INDOLES
PART TWO

Edited by

William J. Houlihan

INDOLES
PART TWO

Edited by

William J. Houlihan

Sandoz-Wander, Inc.
Research and Development Division
Hanover, New Jersey

CONTRIBUTORS

Ronald J. Parry

Department of Chemistry
Stanford University
Stanford, California

James C. Powers

Georgia Institute of Technology
Atlanta, Georgia

Kent Rus[illegible]

Eastman Kodak Co.
Research Laboratories
Rochester, New York

L. R. Smith

Monsanto Co.
St. Louis, Missouri

F. Troxler

Chemical-Pharmaceutical Research Division
Sandoz AG Basle, Switzerland

WILEY-INTERSCIENCE

a division of

JOHN WILEY & SONS, INC. · NEW YORK · LONDON · SYDNEY · TORONTO

Library of Congress Catalog Card Number: 76-154323

ISBN 0-471-37501-2

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

INDOLES

PART TWO

This is the twenty-fifth volume in the series

THE CHEMISTRY OF HETEROCYCLIC COMPOUNDS

THE CHEMISTRY OF HETEROCYCLIC COMPOUNDS

A SERIES OF MONOGRAPHS

ARNOLD WEISSBERGER and EDWARD C. TAYLOR

Editors

The Chemistry of Heterocyclic Compounds

The chemistry of heterocyclic compounds is one of the most complex branches of organic chemistry. It is equally interesting for its theoretical implications, for the diversity of its synthetic procedures, and for the physiological and industrial significance of heterocyclic compounds.

A field of such importance and intrinsic difficulty should be made as readily accessible as possible, and the lack of a modern detailed and comprehensive presentation of heterocyclic chemistry is therefore keenly felt. It is the intention of the present series to fill this gap by expert presentations of the various branches of heterocyclic chemistry. The subdivisions have been designed to cover the field in its entirety by monographs which reflect the importance and the interrelations of the various compounds, and accommodate the specific interests of the authors.

In order to continue to make heterocyclic chemistry as readily accessible as possible, new editions are planned for those areas where the respective volumes in the first edition have become obsolete by overwhelming progress. If, however, the changes are not too great so that the first editions can be brought up-to-date by supplementary volumes, supplements to the respective volumes will be published in the first edition.

ARNOLD WEISSBERGER

Research Laboratories
Eastman Kodak Company
Rochester, New York

EDWARD C. TAYLOR

Princeton University
Princeton, New Jersey

Preface

Indoles Part Two begins the detailed coverage of the preparation, properties, reactions and tabulation of compounds containing an indole nucleus. It starts with a chapter on indole biosynthesis since this was the first and only source of indole preparations during the early years of indole chemistry.

The editor is grateful to Mrs. Maria Fanlo and Mr. Siegfried Wahrmann for library assistance and to Miss Linda Heuser for typing a portion of the manuscript.

WILLIAM J. HOULIHAN

Hanover, New Jersey

Contents

Part Two

Part One

Part Three

CHAPTER III

Biosynthesis of Compounds Containing an Indole Nucleus

RONALD J. PARRY

University Chemical Laboratory, Cambridge, England

I. Introduction

Living plants produce an extraordinarily rich variety of chemical substances, many of which lack any apparent biochemical function. These

metabolites have often proved the delight (and the frustration!) of organic chemists for the challenges of structural and synthetic chemistry which they offer. With the advent of radioactive tracers and the development of more sensitive chemical and spectroscopic tools, the doors leading to a deeper understanding of the chemistry of the plant world have been opened: the investigation of natural product biosynthesis has begun. Of the fruits of a field yet in its infancy, those arising from an examination of the biosynthesis of the naturally-occurring indoles have proved among the most tantalizing, and much may be expected of the future. The present account summarizes our knowledge of the biosynthesis of compounds containing an indole nucleus, and covers the literature through November 1968.

II. Simple Indole Derivatives

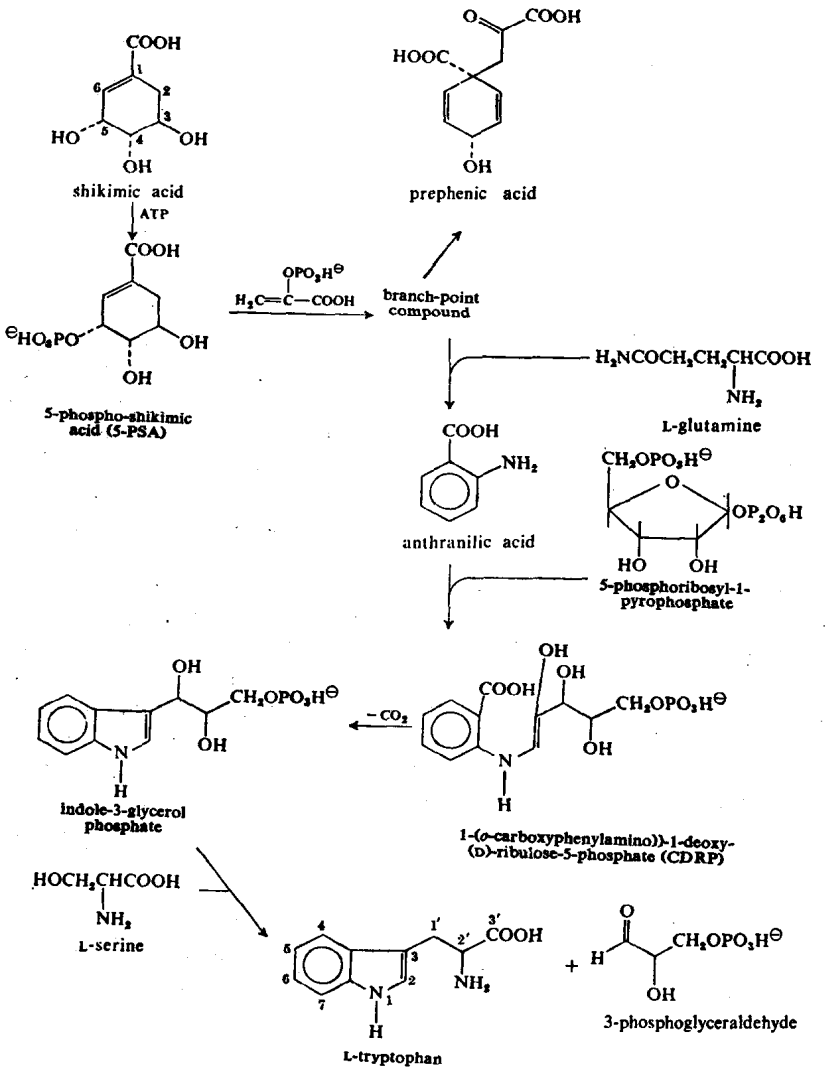
A. Tryptophan

By virtue of its ubiquitous distribution in plant and animal proteins, tryptophan may justifiably be regarded as the most important of the naturally-occurring indoles. Extensive explorations aimed at unraveling the tangled thread of its biosynthesis are a consequence of this importance. These investigations have been limited almost exclusively to microorganisms, and disappointingly few experiments have been conducted with fungi and higher plants.

The subject of tryptophan biosynthesis in microorganisms was carefully reviewed in 1960 by Doy[1]; Scheme 1 summarizes the metabolic picture presented by the experimental evidence available at that time. More recent work has supplied some of the significant detail absent from this picture.

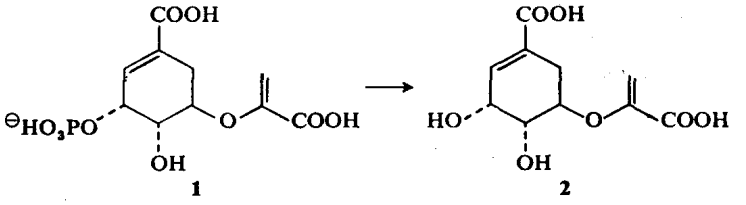
The enzyme which phosphorylates shikimic acid to 5-phosphoshikimic acid (5-PSA) has been isolated from *Escherichia coli* by Fewster.[2] Its optimum pH is 7.0 and it exhibits a requirement for divalent magnesium or manganese. Neither the formation of the enzyme nor its activity is affected by the ultimate products of the aromatic biosynthetic pathway. The same author also reported evidence for the presence of this enzyme system in a variety of microorganisms known to synthesize aromatic amino acids.

One of the most fascinating problems in tryptophan biosynthesis, the nature of the so called branch point compound leading either to prephenic acid or to anthranilic acid, has yielded to the patience of the investigators. Early experimental evidence[3] suggested that at least one additional substance, called Z_1 and formulated[4] as the 5-enolpyruvyl ether of shikimic acid, was produced from 5-PSA before the branch point. Later work by Srinivasan[5]


shikimic acid
ATP
5-phospho-shikimic acid (5-PSA)
branch-point compound
prephenic acid
L-glutamine
anthranilic acid
5-phosphoribosyl-1-pyrophosphate
1-(o-carboxyphenylamino))-1-deoxy-(D)-ribulose-5-phosphate (CDRP)
indole-3-glycerol phosphate
L-serine
L-tryptophan
3-phosphoglyceraldehyde

Scheme 1

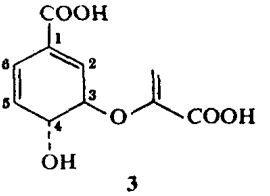
indicated, however, that Z_1 was not in fact an intermediate in the conversion of 5-PSA to anthranilic acid in cellfree extracts of *E. coli*, and Levin and Sprinson[6] found that Z_1 is not converted to prephenic acid by extracts of the same organism. These authors present additional data suggesting that the first product formed from 5-PSA and phosphoenolpyruvic acid is 3-enolpyruvylshikimate-5-phosphate (**1**), which is then dephosphorylated to Z_1, proposed to be the 3-enolpyruvyl ether of shikimic acid (**2**). In the presence



1 **2**

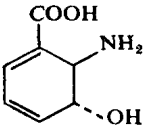
of fluoride ion, the dephosphorylation reaction is inhibited, and Z_1-phosphate, **1**, accumulates. Since prephenate formation from 5-PSA and phosphoenolpyruvic acid could be demonstrated, it follows that Z_1-phosphate is probably the active intermediate leading to prephenate. Evidence implicating Z_1-phosphate as a precursor of anthranilic acid was also forthcoming. Cellfree extracts of *Aerobacter aerogenes*[7] converted shikimic acid or 5-PSA in the presence of phosphoenolpyruvic acid to a substance with properties identical to those of Z_1-phosphate as reported by Levin and Sprinson. By using mutants of the same organism which were unable to convert 5-PSA to Z_1-phosphate, to anthranilic acid, or to phenylpyruvic acids, the formation of these acids in a cell extract containing Z_1-phosphate could be detected. Treatment of the Z_1-phosphate containing extract with alkaline phosphatase followed by acid produced a substance supporting the growth of an *E. coli* mutant requiring shikimic acid. Neither treatment alone produced a growth factor, but either destroyed the substrate for anthranilic acid formation. Addition of fluoride ion improved the yield of anthranilic acid from the substrate. The role of Z_1-phosphate in anthranilic acid biosynthesis in *E. coli* has been studied by Rivera and Srinivasan.[8] Ammonium sulfate or protamine sulfate treatment of a crude anthranilate forming enzyme preparation from an *E. coli* mutant gave two fractions. One of these contained an enzyme, named 3-enolpyruvylshikimate 5-phosphate synthetase, that condensed 5-PSA and phosphoenolpyruvic acid to give Z_1-phosphate. This enzyme fraction further converted Z_1-phosphate to a new, unidentified substance which was itself converted to anthranilic acid by the second enzyme fraction in the presence of L-glutamine, divalent magnesium, nicotinamide adenine dinucleotide (NAD^+), and a nicotinamide adenine dinucleotide, reduced form (NADH) regenerating system. The second fraction was unable to convert Z_1-phosphate to anthranilate.

Gibson and Gibson[9, 10] also reported the presence of a new intermediate in aromatic ring biosynthesis in extracts of an *A. aerogenes* mutant. This substance could be converted by mild chemical treatment into prephenic acid, *p*-hydroxybenzoic acid, and phenylpyruvic acid; enzymically, it was transformed into anthranilic, prephenic, phenylpyruvic, *p*-hydroxyphenylpyruvic, and *p*-hydroxybenzoic acids. On the basis of this evidence, the substance was judged to be the elusive branch point compound, and was named chorismic acid (chorismic = separating). The obtention of a multiply-blocked auxotroph of *A. aerogenes* which accumulated the acid allowed its isolation[11] as the barium salt and formulation[11, 12] as the 3-enolpyruvyl ether of *trans*-3,4-dihydroxycyclohexa-1,5-diene carboxylic acid (**3**). Chorismic acid has also been isolated from a *Saccharomyces cerevisiae* mutant by Lingens and Luck.[13]



3

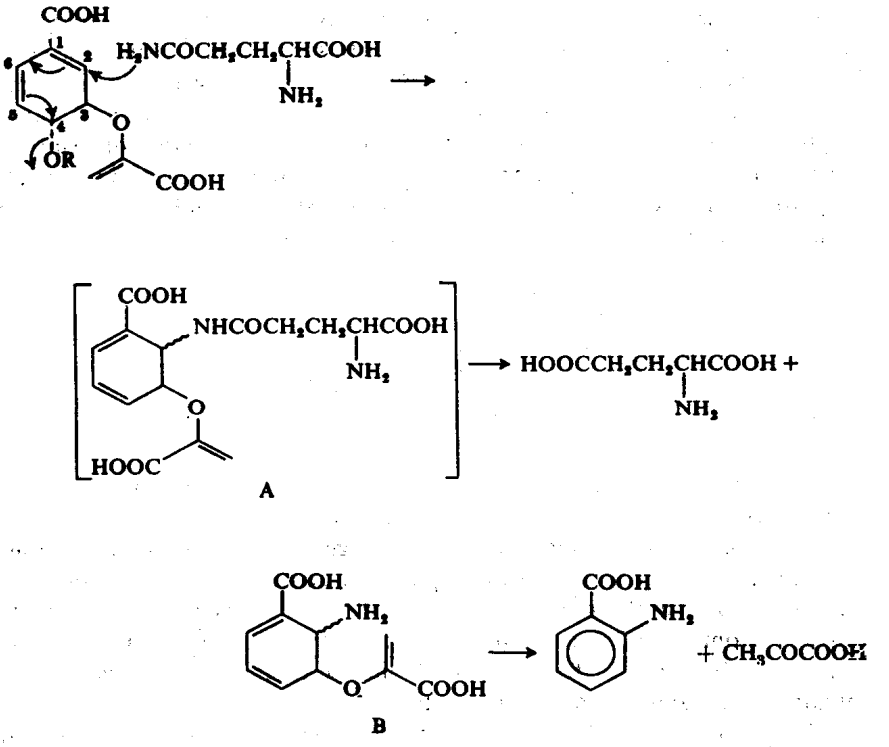
The problem of the conversion of chorismic acid into anthranilic acid is still under investigation. Srinivasan and Rivera,[14] working with *E. coli* mutants, demonstrated that an NADH regenerating system and either divalent magnesium or iron were required in addition to L-glutamine; the amino group of anthranilic acid was found to be derived from the amide nitrogen atom of glutamine. More recently, some very informative results have been obtained[15] by feeding 3,4-^{14}C-glucose to an *E. coli* mutant accumulating anthranilic acid. Earlier isotopic studies of the incorporation of 3,4 ^{14}C-glucose into shikimic acid established that the carboxyl carbon and carbons 3 and 4 become labeled.[16] Utilizing this information, degradation of the labeled anthranilic acid produced from the radioactive glucose showed that the carboxyl group of shikimic acid becomes the carboxyl group of anthranilic acid, and that the amination of chorismic acid occurs at C-2 rather than at C-6. Examination of *trans*-2,3-dihydro-3-hydroxyanthranilic acid (**4**), which has been isolated from *Streptomyces aureofaciens*,[17] as a



4

possible anthranilate precursor in cellfree extracts of *E. coli* gave negative results, indicating that the actual intermediate probably still bears the

enolpyruvyl moiety. These results were rationalized by the scheme shown in Scheme 2. In a recent report, Lingens et al. claim to have isolated a substance corresponding to A from a mutant of *Saccharomyces cerevisiae*.[18]

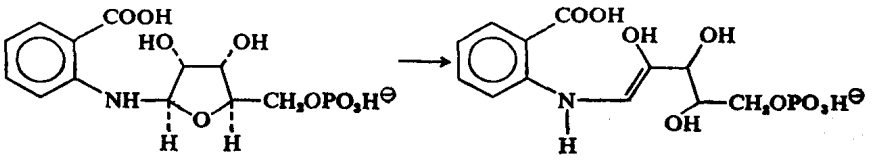


Scheme 2

DeMoss[19] has investigated the formation of anthranilic acid in *Neurospora crassa*. Cellfree extracts will convert shikimic acid into anthranilic acid after an inhibitor that is present has been removed by ammonium sulfate precipitation. The cofactors required for the transformations are identical to those of the bacterial systems studied except that nicotinamide adenine dinucleotide phosphate, reduced form (NADPH) is needed in place of NADH. Omission of L-glutamine from the incubation mixture led to the accumulation of chorismic acid. The enzyme catalyzing the chorismate to anthranilate conversion was purified 83-fold and appeared to be homogeneous. Its activity was completely inhibited by low concentrations of L-tryptophan, and this inhibition was competitively reversed by chorismic acid, suggesting that the conversion of chorismic acid to anthranilic acid is specifically involved in tryptophan biosynthesis.

The L-glutamine requirement does not appear to be obligatory. Gibson et al.[20] have recently recorded anthranilate biosynthesis that does not require glutamine. A strain of *E. coli* was obtained that required both glutamine and tryptophan. Cell suspensions of this organism were able to synthesize anthranilic acid by using glucose as the carbon source and ammonium ions as the only nitrogen source; addition of DON (6-diazo-5-oxo-6-norleucine), a glutamine antagonist to the cell suspensions, caused no inhibition of anthranilate formation. It was suggested that the un-ionized form of ammonia may be available for transfer reactions normally requiring glutamine.

The steps leading from anthranilic acid to tryptophan have been carefully scrutinized. Yanofsky proposed[21] that the first intermediate resulting from the reaction of anthranilic acid and .5-phosphoribosyl-1-pyrophosphate should be *N-o*-carboxyphenylribosylamine-5-phosphate (**5**), also known as *N*-(5′-phosphoribosyl)anthranilic acid or PRA. An Amadori rearrangement[22] of this substance to 1-(*o*-carboxyphenylamino)-1-deoxyribulose-5-phosphate (CDRP) (**6**) was postulated. Early failures to detect the presence of PRA[1, 23]



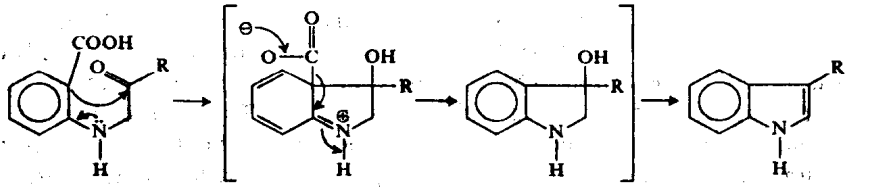
5 **6**

were ascribed to the lability of anthranilic acid glycosylamines.[1, 24] Evidence supporting this instability was subsequently provided by Doy and coworkers.[25] The half-life of synthetic PRA at 37°C in aqueous solution, pH 6, was found to be 6 min. The sensitivity of the substance increased with decreasing pH and decreased with increasing pH. This ease of hydrolysis predicts that mutants blocked between PRA and CDRP will appear to accumulate anthranilic acid unless special precautions are taken. An investigation by Doy et al.[25] confirms this prediction; extracts of certain mutant microorganisms (*E. coli*, *A. aerogenes*, *Salmonella typhimurium*), which in whole cell experiments accumulated anthranilic acid, were found to catalyze a reaction between anthranilic acid and 5-phosphoribosyl-1-pyrophosphate leading to an acid labile substance, less fluorescent than anthranilic acid, and readily hydrolyzing back to that compound. The substance was converted enzymically to indole-3-glycerol phosphate, and was surmized to be PRA. Doy[26] also reported similar behavior in two tryptophan auxotrophs of *Pseudomonas aeruginosa* which are phenotypically identical, i.e., both require indole or tryptophan for growth and accumulate anthranilic acid. The two strains, however, differ genotypically, because one is blocked between anthranilic acid and PRA while the other is blocked between PRA

and CDRP. The rapid hydrolysis of accumulating PRA results in the apparent accumulation of anthranilic acid by the second mutant.

1(-*o*-Carboxyphenylamino)-1-deoxyribulose, the dephosphorylated Amadori product, was originally detected in cell suspensions of *A. aerogenes* and in *E. coli* mutants.[27, 29] Its identity was based upon R_F values, color reactions, absorption spectra, and a DNP derivative as compared with synthetic material. On the basis of Yanofsky's scheme,[21] the substance is most reasonably considered as an artifact, derived from the actual tryptophan precursor by loss of the 5-phosphate group. Smith and Yanofsky[23] have since provided evidence for this by detecting what appears to be the phosphorylated compound in extracts of *E. coli* and *S. typhimurium* mutants; impure CDRP was obtained synthetically from anthranilic acid and the sodium salt of ribose-5-phosphate; its properties compared reasonably well with those of the naturally occurring substance. Both the synthetic and natural compounds were converted to indole-3-glycerol phosphate in the extracts.

The mechanism of the conversion of CDRP to indole-3-glycerol phosphate appears to have been little studied. Smith and Yanofsky[23] prepared the decarboxylated analog of CDRP, but found that it was not transformed into indole-3-glycerol phosphate by appropriate cell extracts. This implies that the decarboxylated substance is not a free intermediate in the reaction. Mechanistically, it seems reasonable to suppose that the decarboxylation may not occur until after ring closure has taken place, so that this biochemical inertness of the decarboxylated analog need occasion no surprise (Eq. 1).



(1)

Studies of the tryptophan synthetase enzyme system have elucidated some of the details of the last step in tryptophan biosynthesis. The enzyme obtained from *E. coli* has been shown to consist of two protein subunits, A and B,[28, 29] which catalyze three reactions[30] (Eqs. 2–4):

$$\text{indole} + \text{L-serine} \rightarrow \text{L-tryptophan} \tag{2}$$

$$\text{indole-3-glycerol phosphate} \rightleftharpoons \text{indole} + \text{3-phosphoglyceraldehyde} \tag{3}$$

$$\text{indole-3-glycerol phosphate} + \text{L-serine} \rightarrow \text{L-tryptophan} + \text{3-phosphoglyceraldehyde} \tag{4}$$

The B subunit will catalyze Reaction (2) in the absence of the A subunit,[31] and the A subunit will catalyze Reaction (3).[28] Reaction (4) only occurs in