ACTA CHEMICA SCANDINAVICA 27 (1973) 2405–2410

Tobacco Chemistry

20. Structures and Syntheses of Three New Tobacco Constituents of Probable Isoprenoid Origin

ARNE J. AASEN, JOSEPH R. HLUBUCEK, SVEN-OLOF ALMQVIST, BJARNE KIMLAND and CURT R. ENZELL

Research Department, Swedish Tobacco Co., Box 17007, S-104 62 Stockholm 17, Sweden

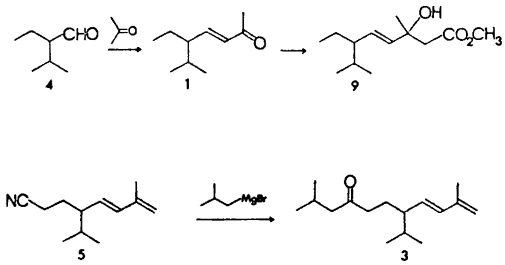
The structures of three new tobacco isolates, 5ξ-isopropyl-3*E*-hepten-2-one (1), 3ξ-hydroxy-3ξ-methyl-6ξ-isopropyl-4*E*-octenoic acid (2), and 2,10-dimethyl-7ξ-isopropyl-8*E*,10-undecadien-4-one (3) have been elucidated on the basis of their spectral data and subsequently confirmed by total syntheses. The possible isoprenoid origin of these compounds is discussed.

Earlier work in this laboratory on the volatiles of Greek tobacco, *Nicotiana tabacum* L., has led to the isolation of three new minor constituents.[1–3] In the present communication we wish to report their structures, which on account of the small amounts available were deduced solely from their spectral data, and syntheses.

*5ξ-Isopropyl-3*E*-hepten-2-one (1).* The presence of a 2-oxo-3*E*-pentenylidene moiety, $CH_3CO–CH=CH–CH$, in this tobacco isolate having the elemental composition $C_{10}H_{18}O$ (accurate mass determination), was apparent from its IR (1677 cm^{-1}, conj. CO; 987 cm^{-1}, *trans* disubst. C=C), UV (223 nm, monosubst. conj. CO), and NMR spectra [δ 2.24, CH_3CO; δ 6.04 (d, J 16 Hz) and δ 6.6 (q, J 8.5 and 16 Hz), $CO–CH=CH–CH$]. The remaining two saturated alkyl residues accounting for C_5H_{12}, attached to the vinylic carbon atom, had to be either an ethyl group and an isopropyl group, or a methyl group and a *sec*-butyl group since the NMR spectrum exhibited one triplet (δ 0.83, J 7 Hz) and two doublets (δ 0.86, J 6.5 Hz and δ 0.91, J 6.5 Hz) in the methyl region. The ethyl-isopropyl alternative was favoured on account of bands at 1361 cm^{-1} and 1369 cm^{-1} indicative of a *gem.* dimethyl grouping and since an allylic methyl group would resonate further downfield, *i.e.* at *ca.* 1 ppm.[4] Synthesis of the racemate of this compound performed by an aldol condensation of 2-isopropylbutyraldehyde (*4*) with acetone followed by *in situ* de-

hydration furnished a product with IR, NMR, and mass spectra identical to those of the tobacco isolate thereby confirming the assigned structure.

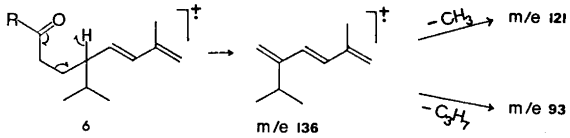
*3ξ-Hydroxy-3ξ-methyl-6ξ-isopropyl-4*E*-octenoic acid (2).* The methyl ester (*9*) of this acid, for which the spectral data were obtained, displayed hydroxyl absorption in the infrared (3500 cm^{-1}) clearly suggesting that it had



the composition $C_{13}H_{24}O_3$ rather than $C_{13}H_{22}O_2$ (M − 18) determined by accurate mass measurement for the heaviest mass, *m/e* 210, observed in the mass spectrum. A three-proton singlet at δ 1.31 was ascribed to a methyl group attached to a tertiary carbinol group which, from the NMR spectrum after addition of $Eu(DPM)_3$, had as further substituents a $-CH_2COOR$ (δ 2.56, s) and a $-CH=CH-CHR_1R_2$ grouping (δ 5.45, m and 1.58, m). The addition of the shift reagent separated the methylene protons of the $-CH_2COOR$ group into an AB system ($r = 1.28$ and 1.21, J 16 Hz) and resolved the olefinic proton multiplet into a doublet ($r = 0.87$, J 16 Hz) and a quartet ($r = 0.68$, J 9 and 16 Hz), thereby revealing the *trans* configuration of the double bond. The partial structure $R_1R_2CH-CH=CH-C(OH,CH_3)-CH_2-COOCH_3$ follows from these data and also that R_1 and R_2 must be two saturated alkyl substituents to account for the remainder of the molecule, C_5H_{12}. Three three-proton signals constituting one triplet and two doublets (δ 0.78, 0.79, 0.84; all J's 7 Hz), left the same two possible sets of alkyl residues as for 5ξ-isopropyl-3*E*-hepten-3-one (1). The ethyl-isopropyl alternative was favoured over that comprising a methyl and *sec*-butyl group for the following reasons: the infrared spectrum suggested the presence of a *gem.* dimethyl grouping (1384 and 1369 cm^{-1}), irradiation of the olefinic proton at $r = 0.68$ simplified the multiplet of the methine proton ($r = 0.26$) to a doublet of triplets demonstrating further spin-spin coupling to three rather than four protons, and none of the methyl signals occurred at sufficiently low field to correspond to an allylic methyl group.[4] It follows therefore that the tobacco constituent is 3ξ-hydroxy-3ξ-methyl-6ξ-isopropyl-4*E*-octenoic acid (*2*), which judging from the NMR spectrum after addition of shift reagent is partly racemic (ratio 1:4). This structural assignment was confirmed by comparison of the methyl ester of the tobacco acid with a synthetic specimen prepared by a Reformatsky reaction of (±)-5-isopropyl-3*E*-hepten-3-one (*1*) with methyl bromoacetate.

*2,10-Dimethyl-7ξ-isopropyl-8*E*,10-undecadien-4-one (3).* A small quantity of this compound was isolated about two years ago but no structure could then be deduced. In the light of subsequent work and deepening interest in the

biogenesis of tobacco nor-isoprenoids, we recently reconsidered its NMR and mass spectra. Due mainly to the presence of signals in the NMR spectrum corresponding to a 2-methyl-1,3E-pentadienylidene moiety (three-proton triplet at δ 1.82, J 1.1 Hz; two-proton narrow multiplet at δ 4.86; one-proton quartet at δ 5.33, J 8 and 16 Hz; one proton doublet at δ 6.06, J 16 Hz), and the striking resemblance of its mass spectrum to that of solanone (*6*, R = CH_3), we were able to postulate two possible structures for this compound, namely 2,10-dimethyl-7ξ-isopropyl-8*E*,10-undecadien-4-one (*3*) or 8ξ-isopropyl-11-methyl-9*E*,11-dodecadien-5-one (*7*).



Scheme 1.

The mass spectral fragmentation of solanone (*6*, R = CH_3) is depicted in Scheme 1 accounting for the formation of the three characteristic ions at *m/e* 93, 121, and 136 which presumably are present irrespective of which saturated alkyl residue (R) is attached to the carbonyl group. The molecular weight of the tobacco compound, 236, corresponds to R = C_4H_9, *i.e.* butyl, isobutyl, *sec*-butyl, or *t*-butyl. The last two possibilities were excluded since the NMR spectrum disclosed the presence of four protons adjacent to the carbonyl group (δ 2.1 – 2.45). Considerable overlap in the methyl region of the NMR spectrum did not enable us to clearly distinguish between the butyl and isobutyl alternatives. Although consideration of the integrated methyl region favoured the latter, the previously advanced hypothesis that the structurally related solanone (*6*, R = CH_3) and other tobacco constituents are degradation products of diterpenoids possessing the thunbergane skeleton * (*10*),[5–7] made the butyl possibility also appear likely. Subsequent synthesis of 8ξ-isopropyl-11-methyl-9*E*,11-dodecadien-5-one (*7*) performed by reacting 2-methyl-5-isopropyl-7-cyano-1,3*E*-heptadiene (*5*) [8] with butylmagnesium bromide in analogy with the last step in Johnson and Nicholson's [8] elegant synthesis of solanone (*6*, R = CH_3), furnished a product which exhibited a mass spectrum nearly identical to that of the tobacco constituent. However, the two specimens separated when co-injected on a capillary GC column and their NMR spectra displayed differences. The isobutyl-isomer was therefore prepared next by the same procedure using isobutylmagnesium bromide. The product displayed NMR and mass spectra which were indistinguishable from those recorded for the tobacco constituent and no separation could be observed when co-injected on a capillary GC column. It follows therefore that the tobacco constituent is 2,10-dimethyl-7ξ-isopropyl-8*E*,10-undecadien-4-one (*3*).

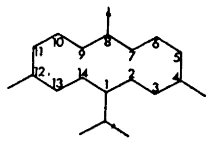
The compound with the *sec*-butyl end group – 3,10-dimethyl-7-isopropyl-8*E*,10-undecadien-4-one (*8*) – was also prepared and its mass spectrum

* Nomenclature according to J. W. Rowe, Oct. 1968; personal communication.

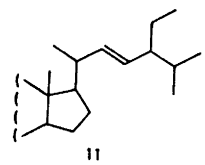
exhibited, in addition to the expected set of ions at *m/e* 93, 121 and 136, an intense *m/e* 57 ion due to cleavage of the C(3) – C(4) bond which is α both to the carbonyl and a disubstituted carbon atom.

Biogenetic considerations

Several tobacco constituents possessing an isopropyl group protruding from a methyl substituted chain, *e.g.* solanone (*6*, R = CH_3), have been postulated to be degradation products of tobacco diterpenoids belonging to the thunbergane (cembrane, *10*) class.[5–7] The skeletons of the present three compounds may formally be derived from that of thunbergane, requiring for the enone (*1*) and the acid (*2*), if not more directly interrelated, cleavages of the 4,5- and 12,13-bonds and of the 6,7- and 12,13-bonds, respectively. The dienone (*3*), which has many structural features in common with solanone (*6*, R = CH_3), would require more elaborate changes such as formation of a bond between C(6) and C(13) and cleavages of the 5,6- and 12,13-bonds, plus cleavage of the 9,10-bond. In view of the multiplicity of functional groups encountered in these macrocyclic diterpenoids and the number of reactions



10



11

that can be envisaged, a variety of routes to *1*, *2*, and *3* may be postulated. However, further examples and preferably a knowledge of the steric relationships, not presently at hand, seem desirable before considering these in detail.

Presently, alternative pathways might also be considered, notably those involving alkylation such as encountered in certain isoprenoids, *e.g.* irones,[9] juvenile hormones,[10] steroids,[11] and C_{50}-carotenoids.[12] Of particular interest in this context is the structural similarity between the side chain of stigmasterol (*11*), a steroid known to occur in tobacco,[13] and the enone (*1*) and the acid (*2*).

EXPERIMENTAL

NMR, IR, UV, and mass spectra were recorded on Varian HA 100D and A60-A, Digilab FTS-14, Perkin-Elmer 257, Beckmann DK-2A and LKB 9000 (70 eV) instruments, respectively. The mass spectra were obtained by gas chromatography in combination with the mass spectrometer (GLC–MS) using steel capillary columns (0.5 mm × 50 m, Handy and Harman grade 316-S) coated with emulphor using the dynamic packing method.[14] Preparative GLC was performed with a Varian 1700 instrument using a 3.2 mm × 3 m glass column packed with 5 % Carbowax 20 M on Chromosorb G. Fractions were collected at −70° in U-shaped teflon tubes (i.d. 5 mm) equipped with an electrostatic precipitator.[15] Accurate mass determinations were carried out at the Laboratory for Mass Spectrometry, Karolinska Institutet, Stockholm. The solvents, silica gel, and drying agents were purified as previously described.[16]

Isolation

5ξ-Isopropyl-3E-hepten-2-one (*1*, 3 mg) was isolated from fraction "B 3" [1] by preparative gas chromatography.[3] MS: 154 (M^+, 3.5) 43 (100), 97 (45), 55 (34), 111 (34), 112 (34), 69 (30), 41 (23), 39 (11), 125 (11), 84 (7); accurate mass determination: $C_{10}H_{18}O$, found 154.1361, calc. 154.1358; λ_{max} (EtOH) 223 nm (ε 14 600); $[\alpha]_D^{20}+4.7°$ (*c* 0.4 in ether); ν(film): 2961 (s), 2931 (m), 2872 (m), 1696 (m), 1677 (s), 1625 (m), 1460 (m), 1387 (w), 1369 (m), 1361 (m), 1255 (s), 1180 (w), 987 (m); $\delta(CDCl_3)$: 0.83 (3 H, t, *J* 7 Hz), 0.86 (3 H, d, *J* 6.5 Hz), 0.91 (3 H, d, *J* 6.5 Hz), 1.1–2.0 (4 H, m), 2.24 (3. H, s), 6.04 (1 H, d, *J* 16 Hz), 6.6 (1 H, quartet, *J* 8.5 and 16 Hz).

3ξ-Hydroxy-3ξ-methyl-6ξ-isopropyl-4E-octenoic acid (*2*, 17 mg) was isolated as the corresponding methyl ester (*9*) from 'Acids' [1] by preparative gas chromatography.[2] MS: M^+ at *m/e* 228 was not seen, 43 (100), 97 (39), 111 (28), 55 (27), 112 (24), 69 (23), 41 (23), 71 (14); accurate mass determination: $C_{13}H_{22}O_2$ (M−18), found 210.1620, calc. 210.1620; ν(film): 3500 (broad), 2959 (s), 2931 (m), 2875 (m), 1730–1720 (s), 1440 (m), 1384 (m), 1369 (m), 1338 (m), 1209 (s), 1175 (m), 1128 (w), 1012 (w), 978 (m), 939 (w) cm^{-1}; $\delta(CDCl_3)$: 0.78 (3 H, t, *J* 7 Hz), 0.79 (3 H, d, *J* 7 Hz), 0.84 (3 H, d, *J* 7 Hz), 1.2–1.8 (4 H, m), 1.31 (3 H, s), 2.56 (2 H, s), 3.66 (3 H, s), 3.82 (OH, broad singlet), 5.42 (2 H, m). Addition of $Eu(DPM)_3$: *r* (relative induced shift ratio) [17] = 1.28 (1 H, d, *J* 16 Hz), r = 1.21 (1 H, d, *J* 16 Hz), *r* = 1 (3 H, s), r = 0.87 (1 H, d, J 16 Hz), *r* = 0.68 (1 H, quartet, *J* 9 and 16 Hz), *r* = 0.26 (1 H, m), *r* = 0.13 (3 H, t, *J* 7 Hz), *r* = 0.11 (6 H, d, *J* 7 Hz), *r* = 0.31 (3 H, s). Extrapolation to zero addition of $Eu(DPM)_3$ indicated the chemical shift of C(6)H having *r* = 0.26 to be δ 1.58. The coupling (*J* 9 Hz) between this proton and C(5)H with *r* = 0.68 was demonstrated in spin decoupling experiments. Irradiation of C(5)H in the LIS-doped sample simplified the multiplet having r = 0.26 to a doublet of triplets. $[\alpha]_D^{25}-0.8°$ (*c* 0.5 in chloroform). The NMR spectra obtained after addition of $Eu(DPM)_3$ indicated a diastereomeric mixture 1:4.

2,10-Dimethyl-7ξ-isopropyl-8E,10-undecadien-4-one (*3*, 3.4 mg) was isolated about two years ago from fraction "C 9" [1] by preparative gas chromatography. MS: *m/e* 236 (M^+, 8), 93 (100), 121 (53), 57 (53), 136 (48), 85 (42), 41 (35), 43 (28), 79 (20), 81 (17); NMR was found identical to that of synthetic *3*.

Synthesis

(±)-5-Isopropyl-3E-hepten-2-one(1). A mixture of (±)-2-isopropylbutyraldehyde (*4*), (3.4 g), acetone (25 ml) and 10 % aqueous KOH solution (3 ml) was refluxed on a steam-bath for 16 h. The cooled reaction mixture was concentrated *in vacuo*, extracted with pentane, and the combined pentane extracts washed with water and dried over sodium sulphate. The solvent was removed under reduced pressure to leave *1* as a straw-coloured liquid (3.9 g, purity: > 95 %, yield 85 %). The physical properties of the product were indistinguishable from those of the natural compound (*vide supra*).

(±)-3-Hydroxy-3-methyl-6-isopropyl-4E-octenoic acid (2). A mixture of (±)-5-isopropyl-3*E*-hepten-2-one (*1*, 1.54 g), methyl bromoacetate (1.68 g) and activated [18] zinc powder (0.7 g) in dry benzene (40 ml) was refluxed with stirring under nitrogen. A vigorous reaction started shortly after refluxing began and the reaction mixture turned a dark green colour. The refluxing and stirring was continued for 1 h after which the reaction mixture was cooled, treated with ice-cold 10 % aqueous acetic acid and extracted with ether. The ether-benzene solution was washed with 10 % aqueous acetic acid, 5 % $NaHCO_3$ solution, water and finally dried over anhydrous sodium sulphate. The solvents were removed *in vacuo* to leave *9* as a pale yellow liquid (2.1 g). The NMR, IR, and mass spectra of a sample purified by chromatography on silica gel were identical to those of the methyl ester isolated from tobacco (*vide supra*).

(±)-2,10-Dimethyl-7-isopropyl-8E,10-undecadien-4-one (3). (±)-2-Methyl-5-isopropyl-7-cyano-1,3*E*-heptadiene [8] (*5*, 405 mg) in dry ether (10 ml) was slowly added to a stirred solution of isobutylmagnesium bromide prepared from isobutyl bromide (2 g) and magnesium (340 mg) in 30 ml dry ether. After the addition, dry benzene (25 ml) was added and the ether removed by distillation. The mixture was refluxed for 24 h after which the solution was cooled and poured over crushed ice to hydrolyze the intermediate ketimine.

After stirring for 4 h at room temperature the benzene layer was removed and the aqueous phase extracted with ether. The solvent was removed by distillation and the residue chromatographed on a silica gel column. Elution with 1 % ether in pentane furnished 85 mg (16 %) of *3*, which was found indistinguishable from the tobacco compound when co-injected on a capillary column. Furthermore, the NMR and mass spectra (*vide supra*) were superimposable. At this stage the tobacco compound was only available in a reference sample of the fraction from which it had been isolated, and an IR spectrum of the natural compound could not be obtained for comparison. λ_{max} (EtOH): 230 nm (ε 23 600); ν(film): 2960 (s), 2875 (s), 1712 (s), 1609 (w), 1469 (m), 1410 (w), 1386 (m), 1369 (s), 1170 (w), 1142 (w), 1068 (w), 1037 (w), 972 (s), 884 (m) cm^{-1}; $\delta(CDCl_3)$: 0.85 (3 H, d), 0.89 (6 H, d), 0.90 (3 H, d), 1.3 – 1.8 (5 H, m), 1.82 (3 H, t, *J* 1.1 Hz), 2.1 – 2.45 (4 H, m), 4.86 (2 H, m), 5.33 (1 H, quartet, *J* 8 and 16 Hz), 6.06 (1 H, d, *J* 16 Hz).

(±)-*8-Isopropyl-11-methyl-9*E,*11-dodecadien-5-one* (*7*). *7* was prepared as described for *3* except that butylmagnesium bromide was used in the Grignard reaction. *7* disclosed a longer retention time than the tobacco compound (*3*) when co-chromatographed on a capillary column. MS: *m/e* 236 (M^+, 17), 93 (100), 136 (66), 121 (65), 85 (37), 57 (34), 41 (26), 79 (15), 43 (14), 81 (13). $\delta(CDCl_3)$: 0.86 (3 H, d), 0.90 (3 H, t), 0.91 (3 H, d), 1.2 – 1.8 (8 H, m), 1.82 (3 H, t, *J* 1.1 Hz), 2.34 (6 H, two triplets, *J* 7 Hz), 4.85 (2 H, m), 5.32 (1 H, quartet, *J* 8 and 16 Hz), 6.07 (1 H, d, *J* 16 Hz).

(±)-*3,10-Dimethyl-7-isopropyl-8*E,*10-undecadien-4-one* (*8*). *8* was synthesized as outlined for *3* except for using *sec*-butylmagnesium bromide in the Grignard coupling. MS: *m/e* 236 (M^+, 9.5), 93 (100), 57 (98), 121 (54), 136 (46), 41 (38), 85 (37), 43 (31), 79 (24), 81 (22). $\delta(CDCl_3)$: 0.7 – 1.1 (12 H, m), 1.2 – 2.0 (6 H, m), 1.82 (3 H, t, *J* 1.1 Hz) 2.36 (2 H, t, *J* 6.5 Hz), 4.85 (2 H, m), 5.32 (1 H, quartet, *J* 8 and 16 Hz), 6.05 (1 H, d, *J* 16 Hz).

Acknowledgements. The authors are indebted to Miss A.-M. Eklund for skilful technical assistance, Prof. K. Olsson, Agricultural College of Sweden, Uppsala, for placing the NMR instrument at their disposal, and Dr. R. R. Johnson, Brown and Williamson Tobacco Corporation, Louisville, Kentucky, for a generous gift of (±)-2-methyl-5-isopropyl-7-cyano-1,3*E*-heptadiene.

REFERENCES

1. Kimland, B., Aasen, A. J. and Enzell, C. R. *Acta Chem. Scand.* **26** (1972) 2177.
2. Kimland, B., Aasen, A. J., Almqvist, S.-O., Arpino, P. and Enzell, C. R. *Phytochemistry* **12** (1973) 835.
3. Hlubucek, J. R., Aasen, A. J., Kimland, B. and Enzell, C. R. *Phytochemistry* **12** (1973) 2555.
4. Slomp, G. and MacKellar, F. *J. Am. Chem. Soc.* **84** (1962) 204.
5. Kinzer, G. W., Page, T. F., Jr. and Johnson, R. R. *J. Org. Chem.* **31** (1966) 1797.
6. Kimland, B., Appleton, R. A., Aasen, A. J., Roeraade, J. and Enzell, C. R. *Phytochemistry* **11** (1972) 309.
7. Roberts, D. L. and Rohde, W. A. *Tobacco Science* **16** (1972) 107.
8. Johnson, R. R. and Nicholson, J. A. *J. Org. Chem.* **30** (1965) 2918.
9. Whiting, D. A. and Harper, S. H. In Coffey, S., Ed., *Rodd's Chemistry of Carbon Compounds*, Elsevier, Amsterdam 1968, Vol. II, part B, p. 153.
10. Karlson, P. In Goodwin, T. W., Ed., *Natural Substances Formed Biologically From Mevalonic Acid*, Academic, London 1970, p. 145.
11. Goad, L. J. *Ibid.* p. 45.
12. Straub, O. In Isler, O., Ed., *Carotenoids*, Birkhäuser, Basel 1971.
13. Carruthers, W. and Plimmer, J. R. *Chem. Ind. (London)* **1959** 48.
14. Dijkstra and de Goey, J. In Desty, D. H., Ed., *Gas Chromatography* 1958, Butterworths, London 1958, p. 56.
15. Kratz, P., Jacobs, M. and Mitzner, B. M. *Analyst* **84** (1959) 671.
16. Appleton, R. A., Enzell, C. R. and Kimland, B. *Beitr. Tabakforsch.* **5** (1970) 266.
17. Wineburg, J. P. and Swern, D. *J. Am. Oil. Chemists' Soc.* **49** (1972) 267.
18. Hauser, C. R. and Breslow, D. S. *Org. Syn. Coll. Vol.* **3** (1955) 408.

Received March 17, 1973.