FEBS Letters 350 (1994) 216–218

FEBS 14380

The nucleotide sequence of a complementary DNA encoding *Flaveria bidentis* carbonic anhydrase

Antonino Cavallaro, Martha Ludwig, James Burnell*

Centre for Molecular Biotechnology, Queensland University of Technology, PO Box 2434, Brisbane Q 4001, Australia

Received 5 July 1994

Abstract

We have isolated and characterised a cDNA clone encoding the cytosolic form of carbonic anhydrase in the leaves of *Flaveria bidentis*, a C_4 dicotyledonous plant. The deduced amino acid sequence is similar to the carbonic anhydrase found in the chloroplasts of C_3 dicotyledonous plants. Western blot analysis of crude leaf extracts of *F. bidentis* indicates that the leader sequence (equivalent to the transit peptide of the chloroplastic form of CA found in C_3 plants) is not removed following translation of mRNA.

Key words: Flaveria bidentis; Carbonic anhydrase; Leader sequence; C_4 plant; Transit peptide

1. Introduction

Carbonic anhydrase (carbonate dehydratase, CA; EC 4.2.1.1) catalyses the reversible hydration of CO_2 according to the reaction $CO_2 + H_2O = HCO_3^- + H^+$. The enzyme has been found in all plants so far examined and represents between 1% and 2% of the total leaf protein [1]. In C_4 plants CA is located in the cytosol of mesophyll cells [2–4] and catalyses the first reaction of the C_4 acid cycle ensuring the supply of bicarbonate for PEP carboxylase, the primary CO_2 fixing enzyme in C_4 plants [5]. In C_3 plants CA is located in the chloroplasts and is thought to facilitate the diffusion of CO_2 from the cytosol into the chloroplasts [6]. Therefore, during the evolution of C_4 plants from their C_3 ancestors both the function and location of carbonic anhydrase has changed.

We are seeking to determine the molecular changes underlying the evolution of C_4 plants from their C_3 ancestors and we are studying the control of expression of carbonic anhydrase in plants within the genus *Flaveria* (Asteracea). This genus contains C_3 and C_4 plants and a number of C_3-C_4 intermediate species (see [7]).

In this paper we report the primary structure of carbonic anhydrase of the C_4 dicotyledonous plant *F. bidentis*. This paper represents the first report of a carbonic anhydrase from a C_4 plant and we present evidence that the protein is synthesised with a leader sequence (analogous to the transit peptide in C_3 plant CA) which is retained.

2. Methods and materials

2.1. Plant material

F. bidentis was grown in soil in a naturally illuminated glasshouse. Silver beet was purchased from the local market.

2.2. Construction and screening of cDNA library

Total RNA was isolated from leaves of *F. bidentis* [8] and poly(A)$^+$ RNA was purified using an oligo-dT cellulose column (Pharmacia, Australia). cDNA was synthesised using a cDNA synthesis kit (Amersham, Australia) and a cDNA library was constructed in λgt11 as described by the supplier (Amersham, Australia). Phage were plated on Y1088 *E. coli* cells resulting in about 5×10^5 independent clones. The cDNA library (3×10^5 of the amplified library) was screened using a 220 bp *Ava*II fragment of *F. brownii* CA cDNA. Positive clones were identified and λgt11 DNA isolated and the insert cDNA was subcloned into pTZ18R (Pharmacia, Australia). The nucleotide sequence of both strands of the isolated cDNA clone was determined by the dideoxy chain termination method modified for double stranded plasmid DNA [9].

2.3. Western blot analysis

Crude extracts were made by grinding 1 g of leaf material in 3 ml of buffer containing 50 mM HEPES-KOH, 10 mM $MgCl_2$, and 1 mM EDTA, pH 7.5, at 0°C. The extracts were filtered through two layers of Miracloth (Calbiochem) and centrifuged at 13K at 4°C for 10 min. SDS-PAGE was conducted using 7.5% polyacrylamide gels according to Laemmli [10]. Proteins were transferred to Immobilin-PQ (Millipore) membranes and Western blotting using anti-spinach CA antiserum was conducted as described previously [11].

3. Results and discussion

The *F. bidentis* cDNA library was screened with a fragment of cDNA encoding CA in *F. brownii* (unpublished data). Two positive clones were isolated and restriction enzyme and Southern analysis indicated that they contained inserts about 1.25 kbp long. The insert from one of these clones was subcloned into pTZ18R and the entire nucleotide sequence of both strands determined. The cDNA insert consisted of an open reading frame of 993 base pairs, 56-nucleotides 5′-untranslated

*Corresponding author. Fax: (61) (7) 864-1534.

SSDI 0014-5793(94)00767-5

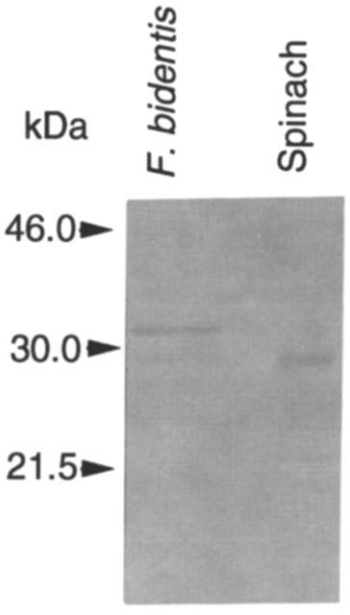

Fig. 1. Western blot analysis of crude leaf extracts from *F. bidentis* and silver beet. Crude leaf extracts were prepared from leaves of *F. bidentis* and silver beet and 1.15 μg of protein was subjected to electrophoresis, transferred to nitrocellulose membrane and treated with anti-spinach CA antiserum. CA bands are indicated with arrows. Molecular mass is indicated on the right hand side.

and 122 nucleotides 3′-untranslated which contains a poly(A) tail (GenBank Accession no. U08398). The entire open reading frame can be translated into 331 amino acid residues resulting in a polypeptide of 35.97 kDa.

It has been recognised that poly(A)$^+$ mRNAs possess a conserved poly(A) addition signal, AATAAA, usually found 10–30 nucleotides upstream from the poly(A) tail [12]. This signal is present in the CA mRNA but is located about 130 bp upstream of the poly(A) tail. A comparison of the amino acid sequence of *F. bidentis* CA with published CA sequences of C_3 plants [13–18] reveals about 70% homology with C_3 dicot CA and about 60% homology with the C_3 monocot CA. CA in C_3 plants is nuclear-encoded and synthesised in the cytosol as an immature protein containing a transit peptide which is removed during transport into the chloroplast. In contrast, CA in C_4 plants is located in the cytosol of mesophyll cells. As mentioned above *F. bidentis* codes for a protein about 36 kDa in size. To determine whether *F. bidentis* CA is subject to post-translational processing, Western blot analysis of a *F. bidentis* crude leaf extract was conducted. Using spinach CA antiserum an immunoreactive band equivalent to a protein with a molecular mass of about 36 kDa was observed (Fig. 1) and this band was significantly larger than the immunoreactive band detected in crude leaf extracts of silver beet (a C_3 dicot). These results suggest that the leader sequence, homologous with the transit peptide of C_3 CA, may not be removed following translation of the protein. However, these results could also be explained by the unlikely possibility that the transit peptide is removed and the mature protein is subject to post-translation modification (e.g. by glycosylation).

Alignment of the *F. bidentis* CA amino acid sequence (deduced from the cDNA nucleotide sequence) with the published sequences of C_3 dicot CAs (see [13–18]) shows three major differences (Fig. 2). The N-terminal amino acid sequence of the immature protein from all known C_3 CAs begins with the amino acid sequence MST. In contrast the N-terminal amino acid sequence of *F. bidentis* (C_4 plant) begins MSA. In addition a leucine residue is located at amino acid residue 59 in *F. bidentis* in contrast to a proline residue which is conserved in all C_3 CAs. The transit peptide cleavage sites have been determined empirically only for spinach [13] and pea [18] and, while the amino acid sequence on the N-side of the cleav-

```
            1                                                  50
F.bidCA     MSAASAFAMNAPSFVNA.SSLKKAST.SARSGVLSARFTCNSSSSSSSSS
SpinCA      MST...ING.CLTSISPSRTQL-NTSTLRPTFIA........N-RVNP--
TobCA       MSTASINS.CLT.ISP-QA----PT..RPVAFA.........RLSN----
ArabCA      MSTAPLSGFFLT-LSPSQ---Q-L-LRTSSTVACLPPASSS---------
PeaCA       MSTSSINGFSLS-LSP-KT-T-RTTL.RPFVFASL......NT-------

           51                                                 100
F.bidCA     ATPPSLIRNELVFAAPAPIITPNWTED.GNESYEEAIDALKKTLIEKGEL
SpinCA      SV-------QP-----------TLK--MA...------A----L-S-----
TobCA       TSV-------P-----T---N-ILR-EMAK----Q--A--E-L-S-----
ArabCA      RSV-T-----P---------A-Y-S-EM-T-A-D---E----L----E--
PeaCA       S-F----QDKP---SSS-----VLR-EM-K.G-D---EE-Q-L-R--T--

          101           *                                    150
F.bidCA     EPVAATRIDQITAQ...AAAPDTKAPFDPVERIKSGFVKFKTEKFVTNPA
SpinCA      -NE--SKVA---SELADGGT-..S-.SY--Q---E--I---K--YEK---
TobCA       G-I--A-V-----ELQSSDGS..-P.-----HM-A--IH-----YEK---
ArabCA      KT---AKVE----ALQTGTSS-K--.-----T--Q--I---K--YE----
PeaCA       KAT--EKVE-----LGTTSSS-GIPKSEAS----T--LH--K--YDK---

          151                                                  200
F.bidCA     LYDELAKGQSPKFMVFACSDSRVCPSHVLDFQPGEAFVVRNVANMVPPFD
SpinCA      --G--S---A--------------------------M---I-----V--
TobCA       --G--S-----------------------N-----------I-----AY-
ArabCA      --G---------Y---------------------D------I--------
PeaCA       --G--------P-------------------------------L---Y-

          201                                                  250
F.bidCA     KTKYSGVGAAVEYAVLHLKVQEIFVIGHSRCGGIKGLMTFPDEGPHSTDF
SpinCA      -D--A-----I---------EN-V-----A--------S---A--TT---
TobCA       --R-------I---------EN-V-----A--------SL-AD-SE--A-
ArabCA      -V--G-----I---------EN-V-----A--------L--LD-NN----
PeaCA       QA--A-T---I---------SN-V-----A-------LS--FD-TY----

          251                                                  300
F.bidCA     IEDWVKVCLPAKSKVVAEHNGTHLDDQCVLCEKEAVNVSLGNLLTYPFVR
SpinCA      ------I-----H--L---GNATFAE--TH--------------------
TobCA       ------IG----A--QG--VDKCFA---TA--------------------
ArabCA      ------I--------IS-LGDSAFE---GR--R-------A---------
PeaCA       --E---IG----A--K-Q-GDAPFAEL-TH-------A------------

          301
F.bidCA     DGLRNKTLALKGGHYDFVNGTFELWALDFGLSSPTSV
SpinCA      ---VK-----Q--Y------S----G-E----PSQ--
TobCA       E--VK---------------G----G-E----PSL--
ArabCA      E--VKG-------Y----K-A----G-E----ETS--
PeaCA       E--V---------Y----K-S----G-E-----TF--
```

Fig. 2. Alignment of the deduced amino acid sequences of *F. bidentis* CA with the deduced amimo acid sequences of C_3 plant CAs. Sequences were aligned using the GCG Pileup multiple alignment program. Dots are included to maximise alignment and dashes indicate homology with the *F. bidentis* amino acid sequence. The N-terminal amino acid sequence of purified spinach CA as determined by amino acid analysis is underlined and the asterisk denotes the presumed processing site of C_3 CAs. F.bidCA = *F. bidentis*; SpinCA = Spinach; TobCA = tobacco; ArabCA = *Arabidopsis*; PeaCA = pea.

age site is highly conserved amongst the published C_3 CAs, the amino acid sequence on the C-terminal side of the cleavage site of *F. bidentis* CA differs from that found in C_3 species. In addition *F. bidentis* CA has a three amino acid deletion one amino acid C-terminal of the processing point of spinach and pea CA. Any or all of these amino acid sequence differences may be responsible for the non-processing of the immature CA in *F. bidentis* mesophyll cells and may explain why CA in the mesophyll cytosol is not processed by the chloroplast and, therefore, why CA is located in the cytosol rather than in the chloroplast.

Only one immunoreactive band was detected during Western blot analysis of crude extracts made in the presence or absence of Triton X-100 indicating that a membrane-bound form of CA is not present in *F. bidentis*, that the anti-spinach CA antiserum does not bind to a membrane-bound form of CA or, that if there is, it must be the same size as the cytosolic form (see [19]).

With respect to the overall amino acid homology between *F. bidentis* and C_3 CAs, less similarity (about 42% between *F. bidentis* and spinach) is found in the transit peptide (leader sequence). This difference is mainly located at the N-terminal end since the level of homology is high in the region closer to the processing point of spinach and other CAs.

References

[1] Okabe, K., Yang, S., Tsuzuki, M. and Miyachi, S. (1984) Plant Sci. Lett. 33, 145–153.
[2] Burnell, J.N. and Hatch, M.D. (1988) Plant Physiol. 86, 1252–1256.
[3] Gutierrez, M., Gracen, V.E. and Edwards, G.E. (1974) Planta 119, 279–300.
[4] Ku, M.S.B. and Edwards, G.E. (1975) Z. Pflanzenphysiol. 77, 16–32.
[5] Hatch, M.D. and Burnell, J.N. (1990) Plant Physiol. 93, 825–828.
[6] Edwards, G.E. and Walker, D.A. (1983) C_3 C_4: Mechanisms, and Cellular and Environmental Regulation of Photosynthesis, Blackwell, Oxford, 542 pp.
[7] Edwards, G.E. and Ku, M.S.B. (1987) in: The Biochemistry of Plants, vol. 10 (Hatch, M.D. and Boardman, N.K. eds.) pp. 275–325, Academic Press, New York.
[8] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) in: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[9] Tabor, S. and Richardson, C.C. (1989) Proc. Natl. Acad. Sci. USA 86, 4076–4080.
[10] Laemmli, U.K. (1970) Nature 227, 680–685.
[11] Burnell, J.N. (1990) Plant Cell Physiol. 31, 423–427.
[12] Nevins, J.R. (1983) Annu. Rev. Biochem. 52, 441–466.
[13] Burnell, J.N., Gibbs, M.J. and Mason, J.G. (1990) Plant Physiol. 92, 37–40
[14] Fawcett, T.W., Browse, J.A., Volokita, M. and Bartlett, S.G. (1990) J. Biol. Chem. 265, 5414–5427.
[15] Majeau, N. and Coleman, J.R. (1991) Plant Physiol. 95, 264–268.
[16] Majeau, N. and Coleman, J.R. (1992) Plant Physiol. 100, 1077–1078.
[17] Raines, C.A., Horsnell, P.R., Holder, C. and Lloyd, J.C. (1992) Plant Molec. Biol. 20, 1143–1148
[18] Roeske, C.A. and Ogren, W.L. (1990) Nucleic Acids Res. 18, 3413.
[19] Utsunomiya, E. and Muto, S. (1993) Physiol Plant 88, 179–190.