Pseudomonas aeruginosa Diaminopimelate Decarboxylase: Evolutionary Relationship with Other Amino Acid Decarboxylases

Christine Martin, * Brigitte Cami, * Patrice Yeh, † Patrick Stragier, ‡ Claude Parsot, § and Jean-Claude Patte*

*Laboratoire de Chimie Bacterienne, Centre National de la Recherche Scientifique, Marseille; †Genetica, Joinville-le-Pont, France; ‡Institut de Microbiologie, Université Paris-Sud, Orsay, France; and §Unite de Biochimie Cellulaire, Institut Pasteur, Paris

The lysA gene encodes meso-diaminopimelate (DAP) decarboxylase (E.C.4.1.1.20), the last enzyme of the lysine biosynthetic pathway in bacteria. We have determined the nucleotide sequence of the lysA gene from Pseudomonas aeruginosa. Comparison of the deduced amino acid sequence of the lysA gene product revealed extensive similarity with the sequences of the functionally equivalent enzymes from Escherichia coli and Corynebacterium glutamicum. Even though both P. aeruginosa and E. coli are Gram-negative bacteria, sequence comparisons indicate a greater similarity between enzymes of P. aeruginosa and the Gram-positive bacterium C. glutamicum than between those of P. aeruginosa and E. coli enzymes. Comparison of DAP decarboxylase with protein sequences present in data bases revealed that bacterial DAP decarboxylases are homologous to mouse (Mus musculus) ornithine decarboxylase (E.C.4.1.1.17), the key enzyme in polyamine biosynthesis in mammals. On the other hand, no similarity was detected between DAP decarboxylases and other bacterial amino acid decarboxylases.

Introduction

Amino acid sequence comparisons have been widely used as a tool for a better understanding of the relationships between structure and function in proteins. Sequence similarities have allowed some rules underlying the evolution and specialization of enzymes catalyzing different reactions in cellular metabolism to be proposed. For example, aspartate and ornithine carbamoyltransferases, involved in the pyrimidine and in the arginine biosynthetic pathways, respectively, present extensive sequence similarity (Van Vliet et al. 1984). A common origin has been proposed for enzymes involved in threonine, tryptophan, and isoleucine biosynthesis, as well as for some methionine and cysteine biosynthetic enzymes (Parsot et al. 1987). As the different steps of these anabolic pathways are the same in various species, the divergence of these homologous enzymes probably began early. Conversely, the lysine biosynthetic pathway is not universal, as it involves two different pathways in bacteria and in fungi.

1. Key words: decarboxylase, DAP decarboxylase, ornithine decarboxylase, Pseudomonas, evolution, pyridoxal phosphate, lysine.

2. Current address: INRA-CNRZ de Theix, 63122 Ceyrat, France.

3. Current address: Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115.

Address for correspondence and reprints: Jean-Claude Patte, Laboratoire de Chimie Bacterienne, CNRS, BP71, 13277 Marseille CX 9, France.

© 1988 by The University of Chicago. All rights reserved.
0737-4038/88/0505-0007$02.00

549
In *Escherichia coli*, lysine biosynthesis is achieved in nine steps from aspartate, the two first steps being common to the methionine, threonine, and isoleucine biosynthetic pathways (see review by Patte 1983). Lysine results from the decarboxylation of meso-diaminopimelate (DAP), which is also a characteristic component of the bacterial cell wall. In the Gram-positive organism *Corynebacterium glutamicum*, steps 5–8 of the *E. coli* pathways are bypassed by the reaction catalyzed by meso-DAP dehydrogenase, but DAP remains the direct precursor of lysine (Misono et al. 1979; Ishino et al. 1984; Yeh et al., 1988a). In *Saccharomyces cerevisiae*, as in other fungi, lysine is synthesized from α-ketoglutarate by the so-called α-aminoadipate pathway, which does not involve DAP (Bhattacharjee and Sinha 1972).

We have recently cloned the *lysA* gene, encoding DAP decarboxylase (EC 4.1.1.20), from *Pseudomonas aeruginosa* (Martin et al. 1986). We present here the determination of the nucleotide sequence of this gene and the comparison of the amino acid sequences of DAP decarboxylases from *E. coli*, *C. glutamicum*, and *P. aeruginosa*. The structural relationships of these bacterial DAP decarboxylases are presented and compared with those of other basic amino acid decarboxylases.

**Material and Methods**

**Strains and Plasmids**

*Escherichia coli* K12 derivatives used were MC 1061 (*araD139 Δara-leu lacX74 galU galK hisD rpsL*); RRI ΔM15 (*leu pro thi rpsL hisD ΔM15 Δm15, F’[lacIΔ lacZ ΔM15 pro^+]*); RRI ΔM15 *lysA* (RRI ΔM15 made Lys^− by transduction of a *lysA*: Km^+ mutation; C. Printz and P. Stragier, unpublished data). The previously isolated pLP218 recombinant plasmid (Martin et al. 1986) was used as a source of *Pseudomonas aeruginosa* *lysA* DNA in the constructions of plasmids described below. Plasmid pLP818 was obtained by inserting the 1.7-kb-long NcoI fragment of plasmid pLP218 (fig. 1) into the NcoI site located downstream from a promoter and a ribosomal binding site in plasmid vector PKK 233-2 (Amann and Brosius 1985). To locate the end of the *lysA* coding sequence, we constructed plasmid pLP220 as follows: the 190-bp-long TaqI fragment of plasmid pLP218 (from position 1147 to 1332) was first inserted into the AccI site of pUC 19 (Vieira and Messing 1982). From the pTA19 recombinant plasmid thus obtained, a 170-bp-long NotI-HindIII fragment (extending from the NotI site of the insert to the HindIII site of the vector polylinker) was purified. This DNA fragment was then inserted between the NotI and HindIII sites of pLP218 in place of the original 3.8-kb-long NotI-HindIII fragment. Therefore, the *P. aeruginosa* DNA ends at the TaqI site (fig. 2, position 1332) in the pLP220 plasmid.

Plasmid pBN72 was constructed in order to verify the *lysA* reading frame at the *SphI* site (fig. 2, position 170). First, a 3.1-kb-long EcoRI-DraI fragment of pNM480 (Minton 1984) containing the *lacZ* gene devoid of its first seven codons was inserted between the EcoRI and EcoRV sites of plasmid pBR322 (Bolivar et al. 1977), giving rise to plasmid pBN480. The 698-bp-long PvuII-SphI fragment of pLP218 was cloned between the SmaI and SphI sites of pUC19 (Vieira and Messing 1982) to give plasmid pUP72. This second construction provided an EcoRI-HindIII DNA fragment encompassing the entire insert, which was then cloned between the EcoRI and HindIII sites located 5′ of *lacZ* in plasmid pBN480. *lysA-lacZ* fusions were also constructed at the *SalI* site (fig. 2, position 607). For this purpose, the product of the digestion of pLP218 by EcoRI and SalI was cloned between the EcoRI and SalI sites of the pNM480, pNM481, and pNM482 plasmids (Minton 1984), leading to pNP5, pNP6, and pNP7, respectively.
P. aeruginosa DAP Decarboxylase

DNA Manipulation

Transformation of E. coli was done according to the method of Dagert and Erlich (1979). Large-scale purification of plasmid DNA was made according to the method of Humphreys et al. (1975). For rapid analysis of recombinant plasmids, the alkaline lysis procedure of Birnboim and Doly (1979) was used.

Restriction enzymes were purchased from New England Biolabs and Bethesda Research Laboratories and were used under the conditions recommended by the manufacturers. Polynucleotide kinase and terminal deoxynucleotidyl transferase were purchased from Boehringer Mannheim. (\gamma^{32}\text{P})\text{ATP} and (\alpha^{32}\text{P}) dideoxyadenosine were purchased from Amersham. Nucleotide sequence of both DNA strands was determined by the procedure described by Maxam and Gilbert (1980).

Results and Discussion

Nucleotide Sequence of the Pseudomonas aeruginosa lysA Gene

We have previously isolated the P. aeruginosa lysA gene on the pLP218 recombinant plasmid (fig. 1) through its ability to complement lysA mutations in both P. aeruginosa and Escherichia coli (Martin et al. 1986). The direction of transcription of lysA as well as its approximate location on pLP218 were then determined using miniMudII 1734 mutagenesis (data not shown) using the procedure described by de Castilho et al. (1984). This led us to determine the nucleotide sequence of the NcoI-SalI fragment according to the sequencing strategy shown in figure 1. Examination of the nucleotide sequence revealed the existence of a large open reading frame in the orientation predicted by miniMudII 1734 experiments. The accuracy of this
The identification of the initiation codon as the ATG at position 22 (fig. 2) was based on the following criteria: (1) the existence upstream of this ATG codon of a convenient ribosomal binding site (5'GAGA 3') fitting with the consensus sequence described in *P. aeruginosa* (Nakai et al. 1983) and (2) the ability of a DNA fragment starting exactly at this ATG codon to complement an *E. coli* lysA strain. A 1.7-kb-long *NcoI* DNA fragment starting at position 20 (see fig. 2) was inserted downstream from an isopropyl-β-thiogalactoside (IPTG)–inducible *trc* promoter and a convenient ribosome binding site. The pLP818 recombinant plasmid thus obtained allowed complementation of an *E. coli* lysA strain. Moreover, addition of IPTG into the growth medium of the transformed strain was lethal, probably as a consequence of the overproduction of DAP decarboxylase and the destruction of the DAP pool. This phenomenon was previously observed when the *E. coli* lysA gene was overexpressed in *E. coli* (P. Stragier and J.-C. Patte, unpublished observations).

The end of the coding sequence was also verified using plasmid pLP220 (see Material and Methods), which carries an insert extending to the *TaqI* site (position 1332). This plasmid transformed an *E. coli* lysA strain to a Lys+ phenotype, a result suggesting the presence of a complete lysA gene in the insert.

The *P. aeruginosa* lysA gene thus encodes a 415-amino-acid polypeptide with a deduced molecular weight of 45,496 daltons. Table 1 shows the codon usage within the *P. aeruginosa* lysA gene. Since the GC content of *P. aeruginosa* is quite high (~67%), G or C occurs in the third-codon position ~90% of the time. However, as already noted (Minton et al. 1984; Hadero and Crawford 1986; Spooner et al. 1986), there is a preference for C over G (see the Thr, Ala, Gly, and Arg codons). Moreover, a strong bias is observed for the UUC phenylalanine codon, the AUC isoleucine codon, the CAC histidine codon, the AAC asparagine codon, the GAC aspartate codon, and the GAG glutamate codon. We also note that arginine is overrepresented (mostly by the CGC codon) relative to lysine, as compared with the *E. coli* proteins.

**Amino Acid Sequence Comparison of Bacterial DAP Decarboxylases**

The lysA genes from *E. coli* (Stragier et al. 1983), *Corynebacterium glutamicum* (Yeh et al. 1988b), and *P. aeruginosa* (present paper) encode 420-, 445-, and 415-amino-acid-residue-long polypeptides, respectively. Amino acid sequences of the three DAP decarboxylases were compared by pairs by using the Needleman and Wunsch algorithm (1970), and the three alignments were then combined into a single one (fig. 3). Excluding the positions where gaps have been introduced in some of the sequences,
Table 1

| Codon Usage in the *lysA* Gene of *Pseudomonas aeruginosa* |
|------------------|-----------|----------------|-----------|-----------|
| TTT              | Phe       | 0              | TCT       | Ser       | 0          | TAT       | Tyr       | 3          | TGT       | Cys   | 0    |
| TTC              | Phe       | 13             | TCC       | Ser       | 8          | TAC       | Tyr       | 12         | TGC       | Cys   | 4    |
| TTA              | Leu       | 0              | TCA       | Ser       | 0          | TAA       | Stop      | 0          | TGA       | Stop  | 1    |
| TIG              | Leu       | 1              | TCG       | Ser       | 4          | TAG       | Stop      | 0          | TGG       | Trp   | 1    |
| CTT              | Leu       | 2              | CCT       | Pro       | 0          | CAT       | His       | 1          | CGT       | Arg   | 4    |
| CTC              | Leu       | 15             | CCC       | Pro       | 3          | CAC       | His       | 10         | CGC       | Arg   | 31   |
| CTA              | Leu       | 1              | CCA       | Pro       | 0          | CAA       | Gln       | 2          | CGA       | Arg   | 0    |
| CTTG             | Leu       | 29             | CCG       | Pro       | 12         | CAG       | Gln       | 9          | CGG       | Arg   | 3    |
| ATT              | Ile       | 1              | ACT       | Thr       | 0          | AAT       | Asn       | 1          | AGT       | Ser   | 1    |
| ATC              | Ile       | 15             | ACC       | Thr       | 12         | AAC       | Asn       | 9          | AGC       | Ser   | 2    |
| ATG              | Ile       | 0              | ACA       | Thr       | 1          | AAA       | Lys       | 3          | AGA       | Arg   | 0    |
| GTT              | Val       | 6              | ACG       | Thr       | 1          | AAG       | Lys       | 7          | AGG       | Arg   | 0    |
| GTG              | Val       | 18             | GCC       | Ala       | 34         | GAC       | Asp       | 24         | GGC       | Gly   | 28   |
| GTA              | Val       | 3              | GCA       | Ala       | 0          | GAA       | Glu       | 7          | GGA       | Gyl   | 0    |
| GTG              | Val       | 16             | GCG       | Ala       | 14         | GAG       | Glu       | 25         | GGG       | Gyl   | 5    |

there are 395 comparable positions, out of which 78 are occupied by identical residues and 35 by conservative replacements (as defined in the legend to fig. 3). Sequence similarities are scattered along the entire length of the polypeptides, suggesting a common evolutionary origin for these three functionally equivalent enzymes.

Examination of pairs of protein sequences revealed a closer relationship between *P. aeruginosa* and *C. glutamicum* DAP decarboxylases (139 identical and 57 similar residues) than between *P. aeruginosa* and *E. coli* enzymes (129 identical and 45 similar residues), the more distantly related sequences being those of *E. coli* and *C. glutamicum* (119 identical and 46 similar residues). This observation was unexpected, as both *E. coli* and *P. aeruginosa* are Gram-negative organisms whereas *C. glutamicum* is a Gram-positive organism. Such a discrepancy with the phylogenetic relationship of these three bacterial species could be due to (1) a horizontal gene transfer, (2) a faster rate of fixation of mutations in *E. coli* generally than in the two other organisms, or (3) the *E. coli* gene being paralogous to the other two. The second proposal is consistent with the observation of a closer relationship between the tryptophan synthase sequences of *P. aeruginosa* (Hadero and Crawford 1986) and those of the Gram-positive organism *B. subtilis* (Henner et al. 1984) than between those of *P. aeruginosa* and those of *E. coli* (Crawford et al. 1980; see the alignment presented in Hadero and Crawford 1986).

DAP decarboxylase, like most of the amino acid decarboxylases, is a pyridoxal phosphate-dependent enzyme. Pyridoxal phosphate is covalently bound to the enzyme by the \( \eta \)-NH2 moiety of a lysine residue. Because of the overall similarity detected between the sequences of *E. coli*, *P. aeruginosa*, and *C. glutamicum* DAP decarboxylase and mouse ornithine decarboxylase.

**FIG. 3.**—Comparison of bacterial DAP decarboxylases and mouse ornithine decarboxylase. The entire sequences of DAP decarboxylases from *Escherichia coli* (E.C; Stragier et al. 1983), *Corynebacterium glutamicum* (C.G; Ych et al. 1986b) and *Pseudomonas aeruginosa* (P.A; present paper), as well as most of the sequence of mouse ornithine decarboxylase (M.M; Kahana and Nathans 1985), have been aligned by introducing gaps indicated by dashes. Numbers in parentheses indicate position of the relevant residue in the original sequences. The equals signs (=) and the plus signs (+) above the alignment indicate positions where all three of the DAP decarboxylase residues are identical and similar, respectively. Accepted alternatives are D-E, R-K, T-S, G-A, F-Y, and L-V. The asterisks (*) and the plus signs (+) above the sequence of ornithine decarboxylate indicate identical and similar residues, respectively, shared by *P. aeruginosa* DAP decarboxylase and mouse ornithine decarboxylase.
ylases, it seems unlikely that different domains would be involved in the binding of pyridoxal phosphate to these three enzymes. Amino acid sequence comparison points to three lysine residues that are conserved among the three sequences (fig. 3) and that thus appear as potential candidates for the binding of the cofactor. A mutation reducing pyridoxal phosphate apparent affinity to *E. coli* DAP decarboxylase has been localized between positions 249 and 362, suggesting a possible involvement of this region in cofactor binding (Stragier et al. 1983). This region of the *E. coli* enzyme contains two lysine residues, one of which is also present at the same place in the sequences of *P. aeruginosa* and *C. glutamicum* DAP decarboxylases. However, this lysine residue is not located in a very conserved region, which it should if it is the active site.

The sequence of the peptide containing the lysine residue bound to pyridoxal phosphate has been determined for several other amino acid decarboxylases (see Tanase et al. 1979), and, in each case, the lysine residue bound to the cofactor was immediately preceded by a histidine residue. Conservation of this histidine residue in the vicinity of the active lysine residue suggested that it could represent an important feature of the pyridoxal phosphate binding site in amino acid decarboxylases. None of the lysine residues conserved between the DAP decarboxylases is preceded by a histidine residue. A His-Lys dipeptide is found only in the sequence of *P. aeruginosa* (fig. 3, position 297), where it is located in a nonconserved region. Although these data do not allow us to determine the localization of the pyridoxal phosphate binding site in DAP decarboxylases, this site is definitely different from the cofactor binding site on other amino acid decarboxylases.

**Comparison of DAP Decarboxylase with Ornithine Decarboxylase**

The amino acid sequence of *P. aeruginosa* DAP decarboxylase was compared with the sequences of all the enzymes present in the National Biomedical Research Foundation protein library by using the computer program FAST (Lipman and Pearson 1985). In addition to the sequence of *E. coli* DAP decarboxylase (similarity value of 126), the sequence of mouse (*Mus musculus*) ornithine decarboxylase (E.C.4.1.1.17; Kahana and Nathans 1985) was resolved from the bulk of other sequences, with a similarity value of 75. To ascertain the statistical significance of the similarity detected between *P. aeruginosa* DAP decarboxylase and mouse ornithine decarboxylase, we compared the two sequences by using the RDF program (Lipman and Pearson 1985). Comparison of DAP decarboxylase with ornithine decarboxylase and with 100 randomly permuted versions of the latter gave, for the comparison of the two initial sequences, a similarity value corresponding to 7.7 SDs above the mean of random sequences. After optimization (Lipman and Pearson 1985), the similarity value rose to 31 SDs above the mean of random sequences, indicating a highly significant degree of relatedness.

The alignment of mouse ornithine decarboxylase (Kahana and Nathans 1985) with the three DAP decarboxylases was then further improved by eye and is shown in figure 3. The simultaneous comparison of the four sequences indicates that 35 positions are occupied by identical residues and that 34 positions are occupied by chemically similar residues. The similarity detected between ornithine and DAP decarboxylases encompasses the entire length of the polypeptides, except for a C-terminal portion of approximately 30 residues that is present only in ornithine decarboxylase. Therefore we conclude that mouse ornithine decarboxylase and DAP decarboxylase are homologous.
Ornithine decarboxylase is involved in the biosynthesis of polyamines and catalyzes a reaction very similar to that of arginine or lysine decarboxylases. There are two different types of ornithine decarboxylases, as well as arginine decarboxylases, in *E. coli*: the so-called biosynthetic (which are expressed constitutively) and degradative (inducible) enzymes (see reviews in Tabor and Tabor 1985; Glansdorff 1987). In addition to the compelling parallels that exist between the properties of the two classes of ornithine and arginine degrading enzymes, there are numerous biochemical similarities between biosynthetic and degradative arginine decarboxylases (Wu and Morris 1973) and between biosynthetic and degradative ornithine decarboxylases (Applebaum et al. 1977). The properties of lysine decarboxylase are also very similar to those of the degradative arginine decarboxylase, including the sequence similarity of the phosphopyridoxal-lysine-containing peptides of these two enzymes (Sabo and Fischer 1974). Elucidation of the sequence of the pyridoxal-phosphate binding site in the *E. coli* degradative ornithine decarboxylase and its comparison with those of arginine and lysine decarboxylases strengthened the hypothesis of a common origin for these five enzymes (Applebaum et al. 1977). The only known complete sequence of a basic amino acid decarboxylase is that of the enterobacterium *Hafnia alvei* lysine decarboxylase (E.C.4.1.1.18; Fecker et al. 1986). Because of the considerable sequence similarity that exists between the *H. alvei* and *E. coli* lysine decarboxylases around the pyridoxal-phosphate bound lysine residues, the *H. alvei* sequence is most certainly very closely related to that of the *E. coli* lysine decarboxylase.

As bacterial ornithine, arginine, and lysine decarboxylases have been proposed to be evolutionarily related (Applebaum et al. 1977) and as we have shown that mouse ornithine decarboxylase is homologous to bacterial DAP decarboxylases, we have compared the sequences of these enzymes to that of *H. alvei* lysine decarboxylase. These sequence comparisons failed to detect any similarity between lysine decarboxylase and either DAP decarboxylases or mouse ornithine decarboxylase. It is also noteworthy that the subunits of bacterial ornithine, lysine, and arginine decarboxylases have a molecular weight of ~80,000, whereas mouse ornithine decarboxylase, as well as bacterial DAP decarboxylases, have a molecular weight of ~50,000. These data suggest that the mouse and *E. coli* ornithine decarboxylases have little, if any, sequence similarity, even though these enzymes are functionally equivalent. In this respect, it is noteworthy that a His-Lys dipeptide is not present in the sequence of mouse ornithine decarboxylase. These observations, i.e., the similarity between bacterial DAP decarboxylases and mouse ornithine decarboxylase as well as the probable lack of similarity between these enzymes and the *E. coli* ornithine decarboxylases, open many questions concerning the origin of mouse ornithine decarboxylase, especially since neither DAP nor lysine is synthesized in mammals. The similarity (90 identical and 66 similar residues) shared by *P. aeruginosa* DAP decarboxylase and mouse ornithine decarboxylase is only slightly less than the similarity detected between the functionally equivalent DAP decarboxylases from *E. coli* and those from *C. glutamicum* (which share 119 identical and 46 similar residues). These data suggest to us that the nuclear gene encoding the cytoplasmic ornithine decarboxylase in mouse could have been recruited from bacteria (via mitochondria?), where it was specifying a DAP decarboxylase.

**Acknowledgments**

We gratefully acknowledge F. Borne for technical assistance and A. Pugsley for careful reading of the manuscript. The nucleotide sequence of the *lysA* gene of *Pseu-
domonas aeruginosa was previously reported at the Fifth International Meeting on Genetics of Industrial Microorganisms, held in Split, Yugoslavia on September 14–20, 1986. This work was supported by grant LP 7221, U.A.1129 from the CNRS, by the Fondation pour la Recherche Médicale Française, and by the University of Aix-Marseille II.

LITERATURE CITED


**Walter M. Fitch, reviewing editor**

Received February 3, 1988; revision received April 28, 1988