (filled in) fragment (2 kb) containing the *EPF1* upstream sequence and first 27 bp of the coding sequence was inserted into the *XbaI*–*SmaI* site of PBI101.2 (Clontech Laboratories) in-frame to the GUS coding sequence. Transformation of *Petunia* was carried out as described by Benfey *et al.* (1989).

Acknowledgements

This work was supported by a grant from Monsanto Co. and The Ministry of Agriculture, Forestry and Fishery of Japan. P.N.B. was supported by a Helen Hay Whitney post-doctoral fellowship.

References

- Aviv, H. and Leder, P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408–1412.
- Baldwin, A.S., Jr, LeClair, K.P., Singh, H. and Sharp, P.A. (1990) *Mol. Cell. Biol.*, **10**, 1406–1414.
- Benfey, P.N. and Chua, N.-H. (1989) *Science*, **244**, 174–181.
- Benfey, P.N., Ren, L. and Chua, N.-H. (1989) *EMBO J.*, **8**, 2195–2202.
- Benfey, P.N., Takatsui, H., Ren, L., Shah, D.M. and Chua, N.-H. (1990) *Plant Cell*, **2**, 849–856.
- Boutry, M. and Chua, N.-H. (1985) *EMBO J.*, **4**, 2159–2165.
- Carey, M., Leatherwood, J. and Ptashne, M. (1990) *Science*, **247**, 710–712.
- Evans, R.M. and Hollenberg, S.M. (1988) *Cell*, **52**, 1–3.
- Fan, C.-M. and Maniatis, T. (1990) *Genes Dev.*, **4**, 29–42.
- Fasano, L., Roeder, L., Core, N., Alexandre, E., Vola, C., Jacq, B. and Kerridge, S. (1991) *Cell*, **64**, 63–79.
- Fromm, H., David, M., Fluhr, R. and Edelman, M. (1985) *EMBO J.*, **4**, 291–295.
- Gasser, C.S., Winter, J.A., Hironaka, C.M. and Shah, D.M. (1988) *J. Biol. Chem.*, **263**, 4280–4289.
- Goff, S.A., Klein, T.M., Roth, B.A., Fromm, M.E., Cone, K.C., Radicella, J.P. and Chandler, V.L. (1990) *EMBO J.*, **9**, 2517–2522.
- Green, P.J., Kay, S.A. and Chua, N.-H. (1987) *EMBO J.*, **6**, 2543–2549.
- Hanas, J.S., Hazuda, D.J., Bogenhagen, D.F., Wu, F.Y.-H. and Wu, C.W. (1983) *J. Biol. Chem.*, **258**, 14120–14125.
- Huynh, T.V., Young, R.A. and Davis, R.W. (1985) In Glover, D.M. (ed.), *DNA Cloning—A Practical Approach*. IRL Press, Oxford, Vol. 1, pp. 49–78.
- Jackson, S.P., MacDonald, J.J., Lees-Miller, S. and Tjian, R. (1990) *Cell*, **63**, 155–165.
- Kadonaga, J.T., Carner, K.R., Masiarz, F.R. and Tjian, R. (1987) *Cell*, **51**, 1079–1090.
- Katagiri, F., Lam, E. and Chua, N.-H. (1989) *Nature*, **340**, 727–730.
- Kuhlemeier, C., Fluhr, R. and Chua, N.-H. (1988) *Mol. Gen. Genet.*, **212**, 405–411.
- Lam, E. and Chua, N.-H. (1989) *Plant Cell*, **1**, 1147–1156.
- Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.*, **65**, 499–560.
- Mikami, K., Tabata, T., Takefumi, K., Nakayama, T. and Iwabuchi, M. (1987) *FEBS Lett.*, **223**, 273–278.

- Paul, J.I., Tavare, J., Denton, R.M. and Steiner, D.F. (1990) *J. Biol. Chem.*, **265**, 13074–13083.
- Pavletich, N. and Pabo, C. (1991) *Science*, **252**, 809–817.
- Pays, E. and Murphy, N.B. (1987) *J. Mol. Biol.*, **197**, 147–148.
- Reuter, G., Giarre, M., Farah, J., Gausz, J., Spierer, A. and Spierer, P. (1990) *Nature*, **344**, 219–223.
- Rosenberg, U.B., Schroder, C., Preiss, A., Kienlin, A., Cote, S., Riede, I. and Jackle, H. (1986) *Nature*, **319**, 336–339.
- Roth, B.A., Goff, S.A., Klein, T.M. and Fromm, M.E. (1991) *Plant Cell*, **3**, 317–325.
- Scheidtmann, K.H., Echle, B. and Walter, G. (1982) *J. Virol.*, **44**, 116–133.
- Singh, H., Lebowitz, J.H., Baldwin, A.S., Jr and Sharp, P.A. (1988) *Cell*, **52**, 415–423.
- Sommer, H. and Saedler, H. (1986) *Mol. Gen. Genet.*, **202**, 429–434.
- Sommer, H., Beltran, J.-P., Huijser, P., Heike, P., Lonnig, W.-E., Saedler, H. and Schwarz-Sommer, Z. (1990) *EMBO J.*, **9**, 605–613.
- Steinrucken, H.C. and Amerhein, N. (1980) *Biochem. Biophys. Res. Commun.*, **94**, 1207–1212.
- Summers, M.D. and Smith, G.E. (1987) *A Manual of Methods of Baculovirus Vectors and Insect Cell Culture Procedures*. Texas Agriculture Experiment Station bulletin no. 1555.
- Tautz, D., Lehman, R., Schnurich, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K. and Jackle, H. (1987) *Nature*, **327**, 383–389.
- Theill, L.E., Castrillo, J.L., Wu, D. and Karin, M. (1989) *Nature*, **342**, 945–948.
- Vialard, J., Lalumiere, M., Verent, T., Briedis, D., Alkhatib, G., Henning, D., Levin, D. and Richardson, C. (1990) *J. Virol.*, **64**, 37–50.
- Wu, R., Wu, T. and Ray, A. (1987) *Methods Enzymol.*, **152**, 343–349.
- Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A. and Meyerowitz, E.M. (1990) *Nature*, **346**, 35–39.

Received on July 1, 1991; revised on October 23, 1991