Isolation and Characterization of a cDNA Coding for Pea Chloroplastic Carbonic Anhydrase

Nathalie Majeau and John R. Coleman*

Centre for Plant Biotechnology, Department of Botany, University of Toronto, Toronto, Ontario, Canada M5S 3B2

ABSTRACT

Using a polyclonal antibody generated against the purified pea (Pisum sativum) carbonic anhydrase (CA) monomeric species, we have isolated and characterized a cDNA coding for this enzyme. Protein sequence analysis was used to confirm the identity of the clone. The presence of a large transit peptide suggests that CA is transported into the chloroplast and then processed to the mature size of approximately 26 kilodaltons. Northern hybridization, using the CA cDNA as a probe of total leaf RNA, revealed a single transcript of 1.45 kilobase pairs. This transcript was not detected in RNA extracted from root or etiolated leaf tissue. Comparison of the deduced amino acid sequence with that of spinach CA showed approximately 68% identity over the length of the nascent protein but with greater similarity observed within the mature protein sequences. In addition, regions of the pea and spinach CA proteins were found to be significantly similar to the Escherichia coli cyanate permease.

One of the more abundant soluble proteins in the leaves of higher plants is the enzyme CA1 (carbonate dehydratase 4.2.1.1), which catalyzes the reversible hydration of CO2 to HCO3-. Although it may represent as much as 1 to 2% of the soluble leaf protein, there are many aspects of higher plant CA structure, localization, regulation, and role in metabolism that are not well understood (16, 18). For example, significant differences appear to exist between monocot- and dicotyle- donous plants with respect to molecular mass of the holoenzyme. Molecular sizes can range from 42 to 45 kD in monocot species to 140 to 250 kD when isolated from a number of dicot species (1, 2, 18). The relatively small range in the observed subunit size (26–34 kD) for all higher plant CA proteins suggests that some variation in subunit number of the native protein occurs. Most localization studies indicate that CA is found in the chloroplasts of C3 plants and within the cytosol of C4 species; however, other reports indicate that additional isozymes and tissue locations are possible (1, 13, 18, 24). Within the C3 chloroplast, CA is thought to speed the dehydration of HCO3- and thus maintain the supply of CO2 for Rubisco (18). It has also been postulated, however, that CA activity could directly facilitate diffusion of CO2 across the chloroplast envelope by maintenance of the equilibrium between the inorganic carbon species (18, 24). In addition, CA may play a role in the buffering capacity of the chloroplast stroma by enhancing the rates of the dehydration/hydration reactions.

As an initial step in the study of CA activity and regulation of expression in higher plants, we have isolated and characterized a cDNA clone coding for pea CA. We have also determined CA transcript abundance and compared the deduced amino acid sequence with another higher plant CA and with an Escherichia coli polypeptide which exhibits some sequence similarity to the pea CA.

MATERIALS AND METHODS

Plant Material

Peas (Pisum sativum var Little Marvel) and corn (Zea mays var Early King) were grown in growth rooms for either 14 or 21 d in vermiculite, with a day length of 16 h, illumination of 700 μE·m−2·s−1 and a day/night temperature regime of 20°C/16°C. Plants were watered every 3 d and provided with a nutrient solution weekly. Etiolated tissue was produced by germination and growth in total darkness using a similar temperature, watering, and nutrient regime.

Isolation of Pea Leaf Carbonic Anhydrase

Leaves and stems (250 g) of 21-d-old, light-grown peas were homogenized in ice-cold extraction buffer (0.3 M Tris-SO4 [pH 8.3] containing 5 mM DTT, 1 mM PMSF, 1 mM benzamidine) using a Waring blender for 1 min and filtered through Miracloth, and the soluble fraction was clarified by centrifugation at 4°C (40 min, 30,000g). The supernatant was collected and subjected to (NH4)2SO4 fractionation. The 30 to 60% saturation precipitate was resuspended, dialyzed against 5 mM Tris-SO4 (pH 9.5) buffer, and then charged onto an affinity column (p-methylaminobenzene sulfonamide substituted Sepharose 4B) synthesized as previously described (26). After extensive washing with 25 mM Tris-SO4 (pH 8.3), the CA containing fraction was eluted with 25 mM Tris-SO4 (pH 8.3) containing 25 mM NaClO4, and then was dialyzed extensively against 5 mM Tris-SO4 (pH 8.3). Fractions containing CA activity, as determined by an electrophoretic assay (25), were analyzed by SDS-PAGE as described previously (6). In situ CA activity, after separation of total soluble pea protein on nondenaturing polyacrylamide gels, was detected by ultraviolet fluorescence of bromocresol purple after exposure to CO2 at low temperature (1).
Production of Antibodies Against Carbonic Anhydrase

Affinity column-purified CA was subjected to SDS-PAGE (7–15% gradient gel), and the CA monomeric species (primarily a 25.5 kD polypeptide and a minor 27.5 kD band) were electrophoresed from the gel slices. After collection of preimmune serum, the isolated CA was combined with Freund's adjuvant and was injected into rabbits using a previously described immunization schedule (5). Western blotting techniques were used to test the specificity of the rabbit polyclonal antisera and to identify the CA monomer in total soluble protein extracts (3).

Isolation and Characterization of cDNA Clones

A λ gt11 cDNA library, synthesized from light-grown, pea leaf mRNA (23), was screened using the antcarbonic anhydrase serum as previously described (3). Inserts from the immunopositive recombinant phage were generated by EcoRI digestion and then subcloned directly into the plasmid vector PBS (Stratagene Inc.). Double-stranded DNA sequencing of the inserts was performed using synthesized oligonucleotide primers and the dideoxy method of Sanger et al. (19). DNA sequence analysis was performed by using the Pustell Sequence Analysis Software (International Biotech. Inc.).

RNA Isolation and Northern Hybridization

Total RNA was isolated from leaves, roots and etiolated leaves of 14-d-old plants that were first frozen and powdered in liquid nitrogen, prior to extraction using the method described in (6). Equal amounts of RNA from each tissue were denatured and electrophoresed on a 1.5% agarose gel containing 0.66 M formaldehyde. Transfer to nitrocellulose, prehybridization and hybridization conditions were as previously described (17). The cloned 0.95 kb EcoRI cDNA insert used as a probe was labeled with [α-32P]dCTP using a random primer procedure (9). The hybridization pattern was determined by autoradiography and the sizes of the stained ribosomal bands and transcripts estimated by comparison with known standards.

Protein Sequence Determination

An internal peptide fragment, generated by cyanogen bromide digestion of affinity column and SDS-PAGE purified CA was sequenced by Edman degradation using an Applied Biosystems Sequenater, model 477A.

RESULTS

Carbonic anhydrase, isolated from the green tissue of Pisum sativum, has been purified by a combination of affinity chromatography and gel electrophoresis. After separation on an SDS-polyacrylamide gel, the protein fraction eluted from the sulfonamide affinity column contained an abundant 25.5 kD polypeptide and a minor band of 27.5 kD (Fig. 1, lane 3). This same protein fraction exhibited maximal levels of CA activity prior to SDS-PAGE. CA activity was also detected after non-denaturing gel electrophoresis of soluble pea polypeptides and staining of the gel with bromocresol purple. The fluorescent band which appears after exposure of the gel to CO2 was excised, electroeluted, and the extracted proteins subjected to SDS-PAGE. The same abundant 25.5 kD polypeptide and minor 27.5 kD protein were observed (data not shown).

The 25.5/27.5 kD doublet was excised from SDS-polyacrylamide gels and used to generate a polyclonal antibody in rabbits. A Western blot of total soluble protein isolated from Pisum sativum and Zea mays is shown in Figure 1. The antibody cross-reacted strongly with a 25.5 kD polypeptide in the total soluble protein isolated from pea (Fig. 1, lane 3). The antibody failed to cross-react with any of the soluble proteins from corn.
Figure 2. Nucleotide and deduced amino acid sequence of the pea leaf carbonic anhydrase cDNA. The amino acids identified by protein sequencing of an internal peptide fragment are underlined. The underlined region at the 3’ end of the cDNA represents a putative polyadenylation signal sequence.

proteins isolated from corn (Fig. 1, lane 4) and from etiolated pea leaf or root tissue (data not shown).

Approximately 200,000 plaques obtained from the pea leaf cDNA library were screened with the pea CA antibody. A total of 10 individual plaques were found to be true immunopositive clones after rescreening, and these clones were also found to exhibit strong sequence homology as determined by Southern hybridization techniques (data not shown). After digestion by EcoRI, the two largest inserts, 0.95 and 1.3 kb, were subcloned into the vector pBS. Partial nucleotide sequence analysis revealed that these two clones were identical. The 1.3 kb fragment was sequenced completely in both directions and is presented along with the deduced amino acid sequence in Figure 2. The sequence encodes an open reading frame for a polypeptide of 35.7 kD and a segment of the deduced amino acid sequence was found to match exactly the amino acid sequence of CNBr-generated peptide fragment isolated from the purified protein (Fig. 2). N-Terminal protein sequencing of the purified polypeptide did not provide conclusive data.

The 1.3 kb cDNA insert was hybridized to total pea RNA isolated from light grown leaf and stem tissue, root tissue, and etiolated leaf tissue (Fig. 3). A single 1.45 kb transcript was identified in the RNA isolated from light grown leaf and stem tissue (Fig. 3, lane 4) that was not present in RNA isolated from etiolated leaf or root tissue (Fig. 3, lanes 5 and 6).

The deduced pea CA amino acid sequence is compared with spinach CA (4, 8) (Fig. 4) and with the deduced amino acid sequence of the cyanate permease gene isolated from Escherichia coli (20) (Fig. 5). The permease sequence was identified as having similar homology with carbonic anhydrase using Genbank and the Pestell sequence analysis programmes. An examination of a number of animal CA amino acid sequences (7, 10, 22) using the same software did not reveal any significant regions of similarity with the pea sequence.

Figure 3. Determination of pea CA transcript abundance. Total RNA isolated from light-grown leaves (lane 1), dark-grown (etiolated) leaves (lane 2), and root tissue (lane 3) was separated by electrophoresis on a 1.5% agarose gel containing 0.66 M formaldehyde, transferred to nitrocellulose, and probed with the 32P-labeled CA cDNA. The ethidium-stained gel (lanes 1–3) and the resulting hybridization profiles (lanes 4–6) are shown. The size of the transcript was determined by its position relative to the mobility of denatured λ-DNA fragments after digestion with HindIII and EcoRI, as well as Escherichia coli ribosomal RNA.
**Growth**

that
tissue
in
Bpi
Bpi KTLALQGGYYDFVNGSFELWGLEYGLSPSQSV

Pea E
SPIITPVLREEMGKGYDEAIEELQKLLREKTELKATAAEKVEQITAQLGTTSSSDGIPKSEASERIKTGFLHFKKEKYDKNPALYGELAKGQSPPFMVFACSDSRVCPSHVLDFQPGKA

Pea
Pea
Pea
Pea MSTSSINGFSLSSLSPAKTSTKRT

FMVRNIANNVPVFDKDKYAGVGAAIEYAVLHLKVENIVVIGHSACGGIKGLMSFPDAGPT

CA
pea
(4, 8)
additional
ition is that the
cytosolic
higher plant carbonic anhydrases (16, 18).

reported
ted during
monomeric
of the
cDNA
by protein sequencing (4),
determined
at the amino acid level
new bands were observed after the protein was electrophoresed under denaturing conditions (2, 15). It is possible that the formation of the smaller, more abundant polypeptide in the doublet is the result of proteolysis despite the presence of protease inhibitors during the isolation procedure. An alternative explanation is that the 27.5 kD polypeptide represents a less abundant, additional CA isoform (possibly cytosolic) with limited antigenicity to the polyclonal antibody. The inability of the pea CA antibody to cross-react with a corn polypeptide supports the view that significant taxonomic diversity exists among higher plant carbonic anhydrases (16, 18). The identity of the two largest cDNA clones detected by cross-reactivity with the pea CA polyclonal antibody was verified after nucleotide sequence analysis and comparison with the amino acid sequence of an internal peptide fragment generated by CNBr digestion of the purified enzyme. In addition, the recently published sequence of a cDNA clone coding for spinach CA (4, 8) was found to be similar to the pea sequence (approximately 76.5% identity at the amino acid level over the length of the mature spinach protein).

Northern analysis of total mRNA indicates that the appearance of the pea CA transcript is light-regulated and leaf-tissue specific. Western blot analysis of green and etiolated leaf tissue using the pea CA antibody supports these observations. Previous studies have reported the presence of CA activity in etiolated or root tissue (11, 13, 18). It is possible that species-specific variation or low light contamination of growth conditions may account for these differences. It is also possible that our cDNA probe or polyclonal antibody are unable to identify additional (nonchloroplastic) pea CA isozymes.

Several regions of the two higher plant CA proteins contain stretches of amino acids which are identical and these may contain functional domains such as zinc, inorganic carbon, or anion binding sites. The plant carbonic anhydrase amino acid sequences were not found to exhibit significant homology to a number of animal CA sequences. In particular, the active site arrays of histidine residues, common to animal CA proteins, were not present (22). It is possible that until the active site configuration of the plant protein is determined, similarities between plant and animal carbonic anhydrases will remain elusive. Although we were unable to determine with certainty the N-terminal sequence of the purified pea CA, the size of the mature pea CA protein as well as N-terminal sequence analysis of the spinach CA indicate that both polypeptides have very long putative transit peptides, in excess of 100 amino acids (12, 14). In addition to the length, an array of eight consecutive Ser residues flanked by Thr residues, is another unusual feature of the pea pre-sequence. It has been suggested that the large number of acidic residues between amino acids 60 and 100 of the spinach polypeptide are unlikely components of a transit peptide and that this portion of the pre-sequence is removed after uptake by the chloroplast and cleavage of the transit peptide (8). Similarly, the pea sequence contains a large number of charged residues within this region (AA 60–104) and the typical Arg residue enrichment of the putative transit peptide C-terminus (immediately upstream of AA 104, Fig. 4) is also not apparent (12). These data support the notion that the site of transit cleavage may reside within the first 60 amino acids of the nascent polypeptide (8).

The observation of similarities between the plant CA and the *Escherichia coli* cyanase permease sequences was very surprising and is not yet fully understood. The permease is responsible for transport of the cyanate ion (OCN−) across the bacterial membrane, after which the cytosolic enzyme cyanate catalyzes its decomposition to bicarbonate and ammonia (20, 21). The structure of the cyanate ion may be sufficiently similar to one of the inorganic carbon species or anions capable of binding to carbonic anhydrase. Cyanate is

**Figure 5.** Comparison of the deduced amino acid sequences for pea CA (CA) and the *E. coli* cyanase permease gene (Per). Numbering for the pea CA mature protein is as described in Figure 3. Numbering starts at the beginning of the open reading frame for the *E. coli* permease gene product. Identity at the amino acid level is indicated by the symbol *"*.
known to be a potent inhibitor of mammalian CA (7 and references within) and does inhibit pea CA activity (JR Coleman, unpublished data). We are currently investigating the relationship between CA activity and the ability to assimilate or metabolize cyanate.

Note Added in Proof

C. A. Roeske and W. L. Ogren have also recently identified and sequenced a pea cDNA clone coding for chlorplast carbonic anhydrase. The nucleotide and deduced amino acid sequences are presented in Nucleic Acids Research 18: 3413 (1990).

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dr. Gloria Coruzzi, The Rockefeller University, for the generous gift of the pea leaf cDNA library.

LITERATURE CITED