Phytochemistry, 1972, Vol. 11, pp. 3485 to 3490. Pergamon Press. Printed in England.

ALKALOIDS OF CIGARETTE SMOKE CONDENSATE

E. V. BROWN and I. AHMAD
Department of Chemistry, University of Kentucky, Lexington, U.S.A.

(*Received* 28 *April* 1972. *Accepted* 24 *July* 1972)

Key Word Index—*Nicotiana tabacum*; Solanaceae; tobacco smoke; alkaloids; 2,6-dimethylquinoline; 2,2′-dipyridyl; *N*-methylanabasine.

Abstract—The high boiling basic portion of tobacco smoke condensate from the IRI research cigarette has been studied. Three previously unreported alkaloids were isolated and their structures conclusively established. Five other alkaloids, previously suggested to be present, have been conclusively characterized by NMR and MS analysis. Two new basic compounds with MWs of 190 and 243 respectively have been detected.

INTRODUCTION

SEVERAL earlier investigators have studied the chemical composition of the basic portion of tobacco smoke condensate and a large number of the compounds have been reported in it.[1,2] However, the characterization of a considerable portion of the basic tar, particularly that containing high boiling bases, still remains to be accomplished. Previous authors have generally used PC or TLC; in our study we have used GLC, NMR and MS spectrometry.

RESULTS

In order to achieve a better resolution of this complex mixture, we found it essential to submit the condensate to liquid column chromatography prior to GLC. The results described here deal with those fractions obtained from the alumina (basic) column chromatography as this column gave the best separation of the several sorbents used (i.e. silica, florisil, Sephadex LH20, cellulose, and magnesia). Quin[3] was the first to use GLC for the isolation of tobacco smoke alkaloids. Later Schmeltz *et al.*[4] and Osman[5] employed GLC using temperature programming. All of these investigators isolated mysomine, nornicotine, and anabasine, but failed to isolate nicotyrine. The methods of identification used by these authors[3,4] were generally comparison of R_fs (by PC) and in certain cases UV and/or IR analysis. In our experience the use of TLC, GLC, and PC individually, as methods for identification, proved very unsatisfactory and misleading because some of the components which appeared as a single uniform peak on their GLC were later resolved into two or three components. We have successfully isolated (by GLC) relatively pure samples of several compounds (as shown in Table 1) and have conclusively identified them. GLC of Fraction II-3 revealed at least 23 components, but the compounds discussed in this paper are those with which we could perform a detailed structural study.

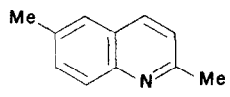
[1] R. A. W. JOHNSTONE and J. R. PLIMMER. *Chem. Rev.* 885 (1959).
[2] R. L. STEDMAN, *Chem. Rev.* 153 (1968).
[3] L. D. QUIN, *J. Org. Chem.* **24**, 911 (1959); **24**, 914 (1959).
[4] I. SCHMELTZ, R. L. STEDMAN, W. J. CHAMBERLAIN and D. BURDICK, *J. Sci. Food Agric.* **15**, 774 (1964).
[5] S. OSMAN and J. BARSON, *Phytochem.* **3**, 587 (1964).

TABLE 1. COMPONENTS IDENTIFIED IN THE HIGH BOILING BASIC FRACTION OF CIGARETTE SMOKE

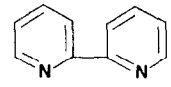
Compound	GLC R_t* (Min) column temp. 240°, HE flow 60 ml/min Unknown	Known	Method of identification	Previously used methods of identification
Myosmine	3·4	—	MS, GLC, NMR	PC
2,6-Dimethyl quinoline (I)	4·4	4·4	MS, GLC, NMR	—
β-Nicotyrine	4·7	4·7	MS, GLC, TLC	PC
2,2′-Bipyridyl (II)	6·7	6·7	MS, GLC, TLC	—
Anabasine	8·4	8·4	MS, GLC, NMR	PC, GLC
Nicotinamide	9·6	9·6	MS, GLC, TLC	Microbiol, PC essay
N-Methylanabasine (III)	13·2	—	MS, NMR	—
Harmane	14·7	14·7	MS, GLC, NMR	IR, UV
Compound MW 190	15·6	—	MS, NMR	—
Compound MW 243	23·0	—	MS	—

* These R_ts are representative of the various peaks collected by GLC using temperature programming according to the conditions described in the Experimental.

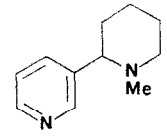
The peak with R_f 3·4 min was identified as myosmine. The MS of this peak showed a parent ion peak at *m/e* 146 in a very large abundance. A loss of H atom gave a less abundant peak at *m/e* 145, which on a subsequent loss of 27 m.u. (due to a molecule of HCN) gave the most abundant peak at *m/e* 118. The other significant peaks appeared at *m/e* 106, 105, 91, 78, 77, 76, 65, 51, 50, 42, 41 and 39. A large abundance of a peak at *m/e* 78 and the behavior of its fragmentation pattern strongly suggests the presence of pyridyl moiety. This is further supported by its NMR spectrum showing characteristic chemical shifts for β-substituted pyridine protons δ7·4 (*m*, H, C_4), δ7·62 (*m*, H, C_3), δ8·4 (*m*, H, C_1 and C_5). The NMR spectrum also showed a broad multiplet at *ca.* δ3·1 due to CH_2 protons of 1-pyrroline. The identification of this compound has previously been claimed by PC comparison.[5]



2, 6-Dimethylquinoline
(I)



2, 2′-Dipyridyl
(II)



N-Methylanabasine
(III)

The peak with R_t 4·4 min was identified as 2,6-dimethylquinoline (I). The presence of dimethylquinoline has been recently reported by Neurath.[6] However, the position of various methyl groups was not assigned. Furthermore, in their MS study, which they used as a tool for identification, only two fragmentation ions *m/e* 156 and 154 were shown. We collected a sufficient amount of the material by repeated chromatography and were able to perform NMR and detailed MS studies. The MS of this peak showed a molecular ion peak at *m/e* 157 which also appeared as the base peak. The other significant ions appeared at *m/e* 142 (loss f Me), 130, 129, 128, 116, 115, 103 and 102. The cracking pattern of this unknown

[6] G. NEURATH and M. DÜNGER, *Beitr. Tabakforsh.* **5** (1), 1 (1969).

was found to be identical to that of a known sample of 2,6-dimethylquinoline. The NMR spectrum showed two sharp singlets for the two–Me groups which are attached at position 2 (δ2·38) and position 6 (δ2·72) respectively. Multiplets due to CH protons at positions 3 and 4 in the pyridine ring were observed at (δ7·2) and (δ7·9) respectively. The NMR spectrum of a known sample of 2,6-dimethylquinoline fits well with that of the unknown sample.

The peak with R_t 4·7 min was identified as β-nicotyrine by NMR and MS. The MS of this peak showed a parent ion peak at *m/e* 158. A substantial peak appeared at *m/e* 157 due to the loss of hydrogen. A loss of HCN was observed giving a peak at *m/e* 130. This in turn lost another 27 m.u. for the loss of HCN which gave a substantial peak at *m/e* 117. The other significant peaks appeared at *m/e* 105, 104, 78, 77, 51 and 50. A characteristic fragmentation pattern showing 78, 77, 51 and 50 suggested the presence of a pyridyl moiety. Although the NMR spectrum was not very sharp (due to the small amount of material) the presence of an *N*-methylpyrrol moiety was suggested by the presence of signals at δ3·26 (*s*, Me attached to ring nitrogen); δ6·1 (*m*, Me proton) and δ6·52 (*m*, C_2 and C_5 protons). Characteristic signals for pyridine protons were also observed. The comparison of the MS of the unknown with that of an authentic sample of β-nicotyrine showed them to be identical. The identification of this compound has been previously claimed by PC.[7]

The peak with a R_t 6·7 min was identified as 2,2′-bipyridyl (II). The mass spectrum of this peak gave a molecular ion peak at *m/e* 156. A loss of a H atom resulted in a less abundant peak at *m/e* 155. The other significant peaks appeared at *m/e* 130, 129, 106, 78, 77 and 51. The peak at *m/e* 129 was formed by the elimination of a molecule of HCN. The fragmentation pattern for the formation of peaks at *m/e* 78, 77 and 51 is characteristic of a pyridyl moiety. The comparison of this spectrum with that of an authentic sample of 2,2′-bipyridyl showed them to be identical. The presence of 2,3′-bipyridyl has previously been reported[8] but 2,2′-bipyridyl has never been shown to be a tobacco or smoke constituent. These two isomers on GLC give very close retention times with a very poor resolution. However, by MS analysis the two isomers could be clearly differentiated. In the case of 2,3′-bipyridyl the peaks at *m/e* 129, 51 and 50 appeared in very large abundance whereas in the case of 2,2′-bipyridyl, the corresponding peaks were extremely small in abundance. The R_t (from GLC) and R_f (from TLC) of the unknown corresponded to that of an authentic sample of 2,2′-bipyridyl. The above data revealed the identity of the peak to be 2,2′-bipyridyl.

The peak with a R_t 8·4 min gave a molecular ion peak at *m/e* 162. The most abundant peak was formed at *m/e* 84 with a loss of 78 m.u. which may be due to the elimination of a pyridyl ion. Nicotine and anabasine both have a MW of 162 but the appearance of peaks at *m/e* 106 and 105 in large abundances ruled out the possibility of the presence of nicotine. The absence of N–Me protons in the NMR spectrum further eliminated the possibility of nicotine. The comparison of the NMR and MS of this unknown with that of an authentic sample of anabasine showed them to be identical.

The peak with a R_t 9·6 min as shown in Table 1 was identified as nicotinamide. The MS of the component corresponding to this peak showed a molecular ion peak at *m/e* 122. An abundant peak at *m/e* 106 was observed which may account for the loss of NH_2. The further loss of 28 units gave a significant peak at *m/e* 78 due to the loss of CO. The other prominent peaks appeared at *m/e* 77, 76, 51 and 42 showing a characteristic fragmentation pattern of a pyridyl moiety. The comparison of the MS of this unknown with that of an authentic

[7] N. IVANOVE, *Compt. Rend. Acad. Bulgaria Sci.* **12**, 317 (1959); *Chem. Abs.* **54**, 93, 200 (1960).
[8] L. D. QUIN, *Tobacco Chemists Research Conference*, Durham, N.C. (1958).

sample of nicotinamide showed them to be identical. Both the unknown and the known samples of nicotinamide gave identical R_t when gas chromatographed using various experimental conditions. This compound has previously been shown to be present[9] but the details on its structural identification are lacking.

The peak with a R_t 13·2 min as shown in Table 1 was identified as *N*-methylanabasine (III). The compound has previously been cited to be present in tobacco,[10] but to our present knowledge there exists no earlier report of its presence in tobacco smoke. The MS of this peak showed a parent ion peak at *m/e* 176 which lost 29 m.u. to give an abundant peak at *m/e* 147. The base peak, which appeared at *m/e* 98, seemed to have resulted from a cleavage of a pyridine–piperidine bond and a loss of a pyridine ring from the parent ion. The pyridyl moiety is further suggested by a characteristic cracking pattern which gave peaks at *m/e* 78, 77, 76, 51 and 42. The other significant peaks at *m/e* 147, 120, 119 and 106 suggested a cracking pattern similar to that of α-substituted piperidine. The NMR spectrum of this unknown suggested the presence of Me protons (δ2·12) which are probably attached to a ring nitrogen. The presence of a broad multiplet centered at ~δ3·15 suggested CH_2 protons at positions 2 and 5 in piperidine ring. A broad multiplet superimposed by the Me signals may be due to CH_2 protons at positions 3 and 4 in the piperidine ring. A characteristic pattern for the pyridine protons was exhibited, δ7·31 (H, C_4, *m*), δ7·75 (H, C_3, *m*) and δ8·6 (H, C_5, *m*). The above data revealed the structural identification of the unknown to be *N*-methylanabasine.

The peak with R_t 14·7 min from GLC was identified as harmane. The compound was previously isolated by the relatively long and tedious process of column and PC.[11,12] However, to our knowledge this is the first report on its isolation by GLC and identification by its MS and NMR analysis. Poindexter[12] has reported its identification based on its IR and UV analysis. However, the close examination of the IR spectra reveals a rather weak and poor resolution (which may, perhaps, be due to insufficient quantities of a purer sample). In our studies we were able to clearly establish the structure of this unknown to be harmane. The MS fragmentation gave a parent ion peak at *m/e* 182 which also appeared as the most abundant ion suggesting the stable nature of the molecule. The $(P-1)^+$ peak was formed in considerable abundance with a loss of H. A peak with a large abundance appeared at *m/e* 154 due to a loss of HCN. This in turn lost another molecule of HCN to give an abundant peak at *m/e* 127. A peak with a small abundance was observed at *m/e* 141 with loss of 41 m.u. from the parent ion, and is probably formed by the loss of CH_3CN from the parent ion. The fragmentation of a known sample of harmane fits well with that of the unknown. The presence of various fragments was proved by comparing our results with those of Coutts *et al.*[13] The NMR spectrum of this unknown showed a singlet at δ2·38 suggesting the presence of a H proton attached to ring nitrogen. A characteristic multiplet for an aromatic proton at ~δ7·3 was observed. The GLC R_t of the unknown was identical to that of a known sample of harmane.

The peak with R_t 15·6 min as shown in Table 1 on its MS analysis showed a parent ion peak at *m/e* 190. The molecular ion appeared in a low concentration which suggests that either the compound has several heteroatoms or is a branched chain. An abundant peak appeared at (P-14) which in turn lost 29 m.u. (probably due to C_2H_5), to give the most

[9] D. A. BUYSUE, J. M. FLOWERS, P. WILDER and M. E. HOBBS, *Science* **124**, 1080 (1956).
[10] E. SPATH and F. KESZTHER, *Chem. Ber.* **70**, 704 (1937).
[11] *A Testa and Testa Ann Direc. Studies Equipment.* SEITA, Sect. **1**, 3, 103 (1965).
[12] E. H. POINDEXTER, JR. and R. D. CARPENTER, *Phys. Chem.* **1**, 215 (1962).
[13] R. T. COUTTS, R. A. LOCOCK and G. W. A. SLYWKA, *Org. Mass Spectos.* **3**, 879 (1970).

abundant peak at *m/e* 147. The ratio of intensities of P+1/P+ suggest the maximum carbon limits to be 12 or 13. The base peak at 147 with a loss of 27 m.u. (due to elimination of HCN) gave an abundant peak at *m/e* 120. The other significant peaks appeared at *m/e* 119, 116, 105, 98, 78, 77, 70, 51, 52, 42 and 39. The fragmentation pattern from the base peak (*m/e* 147) downward is quite analogous to that of nornicotine ($C_9H_{12}N_2$). The NMR spectrum showed weak signals at ca. δ1·3 (*m*, Me), δ2·2 (*m*, Me 3H), δ3·62 (*m*, CH^{Me}_{Me} and a multiplet centered at δ7·56 due to pyridine protons. Due to insufficient quantities of the material, complete identification could not be achieved. However, there appears no previous report showing the presence of a compound with MW 190 in the basic portion of the cigarette smoke condensate.

The peak with the R_t 23·0 min in its MS showed a molecular ion peak at *m/e* 243. A peak with significant abundance appeared at *m/e* 242 due to loss of a hydrogen atom. A loss of 30 m.u. from the parent ion resulted in a less abundant ion peak at *m/e* 204 which in turn lost 28 m.u. to give a large peak at *m/e* 176. The most abundant ion appeared at *m/e* 98. The other significant peaks appeared at *m/e* 71, 69 and 41. Since the parent ion suggests an odd MW, the unknown may have an odd number of nitrogen atoms. The loss of 67 m.u. from the parent ion may be due to C_4H_5N. This transition is supported by the metastable ion. The loss of 78 m.u. from *m/e* 176 to give a base peak may be due to elimination of a pyridyl ion. The cracking pattern of the unknown from its base peak (*m/e* 98) was very analogous to that of methylpiperidine. Taking this information into consideration, one may suggest the molecular formula of the unknown (MW 243) to be $C_{15}H_{21}N_3$. Due to insufficient quantities of the material, complete identification could not be achieved. However, it may be pointed out, there exists no previous report on the isolation of a compound with this MW.

EXPERIMENTAL

Cigarette smoking and condensate preparation. 5000 cigarettes (IR-I), unless otherwise mentioned, were smoked with a Mason smoking machine using the standard conditions of 2 sec a puff with 35 ml puff vol. and at a rate of 1 puff per min. Cigarettes were smoked to an average butt length of 23 mm. The condensate was collected in traps as described by Elmenhorst,[14] and cooled to −78°. The condensate was removed from the traps (using no solvent) and transferred to precooled jars and stored at −80° until it was used. The H_2O content of the condensate as determined by GLC averaged to 22% of the tarr.

Isolation of the basic fraction from the total smoke condensate and its fractionation. The condensate (480 g) was dissolved in 3·0 l. of Et_2O and left standing for about 12 hr with occasional shaking. The Et_2O solution was treated with 1 N HCl. The acid solution was saturated with NaCl and extracted with Et_2O (1·0 l.). The aq. acid layer was then adjusted to pH 10–11 with 1 N NaOH, saturated with NaCl and extracted with Et_2O (500 ml × 5). The Et_2O extract was dried and the solvent was removed giving a brown oil (35·8 g). Distillation on a spinning band column gave two main fractions; (I) low boiling bases (21 g), and (II) high boiling bases (10·5 g) (containing nicotine and other higher boiling bases). Fraction II was chromatographed on a column (1·5 × 45 mm) packed with basic alumina (400 g). The column was eluted with solvents in the order of benzene, benzene–$CHCl_3$ mixture (1:9, 1:4–4:1), $CHCl_3$, $CHCl_3$–MeOH mixtures (1:9, 1:4–4:1), and finally with MeOH. Aliquots of these fractions were examined by GLC and TLC. The fractions showing the presence of similar components were combined, giving 8 main fractions (II–1 to II–8).

GLC. This was carried out using; (1) a Beckman GC 4 instrument equipped with a thermal conductivity detector, and (2) a Hewlett Packard GC 5750 instrument equipped with a thermal conductivity detector and a flame ionization detector. Preparative GLC was carried out on aluminum columns (450 cm × 8 mm) packed with 20% Apiezon L on Chromosorb W (60–80 mesh) or anakrom (60–70 mesh) coated with 2% KOH. The average size of analytical column was (210 cm × 8 mm) using a variety of stationary phases such as Apiezon L, Apiezon M, Apiezon W, SE-30, Versamide, Silicon rubber UC-W98 and Carblowax 20M.

Sample collection and identification of gas chromatographic peaks. Each of the main fractions obtained from the column chromatography was examined by GLC at; (1) isothermal temperature, and (2) by temper ature programming. The fractions (II-2, II-3 and II-5A) showing good resolutions were submitted to pre parative GLC and temp. programmed from 160–240 at 4°/min with constant He flow rate of 60 ml/min.

[14] H. ELMENHORST, *Beit. Tabakforsh.* **3**, 101 (1965).

Several injections were made in order to collect sufficient quantities of various components for their identification purpose. The detector temp. was kept at 290° and injection port temp. 280°. The samples were collected in U-shaped glass capillary tubes cooled in a dry ice-acetone bath.

TLC. TLC was done on precoated plate MN-polygram (silica gel N-HR and aluminum oxide N/UV 254) (20 × 20 cm). Preparative TLC of fraction II-5 was carried out on silica gel plates (20 × 20 cm) using $CHCl_3$–benzene (3:7). 4 prominent bands were located under UV light and were scraped from the plate: The band II-5A (R_f 0·04) was eluted from silica gel with $CHCl_3$–MeOH (4:1) and examined by GLC.

Spectral analysis. MS were run on a Hitachi RMU-6E double focussing mass spectrometer with an ionization potential of 70 eV and inlet temp. of 200°. All the samples were inserted through a direct inlet system. The significant ions at higher *m/e* values as well as ions with a relative abundance of 15% or more of the base peak are quoted as *m/e* values. The NMR spectra were recorded on Varian Associates A60 and T60 instruments using $CDCl_3$ as the solvent, and peak positions are given in δ values with tetramethylsilane as an internal standard.

Authentic samples. 2,6-Dimethylquinoline, nicotinamide, anabasine, β-nicotyrine, harmane and 2,2′-bipyridyl were commercially obtained. All the other isomers of bipyridyl were prepared by literature procedures.[15]

Acknowledgements—This study was carried out under Contract No. 12-14-100-10311(73) with the Agricultural Research Service, U.S. Department of Agriculture, administered by the Eastern Marketing and Research Division, 600 E. Mermaid Lane, Philadelphia, Pennsylvania 19118.

[15] E. V. BROWN, unpublished work.