CHEMICAL SOCIETY SYMPOSIA

BRISTOL 1958

Special Publication No. 12

CHEMICAL SOCIETY SYMPOSIA

BRISTOL 1958

DEVELOPMENTS IN AROMATIC CHEMISTRY

APPLICATIONS OF ELECTRON AND NUCLEAR RESONANCE IN CHEMISTRY

RECENT WORK ON THE INORGANIC CHEMISTRY OF SULPHUR

Special Publication No. 12

LONDON: THE CHEMICAL SOCIETY, BURLINGTON HOUSE, W.1.
1958

DEVELOPMENTS IN AROMATIC CHEMISTRY

Organised by

W. D. OLLIS

THE CHEMICAL SOCIETY

Published by
The Chemical Society, Burlington House, London, W.1

DEVELOPMENTS IN AROMATIC CHEMISTRY

APPLICATIONS OF ELECTRON AND NUCLEAR RESONANCE IN CHEMISTRY

RECENT WORK ON THE INORGANIC CHEMISTRY OF SULPHUR

PRINTED IN ENGLAND
BURTON PRESS LTD., LOUGHBOROUGH,
LEICESTERSHIRE

THE BIOSYNTHESIS OF AROMATIC COMPOUNDS FROM C_1- AND C_2-UNITS

By A. J. BIRCH and HERCHEL SMITH

(THE UNIVERSITY, MANCHESTER)

TWO main routes for the biosynthesis of aromatic compounds are now established. One, based on shikimic acid, depended for its recognition[1] chiefly on the use of mutants of micro-organisms deficient in the ability to produce essential metabolites related to phenylalanine and tryptophan. The other route, based on acetic acid, was recognised[2]* as the result of inductive reasoning concerning the inter-relations between the structures of a number of natural products, chiefly phenols. Other routes exist; *e.g.*, certain phenols can arise from the oxidation of polyisoprenoid substances. Such biosyntheses will not be discussed here. The acetic acid theory in fact correlates a considerable number of natural substances of diverse origins. The basic concepts are few and of an engaging simplicity. The theory requires the following assumptions.

(*a*) Acetic acid units are joined by formal elimination of water in head-to-tail linkage with each other or with naturally occurring carboxylic acids to form β-polyketo-acids.

(*b*) These β-polyketo-acids undergo secondary changes; notably they may cyclise by reactions of the *C*-acylation or aldol-condensation type to form aromatic rings.

(*c*) The carbon skeleton so generated may be modified by the introduction of methyl or *iso*pentenyl or polyisoprenoid groups.

(*d*) Secondary processes of reduction (sometimes accompanied by dehydration) and of oxidation may occur.

* Collie's earlier hypotheses[3] was based on the chemical reactions of 2:4:6-triketones. We were not, however, initially aware of this work.

[1] For a review see Davis, *Adv. Enzymol.*, 1955, **16**, 247.

[2] For a review see Birch, *Fortschr. Chem. org. Naturstoffe*, 1957, **16**, 186. Unless otherwise indicated, the evidence presented in this paper has been given in that review.

[3] Collie, *J.*, 1907, 1806, and earlier papers.

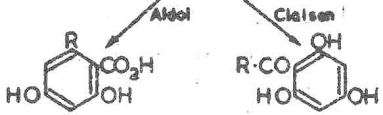
The reactions postulated have laboratory analogies, and in many cases biochemical analogies, or are mechanistically credible on the basis of known biochemical reactions.

The nature of the final products depends on the acid which initiates the synthesis, the number of acetic acid units involved, and the modes of ring closure, as well as the nature and number of the modifying substituents and reactions. The simplest routes to compounds with one aromatic ring are as formulated in (A). The initiating acids are usually straight-chain

(A) $R{\cdot}CO_2H + 3CH_3{\cdot}CO_2H \rightarrow$

$R{\cdot}CO{\cdot}CH_2{\cdot}CO{\cdot}CH_2{\cdot}CO{\cdot}CH_2{\cdot}CO_2H$

Aldol Claisen



Orcinols Acylphloroglucinols

fatty acids, branched-chain acids related to terpenes or amino-acids, or cinnamic acids, leading in the last case to pinosylvin derivatives or flavonoids (and hence to anthocyanins).

Where the number of acetic acid units is less than three, ring closure may occur on to oxygen rather than carbon, giving, for example, α-pyrone derivatives. The initiating acids in these cases are cinnamic or benzoic acid and nicotinic acid.[4] Anthranilic acid may be involved in the synthesis of acridine or quinoline alkaloids.

Where the total number of acid units is greater than four, more complex systems may result: five or more units, naphthalene (flaviolin); eight or more units, anthracene (emodin); intermediate numbers, diphenyl (alternariol)[5] or benzophenone (griseofulvin).

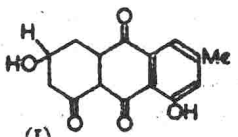
Many natural substances are closely related to those within the scope of the theory but lack expected oxygen atoms (generally as aromatic hydroxyl groups) or contain additional oxygen

[4] Moors, Gottlieb, and Djerassi, *J. Amer. Chem. Soc.*, 1957, **79**, 4506.

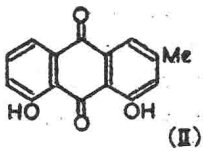
[5] Raistrick, Stickings, and Thomas, *Biochem. J.*, 1953, **55**, 421.

(usually as derivatives of phenolic hydroxyl groups, or as quinonoid oxygen). Additional methyl or *iso*pentenyl groups are found attached to carbon.

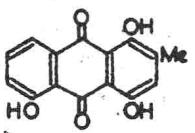
The removal of oxygen* is thought to involve reduction of a carbonyl group in a non-aromatic intermediate followed by dehydration. A biochemical analogy is the production of crotonic acid from acetoacetic acid by way of β-hydroxybutyric acid. Structures of the type envisaged as intermediates occur in some mould metabolites, *e.g.*, flavoskyrin[5,6] (I) which is readily dehydrated to the fully aromatic quinone (II). The addition of



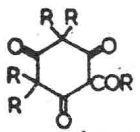
(I)



(II)



(III)



(IV; R=H or Me)

oxygen* in an *ortho*- or *para*-position to an existing hydroxyl group is chemically feasible, and analogous enzymic processes are known. Structures involving such additions are found amongst the natural anthraquinones, benzoquinones, and naphthaquinones, and even as far afield as the non-aromatic macrolide antibiotics. In some cases oxygen has been apparently removed in one place and added in another, *e.g.*, in the anthraquinone (III).

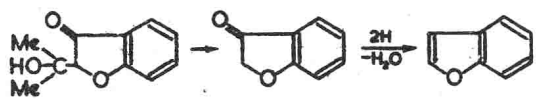
Additional methyl or *iso*pentenyl groups are invariably found in positions which could be derived from the methyl group of acetic acid. The assumption that they are introduced in a stage distinct from the formation of the main skeleton rested initially on the existence of a series of compounds with a common skeleton in which the number of groups varied from 0 to 4 in one ring, as in the triketone (IV). We suggest that the groups may be introduced as the cations which could be gene-

* This term is intended to cover any group attached to the ring by an O–C bond.

[6] Shibata, Murakami, Kitagawa, and Takido, *Proc. Japan Acad.*, 1956, **32**, 356.

rated by carriers of methyl groups (*e.g.*, methionine), thus relating the process to transmethylation on to oxygen, nitrogen, or sulphur by carriers of *iso*pentenyl groups, possibly the hitherto unknown *iso*pentenylhomocysteine or its 4-carboxylic acid. It is not possible to decide whether alkylation occurs at an open-chain or ring-closed stage; the methylene groups of a β-polyketone system and the methine groups of a phenolic ring, particularly of the phloroglucinol type, would possess the required anionoid reactivity. In some related cases, including the macrolide antibiotics, citrinin and sclerotiorin, methylation definitely occurs at an open-chain stage.

Because of its larger size, the *iso*pentenyl group is capable of cyclisations and oxidation leading to more complex products. In our view the unsubstituted furan ring in benzofurans is probably derived from an introduced C_5 unit. We believe that the process may be as illustrated. Experimental evidence in the

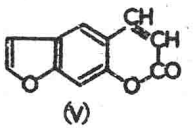


laboratory has been provided to show that this sequence is chemically feasible. The statistical evidence on which it is based consists of occurrences such as these: In the literature[8] we find 22 substances based on psoralen (V). Eighteen of these have unsubstituted furan rings, but 9 contain side chains derived from isopentenyl groups attached to oxygen; two have a 2′:3′-dihydrofuran ring substituted by a hydroxy*iso*propyl group in the 2′-position. One contains a 3′-methoxy-2-*iso*propylfuran system. There are also known similar chromenocoumarins where the furan ring is replaced by 2′:2′-dimethylchromen. In legumes related to *Derris* there occur three substances with a dimethylchroman system (toxicarol, tephrosin, and deguelin), two with a dihydro-2′-*iso*propenylfuran system (sumatrol and rotenone), and two with an unsubstituted furan ring (elliptone and malaccol), all in similar environments. Similar significant occurrences are found in the furanoquinoline alkaloids, which moreover commonly occur together with coumarins containing

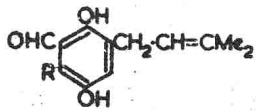
[7] Birch and Nurranabi, unpublished work.
[8] For a review of this and related evidence see Nurranabi, Thesis, Sydney, 1957.

isoprenoid side-chains.[8] The coumarin ring itself we believe to be formed by an oxidative cyclisation from C_6–C_3 (cinnamic acids) or their congeners,[9] rather than from acetic acid as the *meta*-oxygenated groups might suggest.

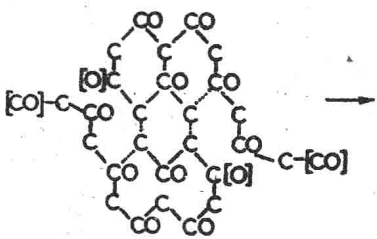
Much of the structural evidence has been reviewed elsewhere.[2] Here we wish to draw attention to several cases not previously mentioned.

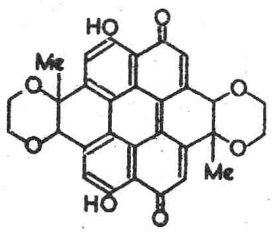


(V)



(VI; R = Me·[CH:CH]$_3$)
(VII: R = Me·[CH$_2$]$_6$)





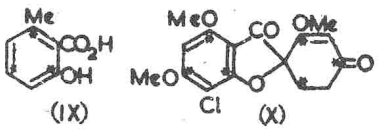
An interesting pair of quinol derivatives[10] is auroglaucin (VI) and flavoglaucin (VII) from a number of moulds. The straight chain must come from four acetic acid units, so the ring is probably derived from three more. In order to explain the other substituents it is necessary to postulate introduction of a C_1 unit (CHO, which could also arise from reduction of CO_2H of acetic acid), a C_5 unit ($Me_2C{:}CH{\cdot}CH_2$), removal of the oxygen atom at the 5-position and introduction of another oxygen atom at the neighbouring position. This complex series of reactions would certainly not be predictable except on the basis of the work on other compounds; it can be tested by tracer experiments.

It is noteworthy that the coronene skeleton of the aphin chromophore[11] (VIII) could be provided by coupling of seven acetic acid units with reductive removal of one oxygen atom.

[9] Haworth, *J.*, 1942, 448.
[10] Pannizi and Nicolaus, *Gazzetta*, 1953, **83**, 774; Quilico, Cardani, and Stagno d'Alcontres, *ibid.*, p. 754, and earlier papers.
[11] Brown, Calderbank, Johnson, Quayle, and Todd, *J.*, 1955, 1144, and earlier papers.

Uses of Hypotheses. – Apart from the unifying correlations which they achieve, one use of biosynthetic hypotheses in organic chemistry is to assist structural determinations by correcting erroneous formulæ and by suggesting which of otherwise equally probable structures should be correct. Many examples of such assistance may be found in the field of polyisoprenoids and alkaloids where the building units are quite large. A number have also become available associated with the acetate hypothesis. The units in this case are small, and the characteristic oxygenation patterns may have been partially obscured by the oxidation–reduction reactions mentioned above; nevertheless it is clear, other things being equal, that a structure which can be derived from an acetic acid chain in which the linkage is head-to-tail throughout is more likely than one where this is not the case; that an oxygenation pattern which corresponds most closely to the theoretical acetate one is the most probable, and that a structure in which introduced groups are attached to carbon atoms originating from methyl groups on the acetic acid theory is more probable than one where this is not the case. With these rules as background the correct one of several possible formulæ has been predicted for eleutherinol, flaviolin, mellein, α- and β-sorigenin, and nalgiovensin.[12] Such uses serve to increase confidence in the essential correctness of the theory.

Tracer Experiments. – The most obvious way of obtaining biochemical evidence relevant to the hypothesis is by feeding experiments with isotopically labelled substrates. The aim of our initial experiments was to provide evidence that acetic acid joined in the predicted manner could furnish all the carbon atoms in selected orcinol and phloroglucinol systems; such findings would indicate that the postulated cyclisation reactions can occur.



Degradation of 6-methylsalicylic acid (IX) obtained by growing a strain of *P. griseofulvum* in a medium containing

[12] Birch and Massy-Westropp, *J.*, 1957, 2215.

$CH_3 \cdot {}^{14}CO_2H$ gave results in accord with the labelling pattern shown.[13] Parallel findings[14] expressed in (X) were obtained with the radioactive griseofulvin obtained from another strain of the same mould, also by means of $CH_3 \cdot {}^{14}CO_2H$. Griseofulvin is of particular importance in that it contains in the same molecule a phloroglucinol and a (potential) orcinol ring. These results have recently been supplemented by work with ${}^{14}CH_3 \cdot CO_2H$ as substrate; atoms previously unlabelled are then labelled and the remainder are only slightly active, indicating[15] only a small redistribution of the label in ${}^{14}CH_3 \cdot CO_2H$. Subsequently it was shown that the phloroglucinol rings in cyanidin from red cabbage[16] and in quercetin from buckwheat[17] arise in the expected fashion.

To determine whether the postulated *C*-alkylations can occur we next studied the biosynthesis of mycophenolic acid (XI) from *P. brevi-compactum* which, according to the hypothesis above probably contains an acetate-derived 4:6-dihydroxyphthalide nucleus bearing an introduced methyl group and a terpenoid chain which can be regarded as the remnant of an oxidised geranyl group.

This view of the biosynthesis was confirmed as follows: (*a*) Feeding of [${}^{14}CH_3$]methionine gave mycophenolic acid containing 77% of the added activity with the isotope located on the carbon atoms designated†.[18] The activity of the *O*-methyl group is rather higher than that of the *C*-methyl group, and this can probably be taken to indicate the presence of unmethylated diphenol at the stage of the addition to the growth-medium of labelled methionine which may be a limiting factor in the synthesis. (*b*) Degradation of the mycophenolic acid obtained by feeding $CH_3 \cdot {}^{14}CO_2H$ gave results in accord with the typical acetate pattern of labelling in the nucleus and the terpenoid pat-

[13] Birch, Massy-Westropp, and Moye, *Austral. J. Chem.*, 1955,8, 539.
[14] Birch, Massy-Westropp, Rickards, and Smith, *J.*, 1958,360.
[15] Birch, Rickards, and Smith, unpublished work.
[16] Grisebach, *Z. Naturforsch.*, 1957,12b, 227.
[17] Underhill, Watkin, and Neish, *Canad. J. Biochem. Biophys.*, 1957,35, 219, 229; Geissman and Swain, *Chem. and Ind.*, 1957, 984.
[18] Birch, English, Massy-Westropp, Slaytor, and Smith, *J.*, 1958, 365.

tern in the side chain.[19] In initial experiments the degree of labelling in the ring and the side-chain were found to be the same: later work gave a definitely higher incorporation into the side chain.

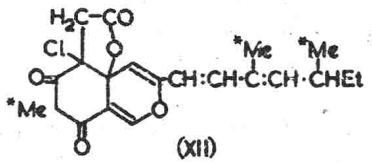
Because of the known biochemical route from formic acid to the methyl group of methionine it is frequently possible to diagnose transmethylation reactions by feeding experiments with the very readily available $H \cdot {}^{14}CO_2H$. This device has been used[20] to detect three introduced methyl groups in sclerotiorin (probable formula XII: labelled atoms denoted*),[21] a metabolite of a strain of *P. multicolor*, into which the tracer was incorporated to an extent of 20%; the remainder of the molecule has been proved to arise from acetic acid.[19] It is particularly interesting that the =CH–O– group of the pyran ring apparently arises from reduction of carboxyl.

Tracer methods, in conjunction with a biosynthetic hypothesis of proved value, may prove useful in structure determination. Some of our preliminary experiments with curvularin produced by a *Curvularia* species may be of interest in this connexion (this work is being carried out with Dr. Musgrave, Aberdeen). This substance is a C_{16} compound which might contain, on the basis of Kuhn–Roth oxidation, one or two *C*-methyl groups. Oxidation of the trimethyl ether gave a fragment (2-carboxy-3:5-dimethoxyphenylacetic acid) with an oxygenation pattern suggesting derivation from acetic acid. In this event the presence of two *C*-methyl groups indicates with a high degree of probability the introduction of one methyl group by transmethylation. The number of acetic acid units would then be seven, one of them decarboxylated. This possibility was eliminated as follows. Feeding $CH_3 \cdot {}^{14}CO_2H$ gave radioactive curvularin into which the tracer had been incorporated to an extent of 8%; Kuhn–Roth oxidation then gave acetic acid having one-eighth of the activity of the curvularin, a result which indicates (*a*) complete derivation from eight acetic acid units and (*b*) lack of an introduced methyl group, which would otherwise have given rise to some unlabelled acetic acid. Degradation

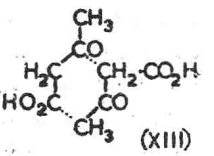
[19] Birch, English, Massy-Westropp, and Smith, *J.*, 1958, 369.
[20] Birch, Fitton, and Smith, unpublished work.
[21] Whalley, personal communication.

of the acetic acid showed all the label to be, as expected, on the carboxyl group. Confirmation of these views is obtained by observing that $H \cdot {}^{14}CO_2H$ was incorporated non-specifically and to an insignificant extent into the metabolite. The probability that curvularin is built by head-to-tail linkage of eight acetic acid units should greatly simplify the structural problem once sufficient degradative evidence has been obtained.

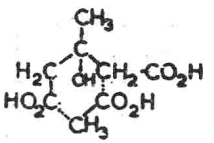
Intermediates. – Once the structural origins of the carbon atoms of a natural product have been elucidated, the question arises whether the biosynthesis involves the simultaneous assembly of the building units which are then united in a concerted process, or the structure is forged piecemeal by way of definite intermediates. Evidence on this point in the field of compounds derived from acetic acid is scanty but on balance favours the latter type of process. In the simplest case, that of 6-methylsalicylic acid, if intermediates are involved, there are four obvious ways for the synthesis to proceed. These are represented in the expressions (XIII)–(XVI). Accordingly, 6-



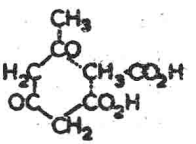
(XII)



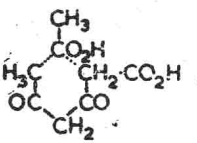
(XIII)



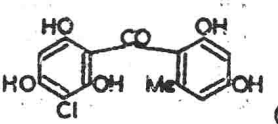
(XIV)



(XV)



(XVI)



(XVII)

methylsalicylic acid was produced from $CH_3 \cdot {}^{14}CO_2H$ in the presence of acetoacetic acid, β-hydroxy-β-methylglutaric acid, and 3:5-dioxohexanoic acid, but in no case was incorporation of radioactivity depressed at appropriate portions of the molecule. These results need not debar the three acids from being intermediates; systems capable of activating them may be absent or