Molecular Microbiology (1988) 2(6), 777–783

Phylogeny of metabolic pathways: *O*-acetylserine sulphydrylase A is homologous to the tryptophan synthase beta subunit

S. Levy* and A. Danchin†
Unité de Régulation de l'Expression Génétique, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris, Cedex 15, France.

Summary

The *cysK* gene of *Escherichia coli* K-12 encoding *O*-acetylserine sulphydrylase A, was cloned and its nucleotide sequence, together with that of the flanking regions, was determined. The deduced amino acid sequence of the carboxy-terminal moiety of *O*-acetylserine sulphydrylase A shows significant similarity to the amino acid sequence of tryptophan synthase beta chain from several organisms. This sequence similarity is likely to reflect the structural homologies of substrates shared by both enzymes. This may indicate that these proteins, although catalysing different reactions in different metabolic pathways, have evolved from a common ancestral gene.

Introduction

Molecular evolution of metabolic pathways is a central issue regarding molecular aspects of the origin of life. Two major working hypotheses have been proposed in this respect. As early as 1945, Horowitz realized that biosynthetic pathways could have been progressively built backwards from the final metabolite of the pathway. Following a different approach, Yčas (1974) and Jensen (1976), in view of the substrate ambiguity exhibited by contemporary enzymes, suggested as an alternative hypothesis that primitive enzymes possessed a very broad specificity permitting utilization of a wide variety of structurally related substrates. This latter hypothesis has been substantiated by the recent discovery of amino acid sequence similarity between enzymes catalysing consecutive steps in the isoleucine biosynthetic pathway (Parsot, 1986), as well as between two enzymes catalysing consecutive steps in the methionine biosynthetic pathway (Belfaiza *et al.*, 1986).

Received 7 June, 1988; revised 14 July, 1988. *Present address: Laboratoire de Génétique, Institut de Biotechnologie, Centre de Recherche de Vitry, 13 Quai Jules Guesde, 94400 Vitry, France. †For correspondence.

The cysteine biosynthetic pathway has been less thoroughly studied than the methionine pathway and so far only one nucleotide sequence of a gene from the cysteine regulon, *cysE*, has been published (Denk and Böck, 1987). The *cysK* gene cloned by Boronat *et al.* (1984) maps at min. 52 on *E. coli* genome. This gene, the sequence of which is presented here, encodes *O*-acetylserine sulphydrylase A, an enzyme which catalyses the synthesis of cysteine from *O*-acetylserine. It reveals, unexpectedly, that the *O*-acetylserine sulphydrylase A amino acid sequence is significantly similar to the amino acid sequence of the beta subunit of tryptophan synthase. The sequence similarity detected correlates well with similarities in the catalytic process of the two enzymes, despite the fact that they catalyse the last step of two different biosynthetic pathways.

Results

Nucleotide sequences of the cysK *gene*

Plasmid pDIA4535, shown in Fig. 1, carries the *lig* gene, the *ptsH* gene and part of the *ptsI* gene. The *cysK* gene has been located by Boronat *et al.* (1984) upstream from *ptsH*, with 90% of the gene lying upstream from the *Hin*dIII restriction site (H_1), and has been shown to be transcribed in the same direction as *ptsH*.

The nucleotide sequence of the 1.9 kb *Cla*I-*Bam*HI (B_1) DNA fragment (Fig. 1) located immediately upstream from the *ptsH* gene is displayed in Fig. 2. An open reading frame, ORF1, located between nucleotides 1 and 588, having a 'standard' codon usage, could correspond to the 3′ end of an unknown truncated gene. A second ORF, 977 nucleotides long, was found 178 nucleotides downstream from ORF1. An ATG codon, located 9 bp downstream from the beginning of this second ORF, is preceded by a sequence that fits well to a consensus ribosome binding site (Shine and Dalgarno, 1974). This ORF could encode a protein of 323 amino acids with a calculated molecular weight of 34450 Daltons. This is in good agreement with the value of 34000 Daltons estimated for purified *O*-acetylserine sulphydrylase A of *Salmonella typhimurium* (Becker *et al.*, 1969). In addition, nucleotide sequencing analysis indicates that truncating the DNA at the *Hin*dIII site (H_1) would lead to a shorter protein of about 30600

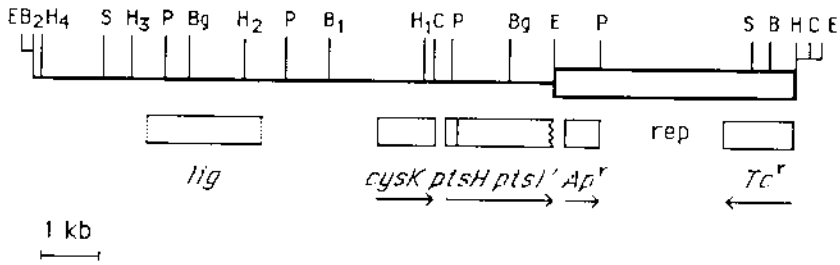


Fig. 1. Restriction map of plasmid pDIA4535. The box represents the vector pBR322 and the line the bacterial DNA insert. Genes are represented by boxes placed under the corresponding DNA segments. The direction of transcription (when known) is indicated by an arrow underneath the box. The symbol 'rep' specifies the plasmid replication region. The *lig* gene was localized by analysis of the ability of different subclones to complement the thermosensitive ligase defect of strain GR501. Uncertainty about the exact limits of the coding region is indicated by the dotted part of the box. The notch at the end of the box and the single quote next to *ptsI* indicate that the gene is truncated. Restriction enzymes sites are as follows: B, *Bam*HI; Bg, *Bgl*II; C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; S, *Sal*I.

Daltons. Such a truncated protein, fused to 9 amino acids derived from the pBR322 DNA sequence, has been observed on an electrophoresis gel by Boronat *et al.* (1984); it had lost the *O*-acetylserine sulphydrylase activity. A further confirmation that this sequence is indeed the *cysK* gene stems from the comparison between the amino acid composition of the protein determined by Boronat *et al.* (1984) and that deduced from the nucleotide sequence (Table 1). It should be noted, however, that in the determination made by these authors the number of phenylalanine residues (14.4) seems significantly higher than the number derived from the sequence (8 residues), but the latter figure fits with the value obtained by Becker *et al.* (1969) for the *S. typhimurium* enzyme (7.2 residues).

Immediately upstream from the gene, a palindromic sequence (nucleotides 630–651) is likely to mark the end of a preceding transcript. Another palindromic sequence, immediately downstream (nucleotides 1770–1797), might act as a terminator. These features suggest that *cysK* might be expressed as a monocistronic operon.

Table 1. Amino acid composition of *O*-acetylserine sulphydrylase A.

Amino acid	Protein *E. coli*[a]	Protein *S. typhimurium*[b]	Predicted *E. coli*[c]
Ala	30.1	30.8	32
Arg	12.7	13.4	13
Asn + Asp	24.7	25.5	25
Cys	0.0	0.7	1
Gln + Glu	37.1	37.5	37
Gly	35.4	32.6	32
His	3.1	3.0	3
Ile	24.5	24.1	27
Leu	31.3	31.6	32
Lys	22.9	25.5	23
Met	4.4	4.9	6
Phe	14.4	7.2	8
Pro	15.2	15.8	15
Ser	22.1	17.0	19
Val	18.8	19.2	19
Thr	21.0	24.0	22
Trp	N.D.	5.0	2
Tyr	5.6	6.4	5

N.D.: not determined. The composition is given as mol per subunit.
a. Boronat *et al.* (1984).
b. Becker *et al.* (1969).
c. This work.

Similarity between the cysK *gene product and the beta subunit of tryptophan synthase*

The amino acid sequence deduced from the *cysK* nucleotide sequence was compared with 6467 protein sequences present in the PSeqIP data bank using the FASTP program (Lipman and Pearson, 1985). A very significant similarity was detected between the second half of the *E. coli O*-acetylserine sulphydrylase A and the second half of the beta subunit of tryptophan synthase from *E. coli*, *S. typhimurium* and *Bacillus subtilis*. This similarity encompasses 157 amino acid residues. It is remarkable that, in this region, the extent of similarity between heterologous enzymes from homologous organisms is similar to the one observed between homologous enzymes from heterologous organisms.

Figure 3 displays the alignment of homologous segments of *E. coli O*-acetylserine sulphydrylase A, *E. coli* tryptophan synthase beta chain (Crawford *et al.*, 1980) and *B. subtilis* tryptophan synthase beta-chain sequences (Henner *et al.*, 1984). In the alignment with the *E. coli* enzyme, 6 gaps have been introduced. Out of the remaining 149 positions compared, 37 are occupied by identical residues (25%) and 39 by conservative replacements (as defined in the legend of Fig. 3). This gives a total score of 51% similarity. In the alignment with the *B. subtilis* enzyme, 7 gaps have been introduced. Out of the remaining 149 positions compared, 39 are occupied by identical residues (26%) and 31 by conservative replacements, resulting in a score of 47% similarity. The most-conserved regions in this alignment are also highly conserved between the beta chains of tryptophan synthase in *E. coli*, *Pseudomonas aeruginosa*, *B. subtilis* and *Saccharomyces cerevisiae* (Hadero and Crawford, 1986).

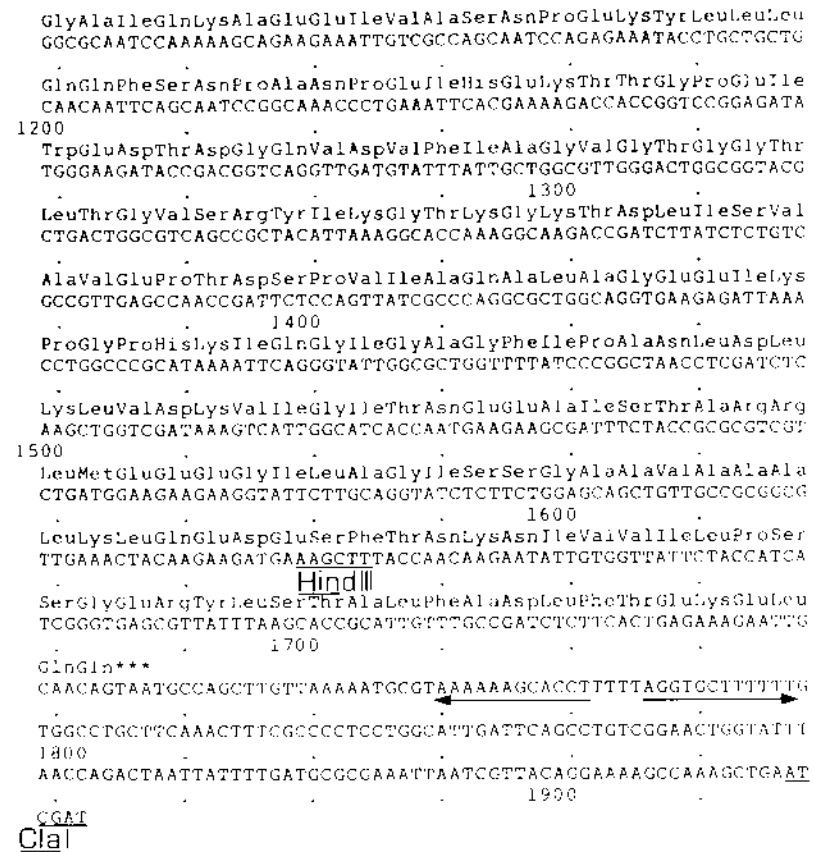


Fig. 2. Nucleotide sequence of the 1917 bp *Bam*HI-*Cla*I of plasmid pDIA4535, including the *cysK* gene and its flanking regions. The deduced amino acid sequence of *O*-acetylserine sulphydrylase A is presented. The AUG start codon proposed for the *cysK* gene is boxed and the ribosome binding site is indicated as SD (Shine and Dalgarno). Two palindromic sequences are underlined with pairs of arrows. These sequences have been submitted to the EMBL/GenBank Data Libraries under the accession number X12615.

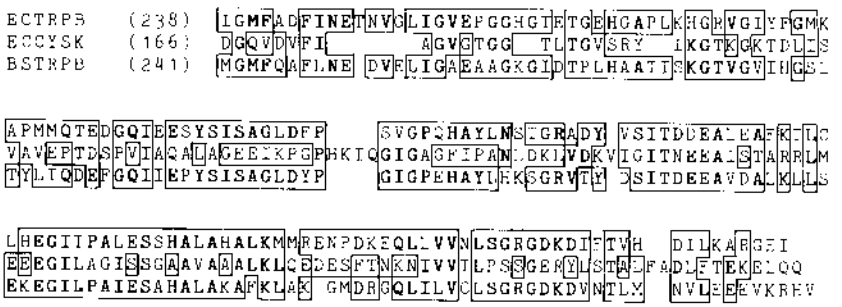


Fig. 3. Comparison of the carboxy-terminal amino acid sequences of the beta chain of tryptophan synthase from *E. coli* (ECTRPB), *B. subtilis* (BSTRPB) and *E. coli* *O*-acetylserine sulphydrylase A (ECCYSK). Single-letter abbreviations are used and gaps have been introduced to permit optimal pairwise alignment. Positions having identical residues are emphasized and boxed; those having conservative residues are boxed. Conservative amino acid substitutions were determined according to the following groupings: I-L-V-M; R-K-H; A-S-T-P-G; Y-F-W; N-Q-D-E. Numbers between brackets indicate the location of each first amino acid of the fragment in the complete sequence.

The pyridoxal phosphate binding site

O-acetylserine sulphydrylase A is a pyridoxal phosphate (PLP) enzyme. It is known that PLP is covalently bound to a lysine residue. A window of 28 amino acids was centred on every lysine residue present in the *cysK* sequence and compared with known PLP binding sites (Tanase *et al.*, 1979; Parsot *et al.*, 1987). As a result, lysine 137 can be proposed as the PLP binding site present in *O*-acetylserine sulphydrylase A. As shown in Fig. 4, the lysine 137 region exhibits significant similarity with the cognate PLP binding site of *E. coli metB* and *metC* gene products. Moreover, a weaker similarity can be detected between this same sequence and the PLP binding sites from enzymes involved in serine and threonine dehydration, as well as tryptophan synthase (Parsot *et al.*, 1987).

Discussion

Understanding the molecular origin of metabolic pathways is a major challenge for the near future. Because of the large number of known DNA sequences, the deduced primary structure of many proteins involved in intermediary metabolism has become available. As a first step, it seems possible to evaluate the hypotheses of Horowitz (1945), as well as the hypotheses of Yčas (1974) and Jensen (1976), who proposed that poorly specific primitive enzymes could have evolved towards greater specificity. In *E. coli*, for instance, biosynthesis of homocysteine from homoserine (catalysed consecutively by the *metB* and *metC* gene products) is strikingly similar to the biosynthesis of cysteine from *O*-acetylserine (catalysed by the *cysK* gene product). In *S. cerevisiae*, a single enzyme (coded by the *MET25* gene) catalyses both reactions. Parsot *et al.* (1987) predicted, accordingly, that the *cysK* gene product would have a common descent with *metB*, *metC* and *MET25* gene products. Analysis of the deduced amino acid sequence of the *cysK* gene shows, however, that this is not the case. Rather, *O*-acetylserine sulphydrylase A is similar to the beta subunit of tryptophan synthase. The similarity applies only to the second half of the sequence, thus suggesting modular enzyme evolution. The observation presented here suggests that tryptophan and cysteine biosynthetic pathways are more closely related than was previously thought. It seems remarkable that these two amino acids, which share sister codons, are both highly sensitive towards oxygen. This suggests that their metabolism might have originated at a time when oxygen was absent from the earth's atmosphere. Thus one might speculate that both amino acids were constituents of very early proteins.

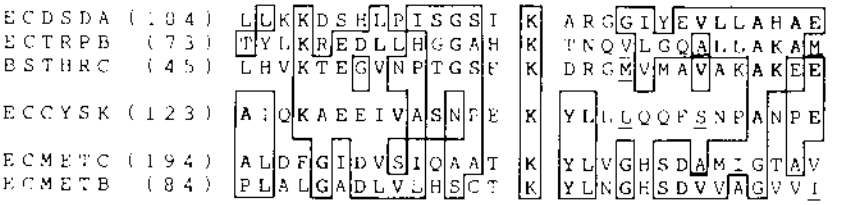
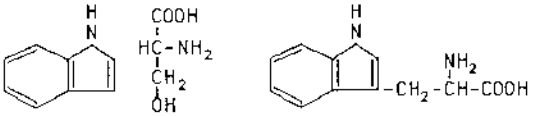
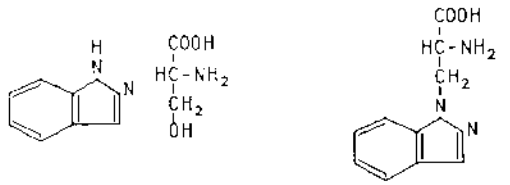


Fig. 4. Comparison of the PLP binding site region from D-serine dehydratase from *E. coli* (ECDSDA), tryptophan synthase from *E. coli* (beta chain) (ECTRPB), threonine synthase from *B. subtilis* (BSTHRC), cystathionine-β-lyase from *E. coli* (ECMETC) and cystathionine-γ-synthase from *E. coli* (ECMETB), with the lysine 137 region from *O*-acetylserine sulphydrylase A of *E. coli* (ECCYSK). Corresponding references can be found in Parsot *et al.* (1986). Same comments as in Fig. 3. When two classes of three amino acids appear simultaneously among the six sequences, one class is boxed while the other one is underlined.

Reactions catalysed by E.coli tryptophan synthase and its subunits.

Reaction	catalysed by
1. Indole-3-glycerolphosphate + L-serine ⟶ L-tryptophan + D-glyceraldehyde 3-phosphate + H_2O	$\alpha_2\beta_2$
2. Indole + L-serine ⟶ L-tryptophan + H_2O 	β_2; $\alpha_2\beta_2$
3 β-Mercaptoethanol + L-serine ⟶ S-hydroxyethyl-L-cysteine + H_2O	β_2; $\alpha_2\beta_2$
4. CH_3SH + L-serine ⟶ SCH_3-L-cysteine + H_2O	β_2
5. H_2S + L-serine ⟶ L-cysteine + H_2O	$\alpha_2\beta_2$
6. Indazole + L-serine ⟶ β-1 indazole-L-alanine 	$\alpha_2\beta_2$

Reactions catalysed by E.coli O-acetylserine sulfhydrylase A.

a. H_2S + O-acetyl-L-serine ⟶ L-cysteine + acetate + H_2O

b. Triazole + O-acetyl-L-serine ⟶ Triazole-L-alanine + acetate + H_2O

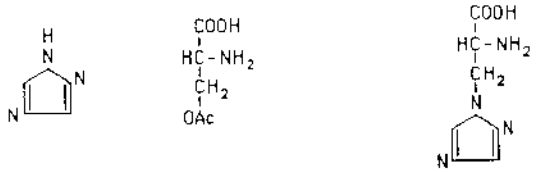


Fig. 5. Reactions catalysed by *E. coli* tryptophan synthase and *O*-acetylserine sulphydrylase.

Analysis of the biochemical properties of the two enzymes indicates that their carboxy-terminal region could bind common substrates. *O*-acetylserine sulphydrylase A is a homodimer catalysing the biosynthesis of cysteine from *O*-acetylserine and sulphide (reaction a, Fig. 5). The activity of this enzyme depends on a PLP cofactor. Tryptophan synthase is a heterodimer catalysing *in vivo* reaction 1 of Fig. 5. Moreover, different combinations of the α- and β subunits can catalyse, *in vitro*, a series of other PLP-dependent reactions; some of these are illustrated in Fig. 5. Thus the dimer β_2 can use substrates other than indole, such as β-mercaptoethanol (reaction 3) (Miles *et al.*, 1968) and CH_3SH (reaction 4) (Kumegaï and Miles, 1971), *in vitro*. In this respect it must be pointed out that Goldberg and Baldwin (1967) have shown that the α_2-β_2 complex was able to produce cysteine from serine and sulphide (reaction 5). A somewhat symmetrical reaction, i.e. synthesis of tryptophan from indole and *O*-acetylserine by *O*-acetylserine sulphydrylase A, has not yet been described. This enzyme, however, can use triazole (reaction b) (Kredich *et al.*, 1975) as a substrate. Triazole is a compound closely related to indazole which can be used by tryptophan synthase (reaction 6) (Tanaka *et al.*, 1986). On the basis of these comparisons, it can be proposed that *O*-acetylserine sulphydrylase A and the beta subunit of tryptophan

synthase could bind both indole and sulphide. The residues in the tryptophan synthase beta subunit which have been identified as interacting with serine belong to the amino-terminal part of the sequence (Miles, 1979). No residues interacting with indole have yet been identified. Sequence similarities reported above suggest that the indole- or sulphide binding region in both enzymes could be carboxy-terminal.

Finally, another region in the *O*-acetylserine sulphydrylase A amino acid sequence exhibits a further set of similarities with enzymes involved in biosynthetic pathways. The sequence shown to represent a PLP binding site could be aligned with both of the homologous sites present in serine, threonine and tryptophan biosynthetic or degradation enzymes, as well as to the methionine PLP sites. The *cysK* gene product might therefore represent a link between these metabolic pathways. Thus it can be proposed that pyridoxal phosphate, bound at a site present in a broad range of primitive enzymes, might have evolved towards different specificities in catalysing reactions involving serine, threonine, tryptophan, cysteine and methionine.

After submission of this article, a paper dealing with the nucleotidic sequence of *E. coli cysK* gene and the identification of the promoter region was published (Byrne *et al.*, 1988). Comparing the two nucleotidic sequences, we found four differences in the region from nucleotide 1320 to 1335. This leads to the following differences in *O*-acetylserine sulphydrylase A sequence: we find amino acids Leu, Ser and Arg, respectively, in positions 182, 186 and 187 instead of Trp, Thr and Pro in their sequence. This may be accounted for by allelic variation in the *cysK* locus of *E. coli*. Furthermore, Byrne *et al.* have mapped the major transcription initiation site for *cysK* gene at nucleotide 744. They proposed a promoter region with a −10 region (TATGCT) located between nucleotides 731 and 736 and a −35 region (TTCCGC) located between nucleotides 709 and 714. This −35 region clearly deviates from the consensus region.

Experimental procedures

Bacterial strains and growth media

E. coli K12 strain GR501 (*lig-251 relA bglB7 thi spoT1*) (Dermody *et al.*, 1979) was used for cloning experiments and to select plasmids able to complement the *lig* mutation. Strain 71-18 (*Δlac-proAB thi supE* (F′ *proAB lacI lacZ ΔM15*)) (Yanish-Perron *et al.*, 1985) was used as a recipient for M13 derivatives. Strain TP2111 (F⁻, *xyl ilvA aroB lac ΔX74 argH1*) (Roy *et al.*, 1983) was used as the source of chromosomal DNA. Growth media were according to Miller (1972) and were supplemented with ampicillin (100 μg ml^{-1}) when appropriate.

DNA cloning and sequencing

DNA preparation, DNA restriction and fragment separation on agarose gels by electrophoresis, ligation and bacterial transformation were performed according to the methods described in Maniatis *et al.* (1982). Cloning of an *Eco*RI DNA fragment, isolated from strain TP2111 chromosome, carrying the *lig* gene into plasmid pBR322 was performed by complementing a strain expressing a temperature-sensitive DNA ligase (GR501). The restriction map of the resulting plasmid, pDIA4535, is presented in Fig. 2. Comparison of this map with the map of the chromosomal DNA fragment of plasmid pDIA3206, carrying the entire *pts* operon (De Reuse *et al.*, 1984), showed that plasmid pDIA4535 also carries the *ptsH* gene and part of the *ptsI* gene. Nucleotide sequences were determined using the dideoxynucleotide chain termination procedure (Sanger *et al.*, 1977) on single-stranded M13 DNA obtained after cloning of the DNA fragments into plasmid pTZ18R. Unidirectional deletions on single-stranded DNA were obtained by following the procedure of Dale *et al.* (1985). All the fragments used for sequencing overlapped and the whole sequence was determined on both strands.

Chemicals

Restriction enzymes, T4 DNA ligase and Klenow fragment of *E. coli* DNA polymerase were used as recommended by the suppliers (Boehringer Mannheim, FRG; Pharmacia, Sweden; or Amersham, UK). Agarose type II for gel electrophoresis was from Litex (Denmark), ampicillin was from Sigma and other chemicals were from Merck (Darmstadt, FRG). The alpha-[^{35}S]-dATP was from Amersham (UK). pTZ18R plasmid vector was purchased from Pharmacia (Sweden) and the cyclone deletion kit was from IBI (USA).

Acknowledgements

We thank H. De Reuse for her constant interest, and C. Zetina and M. Goldberg for helpful discussions on the nature of the protein. We are indebted to P. Marlière and C. Parsot for constructive discussions about protein evolution. We are grateful to R. Fleer for critical reading of the manuscript and for linguistic editing and to F. Doher for expert typing. During this work, S.L. was a recipient of a Rhone Poulenc fellowship. Financial support came from the Centre National de la Recherche Scientifique (UA1129) and the Institut Pasteur (13845).

References

Becker, M.A., Kredich, N.M., and Tomkins, G.M. (1969) The purification and characterization of *O*-acetylserine sulfydrylase A from *Salmonella typhimurium*. *J Biol Chem* **244:** 2418–2427.

Belfaiza, J., Parsot, C., Martel, A., Bouthier de la Tour, C., Margarita, D., Cohen, G.N., and Saint-Girons, I. (1986) Evolution in biosynthetic pathways: two enzymes catalysing consecutive steps in methionine biosynthesis originate from a common ancestor and possess a similar regulatory region. *Proc Natl Acad Sci USA* **83:** 867–871.

Byrne, C.R., Monroe, R.S., Ward, K.A., and Kredich, N.M. (1988) DNA sequences of the *cysK* regions of *Salmonella typhimurium* and *Escherichia coli* and linkage of the *cysK* regions to *ptsH*. *J Bacteriol* **170:** 3150–3157.

Boronat, A., Britton, P., Jones-Mortimer, M.C., Kornberg, H.L., Lee, L.G., Murfitt, D., and Parra, F. (1984) Location on the *Escherichia coli* genome of a gene specifying O-acetylserine (thiol)-lyase. *J Gen Microbiol* **130:** 673–685.

Crawford, I.P., Nichols, B.P., and Yanofsky, C. (1980) Nucleotide sequence of the *trpB* gene in *Escherichia coli* and *Salmonella typhimurium*. *J Mol Biol* **142:** 489–502.

Dale, R.M.K., McClure, B.A., and Houchins, J.P. (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. *Plasmid* **13:** 31–40.

Denk, D., and Böck, A. (1987) L-cysteine biosynthesis in *Escherichia coli*: nucleotide sequence and expression of the serine acetyltransferase (*cysE*) gene from the wild-type and a cysteine-excreting mutant. *J Gen Microbiol* **133:** 515–525.

Dermody, J.J., Robinson, G.T., and Sternglanz, R. (1979) Conditional-lethal deoxyribonucleic acid ligase mutant of *Escherichia coli*. *J Bacteriol* **139:** 701–704.

De Reuse, H., Huttner, E., and Danchin, A. (1984) Analysis of the *ptsH-ptsI-crr* region in *Escherichia coli* K-12: evidence for the existence of a single transcriptional unit. *Gene* **32:** 31–40.

Goldberg, M.E., and Baldwin, R.L. (1967) Interactions between the subunits of the tryptophan synthetase of *Escherichia coli*. Optical properties of an intermediate bound to the α_2 β_2 complex. *Biochemistry* **6:** 2113–2119.

Hadero, A., and Crawford, I.P. (1986) Nucleotide sequence of the genes for tryptophan synthase in *Pseudomonas aeruginosa*. *Mol Biol Evol* **3:** 191–204.

Henner, D.J., Band, L., and Shimotsa, H. (1984) Nucleotide sequence of the *Bacillus subtilis* tryptophan operon. *Gene* **34:** 169–177.

Horowitz, N.H. (1945) On the evolution of biochemical synthesis. *Proc Natl Acad Sci USA* **31:** 153–157.

Jensen, R.A. (1976) Enzyme recruitment in the evolution of new function. *Ann Rev Microbiol* **30:** 409–425.

Kredich, N.M., Foote, L.J., and Hulanicka, M.D. (1975) Studies on the mechanism of inhibition of *Salmonella typhimurium* by 1,2,4-triazole. *J Biol Chem* **250:** 7324–7331.

Kumagai, H., and Miles, E.W. (1971) The B protein of *Escherichia coli* tryptophan synthetase II. New β elimination and β replacement reactions. *Biochem Biophys Res Commun* **44:** 1271–1278.

Lipman, D.J., and Pearson, W.R. (1985) Rapid and sensitive protein similarity searches. *Science* **227:** 1435–1441.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Miles, E.W., Hatanaka, M., and Crawford, I.P. (1968) A new thiol-dependent transamination reaction catalysed by the B protein of *Escherichia coli* tryptophan synthetase. *Biochemistry* **7:** 2742–2753.

Miles, E.W. (1979) Tryptophan synthase: structure, function and subunit interaction. *Adv Enzymol Relat Areas Mol Biol* **49:** 127–186.

Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Parsot, C. (1986) Evolution of biosynthetic pathways: a common ancestor for threonine synthase, threonine dehydratase and D-serine dehydratase. *EMBO J* **5:** 3013–3019.

Parsot, C., Saint-Girons, I., and Cohen, G.N. (1987) Enzyme specialization during the evolution of amino acid biosynthetic pathways. *Micro Sci* **4:** 258–262.

Roy, A., Haziza, C., and Danchin, A. (1983) Regulation of adenylate cyclase synthesis in *Escherichia coli*: nucleotide sequence of the control region. *EMBO J* **2:** 791–797.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74:** 5463–5467.

Shine, J., and Dalgarno, L. (1974) The 3′ terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci USA* **71:** 1342–1346.

Tanaka, H., Tanizawa, K., Arai, T., Saito, K., and Soda, K. (1986) Production of a novel tryptophan analog, β-1-indazole-L-alanine with tryptophan synthase of *Escherichia coli*. *FEBS Letts* **196:** 357–360.

Tanase, S., Kojima, H., and Morino, Y. (1979) Pyridoxal 5′-phosphate binding site of pig heart alanine aminotransferase. *Biochemistry* **18:** 3002–3007.

Yčas, M. (1974) On earlier states of the biochemical system. *J Theor Biol* **44:** 145–160.

Yanish-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequence of M13mp18 and pUC19 vectors. *Gene* **33:** 103–119.