FINAL TECHNICAL REPORT

July 26, 1978

U.S. Department of Energy

Contract # EY-76-S-03-0034, P.A. 133

"Genetic Control of Enzyme Formation"

AT03-76EV70133

Dr. Stanley E. Mills

Principal Investigator

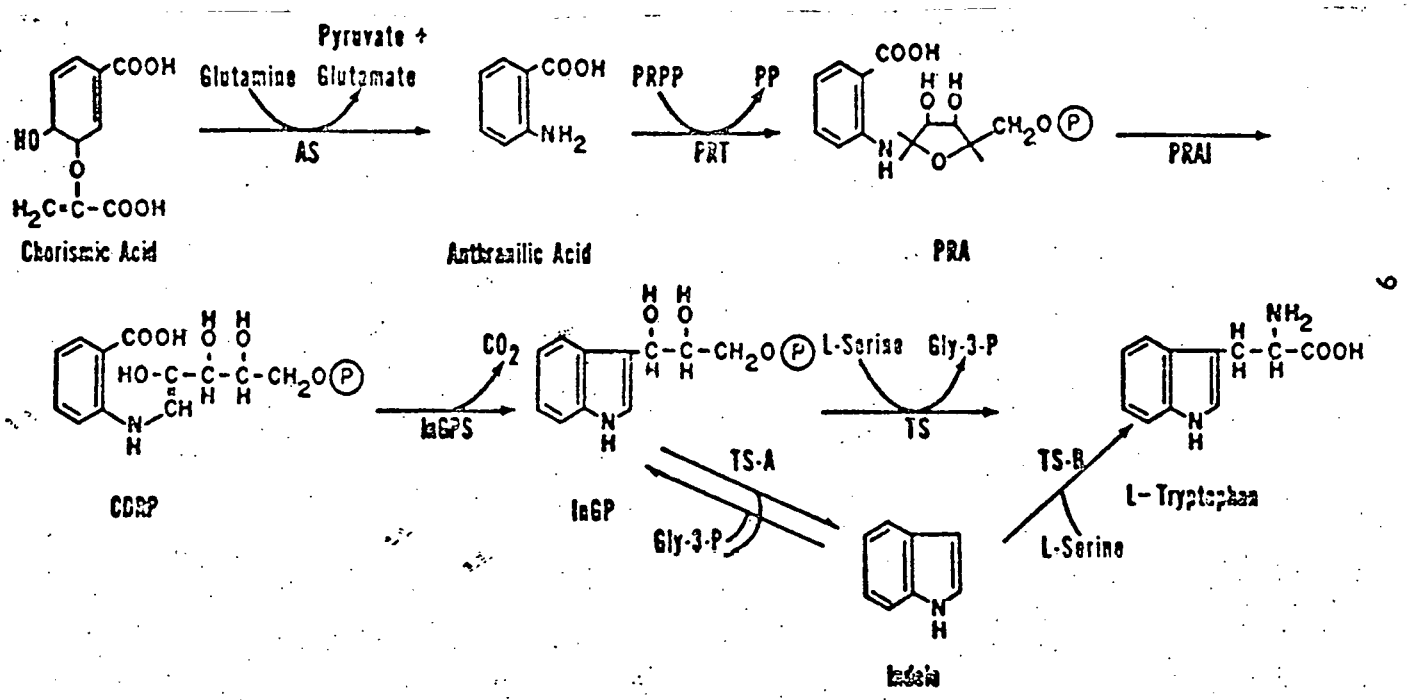
DISCLAIMER

Abstract

The final period of this grant has been productive. We completed, and published, elements of our work on tryptophan biosynthesis in Euglena gracilis and higher plants. The experimental data provide an outline of the general evolution of the pathway. We also completed the latest in our series of structural analyses of the pathway proteins by quantitative immunochemical methods; this work was done with the anthranilate synthase-phosphoribosyl transferase complex in Escherichia coli. And we have begun an examination of the evolution, in the Enterobacteriaceae, of the enzyme activities anthranilate synthase and anthranilate-5-phosphoribosyl-1-pyrophosphate phosphoribosyltransferase. This project has required the purification of several proteins. The results will be discussed in detail below.

Introduction

During the past twenty years much has been learned largely in microorganisms, about the biosynthesis of the amino acid L-tryptophan.



Pathway for L-tryptophan biosynthesis. Abbreviations for the enzymes are: AS, anthranilate synthase; PRT, anthranilate phosphoribosyl transferase; PRAI, phosphoribosylanthranilate isomerase; InGPs, indoleglycerol phosphate synthase; TS, tryptophan synthase; TS-A, tryptophan synthase α protein; TS-B, tryptophan synthase β_2 protein.

The data have contributed to many of the general principles formulated on the biology of cells and organisms. Among these are (1) elements of the relationship between structural genes and their products, including colinearity, the code, and mutation and its consequence; (2) amino acid regulation, exemplified by the diversity of structural gene assemblies of the pathway in microorganisms and the presence and mode of action of both positive and negative control elements.

The work on the pathway enzymes has concerned itself with their sequence, structure, aggregation patterns, subunit interactions and catalytic functions. By far most of the information has been obtained with tryptophan synthase and anthranilate synthase from Escherichia coli and Salmonella typhimurium, members of the Enterobacteriaceae. For example, with tryptophan synthase (formula $\alpha_2\beta_2$) the complete sequence of the α subunit is known for the molecule in Escherichia coli, Salmonella typhimurium and Enterobacter aerogenes. Partial sequences exist for both the α and β subunits for other bacteria in this family as well as for Pseudomonas putida and Bacillus subtilis. There is information on the energy of bonding of the subunits to one another studied under various conditions including substrate and cofactor (pyridoxal phosphate) concentration, amino acid replacement and subunit substitution by homologues from other organisms.

In recent years analysis of the pathway has been extended to organisms other than those traditionally studied, the bacteria and the higher fungi, yeast and Neurospora crassa. In particular, during the tenure of this grant, we have obtained information on the pathway in eucaryotic, photosynthetic organisms ranging from Euglena gracilis to higher plants including representatives of the mono- and dicotyledons.

In our work with microorganisms we have made contributions to some of the categories mentioned above, namely, the structure, assembly, function and relationship of several of the pathway proteins.

Summary of Recent Results

I. Photosynthetic Eucaryotic Organisms

A) Higher Plants

In previous years during the tenure of this grant we had obtained the following information. (1) The chemical steps of the pathway were invariant wherever we looked; this included mono- and dicotyledons, tobacco, corn, peas and beans and the single-celled Euglena, Chlorella and Chlamydomonas. (2) In tobacco (and as shown by others in peas) tryptophan synthase is a heteropolymer. The bacterial form of the enzyme is a heteropolymeric tetramer; the Neurospora enzyme is a homodimer encoded by a single gene. (3) We found feed-back inhibition by tryptophan of the pathway in vivo. (4) We found a DL-tryptophan racemase in tobacco, apparently the first amino acid racemase observed in higher plants.

In our recent work we measured the molecular weights and observed the aggregation patterns of the pathway enzymes in corn and in peas. Major points of interest that emerged were as follows. (1) The enzyme activities were all separate either

on DEAE or on sephadex G-150 chromatography. (2) The molecular weights of the individual enzymes were similar to one another and to those of the bacteria Pseudomonas putida and Bacillus subtilis. There was one exception. Indoleglycerol-phosphate synthase activity in peas eluted at a position corresponding to 52,000 daltons in contrast to the 30,000 daltons characteristic of corn and the two bacterial forms of the enzyme. (3) There was one additional difference between corn and peas. The corn phosphoribosyltransferase had an unusual requirement of 1.0M KCl for stability. (4) The data show conclusively that neither gene fusion nor stable enzyme aggregation are general features of eucaryotic cells.

We also examined in some detail the kinetics of anthranilate synthase of peas and corn. We found that the kinetic behavior, ordered sequential, was analogous to that of the bacterial enzyme and unlike the ping pong kinetics displayed by the Euglena enzyme.

In summary, the evidence to date is that the pathway enzymes in higher plants strongly resemble those of the procaryotic bacteria and not those of the eucaryotic higher fungi and Euglena.

B) Euglena gracilis

We had found earlier that the pathway enzymes of Euglena formed an unprecedented aggregate that included all enzyme activities save for anthranilate synthase. (It is the only example known of a non-separate tryptophan synthase). In our recent work we have purified the aggregate and anthranilate synthase which display the following features.

(1) Anthranilate synthase was purified 10,000-fold. Strong evidence was obtained that the molecule is 80,000 daltons and a single chain. No other case is known yet in nature in which the site of activation of glutamine is covalently linked to the site of activation of chorismate. The ratio of the NH_3 dependent to glutamine dependent activity of the fused enzyme was 0.3, the lowest ratio yet reported. (2) (a) A detailed kinetic study of the enzyme showed that the order of addition of the cofactor, Mg^{2+}, and substrates was, most probably, Mg^{2+}, chorismate, then glutamine. (b) Inhibitor studies clearly revealed the involvement of an - SH group in glutamine activation; this has been reported for the bacterial enzymes. (c) Ping pong kinetics were observed; unique reaction kinetics coupled to a unique structure. (d) This creates a dilemma. Ping pong kinetics demand a different reaction mechanism than ordered sequential kinetics. Does the Euglena enzyme uniquely require a leaving product before the reaction is complete? It was clearly shown that if so, the first leaving product, required by a ping pong mechanism, was not pyruvate. (e) Thus we were led to the possibilities that either we had true ping pong kinetics and the first leaving product was either OH or H_2O or some other inobvious group, or we had pseudo ping pong kinetics, a circumstance that can occur through particular values of the kinetic constants. (3) The 4-enzyme activity aggregate was purified. It was found that all of the activities were covalently linked as a homodimer with a molecular weight of 325,000 ± 20,000 for the enzyme and 155,000 ± 5,000 for the subunit. The sum of the minimum molecular weights of the corresponding activities in bacteria add up to about 150,000. The remarkable accommodation of all four activities on a single chain does

not seem to have demanded much change in the total amount of structural DNA.

The multiple gene fusions of Euglena and those that are becoming evident in the higher fungi show, in addition to their common and unique possession of the α-ketoadipate pathway of lysine biosynthesis, the close evolutionary relationship between these organisms. Further, that gene fusion is the direction of evolution in passing from procaryotic to eucaryotic organisms, proposed by Bonner, De Moss and myself many years ago on the basis of the work in E. coli and Neurospora, is also clearly incorrect. The higher fungi and Euglena represent an evolutionary branch point, interesting and biologically turbulent, but nonetheless a branch point.

With respect to other single-celled photosynthetic organisms, we examined Chlorella and Chlamydomonas; but decided not to do detailed studies when it developed that they showed the unaggregated enzyme pattern similar to the higher plants and the majority of bacteria. These data reinforce the impression that multiple structural gene fusion in amino acid biosynthetic pathways is confined to the higher fungi. The present evidence appears to be that the genes of a given pathway are scattered among the fungal chromosomes. However those genes fuse that evolve contiguously on the same chromosome. This represents an arithmetically immaculate way of regulating the ratio of the synthesis of formerly independent chains.

II. Bacteria

(a) Purification and Enzyme Studies on Microevolution

During this grant period our efforts have been concentrated on the enzymes anthranilatesynthase and phosphoribosyltransferase in the Enterobacteriaceae. It had previously been shown that these two enzymes were aggregated in some but not all of the members of this bacterial family; Escherichia coli is an example. In those bacteria wherein the enzymes were separate, molecular weight determinations in crude extracts revealed similar anthranilate synthases but two different phosphoribosy transferases, one of 68,000 daltons in Proteus and Erwinia and one of 45,000 daltons found in the genus Serratia. In Escherichia coli, Salmonella typhimurium and Enterobacter aerogenes, phosphoribosyl transferase was known to be fused to glutamine amidotransferase, the subunit of anthranilate synthase that activates glutamine in the first reaction of the pathway. The consequence of the fusion was that phosphoribosyl transferase and anthranilate synthase were aggregated. The size of the fused polypeptide was about 65,000 and the two functions had been shown to be associated with different segments of the polypeptide chain, with the glutamine-binding site comprising one third of the chain starting from the amino terminus. Proteolytic digestion of the fused enzyme of either Escherichia coli or Salmonella typhimurium destroyed the phosphoribosyl transferase activity but left intact a fragment, of molecular weight about 23,000, that retained the glutamine-activating site and formed an anthranilate synthase altogether comparable to the enzyme found in Serratia marcescens.

Purification of the Serratia anthranilate synthase revealed the molecular weight of the glutamine amidotransferase subunit as 21,000; 21,000 plus 45,000 for the Serratia phosphoribosyl transferase added up to the molecular weight of the

Escherichia fused enzyme and led to the postulation of a gene fusion.

Outside of the determination of their molecular weight relatively little had been done with the phosphoribosyl transferases. It appeared to us to be an excellent system, within a well defined and well described group of bacteria, to obtain information on the microevolution of an enzyme. We began our studies with the purification of the enzyme from Serratia marcescens and Erwinia carotovora, together with anthranilate synthase from Erwinia since the molecular weight of the glutamine amidotransferase was required. The data are summarized, together with the data from Escherichia, in Table 1.

Table 1

	Gene Products				Enzyme Formula	
	TrpE	Trp(G-D)	TrpG	TrpD	AS	PRT
		proteolysis				
E. coli	65,000	65,000 →	23,000	-	$E_2(G\text{-}D)_2$	
S. marcescens	65,000		21,000	45,000*	E_2G_2*	D*
E. carotarova	65,000*		23,500*	40,000*	E_2G_2*	D_2*

AS = anthranilate synthase PRT = phosphoribosyl transferase
The formulas are given in terms of the gene symbols. G-D represents the fused enzyme.

The starred entries are the proteins purified and analyzed by us.

The amino-terminal amino acid sequences of the two phosphoribosyl transferases are as follows:

```
                 1               5                   10
Erwinia          Met-Glu-Ala-Thr-Leu-Ile-Lys-Pro-Thr-Ile-Phe-Thr-

                 13      15                  20
                 His-Glu-Pro-Ile-Leu-Glu-Lys-Leu-Phe-Lys-Ser-Gln-
Serratia         Met-Glu-Pro-Ile-Leu-Glu-Lys-Leu-Tyr-Arg-Ala-Glu-

                 25                  30                  35
                 Ser-Met-Thr-Gln-Glu-Glu-Ser-?  -Gln-Leu-Phe-
                 Ser-Met-Ser-Gln-Gln-Glu-Ser-Gln-Gln-Leu-Phe-

                 Ala-Ala-Ile-Val
                 Ser-Ala-Ile-Val-Arg-Gly-Glu-Leu-Glu-Pro-Ser-Glu
                                    30                  35
```

As is customary the analysis of the purified proteins brought out some interesting features. The previously reported 67,000 dalton Erwinia phosphoribosyl transferase was found to be a homodimer with a subunit of molecular weight equal to 40,000. While the trpD gene product was smaller than that of Serratia the trpG gene product was found to be larger, 23,500 as compared with 21,000; thus the trpG and trpD gene products of both Erwinia and Serratia add up to the Escherichia fused product. The amino acid analysis showed that of the 25 amino acids in register (neglecting the first methionine in Serratia) 18 are identical and seven, all single base changes, are conservative changes, preserving charge, geometry and solubility. However, and somewhat unexpected, the larger Serratia molecule came into register with amino acid 13 of the small Erwinia gene product. If the molecular weight estimate obtained by dodecylsulfate acrylamide gel electrophoresis proves to be equivalent to an amino acid difference of about 20 to 40 then the Serratia monomer has a carboxy terminus of about 35-55 amino acids longer than the Erwinia subunit.

An outline of the general relationship among the phosphoribosyl transferases and the glutamine amidotransferases is now emerging. A tentative hypothesis can be formulated to try and account for the multiple structural forms of the phosphoribosyl transferase as compared to the relative conservation of the other pathway structural gene products. First, it has been postulated that all the Enterobacteriaceae possess a tryptophan operon similar to that of Escherichia. This implies that the trpG and trpD genes lie adjacent to one another and that the Escherichia trp G-D gene represents a fusion of contiguous genes. If the fusion occurred in a precursor organism to the Enterobacteriaceae then what we see is different points of divorce of the fused gene followed by processing as the family members adapted to their various niches. If fusion occurred after the appearance of the Enterobacteriaceae then the argument might go as follows. It is known that the tryptophan biosynthetic pathway can function without glutamine. The normally aggregated Escherichia trpE gene product alone, as is true of every trpE gene product tested, catalyzes the formation of anthranilate with NH_3 as substrate and trpG mutant strains grow when supplemented with a source of NH_3. TrpG, which introduced glutamine into the pathway, thus may be postulated as a recent arrival. The fluctuation of the phosphoribosyl transferase then mirrors the fact that it is through trpD that the pathway has accommodated to the intrusion of trpG. The other pathway genes, including trpE have been left undistrubed and exhibit the normal intergeneric variation. This working hypothesis forms the basis of our further efforts to isolate additional representative of the enzyme and to learn more about their properties.

(b) The Structure of the Aggregate in Escherichia coli

We examined the structure of the anthranilate synthase - phosphoribosyltransferase aggregate by quantitative mico-complement fixation and I^{125} radiolabelled antigens. We obtained the following information. (1) Component I of the aggregate (formula $(CoI)_2(CoII)_2$) the polypeptide prescribed by trpE and which activates chorismate and NH_3, differs little in tertiary structure whether free or aggregated. (2) Component II, encoded by trpD and the activator of glutamine, anthranilate and PRPP, undergoes marked changes in tertary structure displaying extensive heterogeneity when disaggregated. (3) The glutamine-activating region of CoII, the

region that links the two chains into the aggregate, was found to contribute no antibodies to the antisera (both goat and rabbit). Antibodies to this region were easily demonstrable when unbound Component II was injected into animals. Therefore it appears that the entire region is buried in the complex. This explains the essentially irreversible but non-covalent bonding between the subunit components. Further the immunochemically observed heterogeneity of the isolated trpD gene product explains why it has not yet been reported purified.

III. Purification and Assay procedures.

(1) We developed an improved solid phase antibody method for the purification of proteins. By coupling normal γ-globulin to the appropriate solid support media and using it as a prefilter before the antibody column we were able to obtain, in reasonably good yield, electrophoretically clean protein. We used the method to purify the β_2 component of tryptophan synthase and a mutant form of the β_2 protein.

(2) In examining the enzymes of tryptophan synthesis in plants we were unable to utilize the existing assays for indoleglycerolphosphate synthase. Accordingly, by taking advantage of the fluorescence of the reaction product we devised a rapid and sensitive spectrophotofluorometric assay for the enzyme that is indifferent to the presence of glycerol, often a necessary stabilizing agent where low concentrations of enzyme are sought. With the addition of this method the first four reactions of the pathway can be conveniently and quickly measured by fluorometry.

Symmary

The research we have completed contributed evidence for the following propositions.

(1) Examination of the pathway of tryptophan biosynthesis in representatives of several major taxa reveals that the higher plants share the same enzyme pattern, lack of aggregation, with the eucaryotic algae and the majority of procaryotic bacteria. This occurs in the face of the principle that procaryotic cells generally regulate amino acid synthesis by negative control and eucaryotic cells by positive control. The genetic arrangement of the pathway structural genes in higher plants is unknown.

(2) In the eucaryotic fungi, in which positive control appears to be predominant, the pathway genes are scattered. Those that remain on the same chromosome appear to fuse. In Euglena gracilis, a photosynthetic, physiological nearest-neighbor of the higher fungi, there is no information on the location of the structural genes of the pathway, but they have been reduced to two. This is inferred from the fact that two polypeptide chains comprise the entire pathway (made up of seven chains in the Pseudomonads). This is an extreme method of regulating the polypeptide chains in a biosynthetic pathway.

(3) The use of quantitative immunochemical procedures has revealed some structural details on the anthranilate synthase-ribosyl transferase aggregate in Escherichia coli. It was inferred that the entire proximal portion of the trpD gene product, about 1/3 of the 65,000 dalton polypeptide chain, is entirely buried in the complex. This inference puts into perspective many of the known biochemical properties of the aggregate.

(4) Progress has been made in revealing the details of the three strikingly different forms of the phosphoribosyl transferase in the Enterobacteriaceae. A theorem has been proposed for the microevolution of the corresponding structural gene from a free to the form fused with the gene of the glutamine-activating subunit. The theorem has consequences that can be tested.

I thank those affiliated with ERDA and its predecessor agency, the AEC, for their support. If, by chance, the agency should come to regret their termination of this research, which will continue with the above work plus an investigation of indole acetic acid biosynthesis and control, I would be delighted to resume our relationship.

Publications

1. Hankins, C. N., Largen, M. T. and S. E. Mills (1975). A rapid spectrophotofluorometric assay for indoleglycerol phosphate synthetase. Anal. Biochem. 69, 510-517.

2. Hankins, C. N., Largen, M. T. and S. E. Mills (1976). Some physical characteristics of the enzymex of L-tryptophan biosynthesis in higher plants. Plant Physiol. 57, 101-104.

3. Shannon, L. M. and S. E. Mills (1976). Purification by immunoadsorbtion chromatography of the normal and a mutant form of the β_2 subunit of Escherichia coli tryptophan synthase. Eur. J. Biochem. 63, 563-568.

4. Largen, M., Mills, S. E., Rowe, J. and C. Yanofsky (1976). Purification, subunit structure and partial amino-acid sequence of anthranilate-5-phosphoribosyl pyrophosphate phosphoribosyltransferase from the enteric bacterium Serratia marcescens. Eur. J. Biochem. 67, 31-36.

5. Hankins, C. N. and S. E. Mills (1976). Anthranilate synthase-amidotransferase (Combined)* A novel form of anthranilate synthase from Euglena gracilis. J. Biol. Chem. 251, 7774-7778.

6. Largen, M. and S. E. Mills (1977). Immunochemical Analysis of the anthranilate synthase-anthranilate phosphoribosylpyrophosphate phosphoribosyltransferase aggregate of Escherichia coli. Biochem. 16, 2526-2531.

7. Hankins, C. N. and S. E. Mills (1977) A dimer of a single polypeptide chain catalyzes the terminal four reactions of the L-tryptophan pathway in Euglena gracilis. J. Biol. Chem. 252, 235-239.

8. Largen, M., Mills, S. E., Rose, J. and C. Yanofsky (1978). Purification and properties of a third form of anthranilate-5-phosphoribosylpyrophosphate phosphoribosyltransferase from the Enterobacteriaceae. J. Biol. Chem. 253, 409-412.

In Preparation

1. Hankins, C. N. and S. E. Mills. A kinetic analysis of anthranilate synthase-amidotransferase (Combined) from Euglena gracilis.

2. Schwartz, B. and S. E. Mills. Purification of anthranilate synthase from Erwinia carotovora.

3. Hankins, C. N. and S. E. Mills. A kinetic analysis of anthranilate synthase from Zea mays.

Reprints removed
cp 2/24/81