Identification of *Cis*- and *Trans*-Regulatory Factors Controlling the Expression of the C_4 Phospho*enol*pyruvate Carboxylase Gene of the C_4 Dicot *Flaveria trinervia*

Inaugural-Dissertation

zur

Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine Universität Düsseldorf

vorgelegt von

Meryem Akyildiz

aus Duisburg

Düsseldorf im Juli 2007

Aus dem Institut für Entwicklungs- und Molekularbiologie der Pflanzen
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent:	Prof. Dr. P. Westhoff
Koreferent:	Prof. Dr. G. Groth

Tag der mündlichen Prüfung: 24.09.2007

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 02.07.2007

Contents

3) Meryem Akyildiz, Ming Chang Tsai, Claus Seidel and Peter Westhoff (2007). **Basic leucine zipper proteins interact with MEM1, the mesophyll specificity *cis*-regulatory element of the C_4 phospho*enol*pyruvate carboxylase gene of *Flaveria trinervia*.**

I. General Introduction

I.1 C_4 Photosynthesis

Photosynthesis is the physico-chemical process by which plants, algae and photosynthetic bacteria use light energy to drive the synthesis of organic compounds. In plants, algae and certain types of bacteria, the photosynthetic process results in the release of molecular oxygen and the removal of carbon dioxide (CO_2) from the atmosphere that is used to synthesize carbohydrates. Based on the number of carbon compounds in the first stable molecule, formed from the carbon dioxide fixation process, land plants can be divided into two major photosynthetic types, namely C_3 and C_4.

In plants with C_3 photosynthesis the fixation of carbohydrates is catalyzed by the enzyme ribulose-1,5-bisphosphate carboxylase oxygenase in which carbon dioxide is combined with a five-carbon sugar, ribulose 1,5-bisphosphate, to yield two molecules of a three-carbon compound, 3-phosphoglycerate, hence the name C_3 photosynthesis. However, in low atmospheric carbon dioxide concentrations C_3 photosynthesis is impaired by the lack of carbon dioxide as a substrate in addition to photorespiration (Furbank & Badger, 1983; Ogren, 1984; Furbank & Taylor, 1995; von Caemmerer & Furbank, 2003). At low CO_2 concentrations ribulose-1,5-bisphosphate is oxygenated by the oxygenating function of ribulose-1,5-bisphosphate carboxylase oxygenase producing one molecule of 3-phosphoglycerate and one molecule of phosphoglycolate (Andrews and Lorimer, 1987). Phosphoglycolate is metabolically useless and toxic if it accumulates in the cell. The conversion of phosphoglycolate to useful metabolites is therefore essential for land plants. During this metabolic process, called photorespiration, CO_2 and NH_3 are produced and ATP and reducing equivalents are consumed, thus making photorespiration a wasteful process (Furbank & Taylor, 1995) resulting in the loss of 25% of the carbon entering the pool of phosphoglycolate molecules (Ogren, 1984; Leegood *et al.*, 1995). As a result water- and nitrogen-use efficiencies are low (Black, 1973; Ehleringer & Monson, 1993).

C_4 plants assimilate atmospheric carbon dioxide through a sophisticated addition to the ancient C_3 photosynthetic pathway. This enables C_4 plants to cope well with high temperatures, high light intensities, drought and increased salinity. Important crop plants like maize, sugar cane and sorghum belong to this photosynthetic type. The high photosynthetic capacity of C_4 plants is due to their unique mode of carbon assimilation which involves two different photosynthetic cell types, mesophyll and bundle-sheath cells. Recently this

paradigm, the requirement of two photosynthetic active cell types in the C_4 photosynthesis cycle, was disproved by the identification of single-celled C_4 photosynthesis (Voznesenskaya *et al.*, 2001; Voznesenskaya *et al.*, 2002; Edwards *et al.*, 2004). However, this single-celled C_4 photosynthesis is an exceptional case rather than representing a common type of C_4 photosynthesis.

In C_4 photosynthesis, atmospheric carbon dioxide is first hydrated to bicarbonate (HCO_3^-) by the enzyme carbonic anhydrase (Hatch & Burnell, 1990; Badger & Price, 1994) in the cytosol of mesophyll cells and subsequently fixed into the C_4 acid oxaloacetate with the three-carbon substrate phospho*enol*pyruvate through the enzyme phospho*enol*pyruvate carboxylase (PEPC) (Figure 1). Oxaloacetate is rapidly reduced to malate in the mesophyll chloroplasts by NADP-malate dehydrogenase or transaminated to aspartate in the cytosol by glutamate-aspartate aminotransferase, depending on the C_4 acid-decarboxylating mechanism of C_4 plants (Hatch, 1987). These C_4 compounds are then transported to the bundle-sheath cells where they are decarboxylated to release carbon dioxide by one of the three decarboxylation enzymes: NADP-malic enzyme, NAD-malic enzyme or PEP carboxykinase (Kanai &Edwards, 1999). The consequence of this decarboxylation is that carbon dioxide is concentrated within the bundle-sheath cells. The released carbon dioxide is re-assimilated by the bundle-sheath-specific enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase in the reductive pentose phosphate pathway (Calvin cycle) (Hatch, 1987; Dai *et al.*, 1993; Furbank & Taylor, 1995). The decarboxylation reaction also leaves pyruvate, which is transported back to the mesophyll cells and phosphorylated, in a reaction catalysed by the pyruvate orthophosphate dikinase enzyme, to regenerate phospho*enol*pyruvate at the cost of a phosphorus group and one ATP molecule.

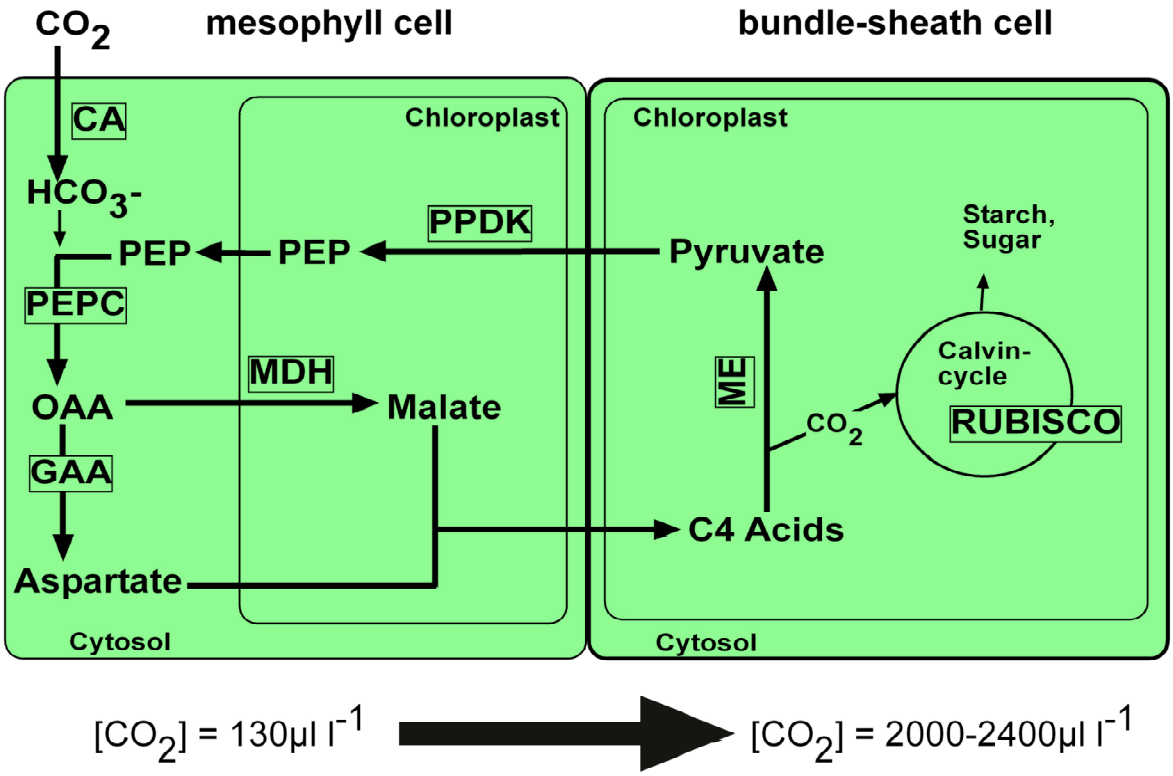


Figure 1. **Schematic presentation of the CO_2 concentrating mechanism in the NADP-malic enzyme type C_4 plant *Flaveria trinervia* (Hatch, 1987).**
CA: carbonic anhydrase; GAA: glutamate-aspartate aminotransferase; MDH: NADP-malate dehydrogenase; ME: NADP-malic enzyme; PEPC: phospho*enol*pyruvate carboxylase; PEP: phospho*enol*pyruvate; PPDK: pyruvate phosphate dikinase; RUBISCO: ribulose-1,5-bisphosphate carboxylase oxygenase.

As a consequence of this carbon dioxide concentration, also referred to as CO_2 pump, in the chloroplasts of bundle-sheath cells ribulose-1,5-bisphosphate carboxylase/oxygenase operates in C_4 plants at high carbon dioxide/oxygen ratios. The competitive inhibition of the carboxylating function of ribulose-1,5-bisphosphate carboxylase/oxygenase by oxygen, which becomes prominent at higher temperatures, is largely excluded and C_4 plants show drastically reduced rates of photorespiration. The CO_2 pump ensures high rates of photosynthesis even at low atmospheric carbon dioxide concentrations, thus C_4 plants are capable to limit the opening of their stomata and thereby minimize water loss through transpiration. As the carbon dioxide pump delivers saturating carbon dioxide concentrations to the site of ribulose-1,5-bisphosphate carboxylase/oxygenase high photosynthetic rates are maintained with three to six times less enzyme than is required in C_3 species. This is reflected in a higher nitrogen use efficiency (Ehleringer & Monson, 1993; Long, 1999).

To achieve an effective CO_2 concentration mechanism in the C_4 photosynthesis the distance between mesophyll and bundle-sheath cells has to decline to allow for rapid diffusion of metabolites (Raghavendra, 1980; Ehleringer *et al.*, 1997). The latter is accomplished by reducing the interveinal distance. Thus the C_4 leaves are highly vascularized, with veins often separated by as few as four photosynthetic active mesophyll cells (Nelson & Langdale, 1992). The leaf thickness is limited in C_4 eudicot plants and is usually smaller than in the leaves of C_3 plants (Hattersley, 1992; McKown & Dengler, 2007). The proportion of the photosynthetic carbon reduction tissue (bundle-sheath cells) area is significantly higher in C_4 dicots than in C_3 relatives. The enhanced photosynthetic carbon reduction tissue volume of C_4 species reflects the physiological requirement for accommodating the numerous and large organelles that are involved in the C_4 cycle and are required for generating high levels of CO_2 around ribulose-1,5-bisphosphate-carboxylase/oxygenase (Hattersley, 1984; Dengler & Nelson, 1999; Kanai & Edwards, 1999).

In C_4 monocots and dicots the vascular system provides the framework around which photosynthetic mesophyll and bundle-sheath cells are arranged. This so called Kranz anatomy (Haberlandt, 1904) is fully evolved before complete C_4 biochemistry is achieved (McKown & Dengler, 2007). This specialized Kranz anatomy is characterized by a well developed layer of bundle-sheath cells surrounding the leaf veins (Figure 2) and containing large quantities of mitochondria, chloroplasts and peroxisomes (Brown & Hattersley, 1989). However, the variants observed in the dicot family *Chenopodiaceae,* e.g. *Binertia cycloptera* and *Borszczowia aralocaspica,* have unusual chlorencyma and lack Kranz anatomy (Freitag & Stichler, 2000; Freitag & Stichler, 2002; Voznesenskaya *et al.*, 2002). The chloroplasts in mesophyll and bundle-sheath cells in NADP-malic enzyme type C_4 species are dimorphic due to differences in enzymes, location of starch (predominantly in the bundle-sheath chloroplasts) and in chloroplast ultrastructure (Woo *et al.*, 1970; Edwards *et al.*, 2001). High grana-containing chloroplasts have higher photosystem II activities and linear electron flow (producing NADPH and ATP), whereas low-grana containing chloroplasts are richer in photosystem I-mediated cyclic electron flow producing ATP. In NADP-malic enzyme type species mesophyll chloroplasts have well developed grana, whereas the bundle-sheath chloroplasts are deficient in grana (Woo *et al.*, 1970; Golbeck *et al*., 1981). However, the photorespiratory enzymes, i.e. glycine decarboxylase, are restricted to bundle-sheath cells (Baldy & Cavalié, 1984). The large amounts of organelles in the bundle-sheath cells of the C_4 plants creates the necessary metabolite sinks for the photorespiratoy metabolism. Due to

the enhanced activity of the photorespiratory enzyme glycine decarboxylase, the photorespired carbon dioxide could be re-assimilated by ribulose-1,5-bisphosphate-carboxylase/oxygenase rather than being lost in the atmosphere (Monson & Moore, 1989).

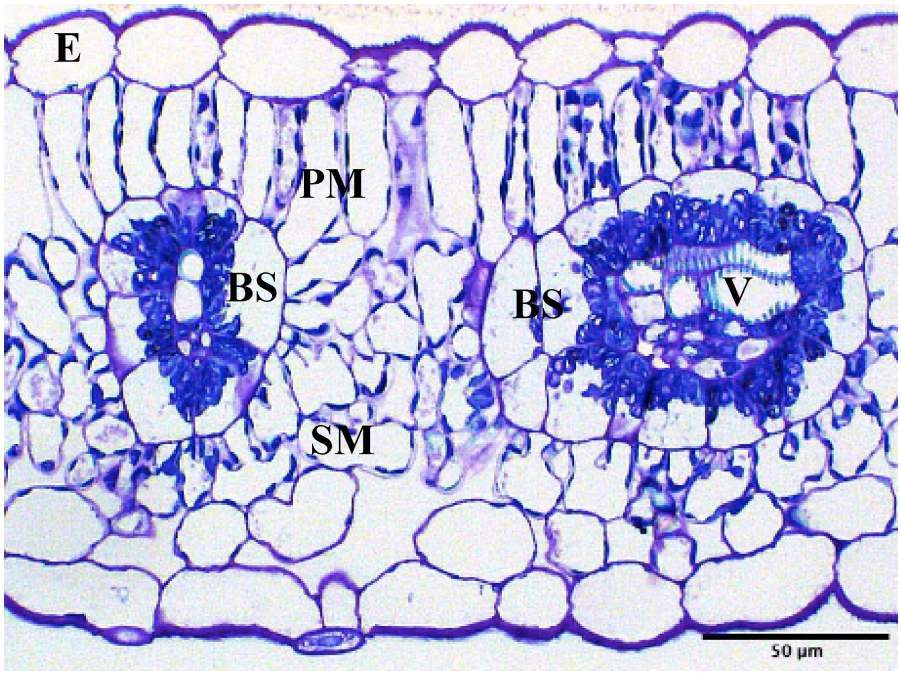


Figure 2. **Cross-section of a *Flaveria bidentis* leaf (McKown *et al*., 2007).**
The vascular bundle is encircled by a cell-layer of bundle-sheath cells, referred to as "Kranz " anatomy. The bundle-sheath cells themselves are surrounded by mesophyll cells. BS: bundle-sheath cells; E: epidermis; PM: palisade mesophyll; SM: spongy mesophyll; V: vascular bundle.

A characteristic of C_4 plants, the division of labour between mesophyll and bundle-sheath cells, is governed by differential gene expression (Nelson & Dengler, 1992; Dengler & Nelson, 1999; Sheen, 1999). In NADP-malic enzyme type C_4 species the phospho*enol*pyruvate carboxylase, NADP-malate dehydrogenase and pyruvate orthophosphate dikinase are specifically expressed in mesophyll cells, whereas the NADP-dependent malic enzyme and ribulose 1,5-bisphosphate carboxylase/oxygenase are expressed exclusively in bundle-sheath cells (Hatch, 1987; Kanai & Edwards, 1999). Regulatory mechanisms acting on the transcriptional level appear to be major the determinants of the cell-specific expression of these genes (Schäffner & Sheen, 1992; Matsuoka *et al.*, 1994; Stockhaus *et al.*, 1997; Onodera *et al.*, 2001; Gowik *et al.*, 2004), however, posttranscriptional control has been reported too (Berry *et al.*, 1986; Kubicki *et al.*, 1994; Rosche & Westhoff, 1995; Brutnell *et al.*, 1999; Lai *et al*., 2002; Patel *et al.*, 2006).

I.2 The Evolution of C_4 Photosynthesis

In terms of geological time, the appearance of C_4 photosynthesis is a relatively recent phenomenon. The earliest definitive evidence of C_4 photosynthesis in grasses is dated to 20-30 million years ago, in the late Oligocene or early Miocene, based on estimations of C_4 age from phylogenetic analyses utilizing molecular clock techniques (Kellogg, 1999). The oldest C_4 fossils are from 12.5 million year old grass leaves, the global expansion of C_4 grasslands however, is dated to 5 to 7 million years ago (Cerling, 1999). It is suggested that this widespread expansion of C_4 photosynthesis is correlated with the global reduction in the atmospheric CO_2 concentration, that suppressed photosynthesis and stimulated photorespiration to such an extent, that the C_4-specific carbon dioxide-concentrating mechanism was selectively favoured (Ehleringer *et al.*, 1991; Ehleringer & Monson, 1993; Ehleringer *et al.*, 1997; Sage *et al.*, 1999). According to this assumption favourable plant carbon balance is the principal driver of the evolution of C_4 photosynthesis. Two factors have contributed to the decreasing atmospheric carbon dioxide levels over geological time: the burying of organic matter as coal, oil and peat, and the silicate-rock weathering. The decreasing atmospheric CO_2 concentration is presumed to favour the expansion of C_4 photosynthesis because of the C_4 advantage of reduced photorespiration (Ehleringer *et al.*, 1991).

Plants exhibiting C_4 photosynthesis occur in at least 19 families of mono- and dicotyledonous plants (Sage, 2004). These families differ phylogenetically from each other. This indicates that C_4 plants must have evolved several times independently from C_3 ancestors during the evolution of angiosperms (Kellogg, 1999). Even within many of the families with C_4 species multiple independent origins are apparent, for example the grass family (*Poaceae*) is estimated to have eleven independent origins of C_4 species, and the dicotyledonous family of the Asteraceae has three independent C_4 lineages (Kellogg, 1999; Giussani *et al.*, 2001). C_4 photosynthesis is currently known to have at least 45 origins in the angiosperms based on molecular phylogenetic data (Sage, 2004). In addition, some genera include species that have features that are between C_3 and C_4 values and may include a combination of characteristics that reflect a partially to nearly complete C_4 photosynthetic cycle (Edwards and Ku, 1987; Monson and Moore, 1989; Bruhl and Perry, 1995; Sage *et al.*, 1999; Monson and Rawsthorne, 2000). These species are classified as photosynthetic "intermediates".

In order to understand the molecular basis of C_4 evolution C_3-C_4 intermediate species, which differ quantitatively in the expression of the C_4 photosynthetic traits, have been examined (Monson & Moore, 1989). At least 27 species of monocot and dicot plants exhibit anatomic and physiological characteristics intermediate between C_3 and C_4 plants. Examples for C_3-C_4 intermediate species occur in the genera *Flaveria, Atriplex, Panicum, Parthenium, Moricanda, Mollugo, Neurachne and Alternanthera* (Monson & Moore, 1989). There is convincing evidence that at least some of these C_3-C_4 intermediate plants in the genus *Flaveria* are true evolutionary intermediates that are progressing from fully expressed C_3 plants towards fully expressed C_4 plants (Figure 3) (Monson & Moore, 1989).

Examinations of C_3-C_4 intermediate plants which reveal little to no C_4 cycling demonstrate that anatomical and vein pattern alterations, that modify photosynthetic tissues to facilitate photorespiratory cycling, occurred prior to the establishment of a C_4 cycle (McKown & Dengler, 2007). As the activity of the C_4 cycle increases it compensates with ribulose-1,5-bisphosphate carboxylase/oxygenase and the C_3 cycle for CO_2 and ATP. To avoid this competition and to integrate the C_3 and C_4 cycles, the expression pattern of most enzymes in the photosynthetic apparatus had to be reorganized. Recapitulating one can say that the following major steps have been taken during the evolution of C_4 photosynthesis: (1) the evolution of C_4-like anatomical traits; (2) compartmentalization of photorespiratory enzymes, i.e. the glycine decarboxylase, in conjunction with an overall elevated level of organelles in the bundle-sheath cells; and (3) elevated phospho*enol*pyruvate carboxylase activities and elevated activities of other enzymes of the C_4 photosynthesis pathway (see Figure 3 for details).

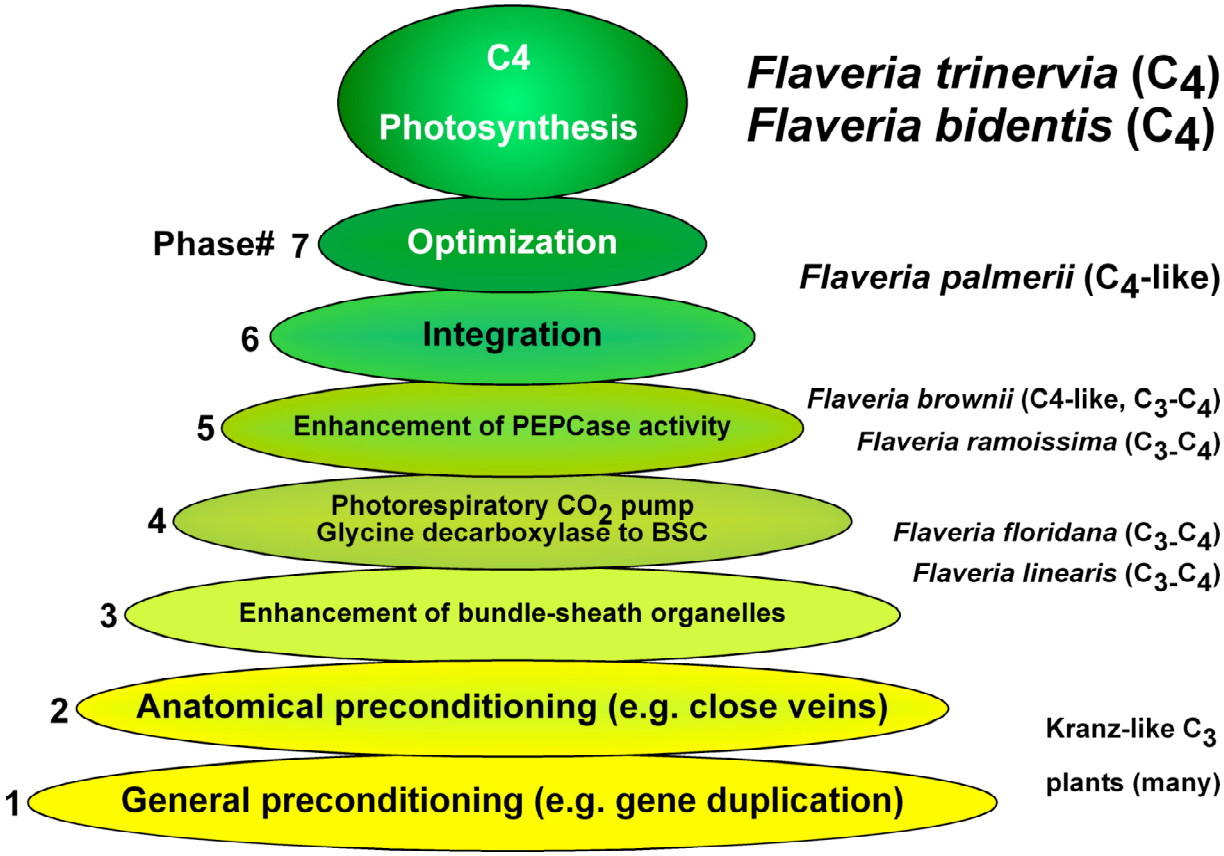


Figure 3. Simplified model of the main phases of C_4 evolution in the genus *Flaveria* (Sage, 2004).
Some intermediate *Flaveria* species are listed beside the phases they represent.

The multiple independent origin of C_4 photosynthesis suggests that the evolution of a C_3 into a C_4 species must have been relatively easy (Ehleringer & Monson, 1993; Ehleringer *et al.*, 1997). The available molecular data on the C_4 cycle enzymes support this point of view. None of the C_4 photosynthetic enzymes are unique to C_4 plants, non-photosynthetic isoforms of these enzymes are also present in C_3 species and in non-photosynthetic tissues of C_4 species where they fulfil basic functions in the plant metabolism (Latzko & Kelly, 1983). The ubiquity of these non-photosynthetic isoforms of the C_4 cycle enzymes in C_3 plants strongly indicates that these C_3 isoforms served as the starting point for the evolution of the C_4 genes (Monson, 1999).

Here we use the genus *Flaveria* as an experimental system. *Flaveria* consists of C_3 and C_4 species and a large number of C_3-C_4 photosynthetic intermediates (Edwards & Ku, 1987). Members of this genus can be grouped into two lines based on genetic hybridization experiments, the number of phyllaries (floral bracts), and molecular phylogenetic studies (Powell, 1978; Kopriva *et al.*, 1996; Marshall *et al.*, 1996; McKown *et al.*, 2005). All C_4 species, in addition to two C_3-C_4 intermediate *Flaveria* species, belong to the clade A,

whereas the clade B consists of the remaining C_3-C_4 intermediate species, with *Flaveria brownii* as the only C_4-like C_3-C_4 intermediate species. The C_3 *Flaveria* species are all basal to intermediate (C_3-C_4 and C_4-like) and fully expressed C_4 *Flaveria* species (Figure 4) (McKown *et al*., 2005).

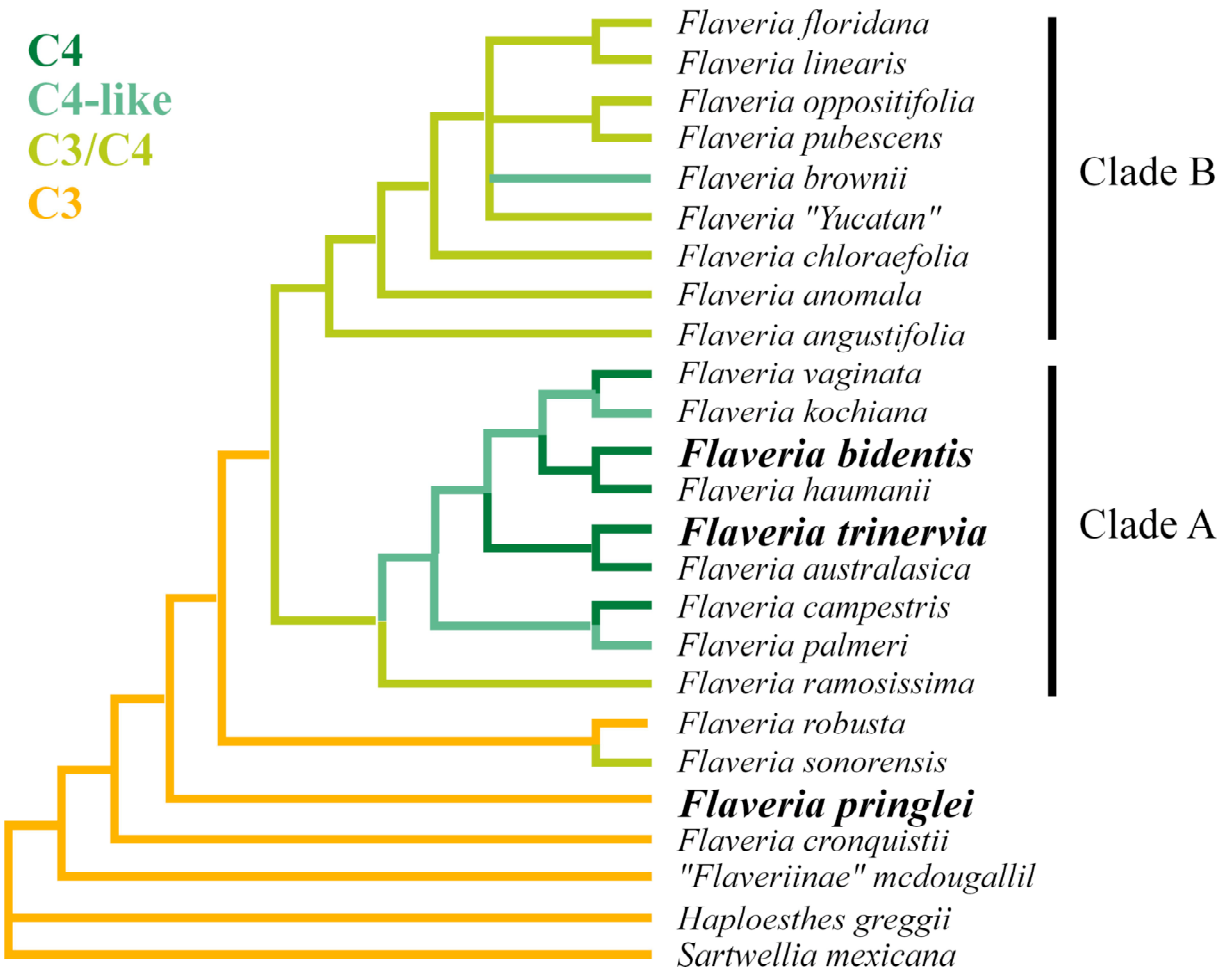


Figure 4. Phylogeny of *Flaveria* based on combined gene and morphological data analysis from *Flaveria*, *Haploestes* and *Sartwellia* species (McKown *et al*., 2005).
Data on morphology, life history and DNA sequences (chloroplastic *trnL-F*, nuclear ITS and ETS) for 21 of 23 known *Flaveria* species were collected. Each data set was analyzed separately and in combination using maximum parsimony and Bayesian analyses. The phylogeny of *Flaveria* is based on the combined analysis of all data (McKown *et al*., 2005). Solid vertical lines represent well-supported clades A and B. Branch colors indicates photosynthetic types: dark green = C_4 photosynthesis; light green = C_4-like photosynthesis; ochre = C_3-C_4 photosynthesis; orange = C_3 photosynthesis.

During the evolution of C_4 photosynthesis three major changes occurred to shape a C_4 isoform gene starting from a C_3 ancestral gene. First, C_4 isoforms genes are highly expressed (Harpster & Taylor, 1986; Hermans & Westhoff, 1990), whereas C_3 isoform genes are only moderately transcribed (Ernst & Westhoff, 1996). Thus, the efficiency of gene

expression had to be increased. Secondly, the C_4 isoform genes had to evolve an organ- and cell-specific expression pattern, because the strict compartmentation of these enzymes is imperative to guarantee the proper functioning of the C_4 cycle (Hatch, 1987). Thirdly it is known, that the C_4 cycle enzymes differ from its C_3 counterparts with respect to kinetic and regulatory characteristics (Bauwe & Chollet, 1986; Svensson *et al.*, 1997; Bläsing *et al.*, 2000; Engelmann *et al.*, 2002; Svensson *et al.*, 2003). Therefore the coding regions had to be changed to achieve the required adaptations of the enzymatic properties. In all the *Flaveria* intermediates that have been investigated, both carbon dioxide assimilating enzymes ribulose-1,5-bisphosphate carboxylase/oxygenase and phospho*enol*pyruvate carboxylase are not entirely compartmentalized between mesophyll and bundle-sheath cells as in fully expressed C_4 plants.

The acquisition of new functions for an old gene, usually through changes in expression patterns and/or functional modifications of the protein it encodes, is known as gene co-option (Ganfornina & Sanchez, 1999; True & Carroll, 2002; Olson, 2006). The gain, loss or modification of morphological features is due predominantly to modifications of spatiotemporal patterns of gene expression. Changes in the expression of a particular gene can result from alterations either in its *cis*-regulatory sequences or from alterations in the employment and function of the transcription factors that control gene expression, or both (Doebley & Lukens, 1998; Irish & Litt, 2005; Gompel *et al.*, 2005; Costa *et al.*, 2005; Yang *et al.*, 2006; Lynch *et al.*, 2006; Wray, 2007). This often involves gene duplication followed by specialization of the resulting paralogous genes into particular functions, however examples of direct co-option without gene duplication has been reported, too (Ganfornina & Sanchez, 1999; Harlin-Cognato *et al.*, 2006). Gene duplications followed by subsequent modifications provide the basic raw material for the evolution of new morphologies, thus they are the primary driving forces in the evolution of genomes. For example, around 29% of the genes in the genome of *Arabidopsis thaliana* are believed to arise from duplication events (Lynch & Conery, 2000). Gene duplications can result from unequal crossing over, retroposition or chromosomal (or genome) duplication (Zhang, 2003). Gene duplications are especially prevalent in plants, and the genomes of most present angiosperm species are the result of a series of segmental or whole genome duplication (polyploidization) events (Wendel, 2000; Vision *et al.*, 2000; Simillion *et al.*, 2002; Vandepoele *et al.*, 2003; Guyot & Keller, 2004; Maere *et al.*, 2005; Wang *et al.*, 2005). At the gene level there are several potential functional fates of duplicated genes: functional redundancy, sub-functionalization, neo-functionalization, or pseudo-gene formation (Ohno, 1970; Lynch & Conery, 2000;

Monson, 2003; Zhang, 2003; Moore & Purugganan, 2005). A. C. Wilson pointed out that 'quantitative mutations affecting enzyme levels may have had a major role in the adaptive metabolic evolution of multicellular organisms' and that 'these quantitative effects result from point mutations in control genes' (Wilson *et al.*, 1977). He believed that these regulatory mutations, which give rise to altered spatial-temporal expression patterns, happened before mutations in protein coding regions result in alterations of the regulatory and catalytic enzyme properties. The phospho*enol*pyruvate carboxylase encoding *ppcA* genes in the genus *Flaveria* are a good model system which allow in an evolution-orientated background to understand how the enzyme changed during transition from C_3 to C_4 photosynthesis, and to enlighten the regulation of gene expression, i.e. to identify the corresponding *cis*- and *trans*- regulatory factors responsible for a C_4-specific gene expression.

Three phospho*enol*pyruvate carboxylase gene classes are present in *Flaveria* (*ppcA, ppcB, ppcC*; Figure 5) with different isoforms taking on specific roles in plant metabolism. The *ppcA* gene class of *Flaveria trinervia* contains two members and encodes the C_4 isoform of phospho*enol*pyruvate carboxylase, while the *ppcB* and *ppcC gene* classes probably consist of only one gene each and code for the non-photosynthetic isoforms (Hermans & Westhoff, 1992; Ernst and Westhoff, 1997). Maximum likelihood analyses of the sequences of the *ppc* genes suggests that the *ppcA* gene class was derived from the *ppcB* gene class by gene duplication (Svensson *et al.*, 2003). And indeed, the same three classes of *ppc* genes of the C_4 plant *Flaveria trinervia* are found in the C_3 plant *Flaveria pringlei* (Hermans & Westhoff, 1990, 1992). One can conclude from these findings that the most recent common ancestor for the C_3 and C_4 *Flaveria* species must have contained the same set of genes found in the present *Flaveria* species (Figure 5).

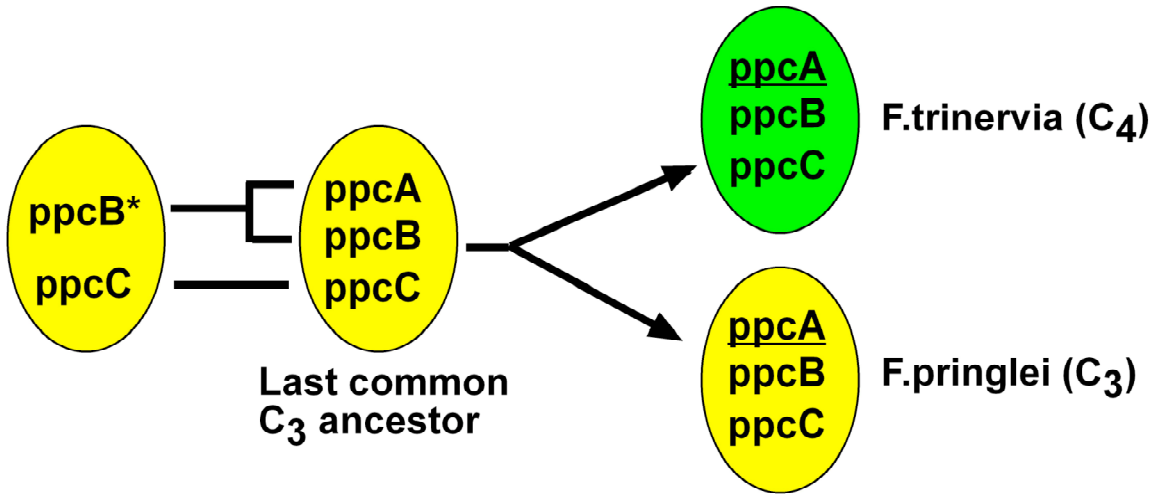


Figure 5. Evolution of phospho*enol*pyruvate carboxylase (PEPC) genes (*ppc*) in the genus *Flaveria* (Bläsing *et al.*, 2002).
Sequence analyses suggest that the *ppcA* gene was formed by gene duplication of a *ppcB*-type PEPC gene (*ppcB**) already present in the last common ancestral C_3 plant that gave rise to *Flaveria pringlei* (C_3) and *Flaveria trinervia* (C_4) The *ppcA* gene class of *Flaveria trinervia* encodes the C_4 PEPC of this species. The orthologous *ppcA* gene of the closely related C_3 species *Flaveria pringlei* is used as a reference C_3 gene for studying the evolution of the C_4 PEPC gene.

I.3 The Phospho*enol*pyruvate Carboxylase

The entry enzyme of the C_4 photosynthesis pathway, the phospho*enol*pyruvate carboylase (EC.4.1.1.31), is a homotetramer with four active sites and a molecular mass of about 100 kDa. The size of the phospho*enol*pyruvate carboxylase polypeptide varies significantly depending on the kind of organism: the approximate number of amino acid residues is 870 (100 kDa), 970 (110 kDa),1010 (116 kDa) and 1150 (134 kDa) for phospho*enol*pyruvate carboxylase genes from bacteria, vascular plants, cyanobacteria and protozoa, respectively (Izui *et al.*, 2004). The phospho*enol*pyruvate carboxylase enzyme is absent from animals, fungi and yeast (Lepiniec *et al.*, 1994; Gehrig *et al.*, 1998). The enzyme was first described by Bandurski and Greiner in 1953 (Bandurski & Greiner, 1953), it catalyzes the fixation of the bicarbonate HCO^{3-} to the receptor phosphoenolpyruvate resulting in the formation of oxaloacetate and inorganic phosphate (O'Leary, 1982; Andreo *et al.*, 1987). Phospho*enol*pyruvate carboxylase is dependent on bivalent cations and prefers Mg^{2+} *in vivo*, but can use Mn^{2+} or Co^{2+} *in vitro* (O'Leary, 1982). The phospho*enol*pyruvate carboxylase reaction is highly exergonic with a $\Delta G°$ of -30 kJ mol^{-1} and the overall reaction is strictly

irreversible. The activity of the phospho*enol*pyruvate carboxylase enzyme is sensitive to various internal and external factors (e.g. light, temperature, pH, metabolic effectors). Phospho*enol*pyruvate carboxylase from C_4 plants is activated by glucose-6-phosphate, a final product of CO_2 fixation (Ting & Osmond, 1973; Uedan & Sugiyama, 1976; Rajagopalan *et al.*, 1994). Another metabolite, L- malate, which is an intermediate product of the carboxylation reaction, is an feedback inhibitor of the phospho*enol*pyruvate carboxylase (Ting & Osmond, 1973; Wedding *et al.*, 1990). Aspartate, another important product of the carboxylation in some C_4 plants, also inhibits the enzyme (Andreo, *et al.*, 1987). Like many of the C_3 and C_4 photosynthetic enzymes, the kinetic and regulatory properties of C_4 phospho*enol*pyruvate carboxylase are modulated by light/dark transitions *in vivo* (Andreo *et al.*, 1987; Jiao & Chollet, 1991).

Besides its basic role in the initial fixation of atmospheric CO_2 during C_4 photosynthesis, phospho*enol*pyruvate carboxylase plays an anaplerotic role in non-photosynthetic tissues of all plants (Latzko & Kelly, 1983), e.g., replenishment of citric acid cycle intermediates (i.e. oxaloacetate and malate), thus providing the carbon skeletons necessary for nitrogen assimilation and amino-acid biosynthesis (Melzer & O'Leary, 1987). In plants, this function is especially important during fruit maturation, seed formation and germination (Lepiniec *et al.*, 1994). Non-photosynthetic isoforms of phospho*enol*pyruvate carboxylase also play specialized roles in guard cell carbon metabolism during stomatal opening (Tarczynski & Outlaw, 1990) and in plant root nodules of N_2-fixing legumes (Chollet *et al.*, 1996; Outlaw, 2002).

The *ppcA* phospho*enol*pyruvate carboxylase gene class in C_4 plants reveals a completely different expression pattern as compared to the C_3 counterpart. The C_4 phospho*enol*pyruvate carboxylase genes reveal a high expression in the mesophyll cells of the leaves and no expression in other plant cells or tissues. The orthologues C_3 phospho*enol*pyruvate carboxylase genes are weakly expressed and their transcripts do not show the strict leaf-specific accumulation pattern as the C_4 phospho*enol*pyruvate carboxylase genes (Hermans & Westhoff, 1990; Stockhaus *et al.*, 1997). The other phospho*enol*pyruvate carboxylase isoforms, *ppcB* and *ppcC*, involved in the basic cell metabolism, reveal also different expression patterns and are only moderately expressed due to their function in roots, stems or leaves (Ernst and Westhoff, 1997). Their expression in leaves shows no mesophyll specificity.

I.4 The *Cis*-Regulatory Modules Required in the Expression of the *ppcA* Phospho*enol*pyruvate Carboxylase Gene in the Genus *Flaveria*

To uncover the evolutionary changes that the C_4 *ppcA* genes of the C_4 *Flaveria* species have undergone during the evolution from C_3 to C_4 photosynthesis the *ppcA* gene class of *Flaveria trinervia* and *Flaveria pringlei* is used since these genes are true evolutionary orthologs (Hermans & Westhoff, 1992). The common ancestor of the current C_3 and C_4 *Flaveria* species was a C_3 plant. If it is assumed that the *ppcA* gene class of *Flaveria pringlei* is not very different from the *ppcA* gene class of the common C_3 ancestor, then the *ppcA* gene class of *Flaveria pringlei* can serve as an reference C_3 gene for studying the evolutionary changes at the promoter level.

One representative of each of the *ppcA* gene class of *Flaveria trinervia* and *Flaveria pringlei*, named *ppcA*1, has been characterized in detail (Hermans & Westhoff, 1990). A conserved segment of about 550 nucleotides which is located 5'-proximal to the translational initiation start codon of both genes reveal the highest similarity (more than 70% sequence identity) (Figure 6A). In *Flaveria trinervia* the proximal region extends to -531 and in *Flaveria pringlei* the proximal region extends to -566 (nucleotide numbers are given with respect to the translation initiation site). Furthermore, the 5'-untranslated regions of both orthologous *ppcA*1 genes contain an intron (*F. trinervia*: -209 to -40; *F. pringlei*: -212 to -33; distances are given with respect to the translation initiation site).

For the examination of promoter/ß-glucuronidase (GUS) reporter gene fusions the C_4 plant *Flaveria bidentis* is used. *Flaveria bidentis* is very closely related to *Flaveria trinervia* (Powell, 1978) and in contrast to *Flaveria trinervia* it can be transformed by *Agrobacterium tumefaciens* (Chitty *et al.*, 1994). The analysis of *ppcA*1 promoter-ß-glucuronidase reporter gene fusions revealed that the 5' promoter region of the *ppcA*1 gene of the C_4 plant *Flaveria trinervia* (-2188 to -1) directs expression of the GUS reporter gene exclusively to the mesophyll cells (Figure 6B), which is designated as the C_4-specific expression pattern. The promoter region of the orthologous non-photosynthetic *ppcA*1 gene of the C_3 plant *Flaveria pringlei* (-2584 to -1) however, induces a very low reporter gene activity, the GUS staining is found mainly in the vascular tissue of leaves and stems as well as in mesophyll cells of transgenic *Flaveria bidentis* plants (Figure 6B) (Stockhaus *et al.*,

1997). This non-cell-specific activity is therefore designated as the C_3-specific expression pattern.

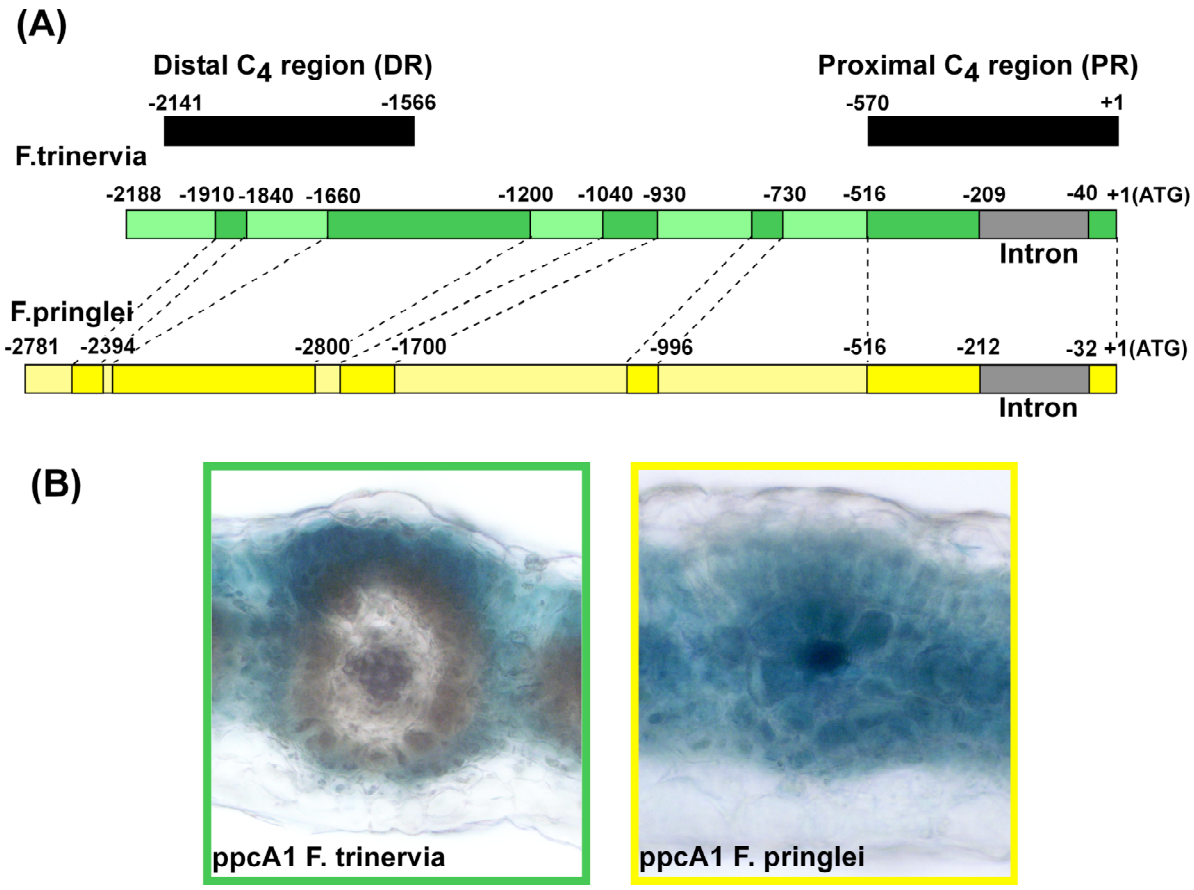


Figure 6. Schematic structure of the 5' flanking regions of the *ppcA*1 genes of *F. trinervia* (C_4) and *F. pringlei* (C_3) and analysis of the *ppcA*1 GUS reporter gene constructs ppcA1Ft and ppcA1Fp in transgenic *F. bidentis* (Westhoff and Gowik, 2004).
(A) Nucleotide positions are numbered with respect to the AUG translational initiation codon. The proximal (PR) and distal region (DR) of the *Flaveria trinervia* promoter are given in black bars. Regions with high similarity between both promoters (60% or more identical nucleotides in a window of 15 bp) are indicated by darker coloured boxes. **(B)** Histochemical analysis of the activities of the *ppcA*1 promoters of *Flaveria trinervia* (C_4, -1 to -2188) and *Flaveria pringlei* (C_3, -1 to -2584) in transgenic *Flaveria bidentis* (C_4). The ß-glucuronidase gene (*GUS*) was used as a reporter gene (Jefferson et al., 1987).

With “loss-of-function” and “gain-of-function” experiments J. Burscheidt (1998) could show that a distal promoter sequence (-1565 to -2188, with reference to the AUG translational start codon) in combination with the proximal 570 bp of the 5' flanking sequences of the *ppcA*1 gene of the C_4 species *Flaveria trinervia* are completely sufficient to allow an elevated expression of the ß-glucuronidase reporter gene in the mesophyll cells of transgenic *Flaveria bidentis* leaves. The nucleotide sequences between -570 and -1566 were found to be essentially dispensable for the *ppcA*1 promoter activity (Gowik *et al.*, 2004). The

expression of the proximal promoter region alone resulted in a very weak expression in leaves, stems and roots in the C_3 plant *Nicotiana tabacum* (Stockhaus *et al.*, 1994) and in the C_4 plant *Flaveria bidentis* (Gowik *et al.*, 2004). These results suggested that the 5' distal region of the C_4 *ppcA* promoter contains *cis*-regulatory elements that are indispensable for a high level of expression and for mesophyll specificity. The C_4 distal region functions both in the correct and the inverse orientation (Gowik *et al.*, 2004), this *cis*-regulatory region therefore shows the typical features of a transcriptional enhancer (Blackwood & Kadonaga, 1998).

To characterize the function of the distal region more precisely, the distal region of the C_4 *ppcA*1 promoter was fused to the C_3 *ppcA*1 promoter (-1853 to -1) from which the distal region (-2538 to -1854) was removed (Gowik *et al.*, 2004). The chimeric promoter construct revealed a low promoter activity and directed expression in the mesophyll and bundle-sheath cells and the vascular bundles, like the 1853 bp of the C_3 *ppcA*1 promoter alone. However, this C_4-C_3 chimerical promoter had acquired a new cell specificity of expression. While the 1853 bp of the C_3 *ppcA*1 promoter revealed the highest activity in the bundle-sheath cells and the vascular tissues and only a low activity in the mesophyll cells as the whole C_3 *ppcA*1 promoter (-2538 to -1), the chimeric promoter revealed a clear enhanced activity in the mesophyll cells and a very low activity in the bundle-sheath cells (Gowik *et al.*, 2004). Thus, the addition of the C_4 distal region of *Flaveria trinervia* to the *ppcA*1 promoter of *Flaveria pringlei* (-1853 to -1) did not only add a mesophyll expression component to the *ppcA*1 promoter part of *Flaveria pringlei* but reduced its original activity in bundle-sheath cells (Gowik *et al.*, 2004).

From the described experiments one can infer that the C_4 proximal region alone is not capable to drive a mesophyll-specifc expression of the *ppcA*1 gene of *Flaveria trinervia*. To achieve a high level of expression in the mesophyll cells, the C_4 distal region has to be combined with the proximal region of the *ppcA*1 gene of either *Flaveria* species. When the distal region of the C_4 *ppcA*1 promoter is combined with the proximal region of the C_3 *ppcA*1 promoter the chimerical promoter still directs a mesophyll-specific expression of the reporter gene. However, the expression strength of this promoter is low indicating that the proximal regions of the *ppcA* promoters are responsible for expression quantity, but not specificity and that the C_4 and C_3 proximal regions differ in expression strength.

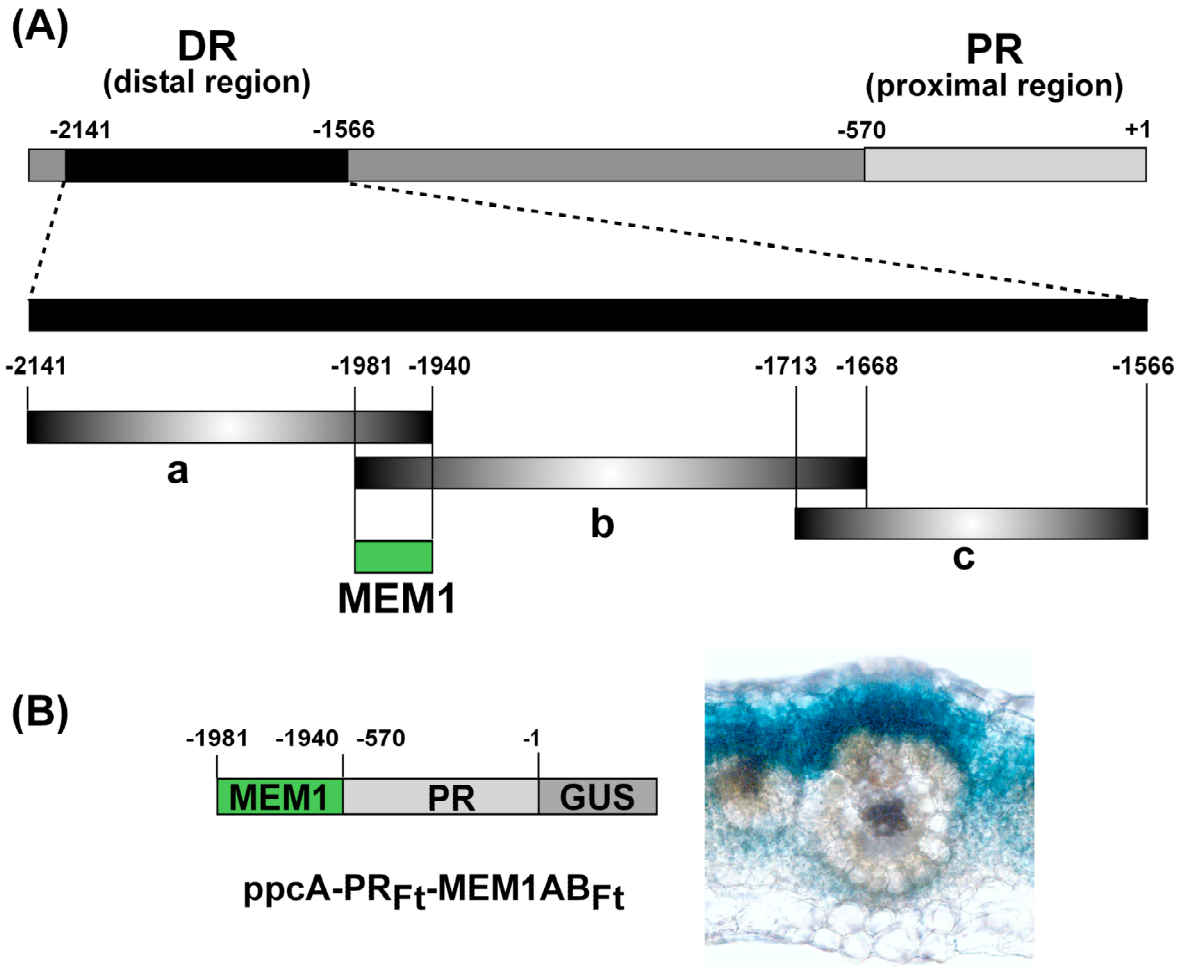


Figure 7. Schematic presentation of the 5'-untranslated region of the C_4 species *Flaveria trinervia* and analysis of the *ppcA* GUS reporter gene construct ppcA-PR_{Ft}-$MEM1AB_{Ft}$ in transgenic *Flaveria bidentis* (Gowik *et al*., 2004).
(A) The dissection of the distal promoter region (DR) into the a-, b- and c-fragment is shown. The MEM1 region (a/b-overlapping part) is shown below (green coloured box). Nucleotide positions are numbered with respect to the AUG translational initiation codon. **(B)** Histochemical localization of the GUS activity in leaf sections of transgenic *Flaveria bidentis* plants transformed with ppcA-PR_{Ft}-$MEM1AB_{Ft}$. The ß-glucuronidase gene (*GUS*) was used as a reporter gene (Jefferson et al., 1987).

To identify the *cis*-regulatory element(s), responsible for the mesophyll-specific expression of the C_4 *ppcA* phospho*enol*pyruvate carboxylase gene, more precisely the C_4 distal region was dissected into three pieces of approximately equal size that overlap by about 50 bp (Figure 7) (Gowik *et al*., 2004). Each fragment of the C_4 distal region, named a to c, was fused with the C_4 proximal region of the *ppcA* promoter of *Flaveria trinervia* and transformed into *Flaveria bidentis*. In contrast to the c-fragment of the C_4 distal region, the a- and b-promoter fragments of the C_4 distal region directed GUS expression of the reporter gene exclusively in the mesophyll cells. Additional analyses showed that the overlapping part of the a- and b-promoter fragments alone led to a mesophyll expression in transgenic

Flaveria bidentis leaves (Gowik *et al*., 2004). Because of these results one can infer that the 41 bp long a-/b-overlapping fragment of the C_4 distal region holds determinants which are responsible for a mesophyll-specific gene expression of the C_4 *ppcA* gene, and for this reason it was named MEM1 for mesophyll expression module 1. Sequence analyses of the 5' regions of the *ppcA* genes of different *Flaveria* species revealed that this element is present in all of the analysed genes (Figure 8). MEM1 consists of two submodules, named A and B, that except in *Flaveria trinervia* are separated by an insertion of 90 to 100 bp (Figure 8). The A-submodules of the C_4 and C_4-like species hold a guanine at their first nucleotide position, while an adenine is present in the A-submodules of the C_3 plants. An additional difference is related to the tetranucleotide CACT that is present in the B-submodules of the C_4 and C_4-like plants but is absent from the B-submodules of the C_3 plants (Figure 8) (Gowik *et al*., 2004).

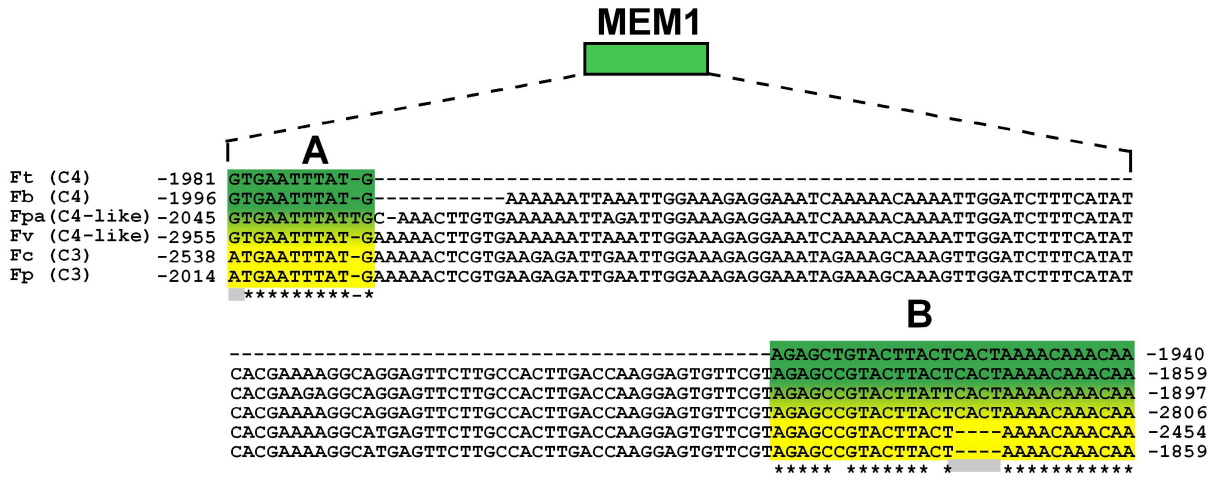


Figure 8. Sequence comparison of MEM1 and its orthologs.
The A- and B-submodules are highlighted by boxes. Asterisks label identical nucleotides in the A- or B-submodule of all MEM1 fragments. Grey bars indicate the single nucleotide difference in A-submodule and the CACT tetranucleotide in B-submodule of MEM1. The numbers of nucleotides refer to the translation initiation codons. Ft: *Flaveria trinervia*; Fb: *F. bidentis*; Fpa: *F. palmerii*; Fv: *F. vaginata*; Fc: *F. cronquistii*; Fp: *F. pringlei.*

The experiments that were performed to elucidate the functional impact of the two C_3-to-C_4 associated nucleotide differences in MEM1 towards the mesophyll-specific expression, and to identify C_4 MEM1 specifically interacting *trans*-regulatory proteins involved in the mesophyll-specific expression of the *ppcA* gene of *Flaveria* are described in the three manuscripts found at the end of this PHD-thesis.

II. Objectives

The objective of the present PHD thesis was (1) to investigate the functional significance of the C_3-to-C_4 associated differences in MEM1 structure and (2) to isolate *trans*-regulatory factors interacting with MEM1.

(1) Manuscripts 1 (Gowik *et al*., 2004) and 2 (Akyildiz *et al*., 2007) describe the experiments to identify the effects of the C_3-to-C_4 associated differences in C_4 and C_3 MEM1 structure on mesophyll specificity of reporter gene expression. To this end various MEM1 versions were constructed and combined with the proximal region of the C_4 *ppcA* promoter. The chimerical promoters were fused to the ß-glucuronidase (GUS) reporter gene and the strength and specificity of reporter gene expression was investigated in transgenic *F. bidentis*.

(2) Manuscript 3 (Akyildiz *et al*., 2007) describe the experiments to isolate DNA-binding proteins interacting with the C_4 MEM1. To isolate these C_4-MEM1-binding proteins the yeast one-hybrid screen was used. In such a screen MEM1 interacting factors can be identified by using the MEM1 as the bait sequence and a cDNA library expressing prey proteins. The prey proteins are fused to the activation domain of the yeast GAL4 transcription factor. The specificity of binding, i.e. C_4 versus C_3 MEM1, was analyzed in the yeast one-hybrid assay using the complete coding sequences of the identified proteins. The physical binding of these proteins to C_4- and C_3-type MEM1 variants was compared by using the electrophoretic mobility shift assay and fluorescence polarization/anisotropy measurements. The electrophoretic mobility shift assay is based on the observation that protein-DNA complexes migrate more slowly than free DNA molecules when subjected to non-denaturing polyacrylamide gel electrophoresis. Because the rate of DNA migration is shifted or retarded upon protein binding, the assay is also referred to as a gel shift or gel retardation assay. Fluorescence polarization is a technique to measure molecular orientation and mobility using polarized light and fluorescent labelled DNA molecule. When excited with polarized light, DNA molecules attached to proteins with high molecular weight emits high level of polarized fluorescence since DNA molecules are slower in rotations compared to DNA molecules attached to smaller proteins.

.

III. Theses

Previous work identified a 41 bp *cis*-regulatory element in the distal region of the promoter of the C_4 *ppcA*1 gene of the C_4 plant *Flaveria trinervia* as being responsible for the mesophyll specificity of gene expression (Gowik *et al*., 2004). This *cis*-regulatory module, named mesophyll expression module 1 (MEM1) consists of two submodules that differ at two positions, a guanine (C_4)-to-adenine (C_3) exchange in the A-submodule and the presence (C_4)/lack (C_3) of the tetranucleotide CACT in the B-submodule.

Work presented here demonstrates:

1) Both C_3-to-C_4 associated nucleotide changes in MEM1 are required for the mesophyll-specific gene expression.
2) MEM1 behaves as an enhancer of mesophyll-specific gene expression and concomitantly as a repressor of expression in bundle-sheath cells and the vascular bundles.
3) The evolution of a C_3- into a C_4-type MEM1 proceeded step by step.
4) MEM1 interacts with basic leucine zipper (bZIP) proteins that are homologous to group I bZIP proteins of *Arabidopsis thaliana*.
5) The three identified bZIP proteins FtbZIP18, FtbZIP29 and FtbZIP51 are considered as *trans*-regulatory factors required for the control of *ppcA* gene regulation in the genus *Flaveria*.

IV.1 Summary

C_4 plants evolved independently several times from C_3 ancestor species. During this evolution towards C_4 photosynthesis the expression programme of the involved genes had to be changed. C_4 photosynthesis is characterized by a division of labour between two different photosynthetic cell types, mesophyll and bundle-sheath cells. The key enzyme of C_4 photosynthesis, the phospho*enol*pyruvate carboxylase (PEPC), is expressed at high levels in leaves but only in mesophyll cells, while the C_3 PEPC gene is expressed at low levels with no apparent cell or organ specificity. Mesophyll expression determinants had been restricted to a 41 bp segment in the distal promoter of the PEPC gene, referred to as mesophyll expression module 1 (MEM1).

In the present PHD-thesis the focus is directed towards the *cis*- and *trans*-regulatory factors by which the differential expression of the PEPC gene (*ppcA*) in the genus *Flaveria* is ensured. This genus contains C_3, C_4 and a large number of C_3-C_4 intermediate species which allow a comparative analysis of ortholog *ppcA* genes of plants that differ in their photosynthetic traits.

The identified MEM1 element in the 5' distal region of the C_4 *ppcA* gene of *F. trinervia* which is responsible for the mesophyll-specific gene expression was analysed in detail *via* promoter-GUS reporter gene experiments. A comparison of this C_4 MEM1 element with that of the C_3 plant *F. pringlei* revealed that both differ in three points. First, the A- and B-submodule of the C_4 MEM1 are contiguous whereas in the C_3 MEM1 they are separated by an insertion of about 100 bp. Secondly, the first nucleotide in the A-submodule is characterized by an adenine (C_3) to guanine (C_4) exchange and the third difference is related to the tetranucleotide CACT that is only present in the B-submodule of the C_4 MEM1. For mesophyll specificity both MEM1-submodules have to be present in the C_4-specific status. The C_4 MEM1 of *F. trinervia* enhances mesophyll expression and concomitantly represses expression in the bundle-sheath cells and vascular bundles. The separation of both MEM1 modules by an insertion does not affect its mesophyll-specific expression.

A search for *F. trinervia* MEM1 corresponding interacting proteins resulted in the identification of three different bZIP proteins, FtbZIP18, 29 and 51, which are highly similar to the *A. thaliana* group I bZIP proteins. Yeast one-hybrid analyses revealed that the FtbZIP proteins interact with C_4-type but not with C_3-type MEM1 elements. However, *in vitro* protein-DNA interaction studies showed that the FtbZIP proteins physically bind to the C_4 and C_3 MEM1 with no difference in binding affinity. This contrasting interaction behaviour

of the FtbZIP proteins with MEM1 *in vitro*, compared to the *in vivo* situation in the yeast´s nucleus, suggests the involvement of additional factor(s). This unknown factor(s), which is already present in yeast, may be responsible for the observed differentiation of the FtbZIP proteins between a C_4- and a C_3-type MEM1 in the yeast system.

IV.2 Zusammenfassung

C_4-Pflanzen haben sich mehrmals unabhängig voneinander aus C_3-Pflanzen entwickelt. Im Verlaufe dieser Evolution zur C_4-Photosynthese wurden unter anderem die Expressionsmuster der involvierten Gene verändert. Die Phospho*enol*pyruvat Carboxylase (PEPC) ist in Blättern von C_4-Pflanzen deutlich höher exprimiert als in Blättern von C_3-Pflanzen. Die C_4-Photosynthese zeichnet sich durch eine arbeitsteilige Organisation zwischen Mesophyll- und Bündelscheidenzellen aus. Die Expression des C_4 Enzyms ist auf Mesophyllzellen beschränkt während das orthologe C_3 Enzym aus *F. pringlei* keine zellspezifische Expression aufweist. *Cis*-regulatorische Determinanten, die eine mesophyll-spezifische Expression des PEPC Gens bewirken, konnten einem 41 bp langen Segment im distalen *ppcA*-Promotor zugeordnet werden. Dieses Element wurde MEM1 für „mesophyll expression module 1“ benannt.

Im Rahmen der vorliegenden Arbeit ist das Interesse auf die beteiligten *cis*- und *trans*-regulatorischen Faktoren gerichtet, die die differentielle Expression des PEPC-Gens (*ppcA*) in der Gattung *Flaveria* steuern. Da in dieser Gattung neben C_3-, C_4- auch C_3-C_4 intermediäre Pflanzen existieren, war eine vergleichende Analyse von orthologen Genen möglich.

Das im distalen Promotor des *ppcA*-Gens aus *F. trinervia* identifizierte MEM1 Element, welches in Fusion mit dem proximalen Bereich für eine mesophyll-spezifische Expression ausreicht, wurde mittels Promotor-GUS-Reportergen-Experimenten eingehend untersucht. Eine vergleichende Sequenzanalyse der MEM1 Elemente aus *F. trinervia* (C_4) und aus *F. pringlei* (C_3) zeigte, dass sich die beiden orthologen Elemente in drei wesentlichen Punkten unterscheiden. Erstens, das A- und B-Submodul des C_4 MEM1 Elementes grenzen einander, während beim C_3 MEM1 Element beide Submodule durch eine Insertion von etwa 100 bp getrennt sind. Zweitens, die erste Base im A-Submodul zeichnet sich durch ein Adenin (C_3) zu Guanin (C_4) Austausch aus und drittens, im B-Submodul unterscheidet sich das C_4 vom C_3 MEM1 Element durch eine Insertion von vier Basenpaaren. Für die mesophyll-spezifische Expression müssen beide MEM1-Submodule im C_4-Status vorhanden sein, und wird durch eine Trennung beider MEM1-Submodule nicht beeinflusst. Das MEM1 Element des *ppcA* Gens aus *F. trinervia* hat eine duale Funktion, es erhöht die Expression in Mesophyllzellen und verhält sich wie ein Repressor-Element das eine Expression in den Bündelscheidenzellen und im vaskulären Gefäßsystem verhindert.

Bei der Suche nach *trans*-regulatorischen Faktoren, die spezifisch an das MEM1 Element des *ppcA*-Gens aus *F. trinervia* binden, wurden drei verschiedene bZIP Proteine,

FtbZIP18, 29 und 51, identifiziert die eine große Ähnlichkeit zu den bZIP Proteinen der Gruppe I aus *A. thaliana* aufweisen. *In vivo* Bindungsstudien im Hefesystem zeigten, dass die FtbZIP Proteine mit dem C_4 MEM1 Element jedoch nicht mit dem orthologen C_3 MEM1 Element interagieren. Folgende *in vitro* Protein-DNA Interaktionsstudien offenbarten jedoch, dass die FtbZIP Proteine in der Lage sind mit vergleichbarer Affinität an ein MEM1 Element des C_4- und C_3-Typs zu binden. Diese Interaktion der FtbZIP Proteine mit C_4 und C_3 MEM1 Elementen *in vitro*, die im Widerspruch mit der *in vivo*-Situation steht, deutet darauf hin, dass mindestens ein weiterer unbekannter Faktor involviert sein muss. Es wird angenommen, dass dieser unbekannte, von der Hefe stammende, Faktor für die unterschiedliche Interaktion von FtbZIP Proteinen an C_4 und C_3 MEM1 Elementen im Hefe Ein-Hybrid System verantwortlich ist.

V. Literature

Andreo, C.S., Gonzales, D.H., and Iglesias, A.A. (1987). Higher plant phosphoenolpyruvate carboxylase. FEBS Lett. **213**: 1-8.

Andrews, T.J., and Lorimer, G.H. (1987) Rubisco: structure, mechanism, and prospects for improvement. Biochemistry of Plants (Hatch, M.D. and Boardman, N.K., eds.) Academic Press, New York **10**: 131-218.

Badger, M.R., and Price, G.D. (1994). The role of carbonic anhydrase in photosynthesis. Annu. Rev. Plant Physiol. in Plant Mol. Biol. **45**: 369-392.

Baldy, P., and Cavalié, G. (1984). Compartmentation of photorespiratory enzymes in a C_4 photosynthesis plant, *Zea mays*. Z. Pflanzenphysiol. **114**: 255-259.

Bandurski, R.S., and Greiner, C.M. (1953). The enzymatic synthesis of oxalacetate from phosphoryl-enolpyruvate and carbon dioxide. J. Biol. Chem. **204**: 781-6.

Bauwe, H., and Chollet, R. (1986). "Kinetic properties of phosphoenolpyruvate carboxylase from C_3, C_4, and C_3-C_4 intermediate species of *Flaveria* (Asteraceae)." Plant Physiol. **82**: 695-699.

Berry, J.O., Nikolau, B.J., Carr, J.P., and Klessig, D.F. (1986). Translational regulation of light-induced ribulose 1,5-bisphosphate carboxylase gene expression in amaranth. Mol. Cell. Biol. **6**: 2347-2353.

Black, C.C.J. (1973). Photosynthetic carbon fixation in relation to net CO_2 uptake. Annu. Rev. Plant Physiol. **24**: 253-286.

Blackwood, E.M., and Kadonaga, J.T. (1998). Going the distance: A current view of enhancer action. Science **281**: 60-63.

Bläsing, O.E., Ernst, K., Streubel, M., Westhoff, P., and Svensson, P. (2002). The non-photosynthetic phosphoenolpyruvate carboxylases of the C4 dicot Flaveria trinervia - implications for the evolution of C4 photosynthesis. Planta **215**: 448-456.

Bläsing, O.E., Westhoff, P., and Svensson, P. (2000). Evolution of C4 phosphoenolpyruvate carboxylase in *Flaveria*, a conserved serine residue in the carboxyl-terminal part of the enzyme is a major determinant for C4-specific characteristics. J. Biol. Chem. **275**: 27917-27923.

Brown, R.H., and Hattersley P.W. (1989). Leaf anatomy of C_3-C_4 species as related to evolution of C_4 photosynthesis. Plant Physiol. **91**: 1543-1550.

Brown, W.V. (1975). Variations in anatomy, associations, and origins of Kranz tissue. Amer. J. Bot. **62**: 395-402.

Bruhl, J.J., and Perry, S. (1995). Photosynthetic pathway-related ultrastructure of C_3, C_4 and C_3-like C_3-C_4 intermediate sedges (Cyperaceae), with special reference to *Eleocharis*. Aust. J. Plant Physiol. **22,** 521-530.

Brutnell, T.P., Sawers, R.J.H., Mant, A., and Langdale, J.A. (1999). Bundle sheath defective2, a novel protein required for post-translational regulation of the *rbcL* gene of maize. Plant Cell. **11**: 849-864.

Burscheidt, J. (1998). Cis-regulatorische Determinanten für mesophyll- und bündelscheidenspezifische Genexpression in C4-Spezies der Gattung *Flaveria* - Die Promotoren der Phosphoenolpyruvat-Carboxylase- und der Glycin-Decarboxylasegene. Math.-Nat. Fakultät. Düsseldorf, Heinrich-Heine-Universität.

Cerling, T.E. (1999). Paleorecords of C4 plants and ecosystems. C4 Plant Biology. R. F. Sage and R. K. Monson. San Diego, Academic Press**:** 445-469.

Chitty, J.A., Furbank, R.T., Marshall, J.S., Chen, Z., and Taylor, W.C. (1994). Genetic transformation of the C_4 plant, *Flaveria bidentis*. Plant J. **6**: 949-956.

Chollet, R., Vidal, J., and O'Leary, M.H. (1996). Phospho*enol*pyruvate carboxylase: A ubiquitous, highly regulated enzyme in plants. Annu. Rev. Plant Physiol. in Plant. Mol. Biol. **47**: 273-298.

Costa, M.M., Fox, S., Hanna, A.I., Baxter, C., and Coen, E. (2005). Evolution of regulatory interactions controlling floral asymmetry. Development **132**: 5093-101.

Dai, Z., Ku, M.S.B., and Edwards, G.E. (1993). C4 photosynthesis. The CO_2-concentrating mechanism and photorespiration. Plant Physiol. **103**: 83-90.

Davies, D.D. (1979). The central role of phosphoenolpyruvate in plant. Annu. Rev. Plant Phsiol. **30**: 131-158

Dengler, N.G., and Nelson, T. (1999). Leaf structure and development in C4 plants. C4 Plant Biology. R. F. Sage and R. K. Monson. San Diego, Academic Press**:** 133-172.

Doebley, J., and Lukens, L. (1998). Transcriptional regulators and the evolution of plant form. Plant Cell **10**: 1075-1082.

Edwards, G.E., Franceschi, V.R., Ku, M.S.B., Voznesenskaya, E.V., Pyankov, V.I., and Andreo, C.S. (2001). Compartmentation of photosynthesis in cells and tissues of C4 plants. J. Exp. Bot. **52**: 577-590.

Edwards, G.E., Franceschi, V.R., and Voznesenskaya, E.V. (2004). Single-cell C-4 photosynthesis versus the dual-cell (Kranz) paradigm. Annu. Rev. Plant. Biol. **55**: 173-196.

Edwards, G. E., and Ku, M.S.B. (1987). Biochemistry of C_3-C_4 intermediates. Biochem. of Plants (Hatch, M.D. and Boardman, N.K., eds.) Academic Press, Inc. **10**: 275-325.

Ehleringer, J.R., Cerling, T.E., and Helliker, B.R. (1997). C_4 photosynthesis, atmospheric CO_2, and climate. Oecologia **112**: 285-299.

Ehleringer, J.R., and Monson, R.K. (1993). Evolutionary and ecological aspects of photosynthetic pathway variation. Annu. Rev. Ecol. Syst. **24**: 411-439.

Ehleringer, J.R., Sage, R.F., Flanagan, L.B., and Pearcy, R.W. (1991). Climate change and the evolution of C_4 photosynthesis. Trends Ecol. Evol. **6**: 95-99.

Engelmann, S., Bläsing, O.E., Gowik, U., Svensson, P., and Westhoff, P. (2003). Molecular evolution of C4 phospho*enol*pyruvate carboxylase in the genus *Flaveria* - a gradual increase from C3 to C4 characteristics. Planta **217**: 717-725.

Engelmann, S., Bläsing, O.E., Westhoff, P., and Svensson, P. (2002). Serine 774 and amino acids 296 to 437 comprise the major C4 determinants of the C4 phosphoenolpyruvate carboxylase of Flaveria trinervia. Febs Lett. **524**: 11-14.

Ernst, K., and Westhoff, P. (1996). The phosphoenolpyruvate carboxylase (*ppc*) gene family of *Flaveria trinervia* (C_4) and *F. pringlei* (C_3): molecular characterization and expression analysis of the *ppcB* and *ppcC* genes. Plant Mol. Biol. **34**: 427-443.

Freitag, H., and Stichler, **W.** (2000). A remarkable new leaf type with unusual photosynthetic tissue in a central asiatic genus of Chenopodiaceae. Plant Biol. **2**: 154-160.

Freitag, H., and Stichler, W. (2002). *Bienertia cycloptera* Bunge ex Boiss., Chenopodiaceae, another C4 plant without Kranz tissues. Plant Biol. **4**: 121-134.

Furbank, R.T., and Badger, M.R. (1983). Photorespiratory characteristics of isolated bundle sheath strands of C_4 monocotyledons. Aust. J. Plant Physiol. **10**: 451-458.

Furbank, R.T., and Hatch, M.D. (1987). Mechanism of C4 photosynthesis. The size and composition of the inorganic carbon pool in bundle sheath cells. Plant Physiol. **85**: 958-964.

Furbank, R.T., and Taylor, W.C. (1995). Regulation of photosynthesis in C_3 and C_4 plants: A molecular approach. Plant Cell **7**: 797-807.

Ganfornina, M.D., and Sanchez, D. (1999). Generation of evolutionary novelty by functional shift. Bioessays **21**: 432-9.

Gehrig, H.H., Heute, V., and Kluge, M. (1998). Toward a better knowledge of the molecular evolution of phosphoenolpyruvate carboxylase by comparison of partial cDNA sequences. J. Mol. Evol. **46**: 107-114.

Giussani, L.M., Cota-Sánchez, J.H., Zuloaga, F.O., and Kellogg, E.A. (2001). A molecular phylogeny of the grass subfamily Panicoideae (Poaceae) shows multiple origins of C4 photosynthesis. Amer. J. Bot. **88**: 1993-2012.

Golbeck, J.H., Martin, I.F., Velthuys, B.R., and Radmer, R. (1981). A critical reassessment of the photosystem II content in bundle sheath chloroplasts of young leaves of Zea mays. Proc. of the Vth Int. Congress of Photosynthesis. Chloroplast Development. J. Akoyunoglou. Philadelphia, Balaban International Science Services. **5:** 533-546.

Gompel, N., Prud'homme, B., Wittkopp, P.J., Kassner, V.A., and Carroll, S.B. (2005). Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila. Nature **433**: 481-7.

Gowik, U., Burscheidt, J., Akyildiz, M., Schlue, U., Koczor, M., Streubel, M., and Westhoff, P. (2004). cis-regulatory elements for mesophyll-specific gene expression in the C4 plant Flaveria trinervia, the promoter of the C4 phosphoenolpyruvate carboxylase gene. Plant Cell **16**: 1077-1090

Guyot, R. and Keller, B. (2004). Ancestral genome duplication in rice. Genome **47**: 610-4.

Haberlandt (1904). Physiologische Pflanzenanatomie. Leipzig: Wilhelm Engelmann.

Harlin-Cognato, A., Hoffman, E.A., and Jones, A.G. (2006). Gene cooption without duplication during the evolution of a male-pregnancy. Proc. Natl. Acad. Sci. USA **103**: 19407-12.

Harpster, M.H., and Taylor, W.C. (1986). Maize phosphoenolpyruvate carboxylase. Cloning and characterization of mRNAs encoding isozymic forms. J. Biol. Chem. **261**: 6132-6136.

Hatch, M.D. (1987). C_4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. Biochim. Biophys. Acta **895**: 81-106.

Hatch, M.D. (1999). C4 photosynthesis: a historical overview. C4 Plant Biology. R. F. Sage and R. K. Monson. San Diego, Academic Press**:** 17-46.

Hatch, M.D., and Burnell, J.N. (1990). Carbonic anhydrase activity in leaves and its role in the first step of C_4 photosynthesis. Plant Physiol. **93**: 825-828.

Hattersley, P.W. (1984). Characterization of C_4 type leaf anatomy in grasses (Poaceae). Mesophyll : bundle sheath area ratios. Ann. Bot. **53**: 163-179.

Hattersley, P.W. (1992). C_4 photosynthetic pathway variation in grasses (Poaceae): its significance for arid and semi-arid lands. Desertified grasslands: their biology and management. G. P. Chapman. London, Academic Press**:** 181-212.

Hattersley, P.W., and Browning, A.J. (1981). Occurence of the suberised lamella in leaves of grasses of different photosynthetic types. I. In parenchymatous bundle sheath and PCR (`Kranz`) sheaths. Protoplasma **109**: 371-333.

Hermans, J., and Westhoff, P. (1990). Analysis of expression and evolutionary relationships of phosphoenolpyruvate carboxylase genes in *Flaveria trinervia* (C_4) and *F. pringlei* (C_3). Mol. Gen. Genet. **224**: 459-468.

Hermans, J., and Westhoff, P. (1992). Homologous genes for the C_4 isoform of phosphoenolpyruvate carboxylase in a C_3- and a C_4-*Flaveria* species. Mol. Gen. Genet. **234**: 275-284.

Irish, V.F., and Litt, A. (2005). Flower development and evolution: gene duplication, diversification and redeployment. Curr. Opin. Genet. Dev. **15**: 454-60.

Izui, K., Matsumura, H., Furumoto, T., and Kai, Y. (2004). Phosphoenolpyruvate carboxylase: A new era of structural biology. Annu. Rev. Plant Biol. **55**: 69-84.

Jiao, J., and Chollet, R. (1991). Posttranslational regulation of phospho*enol*pyruvate carboxylase in C_4 and Crassulacean acid metabolism plants. Plant Physiol. **95**: 981-985.

Kanai, R., and Edwards, G.E. (1999). The biochemistry of C4 photosynthesis. C4 Plant Biology. R. F. Sage and R. K. Monson. San Diego, Academic Press**:** 49-87.

Kellogg, E.A. (1999). Phylogenetic aspects of the evolution of C4 photosynthesis. C4 Plant Biology. R. F. Sage and R. K. Monson. San Diego, Academic Press**:** 411-444.

Kirchhamer, C.V., Yuh, C.H., and Davidson, E.H. (1996). Modular cis-regulatory organization of developmentally expressed genes. Proc. Natl. Acad. Sci. USA **93**: 9322-8.

Kopriva, S., Chu, C.C., and Bauwe, H. (1996). Molecular phylogeny of *Flaveria* as deduced from the analysis of nucleotide sequences encoding the H-protein of the glycine cleavage system. Plant Cell Environ. **19**: 1028-1036.

Ku, M.S.B., Agarie, S., Nomura, M., Fukayama, H., Tsuchida, H., Ono, K., Hirose, S., Toki, S., Miyao, M., and Matsuoka, M. (1999). High-level expression of maize phospho*enol*pyruvate carboxylase in transgenic rice plants. Nat. Biotechnol. **17**: 76-80.

Ku, M.S.B., Wu, J., Dai, Z., Scott, R.A., Chu, C., and Edwards, G.E. (1991). Photosynthetic and photorespiratory characteristics of *Flaveria* species. Plant Physiol. **96**: 518-528.

Kubicki, A., Steinmüller, K., and Westhoff, P. (1994). Differential transcription of plastome-encoded genes in the mesophyll and bundle-sheath chloroplasts of the monocotyledonous NADP-malic enzyme type C4 plants maize and Sorghum. Plant. Mol. Biol. **25**: 669-679.

Lai, L.B., Tausta, S.L., and Nelson, T.M. (2002). Differential regulation of transcripts encoding cytosolic NADP-malic enzyme in C3 and C4 Flaveria species. Plant Physiol. **128**: 140-149.

Langdale, J.A., and Nelson, T. (1991). Spatial regulation of photosynthetic development in C4 plants. Trends Genet. **7**: 191-196.

Langdale, J.A., Zelitch, I., Miller, E., and Nelson, T. (1988). Cell position and light influence C_4 versus C_3 patterns of photosynthetic gene expression in maize. EMBO J. **7**: 3643-3651.

Latzko, E., and Kelly, J. (1983). The multi-faceted function of phosphoenolpyruvate carboxylase in C3 plants. Physiol. Vég. **21**: 805-815.

Leegood, R.C. (2002). C-4 photosynthesis: principles of CO2 concentration and prospects for its introduction into C-3 plants. J. Exp. Bot. **53**: 581-590.

Leegood, R.C., Lea, P.J., Adcock, M.D., and Häusler, R.E. (1995). The regulation and control of photorespiration. J Exp. Bot. **46** Suppl.: 1397-1414.

Leegood, R.C., and Walker, R.P. (1999). Regulation of the C4 pathway. C4 Plant Biology. R. F. Sage and R. K. Monson. San Diego, Academic Press**:** 89-131.

Lepiniec, L., Vidal, J., Chollet, R., Gadal, P., and Crétin, C. (1994). Phosphoenolpyruvate carboxylase: Structure, regulation and evolution. Plant Sci. **99**: 111-124.

Li, B., Pacquit, V., J.A., J., Duff, S.M.G., Maralihalli, G.B., Sarath, G., Condon, S.C., Vidal, J., and Chollet, R. (1997). Structural requirements for phosphorylation of C4-leaf phosphoenolpyruvate carboxylase by its highly regulated protein-serine kinase. A comparative study with synthetic-peptide substrates and mutant target proteins. Aust. J. Plant Phsiol. **24**: 443-449.

Li, J.J., and Herskowitz, I. (1993). Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. Science **262**: 1870-1874.

Long, S.P. (1999). Environmental responses. C4 Plant Biology. R. F. Sage and R. K. Monson. San Diego, Academic Press**:** 215-249.

Lynch, M., and Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. Science **290**: 1151-5.

Lynch, V.J., Roth, J.J., and Wagner, G.P. (2006). Adaptive evolution of Hox-gene homeodomains after cluster duplications. BMC Evol. Biol. **6**: 86.

Maere, S., De Bodt, S., Raes, J., Casneuf, T., Van Montagu, M., Kuiper, M., and Van de Peer, Y. (2005). Modeling gene and genome duplications in eukaryotes. Proc. Natl. Acad. Sci. USA **102**: 5454-9.

Marshall, J.S., Stubbs, J.D., and Taylor, W.C. (1996). Two genes encode highly similar chloroplastic NADP-malic enzymes in *Flaveria* - Implications for the evolution of C_4 photosynthesis. Plant Physiol. **111**: 1251-1261.

Matsuoka, M. (1995). The gene for pyruvate orthophosphate dikinase in C4 plants: structure, regulation and evolution. Plant Cell Physiol. **36**: 937-943.

Matsuoka, M., Kyozuka, J., Shimamoto, K., and Kano-Murakami, Y. (1994). The promoters of two carboxylases in a C_4 plant (maize) direct cell-specific, light-regulated expression in a C_3 plant (rice). Plant J. **6**: 311-319.

McGonigle, B., and Nelson, T. (1995). C_4 isoform of NADP-malate dehydrogenase. cDNA cloning and expression in leaves of C_4, C_3, and C_3-C_4 intermediate species of *Flaveria*. Plant Physiol. **108**: 1119-1126.

McKown, A.D., and Dengler, N.G. (2007). Key innovations in the evolution of Kranz anatomy and C4 vein pattern in Flaveria (Asteraceae). Am. J. Bot. **94**(3): 382-399.

Melzer, E., and O'Leary, M. H. (1987). Anaplerotic CO_2 fixation by phosphoenolpyruvate carboxylase in C(3). Plant Physiol. **84**: 58-60.

Monson, R.K. (1999). "The origins of C4 genes and evolutionary pattern in the C4 metabolic phenotype." C4 Plant Biology. R. F. Sage and R. K. Monson. San Diego, Academic Press**:** 377-410.

Monson, R.K. (2003). Gene duplication, neofunctionalization, and the evolution of C4 photosynthesis. Int. J. Plant Sci. **164** Suppl: 43-54.

Monson, R.K., and Moore, B.D. (1989). On the significance of C_3-C_4 intermediate photosynthesis to the evolution of C_4 photosynthesis. Plant Cell Environ. **12**: 689-699.

Moore, R.C., and Purugganan, M.D. (2005). The evolutionary dynamics of plant duplicate genes. Curr. Opin. Plant Biol. **8**: 122-128.

Nelson, T., and Dengler, N.G. (1992). Photosynthetic tissue differentiation in C_4 plants. Int. J. Plant Sci. **153**: S93-S105.

Nelson, T., and Langdale, J.A. (1989). Patterns of leaf development in C4 plants. Plant Cell **1**: 3-13.

Nelson, T., and Langdale, J.A. (1992). Developmental genetics of C_4 photosynthesis. Annu. Rev. Plant Physiol. in Plant Mol. Biol. **43**: 25-47.

Nomura, M., Katayama, K., Nishimura, A., Ishida, Y., Ohta, S., Komari, T., Miyao-Tokutomi, M., Tajima, S., and Matsuoka, M. (2000). The promoter of *rbcS* in a C3 plant (rice) directs organ-specific, light-dependent expression in a C4 plant (maize), but does not confer bundle sheath cell-specific expression. Plant Mol. Biol. **44**: 99-106.

Nomura, M., Sentoku, N., Nishimura, A., Lin, J.H., Honda, C., Taniguchi, M., Ishida, Y., Ohta, S., Komari, T., Miyao-Tokutomi, M., Kano-Murakami, Y., Tajima, S., Ku, M.S.B., and Matsuoka, M. (2000) a. The evolution of C4 plants: acquisition of *cis*-regulatory sequences in the promoter of C4-type pyruvate, orthophosphate dikinase gene. Plant J. **22**: 211-221.

Nomura, M., Sentoku, N., Tajima, S., and Matsuoka, M. (2000) b. Expression patterns of cytoplasmic pyruvate, orthophosphate dikinase of rice (C_3) and maize (C_4) in a C_3 plant, rice. Aust. J. Plant Physiol. **27**: 343-347.

Ogren, W.L. (1984). Photorespiration: pathways, regulation, and modification. Annu. Rev. Plant Physiol. **35**: 415-442.

Ohno, S. (1970). Evolution by gene duplication. Springer-Verlag, New York.

O'Leary, M.H. (1982). Phosphoenolpyruvate carboxylase: an enzymologists view. Annu. Rev. Plant Physiol. **33**: 297-315.

Olson, E.N. (2006). Gene regulatory networks in the evolution and development of the heart. Science **313**: 1922-7.

Onodera, Y., Suzuki, A., Wu, C.Y., Washida, H., and Takaiwa, F. (2001). A rice functional transcriptional activator, RISBZ1, responsible for endosperm-specific expression of storage protein genes through GCN4 motif. J. Biol. Chem. **276**: 14139-14152.

Outlaw, W.H.J., Du, Z., Xia Meng, F., Aghoram, K., Riddle, K.A., and Chollet, R. (2002). Requirements for activation of the signal-transduction network that leads to regulatory phosphorylation of leaf guard-cell phosphoenolpyruvate carboxylase during fusicoccin-stimulated stomatal opening. Arch. Biochem. Biophys. **407**: 63-71.

Patel, M., Siegel, A.J., and Berry, J.O. (2006). Untranslated regions of FbRbcS1 mRNA mediate bundle sheath cell-specific gene expression in leaves of a C4 plant. J. Biol. Chem. **281**: 25485-91.

Powell, A.M. (1978). Systematics of *Flaveria* (Flaveriinae-Asteraceae). Ann. Mo. Bot. Gard. **65**: 590-636.

Raghavendra, A.S. (1980). Characteristics of plant species intermediate between C_3 and C_4 pathways of photosynthesis: their focus of mechanism and evolution of C_4 syndrome. Photosynthetica **14**: 271-283.

Rajagopalan, A.V., Devi, M.T., and Raghavendra, A.S. (1994). Molecular biology of C_4 phospho*enol*pyruvate carboxylase: Structure, regulation and genetic engineering. Photosynth. Res. **39**: 115-135.

Rosche, E., and Westhoff, P. (1995). Genomic structure and expression of the pyruvate, orthophosphate dikinase gene of the dicotyledonous C_4 plant *Flaveria trinervia* (Asteraceae). Plant Mol. Biol. **29**: 663-678.

Sage, R.F. (2004). The evolution of C4 photosynthesis. New Phytol. **161**: 341-370.

Sage, R.F., Li, M., and Monson, R.K. (1999). The taxonomic distribution of C4 photosynthesis. C4 Plant Biology. R. F. Sage and R. K. Monson. San Diego, Academic Press**:** 551-584.

Schäffner, A.R., and Sheen, J. (1991). Maize *rbcS* promoter activity depends on sequence elements not found in dicot *rbcS* promoters. Plant Cell **3**: 997-1012.

Schäffner, A.R., and Sheen, J. (1992). Maize C_4 photosynthesis involves differential regulation of phosphoenolpyruvate carboxylase genes. Plant J. **2**: 221-232.

Sheen, J. (1991). Molecular mechanisms underlying the differential expression of maize pyruvate, orthophosphate dikinase genes. Plant Cell **3**: 225-245.

Sheen, J. (1999). C4 gene expression. Annu .Rev. Plant. Physiol. in Plant Mol. Biol. **50**: 187-217.

Simillion, C., Vandepoele, K., Van Montagu, M., Zabeau, M., and Van de Peer, Y. (2002). The hidden duplication past of Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA **99**: 13627-32.

Singh, K.B. (1998). Transcriptional regulation in plants: The importance of combinatorial control. Plant Physiol. **118**: 1111-1120.

Stockhaus, J., Poetsch, W., Steinmüller, K., and Westhoff, P. (1994). Evolution of the C_4 phosphoenolpyruvate carboxylase promoter of the C_4 dicot *Flaveria trinervia*: an expression analysis in the C_3 plant tobacco. Mol. Gen. Genet. **245**: 286-293.

Stockhaus, J., Schlue, U., Koczor, M., Chitty, J.A., Taylor, W.C., and Westhoff, P. (1997). The promoter of the gene encoding the C_4 form of phosphoenolpyruvate carboxylase directs mesophyll specific expression in transgenic C_4 *Flaveria* spp. Plant Cell **9**: 479-489.

Svensson, P., Bläsing, O.E., and Westhoff, P. (1997). Evolution of the enzymatic characteristics of C_4 phosphoenolpyruvate carboxylase: a comparison of the orthologous ppcA phosphoenolpyruvate carboxylases of *Flaveria trinervia* (C_4) and *F. pringlei* (C_3). Eur. J. Biochem. **246**: 452-460.

Svensson, P., Bläsing, O.E., and Westhoff, P. (2003). Evolution of C4 phosphoenolpyruvate carboxylase. Arch. Biochem. Biophys. **414**: 180-188.

Tarczynski, M.C., and Outlaw, J.W.H. (1990). Partial characterization of guard-cell phosphoenolpyruvate carboxylase: kinetic datum collection in real time from single-cell activities. Arch. Biochem. Biophys. **280**: 153-158.

Ting, I.P., and Osmond, C.B. (1973). Photosynthetic phosphoenolpyruvate carboxylase. Characteristics of allozymes from leaves of C_3 and C_4 plants. Plant Physiol. **51**: 439-447.

True, J.R., and Carroll, S.B. (2002). Gene co-option in physiological and morphological evolution. Annu. Rev. Cell. Dev. Biol. **18**: 53-80.

Uedan, K., and Sugiyama, T. (1976). Purification and characterization of phosphoenolpyruvate carboxylase from maize leaves. Plant Physiol. **57**: 906-910.

Van de Poele, K., Simillion, C., and Van de Peer, Y. (2003). Evidence that rice and other cereals are ancient aneuploids. Plant Cell **15**: 2192-2202.

Vision, T.J., Brown, D.G., and Tanhksley, S.D. (2000). The origins of genomic duplications in Arabidopsis. Science **290**: 2114-2117.

von Caemmerer, S., and Furbank, R.T. (2003). The C-4 pathway: an efficient CO2 pump. Photosynth. Res. **77**: 191-207.

Voznesenskaya, E.V., Franceschi, V.R., Kiirats, O., Artyusheva, E.G., Freitag, H., and Edwards, G.E. (2002). Proof of C4 photosynthesis without Kranz anatomy in Bienertia cycloptera (Chenopodiaceae). Plant J. **31**: 649-662.

Voznesenskaya, E.V., Franceschi, V.R., Kiirats, O., Freitag, H., and Edwards, G.E. (2001). Kranz anatomy is not essential for terrestrial C-4 plant photosynthesis. Nature **414**: 543-546.

Wang, X., Shi, X., Hao, B., Ge, S., and Luo, J. (2005). Duplication and DNA segmental loss in the rice genome: implications for diploidization. New Phytol. **165**(3): 937-46.

Wedding, R.T., Black, M.K., and Meyer, C.R. (1990). Inhibition of phospho*enol*pyruvate carboxylase by malate. Plant Physiol. **92**: 456-461.

Wendel, J.F. (2000). Genome evolution in polyploids. Plant Mol. Biol. **42**: 225-249.

Westhoff, P., Svensson, P., Ernst, K., Bläsing, O., Burscheidt, J., and Stockhaus, J. (1997). Molecular evolution of C4 phosphoenolpyruvate carboxylase in the genus *Flaveria*. Aust. J. Plant Physiol. **24**: 429-436.

Wilson, A.C., Carlson, S.S., and White, T.J. (1977). Biochemical evolution. Annu. Rev. Biochem. **46**: 573-639.

Windhövel, A., Hein, I., Dabrowa, R., and Stockhaus, J. (2001). Characterization of a novel class of plant homeodomain proteins that bind to the C_4 phospho*enol*pyruvate carboxylase gene of *Flaveria trinervia*. Plant Mol. Biol. **45**: 201-214.

Woo, K.C., Anderson, J.M., Boardman, N.K., Downton, W.J.S., Osmond, C.B., and Thorne, S.W. (1970). Deficient photosystem II in agranal bundle sheath chloroplasts of C_4 plants. Proc. Natl. Acad. Sci. USA **67**: 18-25.

Wray, G.A. (2007). The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. **8**: 206-16.

Wyrich, R., Dreßen, U., Brockmann, S., Streubel, M., Chang, C., Qiang, D., Paterson, A.H., and Westhoff, P. (1998). The molecular basis of C_4 photosynthesis in sorghum: isolation, characterization and RFLP mapping of mesophyll- and bundle-sheath-specific cDNAs obtained by differential screening. Plant Mol. Biol. **37**: 319-335.

Yang, X., Tuskan, G.A., and Cheng, M.Z. (2006). Divergence of the Dof gene families in poplar, Arabidopsis, and rice. Plant Physiol. **142**: 820-830.
Zhang, J. (2003). Evolution by gene duplication: an update. Trends Eco. Evo. **18**: 292-298.

Acknowledgment/Danksagung

Ich danke...

...Prof. Dr. Peter Westhoff, an dessen Institut diese Arbeit entstand und für sein Interesse am Fortgang der Arbeit

...Prof. Dr. Georg Groth für die Übernahme des Koreferates

...Udo Gowik für seine stete Hilfe wenn mal ein kleines Problem auftauchte und für die Hilfe bei den „Papern“

...Prof. Dr. Claus Seidel für die Einführung in die Geheimnisse der Anisotropie

...Maria Koczor, Monika Streubel und Ute Schlue für die Aufzucht und Pflege meiner lieben Flaveria Pflanzen

...Karin Ernst für ihr Korrektur-Lesen

...dem Hauptlabor 1 für eine lustige, chaotische und entspannte Arbeitsatmosphäre

...Monika Streubel für ihre Freundschaft

...dem gesamten Institut Botanik IV. für eine vergnügliche Arbeitsatmosphäre

...meiner Familie für ihre Unterstützung

...und meinem Verlobten Bernhard Stein für seine Aufmunterung und Unterstützung wenn es mal nicht so gut lief

Manuscripts

1) Udo Gowik, Janet Burscheidt, Meryem Akyildiz, Ute Schlue, Maria Koczor, Monika Streubel and Peter Westhoff (2004). ***Cis*-regulatory elements for mesophyll-specific gene expression in the C_4 plant *Flaveria trinervia*, the promoter of the C_4 phospho*enol*pyruvate carboxylase gene**. The Plant Cell 16: 1077-1090.

2) Meryem Akyildiz, Udo Gowik, Maria Koczor, Monika Streubel and Peter Westhoff (2007). **Evolution and function of a *cis*-regulatory module for mesophyll-specific gene expression in the C_4 dicot *Flaveria trinervia***. Submitted to Plant Cell for publication.

3) Meryem Akyildiz, Ming Chang Tsai, Claus Seidel and Peter Westhoff (2007). **Basic leucine zipper proteins interact with MEM1, the mesophyll specificity *cis*-regulatory element of the C_4 phospho*enol*pyruvate carboxylase gene of *Flaveria trinervia*.**

The Plant Cell, Vol. 16, 1077–1090, May 2004, www.plantcell.org © 2004 American Society of Plant Biologists

RESEARCH ARTICLES

cis-Regulatory Elements for Mesophyll-Specific Gene Expression in the C_4 Plant *Flaveria trinervia*, the Promoter of the C_4 Phosphoenolpyruvate Carboxylase Gene [W]

Udo Gowik, Janet Burscheidt, Meryem Akyildiz, Ute Schlue, Maria Koczor, Monika Streubel, and Peter Westhoff[1]

Heinrich-Heine-Universität, Institut für Entwicklungs und Molekularbiologie der Pflanzen, 40225 Düsseldorf, Germany

C_4 photosynthesis depends on the strict compartmentalization of CO_2 assimilatory enzymes. *cis*-regulatory mechanisms are described that ensure mesophyll-specific expression of the gene encoding the C_4 isoform of phosphoenolpyruvate carboxylase (*ppcA1*) of the C_4 dicot *Flaveria trinervia*. To elucidate and understand the anatomy of the C_4 *ppcA1* promoter, detailed promoter/reporter gene studies were performed in the closely related C_4 species *F. bidentis*, revealing that the C_4 promoter contains two regions, a proximal segment up to −570 and a distal part from −1566 to −2141, which are necessary but also sufficient for high mesophyll-specific expression of the β-glucuronidase reporter gene. The distal region behaves as an enhancer-like expression module that can direct mesophyll-specific expression when inserted into the *ppcA1* promoter of the C_3 plant *F. pringlei*. Mesophyll expression determinants were restricted to a 41-bp segment, referred to as mesophyll expression module 1 (*Mem1*). Evolutionary and functional studies identified the tetranucleotide sequence CACT as a key component of *Mem1*.

INTRODUCTION

C_4 plants are characterized by high rates of photosynthesis as well as an efficient use of water and nitrogen resources. This is because of their unique mode of carbon assimilation that concentrates CO_2 at the site of ribulose bisphosphate carboxylase/oxygenase. The functioning of C_4 photosynthesis is dependent upon the strict compartmentation of the CO_2 assimilatory enzymes into two distinct cell types, mesophyll and bundle-sheath cells. The primary carboxylating enzyme, phosphoenolpyruvate carboxylase, accumulates exclusively in the mesophyll cells, and the secondary carboxylase, ribulose bisphosphate carboxylase/oxygenase, and the decarboxylating enzymes, such as NADP-dependent malic enzyme, are restricted to the bundle-sheath cells (Hatch, 1987).

This division of labor between mesophyll and bundle-sheath cells is the result of differential gene expression. In NADP-malic enzyme-type C_4 species, for instance, transcripts for phosphoenolpyruvate carboxylase, pyruvate phosphate dikinase, NADP-malic enzyme, and the small subunit of ribulose bisphosphate carboxylase/oxygenase, accumulate differentially in the two cell types. This differential accumulation is largely because of transcriptional control (Sheen, 1999).

[1] To whom correspondence should be addressed. E-mail west@uni-duesseldorf.de; fax 49-211-81-14871.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Peter Westhoff (west@uni-duesseldorf.de).
[W] Online version contains Web-only data.
Article, publication date, and citation information can be found at www.plantcell.org/cgi/doi/10.1105/tpc.019729.

C_4 plants occur in at least 18 families of monocotyledonous and dicotyledonous plants. This indicates that C_4 plants must have evolved several times independently from C_3 ancestors during the evolution of angiosperms (Kellogg, 1999; Sage et al., 1999). The multiple independent origin of C_4 photosynthesis suggests that the evolution of a C_3 into a C_4 species must have been relatively easy in genetic terms. The available molecular data on the C_4 cycle enzymes support this point of view. None of the C_4 enzymes, phosphoenolpyruvate carboxylase (PEPC), pyruvate orthophosphate dikinase, or NADP-dependent malic enzyme, are unique to C_4 plants. Nonphotosynthetic isoforms of these enzymes are also present in C_3 species and in the nonphotosynthetic tissues of C_4 species. The ubiquitous presence of these nonphotosynthetic isoforms of the C_4 cycle enzymes in C_3 plants indicates that these C_3 isoforms served as the starting point for the evolution of the C_4 genes (reviewed in Monson, 1999).

As a starting point to understanding the molecular basis of the evolution of C_4 genes, we are focusing on the C_4 gene for PEPC and are using the genus Flaveria (Asteraceae) (Powell, 1978) as an experimental system. Flaveria has C_3 and C_4 species and a large number of C_3-C_4 photosynthetic intermediates (reviewed in Edwards and Ku, 1987). These intermediates differ in the expression of the C_4 photosynthetic traits, and there is convincing evidence that at least some of these species are true evolutionary intermediates (Monson and Moore, 1989).

Three major changes must have occurred during C_3-to-C_4 evolution to transform a C_3 PEPC gene into a C_4 gene (reviewed in Westhoff and Gowik, 2004). C_4 PEPC genes are highly expressed (Hermans and Westhoff, 1990; Crétin et al., 1991), whereas C_3 PEPC transcripts generally occur only in moderate

amounts in plant tissues (Crétin et al., 1991; Ernst and Westhoff, 1996). First, therefore, promoter strength had to increase. Second, a mesophyll-specific expression pattern had to evolve because the strict compartmentation of PEPC is imperative for an effectively functioning C_4 cycle (Hatch, 1987). Finally, the metabolic context of C_3 and C_4 PEPCs differs; therefore, the PEPC protein had to change its kinetic and regulatory enzyme properties to meet the metabolic requirements of C_4 photosynthesis (Svensson et al., 2003).

The C_4 PEPCs of C_4 Flaveria species are encoded by the phosphoenolpyruvate carboxylase A (*ppcA*) gene class (Hermans and Westhoff, 1992). *ppcA* orthologous PEPC genes are found in all C_3 and C_3-C_4 intermediate Flaveria species, indicating that this PEPC gene class was already present in the last common ancestor of present C_3 and C_4 Flaveria species (Westhoff and Gowik, 2004). The comparative enzymatic analysis of *ppcA* PEPC proteins from C_3, C_3-C_4 intermediate, and C_4 Flaveria species revealed that the *ppcA* PEPCs of *F. pringlei* (C_3) and *F. trinervia* (C_4) are typical C_3 and C_4 PEPCs, respectively, and that only a few amino acid changes, most notably a C_4 invariant Ser residue in the vicinity of the catalytic site, were responsible for the observed differences in kinetic and regulatory behavior (Svensson et al., 1997; Bläsing et al., 2000). The *ppcA* PEPCs from the C_3-C_4 species *F. pubescens* and *F. brownii* were found to be intermediate, indicating that the *ppcA* PEPCs changed gradually from C_3 to C_4 (Engelmann et al., 2003), and this PEPC gene class could serve as an evolutionary model to unravel the C_4-associated changes in enzyme and gene expression characteristics (Svensson et al., 2003; Westhoff and Gowik, 2004).

Analysis of *ppcA1* promoter/β-glucuronidase (GUS) reporter gene fusions in the C_4 plant *F. bidentis* revealed that the *ppcA1* promoter of *F. trinervia* directs high levels of reporter gene expression in the mesophyll cells (Stockhaus et al., 1997). This demonstrated that the expression of the corresponding gene is largely determined by transcription and that the 2188 bp (with reference to the AUG start codon of the *ppcA1* reading frame) of the 5′ flanking sequences contain all the essential *cis*-regulatory elements for a high and mesophyll-specific expression. By contrast, the 2538 bp (with reference to the AUG start codon) of the 5′ flanking sequences of the *ppcA1* gene of *F. pringlei* were found to be a weak promoter and did not direct any organ- or cell-specific expression (Stockhaus et al., 1997). Both promoters thus exhibited all the attributes that were expected from the accumulation patterns of their corresponding RNAs and proteins (Höfer et al., 1992; Ernst and Westhoff, 1996).

To fully understand the anatomy of the C_4 *ppcA1* promoter and to identify the *cis*-regulatory elements that are essential for its mesophyll specificity, detailed promoter reporter gene analyses with transgenic *F. bidentis* were performed. These experiments revealed that the C_4 promoter contains two regions, a proximal region up to −570 (PR) and a distal region from −1566 to −2141 (DR), which are necessary and sufficient for a high mesophyll-specific expression. The DR behaves as an enhancer-like expression module and is able to confer a mesophyll expression component to the *ppcA1* promoter of *F. pringlei*. Further dissection of the DR identified a 41-bp module (mesophyll expression module 1 [*Mem1*]) that in conjunction with the PR, is sufficient for mesophyll-specific expression. Evolutionary and functional analyses identified the tetranucleotide CACT as a key element of *Mem1*.

RESULTS

The Distal Segment of the C_4 *ppcA1* Promoter Is Required for both Expression Specificity and Quantity and Behaves as an Enhancer-Like Expression Module

In the C_3 plant tobacco (*Nicotiana tabacum*), the C_4 *ppcA1* promoter behaved essentially as a palisade parenchyma–specific promoter. The expression in the palisade parenchyma was lost when the 5′ distal 1618 bp of 5′ distal sequences were removed and the remaining 570 bp of proximal sequences were analyzed for promoter activity (Stockhaus et al., 1994). This finding suggested that the 5′ DR of the promoter contains *cis*-regulatory elements that are absolutely essential for a high level of expression and for mesophyll specificity. To define this distal promoter region precisely, a systematic deletion analysis was performed using the high level of expression in the palisade parenchyma cells of the C_3 plant tobacco as a test system.

When the *ppcA1* starting promoter of 2188 bp (named ppcA-L_{Ft}; Figure 1) was shortened by 623 bp (construct ppcA-1,5_{Ft}; Figure 2A), the expression activity was almost entirely lost. It is highly significant that no palisade parenchyma expression was detected by histochemical staining (data not shown). Further deletion of 5′ promoter sequences (construct ppcA-1,0_{Ft}; Figure 2A) influenced the resulting promoter activity as compared with the ppcA-1,5_{Ft} promoter construct only marginally (Figure 2B). Both the ppcA-1,5_{Ft} and ppcA-1,0_{Ft} constructs showed a lower

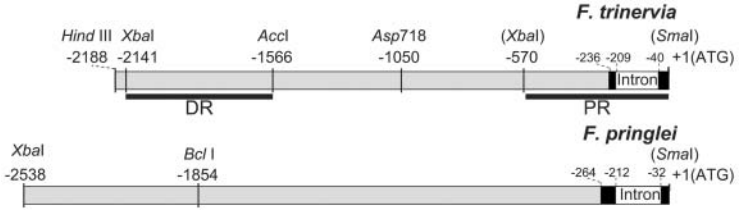


Figure 1. Restriction Maps of the 5′ Flanking Regions of the *ppcA1* Genes of *F. trinervia* and *F. pringlei*.

Nucleotide positions are numbered with respect to the AUG translational initiation codon. The DR and PR of the *ppcA1* promoter of *F. trinervia* are labeled by black bars.

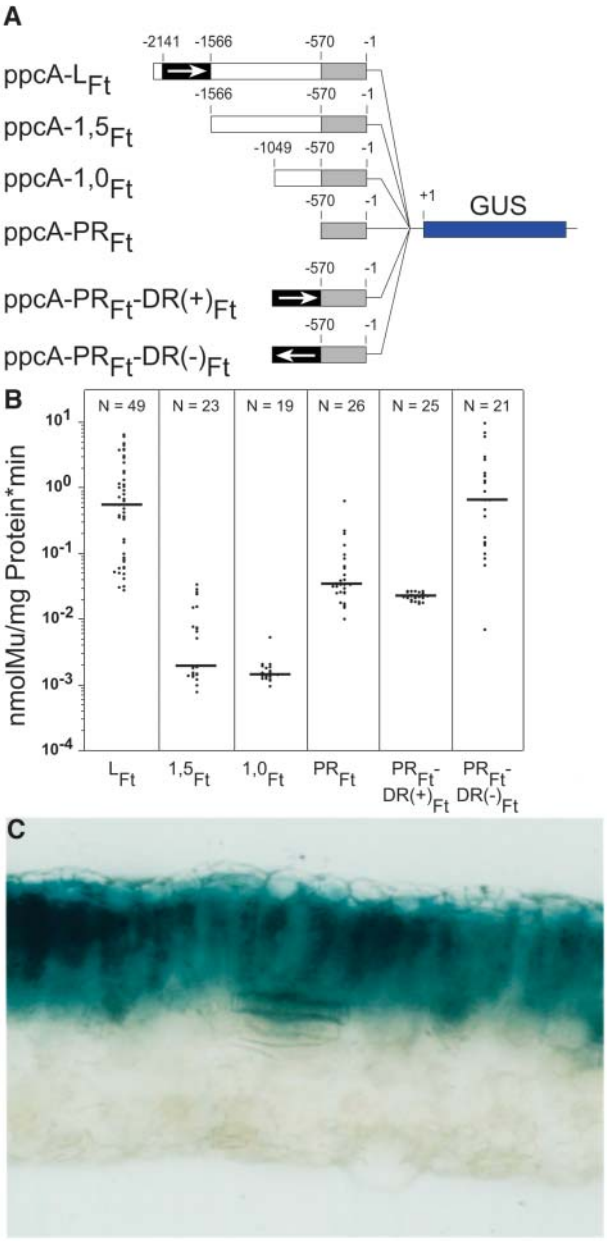


Figure 2. Deletion Analysis of the *ppcA1* Promoter of *F. trinervia* in the C_3 Plant Tobacco.

(A) Structures of the *ppcA1*/*GUS* chimerical genes used for tobacco transformation. The nucleotide numbers refer to the translation initiation codon. The DR of the *ppcA1*-promoter from the C_4 plant *F. trinervia* is indicated by a black box, and the PR is indicated by a gray box.

(B) GUS activities in leaves of transgenic tobacco plants. The median value of the GUS activities is expressed in nanomoles of the reaction product 4-methylumbelliferone (Mu) generated per milligram of protein per minute. The numbers of independent transgenic plants investigated (N) are indicated at the top of each column.

(C) Histochemical localization of GUS activity in a leaf section of a transgenic tobacco plant transformed with the ppcA-PR$_{Ft}$-DR(−)$_{Ft}$ construct. Incubation for 6 h.

promoter activity than the ppcA-PR$_{Ft}$ promoter fragment (Figures 2A and 2B). These observations suggested that the region between base pairs −570 and −1565 appears to contain sequences that reduce promoter activity but that the DR between base pairs −1565 and −2188 is absolutely essential for the C_4 *ppcA1* promoter activity in tobacco.

This fact was tested directly by fusing the distal part between base pairs −1566 and −2141 in direct and inverse orientation with the proximal −570 bp of promoter sequences (constructs ppcA-PR$_{Ft}$-DR(+)$_{Ft}$ and ppcA-PR$_{Ft}$-DR(−)$_{Ft}$; Figure 2A) and by analyzing the promoter activities in transgenic tobacco. The GUS activity of the ppcA-PR$_{Ft}$-DR(−)$_{Ft}$ construct in the leaf was comparable to that of the L$_{Ft}$ chimerical gene, but the activity of the ppcA-PR$_{Ft}$-DR(+)$_{Ft}$ construct was drastically reduced (Figure 2B). For the ppcA-PR$_{Ft}$-DR(−)$_{Ft}$ construct, histochemical analyses showed that this promoter directed a palisade parenchyma–specific expression of the GUS reporter gene (Figure 2C). It follows from these expression analyses with the C_3 plant tobacco that the distal and proximal parts of the promoter will be sufficient for the *ppcA1* promoter activity and that the nucleotide sequences between −570 and −1566 are probably not necessary for its expression specificity.

To verify these conclusions, transformation experiments with the C_4 plant *F. bidentis* were performed. *F. bidentis* is very closely related to *F. trinervia* and is used for these experiments (Stockhaus et al., 1997) because this species, in contrast with *F. trinervia*, may be transformed by *Agrobacterium tumefaciens* (Chitty et al., 1994). Quantitative measurements of GUS activity showed that ppcA-PR$_{Ft}$-DR(−)$_{Ft}$ promoter was approximately half as active as the full C_4 *ppcA1* promoter, whereas the ppcA-PR$_{Ft}$-DR(+)$_{Ft}$ promoter activity was reduced to approximately one-quarter (Figure 3B). Both promoters directed a mesophyll-specific expression of the GUS reporter gene and showed the same expression pattern as the full C_4 *ppcA1* promoter (Stockhaus et al., 1997). Two conclusions were drawn from these experiments. First, the DR and PR of the promoter (Figure 1) are sufficient for an elevated and mesophyll-specific promoter activity (i.e., the nucleotide sequences between −570 and −1566 are essentially dispensable). Second, because the C_4-DR functions both in the correct and the inverse orientation, this *cis*-regulatory region shows the typical features of a transcriptional enhancer (Blackwood and Kadonaga, 1998).

The DR of the C_4 *ppcA1* Promoter Provides Mesophyll Specificity but No Raised Expression Quantity in the Context of the C_3 *ppcA1* Promoter

The C_4-DR could function as a C_4 expression module that confers both height and specificity of expression. If this were true, one should expect that upon transfer of the C_4-DR into the *ppcA1* promoter of the C_3 plant *F. pringlei*, the C_4-C_3 hybrid promoter would behave like a C_4 *ppcA1* promoter and show a high level of expression in the mesophyll cells. To test this, the DR of the C_3 *ppcA1* promoter from nucleotides −2538 to −1854 (Figure 1) was removed, giving rise to ppcA-M$_{Fp}$, and replaced by the C_4-DR in correct orientation (Figure 4A).

It is known from previous work (Stockhaus et al., 1997) that the *ppcA1* promoter of *F. pringlei* (ppcA-L$_{Fp}$) is relatively weak when

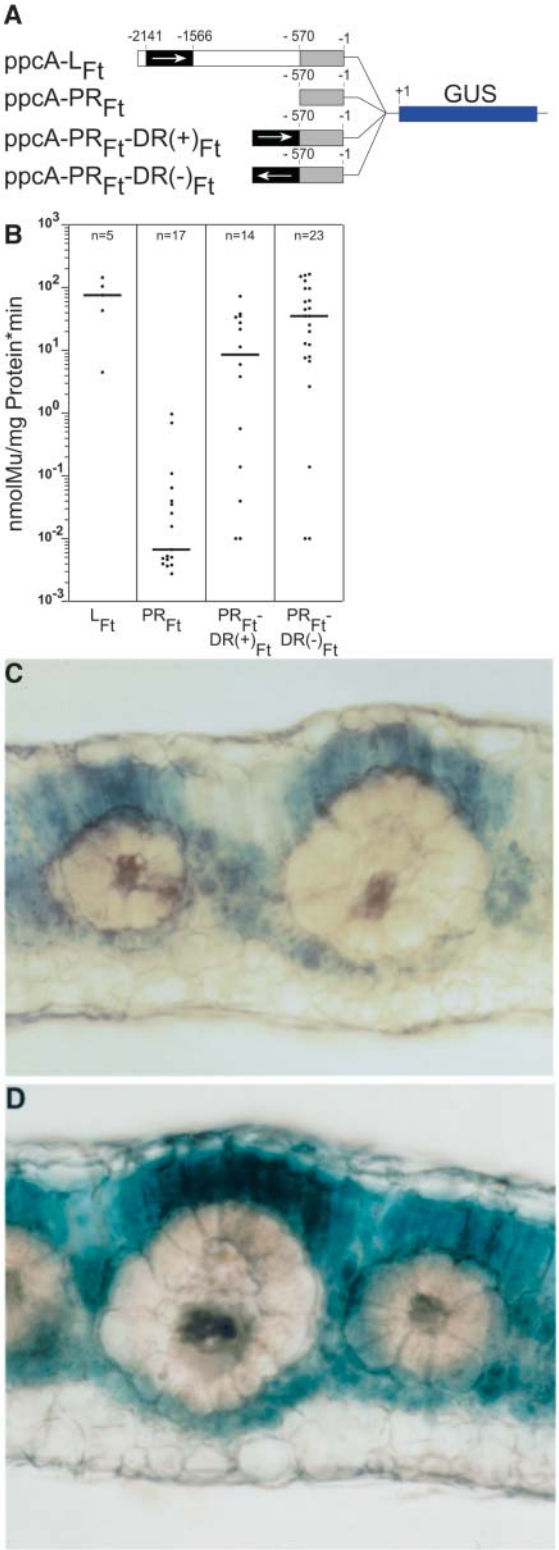


Figure 3. Analysis of the *ppcA1* GUS Reporter Gene Constructs ppcA-PR_{Ft}-$DR(+)_{Ft}$ and ppcA-PR_{Ft}-$DR(-)_{Ft}$ in Transgenic *F. bidentis*.

(A) Structures of the *ppcA1/GUS* chimerical genes.

(B) GUS activities in leaves of transgenic *F. bidentis* plants. The numbers of independent transgenic plants investigated (n) are indicated at the top of each column. Median values are shown. Mu, 4-methylumbelliferone.

(C) and **(D)** Histochemical localization of GUS activity in leaf sections of transgenic *F. bidentis* plants transformed with the ppcA-PR_{Ft}-$DR(+)_{Ft}$ **(C)** or the ppcA-PR_{Ft}-$DR(-)_{Ft}$ construct **(D)**. Incubation time was 20 min in case of the ppcA-PR_{Ft}-$DR(+)_{Ft}$ plant and 10 min in case of the ppcA-PR_{Ft}-$DR(-)_{Ft}$ plant.

compared with the C_4 *ppcA1* promoter and directs expression in all cells of the leaf, including the vascular bundle. The shortened promoter version ppcA-M_{Fp} behaves similarly (Figures 4B and 4C). The addition of the C_4-DR to the ppcA-M_{Fp} (resulting in ppcA-M_{Fp}-DR_{Ft}) causes just a small increase in the expression strength (Figure 4B). However, the in situ analysis of the transgenic plants revealed that the C_4-C_3 chimerical promoter had acquired a mesophyll expression component that was not detectable with the ppcA-M_{Fp} construct (Figure 4E). A visual comparison of the in situ promoter activities of ppcA-M_{Fp} and ppcA-M_{Fp}-DR_{Ft} also suggested that the C_4-DR did not only add a mesophyll expression component to the *ppcA1* promoter part of *F. pringle* but reduced its original activity in bundle-sheath cells and vascular tissue. We concluded from these experiments that the C_4-DR contains mesophyll expression components that are not able to increase the strength of the C_3 *ppcA1* promoter substantially.

Mapping of *cis*-Regulatory Elements in the C_4 DR of the C_4 *ppcA1* Promoter

The C_4-DR consists of 575 bp. To identify the *cis*-regulatory element(s) within the C_4-DR more precisely, this region was dissected into three pieces of approximately equal size that overlap by 50 bp (Figure 5). Each fragment of the C_4-DR, named a to c, was fused with the C_4-PR of the *ppcA1* promoter of *F. trinervia* in the correct orientation, and the resulting constructs, ppcA-PR_{Ft}-DRa_{Ft}, ppcA-PR_{Ft}-DRb_{Ft}, and ppcA-PR_{Ft}-DRc_{Ft} (Figure 6A), were transformed into *F. bidentis.*

The ppcA-PR_{Ft}-DRa_{Ft} and ppcA-PR_{Ft}-DRb_{Ft} promoters directed a clear and reproducible GUS expression in the mesophyll cells (Figures 6C and 6D), although their activities were reduced by ~5 to 10 times when compared with the activity of the ppcA-PR_{Ft}-$DR(+)_{Ft}$ reference promoter (cf Figures 6B and 3B). This indicates that both the ppcA-PR_{Ft}-DRa_{Ft} and ppcA-PR_{Ft}-DRb_{Ft} promoters harbor *cis*-regulatory elements that are sufficient for mesophyll-specific transcription. By contrast, the ppcA-PR_{Ft}-DRc_{Ft} promotor produced a minute amount of GUS activity that is below the level of histochemical detection (Figure 6E) but that is higher than the activity of the ppcA-PR_{Ft} construct (Figure 6B). This suggested that this segment of the C_4-DR might contain some weak transcriptional enhancing element(s). These elements are not essential for mesophyll-specific gene expression, but they may interact with the *cis*-regulatory elements of the a and b parts, thereby increasing their mesophyll enhancer activity.

Subfragments a and b of the C_4-DR were able to direct a mesophyll-specific expression. This implies that either the two

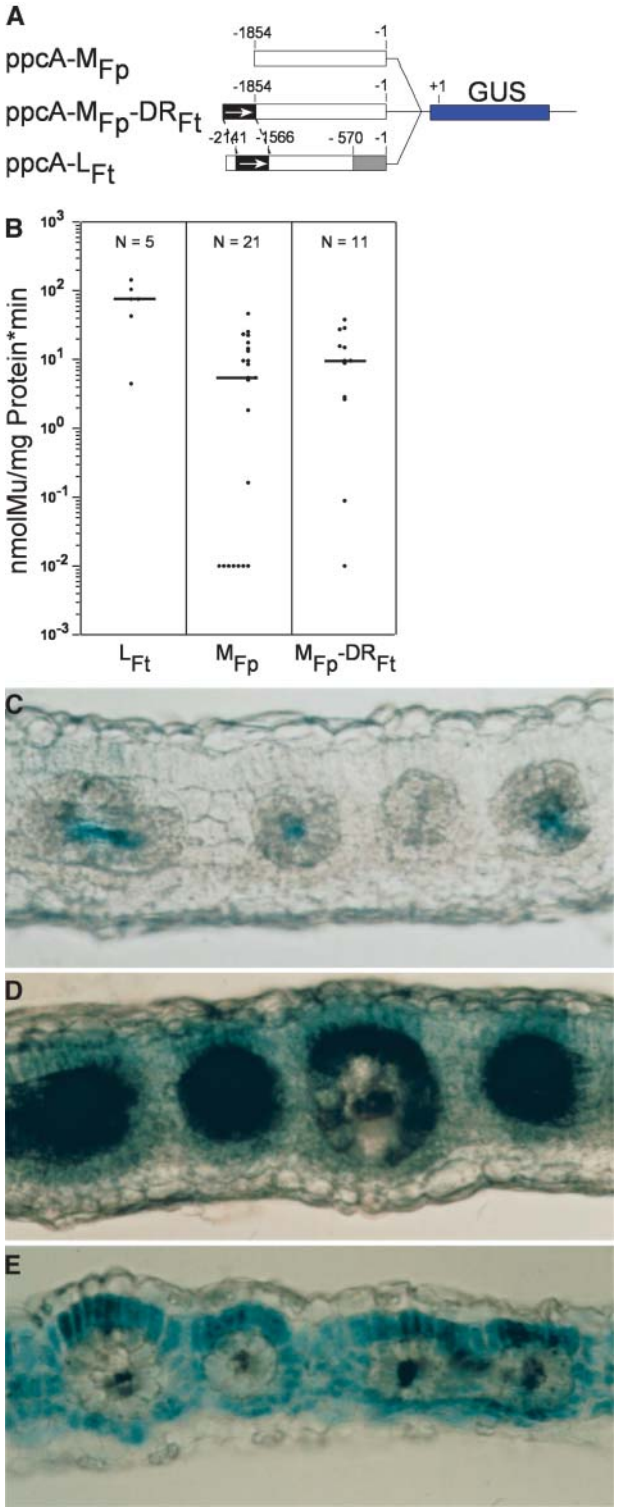


Figure 4. Analysis of the *ppcA1* GUS Reporter Gene Constructs ppcA-M_{Fp} and ppcA-M_{Fp}-DR(+)$_{Ft}$ in Transgenic *F. bidentis*.

(A) Structures of the *ppcA1/GUS* chimerical genes.

(B) GUS activities in leaves of transgenic *F. bidentis* plants. The numbers of independent transgenic plants investigated (N) are indicated at the top of each column. Median values are shown. Mu, 4-methylumbelliferone.

(C) to **(E)** Histochemical localization of GUS activity in leaf sections of transgenic *F. bidentis* plants transformed with ppcA-M_{Fp} **[C]** and **[D]**) or ppcA-M_{Fp}-DR(+)$_{Ft}$ **(E)**. Incubation times were 22 min **(C)**, 48 min **(D)**, and 43 min **(E)**.

segments contain distinct and different mesophyll specificity elements or that the overlapping stretch of 41 bp harbors a *cis*-regulatory element for mesophyll expression. To test these possibilities, one and four tandemly oriented copies of the a/b-overlapping fragment were fused in direct orientation with the PR segment of the C_4 *ppcA1* promoter. The resulting constructs, ppcA-PR_{Ft}-DRa/b_{Ft} and ppcA-PR_{Ft}-DR4a/b_{Ft} (Figure 7A), were analyzed in transgenic *F. bidentis*.

Both constructs exhibited similar expression levels (Figure 7B) and directed a mesophyll-specific expression of the GUS reporter gene (Figures 7C and 7D). It follows that the *a/b* overlapping C_4-DR fragment contains determinants for mesophyll-specific gene expression, and the segment was designated as *Mem1*.

The Tetranucleotide CACT Is Essential for *Mem1* Function

To identify the putative location of mesophyll-specific determinants within the 41-bp *Mem1*, its sequence was compared with the available *ppcA1* promoter sequences of *F. pringlei*. The search resulted in the unambiguous identification of a C_4-DR counterpart in the *F. pringlei* promoter at its very 5′ end. Sequences matching the 30 bp of the 3′ part of *Mem1* (named part B, Figure 8B) were detected in the *F. pringlei* promoter sequence. However, the 11 bp of the 5′ terminal sequences (named part A, Figure 8B) were lacking in the *F. pringlei* sequence. This suggests that the C_3 *ppcA1* promoter of *F. pringlei* contains only the homolog of part B of *Mem1* or, alternatively, that homologous sequences of part A are present in the promoter but have not been detected yet because they are located further upstream. To clarify this ambiguity, the available 5′ flanking sequences of the *ppcA1* gene of *F. pringlei* were extended by vectorette PCR (Siebert et al., 1995). Part A–type sequences were indeed shown to be separated from part B by 108 bp of intervening sequences (Figure 8).

The comparison of *Mem1* and its homolog in *F. pringlei* shows two remarkable features. The A part differs only in one single nucleotide at the very 5′ end (labeled in Figure 8). *Mem1* of *F. trinervia* holds a guanine in this position, and there is an adenine in the *Mem1* homolog of *F. pringlei*. More prominent is the difference in part B. A tetranucleotide (CACT) is present in the *Mem1* of *F. trinervia* but is absent in the *F. pringlei* sequence. The remainder of part B sequences is virtually identical in both promoters.

To elucidate which of the observed differences between the *Mem1* of *F. trinervia* and its homolog in *F. pringlei* are candidates for mesophyll expression determinants, we pursued a comparative approach. The 5′ flanking sequences of *ppcA1*-type genes were isolated by vectorette PCR (Siebert et al., 1995) from another C_4 species of Flaveria (i.e., *F. bidentis*), from two C_4-like plants, *F. palmeri* and *F. vaginata*, and from an additional C_3 species (*F. cronquistii*).

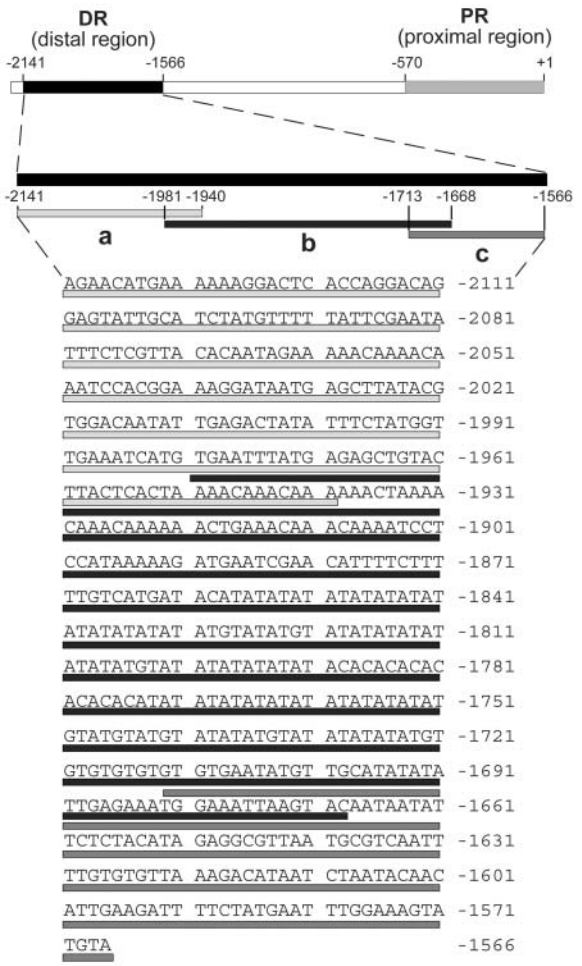


Figure 5. Nucleotide Sequence of the DR of the *ppcA1* Promoter of *F. trinervia* Showing the Location of the Three Subfragments a, b, and c.

A comparison of the 5′ flanking regions identified in each case *Mem1* homologous sequences where the A and B parts were, as in *F. pringlei*, separated by 97 to 108 bp (Figure 8). The A parts of all C_4 and C_4-like species showed a guanine at their first nucleotide position. An adenine was present in the A homologs of the two C_3 species. A more striking C_4-to-C_3 associated difference is found for the tetranucleotide CACT. This assemblage is present in the B parts of all C_4 and C_4-like species but lacking in both C_3 promoters. This suggested that the CACT tetranucleotide is critical for *Mem1* function.

To challenge this assumption, the CACT tetranucleotide was deleted in *Mem1* and the modified *Mem1*, (construct ppcA-PR_{Ft}-DRa/b_{Ft}-ΔCACT; Figure 9A) was tested for its expression profile in transgenic *F. bidentis*. Deletion of the CACT tetranucleotide resulted in a 50-fold drop in promoter activity. The resulting activity of the ppcA-PR_{Ft}-DRa/b_{Ft}-ΔCACT construct was statistically indistinguishable from that of ppcA-PR_{Ft} (Figure 9B). It follows that the CACT tetranucleotide is essential for the quantity of mesophyll expression. Histochemical analysis of the 10 transgenic plants obtained showed that nine plants did not reveal any GUS activity in the mesophyll cells (Figure 9C; plant C is shown here as an example because this plant shows the second highest GUS activity). Only one plant, with the highest promoter activity of all 10 plants (labeled D in Figure 9B), directed a mesophyll-specific expression of the GUS reporter gene. Whether the transgene of this plant has been integrated into the genome close to a mesophyll-specific enhancer and the resulting mesophyll expression of the GUS reporter gene represents an artifact is not known. At this stage of investigation, we have to conclude that the CACT-containing *cis*-regulatory element is necessary for mesophyll expression but may not be sufficient.

DISCUSSION

The C_4 cycle genes are largely regulated by transcription (Sheen, 1999). This type of gene regulation involves sets of *cis*-regulatory modules and their corresponding *trans*-regulatory factors that interact and thereby control the specific expression of the C_4 genes in either mesophyll or bundle-sheath cells. Plants with the C_4 photosynthetic pathway are of polyphyletic origin (Kellogg, 1999; Sage et al., 1999), and the networks that regulate cell type–specific gene expression are also likely to have evolved several times independently. To date, it is not known which types of *cis*- and *trans*-regulatory elements constitute mesophyll or bundle-sheath cell expression modules at the molecular level and how regulatory networks for mesophyll or bundle-sheath cell-specific gene expression have evolved. Here, a scenario is presented that indicates how the mesophyll cell-specific expression of the C_4 PEPC gene in the genus Flaveria may have evolved.

Analysis of chimerical promoter/reporter genes in transgenic *F. bidentis* identified two segments in the 5′ flanking region of the *ppcA1* gene of the C_4 plant *F. trinervia* that are necessary and sufficient for the mesophyll-specific expression of this gene. The DR exhibits enhancer-like properties and, combined with its corresponding PR, confers high levels of mesophyll expression to the reporter gene (Figure 3). The orthologous *ppcA1* promoter of the C_3 plant *F. pringlei* directs neither a high nor a mesophyll-specific expression (Stockhaus et al., 1997). One has to conclude therefore that both the distal and proximal promoter regions of the *ppcA1* gene have been sites of evolutionary actions. All available evidence in this and previous studies (Stockhaus et al., 1997; Windhövel et al., 2001) supports this point of view.

When the DR of the *ppcA1* promoter of *F. trinervia* (C_4 *ppcA1* promoter) is fused to the *ppcA1* promoter of *F. pringlei* (C_3 *ppcA1* promoter), a mesophyll expression component is added to that promoter, but the overall promoter strength does not alter substantially (Figure 4). This may be explained by assuming that the DR of the C_4 *ppcA1* promoter provides mesophyll specificity, while the PR is responsible for quantitative expression. The quantity elements are not present in the C_3 *ppcA1* promoter, and, therefore, the chimerical ppcA-M_{Fp}-DR_{Ft} promoter (Figure 4) does not direct high levels of mesophyll expression. Alternatively, the DR of the C_4 *ppcA1* promoter may contain transcription repressing sequences that reduce *ppcA1* expression in the bundle-sheath cells and the vascular bundle and that thereby relatively increase mesophyll expression. However, the activity

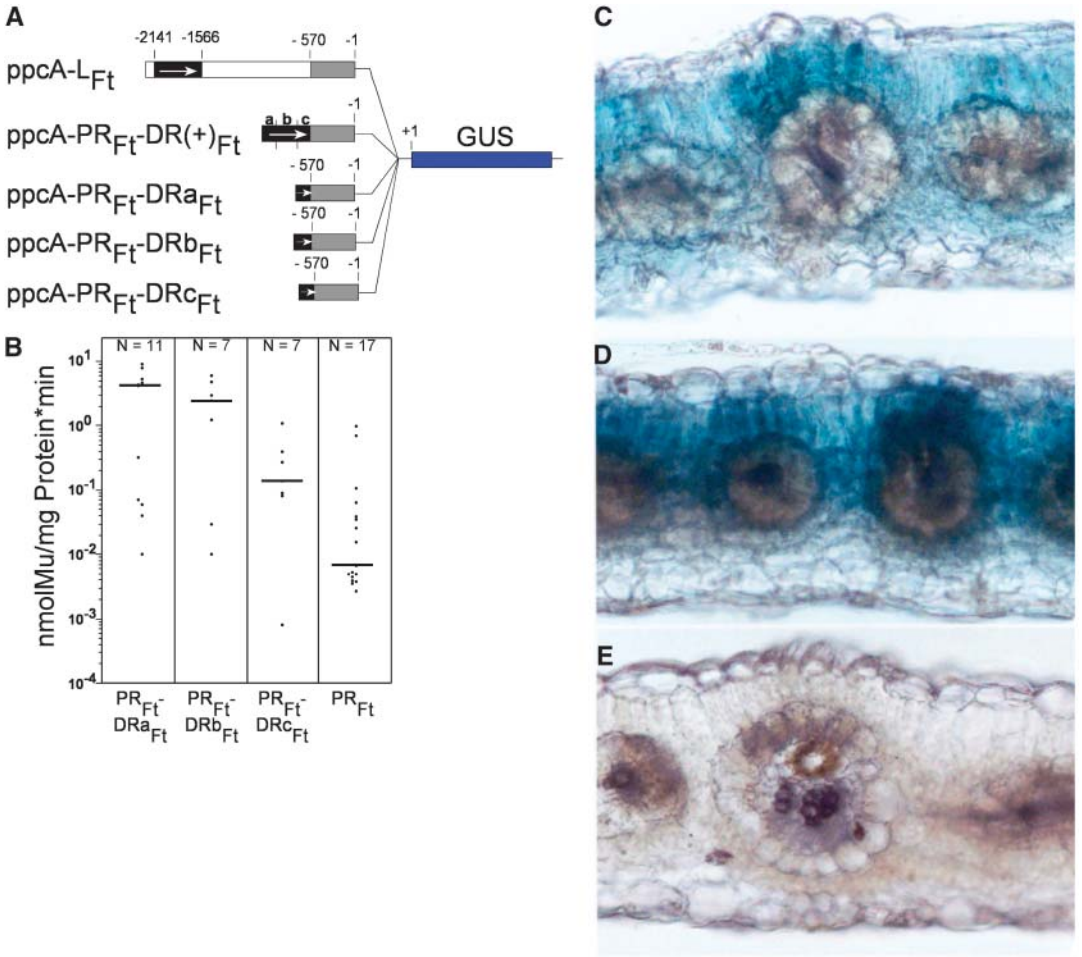


Figure 6. Analysis of the *ppcA1* GUS Reporter Gene Constructs ppcA-PR$_{Ft}$-DRa(+)$_{Ft}$, ppcA-PR$_{Ft}$-DRb(+)$_{Ft}$, and ppcA-PR$_{Ft}$-DRc(+)$_{Ft}$ in Transgenic *F. bidentis*.

(A) Structures of the *ppcA1/GUS* chimerical genes.

(B) GUS activities in leaves of transgenic *F. bidentis* plants. The numbers of independent transgenic plants investigated (N) are indicated at the top of each column. Median values are shown. Mu, 4-methylumbelliferone.

(C) to **(E)** Histochemical localization of GUS activity in leaf sections of transgenic *F. bidentis* plants transformed with the ppcA-PR$_{Ft}$-DRa(+)$_{Ft}$ **(C)**, ppcA-PR$_{Ft}$-DRb(+)$_{Ft}$ **(D)**, or ppcA-PR$_{Ft}$-DRc(+)$_{Ft}$ **(E)**. Incubation times were 6 h **(C)**, 18 h **(D)**, and 48 h **(E)**.

of all constructs containing the C_4 DR or its subfragments is clearly higher than the activity of the PR of the C_4 *ppcA1* promoter alone (Figures 3, 6, and 7). This demonstrates that the C_4 DR contains mesophyll transcriptional enhancer sequences. Whether there are, in addition, bundle-sheath repressing sequences remains an open question.

The proposed attributes of the DRs and PRs of the C_4 *ppcA1* promoter may not be easily identified by experiments. The PR of the C_4 *ppcA1* promoter (C_4-PR) alone shows only a very basic level of expression (Figure 3). This demonstrates that the DR of the C_4 promoter (C_4-DR) is absolutely essential for the C_4-typical high expression potential of the corresponding PR. On the other hand, the C_4-DR does not result in any mesophyll expression when it is fused to the −46 fragment of the 35S promoter of the *Cauliflower mosaic virus* (Burscheidt, 1998). This indicates that the C_4-DR exhibits its mesophyll expression potential only when it is combined with a PR from either the C_4 or the C_3 *ppcA1* promoter. To achieve a high mesophyll-specific expression, the C_4-DR has to be combined with its cognate PR. It has to be concluded, therefore, that the distal and proximal promoter regions do not function as separate modules and act additively but, rather, as a synergistic transcriptional controlling system that evolved together.

Are the C_4-DR and C_4-PR segments the only parts of the 5′ flanking region of the C_4 *ppcA1* gene that are involved in controlling the transcription of that gene? When the ppcA-PR$_{Ft}$-DR$_{Ft}$ construct and its derivatives are compared with that of the full C_4 *ppcA1* promoter (Figures 3 and 6), there is clearly a significant loss in expression quantity. This indicates that the smaller promoter constructs lack quantitative *cis*-regulatory

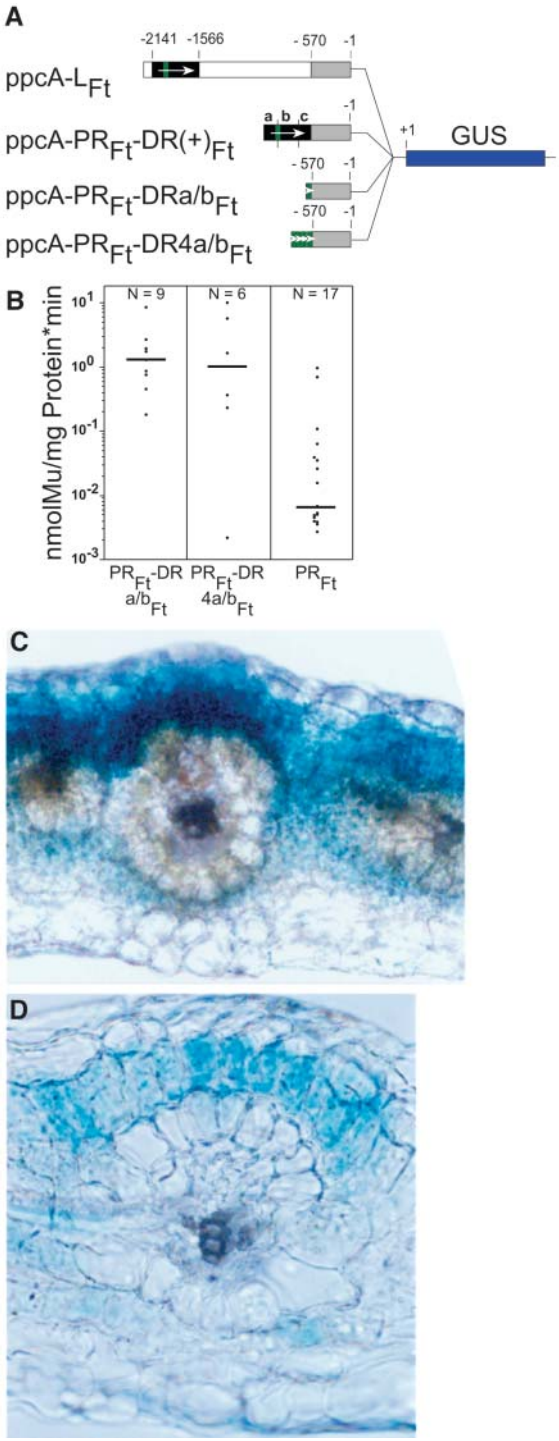


Figure 7. Analysis of the *ppcA1* GUS Reporter Gene Constructs ppcA-PR$_{Ft}$-DRa/b(+)$_{Ft}$ and ppcA-PR$_{Ft}$-DR4a/b(+)$_{Ft}$ in Transgenic *F. bidentis*.

(A) Structures of the *ppcA1/GUS* chimerical genes.

(B) GUS activities in leaves of transgenic *F. bidentis* plants. The numbers of independent transgenic plants investigated (N) are indicated at the top of each column. Median values are shown.

(C) and **(D)** Histochemical localization of GUS activity in leaf sections of transgenic *F. bidentis* plants transformed with ppcA-PR$_{Ft}$-DRa/b(+)$_{Ft}$ **(C)** and ppcA-PR$_{Ft}$-DR4a/b(+)$_{Ft}$ **(D)**. Incubation times were 12 h **(C)** and 6 h **(D)**.

elements that are present in the full promoter. Alternatively, the reduced expression levels of the ppcA-PR$_{Ft}$-DR$_{Ft}$ construct and its derivatives may also be attibutable to the changed distance between the DR and PR segments in these promoters and exhibit topological constraints (Rippe et al., 1995). The difference in expression levels between the ppcA-PR$_{Ft}$-DR(+)$_{Ft}$ and ppcA-PR$_{Ft}$-DR(−)$_{Ft}$ constructs (Figure 3) supports this possibility.

Whether the C_4 *ppcA1* promoter sequences between −570 and −1049 encompasses further C_4-relevant *cis*-regulatory elements remains unclear. The C_4 *ppcA1* promoter deletion experiments with the heterologous C_3 plant tobacco suggest that the intermediate region may contain repressing sequences (cf. the expression levels of the ppcA-1,5$_{Ft}$ and ppcA-1,0$_{Ft}$ constructs with the ppcA-PR$_{Ft}$ promoter; Figure 2). However, the ppcA-1,5$_{Ft}$ and ppcA-1,0$_{Ft}$ constructs have not been analyzed in the homologous C_4 system, and the biological meaning of the tobacco data therefore remains questionable. Even though we cannot exclude that the segment between the DR and PR region contains *cis*-regulatory elements, we conclude that they are most probably only of minor importance. The DR and PR segments are the major and essential *cis*-regulatory modules for the high and mesophyll-specific expression of the C_4 *ppcA1* gene.

To date, the *cis*-regulatory elements of the PR have not been mapped precisely. Using the yeast one-hybrid system, it was found that the PR of the C_4 *ppcA1* promoter interacts with homeobox transcription factors of the zinc finger subclass (Windhövel et al., 2001), whereas the PR of the C_3 *ppcA1* promoter does not contain detectable binding sites for these zinc finger homeobox proteins. At least one binding site is located in the first intron, which is inserted in the 5′ leader region of the C_4 *ppcA1* gene (Windhövel et al., 2001). The in planta significance of the zinc finger homeobox proteins and their exact target sequences need to be investigated.

A *cis*-regulatory module for mesophyll-specific gene expression named *Mem1* has been identified in the DR of the C_4 *ppcA1* promoter. The module is mapped at 41-bp resolution and overlaps with the a and b parts of the C_4-DR segment. Fusing the 41-bp segment to the PR of the C_4 *ppcA1* promoter is sufficient to confer mesophyll-specific expression to the GUS reporter gene. *Mem1*, therefore, has to carry *cis*-regulatory elements for mesophyll-specific gene expression. Whether there are other mesophyll expression elements in the a or b part of the C_4-DR segment is unknown. If these elements exist, they are probably redundant to *Mem1*. Whether *Mem1* harbors also a bundle-sheath repressing element remains an open question and should be investigated in the future.

A comparative analysis with *ppcA1* promoter sequences from other C_4, C_4-like, and C_3 Flaveria species identified *Mem1* homologous sequences in all examined plants. Their comparison hinted at elements for mesophyll-specific *ppcA1* gene expression. The most notable C_4-to-C_3 difference detected between *Mem1* of the C_4/C_4-like plants and its counterpart in the C_3 species is a CACT tetranucleotide (Figure 8). The motif is present

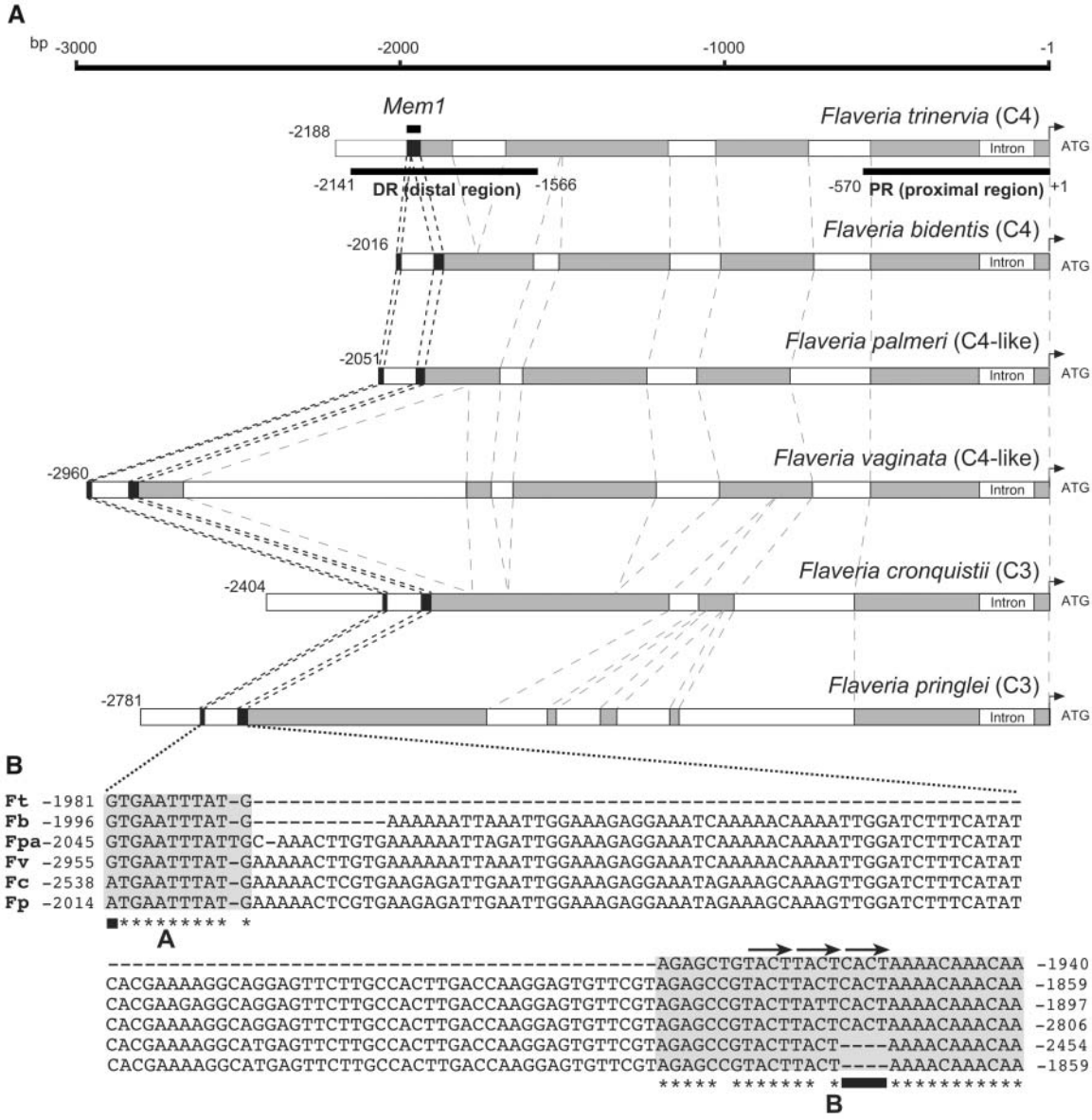


Figure 8. The Structures of *ppcA1* Promoters from C_4, C_4-like, and C_3 Flaveria Species and the Nucleotide Composition of *Mem1* and Its Homologs.

(A) Schematic comparisons of the 5′ flanking sequences of the *ppcA1* genes of the C_4 plants *F. trinervia* and *F. bidentis*, the the C_4-like species *F. palmeri* and *F. vaginata*, and the C_3 plants *F. pringlei* and *F. cronquistii*. The numbers of nucleotides refer to the translation initiation codons. Regions with high similarity between all promoters (60% or more identical nucleotides) are indicated by gray boxes. The position of *Mem1* is indicated by black boxes.

(B) Sequence comparison of *Mem1* and its homologs. The A and B segments are shaded. Asterisks label identical nucleotides in the A or B segment of all promoters. Black bars indicate the single nucleotide difference in A and the CACT tetranucleotide in B. The tandemly duplicated T/CACT repeats are labeled by arrows.

in all *Mem1* sequences of the C_4/C_4-like plants but lacking in the *Mem1* homologs of the C_3 species. The CACT tetranucleotide is found in a sequence segment, the B region of *Mem1*, which is fully conserved in the C_4 and C_3 *ppcA1* promoters. This finding suggested that the CACT motif is essential for mesophyll-specific gene expression. Functional analyses with transgenic plants confirmed this assumption. Deletion of the CACT tetranucleotide from *Mem1* abolished the mesophyll expression of the GUS reporter gene. We conclude that the addition of the CACT tetranucleotide to the C_3 promoter during C_3-to-C_4 evolution created a new *cis*-regulatory element that was necessary for confering mesophyll expression to the promoter.

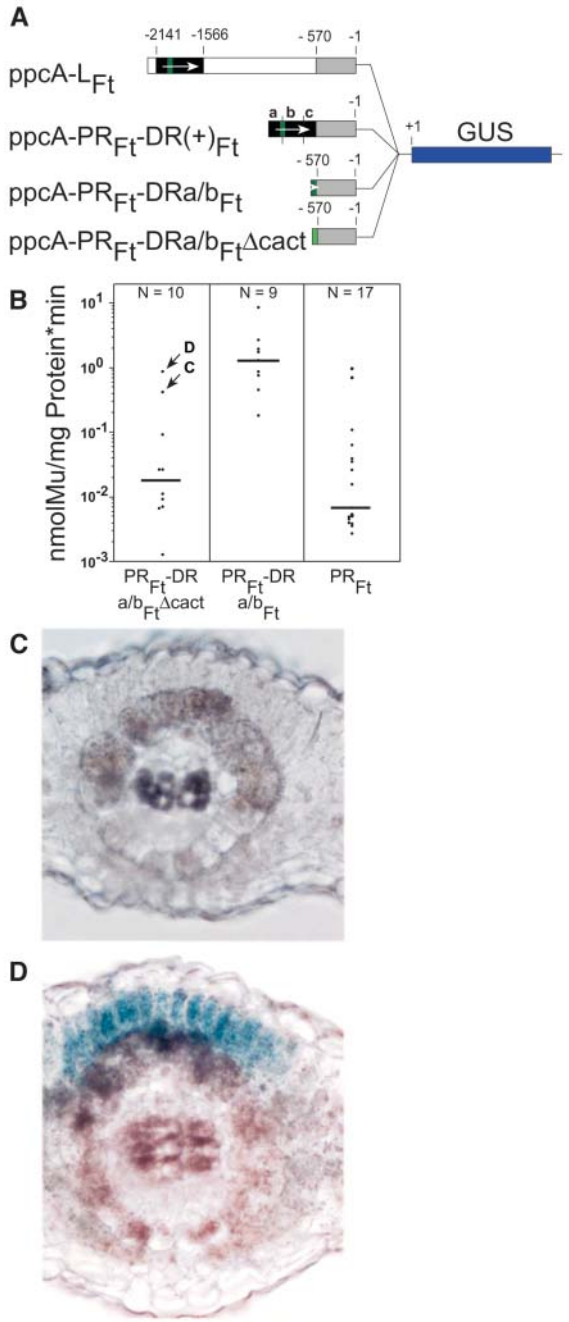


Figure 9. Analysis of the *ppcA1* GUS Reporter Gene Construct ppcA-PR_{Ft}-DRa/b_{Ft}-ΔCACT in Transgenic *F. bidentis*.

(A) Structure of the *ppcA1/GUS* chimerical gene.

(B) GUS activities in leaves of transgenic *F. bidentis* plants. The GUS activity of the ppcA-PR_{Ft}-DRa/b_{Ft} construct was taken from Figure 7B. The numbers of independent transgenic plants investigated (N) are indicated at the top of each column. Median values are shown. Mu, 4-methylumbelliferone.

(C) and **(D)** Histochemical analysis of the GUS activity of the ppcA-PR_{Ft}-DRa/b_{Ft}-ΔCACT driven reporter gene construct in leaf sections of transgenic *F. bidentis.* In **(C)**, the GUS expression pattern of plant C from **(B)** is shown, and in **(D)**, that of plant D is shown. Incubation time for **(C)** was 48 h; incubation time for **(D)** was 24 h.

Table 1. Oligonucleotide Primers Used for the Construction of Chimerical Promoters

FtDEa5′	5′-GGGAAGCTTAGAACATGAAAAAAGGACTC ACCAGG-3′
FtDEa3′	5′-GGGTCTAGATTGTTTGTTTTAGTGAGTAAG-3′
FtDEb5′	5′-GGGAAGCTTGTGAATTTATGAGAGCTGTAC-3′
FtDEb3′	5′-GGGTCTAGAGTACTTAATTTCCATTTCTC-3′
FtDEc5′	5′-GGGAAGCTTTGTGTGTGTGAATATGTTGC-3′
FtDEC3′	5′-GGGTCTAGATACATACTTTCCAAATTCATAG-3′
FtDEa3′-Xho	5′-GGGCTCGAGTTGTTTGTTTTAGTGAGTAAG-3′
FtDEb5′-Sal	5′-GGGGTCGACGTGAATTTATGAGAGCTGTAC-3′
FtDEa/bΔ5′	5′-AGCTTGTGAATTTATGAGAGCTGTACTTACTA AAACAAACAAT-3′
FtDEa/bΔ3′	5′-CTAGATTGTTTGTTTTAGTAAGTACAGCTCTCATAA ATTCACA-3′

Although the deletion of the CACT tetranucleotide reduced promoter activity almost completely, one single transgenic plant out of 10 (i.e., that with the highest activity) expressed the transgene in the mesophyll cells. We do not know whether the expression pattern of this transgene reflects an artifact because of the nearby presence of a mesophyll enhancer within the genome. Therefore, we have to conclude that the CACT-containing *cis*-regulatory element may not be the only *cis*-regulatory element in *Mem1*. The CACT-containing *cis*-element is necessary but may not be sufficient for mesophyll expression.

The CACT tetranucleotide is embedded in a sequence context (TTACTCACTAA) that can form an imperfect palindrome. The palindrome resembles a binding site for a GCN4-like basic leucine zipper transcription factor (Arndt and Fink, 1986; Oñate et al., 1999; Matys et al., 2003). A DNA protein interaction screen with the yeast one-hybrid system (Li and Herskowitz, 1993) supports this notion. Using *Mem1* as a bait, a basic leucine zipper protein was isolated that interacts with *Mem1* of *F. trinervia* but not with the *Mem1* homolog of *F. pringlei* (M. Akyildiz and P. Westhoff, unpublished data). Taken together, all available evidence suggests that the tetranucleotide CACT is part of a *cis*-regulatory element that may be targeted by a basic leucine zipper transcription factor.

How did this novel *cis*-regulatory element evolve? Adjacent to the CACT motif in the 5′ direction, two tandem TACT repeats are observed in all C_4/C_4-like *ppcA1* promoters but also in the C_3 promoters (Figure 8). Short direct repeats are known to be an important source of genetic change in all organisms because replication misalignment may lead to the deletion or addition of repeat units (Bzymek and Lovett, 2001; Li et al., 2002). We propose that such a mechanism was responsible for adding the CACT motif to the C_3 promoter. The addition of this third imperfect repeat unit resulted in the formation of a novel *cis*-regulatory element. This element could be targeted by transcription factors available already in the mesophyll cells, and thereby a new expression pattern was created.

A C_4-to-C_3 associated nucleotide difference was also observed in the A segment of *Mem1* (Figure 8). Whereas all C_4 and C_4-like species have a guanine at the outermost 5′ position of *Mem1*, it is an adenine in the C_3 species. No putative

Table 2. Oligonucleotide Primers Used for Genome Walking

GSP1	5′-CGAATCGATGTAATTTCTCCACATTCCGG-3′
GSP2	5′-TCATACTCAACAAGCTTATCATCCTCAGAA-3′
GSP3Fv	5′-TAAGTCARTCTATGACTCGCGCGTTGTG-3′
GSP4Fv	5′-CGCGTCGACGTAAAAACATTGAAGCCACAY-3′
GSP3Fc	5′-CACGCTTAGCTAAATGGGTAAGTGTAGAG-3′
GSP4Fc	5′-ATGATGTGTTCATGAGTTCATCTGGTTA-3′
GSP3Fpa	5′-CGTTGTGACGGGGCCATCAAATGGA-3′
GSP4Fpa	5′-ATGCGCACGTTGCCGCGTGTAAACTCGT-3′
GSP1Fp	5′-CGCCTCTATGTACAGAGAATACCTTTGTTC-3′
GSP2Fp	5′-GGCTCTACGAACACTCCTTGGTCAAG-3′

transcription factor binding site is detectable in this *Mem1* segment; thus, it is not clear whether this nucleotide difference is of functional importance.

The A segment is contiguous with the B segment only in *Mem1* of *F. trinervia*, but in all other C_4 and C_3 *ppcA1* promoters, the two segments are separated from each other by ~100 bp. This suggests that the contiguous arrangement of the A and B parts is not of functional importance. If the A segment contains a *cis*-regulatory element, then the A and B parts should form separate *cis*-regulatory units.

Which scenario can be envisaged for the evolution of the C_4 *ppcA* promoter in the genus Flaveria? The C_3 reference promoter from *F. pringlei* is weak and does not show any cell specificity. The activity of this promoter is even higher in the bundle-sheath cells and the vascular bundle than in the mesophyll cells. By contrast, the C_4 promoter is strong and is active only in the mesophyll cells. Evolution toward C_4 could therefore have started by increasing promoter strength. This is supported by *ppcA1* mRNA quantification in C_3-C_4 intermediate Flaveria species. Even C_3-C_4 intermediates with a low degree of C_4 trait expression (i.e., *F. chloraefolia*; Edwards and Ku, 1987) show already elevated *ppcA* mRNA amounts (Engelmann et al., 2003). It is reasonable to assume that the increase in *ppcA* expression was restricted to the mesophyll cells. This implies the evolution of mesophyll expression elements, for instance, by modifying rudimentary progenitor elements that were already present in the C_3 promoter (i.e., *Mem1*). The isolation and functional analysis of *ppcA* promoters from C_3-C_4 intermediate species of Flaveria has been initiated and should clarify this point. With a delay or maybe even in parallel the *cis*-regulatory modules for expression in bundle-sheath cells and the vascular bundle had to be inactivated. This could have been achieved by direct mutational modification of these modules and/or by the addition of bundle-sheath repressor elements. Which of these strategies nature has pursued is unknown.

In his review on biochemical evolution, A.C. Wilson (Wilson et al., 1977) pointed out that ''quantitative mutations affecting enzyme levels may have had a major role in the adaptative metabolic evolution of multicellular organisms'' and that ''these quantitative effects can result from point mutations in control genes.'' In the meantime, evolutionary biologists have collected convincing evidence that supports the view that changes in the spatiotemporal expression patterns of genes are the principal mechanism for the evolution of novelty, both in morphological and biochemical traits (Doebley and Lukens, 1998; Carroll, 2000).

Our investigations on the molecular evolution of C_4 PEPC in the genus Flaveria are in line with this concept. The studies show that at the onset of the transition from C_3 to C_4 photosynthesis, the enzyme is still rather C_3-like with respect to its kinetic and regulatory properties; it becomes C_4-like only much later (reviewed in Svensson et al., 2003). By contrast, the expression pattern of the *ppcA1* gene was modified very early in evolution from C_3 to C_4 (Engelmann et al., 2003). The data presented here indicate that comparatively small changes in the nucleotide sequence should be responsible for these changes that give rise to a novel mode of expression. It will be interesting to see whether the evolution of mesophyll- or bundle-sheath cell-specific gene expression in Flaveria always relied upon the same set of *cis*- and *trans*-regulatory elements. The analysis of another mesophyll specifically expressed gene, for instance, carbonic anhydrase (Badger, 2003), would therefore be highly desirable, and the study of bundle-sheath specific gene expression should be initiated. Because the C_4 photosynthetic pathway evolved several times independently (Sage, 2004), it will be even more interesting to investigate whether the various C_4 species pursued similar or different strategies to achieve the same goal, a differential expression of their C_4-photosynthesis associated genes.

METHODS

Construction of Chimerical Promoters

DNA manipulations and cloning were performed according to Sambrook and Russell (2001). All promoter GUS reporter gene fusions used in this work are based on the constructs ppcA-PR_{Ft}, ppcA-L_{Ft}, and ppcA-L_{Fp}, which were formerly designated ppcA-S-Ft, ppcA-L-Ft, and ppcA-L-Fp. Their construction has been described in detail (Stockhaus et al., 1994). In all constructs, the 3′ border of the *ppcA1* 5′ flanking sequences of *Flaveria trinervia* and *F. pringlei* (Figure 1) is located just upstream of the AUG initiation codon. For cloning purposes, a *Sma*I site was added to the 3′ border of each fragment by PCR amplification with an appropriately designed oligonucleotide. The 5′ borders of ppcA-L_{Ft} and ppcA-L_{Fp} are defined by *Hind*III (ppcA-L_{Ft}) or *Xba*I sites (ppcA-L_{Fp}), which occur naturally in these promoter regions (Figure 1). The 5′ border of the ppcA-PR_{Ft} promoter fragment of *F. trinervia* corresponds to nucleotide position −570. A *Xba*I site was added to this border by PCR amplification. All promoter fragments were assembled in pBluescribe M13− (Stratagene Cloning Systems, La Jolla, CA) and confirmed by sequencing. They were excised by *Hind*III/*Sma*I digestion and transferred to *Hind*III/*Sma*I-restricted pBI121 (Clontech Laboratories, Palo Alto, CA) in front of the GUS reporter gene (Stockhaus et al., 1994). The *ppcA1* promoter reporter gene constructs prepared in the course of this study were cloned as described below. All DNA fragments generated by PCR were confirmed by DNA sequencing.

Construction of ppcA-1,5$_{Ft}$ and ppcA1,0$_{Ft}$

The ppcA-L_{Ft} promoter plasmid (pBluescribe M13−) was digested with *Acc*I (−1566) or *Asp*718 (−1049) (Figure 1). Blunt ends were generated by fill-in synthesis with the Klenow fragment of *Escherichia coli* DNA polymerase I followed by ligation of *Hind*III linkers. The *Acc*I and

*Asp*718 restricted plasmids were digested with *Hind*III, the 5′ located ppcA-L_{Ft} promoter fragments were removed by agarose gel electrophoresis, and the *Hind*III ends of the remaining plasmids were religated. After an intermediate cloning step in *E. coli*, the resulting ppcA-$1{,}5_{Ft}$ and ppcA$1{,}0_{Ft}$ promoter regions were excised by *Hind*III/*Sma*I restriction and inserted into pBI121 in front of the GUS reporter gene.

Construction of ppcA-PR_{Ft}-DR(+)$_{Ft}$ and ppcA-PR_{Ft}-DR(−)$_{Ft}$

The ppcA-L_{Ft} promoter plasmid (pBluescribe M13−) was digested with *Acc*I (−1566). Blunt ends were generated by fill-in synthesis, and *Xba*I linkers were ligated. The DNA was restricted with *Xba*I (−2141), and the released 575-bp *Xba*I fragment (named DR) was isolated. The DR fragment was cloned into *Xba*I-digested ppcA-PR_{Ft} pBI121 and led to ppcA-PR_{Ft}-DR(+)$_{Ft}$ (DR inserted in correct orientation) or ppcA-PR_{Ft}-DR(−)$_{Ft}$ (DR in opposite orientation).

Construction of ppcA-M_{Fp}

The ppcA-L_{Fp} promoter plasmid (pBluescribe M13−) was digested with *Bcl*I (−1854) and *Xba*I (−2584) (Figure 1). The released 685-bp *Bcl*I/*Xba*I fragment was removed by agarose gel electrophoresis, and the remaining *ppcA1* promoter plasmid DNA was recovered. Blunt ends were generated by treatment with Klenow polymerase followed by religation of the promoter plasmid. The resulting ppcA-M_{Fp} promoter fragment was excised by *Hind*III/*Sma*I and inserted into pBI121.

Construction of ppcA-M_{Fp}-DR_{Ft}

The ppcA-L_{Fp} promoter plasmid (pBluescribe M13−) was digested with *Bcl*I (−1854), and blunt ends were generated by treatment with Klenow polymerase. The DNA was restricted with *Xba*I (−2584) (Figure 1). The released 685-bp *Bcl*I/*Xba*I fragment was removed by agarose gel electrophoresis, and the 5′ deleted *ppcA1* promoter plasmid DNA was recovered. The DR of the *ppcA1* promoter of *F. trinervia* was isolated by digesting ppcA-L_{Ft} with *Acc*I, creating blunts by treatment with Klenow polymerase, and releasing the DR by restriction with *Xba*I. The DR was ligated with the 5′ deleted *ppcA1* promoter of *F. pringlei*. The resulting ppcA-M_{Fp}-DR_{Ft} promoter was excised by *Hind*III/*Sma*I and inserted into pBI121.

Construction of ppcA-PR_{Ft}-DRa_{Ft}, ppcA-PR_{Ft}-DRb_{Ft}, and ppcA-PR_{Ft}-DRc_{Ft}

The DR of the *ppcA1* promoter of *F. trinervia* was divided into the three overlapping segments a (−2141 to −1940), b (−1981 to −1668), and c (−1713 to −1566). These were amplified by PCR with the ppcA-PR_{Ft}-DR(+)$_{Ft}$ promoter plasmid as template. Each 3′ oligonucleotide primer carried a *Xba*I, and each 5′ oligonucleotide primer carried a *Hind*III site (Table 1). After digestion with *Hind*III and *Xba*I, the resulting PCR products were used to replace the DR fragment in the ppcA-PR_{Ft}-DR(+)$_{Ft}$ construct.

Construction of ppcA-PR_{Ft}-DRa/b_{Ft} and ppcA-PR_{Ft}-DR4a/b_{Ft}

The a/b-overlapping region (−1981 to −1940) was amplified by PCR using the FtDEb5′ and FtDEa3′ primers. After digestion with *Hind*III and *Xba*I, the a/b-fragment was inserted into ppcA-PR_{Ft}-DR(+)$_{Ft}$ to replace the DR fragment. The resulting promoter was named ppcA-PR_{Ft}-DRa/b_{Ft}.

Tandem repeats of the a/b-overlapping region were generated as described by de Pater et al. (1993) using the primers FtDEa3′-Xho and FtDEb5′-Sal (Table 1), which contain *Xho*I and *Sal*I sites instead of *Xba*I and *Hind*III sites. The resulting multimeric DNAs were used as a template for PCR amplification with the FtDEb5′ and FtDEa3′ primers (Table 1). The fragment that contains four tandem repeats of the a/b-overlapping region was isolated by agarose gel electrophoresis, digested with *Hind*III and *Xba*I, and inserted into ppcA-PR_{Ft}-DR(+)$_{Ft}$ to replace the DR fragment. The resulting promoter was named ppcA-PR_{Ft}-DR4a/b_{Ft}.

*Construction of ppcA-PR_{Ft}-DRa/b_{Ft}-Δ*CACT*

The a/b-overlapping region (−1981 to −1940) without the CACT tetranucleotide was generated by annealing the two oligonucleotides FtDEa/bΔ5′ and FtDEa/b3′. The annealed oligonucleotides were inserted into ppcA-PR_{Ft}-DR(+)$_{Ft}$ to replace the DR fragment. The resulting promoter was named ppcA-PR_{Ft}-DRa/b_{Ft}-ΔCACT.

Plant Transformation

The promoter/GUS constructs were introduced by electroporation into the *Agrobacterium tumefaciens* strain AGL1 (Lazo et al., 1991). Tobacco (*Nicotiana tabacum*) plants were transformed as described (Horsch et al., 1985; Stockhaus et al., 1994). The transformation of *F. bidentis* plants was performed according to Chitty et al. (1994). Integration of the chimerical genes into the *F. bidentis* genome was examined by DNA gel blot analysis or by PCR (Stockhaus et al., 1997). In all tested transgenic plants, the hybridizing fragment or the PCR product had the expected size, indicating that the promoter fragment and the GUS gene were linked in the genomic DNA and that each transgenic plant contains at least one copy of the respective chimerical gene.

Measurement of GUS Activity and Histochemical Analysis

Regenerated plants or T1 plants were used for the analysis of the GUS activity. Tobacco plants grown from tissue culture were used for the measurements of GUS activity. For the histochemical analysis, tobacco plants were transferred to soil and grown in a greenhouse. *F. bidentis* were greenhouse plants, 40 to 50 cm tall and before flower initiation. GUS activities were measured quantitatively (Jefferson et al., 1987; Kosugi et al., 1990) or in situ (Stockhaus et al., 1997). The average values of the data are expressed by medians, and the Mann-Whitney U test statistics as implemented in the software package Kaleidagraph 3.6 for Mac OS X (Synergy Software, Reading, PA, www.synergy.com) were used to test whether two data series differ from each other.

DNA Isolation

Nucleic acids were isolated from leaf tissue (Westhoff et al., 1991). DNA was recovered from the 2 M LiCl soluble nucleic acid fraction by isopropanol precipitation. The DNA was dissolved in double-distilled water, and residual RNA was digested by RNase A treatment. After phenol/chloroform extraction, the DNA was precipitated with isopropanol, dissolved in double-distilled water, and stored at −20°C until use.

Isolation of 5′ Flanking Sequences from the *ppcA1* Genes of *F. bidentis*, *F. vaginata*, *F. palmeri*, and *F. cronquistii*

The 5′ flanking regions of *ppcA1* genes of *F. bidentis*, *F. vaginata*, *F. palmeri*, and *F. cronquistii* were isolated from total DNA by vectorette PCR (Siebert et al., 1995) with the Universal Genome Walker Kit (Clontech Laboratories) as recommended by the manufacturer. For each plant species, *Dra*I, *Eco*RV, *Pvu*II, and *Stu*I DNA libraries were constructed. The gene-specific primers for the primary and secondary PCR reactions (GSP1 and GSP2; Table 2) of the first walking step were designed to

hybridize to the very 5′ part of the coding region of the *ppcA1* genes of both *F. trinervia* and *F. pringlei*. The primers were expected to hybridize to the 5′ coding regions of the other *ppcA1* genes as well. In the second walking step, gene-specific primers were designed according to the sequence of the promoter fragments isolated in the first walking step (Table 2). The resulting PCR fragments were cloned into the pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA) and sequenced.

Sequence data from this article have been deposited with the EMBL/GenBank data libraries under the accession numbers X64143 (*F. trinervia*), AY297087 (*F. bidentis*), AY297088 (*F. palmeri*), AY297090 (*F. vaginata*), AY297089 (*F. cronquistii*), and X64144 (*F. pringlei*).

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft initially within the Graduiertenkolleg Molekulare Physiologie and later by SFB 590 Inhärente und adaptive Differenzierungsprozesse at the Heinrich Heine University of Düsseldorf. Additional support from the Fonds der Chemischen Industrie is gratefully acknowledged. We are indebted to Uwe Santore for carefully reading the manuscript.

Received December 3, 2003; accepted February 28, 2004.

REFERENCES

Arndt, K., and Fink, G.R. (1986). GCN4 protein, a positive transcription factor in yeast, binds general promoters at all 5′ TGACTC 3′ sequences. Proc. Natl. Acad. Sci. USA **83,** 8516–8520.

Badger, M. (2003). The roles of carbonic anhydrases in photosynthetic CO2 concentrating mechanisms. Photosyn. Res. **77,** 83–94.

Blackwood, E.M., and Kadonaga, J.T. (1998). Going the distance: A current view of enhancer action. Science **281,** 60–63.

Bläsing, O.E., Westhoff, P., and Svensson, P. (2000). Evolution of C4 phosphoenolpyruvate carboxylase in *Flaveria*, a conserved serine residue in the carboxyl-terminal part of the enzyme is a major determinant for C4-specific characteristics. J. Biol. Chem. **275,** 27917–27923.

Burscheidt, J. (1998). Cis-Regulatorische Determinanten für Mesophyll- und Bündelscheidenspezifische Genexpression in C4-Spezies der Gattung *Flaveria*—Die Promotoren der Phosphoenolpyruvat-Carboxylase- und der Glycin-Decarboxylasegene. Master's thesis (Düsseldorf, Germany: Heinrich-Heine-Universität).

Bzymek, M., and Lovett, S.T. (2001). Instability of repetitive DNA sequences: The role of replication in multiple mechanisms. Proc. Natl. Acad. Sci. USA **98,** 8319–8325.

Carroll, S.B. (2000). Endless forms: The evolution of gene regulation and morphological diversity. Cell **101,** 577–580.

Chitty, J.A., Furbank, R.T., Marshall, J.S., Chen, Z., and Taylor, W.C. (1994). Genetic transformation of the C_4 plant, *Flaveria bidentis*. Plant J. **6,** 949–956.

Crétin, C., Santi, S., Keryer, E., Lepiniec, L., Tagu, D., Vidal, J., and Gadal, P. (1991). The phosphoenolpyruvate carboxylase gene family of *Sorghum*: Promoter structures, amino acid sequences and expression of genes. Gene **99,** 87–94.

de Pater, S., Pham, K., Chua, N.H., Memelink, J., and Kijne, J. (1993). A 22-bp fragment of the pea lectin promoter containing essential TGAC-like motifs confers seed-specific gene expression. Plant Cell **5,** 877–886.

Doebley, J., and Lukens, L. (1998). Transcriptional regulators and the evolution of plant form. Plant Cell **10,** 1075–1082.

Edwards, G.E., and Ku, M.S.B. (1987). Biochemistry of C_3-C_4 intermediates. In The Biochemistry of Plants, Vol. 10, M.D. Hatch and N.K. Boardman, eds (New York: Academic Press), pp. 275–325.

Engelmann, S., Bläsing, O.E., Gowik, U., Svensson, P., and Westhoff, P. (2003). Molecular evolution of C4 phospho*enol*-pyruvate carboxylase in the genus *Flaveria*—A gradual increase from C3 to C4 characteristics. Planta **217,** 717–725.

Ernst, K., and Westhoff, P. (1996). The phosphoenolpyruvate carboxylase (*ppc*) gene family of *Flaveria trinervia* (C_4) and *F. pringlei* (C_3): Molecular characterization and expression analysis of the *ppcB* and *ppcC* genes. Plant Mol. Biol. **34,** 427–443.

Hatch, M.D. (1987). C_4 photosynthesis: A unique blend of modified biochemistry, anatomy and ultrastructure. Biochim. Biophys. Acta **895,** 81–106.

Hermans, J., and Westhoff, P. (1990). Analysis of expression and evolutionary relationships of phosphoenolpyruvate carboxylase genes in *Flaveria trinervia* (C_4) and *F. pringlei* (C_3). Mol. Gen. Genet. **224,** 459–468.

Hermans, J., and Westhoff, P. (1992). Homologous genes for the C_4 isoform of phosphoenolpyruvate carboxylase in a C_3- and a C_4-*Flaveria* species. Mol. Gen. Genet. **234,** 275–284.

Höfer, M.U., Santore, U.J., and Westhoff, P. (1992). Differential accumulation of the 10-, 16- and 23-kDa peripheral components of the water-splitting complex of photosystem II in mesophyll and bundle-sheath chloroplasts of the dicotyledonous C_4 plant *Flaveria trinervia* (Spreng.) C. Mohr. Planta **186,** 304–312.

Horsch, R.B., Fry, F.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G., and Fraley, R.T. (1985). A simple and general method for transferring genes into plants. Science **227,** 1229–1231.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusions: β-Glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6,** 3901–3907.

Kellogg, E.A. (1999). Phylogenetic aspects of the evolution of C4 photosynthesis. In C4 Plant Biology, R.F. Sage and R.K. Monson, eds (San Diego, CA: Academic Press), pp. 411–444.

Kosugi, S., Ohashi, Y., Nakajima, K., and Arai, Y. (1990). An improved assay for β-glucuronidase in transformed cells: Methanol almost completely suppresses a putative endogenous β-glucuronidase activity. Plant Sci. **70,** 133–140.

Lazo, G.R., Stein, P.A., and Ludwig, R.A. (1991). A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. Biotechnology **9,** 963–967.

Li, J.J., and Herskowitz, I. (1993). Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. Science **262,** 1870–1874.

Li, Y.-C., Korol, A.B., Fahima, T., Beiles, A., and Nevo, E. (2002). Microsatellites: Genomic distribution, putative functions and mutational mechanisms: A review. Mol. Ecol. **11,** 2453–2465.

Matys, V., et al. (2003). TRANSFAC: Transcriptional regulation, from patterns to profiles. Nucleic Acids Res. **31,** 374–378.

Monson, R.K. (1999). The origins of C4 genes and evolutionary pattern in the C4 metabolic phenotype. In C4 Plant Biology, R.F. Sage and R.K. Monson, eds (San Diego, CA: Academic Press), pp. 377–410.

Monson, R.K., and Moore, B.D. (1989). On the significance of C_3-C_4 intermediate photosynthesis to the evolution of C_4 photosynthesis. Plant Cell Environ. **12,** 689–699.

Oñate, L., Vicente-Carbajosa, J., Lara, P., Díaz, I., and Carbonero, P. (1999). Barley BLZ2, a seed-specific bZIP protein that interacts with BLZ1 in vivo and activates transcription from the GCN4-like motif of B-hordein promoters in barley endosperm. J. Biol. Chem. **274,** 9175–9182.

Powell, A.M. (1978). Systematics of *Flaveria* (Flaveriinae-Asteraceae). Ann. Mo. Bot. Gard. **65,** 590–636.

Rippe, K., von Hippel, P.H., and Langowski, J. (1995). Action at a distance: DNA-looping and initiation of transcription. Trends Biochem. Sci. **20,** 500–506.

Sage, R.F. (2004). The evolution of C4 photosynthesis. New Phytol. **161,** 341–370.

Sage, R.F., Li, M., and Monson, R.K. (1999). The taxonomic distribution of C4 photosynthesis. In C4 Plant Biology, R.F. Sage and R.K. Monson, eds (San Diego, CA: Academic Press), pp. 551–584.

Sambrook, J., and Russell, D.W. (2001). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Sheen, J. (1999). C4 gene expression. Annu. Rev. Plant Physiol. Plant Mol. Biol. **50,** 187–217.

Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A., and Lukyanov, S.A. (1995). An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res. **23,** 1087–1088.

Stockhaus, J., Poetsch, W., Steinmüller, K., and Westhoff, P. (1994). Evolution of the C_4 phosphoenolpyruvate carboxylase promoter of the C_4 dicot *Flaveria trinervia*: An expression analysis in the C_3 plant tobacco. Mol. Gen. Genet. **245,** 286–293.

Stockhaus, J., Schlue, U., Koczor, M., Chitty, J.A., Taylor, W.C., and Westhoff, P. (1997). The promoter of the gene encoding the C_4 form of phosphoenolpyruvate carboxylase directs mesophyll specific expression in transgenic C_4 *Flaveria* spp. Plant Cell **9,** 479–489.

Svensson, P., Bläsing, O., and Westhoff, P. (1997). Evolution of the enzymatic characteristics of C_4 phosphoenolpyruvate carboxylase: A comparison of the orthologous ppcA phosphoenolpyruvate carboxylases of *Flaveria trinervia* (C_4) and *F. pringlei* (C_3). Eur. J. Biochem. **246,** 452–460.

Svensson, P., Bläsing, O.E., and Westhoff, P. (2003). Evolution of C4 phosphoenolpyruvate carboxylase. Arch. Biochem. Biophys. **414,** 180–188.

Westhoff, P., and Gowik, U. (2004). Evolution of C4 phosphoenolpyruvate carboxylase. Genes and proteins: A case study with the genus *Flaveria*. Ann. Bot. **93,** 1–11.

Westhoff, P., Offermann-Steinhard, K., Höfer, M., Eskins, K., Oswald, A., and Streubel, M. (1991). Differential accumulation of plastid transcripts encoding photosystem II components in the mesophyll and bundle-sheath cells of monocotyledonous NADP-malic enzyme-type C_4 plants. Planta **184,** 377–388.

Wilson, A.C., Carlson, S.S., and White, T.J. (1977). Biochemical evolution. Annu. Rev. Biochem. **46,** 573–639.

Windhövel, A., Hein, I., Dabrowa, R., and Stockhaus, J. (2001). Characterization of a novel class of plant homeodomain proteins that bind to the C_4 phosphoenolpyruvate carboxylase gene of *Flaveria trinervia*. Plant Mol. Biol. **45,** 201–214.

Evolution and Function of a *Cis*-Regulatory Module for Mesophyll-Specific Gene Expression in the C_4 Dicot *Flaveria trinervia*

Meryem Akyildiz, Udo Gowik, Maria Koczor, Monika Streubel and Peter Westhoff*

Heinrich-Heine Universität, Institut für Entwicklungs- und Molekularbiologie der Pflanzen, 40225 Düsseldorf, Germany

Corresponding author: Prof. Dr. Peter Westhoff
Institut für Entwicklungs- und Molekularbiologie der Pflanzen
Heinrich-Heine Universität
Universitätsstrasse 1
40225 Düsseldorf, Germany
E-mail: west@uni-duesseldorf.de
Tel. +49 211 81 12338
Fax. +49 211 81 14871

Key words: C_4 photosynthesis | phospho*enol*pyruvate carboxylase | *Flaveria* | *cis*-regulatory element

Pages: 21
Tables: 1
Figures: 5
Supplemental Figures: 1

Abstract

C_4 photosynthesis presents a sophisticated integration of two complementary cell-types, mesophyll and bundle-sheath cells. It relies on the differential expression of the genes encoding the component enzymes and transporters of this pathway. The entry enzyme of C_4 photosynthesis, phospho*enol*pyruvate carboxylase (PEPC) is exclusively found in mesophyll cells and the expression of the correspondent gene is regulated at the transcriptional level. In the C_4 dicot *Flaveria trinervia*, the mesophyll-specific expression of the C_4 phospho*enol*pyruvate carboxylase gene (*ppcA*) depends on a 41 bp segment in the distal promoter region referred to as MEM1 (mesophyll expression module 1). Here we show that the orthologous MEM1 sequences from the *ppcA* gene from the C_3 species *F. pringlei* direct a completely different expression, reporter gene activity was found in all chlorenchyma cells and the vascular bundles of a C_4 leaf. The orthologous MEM1 sequences of *F. pringlei* differ from MEM1 of *F. trinervia* at two positions, a guanine to adenine exchange and the insertion of the tetranucleotide CACT. Changes at these two positions in the C_3 MEM1-sequence were necessary and sufficient to create a mesophyll specificity element during C_4 evolution. The MEM1 of *F. trinervia* enhances mesophyll expression and concomitantly represses expression in bundle-sheath-cells and vascular bundles.

Introduction

The photosynthetic C_4 cycle is a sophisticated add-on to the C_3 photosynthetic pathway. It is characterized by an initial CO_2 fixation step in the mesophyll cells by the oxygen-insensitive phospho*enol*pyruvate carboxylase (PEPC), resulting in the C_4 acids malate and/or aspartate. These C_4 acids are subsequently transported to neighbouring bundle-sheath cells where they are decarboxylated. The released CO_2 is refixed by ribulose 1,5-bisphosphate carboxylase/oxygenase. Due to this pre-fixation of CO_2 in the mesophyll cells, the photosynthetic C_4 cycle acts as a pump that delivers saturating concentrations of CO_2 at the site of ribulose 1,5-bisphosphate carboxylase/oxygenase in the bundle-sheath cells. As a consequence of this CO_2 concentrating mechanism photorespiration in C_4 plants is minimized and the net photosynthesis rate is increased (Hatch, 1987).

This division of labour between mesophyll and bundle-sheath cells depends on differential gene expression (Nelson and Dengler, 1992). In NADP-malic enzyme type C_4 species, PEPC, NADP-malate dehydrogenase and pyruvate orthophosphate dikinase are specifically expressed in mesophyll cells, whereas the decarboxylating enzymes, for instance NADP-dependent malic enzyme and the secondary carboxylase ribulose 1,5-bisphosphate carboxylase/oxygenase, are expressed exclusively in bundle-sheath cells. The cell-specific expression of these genes is predominantly regulated by transcription (Schäffner and Sheen, 1992), however, posttranscriptional control has been reported too (Patel et al., 2006).

The C_4 pathway evolved independently more than 45 times during the evolution of angiosperms (Sage, 2004). The genes encoding the C_4 isoforms of the C_4 cycle enzymes originated from non-photosynthetic progenitor genes that were already present in C_3 ancestral species. To meet the special requirements of the C_4 photosynthetic pathway, the expression programme of the C_3 progenitor genes had to be changed to a high and selective expression in the mesophyll or bundle-sheath cells of the leaf, and the enzymes themselves had to be adapted to the metabolic and regulatory context of the C_4 cycle (Bauwe and Chollet, 1986).

To gain insight into the evolution of C_4 genes, we are using the entry enzyme of the C_4 cycle, the PEPC, as the model C_4 enzyme/gene (Westhoff and Gowik, 2004), and the dicot genus *Flaveria* (Asteraceae) (Powell, 1978) as the experimental system. This genus contains closely related C_3, C_4 and numerous C_3-C_4 intermediate species (Ku et al., 1991). These C_3-C_4 species differ quantitatively in the expression of C_4 photosynthetic traits, and are considered, at least partly, as evolutionary intermediates (Monson and Moore, 1989).

The photosynthetic PEPCs of C_4 *Flaveria* species are encoded by the *ppcA* gene class whose orthologues are also found in C_3 and C_3-C_4 intermediate species of this genus

(Hermans and Westhoff, 1992). Analysis of *ppcA* promoter/ß-glucuronidase (GUS) reporter gene fusions in the C_4 plant *F. bidentis* revealed that the *ppcA* promoter of the C_4 species *F. trinervia* directs a high expression of the reporter gene in the mesophyll cells. The orthologous *ppcA* promoter of the C_3 plant *F. pringlei*, however, is weak and does not show any apparent cell or organ specificity (Stockhaus et al., 1997).

Detailed promoter reporter gene studies of the C_4 *ppcA* gene in transgenic *F. bidentis* plants indicated that the proximal (-1 to -570) and the distal region (-1566 to -2141) (Gowik et al., 2004) are sufficient for a high mesophyll-specific expression of the GUS reporter gene. While the proximal promoter region mediates a very low basal promoter activity, the distal region confers a mesophyll expression component when fused to the *ppcA* promoter of *F. pringlei*. By dissection of the C_4 distal region a 41 bp module, named MEM1 (mesophyll expression module 1), was identified that together with the C_4 proximal region is sufficient for mesophyll-specific reporter gene expression (Gowik et al., 2004).

MEM1 could be subdivided into two submodules, A and B, of 11 and 30 bp, respectively. The comparison of MEM1 sequences from C_4 *Flaveria* species and of MEM1-homologues from C_3 species of this genus revealed, that the A-submodules of the C_4 and C_4-like species have a guanine at their first nucleotide position, while an adenine is present in the A-submodules of the C_3 plants (Gowik et al., 2004). An additional difference is related to the tetranucleotide CACT that is present in the B-submodules of the C_4 and C_4-like plants but is absent from the B-submodules of the C_3 plants (Gowik et al., 2004).

These data suggested that both nucleotide polymorphisms are involved in determining the mesophyll-specific expression of the C_4 *ppcA* gene (Gowik et al., 2004). Hence the present investigation was initiated to identify at the nucleotide level the determinants for the mesophyll-specific expression of the C_4 *ppcA* PEPC gene. By the analysis of MEM1 deletion- and substitution reporter gene constructs it was found that both submodules of the MEM1 have to be present in the C_4-specific state, i.e. the guanine in the A-submodule and the tetranucleotide CACT in the B-submodule of MEM1, in order to provide a mesophyll-specific expression of the reporter gene. The C_4 MEM1 behaves as an enhancer of mesophyll expression and, in addition, as a repressor of *ppcA* gene expression in the bundle-sheath cells and the vascular bundles.

Results

MEM1 Displays a Transcriptional Enhancing and Repressor Activity

The proximal promoter region of the *ppcA* gene of *F. trinervia* (C_4) (construct FtPR, Fig. 1A) revealed only a very weak activity in previous experiments preventing an unequivocal *in situ* analysis (Gowik et al., 2004). To elucidate the expression-specificity of this basal promoter, this experiment was repeated and *F. bidentis* was re-transformed with this construct and the histochemical activity of the GUS reporter gene was analyzed in the 16 obtained transgenic plants. The majority of the plants showed no GUS staining. In all five stainable plants the GUS reporter gene was found to be expressed in mesophyll and bundle-sheath cells, but also in the vascular tissue (Fig. 1C). Thus the basal *ppcA* promoter directs no cell specificity.

The fusion of the MEM1 module of the C_4 *ppcA* promoter of *F. trinervia* to the proximal promoter region (construct FtPR-FtM, Fig 1A) results in an statistically significant eight fold elevated promoter activity when compared to the activity of the proximal region alone (Fig. 1B), indicating that MEM1 of *F. trinervia* contains a transcriptional enhancing activity. *In situ* analysis revealed a clear mesophyll-specific expression, while no GUS activity could be detected in the bundle-sheath cells and the vascular bundles (Fig. 1C). Taken together the MEM1 of *F. trinervia* enhances not only expression in mesophyll cells but concomitantly represses expression in bundle-sheath cells and vascular tissues. Thus the MEM1 of *F. trinervia* adds a mesophyll expression component to the proximal promoter.

Insertions between the A- and B-Submodules of MEM1 do not Affect Mesophyll Specificity

The MEM1 of *F. trinervia* is unique in that the A- and B-submodules are fused together with no intermediate sequence. In contrast, the A- and B-submodules of MEM1 of the C_4, C_4-like and C_3 species of *Flaveria* are separated by about 90 to 100 bp of intervening sequences (Gowik et al., 2004). To investigate the effect of these spacer sequences on mesophyll expression specificity, the MEM1 of the C_4 plant *F. bidentis* was fused to the proximal region of the *ppcA* promoter of *F. trinervia*, and the resulting promoter reporter gene construct FtPR-FbM (Fig. 1A) was transformed into *F. bidentis*. Histochemical analysis of transgenic plants, carrying the chimerical gene, showed that FtPR-FbM directs a GUS expression in mesophyll cells (Fig. 1E) not distinguishable from that of the FtPR-FtM promoter. The expression strength of FtPR-FbM is about four times higher than that of the proximal promoter

reinforcing that a C_4 MEM1 adds a mesophyll-specific enhancing activity to the proximal promoter. We conclude from these experiments that the spacer between the A- and B-submodules in MEM1 of *F. bidentis* does not contain any *cis*-regulatory element of relevance for the mesophyll specificity of gene expression.

The C_3-type MEM1 of *F. pringlei* does not Confer Mesophyll-Specific Gene Expression

The comparison of MEM1 sequences from *Flaveria* species differing in the mode of photosynthesis and preliminary experiments described in Gowik *et al*. (2004) suggested that the single G/A nucleotide polymorphism in the A-submodule and the insertion/deletion of the CACT tetranucleotide in the B-submodule might be key determinants for mesophyll specificity. If so, a C_3-type MEM1 element when fused to the proximal promoter region of the C_4 *ppcA* promoter should not direct a mesophyll-specific gene expression.

In order to examine the expression specificity of a C_3-type MEM1 we used the orthologous *ppcA* promoter of the C_3 species *F. pringlei*. The previously examined *ppcA* promoter construct of *F. pringlei* (Stockhaus et al., 1997) did not contain the complete MEM1. Hence it was necessary to isolate a *bona fide* complete *ppcA* promoter of *F. pringlei*, i.e. containing the entire MEM1 element, by vectorette PCR. The obtained fragment of 2538 bp was fused to the GUS reporter gene (construct ppcAFp; Fig. 2A), and the expression profile and strength of this promoter were analyzed in transgenic *F. bidentis*. Histochemical analysis of ppcAFp revealed an expression in mesophyll cells and bundle-sheath cells, and in the vascular bundle (Fig. 2C). It follows that this promoter does not show any apparent cell-specific expression. Its expression profile as well as its strength (Fig. 2B) are indistinguishable from the previously analyzed truncated *ppcA* promoter of *F. pringlei* (Stockhaus et al., 1997).

To assess the expression specificity of the C_3-type MEM1 of *F. pringlei* in the context of the proximal region of the C_4 *ppcA* promoter of *F. trinervia*, the corresponding construct FtPR-FpM (Fig. 2A) was prepared and introduced into *F. bidentis*. For comparison, the C_3-type MEM1 was also fused to its native proximal region (FpPR-FpM; Fig. 2A). Both promoter constructs gave rise to the same expression pattern as ppcAFp, i.e. GUS expression was detected in the mesophyll and bundle-sheath cells, and in the vascular tissue (Fig. 2D and E). This suggests that the C_3-C_4 associated sequence polymorphisms in MEM1 are necessary for the mesophyll specificity of gene expression, and that the proximal *ppcA* promoter segment does not interfere with the pattern of expression.

The FpPR-FpM promoter exhibits the same pattern of reporter gene expression as the *bona fide* complete *ppcA* promoter of *F. pringlei* (Fig. 2C and D). However, the FpPR-FpM promoter is about 26 times lower in expression strength than the complete promoter (Fig. 2B). This indicates that the nucleotide sequences between MEM1 and the proximal part of the *ppcA* promoter of *F. pringlei* contain elements that enhance the overall expression of this C_3-type *ppcA* promoter. This is similar to what has been observed when the expression strength of the corresponding constructs of the C_4 *ppcA* promoter were analyzed (Gowik et al., 2004).

Conversion of a C_3 to a C_4 MEM1 and *vice versa* Reveals that two C_3-C_4 Associated Nucleotide Sequence Polymorphisms are Sufficient for Mesophyll Expression

In order to clarify whether the two C_3-to-C_4 associated differences in MEM1 (Gowik et al., 2004) are the only determinants of mesophyll specificity in a C_4-type MEM1, the A- and B-submodules of the C_4 MEM1 of *F. trinervia* were changed into C_3-type MEM1-submodules (FtPR-FtM/A_ΔCACT, Fig. 2A), and *vice versa* the A- and B-submodules of the C_3 MEM1 of *F. pringlei* were changed into C_4-type MEM1-submodules (FtPR-FpM/G_+CACT, Fig. 2A). The C_3-type FtPR-FtM/A_ΔCACT promoter construct revealed the same expression pattern as the C_3-type promoter constructs FpPR-FpM and FtPR-FpM, i.e. the GUS-staining is found in mesophyll and bundle-sheath cells and in the vascular bundles (Fig. 2D to F). The C_4-equivalent FtPR-FpM/G_+CACT promoter construct, however, directed a mesophyll-specific expression behaviour (Fig. 2G).

It follows that the change of the A- and B-submodules of MEM1 from a C_4 into a C_3 state results in the loss of expression specificity. On the other side the change of MEM1 from a C_3 into a C_4 state leads to the acquisition of mesophyll specificity by repressing gene activity in the bundle-sheath cells and the vascular bundles.

The A- and B-Submodules of MEM1 are both Required for Mesophyll-Specific Expression

The two C_4 associated polymorphisms correlate with mesophyll specificity. In order to find out if both C_3-to-C_4 correlated differences are necessary for a mesophyll-specific gene expression, we followed two strategies. First, we divided the *F. trinervia* MEM1 into the 11 bp A-submodule and into the 30 bp B-submodule, and fused each submodule with the C_4 proximal promoter region (constructs FtPR-FtM/Δnt12-41 and FtPR-FtM/Δnt1-11, Fig. 3A).

Secondly, we converted independently the A- and B-submodules of the C_4 MEM1 into C_3-type submodules, and combined each chimeric MEM1 module with the proximal part of the C_4 *ppcA* promoter (constructs FtPR-FtM/A, FtPR-FtM/ΔCACT, Fig. 3A; cf. also Gowik et al., 2004).

Deletion of one submodule (constructs FtPR-FtM/Δnt12-41 and FtPR-FtM/Δnt1-11) caused a loss of mesophyll-specificity (Fig. 3C and D), i.e. the great majority of the GUS-stainable plants expressed the reporter gene in mesophyll and bundle-sheath cells as well in the vascular tissue. The same expression pattern, i.e. the loss of mesophyll specificity, was also observed when one submodule was in the C_3 state, while the other remained in the C_4 state (Fig. 3E and F). We conclude that both C_3-to-C_4 associated nucleotide sequence polymorphisms in MEM1 have to be in the C_4 state for a robust mesophyll-specific gene expression and that one C_4-type submodule is not sufficient.

Evolutionary Origin of MEM1 in the Genus *Flaveria*

The experiments carried out confirm the significance of both C_3-to-C_4 associated nucleotide polymorphisms in MEM1 as being indispensable for a mesophyll-specifc *ppcA* gene expression in C_4 *Flaverias*. A previously performed comparative analysis of *ppcA* gene sequences of two C_4 plants (*F. trinervia, F. bidentis*), two C_4-like plants (*F. palmerii, F. vaginata*) and of two C_3 plants (*F. cronquistii, F. pringlei*) revealed MEM1-like sequences in the *ppcA* gene of all six *Flaveria* species (Gowik et al., 2004). A C_4-specifc guanine at the first position in the A-submodule and the tetranucleotide CACT in the B-submodule is found in the MEM1 of the C_4 and C_4-like plants, whereas in the MEM1 of the two C_3 plants an C_3-specifc adenine is present and the tetranucleotide is absent (Gowik, et al. 2004).

According to the recently published phylogeny of the genus *Flaveria* which was based on both morphological and molecular characters all C_4 and C_4-like species form a distinct clade (clade A), while the C_3 species *F. pringlei* and *F. cronquistii* are basal (McKown et al., 2005). The C_3-C_4 intermediate species of the genus are contained within clade B (McKown et al., 2005). We were interested whether a MEM1 sequence is also found in the *ppcA* genes of C_3-C_4 intermediate *Flaveria* species and which states of MEM1 submodules are present. We selected *F. pubescens* and *F. brownii* the latter of which is considered to be a C_4-like C_3-C_4 intermediate. *PpcA* promoter sequences for *F. brownii* and *F. pubescens* with a length of 4030 bp and 4596 bp, respectively, were isolated, and MEM1 sequences were identified in each of these *ppcA* promoters. In the *F. brownii* promoter MEM1 was located 3830 base pairs

upstream of the translational start, while in the *F. pubescens* promoter MEM1 is positioned 4008 base pairs upstream of the ATG start codon. The MEM1 sequences of both *Flaverias* revealed an intermediate character, they possess a C_3-specifc adenine at the first position in the A-submodules, whereas the B-submodules are of a C_4-type due to the presence of the tetranucleotide CACT (Fig. 4A).

Phylogenetic analysis indicated that the current *ppcA* and *ppcB* genes arose from an ancestral *ppcB*-like gene by gene duplication (Bläsing et al., 2002). In order to enlighten the evolutionary origin of MEM1 further we analysed *ppcB* promoter sequences from the C_4 plant *F. trinervia* and the C_3 plant *F. pringlei* (Ernst and Westhoff, 1996). In both promoters MEM1-like sequences could be identified, hereafter referred to as MEM1*. The MEM1*-modules of the *ppcB* genes of *F. trinervia* and *F. pringlei* are quite similar. Both MEM1*-submodules are in the C_3-state with an adenine at the first nucleotide position of the A-submodules and with no CACT motif in the B-submodules. The A-submodules of MEM1* of the *ppcB* PEPC genes are not contiguous, they are interrupted by an insertion of 13 and 15 bp, respectively (Fig. 4B, supplementary Fig.1).

Discussion

The acquisition of new functions for an old gene, usually through changes in expression patterns and/or functional modifications of the encoded protein, is known as gene co-option and requires gene duplication events (Olson, 2006). Changes in the expression of a particular gene can result from alterations either in its *cis*-regulatory sequences or in the deployment and function of the transcription factors that control gene expression, or both (Love et al., 2007). Evolutionary biologists have collected convincing evidence, which supports the view that changes in the spatiotemporal expression patterns of genes are a principal mechanism for the evolution of novelty, both in morphological and biochemical traits (Doebley and Lukens, 1998). The multiple independent origins of C_4 photosynthesis in the angiosperms provides a good example for studying the evolution of novel morphological and biochemical traits. The evolution of the C_4 cycle enzymes required gene duplication with subsequent diversification through neo-functionalization (Monson, 2003).

In the present study, we have investigated the evolution of a *cis*-regulatory module, MEM1, that is necessary and sufficient for mesophyll-specific gene expression. We have found that changes at two positions in the 41 bp module, i.e. an A-to-G conversion in the A-submodule and the addition of the tetranucleotide CACT in the B-submodule, convert an element with no obvious function into a mesophyll specificity module. The comparison of MEM1 sequences from eight *Flaveria* species of known phylogenetic relationship and different modes of photosynthesis (Fig. 4A) suggests a scenario how an ubiquitously expressed C_3-type PEPC gene was converted into a cell-specifically expressed C_4-type PEPC (Fig. 5). All C_4 and C_4-like species of clade A possess MEM1 sequences with both submodules in the C_4 state, while the basal C_3 species have C_3-type submodules. The MEM1 modules of the C_3-C_4 species of clade B are intermediate with a C_3-type submodule A and a C_4-type submodule B. Based on the clear phylogenetic separation of the two clades one has to assume that the last common ancestor of clades A and B should have had already a C_4-type submodule B, and the insertion of the tetranucleotide CACT in the B-submodule of MEM1 occurred before the formation of the A and the B clade. *F. brownii* is the only C_4-like C_3-C_4 intermediate of the B clade and *F. pubescens* is less advanced than *F. brownii* in progression towards C_4 photosynthesis. These two species may be therefore considered as representative for the degree of C_3-C_4 intermediacy found in *Flaveria*, and we infer from this that MEM1 did not change during further evolution of the B clade species. In contrast, an adenine to guanine exchange in the A submodule of MEM1 led to the evolution of a mesophyll

specificity element in the A clade. We conclude that the C_3-to-C_4 related changes in MEM1 occurred step by step during evolution of C_4 photosynthesis in the genus *Flaveria* (Fig. 5).

While the C_4-type MEM1 acts as an expression module for mesophyll specificity, the function of its C_3 counterpart remains obscure. The incomplete C_3-type MEM1 in the original *ppcA* promoter of *F. pringlei* (Stockhaus et al., 1997) resulted in the same expression pattern as obtained for the complete promoter (Fig. 2C) suggesting that the C_3 MEM1 is dispensable. However, MEM1 is conserved in the two C_3 species *F. pringlei* and *F. cronquistii* (Fig. 4A). This may indicate that this sequence element is functional. Interestingly, MEM1-like (MEM1*) sequences can also be found in the non-photosynthetic *ppcB* PEPC genes of *F. pringlei* and *F. trinervia* (Ernst and Westhoff, 1996) (Fig. 4B). While the complete *ppcA* and *ppcB* promotors of *F. pringlei* and *F. trinervia* share 19 - 22% identical nucleotides, 75 - 82% sequence similarity is found if only the MEM1 and MEM1* sequences of these promoters are compared. This conservation of MEM1 and MEM1* sequences suggests a function of these elements also in the promoters of the *ppcB* and the non-photosynthetic *ppcA* genes even if this function remains obscure so far.

The C_4 MEM1 is a *cis*-regulatory module with a dual function. It represses the expression of the linked gene in bundle-sheath cells and the vascular tissue and concomitantly enhances transcription in the mesophyll cells. Interestingly the repressing function of C_4 MEM1 acts both in the bundle-sheath cells and the vascular tissue suggesting that with respect to gene expression both tissues are partly co-regulated. How transcriptional enhancing and repressing are explained mechanistically is unclear at present. The two polymorphic sites are 25 bp apart in *F. trinervia* but are separated by 122 bp in *F. bidentis*. Since both MEM1 modules direct the same expression specificity, the distance of the two polymorphic sites is not relevant. One may infer that *trans*-regulatory factors bind separately to the A- and B-submodules of MEM1. Preliminary analysis with the yeast one-hybrid system revealed that *trans*-regulatory factors of the basic leucine zipper (bZIP) protein family interact strongly with the C_4-type MEM1 while interaction with the C_3-type MEM1 is relatively weak (M. Akyildiz & P. Westhoff, unpublished data). The binding sites of these bZIP-type proteins within MEM1 have not been determined and their *in vivo* relevance for the control of expression of the *ppcA* gene in C_4 *Flaveria* species has not been investigated yet. It is also not known whether other *trans*-regulatory factors may interact with these bZIP proteins and are required for the functioning of MEM1 (Després et al., 2000).

The data presented here show that small changes in nucleotide sequence were sufficient to create a novel mode of gene expression. Since such small changes are likely to occur in

plant genomes quite easily, it is conceivable that the compartmentalized gene expression in C_4 plants arose many times independently during the evolution of angiosperms. For the first time in the study of C_4 photosynthesis the molecular anatomy and evolution of a *cis*-regulatory module for cell-specific gene expression has been elucidated at the nucleotide level. It will be interesting to know whether MEM1 represents an universal *cis*-regulatory module for mesophyll-specific gene expression in *Flaveria*. The C_4 carbonic anhydrase of *Flaveria* is a good choice to answer this question (Burnell and Hatch, 1988). It may be even more interesting to investigate how mesophyll-specific gene expression was achieved in other families of the angiosperms that evolved C_4 species. The genomes of the Brassicaceae are presently intensively studied, and therefore the genus *Cleome* with its C_4 and C_3 species might be a good model system for a comparative analysis at the genome level (Brown et al., 2005).

Methods

DNA manipulations were carried out according to Sambrook and Russell (Sambrook and Russell, 2001). All DNA fragments created by PCR were confirmed by DNA sequencing. Plasmid pBluescript II SK(+) (pBIISK(+); Stratagene, La Jolla, California) was used for standard cloning in *Escherichia coli* (Sambrook and Russell, 2001).

Construction of a Complete Promoter of *F. pringlei* (*ppcA*Fp)

Inspection of the previously used *ppcA* promoter of *F. pringlei* (named ppcA-L-Fp, from nucleotides -1 to -2583, Stockhaus et al., 1994) revealed that the A-submodule of MEM1 was lacking. For the generation of a complete *ppcA* promoter of *F. pringlei* (from nucleotides -1 to -2781) a PCR reaction with the oligonucleotide primers Fp-5'HindIII/Fp-3'XhoI (Table 1) and genomic DNA of *F. pringlei* as template was performed. The resulting DNA fragment was sub-cloned into pBIISK(+). After digestion with *Hind*III/*Hpa*I the released fragment was inserted into the *Hind*III/*Hpa*I digested vector ppcA-L-Fp (Stockhaus et al., 1994).

Construction of the Fusion of MEM1 of *F. pringlei* with the Proximal Region of the *ppcA* Promoter of *F. pringlei* (FpPR-FpM)

For the generation of this construct a PCR reaction with the oligonucleotide primers FpDEab5'HindIII/FpDEab3'XbaI (Table 1) and genomic DNA of *F. pringlei* as template was performed. The resulting DNA fragment was sub-cloned into pBIISK(+). After digestion with *Hind*III/*Xba*I, the resulting fragment was inserted into the *Hind*III/*Xba*I digested vector ppcA-PR_{Fp} (Stockhaus et al., 1994).

Fusions of MEM1 Variants with the C_4 Proximal Region of the *ppcA* Promoter of *F. trinervia*

The following promoter GUS reporter gene fusions are based on the construct ppcA-PR_{Ft}-$DR(+)_{Ft}$ (Gowik et al., 2004). The constructs with native MEM1 of *F. trinervia* (ppcA-PR_{Ft}-$DRab_{Ft}$) and the MEM1 version lacking the CACT tetranucleotide in the B-submodule (ppcA-PR_{Ft}-$DRab_{Ft}$-Δcact) have been described (Gowik et al., 2004). For reasons of nomenclature they were renamed to FtPR-FtM and FtPR-FtM/ΔCACT, respectively. For the generation of the two deletion-constructs FtPR-FtM/Δnt1-11 and FtPR-FtM/Δnt12-41, and for

the two substitution-constructs FtPR-FtM/A and FtPR-FtM/A_ΔCACT, the respective oligonucleotides were synthesized (Table 1) and annealed. The annealed oligonucleotides were digested with *Hind*III/*Xba*I and inserted into ppcA-PR_{Ft}-$DR(+)_{Ft}$ to replace the DR fragment.

In order to fuse MEM1 of *F. bidentis* (-1859 to -1996), the MEM1 of *F. pringlei* (-2454 to -2538) and the "C_4-converted" MEM1 of *F. pringlei* to the proximal region of the *ppcA* promoter of *F. trinervia*, PCR reactions were performed using the primers as depicted in Table 1 and genomic DNA of *F. bidentis* and *F. pringlei* as templates. The resulting DNA fragments were digested with *Hind*III/*Xba*I and sub-cloned into pBIISK(+). After an additional digestion with *Hind*III/*Xba*I, the released MEM1-fragments were inserted into ppcA-PR_{Ft}-$DR(+)_{Ft}$ to replace the DR fragment. The constructs were named FtPR-FbM, FtPR-FpM, and FtPR-FpM/G_+CACT, respectively.

Isolation of 5' Flanking Sequences from the *ppcA*-PEPC Genes of *F. brownii* and *F. pubescens*

The 5' flanking regions of *ppcA* genes of *F. brownii* and *F. pubescens* were isolated from total DNA by vectorette PCR (Siebert et al., 1995) with the Universal Genome Walker Kit (Clontech, Mountain View, California) as described in Gowik *et al.* (2004).

Generation of Transgenic *Flaveria bidentis*

The promoter/GUS reporter gene constructs were introduced into the *Agrobacterium tumefaciens* strain AGL1 (Lazo et al., 1991) *via* electroporation. *F. bidentis* was transformed as described (Chitty et al., 1994). Integration of the chimerical genes into the *F. bidentis* genome was confirmed by PCR.

Measurement of GUS Activity and Histochemical Analysis

T_o plants of *F. bidentis* were used for quantitative and histochemical analysis of the GUS reporter gene. For the histochemical analysis mature *F. bidentis* plants, grown in the greenhouse up to 40/50 cm and before flower initiation, were used (Stockhaus et al., 1997). GUS activities were measured quantitatively as described (Jefferson, 1987). The average values of the data are expressed by medians, and the Mann-Whitney U test statistics as

implemented in the software package Kaleidagraph Version 3.6 for Mac OS X (Synergy Software, Reading, PA www.synergy.com) was used to test whether two data series differ from each other.

Accession numbers

Sequence data have been deposited with the EMBL/GenBank data libraries under the accession numbers: EF522173 (5' upstream region of the *ppcA* gene of *F. brownii*) and EF522174 (5' upstream region of the *ppcA* gene of *F. pubescens*).

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. Nucleotide sequence alignment of the 5'-upstream regions of the *ppcA* genes of *F. trinervia*, *F. bidentis*, *F. brownii*, *F. pubescens*, *F. cronquistii* and *F. pringlei* and the *ppcB* genes of *F. trinervia* and *F. pringlei*.

Acknowledgments

Work described here was supported by the Deutsche Forschungsgemeinschaft (DFG) through the Sonderforschungsbereich 590 at the Heinrich Heine University of Düsseldorf. We are very grateful to our gardeners for the careful cultivation of the plants in the greenhouse. We are indebted to K. Ernst and K. Meierhoff for carefully reading the manuscript.

Footnotes

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Peter Westhoff (west@uni-duesseldorf.de).

Abbreviations

MEM1, mesophyll expression module 1; PEPC, phosphoeno*l*pyruvate carboxylase; PEP, phosphoeno*l*pyruvate; GUS, ß-glucuronidase; Ft, *Flaveria trinervia*; Fp, *Flaveria pringlei*; Fb, *Flaveria bidentis*; DR, distal region; PR, proximal region

References

Bauwe, H., and Chollet, R. (1986). Kinetic properties of phosphoenolpyruvate carboxylase from C_3, C_4, and C_3-C_4 intermediate species of *Flaveria* (Asteraceae). Plant Physiol. **82,** 695-699.

Bläsing, O.E., Ernst, K., Streubel, M., Westhoff, P., and Svensson, P. (2002). The non-photosynthetic phosphoenolpyruvate carboxylases of the C4 dicot Flaveria trinervia - implications for the evolution of C4 photosynthesis. Planta **215,** 448-456.

Brown, N.J., Parsley, K., and Hibberd, J.M. (2005). The future of C4 research-maize, Flaveria or Cleome? Trends Plant Sci. **10,** 215-221.

Burnell, J.N., and Hatch, M.D. (1988). Low bundle sheath carbonic anhydrase is apparently essential for effective C4 pathway operation. Plant Physiol. **86,** 1252-1256.

Chitty, J.A., Furbank, R.T., Marshall, J.S., Chen, Z., and Taylor, W.C. (1994). Genetic transformation of the C_4 plant, *Flaveria bidentis*. Plant J. **6,** 949-956.

Després, C., DeLong, C., Glaze, S., Liu, E., and Fobert, P.R. (2000). The Arabidopsis NPR1/NIM1 protein enhances the DNA binding activity of a subgroup of the TGA family of bZIP transcription factors. Plant Cell **12,** 279-290.

Doebley, J., and Lukens, L. (1998). Transcriptional regulators and the evolution of plant form. Plant Cell **10,** 1075-1082.

Ernst, K., and Westhoff, P. (1996). The phosphoenolpyruvate carboxylase (*ppc*) gene family of *Flaveria trinervia* (C_4) and *F. pringlei* (C_3): molecular characterization and expression analysis of the *ppcB* and *ppcC* genes. Plant Mol. Biol. **34,** 427-443.

Gowik, U., Burscheidt, J., Akyildiz, M., Schlue, U., Koczor, M., Streubel, M., Westhoff, P. (2004). cis-Regulatory elements for mesophyll-specific gene expression in the C4 plant Flaveria trinervia, the promoter of the C4 phosphoenolpyruvate carboxylase gene. Plant Cell. **16,** 1077-1090.

Hatch, M.D. (1987). C_4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. Biochim. Biophys. Acta **895,** 81-106.

Hermans, J., and Westhoff, P. (1992). Homologous genes for the C_4 isoform of phosphoenolpyruvate carboxylase in a C_3- and a C_4-*Flaveria* species. Mol. Gen. Genet. **234,** 275-284.

Jefferson, R.A. (1987). Assaying chimeric genes in plants: The GUS gene fusion system. Plant Mol. Biol. Rep. **5,** 387-405.

Ku, M.S.B., Wu, J., Dai, Z., Scott, R.A., Chu, C., and Edwards, G.E. (1991). Photosynthetic and photorespiratory characteristics of *Flaveria* species. Plant Physiol. **96,** 518-528.

Lazo, G.R., Stein, P.A., and Ludwig, R.A. (1991). A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. Bio. Technol. **9,** 963-967.

Love, A.C., Andrews, M.E., and Raff, R.A. (2007). Gene expression patterns in a novel animal appendage: the sea urchin. Evol. Dev. **9,** 51-68.

McKown, A.D., Moncalvo, J.M., and Dengler, N.G. (2005). Phylogeny of Flaveria (Asteraceae) and of C4 Photosynthesis Evolution. Am. J. Bot. **92,** 1911-1928.

Monson, R.K. (2003). Gene duplication, neofunctionalization, and the evolution of C4 photosynthesis. Int. J. Plant Sci. **164 Suppl,** S43-S54.

Monson, R.K., and Moore, B.D. (1989). On the significance of C_3-C_4 intermediate photosynthesis to the evolution of C_4 photosynthesis. Plant Cell Environ. **12,** 689-699.

Nelson, T., and Dengler, N.G. (1992). Photosynthetic tissue differentiation in C_4 plants. Int. J. Plant Sci. **153,** S93-S105.

Olson, E.N. (2006). Gene regulatory networks in the evolution and development of the heart. Science **313,** 1922-1927.

Patel, M., Siegel, A.J., and Berry, J.O. (2006). Untranslated regions of FbRbcS1 mRNA mediate bundle sheath cell-specific gene expression. J. Biol. Chem. **281,** 25485-25491.

Powell, A.M. (1978). Systematics of *Flaveria* (Flaveriinae-Asteraceae). Ann. Mo. Bot. Gard. **65,** 590-636.

Sage, R.F. (2004). The evolution of C4 photosynthesis. New Phytol. **161,** 341-370.

Sambrook, J., and Russell, D.W. (2001). Molecular Cloning. A Laboratory Manual. (Cold Spring Harbor Laboratory: Cold Spring Harbor Laboratory Press).

Schäffner, A.R., and Sheen, J. (1992). Maize C_4 photosynthesis involves differential regulation of phosphoenolpyruvate carboxylase genes. Plant J. **2,** 221-232.

Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A., and Lukyanov, S.A. (1995). An improved PCR method for walking in uncloned genomic DNA. Nucl. Acid. Res. **23,** 1087-1088.

Stockhaus, J., Poetsch, W., Steinmüller, K., and Westhoff, P. (1994). Evolution of the C_4 phosphoenolpyruvate carboxylase promoter of the C_4 dicot *Flaveria trinervia*: an expression analysis in the C_3 plant tobacco. Mol. Gen. Genet. **245,** 286-293.

Stockhaus, J., Schlue, U., Koczor, M., Chitty, J.A., Taylor, W.C., and Westhoff, P. (1997). The promoter of the gene encoding the C_4 form of phosphoenolpyruvate carboxylase directs mesophyll specific expression in transgenic C_4 *Flaveria* spp. Plant Cell **9,** 479-489.

Westhoff, P., and Gowik, U. (2004). Evolution of C4 phospho*enol*pyruvate carboxylase genes and proteins: a case study with the genus *Flaveria*. Ann. Bot. **93,** 1-11.

Tables

Table 1. Oligonucleotides Used for the Generation of MEM1 Reporter Gene Vectors.

Promoter-GUS-Reporter Gene Vectors	Oligonucleotide/ Primer Designation	Sequence 5' to 3'
	FtDEΔAB5'HindIII	**AGCTT**AGAGCTGTACTTACTCACTAAAACAAACAA**T**
FtPR-FtM/Δnt1-11	FtDEΔAB3'XbaI	**CTAGA**TTGTTTGTTTTAGTGAGTAAGTACAGCTCT**A**
	FtDEAΔB5'HindIII	**AGCTT**GTGAATTTATG**T**
FtPR-FtM/Δnt12-41	FtDEAΔB3'XbaI	**CTAGA**CATAAATTCAC**A**
	FtDEaB5'HindIII	**AGCTT**ATGAATTTATGAGAGCTGTACTTACTCACTAAAACAAACAA**T**
FtPR-FtM/A	FtDEaB3'XbaI	**CTAGA**TTGTTTGTTTTAGTGAGTAAGTACAGCTCTCATAAATTCAT**A**
	FtDEAb5'HindIII	**AGCTT**GTGAATTTATGAGAGCTGTACTTACTAAAACAAACAA**T**
FtPR-FtM/ΔCACT	FtDEAb3'XbaI	**CTAGA**TTGTTTGTTTTAGTAAGTACAGCTCTCATAAATTCAC**A**
	FtDEab5'HindIII	**AGCTT**ATGAATTTATGAGAGCTGTACTTACTAAAACAAACAA**T**
FtPR-FtM/A_ΔCACT	FtDEab3'XbaI	**CTAGA**TTGTTTGTTTTAGTAAGTACAGCTCTCATAAATTCAT**A**
	FbDEAB5'HindIII	GGG**AAGCTT**GTGAATTTATGAAAAAATTAAATTGGAAAGAGG
FtPR-FbM	FbDEAB3'XbaI	GGG**TCTAGA**TTGTTTGTTTTAGTGAGTAAGTACGGCTCTACGAACAC
	Fp5'HindIII	GGG**AAGCTT**TTTCTTTTGTATTTGTTATTGTTTACG
ppcAFp	Fp3'HpaI	GGG**GTTAAC**GCCTCTATGTACAGAGAATACC
FpPR-FpM	FpDEab5'HindIII	GGG**AAGCTT**ATGAATTTATGAAAAACTCGTG
FtPR-FpM	FpDEab3'XbaI	GGG**TCTAGA**TTGTTTGTTTTAGTAAGTACG
	FpDEAB5'HindIII	GGG**AAGCTT**GTGAATTTATGAAAAACTCGTGAAGAG
FtPR-FpM/G_+CACT	FbDEAB3'XbaI	GGG**TCTAGA**TTGTTTGTTTTAGTGAGTAAGTACGGCTCTACGAACAC

Restriction sites of *Hind*III, *Xba*I and *Hpa*I are given in bold letters.

Figure Legends

Figure 1. Analysis of the *ppcA*1 GUS Reporter Gene Constructs FtPR, FtPR-FtM and FtPR-FbM in Transgenic *F. bidentis.*

(A) Schematic presentation of the *ppcA*/GUS chimerical genes. Nucleotide numbers refer to the translation initiation codon. The Ft MEM1 region is indicated by light green boxes, the proximal region (PR) is indicated by dark green boxes. The state of the C_3-to-C_4 associated polymorphisms in the A-submodule (G or A) and the B-submodule (presence or absence of CACT) of MEM1 is indicated.

(B) GUS activities in leaves of transgenic *F. bidentis* plants. The median value of the GUS activities is expressed in nanomoles of the reaction product 4-methylumbellliferone (MU) generated per milligram of protein per minute. The numbers of independent transgenic plants investigated (N) are indicated at the top of each column.

(C) to **(E)** Histochemical localisation of GUS activity in leaf sections of transgenic *F. bidentis* plants transformed with the FtPR (C), FtPR-FtM **(D)** or the construct FtPR-FbM **(E)**. Incubation times were 25 h **(C),** 12 h **(D)** and 24 h **(E)**.

Figure 2. Analysis of the *ppcA* GUS Reporter Gene Constructs ppcAFp, FpPR-FpM, FtPR-FpM, FtPR-FtM/A_ΔCACT and FtPR-FpM/G_+CACT in Transgenic *F. bidentis.*

(A) Schematic presentation of the *ppcA*1/GUS chimerical genes. Nucleotide numbers refer to the translation initiation codon. The Ft MEM1 region is indicated by light green boxes, the Fp MEM1 region is indicated by light yellow boxes. The Ft proximal region (PR) is indicated by dark green boxes, the Fp PR is indicated by dark yellow boxes. The state of the C_3-to-C_4 associated polymorphisms in the A-submodule (G or A) and the B-submodule (presence or absence of CACT) of MEM1 is indicated.

(B) GUS activities in leaves of transgenic *F. bidentis* plants. The numbers of independent transgenic plants investigated (N) are indicated at the top of each column. Median values are shown. Mu, 4-methylumbellliferone.

(C) to **(G)** Histochemical localisation of GUS activity in leaf sections of transgenic *F. bidentis* plants transformed with ppcAFp **(C)**, FpPR-FpM **(D)**, FtPR-FpM **(E)**, FtPR-FtM/A_ΔCACT **(F)** and FtPR-FpM/G_+CACT **(G)**. Incubation times were 2 h **(C)**, 3.5 h **(D)**, 18 h **(E)**, 17 h **(F)** and 3 h **(G)**.

Figure 3. Analysis of the *ppcA* GUS Reporter Gene Constructs FtPR-FtM/Δnt12-41, FtPR-FtM/Δnt1-11, FtPR-FtM/A and FtPR-FtM/ΔCACT in Transgenic *F. bidentis.*

(A) Schematic presentation of the *ppcA*/GUS chimerical genes. Nucleotide numbers refer to the translation initiation codon. The Ft MEM1 region is indicated by dark green boxes, the Fp MEM1 region is indicated by dark yellow boxes. The Ft proximal region (PR) is indicated by a light green box. The state of the C_3-to-C_4 associated polymorphisms in the A-submodule (G or A) and the B-submodule (presence or absence of CACT) of MEM1 is indicated.

(B) GUS activities in leaves of transgenic *F. bidentis* plants. The numbers of independent transgenic plants investigated (N) are indicated at the top of each column. Median values are shown. Mu, 4-methylumbellliferone.

(C) to **(F)** Histochemical localisation of GUS activity in leaf sections of transgenic *F. bidentis* plants transformed with FtPR-FtM/Δnt12-41 **(C)**, FtPR-FtM/Δnt1-11 **(D)**, FtPR-FtM/A **(E)** and FtPR-FtM/ΔCACT **(F)**. Incubation times were 21 h **(C)**, **(D)**, **(F)** and 24 h **(E)**. Note that some FtPR-FtM/ΔCACT plants have been already analyzed in Gowik et al., (2004). In the course study they were re-analyzed by extending the staining period to two days in order to increase the sensitivity of detection. In addition new plants were generated.

Figure 4. Comparisons of MEM1 and MEM1-Like Sequences

(A) MEM1 sequences of the *ppcA* promoter from C_4, C_4-like, C_3-C_4-intermediate and C_3 *Flaveria* species. (Ft: *F. trinervia,* Fb: *F. bidentis*, Fpa: *F. palmerii,* Fv: F. vaginata, Fbr: *F. brownii,* Fpu: *F. pubescens*, Fc: *F. cronquistii*, Fp: *F. pringlei*)

(B) Comparison of MEM1 sequences from the *ppcA* promoters of *F. trinervia* (Ft) and *F. pringlei* (Fp) with their MEM1-like counterparts from the *ppcB* genes.

The MEM1 A- and B-submodules are highlighted by boxes. Asterisks label identical nucleotides in the A- or B-submodule of all promoters. Grey bars indicate the single nucleotide difference in the A- and the insertion/deletion of the CACT tetranucleotide in the B-submodule. For a comparison of the whole promoters see supplement figure 1.

Figure 5. Model of MEM1 Evolution in the Genus *Flaveria*.

The model relies on the phylogeny of the genus *Flaveria* which is based on morphological and molecular data sets (McKown *et al*., 2005).

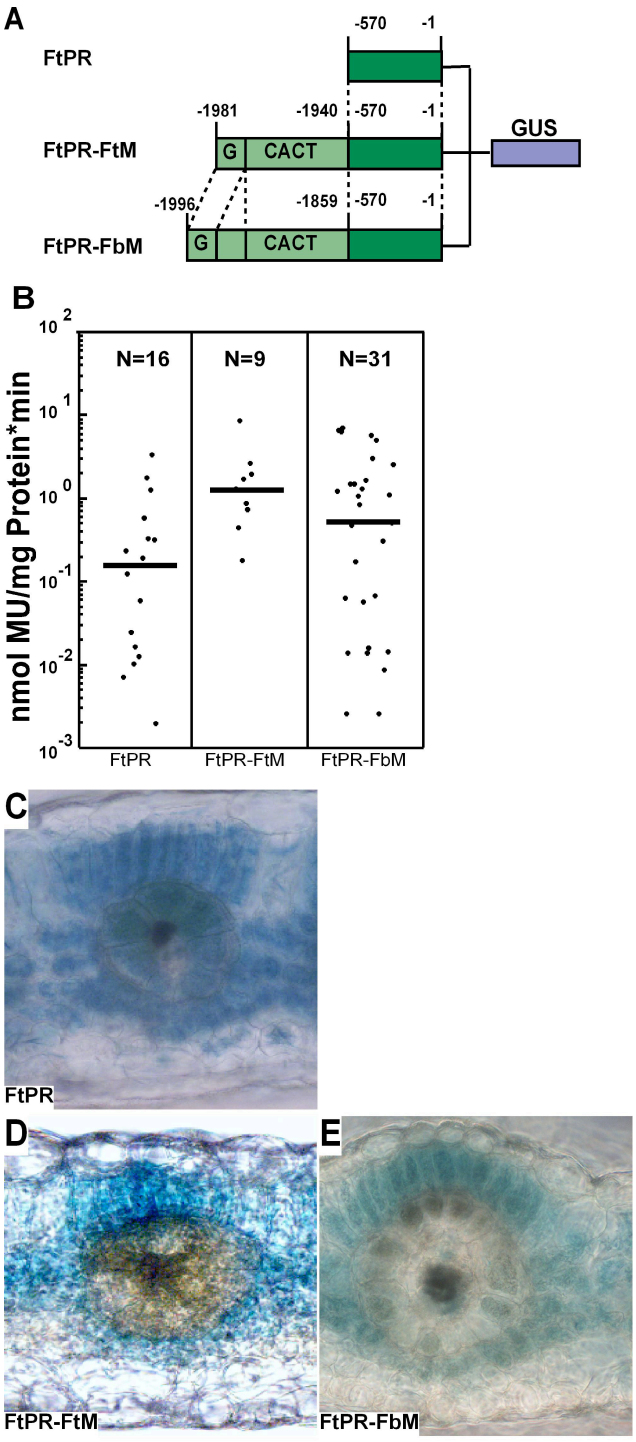

A
FtPR
-570
-1
FtPR-FtM
-1981
-1940
-570
-1
G
CACT
GUS
FtPR-FbM
-1996
-1859
-570
-1
G
CACT
B
nmol MU/mg Protein*min
N=16
N=9
N=31
FtPR
FtPR-FtM
FtPR-FbM
C
FtPR
D
FtPR-FtM
E
FtPR-FbM

Figure 1.

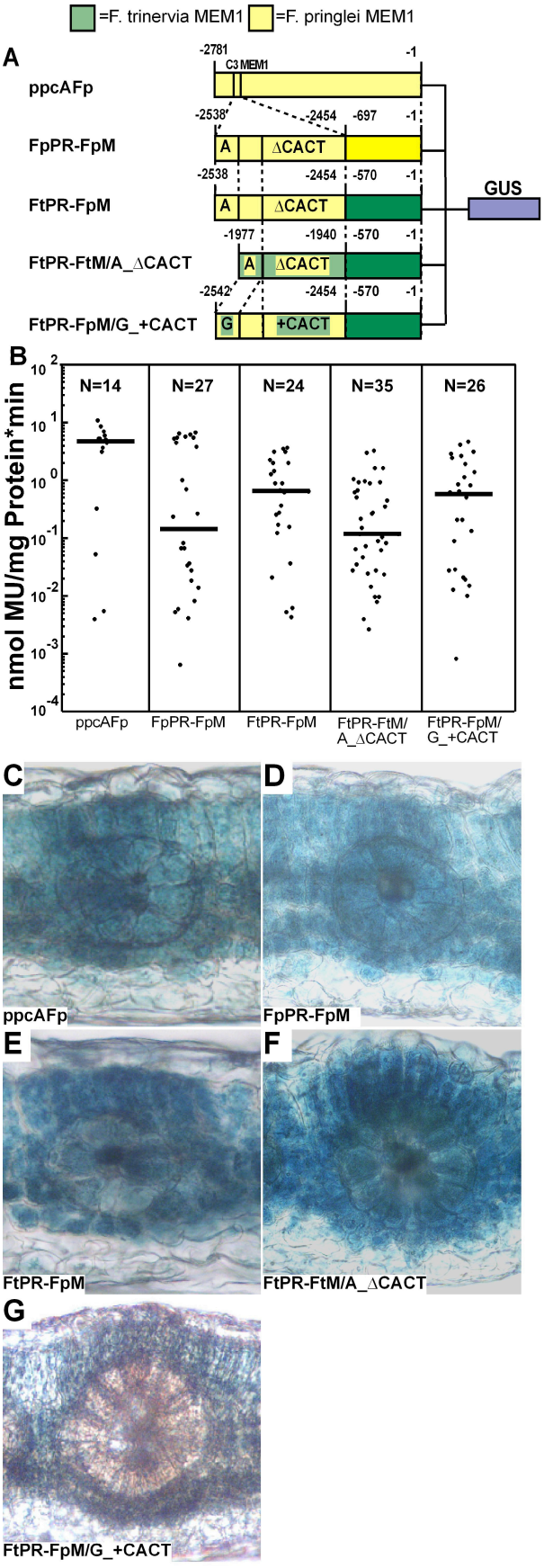

=F. trinervia MEM1
=F. pringlei MEM1
A
ppcAFp
-2781
C3 MEM1
-1
FpPR-FpM
-2538
-2454
-697
-1
A
ΔCACT
FtPR-FpM
-2538
-2454
-570
-1
GUS
FtPR-FtM/A_ΔCACT
-1977
-1940
-570
-1
FtPR-FpM/G_+CACT
-2542
-2454
-570
-1
G
+CACT
B
nmol MU/mg Protein*min
N=14
N=27
N=24
N=35
N=26
ppcAFp
FpPR-FpM
FtPR-FpM
FtPR-FtM/ A_ΔCACT
FtPR-FpM/ G_+CACT
C
ppcAFp
D
FpPR-FpM
E
FtPR-FpM
F
FtPR-FtM/A_ΔCACT
G
FtPR-FpM/G_+CACT

Figure 2.

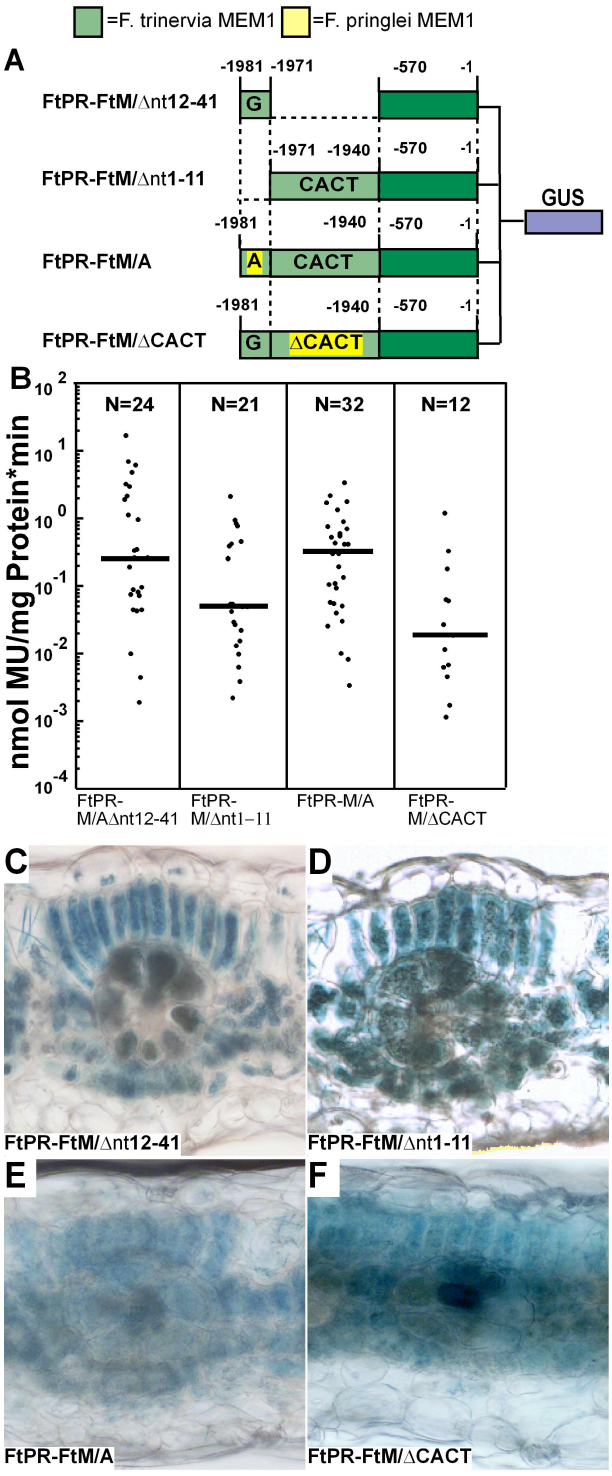

=F. trinervia MEM1
=F. pringlei MEM1
A
-1981 -1971 -570 -1
FtPR-FtM/Δnt12-41
G
-1971 -1940 -570 -1
FtPR-FtM/Δnt1-11
CACT
GUS
-1981 -1940 -570 -1
FtPR-FtM/A
A
CACT
-1981 -1940 -570 -1
FtPR-FtM/ΔCACT
G
ΔCACT
B
nmol MU/mg Protein*min
N=24
N=21
N=32
N=12
FtPR-M/AΔnt12-41
FtPR-M/Δnt1−11
FtPR-M/A
FtPR-M/ΔCACT
C
FtPR-FtM/Δnt12-41
D
FtPR-FtM/Δnt1-11
E
FtPR-FtM/A
F
FtPR-FtM/ΔCACT

Figure 3.

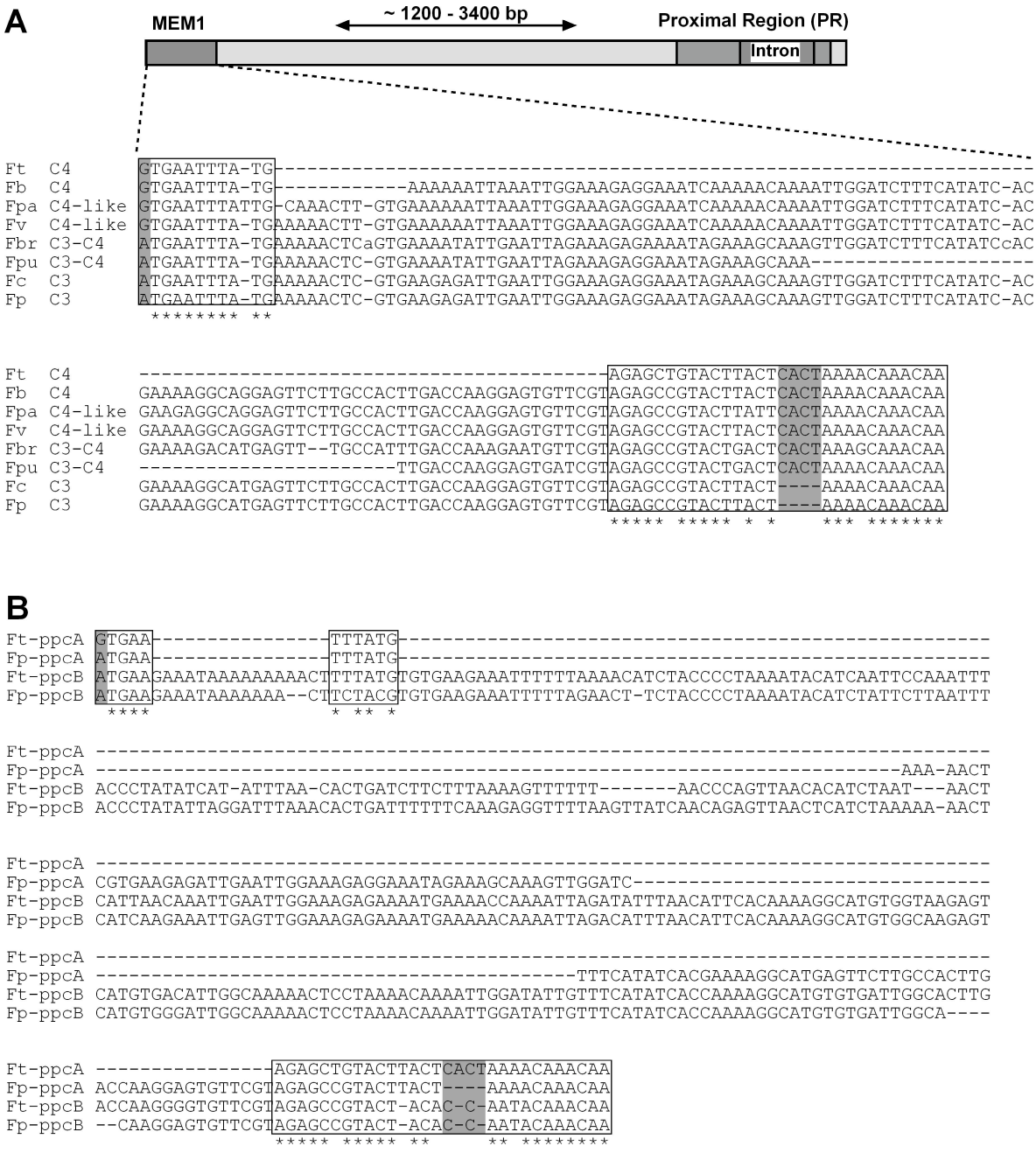

A
MEM1
~ 1200 - 3400 bp
Proximal Region (PR)
Intron
Ft C4
Fb C4
Fpa C4-like
Fv C4-like
Fbr C3-C4
Fpu C3-C4
Fc C3
Fp C3
B
Ft-ppcA
Fp-ppcA
Ft-ppcB
Fp-ppcB

Figure 4.

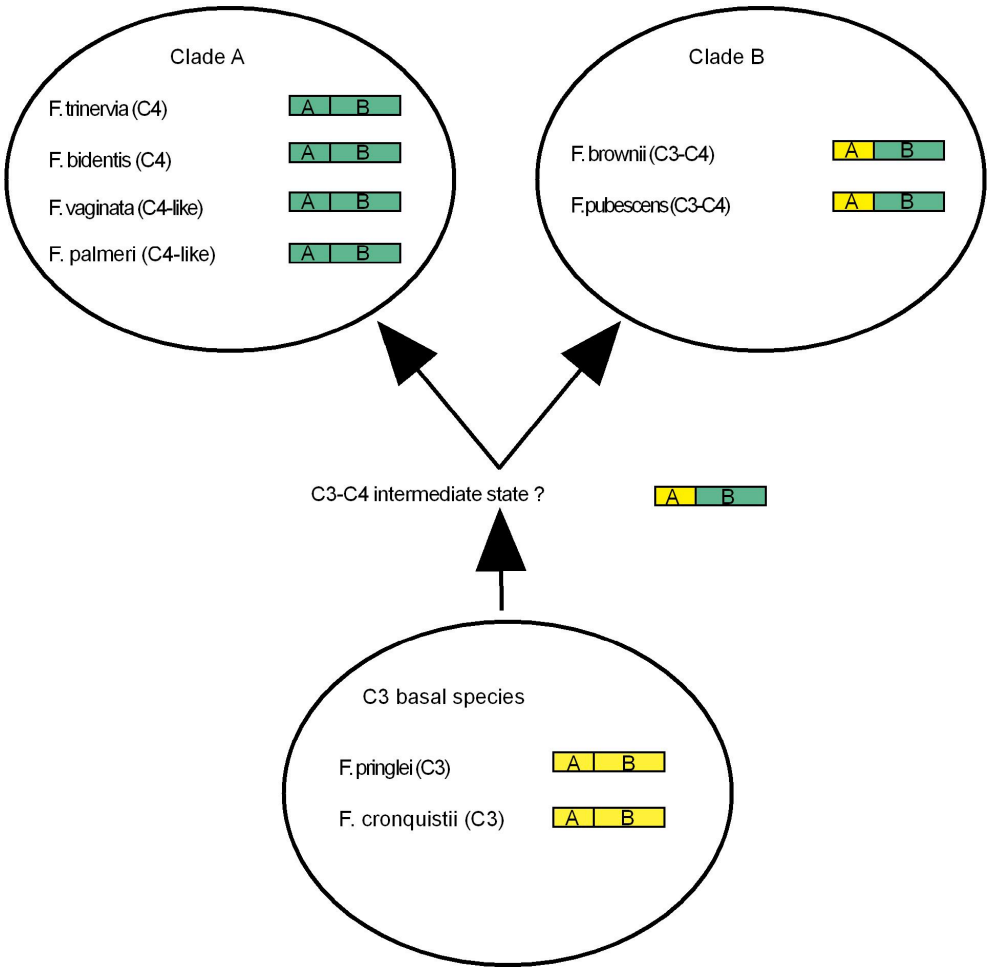
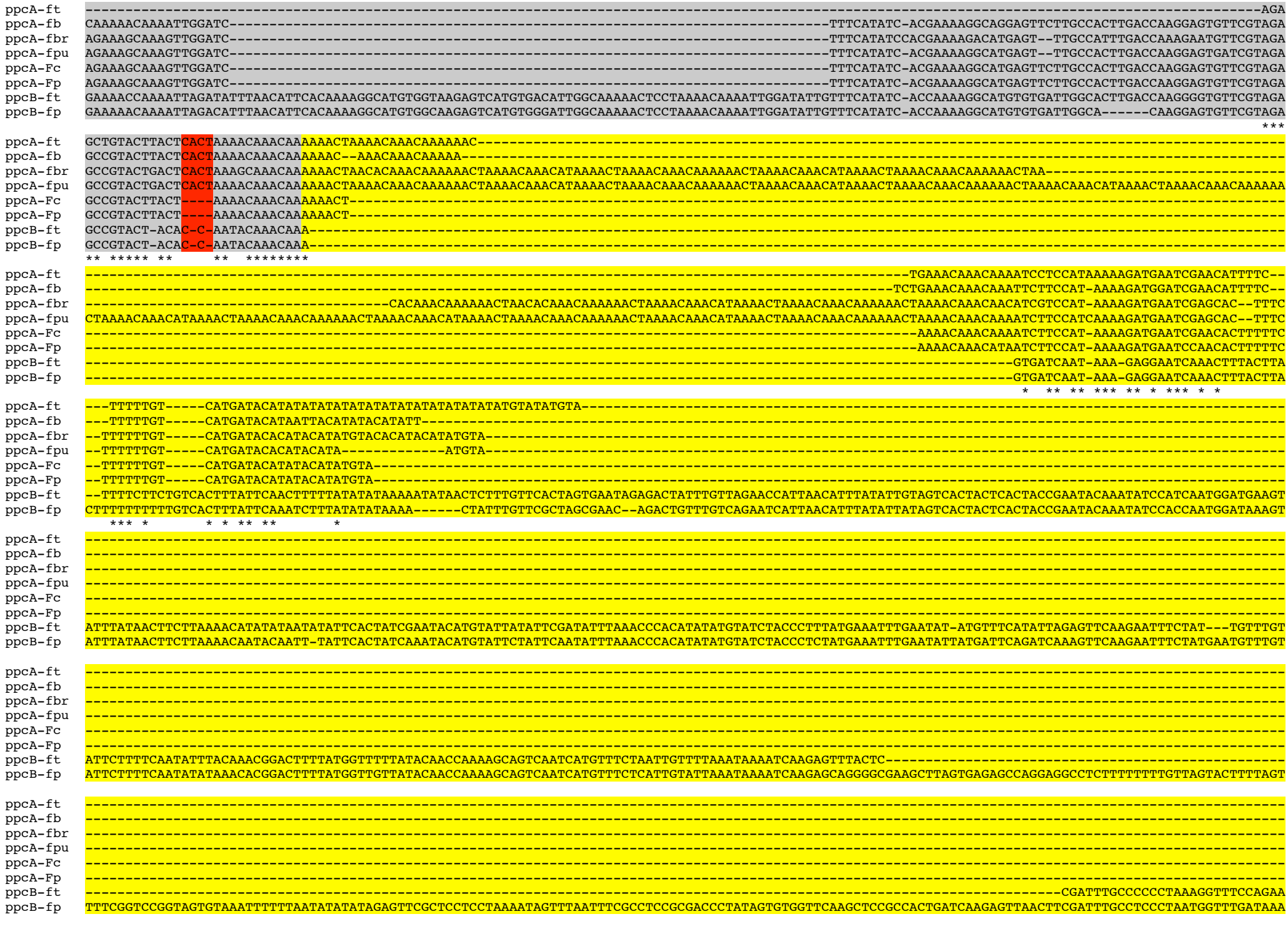
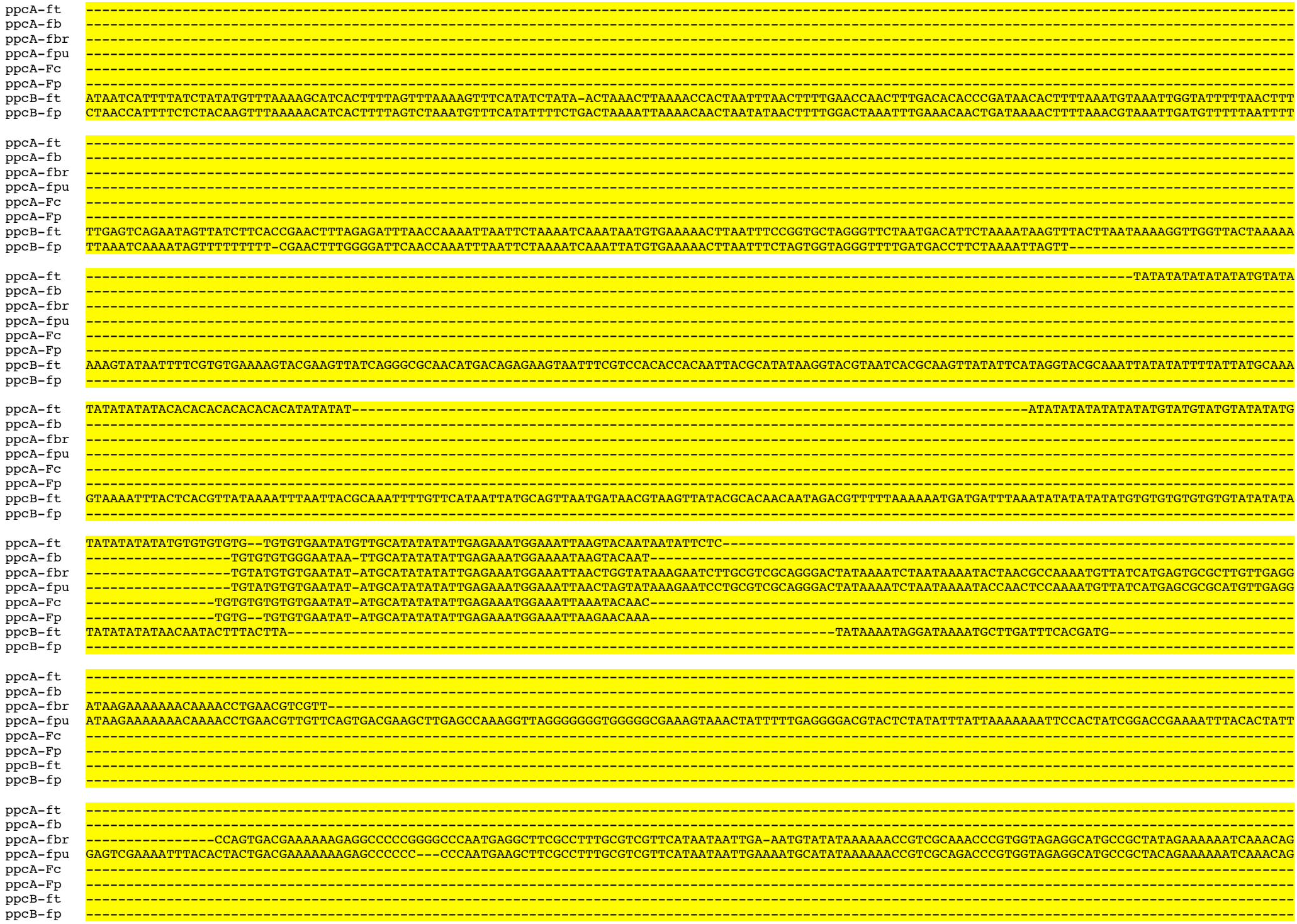

Clade A
F. trinervia (C4)
A B
F. bidentis (C4)
A B
F. vaginata (C4-like)
A B
F. palmeri (C4-like)
A B
Clade B
F. brownii (C3-C4)
A B
F. pubescens (C3-C4)
A B
C3-C4 intermediate state ?
A B
C3 basal species
F. pringlei (C3)
A B
F. cronquistii (C3)
A B

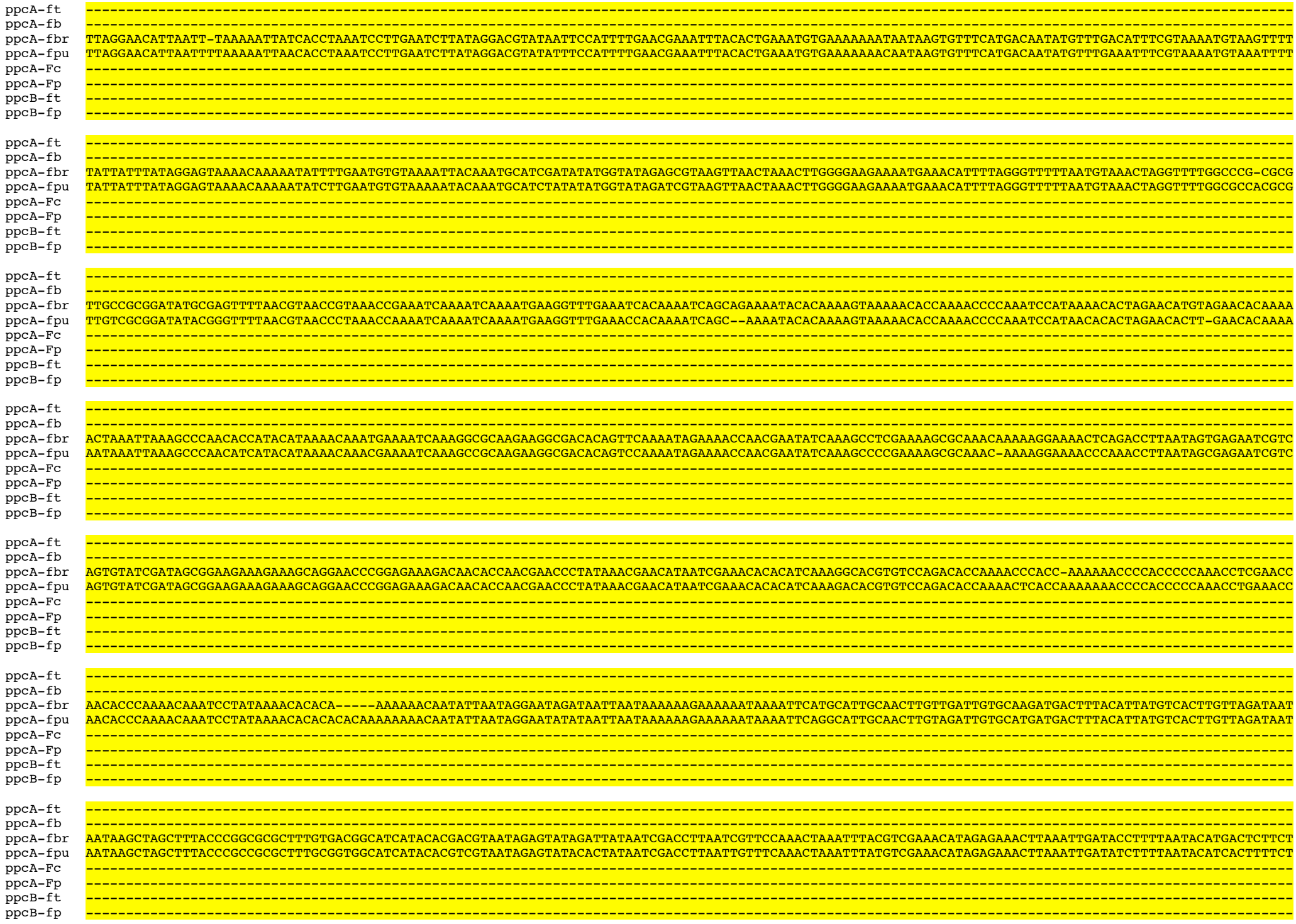
Figure 5.

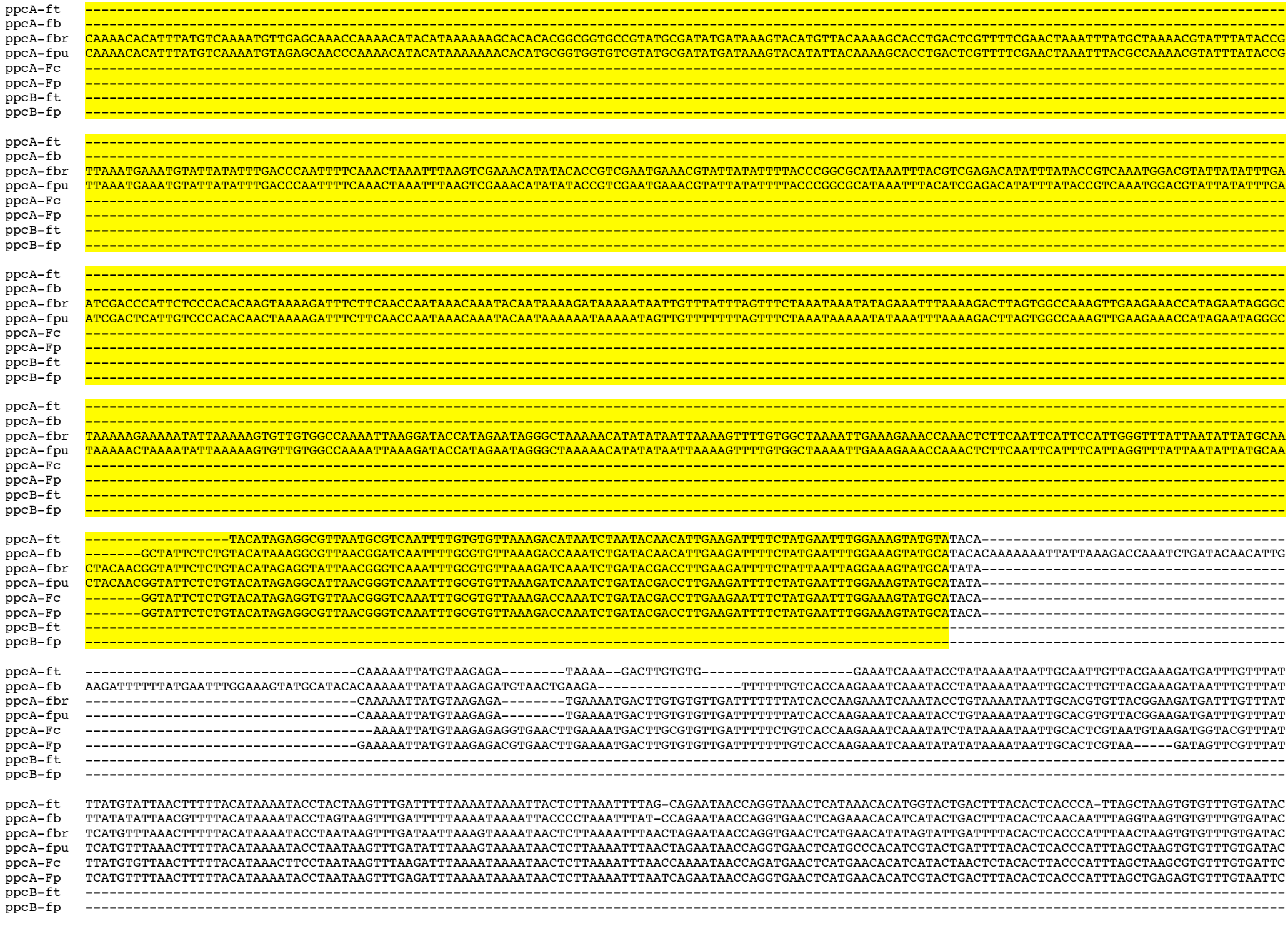

ppcA-ft
ppcA-fb
ppcA-fbr
ppcA-fpu
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp


ppcA-ft
ppcA-fb
ppcA-fbr
ppcA-fpu
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp

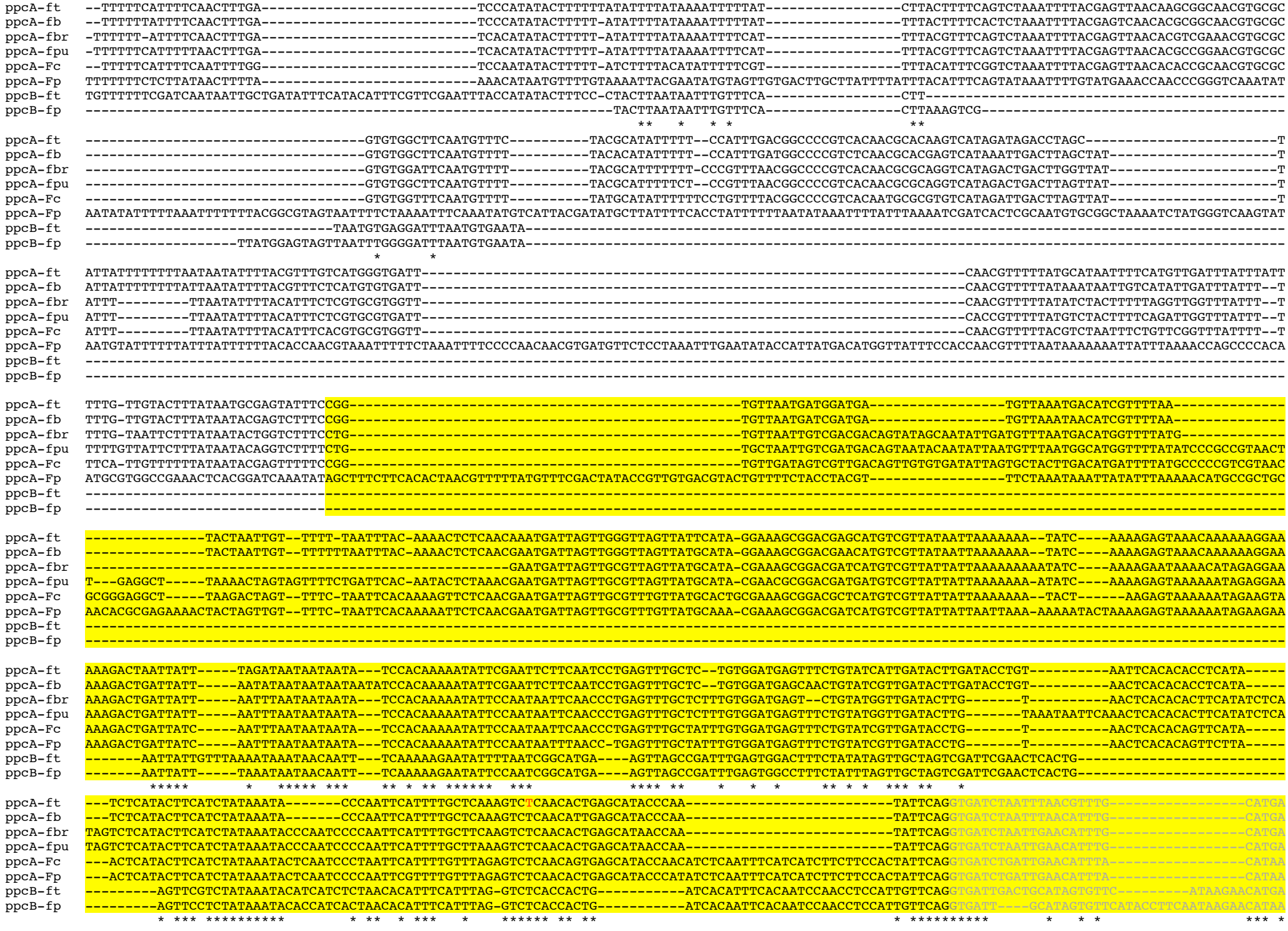

ppcA-ft
ppcA-fb
ppcA-fbr
ppcA-fpu
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp


ppcA-ft
ppcA-fb
ppcA-fbr
ppcA-fpu
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp


ppcA-ft
ppcA-fb
ppcA-fbr TTAGGAACATTAATT-TAAAAATTATCACCTAAATCCTTGAATCTTATAGGACGTATAATTCCATTTTGAACGAAATTTACACTGAAATGTGAAAAAAATAATAAGTGTTTCATGACAATATGTTTGACATTTCGTAAAATGTAAGTTTT
ppcA-fpu TTAGGAACATTAATTTTAAAAATTAACACCTAAATCCTTGAATCTTATAGGACGTATATTTCCATTTTGAACGAAATTTACACTGAAATGTGAAAAAAACACAATAAGTGTTTCATGACAATATGTTTGAAATTTCGTAAAATGTAAATTTT
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp
ppcA-ft
ppcA-fb
ppcA-fbr TATTATTTATAGGAGTAAAACAAAAATATTTTGAATGTGTAAAATTACAAATGCATCGATATATGGTATAGAGCGTAAGTTAACTAAACTTGGGGAAGAAAATGAAACATTTTAGGGTTTTTAATGTAAACTAGGTTTTGGCCCG-CGCG
ppcA-fpu TATTATTTATAGGAGTAAAACAAAAATATCTTGAATGTGTAAAAATACAAATGCATCTATATATGGTATAGATCGTAAGTTAACTAAACTTGGGGAAGAAAATGAAACATTTTAGGGTTTTTAATGTAAACTAGGTTTTGGCGCCACGCG
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp
ppcA-ft
ppcA-fb
ppcA-fbr TTGCCGCGGATATGCGAGTTTTAACGTAACCGTAAACCGAAATCAAAATCAAAATGAAGGTTTGAAATCACAAAATCAGCAGAAAATACACAAAAGTAAAAACACCAAAACCCCAAATCCATAAAACACTAGAACATGTAGAACACAAAA
ppcA-fpu TTGTCGCGGATATACGGGTTTTAACGTAACCCTAAACCAAAATCAAAATCAAAATGAAGGTTTGAAACCACAAAATCAGC--AAAATACACAAAAGTAAAAACACCAAAACCCCAAATCCATAACACACTAGAACACTT-GAACACAAAA
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp
ppcA-ft
ppcA-fb
ppcA-fbr ACTAAATTAAAGCCCAACACCATACATAAAACAAATGAAAATCAAAGGCGCAAGAAGGCGACACAGTTCAAAATAGAAAACCAACGAATATCAAAGCCTCGAAAAGCGCAAACAAAAAGGAAAACTCAGACCTTAATAGTGAGAATCGTC
ppcA-fpu AATAAATTAAAGCCCAACATCATACATAAAACAAACGAAAATCAAAGCCGCAAGAAGGCGACACAGTCCAAAATAGAAAACCAACGAATATCAAAGCCCCGAAAAGCGCAAAC-AAAAGGAAAACCCAAACCTTAATAGCGAGAATCGTC
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp
ppcA-ft
ppcA-fb
ppcA-fbr AGTGTATCGATAGCGGAAGAAAGAAAGCAGGAACCCGGAGAAAGACAACACCAACGAACCCTATAAACGAACATAATCGAAACACACATCAAAGGCACGTGTCCAGACACCAAAACCCACC-AAAAAACCCCACCCCCAAACCTCGAACC
ppcA-fpu AGTGTATCGATAGCGGAAGAAAGAAAGCAGGAACCCGGAGAAAGACAACACCAACGAACCCTATAAACGAACATAATCGAAACACACATCAAAGACACGTGTCCAGACACCAAAACTCACCAAAAAAACCCCACCCCCAAACCTGAAACC
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp
ppcA-ft
ppcA-fb
ppcA-fbr AACACCCAAAACAAATCCTATAAAACACACA-----AAAAAACAATATTAATAGGAATAGATAATTAATAAAAAAGAAAAAATAAAATTCATGCATTGCAACTTGTTGATTGTGCAAGATGACTTTACATTATGTCACTTGTTAGATAAT
ppcA-fpu AACACCCAAAACAAATCCTATAAAACACACACACAAAAAAAACAATATTAATAGGAATATATAATTAATAAAAAAGAAAAAATAAAATTCAGGCATTGCAACTTGTAGATTGTGCATGATGACTTTACATTATGTCACTTGTTAGATAAT
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp
ppcA-ft
ppcA-fb
ppcA-fbr AATAAGCTAGCTTTACCCGGCGCGCTTTGTGACGGCATCATACACGACGTAATAGAGTATAGATTATAATCGACCTTAATCGTTCCAAACTAAATTTACGTCGAAACATAGAGAAACTTAAATTGATACCTTTTAATACATGACTCTTCT
ppcA-fpu AATAAGCTAGCTTTACCCGCCGCGCTTTGCGGTGGCATCATACACGTCGTAATAGAGTATACACTATAATCGACCTTAATTGTTTCAAACTAAATTTATGTCGAAACATAGAGAAACTTAAATTGATATCTTTTAATACATCACTTTTCT
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp


ppcA-ft
ppcA-fb
ppcA-fbr
ppcA-fpu
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp


ppcA-ft
ppcA-fb
ppcA-fbr
ppcA-fpu
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp


ppcA-ft
ppcA-fb
ppcA-fbr
ppcA-fpu
ppcA-Fc
ppcA-Fp
ppcB-ft
ppcB-fp

Supplementary Fig. 1 Nucleotide sequence alignment of the 5'-upstream regions of the *ppcA* genes of *F. trinervia*, *F. bidentis*, *F. brownii*, *F. pubescens*, *F. cronquistii* and *F. pringlei* and the *ppcB* genes of *F. trinervia* and *F. pringlei*. The *ppcB* promoters of *F. trinervia* and *F. pringlei* are quite similar, they share about 66% identical nucleotides and contain a MEM1-like sequence. Identical positions in all *ppc* genes are marked by an asterisk. The start site of the *F. trinervia ppc* transcript is indicated by the red thymidine nucleotide, the position of the distal and proximal promoter region by the yellow boxes and the position of the intron in the 5' untranslated leader of the *ppc* genes by grey nucleotides. The MEM1 region is indicated by the grey box, the first nucleotide at the 5' region in MEM1 and the CACT tetranucleotide are indicated by red boxes.

Basic Leucine Zipper Proteins Interact with MEM1, the Mesophyll Specificity *Cis*-Regulatory Element of the C_4 Phospho*enol*pyruvate Carboxylase Gene of *Flaveria trinervia*

Meryem Akyildiz[*], Ming Chang Tsai[‡], Claus Seidel[‡] and Peter Westhoff[*†]

[*]Heinrich-Heine-Universität, Institut für Entwicklungs- und Molekularbiologie der Pflanzen, 40225 Düsseldorf, Germany

[‡]Heinrich-Heine-Universität, Institut für Molekulare Physikalische Chemie, 40225 Düsseldorf, Germany

Corresponding author: Prof. Dr. Peter Westhoff

Institut für Entwicklungs- und Molekularbiologie der Pflanzen

Heinrich-Heine Universität

Universitätsstrasse 1

40225 Düsseldorf, Germany

E-mail: west@uni-duesseldorf.de

Tel. +49 211 81 12338

Fax. +49 211 81 14871

Key words: C_4 photosynthesis | phospho*enol*pyruvate carboxylase | *Flaveria* | *cis*-regulatory element | *trans*-regulatory element | bZIP | fluorescence poalrization/anisotropy | yeast one-hybrid | electrophoretic mobility shift assay | fluorescence polarization/anisotropy

Pages: 36

Tables: 2

Figures: 6

Abstract

C_4 photosynthesis depends on a division of labour between two photosynthetic cell types, mesophyll and bundle-sheath cells. The two involved CO_2-assimilatory enzymes are strictly compartmentalized, while phospho*enol*pyruvate carboxylase (PEPC) is expressed only in mesophyll cells, the expression of ribulose-1,5-bisphosphate carboxylase/oxygenase is restricted to bundle-sheath cells. Previous studies identified a 41 bp *cis*-regulatory element, MEM1 (mesophyll expression module 1), in the distal promoter region of PEPC of the C_4 plant *Flaveria trinervia*, as being the major determinant for mesophyll-specific PEPC expression. MEM1 enhances PEPC expression in the mesophyll cells and acts as a repressor element suppressing PEPC expression in bundle-sheath cells and in the vascular bundles. In a yeast one-hybrid screen of a *F. trinervia* leaf cDNA library, using the C_4 MEM1 as bait, we identified three proteins of the basic leucine zipper family (bZIP), named FtbZIP18, 29 and 51, as potential binding partners. These bZIP proteins are highly similar to members of the group I of bZIP proteins of *Arabidopsis thaliana*. In protein-DNA interaction studies using the yeast one-hybrid system the FtbZIP proteins interacted strongly with the C_4 version of MEM1, while only a minor interaction was observed with the C_3 MEM1 variant. In contrast, *in vitro* analysis of the interaction of the FtbZIP proteins with MEM1 by electrophoretic mobility shift assays and fluorescence polarization/anisotropy measurements revealed no difference in the binding affinity of FtbZIP with either a C_4 or a C_3 MEM1. This indicates the involvement of other factors cooperating with FtbZIP proteins to bring about the required mesophyll specificity.

Introduction

C_4 plants evolved a unique biochemical mechanism to overcome the limitations of low CO_2 concentration and reduced photosynthetic efficiency associated with photorespiration caused by ribulose-1,5-bisphosphate carboxylase/oxygenase. The photosynthetic C_4 cycle acts as a pump that concentrates CO_2 at the site of ribulose-1,5-bisphosphate carboxylase/oxygenase. As a consequence the competitive inhibition of ribulose-1,5-bisphosphate carboxylase/oxygenase by oxygen is largely excluded and the net photosynthesis rate is increased (Edwards and Walker, 1983). The concentration of CO_2 in C_4 plant species is achieved by the metabolic interaction of two morphologically and functionally distinct cell types, mesophyll and bundle-sheath cells, and relies on the correct compartmentalization of the CO_2 assimilatory enzymes of the C_4 and the C_3 photosynthetic carbon cycle (Hatch and Oliver, 1978; Hatch, 1987).

Phospho*enol*pyruvate carboxylase (PEPC) is located in the mesophyll cells and catalyzes the primary fixation of CO_2 into the C_4 acid oxaloacetate, hence the name C_4 photosynthesis (O'Leary, 1982). In the NADP-malic enzyme subgroup of C_4 species oxaloacetate is subsequently reduced by NADP-malate dehydrogenase to form malate, which is transported to bundle-sheath cells where it is decarboxylated. The released CO_2 is finally channelled into the Calvin-Benson cycle and refixed by the bundle-sheath-specific enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (Hatch and Osmond, 1976; Edwards and Walker, 1983; Hatch, 1987).

This division of work between mesophyll and bundle-sheath cells depends on differential gene expression (Nelson and Dengler, 1992). In NADP-malic enzyme type C_4 species, transcripts for PEPC, NADP-malate dehydrogenase and pyruvate orthophosphate dikinase accumulate in mesophyll cells, whereas the enzymes NADP-dependent malic enzyme and ribulose-1,5-bisphosphate carboxylase/oxygenase are expressed exclusively in bundle-sheath cells. Regulatory mechanisms acting on the transcriptional level appear to be major determinants of the cell-specific expression of the respective genes (Schäffner and Sheen, 1992, Matsuoka et al., 1994; Stockhaus et al., 1997; Onodera et al., 2001; Gowik, 2004), however, posttranscriptional control has been reported too (Berry et al., 1986; Kubicki et al., 1994; Rosche and Westhoff, 1995; Brutnell et al., 1999; Lai et al., 2002; Patel et al., 2006).

C_4 plants evolved from C_3 ancestor species and this transition occurred independently several times during the evolution of angiosperms (Kellogg, 1999; Sage, 2004). The polyphyletic origin of C_4 plants suggests that the evolution of a C_3 into a C_4 species must

have been accomplished quite easily in genetic terms (Ehleringer and Monson, 1993; Ehleringer et al., 1997;). All the enzymes of the C_4 metabolism are already present in C_3 plants where they are involved in the basic cell metabolism (Latzko and Kelly, 1983). These ancestral C_3 genes have served as the basis for the evolution of the C_4 isoform genes (Moore, 1982; Monson and Moore, 1989; Monson, 1999).

To meet the special requirements of the C_4 photosynthetic pathway, the expression programme of the C_3 progenitor genes had to be changed to a high and selective expression in the mesophyll or bundle-sheath cells of the leaf. To gain insight into the evolution of C_4 genes, we are using the entry enzyme of the C_4 cycle PEPC as the model C_4 enzyme/gene and the dicot genus *Flaveria* (Asteraceae) (Powell, 1978) as the experimental system (Westhoff and Gowik, 2004).

The photosynthetic PEPCs of C_4 *Flaveria* species are encoded by the *ppcA* gene class whose orthologues are also found in C_3 and C_3-C_4 intermediate species of this genus (Hermans and Westhoff, 1992). Analysis of *ppcA* promoter/ß-glucuronidase reporter gene fusions in the C_4 plant *F. bidentis* revealed that the *ppcA* promoter of the C_4 plant *F. trinervia* directs a high expression of the reporter gene exclusively in the mesophyll cells. The orthologous *ppcA* promoter of the C_3 plant *F. pringlei*, however, is weak and does not show any apparent cell or organ specificity (Stockhaus et al., 1997). Combination of the proximal (-1 to -570) and the distal promoter region (-1566 to -2141) of the *ppcA* gene of *F. trinervia* is sufficient for a high mesophyll-specific expression of the reporter gene. While the proximal promoter region alone mediates a very low basal promoter activity with no cell specificity, the distal region confers mesophyll specificity and acts as an enhancer when fused to the proximal region. By dissection of the C_4 distal region a 41 bp module named MEM1 (mesophyll expression module 1) was identified that together with the C_4 proximal region is sufficient for mesophyll-specific reporter gene expression (Gowik, 2004).

MEM1 consists of two submodules, A and B, of 11 and 30 bp, respectively. The MEM1-module of the C_4 *ppcA* gene of *F. trinervia* is unique in that the A- and B-submodules are fused together with no intermediate sequence. In contrast, the A- and B-submodules of MEM1 of the closely related C_4 species *F. bidentis*, of the two C_4-like plants *F. palmerii* and *F. vaginata* and of the two C_3 species *F. cronquistii* and *F. pringlei* are separated by about 90 to 100 bp of intervening sequences (Fig. 1A). This insertion is of no significance for the mesophyll specificity of C_4 *ppcA* gene expression (M. Akyildiz *et al.,* 2007; paper submitted to Plant Cell for publication). Apart from that disparity the MEM1 of the two C_4 and two C_4-like plants and of the two C_3 plants differ only in two positions. The first nucleotide of the A-

submodule of the C_4 MEM1 is characterized by a guanine while the C_3 MEM1 holds an adenine at that position. The second difference concerns the insertion of a CACT tetranucleotide in the B-submodule of all C_4 and C_4-like MEM1 elements (Fig 1A). Analyses of *F. trinervia* MEM1/ß-glucuronidase reporter gene fusions in transgenic *F. bidentis* plants revealed that both MEM1-submodules have to be in the C_4-state in order to achieve mesophyll-specific reporter gene expression. The C_4 MEM1 functions as an enhancer of mesophyll expression and in addition as a repressor of gene expression in the bundle-sheath cells and the vascular bundles (M. Akyildiz *et al.,* 2007; paper submitted to Plant Cell for publication).

The present work was initiated to identify C_4 MEM1 interacting proteins that might be responsible for mediating mesophyll-specific expression. Hence we performed a yeast one-hybrid screen to identify such proteins. Here we report the identification of proteins of the basic leucine-zipper (bZIP) protein family which are highly similar to the *Arabidopsis thaliana* bZIP protein group I (Jakoby et al., 2002). We show that the identified FtbZIP proteins interact with the C_4 MEM1 but not with the C_3 MEM1, and that the FtbZIP-MEM1 interaction occurs irrespective whether the A- and B-submodules are separated. However, *in vitro* protein-DNA interaction studies by using electrophoretic mobility shift assays and fluorescence polarization/anisotropy measurements revealed that the FtbZIP proteins physically interact with a C_4 and C_3 MEM1 with no difference in binding affinity. This *in vitro* interaction behaviour of the FtbZIP proteins with MEM1 contrasts the *in vivo* situation in the yeast´s nucleus and suggests the involvement of additional factors in C_4 MEM1 function.

Results

One-Hybrid Screen with a C_4 MEM1

To identify proteins that interact with the MEM1 element of the C_4 *ppcA* gene of *F. trinervia* a yeast one-hybrid screen (Wang and Reed, 1993; Li and Herskowitz, 1993) was performed. In the initial screen, i.e. before MEM1 was confirmed as the final *cis*-regulatory element for mesophyll-specific gene expression, sub-fragments FtDRa and FtDRb of the distal promoter region (Fig. 1A) of the *ppcA* gene of *F. trinervia* should be used as bait sequences. Hence the FtDRa- and FtDRb-fragments were cloned independently into the pHIS-i and pLacZ-i reporter vectors. Both constructs were transformed sequentially into the yeast strain YM4271, resulting in the reporter strains FtDRa-*lacZ*/*His3* and FtDRb-*lacZ*/*His3*. A cDNA expression library from *F. trinervia* leaves in the vector pAD-GAL4 (Windhövel et al., 2001) was used as a prey in the yeast one-hybrid screen.

Screening of 3 x 10^{-6} (FtDRa) and 4 x 10^{-6} (FtDRb) yeast transformants resulted in the identification of a total of 46 yeast clones that showed an activated transcription of the *HIS3*- and *LacZ*-reporter genes driven by the FtDRa- and FtDRb-bait promoter elements. Re-transformation of the isolated prey plasmids of the 46 yeast clones into the FtDRa- and FtDRb- reporter strains resulted only in the case of two plasmids in the activation of both the *HIS3* and *LacZ* reporter genes, and hence confirmed the specificity of interaction. The cDNA inserts of all 46 pAD-GAL4 plasmids were sequenced, and the cDNA sequence of the two plasmids revealing a reporter gene activation after re-transformation turned out to represent two different sequences, both derived from the same gene. Blast searches indicated that both cDNA sequences are highly similar to the gene *At1g43700* of *Arabidopsis thaliana* which belongs to the group I of the bZIP proteins of *A. thaliana* (Jakoby et al., 2002). *At1g43700* encodes a basic leucine zipper (bZIP) protein named AtbZIP51. The identified gene of *F. trinervia* was correspondingly named FtbZIP51. The remaining 44 sequences turned out to be false-positives since they were not able to activate transcription of both the *HIS3* and *LacZ* reporter genes after re-transformation into the FtDRa- and FtDRb-*lacZ*/*His3* reporter strains.

A second and more extended screen, after identification of MEM1 as the mesophyll specificity element, was performed with the MEM1 sequence of the *ppcA* gene of *F. trinervia* as a bait (Fig. 1A). A total of 5,9 x 10^6 yeast transformants were screened resulting in the identification of 623 yeast colonies. In order to categorize the positive clones the cDNA-inserts of their pAD-GAL4 plasmids were amplified by PCR, spotted onto nylon membrane and subjected to sequential hybridizations. In the first round of hybridization both FtbZIP51

sequences, isolated in the first screen, were used as probes resulting in 15 strong hybridization signals. Sequence analysis showed that all cDNA sequences were derived from the FtbZIP51 gene. They could be assigned to eight different classes with respect to the size and location of the cDNA fragment within the structural gene. In the second round of hybridization the filters were hybridized with a mixture of the 44 false-positive sequences that were identified in the first one-hybrid screen in order to identify most of the false-positive clones. The cDNA inserts of 483 clones hybridized to this probe of false-positive sequences and were discarded. The cDNA inserts of the remaining 125 clones, that revealed a faint hybridization to the false-positive probe, were sequenced. Database searches suggested that 99 additional clones were false-positives. The cDNAs of all the remaining 26 clones turned out to encode bZIP proteins belonging to the group I of the bZIP proteins of *A. thaliana* (Fig. 3) (Jakoby et al., 2002). Twelve of the bZIP- sequences, representing four different cDNA inserts, were highly similar to the gene *At2g40620*, and fourteen sequences which could be assigned to five distinct cDNA classes matched to the gene *At4g38900*. *At2g40620* and *At4g38900* encode the bZIP proteins AtbZIP18 and AtbZIP29, respectively (Jakoby et al., 2002). Hence, the corresponding bZIP proteins of *F. trinervia* were named FtbZIP18 and FtbZIP29.

The plasmids of the identified 41 bZIP-clones were isolated and re-transformed into the reporter strain FtM-*LacZ/HIS3*. In all cases the FtM (C_4)-driven reporter genes were activated. Thus all 41 FtbZIP sequences interacted with the C_4 MEM1 of *F. trinervia* (Fig. 2B, data are shown only for one clone of FtbZIP18 [clone F6E2], FtbZIP29 [clone F1A8] and FtbZIP51 [clone F4F8]).

Interaction of FtbZIP Proteins with C_4- and C_3-Type MEM1 Elements

In planta transformation experiments had shown that the native C_4-type MEM1 elements (containing a guanine at the first nucleotide position in the A-submodule and a CACT tetranucleotide in the B-submodule, Fig. 1A and B) of *F. trinervia* (FtM) and *F. bidentis* (FbM) function as mesophyll specificity elements while the native C_3-type MEM1 of *F. pringlei* (FpM) (containing an adenine at the first nucleotide position in the A-submodule and lack the CACT tetranucleotide in the B-submodule, Fig. 1A and B) as well as the synthetic C_3-type version of FtM (FtM/A_ΔCACT, characterized by a G-to-A-exchange in the A-submodule and by the deletion of the CACT tetranucleotide in the B-submodule; Fig. 1B) do not show any cell specificity (M. Akyildiz *et al.,* 2007, submitted to Plant Cell for

publication). We therefore wanted to know whether the isolated FtbZIP proteins differentiated in the yeast system between C_4- and C_3-type MEM1 elements with regard to interaction. All 41 FtbZIP prey plasmids were therefore transformed into the FpM (C_3) reporter strain (Fig. 1B) and the yeast cells were assayed for histidine prototrophy and ß-galactosidase activity. None of the FtbZIP prey-sequences was able of activating the *HIS3* and *LacZ* reporter genes in the FpM (C_3) reporter strain to a detectable level. This result contrasts the situation observed in the FtM (C_4) reporter strain (see above), indicating that the identified FtbZIP protein fragments were not able to interact with the C_3 MEM1 of *F. pringlei* (Fig. 2B).

Isolation and Comparison of Complete Coding Sequences of FtbZIP18, 29 and 51

To analyze the interaction specificity of the FtbZIP18, 29 and 51 proteins with MEM1 more precisely and to confirm the phylogenetic relationships of these proteins, complete coding sequences were isolated from a cDNA library (see Materials and Methods). The coding sequences could be translated into 336 (FtbZIP18, clone F6E2), 591 (FtbZIP29, clone F1A8) and 334 (FtbZIP51, clone F4F8) amino acids with predicted molecular masses of 37 kDa (FtbZIP18 and 51) and 65 kDa (FtbZIP29), respectively. The comparison of their amino acid sequences with that of the corresponding bZIP proteins of *A. thaliana* (AtbZIP18, 29 and 51) showed that all six bZIP proteins are highly conserved in the basic and leucine zipper regions, and differ in the sequences located towards their amino-termini as well as their carboxy-termini (Fig. 3A).

The initial BLAST searches (see above) had indicated that the isolated FtbZIP proteins revealed the highest sequence similarity to the group I bZIP proteins AtbZIP18, AtbZIP29 and AtbZIP51, respectively. To confirm this assignment the complete coding sequences of FtbZIP18, 29 and 51 were aligned with all group I AtbZIP proteins, and the group S AtbZIP60 protein was used as an out-group. The dendrogram that was obtained by applying the neighbour-joining method as implemented in Clustal W (Thompson et al., 1994) confirms the assignment of the FtbZIP proteins as predicted from the BLAST searches (Fig. 3B).

Interaction of Full-Size FtbZIP Proteins with C_4- and C_3-Type MEM1 Elements

In order to study the interaction properties of the FtbZIP proteins to MEM1 in more detail the isolated full-size FtbZIP18 (clone F6E2), FtbZIP29 (clone F1A8) and FtbZIP51 (clone F4F8)

sequences were fused in frame at their carboxy-termini to the GAL4 activation domain. Transformation of these FtbZIP18, 29 and 51 prey plasmids into native C_4-type reporter strains, FtM and FbM (Fig. 1A and B), resulted in *HIS3* and *LacZ* reporter gene activity in FtM and FbM only in case of FtbZIP18 and 51, but not with FtbZIP29 (Fig. 2B). Activation of the *HIS3* gene in FtM was stronger than its activation in FbM as deduced from the growth characteristics of the yeast transformants (Table 2). This finding indicated that the full-length FtbZIP18 and 51 proteins, when fused to the GAL4 activation domain, were able to interact with a C_4-type MEM1, while the full-size FtbZIP29 fusion protein was not in contrast to its partial variants (see above). FtbZIP29 was therefore omitted from the further yeast one-hybrid interaction analysis. Yeast growth was remarkably stronger by the FtM driven reporter gene as compared to the FbM driven reporter gene (Table 2), indicating that an insertion between the A- and B-submodules of MEM1 (Fig. 1A and B) has a negative effect on interaction with FtbZIP18 and 51. An reduced yeast growth was also observed after transformation of FtbZIP18 and 51 into the synthetic C_4-type FpM driven reporter strain (FpM/G_+CACT, characterized by an A-to-G-exchange in the A-submodule and by the presence of the CACT tetranucleotide in the B-submodule) carrying the insertion of the MEM1 of *F. pringlei* between the A- and B-submodules (Fig. 1B and Table 2).

Transformation of the FtbZIP18 and 51 fusion plasmids into the native FpM (C_3) reporter strain revealed no substantial *HIS3* and *LacZ* reporter gene activity at all, indicating that the full-size FtbZIP18 and 51 proteins, just like their partial variants, were not able to interact with the C_3-type MEM1 of *F. pringlei* (Table 2). To verify this result FtbZIP18 and 51 were transformed into the synthetic C_3-type FtM reporter strain (FtM/A_ΔCACT) in which the A- and B-submodules are contiguous in contrast to the situation in the FpM reporter strain (Fig 1B). As to be expected, no activation of both the *HIS3* and *LacZ* reporter genes was detectable (Table 2). This indicates that the separation of the A- and B-submodules of MEM1 by an insertion is not responsible for the lack of interaction of FtbZIP18 and 51 with a C_3-type MEM1. It follows that FtbZIP18 and 51 interact with a C_4-type but not with a C_3-type MEM1, and that the intervening sequence between the A- and B-submodules, if present, has a minor influence on the interaction.

In vivo promoter-ß-glucuronidase reporter gene experiments in transgenic *F. bidentis* had shown that the C_3-C_4 chimerical versions of MEM1 do not result in mesophyll-specific expression of the reporter gene, but in an expression in the mesophyll cells, the bundle sheath and in the vascular bundles (M. Akyildiz *et al.,* 2007; paper submitted to Plant Cell for publication). This finding demonstrated that both submodules have to be in the C_4-state in

order to achieve mesophyll-specific gene expression. To test weather FtbZIP18 and 51 are able to interact with a C_3-C_4 chimerical MEM1 two additional synthetic reporter strains (FtM/A and FtM/_ΔCACT) were tested. Switch of either the A- or the B-MEM1 submodule of the C_4-type reporter strain FtM from the C_4 into the C_3 state (FtM/A and FtM/_ΔCACT) (Fig. 1B) did not abolish the interaction of FtbZIP18 and 51 with the chimerical MEM1 (Table 2). This suggests that the FtbZIP18 and 51 proteins are able of interacting with both the A- and the B-submodules of a C_4 MEM1 and that one submodule in the C_4-state is sufficient for FtbZIP-binding in yeast.

FtbZIP18 Interact *in vitro* with the MEM1 of both *F. trinervia* (C_4) and *F. pringlei* (C_3)

The yeast one-hybrid system allows to assay for the interaction of a putative *trans*-regulatory protein with its cognate DNA element *in vivo* in the heterologous context of the yeast nuclear environment. While a positive interaction of a putative *trans*-regulatory protein with its *cis*-regulatory element in the yeast one-hybrid assay is likely to involve binding of that protein to its recognition sequence, it cannot be excluded that yeast proteins interfere positively or negatively with this interaction. We therefore wanted to investigate whether the identified FtbZIP proteins physically interact with MEM1 and whether the C_4- and C_3-type MEM1 differ in their binding.

To assay directly for the physical binding of the FtbZIP proteins to MEM1 purified FtbZIP proteins (Fig. 4) were incubated with various MEM1 variants and the protein-DNA interaction was monitored by electrophoretic mobility shift assay measurements. Since FtbZIP18 exhibited the strongest interaction with C_4 MEM1 elements in the yeast one-hybrid assay, this protein was selected for studying its binding characteristics *in vitro.*

In the gel retardation assay FtbZIP18 protein binds to both a native C_4- (FtM) and a native C_3-type (FpM) MEM1 (Fig. 5A). In both cases the shifted protein-DNA complex becomes detectable at about the same protein concentration suggesting that FtbZIP18 binds to C_4- and C_3-type MEM1 elements with a similar affinity. This was confirmed by analyzing the physical binding of FtbZIP18 with another native C_4-type MEM1, FbM, and with the synthetic C_4-type version of FpM (FpM/G_+CACT) (Fig. 5B). To confirm the results of the gel retardation experiments and to estimate the degree of binding of FtbZIP18 to MEM1 fluorescence polarization/anisotropy measurements were carried out. This method allows to calculate dissociation constants of protein-DNA complexes under constitutions of equilibrium (Heyduk et al., 1996). To exclude the influence of an insertion between the A- and B-

submodules of MEM1 the C_4-type MEM1 of *F. trinervia* (FtM) was compared with its synthetic C_3-type version (FtM/A_ΔCACT; Fig. 1B). The binding reactions were titrated by successive addition of FtbZIP18 protein-solution, and the fluorescence anisotropy was measured for each titration point (Fig. 6). The calculated dissociation constants (K_d) for the interaction of FtbZIP18 with FtM (K_d = 0.92 µM) and FtM/A_ΔCACT (K_d = 0.73 µM) did not differ substantially indicating that FtbZIP18 binds to C_4- and C_3-type MEM1 elements with almost the same affinity. The results obtained by the gel retardation and fluorescence polarization/anisotropy measurements demonstrate that FtbZIP18 protein physically binds with an equal affinity to both C_4- and C_3-type MEM1 elements.

Discussion

Previous analyses had shown that the C_4 MEM1 *cis*-regulatory element of the C_4 *ppcA* gene of *F. trinervia* is sufficient for mesophyll-specific expression (M. Akyildiz *et al.,* 2007; paper submitted to Plant Cell for publication). We therefore searched for proteins interacting with this MEM1. In a yeast one-hybrid screen for *F. trinervia* MEM1 interacting proteins we identified three different bZIP proteins, FtbZIP18, 29 and 51, which are similar to the group I bZIP proteins of *A. thaliana* (Fig. 3) (Jakoby et al., 2002). In general, plant bZIP proteins are known to be involved in many processes that are critical to the function of the plant organism, e.g. seed maturation (Lara et al., 2003), pathogen defense (Ndamukong et al., 2007) and flower development (Abe et al., 2005). A well studied example of group I homolog bZIP proteins is RSG (for repression of shoot growth) from *N. tabacum*. The RSG gene is specifically expressed in the phloem and regulates one of the genes that encodes enzymes for the gibberellin biosynthesis (Fukazawa et al., 2000). However, in contrast to the RSG protein from *N. tabacum*, none of these class I bZIP proteins of *A. thaliana* has been characterized to date and their function remains unknown.

A common feature of the identified FtbZIP proteins and their related AtbZIP proteins is that they possess a lysine in the basic region at position -10 relative to the first leucine residue in the leucine zipper region (Fig. 4A). The characteristic lysine in the basic region of the I-type FtbZIP and AtbZIP proteins distinguishes them from other plant bZIP proteins. This lysine in the basic domain replaces the highly conserved arginine (Fukazawa et al., 2000). It is believed that this amino acid exchange might determine the specific binding site requirements of group I bZIP proteins because it correlates with a higher affinity to non-palindromic binding sites (Aukerman et al., 1991; Suckow et al., 1994; Ringli and Keller, 1998). The corresponding binding-site of the identified FtbZIP proteins in MEM1 is unknown. Based on the results of *in planta* transformation experiments in transgenic *F. bidentis* (M. Akyildiz *et al.,* 2007; paper submitted to Plant Cell for publication) and of the yeast one-hybrid experiments we assume that two binding-sites are located in MEM1, one in the A- and the other in the B-submodule of MEM1 (Fig. 1A). The CACT-containing DNA-binding sequence in the B-submodule of MEM1 of the two C_4 species *F. trinervia* and *F. bidentis* and of the two C_4-like plants *F. palmerii* and *F. vaginata* is embedded in a sequence context which forms an imperfect palindrome, and resembles the binding-site for a GCN4-like bZIP factor which represents a palindrome (Arndt and Fink, 1986; Gowik, 2004). The other supposed binding-site in the A-submodule is represented by the tetranucleotide GTGA which is reverse complementary to the TCAC-motif. This TCAC-motif is a component part

of the binding-sequence in the B-submodule of MEM1. These two hypothesized binding sequences are only found in the A- and B-MEM1 submodules of the C_4- and C_4-like plants and are absent in the MEM1 submodules of the two C_3 plants *F. cronquistii* and *F. pringlei* (Fig. 1A).

Depending on the amino acid sequence of the leucine zipper region, which is both necessary and sufficient for dimerization, the bZIP proteins generally functions as either homo- or heterodimers (Landschulz et al., 1988; Vinson et al., 1989; Deppmann et al., 2004). A comparative analysis of the *A. thaliana* bZIP motifs revealed that the FtbZIP homolog group I proteins AtbZIP18, AtbZIP29 and AtbZIP51 are predicted to form homo- as well as heterodimers with themselves and each other (Deppmann et al., 2004; Deppmann et al., 2006). An additional common characteristic of the identified FtbZIP proteins is that they possess ten ’heptads‘ (for definition see Fig. 3) in their leucine zipper region as well as the counterparts in *A. thaliana* (Fig. 4A), whereas leucine zippers of common bZIP proteins have only three to seven heptads (Landschulz et al., 1988). The proteins AtbZIP18/29/51 and FtbZIP18/29/51 share the same amino acid replacements in the leucine zipper. The seventh leucine is replaced by an methionine, except for AtbZIP51, and the leucine of the eighth heptad is replaced by an arginine (Fig. 3A). Whereas the replacement of the uncharged amino acid leucine by the non-polar and uncharged, respectively, amino acids methionine, isoleucine, valine and phenylalanine is well known (Jakoby et al., 2002), the purpose of the substitution of an leucine by the basic amino acid arginine in the eighth heptad is currently unknown.

Using the yeast one-hybrid system for protein-DNA interaction studies we showed that all prey proteins comprising partial sequences of FtbZIP18, 29 and 51 fused to the GAL4 activation domain interacted with the C_4 MEM1 bait sequence of *F. trinervia* and *F. bidentis* but not with the C_3 MEM1 of *F. pringlei* (Fig. 2B). The same differential interaction pattern was also observable with the prey proteins containing the corresponding complete coding region of FtbZIP18 and 51, whereas the full-size FtbZIP29 protein revealed no interaction with either C_4 and C_3 MEM1 bait sequences (Fig. 2B and Table 2). The proteins FtbZIP18 and 51 comprise of 336 and 334 amino acids whereas FtbZIP29 encodes a protein of 591 amino acids (Fig. 3A). Sequence analysis of the truncated versions of FtbZIP18, 29 and 51 showed that all three proteins contain a complete basic and leucine zipper region, and all lack amino acids only at the amino-terminus (FtbZIP18: bp 1 to 101; FtbZIP29: bp 1 to 140 and FtbZIP51: bp 1 to 40) (Fig. 2A). The absence of an interaction between MEM1 and the full-length protein FtbZIP29 may be due to a negative influence of the amino terminus, of which

140 bp are absent in the truncated version of protein FtbZIP29 (Fig. 2A). The interaction of the FtbZIP proteins with the C_4 MEM1 of *F. trinervia* (FtM) caused a strong growth of the yeast transformants, while their interaction with MEM1 containing an insertion between the A- and B-submodules, like the C_4 MEM1 of *F. bidentis* (FbM) and the synthetic C_4-type MEM1 of *F. pringlei* (FpM/G_+CACT), respectively, resulted in a reduced growth of the yeast transformants (Table 2). Quantitative measurements of promoter-ß-glucuronidase reporter genes in transgenic *F.* bidentis revealed that the presence of an insertion between the A- and B-submodules of MEM1 effected a two to three fold reduction of reporter gene activity, indicating that an insertion between the A- and B-submodules of MEM1 may interfere with the binding of a protein to MEM1 (M. Akyildiz *et al.,* 2007; paper submitted to Plant Cell for publication).

Subsequent *in vitro* protein-DNA interaction studies revealed that the interaction between purified FtbZIP18 protein and a C_4- and a C_3-type MEM1 exhibited almost no difference in binding affinity. The FtbZIP18 proteins showed no differential interaction with C_4- and C_3-type MEM1 elements as has been observed in the yeast system (Fig. 5A, B and 6). A possible explanation for the observed differential interaction of the FtbZIP proteins with a C_4- and a C_3-type MEM1 in yeast could be due to additional factors already present in the yeast's nucleus. This putative factor may interact with the FtbZIP proteins or with MEM1 and the FtbZIP proteins. The DNA-binding and/or the regulatory specificity of the identified FtbZIP proteins may be changed upon interaction with other proteins. A well investigated example is given by the mammalian Fos and Jun bZIP family proteins that function as dimeric transcription factors (Curran and Franza, 1988). Jun proteins form both homo- and heterodimers, whereas Fos proteins do not form homodimers and require heterodimerization to bind DNA (Abate et al., 1993; Kerppola and Curran, 1994a). Both, Jun/Jun and Fos/Jun, bind to AP-1 regulatory elements, however, upon heterodimerization between Fos, Jun and members of the MAF group of bZIP proteins they are targeted to composite regulatory element sites consisting of AP-1 and Maf half-sites (Kerppola and Curran, 1994b). The interaction of Fos-Jun with structurally unrelated DNA-binding proteins can further increase the combinatorial potential of Fos-Jun dependent transcription regulation (Bassuk and Leiden, 1995; Zhang et al., 1998; Chen et al., 1998). Another possibility of modulating the DNA-binding and *trans*-activation specificity of bZIP proteins might be mediated through the interaction with non-transcription factor proteins like members of the high-mobility group (Wissmuller et al., 2006), glutaredoxin family (Ndamukong et al., 2007) and with protein kinases (Choi et al., 2005), (Ahn et al., 2006). A well characterized example in *A. thaliana* is

presented by the bZIP protein FD that binds to the promoters of floral genes. FD interacts with *FT* (*Flowering Locus T*) and *TFL1* (*Terminal Flower1*), which reveal homologies to human Raf kinase inhibitor proteins. Whereas the interaction of FD with FT induces flowering, its interaction with TFL1 represses flowering (Ahn et al., 2006).

Based on the finding that the MEM1 *cis*-regulatory element functions as an enhancer of *ppcA* gene expression in mesophyll cells and concomitantly as a repressor of *ppcA* gene expression in bundle sheath cells and vascular tissues of *Flaveria* (Akyildiz *et al.,* submitted to Plant Cell for publication) there are multitude of possibilities by which the identified FtbZIP proteins may be involved in the regulation of this mesophyll-specific expression of the *ppcA* gene in *Flaveria*. We know from the *in vitro* experiments that the protein FtbZIP18 physically binds to a MEM1 element of either a C_4- and a C_3-type (Fig. 5A, B and 6), but for the differentiation between the two elements *in vivo* they demand the involvement of another factor. The interaction of this unknown factor expands the functional versatility of the FtbZIP proteins, it brings about the required specificity to differentiate between a C_4- and a C_3-type MEM1.

The data presented here indicate that the identified FtbZIP proteins bind *in vitro* at the MEM1 *cis*-regulatory element of either a C_4- and a C_3-type (Fig. 5A, B and 6), while *in vivo* they differentiate between these two MEM1 types (Fig. 2 and Table2). The C_3- and C_4-type MEM1 differ only in two positions, an adenine (C_3) to guanine (C_4) exchange and the absence (C_3) /presence (C_4) of the tetranucleotide CACT (Fig. 1A). *In planta* transformation experiments had confirmed the relevance of the C_4-type MEM1 element as a mesophyll specificity element, while the C_3-type MEM1 revealed no cell specificity (M. Akyildiz *et al.,* 2007; paper submitted to Plant Cell for publication). It is supposed that this differential binding of the FtbZIP proteins at MEM1 *in vivo* is dependent on the interaction between the FtbZIP proteins and an unknown factor. It has to be detected in a yeast two-hybrid assay with which factor(s) the FtbZIP proteins interact (Bartel and Fields, 1995; Fields and Bartel, 2001). However, it remains an open question if the FtbZIP proteins are the relevant proteins involved in the mesophyll-specific expression of the *ppcA* gene of *Flaveria*. One can not exclude that another MEM1-binding protein was not detected by the yeast one-hybrid screen because the protein was toxic for the yeast cells. For this reason the *in planta* relevance of the identified FtbZIP proteins towards their participation in C_4 *ppcA* gene expression had to be examined. For this purpose reverse genetic approaches based on RNAi-triggered gene inactivation (Waterhouse and Helliwell, 2003) or gain-of-function experiments which ectopically express the putative transcription factor (Schwechheimer et al., 1998) are a good

choice to answer this question. In addition *in situ* hybridization experiments (Uhl, 1989) should display where the mRNA of the proposed transcription factors are expressed.

Materials and Methods

DNA manipulations were carried out according to Sambrook and Russell (Sambrook and Russell, 2001). All DNA fragments created by PCR were confirmed by DNA sequencing.

Generation of FtDRa-, FtDRb- and MEM1-Containing Promoter-Reporter Constructs

The distal region of the *ppcA*1 promoter of *F. trinervia* was divided into three overlapping segments FtDRa (-2141 to -1940), FtDRb (-1981 to -1668) and FtDRc (-1713 to -1566) (Fig. 1A). FtDRa and FtDRb were used to replace the distal region in the ppcA-PR_{Ft}-$DR(+)_{Ft}$ construct (constructs ppcA-PR_{Ft}-DRa_{Ft} and ppcA-PR_{Ft}-DRb_{Ft}) as described in Gowik *et al*. (2004). For the generation of the FtDRa- and FtDRb- containing promoter reporter constructs a PCR with primers depicted in Table 1A and constructs ppcA-PR_{Ft}-DRa_{Ft} and ppcA-PR_{Ft}-DRb_{Ft} as template was performed.

The various MEM1 variants (Fig. 1B) were either constructed by assembling synthetic oligonucleotides (1) or by PCR (2) using genomic DNA of *F. trinervia* as template (Table 1B). The FtDRa, FtDRb and MEM1 sequences were designed to contain an *Eco*RI site at their 5' termini and a *Xba*I or *Xho*I site, respectively, at their 3' termini. FtDRa-, FtDRb- and MEM1-fragments were inserted into the *Eco*RI/*Xba*I digested pHIS-i and the *Eco*RI/*Xho*I digested pLacZ-i reporter vectors (Clontech, Mountain View, California).

In order to suppress excessive basal *HIS3* gene expression in case of the MEM1-containing reporter vector pHIS-i, a 200 bp spacer fragment was inserted into the *Xba*I restriction site, 3' behind the MEM1 fragment. The spacer segment was taken from the coding region of *Hcf136* which encodes a hydrophilic assembly factor of photosystem II in *Arabidopsis thaliana* (Meurer et al., 1998). For the generation of the spacer a PCR reaction with the primers HCF5'-XbaI (5'GGGTCTAGACTCTGTTTCTCCTCG3')/ HCF3'-XbaI (5'GGGTCTAGAACTGTTCATCAGCTC3') and *HCF136* cDNA (Meurer et al., 1998) as template was performed. After digestion with *Xba*I the resulting HCF-spacer was inserted into *Xba*I digested pHIS-i plasmid.

Construction of a cDNA Library

Total RNA was isolated as described in Logemann *et al*. (1987) (Logemann et al., 1987), and poly(A)$^+$ RNA was purified using Oligotex beads (Qiagen, Hilden, Germany). The cDNA library was constructed from poly(A)$^+$-RNA isolated from leaves of 2 month old *F. trinervia*.

The cDNAs were synthesized following the instruction manual of the pBluescript® II XR cDNA Library Construction Kit (Stratagene, La Jolla, California, USA). The synthesized cDNA was size-fractionated on a 1,2% (w/v) agarose gel, and the cDNA fraction from 0,4 to 4 kb was isolated with the Min Elute Gel Extraction Kit (Qiagen, Hilden, Germany). The cDNAs were ligated into *Eco*RI/*Xho*I digested pBIISK(+) vector (Stratagene, La Jolla, California, USA) and the cDNA-plasmids were transformed into XL10-Gold ultracompetent *E. coli* cells (Stratagene, La Jolla, California, USA). A total of about 300.000 independently transformed *E. coli* cells were plated on Luria Broth (LB) agar (supplemented with 100 µg ampicillin/ml) to a density of about 1200 colonies per plate (12 cm x 12 cm). Colonies were washed off from the plates with 10 ml per plate of LB medium resulting in 192 pools of *E. coli* cDNA clones. Aliquots from each of the 192 cDNA library pools were supplemented with glycerol (30%, v/v) and stored at –80°C until use. The remainder of each pool was independently used for plasmid isolation (Holmes and Quigley, 1981). Plasmid DNAs were dissolved in 100 µl TE-buffer (10 mM Tris-HCl, 1 mM EDTA, pH: 8.0) and stored in 96er micro-plates at –20°C.

Yeast One-Hybrid Screening

The dual yeast one-hybrid reporter strains were generated by the integration of the reporter vectors pHIS-i and pLacZ-i independently into the genomic DNA of the same yeast host strain YM4271 (Clontech, Mountain View, California) at the *his3*- or *ura3*- loci, respectively, *via* homologous recombination (Liu et al., 1993). Integration of pLacZ-i into the *ura3*-locus confers an constitutive Ura^+ phenotype to the transformants, while integration of pHIS-i into the *his3*-locus results in a $His3^+$ phenotype only if the bait *cis*-regulatory element in front of the *HIS3* gene interacts with a prey protein. The plasmids pHIS-i and pLacZ-i (Alexandre et al., 1993) were first linearized by restriction, pHIS-i with *Xho*I and pLacZ-i with *Nco*I, following transformation into the host strain YM4271 through a lithium acetate-based method as described by Gietz and Woods (2001, 2002). The resulting MEM1-*lacZ*/-*His3* reporter strains were verified by PCR.

Putative MEM1-interacting proteins were isolated by inserting a cDNA library of *F. trinervia* into the reporter vector pAD-GAL4 (Windhövel et al., 2001). In this vector the cDNA fragments are fused at their carboxy-termini to the activation domain of the yeast transcription factor GAL4, and expression of the fusion proteins is driven by the yeast alcohol dehydrogenase (ADH1) promoter (Giniger et al., 1985). Yeast cells transformed with the

reporter vector pAD-GAL4 (Clontech, Mountain View, California) are capable of growing on minimal medium lacking leucine. For the analysis of MEM1-driven promoter activation, the reporter strains were plated on minimal medium lacking histidine and leucine. The leaky expression of *HIS3* was controlled by adding 3-amino-1,2,4-triazole to the medium. The optimal concentration of 3-amino-1,2,4-triazole was determined for each reporter strain before the screen as described in Meijer *et al.* (1998). In case of the FbM reporter strain the minimal medium was supplemented with 15 mM 3-amino-1,2,4-triazole, in the case of all other reporter strains 50 mM 3-amino-1,2,4-triazole were necessary to prevent basal *HIS3* activity.

ß-Galactosidase Colony-Lift Filter Assay

The putative positive yeast clones were grown for three days at 30°C and lifted onto Whatman filter membranes (# 1, Whatman Int. Ltd., Maidstone England). Cells were lysed by freezing in liquid nitrogen for 20 sec., thawed at room temperature for several minutes, and placed onto filters that were pre-soaked with Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM ß-mercaptoethanol) containing X-gal (5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside, 1 mg/ml) as substrate (Breeden and Nasmyth, 1985).

Hybridization Analysis of the Yeast One-Hybrid Library

To classify positive clones from the various primary yeast one-hybrid screens and to prevent the sequencing of all cDNA inserts from the positive clones, sequencing was serially combined with hybridization analysis of membrane-arranged cDNA inserts.

The cDNA-inserts of the pAD-GAL4 plasmids of all positive yeast clones were amplified by PCR with the primers T7 (5'CAGGATGCACAGTTGAAGTG3') and pGAD (5'TGAAGATACCCCACCAAAC3'), and the resulting PCR products were spotted by a multi-blot-replicator (96 slot pins, VP46852, V&P Scientific Inc. San Diego, USA) onto Hybond N^+ membranes (Amersham Biosciences, Little Chalfont, UK). The spotted DNAs were fixed to the membranes by baking for two hours at 80°C. Hybridizations were carried out for 16 h at 64°C in 7% (w/v) SDS, 2.5 mM EDTA, 250 mM Na_2HPO_4, pH 7.2 with $\alpha[^{32}P]$-labelled probes (Megaprime™ DNA labelling systems, Amersham Biosciences, Little Chalfont, UK). After hybridization, the filters were washed at the same temperature for 15 min in 1 x SSC/1% (w/v) SDS and twice each in 0.5 x SSC/1% (w/v) SDS and 0.1 x

SSC/0.5% (w/v) SDS [20 x SSC: 3 M NaCl, 0.3 M NaCitrate, pH 7.4]. Hybridization signals were visualized by phosphor-imaging using a BioImager (BioImaging Analyzer BAS-1800, program: Image Reader V 1.4E [Fuji Photo Film Co., Ltd.]). Signals were quantified using the program Image Gauge 3.0 (Fuji Photo Film Co., Ltd. Düsseldorf).

Isolation of Full-Length FtbZIP cDNAs

To unequivocally distinguish between FtbZIP18, 29 and 51 sequences primers were designed (18-1/-2, 29-1/-4 and 51-1/-2, Table 1C) that allowed a specific PCR amplification of the three FtbZIP sequences from the cDNA library. All pools of the cDNA library were screened by PCR using these primers for the presence of FtbZIP18, 29 and 51 sequences. Once a pool had been identified as containing FtbZIP cDNA, the length of the respective FtbZIP 5' cDNA part was estimated by PCR amplification. The T3 primer (5' AATTAACCCTCACTAAAGGG 3') was used as the universal primer for directing DNA synthesis from the 5' end of the cDNA insert into the coding part and specific primers for each FtbZIP sequence (Table 1C, primers 18-2, 29-4 and 51-2) to drive synthesis outwards the FtbZIP coding part towards the 5' end. Pools with putative full-size FtbZIP cDNAs were selected for isolating the respective cDNA clone by colony hybridization (Sambrook and Russell, 2001). The putative full-size FtbZIP cDNAs were sequenced in order to confirm that they contain a complete coding region frame.

Expression of Proteins in *Escherichia coli*

The coding sequences of FtbZIP18, FtbZIP29 and FtbZIP51 (from start- (ATG) to stop-codon (TAA [FtbZIP18]/TAG [FtbZIP29]/TGA [FtbZIP51])) were amplified with the respective primers 18pET5'EcoRI/18pET3'NotI, 29pET5'EcoRI/29pET3'NotI and 51pET5'EcoRI/ 51pET3'NotI (Table 1C) using FtbZIP18, 29 and 51 full-length cDNA clones as template. The respective PCR products were digested with *Eco*RI/*Not*I (FtbZIP18 and 51) and *Bam*HI/*Not*I (FtbZIP29) and cloned in frame into the appropriate digested pET-21d (+) vector (Novagen, Merck KGaA, Darmstadt, Germany), which expresses proteins with an N-terminal T7-Tag and a C-terminal His_6-Tag. *E. coli* BL21 (DE3) pLys-S (Novagen, Merck KGaA, Darmstadt, Germany) were used as a host (Studier et al., 1990).

For large scale preparations eight liter of LB-medium (100 µg/µl ampicillin) were inoculated with an overnight culture of the appropriate plasmid in BL21 (DE3) pLys-S to an

OD_{600} of 0.02. The culture was incubated at 37°C until an OD_{600} of 0.6 was reached. Then IPTG (isopropyl-thio-ß-D-galactopyranosid) was added to 0.5 mM, and incubation was continued at 20°C for four to five hours. Cells were collected by centrifugation at 4°C and cleared lysates of the bacterial cells were prepared under denaturing conditions (8 M urea) following the protocol of Qiagen (The QIAexpressionist 06/2003, protocol10) (Qiagen, Hilden, Germany). The His-tagged proteins were purified on a Ni-nitrilotriacetic acid resin according to the manufacturers instructions (The QIAexpressionist 06/2003, protocol17).

Proteins were refolded by removing the denaturant by dialysis. In detail the protein solution was transferred into dialysis tubing (SnakeSkin®, MWCO 10 kDa; Pierce, Rockford, Illinois, USA) and dialysed over night at 4°C against a 200 fold volume of 20 mM Hepes-KOH, 100 mM KCl, 0.1 mM EDTA and 10 % glycerin (v/v), pH 7.8. The dialysis buffer was exchanged six times. Following dialysis the protein-solution was stored at –80°C until use. Protein concentrations were determined by the Bradford assay (Bio-Rad, Hercules/California, USA) (Bradford, 1976). The integrity of the proteins was checked by SDS-polyacrylamide gel electrophoresis according to Laemmli (1970) (Laemmli, 1970). An aliquot of the FtbZIP proteins were precipitated with 15% (v/v) trichloroacetic acid for 30 min on ice following centrifugation for 10 min at 10.000 Upm and 4°C to pellet precipitated FtbZIP protein. The FtbZIP proteins were washed twice with 500 µl ice-cold acetone (80%). After drying for 5-10 min in a 95°C heat block to drive off acetone the protein was dissolved in 0.1 M Na_2CO_3, 10% (w/v) sucrose and 50 mM DTT and incubated on ice for 30 min.

Protein-DNA Interaction Studies in Yeast

For *in vivo* protein-MEM1 interaction studies in yeast the full-length FtbZIP18-, FtbZIP29- and FtbZIP51-cDNAs were cloned into the yeast reporter vector pAD-GAL4 in front of the sequence encoding the activation domain of the GAL4 transcription factor (Guthrie and Fink, 1991). The FtbZIP coding sequences were PCR-amplified by using the primers 18pAD5'EcoRI/18pAD3'XhoI, 29pAD5'BamHI/29pAD3'XhoI and 51pAD5'EcoRI/51pAD3'XhoI (Table 1C). The resulting PCR fragments were digested with *Eco*RI/*Xho*I (FtbZIP18/51) and *Bam*HI/*Xho*I (FtbZIP29) and inserted into *Eco*RI/*Xho*I and *Bam*HI/*Xho*I digested pAD-GAL4. The fusion-proteins FtbZIP18, 29 and 51 consist of the respective protein fused to the activation domain of the yeast transcription factor GAL4 (Guthrie and Fink, 1991) at their carboxy-termini. The FtbZIP pAD-GAL4 plasmids were then transformed into the various MEM1 promoter strains (Fig. 1B).

Generation of MEM1 Promoter Variants for *In Vitro* Protein-DNA Interaction Studies

DNA fragments used for electrophoretic mobility shift assays and fluorescence polarization/anisotropy measurements were labelled with the fluorophor carboxytetramethyl-rhodamine (TAMRA; Fa. Purimex, Grebenstein, Germany). The fluorophor was linked covalently to the 5' end of the oligonucleotides and primers *via* a C6-aminolinker (Purimex, Grebenstein, Germany). The various MEM1 variants (Fig. 1B and Table 1D) were either constructed by assembling synthetic oligonucleotides (1) or by PCR (2) using cloned promoter fragments as templates. The resulting fluorophor-labelled DNA fragments were purified by electrophoresis in 5% polyacrylamide gels (acrylamide: bisacrylamide ratio 46:1) prepared in TBE buffer (90 mM Tris, 90 mM borate, 1 mM EDTA, pH 8.3) followed by elution from the gel slices in 300 mM NaCl, 30 mM TrisCl pH 8.0 and 0.3 mM EDTA overnight. After a phenol/chloroform extraction the TAMRA-labelled DNA fragments were precipitated with 3 vol. ethanol, dissolved in double-distilled water and stored at -20°C until use.

Electrophoretic Mobility Shift DNA-Binding Experiments

Binding reactions (30 µl) contained 500 fmol of fluorophor-labelled DNA probe, 20 mM Hepes-KOH pH 7.8, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT and 10% (v/v) glycerol. In order to reduce non-specific binding of proteins to the labelled DNA fragments, the nucleic acid polymer poly(dIdC)-poly(dIdC) was added to the binding assays. The required amount of poly(dIdC) was determined empirically for each DNA-binding activity, and the optimum amount was found to be 0.3 µg/µl. Reactions were started by addition of protein, incubated at room temperature for 60 min and loaded on pre-electrophoresed 5% polyacrylamide gels (acrylamide : bisacrylamide ratio 46:1) prepared with TGE buffer (25 mM Tris, 0.19 mM glycine, 4 mM $MgCl_2$, 1 mM EDTA, ph 8.7). The gels were supplemented with 0.5 mM DTT and 3% (w/v) glycerol in order to stabilize protein-DNA interactions. Electrophoresis was carried out for four hours at 10 V/cm. Gels were laid on a Fluor-stage (Fuji Photo Film Co., Ltd. Düsseldorf) and visualized by fluorescence imaging using a BioImager (BioImaging Analyzer BAS-1800, program: Image Reader V 1.4E [Fuji Photo Film Co., Ltd.]) at λ_{ex} of 532 nm. Signals were quantified using the program Image Gauge 3.0 (Fuji Photo Film Co., Ltd. Düsseldorf).

Fluorescence Anisotropy Analysis and Estimation of Dissociation Constants

Fluorescence anisotropy/Fluorescence polarization is a spectroscopic technique that can be used to monitor the assembly of transcription factors with DNA molecules (Lundblad et al., 1996). It is performed in solution and can provide a true equilibrium measurement of binding (Lakowicz, 1980). If a fluorescent molecule is excited with polarized light, it will emit light of the same polarization assuming the molecule does not rotate during its emission lifetime. Depolarization, or decrease in light being emitted in the same polarization direction, occurs when the molecule rotates during its emission lifetime (Lundblad et al., 1996; Nasir and Jolley, 1999). If the fluorophore-labelled DNA fragment (D) is not bound by any protein (P), i. e. the molecule is small, it will rotate rapidly, thus the value of anisotropy will be low. But if P binds with a high affinity to D, i. e. the molecule is large, the tumbling rate of D is reduced and the anisotropy increases.

Fluorescence polarization/anisotropy measurements were performed with a Fluorolog-3 spectrofluorometer (HORIBA Jobin Yvon Inc., Edison, New Jersey, USA) and with the following settings: emission acquisition: scan start/end: 580 nm/590 nm; increment: 1 nm; integration time:4 s/nm; excitation wavelength: 555 nm; slits: excitation (5.0 nm)/ emission (8.0 nm). All measurements were taken in a 2x10 mm quartz cuvette at room temperature. The instrumental *G*-factor was measured before each experiment with a DNA sample. The initial concentration of labelled DNA was 50 nM in 250 µl of protein buffer (25 mM Hepes-KOH pH 7.8, 100 mM KCl; 0.1 mM EDTA; 10% (v/v) glycerol) in the presence of BSA (0.2 µg/µl) to decrease the amount of unspecific FtbZIP protein binding to MEM1-DNA. After the blank was read, protein solutions were added to DNA samples in 2 µl to 65 µl steps with a starting protein concentration of 0.0025 µM to a final concentration of 3.6 µM. Anisotropy values represent average values over the specified wavelength region (580 to 590 nm). Each anisotropy value was measured twice and the average was taken. Fluorescence intensities were also recorded and corrected for dilution. The total fluorescence intensity showed no noticeable dependence on the protein concentration.

Dissociation constants (K_d) were calculated based on the models and equations of Riggs *et al.* (1970) (Riggs et al., 1970). If a protein-DNA complex is in equilibrium with the free protein (P_f) and DNA (D_f) the dissociation constant (K_d) is defined by equation 1:

$$K_d = \frac{[P_f]\ [D_f]}{[PD]} \qquad (1)$$

where $[P_f]$, $[D_f]$ and [PD] represent the concentrations of free protein, free DNA and of the protein-DNA complex, respectively. The concentration of free protein and free DNA are given by the following equations:

$$[P_f] = [P]_0 - [PD] \quad (2)$$

$$[D_f] = [D]_0 - [PD] \quad (3)$$

where the subscript "0" refers to the total (added) concentration. The concentration of the protein-DNA complex is then given by equation 4:

$$[DP] = \frac{1}{2}\left([D]_0 + [P]_0 + K_d - \sqrt{([D]_0 + [P]_0 + K_d)^2 - 4 \quad x \quad [D]_0 \quad x \quad [P]_0}\right) \quad (4)$$

If it is assumed that one protein molecule binds to the DNA the data are fitted by using the following hyperbolic model:

$$r = r_D + \frac{[PD]}{[D]_0}(r_{DP} - r_D) \quad (5)$$

where r is the measured fluorescence anisotropy, and r_D and r_{PD} denote the fluorescence anisotropy of labelled DNA in the absence of protein, and the anisotropy of the protein-DNA complex, respectively.

Computer Analyses

DNA and protein sequence analyses were performed with MacMolly Tetra (Schoeneberg et al., 1994). Database searches were carried out with the Blast program at the National Center for Biotechnology Information (NCBI) (Altschul et al., 1997), Munich Information Center for Protein Sequences (MIPS) (Altschul et al., 1994) and Arabidopsis Information Resource (TAIR)- (Garcia-Hernandez et al., 2002). The sequence alignments and phylogenetic trees were created with the program Clustal W 1.8 (Thompson et al., 1994).

Acknowledgments

Work described here was supported by the Deutsche Forschungsgemeinschaft (DFG) through the Sonderforschungsbereich 590 at the Heinrich Heine University of Düsseldorf. We are indebted to K. Ernst and U. Gowik for carefully reading the manuscript.

Footnotes

†To whom correspondence should be addressed at: Heinrich-Heine Universität, Institut für Entwicklungs- und Molekularbiologie der Pflanzen, 40225 Düsseldorf, Germany. E-mail: west@uni-duesseldorf.de

Abbreviations

bZIP, basic leucine zipper; MEM1, mesophyll expression module 1; PEPC, phosphoeno*l*pyruvate carboxylase; Ft, *Flaveria trinervia*; Fp, *Flaveria pringlei*; Fb, *Flaveria bidentis*; DR, distal region; PR, proximal region]

References

Abate, C., Baker, S.J., Lees-Miller, S.P., Anderson, C.W., Marshak, D.R., and Curran, T. (1993). Dimerization and DNA binding alter phosphorylation of Fos and Jun. Proc Natl. Acad. Sci. USA **90,** 6766-6770.

Abe, M., Kobayashi, T., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., Ichinoki, H., Notaguchi, M., Goto, K., and Araki, T. (2005). FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. Science **309,** 1052-1056.

Ahn, J.H., Miller, D., Winter, V.J., Banfield, M.J., Lee, J.H., Yoo, S.Y., Henz, S.R., Brady, R.L., and Weigel, D. (2006). A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. Embo J. **25,** 605-614.

Alexandre, C., Grueneberg, D.A., and Gilman, M.Z. (1993). Studying heterologous transcription factors in yeast. Methods Enzymol. **5,** 147-155.

Altschul, S.F., Boguski, M.S., Gish, W., and Wootton, J.C. (1994). Issues in searching molecular sequence databases. Nat. Genet .**6,** 119-129.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J.H., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acid. Res. **25,** 3389-3402.

Arndt, K., and Fink, G.R. (1986). GCN4 protein, a positive transcription factor in yeast, binds general promoters at all 5' TGACTC 3' sequences. Proc. Nat. Acad. Sci. USA **83,** 8516-8520.

Aukerman, M.J., Schmidt, R.J., Burr, B., and Burr, F.A. (1991). An arginine to lysine substitution in the bZIP domain of an *opaque-2* mutant in maize abolishes specific DNA binding. Genes. Dev. **5,** 310-320.

Bartel, P.L., and Fields, S. (1995). Analyzing protein-protein interactions using two-hybrid system. Methods. Enzymol. **254,** 241-263.

Bassuk, A.G., and Leiden, J.M. (1995). A direct physical association between ETS and AP-1 transcription factors in normal human T cells. Immunity **3,** 223-237.

Berry, J.O., Nikolau, B.J., Carr, J.P., and Klessig, D.F. (1986). Translational regulation of light-induced ribulose 1,5-bisphosphate carboxylase gene expression in amaranth. Mol. Cell. Biol. **6,** 2347-2353.

Bradford, M.M. (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.

Breeden, L., and Nasmyth, K. (1985). Regulation of the yeast HO gene. Cold Spring Harb. Symp. Quant. Biol. **50,** 643-650.

Brutnell, T.P., Sawers, R.J.H., Mant, A., and Langdale, J.A. (1999). Bundle sheath defective2, a novel protein required for post-translational regulation of the *rbcL* gene of maize. Plant Cell **11,** 849-864.

Chen, L., Glover, J.N., Hogan, P.G., Rao, A., and Harrison, S.C. (1998). Structure of the DNA-binding domains from NFAT, Fos and Jun bound specifically to DNA. Nature **392,** 42-48.

Choi, H.-I., Park, H.-J., Park, J.-H., Kim, S., Im, M.-Y., Seo, H.-H., Kim, Y.-W., Hwang, I., and Kim, S.-Y. (2005). Arabidopsis calcium-dependent protein kinase AtCPK32 interacts with ABF4, a transcriptional regulator of abscisic acid-responsive gene expression, and modulates its activity. Plant Physiol. **139,** 1750-1761.

Curran, T., and Franza, B.R., Jr. (1988). Fos and Jun: the AP-1 connection. Cell **55,** 395-397.

Deppmann, C.D., Alvania, R.S., and Taparowsky, E.J. (2006). Cross-species annotation of basic leucine zipper factor interactions. Mol. Biol. Evol. **23,** 1480-1492.

Deppmann, C.D., Acharya, A., Rishi, V., Wobbes, B., Smeekens, S., Taparowsky, E.J., and Vinson, C. (2004). Dimerization specificity of all 67 B-ZIP motifs in Arabidopsis thaliana: a comparison to Homo sapiens B-ZIP motifs. Nucl. Acid. Res. **32,** 3435-3445.

Edwards, G.E., and Walker, D.A. (1983). C_3, C_4: Mechanism, and Cellular and Environmental Regulation, of Photosynthesis. (Oxford-London: Blackwell Scientific Publications).

Ehleringer, J.R., and Monson, R.K. (1993). Evolutionary and ecological aspects of photosynthetic pathway variation. Annu. Rev. Ecol. Syst. **24,** 411-439.

Ehleringer, J.R., Cerling, T.E., and Helliker, B.R. (1997). C_4 photosynthesis, atmospheric CO_2, and climate. Oecologia **112,** 285-299.

Fields, S., and Bartel, P.L. (2001). The two-hybrid system. A personal view. Methods. Mol. Biol. **177,** 3-8.

Fukazawa, J., Sakai, T., Ishida, S., Yamaguchi, I., Kamiya, Y., and Takahashi, Y. (2000). Repression of shoot growth, a bZIP transcriptional activator, regulates cell elongation by controlling the level of gibberellins. Plant Cell **12,** 901-915.

Garcia-Hernandez, M., Berardini, T.Z., Chen, G., Crist, D., Doyle, A., Huala, E., Knee, E., Lambrecht, M., Miller, N., Mueller, L.A., Mundodi, S., Reiser, L., Rhee, S.Y., Scholl, R., Tacklind, J., Weems, D.C., Wu, Y., Xu, I., Yoo, D., Yoon, J., and Zhang, P. (2002). TAIR: a resource for integrated Arabidopsis data. Funct. Integr. Genomics **2,** 239-253.

Gietz, R.D., and Woods, R.A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/PEG method. Methods. Enzymol. **350,** 87-96.

Giniger, E., Varnum, S.M., and Ptashne, M. (1985). Specific DNA binding of GAL4, a positive regulatory protein of yeast. Cell **40,** 767-774.

Gowik, U., Burscheidt, J., Akyildiz, M., Schlue, U., Koczor, M., Streubel, M., Westhoff, P. (2004). cis-regulatory elements for mesophyll-specific gene expression in the C4 plant Flaveria trinervia, the promoter of the C4 phosphoenolpyruvate carboxylase gene. Plant Cell **16,** 1077-1090.

Guthrie, C., and Fink, G.R. (1991). Guide to yeast genetics and molecular biology. Methods. Enzymol. **194,** 1-932.

Hatch, M.D. (1987). C_4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. Biochim. Biophys. Acta **895,** 81-106.

Hatch, M.D., and Osmond, C.B. (1976). Compartmentation and transport in C4 photosynthesis. In Transport in Plants. III. Intracellular Interactions and Transport Processes. Encyclopedia of Plant Physiol, Vol. 3, C.R. Stocking and U. Heber, eds (Berlin: Springer Verlag), pp. 144-184.

Hatch, M.D., and Oliver, I.R. (1978). Activation and inactivation of phosphoenolpyruvate carboxylase (EC-4.1.1.31) in leaf extracts from C4 species. Aust. J. Plant. Physiol. **5,** 571-580.

Hermans, J., and Westhoff, P. (1992). Homologous genes for the C_4 isoform of phosphoenolpyruvate carboxylase in a C_3- and a C_4-*Flaveria* species. Mol. Gen. Genet. **234,** 275-284.

Heyduk, T., Ma, Y., Tang, H., and Ebright, R.H. (1996). Fluorescence anisotropy: rapid, quantitative assay for protein-DNA and protein-protein interaction. Methods. Enzymol. **274,** 492-503.

Holmes, D.S., and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114,** 193-197.

Jakoby, M., Weisshaar, B., Dröge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., and Parcy, F. (2002). bZIP transcription factors in Arabidopsis. Trends Plant Sci. **7,** 106-111.

Kellogg, E.A. (1999). Phylogenetic aspects of the evolution of C4 photosynthesis. In C4 Plant Biology, R.F. Sage and R.K. Monson, eds (San Diego: Academic Press), pp. 411-444.

Kerppola, T.K., and Curran, T. (1994a). Maf and Nrl can bind to AP-1 sites and form heterodimers with Fos and Jun. Oncogene **9,** 675-684.

Kerppola, T.K., and Curran, T. (1994b). A conserved region adjacent to the basic domain is required for recognition of an extended DNA binding site by Maf/Nrl family proteins. Oncogene **9,** 3149-3158.

Kubicki, A., Steinmüller, K., and Westhoff, P. (1994). Differential transcription of plastome-encoded genes in the mesophyll and bundle-sheath chloroplasts of the monocotyledonous NADP-malic enzyme type C4 plants maize and Sorghum. Plant Mol. Biol. **25,** 669-679.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227,** 680-684.

Lai, L.B., Tausta, S.L., and Nelson, T.M. (2002). Differential regulation of transcripts encoding cytosolic NADP-malic enzyme in C3 and C4 Flaveria species. Plant Physiol. **128,** 140-149.

Lakowicz, J.R. (1980). Fluorescence spectroscopic investigations of the dynamic properties of proteins, membranes and nucleic acids. J. Biochem. Biophys. Methods **2,** 91-119.

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

Lara, P., Onate-Sanchez, L., Abraham, Z., Ferrandiz, C., Diaz, I., Carbonero, P., and Vicente-Carbajosa, J. (2003). Synergistic activation of seed storage protein gene expression in Arabidopsis by ABI3 and two bZIPs related to OPAQUE2. J. Biol. Chem. **278,** 21003-21011.

Latzko, E., and Kelly, J. (1983). The multi-faceted function of phosphoenolpyruvate carboxylase in C3 plants. Physiol. Vég. **21,** 805-815.

Li, J.J., and Herskowitz, I. (1993). Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. Science **262,** 1870-1874.

Liu, Y., Wilson, T.C., Milbrandt, J., and Johnston, M. (1993). Identifying DNA-binding sites and analysing DNA-binding domains using a yeast selection system. Methods: A companion to Methods Enzymol. **5,** 125-137.

Logemann, J., Schell, J., and Willmitzer, L. (1987). Improved method for the isolation of RNA from plant tissues. Anal. Biochem. **163,** 16-20.

Lundblad, J.R., Laurance, M., and Goodman, R.H. (1996). Fluorescence polarization analysis of protein-DNA and protein-protein. Mol. Endocrinol. **10,** 607-612.

Matsuoka, M., Kyozuka, J., Shimamoto, K., and Kano-Murakami, Y. (1994). The promoters of two carboxylases in a C_4 plant (maize) direct cell-specific, light-regulated expression in a C_3 plant (rice). Plant J. **6,** 311-319.

Meijer, A.H., Ouwerkerk, P.B., and Hoge, J.H. (1998). Vectors for transcription factor cloning and target site identification by means of genetic selection in yeast. Yeast **14,** 1407-1415.

Meurer, J., Plücken, H., Kowallik, K.V., and Westhoff, P. (1998). A nuclear-encoded protein of prokaryotic origin is essential for the stability of photosystem II in *Arabidopsis thaliana*. EMBO J. **17,** 5286-5297.

Monson, R.K. (1999). The origins of C4 genes and evolutionary pattern in the C4 metabolic phenotype. In C4 Plant Biology, R.F. Sage and R.K. Monson, eds (San Diego: Academic Press), pp. 377-410.

Monson, R.K., and Moore, B.d. (1989). On the significance of C_3-C_4 intermediate photosynthesis to the evolution of C_4 photosynthesis. Plant Cell Environ **12,** 689-699.

Moore, P.D. (1982). Evolution of photosynthetic pathways in flowering plants. Nature **295,** 647-648.

Nasir, M.S., and Jolley, M.E. (1999). Fluorescence polarization: an analytical tool for immunoassay and drug discovery. Comb. Chem. High Throughput Screen **2,** 177-190.

Ndamukong, I., Abdallat, A.A., Thurow, C., Fode, B., Zander, M., Weigel, R., and Gatz, C. (2007). SA-inducible Arabidopsis glutaredoxin interacts with TGA factors and TGA factors and suppresses JA-responsive PDF1.2 transcription. Plant J. **50,** 128-139.

Nelson, T., and Dengler, N.G. (1992). Photosynthetic tissue differentiation in C_4 plants. Int. J. Plant Sci. **153,** S93-S105.

O'Leary, M.H. (1982). Phosphoenolpyruvate carboxylase: an enzymologisttτs view. Annu. Rev. Plant Physiol. **33,** 297-315.

Onodera, Y., Suzuki, A., Wu, C.Y., Washida, H., and Takaiwa, F. (2001). A rice functional transcriptional activator, RISBZ1, responsible for endosperm-specific expression of storage protein genes through GCN4 motif. J. Biol. Chem. **276,** 14139-14152.

Patel, M., Siegel, A.J., and Berry, J.O. (2006). Untranslated regions of FbRbcS1 mRNA mediate bundle sheath cell-specific gene expression in leaves of a C4 plant. J. Biol. Chem. **281,** 25485-25491.

Powell, A.M. (1978). Systematics of *Flaveria* (Flaveriinae-Asteraceae). Ann. Mo. Bot. Gard. **65,** 590-636.

Riggs, A.D., Suzuki, H., and Bourgeois, S. (1970). Lac repressor-operator interaction. I. Equilibrium studies. J. Mol. Biol. **48,** 67-83.

Ringli, C., and Keller, B. (1998). Specific interaction of the tomato bZIP transcription factor VSF-1 with a non-palindromic DNA sequence that controls vascular gene expression. Plant Mol. Biol. **37,** 977-988.

Rosche, E., and Westhoff, P. (1995). Genomic structure and expression of the pyruvate,orthophosphate dikinase gene of the dicotyledonous C_4 plant *Flaveria trinervia* (Asteraceae). Plant Mol. Biol. **29,** 663-678.

Sage, R.F. (2004). The evolution of C4 photosynthesis. New Phytol. **161,** 341-370.

Sambrook, J., and Russell, D.W. (2001). Molecular Cloning. A Laboratory Manual. (Cold Spring Harbor Laboratory: Cold Spring Harbor Laboratory Press).

Schäffner, A.R., and Sheen, J. (1992). Maize C_4 photosynthesis involves differential regulation of phosphoenolpyruvate carboxylase genes. Plant J. **2,** 221-232.

Schoeneberg, U., Vahrson, W., Priedemuth, U., and Wittig, B. (1994). Analysis and interpretation of DNA and protein sequences using MacMolly Tetra. KAROI-Verlag Bornemann, Bielefeld, Germany.

Schwechheimer, C., Zourelidou, M., and Bevan, M.W. (1998). Plant transcription factor studies. Annu. Rev. Plant. Physiol. in Plant. Mol. Biol. **49,** 127-150.

Stockhaus, J., Schlue, U., Koczor, M., Chitty, J.A., Taylor, W.C., and Westhoff, P. (1997). The promoter of the gene encoding the C_4 form of phosphoenolpyruvate carboxylase directs mesophyll specific expression in transgenic C_4 *Flaveria* spp. Plant Cell **9,** 479-489.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. Meth. Enzymol. **185,** 60-88.

Suckow, M., Schwamborn, K., Kisters-Woike, B., von Wilcken-Bergmann, B., and Muller-Hill, B. (1994). Replacement of invariant bZip residues within the basic region

of the yeast transcriptional activator GCN4 can change its DNA binding specificity. Nucleic Acids Res. **22,** 4395-4404.

Syrovy, I., and Hodny, Z. (1991). Staining and quantification of proteins separated by polyacrylamide gel electrophoresis. J. Chromatogr. **569,** 175-196.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acid Res. **22,** 4673-4680.

Uhl, G.R. (1989). In situ hybridization: quantitation using radiolabeled hybridization probes. Meth. Enzymol. **168,** 741-752.

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

Wang, M.M., and Reed, R.R. (1993). Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. Nature **364,** 121-126.

Waterhouse, P.M., and Helliwell, C.A. (2003). Exploring plant genomes by RNA-induced gene silencing. Nat. Rev. Genet. **4,** 29-38.

Westhoff, P., and Gowik, U. (2004). Evolution of C4 phospho*enol*pyruvate carboxylase – genes and proteins: a case study with the genus *Flaveria*. Ann. Bot. **93,** 1-11.

Windhövel, A., Hein, I., Dabrowa, R., and Stockhaus, J. (2001). Characterization of a novel class of plant homeodomain proteins that bind to the C_4 phospho*enol*pyruvate carboxylase gene of *Flaveria trinervia*. Plant Mol. Biol. **45,** 201-214.

Wissmuller, S., Kosian, T., Wolf, M., Finzsch, M., and Wegner, M. (2006). The high-mobility-group domain of Sox proteins interacts with DNA-binding domains of many transcription factors. Nucleic Acids Res. **34,** 1735-1744.

Woods, R.A., and Gietz, R.D. (2001). High-efficiency transformation of plasmid DNA into yeast. Methods Mol. Biol. **177,** 85-97.

Zhang, Y., Feng, X.H., and Derynck, R. (1998). Smad3 and Smad4 cooperate with c-Jun/c-Fos to mediate TGF-beta-induced transcription. Nature **394,** 909-913.

Tables

Table 1. Listing of Used Oligonucleotides and Primers

	Oligonucleotide/ primer designation	Sequence 5' to 3'
(A) Primers/oligonucleotides for the construction of FtDRa/FtDRb promoter fusions		
	Ft-DE-A5'	GGG**GAATTC**AGAACATGAAAAAAGGACTCACCAGG
	Ft-DE-A3'	GGG**CTCGAG**TTGTTTGTTTTAGTGAGTAAGTACGGCTCTACGAACAC
FtDRa	Ft-DE-A3'	GGG**TCTAGA**TTGTTTGTTTTAGTGAGTAAGTACGGCTCTACGAACAC
	Ft-DE-B5'	GGG**GAATTC**GTGAATTTATGAGAGCTGTAC
	Ft-DE-B3'	GGG**CTCGAG**GTACTTAATTTCCATTTCTC
FtDRb	Ft-DE-B3'	GGG**TCTAGA**GTACTTAATTTCCATTTCTC
(B) Primers/oligonucleotides for the construction of MEM1 promoter fusions		
	FtM5'EcoRI-XbaI/-XhoI	**AATTC**GTGAATTTATGAGAGCTGTACTTACTCACTAAAACAAACAA**T**/**C**
	FtM3'EcoRI-XbaI/-XhoI	**CTAGA**/**TCGAG**TTGTTTGTTTTAGTGAGTAAGTACAGCTCTCATAAATTCAC**G**
FtM (1)		
	FtMAb5'EcoRI-XbaI/-XhoI	**AATTC**GTGAATTTATGAGAGCTGTACTTACTAAAACAAACAA**T**/**C**
FtM/_ΔCACT (1)	FtMAb3'EcoRI-XbaI/-XhoI	**CTAGA/TCGAG**TTGTTTGTTTTAGTAAGTACAGCTCTCATAAATTCAC**G**
	FtMab5'EcoRI-XbaI/XhoI	**AATTC**ATGAATTTATGAGAGCTGTACTTACTAAAACAAACAA**T**/**C**
FtM/A_ΔCACT (1)	FtMab3'EcoRI-XbaI/XhoI	**CTAGA**/**TCGAG**TTGTTTGTTTTAGTAAGTACAGCTCTCATAAATTCAT**G**
	FtMaB5'EcoRI-XbaI/-XhoI	**AATTC**ATGAATTTATGAGAGCTGTACTTACTCACTAAAACAAACAA**T**/**C**
	FtMaB3'EcoRI-XbaI/-XhoI	**CTAGA/TCGAG**TTGTTTGTTTTAGTGAGTAAGTACAGCTCTCATAAATTCAT **G**
FtM/A (1)		
	FbM5'EcoRI	GGG**GAATTC**GTGAATTTATGAAAAAATTAAATTGGAAAGAGG
	FpMaB3'XbaI	GGG**TCTAGA**TTGTTTGTTTTAGTGAGTAAGTACGGCTCTACGAACAC
FbM (2)	FbM3'EcoRI-XhoI	GGG**CTCGAG**TTGTTTGTTTTAGTGAGTAAGTACGGCTCTACGAACAC
	FpM5'EcoRI	GGG**GAATTC**ATGAATTTATGAAAAACTCGT
FpM (2)	FpM3'XbaI/-XhoI	GGG**TCTAGA**/GGG**CTCGAG**TTGTTTGTTTTAGTAAGTACG
	FpMaB5'EcoRI	GGG**GAATTC**ATGAATTTATGAAAAACTCGTG
	FpMaB3'XbaI	GGG**TCTAGA**TTGTTTGTTTTAGTGAGTAAGTACGGCTCTACGAACAC
FpM/+CACT (2)	FpMaB3'XhoI	GGG**CTCGAG**TTGTTTGTTTTAGTGAGTAAGTACGGCTCTACGAACAC
	FpMAb5'EcoRI	GGG**GAATTC**GTGAATTTATGAAAAACTCGTG
	FMab3'XbaI	GGG**TCTAGA**TTGTTTGTTTTAGTAAGTACG
FpM/G (2)	FpMaB3'XhoI	GGG**CTCGAG**TTGTTTGTTTTAGTGAGTAAGTACGGCTCTACGAACAC
	FpMAb5'EcoRI	GGG**GAATTC**GTGAATTTATGAAAAACTCGTG
FpM/G_+CACT	FpMaB3'XbaI	GGG**TCTAGA**TTGTTTGTTTTAGTGAGTAAGTACGGCTCTACGAACAC
(2)	FpMaB3'XhoI	GGG**CTCGAG**TTGTTTGTTTTAGTGAGTAAGTACGGCTCTACGAACAC

Table 1. (continued)

(C) Primers for the isolation and cloning of FtbZIP sequences		
	18pAD5'EcoRI	GATC**GAATTC**ATGCAAGATCCAAACCCTAACCCGATGAAGGCTC
	18pAD3'XhoI	CTAG**CTCGAG**GAACGTACTGCTGCTTTCACTGGCGGAGATTGAAG
	18-1	ATGGATATGGACAAGCTCCG
	18-2	TTCAGTGCCTCATTCAGAGC
	18-pET5'EcoRI	GATC**GAATTC**GAATGCAAGATCCAAACCCTAACC
FtbZIP 18	18-pET3'NotI	CTAG**GCGGCCGC**GAACGTACTGCTGCTTTCACTG
	29pAD5'BamHI	GATC**GGATCC**ATGGGTGATACTGAAGGGGGA
	29pAD3'XhoI	CTAG**CTCGAG**CTGCTTCGAATCAGGTTTGGTGGCTGCACCGCC
	29-1	TCTCTACTATACTGCAGTCG
	29-4	TCGGTTAGAGCTATCTCAGC
	29-pET5'BamHI	GATC**GGATCC**GAATGGGTGATACTGAAGGGGGA
FtbZIP 29	29-pET3'NotI	CTAG**GCGGCCGC**CTGCTTCGAATCAGGTTTGGTG
	51pAD5'EcoRI	GATC**GAATTC**ATGGACCCAAAGTTCGCCGGAAAACCAATTCCC
	51pAD3'XhoI	CTAG**CTCGAG**ACTGAAGTCCATGAAGGTAGGTTTGAGTTGGCTG
	51-1	TGATTTCTTTGATGAGTTGG
	51-2	TAGTTCATTCCATTAAGGAG
	51-pET5'EcoRI	GATC**GAATTC**GAATGGACCCAAAGTTCGCCGGA
FtbZIP 51	51-pET3'NotI	CTAG**GCGGCCGC**ACTGAAGTCCATGAAGGTAGG
(D) Primers/oligonucleotides for the fluorophore-labelling of MEM1 variants		
	Ft5'TAMRA	d((TAMRA)TTGTTTGTTTTAGTGAGTAAGTACAGCTCTCATAAATTCAC)
FtM (1)	Ft3'	GTGAATTTATGAGAGCTGTACTTACTCACTAAAACAAACAA
FtM/A_ΔCACT	Ft_{ab}5'TAMRA	d((TAMRA)TTGTTTGTTTTAGTAAGTACAGCTCTCATAAATTCAT)
(1)	Ft_{ab}3'	ATGAATTTATGAGAGCTGTACTTACTAAAACAAACAA
	Fb5'TAMRA	d((TAMRA)TTGTTTGTTTTAGTGAGTAAG)
FbM (2)	Fb3'	GTGAATTTATGAAAAAATTAAATTGGAAAGAGG
	Fp5'TAMRA	d((TAMRA)TTGTTTGTTTTAGTAAGTAC)
FpM (2)	Fp3'	ATGAATTTATGAAAAACTCG
FpM/G_+CACT	Fb5'TAMRA	d((TAMRA)TTGTTTGTTTTAGTGAGTAAG)
(2)	Fp_{Ab}3'	GTGAATTTATGAAAAACTCG

The *Eco*RI, *Xba*I, *Xho*I, *Not*I and *Bam*HI restriction sites are given in bold letters.

Table 2. Tabulation of the Analysis of MEM1-FtbZIP Interactions by Use of the Yeast One-Hybrid System.

	FtbZIP18		FtbZIP51	
MEM1 reporter strains	Growth on SD-Media	LacZ positive	Growth on SD-Media	LacZ positive
FtM-*lacZ*/*His3*	(*)++++	yes	+	yes
FtM/A_ΔCACT-*lacZ*/*His3*	—	—	—	—
FtM/A-*lacZ*/*His3*	++++	yes	+	yes
FtM/_ΔCACT-*lacZ*/*His3*	++++	yes	+++	yes
FbM-*lacZ*/*His3*	++	yes	+	yes
FpM-*lacZ*/*His3*	—	—	—	—
FpM/G_+CACT-*lacZ*/*His3*	++	yes	+	yes

The results of the interaction between full-length FtbZIP proteins and the respective MEM1 reporter strains is shown. The + symbol reflects the growth characteristics of the transformants per 1 µg of transformed plasmid-DNA. ++++ = ≥ 200 clones; +++ = 150 clones; ++ = 70-100 clones; + = 30-70 clones; (*) = Ø 2-3 mm, all other yeast clones: Ø 1-2 mm. The transformants were plated on synthetical dropout (SD) medium lacking histidine and leucine supplemented with 60 mM 3-AT, except for FbM-*lacZ*/*His3* the plates were supplemented with 15 mM 3-AT. The plates were incubated for 2 weeks at 30°C.

Figure Legends

Figure 1. The Molecular Anatomy of the *ppcA* Promoter of the C_4 Plant *Flaveria trinervia* and the Composition of the MEM1-Containing *Cis*-Regulatory Sequences Used in the Yeast One-Hybrid System.

(A) The structure of the distal region of the *ppcA* promoter of *F. trinervia* showing the location of the *cis*-regulatory sequences (FtDRa, FtDRb and MEM1) used as bait sequences in the yeast one-hybrid screen. The nucleotide sequences of two C_4 MEM1 (*F. trinervia*, *F. bidentis*), two C_4-like MEM1 (*F. palmerii, F. vaginata*) and two C_3 MEM1 (*F. cronquistii, F. pringlei*) are depicted, and the A- and B-submodules are highlighted by boxes. C_4-type submodules are in green, C_3-type submodules in yellow. Asterisks label identical nucleotides in the MEM1 modules of all *ppcA* promoters analyzed to date (Gowik *et al.,* 2004). Red bars indicate the single nucleotide difference in the A- and the insertion/deletion of the CACT tetranucleotide in the B-submodule. The CACT-containing DNA-binding sequence in the B- and the revers complementary version of the TCAC tetranucleotide in the A-submodules of the C_4 and C_4-like MEM1 (GTGA) are highlighted by a box. The TCAC-motif in the B-submodules of the C_4 and C_4-like MEM1 are underlined. The DNA-binding sequence of the transcription factor GCN4 is shown in italics.

(B) Schematic presentation of the MEM1 variants and of the *His3*/*LacZ*-promoter reporter constructs used in the yeast one-hybrid assay. The state of the submodule is colour-coded, green denotes a C_4-type submodule and yellow a C_3-type submodule. M: MEM1, Ft: *F. trinervia*, Fb: *F. bidentis,* Fp: *F. palmerii*, Fv: *F. vaginata*, Fc: *F. cronquistii*, Fp: *F. pringlei.*

Figure 2. Schematic Illustration of the MEM1 Interacting Proteins FtbZIP18, 29 and 51 and Comparison of their Interaction with C_4- and C_3-type MEM1 Modules.

(A) Domain structures of the identified proteins FtbZIP18, 29 and 51. The respective full-size and truncated version of each protein FtbZIP18 (clone F6E2), FtbZIP29 (clone F1A8) and FtbZIP51 (clone F4F8) are shown, parenthesis refer to the designation of the clones obtained by the yeast one-hybrid screening. Numbers refer to the amino acids of the proteins and the basic region and leucine zipper are indicated. The amino and carboxy terminus are indicated by N and C. Black bars below each protein indicate the region of the truncated proteins. The cDNA inserts of the truncated proteins FtbZIP18, 29 and 51 encode amino acids 102 to 336, 405 to 591 and 41 to 334, respectively, of the corresponding complete proteins.

(B) Transcriptional activation function of truncated and complete FtbZIP18, 29 and 51 proteins. The left block show the results of the truncated FtbZIP proteins (FtbZIP18tr, 29tr and 51tr) and the right block the results of the complete FtbZIP proteins (FtbZIP18, 29 and 51). The C_4-type MEM1 is from *F. trinervia* (FtM), the C_3-type MEM1 from *F. pringlei* (FpM). The left panel in each block shows the *in situ* staining for ß-galactosidase activity, the right panel illustrates the growth of the strains under selective conditions, i.e. in the absence of histidine. Yeast colonies were allowed to grow for two days at 30°C. FtM/FpM strains transformed with the pAD-GAL4 vector were used as negative controls. Representative results out of twenty repetitions are shown.

Figure 3. Alignment of the Amino Acid Sequences of the Ft and AtbZIP Proteins 18, 29 and 51 from *F. trinervia* and *A. thaliana* and the Putative Phylogenetic Relationship of FtbZIP18, 29 and 51 with bZIP Proteins from *A. thaliana.*

(A) Amino acid sequence alignment of FtbZIP and AtbZIP proteins. The amino acid sequences of the proteins were aligned by using Clustal W (Thompson et al., 1994). Identical amino acids are marked by asterisks. The basic region is highlighted by red colour, the leucine zipper segment by turquoise. The asparagine (N) and lysine (K) residues that are highly conserved in all bZIP proteins and the amino acid residues at position d within the repeating unit of the leucine zipper domain are labelled by bold letters. The exchange of the leucines in the leucine zipper by methionine (M) and arginine (R) are indicated by red bold letters. The numeration of the bZIP region refers to the first leucine residue in the leucine-zipper region. One repeating unit of a leucine zipper (a to g) is referred to as heptad (Landschulz et al., 1988). The leucines are sometimes replaced by the amino acids phenylalanine (F), isoleucine (I), and methionine (M).

(B) Phylogenetic analysis of FtbZIP proteins. The amino acid sequences of all bZIP proteins of group I of *A. thaliana:* AtbZIP18 (AY0744269), AtbZIP29 (AF401297), AtbZIP30 (AF401298), AtbZIP31 (AF401301), AtbZIP32 (AV566578), AtbZIP33 (-), AtbZIP51 (AF225983), AtbZIP52 (AJ419852/53), AtbZIP59 (X61031), AtbZIP69 (AJ419854), AtbZIP71 (-), AtbZIP73 (-) and AtbZIP74 (-)) and that of AtbZIP60 (AY045964), a bZIP protein of group S of *A. thaliana* (Jakoby *et al.* (2002), were aligned with FtbZIP18, 29 and 51 protein sequences by using Clustal W. The phylogenetic tree was constructed by using the same program (Thompson et al., 1994). Gene bank accession numbers, if known, are given in parenthesis.

Figure 4. SDS-Polyacrylamide Analysis of Purified Recombinant bZIP Proteins FtbZIP18, FtbZIP29 and FtbZIP51.

An aliquot of the His-tagged FtbZIP proteins, purified by affinity chromatography on Ni-nitrilotriacetic acid resin, corresponding to 10 µg of protein was separated on a 10% polyacrylamide gel (acrylamide : bisacrylamide ratio 46:1) and run according to Laemmli (1970) (Laemmli, 1970). The gels were stained by Coomassie Brilliant Blue as described in Syrovy and Hodny (1991).

Figure 5. Electrophoretic Mobility Shift Analysis of the Binding of FtbZIP18 to C_4- and C_3-type MEM1.

(A) Binding of FtbZIP18 to FtM and FpM. Binding reactions were performed at a constant concentration of DNA (500 fmol) and increasing amounts of protein (0-120 ng) as described in Materials and Methods.

(B) Binding of FtbZIP18 to FtM (C_4), FtM/A_ΔCACT (C_3), FbM (C_4), FpM (C_3) and FpM/G_+CACT (C_4). Binding reactions were performed in the absence (-) or presence (+) of 80 ng of FtbZIP18 protein. The positions of the protein-DNA complex and free DNA are indicated. Note the different sizes of the MEM1 elements FtM and FtM/A_ΔCACT on one, and of FbM, FpM and FpM/G_+CACT on the other side which is due to the absence or presence of a spacer segment of about 100 bp between the A- and B-submodules.

Figure 6. Binding of FtbZIP18 to FtM and FtM/A_ΔCACT, Detected by Fluorescence Polarization/Anisotropy.

Binding of Ft-bZIP18 to FtM (closed squares), FtM/A_ΔCACT (closed circles) was measured as described in Materials and Methods, and the data were fitted following the hyperbolic model. The best fit for the binding of FtbZIP18 to FtM is plotted as a solid line, that to FtM/A_ΔCACT as a dashed line.

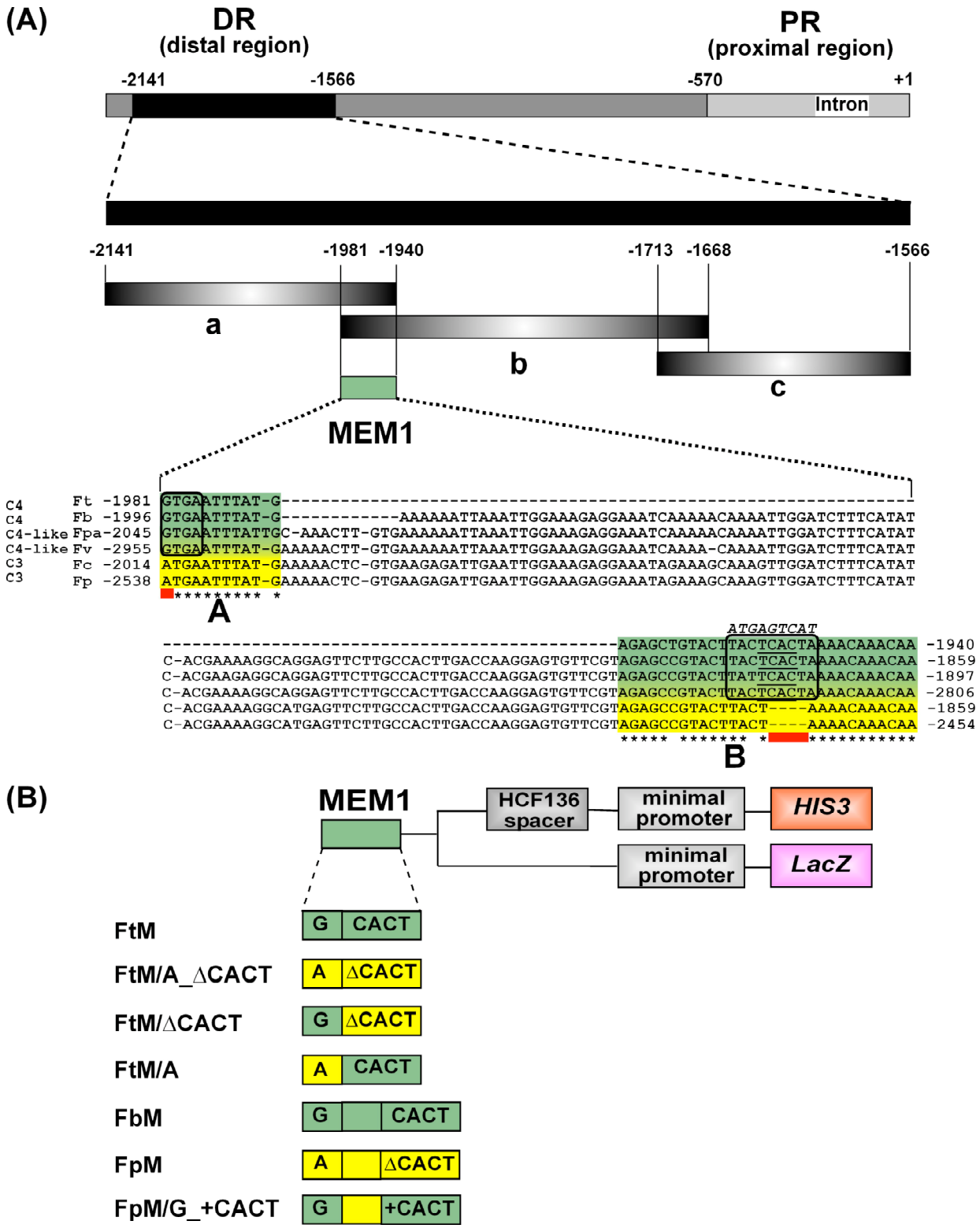

(A)
DR
(distal region)
PR
(proximal region)
-2141
-1566
-570
+1
Intron
-2141
-1981
-1940
-1713
-1668
-1566
a
b
c
MEM1
C4 Ft -1981 GTGAATTTAT-G----------
C4 Fb -1996 GTGAATTTAT-G----------AAAAAATTAAATTGGAAAGAGGAAATCAAAAACAAAATTGGATCTTTCATAT
C4-like Fpa-2045 GTGAATTTATTGC-AAACTT-GTGAAAAAATTAAATTGGAAAGAGGAAATCAAAAACAAAATTGGATCTTTCATAT
C4-like Fv -2955 GTGAATTTAT-GAAAAACTT-GTGAAAAAATTAAATTGGAAAGAGGAAATCAAAA-CAAAATTGGATCTTTCATAT
C3 Fc -2014 ATGAATTTAT-GAAAAACTC-GTGAAGAGATTGAATTGGAAAGAGGAAATAGAAAGCAAAGTTGGATCTTTCATAT
C3 Fp -2538 ATGAATTTAT-GAAAAACTC-GTGAAGAGATTGAATTGGAAAGAGGAAATAGAAAGCAAAGTTGGATCTTTCATAT
A
ATGAGTCAT
----------AGAGCTGTACTTACTCACTAAAACAAACAA -1940
C-ACGAAAAGGCAGGAGTTCTTGCCACTTGACCAAGGAGTGTTCGTAGAGCCGTACTTACTCACTAAAACAAACAA -1859
C-ACGAAGAGGCAGGAGTTCTTGCCACTTGACCAAGGAGTGTTCGTAGAGCCGTACTTATTCACTAAAACAAACAA -1897
C-ACGAAAAGGCAGGAGTTCTTGCCACTTGACCAAGGAGTGTTCGTAGAGCCGTACTTACTCACTAAAACAAACAA -2806
C-ACGAAAAGGCATGAGTTCTTGCCACTTGACCAAGGAGTGTTCGTAGAGCCGTACTTACT----AAAACAAACAA -1859
C-ACGAAAAGGCATGAGTTCTTGCCACTTGACCAAGGAGTGTTCGTAGAGCCGTACTTACT----AAAACAAACAA -2454
B
(B)
MEM1
HCF136 spacer
minimal promoter
HIS3
minimal promoter
LacZ
FtM
FtM/A_ΔCACT
FtM/ΔCACT
FtM/A
FbM
FpM
FpM/G_+CACT
G
CACT
A
ΔCACT
+CACT

Figure 1.

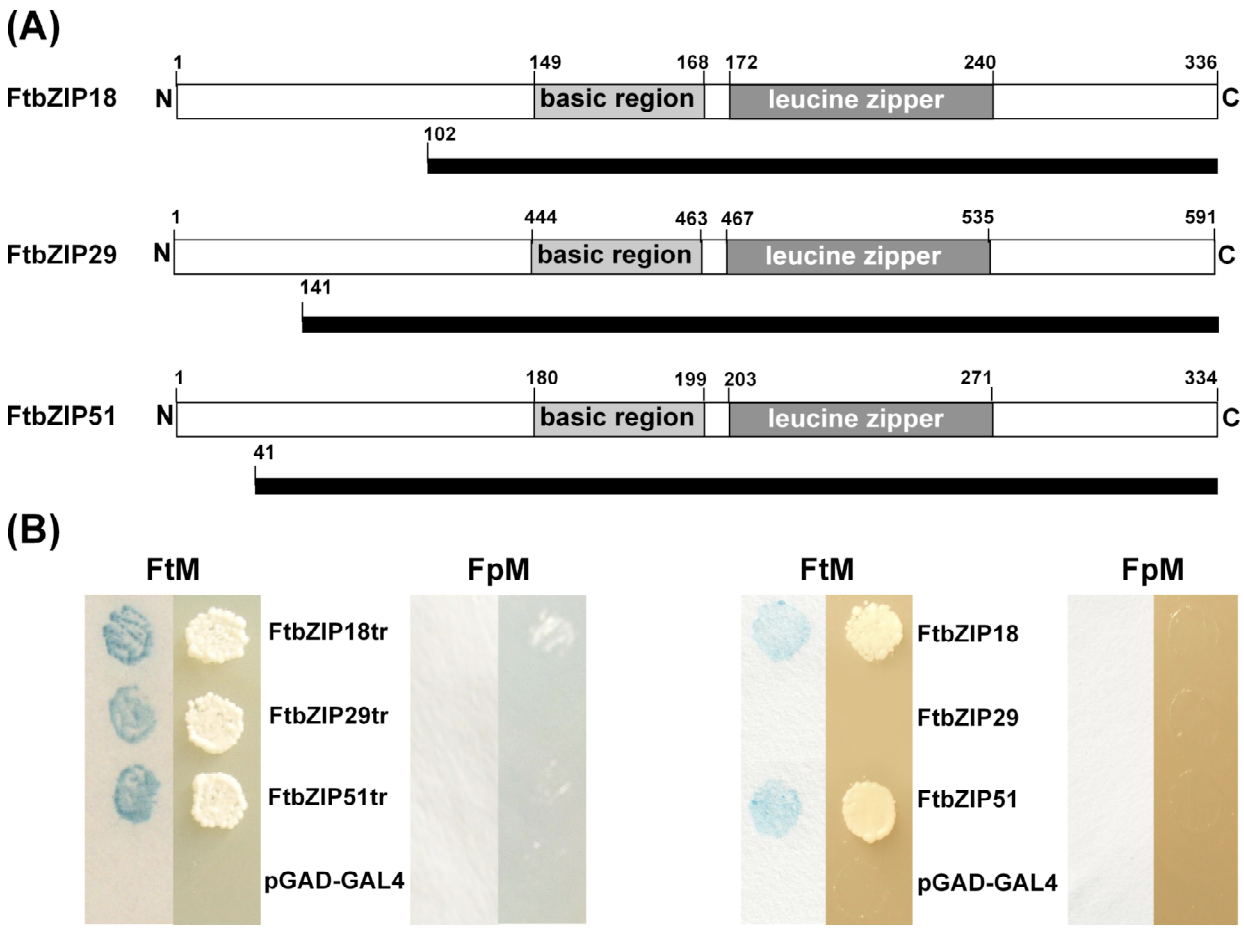

(A)
FtbZIP18
N
1
149
168
172
240
336
basic region
leucine zipper
C
102
FtbZIP29
N
1
444
463
467
535
591
basic region
leucine zipper
C
141
FtbZIP51
N
1
180
199
203
271
334
basic region
leucine zipper
C
41
(B)
FtM
FpM
FtbZIP18tr
FtbZIP29tr
FtbZIP51tr
pGAD-GAL4
FtM
FpM
FtbZIP18
FtbZIP29
FtbZIP51
pGAD-GAL4

Figure 2.

(A)

```
FTBZIP18   MQD-----------------------------------------------------------------------------------------------------
ATBZIP18   MEDP----------------------------------------------------------------------------------------------------
FTBZIP29   MGDTEGGNTDMIQRVQSSFGTSSSSVPKHQTQNPLTFNQLDIPQSNAQFRATQNRQFSPNYNVDNSSKRVGIPPSHPQF-PPVSPYSQIPVTQALGNTHK
ATBZIP29   MGDTEKCNSDMIQRLHSSFGTTSSSIPKNPISQLDL---------NPNFIRSSAPQFSKPFSD--SGKRIGVPPSHPNLIPPTSPFSQIPTTRQPGSHNF
FTBZIP51   MDPKFAGKP-------------------------------------------------------------IPTSHYHGGRR-----------------
ATBZIP51   MEG-----------------------------------------------------------------------GGRG-----------------
           *

FTBZIP18   ----------------------------------------------------------PNPNPMKAPPFPF-RGSH-----------
ATBZIP18   -----------------SNPQP-NQSNLSQCPPLATAPT-----------------------------------PAPV-RGPY-----------
FTBZIP29   IGGSQSFNHRQGPSHSRSLSQP-SFFPLDSLPPLSPSPYRDSPSSRSSDQAADDVLMDEHDGSSNSNSNSNPNPHSLLPPSSPFGRGSLTRTGESLPPRK
ATBZIP29   NPG--------GANHSRSMSQPNSFFSFDSLPPLSPSPFRDH----------DVSMEDRDSGVFNSNHSLPPSPFTRCNSTSSSSL----RVGESLPPRK
FTBZIP51   ----------------------NDIDQM-----------------------------------------------PETPT-RGAH-----------
ATBZIP51   --------------------P-NQTILSEIEHM-----------------------------------------PEAPRQRISH-----------

FTBZIP18   -HRRAHSEVNYRLPDDLDLVSDTFDAPSG---------------------------------SFE----------DLGSEDDLFCTYMDMDKLRSNFTD
ATBZIP18   -HRRAHSEVQFRLPEDLDLSE-----PFG---------------------------------GFD----------ELGSEDDLFCSYMDIEKLGSG---
FTBZIP29   AHRRSNSDIPFGFSTILQSSP-----PLIPLRNPSTDRAAPSNSQGSKPIQLVKRESMWEKSGNEGDAEGMGERKSEGEVVDDLFSAYMNLDNLDTL---
ATBZIP29   SHRRSNSDIPSGF-----NSM-----PLIPPRPLERSFSGGECADWSKSNPFVKKESSCEREGVG----------EREAMDDLFSAYMNLENIDVL---
FTBZIP51   -HRRTQSETFFPFPDEDILLEDVVAD-------------------------------------------------------------FNFAGIDLP---
ATBZIP51   -HRRARSETFFSGESIDDLLLFDPSD-------------------------------------------------------------IDFSSLDFL---
            ***  *

FTBZIP18   AAAGGALDSGRIANAA-----------------------------------------AGAAAEENGGGDGEKTST----------------RP
ATBZIP18   ---------SGS-ASDSAGPSAPRSDNPFSAEN-------------------------------------GGAEAGNS----------------RP
FTBZIP29   ---------NSS-GTDEKQGTENREDLD-SRASGTKTTGADSSDNEATSSMNESGKHLRNSGISSVANKREGVK-RSAGGDI----------------AP
ATBZIP29   ---------NSSEADDSKNGNENRDDMESSRASGTKTNGSD------TEGESSSVNESANNNMNSSGEKRESVKRRAAGGDI----------------AP
FTBZIP51   -----SLSSDAPAPTG-----------------------------------------------DSSQEKDSSGGKPTARKT----------------AG
ATBZIP51   --------NAPPPPQ-------------------------------------------------QSQQQPQASPMSVDSEETSSNGVVPPNSLPPKPEAR

FTBZIP18   RHRHSNSVDSPSIFRETI-----------------------------------------------EAKKA----MAPDKLAELWTIDPKRAK------R
ATBZIP18   RHRHSLSVDGSS-TLESI-----------------------------------------------EAKKA----MAPDKLAELWVVDPKRAK------R
FTBZIP29   TTRHYRSVSMDS-FMERMNFGDESPKLPPSPGGQIGQL--SPNNSIDSNSNT-FSLEFGNGEFTGAELKKI----MANEKLAEIALTDPKRAK------R
ATBZIP29   TTRHYRSVSVDSCFMEKLSFGDESLKPPSPSPGSMSRKV--SPTNSVDGNSGAAFSIEFNNGEFTAAEMKKI----MANDKLAEMAMSDPKRVKRNDPLFR
FTBZIP51   YGTHLRSLSVGSDFFDELGLSS----AVEAEKTGDGGYRHRRSGSTDGSAATSFEGDSVLMLLDNS--KKA----LAPDKLAELSLIDPKRAK------R
ATBZIP51   FGRHVRSFSVDSDFFDDLGVTEEKFIATSSGEKKKGNHHHSRSNSMDGEMSSASFNIESILASVSGKDSGKKNMGMGGDRLAELALLDPKRAK------R
             *                                                                                  ***    **** *      *

                -18     -10       +1     +7     +14    +21    +28    +35    +42    +49    +56    +63
                          abcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefg
FTBZIP18   ILANRQSAARSKERKARYISELERKVQTLQTEATTLSAQLTLFQRDTTGLSSENTELKLRLQAMEQQAQLRDALNEALKQEVERLRVATGEIASCSDTYN
ATBZIP18   IIANRQSAARSKERKARYILELERKVQTLQTEATTLSAQLSLFQRDTTGLSSENTELKLRLQVMEQQAKLRDALNEQLKKEVERLKFATGEVSP-ADAYN
FTBZIP29   ILANRQSAARSKERKMRYITELEHKVQTLQTEATTLSAQLTLLQRDSAGLTSQNNELKFRLQAMEQQSQLRDALNEALTAEVHRLKMTNAELN-------
ATBZIP29   ILANRQSAARSKERKMRYIVELEHKVQTLQTEATTLSAQLTLLQRDMMGLTNQNNELKFRLQAMEQQARLRDALNEALNGEVQRLKLAIGESS-------
FTBZIP51   ILANRQSAARSKERKTRYTSELEKKVKTLQTEATTLSAQVTKLQRETSDITTENKELKLRLEAMEQHAHLRDALNEALREEVNRLKLEAGQPPLL-----
ATBZIP51   ILANRQSAARSKERKIRYTGELERKVQTLQNEATTLSAQVTMLQRGTSELNTENKHLKMRLQALEQQAELRDALNEALRDELNRLKVVAGEIPQG-----
           *************** **   *** ** *** ********    **         * *** **     *   ******* *  *    *

FTBZIP18   LGMHHVPYNQP-----------------NL-------------------FTNQHQPDPQKYQQSNHHRNHPFLAATH--------ETLQQDQLNRFQGLD
ATBZIP18   LGMAHMQYQQQPQQSFFQHHHQQQTDAQNLQQMTHQF---------HLFQPNNNQNQSSRTNPPTAHQLMHHATSNAPAQSHSYSEAMHEDHLGRLQGLD
FTBZIP29   ------------GDAAKFSQL-SISPQMFQLQ----------------QQQHAHQMQHQNQQQSHQQN---GGAATKPDSKQ-----------------
ATBZIP29   ------------QNESERSKMQSLNAEMFQQL----------------NISQLRQQPQQMQQQSHQQNHQNGTMATKSESNE-----------------
FTBZIP51   ------------------------NGMNYNASLPPRYSSNTQPLHHFASPNAQQPQPQNTQMPNSNNISQLK----PTF-----------------MD
ATBZIP51   ------------------------NGNSYNRA----------------QFSSQQSAMNQFGNKTNQQMSTNGQPSLPSY-----------------MD

FTBZIP18   ISNRGGGSHTVKTEVPSISASESSSTF
ATBZIP18   ISSCGRGSNFGRSDTVSESSSTM----
FTBZIP29   ---------------------------
ATBZIP29   ---------------------------
FTBZIP51   FS-------------------------
ATBZIP51   FTKRG----------------------
```

(B)

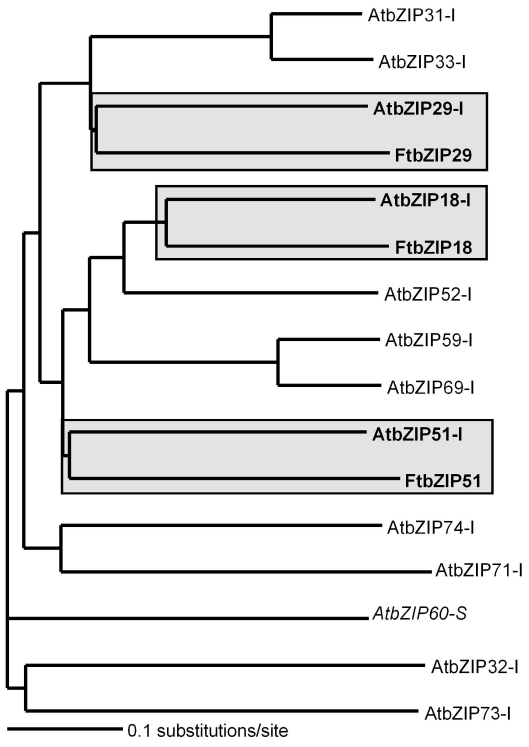


Figure 3.

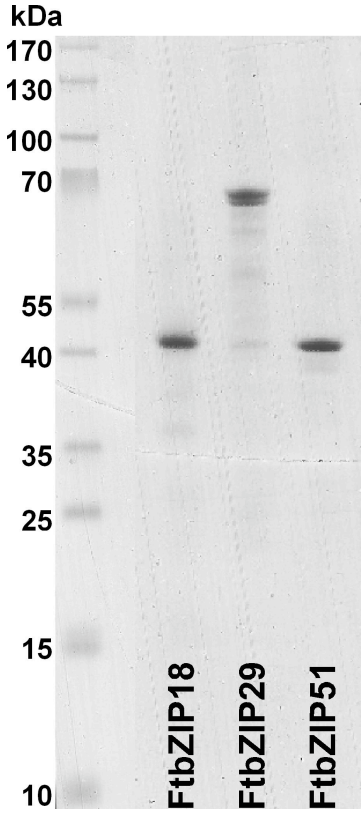

kDa
170
130
100
70
55
40
35
25
15
10
FtbZIP18
FtbZIP29
FtbZIP51

Figure 4.

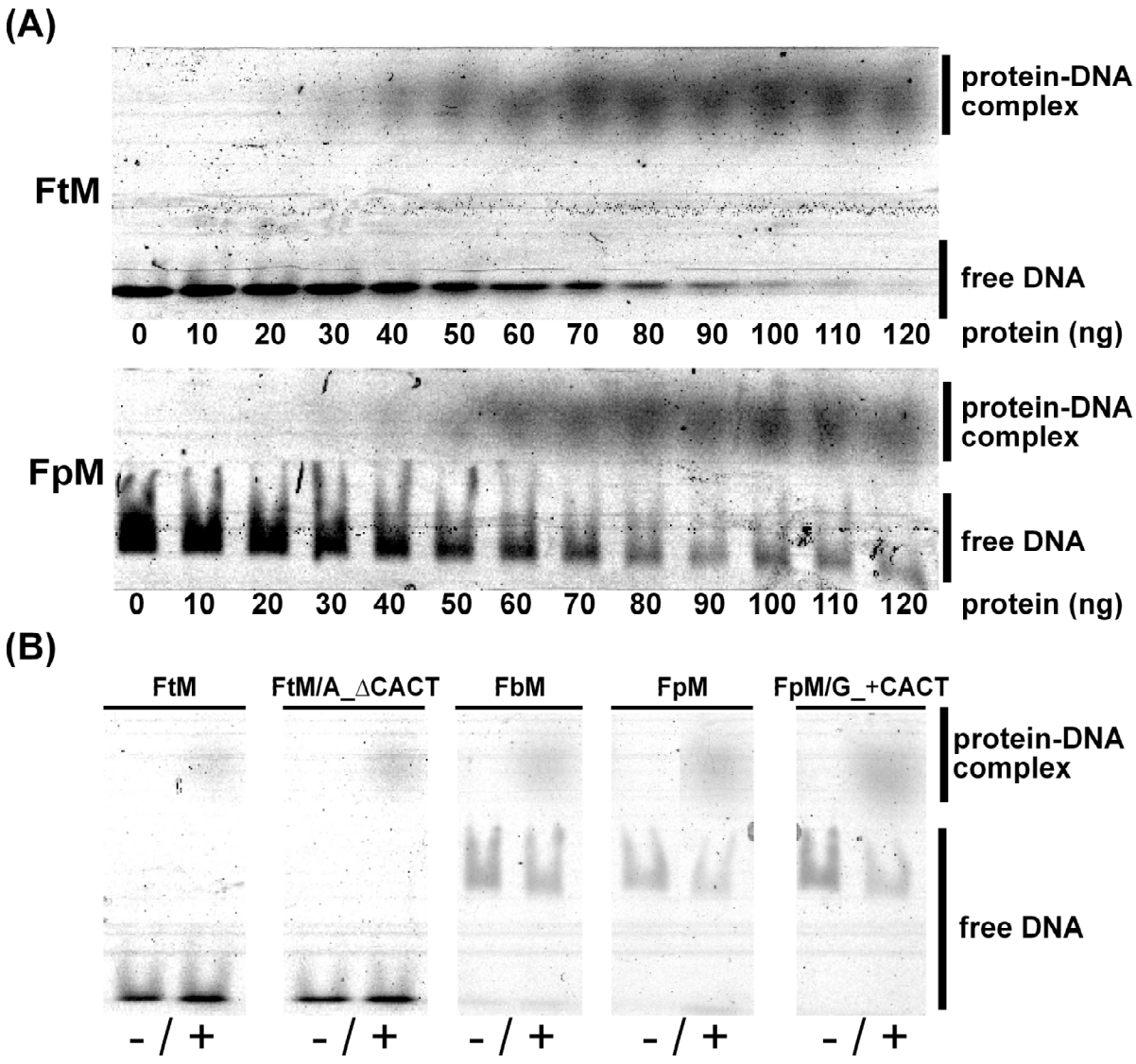

(A)
FtM
protein-DNA complex
free DNA
0 10 20 30 40 50 60 70 80 90 100 110 120
protein (ng)
FpM
protein-DNA complex
free DNA
0 10 20 30 40 50 60 70 80 90 100 110 120
protein (ng)
(B)
FtM
FtM/A_ΔCACT
FbM
FpM
FpM/G_+CACT
protein-DNA complex
free DNA
- / +
- / +
- / +
- / +
- / +

Figure 5.

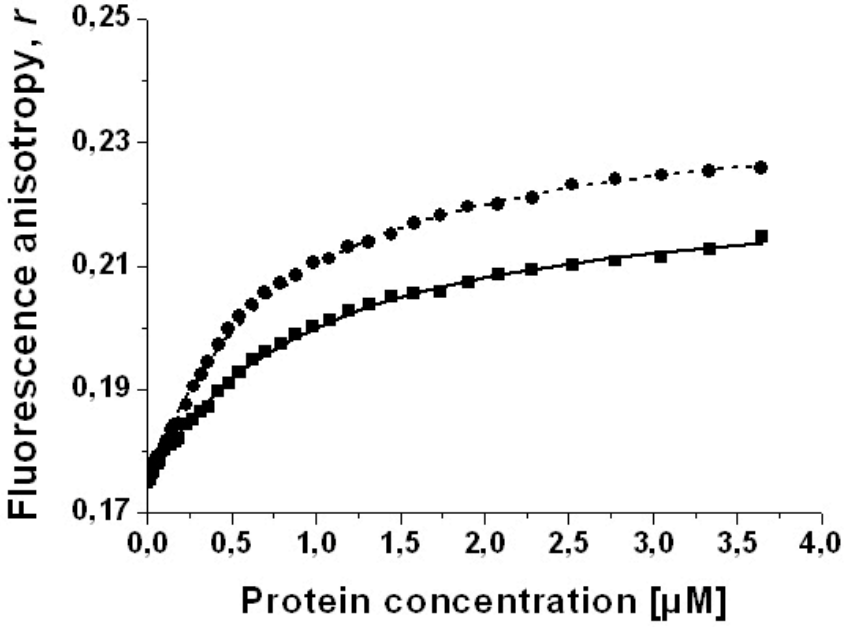

Fluorescence anisotropy, r
0,25
0,23
0,21
0,19
0,17
0,0
0,5
1,0
1,5
2,0
2,5
3,0
3,5
4,0
Protein concentration [μM]

Figure 6.

Supplemental Data

Complete cDNA sequences of the isolated FtbZIP proteins FtbZIP18 (clone F6E2), FtbZIP29 (clone F1A8) and FtbZIP51 (clone (F4F8). The respective start-/stop-codons and the polyA-tail are shown in bold letters.

FtbZIP18 cDNA

```
GGCACGAGGC TCCTCTTTGC TCTGAAGATC ACACACACAC AGAGCGAGAG AGACGCAAAC     60
GCA**ATG**CAAG ATCCAAACCC TAACCCGATG AAGGCTCCAC CATTTCCGTT CAGAGGCTCT    120
CACCACCGGA GAGCCCACTC CGAGGTCAAC TACCGCCTCC CCGACGACCT AGATCTGGTT    180
TCCGACACCT TCGATGCGCC TTCTGGAAGC TTCGAGGACC TCGGATCTGA AGATGATCTC    240
TTCTGCACTT ACATGGATAT GGACAAGCTC CGATCCAACT TCACTGATGC CGCTGCTGGA    300
GGAGCGCTCG ACTCCGGCCG GATCGCCAAT GCCGCTGCCG GTGCCGCTGC GGAGGAGAAT    360
GGAGGTGGTG ATGGTGAGAA GACTAGTACC AGACCTCGAC ATCGGCATAG TAATTCGGTG    420
GATAGTCCGA GTATTTTCAG GGAGACGATT GAGGCGAAGA AGGCGATGGC TCCTGATAAG    480
CTTGCTGAAT TGTGGACTAT TGATCCCAAA CGAGCTAAGA GGATTCTGGC CAATCGACAA    540
TCTGCTGCTC GGTCAAAAGA GAGGAAAGCC CGCTATATAT CCGAGCTTGA GAGAAAAGTT    600
CAAACCCTAC AAACAGAAGC AACCACTCTT TCAGCGCAAT TGACTCTGTT TCAGAGGGAT    660
ACTACTGGCC TATCTTCTGA AAACACAGAG CTTAAGCTCC GATTACAAGC AATGGAACAA    720
CAAGCTCAGT TACGGGATGC TCTGAATGAG GCACTGAAGC AAGAAGTGGA AAGGCTGAGA    780
GTTGCAACTG GAGAAATAGC AAGCTGTTCA GATACATACA ATTTAGGAAT GCACCATGTT    840
CCCTATAACC AACCAAACTT GTTCACAAAC CAACACCAAC CCGATCCCCA AAAATACCAA    900
CAGTCAAACC ACCATCGCAA CCACCCTTTT CTTGCAGCCA CCCATGAAAC GTTGCAGCAG    960
GATCAACTTA ACCGCTTCCA GGGTCTTGAC ATTAGTAACA GAGGTGGTGG TTCTCATACT   1020
GTCAAAACCG AAGTCCCTTC AATCTCCGCC AGTGAAAGCA GCAGTACGTT C**TGA**TCCCCA   1080
ACATCTACAT TTTCACCATT ACATGCATTT AGTTGCTATT TATGTTCCTA TATCTTTCAA   1140
CTGTTCATAG TTTGACACCT CATTTAGCAC CCACCGATCA ACACGCACGA ATCCTTTCAA   1200
AATTCTATTT GTTTATTGTC ACATGGTTAG GGCTTTAGGT TTAACAAGTG TCGTTCTTTA   1260
TTCTTGTTGT CCTGTTGTTA TGTTACTGAA TTGCATCATC TTGTGCTATT CCCAAAAGTC   1320
CTGTTGTAGA GTGTGTTATT CATCCAACTC TTACCATTGT TGTGCATGAC GATTATTGTG   1380
TTGATTTATT AAACTAATTT ATTTTCAGTT TACTATC**AAA AAAAAAAAAA AAAAAA**      1436
```

FtbZIP29 cDNA

GGCACGAGGA TTTCTTCACT TTCTAGAGAG AAGAAAAACA CACACACACA CATACAGATA 60
CAATTACACA CCATTAGGGC GGAACAATCT CTGCTACACT CAAATTCTCC GGTAAACTGG 120
TTTTTCGTAC ATTGTTTCGT CATTTGTGAG CGATTGAGCT GAAAGGTGTG ATTAGGGTTT 180
TGAATTGGTT GAATTGATTG ATTAGTGTCG GTGATTTCGG TGGAGAATTG GGGGAATTGA 240
GAGGAA**ATG**G GTGATACTGA AGGGGGAAAT ACTGATATGA TTCAGAGAGT TCAATCTTCG 300
TTTGGGACTT CATCTTCATC AGTTCCTAAA CATCAAACGC AAAACCCTCT TACGTTTAAT 360
CAACTTGATA TACCTCAATC GAACGCACAG TTTCGTGCTA CTCAGAATCG ACAGTTTTCC 420
CCTAATTATA ATGTTGATAA TAGCAGTAAA AGAGTCGGAA TCCCACCGTC TCATCCTCAG 480
TTTCCTCCGG TGTCTCCGTA CTCACAGATT CCGGTGACCC AAGCTTTAGG GAACACTCAT 540
AAAATTGGGG GTTCACAGAG TTTTAATCAT AGGCAAGGGC CTTCGCATTC GCGATCGTTA 600
TCGCAGCCTT CGTTTTTCCC TCTTGATTCC CTTCCGCCTT TGAGCCCATC ACCGTATCGC 660
GATTCTCCTT CGTCGCGTTC CTCTGATCAA GCTGCAGATG ATGTATTAAT GGATGAACAT 720
GATGGAAGTT CGAATTCGAA CTCGAATTCG AATCCGAATC CTCACTCTTT ATTGCCGCCT 780
TCTTCCCCTT TTGGTAGGGG GAGTTTGACA CGAACTGGTG AGAGTCTTCC ACCACGTAAG 840
GCTCATAGGC GGTCGAATAG TGATATTCCA TTTGGGTTCT CTACTATACT GCAGTCGTCG 900
CCACCGTTGA TTCCTTTGAG AAACCCGAGT ACAGACAGGG CTGCACCCTC CAACAGTCAA 960
GGGTCAAAGC CGATTCAGTT GGTGAAACGG GAATCGATGT GGGAGAAAAG TGGTAATGAA 1020
GGTGATGCTG AAGGAATGGG AGAGAGGAAA TCTGAAGGGG AAGTTGTAGA CGATTTGTTT 1080
TCTGCTTATA TGAACTTGGA CAATCTTGAT ACATTGAACT CATCGGGAAC AGATGAAAAA 1140
CAAGGAACCG AGAATCGGGA AGATTTGGAT AGCAGAGCTA GTGGTACAAA GACGACTGGT 1200
GCTGATAGCA GTGATAACGA AGCTACAAGT AGTATGAATG AAAGCGGGAA ACATCTGCGA 1260
AATTCGGGAA TCAGTTCGGT TGCTAATAAA AGAGAAGGGG TTAAAAGGAG TGCAGGTGGA 1320
GACATTGCTC CAACTACAAG GCATTATAGA AGTGTTTCAA TGGATAGTTT TATGGAGAGG 1380
ATGAACTTTG GTGATGAATC ACCTAAGCTC CCTCCTTCCC CTGGTGGACA AATTGGTCAA 1440
CTGTCACCTA ATAATTCTAT TGACTCGAAT TCAAATACGT TCAGCTTAGA GTTTGGTAAT 1500
GGTGAGTTTA CTGGAGCTGA ACTCAAGAAA ATCATGGCGA ATGAGAAACT TGCTGAGATA 1560
GCTCTAACCG ATCCAAAACG AGCTAAGAGG ATCTTGGCTA ACAGACAGTC TGCTGCTCGA 1620
TCAAAAGAGC GTAAAATGCG TTACATTACA GAACTAGAAC ACAAGGTTCA AACTTTACAG 1680
ACCGAAGCAA CAACACTGTC TGCACAACTT ACCTTACTGC AGAGGGATTC AGCCGGGCTC 1740
ACCAGTCAAA ACAATGAGCT AAAGTTTCGT CTGCAGGCTA TGGAACAGCA GTCACAGCTT 1800
CGAGATGCTT TAAACGAGGC TTTGACCGCC GAGGTTCATC GTTTGAAAAT GACTAACGCT 1860
GAACTAAATG GGGATGCTGC TAAGTTCTCT CAACTCTCAA TCAGCCCACA AATGTTCCAG 1920
TTACAGCAGC AACAACACGC TCATCAAATG CAACACCAAA ACCAGCAGCA GTCGCACCAG 1980
CAAAATGGCG GTGCAGCCAC CAAACCTGAT TCGAAGCAG**T AG**GGCCGAGT GGCCCCTAAT 2040
AACCTTGAGG TGCTGAGTTT GTTCATACAA TCACTATGAT GTGTTGATTT ATATCCCAGG 2100
TGCATTCGTT CATTTGTTTT TTGTGTCTTC TGGGAATGGG TTCTGTTAGG AACTTCTAAT 2160
ATATTCTATC AAGTGTTTTA GTCACTTATA TTATATATAT AGAGACAGAG AGGTTATAAT 2220
ATGAACTTAA ATATAGTAAG AACGGTAAGA ACTATTAAGG GCCGCATTTC AAATGCGAGT 2280
GAAACTTGAA TGCATTTGAA GGGAATTTAG ATTTATTTGA ATTTAAATTT CACCTTTTGA 2340
GATGGGGCTC ATATTAGTTC TTACAATTTT TATTTATATT TAAGTTTCTA TTTGATCTTG 2400
TATGTATACA CACTACTTGG TGTTTTACAA CTATTTATAT TTTGTTCATG ATCTTGATAT 2460
TCATTGGATT TGAACTTGTT CTTGTT**AAAA A** 2491

FtbZIP51 cDNA

GGCACGAGGT AATAACCATC TCCCCCAATC GGTAAACCTT AACACCCAAA ACCACCACCC 60

TCCGGCGAAA TCACCACCAG CTCCCACCGC CGCCGACTCA ACCTTCACCG CATG**ATG**GAC 120

CCAAAGTTCG CCGGAAAACC AATTCCCACC TCCCACTACC ACGGTGGTCG CCGGAACGAC 180

ATCGATCAGA TGCCCGAAAC ACCCACTCGA GGAGCCCACC ACCGCCGGAC CCAATCGGAG 240

ACCTTCTTCC CCTTCCCCGA CGAAGATATT CTCCTCGAAG ACGTCGTTGC GGACTTCAAT 300

TTCGCCGGAA TTGACCTCCC GTCACTCTCC TCCGACGCCC CTGCACCCAC CGGAGATTCA 360

TCTCAGGAGA AAGATTCCTC CGGCGGTAAG CCGACGGCGA GAAAGACGGC GGGGTATGGG 420

ACCCATCTCA GGAGTCTCTC TGTGGGGTCT GATTTCTTTG ATGAGTTGGG GCTGAGCTCC 480

GCCGTGGAGG CGGAGAAGAC CGGCGACGGT GGGTATCGTC ACCGGCGAAG TGGTTCGACC 540

GATGGGTCTG CTGCGACGTC GTTTGAAGGG GATTCTGTGT TGATGTTGCT TGATAACTCG 600

AAGAAAGCAC TTGCACCTGA TAAGTTAGCT GAGCTTTCGT TGATTGATCC CAAAAGAGCT 660

AAAAGGATTC TTGCGAATAG GCAGTCTGCA GCGCGATCGA AGGAGCGAAA GACGCGGTAT 720

ACCAGTGAGT TGGAGAAGAA GGTCAAGACC CTGCAGACTG AAGCTACTAC CCTCTCTGCA 780

CAAGTCACTA AACTGCAGCG GGAAACTAGT GACATAACGA CTGAAAACAA GGAGCTCAAA 840

CTGAGGTTAG AAGCAATGGA ACAGCATGCC CACCTCAGAG ATGCTTTAAA TGAAGCACTA 900

AGAGAGGAAG TTAATCGTCT TAAACTTGAG GCGGGTCAAC CTCCTCTCCT TAATGGAATG 960

AACTACAATG CATCATTACC GCCTCGATAC TCGTCAAACA CCCAACCACT TCATCACTTC 1020

GCAAGCCCGA ATGCTCAACA GCCACAGCCA CAAAATACTC AAATGCCCAA TTCAAATAAC 1080

ATCAGCCAAC TCAAACCTAC CTTCATGGAC TTCAGT**TAA**A TTTAAAAACA TACACAAGGT 1140

ACAATTGTAT CTTCTTGTAG ATAAGTACGC GATATGGGAT TGTATACAAT TATCGGTATA 1200

GAGTGTTATG GCGCAACATT TACACAAATA CATGTTTTCA TATTGGTCAG CTTTGTAACG 1260

GTTTGGGAGC TTGGTATTTG TACATAGGTT GTATCCATTT TATGTTACTG TTGTAGGATA 1320

TATATATATA TACACTGGTT TACAGGGAAC TTATAGAAT**A AAAAAAAAAA AAAAAAAA** 1378