Evolution of the Phosphoenolpyruvate Carboxylase Protein Kinase Family in C\textsubscript{3} and C\textsubscript{4} Flaveria spp.\textsuperscript{1}[W][OPEN]

Sophia H. Aldous, Sean E. Weise, Thomas D. Sharkey, Daniel M. Waldera-Lupa, Kai Stühler, Julia Mallmann, Georg Groth, Udo Gowik, Peter Westhoff, and Borjana Arsova\textsuperscript{*}


The key enzyme for C\textsubscript{4} photosynthesis, Phosphoenolpyruvate Carboxylase (PEPC), evolved from nonphotosynthetic PEPC found in C\textsubscript{3} ancestors. In all plants, PEPC is phosphorylated by Phosphoenolpyruvate Carboxylase Protein Kinase (PPCK). However, differences in the phosphorylation pattern exist among plants with these photosynthetic types, and it is still not clear if they are due to interspecies differences or depend on photosynthetic type. The genus Flaveria contains closely related C\textsubscript{3}, C\textsubscript{3}-C\textsubscript{4} intermediate, and C\textsubscript{4} species, which are evolutionarily young and thus well suited for comparative analysis. To characterize the evolutionary differences in PPCK between plants with C\textsubscript{3} and C\textsubscript{4} photosynthesis, transcriptome libraries from nine Flaveria spp. were used, and a two-member PPCK family (PPCKA and PPCKB) was identified. Sequence analysis identified a number of C\textsubscript{3} and C\textsubscript{4}-specific residues with various occurrences in the intermediates. Quantitative analysis of transcriptome data revealed that PPCKA and PPCKB exhibit inverse diel expression patterns and that C\textsubscript{3} and C\textsubscript{4} Flaveria spp. differ in the expression levels of these genes. PPCKA has maximal expression levels during the day, whereas PPCKB has maximal expression during the night. Phosphorylation patterns of PEPC varied among C\textsubscript{3} and C\textsubscript{4} Flaveria spp. too, with PEPC from the C\textsubscript{4} species being predominantly phosphorylated throughout the day, while in the C\textsubscript{3} species the phosphorylation level was maintained during the entire 24 h. Since C\textsubscript{4} Flaveria spp. evolved from C\textsubscript{3} ancestors, this work links the evolutionary changes in sequence, PPCK expression, and phosphorylation pattern to an evolutionary phase shift of kinase activity from a C\textsubscript{3} to a C\textsubscript{4} mode.

C\textsubscript{4} plants evolved from C\textsubscript{3} plants by developing a spatial separation for the process of carbon fixation in the leaves and carrying it out in two cell types, mesophyll and bundle-sheath cells. This special leaf anatomy is known as Kranz anatomy and includes enlarged, chloroplast-rich bundle-sheath cells around the closely spaced veins to ensure an intense contact between the mesophyll and bundle-sheath cells (Sage et al., 2012). During C\textsubscript{4} photosynthesis, atmospheric CO\textsubscript{2} is initially fixed by Phosphoenolpyruvate Carboxylase (PEPC) in the mesophyll cells. Here, PEPC catalyzes the $\beta$-carboxylation of phosphoenolpyruvate, using HCO\textsubscript{3}\textsuperscript{-} as a substrate, leading to the formation of the four-carbon organic acid oxaloacetate. Oxaloacetate is reduced to malate or transaminated to Asp and transported into the bundle-sheath cells, where CO\textsubscript{2} is released by a decarboxylating enzyme (Edwards and Walker, 1983). By this process, CO\textsubscript{2} is effectively concentrated in the bundle-sheath cells, bringing its concentration up to 1,500 $\mu$L L\textsuperscript{-1} (Hatch, 1987; Sage et al., 2012). CO\textsubscript{2} is refixed by Rubisco, which in C\textsubscript{4} plants is restricted to bundle-sheath cells, and further metabolized by the Calvin-Benson cycle (Sage, 2004). This characteristic allows C\textsubscript{4} plants to survive in more extreme environments, where heat and the lack of water cause the closing of stomata and result in low CO\textsubscript{2} concentrations in the intercellular spaces around the mesophyll cells (Ehleringer et al., 1997).

However, the evolution of C\textsubscript{4} photosynthesis involved more than just the development of Kranz anatomy. It included changes in the transcriptional regulation of genes that encode components of the C\textsubscript{4} pathway and the adjustment of enzyme properties and their regulation (Sheen, 1999; Sage, 2004; Akylidiz et al., 2007; Engelmann et al., 2008; Aubry et al., 2011; Brown et al., 2011; Ludwig, 2011; Wiludda et al., 2012; Paulus et al., 2013a, 2013b). In the process of evolution toward C\textsubscript{4} photosynthesis, PEPC was influenced by a series of single-amino acid exchanges that raised the Michaelis-Menten constant for the substrate phosphoenolpyruvate (Bläsing et al., 2000) and lowered the sensitivity to malate and Asp (Engelmann et al., 2002; Paulus et al., 2013a, 2013b). So far, there is no information if such exchanges have also occurred in the regulatory protein Phosphoenolpyruvate Carboxylase Protein Kinase.
(PPCK), whose phosphorylation of PEPC changes the kinetic and regulatory properties of the enzyme (Nimmo, 2003). In order to shed light on two questions (What are the differences in PEPC regulation, by phosphorylation, between plants with C₃ and C₄ photosynthesis? And how did PEPC phosphorylation in C₄ plants evolve to the state it is in today?), we turned to species in the genus *Flaveria*.

Although the first C₄ plants originated about 30 million years ago (Pagani et al., 2005; Tipple and Pagani, 2007), the much younger genus *Flaveria* is the preferred model for studying the evolution of C₄ photosynthesis. Having evolved in the last 3 million years, the species contained within this genus still retain a high level of similarity and allow the study of changes driven predominantly by the evolution of C₄ photosynthesis (Christin et al., 2011; Sage et al., 2012). Most importantly, the genus contains C₄, C₃-C₄ intermediate, and C₃ species, such as *Flaveria trinervia*, *Flaveria ramosissima*, and *Flaveria pingleri*, respectively (Sage, 2004; McKown et al., 2005). This palette of species with various photosynthetic types allows investigation of the gradual development of C₄-related characteristics by comparative analysis.

Contrary to its photosynthetic function in C₄ plants, in C₃ plants PEPC is nonphotosynthetic and involved in diverse functions, such as replenishment of the tricarboxylic acid cycle, carbon-nitrogen interactions, carbon storage, and pH maintenance (O’Leary et al., 2011). This multifaceted role of PEPC in plant metabolism is underlined by a complex regulation system (O’Leary et al., 2011).

In order to carry out all these functions, PEPC in all investigated angiosperms is encoded by small gene families (Lepiniec et al., 1994). In *Flaveria* spp., the PEPC gene family is composed of four genes, *ppcA* to *ppcD*, with the latter being a bacterial PEPC and out of the scope of this article (Hermans and Westhoff, 1990, 1992). The *ppcA* gene of *C₄ Flaveria* spp. encodes the C₄-type photosynthetic PEPC isofrom, which is highly expressed in the mesophyll cells of *C₄ Flaveria* spp., while its evolutionary ortholog in *C₃ Flaveria* spp. encodes a typical nonphotosynthetic C₃-type PEPC (Hermans and Westhoff, 1992; Svensson et al., 2003). *ppcB* and *ppcC* of both *C₃* and *C₄ Flaveria* spp. encode nonphotosynthetic PEPC isofroms with a ubiquitous expression pattern (Ernst and Westhoff, 1997; Svensson et al., 2003).

As a primary enzyme in *C₃* photosynthesis or primary metabolism in *C₄* plants, PEPC is controlled both metabolically and posttranslationally. Malate, Asp, and oxaloacetate function as negative feedback regulators, while the PEPC activity increases in the presence of triose and hexose phosphate (Rajagopalan et al., 1994; Law and Plaxton, 1997; Bläsing et al., 2002; Svensson et al., 2003; Takahashi-Terada et al., 2005; Jacobs et al., 2008). The *C₄ ppcA* PEPC isoform has been shown to have a lower sensitivity for the allosteric inhibitors mentioned above, a characteristic that is explained by a single-amino acid mutation from Arg-884 to Gly (Paulus et al., 2013a, 2013b).

Both phosphorylation and ubiquitination have been indicated as posttranslational modifications that affect PEPC activity. Monoubiquitination of PEPC has been identified only recently; it was found to be tissue specific and to influence the feedback regulation of PEPC by various metabolites (Ageotsuma et al., 2005; Uhrig et al., 2008; O’Leary et al., 2011; Shane et al., 2013).

Phosphorylation of the photosynthetic PEPC of *C₄* and Crassulacean acid metabolism (CAM) plants has been known for decades (Budde and Chollet, 1986; Nimmo et al., 1986; Jiao and Chollet, 1989). Phosphorylation of PEPC was shown to be induced by several factors, such as light or the availability of nitrogen and phosphorus (Leport et al., 1996). Diel regulation of phosphorylation is evident for CAM plants, Arabidopsis (*Arabidopsis thaliana*), maize (*Zea mays*), sorghum (*Sorghum bicolor*), *Digitaria sanguinalis*, and wheat (*Triticum aestivum*; Jiao and Chollet, 1988; Echewarría et al., 1990; McNaughton et al., 1991; Duff and Chollet, 1995; Giglioli-Guivarč’ et al., 1996; Nimmo, 2000; Fontaine et al., 2002). However, these same studies outline some crucial differences among species when grouped by photosynthesis type. In *C₄* plants, the phosphorylation of PEPC is induced by light and decreased in the dark (Úeno et al., 2000). In CAM plants, both the expression of the *ppc* transcript and the phosphorylation level of PEPC peak at night (Nimmo, 2000), while in *C₃* plants, mixed results were found (Ván Quy et al., 1991; Duff and Chollet, 1995; Li et al., 1996; Fukayama et al., 2006; Meimoun et al., 2009).

In the context of the changes occurring during *C₄* evolution, it is crucial to understand whether the above differences result from the different photosynthetic types among the studied species or are because of species individuality. The effects of malate and Glc-6-P on PEPC activity are modulated by phosphorylation, which decreases the inhibition of PEPC by malate and renders the enzyme more sensitive to the activator Glc-6-P. Thus, phosphorylation seems to broaden the conditions under which PEPC can be active and seems to decrease the *Kₘ* while the *Vₘₐₓ* of PEPC is only modestly affected (Duff et al., 1995; Tovar-Méndez et al., 2000; Takahashi-Terada et al., 2005). Interestingly, the knockdown of PEPC phosphorylation in *Flaveria bidentis* (*C₄*) by RNA interference inhibition of PPCK did not affect the CO₂ assimilation rate, although the response to malate was observed as in previous studies (Furumoto et al., 2007). The latter indicates that phosphorylation is probably required to adjust PEPC activity in response to various signals or fluctuating conditions or to potentially mediate the interaction with another protein, such as 14-3-3 proteins (O’Leary et al., 2011; Grieco et al., 2012). The implicated phosphorylation site is located on the N terminus of PEPC and has been identified as Ser-6, Ser-8, Ser-11, or Ser-15 among various plants (Jiao and Chollet, 1990; Wang et al., 1992; Lepiniec et al., 1994; Chollet et al., 1996; Vidal and Chollet, 1997; Tripodi et al., 2005).

PEPC phosphorylation is carried out by a kinase named PPCK found to range in size from 30 to 39 kD in diverse species (Chollet et al., 1996). PPCK is most similar to a Ca²⁺-independent Ser/Thr protein kinase. It has been classified as belonging to the Ca²⁺-sensing calmodulin-like regulated calcium-dependent protein kinase SNF1-related kinase superfamily (Hartwell et al., 1996, 1999;
Halford and Hardie, 1998; Hrabak et al., 2003). The kinase is composed of a single core domain without broader N- and C-terminal extensions. The interaction between the kinase catalytic domain and the substrate has been studied for other CDP kinases whose catalytic domain is composed of two lobes: an N-terminal lobe containing β-sheets and a C-terminal lobe composed of α-helices. It is thought that the ATP molecule is bound in a cleft between these two lobes, while the substrate binds along the cleft (Ubersax and Ferrell, 2007). On the substrate protein, the amino acids flanking the phosphorylation site, probably up to four residues on either side, are thought to contribute to substrate recognition. However, residues farther away from the phosphorylation site can interact with portions of the kinase just outside the active site (Ubersax and Ferrell, 2007). Echevarria and Vidal (2003) list the consensus domain acid-base-X-X-S-I-D-A-Q-L-R as characteristic for phosphorylation by PPCK.

The protein amount of PPCK seems to be regulated at the level of transcription and protein synthesis/turnover in C₄ maize and seeds from C₃ plants (Jiao et al., 1991; Hartwell et al., 1999; Osuna et al., 1999; Echevarria and Vidal, 2003). In addition, the activity and/or expression of the PPCK protein are known to be influenced by the reduction of disulfide bonds (Saze et al., 2001; Tsuchida et al., 2001; Nimmo, 2003), a proteinaceous inhibitor (Nimmo, 2000), and inhibition by malate (Borland et al., 2001; Nimmo, 2003), a proteinaceous inhibitor (Nimmo, 2000), and inhibition by malate (Borland et al., 2001; Nimmo, 2003). Echevarria and Vidal (2003) list the consensus domain acid-base-X-X-S-I-D-A-Q-L-R as characteristic for phosphorylation by PPCK.

PPCK proteins were found to be encoded by a small gene family (PPCK), which in most investigated species is composed of at least two isoforms (Fontaine et al., 2002; Nimmo, 2003; Fukuyama et al., 2006; Shenton et al., 2006). The genus Flaveria is an exception, with only one characterized PPCK gene, encoding a 31.8-kD protein identified in F. trinervia and F. bidentis (Tsuchida et al., 2001; Nimmo, 2003), a proteinaceous inhibitor (Nimmo, 2000), and inhibition by malate (Borland et al., 2001; Nimmo, 2003). Echevarria and Vidal (2003) list the consensus domain acid-base-X-X-S-I-D-A-Q-L-R as characteristic for phosphorylation by PPCK.

RESULTS

PPCK Is Encoded by a Two-Member Gene Family in Flaveria spp., and Both Genes Were Altered in the Process of C₄ Evolution

To investigate the number of PPCK isoforms in Flaveria spp., previous leaf transcriptome data for the two C₃ species F. pringlei and Flaveria robusta, the C₅-C₄ intermediate F. ramosissima, and the two C₄ species F. trinervia and F. bidentis (Gowik et al., 2011) were complemented by RNA sequencing (RNAseq) data that included four additional C₃-C₄ intermediates, namely Flaveria chilonefolia, Flaveria pubescens, Flaveria anomala, and Flaveria brownii (Mallmann et al., 2014). In all nine Flaveria spp., we consistently found two PPCK isoforms (Supplemental File S1). A phylogenetic analysis of the PPCK proteins showed that the two isoforms clearly differ from each other and belong to two separate branches of a maximum-likelihood phylogram (Fig. 1). We will thus refer to them as PPCKA and PPCKB. The PPCKA isoform branching reflects the photosynthetic type of the species. At the base of the branch are the sequences from the C₃ plants, and the evolutionary time increases toward the sequence form of the C₄ species. In the case of PPCKB, there is high similarity between the PPCKB isoforms from the species with C₃ and intermediate photosynthesis types. These results are similar to previous phylogenetic analyses of PEPC isoforms in Flaveria spp. (Svensson et al., 2003; Fig. 1).

Extending the analysis to a selection of known PPCKs from C₃, CAM, and C₄ species, we found that the PPCKs cluster primarily based on their systematic relationships rather than photosynthetic type or PPCK isoform (Supplemental File S2). All PPCK isoforms from the genus Flaveria cluster together, just as the PPCKs from Magnoliophyta and the rosids do. Second, within the systematics groups, further organization driven by PPCK isoform is observed with Flaveria spp., Fabales (Lotus and Soja spp.), and the order Brassicales (Arabidopsis, C₅ Cleome gynandra, C₄ Bräutigam et al., 2011). Similarly, in the family Poaceae, here represented by maize (C₄), sorghum (C₄), Setaria italica (C₄), and rice (Oryza sativa; C₃), all PPCKs are grouped together, but within this clade the subdivision is driven by the high similarity of PPCK1 proteins. These data indicate that the PPCK isoforms evolved independently, after speciation of the highly divergent species (Supplemental File S2).

In order to pinpoint any evolutionary changes among the Flaveria spp. that were associated with the transition from C₃ to C₄ photosynthesis in this genus, we compared the amino acid differences in the PPCK sequences. Amino acid exchanges of interest were de

In line with the parallel evolution of the PPCK isoforms seen in the phylogenetic analysis (Supplemental File S2), these amino acid exchanges seem characteristic for the genus Flaveria (Table I; Supplemental File S3). A comparison of sequences from the order Brassicales and the family Poaceae, both of which have C₃ and C₄ representatives, shows a similar trend of C₃-to-C₄ exchanges primarily for PPCK1 (Supplemental File S3). However, it must be noted that for Brassicales, only single PPCK1.
representatives from C₃ and C₄ species are available, and in Poaceae, we have a single C₃ PPCK1 representative from rice (Supplemental File S2). It is thus difficult to conclude if the observed differences are conserved among related species with the same type of photosynthesis or if we are seeing species-specific characteristics. The presence of the three PPCK1 sequences in Poaceae originating from C₄ species that all show the same amino acid difference from the rice PPCK1 sequence supports the hypothesis that we are observing C₃ to C₄-related changes, as in the genus Flaveria (Table I).

In order to visualize the distribution of these exchanges in Flaveria spp., we first summarized the mutations in a schematic model (Fig. 2B) and could observe that most of the conserved differences between C₃ and C₄ species are in the C-terminal part of the kinase. Subsequently, the tertiary structure of PPCKA from F. pringlei (C₃) and F. trinervia (C₄) was modeled using SWISS-MODEL (de Brevern et al., 2000; Joseph et al., 2010; Benkert et al., 2011); the models were aligned in iPBA (Gelly et al., 2011) and displayed using PyMol version 1.3. The sequence identity of the template (Protein Data Bank identifier 3ku2.1.A) used for modeling was 38.49% and 39.62% for F. pringlei (C₃) and F. trinervia (C₄), respectively (Fig. 2C). The structure of the kinase was composed of N- and C-terminal lobes with an unstructured protein region between them. The cluster of amino acid exchanges on positions 135, 160, 163, 165, and 185 was found on the unstructured protein region in the three-dimensional model, while most other exchanges were located on the C-terminal lobe (Fig. 2, B and C).

For PPCKB, only five amino acid exchanges were found among C₃ and C₄ species according to the above criteria. I₈₇M, S₁₀₉A, and I₂₃₃M are found in some intermediates and the C₄ species (type 2), while T₄₉I and S₁₂₃G are found in the C₄ species only (type 3; Fig. 2A; Table I). The smaller number of differences in this isoform, which are related to the photosynthetic type of the species, illustrates that higher evolutionary pressure was put on optimization of the structure of PPCKA.

**PPCK Genes Have Diel Expression Patterns in Flaveria spp., and Their Levels Vary between C₃ and C₄ Species**

The differences in the protein sequences among PPCK isoforms led to the question of whether changes in the regulation of transcript abundance were also present among species with C₃ and C₄ photosynthesis. Therefore, we investigated PPCK transcript levels by a 24-h harvest experiment and profiled the samples by Illumina RNaseq. Plants from F. trinervia (C₄) and F. pringlei (C₃) were grown in a 10/14-h (light/dark) period, and harvesting was performed every 4 h from the onset of light (Fig. 3).
Figure 2. Amino acid exchanges between *Flaveria* spp. with C₃ and C₄ photosynthesis. A, Multiple sequence alignment of PPCK A and B isoforms from nine *Flaveria* spp. using MUSCLE (Edgar, 2004). B, Schematic model of PPCKA from *F. pringlei*, with marked positions of amino acid exchange among C₃ and C₄ species for both isoforms. C, Three-dimensional model of PPCKA from *F. trinervia* (blue) aligned on the model from *F. pringlei* (red); the positions of the amino acid exchanges are...
Figure 3 depicts that PPCKA and PPCKB transcripts exhibit strikingly different patterns of accumulation. Both PPCKA and PPCKB transcripts, in *F. trinervia* (C₄) as well as in *F. pringlei* (C₃), fluctuate with a phase shift of 12 h. Both transcripts can be classified as having diel regulation, since a significant difference in the P values (P ≤ 0.05) and at least a 3-fold change can be simultaneously observed between a reference time (4 h of light) and three consecutive time points (Facella et al., 2008). However, PPCKA and PPCKB transcripts differ in the temporal order of maxima and minima. While PPCKA transcript amounts of *F. trinervia* and *F. pringlei* reach their maximum during the day, PPCKB transcripts peak in the dark.

The abundance of PPCKA transcripts differs between species with varying photosynthesis types. At the time of expression maximum (4 h after the onset of light), the amounts of PPCKA transcripts are at least 10-fold higher in *F. trinervia* (C₄) than in *F. pringlei* (C₃). As can be seen in Figure 3B, the relative levels in *F. pringlei* seem to drop a bit faster than in the C₄ species, resulting in higher differences in later time points (Fig. 3; Supplemental File S4).

This increase in PPCK expression was only found in PPCKA. In the case of the PPCKB transcript, we found that the amounts in *F. pringlei* (C₃) are generally higher than in *F. trinervia* (C₄) and that the differences among species are less expressed (Fig. 3C; Supplemental File S4). During the night, at the time of the expression maximum for PPCKB, this transcript is 2 to 6 times higher in the C₃ species. The only time point that showed higher levels of PPCKB in the C₄ species was at the end of night at 0/24 h (Fig. 3D).

Comparison of the isoforms transcripts within each species shows that, at midday, the level of PPCKA in *F. trinervia* (C₄) is more than 160 times higher than the level of PPCKB, whereas in *F. pringlei* (C₃), the PPCKA transcript is only 5 times more abundant than PPCKB. During the night, PPCKB is twice as abundant as PPCKA in *F. trinervia* (C₄) and 44 times more abundant in *F. pringlei* (C₃).

In the context of the evolution of C₃ photosynthesis, it is the PPCKA isoform that shows the characteristic increase in transcription seen for many transcripts involved in this process (Gowik et al., 2011; Christin et al., 2013), suggesting that the PPCKA isoform is responsible for the daytime phosphorylation of PEPC. On the other hand, in C₄ species, there is a decrease in the amount of PPCKB transcripts during the night, while this isoform is still relatively highly expressed in C₃ species. This supports an hypothesis for the differential importance of these isoforms in the two compared species, which might be reflected in the levels of PEPC phosphorylation.

The Phosphorylation Pattern of PEPC from Species with C₃ and C₄ Photosynthesis Is Consistent with the Characteristic Expression Patterns of the PPCK Isoforms

To confirm the above hypothesis, we started two approaches: (1) investigation of the protein levels of PPCKA, and (2) monitoring of the in vivo phosphorylation state of PEPC in *F. trinervia* (C₄) and *F. pringlei* (C₃). In the first approach, we generated antibodies against PPCKA isoforms, which showed satisfactory specificity; however, due to their very low sensitivity, they were not usable in a complex plant extract and could only be applied against higher amounts of recombinant protein (Supplemental File S5). We then attempted to monitor the kinase protein using selected reaction monitoring (SRM; Lange et al., 2008), but the kinase levels were below the detection limit as well.

Consequently, we took the transcript levels of PPCK as a proxy for the protein amount and proceeded to monitor the in vivo phosphorylation state of the substrate using the phosphorylated and nonphosphorylated versions of the peptide, which contain the Ser-11 phosphorylation sites LA(pS)IDAQLR and LASIDAQLR, respectively (Svensson et al., 1997; O’Leary et al., 2011; Supplemental File S6). Using the plants harvested over the course of 24 h, we extracted proteins and spiked known concentrations of the labeled synthetic peptides mentioned above. The signal from the standard was used to determine the amount of the native PEPC peptide, following which the ratio of phosphorylated to nonphosphorylated PEPC was calculated.

Figure 2. (Continued.) marked in yellow and numbered according to the alignment shown in A. The protein kinase catalytic core domain starts at amino acid position 15 at the N terminus and ends at position 268 at the C terminus (according to BAB71853 and BAF4832). Amino acids involved in ATP binding are indicated in red, and those involved in substrate binding are marked in light blue. Amino acids in green are involved in ATP and substrate binding; dark blue boxes refer to the A-loop. Highlighted background or stars indicate amino acid residues that differ between species with C₃ and C₄ photosynthesis. PPCK sequences from C₄ species were taken from the accessions BAB71853 and BAF4832. The remaining *Flaveria* spp. were cloned and/or obtained from RNAseq. Sequences used for the alignment and modeling are listed in Supplemental File S1. Species are indicated as follows: C₄ species: *F. trinervia* (Ft) and *F. bidentis* (Fb); C₃-like species: *F. brownii* (Fbr); C₃ species: *F. ramosissima* (Fra), *F. anomala* (Fa), *F. pubescens* (Fpu), and *F. chloraeoloides* (Fch); and C₄ species: *F. pringlei* (Fp) and *F. robusta* (Fro). The three-dimensional models are calculated using SWISS-MODEL (Schwede et al., 2003; Arnold et al., 2006; Kiefer et al., 2009) using the sequences for FtPPCKA and FpPPCKA listed in Supplemental File S1. PPCKAs from *F. anomala* (Fch) and *F. robusta* (Fro) were modeled onto a calmodulin domain protein kinase 1 from *Toxoplasma gondii*, with sequence similarities of 0.4 (sequence identity of 38.49%) and 0.4 (sequence identity of 39.62%), respectively. The models with highest Qualitative Model Energy Analysis values (P ≤ 0.85 and 0.78 for FpPPCKA and FtPPCKA, respectively) were aligned in iPBA (de Brevern et al., 2000; Joseph et al., 2010; Benkert et al., 2011), and the resulting file was displayed in PyMol version 1.3 (http://www.pymol.org/).
The results in Figure 4 show one crucial difference between F. trinervia (C4) and F. pringlei (C3): while the phosphorylation of the PEPC from the species with C4 photosynthesis peaks during the day, the PEPC from the C3 species does not; instead, it maintains a certain level of phosphorylation throughout the entire day, and this level drops only briefly after the end of night.

Comparison of the phosphorylation trend for PEPC from F. trinervia (C4) with the transcript levels of PPCKA and PPCKB from this species shows that the peak of phosphorylation closely matches the accumulation pattern of the PPCKA transcript, while at night the phosphorylation levels are negligible (Fig. 4A; Supplemental File S4).

Comparison of the phosphorylation trend for PEPC from F. pringlei (C3) with the transcript levels of PPCKA and PPCKB, and the maximum phosphorylation point is during the dark period (Fig. 4B; Supplemental File S4).

The accession numbers for the sequences are listed in Supplemental File S2, except for C. gynandra; the sequence for C. gynandra is included at the end of Supplemental File S1.

<table>
<thead>
<tr>
<th>Table 1. Amino acid exchanges in PPCK isoforms occurring among C3 and C4 Flaveria spp. compared with C3 and C4 species from Brassicales and Poaceae</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino acids are listed using standard single-letter code, with positions marked according to the F. pringlei PPCKA sequence. An alignment of all sequences is presented in Supplemental File S3. Boldface letters indicate amino acid exchanges found in some intermediate species and species with C4 photosynthesis; italic letters indicate amino acid exchanges found only in species with C3 photosynthesis. Species indicated with single asterisks perform C3 photosynthesis; those indicated with double asterisks perform C4 photosynthesis; all others are C3-C4 intermediates and the C4-like F. brownii (Fbr). Dashes indicate that sequence information is missing or a residue is not present. Species abbreviations are as follows: Arabidopsis (At), C. gynandra (Cg), maize (Zm), rice (Os), S. italica (Si), and sorghum (Sb). The accession numbers for the sequences are listed in Supplemental File S2, except for C. gynandra; the sequence for C. gynandra is included at the end of Supplemental File S1.</td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td><strong>Species</strong></td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Flaveria spp.</td>
</tr>
<tr>
<td>Fp*</td>
</tr>
<tr>
<td>Fro*</td>
</tr>
<tr>
<td>Fc</td>
</tr>
<tr>
<td>Fa</td>
</tr>
<tr>
<td>Fpu</td>
</tr>
<tr>
<td>Fra</td>
</tr>
<tr>
<td>Fbr</td>
</tr>
<tr>
<td>Flh**</td>
</tr>
<tr>
<td>Ft*</td>
</tr>
<tr>
<td>Fp*</td>
</tr>
<tr>
<td>Fro*</td>
</tr>
<tr>
<td>Fc</td>
</tr>
<tr>
<td>Fpu</td>
</tr>
<tr>
<td>Fra</td>
</tr>
<tr>
<td>Flr</td>
</tr>
<tr>
<td>Fly**</td>
</tr>
<tr>
<td>Ft**</td>
</tr>
<tr>
<td>Brassicales</td>
</tr>
<tr>
<td>At*</td>
</tr>
<tr>
<td>Cg**</td>
</tr>
<tr>
<td>At*</td>
</tr>
<tr>
<td>Poaceae</td>
</tr>
<tr>
<td>Os*</td>
</tr>
<tr>
<td>Zm**</td>
</tr>
<tr>
<td>Sh**</td>
</tr>
<tr>
<td>Si**</td>
</tr>
<tr>
<td>Bd*</td>
</tr>
<tr>
<td>Bd*</td>
</tr>
<tr>
<td>Os*</td>
</tr>
<tr>
<td>Sh*</td>
</tr>
<tr>
<td>Zm**</td>
</tr>
<tr>
<td>Si**</td>
</tr>
<tr>
<td>Os*</td>
</tr>
<tr>
<td>Bd*</td>
</tr>
<tr>
<td>Sh**</td>
</tr>
<tr>
<td>Zm**</td>
</tr>
<tr>
<td>Si**</td>
</tr>
</tbody>
</table>

The accession numbers for the sequences are listed in Supplemental File S2, except for C. gynandra; the sequence for C. gynandra is included at the end of Supplemental File S1.
In the context of C₄ evolution, these results show that the Flaveria spp. plants with C₄ photosynthesis probably abolished the phosphorylation of PEPC in the dark, while this action is still important in C₃ species, where PEPC does not play a photosynthesis-related role but functions in other primary metabolic pathways.

Functional Characterization of Recombinant PPCKA Proteins Shows Cross Phosphorylation of PEPC among Flaveria spp. and Indicates Changes in Activity

The first part of “Results” outlined the changes that have occurred in the amino acid sequence of PPCKA, which we believe is responsible for the phosphorylation of the photosynthetic PEPC isoform in F. trinervia. As can be seen in Figure 2B, many of these amino acid exchanges occurred near or at sites that are implicated in binding either ATP or PEPC. Thus, we investigated if these changes are sufficient to abolish substrate recognition between species with C₃ and C₄ photosynthesis.

To this aim, we first isolated the PPCKA complementary DNAs (cDNAs) from F. pringlei (C₃) and F. trinervia (C₄), based on existing studies (Tsuchida et al., 2001) and our sequencing data, and purified the recombinant PPCKA proteins to near homogeneity. In two independent experiments, either recombinant ppcA PEPC from F. pringlei (C₃) or F. trinervia (C₄) was used as a substrate (obtained from Paulus et al., 2013b). Regardless of the origin of PEPC, there was successful phosphorylation by both F. pringlei and F. trinervia PPCKA recombinant proteins (Fig. 5). Consequently, the interspecies cross phosphorylation is not a trait restricted only to the C₄ species.

We found that the signals from the combination of F. trinervia (C₄) PPCK and PEPC were much stronger than the signals resulting from the phosphorylation of the F. trinervia (C₄) PEPC with the F. pringlei (C₃) PPCK (Fig. 5B). In comparison, even though both kinases could
successfully phosphorylate the PEPC from *F. pringlei* (C₃), the signals were of similar intensity (Fig. 5A).

Consequently, although the PPCKs can cross-phosphorylate among species, the stronger signals from the *F. trinervia* (C₄) PPCKA and PEPC suggest that the changes in both proteins are important for the observed difference in phosphorylation.

To identify possible changes near or on the substrate recognition site, we investigated the N terminus of PEPC in the nine sequenced *Flaveria* spp. and the PEPCs from Brassicales, Amaranthaceae, and Poaceae, all of which contain C₃ and C₄ species. The conserved motif from this interspecies comparison was identified as E-K-X-X-S-I-D-A-Q-L-R (Fig. 6A; Supplemental File S7; Bailey and Elkan, 1994; Echevarria and Vidal, 2003).

Focusing on the genus *Flaveria*, where the N-terminal region of the photosynthetic *ppcA* PEPC is well conserved, we found a single L6V exchange, which is different between the C₃ and C₄ *Flaveria* spp. (Fig. 6B). The species in the genus are phylogenetically subdivided into two clades, A and B, with the C₄ species belonging to clade A, whereas the furthest evolved species in clade B is the C₃-like *F. brownii* (McKown et al., 2005; Gowik et al., 2011). The exchange L6V was found in the C₄ species *F. trinervia* (C₄) and *F. bidentis* (C₄), in clade A, as well as in *F. brownii* (C₃-like) and the C₃-C₄ intermediate species *F. pubescens* and *F. anomala*, all belonging to clade B. Thus, Val has been introduced in the further evolved C₃-C₄ intermediates and the C₄ species in two independent evolutionary paths. Furthermore, the same type of exchange was observed in the C₄ species in Amaranthaceae, which represent a totally different origin of C₄ photosynthesis, but was not found in the Brassicales or Poaceae (Supplemental File S7).

Surprisingly, a second amino acid exchange, Q15H, was discovered in the C₄ nonphotosynthetic *ppcB* PEPC isoforms in *Flaveria* spp. (Fig. 6A; Supplemental File S7).

Using the recently deduced crystal structure of PEPC from plants (Paulus et al., 2013b), we used virtual mutagenesis to display the above-mentioned C₄-related changes (L6V and Q15H) onto the *F. pringlei* PEPC tetramer (Fig. 6C). This in silico investigation shows that, in both cases, the amino acids lie on the surface of the protein and are exposed to the solvent. Consequently, they could be accessible to interact with a PPCK. These findings show that the C₄-related changes, which have been identified to take place in PEPC, also extend to its N terminus.

**DISCUSSION**

This study underlines the existence of a two-member PPCK family in the genus *Flaveria* whose isoforms are expressed in an inverse day/night manner. PPCKA underwent changes characteristic for proteins recruited to C₄ photosynthesis in *F. trinervia* (Gowik et al., 2011; Christin et al., 2013), while PPCKB seems to have a higher importance in the C₃ *Flaveria* spp.

**Parallel Evolution of PEPC and PPCK Proteins**

Phylogenetic analyses for PPCK (Supplemental File S2) and PEPC (Bläsing et al., 2002; Svensson et al., 2003) show that, in both cases, the ancestor proteins diverged at the time of speciation and the respective protein isoforms evolved in parallel within each species. This is in agreement with the polyphyletic evolution of C₄ photosynthesis, where multiple independent origins led to the full C₄ photosynthesis (Sage et al., 2012).
PPCK isoforms and the \textit{ppcA} PEPC from species with \textit{C}_4 photosynthesis are always found on branches marking the longest evolutionary distance. This illustrates (1) the ongoing evolutionary process of both kinase and substrate (Wang et al., 2009) and (2) the presence of a large number of \textit{C}_4-related amino acid exchanges, as we were able to demonstrate with the case of PPCKA (Fig. 2A). PPCKA and \textit{ppcA} PEPC are thus indicated as the kinase-substrate pair that was under strong evolutionary pressure.

In contrast, PPCKB does not show large differences between species with \textit{C}_3 and \textit{C}_4 photosynthesis (Fig. 1), nor do the \textit{ppcB} and \textit{ppcC} PEPC isoforms from \textit{F. trinervia} (\textit{C}_4), which is marked by shorter evolutionary distance in previous studies (Bläsing et al., 2002; Svensson et al., 2003).

Changes in the Quantity of PPCK Transcripts in Leaves of \textit{C}_3 and \textit{C}_4 Flaveria spp.: Potential Mechanisms for the Regulation of Transcript Stability and Transcription Rhythms

Suitable expression patterns seem to be one of the prerequisites necessary for the recruitment of genes in the \textit{C}_4 pathway (Christin et al., 2013). We showed that PPCKA was present in the \textit{C}_3 Flaveria spp. that diverged earlier in evolution, and already there it has a daytime maximum in transcript accumulation. This expression pattern is much more suitable for \textit{C}_4 photosynthesis than the nighttime peak in transcription found for PPCKB. Similar to the case of \textit{ppcA} (Ernst and Westhoff, 1997), PPCKA also showed the characteristic increase in expression between \textit{C}_3 and \textit{C}_4 species. Taken together with the phylogenetic analysis, this change in transcription supports the conclusion that PPCKA is the isoform that was recruited in the transition from \textit{C}_3 to \textit{C}_4 photosynthesis mode.

PPCKB has opposite characteristics from PPCKA. In addition to the maximal transcript accumulation at night in both species, a decrease in transcript levels is observed from the \textit{C}_3 to the \textit{C}_4 species (Fig. 3). Thus, it appears that this isoform has decreased importance in the leaves of the newly evolved \textit{C}_4 Flaveria spp., which was also reflected in the decreased phosphorylation levels of PEPC during the night. On the other hand, PPCKB is fully functional in the \textit{C}_3 species, where nighttime phosphorylation is relatively abundant. As a side finding, the existence of daytime- and nighttime-expressed PPCK in the \textit{C}_3 species, especially if this expression pattern for the PPCK family is confirmed for other \textit{C}_3 plants, is an explanation for the varying PEPC phosphorylation results observed among various \textit{C}_3 plants (Fukayama et al., 2006; Meimoun et al., 2009).

PEPC has been found to be under the influence of the circadian clock in a number of plants (O'Leary et al., 2011). The closely related function of PPCKA in Flaveria spp., in addition to the similar diel expression pattern in \textit{C}_4 photosynthesis, allows for speculation that the same will be the case for this kinase. Indeed, extended light experiments in CAM plants show such regulation for PPCK (Taybi et al., 2000), but no evidence for circadian regulation has been found in maize (\textit{C}_4; Shenton et al., 2006). In soybean (\textit{Glycine max}; \textit{C}_3), PPCK has been found under circadian regulation in leaves but not in roots (Sullivan et al., 2005), and in rice (\textit{C}_3), it appears that there

---

**Figure 5.** In vitro phosphorylation of \textit{ppcA} PEPC protein from Flaveria spp. Recombinant \textit{ppcA} PEPC proteins from species with \textit{C}_3 and \textit{C}_4 photosynthesis were phosphorylated in vitro using radioactive ATP and recombinant PPCK from two \textit{Flaveria} spp. PPCKs are not specific to the PEPC from the respective species and can cross-phosphorylate in vitro. A, The phosphorylation of the \textit{F. pringlei} \textit{ppcA} PEPC (1 \textmu g) was tested, with equal amounts of purified recombinant kinase based on Supplemental File S5: FpPPCKA (25 \mu g of total protein) and FpPPCKA (1 \mu g of total protein). B, Indications of higher activity by the PPCK from \textit{C}_4 species appeared only when FpPEPC (3 \mu g) was used as a substrate: FpPPCK (25 \mu g of total protein) and FpPPCK (1 \mu g of total protein). Each lane represents an independent PPCK expression and purification. Controls are as follows: C1, temperature-deactivated PPCK; C2, no PEPC; C3, no FpPPCK; C4, no FpPPCK.
is no circadian regulation (Fukayama et al., 2006). Additional experiments are needed to test whether the accumulation of PPCKA and PPCKB transcripts in Flaveria spp. is controlled by the circadian clock. An analysis of cis-regulatory elements in the PPCK promoter regions would also be useful to investigate any regulatory differences between Flaveria spp. with varying photosynthetic types (Mockler et al., 2007).

The main mechanisms influencing PPCKA transcript levels could be (1) increased transcription rate and/or (2) increased transcript stability. We inspected the 3' untranslated region (UTR) from the PPCK isoforms from the available data for the presence of the 10 most significant destabilization or stabilization 6-mers (Tsuchida et al., 2001; Narsai et al., 2007). On average, 80% of all identified 6-mers in the PPCKA 3' UTR were destabilization motifs implicated in rapid RNA turnover, while this number was a little lower (72%) for PPCKB (Supplemental File S1; disregarding the partial UTR sequences). Thus, in both cases, we found indications for rapid transcript turnover that agree with a previously postulated theory (Tsuchida et al., 2001).

The Evolutionary Changes in PPCKA Proteins and Their Potential Functional Significance

Eight of the 13 amino acid exchanges conserved among the PPCKA proteins from C₃ and C₄ plants are concentrated in the region of amino acids 135 to 217 (Fig. 2). Some amino acids in this region are tentatively indicated to participate in ATP and substrate binding based on the National Center for Biotechnology Information Conserved Domains Database (GenBank accession nos. BAB71853.1 and BAF48321.1; accession no. cd05123; Marchler-Bauer et al., 2013). This is an indication of the strong evolutionary pressure to optimize the kinetic properties and/or the activation mechanism of the kinase.

Some of the potentially important differences are at positions 147, 160 to 165, and 211. At position 147 there is a Ser-to-Gly exchange. In contrast to the nonpolar amino acid Gly, Ser is a polar amino acid that can be subject to phosphorylation. Moreover, intramolecular or intermolecular hydrogen bonds can be formed with the Ser side chain. At positions 160, 163, and 165, three amino acid exchanges were identified. The first two are exchanges among amino acids with similar chemical properties, but the last exchange, at position 165, replaces Gly, a small nonpolar amino acid, with Arg, a larger basic (polar) amino acid, whose side chain is fully protonated at neutral pH, consequently bringing in positive charge at this position (Taniguchi, 2010). This cluster of point mutations is followed by an Arg-to-Met exchange at position 185. In the three-dimensional structure, Arg-185 is found opposite the region 160 to 165. Together with a third amino acid exchange at position 135, these amino acid exchanges are predicted to lie in an unstructured protein region, partly annotated as the activation loop (A-loop; Fig. 2). Current structural annotations state that phosphorylation of the A-loop results in transition from a disordered to an ordered state of the kinase and is part of the kinase activation mechanism (accession no. cd05123; Marchler-Bauer et al., 2013). As the putative A-loop contains Ser (positions 155 and 168), Thr (positions 166 and 173), and Tyr

Figure 6. Analysis of the phosphorylation motif on the N terminus of PEPC. The first 18 to 21 amino acids (based on the sequence alignment in Supplemental File S7) were analyzed for conserved motifs using the online tool MEME (http://meme.nbcr.net/meme/; Bailey and Elkan, 1994). A, The consensus motif E-K-X-X-S-I-D-A-Q-L-R was found in 50 of the 56 submitted sequences from Flaveria spp., Brassicaceae, Amaranthaceae, and Poaceae. Several C₃- to C₄ amino acid exchanges can be observed in front of the phosphorylated Ser (Supplemental File S7). B, The consensus sequence for Flaveria shows the exchanges L6V (found in C₄ ppcA PEPC) and Q15H (found in C₄ ppcB PEPCs), which differ between C₃ and C₄ species. C, Virtual mutagenesis was applied to the crystal structure of the PEPC tetramer from F. pringlei, which starts at position Leu-6. The applied mutations L6V and Q15H lie on the surface of the molecule and are accessible to the solute. Shades of gray indicate the individual monomers as part of the PEPC tetramer; Val, cyan; Ser-11, yellow; His, green.
(position 175), which are all susceptible to phosphorylation, the nearby exchanges might influence the interaction with another regulating protein. Toward the C terminus, at position 211, instead of a Tyr in C₃ plants, which has been indicated to play a part in substrate binding, is a His (a basic amino acid) in C₄ plants, which is known to function as an acid/base catalyst (Taniguchi, 2010). However, we also must bear in mind that the crystal structure onto which PPCKA was modeled has sequence identity of about 40% with the Flaveria spp. PPCKAs and that PPCK has not been crystalized from any plant species. The latter will be necessary for proper determination of the influence of these amino acid exchanges in the tertiary structure of PPCK.

If we assume that the evolutionary trajectories are accompanied by gains of fitness and that the evolution of the PPCK gene family optimized the C₄ evolutionary processes (Sage et al., 2012; Heckmann et al., 2013), then amino acid exchanges observed in the PPCKA sequences from various species could indicate the order of structure optimization. Starting from a common ancestor in the genus Flaveria, species with C₃ photosynthesis diverged earlier, whereas the C₄ species diverged later. If an amino acid exchange is conserved among all intermediates and the C₄ species (type 1), then this point mutation probably occurred before those found only in the C₃ species (type 3; Table 1). Thus, in the case of PPCKA, we believe that the exchanges E80D, G160A, G165R, R185M, and L257F (type 1) probably precede the amino acid exchanges A4T, I135L, S147G, Y211H, E217D, and E273K (type 3). In the two cases where the C₄ residue is found in some but not all intermediate species (E163D and S270N; type 2), it is difficult to determine what happened. Creating a series of mutated PPCKA proteins will be needed to investigate the importance of each of these amino acid exchanges for the kinase enzymatic activity.

The recombinant PPCKA experiments presented here show that these amino acid differences are not enough to abolish substrate recognition among the species in the genus Flaveria (Fig. 5), which is in agreement with previous findings where PPCK proteins from several plants can phosphorylate exogenous PEPC (Li and Chollet, 1994; Tsuchida et al., 2001; Ermolova et al., 2003). Since a higher signal results from the C₄ form of both kinase and substrate, it is the parallel optimization of these molecules that is responsible for this increased activity (Fig. 5B). In this context, the so-far ignored amino acid differences in the N terminus of PEPC, and in the vicinity of the phosphorylated Ser (Fig. 6; Supplemental File S7), gain in importance and will require further study (Echevarria and Vidal, 2003). Simulation of the C₂-to-C₄ exchanges on the N-terminal part of PEPC that take into account that PEPC is a tetramer show that, despite the macromolecular conformation, these amino acids would still be accessible for protein-protein interaction (Fig. 6C; Paulus et al., 2013b). The L6V exchange on the N terminus of ppcA PEPC, characteristic for the intermediate and C₄ species, is one of two exchanges in the otherwise conserved N-terminal region in Flaveria spp. (Supplemental File S7). The side chains of these two amino acids have similar properties, but the smaller side chain from Val might allow easier uncoupling of the kinase from the substrate, by providing larger separation between the PEPC- and PPCK-interacting residues. Consequently, this might be one of the reasons for observing a stronger signal in the in vitro phosphorylation assay between the C₄ recombinant PEPC and PPCK proteins. ppcB PEPC, on the other hand, showed a Q15H exchange only in the C₄ species (Fig. 6). This introduces a positive charge close to Ser-11, but the true consequence of this exchange will only be known after further investigation, as it will depend on the interacting residue from PPCK. Cocrystallization of both proteins should uncover the amino acid residues involved in the protein-protein interaction and allow the investigation of the observed amino acid exchanges between isoforms from C₃ and C₄ plants.

Although the evolution of the photosynthetic PEPC from C₄ Flaveria spp. has been studied extensively and the influence of crucial amino acid exchanges on enzymatic activity has been clarified (Bläsing et al., 2000; Engelmann et al., 2002, 2008; Paulus et al., 2013a, 2013b), no such information exists for PPCK at this time. Additionally, as the cumulative findings from previous and these studies reinforce the differences in phosphorylation patterns of PEPC from C₃ and C₄ species, and at a time with ongoing efforts for engineering C₃ traits in C₄ crops, we can only stress the need for detailed knowledge of the regulation of one of the key enzymes of the C₄ pathway.

MATERIALS AND METHODS

Plant Material

Plants were initially grown in greenhouse conditions, and before harvest for the Illumina and SRM experiment, they were acclimated for the last 10 d in an open phytocam at 120 μmol quanta m⁻² s⁻¹ irradiance with a 10-h photoperiod, at 19°C night and 22°C day. Leaf blades were harvested from wild-type 4- to 6-week-old plants from Flaveria trinerea, Flaveria ramosissima, and Flaveria pringlei. Leaf samples were collected at 4-h intervals over a 24-h period. Plant material was immediately frozen in liquid nitrogen for further storage at −80°C. Three separate biological replicates for each time point were ground in liquid nitrogen prior to the respective RNA and protein preparations.

Library Construction and Illumina Sequencing

F. trinerea and F. pringlei were used for transcriptome sequencing. Total RNA was extracted from Flaveria spp. leaves using TRIzol Reagent according to the protocol for plant tissue (Bioline) transcriptome sequencing. Digestion with DNase was performed for 15 min. The RNA was treated with phenol and chloroform and precipitated overnight with a sodium acetate/isopropl alcohol solution. The RNA was washed with 70% (v/v) ethanol and dissolved in water. The RNA concentration was determined with the nanodrop (Peqlab Biotechnologie).

Total RNA was set to a final concentration of 1 μg in 50 μL as a starting volume. The DNA libraries were generated according to the manufacturer’s TruSeq RNA Sample Preparation Kit via the Low-Throughput Protocol (Illumina catalog no. RS-930-2001, part no. 15008136 Rev. A, November 2010). RNA aliquots and libraries were validated for qualification and quantitation with the Agilent2100 bioanalyzer (Agilent Technologies). Validated DNA libraries were normalized and pooled to a final concentration of 2 nM. Clusters were generated with the TruSeq SR Cluster Kit version 2 according to the Reagent Preparation Guide with the Illumina cBot device. Single-read sequencing was performed with the Illumina HiSeq 2000.
Transcriptome Mapping and Analysis

The PKCK sequences for various species were obtained by aligning contigs from two de novo assembly pipelines, CLC Genomics Workbench (version 6.5) combined with the CLC Genomics Server (version 5.5; http://www.clcbio.com) and Velvet/Oases (version 1.2.08/0.2.08; Schulz et al., 2012), using the automatic k-mer determination setting with CLC and a k-mer length of 25 bp with Velvet/Oases. These were then combined and aligned to the F. trinervia cDNA (PKCK-A accession no. AB227061), allowing us to combine the contigs of the two pipelines for several species and ensure the largest possible coverage of the PKCK transcript. In case of PKCB, we used a contig generated by de novo assembly of 454 and Illumina reads (Gowik et al., 2011; J. Mallmann, D. Heckmann, A. Beutig, M. J. Lercher, A. P. M. Weber, P. Westhoff, and U. Gowik, unpublished data). To obtain the abundance of reads for the kinase, clean Illumina reads were aligned against F. trinervia cDNA sequences either obtained from public databases or during this work (Supplemental File S1). Reads were mapped with Bowtie2 (version 2.0.6) using the end-to-end mode with sensitive settings and allowing one mismatch per seed for multiseedling (Langmead and Salzberg, 2012). We counted the number of unambiguous best hits per transcript and normalized these values to reads per million mapped reads and kilobase transcript length (RPKM).

Means and s.e. were calculated from three biological replicates. The data were evaluated with a one-way ANOVA by multiple comparison using a Tukey test (multiplicity-adjusted P ≤ 0.05). In parallel, with a two-tailed paired Student's t test, we obtained significance for transcripts of two time points, with major differences in abundance (P ≤ 0.05). Statistical analyses were performed with the program GraphPad Prism 6 (GraphPad Software).

Isolation, Cloning, and Sequencing of PKCK for Recombinant Protein Purification

Total RNA was extracted from F. pringlei and F. trinervia leaves using the RNeasy Plant Mini Purification Kit (Qiagen), including the RNaese-Free DNase Set (Qiagen). The RNA concentration was determined with the nanodrop (Peqlab Biotechnologie). The synthesis of 5' and 3' RACE-ready first-strand DNAs was carried out according to the SMARTer RACE cDNA Amplification Kit (Clontech Laboratories) with 1 μg of total RNA using SMARTer Reverse Transcription (Clontech Laboratories). The 5' and 3' RACE-PCRs were performed using Universal Primer A Mix (Clontech Laboratories) and the Smart II-A/5'RACE oligonucleotides, respectively (Supplemental File S8). The gene-specific oligonucleotides were selected according to the available PKCK sequence from F. trinervia (Genbank accession no. AB065100). The noncoding regions of PKCK from F. pringlei were identified with the PKCK 5' end (PEPC-PK-Fp_5'RACE) and PKCK 3' end (PEPC-PK-Fp_3'RACE) oligonucleotides. PCR was performed with the Advantage 2 Polymerase Mix according to Clontech Laboratories. The obtained DNA fragments were cloned with the CloneJET PCR Cloning Kit (Fermentas). Several independent clones were verified via PCR with the oligonucleotides pETJ-fw and pETJ-rev (Supplemental File S8). Potential positive clones were verified by DNA sequencing (LG Genomics). The cDNA full-length PKCK gene from Flaveria spp. was amplified with the oligonucleotides PEPC-PK-CDs_Fp, PEPC-PK-CDs_Fra, and PEPC-PK-CDs_Fl in Supplemental File S8. Additional restriction sites were amplified to PKCK with the oligonucleotides PEPC-PK-Fp-3'Xhol and PEPC-PK-Fp-5'BamHI (F. pringlei) and PEPC-PK-Fp-5'NdeI and PEPC-PK-Fp-3'BamHI (F. trinervia), as listed in Supplemental File S8. All PCRs were carried out by means of the Phusion High-Fidelity DNA Polymerase (New England Biolabs). PCR products were sequenced, digested with Xhol/BamHI or NdeI/BamHI, and cloned into the corresponding Xhol/BamHI-digested or Xhol/ NdeI-digested pETE15b (+) vector (pET system; Novagen) together with the coding sequence of an N-terminal 6×His tag and Tobacco Etch Virus cleavage site. All generated plasmids were verified by DNA sequencing.

Alignments and Phylogenetic Analysis

Calculations of sequence identity between the PKCK proteins from Flaveria spp. were performed with the Web site http://emboss.bioinformatics.nl. The phylogenetic analysis was performed using MEGA5 (Tamura et al., 2011; Hall, 2013), as listed in the figures. Predicted protein parameters for PKCK, which has been isolated with RACE-PCR, were acquired using the ExPaSy program (http://web.expasy.org/protparam).

Heterologous Expression of Recombinant PKCK in Escherichia coli

The QIAXexpression protocol was used for the heterologous expression of native protein (Qiagen). E. coli BL21 (DE3) cells (Agilent Technologies) were transformed with PKCK-pETE15b. Culture growth was performed using 2YT medium (5 g L⁻¹ NaCl, 10 g L⁻¹ yeast extract, and 16 g L⁻¹ peptone) with ampicillin at 37°C after reaching an optical density of 0.7. Expression of recombinant proteins was induced by the addition of 500 μM isopropyl β-D-thiogalactopyranoside, and cells continued to grow for 5 h at 20°C.

Affinity Purification of the Recombinant PKCK Proteins

Recombinant PKCK protein was purified by nickel-nitrotriacetic acid affinity chromatography according to the manufacturer's instructions (Protein; Macherey-Nagel). In short, intact E. coli cells were incubated on ice and subsequently resuspended in lysis buffer (50 mM NaH₂PO₄, 300 mM NaCl, and 1 mg mL⁻¹ lysozyme, pH 8.0). The cells were disrupted by sonification. Native recombinant proteins were purified with a preequilibrated nickel-nitrotriacetic acid affinity column (Protein; Macherey-Nagel) on the Äktapure Plus system (GE Healthcare). In order to protect protein integrity, a gradient from 0 to 500 mM imidazole in 50 mL steps was used. Proteins were concentrated with the Amicon Ultra-10 centrifugal filter units (10-kD cutoff; Millipore) and desalted and buffered in 50 mM Tris and 5 mM MgCl₂, pH 8, using PD-10 columns (GE Healthcare).

The concentrations of purified recombinant proteins were determined with a standard procedure (Bio-Rad; Bradford, 1976) and examined by Coomassie Blue staining and immunoblotting by SDS-PAGE on a 12% polyacrylamide gel (Laemmli, 1970). Signals were quantified using the program Multi Gauge 3.0 (Fuji Photo Film). The blotted nitrocellulose membrane was incubated with the primary (anti-PKCK) antibody in a dilution of 1:400 for 12 h at 4°C. Immunoblot analysis was examined using the reagent of the chemiluminescence assay (SuperSignal West; Thermo Scientific).

After investigating the protein purity using SDS-PAGE and Coomassie Blue staining, the intensities of the kinase bands from F. pringlei were compared with those from F. trinervia, and the protein content used for the in vitro phosphorylation assays was adjusted accordingly to equalize the kinase amounts (Supplemental File S5).

Enzymatic Activity of PKCK

The in vitro phosphorylation activity of recombinant PKCK was detected by labeling recombinant PEPCK with ³²P in a 15-μL reaction mixture containing 50 mM Tris-HCl, 5 mM MgCl₂, 1 mM EGTA, and 1 mM dithiothreitol, pH 8. After 20 min of incubation on ice, radiolabeled ATP was added as a final step to the reaction at a concentration of 1 μCi mmol⁻¹ [³²P]ATP per reaction. The reaction was run for 15 min at 30°C and terminated by the addition of 2× SDS loading buffer (Laemmli, 1970). Samples were boiled at 100°C for 1 min and centrifuged at 10,000g for 10 min before loading on a 10% SDS polyacrylamide gel. Following electrophoresis, the radioactive bands were autographed visualized on a FLA3000 Bio-Imaging Analyzer (Fuji Photo Film).

PPCK Antibody Synthesis

Peptides conserved among the PKCK from F. pringlei, F. ramosissima, and F. trinervia were considered for the antibody synthesis. Rabbit polyclonal PKCK antibodies were raised against the peptides NH₂-CLQKEPKILHILG-CONH₂, PKCK residues 53 to 65, and NH₂-CRGLGADFDPDV-NH₂-CONH₂, PKCK residues 127 to 140, by Agrisera. An additional anti-PKCK antibody was synthesized by Agrisera that had been obtained from a mix of purified recombinant PKCK proteins from the same Flaveria spp. Unfortunately, although specific, the antibody sensitivity was determined to be above 0.6 μg, which was sufficient for recombinant protein work but not sensitive enough to detect PKCK in plant extracts.

SRM: Peptide Selection, Sample Preparation, Liquid Chromatography-Tandem Mass Spectrometry, and Quantitation

The amino acid sequences of PKCK and PEPCK were aligned, and in silico tryptic peptides conserved among F. trinervia and F. pringlei were considered for SRM. Initial testing was performed on the recombinant proteins for F. trinervia and F. pringlei both in unmodified and phosphorylated states. For the phosphorylated PEPCK test, proteins were phosphorylated as above, using 1 μg of kinase to 10 μg of PEPCK in a 7.5-μL reaction and nonradioactive ATP.
The peptides that showed good properties in the mass spectrometer were checked against the *Flaviera* spp. transcipitomes (BLASTp) supplemented with the experimentally determined sequences of PEPC and PPCK as a control. The BLASTp was performed with increasing e-values until the control peptide was found. The peptide covering the PEPC phosphorylation site is conserved among the three isoforms in *F. trinervia* (C2, Ernst and Westhoff, 1997; Supplemental File S6). The peptides then underwent a second round of selection using plant extract, and those peptides that could be successfully monitored in this complex sample were obtained as heavy peptides (1C and 15N) from Thermo Fisher Scientific. None of the peptides from PPCK could be monitored in complex plant extracts using our setup. Proteins were extracted from *F. trinervia* and *F. pringlei* as already described (Wille et al., 2011) with minor modifications. In short, 300 mg of leaf material was treated with 600 μL of buffer containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, protease inhibitor cocktail (Sigma-Aldrich), phosphatase inhibitor cocktail (Serva), and 1% (w/v) polyvinylpolypyrrolidone. Proteins were precipitated overnight at −20°C using a 5% volume of ice-cold acetone (Sigma-Aldrich). The protein pellet was dissolved in 6 M urea and 2 M thiourea (Merck and Sigma-Aldrich, respectively), and 10 μg of protein was digested and desalted as described by Arsova et al. (2012). Prior to desalting, the synthetic peptides for PEPC (listed in Supplemental File S6) were added to the samples in species-specific amounts. Per 500 ng of total protein, the amounts of synthetic peptides were as follows: *F. trinervia* (nonphosphorylated, 10 fmol; phosphopeptide, 5 fmol); *F. pringlei* (nonphosphorylated, 2.5 fmol; phosphopeptide, 1.25 fmol).

The tryptic digests were measured using the ultra-HPLC system UltraMate 3000 RSLChano in combination with a TSQ Vantage triple quadrupole mass spectrometer (both Thermo Fisher Scientific). A total of 500 ng of digested plant protein was loaded at a rate of 20 μL min⁻¹ and separated on a 20-min gradient (5%–30% B) where solution A is 0.1% formic acid in water and solution B is 80% acetonitrile, 10% trifluoroethanol, and 0.1% formic acid in water), using a column of 75-μm i.d., 15-cm length, C18, and 2-μm particle size (Accliam PepMap RSLC; Thermo Fisher Scientific). Samples were sprayed into the mass spectrometer by electrospray ionization. The resolution of Q1 and Q3 was set to 0.7 u full width at half maximum. Cycle time was 1 s (dwell time was more than 50 ms for each transition). Peptides were fragmented using argon gas for collision-induced dissociation at collision energies specific for each peptide (Supplemental File S6). Instrument scan mode was i.d., and three primary and two secondary fragment ions per peptide were collected; allowed time was 0.1 min, threshold intensity was 300 counts, and dynamic exclusion was activated after three fragmentation events for 0.3 min. Data were analyzed using FinnFpint 1.3 (Thermo Fisher Scientific), and the signal-to-noise ratio was set at less than 10. In general, three transitions per peptide for each light and heavy form were used for peak area summation [except *Flaviera pringlei* peptide LASIDAQLR (Phos), which had two transitions per peptide]. The phosphorylation of PEPC on Ser-11 was monitored to their recruitment into the C4 pathway. J Exp Bot 62: 3049–3059.

**Supplemental Data**

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

**Supplemental File S1. Sequences.**

**Supplemental File S2. Phylogenetic analysis of plant PPCKs.**

**Supplemental File S3. Alignment of PPCKs from various species.**

**Supplemental File S4. RNAseq and SRM data.**

**Supplemental File S5. Recombinant PPCKA expression and anti-PPCK antibodies.**

**Supplemental File S6. SRM peptide information.**

**Supplemental File S7. PEPC N-terminal alignment.**

**Supplemental File S8. Primers.**

**LITERATURE CITED**


Evolution of the PPCK Family in Flaveria spp.


