Phytochemistry, 1980, Vol. 19, pp. 1771–1775. Pergamon Press Ltd. Printed in England

MASS-SPECTROMETRIC EVIDENCE FOR QUINONOID–LYSINE COUPLING PRODUCTS IN CIGAR PROTEIN

JOHN EAGLES*, JOHN F. MARCH† and RICHARD L. M. SYNGE*‡

* Agricultural Research Council's Food Research Institute, Colney Lane, Norwich, NR4 7UA, U.K., † John Innes Institute, Colney Lane, Norwich, NR4 7UH, U.K.

(*Received* 7 *December* 1979)

Key Word Index *Nicotiana tabacum*, Solanaceae; tobacco; hydrogenation; hydrogenolysis; lactamization; amino acids; chlorogenic acid; hydantoins; lysine.

Abstract—On subjecting hydrolysates of hydrogenated cigar protein to derivatization and preparative GLC, two fractions were obtained which had volatilities and mass spectra consistent with an origin from oxidative-condensation products of chlorogenic acid with lysine residues. The mass-spectral evidence was chiefly from high-resolution mass spectrometry of the heaviest fragment ions; there was also high-resolution comparison of the lighter fragment ions with those from relevant model compounds. New compounds prepared were the hydantoin from 6-(1-*trans*-octahydroquinol-2-onyl)-DL-norleucine and the heptafluorobutyrylated *n*-propyl esters of four other less-common amino acids. Details of the mass spectra are provided in a Supplementary Publication § or have been sent to the Mass Spectrometry Data Centre.‖

INTRODUCTION

Attempts at stabilization of quinonoid–amine coupling products by hydrogenation have been described in previous papers [1, 2]. In protein from tobacco leaves which had undergone aerobic autolysis, circumstantial evidence for coupling with chlorogenic acid moieties was obtained by recognition of quinic acid among the hydrolysis products of the altered protein [2, 3]. Evidence was also obtained [2] for the appearance of novel components in hydrolysates of hydrogenated tobacco-leaf protein after the leaves had undergone aerobic autolysis or curing for cigar manufacture.

The expected coupling product of chlorogenic acid with the ε-NH_2 group of a lysine residue of a protein would be **1**, although further oligomeric self-condensation products of oxidizing caffeic acid moieties could be expected [4–7]. The expected hydrogenation product from **1** would be **2**, although simultaneous hydrogenolysis could (for a proportion of molecules) remove one or both of the –OH groups on the new cyclohexane ring (or break off that ring from N^{ε} of the lysine residue) [1, 8].

We examined by GC–MS the novel components found after derivatization of hydrolysates of hydrogenated cigar protein (ref. [2]: SUP 90038, Annex 1). Two of the enhanced zones gave low-resolution mass spectra suggesting that they might contain lysine derivatives. In particular, they showed periodicities of fragments ($m/e < 112$) suggestive of oligomethylenediamine chains [9, 10]; furthermore, of the 20 most intense ions ($m/e \geq 112$) in the low-resolution mass spectrum of derivatized lysine (**3**), 14 coincided with strong ions given by both of the zones, a further two being contributed by each zone separately. The present paper deals with these zones, and includes a comparison of their mass spectra with those of some relevant model compounds. No conclusive identifications of the substances have been achieved, but our experiments strongly suggest that they have arisen by coupling of phenylpropanoid moieties to lysine residues. As we are discontinuing the work, it is proper to publish what we have observed.

RESULTS AND DISCUSSION

Preparative GLC [11, 12] was done on a hydrolysate of hydrogenated cigar protein after *n*-propyl esterification followed by heptafluorobutyrylation [2, 12]. Cut 7 (the derivatized-lysine (**3**) zone) was studied as a control. Our main objects of study were Cut 10 (in the 'arginine' position) and Cut 12 (covering the last two 'octahydrotryptophan' positions) (for chromatogram profile see ref. [2]: SUP 90038, Annex 1, Fig. 9). The three Cuts were subjected to low- and high-resolution mass spectrometry. We did not proceed, as previously suggested [2], to hydrolyse and rederivatize the fractions, as relevant elementary formulae could be computed for most of those ions not being mass-deficient fluorocarbons, etc.

The heaviest ions from Cut 10 suggested that their parent compound might have resulted by loss of two –OH groups from **2** by hydrogenolysis or dehydration. The volatility of the derivative also seemed greater than would have been expected for the derivatized hydrolysis

‡ Present address: School of Chemical Sciences, University of East Anglia, Norwich, NR4 7TJ, U.K.

§ SUP 90045 (80 pp.) deposited at the British Lending Library, Boston Spa, Yorkshire, LS23 7BQ, U.K., from whom copies can be obtained.

‖ U.K. Chemical Information Service, The University, Nottingham, NG7 2RD, U.K.

product from **2**. Accordingly, it seemed worth synthesizing **4** for comparison. This was attempted by hydrolysing DL-*trans*-octahydroquinolone (**6**) to DL-*trans*-3-(2′-aminocyclohexyl)propionic acid (**7**) and coupling that with DL-5-(δ-bromobutyl)hydantoin. The coupling reaction was accompanied by lactamization to give **9**. Acid hydrolysis of **9**, under the severe conditions [13] needed to break up the hydantoin ring, gave lysine as the only recognizable amino acid-like product. The recommended conditions for alkaline hydrolysis [13] gave a mixture of lysine and a product which, from its mass spectrum after derivatization, seemed likely to be the ureido acid **10**; this suggested that only partial hydrolysis of the hydantoin ring had been effected.

Thus, although we were frustrated in making direct mass-spectrometric comparisons of the cigar substances with **5**, we were able to make high-resolution comparisons of the electron-impact fragments from Cuts 10 and 12 with those from **3**, **8**, **11** and derivatized DL-N^{ϵ}-cyclohexyllysine (**13**). Derivatized mixed diastereoisomers of octahydrotryptophan (**15**) also had to be studied, since some of them were present in Cut 12; likewise, they shared a number of structural features with the model compounds.

The low-resolution mass spectrum for derivatized arginine differed so much in outline and details from that for Cut 10 that we omitted subjecting derivatized arginine to high-resolution study. Autolytic processes during curing of the leaf may have degraded most of the arginine residues present to ornithine.

Low-resolution mass spectra of the derivatized 'protein' amino acids (with a few others) (cf. ref. [12]) have been sent by J. E. and J. F. M. to the Mass Spectrometry Data Centre, U.K. Chemical Information Service, The University, Nottingham, NG7 2RD, U.K. Although systematic fragmentation schemes have not been attempted in this or the earlier work with this series of derivatives, we have noticed that molecular-ion intensities may be variable and that the most facile losses are of $-COOC_3H_7$ (-87), HOC_3H_7 (-60) and $-COC_3F_7$ (-197 or $(+H)-196$). Individual F atoms (-19) may also occasionally be lost. The mass spectra tend, therefore, to contain little or no information in their high-mass regions.

Low-resolution mass spectra for Cuts 7, 10 and 12 (as well as for **9** and our mixture of **11** with **3**) are given in Table 1 of the Supplementary Publication. Those for **6**, **8**, **13** and **15** have been sent to the Mass Spectrometry Data Centre.

High-resolution mass spectra for Cuts 10 and 12 and compounds **3**, **8**, **11** (with admixed **3**), **13** and **15** are listed in Tables 2–8 of the Supplementary Publication. Those ions ($m/e \geq 112$) found in Cuts 10 and 12 for which relevant atomic formulae could be computed are shown in a special table (Table 9 of the Supplementary Publication), which also shows the occurrence of isomeric fragments (and of isomeric fragments $\pm H$ or H_2) in the five model compounds. **11** yielded 21 such fragments in common with Cuts 10 and/or 12, but not with **3**.

From that table, we have noted the following further points of interest:

Largest fragments. For Cuts 10 and 12, these have been extracted into Table 1. The *m/e* 561 and 562 fragments, from **13** and Cut 10 respectively, and the 618 and 620 fragments, from **11** and Cut 12 respectively have special interest, as perhaps indicating shared substructures.

Ion m/e 280.0572 ($C_9H_9NOF_7^+$) *and other lysine-derived fragments.* This ion (with associated 281.0606 ($^{12}C_8{}^{13}CH_9NOF_7^+$)) constitutes the base peak for derivatized lysine (**3**) and occurs also, strongly, in **13** and, very weakly, in **8**. It is absent from derivatized octahydrotryptophan (**15**). It is prominent in the high-resolution mass spectra for both Cuts 10 and 12. These Cuts, taken together, also yielded more than 20 other fragments in common with **3**; of those, 15 were yielded by both Cuts. As Cuts 10 and 12 emerged from GLC much later than Cut 7 (shown to be derivatized lysine (**3**) by

Table 1. The high-mass ions found in Cuts 10 and 12 to which relevant atomic formulae could be assigned by high-resolution mass spectrometry (extracted from Supplementary Publication, Table 9)

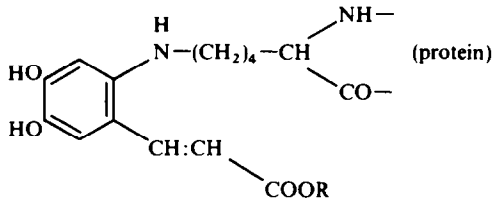
m/e (calc.) for atomic formula	C H N O F	Found deviation from calc. value, in millimass units (mDaltons) Cut 10	Cut 12*
777.0879	21 18 2 5 21†	—	+7.3
627.1470	23 23 1 3 14		+1.5
620.0992	19 18 2 5 14		−1.4‡
562.1301	18 20 2 2 14	−2.2§	—
561.0496	16 11 2 4 14	—	−4.6
560.1270	19 20 1 2 14	...	+0.2
560.0418	16 10 2 4 14		+1.8
543.1317	18 20 2 2 13	−3.7	
542.1239	18 19 2 2 13	−2.1	
536.0981	15 17 2 4 13	+4.5	
483.1882	22 26 2 2 7	...	−4.2
366.1542	14 21 2 1 7	−1.2	...
366.0940	13 15 1 3 7		−4.9
365.1464	14 20 2 1 7	+3.7	
364.1386	14 19 2 1 7	−2.2	

* Excluding ions attributable to derivatized octahydrotryptophan (**15**) (Supplementary Publication, Table 8).

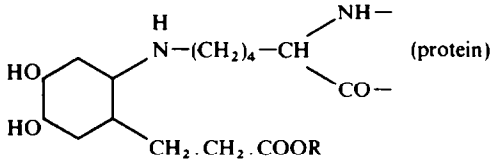
† Formula by inspection from tables, after subtraction of F_{21}.

‡ *m/e* = 618.0835 (calc. for $C_{19}H_{16}N_2O_5F_{14}$) was found (+0.3 millimass units) as a fragment of **11**.

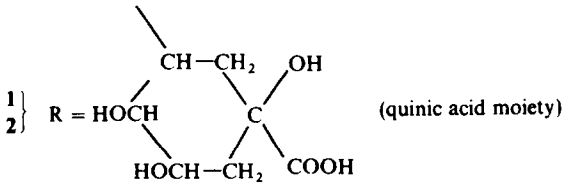
§ *m/e* = 561.1223 (calc. for $C_{18}H_{19}N_2O_2F_{14}$) was found (+4.0 millimass units) as a fragment of **13**.

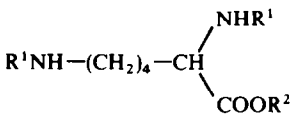


1

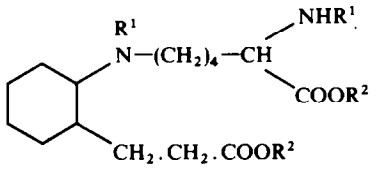


2

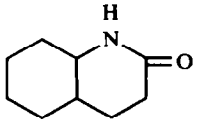




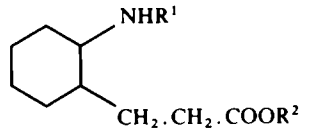
3 $R^1 = C_3F_7CO-$; $R^2 = C_3H_7-$



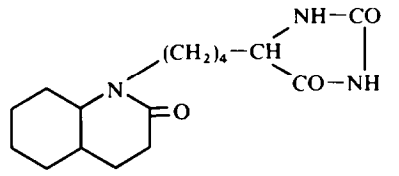
4 $R^1 = R^2 = H-$
5 $R^1 = C_3F_7CO-$; $R^2 = C_3H_7-$



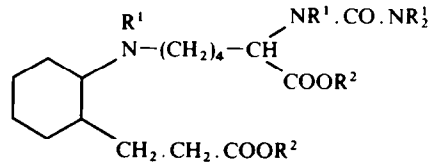
6



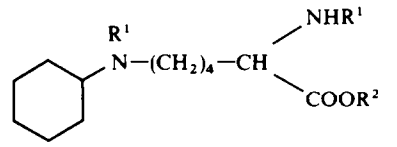
7 $R^1 = R^2 = H-$
8 $R^1 = C_3F_7CO-$; $R^2 = C_3H_7-$



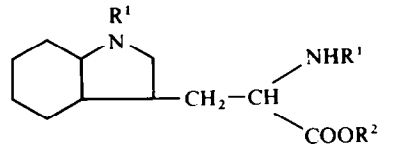
9



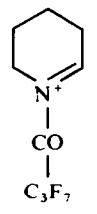
10 $R^1 = R^2 = H-$
11 $R^1 = C_3F_7CO-$; $R^2 = C_3H_7-$



12 $R^1 = R^2 = H-$
13 $R^1 = C_3F_7CO-$; $R^2 = C_3H_7-$



14 $R^1 = R^2 = H-$
15 $R^1 = C_3F_7CO-$; $R_2 = C_3H_7-$



16

chromatographic behaviour and low- and high-resolution mass spectrometry), it is reasonable to conclude that Cuts 10 and 12 contain lysine substitution products. None of the other derivatized 'protein' amino acids gave the $C_9H_9NOF_7^+$ ion, presumably **16**, deriving from the pentamethylenediamine (cadaverine) substructure [9, 10].

C_8 and C_9 fragments. Cuts 10 and 12 gave, taken together, 27 different fragments in the isomer families C_9H_{8-18}, $C_8H_{8-14}N$, $C_9H_{6-16}N$, $C_{12}H_{15}NOF_7$ and $C_{13}H_{11-18}NOF_7$. These fragments were also given by **8** (8 occurrences), **11** (2 occurrences) and **15** (16 occurrences) but showed only one occurrence from **13** and one (doubtful) from **3**. These fragments are consistent with direct attachment of a C_9 moiety (3-cyclohexylpropionic acid) or a C_8 moiety (its decarboxylation product) to N^{ϵ} of a lysine residue.

It is reasonable to conclude that Cut 10 may contain **5**—all its fragments listed in Table 1 are derivable from **5** (M = 776, $C_{29}H_{38}N_2O_6F_{14}$). The higher fragments from Cut 12 (not attributable to the derivatized octahydrotryptophan (**15**) also present in Cut 12) had generally more O atoms, and their parent compound may therefore have retained perhaps one –OH group on the cyclohexane ring. Such a structure, O-heptafluorobutyrylated (M = 988, $C_{33}H_{37}N_2O_8F_{21}$), would explain the somewhat lower volatility of Cut 12 and also the apparent occurrence of an F_{21} (*tris*-heptafluorobutyryl) fragment (*m/e* = 777.0879, Table 1). In general, the volatilities of Cuts 10 and 12 seem consistent with these structures, although a systematic study relating the volatility of such derivatized compounds, including those from the 'protein' amino acids, remains to be made.

Finally, we must stress that the derivatives which we have studied in this work have forced us near to the limits both of GLC volatility and of mass-spectrometric mass range. We recommend that, in any future study along these lines, hydrogenation products should first be separated by liquid chromatography or electrophoresis, and then derivatized for mass-spectrometry with substituents of lower molecular weight [14]. Our present techniques have probably missed the fully hydroxylated condensation products, such as should result from hydrolysis of **2**, as well as any oligomeric products of further oxidative condensation [4–7]. We should also repeat the earlier recommendation [2] that rhodium deposited on charcoal rather than on alumina may prove to be a preferable hydrogenation catalyst. Ruthenium and related metals also merit investigation as catalysts.

EXPERIMENTAL

Materials. The derivatized hydrolysate of hydrogenated cigar protein had been prepared and studied previously [2]. L-Lysine monohydrochloride and other chemicals not specified below were commercial preparations. DL-5-(δ-Bromobutyl)hydantoin, **12** · HBr and **14** (cf. ref. [8]) were available from previous work [1]. **6** and **7** · HCl were prepared according to ref. [15]. To prepare **9**, **7** · HCl (50 mg) was dissolved in MeOH (0.5 ml). The bromobutylhydantoin (70 mg) was added, and also dissolved. K_2CO_3 (50 mg) was then added, followed by *ca* 10 drops of H_2O, which achieved soln. The mixture was stoppered and kept at room temp. for 3 days. Low-resolution MS on the mixture indicated significant formation of **9**, but also lactamization to **6**. Further K_2CO_3 was added and the mixture kept for a further 4 days, though with little change in its MS. About half the MeOH present was then evapd *in vacuo*: the mixture began to crystallize, was saturated with CO_2 and kept at 0–2° overnight. The crystals were filtered off and washed with H_2O at 0–2°, then dried on the filter funnel (75 mg). Most of this material proved to be **6**, which was dissolved away by washing the crystals with MeOH at 0–2°, leaving a residue on the funnel of chunky crystals (5 mg). Low-resolution MS indicated this to be **9** (strong M + 1 = 308—see Table 1 of Supplementary Publication).

An attempt to hydrolyse **9** to **4** by refluxing in a mixture of 10 M HCl (6 vol.) and acetic acid (4 vol.) for 96 hr [13] was unsuccessful—GLC after derivatization of the evapd hydrolysate (see below) indicated only lysine. Another portion of **9** was heated in a sealed vessel in 0.1 M NaOH at 110° for 24 hr [13]. The resulting mixture was acidified with a slight excess of HCl, evapd to dryness *in vacuo* and derivatized. GLC on a portion of the derivatized hydrolysate showed a zone in the 'lysine' position and a later zone, of similar size, close to the 'arginine' position. Direct-insertion low-resolution MS on the derivatized hydrolysate (Table 1 of Supplementary Publication) showed an intense highest-mass ion (*m/e* 1082), a plausible fragment (with loss of 87 + 42) from **11** ($C_{38}H_{37}N_3O_9F_{28}$, M = 1211). The highest-mass ion on which high-resolution MS (Table 6 of Supplementary Publication) could be achieved had *m/e* 835.0947 (calc. for $C_{27}H_{20}N_2O_4F_{21}$, 835.1087). It was surprising that such a high MW compound as **11** should give a GLC zone in the 'arginine' position, being thus substantially more volatile than derivatized homocitrulline [2]. Perhaps, on injection into the GLC column, the hydantoin or lactam rings of **9**, or both, had been re-formed, with thermal elimination from **11** of one or two molecules of *n*-propyl heptafluorobutyrate. (When nopalinic acid was similarly derivatized, it was implied [16] that lactamization had occurred.) The tendency to cyclization in derivatized homocitrulline would be less, since its ureido group is not α- but ε- to its carboxy group.

3, **8**, **13** and **15** were prepared, respectively, from L-lysine monohydrochloride, **7**, **12** and **14** by derivatization as described below.

Derivatization. This was carried out as described previously [11, 12], except that the 2,6-di-*t*-butyl-*p*-cresol (BHT) and acetic anhydride were excluded from the acylation step [2].

Chromatography. Prep. GLC was done on a 5.5 m × 4 mm i.d. glass column as described before [11]. The preparative column was packed with a binary phase composed of 3.33% OV-101 and 0.67% OV-7 on Diatomite MQ support material (90–100 mesh/in. size). All these materials were obtained from J. J.'s Chromatography Ltd., King's Lynn, Norfolk, U.K. The binary phase retained the same ratio of methyl silicones to phenyl silicones as used for the analytical separations [12]. The increased loading for the preparative column was found necessary in order to reproduce the performance of the analytical column. For the separations, an initial column temp. of 110° was used; after 7 min it was increased at 1.5°/min to final temp. 270°. The GLC column described above enabled a vol. of up to 70 μl to be injected, and individual components of as much as 40 μg were chromatographed without causing skewed or distorted peaks. Repeated chromatography of the sample in several portions enabled fractions cut to be built up in a single vessel for subsequent MS examination.

With the exception of **3** and **13**, derivatized synthetic compounds were examined only on the analytical GLC column. The approximate elution temps. for **3**, **8**, **15** and **13** were 182°, 195° (close to the 'arginine' position), 213–218° (the diastereoisomers showed 3 peaks) and 236° (close to the 'cystine' position) respectively. For behaviour of **11**, see above.

It may be significant that these compounds were not all examined by prep. GLC, so any changes due to that treatment will not have been observed.

Low-resolution MS. Samples were introduced by means of the direct-insertion probe into an MS902 mass spectrometer (Kratos

Ltd., Manchester). A resolution of 1000 (10% valley definition), a source temp. of 200° and ionizing energy of 70 eV were used. The spectra were converted to mass and intensity listings using a DS50SM data system (Kratos Ltd., Manchester).

High-resolution MS. This was carried out with the same conditions, except that the resolution was changed to 10 000 (10% valley definition) and perfluorokerosene, as reference, was introduced into the source at the same time as the sample. Atomic compositions were not computed for the mass-deficient, mainly fluorocarbon ions, which included the mass-reference ions as well as some fragments from the samples. However, all ions observed are listed in Tables 2–8 of the Supplementary Publication. When computing atomic compositions, the maximum Nos. of atoms were limited to 40 × C, 3 × N, 8 × O and 14 × F. The error was restricted to *ca* 10 millimass units. When considering atomic compositions, formulae with $7n$ or $(7n - 1)$ × F atoms ($n = 0, 1, 2$ or 3) were taken most seriously. In these, at least n × O atoms had to be present and the C:H ratio had to be realistic.

Acknowledgements—We thank Mrs. Vilas K. Newby for skilled help with the experiments, Mr. Barry J. Gordon for running the MS, Mr. Ron Self for valued advice and criticism and Miss Carol Palfrey for her patience and care in laying out the Supplementary Publication.

REFERENCES

1. Davies, R., Laird, W. M. and Synge, R. L. M. (1975) *Phytochemistry* **14**, 1591.
2. Laird, W. M., March, J. F., Newby, V. K. and Synge, R. L. M. (1979) *Phytochemistry* **18**, 1501.
3. Davies, A. M. C., Newby, V. K. and Synge, R. L. M. (1978) *J. Sci. Food Agric.* **29**, 33.
4. Pierpoint, W. S. (1978) *C. R. Groupe Polyphénols* **8**, 31.
5. Pierpoint, W. S., Ireland, R. J. and Carpenter, J. M. (1977) *Phytochemistry* **16**, 29.
6. Davies, R. snd Synge, R. L. M. (1978) *Agricultural Research Council: Food Research Institute, Report for the Year* 1976, p. 35. Norwich, U.K.
7. Ott, R., Pinter, E. and Kajtna, P. (1979) *Monatsh. Chem.* **110**, 51.
8. Wieland, T. and Gebert, U. (1966) *Liebigs Ann.* **700**, 157.
9. Veith, H. J., Guggisberg, A. and Hesse, M. (1971) *Helv. Chim. Acta* **54**, 653.
10. Schöpp, E. and Hesse, M. (1973) *Helv. Chim. Acta.* **56**, 124.
11. March, J. F. (1979) *J. Chromatogr.* **172**, 366.
12. March, J. F. (1975) *Analyt. Biochem.* **69**, 420.
13. Stark, G. R. and Smyth, D. G. (1963) *J. Biol. Chem.* **238**, 214.
14. Self, R. and Synge, R. L. M. (1977) *Agricultural Research Council: Food Research Institute, Report for the Year* 1975, p. 34. Norwich, U.K.
15. Kost, A. N., Shchegoleva, T. A. and Yudin, L. G. (1955) *Zh. Obshch. Khim.* **25**, 2464.
16. Firmin, J. L. and Fenwick, R. G. (1977) *Phytochemistry* **16**, 761.