Diaminopimelate Decarboxylase from Arabidopsis
Contains Motifs for Pyridoxal-5'-phosphate and Substrate

Jung-Sup Kim
Faculty of Biotechnology, Jeju National University, Jeju 690-756, Korea

Abstract: Diaminopimelate decarboxylase (DAPDC, EC 4.1.1.20) is an enzyme catalyzing the formation of lysine (Lys) from diaminopimelate (DAP) by pyridoxal-5'-phosphate (PLP)-dependent decarboxylation in Lys biosynthetic pathway. The genes for DAPDC were reported in many bacteria, but not in eukaryotes including plants. The present study reported the characterization of a gene for DAPDC from Arabidopsis thaliana (DAFDC1). Sequence analysis of clones was revealed a full-length open reading frame for DAPDC1 which is 1455 bp and encoded 484 amino acid (AA). The DAPDC1 protein contains 2 consensus binding site for PLP as a cofactor and DAP as substrate and has a sequence at the amino terminus that resembles a transit peptide for localization to chloroplast. The predicted amino acid sequence of DAPDC1 is highly homologous to that of the enzymes for DAPDC encoded by lysA of many bacteria. DAPDC1 is ubiquitously expressed in Arabidopsis.

Key words: Arabidopsis, diaminopimelate decarboxylase, lysA, PLP, DAP

INTRODUCTION

Lysine (Lys) is an essential amino acid in animals including humans. The biosynthetic pathway of Lys including methionine and threonine is initiated from aspartate called aspartate family pathway (Fig. 1). Through six more following steps, Diaminopimelate (DAP) is synthesized. Finally, Lys is formed by Pyridoxal-5'-phosphate (PLP)-dependent decarboxylation of DAP by DAP decarboxylase (DAPDC, EC 4.1.1.20) as shown in Fig. 1. The genes, lysA, encoding DAPDC were cloned and characterized from many bacteria (Cremer et al., 1988; Martin et al., 1988; Yamamoto et al., 1991; Mills and Flickinger, 1993). The DAPDC is well characterized in Mycobacterium tuberculosis and is needed PLP for enzyme activity (Gokulan et al., 2003). The DAPDC protein is composed of 406, 445 and 439 amino acids in N. meningitides, C. glutamicum and B. subtilis (Parkhill et al., 2000; Cremer et al., 1988; Yamamoto et al., 1991), respectively.

The binding motif for PLP, a cofactor of DAPDC, is well conserved as a sequence “[FY]-[PA]-x-K-[SACV]-[NHCLFW]-x(4)-[LIVMF]-[LIVMVT]-x(2)-[LIVMA]-x(3)-[GTE]” in bacteria. The covalent bond between DAPDC and the cofactor is formed via formation of an internal aldimine (Gokulan et al., 2003).

Many antibiotics or herbicides for killing microorganisms or plants, respectively, are targeted to a specific enzyme in amino acid biosynthesis (Girodean et al., 1986; Kelland et al., 1986). The several toxic analogs of DAP, which is a substrate of DAPDC, have well been used as antibiotics (Kelland et al., 1986; Ray et al., 2002). The principle of antibiotics is that DAP analogs are toxic using as competitive inhibitors and finally cause to death because Lys deprivation was resulted by inhibition of DAPDC enzyme activity (Kelland et al., 1986).

The genes for DAPDC were well characterized in bacteria, but not in plants. Only partial sequence of probable DAPDC was reported to overexpress by interaction between Arabidopsis plant and nematode
Meloidogyne incognita (Vercauteren et al., 2001). The characterization of another isotype gene for DAPDC from *A. thaliana* to investigate Lys biosynthetic pathway in plants was reported in the present study.

**MATERIALS AND METHODS**

**Sequencing analysis:** An EST clone (GenBank Accession Number AV826463, cDNA ID RAFL08-15-L22) was obtained from the Arabidopsis Biological Resource Center (ABRC). The clone was derived from *A. thaliana* cDNA library (Seki et al., 2004) from rosette plants prepared in FLC-1 vector. DNA sequencing was performed by an automatic sequencer (A1Express DNA sequencer, Pharmacia Biotech. Inc., UK) with synthetic oligonucleotide primers. Nucleotide sequences and amino acid sequences were compared with sequences present in the GenBank and EMBL databases and analyzed using BLAST (Wheeler et al., 2001) and Clustal W multiple sequence alignment program (Thompson et al., 1994) or Biology WorkBench 3.2 (http://workbench.sdsc.edu; San Diego Supercomputer Center, University of California San Diego). Comparison of sequences was performed at the nucleotide and amino acid level. Motifs were searched by GenomeNet Computation Service at Kyoto University (http://www.genome.ad.jp).

**Reverse Transcription-polymerase Chain Reaction (RT-PCR):** RT-PCR analysis (Sambrook and Russell, 2001) was performed as follows. Total RNA from roots, leaves, stalks and flowers in 45 day-old Arabidopsis was extracted using the TRI-reagent kit (Molecular Research Center, USA).

An RT-PCR kit (Titan one tube RT-PCT system, Boehringer Mannheim) was used to investigate expression of *DAPDC1* transcripts in various tissues. Reverse Transcription (RT) reaction of 1°C total RNA was carried out as described by the manufacturer. RT was performed at 42°C for 60 min in a PCR system 9700 (Perkin Elmer, USA). After the cDNA synthesis was complete, the denatured cDNA was amplified by using designed primers from *DAPDC1*-specific sequence or tubulin2 sequence: DAP1 (5'-ACACCAACTGAAAGAACCAACAC-3') and DAP2 (5'-CGATGTTGAACTGATTAATCGT-3') or tub2-1 (5'-TCACCTTCTCTCTCGATT-3') and tub2-2 (5'-TCACCTTCTCTCGATT-3'). The polymerase chain reaction was performed using the PCR system 9700 for 30 cycles with a program 94°C for 1 min, 56°C for 2 min and 72°C for 2 min, with 1 nM primers. PCR products were analyzed on 1%(w/v) agarose gel.

**RESULTS AND DISCUSSION**

The EST clone (GenBank Accession Number AV826463) was determined the nucleotide sequence using designed primers. The nucleotide sequence data was deposited in the GenBank Databases under the accession AF227913. The sequence, *DAPDC1* cDNA, was contained a full-length open reading frame consisted in 1455 bp and encoded a protein of approximately 53.6 kDa. The expected isoelectric point of the protein is 6.54. The predicted amino acid sequence of DAPDC1 is homologous to the DAPDC sequences from bacteria which showed similarities with 60, 56 and 56% to proteins from *Aquifex aeolicus* (GenBank Accession Number AAC07209), *Rubrobacter xylanophilus* (GenBank Accession Number EAN36364) and *Shewanella oneidensis* (GenBank Accession Number AAN5278), respectively. DAPDC1 also shows significant homology with Ornithine Decarboxylase (OmnDC) and arginine Decarboxylase (ArgDC), members of PLP-requiring enzymes.

By the analysis of motif with amino acid sequence of DAPDC1, it was revealed that there is a signature-binding motif for a PLP in the N-terminal region (122-140) of the DAPDC1 (Fig. 2). The motif sequence (YNKANNLKILEHRLRSA) for a PLP is highly homologous to the consensus "[FG]-[PA]-x-[K]-[SACV]-[NHCLFW]-x4-[LVMF][LVMTA]-x2[LIVM]-x3-[QTE]" which underlined amino acids are well conserved. The binding motif for a PLP is present and well conserved in bacterial DAPDCs and Omn/Arg decarboxylases family using PLP as a cofactor. The exact PLP binding site is thought to be K-125 identified by comparison with the binding site of *Mycobacterium tuberculosis* LysA as shown in Fig. 2 (Gokulan et al., 2003). The finding suggests that the DAPDC1 product uses PLP as a cofactor for enzyme activity. Another signature binding motif for substrate exists in the middle region (293-306) of the DAPDC1 (Fig. 2). The motif sequence (GFEVSIGNGGGLG) for DAP is highly homologous to the consensus "[gSA]-x2(2,6)-[LIVMSCP]-x2(2,[LIVMF]-[DANS]-[LVMCA]-G[G][LVMFY]-[GSTPCEF]" which underlined amino acids are well conserved. The binding motif for a DAP/Omn/Arg is present and well conserved in bacterial DAP/Omn/Arg decarboxylases family. Phylogenetic analysis of the related sequences is revealed further that plant DAPDCs are evolved from ancestral bacterial DAPDCs along with OmnDCs and ArgDCs (Fig. 3).

The size of DAPDC1 consisted of 484 amino acids are slightly larger than DAPDCs from bacteria. The amino terminal ~60 amino acids of the DAPDC1 product are not
Fig 2: Multiple sequence alignment of DAPDC1 and DAPDC-related proteins from Arabidopsis (AtDAPDC1 & AtDAPDC2) and several bacteria, were aligned using ClustalW and a color-coded plot was generated using BlastPS program. Yellow background indicates 100% conserved residues of the amino acids, green is identical residues, blue is similar residues. GenBank accession numbers are as follows: AF227913 (AtDAPDC1, this study), NP-2682 (AtDAPDC2), AAC07209 (AtDAPDC from Arabiscus cephalus), EAB74394 (RdDAPDC from Rubrobacter xylanophilus), AAB37278 (SoDAPDC from Synechococcus elongatus), EAL79931 (AdoDAPDC from Arthrobacter dehalogenans), AAB04338 (AadDAPDC from Arthrobacter fulgidus), EAL79838 (CtDAPDC from Chromatidobacter solanum), ZP-00141758 (PaDAPDC from Pseudomonas aeruginosa) and CAE837 (PaDAPDC from Neisseria meningitidis). The binding motif for PLP and DAP is indicated including consensus sequences.
Similar to any part of the bacterial enzymes (Fig. 2). This region contains a high concentration of alanine and hydroxyl amino acids such as serine, 11 over 37 amino acid residues, features of a transit or signal sequences for protein transport into plastids (Nakai and Kanesawa, 1992). Organelar transit peptides are usually cleaved from the precursor protein and are not necessary for attainment of catalytic activity of the mature protein such as cystathionine γ-synthase targeted to chloroplasts (Kim and Leustek, 1996). In this respect it is suggested that the DAPDC1 is synthesized in cytosol and transported into chloroplasts and has become active after cleavage of transit peptide.

Through the analysis of amino acid homology in database with DAPDC1, it is suggested that there are functional DAPDC homologs in plants. Database analysis shows that DAPDC1 sequence is identical to a genomic region located in chromosome III in A. thaliana (At3g14390) and consisted of 8 exons. The other sequence from Arabidopsis found by database search with DAPDC1, DAPDC2 (AT5g11880, GenBank Accession Number AY049266), has high homology with the identity of 93% to DAPDC1 amino acid sequence. Another homologous sequence, DAPDC2, is located in chromosome V in A. thaliana (AT5g11880, GenBank Accession Number AY049266). The genomic organization between the genes encoding DAPDC1 and DAPDC2 is highly similar in size and location of exons and introns. The 95% similarity in amino acid sequence level in addition to highly similar genomic structure between the two genes suggests that the products for DAPDC1 and DAPDC2 are the isotype performing similar function on the Lys biosynthetic pathway in plants. Previous report of partial sequence about DAPDC2 was revealed as DAPDC2 gene which is reported to be overexpressed by interaction between Arabidopsis plant and nematode M. incognita (Vercauteren et al., 2001). Another sequence from rice found by database search with DAPDC1, OsDAPDC1 (GenBank Accession Number AK067100), is highly homologous to DAPDC1 amino acid sequence with 82% similarity. It also contains chloroplast transit peptide, PLP and DAP binding motifs on the corresponding regions compared to the amino acid sequence of DAPDC1 from Arabidopsis.

To investigate the spatial expression pattern of DAPDC1 transcripts in various tissues of Arabidopsis plants, RT-PCR was performed with DAPI and DAP2 primer sets. Total RNA was isolated from leaves, stolts, roots and flowers of mature plants. An expected 0.8 kb fragment as RT-PCR product from all tissues tested was detected on a gel (Fig. 4). The RT-PCR analysis indicate that DAPDC1 is expressed in all tissues, including leaves,
stalks, roots and flowers, an expected result based on the assumption that Lys biosynthetic pathway is active in all tissues of plants. The expression level of DAPDC1 was similar in all tissues tested.

It has been reported the functional characterization of DAPDC encoded by lysA gene in many bacteria, but not in plants. Here I report the isolation of a gene encoding the A. thaliana DAPDC (DAPDC1). Following determination of the nucleotide sequences, the encoded protein was analyzed a putative novel enzyme for DAPDC in plants. The full-length cDNA encodes a 484-amino acid protein that is highly homologous to bacterial LysA function as DAPDC. There is another homologous gene with DAPDC1, DAPDC2, in chromosome V of Arabidopsis which has higher identity over 90% in nucleotide and amino acid sequence level which they assume to be perform similar functions in plants.

The findings also suggest that there is a protein performing similar function as DAPDC1 in plants such as A. thaliana and O. sativa. The DAPDC1 protein from Arabidopsis is consisted of 484 amino acids which are slightly larger in size compared to that from bacteria, which suggest there is a transit peptide to plastids in the amino terminal of the protein.

Currently, I am screening T-DNA insertion mutants in which the expression of each Arabidopsis DAPDC gene is knockout. This approach would provide some important clues into the substrate specificity and physiological function of this noble enzyme for amino acid metabolism in plants.

Recently, an analysis of Lys biosynthetic pathway in the genome of Arabidopsis was reported including two DAPDC genes which were complemented in lysA mutants of E. coli (Hudson et al., 2005). Our report about cloning and characterization of a cDNA encoding DAPDC1 from Arabidopsis confirm the predictions and extend to motifs and expression patterns in plant tissues. It would be a starting point in molecular level to investigate Lys biosynthesis in plants and to screen herbicides and antibiotics or to apply further to modify nutritional composition by metabolic engineering with overexpression or suppression in important crops such as rice.

ACKNOWLEDGEMENTS

I wish to thank ABRC at the Ohio State University for providing an EST clone (RAFL08-15L22). This work was supported by a grant from BioGreen21 Program, Rural Development Administration, Republic of Korea.

REFERENCES


