Asian Journal of Plant Sciences

ISSN 1682-3974

science alert

ANSInet
an open access publisher
http://ansinet.com

Asian Journal of Plant Sciences 5 (2): 260-265, 2006
ISSN 1682-3974

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

Jung-Sup Kim
Faculty of Biotechnology, Cheju 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* (*DAPDC1*). Sequence analysis of clones was revealed a full-length open reading frame for *DAPDC1* which is 1455 bp and encoded 484 amino acids. The DAPDC1 protein contains a 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]-[LIVMTA]-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 (Girodeau *et al*., 1986; Kelland *et al*., 1986). The several toxic analogs of DAP, which is a substrate of DAPDC,

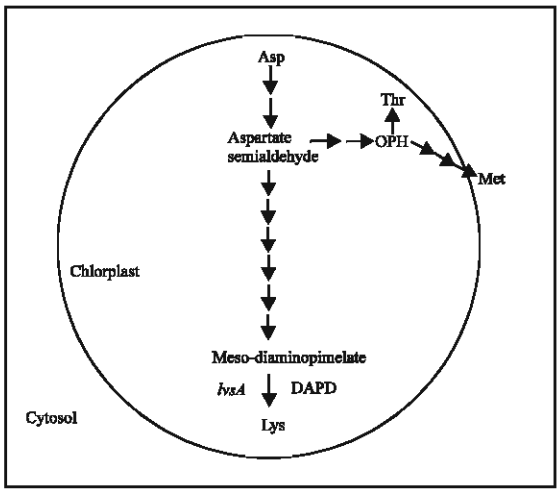

Fig. 1: The pathway of lysine biosynthesis in plants
The abbreviation of chemicals are as follows; Asp, aspartate; Thr, threonine; Met, methionine; Lys, lysine; OPH, O-phosphohomoserine

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 (A1Fexpress 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'-ACACCAACTGAAGAACCAAAC-3') and DAP2 (5'-CGGATCTCGTCAATGTATTCT-3') or tub2-1 (5'-CTCAAGAGGTTCTCAGCAGTA-3') and tub2-2 (5'-TCACCTTCTTCATCCGCAGTT-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 AAN57278), respectively. DAPDC1 also shows significant homology with Ornithine Decarboxylase (OrnDC) 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 (YAIKANNNLKILEHLRSLG) for a PLP is highly homologous to the consensus "[FY]-[PA]-x-K-[SACV]-[NHCLFW]-x(4)-[LIVMF]-[LIVMTA]-x(2)-[LIVMA]-x(3)-[GTE]" which underlined amino acids are well conserved. The binding motif for a PLP is present and well conserved in bacterial DAPDCs and Orn/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 (GFEVSYLNIGGGLG) for DAP is highly homologous to the consensus "[GSA]-x(2,6)-[LIVMSCP]-x(2)-[LIVMF]-[DNS]-[LIVMCA]-G-G-G-[LIVMFY]-[GSTPCEQ]"which underlined amino acids are well conserved. The binding motif for a DAP/Orn/Arg is present and well conserved in bacterial DAP/Orn/Arg decarboxylases family. Phylogenetic analysis of the related sequences is revealed further that plant DAPDCs are evolved from ancestral bacterial DAPDCs along with OrnDCs 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

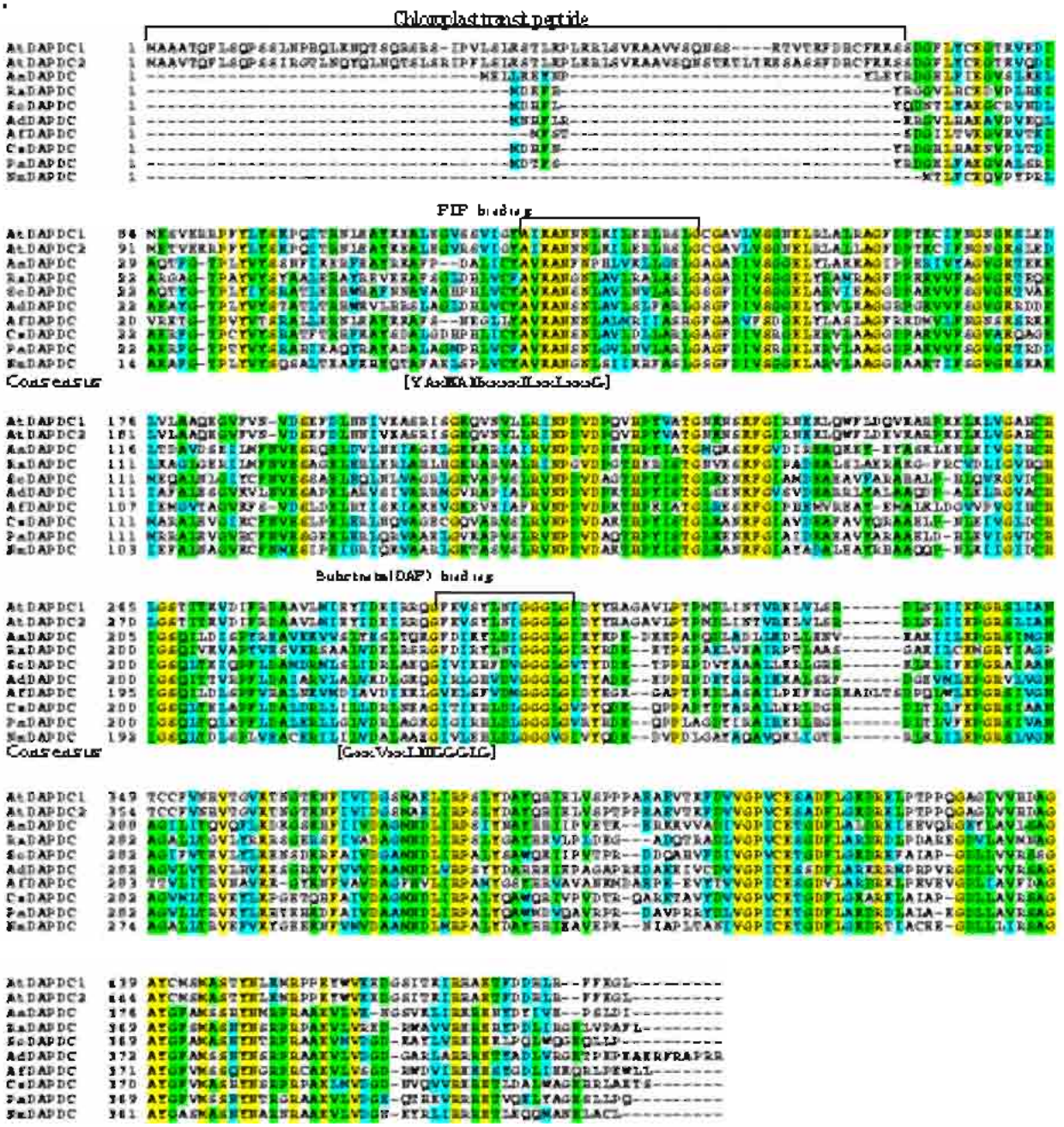

Fig. 2: Multiple sequence alignment of DAPDC1 and DAPDC-related proteins from *Arabidopsis* (AtDAPDC1 a AtDAPDC2) and several bacteria, were aligned using ClustalW and a color-coded plot was generated using Boxsha program. Yellow background indicates 100% conserved residues of the amino acids, green is identical residues, a blue is similar residues. GenBank accession numbers are as follows; AF227913 (AtDAPDC1, this study), NP-5682 (AtDAPDC2), AAC07209 (AaDAPDC from *Aquifex aeolicus*), EAN36364 (RxDAPDC from *Rubrobac xylanophilus*), AAN57278 (SoDAPDC from *Shewanella oneidensis*), EAL78951 (AdDAPDC from *Anaeromyxobac dehalogenans*), AAE90438 (AfDAPDC from *Archaeoglobus fulgidus*), EAM22808 (CsDAPDC fr *Chromohalobacter salexigens*), ZP-00141757 (PaDAPDC from *Pseudomonas aeruginosa*) and CAB837 (NmDAPDC from *Neisseria meningitidis*). The binding motif for PLP and DAP is indicated including consens sequences

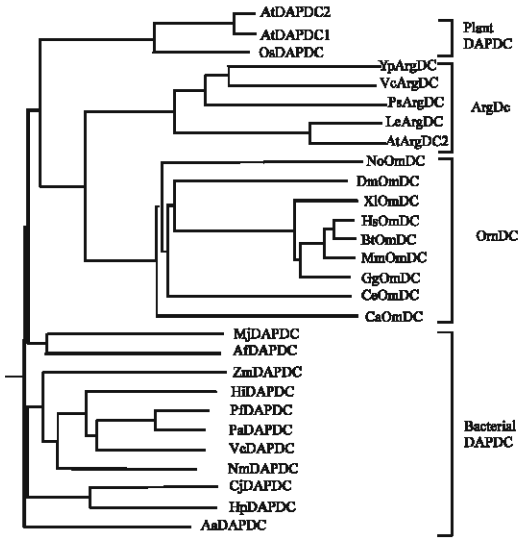

Fig. 3: Phylogenetic tree

Phylogenetic analysis of DAPDC related proteins was performed by the neighbor-joining method using midpoint rooting option with PHYLIP program. Accession numbers are as follows: AF227913 (AtDAPDC1, this study), AAM98306 (AtDAPDC2 from *Arabidopsis thaliana*), XP_465375 (OsDAPDC from *Oryza sativa*), AAS63667 (YpArgDC from *Yersinia pestis*), AAF96713 (VcArgDC from *Vibrio cholerae*), AA058271 (PsArgDC from *Pseudomonas syringae* pv. Tomato), CAI39242 (LeArgDC from *Lycopersicon esculentum*), O23141 (AtArgDC2 from *Arabidopsis thaliana*), P27121 (NcOrnDC from *Neurospora crassa*), P40807 (DmOrnDC from *Drosophila melanogaster*), P27120 (XlOrnDC from *Xenopus laevis*), P11926 (HsOrnDC from *Homo sapiens*), P27117 (BtOrnDC from *Bos taurus*), P00860 (MmOrnDC from *Mus musculus*), P27118 (GgOrnDC from *Gallus gallus*), P41931 (CeOrnDC from *Caenorhabditis elegans*), P78599 (CaOrnDC *from Candida albicans*), AAB99100 (MjDAPDC from *Methanococcus jannaschii*), AAB90438 (AfDAPDC from *Archaeoglobus fulgidus*), AAY90392 (ZmDAPDC from *Zymomonas mobilis*), AAX87778 (HiDAPDC from *Haemophilus influenzae*), AAY95201 (PfDAPDC from *Pseudomonas fluorescens*), ZP_00141757 (PaDAPDC from *Pseudomonas aeruginosa*), AAF93302 (VcDAPDC from *Vibrio cholerae*), CAB83764 (NmDAPDC from *Neisseria meningitidis*), CAB72781 (CjDAPDC from *Campylobacter jejuni*), AAD05866 (HpDAPDC from *Helicobacter pylori* J99), AAC07209 (AaDAPDC from *Aquifex aeolicus*)

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 Kanehisa, 1992). Organellar 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 *DAPDC* 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 DAP1 and DAP2 primer sets. Total RNA was isolated from leaves, stalks, 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,

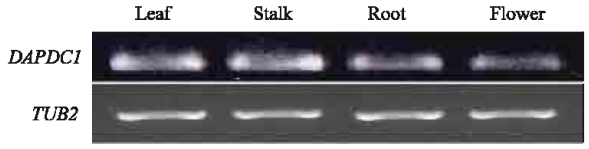

Fig. 4: Analysis of *DAPDC1* expression in various tissues by RT-PCR. The tubulin gene was amplified as a quantitative control

stalks, roots and flowers, an expected results based on 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 functioning 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 reports 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 (RAFL08-15-L22). This work was supported by a grant from BioGreen21 Program, Rural Development Administration, Republic of Korea.

REFERENCES

Cremer, J., C. Treptow, L. Eggeling and H. Sahm, 1988. Regulation of enzymes of lysine biosynthesis in *Corynebacterium glutamicum*. J. Gen. Microbiol., 134: 3221-3229.

Girodeau, J.M., C. Agouridas, M. Masson, R. Pineau and F. Le Goffic, 1986. The lysine pathway as a target for a new genera of synthetic antibacterial antibiotics? J. Med. Chem., 29: 1023-1030.

Gokulan, K., B. Rupp, M.S. Jr. Pavelka, W.R. Jr. Jacobs and J.C. Sacchettini, 2003. Crystal structure of *Mycobacterium tuberculosis* diaminopimelate decarboxylase, an essential enzyme in bacterial lysine biosynthesis. J. Biol. Chem., 278: 18588-18596.

Hudson, A.O., C. Bless, P. Macedo, S.P. Chatterjee and B.K. Singh *et al*., 2005. Biosynthesis of lysine in plants: Evidence for a variant of the known bacterial pathways. Biochim. Biophys. Acta, 1721: 27-36.

Kelland, J.G., L.D. Arnold, M.M. Palcic, M.A. Pickard and J.C. Vederas, 1986. Analogs of diaminopimelic acid as inhibitors of meso-diaminopimelate decarboxylase from *Bacillus sphaericus* and wheat germ. J. Biol. Chem., 261: 13216-13223.

Kim, J. and T. Leustek, 1996. Cloning and analysis of the gene for cystathionine gamma-synthase from Arabidopsis thaliana. Plant Mol. Biol., 32: 1117-1124.

Martin, C., B. Cami, P. Yeh, P. Stragier, C. Parsot and J.C. Patte, 1988. *Pseudomonas aeruginosa* diaminopimelate decarboxylase: Evolutionary relationship with other amino acid decarboxylases. Mol. Biol. Evol., 5: 549-559.

Mills, D.A. and M.C. Flickinger, 1993. Cloning and sequence analysis of the meso-diaminopimelate decarboxylase gene from *Bacillus methanolicus* MGA3 and comparison to other decarboxylase genes. Applied Environ. Microbiol., 59: 2927-2937.

Nakai, K. and M. Kanehisa, 1992. A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics, 14: 897-911.

Parkhill, J., M. Achtman, K.D. James, S.D. Bentley and C. Churcher *et al*., 2000. Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491. Nature, 404: 502-506.

Ray, S.S., J.B. Bonanno, K.R. Rajashankar, M.G. Pinho and G. He *et al*., 2002. Cocrystal structures of diaminopimelate decarboxylase: mechanism, evolution and inhibition of an antibiotic resistance accessory factor. Structure (Camb), 10: 1499-1508.

Sambrook, J. and R.W. Russell, 2001. Molecular Cloning: A Laboratory Manual, 3rd Edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Seki, M., M. Satou, T. Sakurai, K. Akiyama and K. Iida *et al*., 2004. RIKEN Arabidopsis full-length (RAFL) cDNA and its applications for expression profiling under abiotic stress conditions. J. Exp. Bot., 55: 213-223.

Thompson, J.D., D.G. Higgins and T.J. Gibson, 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res., 22: 4673-4680.

Vercauteren, I., E. Van Der Schueren, M. Van Montagu and G. Gheysen, 2001. *Arabidopsis thaliana* genes expressed in the early compatible interaction with root-knot nematodes. Mol. Plant Microbe Interact., 14: 288-299.

Wheeler, D.L., D.M. Church, A.E. Lash, D.D. Leipe and T.L. Madden *et al*., 2001. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res., 29: 11-16.

Yamamoto, J., M. Shimizu and K. Yamane, 1991. Molecular cloning and analysis of nucleotide sequence of the *Bacillus subtilis lysA* gene region using *B. subtilis* phage vectors and a multi-copy plasmid, Pub1.10. Agric. Biol. Chem., 55: 1615-1626.