JOURNAL OF BACTERIOLOGY, Oct. 1971, p. 248-253

Vol. 108, No. 1
Printed in U.S.A.

Pseudomonas putida Tryptophan Synthetase: Partial Sequence of the α Subunit

IRVING P. CRAWFORD AND CHARLES YANOFSKY

Department of Microbiology, Scripps Clinic and Research Foundation, La Jolla, California 92037, and Department of Biological Sciences, Stanford University, Stanford, California 94305

Received for publication 7 April 1971

There is 50% identity in the sequences of the first 50 residues of the α chains of *Escherichia coli* and *Pseudomonas putida*. No deletions or additions of residues are found in this region, except for the N-terminal methionine residue which is missing in the polypeptide isolated from *P. putida*. Most of the residues which differ are chemically dissimilar, and half of them are specified by codons which differ by more than a single base. The two residues known by mutational analysis to be essential for catalysis in *E. coli* are conserved in *P. putida*. The potential taxonomic usefulness of information of this sort is analyzed.

A great many dissimilarities exist in the genes and enzymes of the tryptophan pathway in pseudomonads and enterobacteria. All enzymes in the pathway leading from chorismic acid to tryptophan in *Escherichia coli* and other enterobacteria are specified by a single cluster of five genes, the tyrptophan operon (27, 26). *Pseudomonas putida*, on the other hand, has seven structural genes concerned with this pathway; these are disposed in three groups on the chromosome and do not constitute even one regulon (14). There is evidence that this disposition is the same for *P. aeruginosa* (10). Figure 1 diagrams these *Pseudomonas trp* genes and the enzymes they specify.

The two loci controlling the α and β chains of tryptophan synthetase (EC 4.2.1.20), the last enzyme of the pathway, are regulated in pseudomonads through induction by their substrate, indolyl-3-glycerolphosphate, not through end product repression, as occurs in their enterobacterial counterparts (4). Recently the α and β_2 subunits of tryptophan synthetase were purified from extracts of induced *P. putida* auxotrophs (9*a*, 16). Because of the differences in chromosomal position and regulation of the genes encoding them, we wished to compare these proteins with their better known *E. coli* homologues.

Previous studies showed that in molecular weight, subunit structure, and general catalytic features, the tryptophan synthetases of *P. putida* and *E. coli* are very similar (9, 9*a*, 16). The *Pseudomonas* enzyme differs somewhat from that of *E. coli* in certain catalytic parameters, such as the rate of side reactions and response to cations (9*a*). No interspecies recognition of the α-β or β-β chain binding sites has been demonstrated, and the immunological cross-reaction between the β_2 subunits is very slight (16). In contrast, all enterobacterial α and β_2 subunits recognize each other very well (1, 5), and strong immunological cross-reactions have been demonstrated among the α subunits of the group (17).

The first segment of primary structure available for comparison between the enzymes of *P. putida* and *E. coli* was the 23-residue region surrounding the β chain lysine that binds pyridoxal-5'-phosphate (R. Maurer and I. P. Crawford, J. Biol. Chem., *in press*). Here striking similarity was found, with 15 of the residues identical in the two species, three others changed to chemically similar amino acids, and all ionized groups conserved. It would be surprising if all the amino acids in this 23-residue sequence were concerned with cofactor or substrate binding and activation. But because of the possibility that some of the residues near the coenzyme are invariant for functional reasons, it seemed desirable to sequence another segment of the polypeptide chain not known to be proximate to a catalytic center. For this purpose we chose the N-terminal segment of the α chain. [The N-terminus of the β chain is blocked in both organisms (16); therefore this chain cannot be used in sequential analysis by the Edman procedure (7).] The entire sequence of the *E. coli* α chain is known (24). Only two sites of mutational substitution leading to inactive proteins have been found in the first 50 residues [C. Yanofsky, *unpublished data*; (24)]. This communication presents the result of sequential Edman degradation of the intact *P. pu-*

tida α chain and compares the findings with the *E. coli* sequence. Our aim was to decide whether the tryptophan synthetase genes in these two organisms are evolutionarily related, that is, whether they are homologous or analogous (11). If they are homologous, a genetic event involving translocation to a new position and attachment to or detachment from a new regulatory element would be sufficient to account for the differences in the two bacteria. If these genes are not clearly homologous, the possibility of mutational alteration of function of some previously unrelated enzyme to fill the need occasioned by a deleted, inactivated, or not-yet-evolved tryptophan synthetase segment must be weighed.

MATERIALS AND METHODS

Protein purification. The *P. putida* α subunit was purified, by a method described in an accompanying paper (*9a*), from an extract of induced cells of the auxotroph *trpF661* (biotype A, reference 14). The yield from 375 g (wet weight) of cells was about 60 mg of protein that was essentially homogeneous on acrylamide gel electrophoresis.

Automatic Edman degradation. Edman degradations were performed automatically (7) with a Beckman Sequencer. All solvents and reagents were purchased from Beckman Instruments. The procedures employed are described in the Beckman Sequencer Manual.

Conversion of thiazolinones. The procedures recommended in the Beckman Sequencer Manual were followed with minor changes. When an acid or amide was suspected, the acid conversion was performed for 3 rather than 10 min at 80 C. When proline or threonine was suspected, acid conversion was carried out for 10 min at 50 C.

Gas-liquid chromatography. Gas-liquid chromatography was performed with a Beckman GC45 gas chromatograph and glass columns (2 mm by 1.2 m). The procedures employed are described in the Beckman Sequencer Manual and elsewhere [J. J. Pisano and T. J. Bronzert, Fed. Proc. **24**:916, 1970; (18)]. The supports used were DC560 and 2/3 CFC (a mixture of SP-400, OV-210, and OV-225). The 2/3 CFC support was kindly supplied by J. Pisano of the National Institutes of Health. A temperature program was used for the separation of the phenylthiohydantoin (PTH)-amino acids. Where necessary, samples were silylated by injecting directly onto the column 1 μliter of sample and 1 μliter of *N,O*-bis (trimethylsilyl)-acetamide (Pierce Chemical Co.) contained in the same syringe.

Thin-layer chromatography. Thin-layer chromatographic identification of PTH-amino acids was performed as described by Jeppsson and Sjöquist (15). Solvent system IV (50 ml of heptane, 30 ml of *n*-butanol, and 9 ml of 25% formic acid) and Eastman chromatographic sheets (type 6060 with fluorescent indicator) were employed. The sheets were activated by heating at 100 C for 30 min, and were equilibrated in the chromatographic jar for 30 min before being placed in the developing solvent. Chromatography was terminated when the solvent front reached the middle of the sheet. PTH-amino acids were identified by observing quenching of fluorescence under ultraviolet light. Aspartic and glutamic acids and their amides were clearly distinguished by these procedures.

Hydrolysis of PTH-amino acids. The sample was dried in a hydrolysis tube and 0.3 ml of 5.7 N HCl containing 10^{-3} M β-mercaptoethanol was added. Vacuum was applied while the solution was freezing in a dry ice-acetone bath. The tube was sealed under vacuum and placed in a 130 C oven for 24 hr. The hydrolysis tube was cracked, and the contents were dried in a vacuum desiccator over NaOH pellets. The dried sample was dissolved in *p*H 2.2 buffer, and samples were taken for analysis of amino acid content on a Beckman 121 automatic amino acid analyzer.

Identification of PTH-amino acids. The following PTH-amino acids were unambiguously identified by direct injection onto columns of 2/3 CFC or DC560: serine, alanine, valine, threonine, proline, glycine, isoleucine, leucine, methionine, and phenylalanine. Argi-

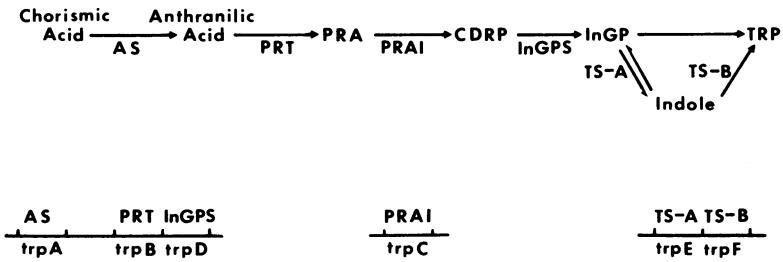

FIG. 1. *Tryptophan pathway organization in Pseudomonas putida. Intermediates in the pathway are shown in the upper line; genes controlling the enzymes of the pathway are on the lower line. Abbreviations: PRA, N-5'-phosphoribosylanthranilic acid; CDRP, 1-(o-carboxyphenylamino)-1-deoxyribulose-5-phosphate; InGP, indolyl-3-glycerolphosphate; TRP, L-tryptophan; AS, anthranilate synthetase; PRT, phosphoribosyl transferase; PRAI, phosphoribosylanthranilate isomerase; InGPS, indoleglycerolphosphate synthetase; TS-A, α subunit of tryptophan synthetase; TS-B, β_2 subunit of tryptophan synthetase. In addition to the six trp genes controlling the enzymes indicated, a seventh exists for the small, glutamine-binding subunit of AS (19); as this polypeptide is regulated coordinately with the larger AS subunit controlled by the trpA gene, it may lie within the trpABD cluster.*

nine, histidine, and lysine were identified by liquid amino acid analysis with a Beckman amino acid analyzer after hydrolysis of the aqueous (arginine, histidine, and sometimes lysine) or ethyl acetate (lysine) fraction obtained after acid conversion. Aspartic and glutamic acids and their amides were identified as the PTH-acids by on-column silylation. The acids and amides were also distinguished by thin-layer chromatography. Proline and threonine were distinguished by silylation and by hydrolysis of PTH-amino acids and liquid amino acid analysis. PTH-threonine is destroyed by hydrolysis, but proline is readily recovered from PTH-proline. PTH-tyrosine was identified by silylation. The identification of particular residues was occasionally confirmed by silylation, or liquid amino acid analysis, or both. The PTH-amino acid standards were purchased from Pierce Chemical Co.

RESULTS

Amino acid sequence of residues 1 to 50 of the isolated Pseudomonas putida α chain. A sample of purified α chain was dialyzed overnight against 0.001 N NH_4OH and then for 8 hr against water. The sample was lyophilized and 250 nmoles of dry protein was added to the reaction cup of a Beckman Sequencer. The sample was dissolved in *n*-heptafluorobutyric acid and vacuum dried on the walls of the spinning reaction cup. Fifty cycles of degradation were performed automatically. A second run was performed identically except that samples assumed to contain the thiazolinones of threonine, proline, aspartic acid, glutamic acid, asparagine, and glutamine were converted to the respective thiohydantoins as described above. Identification procedures employed and deduced residues are presented in Table 1. A comparison between the *P. putida* and *E. coli* sequences is presented in Fig. 2. When the first residue of *P. putida* is aligned with the second residue (Gln) of *E. coli*, a homologous alignment is seen for the entire segment, with half the residues identical for the two species.

DISCUSSION

The aerobic pseudomonads (22) and the enterobacteria (8) form two well defined, well characterized families within the bacterial kingdom. Numerous methods may be used to decide whether a given strain belongs to one or the other of these groups and to assort the members of a given group into named genera and species. Computer-aided numerical taxonomy, nucleic acid analysis and hybridization, genetic analysis by episomal transfer, immunological procedures, and the study of macromolecular aggregates and subunit associations all seem suitable for this task. None of these methods is entirely satisfactory for the determination of natural relationships between these and other major bacterial families, however. The methods listed, though suitable for intragroup taxonomy, seem to possess a resolving power too great for determining intergroup relationships. The comparative analysis of amino acid sequences, though tedious, may have the correct discriminatory capacity to indicate natural relationships among the major bacterial groups, allowing us to reconstruct the probable paths of evolutionary divergence that have resulted in the microbial diversity we know.

Fortunately there is a considerable body of evidence on the divergence of protein sequences among eukaryotic organisms whose evolutionary relationships are not seriously in question (6, 12, 21). Assuming that changes in amino acid sequence occur by point mutations or chromosomal rearrangements as some function of evolutionary time, a dendritic "tree" bearing close resemblance to the course of evolution as decided by zoological and paleontological methods can be constructed (6, 12, 13). Although the genetic apparatus of bacteria differs in form somewhat from that of eukaryotes, we have no reason to suspect that evolutionary processes will act differently upon it. In discussing our results, therefore, we should be able to draw parallels with the more extensive studies performed in higher organisms.

Neither the 50-residue sequence reported here nor the 23-residue pyridoxyl peptide of the β chain (R. Maurer and I. P. Crawford, J. Biol. Chem., *in press*) shows evidence of deletions, additions, or inversions within the genetic segments encoding these regions. The absence of the N-terminal methionyl residue in *P. putida* is probably attributable to action of an aminopeptidase within the cell, able to cleave the methionyl-serine bond (3). The methionylglutamine sequence at the N-terminus of the *E. coli* α chain may be resistant to that organism's intracellular peptidases. The rarity of glutamic acid or glutamine as an N-terminus of soluble proteins in *E. coli* (3) is in agreement with this explanation.

In the first 50 residues of the α chain the degree of conservation is not as great as that found in the 23-residue pyridoxyl peptide of the β chain. This may reflect the known proximity of the phosphopyridoxyl-lysine segment to the active site, or it may be fortuitous. Of the 24 α chain residues that are different, 13 cannot be attained by a single base substitution in the codon, giving 37 as the minimal number of mutations required to convert this segment of the *E. coli* α chain to the *P. putida* version or vice versa. In contrast, several of the enterobacterial α chains studied have 12% or less of the first 50 residues altered, and all of the differences could

TABLE 1. *Amino acid sequence of residues 1 to 50 of the Pseudomonas putida α chain*

Step	Deduced residue	Run no. 1						Run no. 2					
		Gas chromatography				TLC[b]	121[c]	Gas chromatography				TLC[b]	121[c]
		2/3 CFC		DC 560				2/3 CFC		DC 560			
		−S	+S[a]	−S	+S[a]			−S	+S[a]	−S	+S[a]		
1	Ser	Ser	Ser					Ser	Ser				
2	Arg	—O—	—O—				Arg	—O—	—O—				Arg
3	Leu	Leu						Leu					
4	Glu	—O—	Glu			Glu		—O—	Glu	—O—	Glu	Glu	
5	Gln	—O—	Glu			Gln		—O—	Glu	—O—	Gln	Gln	
6	Arg	—O—	—O—				Arg	—O—	—O—				Arg
7	Phe	Phe						Phe					
8	Ala	Ala						Ala					
9	Glu	—O—	Glu			Glu		—O—	Glu	—O—	Glu	Glu	
10	Leu	Leu						Leu					
11	Lys	—O—	—O—				Lys	—O—	—O—				Lys
12	Ala	Ala						Ala					
13	Glu	—O—	Glu			Glu		—O—	Glu	—O—	Glu	Glu	
14	Gly	Gly	Gly					Gly		Gly			
15	Arg	—O—	—O—				Arg	—O—	—O—				Arg
16	Ser	Ser	Ser					Ser	Ser	Ser	Ser		
17	Ala	Ala						Ala					
18	Leu	Leu						Leu		Leu			
19	Val	Val					Val	Val					
20	Ile	Ile					Ile	Ile					
21	Phe	Phe						Phe					
22	Val	Val		Val				Val					
23	Thr	—O—	—O—	—O—	Thr		(Thr)[d]	—O—	—O—	—O—	Thr		
24	Ala	Ala		Ala				Ala					
25	Gly	Gly						—[e]	—[e]				
26	Asp	—O—	Asp			Asp	Asp	—O—	Asp	—O—	Asp	Asp	
27	Pro	Pro-Thr	Pro				Pro	—[e]	—[e]				
28	Gly	Gly	Gly					Gly					
29	Tyr	—O—	Tyr				Tyr	—O—	Tyr				
30	Asp	—O—	Asp			Asp		—O—	Asp	—O—	Asp	Asp	
31	Ala	Ala						Ala		Ala			
32	Ser	Ser	Ser					Ser		Ser	Ser		
33	Leu	Leu						Leu		Leu			
34	Gln	—O—	Glu			Gln		—O—	Glu			Gln	
35	Ile	Ile						Ile					
36	Leu	Leu					Leu	Leu					
37	Lys	—O—	—O—				Lys	—O—	—O—				Lys
38	Gly	Gly						Gly					
39	Leu	Leu						Leu					
40	Pro	Pro-Thr	?				Pro	Pro-Thr		Pro-Thr	Pro		Pro
41	Ala	Ala						Ala					
42	Ala	Ala					Ala	Ala					
43	Gly	Gly					Gly	Gly					
44	Ala	Ala					Ala	Ala					
45	Asp	—O—	Asp?			Asp	Asp	—O—	Asp			Asp	
46	Val	Val						Val					
47	Ile	Ile					Ile	Ile					
48	Glu	—O—	Glu			Glu		—O—	Glu			Glu	
49	Leu	Leu						Leu					
50	Gly	Gly						Gly					

[a] On-column silylation of sample.
[b] Thin-layer chromatography.
[c] Analysis of hydrolyzed sample on Beckman 121 amino acid analyzer.
[d] Threonine and proline both absent, therefore residue is threonine.
[e] Sample lost.

have arisen as a result of single base substitutions (C. Yanofsky and S. Li, *manuscript in preparation*).

The two residues known to mutate to enzymatically inactive missense proteins in this region of the *E. coli* protein are Phe-22 (C. Yanofsky, *unpublished data*) and Glu-49 (25). It is interesting to note that these residues are conserved in *P. putida*, and that they lie adjacent to other conserved residues.

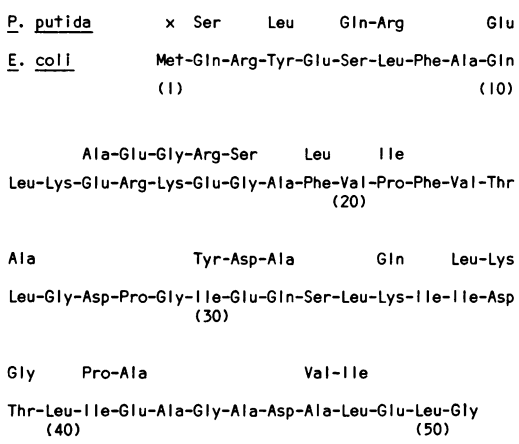

FIG. 2. *Alignment of the first 50 residues of the α chains of Escherichia coli and Pseudomonas putida tryptophan synthetases. Where no residue is shown for the P. putida sequence, the residue found was identical to that in E. coli; × indicates that in the protein analyzed there is no P. putida residue corresponding to the one in E. coli. The sequence presented for the E. coli protein differs from the previously published one (24) in having an additional isoleucine immediately preceding the one at position 36 (C. Yanofsky, unpublished data); i.e., the previous sequence Lys-35-Ile-36-Asp-37 is now known to be Lys-35-Ile-36-Ile-37-Asp-38.*

Those side chains bearing a charge are not particularly well conserved in the region studied; 9 of 16 positions charged in one or the other protein show a charge change, and 3 of those are charge reversals, either Arg ↔ Glu or Lys ↔ Asp. This and the general pattern of amino acid changes observed is reminiscent of the differences between bovine trypsin and chymotrysinogen A (43% of the residues conserved) (21). "Internal" residues inaccessible to solvent constitute most of the residues conserved in these mammalian pancreatic serine proteases; 71% of the "internal" residues and only 19% of the "external" ones are conserved (or chemically similar) among these proteases (21). It is tempting to suppose that the majority of the 24 residues changed in the first portion of the α chain of tryptophan synthetase will lie on the surface of the molecule; certainly most of the 20 "chemically dissimilar" changes (21) may prove to do so. (Changes at positions 10, 31, 37, and 48 are to chemically similar amino acids.) However, two instances of "balanced" alterations of enzymatically important residues already have been observed in the α chain of *E. coli* (25). In these cases a substitution of a dissimilar residue at one position in the sequence can be compensated for by a change of another residue some 35 residues away in the primary sequence, showing that not all chemically dissimilar changes need be external. Nevertheless, it will be interesting to see if the crystallographic studies of the α subunit (20) can confirm the external location of many of the α chain residues that differ in *P. putida* and *E. coli*, and thus account for the observed lack of immunological cross-reactivity and interspecies α-β binding (9*a*, 16).

We do not suggest that all major bacterial taxa can or will be placed in a "natural" or evolutionary order solely by an examination of the first 50 residues of the α chain of tryptophan synthetase. To be believed, results of this sort must be found consistently with a number of proteins of different types and functions. Already work has begun on the rubredoxins (2), and, as techniques improve, it should be possible to study many other polypeptides. Past pitfalls encountered in analyzing the natural relationships among the bacteria are many and well documented (24). Nonetheless, the approach of sequencing homologous proteins, provided homology can be proved (11), is an intellectually satisfying strategy for resolving some long-standing issues. Additional studies on the sequence of the tryptophan synthetase molecules of other bacteria as well as on other proteins would seem to be imperative.

ACKNOWLEDGMENTS

We are indebted to William Lee for help in the purification of the protein, Shoei-Lung Li for aid in the identification of PTH-amino acids, and Susan Stasiowski for excellent technical assistance.

These studies were supported by grants from the National Science Foundation, Public Health Service, and American Heart Association. C. Y. is a Career Investigator of the American Heart Association.

LITERATURE CITED

1. Balbinder, E. 1964. Intergeneric complementation between A and B components of bacterial tryptophan synthetases. Biochem. Biophys. Res. Commun. **17**:770-774.
2. Benson, A., K. Tomoeda, J. Chang, G. Matsueda, E. T. Lade, M. J. Coon, and K. T. Yasunobu. 1971. Evolutionary and phylogenetic relationships of rubredoxin-containing microbes. Biochem. Biophys. Res. Commun. **42**:640-646.
3. Brown, J. L. 1970. The N-terminal region of soluble proteins from procaryotes and eucaryotes. Biochim. Biophys. Acta **221**:480-488.
4. Crawford, I. P., and I. C. Gunsalus. 1966. Inducibility of tryptophan synthetase in *Pseudomonas putida*. Proc. Nat. Acad. Sci. U.S.A. **56**:717-724.
5. Creighton, T. E., D. R. Helinski, R. L. Somerville, and C. Yanofsky. 1966. Comparison of the tryptophan synthetase α subunits of several species of *Enterobacteriaceae*. J. Bacteriol. **91**:1819-1826.
6. Dayhoff, M. O. 1969. Atlas of protein sequence and structure. National Biomedical Res. Found., Silver Spring, Md.
7. Edman, P., and G. Begg. 1967. A protein sequenator. Eur. J. Biochem. **1**:80-91.
8. Edwards, P. R., and W. H. Ewing. 1962. Identification of *Enterobacteriaceae*. Burgess Publ. Co., Minneapolis, Minn.

9. Enatsu, T., and I. P. Crawford. 1968. Enzymes of the tyrptophan synthetic pathway in *Pseudomonas putida*. J. Bacteriol. **95**:107–112.
9*a*. Enatsu, T., and I. P. Crawford. 1971. Pseudomonas putida tryptophan synthetase. J. Bacteriol. **108**:248–253.
10. Fargie, B., and B. W. Holloway. 1965. Absence of clustering of functionally related genes in *Pseudomonas aeruginosa*. Genet. Res. **6**:284–299.
11. Fitch, W. M. 1970. Distinguishing homologous from analogous proteins. Syst. Zool. **19**:99–113.
12. Fitch, W. M., and E. Margoliash, 1966. Construction of phylogenetic trees. Science **155**:279–284.
13. Fitch, W. M., and E. Margoliash. 1968. The construction of phylogenetic trees. II. How well do they reflect past history? Brookhaven Symp. Biol. **21**:217–241.
14. Gunsalus, I. C., C. F. Gunsalus, A. M. Chakrabarty, S. Sikes, and I. P. Crawford. 1968. Fine structure mapping of the tryptophan genes in *Pseudomonas putida*. Genetics **60**:419–435.
15. Jeppsson, J., and J. Sjöquist. 1967. Thin-layer chromatography of PTH amino acids. Anal. Biochem. **18**:264–269.
16. Maurer, R., and I. P. Crawford. 1971. Properties and subunit structure of the B component of *Pseudomonas putida* tryptophan synthetase. Arch. Biochem. Biophys. **144**:193–203.
17. Murphy, T. M., and S. E. Mills. 1968. Immunochemical comparisons of mutant and wild-type α-subunits of tryptophan synthetase. Arch. Biochem. Biophys. **127**:7–16.
18. Pisano, J. J., and T. J. Bronzert. 1969. Analysis of amino acid phenylhydantoins by gas chromatography. J. Biol. Chem. **224**:5597–5607.
19. Queener, S. F., and I. C. Gunsalus. 1970. Anthranilate synthetase enzyme system and complementation in *Pseudomonas* species. Proc. Nat. Acad. Sci. U.S.A. **67**:1225–1237.
20. Schultz, G. E., and T. E. Creighton. 1969. Preliminary X-ray diffraction study of the wild-type and a mutationally altered tryptophan synthetase α-subunit. Eur. J. Biochem. **10**:195–197.
21. Shotton, D. M., and B. S. Hartley. 1970. Amino acid sequence of porcine pancreatic elastase and its homologies with other serine proteases. Nature (London) **225**:802–806.
22. Stanier, R. Y., N. J. Palleroni, and M. Doudoroff. 1966. The aerobic pseudomonads: a taxonomic study. J. Gen. Microbiol. **43**:159–271.
23. Van Neil, C. B. 1946. The classification and natural relationships of bacteria. Cold Spring Harbor Symp. Quant. Biol. **11**:285–300.
24. Yanofsky, C., G. R. Drapeau, J. R. Guest, and B. C. Carlton. 1967. The complete amino acid sequence of the tryptophan synthetase A protein (α subunit) and its colinear relationship with the genetic map of the gene. Proc. Nat. Acad. Sci. U.S.A. **57**:296–298.
25. Yanofsky, C., V. Horn, and D. Thorpe. 1964. Protein structure relationships revealed by mutational analysis. Science **146**:1593–1594.
26. Yanofsky, C., and J. Ito. 1966. Nonsense codons and polarity in the tryptophan operon. J. Mol. Biol. **21**:313–334.
27. Yanofsky, C., and E. S. Lennox. 1959. Transduction and recombination study of linkage relationships among the genes controlling tryptophan synthesis in *Escherichia coli*. Virology **8**:425–447.