Volume 292, number 1,2, 133–136 FEBS 10333 November 1991

ADONIS 001457939101041K

Multiple cDNAs of phosphoenolpyruvate carboxylase in the C_4 dicot *Flaveria trinervia*

Winfried Poetsch, Jürgen Hermans and Peter Westhoff

Institut für Entwicklungs- und Molekularbiologie der Pflanzen, Heinrich-Heine-Universität, Düsseldorf, D-4000 Düsseldorf 1, Germany

Received 3 August 1991; revised version received 9 September 1991

We have isolated and characterized cDNA clones for the leaf-specific C_4-phosphoenolpyruvate carboxylase (PEPCase) from the dicotyledonous C_4 plant *Flaveria trinervia*. The isolation of multiple cDNAs indicates that in this plant the C_4 isoform is encoded by a small subgroup of the PEPCase gene family. The deduced amino acid sequence reveals a higher degree of similarity to the CAM and C_3 isozymes of the dicotyledonous, facultative CAM plant *Mesembryanthemum crystallinum* than to the C_4 PEPCases of monocotyledonous origin.

Phosphoenolpyruvate carboxylase; C_4-plant; *Flaveria*; Gene family

1. INTRODUCTION

Phosphoenolpyruvate carboxylase (EC 4.1.1.31; PEPCase) is a key enzyme of the C_4 photosynthetic pathway. It is located in the mesophyll cells and catalyzes the primary fixation of CO_2 to yield oxaloacetate. The active enzyme consists of four identical subunits, each of about 110 kDa molecular mass. The C_4 enzyme is allosterically activated by glucose 6-phosphate, a final product of CO_2 assimilation, while the immediate products of the carboxylation reaction, i.e. malate or aspartate, act as feedback-inhibitors [1,2]. The PEPCase of C_4 plants is also subject to a light/dark regulation which is mediated by reversible phosphorylation of a single serine residue in the amino-terminal part of the enzyme [3].

PEPCase activity is not only connected with C_4 syndrome. Different isoforms have been detected in various photosynthetic and non-photosynthetic tissues of C_4 as well as C_3 plants and are involved in the basic cell metabolism [5]. These isoforms can be distinguished by chromatographic, immunological and kinetic criteria [1,4]. Molecular analysis shows that the various PEPCase isoforms are products of a multi-gene family [6–8].

We are interested in understanding the molecular events underlying the evolution of C_4 from C_3 plants and are using the PEPCase gene family in the genus *Flaveria* as a model system [7]. This genus is characterized by the concomitant occurrence of C_3 and C_4 plants as well as by a broad spectrum of C_3–C_4 intermediate species. It thus offers the unique possibility to investigate the molecular basis of cell-specific gene expression in C_4 photosynthesis. In this paper we report, for the first time, the primary structure of a C_4-specific PEPCase from a dicotyledonous C_4 plant.

Correspondence address: P. Westhoff, Institut für Entwicklungs- und Molekularbiologie der Pflanzen, Heinrich-Heine-Universität, Düsseldorf, D-4000 Düsseldorf 1, Germany. Fax: (49) (211) 311 3085.

2. MATERIALS AND METHODS

Growth of plants, construction and screening of cDNA libraries, nucleotide sequencing and Northern analysis have been described [9]. Mesophyll and bundle-sheath RNA was isolated according to Höfer et al. [10]. Antiserum against the C_4 PEPCase of *Sorghum bicolor* was kindly provided by J. Vidal (Université Paris-Sud, France).

3. RESULTS AND DISCUSSION

3.1. *Isolation of multiple PEPCase cDNA clones*

A leaf-specific cDNA expression library from *F. trinervia* was screened with an antiserum against the C_4 PEPCase of the monocotyledonous C_4 plant *S. bicolor*. Five cDNA clones of different sizes were isolated. The three largest cDNA inserts of 3.2, 1.82 and 0.86 kbp size were selected for further analysis and subcloned into pBluescript KS I (Stratagene, San Diego, USA). The resulting plasmids were named pcFtppc1-1, pcFtppc8-2 and pcFtppc5-5, respectively.

Sequence analysis revealed that the carboxyterminal and the 3′-untranslated regions of the three clones were strikingly similar, although not identical (Fig. 1). The clones differ by base pair substitutions at one position in the coding and at 8 positions in the 3′ non-coding sequence, respectively (Fig. 1). The sequence downstream the translation termination codon (TAG) is shorter in pcFtppc5-5 than in the two other cDNAs. Similar results have been reported for other PEPCases

```
pcFtppc1-1   GAGGGTGATCCCTACTTGAAACAAGGAATCAGGCTGCGTGATCCGTACATCACAACCTTGAATGTATGCCAAGCTTATACCCTAAAGAGGATCCGTGACCCGAAC
pcFtppc8-2   ..........................................................................................................
pcFtppc5-5   ..................G.......................................................................................

pcFtppc1-1   TATCATGTGACATTAAGGCCTCATATTTCTAAAGAATATGCCGCTGAGCCGAGCAAACCAGCTGATGAGCTTATCCACCTGAACCCAACCAGCGAGTACGCACCC
pcFtppc8-2   ..........................................................................................................
pcFtppc5-5   ..........................................................................................................
                                                                                    Stop
pcFtppc1-1   GGTTTGGAGGACACGCTCATCTTGACCATGAAAGGGATTGCTGCTGGAATGCAGAACACCGGTTAGGT-AACCGTTATGCAAGTATGTATTTATCTTTT-TTTCA
pcFtppc8-2   ..........................................................T.....................CA.....
pcFtppc5-5   ..........................................................T.....................T.....

pcFtppc1-1   TGCTTTTGGATGTTGTGACATGCCTTGTTTACTGCATGTGTTCTTGAGTGTTTGTATGCATCATCAATCTATGCATTTCAGAAGTTGTACTGTGTGCATCTTGAT
pcFtppc8-2   ..............C...........................................................................A.....
pcFtppc5-5   ..............C...........................................................................A.....
                    Poly(A)-I                                          Poly(A)-II
pcFtppc1-1   CGTGGTTTATCCAATAAACTAAAGTCTTA-TTGTACACTTGACTACAAGTGATTATATAAATTGTTGTGCTTTCTACTTAAAAAAAAA
pcFtppc8-2   ..........T.....A....C.....................T.....................AAAAAAAAA(A)n
pcFtppc5-5   ..........T.....A....C.....................AAAAAA
```

Fig. 1. Nucleotide sequence alignment of the 3′ regions of PEPCase cDNA clones pcFtppc1-1, pcFtppc8-2, pcFtppc5-5 of *F. trinervia*. Identical residues are represented by dots; the stop codon and the two putative polyadenylation signals are underlined.

[11,12] and could be explained by alternate usage of the two polyadenylation signals caused by differences in the polyadenylation efficiencies of the motives per se [13] or by variations of the sequence context (see the first polyadenylation motif in Fig. 1). The isolation of multiple cDNA clones indicates that in *F. trinervia* the C_4 isoform of PEPCase is encoded by a separate subfamily. This is further supported by Southern analysis and the isolation of multiple genomic clones for this PEPCase isoform [7, Hermans and Westhoff, in preparation]. The finding with *F. trinervia* is in contrast to maize and *Sorghum* where the C_4 PEPCase is encoded by a single-copy gene [6,7].

3.2. *Expression analysis*

The expression characteristics of the encoded genes were investigated by Northern analysis using the largest cDNA clone, pcFtppc1-1, as a representative probe. The hybridization was performed under stringent conditions in order to distinguish between the various members of the PEPCase gene family of *F. trinervia* [7]. Fig. 2 shows that this cDNA clone detects a single RNA species, 3.3 kb in size. The transcript is abundant in the mesophyll cells, but cannot be detected in bundle-sheath tissue or in roots, stems and flowers. The mesophyll-specific accumulation of the transcript confirms that pcFtppc1-1 encodes the C_4 PEPCase isoform.

3.3. *Sequence analysis of pcFtppc1-1 and the predicted protein*

The entire nucleotide sequence of pcFtppc1-1 was determined. The nucleotide sequence data reported in this paper will appear in the EMBL Nucleotide Sequence Databases under the accession number X61304. The sequence contains a long open reading frame of 2898 bp, 50 nucleotides of 5′ non-translated sequence, a 222 bp 3′ non-coding region and a 9 nucleotide poly(A) tail (Fig. 3). The first ATG codon of the open reading frame is located in a sequence context which meets the requirement of a eukaryotic translational initiation site of plants [14,15].

The open reading frame can be translated into 966 amino acid residues resulting in a protein of about 110 kDa in size. Sequence comparisons with the C_4 PEPCases of the monocotyledonous C_4 plants *Zea mays* and *S. bicolor* and the CAM and C_3 isoforms of the dicotyledonous, facultative CAM plant *Mesembryanthemum crystallinum* reveal 60% identical residues between the five proteins (Fig. 3). The variable positions are nearly equally distributed.

A closer inspection of the protein alignment identifies six conserved cysteine residues which may be involved in subunit/subunit interactions [2]. Furthermore, a spe-

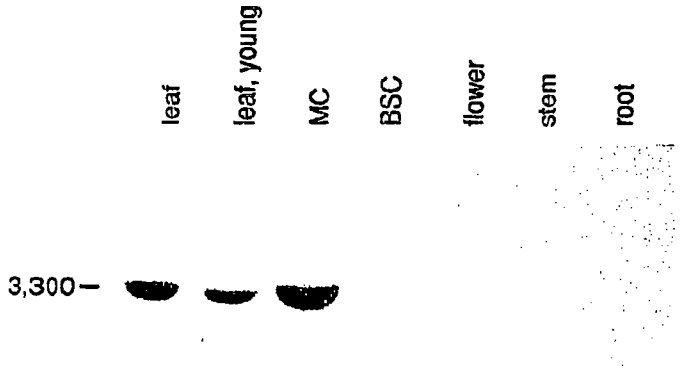

Fig. 2. Transcript analysis with pcFtppc1-1. Total RNA (10 μg per lane) isolated from young and mature leaves, separated mesophyll (MC) and bundle-sheath cells (BSC), flowers, stems and roots was analyzed by Northern hybridization as described [7]. The size of the transcript is given in nucleotides. fluorographs were exposed for 12 h.

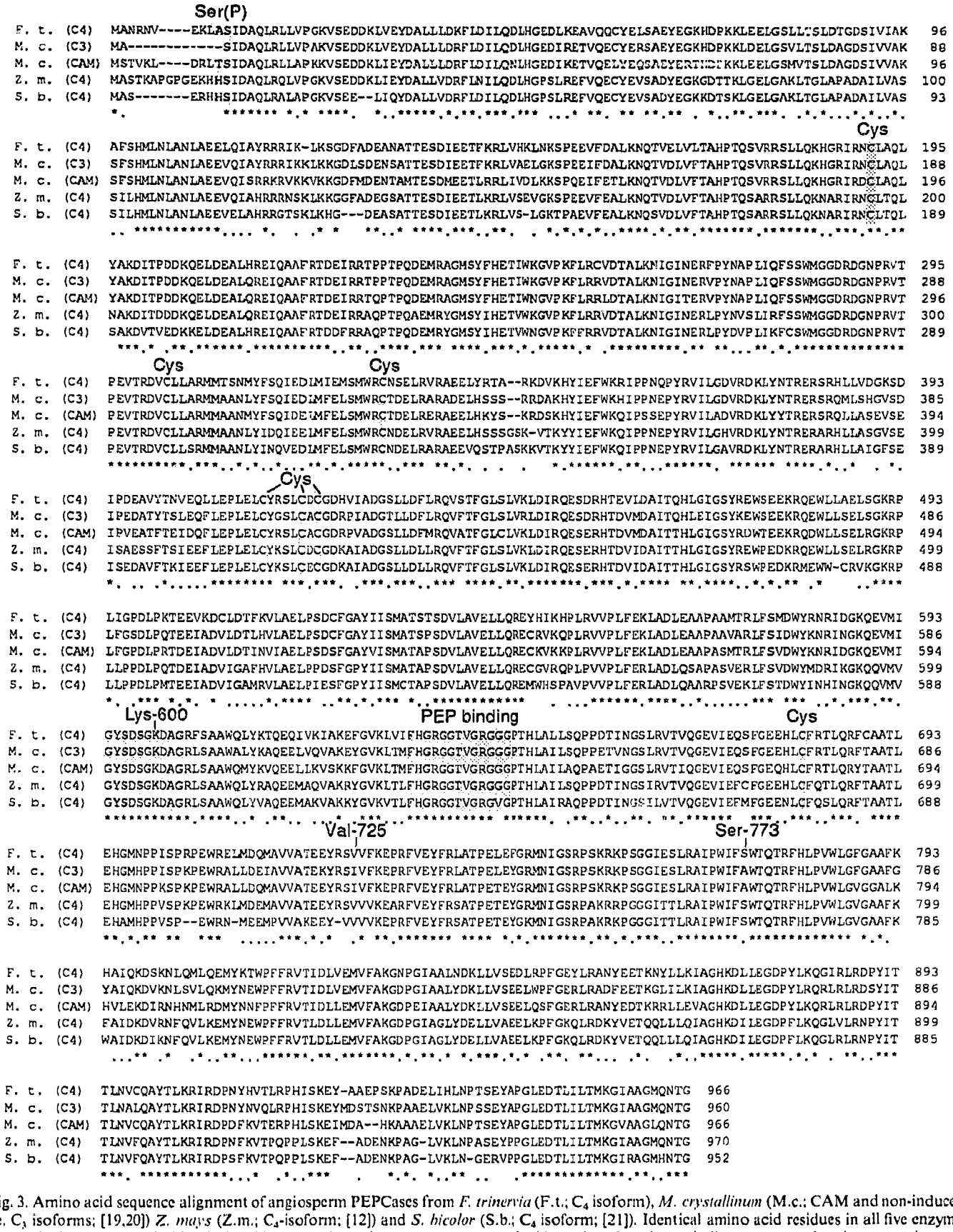

Fig. 3. Amino acid sequence alignment of angiosperm PEPCases from *F. trinervia* (F.t.; C_4 isoform), *M. crystallinum* (M.c.; CAM and non-induced, i.e. C_3 isoforms; [19,20]) *Z. mays* (Z.m.; C_4-isoform; [12]) and *S. bicolor* (S.b.; C_4 isoform; [21]). Identical amino acid residues in all five enzymes are marked by an asterisk, identical residues in at least three enzymes are marked by a dot. Grey boxes indicate conserved regions or amino acid residues whose putative functions are discussed in the text.

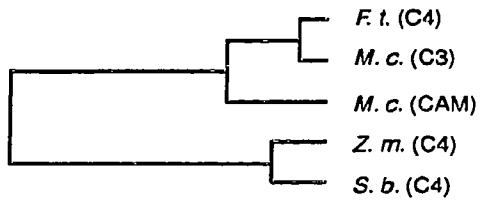

Fig. 4. The putative phylogenetic relationship of the *F. trinervia* (F.t.), *M. crystallinum* (M.c.), *Z. mays* (Z.m.) and *S. bicolor* (S.b.) PEPCases. The tree was constructed by using the program CLUSTAL of the PC gene software package (vers. 5.16, IntelliGenetics, Geel, Belgium).

cies-invariant lysine residue (lysine-600 of the *F. trinervia* protein) can be detected. The regions surrounding this position (residues 594–603 and 632–645 of the *F. trinervia* protein) are highly conserved among all eukaryotic and prokaryotic PEPCases analyzed so far (Fig. 3). There is evidence that this part of the enzyme is involved in phosphoenolpyruvate binding and catalytic activity [2,17]. The serine-15 residue of the maize PEPCase undergoes reversible phosphorylation during light/dark transitions which is correlated with changes in the activation state of the enzyme [3]. The structural motif Lys/Arg-X-X-Ser of this phosphorylation site is also present at the corresponding position of the *F. trinervia* enzyme. Thus, this element is shared by the C_4 and CAM isoforms but is lacking in the C_3-isoenzyme [3,18].

Surprisingly, only two amino acid residues, valine-725 and serine-773, are specific for both the mono- and dicotyledonous C_4 PEPCases. These residues are missing in the CAM and the C_3 enzyme (Fig. 3). The significance of this finding is unknown at present.

To quantitate the relationship between the various enzymes a phylogenetic tree was constructed. Fig. 6 shows that the C_4 PEPCase of *F. trinervia* is more closely related to the C_3 and CAM isoforms of *M. crystallinum* than to the monocotyledonous C_4 PEPCases. This reinforces the view that the mono- and dicotyledonous C4 plants evolved independently from each other after the monocot–dicot divergence [16].

Acknowledgements: We thank Dr J. Vidal (Université Paris-Sud, France) for kindly providing antisera to *Sorghum* PEPCase and M. Höfer for screening the cDNA library. This work was supported by grants from the Bundesminister für Forschung und Technologie, the Deutsche Forschungsgemeinschaft (SFB 189) and the Fonds der Chemischen Industrie.

REFERENCES

[1] O'Leary, M.H. (1982) Annu. Rev. Plant Physiol. 33, 297–315.
[2] Andreo, C.S., Gonzales, D.H. and Iglesias, A.A. (1987) FEBS Lett. 213, 1–8.
[3] Jiao, J. and Chollet, R. (1991) Plant Physiol. 95, 981–985.
[4] Ting, I.P. and Osmond, C.B. (1973) Plant Physiol. 51, 448–453.
[5] Latzko, E. and Kelly, J. (1983) Physiol. Veg. 21, 805–815.
[6] Harpster, M.H. and Taylor, W.C. (1986) J. Biol. Chem. 261, 6132–6136.
[7] Hermans, J. and Westhoff, P. (1990) Mol. Gen. Genet. 224, 459–468.
[8] Crétin, C., Santi, S., Keryer, E., Lepiniec, L., Tagu, D., Vidal, J. and Gadal, P. (1991) Gene 99, 87–94.
[9] Börsch, D. and Westhoff, P. (1990) FEBS Lett. 273, 111–115.
[10] Höfer, M., Santore, U. and Westhoff, P. (1991) Planta (in press).
[11] Yanagisawa, S., Izui, K., Yamaguchi, Y., Shigesada, K. and Katsuki, H. (1988) FEBS Lett. 229, 107–110.
[12] Thomas, M., Crétin, C., Vidal, J., Keryer, E., Gadal, P. and Moesinger, E. (1990) Plant Sci. 69, 65–78.
[13] Ingelbrecht, I.L.W., Herman, L.M.F., Dekeyser, R.A., Van Montagu, M.C. and Depicker, A.G. (1989) Plant Cell 1, 671–680.
[14] Lütcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F. and Scheele, G.A. (1987) EMBO J. 6, 43–48.
[15] Joshi, C.P. (1987) Nucleic Acids Res. 15, 6643–6653.
[16] Monson, R.K. and Moore, B.D. (1989) Plant Cell Environ. 12, 689–699.
[17] Jiao, J.-A., Podestá, F.E., Chollet, R., O'Leary, M.H. and Andreo, C.S. (1990) Biochim. Biophys. Acta 1041, 291–295.
[18] Carter, P.J., Nimmo, H.G., Fewson, C.A. and Wilkins, M.B. (1991) EMBO J. 10, 2063–2068.
[19] Cushman, J.C. and Bohnert, H.J. (1989) Nucleic Acids Res. 17, 6743–6744.
[20] Cushman, J.C. and Bohnert, H.J. (1989) Nucleic Acids Res. 17, 6745–6746.
[21] Crétin, C., Keryer, E., Tagu, D., Lepiniec, L., Vidal, J. and Gadal, P. (1990) Nucleic Acids Res. 18, 658.