Pure & Appl. Chem., Vol. 50, pp.1005-1014.
Pergamon Press Ltd. 1978. Printed in Great Britain.
© IUPAC
0033-4545/78/1001-1005 $02.00/0

CHANCE AND DESIGN IN BIOSYNTHESIS

Arthur J. Birch

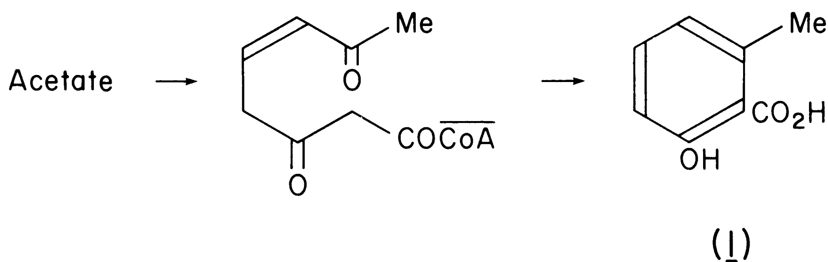
Research School of Chemistry, The Australian National University,
Canberra, A.C.T. 2600, Australia.

Abstract - Biosynthetic hypotheses of organic chemists found practical applications initially in assisting the determinations of the structures of natural products. They later contributed to biochemical work by suggesting what incorporations of isotopically labelled precursors should be examined. They also postulated biochemical mechanisms, including predictions such as C-methylation from methionine and the involvement of carbonium ion cyclisations in terpenoid biosynthesis. The incorporations of precursors specifically labelled with [^{14}C] [^{13}C] or [^{15}N] have further simplified the structure-determinations of some fungal metabolites. Examples illustrating these various facets are discussed, including rosenonolactone (2), echinulin (12), brevianamide-A (14) and phomazarin (23). Finally, a knowledge of biosynthetic pathways can permit their manipulation, and the biosynthesis of novobiocin (24 R=Me) is considered and the knowledge is used to produce an antibiotically active analogue (24 R=Cℓ).

Biosynthesis as a science grew from two separate viewpoints: the speculations of organic chemists based on synthetic mechanisms and the findings of biochemists. When I began work in the area in 1951, from the organic side, there was no connection between the two approaches, although work then in progress on fatty acids and steroids indicated that the connection was imminent. I have discussed elsewhere (Ref. 1) some aspects of the mutual incomprehension and suspicion of chemists and biochemists in earlier years.

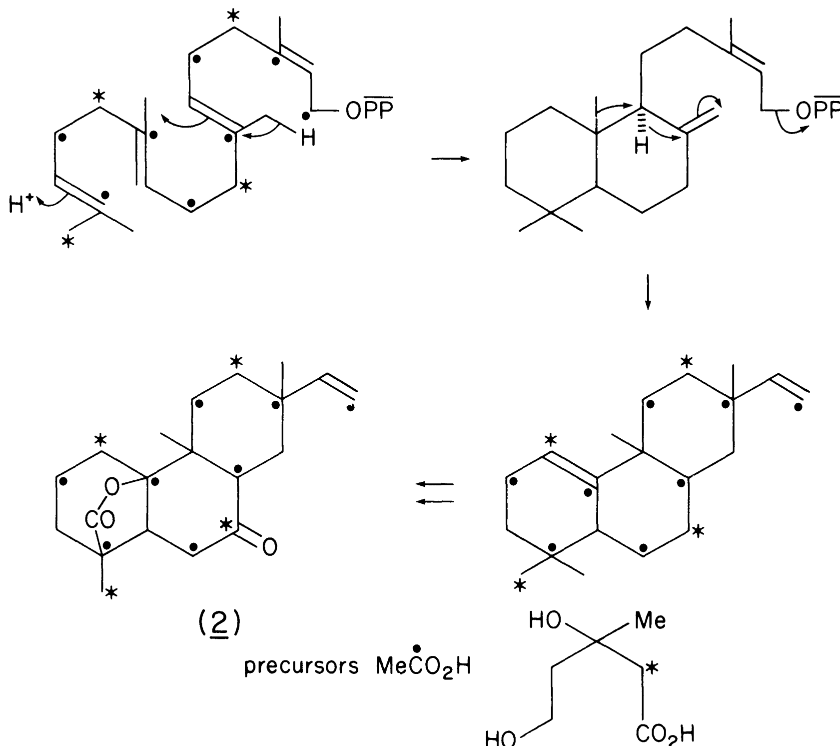
My polyketide hypothesis was formulated from 1951 onwards (Ref. 2) as an extrapolation of the then known biochemical origin of fatty acids from acetyl coenzyme-A. I tested the hypothesis by biochemical experiments using [^{14}C] techniques. The firm definition in 1955 (Ref. 3) of the origin of 6-methylsalicylic acid (1) from acetate, by location of the [^{14}C]-labelling pattern from [^{14}C]-acetate, was the first clear biochemical support for any organic chemical speculation. It was not until 1959 that Battersby (Ref. 4), for example, provided similar firm evidence for the origin of the benzylisoquinoline alkaloids according to the hypothesis of Winterstein and Trier (Ref. 5).

Our radiotracer work defined not only the origin of 6-methylsalicylic acid, but by implication that of all the other polyketides, the known number of which is now over 3000 (Ref. 6). The list contains most of the depsides and depsidones of lichens and the flavonoids and anthocyanins. Much subsequent biochemical work enlarged the scheme, and underlined its close resemblance to the fatty acid route.

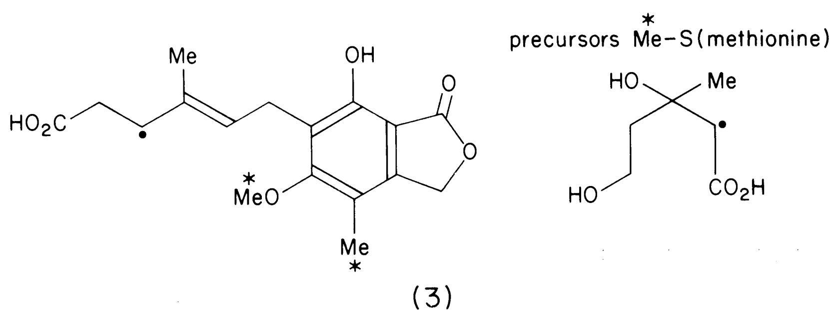


Chemical hypotheses involve all aspects of molecular transformations and the required chemical mechanisms, including how biosynthetic units are joined, and how the final molecules are formed by modifications of initial skeletons. Major hypotheses concerned with the latter aspect include carbonium-ion cyclisations and rearrangements of terpenoid chains. Various workers had made suggestions, notably Ruzicka (Ref. 7) for triterpenes and Wenkert (Ref. 8) for diterpenes, which were based on extrapolations of classical carbonium-ion chemistry.

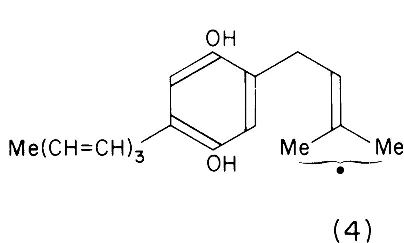
In 1958 we provided (Ref. 9) simultaneously with the Arigoni group (Ref. 10) the first biochemical proof of the validity of the idea of concerted cyclisation by examining the origins of the carbon atoms in the skeleton of rosenonolactone (2). Notably the lack of randomisation of the label from 2[^{14}C]-mevalonic acid between the carbons of what had begun as the terminal gem-dimethyl group of the diterpene chain conclusively proved the concerted nature of the cyclisation. Also the shift of Me from between rings A-B to between rings B-C confirmed an example of the expected carbonium-ion rearrangements. With the more complex example of the tetracyclic nor-diterpene gibberellic acid, our investigation (Ref. 9) carried out at the same time as that of rosenonolactone confirmed the validity of Wenkert's hypothesis. Henceforth there was a full confidence that theories of classical carbonium-ion chemistry could be applied in considering the origins of polyterpenoids. This conclusion was also linked to our demonstration, for the first time, that in the open-chain terpene, methylgeraniolene, from the mould product mycelianamide, the label from the 2-position of mevalonate appears only in that Me of the gem-dimethyl group which is *trans* to the chain (Ref. 11).



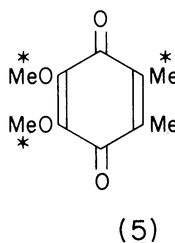
In 1954 (Ref. 12) I postulated biosynthetic processes involving introduction of Me or of isoprenoid chains as cations, on to nucleophilic carbon atoms in pre-formed skeletons. The suggestion involved active methionine for Me, and a source, then unknown, of terpenoid cations. The first support for the suggestions was provided by examination of mycophenolic acid (3) (Ref. 13). The degraded terpene side-chain was shown to come from mevalonic acid, and the nuclear Me from methionine. Later, in 1958, other substances including auroglaucin (4) (Ref. 14) and aurantiogliocladin (5) (Ref. 15) were proved to have the expected incorporations, and were correctly considered at the time to be models for the processes involved in ubiquinone biosynthesis, a suggestion later confirmed by other workers. These results and ideas led us in 1960 (Ref. 16) to postulate correctly substance (6) as the precursor of the ergot alkaloids, and subsequently led us to the work on echinulin and brevianamides discussed later.



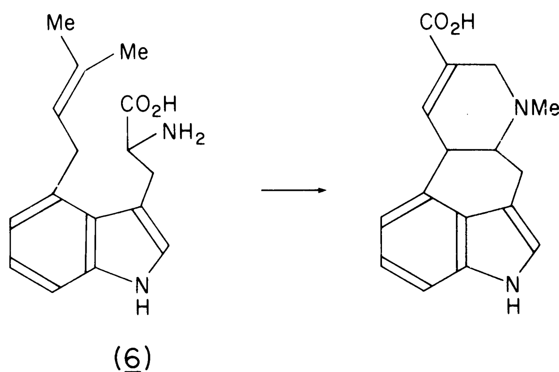
(3)



(4)



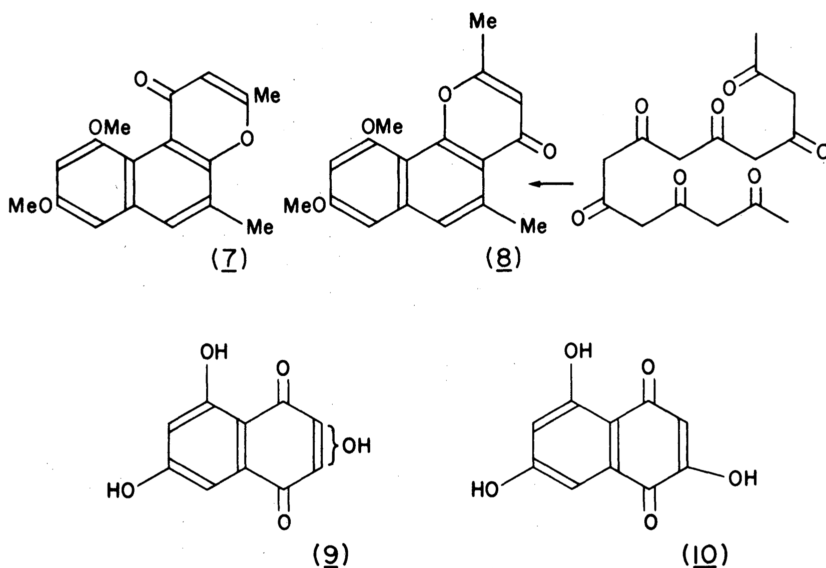
(5)



(6)

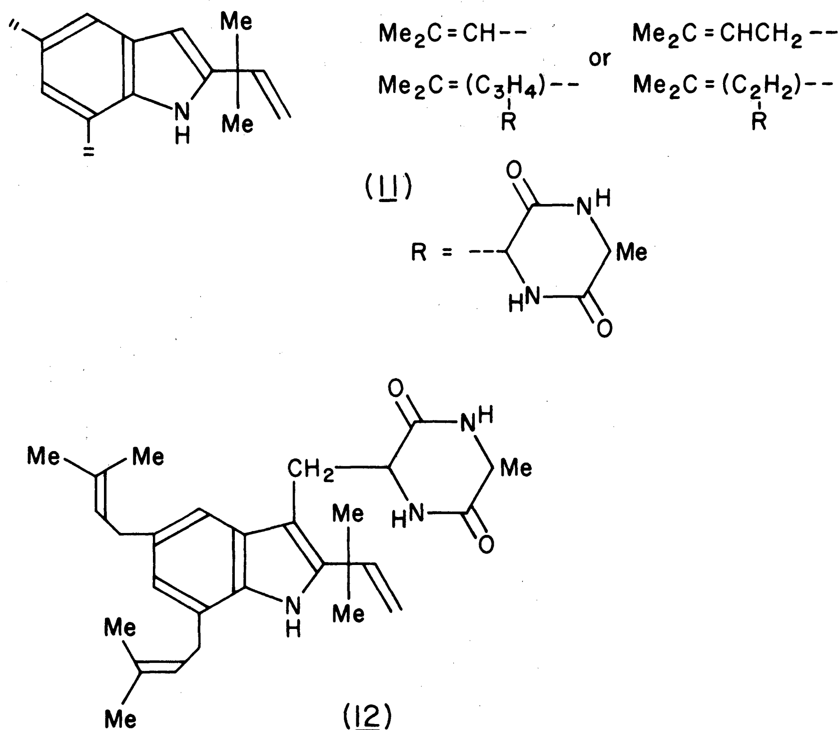
Before considering examples it is necessary to examine some fundamental interactions of chemical and biochemical ideas in relation to skeletal origins. There are many examples of biogenetic structure-determinations, notably of terpenes through the isoprene rule. Early examples in the polyketide area, relevant to the phomazarin case discussed later, were eleutherinol and flaviolin. The former was formulated as (7) by Schmid (Ref. 17), but inspection of the formula shows that if the molecule is a normal polyketide the formula should be (8). This was confirmed (Ref. 18). A related type of problem was represented by the necessity to complete the structures of flaviolin expressed on the basis of chemical work with the uncertainty shown in (9) (Ref. 19). If the substance is a polyketide as suggested by (9) the structure should be (10). This was shown to be correct (Ref. 20). A number of such instances carried out before the isotopic tracer work, convinced me of the validity of the polyketide hypothesis.

To speculate on structure in this way requires some biosynthetic information. This can be derived from chemical work as in (8) or (10), or by direct precursor examinations with the availability of isotopic incorporation techniques. Such incorporations not only indicate the units involved in a molecule, but labelling patterns produced by specifically labelled precursors can indicate the relations of different units and therefore of different parts of the skeleton, to each other. The earliest and still probably most complex examples of this approach assisted the structure-determination of the macrolide antibiotic nystatin (Ref. 21). With the advent of NMR techniques which can deal with [^{13}C] and [^{15}N]the practical task has been much facilitated but the principles are the same.



<u>Some indole derivatives from fungi</u>

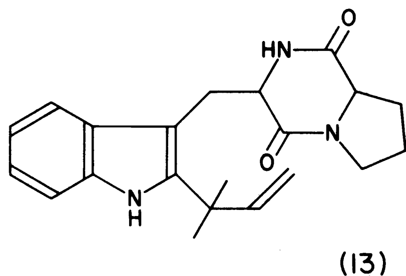
The structure of echinulin (12) was successfully defined by us (Ref. 22) on the basis of incorporations of tryptophan, [^{14}C]-mevalonic acid following biosynthetic speculation, based on some chemical information expressed at the time as (11) from the work of Quilico (Ref. 23). This involved revision of the empirical formula by adding CH_2. The same structure was independently suggested by Quilico on chemical grounds. The important feature for our present purpose is the presence of an indole nucleus with a reversed terpene unit attached to the 2-position.



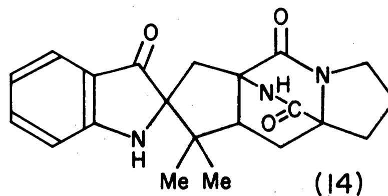
In studying the pigments of *Penicillium brevicompactum* we realised immediately that they belong to a new structural series. They were also available only in mg quantities. The first studied, brevianamide-A, has the formula $C_{21}H_{23}O_3N_3$ and a highly condensed ring-system. This fact, in combination with the presence of 3 O and 3 N made detailed interpretations in structural terms of both NMR and mass spectra very difficult. Outstanding features were chemical evidence for a ψ-indoxyl system, and the loss of C_5H_9 in the mass spectrum, which, combined with NMR evidence for CMe_2, suggested the presence of a terpene unit of some kind. The infra-red spectrum also suggested the presence of a diketopiperazine ring. The combination of suggestions: indole, terpene, diketopiperazine, led to considerations of structures of types based on the echinulin model. The chief problem was to explain the number of rings (six) which had to be present, simply from arithematical considerations of unsaturation. This feature seemed almost certainly linked to the modification of the diketopiperazine structure which prevented any amino-acids being detectable after hydrolysis. The suggestion of proline as the probable amino acid unit involved would help solve two difficulties: it provides one ring, and takes account of the fact that no more Me are visible in the NMR spectrum. The possible very hypothetical precursor at this stage is (13). The lack of the terpene double bond and of the two diketopiperazine H (from NMR) suggested the hypothetical ring-system of (14), in a possible structure for brevianamide-A. Although up to this stage we had done our best to check NMR spectra, for instance for the presence of the proline ring, the structure was still entirely hypothetical.

To test the hypothesis we examined the incorporation by the organism of [^{14}C]-tryptophan, mevalonate, acetate and proline into brevianamide-A. Incorporation occurred to a satisfactory extent although in the absence of known degradations we were not able to check its specificity. The result gave confidence in (14) and enabled detailed interpretations of NMR and mass-spectra to be made (Ref. 24).

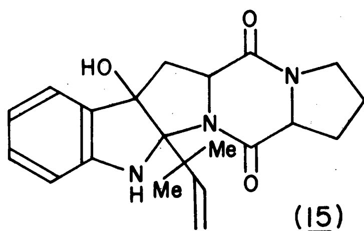
For details, the original paper should be consulted, but the point to be made is that the physical evidence was too complex to lead easily to the correct structure. Having arrived at the postulated biosynthetic structure, however, it could then be readily shown that the existing physical evidence accorded with it, and that it was possible to make leading suggestions to test it further. One example of interpretation is the loss of C_5H_9 as the major peak in the mass spectrum. This loss involves cleavage of three C-C bonds which would not normally be considered as probable. Inspection of (14) shows, however, that it should lead very readily, for structural reasons, to such cleavages. A minor product of the mould was later shown to be (15) (Ref. 25) closely related to the hypothetical precursor (13) by a simple oxidation step. Specific incorporation into the brevianamides of the precursor cyclo-L-tryptophyl-L-proline was later demonstrated (Ref. 26). The terpene unit must therefore have been introduced after formation of the diketopiperazine ring.



(13)



(14)



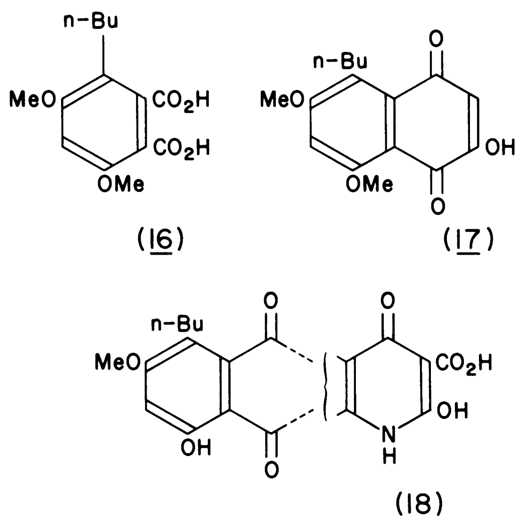
(15)

In more recent years the use of [^{13}C]-labelled precursors with NMR might have led more easily to similar results, since the specific incorporations could be detected without the necessity for first providing some chemical degradations.

Despite the convenience of [^{13}C] in avoiding the necessity for degradation experiments, and its power to indicate the linking of carbons in products in relation to their positions in precursors, there is still a place for [^{14}C]. The radioactive precursors are relatively unaffected by dilution, and their use before [^{13}C] incorporation is helpful. Also, because of the much larger amounts of precursors required, such as [^{13}C]-acetate, their use tends to produce effects which may vary from inhibition of fungal growth, to extensive randomisation of label due to massive involvement in metabolic cycles.

Phomazarin

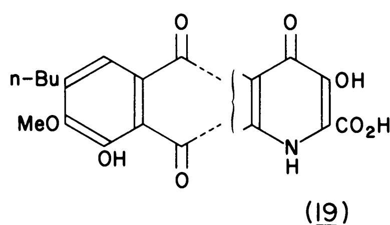
This purple mould metabolite, $C_{19}H_{17}O_8N$, was isolated by Kögl (Ref. 27). It is a carboxylic acid, with quinonoid properties, containing three OH and one OMe. By oxidation and methylation Kögl obtained what he claimed to be the phthalic acid (16), with a minor product thought to be (17), leading to (18) as a possible formula for phomazarin.



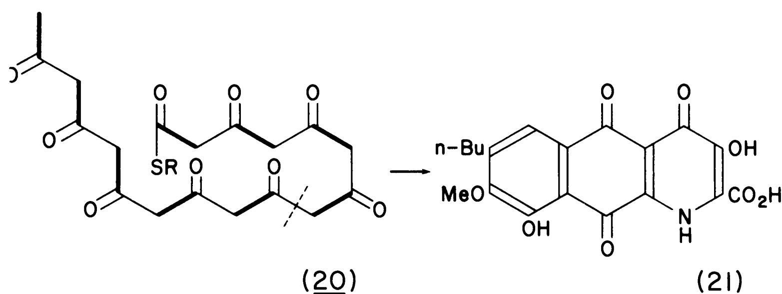
Our attention was first attracted to the substance because of the ambiguity in placing the terminal rings relative to each other. The situation is reminiscent of that of eleutherinol (8) and flaviolin (10) discussed, and if, as seemed probable, phomazarin is a polyketide, a decision might be possible on that basis. We commenced therefore both [^{14}C]-acetate incorporations, and further chemical work. The tracer work suggested (Ref.28) that the molecule comes from 9 acetate units, and that the CO_2H is derived from Me of acetate. This could have settled the question, but in the meantime, it became evident that some of Kögl's chemistry was highly questionable. It is not necessary to discuss these questions here, since the revisions involved standard chemical and spectroscopic procedures. They resulted in a revised nucleus (19) with the same problem of orientation of the end rings. The polyketide origin and the specific origin of CO_2H from acetate Me (Ref. 28) suggested as one possibility a precursor polyketide chain (20) and hence (21) for phomazarin rather than (22).

At this stage problems arose because interpretations of infra-red spectra seemed to favour (22). The carbonyl groups of the quinone have characteristic absorption positions according to structure, and in particular there was no carbonyl absorption observed which could be related to the quinone carbonyl of (21) which is uninvolved in hydrogen-bonding. Similar difficulties were encountered with phomazarin derivatives. Nevertheless the biosynthetic information favoured (21) so strongly that the spectroscopic conclusion was not accepted and a solution to the anomaly was sought. This was found on realising that the pyridone ring structure of (21) and (22) is an unconfirmed assumption, and that this ring might have a pyridol structure (23) instead. On this assumption the infra-red spectra can be reconciled.

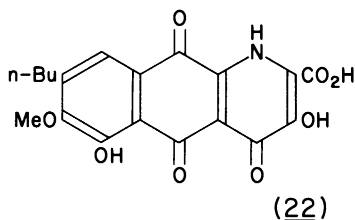
Incorporation of [^{13}C]-acetate and [^{15}N] as nitrate gave products which illuminated both structure and biosynthesis.



(19)



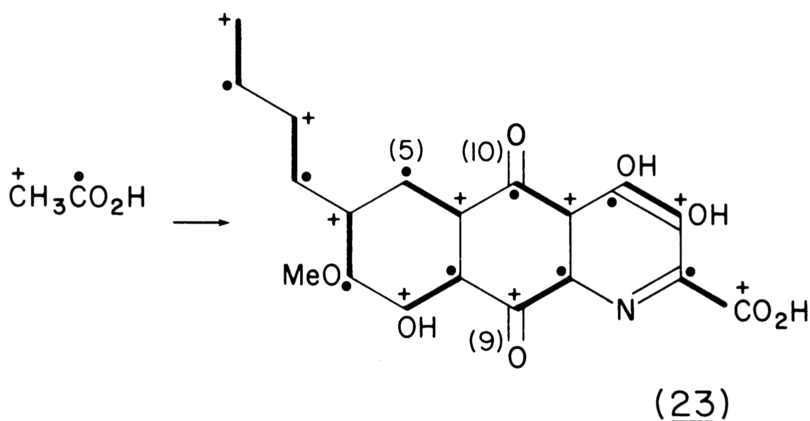
(20) (21)



(22)

The fully proton-coupled [^{13}C] NMR spectrum of the fully methylated phomazarin methyl ester was examined. The C-10 of (23) at 181.3 ppm showed a coupling of 4.4 Hz to the proton at C-5. In the proton-noise-decoupled spectrum of the [^{15}N] enriched compound, the C-9 at 178.9 ppm, but not the C-10, showed a [^{13}C] [^{15}N] coupling of 7.8 Hz, confirming (23). The opposite orientation would have C-9 coupled to both the proton on C-5 and the N. The [^{15}N] enriched substance also showed no direct [^{1}H] [^{15}N] coupling, ruling out the pyridone structure containing NH.

The spectra of the phomazarin resulting from incorporation of 1[^{13}C], 2[^{13}C], and 1,2[^{13}C]-acetate supported the labelling patterns shown in (20) and (23) indicating the formation of phomazarin from the one polyketide chain (20) with subsequent cleavage and an amination process (Ref. 29).



(23)

Novobiocin

A final topic concerns the use that may be made of a knowledge of biosynthetic pathways to tailor new metabolites, particularly antibiotics. I considered the theory of this at a conference in 1963 (Ref. 30) and quote from this. "The factors which apply to the stimulation of formation of a natural antibiotic by providing a normal intermediate should apply also to the production of a new antibiotic by feeding a modified intermediate, with an important addition. Competition will occur between the normal endogenous intermediate, to give the normal antibiotic, and the abnormal exogenous intermediate, to give a new one. If the enzymes can deal at all with the new series they might be expected to do so more slowly, although this might be countered in practice by a high concentration of the added compound. To remove this competition it should, in principle, be possible to mutate the organism so that it cannot complete the synthesis of the antibiotic unless given the normal precursor, which the mutant therefore cannot make, but can utilize. Provided it can utilize at all the altered intermediate, the mutant should then be able to complete the synthesis of the new antibiotic with no competitive formation of the natural one".

Such mutants can readily be found by treating inactive mutants with the metabolite, and choosing those which can then make antibiotic. An important consideration is that any antibiotic analogue can be directly examined for activity without rigorous purification, since it does not contain the normal antibiotic. In the case to be described the biological work was kindly carried out at Boots Pure Drug Company, Nottingham, by Drs P. Macey and M. Lumb.

Our attention was attracted to novobiocin (24 R=Me) as a vehicle to test the idea, for two reasons. One is that as a 'just-failed' antibiotic, some improvement of its properties might be critical to practical use. The second is that it is built of three rather separate molecular fragments, and a biosynthetic sequence involving isolable intermediates might render it particularly suitable to the mutant approach. Apart from the sugar, which was found by us to be added last, the rest of the molecule was obviously readily susceptible to synthesis, as were a range of analogues.

The use of [^{14}C]-tyrosine showed that both the acyl fragment A of (24 R=Me) and the coumarin ring B arise, as expected, through the shikimic acid route and the nuclear Me group from methionine (Ref. 31) (Ref 32).

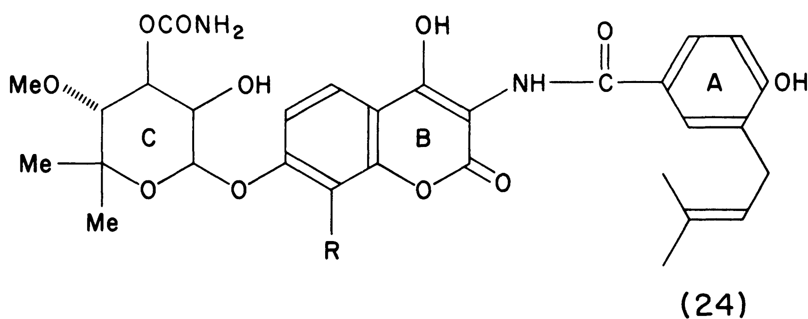
4-Hydroxybenzoic acid, although incorporated into an analogue of novobiocin does not have the C_5-unit introduced, and this unit is, surprisingly, not labelled from [^{14}C]-mevalonic acid or from acetate.

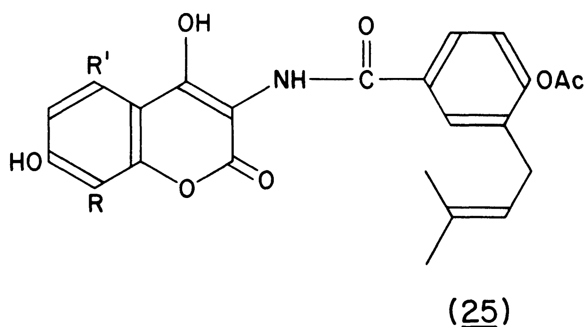
The sugar C was shown to arise directly from glucose, using 1[^{14}C] and 6[^{14}C]-glucose, one Me being introduced stereospecifically from methionine (Ref. 33). This was the first demonstration of C-methylation of a sugar. The sugar fragment, noviose, was not re-incorporated into the antibiotic.

The high incorporation of "novobiocic acid" (25 R=Me, R'=H) indicated that the organism could efficiently attach the correct sugar to this intermediate. Even more interestingly, the synthetic compound (25 R,R'=H) was incorporated both into novobiocin (24 R=Me) and a desmethylnovobiocin (24 R=H). The latter dominated if the intermediate were fed rapidly, and novobiocin itself if fed slowly. This led to two interesting conclusions, one that the C-methylation occurs at a very late stage of biosynthesis (Ref. 32) the other that the specificity of the enzyme which adds noviose is not very sensitive to the nature of the R group. Accordingly other intermediates with different R-groups were examined. These were conveniently fed as the acetyl derivatives of the 4-hydroxybenzoyl fragment, since it had previously been found that the nature of this acyl group makes no difference to the ability of the organism to add the noviose fragment. The acetate is lost in the fermentation.

Precursors (25 R=Et,Pr,Cℓ) were converted into analogues of novobiocin. The first two were not examined in detail, since the antibiotic activity seemed to be very low or non-existent. The intermediates (25 R,R'=Me and R=H,R'=Me) yielded active antibiotic, the latter producing two products, one identical with that from the former precursor, again indicating C-methylation even in an unnatural intermediate.

The most interesting analogue was the 8-chloro-compound (24 R=Cℓ) from (25 R=Cℓ,R'=H) which was isolated and examined in detail. It proved to be a powerful antibiotic, of the same order as novobiocin itself, although somewhat less active against some organisms. For example, inhibiting concentrations (μg/mℓ) against a few organisms are (novobiocin in brackets): *Streptococcus pyrogenes* 17 (2); *Proteus vulgaris* 64 (32); *Escherichia coli* 256 (128); *Aerobacter aerogenes* 4 (4); *Klebsiella pneumonia* 16 (16) (Ref. 34).
We are grateful to Dr B.P. Vaterlaus (Roche) for a gift of (25 R=Cℓ,R'=H).





The mutant which is blocked in the synthesis of a aminocoumarin fragment, accumulates isopentenyl-4-hydroxybenzoic acid. By a curious coincidence the nucleus with R=Cℓ was later found in a natural antibiotic with a rather different sugar (chlorobiocin) (Ref. 35). Using (25 R=Cℓ,R'=H) generated from this antibiotic, Dr O.K. Sebek (of the Upjohn Company, Kalamazoo) has confirmed our production of (24 R=Cℓ) using a similar mutant. I am grateful to him for unpublished information.

This work was discussed in lectures, initially at the 6th International Congress of Biochemistry in New York in 1964. Later workers, apparently independently since they do not quote our earlier discussion, have used mutants in the same way for other antibiotics formed by junction of cyclitols and sugars (e.g. Ref. 36). I am grateful to Drs D.T. Connor, A.R. Jones, P.W. Holloway, and R.W. Rickards for the work described here.

These stories all concern the intervention of classical organic chemical ideas. The future of biosynthesis probably lies more with detailed examinations of enzyme mechanisms by organic, inorganic and physical chemists.

REFERENCES

1. A.J. Birch, Interdiscip. Sci. Rev. 1, 215-233 (1976).
2. A.J. Birch and F.W. Donovan, Aust. J. Chem. 6, 360-368 (1953).
3. A.J. Birch, R.A. Massy-Westropp and C.J. Moye, Chem. Ind. 683-684 (1955); A.J. Birch, R.A. Massy-Westropp and C.J. Moye, Aust. J. Chem. 8, 539-544 (1955).
4. A.R. Battersby and B.J.T. Harper, Proc. Chem. Soc. 152 (1959).
5. E. Winterstein and G. Trier, Die Alkaloide, Gebr. Bornträger, Berlin (1910).
6. T.K. Devon and A.I. Scott, Handbook of Naturally Occurring Compounds, Volume 1, Academic Press, New York (1975).
7. L. Ruzicka in Perspectives in Organic Chemistry (Ed. A. Todd), Interscience, New York (1956).
8. E. Wenkert, Chem. Ind. 282-284 (1955).
9. A.J. Birch, R.W. Rickards, H. Smith, A. Harris and W.B. Whalley, Proc. Chem. Soc. 223 (1958); A.J. Birch, R.W. Rickards, H. Smith, A. Harris and W.B. Whalley, Tetrahedron 7, 241-251 (1959).
10. J.J. Britt and D. Arigoni, Proc. Chem. Soc. 224-225 (1958).
11. A.J. Birch, M. Kocor, N. Sheppard, and J. Winter, J. Chem. Soc. 1502-1505 (1962).

12. A.J. Birch, P. Elliott and A.R. Penfold, Aust. J. Chem. 7, 169-172 (1954).
13. A.J. Birch, R.J. English, R.A. Massy-Westropp, M. Slaytor and H. Smith, Proc. Chem. Soc. 204 (1957); A.J. Birch, R.J. English, R.A. Massy-Westropp and H. Smith, Proc. Chem. Soc. 233 (1957).
14. A.J. Birch, J. Schofield and H. Smith, Chem. Ind. 1321 (1958).
15. A.J. Birch, R.I. Fryer and H. Smith, Proc. Chem. Soc. 343 (1958).
16. A.J. Birch, Chem. Weekblad 56, 597-602 (1960).
17. A. Ebnöther, T.M. Myer and H. Schmid, Helv. Chim. Acta 35, 910-928 (1952).
18. A.J. Birch and F.W. Donovan, Aust. J. Chem. 6, 373-378 (1953).
19. B.D. Astill and J.C. Roberts, J. Chem. Soc. 3302-3307 (1953).
20. A.J. Birch and F.W. Donovan, Aust. J. Chem. 8, 529-533 (1955).
21. A.J. Birch, C.W. Holzapfel, R.W. Rickards, C. Djerassi, M. Suzuki, J. Westley, J.D. Dutcher and R. Thomas, Tetrahedron Lett. 23, 1485-1490 (1964); A.J. Birch, C.W. Holzapfel, R.W. Rickards, C. Djerassi, P.C. Seidel, M. Suzuki, J.W. Westley and J.D. Dutcher, Tetrahedron Lett. 23, 1491-1497 (1964).
22. A.J. Birch, G.E. Blance, S. David and H. Smith, J. Chem. Soc. 3128-3131 (1961).
23. A. Quilico, C. Cardoni and F. Piozzi, Gazz. chim. ital. 85, 3-33 (1955); A. Quilico, C. Cardoni and F. Piozzi, Gazz. chim. ital. 86, 211-233 (1956); A. Quilico, F. Piozzi and C. Cardoni, Gazz. chim. ital. 88, 125-148 (1958).
24. A.J. Birch and J.J. Wright, Tetrahedron 26, 2329-2344 (1970).
25. A.J. Birch and R.A. Russell, Tetrahedron 28, 2999-3008 (1972).
26. A.J. Birch, J. Baldas and R.A. Russell, J. Chem. Soc. Perkin Trans I, 50-52 (1974).
27. F. Kögl, G.C. van Wessem and O.I. Elsbach, Rec. Trav. Chim. 64, 23-29 (1945).
28. A.J. Birch, R.I. Fryer, P.J. Thomson and H. Smith, Nature 190, 441 (1961).
29. A.J. Birch, R. Effenberger, R.W. Rickards and T.J. Simpson, Tetrahedron Lett. 27, 2371-2374 (1976).
30. A.J. Birch, Pure Appl. Chem. 7, 527-537 (1963).
31. C.A. Bunton, G.W. Kenner, M.J.T. Robinson and B.R. Webster, Tetrahedron 19, 1001-1010 (1963).
32. A.J. Birch, D.W. Cameron, P.W. Holloway and R.W. Rickards, Tetrahedron Lett. 25, 26-31 (1960).
33. A.J. Birch, P.W. Holloway and R.W. Rickards, Biochim. Biophys. Acta 57, 143-145 (1962).
34. Personal Communication, P. Macey and M. Lumb.
35. Rhone Poulenc SA, French Pat. 1968, 139878, 160462.
36. W.T. Thomas, K.L. Rinehart and D. Gottlieb, Proc. Nat. Acad. Sci. U.S. 63, 198-204 (1969).