Volume 258, number 2, 335-338 FEB 07906 December 1989

Expression in *Escherichia coli* of catalytically active phenylalanine ammonia-lyase from parsley

Wolfgang Schulz, Hans-Georg Eiben and Klaus Hahlbrock

Max-Planck-Institut fur Züchtungsforschung, Abteilung Biochemie, D-5000 Koln 30, FRG

Received 4 September 1989, revised version received 10 October 1989

In parsley (*Petroselinum crispum*), phenylalanine ammonia-lyase (PAL) is encoded by 4 structurally similar genes The nucleotide sequence of a near full-length cDNA and the deduced amino acid sequence of PAL-4 are presented and compared with the corresponding sequences of PAL-1, a previously described representative of the gene family Transformation of *Escherichia coli* cells with PAL-1 or PAL-4 cDNA yielded catalytically active PAL, suggesting that the catalytic center of the enzyme is formed spontaneously rather than by a plant-specific mechanism

Phenylpropanoid metabolism, Homotetrameric enzyme, Phenylalanine ammonia-lyase cDNA, Nucleotide sequence, Amino acid sequence, (*Petroselinum crispum*)

1. INTRODUCTION

Phenylalanine ammonia-lyase (PAL) is one of the most extensively studied enzymes in higher plants [1,2]. It catalyzes the conversion of L-phenylalanine to *trans*-cinnamic acid, the first committed step in general phenylpropanoid metabolism which supplies the phenylpropanoid precursors for numerous plant-specific branch pathways [2]. The key reaction is the elimination of ammonia through the catalytic action of a dehydroalanine residue in the active center [1] (fig.1). However, despite nearly 30 years of extensive work on the catalytic and regulatory properties of PAL, starting with its first description by Koukol and Conn [3], the mechanism of formation of this unusual amino acid derivative has so far remained unknown. Dehydroalanine could be formed, probably from a serine or cysteine residue in the nascent or finished polypeptide chain [1], either spontaneously as a result of inherent features of the PAL protein or by a mechanism acting upon the enzyme.

To address this question, we have expressed PAL cDNA from parsley in *E. coli*, an organism not normally producing PAL and therefore most likely not possessing a specific mechanism to convert the putative, inactive precursor to catalytically active PAL. This approach became feasible with the recent structural elucidation of the PAL gene family in parsley [4] and the possibility of using PAL cDNA in transformation studies. PAL is encoded in parsley by a small family of 4 genes, 3 of which (PAL-1, -2 and -3) have been analyzed to various degrees [4]. Here, we first report some structural features of the fourth member of this gene family, PAL-4, and then present evidence for the formation of catalytically active, homotetrameric PAL in *E. coli*.

Correspondence address: K Hahlbrock, Max-Planck-Institut fur Zuchtungsforschung, D-5000 Koln 30, FRG

2. MATERIALS AND METHODS

2.1 *PAL-1 and PAL-4 cDNAs*

A cDNA library, derived from poly(A)$^+$ RNA from elicitor-treated parsley (*Petroselinum crispum* L.) cells [5], was screened with a genomic *Hin*cII fragment containing about 1200 bp around the transcription start site of PAL-1 [4]. Fifteen positive clones were obtained from 250000 recombinants, two of which were selected and further analyzed. One comprised the entire PAL coding region, differed in sequence from PAL-1, -2 and -3 [4], and was designated PAL-4 The other consisted of 780 bp which were identical in sequence to parts of the PAL-1 exons. Its 3'-end matched the internal *Eco*RI site [4], indicating an incomplete protection of this site during cDNA synthesis This cDNA was ligated at the 3'-end with the missing portion obtained from another PAL-1 cDNA [4] which was incomplete at the 5'-end but extended to the poly(A) tail at the 3'-end. The resulting cDNA comprised the entire coding sequence and was identical in sequence with the corresponding regions of the PAL-1 gene [4]

2 2 *Expression in* E coli

The two cDNAs, PAL-1 and PAL-4, were rendered suitable for expression by inserting the synthetic oligonucleotide 5'-TCTAGAAAT-

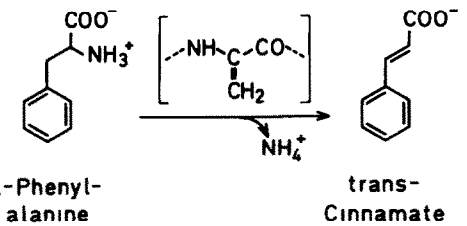

Fig 1 Reaction catalyzed by phenylalanine ammonia-lyase

Published by Elsevier Science Publishers B V. (Biomedical Division)
00145793/89/$3 50

AATTTTGTTTAACTTTAAGAAGGAGATA-3′ into the unique *Sna*B1 or *Hpa*I restriction site, respectively, near the translation start sites After recloning in the expression vector pET-3b [6] and selection in *E coli* HMS 174, *E coli* BL21(DE3) cells were transformed with the recombinant plasmid for expression studies [7]

2 3. *Gel electrophoresis and immunoblotting*
Both methods were carried out as described previously [8]

2 4 *Nucleotide sequence analysis*
The chain-termination method [9,10] was used to determine nu-

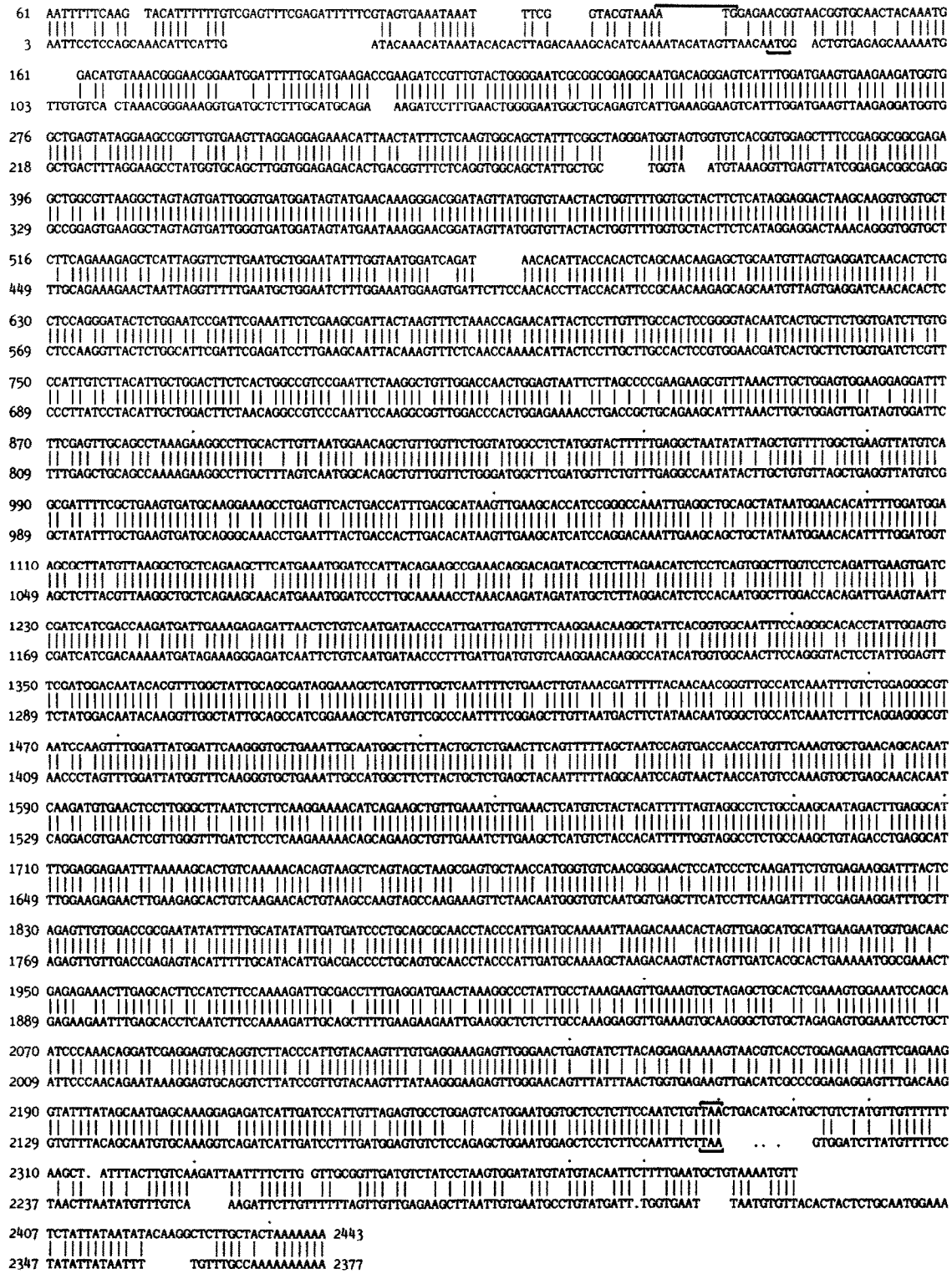

Fig 2 Comparison of nucleotide sequences of the PAL-1 (upper line) and PAL-4 cDNAs used for expression studies in *E coli* Translation start and stop codons are indicated by brackets

cleotide sequences in the Bluescribe Vector (Vector Cloning Systems, San Diego, USA).

2.5. *Enzyme assay*

PAL activity was determined spectrophotometrically [11] or by HPLC analysis of the reaction product, cinnamate. For HPLC (HP OD-552, 5 μm; reversed phase; linear gradient from 5% solvent B in A to 100% B, using 5% aqueous acetic acid as solvent A and 20% aqueous acetic acid, containing 25% acetonitrile, as solvent B), a 20-μl portion of the assay mixture was acidified with 1.5 μl 85% aqueous phosphoric acid, cinnamic acid was extracted with 3 × 200 μl ethylacetate, the solvent evaporated in a Speedvac, and the residue dissolved in 100 μl methanol.

2.6. *Sizing column*

Fractionation on Superose-6 (Pharmacia, Freiburg, FRG) was performed in 50 mM Tris-HCl, pH 8.5, 0.1 M KCl at a flow rate of about 0.5 ml/min.

3. RESULTS AND DISCUSSION

Two PAL cDNAs with full coding length, representing the PAL-1 and PAL-4 genes [4], were used for expression studies in *E. coli*. The PAL-1 cDNA was completely identical in nucleotide sequence with the corresponding portions of the previously described PAL-1 gene [4]. This sequence is compared in fig.2 with that of the new PAL-4 cDNA. The overall similarity is 79% at the nucleotide and 95% at the deduced amino acid levels. However, the areas with sequence deviations are unevenly distributed. Relatively high degrees of mismatching occur near the 5′- and 3′-ends, whereas the inner parts of the coding sequences are highly conserved.

So far, nucleotide and amino acid sequences of PAL-2 and PAL-3 have been established only for the 3′ halves of incomplete cDNAs [4]. In this region, the degree of similarity is about the same for any combination of two of the 4 PAL genes. The deduced PAL-4 protein has an $M_r = 77\,339$ and comprises 712 amino acids, as compared with $M_r = 77\,722$ and 716 amino acids for PAL-1 [4].

The two available, near full-length PAL-1 and PAL-4 cDNAs were inserted in the vector pET-3b and expressed in *E. coli*. Since the rate of expression was higher for PAL-1 than PAL-4, the following presentation of results focusses mainly on the production and some properties of the PAL-1 protein. An expression system was chosen which was expected to yield high levels of product following induction with IPTG [7].

Fig.3A and C shows, on denaturing SDS-polyacrylamide gels, the patterns of total Coomassie blue stainable proteins isolated from *E. coli* cells that had been transformed either with the vector alone or with vector containing the PAL-1 cDNA (*E. coli*$^{PAL-1}$). In the latter case, a single new protein of $M_r \sim 77\,000$ appeared which strongly reacted with a rabbit antiserum raised against parsley PAL (fig.3B and D). This protein was present in similar amounts in both IPTG-treated and untreated cells. Its cellular concentration did not change with time from 30 to 150 min after addition of

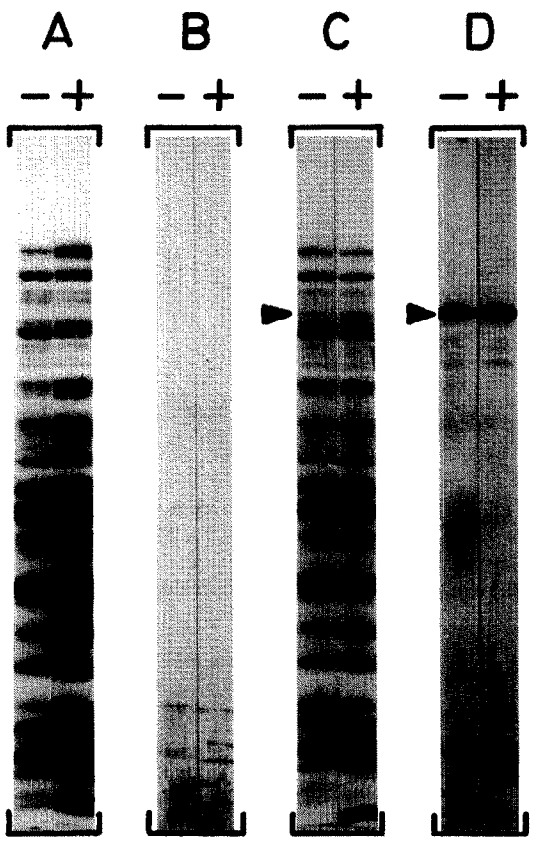

Fig.3. Analysis of protein extracts from untreated (−) and 2.5-h IPTG-treated (+) *E. coli* cells. The cells had been transformed either with the unmodified vector pET-3b (A, B) or with vector containing PAL-1 cDNA (C, D). Protein (approx. 40 μg/lane) was identified by staining with Coomassie blue (A, C) or immunoblotting with an antiserum specific for parsley PAL (B, D). Arrowhead = PAL.

IPTG and was low if compared with that of a different plant protein [12] expressed in the same system in a parallel experiment. The putative PAL protein from *E. coli*$^{PAL-1}$ cells had the same electrophoretic mobility on an SDS-polyacylamide gel as authentic PAL from parsley cells, as tested separately by immunoblotting.

The occurrence of a considerable amount of PAL protein in untreated *E. coli*$^{PAL-1}$ cells was surprising. One possible reason could be a high stability of the protein in the bacterial cells. The protein may thus accumulate to a certain extent, despite its low rate of uninduced synthesis, as a consequence of its slow degradation. The apparent lack of induction by IPTG is probably not due to a defect in the expression system, because we observed an immediate cessation of growth after the application of IPTG. The following results rather suggest that the expressed protein is catalytically active and may prevent its own production by depletion of the intracellular phenylalanine pool. In addition, the product, cinnamate, may be toxic for the cells.

Crude protein extracts from *E. coli*$^{PAL-1}$, but not from untransformed *E. coli* cells, showed PAL activity. This was demonstrated by both the standard optical test

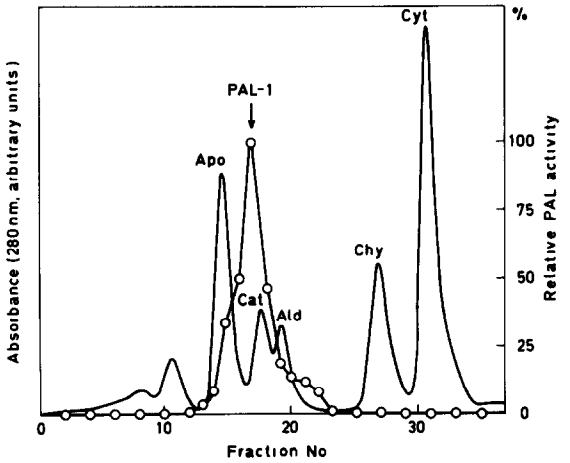

Fig.4. Elution profile of PAL-1, produced in *E coli*, from a Superose-6 column Reference proteins for M_r estimation were apoferritin (Apo; M_r = 440 000), catalase (Cat; 240 000), aldolase (Ald, 150 000), chymotrypsin (Chy; 25 000) and cytochrome *c* (Cyt; 12 500)

measuring the increase in absorbance at 290 nm [11] and the direct identification of cinnamate as the reaction product by HPLC. In agreement with fig.3, no appreciable difference in PAL activity was found in IPTG-treated and untreated cells.

Essentially the same results were obtained with *E. coli*$^{PAL-4}$ cells (transformed in several independent experiments with the vector containing PAL-4 cDNA), except that the amounts of PAL protein and PAL activity were considerably lower. The specific activity in crude extracts was about 900 μkat/kg for PAL-1 and 30 μkat/kg for PAL-4 from *E. coli* cells, as compared with 100 μkat/kg for PAL from elicitor-stimulated parsley cells [13]. The reason for this differential behavior of *E. coli*$^{PAL-1}$ and *E. coli*$^{PAL-4}$ is unknown.

The PAL activity from *E. coli*$^{PAL-1}$ was enriched about 50-fold from a cleared lysate by anion exchange chromatography on DEAE-cellulose and monoQ (FPLC). Similar to PAL from parsley [11,14], the purified enzyme was not very stable, and the purification procedure will have to be improved for more detailed studies of the protein. However, one purpose of this investigation was to find out whether PAL-1 forms homooligomeric aggregates and this was achieved by passing the partially purified enzyme through a sizing column of Superose-6 (fig.4). Using appropriate marker proteins for calibration, we estimated an apparent $M_r \sim 300\,000$ for the eluted enzyme. This indicates that the catalytically active PAL-1 from *E. coli*$^{PAL-1}$ is a homotetrameric protein, in agreement with the tetrameric nature of purified PAL from various plant sources, including parsley [1,11].

In summary, we have demonstrated the production of catalytically active PAL in transformed bacteria and conclude from our results: (a) that formation of the active center of PAL does not require an additional, plant-specific mechanism, but is probably a spontaneous event; and (b) that PAL subunits readily form a homotetrameric enzyme, thus excluding the possibility that multiple gene copies in plants [4,15] are required for the formation of the quarternary enzyme structure.

Acknowledgements We thank Ms E Logemann for excellent technical assistance, Dr F W Studier for a gift of the expression system, and R Boettge, Dr A Gierl, H Keller, Dr R Lois, Dr D Scheel and Dr I Somssich for valuable help during various stages of this work

REFERENCES

[1] Hanson, K R and Havir, E.A (1981) in The Biochemistry of Plants, vol 7 (Stumpf, P K and Conn, E.E eds) pp 577–625, Academic Press, New York
[2] Hahlbrock, K and Scheel, D (1989) Annu Rev Plant Physiol Mol Biol 40, 237–369
[3] Koukol, J. and Conn, E.E (1961) J Biol. Chem 236, 2692–2698
[4] Lois, R , Dietrich, A., Hahlbrock, K and Schulz, W (1989) EMBO J 8, 1641–1648
[5] Lozoya, E , Hoffmann, H., Douglas, C , Schulz, W , Scheel, D and Hahlbrock, K (1988) Eur J Biochem. 176, 661–667
[6] Rosenberg, A.H , Lade, B.N , Chui, D -S , Lin, S -W , Dunn, J.J and Studier, W F (1987) Gene 56, 125–135
[7] Studier, F W and Moffatt, B A (1986) J Mol Biol 189, 113–130.
[8] Kombrink, E., Schroder, M and Hahlbrock, K (1988) Proc Natl. Acad Sci USA 85, 782–786
[9] Sanger, F , Nicklen, S. and Coulson, A R (1977) Proc Natl Acad Sci USA 74, 5463–5467
[10] Chen, E J and Seeburg, P.H. (1985) DNA 4, 165–170
[11] Zimmermann, A. and Hahlbrock, K (1975) Arch Biochem Biophys 166, 54–62
[12] Gierl, A , Luttıcke, St. and Saedler, H (1988) EMBO J 7, 4045–4053
[13] Hauffe, K D , Hahlbrock, K and Scheel, D (1986) Z Naturforsch 41c, 228–239.
[14] Schroder, J , Betz, B and Hahlbrock, K (1976) Eur J Biochem 67, 527–541
[15] Cramer, C L , Edwards, K , Dron, M., Liang, X., Dildine, S L , Bolwell, C P , Dixon, R A , Lamb, Ch J and Schuch, W (1989) Plant Mol Biol 12, 367–383