Stifel, F. B., Herman, R. H., and Rosensweig, N. S. (1970b), *Biochim. Biophys. Acta 222*, 65.
Susor, W. A., and Rutter, W. J. (1971), *Anal. Biochem. 43*, 147.
Szepesi, B., and Freedland, R. A. (1969), *Proc. Soc. Exp. Biol. Med. 132*, 489.
Tanaka, T., Harano, Y., Morimura, H., and Mori, R. (1965), *Biochem. Biophys. Res. Commun. 21*, 55.
Tanaka, T., Harano, Y., Sue, F., and Morimura, H. (1967), *J. Biochem. (Tokyo) 62*, 71.
Taylor, C. B., and Bailey, E. (1967), *Biochem. J. 102*, 32c.
Tyrrell, J. B., and Anderson, J. W. (1971), *Endocrinology 89*, 1178.
Weber, G., Lea, M. A., Convery, H. J. H., and Stamm, N. B. (1967), *Advan. Enzyme Regul. 5*, 257.
Wilkinson, H. J. (1970), Isoenzymes, 2nd ed, Philadelphia, Pa.. J. B. Lippincott Co.

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†

Shoei-Lung Li, Joan Hanlon, and Charles Yanofsky*

ABSTRACT: Amino-terminal sequences are presented for polypeptide components I and II of the anthranilate synthetase of *Escherichia coli*, *Salmonella typhimurium*, and *Serratia marcescens*. The respective amino-terminal sequences are homologous in all three species despite the fact that component II of *E. coli* and *S. typhimurium* is three times as large as component II of *S. marcescens*, *ca.* 60,000–65,000 *vs. ca.* 21,000 and that component II of the former species catalyzes the second or anthranilate phosphoribosyltransferase reaction of tryptophan biosynthesis while component II of *S. marcescens* does not.

Anthranilate synthetase (ASase)[1] catalyzes the initial reaction unique to tryptophan biosynthesis: chorismate + L-glutamine → anthranilate + L-glutamate + pyruvate (Zalkin, 1973). In all bacterial species studied to date ASase exists as an enzyme complex containing two nonidentical polypeptide chains (Zalkin, 1973) designated components I and II (Ito and Yanofsky, 1966). Two general types of ASase complex have been observed in bacteria. One type also catalyzes the second or anthranilate 5-phosphoribosylpyrophosphate phosphoribosyltransferase (PRTase) reaction of tryptophan biosynthesis: anthranilate + 5-phosphoribosyl-1-pyrophosphate → 5-phosphoribosylanthranilate + pyrophosphate. Such complexes, present in *Escherichia coli* (Ito and Yanofsky, 1966, 1969; Ito *et al.*, 1969), *Salmonella typhimurium* (Bauerle and Margolin, 1966; Henderson *et al.*, 1970a,b; Nagano and Zalkin, 1970; Hwang and Zalkin, 1971; Henderson and Zalkin, 1971), and *Aerobacter aerogenes* (Egan and Gibson, 1966), are composed of polypeptides of similar size. The second type of ASase complex, present in *Serratia marcescens* (Robb and Belser, 1972; Zalkin and Hwang, 1971; Robb *et al.*, 1971), *Bacillus subtilis* (Kane and Jensen, 1970; Kane *et al.*, 1972), *Acinetobacter calcoaceticus* (Sawula and Crawford, 1972, 1973), *Chromobacterium violaceum* (Wegman and Crawford, 1968), and various *Pseudomonas* species (Queener and Gunsalus, 1970; Queener *et al.*, 1973) does not have PRTase activity and consists of two polypeptides of very different sizes. Component I of both types of complex can convert chorismate to anthranilate in the absence of component II when L-glutamine is replaced by ammonia as amino donor, *i.e.*, component II provides the glutamine amidotransferase function which allows glutamine to serve as the amino source in anthranilate formation. In *E. coli* and organisms with a similar ASase complex, component II also provides the PRTase activity of the complex. Genetic and biochemical studies have shown that the amino-terminal third of component II of *E. coli* (Yanofsky *et al.*, 1971) and *S. typhimurium* (Hwang and Zalkin, 1971; Grieshaber and Baurele, 1972) provides the glutamine amidotransferase function while the carboxyl-terminal two-thirds of the polypeptide is sufficient to perform the PRTase reaction (Jackson and Yanofsky, 1974). Component II of these organisms has a mol wt of 60,000–65,000 (Ito *et al.*, 1969; Henderson and Zalkin, 1971). By contrast component II's of *S. marcescens* and organisms with the second type of ASase have molecular weights of 14,000–21,000 (Kane and Jensen, 1970; Queener and Gunsalus, 1970; Zalkin and Hwang, 1971; Robb and Belser, 1972; Queener *et al.*, 1973; Sawula and Crawford, 1973) and lack PRTase activity. In view of the differences between the ASase complexes of the bacterial species mentioned, it was of interest to examine the primary structures of the respective components for evidence of homology. In this paper we describe the isolation and separation of ASase components I and II of *E. coli*, *S. typhimurium*, and *S. marcescens* and amino-terminal sequence determinations by automatic Edman degradation. A preliminary report of the

† From the Department of Biological Sciences, Stanford University, Stanford, California 94305. *Received November 28, 1973*. These studies were supported by grants from the National Science Foundation (GB-36967) and the U. S. Public Health Service (GM09738). C. Y. is a Career Investigator of the American Heart Association.

[1] Abbreviations used are: ASase, anthranilate synthetase; PRTase, anthranilate 5-phosphoribosylpyrophosphate phosphoribosyltransferase; TSase, tryptophan synthetase; Quadrol, *N,N,N',N'*-tetrakis(2-hydroxypropyl)ethylenediamine; PhNCS, phenylthiohydantoin.

component II sequences will appear elsewhere (Li *et al.*, 1974).

Experimental Procedures

Growth of Bacteria. Cultures (100 l.) of *E. coli* strain *trpA2/ColVB trpA2*, *S. typhimurium* strain St-13, or *S. marcescens* strain SM6-*trpA3* were grown in single strength minimal medium (Vogel and Bonner, 1956) containing 3.5 μg of indole/ml, 0.01% acid-hydrolyzed casein, and 0.5% glucose in a New Brunswick Scientific Co. Fermacel until the indole was exhausted and the cells were derepressed. The cells were collected by centrifugation in a Sharples centrifuge and suspended by blending in 1.5 vol of buffer containing 0.05 M potassium phosphate (pH 7.5), 0.1 mM EDTA, 0.1 mM dithiothreitol, 0.05 mM L-tryptophan, and 20% glycerol by volume (buffer A). The cells were disrupted in a Manton-Gaulin homogenizer. The suspension was centrifuged for 90 min at 30,000*g* in a Sorvall RC2B centrifuge and the clarified extract was retained.

Protein Purification. ANTHRANILATE SYNTHETASE OF *E. coli*. The procedure employed is based on those of Hwang and Zalkin (1971) and Ito *et al.* (1969). All steps were performed at 4°.

Step 1. The protein concentration of the cell extract was determined (Lowry *et al.*, 1951) and adjusted to 15–20 mg/ml with buffer A. A 2% solution of protamine sulfate in buffer A was added to give a final concentration of 0.17 mg of protamine sulfate/mg of extract protein. The mixture was stirred for 10 min and the precipitate removed by centrifugation for 15 min at 30,000*g*.

Step 2. Solid ammonium sulfate was added (25.3 g/100 ml of supernatant) and the mixture stirred for 40 min. The precipitate was collected by centrifugation for 40 min at 30,000*g*.

Step 3. The precipitate from step 2 was resuspended in 0.1 M potassium phosphate (pH 7.4) containing 0.1 mM dithiothreitol and 1 mM $MgCl_2$ at a protein concentration of 25–30 mg/ml. A lipase solution (25 mg/ml) in the same buffer was added to the extract to give a final concentration of 0.1 mg of lipase/mg of protein. (The proteolytic activity of the crude lipase preparation is responsible for the action on the anthranilate synthetase complex (Hwang and Zalkin, 1971).) The mixture was incubated for 30 min at 37°. The small precipitate which formed was removed by centrifugation. Glycerol was added to 15% by volume, and the preparation frozen for convenience.

Step 4. If the volume of the thawed preparation was greater than 120 ml, the protein was concentrated by precipitation with ammonium sulfate (33 g/100 ml). The protein solution was dialyzed for a minimum of 2 hr against 0.05 M potassium phosphate (pH 7.0), 0.1 mM dithiothreitol, 0.1 mM EDTA, and 15% by volume glycerol (buffer B) and applied to a 7 × 50 cm column of Bio-Rad P-100 equilibrated and run with the same buffer. The column fractions were assayed and the active fractions pooled and precipitated with ammonium sulfate (33 g/100 ml), and the precipitate dissolved in the minimal volume of buffer B.

Step 5. The final fraction from step 4 was dialyzed against buffer B for 6 hr with three changes (1-l. volumes). The dialyzed sample of up to 40 ml was loaded onto a DEAE-cellulose column (2.3 × 120 cm) equilibrated with buffer B, and was eluted with a 1-l. linear gradient from 0.1 to 0.5 M KCl in buffer B. High specific activity fractions were pooled and precipitated with ammonium sulfate (33 g/100 ml), and the precipitate collected by centrifugation. The precipitate was dissolved in 0.05 M potassium phosphate (pH 7.0), 0.1 mM dithiothreitol, 0.1 mM EDTA, and 10% glycerol by volume and applied to and eluted from an Agarose column (2.8 × 100 cm, Bio-Rad 1.5 m) equilibrated with the same buffer. High specific activity fractions were pooled and precipitated with ammonium sulfate (33 g/100 ml) and the precipitate collected by centrifugation. The precipitate was dissolved in 0.01 N NH_4OH.

Step 6. The final fraction from step 5 was dialyzed overnight against 0.01 N NH_4OH, and 99% formic acid was added to the sample to 30% by volume. The sample containing *ca.* 50 mg in 4 ml was applied to a Sephadex G-100 column (2.2 × 45 cm) equilibrated and eluted with 30% formic acid containing 1 mM β-mercaptoethanol. Absorbance was measured at 280 nm and selected fractions were taken for examination by gel electrophoresis.

ANTHRANILATE SYNTHETASE OF *S. typhimurium*. Purification was performed exactly as with *E. coli* extracts except that protamine sulfate addition at step 1 was omitted.

ANTHRANILATE SYNTHETASE OF *S. marcescens*. Steps 1–5 were performed as described above except that the lipase treatment in step 3 was omitted. The final fraction from step 5 was dialyzed overnight against 0.01 N NH_4OH and lyophilized. The protein was treated with 4-vinylpyridine and dialyzed against 0.01 M acetic acid as described (Friedman *et al.*, 1970; Hermodsen *et al.*, 1972). The sample was lyophilized and the powder dissolved in 60% formic acid containing 1 mM β-mercaptoethanol. The sample (20 mg in 3 ml) was loaded onto a column of Sephadex G-100 (2.2 × 45 cm) equilibrated and run with 60% formic acid containing 1 mM β-mercaptoethanol. Absorbance of fractions was measured at 280 nm and selected fractions were examined for purity by gel electrophoresis.

Gel Electrophoresis. Polyacrylamide gel electrophoresis in sodium dodecyl sulfate was performed as described (Weber and Osborn, 1969). The reference proteins used were sperm whale apomyoglobin (mol wt 17,199) and the tryptophan synthetase (TSase) α chain of *E. coli* (mol wt 28,722). Gels were stained with Coomassie Brilliant Blue, destained, and scanned with a densitometer at a wavelength of 600 nm.

Automatic Edman Degradation. Edman degradations were performed automatically using a Beckman Sequencer Model 890 with either the regular or the "undercut" reaction cup (Edman and Begg, 1967). The coupling buffer was either 1 M Quadrol or *N,N*-dimethylallylamine, and acid cleavage was performed once or twice. The programs used in sequence determinations have the following Beckman designations: Regular Quadrol D-X10, Fast Quadrol 072172, Regular *N,N*-dimethylallylamine 032671, and Fast *N,N*-dimethylallylamine 090872. Nitrogen flush was employed at fine vacuum steps. The proteins were prepared for sequence analysis by dialysis overnight against 1 mM NH_4OH and then for 6–8 hr against deionized water, and lyophilized. Cysteine and cystine residues of the proteins were occasionally oxidized with performic acid (Hirs, 1956) or reduced and allowed to react with 4-vinylpyridine (Friedman *et al.*, 1970; Hermodsen *et al.*, 1972). A lyophilized sample of native or modified protein was transferred into the reaction cup, and a small amount of HFBA was delivered to dissolve the protein. The sample was distributed on the wall of the spinning cup at low speed and dried through restricted vacuum, rough vacuum, nitrogen flush, and fine vacuum steps. In order to remove all the HFBA from the delivery tube, protein film, and reaction chamber, the protein was extracted with butyl chloride and then dried under fine vacuum for several hours before starting the auto-

TABLE I: Procedures Employed in the Identification of PhNCS Amino Acids.

	Gas–Liquid Chromatography[b]						
	SP 400[c] Silylation				AA Anal. 121[d]		
>PhNCS[a]	None	On Column	Off Column	CFC[c]	HCl	HI	Tlc
Ala(A)	A[e]	A	A		A	A	
Val(V)	V	V	V		V	V	
Met(M)	M	M	M		M	–	
Phe(F)	F	F	F		F	F	
Tyr(Y)	Y	Y	Y		Y	Y	
Trp(W)	W	W	W		–	W*[g]	
Leu(L)	LI	L	L	L	L	L	
Ile(I)	LI	I	I	I	I	I	
Ser(S)	S	S	S		–	S*[g]	
Gly(G)	G	G	G		G	G	
Thr(T)	TP	T	T		–	T*[g]	T
Pro(P)	TP	P	P		P	P	P
Asp(D)	–	D	D		DN	DN	D
Asn(N)	(N)[f]	N	N		DN	DN	N
Glu(E)	–	E	E		EQ	EQ	E
Gln(Q)	(Q)[f]	Q	Q		EQ	EQ	Q
Lys(K)	(K)[f]	(K)[f]	K		K	K	
His(H)	–	–	–		H	H	
Arg(R)	–	–	–		R	R	
Cys(C)	–	–	–		C	C	

[a] The single and three letter abbreviations for the amino acids are those of Dayhoff (1972). [b] Only those PhNCS amino acids extracted with ethyl acetate following conversion were analyzed by gas–liquid chromatography. [c] SP400 and CFC are the supports present in the column employed in gas–liquid chromatography. [d] 121 indicates amino acid analysis of HCl- or HI-hydrolyzed PhNCS derivatives using a Beckman Model 121 analyzer. [e] Where a PhNCS derivative was identified unambiguously by a particular procedure, the single letter abbreviation is listed. Where we could not distinguish between two PhNCS amino acids the single letter abbreviations of both are given. A dash (–) indicates that the sample was analyzed but no PhNCS derivative was detected. If there is no letter or dash the analysis was not performed. [f] (N), (Q), and (K) indicate that identification of Asn-, Gln-, or Lys>PhNCS was difficult because of low yields. [g] After HI hydrolysis Ser- and Thr>PhNCS were identified as alanine and α-aminobutyric acid, respectively, and Trp>PhNCS was recovered as glycine and alanine (Smithies *et al.*, 1971).

matic program. This was extremely important when *N,N*-dimethylallylamine buffer was used. The above procedure is essentially the sample preroutine program recommended by Beckman Instruments. It was our experience that the fast Quadrol program with the "undercut" reaction cup gave results comparable to those obtained with the regular Quadrol program. For example, 50 residues of the TSase α chain (mol wt 29,000) and 60 residues of ASase component II (mol wt ~23,000) were routinely sequenced using either program. Background levels of PhNCS amino acids, presumably due to nonspecific cleavage, were proportional to the size of the protein being degraded. (The background level is one of major factors which limits the number of residues which can be sequenced.) Only 25 residues could be identified in sequence analyses with ASase component I, a protein with a mol wt of about 62,000. With this and other large proteins the *N,N*-dimethylallylamine single cleavage program gave better results than did the regular Quadrol double cleavage program. Single cleavage probably results in less nonspecific cleavage thereby reducing the background. In addition, the *N,N*-dimethylallylamine program does not contain an ethyl acetate wash; thus, complications caused by separation of the protein film from the wall of the reaction cup and by Quadrol retention are avoided. However, when *N,N*-dimethylallylamine buffer is used histidine residues present difficulties because the histidinyl bond is very unstable in *N,N*-dimethylallylamine buffer and partial cleavage of the polypeptide chain is observed one step earlier than expected (Thomsen *et al.*, 1972; Li and Yanofsky, 1973a). Substitution of cysteine and cystine residues with 4-vinylpyridine (Friedman *et al.*, 1970; Hermodsen *et al.*, 1972) increased the solubility of the protein sample in the coupling buffers (Quadrol and *N,N*-dimethylallylamine) and the cleavage acid, resulting in better adhesion of the protein to the wall of the reaction cup. The Quadrol program could then be used successfully with large proteins.

Identification of PhNCS Amino Acids. Conversion of the thiazolinones to phenylthiohydantoins (PhNCS) and their subsequent extraction were carried out as described by Edman and Begg (1967). Most of the silylated and nonsilylated PhNCS derivatives were identified by gas–liquid chromatography using a Beckman GC-45 and/or a Hewlett-Packard Model 7610 equipped with an automatic sample injection system. The procedures employed were essentially those of Pisano and Bronzert (1969). The glass columns (2 mm i.d. × 4 ft) were filled with 10% SP400. A temperature shift from 160 to 280° was programmed, either linearly for the GC-45 or stepwise for the H-P 7610. One aliquot (4%) was injected onto a SP400 column without silylation. Another aliquot was silylated either "on column" (4%) (Li and Yanofsky, 1972) or "off column" (20%) (Hermodsen *et al.*, 1972) and suitable fractions injected. These procedures allowed the identification of most of the PhNCS derivatives. Figure 1 shows the relative

TABLE II: Sequential Edman Degradation of Anthranilate Synthetase Component I of *E. coli*, *S. typhimurium*, and *S. marcescens*.

Step No.	*E. coli*[a] Deduced Residue	SP400 −S[b]	SP400 +S[b]	121 HCl	121 HI	Tlc	*S. typhimurium*[a] Deduced Residue	SP400 −S[b]	SP400 +S[b]	121 HI	*S. marcescens*[a] Deduced Residue	SP400 −S[b]	SP400 +S[b]	121 HI
1	Met	M	M		–		Met	M	M		Met	M	M	–
2	Gln	–	Q		E		Gln	–	Q		Asn	–	N	D
3	Thr	PT[c]	T		T*[d]	T	Thr	PT	T	T*	Thr	PT	T	T*
4	Gln	–	Q		E		Pro	PT	P	P	Lys	–	K	K
5	Lys	–	K		K		Lys	–	K	K	Pro	PT	P	P
6	Pro	PT	P	P	P	P	Pro	PT	P	P	His	–	–	H
7	Thr	PT	T	–	T*	T	Thr	PT	T	T*	Leu	LI	L	L
8	Leu	LI[c]	L	L	L		Leu	LI	L		Thr	PT	T	T*
9	Glu	–	E		E		Glu	–	E		Leu	LI	L	L
10	Leu	LI	L		L		Leu	LI	L		Leu	LI	L	L
11	Leu	LI	L		L		Leu	LI	L		Lys	–	K	K
12	Thr	PT	T	–	T*	T	Thr	PT	T	T*	Val	V	V	V
13	Ser	S	S	–	S*[d]	S	Ser	S	S	S*	Gln	–	Q	E
14	Glu	–	E	E	E		Asp	–	D	D	Ala	A	A	A
15	Gly	G	G		G		Ala	A	A		Ser	S	S	S*
16	Ala	A	A	A	A		Ala	A	A		Tyr	Y	Y	Y
17	Tyr	Y	Y	Y	Y		Tyr	Y	Y		Arg	–	–	R
18	Arg	–	–	R	R		Arg	–	–	R	Gly	G	G	G
19	Asp	–	D		D		Glu	–	E	E	Asn	–	N	D
20	Asn	–	N		D		Asn	–	N	D	Pro	PT	P	P
21	Pro	PT	P	P	P		Pro	PT	P	P	Thr	PT	T	T*
22	Thr	PT	T		T*		Thr	PT	T	T*	Thr	PT	T	T*
23	Ala	A	A		A		Ala	A	A	A	Leu	LI	L	L
24	Leu	LI	L		L		Leu	LI	L	L	Phe	F	F	F
25	Phe	F	F		F		Phe	F	F		His	–	–	H

[a] For an explanation of other abbreviations see the legend to Table I. The *E. coli* summary represents data from seven degradation runs. Five were performed with the regular Quadrol program and two were performed with the fast *N,N*-dimethylallylamine program. The *S. typhimurium* data were obtained in a single degradation run with the fast *N,N*-dimethylallylamine program. The *S. marcescens* data were obtained in two runs, one with the fast *N,N*-dimethylallylamine program and the other with the fast Quadrol program. [b] +S and −S indicate the analyses of silylated and unsilylated PhNCS derivatives, respectively. [c] PT and LI indicate coelution of the unsilylated PhNCS derivatives of proline and threonine and leucine and isoleucine, respectively. [d] S* and T* indicate that Ser- and Thr->PhNCS were identified as alanine and α-aminobutyric acid after HI hydrolysis.

retention times of silylated and nonsilylated PhNCS derivatives on SP400 columns in the Beckman GC-45. The yield of PhNCS derivatives was estimated by comparing either the peak heights or the peak areas with those of standard PhNCS amino acids.

The remainder of the ethyl acetate solution of each sample and the aqueous layer from each step were hydrolyzed separately with either 5.7 N HCl (Li and Yanofsky, 1972) or 56% HI (Smithies *et al.*, 1971) in evacuated sealed tubes at 130° in a "Temp-Block" heater for 20 hr. Hydrolysis tubes were cracked, and samples dried in a vacuum dessicator over NaOH pellets. Free amino acids were analyzed with a Beckman Model 121 automatic analyzer.

As an aid to others engaged in automatic sequence analyses we have summarized our experiences with the various identification procedures for PhNCS amino acids (Table I). Ala-, Val-, Met-, Phe-, Tyr-, Trp-, Leu-, and Ile>PhNCS could be unequivocally identified by gas chromatography. Ser>PhNCS was difficult to detect by gas chromatography because of its instability and the appearance of multiple silylated peaks. Identification of Ser>PhNCS was straightforward by amino acid analysis following conversion to alanine by HI hydrolysis (Smithies *et al.*, 1971). Gly>PhNCS generally could be identified by gas chromatography. It was silylated poorly if at all

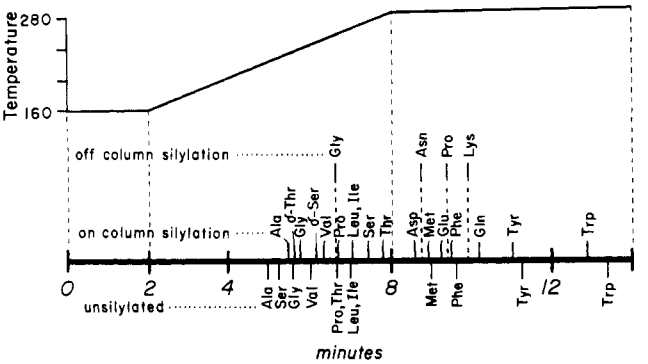
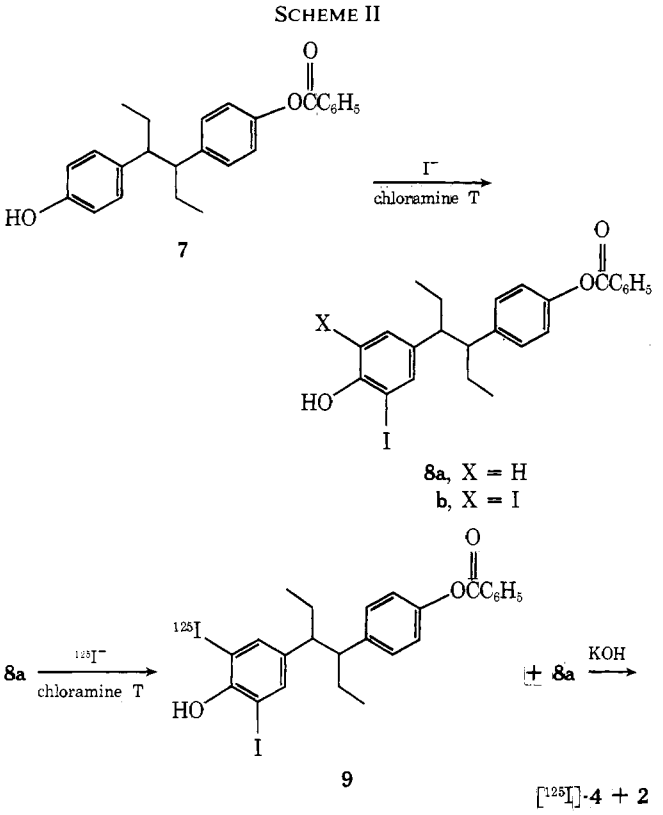


FIGURE 1: Relative retention times of silylated and unsilylated PhNCS amino acid derivatives. Glass columns (4 ft × 2 mm i.d.) of a Beckman GC-45 were filled with 10% SP400. The column temperature was programmed as indicated in the figure: 160° for 2 min, linear increase of 120° in 6 min, held at 280° for 6 min. The temperatures at the inlet, detector line, and detector were 280, 300, and 310°, respectively; make-up gas flows of He, H_2, and air were 50, 60, and 300 cm³/min, respectively, and the carrier He flow was 70 cm³/min. The position of some unsilylated PhNCS derivatives (Asn-, Gln-, and His>PhNCS) are not indicated in the figure because of the low yields generally obtained. Only those PhNCS amino acids which have different retention times following "off column silylation" are given for off column silylation.

SCHEME II



estrol by the iodide–chloramine T method gave the usual mixture of iodinated derivatives. Up to 40% of [^{3}H]monoiodohexestrol ([^{3}H]-**2**) with radiochemical purity in excess of 95% and smaller amounts of the ^{3}H-labeled derivatives **3–6** could be obtained from this reaction by liquid chromatography. Radiochemical purity of [^{3}H]-**2** was again established by cocrystallization and TLC.

Ultraviolet Spectra of Iodinated p-Cresol and Hexestrol Derivatives. The ultraviolet spectra of hexestrol and the iodohexestrol derivatives **2–6**, showing the B-band region, are given in Figure 3. The relationship between structure and spectral appearance can be appreciated by considering first the spectra of *p*-cresol (A), 2-iodo-*p*-cresol (B), and 2,6-diiodo-*p*-cresol (C), which illustrate the absorbance of the parent chromophores. Successive iodine substitution causes changes consistent with the introduction of electron donating substituents onto the *p*-cresol chromophore (Pasto and Johnson, 1969). The λ_{max} of the B band shifts to longer wavelengths, and a shoulder on the long wavelength side becomes more prominent, developing into a second maxima in C; intensification and bathochromic shifts in the E-band region (ca. 200 nm) causes a bathochromic shift and an elevation of the trough with region 240–260 nm. Furthermore, a long wavelength tail develops, especially in C, where it extends to 340–350 nm. The spectra of the compounds **1–6** can be considered simply as combinations of the chromophore types A, B, and C.

The Solution Photochemistry of I-Hex and I$_2$u-Hex. A series of studies by Wolf and Kharash (1965) established that the photolysis of *o*-iodophenols at 254 nm resulted in deiodination, as evidenced by the production of free iodine and the formation of products resulting from reaction with solvent, reduction in alcohols, and phenylation in benzene. These reactions are presumed to be initiated by a photohomolytic scission of the carbon–iodine bond.

We have investigated the photochemistry of I-Hex (**2**)

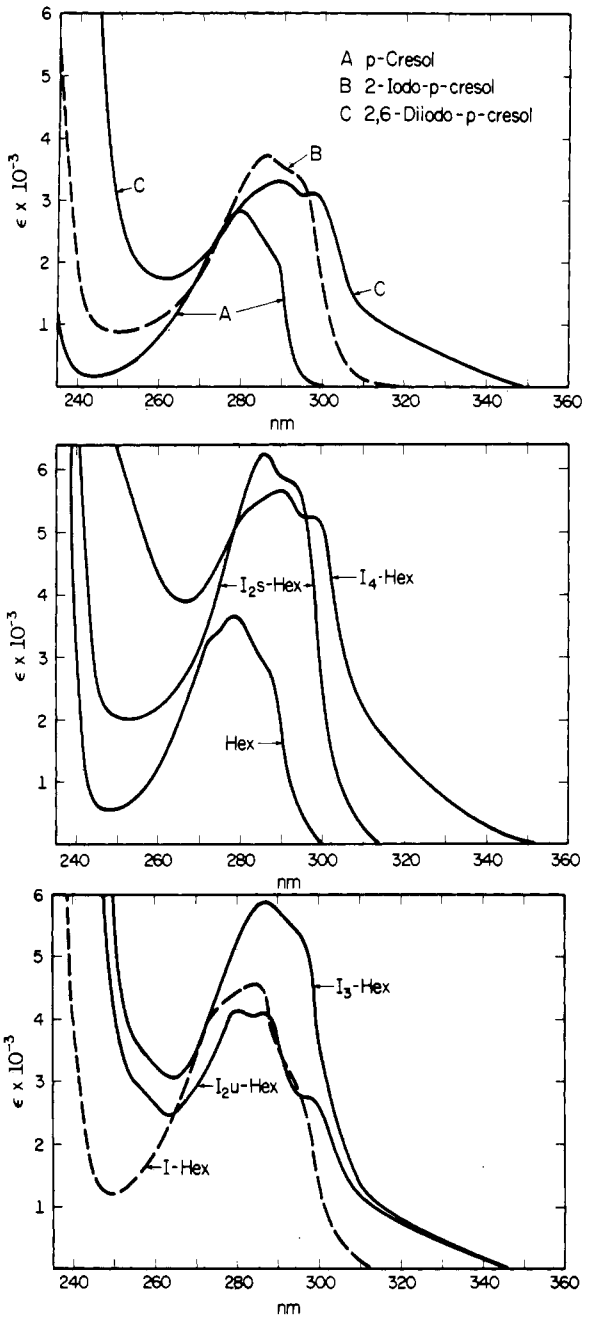


FIGURE 3: Ultraviolet spectra of iodinated hexestrols and *p*-cresols. The ultraviolet spectra of *p*-cresol (A), 2-iodo-*p*-cresol (B), 2,6-diiodo-*p*-cresol (C), and the hexestrols (**1–6**) were measured in 95% ethanol at 50–200 μM.

and I$_2$u-Hex (**4**), being illustrative of the reaction of chromophores B and C, in methanol and benzene at 254 and >315 nm (Scheme III). Irradiation of I$_2$u-Hex at 254 nm in methanol causes rapid reduction to I-Hex; further reduction to Hex proceeded more slowly, as expected on the basis of the relative molar absorptivities of I$_2$u-Hex and I-Hex at 254 nm (cf. Figure 3). Irradiation at >315 nm permitted selective conversion of I$_2$u-Hex into I-Hex, without significant further reduction. Up to 30% of phenylhexestrol (**10**) could be isolated (as the dibenzoate) upon irradiation of I-Hex in benzene at 254 nm.

The Photoreactivity of Iodinated Hexestrols with the Estrogen Binding Protein from Rat Uterus. We have recently described a general method for detecting photolytic reactions of chromophoric ligands bound to the estrogen

FOOTNOTES TO TABLE III

[a] For an explanation of the other abbreviations see the legends to Tables I and II. The *E. coli* summary represents data from four runs, one with the regular Quadrol program, two with the fast *N,N*-dimethylallylamine program, and one with the fast Quadrol program. The data for *S. typhimurium* and *S. marcescens* were from single runs with the fast Quadrol program. [b] W* indicates that Trp>PhNCS was recovered as glycine and alanine after HI hydrolysis. [c] No amino acid could be identified.

by "on column" silylation and appeared between δ-Thr>PhNCS and δ-Ser>PhNCS. Following "off column" silylation most of the Gly>PhNCS was silylated and appeared at the position of nonsilylated Pro>PhNCS. Gly>PhNCS was positively identified as free glycine following acid hydrolysis. Thr>PhNCS and Pro>PhNCS elute at the same position on SP400. Following "on column" silylation multiple peaks of Thr>PhNCS were observed while Pro>PhNCS was not silylated. Following "off column" silylation, silylated Pro>PhNCS appeared between silylated Glu>PhNCS and Phe>PhNCS. Thr>PhNCS was positively identified by conversion to α-aminobutyric acid by HI hydrolysis. The addition of thiols to the chlorobutane increased the recovery of Ser>PhNCS and Thr>PhNCS. In our experience dithioerythritol (15 mg/l.) was superior to ethanethiol, 1,4-butanedithiol, or β-mercaptoethyl ether. HI hydrolysis of PhNCS derivatives for 20 hr at 130° gave the highest yields of alanine and α-aminobutyric acid from Ser>PhNCS and Thr>PhNCS, respectively, although the yield from some PhNCS amino acids such as Val>PhNCS was relatively low. Asp>PhNCS, >AsnPhNCS, Glu>PhNCS, and Gln>PhNCS were readily identified when silylated. However, since Asp>PhNCS and Glu>PhNCS were formed from their amides during the conversion step (we observed about 50% deamination when the conversion was carried out at 80° for 10 min) it was essential that the presence of the amides be excluded in distinguishing between acid and amide. The extent of deamination was reduced appreciably if conversion was performed at 80° for 3 min (Dr. Jack Ohms, personal communication). Also, as SP400 col-

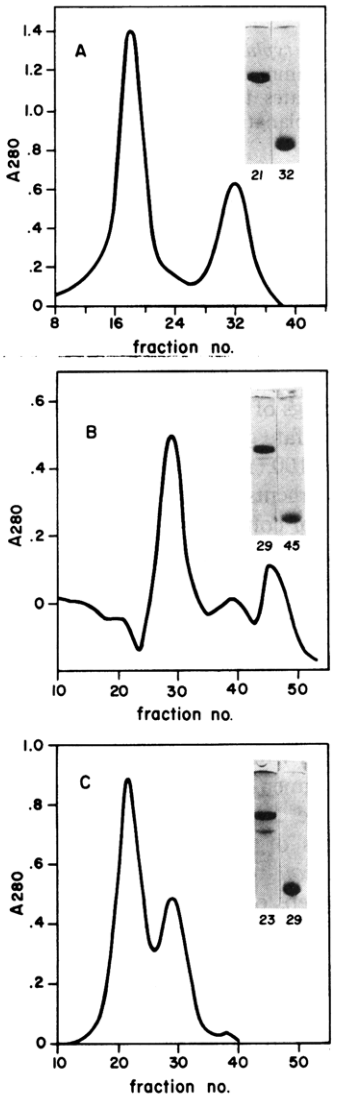


FIGURE 2: Separation of anthranilate synthetase components on Sephadex G-100. Samples were eluted with 30 or 60% formic acid, containing β-mercaptoethanol as described under Experimental Procedures. The sodium dodecyl sulfate gel patterns of selected fractions are indicated in the inset (run from top to bottom as pictured: (A) separation of the component I and component II fragments of lipase-treated ASase of *E. coli* (54 mg loaded in 4 ml); (B) separation of the component I and component II fragments of lipase-treated ASase of *S. typhimurium* (20 mg loaded in 3 ml); (C) separation of vinylpyridine-treated ASase components I and II of *S. marcescens* (50 mg loaded in 4 ml).

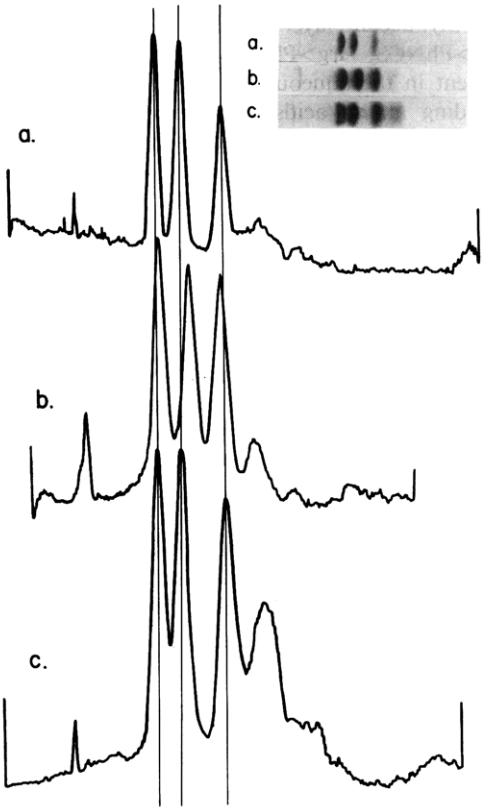


FIGURE 3: Sodium dodecyl sulfate gel electrophoresis of ASase component I and component II fragments from various sources. Electrophoretic gel patterns and densitometer tracing are presented. Electrophoresis was from left to right as pictured. The slowest moving (left-most) band is the TSase α chain of *E. coli* with a mol wt of 28,722, and the fastest moving (right-most) band is sperm whale apomyoglobin, with a mol wt of 17,199. The middle bands are the three different ASase component II's as indicated. The gels were scanned at a densitometer wavelength of 600 nm. Source of component II or component II fragment: (a) *E. coli;* (b) *S. marcescens;* (c) *S. typhimurium.*

	1			5					10					15	16			20					25
E. coli	Met	Gln	Thr	Gln	Lys	Pro	Thr	Leu	Glu	Leu	Leu	Thr	Ser	Glu	Gly	Ala	Tyr	Arg	Asp	Asn	Pro	Thr	Ala	Leu	Phe	X
S. typhimurium	–	–	–	Pro	–	–	–	–	–	–	–	–	–	Asp	Ala	–	–	–	Glu	–	–	–	–	–	–	X
S. marcescens	–	Asn	–	Δ	–	–	His	–	Thr	–	–	Lys	Val	Gln	Ala	Ser	–	–	Gly	–	–	–	Thr	–	–	His
	1					5					10					15					20				25

FIGURE 4: Comparison of amino acid sequences of ASase component I of *E. coli*, *S. typhimurium*, and *S. marcescens*. A residue which differs from that of *E. coli* is given, while identity to an *E. coli* residue is indicated by a dash (–). A Δ signifies the deletion of an amino acid. X indicates that no attempt was made to identify this residue.

	5								10								15								20	
E. coli	Ala	Asp	Ile	Leu	Leu	Leu	Asp	Asn	Ile	Asp	Ser	Phe	Thr	Tyr	Asn	Leu	Ala	Asp	Gln	Leu						
S. typhimurium	–	–	–	–	–	–	–	–	–	–	–	–	–	Trp	–	–	–	–	–	–						
S. marcescens	–	–	–	–	–	–	–	–	Val	–	–	–	–	–	–	–	Val	–	–	–						
	25								30								35								40	
E. coli	Arg	Ser	Asn	Gly	His	Asn	Val	Val	Ile	Tyr	Arg	Asn	His	Ile	Pro	Ala	Gln	Thr	Leu	Ile						
S. typhimurium	–	Thr	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–						
S. marcescens	–	Ala	Ser	–	–	Gln	–	–	–	–	–	–	Gln	–	Gly	–	–	Val	Ile	–						
	45								50								55								60	
E. coli	Glu	Arg	Leu	Ala	Thr	Met	Ser	Asn	Pro	Val	Leu	Met	Leu	Ser	Pro	Gly	Pro	Gly	Val	Pro	Ala					
S. typhimurium	Asp	–	–	–	–	–	?	–	–	–	–	–	X	X	X	X	X	X	X	X	X					
S. marcescens	–	–	–	Gln	Gln	–	?	Gln	–	–	–	–	–	–	–	–	–	–	Ala	–	Val					

FIGURE 5: Comparison of amino acid sequences of ASase component II of *E. coli*, *S. typhimurium*, and *S. marcesens*. Automatic Edman degradation of the ASase component II fragment of *S. typhimurium* was programmed for 52 cycles only, while 61 residues were analyzed with the proteins of *E. coli* and *S. marcescens*. A question mark indicates that no amino acid could be identified (see residue no. 47 in the sequences of *S. typhimurium* and *S. marcescens*). For an explanation of other symbols see the legend to Figure 4.

umns age, recoveries of the PhNCS amides are reduced. Lys>PhNCS could be detected following "off column" silylation, or by hydrolysis and amino acid analysis. Lys>PhNCS was occasionally found in the aqueous layer rather than the ethyl acetate layer following conversion and extraction. His>PhNCS, Arg>PhNCS, and PhNCS cysteic acids were present in the aqueous layer following extraction; the corresponding amino acids were identified following acid hydrolysis. As indicated in Table I, thin layer chromatography (tlc) was used occasionally to confirm the identification of certain PhNCS amino acids. More extensive discussions of these and other procedures for the identification of PhNCS amino acids are presented elsewhere (Edman and Begg, 1967; Pisano and Bronzert, 1969; Edman, 1970; Smithies *et al.*, 1971; Hermodsen *et al.*, 1972; Li and Yanofsky, 1972; Hood *et al.*, 1973).

Chemicals. All reagents and solvents for the Sequencer and the "10%" SP400 support for gas–liquid chromatography were purchased from Beckman Instruments. (The 2/3 CFC support was kindly provided by Dr. J. Pisano of the National Institutes of Health.) 4-Vinylpyridine was a gift from the Riley Tar Co. *N,O-Bis*(trimethylsilyl)acetamide was purchased from Pierce Chemical Co. Hydriodic acid (56.3% HI; preservative hypophosphoric acid) was from Fisher Chemical. Porcine pancreatic lipase (B grade) was from Calbiochem.

Results

Protein Purification and Subunit Separation. The purification procedures described under Experimental Procedures yielded highly purified proteins which showed two major bands following sodium dodecyl sulfate gel electrophoresis. The elution profiles obtained following separation of the subunits on Sephadex G-100 and the corresponding gel patterns of selected fractions are shown in Figure 2. Peak fractions from G-100 columns were combined and prepared for sequence analyses. Single PhNCS amino acids were released at each of the first few steps of Edman degradation attesting to the purity of the protein fractions.

Molecular Weight Estimation. The molecular weights of ASase component I of each of the three species was roughly estimated at 60,000–65,000, in agreement with the results of other studies (Ito *et al.*, 1969; Hwang and Zalkin, 1971; Zalkin and Hwang, 1971). The molecular weights of component II of *S. marcescens* and the component II fragments of *E. coli* and *S. typhimurium* were carefully determined with sperm whale apomyoblobin and *E. coli* TSase α chain as internal reference proteins. Figure 3 presents photographs and densitometer tracings of the protein bands observed following sodium dodecyl sulfate gel electrophoresis. Molecular weights of 23,000 and 23,100, respectively, were calculated for the component II fragments of *E. coli* and *S. typhimurium*. The molecular weight of component II of *S. marcescens* was estimated at 21,400. Previous estimates for the component II of *S. marcescens* and the component II fragment of *S. typhimurium* were 21,000 (Zalkin and Hwang, 1971) and 24,000 (Grieshaber and Bauerle, 1972), respectively.

Amino-Terminal Sequences of ASase Components I and II. The results obtained in sequence analyses with ASase component I of each of the three species are summarized in Table II, and those obtained with component II of *S. marcescens* and component II fragments of *E. coli* and *S. typhimurium* are presented in Table III. To determine whether the amino-terminal sequences of the component I and component II fragments of *E. coli* represent the sequences of the intact polypeptides, 15 steps of Edman degradation were performed with the purified ASase component I–component II complex of this organism. The sequences obtained (data not presented) were identical with the sum of the amino-terminal sequences of component I and component II fragments isolated following protease treatment. In view of the homology of the *E. coli* and *S. typhimurium* sequences (Tables II and III) it is likely that the amino-terminal sequences of the component I and component II fragments of *S. typhimurium* are identical with those of untreated components I and II of this organism.

Discussion

The ASase component I's of *E. coli*, *S. typhimurium*, and *S. marcescens* are similar in size and function (Zalkin, 1973). The component II's differ appreciably in that those of *E. coli* and *S. typhimurium* catalyze the PRTase reaction of tryptophan

TABLE IV: Presumed Base Differences in the Genetic Regions Specifying the Initial Portions of Anthranilate Synthetase Components I and II and Tryptophan Synthetase α Chain of *E. coli*, *S. typhimurium*, and *S. marcescens*.

		ASase I	ASase II	TSase α
E. coli vs. S. typhimurium	No. of residues compared	25	51	28
	No. of residue differences	4 (16%)	3 (6%)	5 (18%)
	AA differences explicable by			
	Single base change	4	2	5
	Double base change	0	1	0
	Deletion[a]	0	0	0
	Minumum base differences[b,c]	4 (5.3%)	4 (2.6%)	5 (6.0%)
E. coli vs. S. marcescens	No. of residues compared	25	60	28
	No. of residue differences	11 (44%)	14 (23%)	9 (32%)
	AA differences explicable by			
	Single base change	6	8	5
	Double base change	4	6	4
	Deletion	1	0	0
	Minimum base differences	15 (20.0%)	20 (11.1%)	13 (14.5%)
S. typhimurium vs. S. marcescens	No. of residues compared	25	51	28
	No. of residue differences	10 (40%)	14 (27%)	8 (29%)
	AA differences explicable by			
	Single base change	4	7	4
	Double base change	5	7	4
	Deletion	1	0	0
	Minimum base differences	15 (20.0%)	21 (13.7%)	12 (14.3%)

[a] The deletion of a single amino acid is regarded as a one-base change since it probably resulted from a single mutational event. [b] Minimum base differences refer to the minimum number of base changes required to account for the observed amino acid differences. [c] The number of amino acid residue differences between the total sequences of the α chains of *E. coli* and *S. typhimurium* is 44/268 (16.4%) and the minimum base difference is 45 (6.3%) (Li *et al.*, 1973a).

biosynthesis while that of *S. marcescens* does not (Zalkin, 1973). In addition, component II of the latter species is one-third the size of the component II's of the other two (Zalkin, 1973). The component II's of all three species provide the glutamine amidotransferase function to the ASase complex. In some organisms in which ASase component II resembles that of *S. marcescens* in size and function, ASase component II also appears to provide glutamine amidotransferase activity in an enzyme complex involved in *p*-aminobenzoate synthesis (Kane and Jensen, 1970; Kane *et al.*, 1972; Sawula and Crawford, 1972, 1973).

The differences in the ASase components prompted this investigation of the extent of sequence homology of components I and II of *E. coli*, *S. typhimurium*, and *S. marcescens*. The amino-terminal sequences of the various components were determined and are aligned for comparison in Figures 4 and 5. These are the first partial sequences determined for these polypeptide chains. It is apparent that the respective polypeptides are homologous in all three species, despite the major differences in size and function between the component II's of the three species. It is noteworthy that mild proteolytic treatment of the ASase complexes of *E. coli* and *S. typhimurium* cleaves the carboxyl-terminal two-thirds of component II, the region responsible for PRTase activity, and produces a modified ASase complex resembling that of *S. marcescens* in amino-terminal sequences, polypeptide sizes, and enzyme activity. As discussed elsewhere (Zalkin and Hwang, 1971; Grieshaber and Bauerle, 1972; Li *et al.*, 1974) these and other findings suggest that the bifunctional component II of *E. coli* and *S. typhimurium* arose as a consequence of the fusion of a gene for a glutamine amidotransferase and the gene specifying PRTase.

The extent of divergence of the amino acid sequences of components I and II is analyzed in Table IV. For comparison divergence of the TSase α-chain sequences of the same bacterial species (Guest *et al.*, 1967; Li and Yanofsky, 1973a,b; Li *et al.*, 1973b) is presented. It is evident that *E. coli* and *S. typhimurium* are more closely related to each other than to *S. marcescens*. Component II shows somewhat less variation than the other two polypeptides suggesting that genes specifying proteins concerned with different reactions in the tryptophan pathway may have diverged to different extents. Results of mRNA–DNA hybridization studies (Denney and Yanofsky, 1972) with *trp* mRNA from the same bacterial species support this conclusion. The apparent slower rate of evolutionary change of component II can be explained by more stringent structural requirements for function or by the invasion of two of the species by an amidotransferase gene from the third, as a relatively late evolutionary event.

Knowledge of the amino-terminal sequence of ASase component I of *E. coli* (component I is specified by the most operator-proximal structural gene of the *trp* operon) and reference to the genetic code permit the prediction of the approximate nucleotide sequence of the corresponding segment of the *trp* operon. Nucleotide sequences of messenger RNA transcribed from the operator-proximal segment of the operon have been isolated (Cohen *et al.*, 1973; Bronson *et al.*, 1973) and a sequence which corresponds to the first 11 amino residues of component I has been reported (Bronson *et al.*, 1973).

Acknowledgment

The author are indebted to Howard Zalkin for his advice concerning the purification of the various anthranilate synthetases. We are also indebted to Drs. Kenneth Walsh, Lee

Hood, and Jack Ohms for generously informing us of their experiences with automatic sequencing methodology.

References

Bauerle, R. H., and Margolin, P. (1966), *Cold Spring Harbor Symp. Quant. Biol. 31*, 203.

Bronson, M. J., Squires, C., and Yanofsky, C. (1973), *Proc. Nat. Acad. Sci. U. S. 70*, 2335.

Cohen, P. T., Yaniv, M., and Yanofsky, C. (1973), *J. Mol. Biol. 74*, 163.

Dayhoff, M. O. (1972), Atlas of Protein Sequence and Structure, Silver Spring, Md., National Biomedical Research Foundation.

Denney, R. M., and Yanofsky, C. (1972), *J. Mol. Biol. 64*, 319.

Edman, P. (1970), *in* Protein Sequence Determinations, Needleman, S. B., Ed., New York, N. Y., Springer-Verlag, p 211.

Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1*, 80.

Egan, A. F., and Gibson, F. (1966), *Biochim. Biophys. Acta 130*, 276.

Friedman, M., Krull, L. H., and Cavins, J. F. (1970), *J. Biol. Chem. 245*, 3868.

Grieshaber, M., and Bauerle, R. (1972), *Nature (London), New Biol. 236*, 232.

Guest, J. F., Drapeau, G. R., Carlton, B. C., and Yanofsky, C. (1967), *J. Biol. Chem. 242*, 5442.

Henderson, E. J., Nagano, H., Zalkin, H., and Hwang, L. H. (1970a), *J. Biol. Chem. 245*, 1416.

Henderson, E. J., and Zalkin, H. (1971), *J. Biol. Chem. 246*, 6891.

Henderson, E. J., Zalkin, H., and Hwang, L. H. (1970b), *J. Biol. Chem. 245*, 1424.

Hermodsen, M. A., Erisson, L. H., Titani, K., Neurath, H., and Walsh, K. A. (1972), *Biochemistry 11*, 4493.

Hirs, C. H. W. (1956), *J. Biol. Chem. 219*, 611.

Hood, L., McKean, D., Farnsworth, V., and Potter, M. (1973), *Biochemistry 12*, 741.

Hwang, L. H., and Zalkin, H. (1971), *J. Biol. Chem. 246*, 2338.

Ito, J., Cox, E. C., and Yanofsky, C. (1969), *J. Bacteriol. 97*, 725.

Ito, J., and Yanofsky, C. (1966), *J. Biol. Chem. 241*, 4112.

Ito, J., and Yanofsky, C. (1969), *J. Bacteriol. 97*, 734.

Jackson, E., and Yanofsky, C. (1974), *J. Bacteriol. 117*, 502.

Kane, J. F., Holmes, W. M., and Jense, R. A. (1972), *J. Biol. Chem. 247*, 1587.

Kane, J. F., and Jensen, R. A. (1970), *Biochem. Biophys. Res. Commun. 41*, 328.

Li, S. L., Denney, R. M., and Yanofsky, C. (1973a), *Proc. Nat. Acad. Sci. U. S. 70*, 1112.

Li, S. L., Drapeau, G. R., and Yanofsky, C. (1973b), *J. Bacteriol. 113*, 1507.

Li, S. L., Hanlon, J., and Yanofsky, C. (1974), *Nature (London)* (in press).

Li, S. L., and Yanofsky, C. (1972), *J. Biol. Chem. 247*, 1031.

Li, S. L., and Yanofsky, C. (1973a), *J. Biol. Chem. 248*, 1830.

Li, S. L., and Yanofsky, C. (1973b), *J. Biol. Chem. 248*, 1837.

Lowry, O. H., Rosebough, N. J., Farr, A. L., and Randall, R. J. (1951), *J. Biol. Chem. 193*, 265.

Nagano, H., and Zalkin, H. (1970), *J. Biol. Chem. 245*, 3097.

Pisano, J. J., and Bronzert, T. J. (1969), *J. Biol. Chem. 244*, 5597.

Queener, S. F., and Gunsalus, I. C. (1970), *Proc. Nat. Acad. Sci. U. S. 67*, 1225.

Queener, S. W., Queener, S. F., Meeks, J. R., and Gunsalus, I. C. (1973), *J. Biol. Chem. 248*, 151.

Robb, F., and Belser, W. L. (1972), *Biochim. Biophys. Acta 285*, 243.

Robb, F., Hutchinson, M. A., and Belser, W. L. (1971), *J. Biol. Chem. 246*, 6908.

Sawula, R. V., and Crawford, I. P. (1972), *J. Bacteriol. 112*, 797.

Sawula, R. V., and Crawford, I. P. (1973), *J. Biol. Chem. 248*, 2573.

Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G., and Ballantyne, D. L. (1971), *Biochemistry 10*, 4912.

Thomsen, J., Kristiansen, K., Brunfeldt, K., and Sundby, F. (1972), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 21*, 315.

Vogel, H., and Bonner, D. (1956), *J. Biol. Chem. 218*, 97.

Weber, K., and Osborn, M. (1969), *J. Biol. Chem. 244*, 4406.

Wegman, J., and Crawford, I. P. (1968), *J. Bacteriol. 94*, 2325.

Yanofsky, C., Horn, V., Bonner, M., and Stasiowski, S. (1971), *Genetics 69*, 409.

Zalkin, H. (1973), *in* Advances in Enzymology, Vol. 38, Meister, A., Ed., New York, N. Y., Wiley, p 1.

Zalkin, H., and Hwang, L. H. (1971), *J. Biol. Chem. 246*, 6899.