Phytochemistry, 1980, Vol. 19, pp. 1509–1512. © Pergamon Press Ltd. Printed in England. 0031-9422/80/0701-1509 $02.00/0

THE IDENTIFICATION OF CHOLESTEROL AND OTHER STEROIDS IN *EUPHORBIA PULCHERIMMA*

BERNARD C. SEKULA and WILLIAM R. NES

Department of Biological Sciences, Drexel University, Philadelphia, PA 19104, U.S.A.

(*Received* 6 *September* 1979)

Key Word Index—*Euphorbia pulcherimma*; Euphorbiaceae; sterols; triterpenoids; nuclear magnetic resonance spectroscopy; cholesterol; sitosterol; campesterol; dihydrobrassicasterol; lanosterol; dihydrolanosterol.

Abstract—The 4-desmethyl tetracycles of the whole poinsettia plant (*Euphorbia pulcherimma*) less roots amounted to 0.07% of the wet wt and were shown by ^{1}H NMR spectroscopy to be steroids and not euphoids. The most abundant component was cholesterol, constituting half the mixture, followed in order of decreasing concentration by 24α-ethylcholesterol (sitosterol), 24α-methylcholesterol (campesterol) and 24β-methylcholesterol (22-dihydrobrassicasterol). The relative amount of cholesterol in this plant is the highest found so far in a tracheophyte. The 4,4-dimethyl compounds (0.1% of wet wt) included lanosterol (5%), 24-dihydrolansterol (5%), β-amyrin (25%), germanicol (50%), an unidentified pentacyclic triterpenoid (8%) and two or more (7%) unidentified components. Both the 4,4-dimethyl- and the 4-desmethylsterols were in the configurational series with a 20α-H-atom. Dihydrolanosterol and lanosterol are the probable intermediates from cycloartenol to cholesterol and 24-alkylcholesterol, respectively. Such a sequence would differ from that operating in most angiosperms, where the alkylation is thought to precede the opening of the 9,19-cyclopropane ring.

INTRODUCTION

In the genus *Euphorbia*, squalene oxide leads both to 4,4-dimethyleuphoids and to 4,4-dimethylsteroids [1–4], but it is not clear from the literature whether both or only one of these is metabolized further to 4-desmethyl compounds. The existence of 4,4-dimethyleuphoids with a 24-methylene group, as in the case of euphorbol in *E. triangularis*, *E. ingenus*, and *E. resinifera* [4], shows that C-24 alkylation can actually occur after cyclization to the euphoidal structure. It has recently become possible to examine the 4-desmethyl problem spectroscopically due to the demonstration that euphoids and steroids exhibit different ^{1}H NMR spectra [5–7]. In particular, the signal from the C-18 protons is shifted downfield by 0.5–1.0 ppm in the euphoid series. This signal is reported, for instance, at 0.70 ppm for lanosterol and at 0.76 and 0.88 for the C-20 epimeric pair of euphoids, euphol and tirucallol.

The poinsettia plant (*Euphorbia pulcherimma*) was chosen for study, because, though of Mexican origin, it is readily available in the U.S.A. as a decorative plant at Christmas time. Furthermore, Khastgir and Pradhan [8] and Dominguez *et al.* [9] examined specimens and reported the presence of 4-desmethyl material which they believed to be sitosterol. In our hands, it was possible to resolve the 4-desmethyl fraction chromatographically and spectroscopically into four components, all of which were Δ^5-sterols (including sitosterol) by the ^{1}H NMR criterion. The 4,4-dimethyl fraction was similarly resolved into several components, among which were lanosterol and 24-dihydrolanosterol, which had not been previously recognized in this plant.

RESULTS AND DISCUSSION

The various polycycles, obtained after alkaline hydrolysis of an acetone extract, were first separated into 4,4-dimethyl and 4-desmethyl fractions by chromatography on alumina. The 4-desmethyl compounds were then chromatographed on lipophilic Sephadex, yielding three fractions. Each of these was examined by ^{1}H NMR at 220 MHz (Table 1). The most polar fractions had the spectrum and GLC retention time expected of cholesterol [10, 11]. The next fraction showed a single GLC peak with the retention time expected of 24-methylcholesterol and a ^{1}H NMR spectrum which was clearly composed of signals from the two epimers at C-24 (campesterol and 24-dihydrobrassicasterol) [10]. Traces of stigmasterol were also suggested by the ^{1}H NMR spectrum. The least polar fraction was stereochemically pure 24α-ethylcholesterol (sitosterol) by ^{1}H NMR and GLC [10]. No evidence for any contaminating euphoidal analogs [5–7] was found. The amounts of the steroids determined by GLC were: cholesterol, 49%; 24α- and 24β-methylcholesterol, 14%, in a ratio (by ^{1}H NMR) of *ca* 2:1, respectively; and 24α-ethylcholesterol (with no 24β-epimer), 37%. The specimens of *E. pulcherimma* which we used, incidentally, have the highest percentage of cholesterol in the sterol mixture of any plant of which we are aware [12–15]. Despite this unusual character, *E. pulcherimma* clearly belongs to the 'main line' of

Table 1. ^{1}H NMR spectral values for steroids of *E. pulcherimma**

	C-18	C-19	C-21	C-26 and C-27	C-28	C-29	C-30(α)	C-31(β)	C-32
Lanosterol	0.69 *s*	1.00 *s*	0.91 *d*	1.68 *s* 1.60 *s*	—	—	0.98 *s*	0.81 *s*	0.88 *s*
24-Dihydro-lanosterol	0.69 *s*	1.00 *s*	0.89 *d*	0.87 *d*	—	—	0.98 *s*	0.81 *s*	0.88 *s*
Cholesterol	0.68 *s*	1.02 *s*	0.92 *d*	0.87 *d*	—	—	—	—	—
24α-Methyl-cholesterol	0.68 *s*	1.02 *s*	0.923 *d*	0.78 *d* 0.86 d	0.81 *d*	—	—	—	
24β-Methyl-cholesterol	0.68 s	1.02 *s*	0.936 *d*	0.78 *d* 0.86 d	0.79 *d*	—	—	—	—
24α-Ethyl-cholesterol	0.68 *s*	1.02 *s*	0.93 *d*	0.82 *d* 0.84 *d*	—	0.85 *t*	—	—	—

*Position numbers for carbon atoms are those of Nes and McKean [13]. Spectra of 4,4-dimethylsteroids are at 360 MHz and of 4-desmethylsteroids at 220 MHz. Reference spectra at 220 MHz of authentic cholesterol and its 24α-ethyl derivative were identical to those from the respective poinsettia steroids. The spectrum of an authentic mixture of 24α- and 24β-methylcholesterol was the same as that of the poinsettia component except for variation due to the relative amounts of epimers and traces of stigmasterol. Doublets had *J* values of 6.5 Hz. The triplet for C-29 had a *J* value of 7.5 Hz.

tracheophytes, i.e. those in which the sterol mixture is composed principally of Δ^5-sterols in the homologous series (24-H, 24-Me, 24-Et) where most of the 24-alkyl component have the α-configuration [14]. Although the relative amount of cholesterol and its 24-alkyl derivatives may vary, most (but not all) angiosperms studied so far are of the sort just described, which we have called 'Category I-A' [14].In such plants the 24-ethyl component has always exclusively the 24α-configuration, as is true in the present case with *E. pulcherimma*. Inversion of the configuration at C-24 produces a clear shift in the ^{1}H NMR triplet for C-29 which is most readily seen in the position of the left-hand branch of the triplet [10, 11]. Not only was there no shift in this peak in the 24α-ethylcholesterol from *E. pulcherimma*, but there was also no shoulder at the position expected for the β-epimer.

After removal of fatty alcohols the 4,4-dimethyl material was rechromatographed on alumina, yielding several fractions enriched in components with the same GLC retention times as lanosterol and 24-dihydrolanosterol. These were finally resolved by chromatography on lipophilic Sephadex. The most polar fraction was shown to be lanosterol and the least polar fraction 24-dihydrolanosterol by ^{1}H NMR (Table 1) and MS. The two steroids were present in nearly equal amounts, as measured by GLC prior to final chromatography. Moving more slowly on Al_2O_3 was material with the GLC retention time (RR_t 1.61) of β-amyrin, accompanied by minor components. Following these materials was a substance with RR_t 2.42, representing 8% of the 4,4-dimethyl fraction. The latter was obtained as a pure component by crystallization (mp 216–218°). This was presumably the 'pseudotaraxasterol' of mp 220–222° obtained by Dominguez *et al.* [9]. Six singlets and two doublets in the ^{1}H NMR spectrum, together with a MW of 426 from MS, indicated that the compound was a pentacyclic triterpenoid containing two methyl groups, each of which was on a CH group, as in α-amyrin, but this excludes pseudotaraxasterol. However, it was not α-amyrin since this triterpene melts at 186° [16]. The compound may be the analog of α-amyrin in the taraxerol or δ-amyrin series, but its structure was not pursued further. It is worth noting that in a review of the constituents of *Euphorbia* sp [1], α-amyrin is listed as present in *E. pulcherimma*. However, we have carefully examined the literature cited [8] and find no reference to α-amyrin.

The material with RR_t 1.61 had mp 168–170° and was the major component (75% by GLC) of the 4,4-dimethyl fraction. This presumably corresponded to the germanicol (mp 168–170° crude; mp 175–176° purified) reported earlier as the principal component of the dimethyl fraction [8, 9]. ^{1}H NMR at 360 MHz showed our material to be a mixture. There were twelve singlets below δ 1.3, despite a MW of 426 by MS, as expected of a pentacyclic triterpenoid. There should have been eight or fewer singlets. Among the lesser ^{1}H NMR peaks were all those found in an authentic sample of β-amyrin. Unfortunately, authentic germanicol was not available to us. It probably represented much of what was in the mixture with RR_t 1.61. The chemical evidence presented previously [8, 9] for its presence was rather convincing. Our material (RR_t 1.61) showed no doublets or triplets in the ^{1}H NMR spectra, indicating the essential absence of triterpenoids, e.g. α-amyrin, in which 1,2-methyl migration has occurred.

The ^{1}H NMR spectra of the steroids (cholesterol and its homologs and lanosterol and its dihydro derivative) permitted elucidation of the configuration at C-20, since we have shown that 20α- and 20β-epimers exhibit a marked difference in the signal from the C-21 protons [17]. In the usual (normal) series, a 20α-H is present and the C-21 signal is at δ 0.9, while in the 20-*epi*- (20-*iso*) series the value is 0.1 ppm less. This difference is quite large. Thus, in 20-epicholesterol the C-21 doublet is upfield from the doublets for C-26, 27 but downfield from the latter in

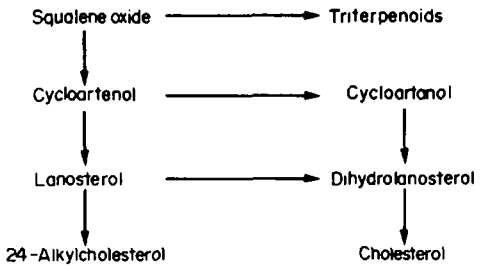

Fig. 1. Proposed pathway to sterols in *E. pulcherimma*.

cholesterol. It will be seen from Table 1 that the signal from C-21 was clearly emanating from the normal series in all cases.

E. pulcherimma was previously thought to be unusual in that it appeared to be one of only a few species of the genus which lacked 4,4-dimethyl tetracycles [1]. Most *Euphorbia* species have been found to contain 4,4-dimethylsteroids, 4,4-dimethyleuphoids or frequently both. Our own work, which documents the presence of 4,4-dimethylsteroids, brings *E. pulcherimma* into congruence with the majority of species. Lanosterol and dihydrolanosterol amounted to about 5% each of the 4,4-dimethyl fraction. If euphol or tirucallol were present, the amounts must be much less than this. They emerge in the lanosterol–dihydrolanosterol region during GLC [1], but were not detected in our 4,4-dimethyl fraction. On the other hand, the Δ^7-euphoids (e.g. butyrospermol) and 24-methylene-Δ^8-euphoids (e.g. euphorbol) would have moved in GLC under the large, broad RR_t 1.61 peak and would have been masked by the latter. In *Euphorbia* species euphoids seem to be concentrated in the latex. Since the proportion of latex to whole plant is small in *E. pulcherimma*, this may also explain why we failed to detect euphoids.

The presence of approximately equal amounts of lanosterol and dihydrolanosterol in *E. pulcherimma* is mirrored by the 1:1 ratio of cholesterol and its 24-alkyl derivatives. This suggests that control of the reduction–alkylation bifurcation [12, 13] resides at the 4,4-dimethyl stage of biosynthesis, as indicated in Fig. 1. Although we did not observe any cycloartenol, it has been found in other lanosterol-containing euphorbs [1], and the presence of a 9,19-cyclopropane, Δ^8-isomerase has been demonstrated [18].

EXPERIMENTAL

GLC was carried out on XE-60 at 235°. RR_t's are given relative to cholesterol. Columns of Al_2O_3 were deactivated with H_2O, as indicated. Et_2O in hexane was used for gradient elution. Chromatography on lipophilic Sephadex (Lipidex 5000) employed 5% hexane in MeOH as the mobile phase. A Kofler hot stage was used for mps and all crystallizations were from MeOH. The ^1H NMR spectra of the 4-desmethylsterols were determined at 220 MHz, and 4,4-dimethyl compounds were examined at 360 MHz. MS (EI) were performed by Morgan–Schaffer of Montreal, Canada.

Whole *E. pulcherimma* plants (3 kg wet wt) without roots were extracted for 24 hr with Me_2CO in a Soxhlet apparatus. The residue, after removal of the solvent, was hydrolysed in refluxing 10% methanolic KOH. The Et_2O soluble material was chromatographed on Al_2O_3 (10% H_2O) and gave 4,4-dimethyl compounds, containing fatty alcohol and 2.2 g (0.07%) of 4-desmethyl compounds. An aliquot (50 mg) of the latter was separated into 3 fractions on lipophilic Sephadex, as described in the text along with the ^1H NMR spectra (Table 1). The 4,4-dimethyl fraction was found to contain fatty alcohols, which were concd in the first 2 of 4 dimethyl fractions. They will be described in another publication. Most of the fatty alcohols (575 mg) were precipitated by adding hot hexane and cooling. The hexane-soluble material (3 g) was rechromatographed on Al_2O_3 (4% H_2O). Lanosterol and 24-dihydrolanosterol moved somewhat faster than material with RR_t 1.61, which in turn moved somewhat faster than material with RR_t 2.42. The material in the first group of fractions mentioned was crystallized until only lanosterol and dihydrolanosterol were evident by GLC. These two substances (30 mg in a lanosterol to dihydrolanosterol ratio 4 to 3) were then separated from each other on lipophilic Sephadex. Lanosterol moved the faster and showed a ^1H NMR spectrum as in Table 1, which agreed with the literature [5–7]; MS *m/e* (rel. int.); 426 (43, M^+), 411 (100, $M^+ - Me$), 393 (50, $M^+ - Me - H_2O$), 259 (22, $M^+ - Me$ – side-chain $- C_3H_5$), 257 (13), 255 (15), 243 (10), 241 (24), 229 (18), and 215 (22). Dihydrolanosterol yielded a ^1H NMR spectrum as in Table 1, which agreed with the literature [6, 7]; MS *m/e* (rel. int.): 428 (22, M^+), 413 (100, $M^+ - Me$), 395 (57, $M^+ - Me - H_2O$), 261 (10, $M^+ - Me$ – side-chain $- C_3H_3$), 259 (15, $M^+ - Me$ – side-chain $- C_3H_5$), 255 (10), 247 (12), 243 (10), 241 (15), 229 (16), and 215 (19).

The second group of fractions from the Al_2O_3 column, which were enriched in the triterpenoid of RR_t 1.61, were successively recrystallized until only a GLC peak of RR_t 1.61 was present, giving mp 168–170°, ^1H NMR: δ (all singlets) 0.74, 0.77, 0.79, 0.83, 0.87, 0.88, 0.94, 0.97, 1.00, 1.02, 1.08 and 1.13. The peak at 0.94 was the strongest, representing at least two singlets in the major component. The peaks (mostly minor) at 0.79, 0.83, 0.87, 0.88, 0.94, 1.00, 1.02 and 1.13 corresponded with those described for β-amyrin [19], which has been reported in this plant [9]. The major peaks at 0.74, 0.77, 0.87, 0.88, 0.94, 0.97 and 1.08 probably represent germanicol, which has been reported to be the major triterpenoid in the plant [8, 9]. Based on our ^1H NMR spectrum, the β-amyrin to germanicol ratio was about 1 to 2. MS *m/e* (rel. int.): 426 (97, M^+), 411 (100, $M^+ - Me$), 393(21, $M^+ - Me - H_2O$), 272 (21), 271 (18), 259 (18), 258 (18), 257 (41), 255 (12), 243 (24), 241 (18), 232 (21), 231 (74), 230 (15), 229 (35), 220 (21), 219 (132), 218 (629), 217 (32), 215 (27), 208 (34), 207 (129), 206 (74), 205 (188), 204 (394), 203 (323), and 201 (41).

The third fraction, enriched in the component with RR_t 2.42, after crystallization melted at 216–218°; MS *m/e* (rel. int.): 426 (25, M^+), 411 (100, $M^+ - Me$), 393 (25, $M^+ - Me - H_2O$), 273 (19), 271 (17), 259 (102), 255 (20), 247 (16), 243 (22), 241 (52), 229 (19), 215 (12), 205 (14), 203 (12), and 201 (13); ^1H NMR: δ 0.73, 0.75 (both singlets not two coincident doublets, since separation only 5 Hz), 0.83 (*d*, $J = 6$ Hz), 0.85 (*s*), 0.89 (*s*), 0.90 (*d*, $J = 6$ Hz), 0.96 (*s*), and 0.99 (*s*).

Acknowledgements—The support of NIH Grant No. AM-12172 is gratefully acknowledged as is the opportunity to use the NIH-supported ^1H NMR facility at the University of Pennsylvania. We should also like to thank Fr. Ralph Chiefo of the Mater Dolorosa Church in Philadelphia for a gift of poinsettia plants, and Mr. John M. Joseph for determining the ^1H NMR spectra.

REFERENCES

1. Ponsinet, G. and Ourisson, G. (1968) *Phytochemistry* **7,** 89.
2. Nielsen, P. E., Nishimura, H., Liang, Y. and Calvin, M. (1979) *Phytochemistry* **18,** 103.
3. Pant, P. and Rastogi, R. P. (1979) *Phytochemistry* **18,** 1095.
4. Ourisson, G., Crabbé, P. and Rodig, O. (1964) *Tetracyclic Triterpenes*, p. 160. Holden-Day, San Francisco.
5. Piatak, D. M. and Reimann, K. A. (1972) *Tetrahedron Letters*, 4525.
6. Itoh, T., Tumura, T. and Matsumoto, T. (1976) *Steroids* **27,** 275.
7. Itoh, T., Tamura, T. and Matsumoto, T. (1976) *Lipids* **11,** 434.
8. Khastgir, H. N., and Pradhan, B. P. (1967) *J. Indian Chem. Soc.* **44,** 159.
9. Dominguez, X. A., Delgado, J. G., Maffey, M. L., Mares, J. G. and Rombold, C. (1967) *J. Pharm. Sci.* **56,** 1184.
10. Nes, W. R., Krevitz, K., and Behzadan, S. (1976) *Lipids* **11,** 118.
11. Rubinstein, I., Goad, L. J., Clague, A. D. H. and Mulheirn, L. J. (1976) *Phytochemistry* **15,** 195.
12. Nes, W. R. (1977) *Adv. Lipid Res.* **15,** 233.
13. Nes, W. R. and McKean, M. L. (1977) *Biochemistry of Steroids and Other Isopentenoids.* University Park Press, Baltimore.
14. Nes, W. R., Krevitz, K., Joseph, J., Nes, W. D., Harris, B., Gibbons, G. F. and Patterson, G. W. (1977) *Lipids* **12,** 511.
15. Goad, L. J. (1977) in *Lipids and Lipid Polymers in Higher Plants* (Tevini, M. and Lichtenthaler, H. K., eds.) p. 146. Springer, Berlin.
16. Karrer, W. (1958) *Konstitution und Vorkommen der Organischen Pflanzenstoffe*, p. 821. Birkhäuser, Basel.
17. Nes, W. R., Varkey, T. E. and Krevitz, K. (1977) *J. Am. Chem. Soc.* **99,** 260.
18. Posinet, G. and Ourisson, G. (1967) *Phytochemistry* **6,** 1235.
19. Tursch, B., Savoir, R., Ottinger, R. and Chiurdoglu (1976) *Tetrahedron Letters* 539.