Phytochemistry, 1981, Vol. 20, pp. 121–126. © Pergamon Press Ltd. Printed in England. 0031–9422/81/0101–0121 $02.00/0

APPLICATION OF MASS SPECTROMETRY IN STRUCTURAL PROBLEMS IN TRITERPENES

L. OGUNKOYA

Department of Chemistry, University of Ife, Ile-Ife, Nigeria

(*Revised received* 10 *March* 1980)

Key Word Index—Triterpenes; mass spectra; structure.

Abstract—Mass spectral data on 105 pentacyclic triterpenoids have been collated and analysed. From this analysis, it has been possible to identify base peaks which are characteristic of specific stereoskeletons.

INTRODUCTION

Since the publication of Djerassi's paper [1] in 1963, mass spectroscopy has become the most powerful tool for the elucidation of the skeleton and especially for the location of double bonds in pentacyclic triterpenoids. The work reported in Djerassi's paper covered triterpenes based on actual or modified oleanane, ursane or lupane skeletons with double bonds located in various positions. This work has been particularly valuable in firmly establishing the retro-Diels–Alder fragmentation pattern for C-12 unsaturated oleanenes and ursenes. However, their findings on other stereoskeletal types were less definitive due essentially to the limited number of appropriate compounds then available. For example, compounds with C-7, C-8 and C-9(11) unsaturation which were examined possess either the oleanane or the ursane stereoskeleton only; no triterpenoid compounds with C-5 unsaturation were known then, neither were any with arborane or fernane skeletons studied. However, in the last 16 years, a number of new triterpenoids with these stereoskeletons have been discovered and the high resolution MS on a large number of them have been recorded.

During the course of our studies on the structures of a number of pentacyclic triterpenoids, a picture of a useful correlation between the fragmentation pattern of many of these compounds with their stereoskeletons slowly emerged. It must be pointed out that a number of earlier workers have independently utilized the generalizations in this paper (see below) in structural elucidation of triterpenoids. But there has been no attempt after Djerassi's work [1] to collate and correlate the MS data on a wide range of the compounds for use as a quick means of identification of different structural types of pentacyclic triterpenodis. This is the objective of the present work.

RESULTS AND DISCUSSION

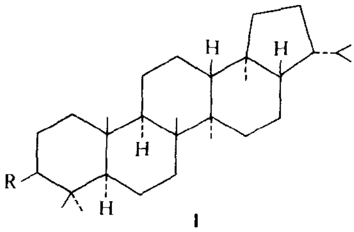
In this work, MS data have been collated for pentacyclic triterpenoid compounds based on hopane (**1**), arborane (**2a**), fernane (**2b**), lupane (**3**), oleanane (**4**) and ursane (**5**) stereoskeletons and with C=C located essentially at C-5, C-7, C-8, C-9(11), C-12, outside ring E and in a few other positions, these being the only structural types known for which characteristic MS features are discernible. In the exercise, all compounds with unsaturation at C-5, C-7, C-8 and C-9(11) for which high resolution MS data are either available in the literature up to 1978 or obtained by the author have been considered. With C-12 unsaturated compounds, only representative examples relevant to the discussion have been considered, because this class of compounds has been extensively studied earlier [1]. A wide range of compounds with lupane or hopane skeletons were examined in order to obtain a coherent picture. No attempt has been made either to study or to propose original mechanisms of fragmentations. Wherever possible, the structures of relevent ions giving rise to the base peak have been given or proposed in some cases.

It is not considered necessary to tabulate all the MS data on every compound. Only the significant peaks (in most cases, the ones with the highest relative abundance in the normally diagnostic high mass region) have been mentioned. References are given for sources of information in all cases. Structures of compounds are not given since these can be easily obtained from the literature references cited, except in cases where the structure helps to clarify the point under discussion.

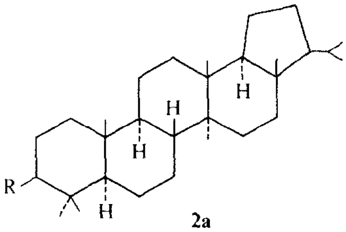
From the present studies, it has been found that armed with a knowledge of the *m*/*e* values of the molecular ion and the base peak plus one or two additional peaks, it is possible to quickly assign to a given compound one of the following probable structural types as a working hypothesis: (i) C-5 unsaturated arborane, fernane, oleanane or lupane skeleton; (ii) C-7, C-8 and C-9(11) unsaturated arborane or fernane skeleton; (iii) C-7, C-8 and C-9(11) unsaturated oleanane or ursane skeleton; (iv) C-12 unsaturated oleanane or ursane skeleton; and (v) lupane or hopane skeleton. It is also found that, in cases where there is substitution in the molecule, it is possible, by a careful analysis of the MS data, to locate approximately the position of such functional groups. The different structural types will now be discussed.

1. C-5 *unsaturated compounds*

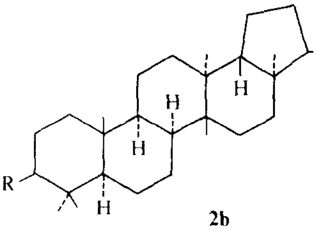
There are three stereoskeletal types known with C-5 unsaturation. These are the fernane-type structures exemplified by adianene [2] and its derivatives, simiarenol and episimiarenol [3–5] and their acetates and ketones, the oleanane type exemplified by glutinol [6] and its



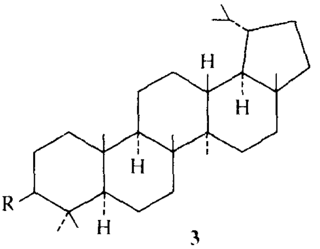
1



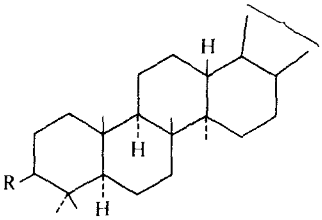
2a

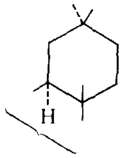


2b

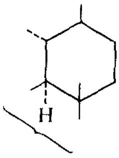


3

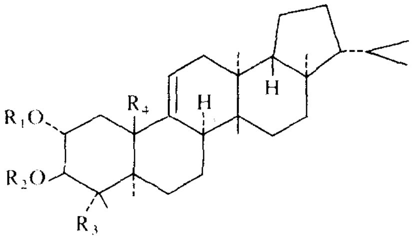




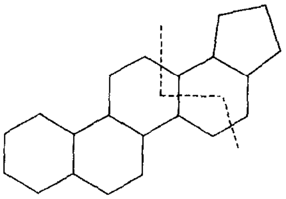
4



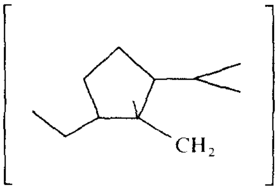
5



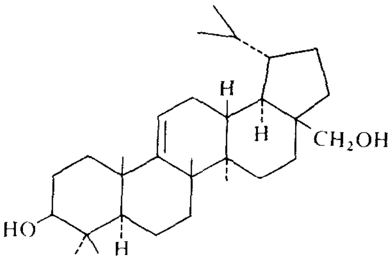
6a $R_1 = R_2 = H; R_3 = COOMe; R_4 = Me$
6b $R_1 = R_2 = Ac; R_3 = COOMe; R_4 = Me$
6c $R_1 = R_2 = H, R_3 = R_4 = COOMe$
6d $R_1 = R_2 = H; R_3 = CH_2OH; R_4 = COOMe$



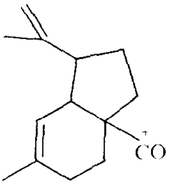
7



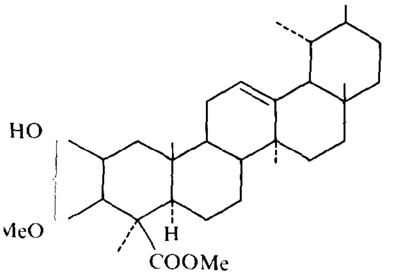
8



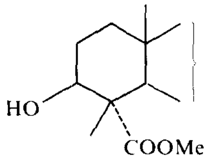
9



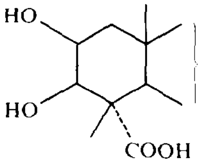
10



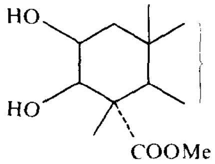
11a



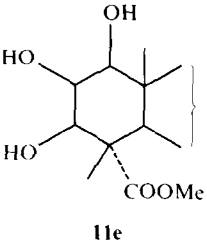
11b



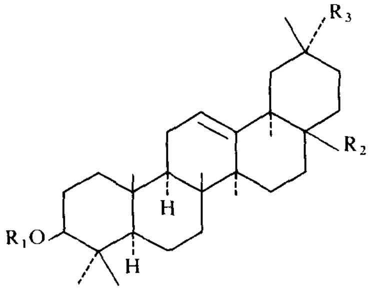
11c



11d

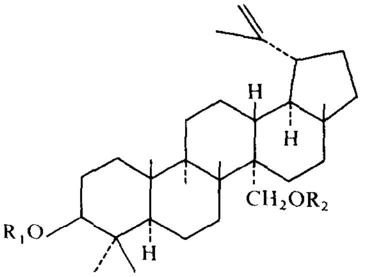


11e



12 R_1 = Ac; R_2 = COOMe; R_3 = CH_2OAc:

13 R_1 = H; R_2 = COOMe; R_3 = COOMe

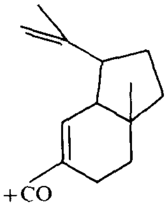


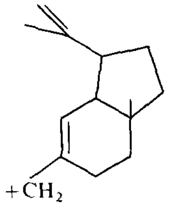
14a R_1 = R_2 = H

14b R_1 = Ac; R_2 = H

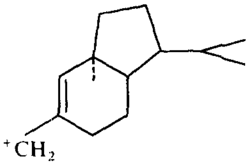
14c R_1 = R_2 = Ac

14d =

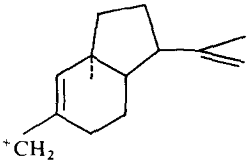




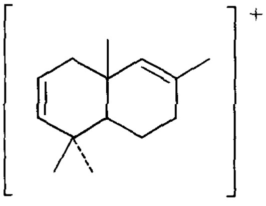
15



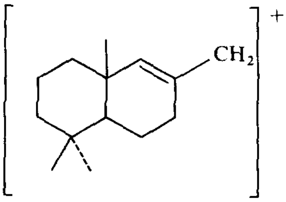
16



17



18



19

acetate and the more recently discovered lupane type exemplified by guimarenol [7] and its acetate and ketone. All these compounds (with the exception of adianene with a base peak at *m/e* 259) give a base peak at *m/e* 274, with strong peaks at *m/e* 259 (274 – Me); and at *m/e* 231 (274 – Pr) for those with isopropyl group in ring E (like the arborane, fernane and lupane skeletons). It is generally accepted [2–7] that, in the MS of these compounds, the base peak must be due to the ion resulting directly from retro-Diels–Alder fission of ring B in which the charge remains with the monoalkene instead of the diene [1].

Thus a base peak at *m/e* 274 or 259 may be regarded as suggestive of C-5 unsaturated pentacyclic triterpenoids in which there is no further substitution in any of rings C, D and E. With substitution in any of these rings, it is to be expected that the base peak will be appropriately shifted to higher *m/e* values.

2. C-7, C-8 *and* C-9(11) *unsaturated compounds*

There are two main classes of pentacyclic triterpenoids known with C-7, C-8 and C-9(11) unsaturation. One class (A) consists of compounds with arborane and fernane skeletons. These include fernenol, arborinol, isoarborinol [8–13], trematol [14, 15], retigenic acids [16] and their functional derivatives together with their derived hydrocarbons, all these being C-9(11) unsaturated compounds. The other members of this class include motiol [2] and hortensenol (isomotiol) [17, 18] which are unsaturated at C-7 and C-8, respectively. Motiol and hortensenol are the only known pentacyclic triterpenoids with C-7 and C-8 unsaturation and of these, spectral data are available only for hortensenol. It has, so far, not been possible to obtain a sample of motiol for spectral analysis. However useful information has been obtained (H. Ageta, private communication) on the derived hydrocarbon, fern-7-ene. The second class (B) of this group consists of compounds with either the ursane or oleanane skeleton exemplified by baurenol, isobaurenol, multiflorenol, isomultifluorenol, their acetates and ketones [8–13, 19].

Class A. In the 28 compounds with arborane and fernane skeletons) considered, the high resolution MS data were very similar. Many of them showed almost identical fragmentation patterns with differences only in the relative abundance at certain *m/e* values. These observations were confirmed by Ageta (private communication) who reported that the MS of fern-7-ene, fern-8-ene and fern-9(11)-ene could not be distinguished under various operational conditions. The most prominent peaks in all cases occur at *m/e* $M^+ - 167$ and $M^+ - Me$, the former being the base peak in most cases. Even in the MS of the four derivatives of retigenic acid (**6a–d**) (a highly substituted C-9(11) unsaturated fernene-type structure) the peaks at $M^+ - 167$ and $M^+ - Me$ are invariably among the strongest peaks, with the addition of other strong peaks arising from loss of HOAc or H_2O from either M^+ or from $M^+ - 167$ [16].

It has been earlier postulated [8–13, 16] that the base peak ion (i.e. at *m/e* $M^+ - 167$) arises from the cleavage (see **7**) across the C-12 to C-13, C-13 to C-14 and C-15 to C-16 bonds resulting in the loss of the species (**8**) from the molecular ion and leaving the base peak ion containing rings A, B and C (in part). This postulate has been found to be consistent with the findings in the present studies.

One apparent exception to the above general fragmentation pattern is with the C-9(11) unsaturated compound (**9**) in which the base peak occurs at *m/e* 203 [20]. It should, however, be noted that this compound has a lupane skeleton (and not an arborane or fernane type like the other members of the class) similar to compounds **14a**, **b** and **c** (see below), each of which has the base peak at *m/e* 203. A possible structure for the base peak ion from compound (**9**) would be the species (**10**) (see also **14d**). Quite clearly, the C-13 methyl has a pronounced effect on the mode of fragmentation of these molecules.

However, in compounds in which the molecule is oxygenated at C-12, the fragmentation follows an entirely different pattern. This is exemplified by the MS of the three fernane-type C-9(11)-unsaturated triterpenoids and their derivatives isolated from the lichen *Xanthoria resendei* [21]. None of these compounds showed a base peak at either $M^+ - 167$ or $M^+ - Me$. In fact, only 3,12-diketofern-9(11)-ene showed relatively strong peaks at these *m/e* values. In general, there is no discernible uniform mode of fragmentation in any of these compounds.

In the light of the foregoing results, it may be concluded that the presence of strong peaks at *m/e* $M^+ - 167$ and $M^+ - Me$ (the former being the base peak in most cases) is indicative of C-7, C-8 and C-9(11) unsaturated pentacyclic triterpenoids with the arborane and fernane skeletons and no substitution at C-12.

While it is relatively easy to distinguish C-8 unsaturated compounds from C-7 and C-9(11) ones by the aid of ^{1}H NMR (vinyl proton absorption), and especially by the more recently developed method of molecular rotation differences [22, 23], we have not been able to distinguish C-7 from C-9(11) unsaturated compounds with equal ease. However, recent findings by Ageta (private communication) showed that derivatives of fern-9(11)-ene and fern-7-ene can be distinguished by the exact assignment of the eight Me signals in their high resolution ^{1}H NMR spectra.

Class B. The MS of C-7, C-8 and C-9(11) unsaturated compounds with the oleanane or ursane skeleton present an entirely different fragmentation pattern. MS data are available for only nine members of this class and all are derivatives of baurenol (Δ^7), multifluorenol (Δ^7), isobaurenol (Δ^8) and isomultifluorenol (Δ^8) [1, 8–13, 19]. Of these, six showed a base peak at $M^+ - 179$ with strong peaks variously at *m/e* $M^+ - 167$ and 205 and strangely at *m/e* 218 (as base peak) in multiflorenone. It is therefore apparent that these compounds fragment by a mechanism which is different from that of the arborane and fernane types. Furthermore, the C-12 oxygenated derivatives of C-9(11) unsaturated oleanane compounds give base peaks which are believed to arise by retro-Diels–Alder decomposition of ring C [1].

3. C-12 *unsaturated compounds*

Djerassi and his co-workers firmly established [1] the retro-Diels–Alder reaction to be responsible for the prominent fragmentation of triterpenoids with the oleanane and ursene skeletons leading to the formation of the base peak ion at *m/e* 218 in which the charge remains with the diene. For example, each of commic acids A, B, C, D and E (**11a–e**) shows a base peak at *m/e* 218 (203 + Me) [24]. However, the position of the base peak in compounds in which there are substituents in rings D and E does not appear to follow a regular pattern. For example, whereas methyl ursolate [25] and 2α-hydroxyuval [26] show a base peak at *m/e* 203 as a result of loss of COOMe and COOH from rings D and E during electron impact, compounds like karachic acid [27], 2α,3β-dihydroxyurs-12-en-28-oic acid, asiatic acid and its methyl ester [28], caulophyllogenin [29] and 3β,20α-diacetoxyurs-12-ene [30] show their base peaks at *m/e* 248 (203 + COOH), 248, 248, 262 (203 + COOMe), 264 (203 + COOH + OH) and 276 (203 + OAc), respectively. In these latter cases, unlike the former, the carboxyl, carbomethoxy and hydroxyl groups were retained in the base peak ion.

Thus with this class of compounds, the diagnostic base peak is at *m/e* 218, when the compounds contain no substituents in rings D and E. However, any other base peak which can be easily related to 203 by the addition of the masses of simple functional groups, like COOH, COOMe, OAc, OH, etc., are equally indicative of C-12 unsaturated pentacyclic triterpenoids with the oleanane or ursane skeleton. This later operation has a useful application because it may be used to detect the presence of such functional groups in rings D and E. Thus, a pentacyclic triterpenoid carboxylic acid which shows a base peak at *m/e* 248 could probably be a C-12 unsaturated compound containing a COOH group in ring D or E.

On the basis of the above generalizations, it is difficult to reconcile the base peaks recorded for methyl queretarvate diacetate (**12**) (*m/e* 201) [31] and dimethyl serratagenate (**13**) (*m/e* 187) [32] with their structures.

4. *Compounds with lupane and hopane skeletons*

MS data on a wide variety of pentacyclic triterpenoids based on the lupane and hopane skeletons were considered. Examples of those with lupane skeleton include lupeol and lupenyl acetate, 2α-hydroxylupeol [33], three derivatives of lupene [34], seven derivatives of betullinic and epibetullinic acids [31, 35], various derivatives of ceanothic [35, 38], alphitolic, colubrinic and granulosic acids [35], glochidol and its acetate [37] and derivatives of spinosic acid [38]. Examples of those with a hopane skeleton include 6β,22-dihydroxyhopane, 7β,22-dihydroxyhopane, 22-hydroxy-7-oxohopane, (6-2H_2)-7-oxohopane and 15α,22-dihydroxyhopane [39].

It was found that virtually all these compounds are characterized by an intense peak at *m/e* 189 irrespective of the nature of substitution in rings A, B, D and E. In cases where there is an *iso*-propyl group instead of an *iso*-propenyl group in ring E, this peak occurs at *m/e* 191. In most cases, the peak at *m/e* 189 or 191 is either the base peak, or is one of the most intense peaks in the high mass region of the spectrum. For example, in glochidol [37], the molecular ion provides the most intense peak followed by the ion $M^+ -$ Me and then the peak at *m/e* 189, whereas in glochidyl acetate, $M^+ -$ HOAc provides the most intense peak followed by the peak at *m/e* 189. In lupene-2α,3α-diol [34], the base peak occurs at *m/e* 205 (which is a common fragment in most pentacyclic triterpenes [1] and hence of no diagnostic value) followed by the peak at *m/e* 189. The observed fragmentation of these lupene and lupane derivatives agrees with an earlier observation of Djerassi [1] with saturated lupanes and therefore appears as an exception to the general fragmentation pattern of this class of compounds. An intense peak at *m/e* 189 may therefore be regarded as being reasonably indicative of this class of pentacyclic triterpenoids.

In cases where the C-14 methyl is functionalized, the base peak has been found to be shifted accordingly. For example, each of the compounds **14a**, **b** and **c** shows a base peak at *m/e* 203 (189 + O − 2H) which could be provided by an ion of structure **14d** in each case, in preference to the structure proposed by the authors [30]. The origin of the fragment ion at *m/e* 189 or 191 in these compounds is still not completely clear. It has been postulated [1, 30, 40] that this ion may arise from the fission of the C-8 to C-14 and C-12 to C-13 bonds followed by H-transfer to give species containing rings D and E such as **15** in the lupane series or **16** and **17** in the hopane series. It may also arise from the fission of the C-9 to C-11 and C-8 to C-14 bonds to give species containing rings A and B such as **18** or **19**.

The proposed structures **15**, **16** and **17** containing rings D and E are difficult to rationalize as the base peak ions from compounds in which the C-17 methyl is functionalized, as in betullinic acid derivatives and also in related compounds with contracted ring A such as the derivatives of colubrinic, ceanothic and granulosic acids in all of which the peak at *m/e* 189 is either the base peak or one of the most intense peaks in the MS. This is because since most of these compounds give the base peak at *m/e* 189, they would be expected (if the base peak ion still retains rings D and E) to have lost the carboxyl or carboalkoxy group on electron impact, in which case the base peak should occur at *m/e* 175 (instead of 189), unless there has been a Me migration from C-8 to either C-17 or C-18, which is quite an unlikely supposition. With these compounds, therefore, it is conceivable that the base peak is provided by an ion with structures similar to either **18** or **19** contrary to the earlier proposition [35]. There is, therefore, a need for more detailed mechanistic studies into the mode of fragmentation of molecules in this class of compounds.

CONCLUSION

This paper has attempted to provide quick answers to the following questions: If the MS of a pentacyclic triterpenoid shows M^+ at *m/e x* and a base peak at *m/e y* what is the probable stereoskeleton, the probable position of unsaturation (if any) and the possible locations of any additional functional groups in the molecule? Helpful answers have been provided for compounds of various stereoskeletons which are unsaturated at C-5, C-7, C-8, C-9(11), C-12 and for the lupane and hopane skeletons. Further work aimed at possible differentiation between C-7, C-8 and C-9(11) unsaturated compounds is in progress. The data recorded by me and also available in the literature for saturated analogues of the skeletal types discussed, for stereoskeletons unsaturated at C-13(18), C-14 and other positions not discussed and for compounds based on cycloartane, friedelane and serratane skeletons do not yet present a uniform picture for useful generalizations.

It should be noted that the generalizations made in this work have been based on empirical data published on selected compounds known up to 1978. Naturally, not all known compounds have been considered and fresh data on new compounds will be forthcoming.

Furthermore, it is important to realise that while some skeletal types generally give strong signals (in many cases as base-peaks) at certain characteristic *m/e* values, it is not uncommon to observe signals (though usually with relatively lower relative abundance) at the same *m/e* values in the MS of other skeletal types. For example, the base peak at *m/e* 189 for hopane and lupane skeletons might also be given by other pentacyclic triterpenoids which differ from the hopane and lupane skeletons in the position and nature of substitution in ring D and especially E. Consequently, the generalizations should be applied with caution and used more as a guide than as a definitive diagnosis in the solution structural problems.

Finally, the generalizations in the paper refer specifically to pentacyclic triterpenoids, but it should be noted that some tetracyclic ones could, in certain cases, conceivably give strong signals at the same important *m/e* values. This possible coincidence should not pose a serious problem for two reasons. Firstly, tetracyclic compounds could be readily distinguished from pentacylic ones by the combined use of molecular formulae, NMR, IR, UV, tests for unsaturation and any others considered appropriate. Secondly, tetracyclic triterpenoids do not appear to give characteristic signals arising from cleavage of ring-bonds like the pentacyclic ones because of the fragmentation of the side-chain on C-20.

Acknowledgement—I thank Dr. E. S. Waight, Department of Chemistry, Imperial College, London, for the determination of mass spectra.

REFERENCES

1. Budzikiewicz, H., Wilson, J. M. and Djerassi, C. (1963) *J. Am. Chem. Soc.* **85**, 3688.
2. Nakamura, S., Yamada, T., Wada, H., Inoue, Y., Goto, T. and Hirata, Y. (1965) *Tetrahedron Letters* 2017.
3. Alpin, R. T., Arthur, H. R. and Hui, W. H., (1966) *J. Chem. Soc. C* 1251.
4. Budzikiewicz, H., Brauman, J. I. and Djerassi, C. (1965) *Tetrahedron* **21**, 1855.
5. Ogunkoya, L. Olubajo, O. O. and Sondha, D. S. (1972) *Phytochemistry* **11**, 3093: (1973) *Phytochemistry* **12**, 732.
6. Alpin, R. T., Arthur, H. R. and Hui, W. H. (1966) *J. Chem. Soc. C* 1251.
7. Gonzalez, A. G., Jerez, F. G. and Escalona, M. C. (1973) *An. Quim.* **69**, 921.
8. Nishimoto, K., Ito, M., Natori, S. and Ohmoto, T. (1968) *Tetrahedron* **24**, 735.
9. Bryce, T. A., Martin-Smith, M., Osske, G., Schreiber, K. and Subramanian, G. (1967) *Tetrahedron* **23**, 1283.
10. Wada, H., Goto, G., Goto, T. and Hirata, Y. (1966) *Tetrahedron Letters* 3461.
11. Gonzalez, A. G., Betancor, C., Hernandez, R. and Salazar, J. A. (1976) *Phytochemistry* **15**, 1996.
12. Palmer, M. A. and Bowden, B. N. (1977) *Phytochemistry* **16**, 459.
13. Gunasekera, S. P., Kumar, V., Sultanbawa, M. U. S. and Balasubramanian, S. (1977) *Phytochemistry* **16**, 923.
14. Ogunkoya, L., Olubajo, O. O. and Sondha, D. S. (1977) *Phytochemistry* **16**, 1606.
15. Obafemi, C. A., Ogunkoya, L., Quartey, J. A. K. and Waight, E. S. (1979) *Phytochemistry* **18**, 496.
16. Takahashi, R., Chiang, H. C., Aimi, N., Tanaka, O. and Shibata, S. (1972) *Phytochemistry* **11**, 2039.
17. Singh, H., Kapoor, V. K., Piozzi, F., Passannanti, S. and Paternostro, M. (1978) *Phytochemistry* **17**, 154.
18. McCandlish, L. E. and Stout, G. H. (1976) *Acta Crystallogr.* **B32**, 1788.
19. Talapatra, S. K., Sengupta, S. and Talapatra, B. (1968) *Tetrahedron Letters*, 5963.
20. Bohlmann, F., *et al.* (1977) *Phytochemistry* **16**, 965.
21. Gonzalez, A. G., Martin J. D. and Perez, C. (1974) *Phytochemistry* **13**, 1547.
22. Ogunkoya, L. (1977) *Tetrahedron* **33**, 3321.
23. Ogunkoya, L. (1978) *Phytochemistry* **17**, 1343.
24. Thomas, A. F. and Willham, B. (1964) *Tetrahedron Letters* 3177.
26. Marco, J. A., Parareda, J. S., Seoane, E., Abarca, B. and Sendra, J. M. (1978) *Phytochemistry* **17**, 1438.
26. Younes, M. E. (1975) *Phytochemistry* **14**, 592.
27. Khan, M. A. and Rahman, A. U. (1975) *Phytochemistry* **14**, 789.
28. Bandaranayake, W. M., *et al.* (1975) *Phytochemistry* **14**, 2043.
29. Strigina, L. I., Chetyrina, N. S., Isakoro, V. V., Dzizenko, A. K. and Elyakoo, G. B. (1974) *Phytochemistry* **13**, 479.
30. Hui, W. H. and Li, M. M. (1976) *J. Chem. Soc. Perkin Trans.* 1, 23.
31. Herz, W., Santhanam, P. S. and Wahlberg, I. (1972) *Phytochemistry* **11**, 3061.
32. Rangaswami, S. and Sarangan, S. (1969) *Tetrahedron* **25**, 3701.
33. Kumar, N. and Seshadri, T. R. (1976) *Phytochemistry* **15**, 1417.
34. Kumar, N. and Seshadri, T. R. (1975) *Phytochemistry* **14**, 521.
35. Roitman, J. N. and Jurd, L. (1978) *Phytochemistry* **17**, 491.
36. Eade, R. A., Grant, P. K., McGrath, M. J. A., Simes, J. J. H and Wootton, W. (1971) *Aust. J. Chem.* **24**, 621.
37. Hui, W. H. and Li, M. M. (1976) *Phytochemistry* **15**, 561.
38. Reddy, G. C. S., Ayenga, K. N. N. and Rangaswami, S. (1975) *Phytochemistry* **14**, 307.
39. Corbett, R. E. and Young, H. (1966) *J. Chem. Soc. C* 1556, 1564.
40. Galbraith, M. N., Miller, C. J., Rawson, J. W. C., Ritchie, E., Shannon, J. S. and Taylor, W. C. (1965) *Aust. J. Chem.* **18**, 226.