JOURNAL OF BACTERIOLOGY, Aug. 1975, p. 761-764

Vol. 123, No. 2
Printed in U.S.A.

Amino-Terminal Sequences of Indoleglycerol Phosphate Synthetase of *Escherichia coli* and *Salmonella typhimurium*

STEVEN SHOEI-LUNG LI,[1] JOAN HANLON, AND CHARLES YANOFSKY*
Department of Biological Sciences, Stanford University, Stanford, California 94305

Received for publication 13 May 1975

The partial sequences of the first 40 residues of indoleglycerol phosphate synthetase of *Escherichia coli* and *Salmonella typhimurium* were determined, and three amino acid differences were observed among the 38 residues compared.

In *Escherichia coli* and *Salmonella typhimurium* the genetic information for the tryptophan biosynthetic enzymes is organized, transcribed, and translated in an identical or very similar fashion (1, 4, 23, 24). Evolutionary divergence of the tryptophan operons of these two bacterial species has been assessed by examination of nucleotide sequence homology and amino acid sequence variation (5, 13). The primary structure of the tryptophan synthetase α chains (9, 15, 16) and the amino-terminal sequences of anthranilate synthetase components I and II (14) have been compared. The peptide patterns of the phosphoribosyl anthranilate isomerase-indoleglycerol phosphate synthetases of the two species have also been compared (17). Genetic and biochemical studies have indicated that the amino-terminal half of this bifunctional polypeptide chain is primarily responsible for the synthetase reaction whereas the carboxyl-terminal half is principally concerned with the isomerase reaction (2, 17, 20, 24). In some bacterial species these activities reside in separate polypeptide chains (7, 12, 19, 22). In this paper we compare the amino-terminal 40 residues of the indoleglycerol phosphate synthetases of *E. coli* and *S. typhimurium*.

(i) *E. coli* strain *trpA2*/F'colVB *trpA2* and *S. typhimurium* auxotroph *trpD13* (*E. coli* nomenclature) were used in this investigation.

(ii) Protein purification. The phosphoribosyl anthranilate isomerase-indoleglycerol phosphate synthetase of *E. coli* and *S. typhimurium* were purified essentially as described (2). The protein samples were oxidized with performic acid (11) or substituted by treatment with 4-vinyl-pyridine (8).

Edman degradations with the purified indoleglycerol phosphate synthetase were performed automatically with a Beckman sequencer (6). Phenylthiohydantoin-amino acids were identified by gas-liquid chromatography (18) and/or amino acid analysis after hydrolysis of phenylthiohydantoin-derivatives with HCl or HI (21). The procedures used are described in detail elsewhere (10, 14).

We are indebted to the Reilly Tar Company for the gift of 4-vinyl-pyridine.

The amino acid sequence of the first 40 residues of indoleglycerol phosphate synthetase of *E. coli* was deduced from six runs on the Beckman sequencer, and the results obtained are summarized in Table 1. It should be noted that the residues at positions 23, 33, and 38 were identified with some uncertainty. The Gln at 23 follows Gln at 21 and at 22 and although there was very little step overlap in our sequencer runs, it is difficult to be absolutely certain of this assignment. Similarly the relatively high background for Val and Ala at the late cycles makes the assignments of Val at 33 and Ala at 38 somewhat questionable, although probable.

Thirty-eight of the first 40 residues of the indoleglycerol phosphate synthetase of *S. typhimurium* were identified in three runs on the Beckman sequencer. The residue at position 23 was identified with some uncertainty and no assignment could be made at position 33 or 37.

The amino-terminal sequences of the first 40 residues of the indoleglycerol phosphate synthetases of *E. coli* and *S. typhimurium* are compared in Fig. 1. It is apparent that the partial sequences of the indoleglycerol phosphate synthetase of these two bacterial species are homologous, and, in fact, the sequences of the first 30 residues are identical. The extent of divergence of the amino-terminal sequences of indoleglycerol phosphate synthetase between these two bacterial species is analyzed in Table 2. For comparison the sequence divergence of anthranilate synthetase components I and II (14) and tryptophan synthetase α chains of the same bacterial species are also presented (9, 16). Among the 38 residues of indoleglycerol phos-

[1] Present address: Department of Microbiology, Mount Sinai School of Medicine, New York, N.Y. 10029.

TABLE 1. *Sequential degradation of indoleglycerol phosphate synthetase of E. coli and S. typhimurium*

Step[a] no.	Deduced residue	*E. coli*					Deduced residue	*S. typhimurium*		
		SP 400[b]		CFC[b]	121[c]			SP 400[b]		121[c]
		−S	+S		HCl	Hl		−S	+S	Hl
1	Met[d]	M	M				Met	M	M	
2	Gln	−[d]	Q				Gln	—	Q	
3	Thr	PT[d]	T				Thr	PT	T	
4	Val	V	V				Val	V	V	
5	Leu	LI[d]	L	L			Leu	LI	L	
6	Ala	A	A		A		Ala	A	A	
7	Lys	—	K		K		Lys	—	K	
8	Ile	LI	I	I			Ile	LI	I	
9	Val	V	V				Val	V	V	
10	Ala	A	A				Ala	A	A	
11	Asp	—	D		D		Asp	—	D	
12	Lys	—	K		K		Lys	—	K	
13	Ala	A	A				Ala	A	A	
14	Ile	LI	I	I			Ile	LI	I	
15	Trp	W	W				Trp	W	W	
16	Val	V	V				Val	V	V	
17	Glu	—	E				Glu	—	E	
18	Ala	A	A				Ala	A	A	
19	Arg	—	—		R		Arg	—	—	R
20	Lys	—	K		K		Lys	—	K	K
21	Gln	—	Q		E	E	Gln	—	Q	E
22	Gln	—	Q		E	E	Gln	—	Q	E
23	Gln?	—	Q?[d]		E?	E?	Gln?	—	Q?	E?
24	Pro	PT	P			P	Pro	PT	P	P
25	Leu	LI	L	L		L	Leu	LI	L	L
26	Ala	A	A		A	A	Ala	A	A	A
27	Ser	S	S		—	S[d]	Ser	S	S	S[d]
28	Phe	F	F		F	F	Phe	F	F	F
29	Gln	—	Q		E	E	Gln	—	Q	E
30	Asn	—	N		D	D	Asn	—	N	D
31	Glu	—	E		E	E	Glu	—	E	E
32	Val	V	V		V	V	Gly	G	G	G
33	Val?	V?	V?		V?	V?	?	?	?	?
34	Pro	PT	P		P	P	Pro	PT	P	P
35	Val	V	V		V	V	Val	V	V	V
36	Asn	—	N		D	D	Thr	PT	T	T[d]
37	Arg	—	—		R	R	?	?	?	?
38	Ala?	A?	A?		A?	A?	Gly	G	G	G
39	Phe	F	F		F	F	Phe	F	F	F
40	Tyr	Y	Y		Y	Y	Tyr	Y	Y	Y

[a] Average repetitive yield based on the stable amino acids was between 96 and 98%.

[b] Only phenylthiohydantoin-amino acids extracted with ethyl acetate after conversion were analyzed on columns of SP400 and CFC in gas-liquid chromatography. +S, Silylated derivatives; −S, unsilylated derivatives.

[c] Amino acid analysis after hydrolysis of phenylthiohydantoin-amino acids with 5.7 N HCl or 56% HI.

[d] Single and three-letter abbreviations are those of Dayhoff (3). −, No phenylthiohydanoin (PTH)-amino acids detected. LI, PTH-amino acids of leucine and isoleucine cannot be differentiated unambiguously; PT, PTH-amino acids of proline and threonine cannot be differentiated unambiguously. S and T, Identified as alanine and α-amino butyric acid after HI hydrolysis. ?, Uncertainty in identification.

phate synthetase compared three amino acid differences are observed, all of which can be explained by single base changes. It is evident that the indoleglycerol phosphate synthetase sequences compared indicate the same extent of divergence as was observed with the respective anthranilate synthetase component II, and both indoleglycerol phosphate synthetase and anthranilate synthetase component II show somewhat less variation than observed with anthra-

	5					10					15					20				
E. coli	Met	Gln	Thr	Val	Leu	Ala	Lys	Ile	Val	Ala	Asp	Lys	Ala	Ile	Trp	Val	Glu	Ala	Arg	Lys
S. typhimurium	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	25					30					35					40				
E. coli	Gln	Gln	Gln*	Pro	Leu	Ala	Ser	Phe	Gln	Asn	Glu	Val	Val*	Pro	Val	Asn	Arg	Ala*	Phe	Tyr
S. typhimurium	—	—	—*	—	—	—	—	—	—	—	—	Gly	?	—	—	Thr	?	Gly	—	—

FIG. 1. *The amino-terminal sequences of indoleglycerol phosphate synthetase of E. coli and S. typhimurium. *, Identification not certain; ?, no positive identification.*

TABLE 2. *Presumed minimum base differences in the genetic regions specifying the initial portions of anthranilate synthetase components I and II and indoleglycerol phosphate synthetase and the entire tryptophan synthetase α chains of E. coli and S. typhimirium*[a]

Determinants	ASase component I	ASase component II	InGPase	TSase α chain
No. of residues compared	25	51	38	268
No. of residue differences	4 (16.0%)	3 (5.9%)	3 (7.9%)	44 (16.4%)
AA differences explicable by:				
Single base change	4	2	3	43
Double base change	0	1	0	1
Minimum base differences	4 (5.3%)	4 (2.6%)	3 (2.6%)	45 (6.3%)

[a] ASase, Anthranilate synthetase; InGPase, indoleglycerol phosphate synthetase; TSase, tryptophan synthetase; AA, amino acid.

nilate synthetase component I and the tryptophan synthetase α chains. This result suggests that genes specifying polypeptides concerned with different reactions in the tryptophan biosynthetic pathway may have diverged to different extents; however, it should be noted that to date only the amino-terminal segments of three of the four polypeptides have been compared. It is possible that the positions of amino acid differences are not randomly distributed throughout all of these polypeptides. In this regard, in the tryptophan synthetase α chain of the two species, the positions of differences are distributed equally throughout the polypeptide, with the exception of the carboxyl-terminal region. It may be significant that the two *trp* operon polypeptides which we find to have diverged the least are bifunctional in both *E. coli* and *S. typhimurium* and are each represented in some bacterial species by two separate polypeptide chains.

We are greatly indebted to Susan Stasiowski for technical assistance.

These studies were supported by National Science Foundation grant GB-36967X and United States Public Health Service grant GM-09738 from the National Institute of General Medical Sciences. C.Y. is a Career Investigator of the American Heart Association.

LITERATURE CITED

1. Blume, A. J., and E. Balbinder. 1966. The tryptophan operon of *Salmonella typhimurium*. Fine structure analysis by deletion mapping and abortive transduction. Genetics **53**:577-592.
2. Creighton, T. E. 1970. N-(5'-phosphoribosyl) anthranilate isomerase-indole-3-glycerol phosphate synthetase of tryptophan biosynthesis: relationship between the two activities of the enzyme from *Escherichia coli*. Biochem. J. **120**:699-707.
3. Dayhoff, M. O. 1972. Atlas of protein sequence and structure. National Biomedical Research Foundation, Silver Spring, Md.
4. Demerec, M., and K. New. 1965. Genetic divergence in *Salmonella typhimurium*, *S. montevideo* and *Escherichia coli*. Biochem. Biophys. Res. Commun. **18**:652-655.
5. Denney, R. M., and C. Yanofsky. 1972. Detection of tryptophan messenger RNA in several bacterial species and examination of the properties of heterologous DNA-RNA hybrids. J. Mol. Biol. **64**:319-339.
6. Edman, P., and G. Begg. 1967. A protein sequenator. Eur. J. Biochem. **1**:80-91.
7. Enatsu, T., and I. P. Crawford. 1968. Enzymes of the tryptophan synthetic pathway in *Pseudomonas putida*. J. Bacteriol. **95**:107-112.
8. Friedman, M., L. H. Krull, and J. F. Cavins. 1970. The chromatographic determination of cystine and cysteine residues in proteins as S-β-(4-pyridylethyl) cysteine. J. Biol. Chem. **245**:3868-3871.
9. Guest, J. R., G. R. Drapeau, B. C. Carlton, and C. Yanofsky. 1967. The amino acid sequence of the A protein (α subunit) of the tryptophan synthetase of *Escherichia coli*. V. Order of tryptic peptides and the complete amino acid sequence. J. Biol. Chem. **242**:5442-5446.
10. Hermodson, M. A., L. H. Ericsson, K. Titani, H. Neurath, and K. A. Walsh. 1972. Application of sequenator analyses to the study of proteins. Biochemistry **11**:4493-4501.
11. Hirs, C. H. W. 1956. The oxidation of ribonuclease with performic acid. J. Biol. Chem. **219**:611-621.
12. Hoch, S. O., C. Anagnostopoulos, and I. P. Crawford. 1969. Enzymes of the tryptophan operon of *Bacillus subtilis*. Biochem. Biophys. Res. Commun. **35**:838-844.
13. Li, S. L., R. M. Denney, and C. Yanofsky. 1973. Nucleotide sequence divergence in the α-chain-structural genes of tryptophan synthetase from *Escherichia coli*, *Salmonella typhimurium*, and *Aerobacter aerogenes*. Proc. Natl. Acad. Sci. U.S.A. **70**:1112-1116.

14. Li, S. L., J. Hanlon, and C. Yanofsky. 1974. Separation of anthranilate synthetase components I and II of *Escherichia coli*, *Salmonella typhimurium*, and *Serratia marcescens* and determination of their amino-terminal sequences by automatic Edman degradation. Biochemistry **13:**1736–1744.
15. Li, S. L., and C. Yanofsky. 1972. Amino acid sequences of fifty residues from the amino termini of the tryptophan synthetase α chains of several Enterobacteria. J. Biol. Chem. **247:**1031–1037.
16. Li, S. L., and C. Yanofsky. 1973. Amino acid sequence studies with the tryptophan synthetase α chain of *Salmonella typhimurium*. J. Biol. Chem. **248:**1830–1836.
17. McQuade, J. F., III, and T. E. Creighton. 1970. Purification and comparison of the N-(5′-phosphoribosyl) anthranilic acid isomerase/indole-3-glycerol phosphate synthetase of tryptophan biosynthesis from three species of Enterobacteriaceae. Eur. J. Biochem. **16:**199–207.
18. Pisano, J. J., T. J. Bronzert, and H. B. Brewer, Jr. 1972. Advances in the gas chromatographic analysis of amino acid phenyl- and methylthiohydantoins. Anal. Biochem. **45:**43–59.
19. Sawula, R. V., and I. P. Crawford. 1972. Mapping of the tryptophan genes of *Acinetobacter calcoaceticus* by transformation. J. Bacteriol. **112:**797–805.
20. Smith, O. H. 1967. Structure of the *trpC* cistron specifying indole-glycerol phosphate synthetase, and its location in the tryptophan operon of *Escherichia coli*. Genetics **57:**95–105.
21. Smithies, O., D. Gibson, E. M. Fanning, R. M. Goodfliesh, J. G. Gilman, and D. L. Ballantyne. 1971. Quantitative procedures for use with the Edman-Begg sequenator. Partial sequences of two unusual immunoglobulin light chains, Rzf and Sac. Biochemistry **10:**4912–4921.
22. Wegman, J., and I. P. Crawford. 1968. Tryptophan synthetic pathway and its regulation in *Chromobacterium violaceum*. J. Bacteriol. **95:**2325–2335.
23. Yanofsky, C. 1971. Tryptophan biosynthesis in *Escherichia coli:* genetic determination of the proteins involved. J. Am. Med. Assoc. **218:**1026–1035.
24. Yanofsky, C., V. Horn, M. Bonner, and S. Stasiowski. 1971. Polarity and enzyme functions in mutants of the first three genes of the tryptophan operon of *Escherichia coli*. Genetics **69:**409–433.