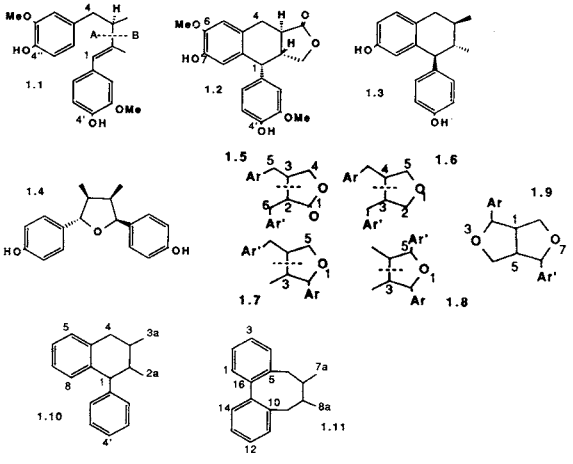
1

Introduction

The lignans form a group of plant phenols whose structure is determined by the union of two cinnamic acid residues or their biogenetic equivalents.

This unifying definition was first made by R.D. Howarth (1936) and is illustrated by the structures (**1.1**) and (**1.2**), in which the bar line separates the two *β*,*β*-linked cinnamic residues. Thus guaiaretic 'acid' (**1.1**), with the terminal groups fully reduced, is structurally one of the simplest members of the group; whilst conidendrin (**1.2**; Lindsey, 1892) is more complex in that the terminal groups differ and are also at higher oxidation levels. Other common variations occur naturally in which the aromatic substituents are modified in whole or in part to methylenedioxy and also where additional oxysubstitution takes place with or without O-methylation. The orientation of these substituents is also subject to variation. No lignan has been isolated with an unsubstituted phenyl ring: attenuol (**2.290**; Joshi



et al., 1978), the compound (**1.3**; Vieira *et al.*, 1983) and a group of furans (e.g. **1.4**) isolated from *Krameria cystisoides* (Achenbach *et al.*, 1987) are the only monosubstituted examples. Removal of phenolic hydroxyl groups *in vitro* is known to be a difficult energy-demanding process. One is not therefore surprised to find that although reduction with deoxygenation proceeds readily *in vivo* in the side chain, it does not do so in the aryl groups.

The ubiquity of aromatic oxysubstituents especially those *para*-related to the side chain gave the lead to Erdtmann (1933) who postulated a mode of lignan biogenesis. He proposed that ionisation of the phenolic hydroxyl group followed by a single electron transfer produces a mesomeric radical (**1.12**; Scheme 1.1).

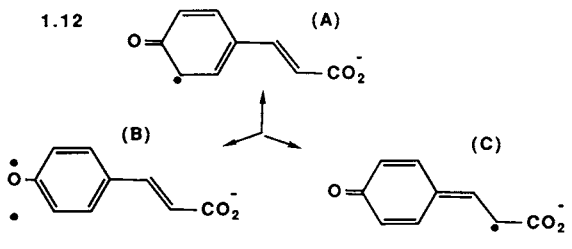
This may dimerise to form structures such as (**1.1**). The point of insertion of substituents (X, Scheme 1.2) and of modification of the terminal groups are still largely matters of conjecture; they will be discussed more fully in Chapter 7.

The principal lignan structures that may be formed as linked dimers are shown (**1.5–1.11**). They are all known with a variety of oxysubstitution patterns in the aromatic rings and hydroxyl and glycosyl substituents may also be inserted in the side chain.

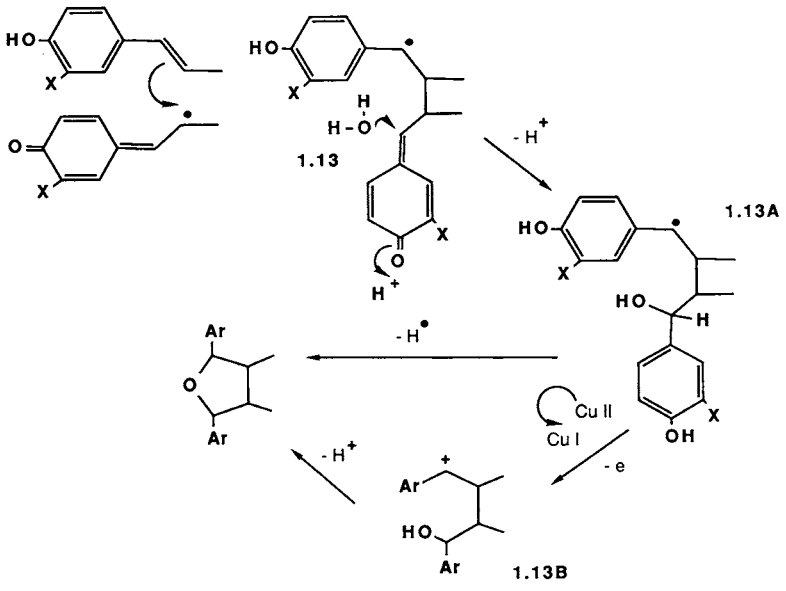
One possible biogenetic route to furans (**1.8**) is the reductive hydroxylation of a quinone methide (**1.13**) formed during initial dimerisation, followed by intramolecular attack upon the inserted −OH group (in **1.13A**) by the residual radical at the γ-position. Lignan alcohols hydroxylated at benzylic positions in this way are known.

There is a limited amount of evidence (Chapter 7) that oxygenation of the aromatic rings is complete before coupling takes place *in vivo*, but by way of illustration one can envisage insertion of a phenolic hydroxyl group as shown in Scheme 1.3. Here the combination of acid or enzymic catalysis together with nucleophilic attack by a water molecule upon the *bis*-quinone methide (**1.14**) leads to the formation of a tetralin (**1.15**). This

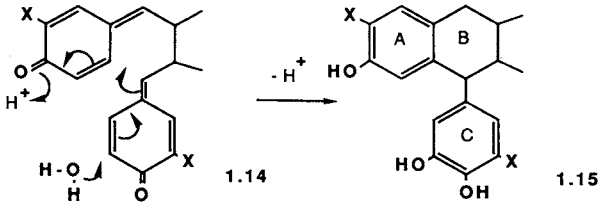
Scheme 1.1 The mesomeric cinnamate radical



Scheme 1.2 Hypothetical mechanism for the elaboration of a coupled dimer



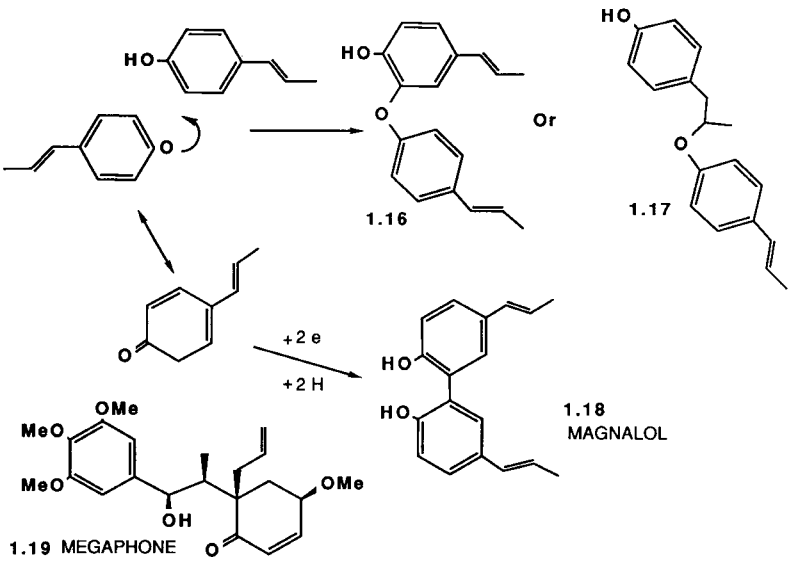
Scheme 1.3 A possible cyclisation mode for a bis-quinone methide



pattern of pendent ring C-substitution occurs in podophyllotoxin (**1.36**; X = H), burseran (**1.30**) and epimagnolin (**1.31**) and in numerous other lignans. It could also be maintained that these postulated pathways follow the transfer of a second electron from **1.12** or **1.13** to an acceptor such as Cu(II), with subsequent changes proceeding through a mesomerically stabilised carbocation. It is also evident that either intermediate type may dimerise in other ways and so give rise to structures such as **1.16–1.18.**

These three dimeric structures do occur naturally, together with some dozen others none of which are β,β-linked and which are known as neolignans (Gottlieb, 1978; Whiting, 1987). The structure **1.18** is found in nature as magnalol; it was isolated in 1983 from *Sassafras taiwanensis* (Watanabe *et al.*, 1983) and shows depressant activity on the central nervous system

Scheme 1.4 Alternative linking modes for the cinnamate radical



(CNS). Another product of the attack of the cinnamyl precursor by the mesomeric radical is the structure **1.19** which corresponds to megaphone (Kupchan *et al.*, 1978), which is active against human carcinoma of the nasopharynx.

The neolignans, although more diverse in structural type, are less numerous than the classical β,β-linked lignans and to date they have been isolated only from the *Magnoliales* and *Piperales*. The monomeric allyl- and propenylphenols are also found in plants of these orders (Gottlieb, 1974). In view of the existence of thorough reviews (Gottlieb, 1978; Whiting, 1987) and of the need to restrict this text to reasonable bounds the neolignans have been given only a limited mention. It should however be pointed out that the material on nomenclature is relevant to them as also is much of the discussion on structure determination in Chapter 6.

Nomenclature

The problem may be defined by reference to structures **1.1**, **1.2** and **1.5–1.11** in which the carbon atoms are numbered in a manner which many authors have accepted. Unfortunately even within the distinct groups and subgroups of lignans a diversity of other systems has been used. Thus the numeration of the C2–C3 linked butanes (as in **1.1**) is often associated with chemically related cyclised compounds such as **1.6–1.8** in which the heteroatom has priority and hence the β-carbons are now desig-

nated C3–C4. This lack of correspondence extends to the butyrolactones (**1.5**) where the writer has found four different systems in use, including those which accord priority to the endocyclic oxygen atom and others where this goes to the carbonyl carbon. It should also be noted that when the aryl groups are differently situated their substituent locants will generate two sets of primed numbers (cf. **1.1**).

In tetralins (**1.10**) the common practice is shown but in many substituted compounds priority is allocated at the point of substitution rather than at the point of attachment of the pendent ring. Furofurans are commonly described as 2,6-diaryl-3,7-dioxabicyclo-[3,3,0]-octanes with priority given to the bridgehead carbon atom, yet this is at variance with the furans (cf. **1.6–1.8**) despite the close chemical and biochemical relationship. Several conflicting methods have been used for dibenzocyclo-octadienes (**1.11**) and the inescapable conclusion is that there is need for a system which relates to the β,β-linkage common to all lignans and which will stand without modification during interconversions between compounds in different lignan groups.

A suitable scheme was proposed in 1961 by Freudenberg and Weinges where the two cinnamyl residues are numbered as in **1.21** and where the 8,8′-linkage is assigned to all lignan classes as a matter of definition. Although this system has been commended by other authors it has not won general acceptance possibly due to the proposed retention of trivial names such as pipero – (for methylenedioxyphenyl) and guaia – (for 4-hydroxy-3-methoxyphenyl) in conjunction with systematic usage. There is now urgent need for revision owing to the welcome subsequent increase in the number and structural diversity of known compounds (see Figure 2.1). At the time of writing a modification of the Freudenberg and Weinges system is being considered by a IUPAC nomenclature committee. This has attracted favourable comment from numerous workers in the field and has been adopted in the writing of this book. As a result it was necessary to change the numeration chosen by many authors in order to interrelate their findings and the writer in apologising for this hopes that they will see the end as justifying this means.

The proposals (Moss, 1987) now being considered set out rules for assigning a priority for unprimed locants; these are to be taken in the following sequence until a distinction can be made:

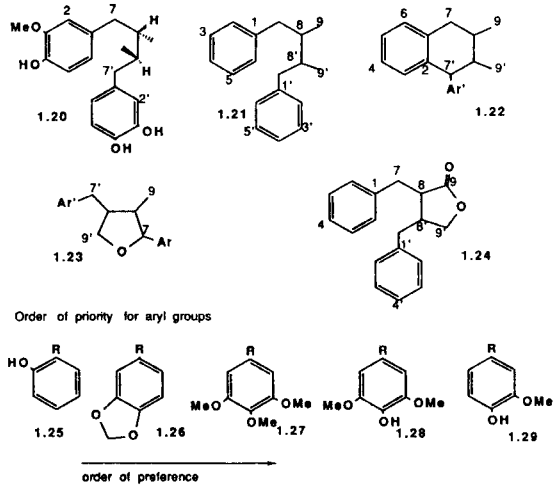
(a) A phenyl group which is fused to another carbocyclic ring has unprimed locants (cf. **1.22**).
(b) A phenyl group directly substituted on a heterocyclic ring has unprimed locants (cf. **1.23**).

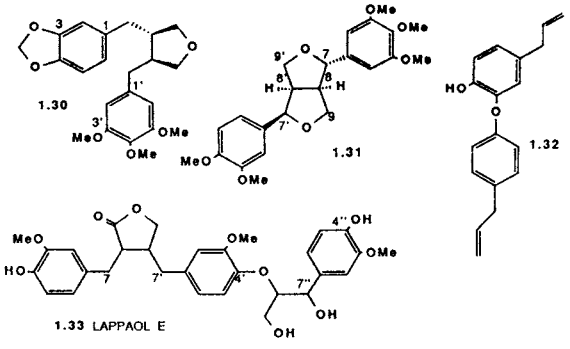
(c) Where both phenyl groups are similarly placed the priority is to be decided using the sequence rules (IUPAC, 1974).

The distinction is clear for guaiaretic acid (**1.1**) since the priority goes to the aryl group associated with the functionalised side chain and guaiaretic acid is described as 3,3′-dimethoxylign-7-ene-4,4′-diol, whilst α-conidendrin (**1.2**) becomes 4,4′-dihydroxy-3,3′-dimethoxy-2,7′-cyclolignan-9,9′-olide. Note that the carbonyl group of the lactone is placed at C-9, the locant 9 preceding 9′. The preferred lactone group is written as a suffix here and it supplants the hydroxyl group which had pride of place in guaiaretic acid. In the open chain lactone **1.24** the priority follows the principal function at C-9.

In structure **1.23** the priority is clear but for burseran (**1.30**) the sequence rule must be invoked. The aryl group with the substitution pattern **1.26** is preferred to the 3,4,5-trimethoxylated residue (**1.27**) and so burseran is described as (8α,8β)-3,4-methylenedioxy-3′,4′,5′-trimethoxy-9,9′-epoxylignan. This is established by taking the atoms in sequence from C-1 when a difference first appears beyond the first ether oxygen. The priority then goes to the methylene ether group at C-3 with a substituent set of (O, H, H) rather than to the methyl group at C-3′ whose set is (H, H, H). The same analysis defines a lower priority to the dihydroxyphenyl group in **1.20** and so it carries primed locants.

If the furofuran lignan (+)-epimagnolin (**1.31**) is written with α-hydrogen atoms at the ring junctions it is then described as −(7α,7′β)-3,3′,4,4′,5-pentamethoxy-7,9′:7′,9-diepoxylignan. Owing to the symmetry of this model its rotation leads to ambiguity, which can be avoided by always





writing the aryl substituents at two and eight o'clock and by allocating the former position to the principal group where a difference exists. The convention of designating α-substituents as below the plane and β-substituents as above the plane can now be adopted (see also Chapter 2, p. 15).

In the present context where lignans are distinguished from neolignans it is not necessary to specify their 8,8′-coupling. Should the distinction between the two be set aside and the term lignan be broadened so as to cover all classes this simplification would not be admissible.

Although the neolignans are not covered in detail in the text, it should be pointed out that they are equally well described using this system. Thus isomagnolol (**1.32**) becomes 3,4′-epoxylign-8,8′-diene-4-ol, whilst megaphone (**1.19**) is 3,3′4,5-tetramethoxy-8,1′-lign-8′-ene-7-ol-6′-one.

It is also possible to extend the system so as to describe sesqui- and dineolignans. Lappaol E (**1.33**) becomes 4,4″,7″,9″-tetrahydroxy-3,3′3″-trimethoxy-4′,8″-epoxy-8,8′-sesquineolignan-9,9′-olide.

An ambiguity of nomenclature has been introduced by Gottlieb (1978) whose group has made a substantial contribution to the chemistry of neolignans. They have chosen to define neolignans as compounds formed by the oxidative dimerisation of allyl or propenylphenols. This was based on a speculation that this coupling mode followed a biogenetic path distinct from the formation of lignans from cinnamyl alcohols (pp. 3, 278). One especially difficult consequence of regarding neolignans as dimers of allylphenols is that all tetralin lignans of the type (**1.15**, Scheme 1.3) would require redefinition. There is therefore a preference for an earlier proposal by Gottlieb (1972) which is consistent with Howarth's original definition and which maintains the unique 8,8′-linkage of lignans.

Additional illustrations of the application of the modified Freudenberg system are given in the registry of naturally occurring lignans in Chapter 2.

Absolute configuration

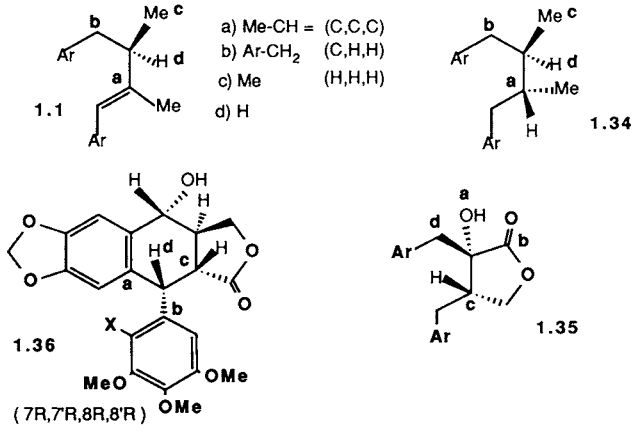
The assignments to lignans are now largely established as a result of X-ray and circular dichroism (CD) studies which are mentioned at appropriate places in the text. These results interrelate with many chemical correlations some of which are discussed at the end of this section.

Absolute configuration is assigned using the sequence rules of Cahn *et al.* (1966) and in these terms guaiaretic acid (**1.1**) becomes (7*E*)-(8′*R*)-3,3′-dimethoxylign-7-ene-4,4′-diol. The order of priority of the C-8′ substituents is shown in Scheme 1.5 and when viewed with the ligand of least atomic number (d = H) to the rear the preference path a > b > c follows in a clockwise (*R*) sense. Note that in accordance with the rules the methyl group is of lower priority than the substituted carbon atoms; in guaiaretic acid the alkene carbon is preferred to the alkane. In the important example of (−)-dihydroguaiaretic acid (**1.34**), the priority order is the same (8*R*,8′*R*), since that of (a) with its substituent group (C, C, H) is greater than (b; C, H, H).

The lactone (−)-dimethylmatairesinol (**1.39**, Scheme 1.6) also has the (8*R*,8′*R*) configuration as the highest priorities at C-9 and C-9′ are determined by the greater atomic weight of the oxygen atoms. However, contradictions which are more apparent than real can arise from the operation of the rules. Thus, although the carbon skeleton of (+)-trachelogenin (**1.35**) is superposable with that of dimethylmatairesinol, the insertion of the hydroxy group at C-8 requires that it be designated as (8*S*,8′*S*) because the sequence of substituents a > b > c is now counterclockwise.

Attention must be drawn to the important aryltetralin class which includes podophyllotoxin (**1.36**; X = H): (7*R*,7′*R*,8*R*,8′*R*)-7-hydroxy-

Scheme 1.5 Assignment of priority by the sequence rules



3′,4′,5′-trimethoxy-4,5-methylenedioxy-2,7′-cyclolignan-9′,9-olide. In determining the order of substituent priority for C-7′ each benzene ring carbon is taken as having three substituent carbon atoms – (C, C, C), where that of the fused ring ranks first because it is *ortho*-carbon substituted. The last priority goes to C-8′, as it has the lowest ranked substituent group (C, C, H). The other assignments are straightforward although note that with regard to C-8 the attached carbons have the sets:

C-7 (O, **C**, H)
C-9 (O, **H**, H)
C-8′(C, C, H)

and therefore the priority for C-7 is determined by the atoms of second preference **C** > **H**.

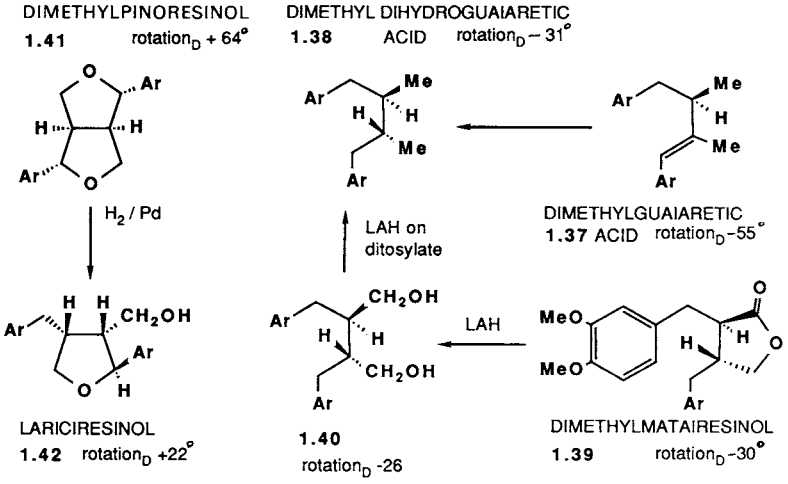
In applying the sequence rules to lignans one must always be wary of the effect of the insertion of additional substituents. As we saw above, the side chain insertion in (−)-trachelogenin led to an apparent anomaly and a similar situation can arise if the status of an aryl group is changed. Should podophyllotoxin be substituted at the 2′-position as in the chloro-derivative (**1.36**; X = Cl), then the priority now goes to the pendent ring and so this compound is properly described as 7′-S.

Correlation of absolute configurations

A whole matrix of chemical transformation exists which interrelates the lignan classes and which links them to guaiaretic acid (**1.1**) and many of these have been recorded (Hearon and MacGregor, 1955). Some examples of the methods used are shown in Scheme 1.6. *R*(−)-Dimethylguaiaretic acid (**1.37**) on hydrogenation (Schroeter *et al.*, 1918) gave a dihydrocompound (**1.38**) together with the isomeric *meso*-addition product. (−)-Dimethylmatairesinol (**1.39**) was then shown (Schrecker and Hartwell, 1955) to have the same absolute configuration at C-8 by reduction with lithium aluminium hydride to give the *trans*-diol (**1.40**). More vigorous treatment of the ditosylate of this compound with the hydride afforded dimethyldihydroguaiaretic acid (**1.38**) as a common product.

The diol (**1.40**) was also a key substance in the correlation of the absolute configuration of furofurans with other lignans. (+)-Dimethylpinoresinol (**1.41**) gave a furan, (+)-lariciresinol (**1.42**), on initial hydrogenation with a palladium catalyst (Howarth and Woodcock, 1939). Correlation of both these classes is possible, since continued hydrogenation afforded the reference compound (**1.40**; Freudenberg and Dietrich, 1953). Fully alkylated lignans are preferred in work of this kind for they are less susceptible to oxidation than the parent phenols and there is no risk of the

Scheme 1.6 Chemical correlation of absolute configuration



separation of phenolate salts during hydride reductions. Other examples of chemical methods of correlation of configuration will be given in further chapters (pp. 213, 228–30).

More modern techniques for assigning absolute configuration include those of optical rotatory dispersion and circular dichroism; numerous examples are included in the series published by Klyne and his collaborators (Hulbert *et al.*, 1981). The earlier work has also been substantiated recently by the application of X-ray crystallography to lignans (pp. 37, 49, 199, 230).

The techniques for establishing relative configurations depend very largely on proton magnetic resonance and the nuclear Overhauser effect and they are discussed in detail in Chapter 6.

References

Achenbach, H., Grob, J., Dominguez, X.A., Cano, G., Star, J.V., Brussolo L. del C., Munoz, G., Salgado, F. and Lopez, L. (1987). Lignans, neolignans and norlignans from *Krameria cystisoides*. *Phytochemistry*, **26**, 1159–63.

Cahn, R.S., Ingold, C.K. and Prelog, V. (1966). Specification of molecular chirality. *Angew. Chem. Intern. Edit.* **5**, 385–415.

Erdtman, H. (1933). Dehydrogenation of conifer substances. II. Dehydrodi-isoeugenol. *Liebigs Ann. Chem.* **503**, 283–94.

Freudenberg, K. and Dietrich, H. (1953). On syringaresinol, an oxidation product of sinapyl alcohol. *Chem. Ber.* **86**, 4–10.

Freudenberg, K. and Weinges, K. (1961). A system of nomenclature for lignans. *Tetrahedron* **15**, 115–28.