THE CHEMISTRY OF THE APHID COLOURING MATTERS

LORD TODD

University Chemical Laboratory, Cambridge, U.K.

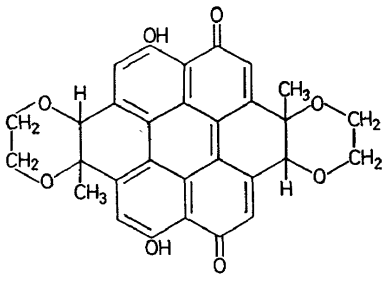
Some fifteen years ago, for reasons into which I shall not enter here, we commenced in Cambridge a study of the colouring matters present in the haemolymph of dark-coloured insects belonging to the family *Aphididae*. Our preliminary studies were made on the common bean aphid (*Aphis fabae*) which is widely distributed in Europe where its most common summer host is the cultivated broad bean *Vicia faba*, but they were later extended to cover some twenty species of dark-coloured aphids from which the same or closely related colouring matters were obtained. We do not as yet know the exact distribution of these colouring matters—to which the general name aphins has been given—among the *Hemiptera* but they are apparently not confined to the *Aphididae* but occur also in at least one member of the *Phylloxeridae* (*Adelges strobii*).

Investigation of these colouring matters quickly revealed a complex and interesting situation. In the blood of the living insect there is present a golden yellow glucosidic pigment, *protoaphin*, which shows a reddish-purple colour in alkaline solution. On the death of the insect or even upon injury protoaphin is rapidly converted by a specific enzyme system present in all these insects into a yellow fat-soluble substance *xanthoaphin* which shows a strong green fluorescence; this change is accompanied by loss of glucose. Xanthoaphin itself is unstable and passes on standing, or very rapidly in presence of acids or alkalis, into the similarly unstable orange *chrysoaphin*. In similar fashion the latter undergoes conversion to the red, strongly fluorescent *erythroaphin*, the relatively stable end product of the aphin series. This series of changes from protoaphin to erythroaphin we found in all the species of dark aphids we examined, but the colouring matters fell into two series, which differ stereochemically from each other and which we describe as the aphins-*fb* and the aphins-*sl*. The descriptions *-fb* and *-sl* derive from the name of the particular aphid species from which the pigments were first isolated, *i.e.* the aphins-*fb* from the bean aphid *Aphis fabae* and the aphins-*sl* from the willow aphid *Tuberolachnus salignus*. There exist, then, a protoaphin-*fb*, xanthoaphin-*fb etc.*, and a stereoisomeric series protoaphin-*sl*, xanthoaphin-*sl etc.* The relationships between these series are set out in the following table:

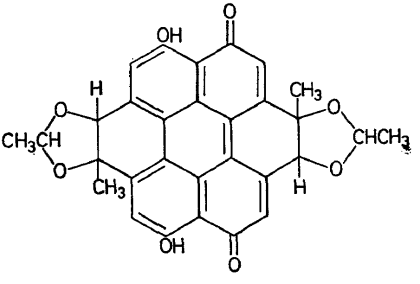
Protoaphin-*fb*	$C_{36}H_{38}O_{16}$	Protoaphin-*sl*
↓	↓	↓
Xanthoaphin-*fb*	$C_{30}H_{26}O_{10}$	Xanthoaphin-*sl*
↓	↓	↓
Chrysoaphin-*fb*	$C_{30}H_{24}O_{9}$	Chrysoaphin-*sl*-I Chrysoaphin-*sl*-II
↓	↓	↓
Erythroaphin-*fb*	$C_{30}H_{22}O_{8}$	Erythroaphin-*sl*

On this table the following observations should be noted. Corresponding pigments in the two series are stereoisomeric and indeed erythroaphin-*sl* can be converted irreversibly to erythroaphin-*fb* by means of alkali. The occurrence of two isomeric chrysoaphins in the *sl*-series is also stereochemical in origin and will be discussed later. It should also be emphasized here that the formation of xanthoaphin from protoaphin is much more complex than a simple removal of glucose from the latter. Although xanthoaphin is readily obtained from protoaphin by enzymic action, acid hydrolysis of protoaphin gives at best only traces of xanthoaphin, the main product being a brown, partly polymeric material which cannot be converted to erythroaphin.

Work on the isolation and properties of the aphins has already been published in a series of papers[1] and will not be discussed here. At the beginning of our studies we decided to attempt to establish the structure of the stable end-product erythroaphin and, thereafter, to investigate more closely the labile precursors. An extensive series of investigations led us in 1954 to advance two possible structures for erythroaphin of which (I) was preferred as the pigment was more stable to acid than would have been expected on the basis of structure (II); only by very vigorous acid treatment can acetaldehyde be produced from erythroaphin.

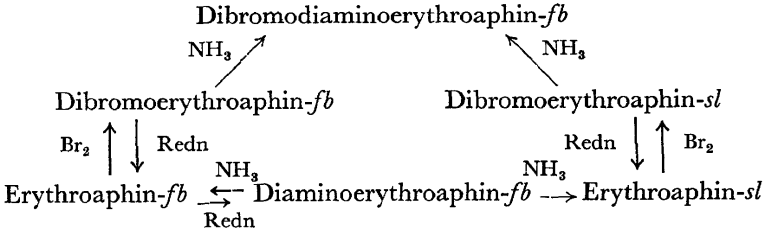


(I)



(II)

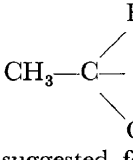
Structure (I) could not, however, be regarded as established. The presence of the perylene skeleton was certain—erythroaphin is without doubt a derivative of 4,9-dihydroxyperylene-3,10-quinone. Erythroaphin contains in addition 4 ether-oxygens, gives acetaldehyde on heating with 60 per cent sulphuric acid and yields nearly 4 mols acetic acid on Kuhn–Roth oxidation. The validity of structure (I), however, rested on the fundamental assumption that in erythroaphin there were four substitutable positions in the perylene nucleus. This basic assumption seemed at the time reasonable in view of the reactions of the erythroaphins, *e.g.* those shown below:



These reactions are precisely what one would expect of a dihydroxyperylene-quinone with 4 free positions in the nucleus, and the properties of the bromine atoms and amino-groups in these derivatives seemed to accord well with this view. If the assumption of 4 free " aromatic " positions in erythroaphin is made, then one is forced to structure (I) or (II) for the pigment.

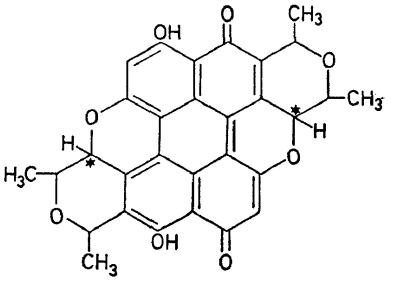
It was clear, however, that further investigation was necessary, for there were features of the chemistry of erythroaphin which were not too easily equated with structure (I). Thus, on the basis of this structure, the production of almost 4 mols acetic acid on Kuhn–Roth oxidation was surprising. Moreover the ultra-violet absorption spectrum of tetra-acetyldihydroerythroaphin, although very similar in form to that of perylene, nevertheless shows a large bathochromic shift of nearly 50 mμ; such a shift would be more readily explicable if erythroaphin contained one or even two substituents attached through oxygen to the perylene ring system. Furthermore, in considering the precursors of erythroaphin, although it was possible to derive plausible structures for chrysoaphin and xanthoaphin, it was virtually impossible to derive a reasonable formula for protoaphin (which spectroscopically appeared to be a naphthaquinone derivative) compatible with (I). We have, therefore, undertaken during the past few years, a complete re-investigation of the problem and we can now present a solution to the whole aphin problem in which all the chemical and stereochemical features of the pigments can be regarded as established with the exception of the configuration at the glycosidic centre in protoaphin.

We commenced these new investigations by applying the techniques of infra-red and proton magnetic resonance spectroscopy to the determination of the number of unsubstituted positions in the perylene nucleus of erythroaphin. The results obtained showed clearly that there were only two such positions; dibromoerythroaphin, for example, contains no aromatic or ethylenic CH-groups. It was further established that each of the 4 side-chain methyl groups in erythroaphin is attached to a carbon atom bearing one hydrogen and one oxygen atom, *i.e.* that there are four groupings

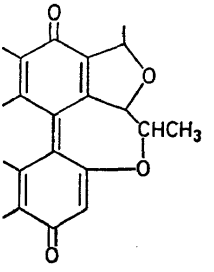


in the molecule. A substantial number of formulae might be suggested for erythroaphin on the basis of these facts, but taking other chemical evidence into account structures (III) and (IV) seemed perhaps the most likely (the seven-membered ring in (IV) can be shown to be stereochemically feasible on models). It may be observed that the dimethyldihydropyran rings present in formula (III) are known to occur in the plant products eleutherin and isoeleutherin whose structures have been established by Schmid and his co-workers[2].

Parallel with these studies on erythroaphin, using physical methods, we also undertook a study of the chemical degradation of protoaphin-*fb*. On treatment with sodium dithionite the molecule split into two parts. The reaction mixture oxidized rapidly in air, and, thereafter, the two breakdown products could be isolated in crystalline form. One of these, substance A,



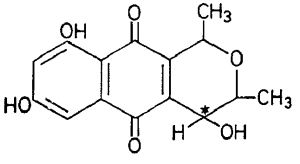
(III)



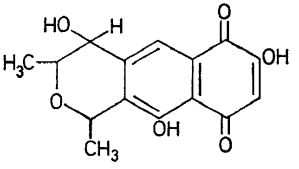
(IV)

was an orange-yellow quinone of formula $C_{15}H_{14}O_6$ and the second a phenolic glycoside B, $C_{21}H_{26}O_{10}$, which was hydrolysed by the crude enzyme isolated from *A. fabae* to give glucose (1 mol) and an unstable, readily oxidized phenol. Substance A was a derivative of 5,7-dihydroxy-naphtha-1,4-quinone, bearing carbon substituents in positions 2 and 3; it contained an aliphatic hydroxyl group, one ethereal oxygen and two C—CH_3 groupings whose identity of location with those in erythroaphin, eleutherin and isoeleutherin was apparent from proton magnetic resonance studies. These facts, together with the results of oxidative degradation, established the formula of substance A as (V).

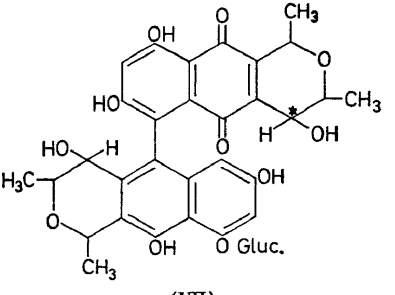
$$C_{36}H_{38}O_{16} \longrightarrow \underset{\text{A}}{C_{15}H_{14}O_6} + \underset{\text{B}}{C_{21}H_{26}O_{10}}$$



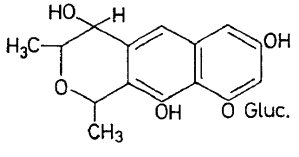
(V)



(Va)



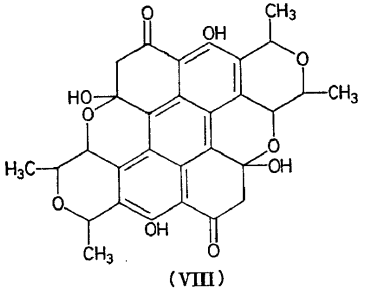
(VII)



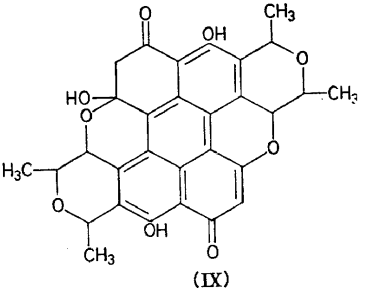
(VI)

On oxidizing the glucoside B with Fremy's salt two quinones were produced, only one of which still contained glucose; this on enzymic hydrolysis yielded glucose and substance A (V). The second product, which was sugar-free, had the properties of a derivative of 2,5-dihydroxynaphtha-1,4-quinone. Since (V) can be produced from B this sugar-free quinone must have structure (Va) and hence the glycoside B can be formulated as (VI).

The absorption spectrum of protoaphin represents a simple summation of the spectra of A and B and we therefore formulate protoaphin as (VII). The reductive fission of protoaphin with dithionite has its analogy in the chemistry of the fungal pigment skyrin[3]. Given that (VII) represents protoaphin, what is the structure of xanthoaphin $C_{30}H_{26}O_{10}$ which differs formally from (VII) by loss of 1 mol glucose and 1 mol water? Chemical degradative studies on xanthoaphin are well-nigh impossible since it changes with such extreme ease into chrysoaphin and thence into erythroaphin, but we were able to establish by spectroscopic methods that it is in fact an anthracene derivative. On this basis and bearing in mind the various interconversions observed in the pigments of the skyrin group[4] we propose for xanthoaphin structure (VIII).



(VIII)



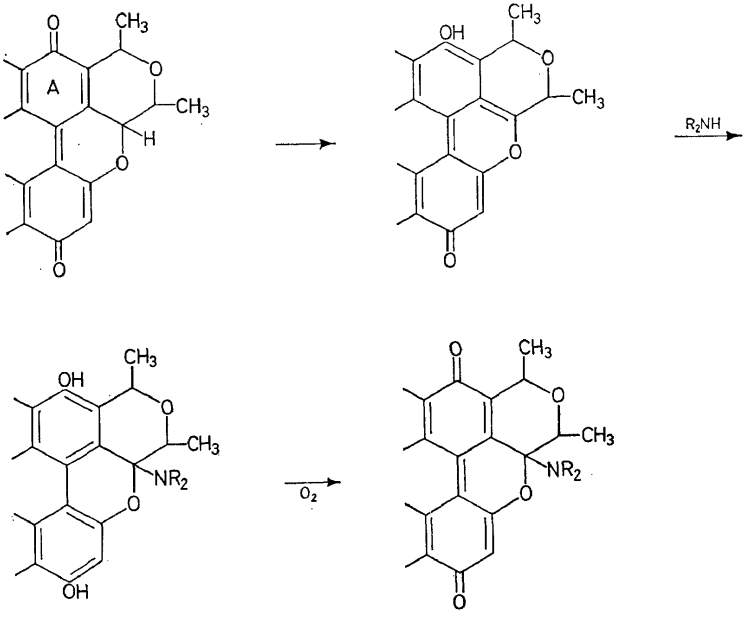
(IX)

It is clear from formulae (VII) and (VIII) that the relative stability of protoaphin depends on the presence in it of the glucose residue. Only when it is removed can the necessary condensation between the quinone portion of protoaphin and the lower phenolic ring occur. Similarly, the failure of acid hydrolysis to yield more than traces of xanthoaphin from protoaphin is understandable; for it is probable that in solution the protoaphin molecule will adopt a configuration in which the positions in the two rings which must be joined to form xanthoaphin are not in close proximity to one another. Presumably part of the function of the enzyme system present in the insect is to hold the protoaphin molecule in such a configuration that ring closure can occur readily after the glucose has been removed. Whether this ring-closure precedes or follows the ring-closure by acetalization on either side of the molecule is at present unknown. The instability of xanthoaphin itself is also readily understood; loss of 1 mol water converts it to chrysoaphin (IX) and this by loss of a further molecule of water passes into the stable erythoaphin (III).

We must now consider how, on the basis of these formulae, we can explain the remarkable amination reactions of the erythroaphins and the stereochemical relationship between compounds of the *fb* and *sl* series. Since the

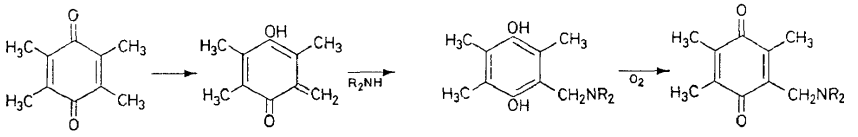
amination of erythroaphin-*fb* or -*sl* leads invariably to the formation of a derivative of the *fb* series, the two problems are clearly interrelated. It will be convenient therefore to discuss first the amination of erythroaphin and then the difference between the *fb* and *sl* isomers.

As already mentioned, erythroaphin yields on bromination a dibromo-erythroaphin in which both bromine atoms are located in the perylene nucleus. This substance, on standing with piperidine, gives a dibromo-dipiperidinoerythroaphin, in which the piperidine residues must be located somewhere in the non-aromatic part of the molecule. The piperidino-groups are not however introduced by simple substitution, for their introduction is accompanied by the uptake of two molecules of oxygen; in the absence of oxygen no dibromo-dipiperidinoerythroaphin is produced. This introduction of piperidino-groups thus parallels the well-known reaction of amines with quinones, in which addition of amine to the conjugated system occurs followed by aerial oxidation of the initially formed hydroquinone. In our view the reaction in the case of erythroaphin (or dibromo-erythroaphin) is of this type and involves a tautomeric form of the molecule as set out in the following scheme (partial formulae are used for the sake of brevity):



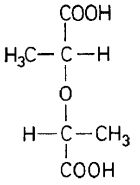
This formulation of the amino- and piperidino-compounds is fully supported by their proton magnetic resonance spectra, but we have also found some interesting analogies to their formation. It seemed *a priori* likely that the remarkable amination reaction depends on the fact that in erythroaphin ring A (above scheme) is a fully substituted quinone system. In that case we would expect similar reactions to occur with simpler fully substituted quinones such as duroquinone or 2,3-dimethylnaphthaquinone.

This we find to be the case; these and similar quinones react readily with piperidine, piperidino-groups being introduced into the side chains. Thus with duroquinone four piperidino-groups can be successively introduced according to the following scheme:

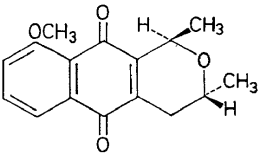


It has already been mentioned that amination or hydroxylation of erythroaphin-*sl* is accompanied by an irreversible conversion to the *fb* series, and a similar conversion without substitution can be brought about by alkali alone. These facts are most readily explained by assuming an inversion of configuration at one or both the centres marked with an asterisk in formula (III). Easy racemization at such centres in a complex quinone would be expected on general grounds; that inversion rather than racemization occurs indicates that one configuration in erythroaphin is considerably more stable than the other. It can reasonably be concluded too, that the only difference between erythroaphin-*fb* and erythroaphin-*sl* is in the configuration at one or both of these centres.

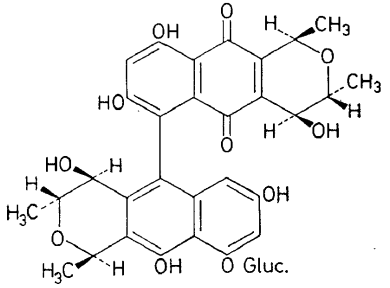
A complete solution of the stereochemical problem was provided by further study of the chemistry of protoaphin-*fb* and protoaphin-*sl*. On fission with dithionite both pigments yield one and the same glucoside B (VI); the quinones A (V) from the two protoaphins are however stereoisomeric. It will be recalled that glucoside B (V) (obtained from either protoaphin-*fb* or -*sl*) yields by oxidation and subsequent removal of glucose the stereoisomer of A (V) which is produced directly from protoaphin-*fb* by the dithionite fission. It follows, therefore, that the only difference between protoaphin-*fb* and protoaphin-*sl* as well as in the two quinones (V) must lie in the configuration at the carbon atom marked with an asterisk in formulae (V) and (VII). By application of proton magnetic resonance spectroscopy it was concluded that in glucoside B and in the quinone A from protoaphin-*fb* the alcoholic hydroxyl group is situated *trans* with respect to the neighbouring methyl group; in the A stereoisomer from protoaphin-*sl* it has the *cis* orientation. In order completely to clarify the stereochemistry of the aphins (apart from the configuration at the glycosidic centre in the protoaphins) it is necessary to know the orientation of the methyl groups in the molecule of protoaphin. This was settled by gentle chromic acid oxidation of the quinone A; the quinone residue was destroyed and the optically active dilactic acid (X) was isolated whose absolute configuration is already known. It follows, therefore, that the methyl groups in each of the dihydropyran rings in the aphins are in the *trans* relationship to one another, *i.e.* the orientation of the methyl groups is the same as that found in isoeleutherin (XI). As a result we can now finally formulate protoaphin-*fb* as (XII) and protoaphin-*sl* as (XIII).



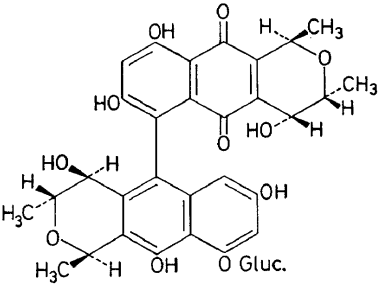
(X)



(XI)

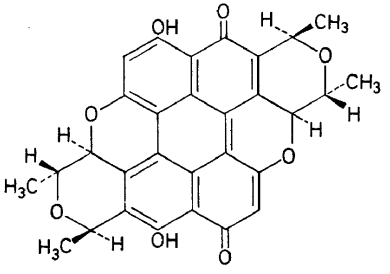


(XII)

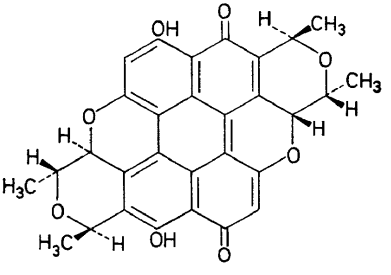


(XIII)

If, in the light of formulae (XII) and (XIII), we now consider the xanthoaphin formula (VIII) it is clear that from protoaphin-*fb* only one xanthoaphin-*fb* can be produced and from this by loss of one molecule of water only one chrysoaphin-*fb* (IX). Protoaphin-*sl* similarly can yield only one xanthoaphin-*sl* but this must be unsymmetrical and could thus give rise to two different chrysoaphins-*sl* according to which side of the molecule (VIII) suffered loss of water; both these chrysoaphins would, of course, yield one and the same erythroaphin-*sl*. In fact, the two postulated chrysoaphins-*sl* have been isolated. Finally, we can now formulate erythroaphin-*fb* as (XIV) and erythroaphin-*sl* as (XV).



(XIV)



(XV)

The orientation of the two important " angular " hydrogen atoms in (XIV) and (XV) rests on evidence from proton magnetic resonance spectra of protoaphin derivatives. It is possible on this basis to understand the ready irreversible conversion of *sl* to *fb* derivatives in the erythroaphin group; for in erythroaphin-*fb* the oxygen bridges to the aromatic system and

the neighbouring methyl groups on both sides of the molecule are situated equatorially with respect to the dihydropyran ring systems, making the *fb* arrangement thermodynamically the more stable.

On the basis of formulae (XIV) and (XV) it is clear that a third stereoisomeric erythroaphin should exist in which both " angular hydrogens " are *trans* to the adjacent methyl groups. This third isomer has not so far been found in nature, but it has been prepared in the laboratory. When solutions of tetra-acetyldihydroerythroaphin-*fb* or -*sl* in inert solvents are irradiated with ultra-violet light, an equilibrium mixture is produced from which three distinct tetra-acetyldihydroerythroaphins can be isolated yielding on hydrolysis and oxidation respectively erythroaphin-*fb* and erythroaphin-*sl* and a new erythroaphin which as expected is readily convertible to erythroaphin-*fb* with alkali.

The formulae here proposed for the aphid colouring matters of the *fb* and *sl* series are in full agreement with all experimental observations and can be regarded as reasonably established, although final proof of their correctness must await total synthesis. The total synthesis of the aphins poses a number of interesting problems and we hope to report later the results of our efforts in this direction. Viewed as a whole the aphin problem is striking in its complexity and the researches to date leave many interesting questions still unanswered. The series of changes leading from protoaphin to erythroaphin may well represent the method used in nature for the synthesis of perylene derivatives, but we have as yet no knowledge of the biosynthesis of the protoaphins. They probably arise by the oxidative coupling of two naphthalene derivatives, but whether the insect synthesises the latter from acetate units or whether it obtains preformed naphthalene derivatives from the plant sap on which it feeds is unknown. It should be mentioned that pigments other than protoaphins occur in certain aphid species. Some of these are at present under investigation in Cambridge, but it is as yet too early to discuss their structure. Nothing is yet known of the function of protoaphin in the haemolymph, although in some species it occurs in very large amount (1–2 per cent by weight) and it is difficult to believe that it is simply a waste product. It is clear that much remains to be done and that the colouring matters of the *Aphididae* offer a fascinating field of study for chemists and biochemists.

Finally, I would mention that most of the newer work here reported on the structure and stereochemistry of the aphins and the course of amination in erythroaphin derivatives is as yet unpublished. The completion of this work is due mainly to the efforts of my colleagues Dr D. W. Cameron, Mr P. M. Scott and Mr D. G. I. Kingston. To them and to Dr N. Sheppard, who provided invaluable help on the spectroscopic side, I am deeply indebted.

References

[1] A. R. Todd *et al. Nature* **162**, 79 (1948); *J. Chem. Soc.* **1950**, 477, 485, 3304; **1951**, 2633; **1953**, 4925, 4928; **1954**, 107, 1280, 1285, 2378; **1955**, 954, 959, 1144; **1956**, 4091.
[2] H. Schmid and A. Ebnöther. *Helv. Chim. Acta.* **34**, 561, 1041 (1951).
[3] S. Shibata, O. Tanaka, and I. Kitagawa. *Pharm. Bull.* (*Tokyo*) **3**, 278 (1955); B. H. Howard and H. Raistrick. *Biochem. J.* **56**, 56 (1954)
[4] O. Tanaka. *Chem. & Pharm. Bull.* (*Tokyo*) **6**, 203 (1958).