Short communication

Analysis of two linked genes coding for the acyl carrier protein (ACP) from Arabidopsis thaliana (columbia)

Gayle Lamppa* and Colleen Jacks1
Department of Molecular Genetics and Cell Biology, The University of Chicago, 920 East 58th St., Chicago, IL 60637, USA; 1 Current address: Department of Biology, Gustavus Adolphus College, St. Peter, MN 56082, USA

Received 20 September 1990; accepted in revised form 6 November 1990

Key words: acyl carrier protein, gene structure and expression, Arabidopsis thaliana (columbia)

Abstract

Two linked genes, A1 and A2, coding for nearly identical isoforms of the acyl carrier protein (ACP) were isolated from an Arabidopsis thaliana (columbia) genomic library and sequenced. The amino acids deduced from the nucleotide sequence of the two genes indicate they encode distinct transit peptides, but the mature proteins are the same except for residue 79. Both genes are predicted to contain three introns in similar positions, although they differ in sequence and length. The introns interrupt regions coding for a) the transit peptide, b) the junction of the transit peptide and mature protein, and c) the highly conserved domain surrounding serine 38 to which the phosphopantetheine is attached. Primer extension analysis indicates that at least A1 is active in young plants.

The acyl carrier protein (ACP) is a key component in de novo fatty acid biosynthesis. It is a small acidic protein with a 4'-phosphopantetheine prosthetic group, attached through a serine located 38 residues from the N-terminus, to which growing fatty acid chains are covalently bound. In higher plants the major site of fatty acid biosynthesis appears to be the chloroplast (see for review Ohlrogge, 1987). Thus, ACP is synthesized as a precursor, imported into the organelle, and processed to its mature form. Recently, it was demonstrated that the chloroplast contains a holo-ACP synthase activity capable of transferring the prosthetic group from CoA, suggesting that production of the functional moiety occurs after translation into the organelle (Fernandez and Lamppa, 1990). At least two isoforms of ACP have been detected by SDS-PAGE and peptide sequence analysis (Hoj and Svendsen, 1984; Ohlrogge and Kuo, 1985), but characterization of cDNA from Brassica (Safford et al., 1988; Rose et al., 1987), barley (Hansen and von Wettstein-Knowles, 1989) and spinach (Scherer and Knauf, 1987) indicates that other variants also exist in the cell.

An Arabidopsis thaliana (columbia) cosmid genomic library (Olszewki et al., 1988) was screened with a Brassica campestris seed-specific cDNA clone (Rose et al., 1987) coding for the ACP precursor. Cosmid DNA from a positive clone was digested with the enzymes shown in Fig. 1, and was predicted to contain ca. 18 kb of genomic DNA. Southern blot analysis using the Brassica cDNA insert (570 nt) as a probe showed that multiple bands hybridized in almost all of the restriction enzyme profiles. For example, in the
Hind III profile (Fig. 1A, lane 1) fragments of 3.9, 2.2 (open arrow) and 1.6 kb (closed arrow) hybridized with the cDNA insert that contains the entire coding sequence of the ACP precursor. The simplest pattern observed was after the Pst I/Sac I double digestion (Fig. 1A, lane 11), which showed two bands hybridizing with sizes of 3.7 (closed arrow) and 4.3 kb (open arrow). The complex hybridization profiles suggested the presence of several introns in a single ACP gene, or multiple ACP genes with internal restriction enzyme sites on the 18 kb genomic insert. To distinguish between these two possibilities, a degenerate 21-mer oligonucleotide was synthesized (see Fig. 1, legend), complementary to the sequence coding for residues DTVEIVM, which begin immediately carboxy to serine 38, and are highly conserved. In most profiles (Fig. 1B) the oligonucleotide hybridized to only two bands which were a subset of those that hybridized with the cDNA, strongly supporting the conclusion that the 18 kb insert contained two linked genes coding for ACP. The relative intensity of the hybridizing bands was about equal, in contrast to the pattern observed when the Brassica cDNA was used as a probe, suggesting that the Brassica sequence may be more distantly related to one of the two genes contained on these fragments. To substantiate this assessment, a partial restriction enzyme

Fig. 1. Detection of two linked genes, A1 and A2, coding for ACP. DNA was isolated from a cosmid clone that hybridized with a Brassica cDNA coding for ACP. The DNA was digested with the following restriction enzymes, as indicated, in lanes: 1) Hind III, 2) Hind III/Eco RI, 3) Hind III/Pst I, 4) Pst I, 5) Pst I/Eco RI, 6) Eco RI, 7) Eco RI/Sac I, 8) Eco RI/Pst I, 9) Kpn I/Sac I/Eco RI, 10) Kpn I/Sac I, 11) Pst I/Sac I, 12) Sac I, 13) Sac I/Hind III, 14) Kpn I/Hind III, 15) Kpn I, 16) Kpn I/Eco RI, 17) Kpn I/Pst I, and 18) Kpn I/Sac I. Southern blot analysis was performed using either the radiolabeled Brassica cDNA insert (panel A), or a 21-mer oligonucleotide (panel B). The sequence of the degenerate oligonucleotide is: CATNAC(A/G/T)AT(T/C)TC(T/C)ACNGT(A/G)TC. The cDNA was hybridized in 3 x SSC (48 °C), and washed in 0.5 x SSC, 0.1% SDS at 55 °C. The oligonucleotide was hybridized in 3 x SSC at 42 °C, and washed in 2 x SSC, 0.1% SDS at 33 °C. In the Hind III and Pst I/Sac I profiles, the large arrows point to the fragments carrying part of either A1 (open) or A2 (closed), as assessed by their relative hybridization with the Brassica cDNA. The arrowheads in panel A (left side) indicate the position of lambda Hind III/Eco RI markers, with sizes (top to bottom) of 21.7, 5.15/5.0, 4.27, 3.48, 1.98, 1.9, 1.59, 1.37, 0.94, 0.83, 0.56, 0.14 kb. In panel C, a partial restriction enzyme map is shown with the orientation of each gene, and its basic structure. In addition to the Hind III-Sac I fragment, for gene A1, 990 nt were sequenced upstream of the internal Hind III site, and 185 nt 3' of the Sac I site. For gene A2, 760 nt were sequenced 5' of the internal Hind III site, as well as 370 nt 3' of the Sac I site. The exon boxes represent the transit peptide (dotted) and mature (black) regions of the ACP precursor. The numbers refer to the approximate size (kb) of each intron. The position of the TATA box relative to the start codon is indicated in nt. The distances between the Pst I and most distal Hind III site are ca. 5.6 kb and 5.7 kb for A1 and A2, respectively.