Beiträge zur Tabakforschung International · Volume 14 · No. 5 · October 1989

Evaluation of a Tobacco Fractionation Procedure using Pyrolysis Mass Spectrometry combined with Multivariate Analysis*

by

Martin A. Scheijen, Bea Brandt-de Boer and Jaap J. Boon

Unit for Mass Spectrometry of Macromolecular Systems, FOM Institute for Atomic and Molecular Physics, Amsterdam, The Netherlands

and

Werner Hass and Volker Heemann

Research and Development, B·A·T Cigarettenfabriken GmbH, Hamburg, Federal Republic of Germany

SUMMARY

This paper presents an evaluation of a fractionation procedure for use with tobacco. Correlation coefficients calculated from the weights of the polymer fractions obtained and data obtained with classical wet-chemical methods show that these parameters have a low resolving power, which precludes any detailed distinction between tobaccos and tobacco-derived fractions. Pyrolysis mass spectrometry combined with multivariate analysis is presented as a promising approach for investigating the variability in the chemical composition of tobacco. Multivariate analysis performed on the pyrolysis mass-spectrometric fingerprints of all tobacco fractions indicates that the differences between the individual tobaccos are smaller than the differences induced by the fractionation procedure. Multivariate analysis of subfiles of the pyrolysis mass-spectrometric fingerprints of separate polymer fractions is an excellent method for classifying tobaccos. Leaf and stem material can easily be distinguished. A further differentiation is made between the flue-cured and Burley-type tobaccos. The strong clustering of tobacco samples after treatment with potassium hydroxide points to a uniform cell-wall skeleton. The greater variability in the ethanolized tobacco samples, and samples subjected to hot-water extraction or amylase digestion, is related to the cytosol characteristics. These fractions appear to reflect both the dissimilarities between the distinct tobacco types and phenotypic variations due to differences in cultural management. Pyrolysis mass spectrometry of the Klason lignin residues points to a large variety of chemical constituents unrelated to lignin. Pyrolysis gas-chromatography mass-spectrometric data on the Klason lignin residues of a Burley tobacco showed that the correlation made between the organic-nitrogen content and the protein content of this fraction might be misleading. As a consequence the maximum value for the lignin content of this tobacco is estimated incorrectly.

* Received: 21st June 1988 — accepted: 5th December 1988.

ZUSAMMENFASSUNG

Gewichtsdaten, die für die polymeren Fraktionen und andere Inhaltsstoffe des Tabaks unter Einsatz einer klassischen naßchemischen Fraktionierungsmethode gewonnen wurden, erwiesen sich auf der Basis der errechneten Korrelationskoeffizienten als nur bedingt aussagefähig. Sie eignen sich nicht dazu, unterschiedliche Merkmale zwischen Tabaken und Tabakfraktionen eindeutig zu erfassen. Pyrolyse-Massenspektrometrie in Verbindung mit Multivarianzanalyse wird als vielversprechende Methode zur Identifizierung von Unterschieden in der chemischen Zusammensetzung von Tabaken vorgestellt. Aus der Untersuchung der Pyrolyse-Massenspektren (fingerprints) aller Tabakfraktionen ergibt sich, daß die Unterschiede zwischen den einzelnen Tabaktypen geringer sind als die Unterschiede, die sich durch die verwendete Fraktionierungsmethode ergeben. Die an den Pyrolyse-Massenspektren der einzelnen polymeren Fraktionen getrennt durchgeführte Multivarianzanalyse ist eine ausgezeichnete Methode zur Klassifizierung von Tabaken. Blatt- und Rippenmaterial ist leicht unterscheidbar; auch werden Burley- und Virgin-Tabake deutlich voneinander unterschieden. Die nach Behandlung mit Kaliumhydroxid unter den Proben zu verzeichnende starke Clusterbildung deutet auf eine einheitliche Struktur der Zellwand bei allen Tabaken hin. Nach Ethanolextraktion und nachfolgend Heißwasserextraktion bzw. Amylaseaufschluß zeigen die jeweiligen Rückstände eine ausgeprägte Variabilität, die auf charakteristische Unterschiede im Cytoplasma zurückzuführen ist. Die auf diese Weise gewonnenen Rückstände spiegeln sowohl die Unterschiedlichkeit zwischen den Tabaktypen selbst als auch phänotypische, auf verschiedenen Kulturbedingungen beruhende Abweichungen wider. Die Untersuchung der Lignin-Rückstände (Klason) mittels Pyrolyse-Massenspektrometrie läßt viele weitere Signale erkennen, die nicht von Lignin herrühren. Bei einem Burley-Tabak ergab sich aus der Untersuchung der Lignin-Rückstände (Klason) mittels Pyrolyse-Gaschromatographie in Verbindung mit Massenspektrometrie unter anderem, daß der Proteingehalt dieser Fraktion durch Bestimmung des organischen Stickstoffs nicht eindeutig ermittelbar ist. Infolgedessen kann der maximale Ligningehalt dieses Tabaks falsch berechnet werden.

RESUME

Les coefficients de corrélation établis à partir des poids des fractions polymères et d'autres composants du tabac obtenus au moyen d'un procédé classique de fractionnement par voie humide se révèlent peu concluants et ne permettent pas de faire nettement la distinction entre les tabacs et les fractions de tabac. L'utilisation de la pyrolyse / spectrométrie de masse en liaison avec une méthode d'analyse faisant intervenir plusieurs variables (multivariance) se révèle une approche prometteuse pour l'étude des différences de composition chimique des tabacs. L'analyse des spectres de masse (fingerprints) obtenus après pyrolyse de toutes les fractions de tabac montre que les différences existant entre les divers types de tabac sont plus faibles que celles qui sont induites par la méthode de fractionnement. L'analyse (multivariance) réalisée séparément sur les spectres de masse des différentes fractions polymères est une excellente méthode pour la classification des tabacs. Les côtes et les feuilles se distinguent facilement; le tabac Burley et le tabac «flue-cured» sont aussi nettement différenciés. La formation marquée de grappes (cluster) parmi les échantillons après traitement par l'hydroxyde de potassium indique que la structure de la paroi cellulaire semble être la même chez tous les tabacs. Après extraction par l'éthanol suivie d'une extraction par l'eau chaude ou d'un traitement par l'amylase, on obtient dans chaque cas des résidus présentant une variabilité marquée qui est imputable à des différences caractéristiques au niveau du cytoplasme. Les résidus ainsi obtenus reflètent tant la différence entre les types de tabac que des variations d'ordre phénotypique liées aux différentes conditions de culture. L'étude des résidus de lignine (Klason) à l'aide de la pyrolyse / spectrométrie de masse permet de détecter de nombreux autres signaux qui ne proviennent pas de la lignine. Dans le cas d'un tabac Burley, l'étude des résidus de lignine (Klason) par pyrolyse / chromatographie en phase gazeuse en liaison avec la spectrométrie de masse a montré entre autre que la teneur en protéines de cette fraction ne pouvait être déterminée sans équivoque à partir du dosage de l'azote organique. Il en résulte que la valeur calculée pour la teneur maximale en lignine de ce tabac peut être fausse.

Abbreviations used

Py-MS:	Curie-point pyrolysis mass spectrometry,
Py-GC:	Curie-point pyrolysis gas chromatography,
Py-GC-MS:	Curie-point pyrolysis gas chromatography combined with mass spectrometry,
ER:	Residue after extraction with 80% ethanol,
AR:	Residue after amylase treatment of ER,
OAR:	Blank run on AR, omitting the enzyme,
KR:	Residue after 0.1 N KOH extraction,
KLR:	Klason lignin residue.

INTRODUCTION

At the molecular level the chemical composition of tobacco, a complex biomatrix which is widely used for the industrial production of smoking materials, is still largely unknown despite major research efforts in low molecular-weight volatiles and in the chemistry of tobacco smoke (13, 14, 31, 34). Modern analytical techniques, such as gas chromatography and gas chromatography combined with mass spectrometry, have become important tools in investigating the chemical constituents of tobacco and tobacco smoke (12, 27, 31, 35). HALKET and SCHULTEN (17) used field-ionization mass spectrometry to profile three different commercial tobacco blends, and were able to carry out a rapid screening of the tobacco samples. Mass-spectrometric techniques appear to be a promising approach for investigating the chemical composition of tobacco and tobacco smoke.

Recently, fractionation procedures adapted from plant science have been applied to the characterization of various tobacco types (2, 3) in order to improve understanding of its molecular architecture. This new interest in the macromolecular system of tobacco is also fostered by the new techniques in plant biotechnology which will eventually allow the chemical composition of plants to be steered.

The fractionation procedure developed by BOKELMAN et al. (3) is based on sequential extractions with solvents, acids and bases and is assumed to result in certain well-defined cell-wall polymer fractions. Important assumptions in any fractionation approach to the analysis of complex biomaterial are the repeatability of the fractionation process and the comparability of the fractionation data for different tobaccos and tobacco-derived products. Therefore the fractions obtained must be monitored using analytical techniques. So far, a number of analytical results such as weight of the fractions, total-sugar and reducing-sugar content, uronic-acid content, nitrate, chloride, and nicotine concentrations have been obtained to characterize these BOKELMAN fractions. These data are mainly of empirical use and do not immediately pertain to the molecular details of the composition of the macromolecular systems which are present in tobacco-leaf and stem material. Hydrolysis of particular fractions will certainly yield information on sugars and amino acids, but the lignin composition cannot be studied in this way nor is any insight provided into conjugate polymers. What is evidently required is an analytical method which is non-discriminatory and which provides an informative profile of all the polymer systems within a plant-material sample.

For this purpose we have explored the potential of Curie-point pyrolysis mass spectrometry [Py-MS] and Curie-point pyrolysis gas chromatography combined with mass spectrometry [Py-GC-MS] for characterizing flue-cured, Burley and Oriental tobaccos. Both analytical pyrolysis methods visualize organic molecules obtained by thermal dissociation from the polymer framework, molecules which often retain the characteristics of the monomeric units of the original architecture (9, 11). Py-GC-MS is used for the chemical identification of these products, whereas Py-MS, which basically simply omits the GC separation, yields highly condensed fingerprinting information on the whole pyrolyzate in one mass spectrum. Owing to the speed of analysis (20—30 analyses per hour) and the high degree of automation in the FOMautopyms instrument used, large datasets, including replicate analyses, are built up in a short time (4). These datasets are analyzed further in multivariate analysis using a factor-discriminant procedure especially developed for Py-MS (19), a procedure which classifies the spectra and describes the similarity and dissimilarity of the samples in a quantitative way.

This paper describes the comparison of the Py-MS fingerprints of a wide selection of tobaccos and polymer fractions derived from them in order to assess the chemical nature of the fractions and their variability. Residues from consecutive extraction with 80% ethanol [ER], amylase treatment [AR] and blank procedure [OAR], base extraction [KR] and residues from the Klason lignin preparation method [KLR] obtained according to BOKELMAN et al. (3), were subjected to Py-MS analysis with the FOM-autopyms. Multivariate analysis of the Py-MS data was performed on a "total" file with every fraction of every tobacco and on subfiles with one type of tobacco fraction. Relevant Py-GC-MS data are presented to aid in the chemical interpretation of the Py-MS data.

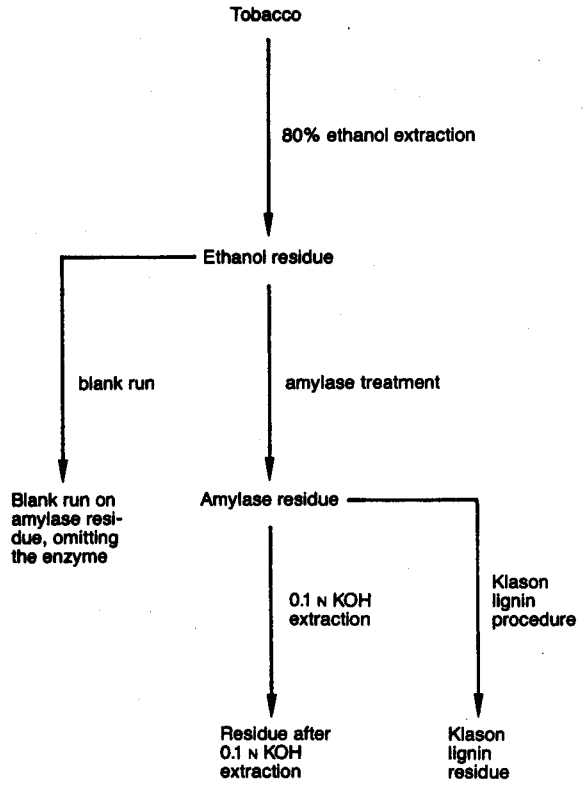
EXPERIMENTAL

Fractionation Procedure

Cured, aged, uncased tobaccos from different growing regions were used in this study. Lamina of eleven flue-cured and three Burley tobaccos from seven different countries, leaves of two Oriental tobaccos, and two flue-cured stems were investigated (Table 1). All the tobacco samples were ground to pass a 20 mesh screen and were fractionated according to BOKELMAN et al. (3), as shown in Figure 1, with only small modifications. The tobacco samples were extracted in a Soxhlet apparatus with 80% aqueous ethanol for 48 h. The residue was treated with amylase of *Bacillus ssp.* (Sigma Chemical Company, type II-A, E.C. 3.2.1.1) at 80 °C. A blank run under the same conditions without the enzyme was also performed. The amylase residue was subjected to alkaline extraction with 0.1 N KOH containing 0.1% (w/v) $NaBH_4$ and 0.03% (v/v) *n*-octanol. The Klason lignin residue was obtained from the amylase residue after treatment with 72% sulfuric acid.

Nicotine, total and reducing sugars, and nitrate were determined with a Technicon AutoAnalyzer II, using colometric methods as described by FINSTER et al. (15). Starch was determined from the ethanol residue using the glucose-oxidase method developed by OAKLEY (24). The amount of pectin was calculated from the uronic-acid content of the ethanol residue which itself was determined using a carbazol-based colometric method

Figure 1.
Tobacco fractionation procedure as designed by BOKELMAN et al. (3).



first described by MCCOMB (22). The content of organic nitrogen was determined from the amylase residue with a CHN-analyzer. All analyses were performed in duplicate.

Pyrolysis Mass Spectrometry

The automated pyrolysis mass spectrometer (FOMautopyms) used has been described in detail by BOON et al. (8). In short, it consists of a sample-changing device, a Curie-point pyrolysis reactor (FOM-3LX) with a high-frequency generator (Fischer Labor- und Verfahrenstechnik, 1.1 MHz, 1.5 kW), a quadrupole mass analyzer (Balzers QMA 150/QMG 511) combined with an ion-counting detection device and an on-line minicomputer which is connected to a central computer (VAX 150/785). Pyrolysis is accompanied by the inductive heating of a ferromagnetic wire coated with sample material and inserted into a glass tube. The volatiles enter the ion source through a heated expansion chamber, are ionized by low-energy electron impact ionization, and mass analyzed. Repetitive scanning and signal averaging are applied.
In a typical experiment plant material was suspended in water. An aliquot (approx. 20 µg) was transferred to a ferromagnetic wire, and dried *in vacuo*. Analysis conditions were as follows: batch size for sample exchanger, 36; Curie-point temperature, 610 °C; temperature-rise time, 0.1 s (5000 K/s); total pyrolysis time, 0.9 s; temperature of the pyrolysis chamber elevated (0.5 A setting); equilibration time in the pyrolysis chamber, 20 s; expansion-chamber temperature, 200 °C; ionization at 15 eV; mass range, 20—220 a.m.u.; scan speed, 10 scans / s; total number of averaged spectra, 200. All samples were analyzed in triplicate.

Multivariate Analysis

The pyrolysis mass spectra, expressed in ion counts / mass channel, are normalized by expressing the mass intensities as a percentage of the total ion counts. To interpret the overall spectral differences within the datasets multivariate analysis of the spectra is performed, using a modified ARTHUR package (Infometrix). The modifications and expansion of this package with linear discriminant analysis have been described by HOOGERBRUGGE et al. (19). After defining the file (training set), an overall spectrum is calculated (zero point), which serves as the reference point for comparing the individual spectra. This spectrum is also used for scaling the datasets. Positive and negative differences with respect to the average spectrum are calculated and expressed in factor scores. Covariant mass peaks are linearly combined to new, independent variables (discriminant functions). The dissimilarity between the categories (groups of multiplicate spectra) is qualitatively expressed in these discriminant functions, which are represented graphically by reconstructed mass spectra. Dissimilarity is quantitatively expressed in discriminant-function scores, which can be plotted as score curves or maps (two dimensions). The Euclidean distances between the samples indicate differences in the distribution of mass peaks in the Py-MS spectra, which in turn point to differences in the chemical composition of the samples. Discriminant functions describe in general a combination of chemical substances expressed in mass peaks, which are used for optimal differentiation. The D scores can be considered as the relative concentration of these substances.
The mass-spectral differences between individual samples or clusters of samples in the dataset can also be visualized by graphical rotation of the calculated discriminant functions (37). This mathematical procedure rotates a discriminant function by a desired angle in a two-dimensional plane parallel to the relative positions of the samples. The reconstructed mass spectra of the rotated discriminant function contain the qualitative and quantitative spectral differences between the Py-MS fingerprints of the samples or sample clusters.

Pyrolysis Gas Chromatography combined with Mass Spectrometry

Pyrolysis gas chromatography combined with mass spectrometry [Py-GC-MS] was performed with a Pack-

Table 1.

Weight data on polymer fractions and on selected chemical compounds of the tobacco samples analyzed.

		Residue after 80% ethanol extraction (v/v) [ER] (%)	Residue after amylase treatment of ER [AR] (%)	Pectin content (%)	Starch content (%)	Organic–nitrogen content (%)	Nicotine content (%)	Nitrate content (%)	Total–sugars content (%)	Reducing–sugars content (%)
Flue-cured tobacco:										
1	U.S.A. 1	50.8	31.6	6.8	2.8	1.024	3.3	0.05	22.0	16.4
2	U.S.A. 2	51.7	36.4	7.3	2.7	0.832	2.8	0.20	19.1	13.4
3	U.S.A. 3	48.0	32.4	7.4	2.5	0.912	2.9	0.17	18.2	14.5
4	Brazil 1	51.0	36.8	2.2	1.2	0.944	2.3	0.05	24.5	20.2
5	Brazil 2	50.2	34.4	1.6	2.3	0.880	3.7	0.10	21.0	15.5
6	Canada 1	40.8	25.3	1.6	2.6	0.720	2.1	0.03	30.0	21.4
7	Taiwan 1	51.7	30.8	1.2	1.2	0.944	3.3	0.10	20.4	14.6
8	Taiwan 2	58.5	32.0	1.5	1.8	1.066	3.3	0.20	17.9	14.2
9	China 1	51.6	28.4	9.4	3.1	0.784	1.1	0.02	27.6	18.2
10	Malaysia 1	75.7	48.3	5.8	0.1	1.536	1.3	1.20	7.5	4.6
11	Malaysia 2	71.6	50.1	3.8	0.1	1.840	0.9	0.80	4.0	1.3
Burley tobacco:										
12	South Korea 1	71.5	47.4	3.9	0.1	1.152	1.7	2.10	3.2	1.2
13	U.S.A. 4	67.0	46.8	4.2	0.1	1.104	3.6	4.10	3.8	1.6
14	Brazil 3	76.0	50.4	4.2	0.1	1.360	3.0	1.70	3.0	1.4
Oriental tobacco:										
15	Turkey 1	54.7	35.0	1.7	2.0	0.944	0.7	0.10	20.8	15.1
16	Turkey 2	55.5	35.8	1.0	1.8	0.992	0.8	0.20	18.6	13.6
Stem material:										
17	U.S.A. 3S	67.7	48.3	4.7	0.8	0.448	0.9	1.09	13.6	10.5
18	China 1S	55.7	42.6	2.8	0.4	0.432	0.3	0.17	18.7	17.2

Table 2.

Correlation study on the weights of the polymer fractions and of selected chemical compounds, compiled in Table 1.

The numbers on the horizontal axis correspond to the digital codes on the vertical axis.

(x: correlation coefficient ≧ 0.98, o: correlation coefficient ≧ 0.99.)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Flue-cured tobacco:																		
1	U.S.A. 1																	
2	U.S.A. 2	o																
3	U.S.A. 3	o	o															
4	Brazil 1	o	×	×														
5	Brazil 2	o	×	o	o													
6	Canada 1																	
7	Taiwan 1	o	×	o	o	o												
8	Taiwan 2	×	×	×		o		o										
9	China 1	×																
10	Malaysia 1									o								
11	Malaysia 2										o							
Burley tobacco:																		
12	South Korea 1										o	o						
13	U.S.A. 4										o	o	o					
14	Brazil 3										o	o	o	o				
Oriental tabacco:																		
15	Turkey 1	o	o	o	o	o	o		o	o								
16	Turkey 2	o	o	o	×	o		o	o							o		
Stem material:																		
17	U.S.A. 3S		×	×							o	×	×	×	×		×	
18	China 1S	×	o	o	×	o		×	×						o	o		×

ard 438-S gas chromatograph coupled with a JEOL DX-303/DA-5000 double-focusing mass spectrometer. The pyrolysis unit (FOM 3-LX) was described earlier in this paper. The pyrolysis conditions were as follows: Curie-point temperature, 510 °C; temperature-rise time, 0.1 s (5000 K/s); total pyrolysis time, 4.0 s; temperature of the pyrolysis chamber, 180 °C. The unit is flushed with helium, which is also used as carrier gas. The pyrolyzate was directly introduced into a capillary column. The column used was a 50 m CP Sil 5 CB fused-silica capillary column (inside diameter, 0.32 mm; film thickness, 1 µm; Chrompack). The GC oven was kept at 35 °C during pyrolysis and subsequently heated to 300 °C at a rate of 4 °/min. Compounds were ionized at 15 eV electron impact energy. The acceleration energy was 3 kV. The scan speed was 0.4 s / decade (mass range: 20—500 a.m.u.). The software package in the JEOL data system was used without modifications.

RESULTS

Bokelman *Fractionation Data on the Tobacco Samples*

Selected values for the wet-chemical analyses which were performed on the different tobacco samples are listed in Table 1. In general the data reflect the well-known general trends for the different tobacco types. The weights of the extraction residues are of a similar order of magnitude as published by Bokelman et al. (3).

In addition the data reflect a great variability even within the different tobacco types, which makes a classification of the tobaccos difficult. Therefore the correlation coefficients for these wet-chemical data were calculated to evaluate the usefulness of these variables for classifying different tobaccos and tobacco types. The results of the correlation study are presented in Table 2. The numbers on the horizontal axis correspond to the digital codes on the vertical axis. Only correlation coefficients higher than 0.98 are given here. This indicates that the spread in the weight data of the polymer fractions and wet-chemical method is limited. It is clear that the flue-cured tobaccos are heterogeneous. The Canadian grade, Canada 1 does not correlate with any other sample, pointing to substantial differences between Canada 1 and all other tobaccos. The Malaysia grades 1 and 2 do not correlate with the other flue-cured tobaccos but do correlate strongly with the Burley tobaccos. The correlation coefficients for the Burley tobaccos are very high (≥ 0.99) and these do not correlate with any other tobaccos except for the Malaysian tobaccos, again indicating their specific composition. For the Malaysian tobaccos this can only be explained by a cultural-management system which differs from the normal procedure used for flue-cured tobaccos. A more detailed differentiation within the flue-cured tobaccos, or between the Oriental and flue-cured tobaccos cannot be made. The stem material shows a divergent result. Sample U.S.A. 3S correlates with only a few of the flue-cured tobaccos, whereas sample China 1S shows a very high correlation with most flue-cured and Oriental tobaccos.

In general it can be stated that the parameters used in this study may be suitable for distinguishing between different tobacco types (flue-cured and Burley), but do not reflect possible characteristic features of the individual tobaccos with sufficient accuracy.

Pyrolysis Mass Spectrometry of Tobacco

We found in our laboratory that Curie-point pyrolysis data from Py-MS, Py-GC, and Py-GC-MS systems match quite well. In pyrolysis mass spectrometry, which is used as profiling mode, the volatiles enter an expansion chamber after pyrolysis. Condensation of volatiles inside the pyrolysis unit is prevented by a heating device located between the high-frequency coil and the glass sample tube. The volatiles enter the ion source through a hole in the expansion chamber. The temperature of the expansion chamber is kept constant at 200 °C, which is an important modification compared to earlier models (22). The fragments produced in the ion source are mass analyzed. This precludes time-resolved data. In Py-GC and Py-GC-MS the pyrolyzate is directly introduced into a high-resolution GC column. In the FOMautopyms a pyrolysis chamber with controlled temperature has been introduced (Fig. 2), which reduces the condensation of less-volatile parts of the pyrolyzate and results in excellent comparability of the Py-GC-MS and Py-MS data (5). Py-MS data are obtained by low-energy ionization to increase the number of molecular ions and suppress fragment ions in the spectra.

Figure 3 shows typical pyrolysis low-energy electron-impact mass spectra of the ER, AR, OAR, KR, and KLR fractions of a typical flue-cured tobacco, U.S.A. 3, and a typical Burley tobacco, U.S.A. 4. Comparison of the fingerprints immediately reveals significant differences in the composition of the residues, visualized by the different distributions of mass peaks. For interpretation of these pyrolysis mass spectra in terms of chem-

Figure 2.
FOM 3-LX pyrolysis unit as designed by the FOM Institute in Amsterdam, The Netherlands (8).

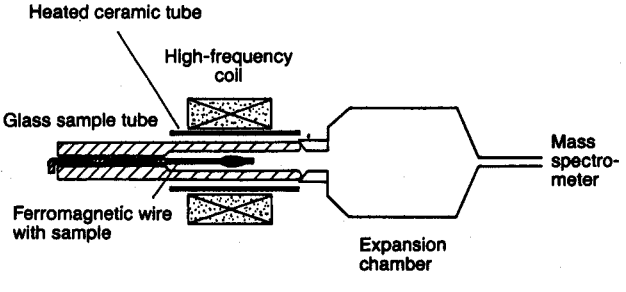
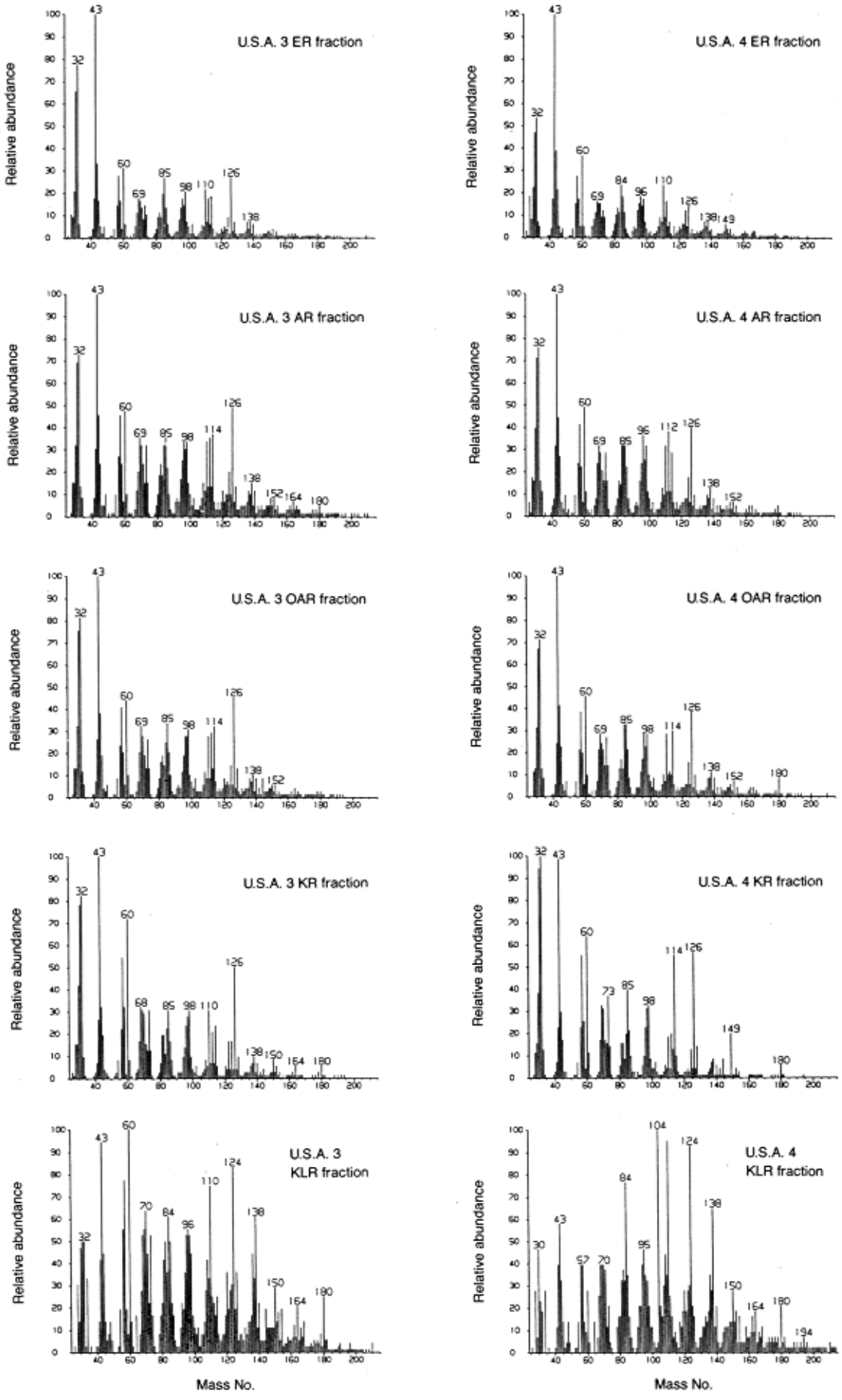


Figure 3.
15 eV electron impact Py-MS fingerprints of the ER, AR, OAR, KR and KLR fractions of a flue-cured (U.S.A. 3) and a Burley (U.S.A. 4) tobacco.



Py-MS: Curie-point pyrolysis mass spectrometry,
ER: ethanol residue,
AR: amylase residue,
OAR: blank run on AR, omitting the enzyme,
KR: residue after 0.1 N KOH extraction,
KLR: Klason lignin residue.

ical substances we refer to Table 3, which compiles characteristic mass peaks selected from Py-MS spectra of reference compounds and biopolymers (10). Owing to the complexity of the pyrolyzate only a few specific mass peaks are present. Most polymer systems generate several chemical compounds upon pyrolysis and this results in certain distributions in mass-peak intensities. These distribution patterns are also used for a tentative chemical interpretation. Mass spectra are compared with Py-MS data on purified polymers (21) and plant material (6, 7, 8, 10, 11, 16, 20, 21, 25, 26).

The Py-MS fingerprints of the tobacco sample after treatment with 72% sulfuric acid (KLR fraction) show markers for lignin (*m/z* 124, 138, 150, 152, 164, 180, 194). In the Py-MS fingerprint of the KLR fraction of U.S.A. 3, additional markers for polysaccharides (*m/z* 57, 60, and 73) and nicotine (*m/z* 84) are found. The latter compound is also present in the KLR fraction of U.S.A. 4 together with dihydroxybenzene (*m/z* 110). The KR fractions reveal mass peaks characteristic of pentoses (*m/z* 114), hexoses (*m/z* 126, 144) and anhydro sugars (*m/z* 57, 60, 73). The mass peaks 85, 32 and 31 point to residual pectins (1). Most of these mass peaks are also found in the Py-MS fingerprints of the KR fractions. The chemical significance of the Py-MS fingerprints of the ER and AR fractions are less obvious because the chemical information obtained is too dense to be properly interpreted.

Py-MS data reflect the whole pyrolyzate in a very condensed form. (Gas-)chromatographic separation is required to obtain the chemical significance of the mass peaks. With pyrolysis gas chromatography combined with mass spectrometry [Py-GC-MS] the pyrolyzate is separated in a high-resolution capillary column prior to mass-spectrometric analysis of the purified compounds. Only preliminary data are given here. The partial ion-current trace (elution-temperature range from 115 to 275 °C) of Py-GC-MS data of the ER, KR, and KLR fractions of U.S.A. 4, a Burley tobacco, is shown in Figure 4. The chemical compounds identified in these pyrolyzates are shown in Table 4. The main peaks in the Py-GC-MS data of the KLR fraction (Fig. 4C) are monomethoxyphenol (guaiacyl) and dimethoxyphenol (syringyl) derivatives (peak Nos. 14, 23, 26, 27, 32, 34, 39, 41, 42, 44, 51, 57 and 58). Identification of these compounds is supported by comparison with high-resolution GC - MS data on beech wood (8, 16).

The most abundant peak in the Py-GC-MS trace of the KLR fraction of U.S.A. 4 (Fig. 4C) represents nicotine, which appears in the Py-MS fingerprints as mass peaks 162 (molecular ion) and 84 (fragment peak). Several other alkaloid markers are observed: 3-ethylpyridine (peak No. 2), 3-vinylpyridine (peak No. 3), myosmine (peak No. 40), and 5-phenyl-3-methylpyrazole (peak No. 44). These compounds appear in the Py-MS data as *m/z* 146 (molecular ion), 145 and 118 (fragment peaks) for myosmine and as *m/z* 158 (molecular ion) and 157 (fragment peak) for the pyrazole derivative (31).

The KR fraction is clearly of a polysaccharide nature.

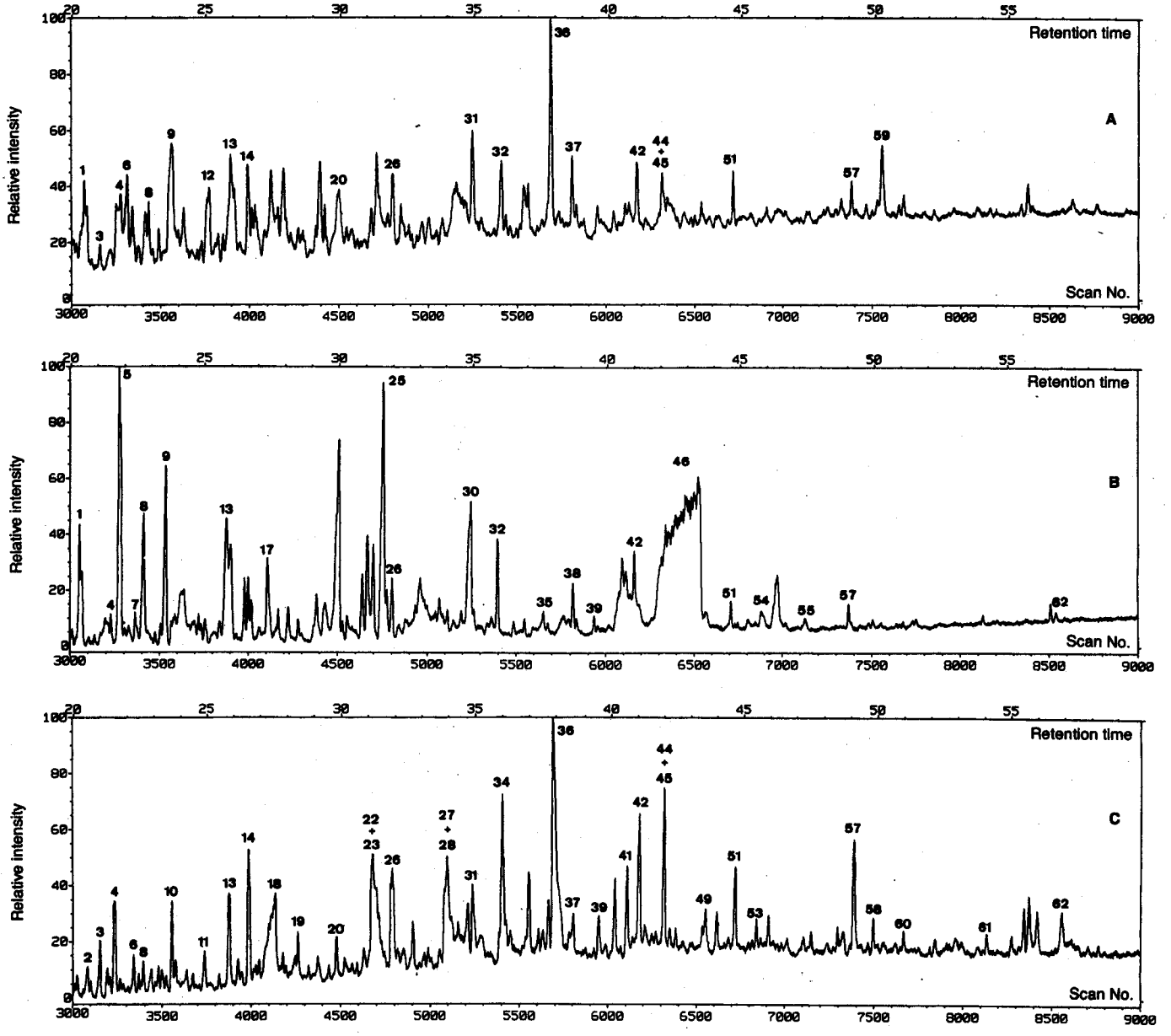
Table 3.
List of low-energy pyrolysis mass peaks found for certain chemical constituents in plant material, adapted from Boon et al. (9).

Standard compound	Mass peaks
Neutral polysaccharides	31, 32, 43, 55, 58, 60, 72, 74, 82
Cellulose, amylose, glycogen	57, 60, 73, 85, 86, 96, 98, 100, 102, 110, 112, 126, 144
Pentoses	114, 85, 58
Hexoses	126, 144
Deoxyhexoses	128
Nicotine	84
Protein	17, 34, 41, 48, 55, 67, 68, 69, 81, 83, 92, 94, 108, 117, 131, 154
Lignin	94, 108, 120, 122, 124, 138, 150, 152, 154, 164, 168, 178, 180, 182, 194, 196, 208, 210, 212
Phenolic acid esters	120, 136, 150, 180
Phospholipids	58, 59, 71, 89

The main peaks in the partial ion current (Fig. 4B) represent 5-methyl-2-furaldehyde (peak No. 1), 4-hydroxy-5,6-dihydro-2*H*-pyran-2-one (peak No. 5), 2-hydroxy-3-methyl-2-cyclopentenone (peak No. 8), 3-hydroxy-2-methyl-2-cyclopentenone (peak No. 9) 5-hydroxymethyl-2-furaldehyde (peak No. 25), 2-hydroxymethyl-5-hydroxy-2,3-dihydro-4*H*-pyran-4-one (peak No. 30), and 1,6-anhydro-β-D-glucopyranose (peak No. 46). These compounds appear in the Py-MS fingerprints as *m/z* 114 (molecular ion) for 4-hydroxy-5,6-dihydro-2*H*-pyran-2-one, *m/z* 112 (molecular ion) for both cyclopentenone derivatives, *m/z* 126 (molecular ion) and *m/z* 97 (fragment peak) for 5-hydroxymethyl-2-furaldehyde, *m/z* 144 (molecular ion) for 2-hydroxymethyl-5-hydroxy-2,3-dihydro-4*H*-pyran-4-one, and *m/z* 57, 60, 73 (fragment peaks) for 1,6-anhydro-β-D-glucopyranose. Among these saccharide markers several markers for lignin are observed in the Py-GC-MS data (peak Nos. 26, 32, 35, 39, 42, 51, 54, 55 and 57). Pyrolysis gas chromatography [Py-GC] with simultaneous NPD and FID revealed only a few nitrogen-containing compounds of very low abundance in this fraction.

The most abundant peaks in the partial ion current of the ER fraction (Fig. 4A) represent 3-hydroxy-2-methyl-2-cyclopentenone (peak No. 9), 4-methylphenol (peak No. 13), 2-methoxyphenol (peak No. 14), indole (peak No. 31), 4-vinyl-2-methoxyphenol (peak No. 32), nicotine (peak No. 36), methylindole (peak No. 37), 5-phenyl-3-methylpyrazole (peak No. 44), 4-acetyl-2-methoxyphenol (peak No. 45), 4-vinyl-2,6-dimethoxy-

Figure 4.
Partial ion current (elution temperature from 115 to 275 °C) **15 eV electron impact Py-GC-MS analysis of the ER** (see A), **KR** (see B) **and KLR** (see C) **fractions of tobacco U.S.A. 4.**
The peak Nos. refer to the corresponding numbers in Table 4.



Py-GC-MS: Curie-point pyrolysis gas chromatography combined with mass spectrometry,
ER: ethanol residue,
KR: residue after 0.1 N KOH extraction,
KLR: Klason lignin residue.

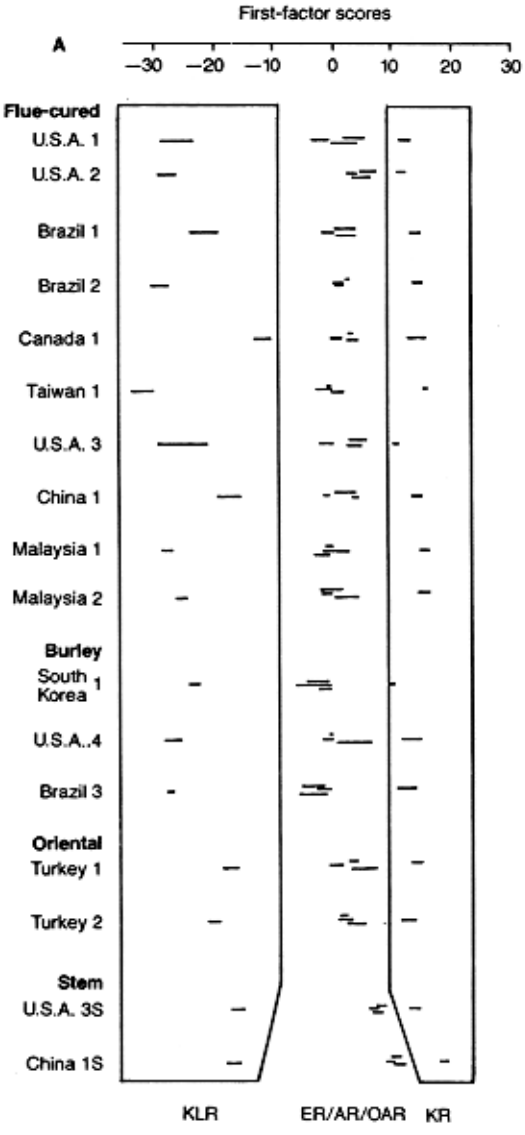
Table 4.

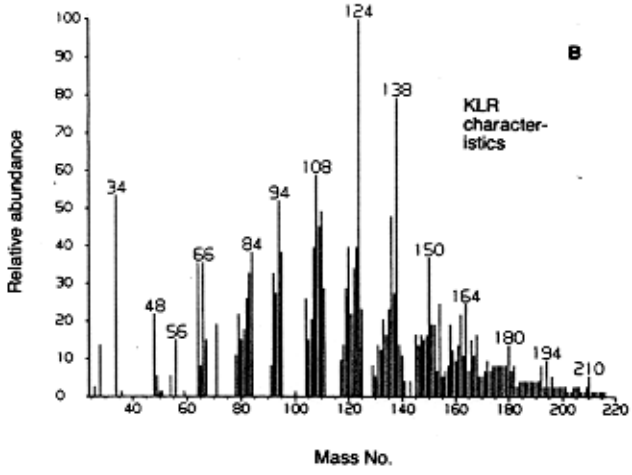
Compilation of the chemical compounds found in the ER, KR and KLR fractions of tobacco U.S.A. 4. The numbers in the first column correspond to the peak numbers in Figure 4.

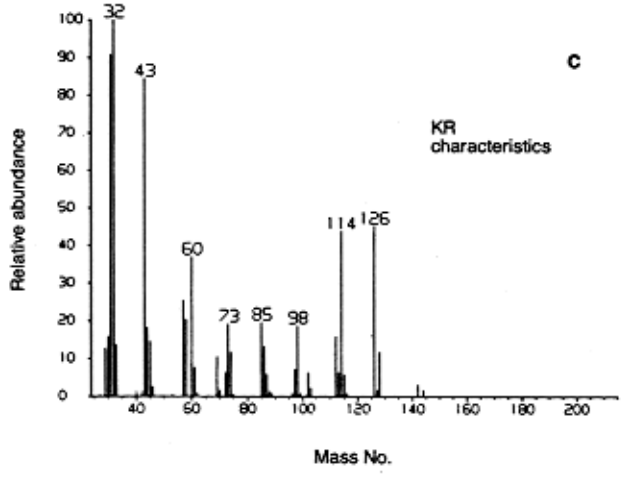
Peak	$M^{+\bullet}$	Scan	Elution temperature (deg. C)	Compound name	Residue after 80% ethanol extraction [ER]	Residue after 0.1 N KOH extraction [KR]	Klason lignin residue [KLR]
1	110	3053	116	5-methyl-2-furaldehyde	×	×	
2	107	3099	117	3-ethylpyridine			×
3	105	3150	119	3-vinylpyridine	×		×
4	94	3230	121	phenol	×	×	×
5	114	3274	122	4-hydroxy-5,6-dihydro-2*H*-pyran-2-one		×	
6	109	3338	124	3-methoxypyridine	×		×
7	112	3365	125	furfurylmethylester		×	
8	112	3427	126	2-hydroxy-3-methylcyclopent-2-en-1-one	×	×	×
9	112	3535	129	3-hydroxy-2-methylcyclopent-2-en-1-one	×	×	
10	107	3555	129	2-ethylpyridine			×
11	85	3738	135	tetrahydropyrrol-2-one			×
12	108	3772	135	methylphenol	×		
13	108	3873	138	4-methylphenol	×	×	×
14	124	3986	141	2-methoxyphenol	×	×	×
15	126	3999	141	levoglucosenone		×	
16	121	4026	142	3-acetylpyridine			×
17	126	4018	144	3-hydroxy-2-methyl-4*H*-pyran-4-one		×	
18	95	4134	145	3-pyridinone			×
19	121	4260	148	ethyl-methylpyridine			×
20	122	4501	155	4-ethylphenol	×		×
21	142	4639	158	3,5-dihydroxy-2-methylpyran-4-one		×	
22	110	4677	159	1,2-dihydroxybenzene		×	×
23	138	4681	159	4-methyl-2-methoxyphenol	×	×	×
24	144	4698	160	1,4:3,6-dianhydro-α-D-glucopyranose		×	
25	126	4753	161	5-hydroxymethyl-2-furaldehyde		×	
26	122	4788	162	4-vinylphenol	×	×	×
27	140	5086	170	4-hydroxy-2-methoxyphenol			×
28	110	5086	170	1,4-dihydroxybenzene			×
29	152	5212	174	4-ethyl-2-methoxyphenol			×
30	144	5246	175	2-hydroxymethyl-5-hydroxy-2,3-dihydro-4*H*-pyran-4-one		×	
31	117	5249	175	indole	×		×
32	150	5405	179	4-vinyl-2-methoxyphenol	×	×	×
33	124	5485	181	1,4-dihydroxy-2-methylbenzene		×	
34	154	5551	183	2,6-dimethoxyphenol	×	×	×
35	164	5652	185	4-(prop-1-enyl)-2-methoxyphenol		×	×
36	162	5687	186	nicotine	×		×
37	131	5803	189	methylindole	×		×
38	152	5815	190	3-methoxy-2-hydroxybenzaldehyde		×	
39	164	5946	193	4-(prop-*cis*-2-enyl)-2-methoxyphenol	×		×
40	146	6038	195	myosmine	×		×
41	168	6104	197	4-methyl-2,6-dimethoxyphenol			×
42	164	6171	199	4-(prop-*trans*-2-enyl)-2-methoxyphenol	×	×	×
43	155	6270	202	4-phenylpyridine			×
44	158	6314	203	5-phenyl-3-methylpyrazole	×		×
45	166	6316	203	4-acetyl-2-methoxyphenol	×		×
46	162	6460	209	1,6-anhydro-β-D-glucopyranose		×	
47	182	6531	209	4-ethyl-2,6-dimethoxyphenol			×
48	180	6531	209	4-(propan-2-one)-2-methoxyphenol			×

cont'd. on p. 273

Figure 5.
A: First-factor scores resulting from factor analysis of the Py-MS fingerprints of the ER, AR, OAR, KR and KLR fractions of the tobacco samples listed in Table 1. Figures B and C represent the reconstructed mass spectra of the first principal component, revealing characteristics for the KLR (see B) and KR (see C) fractions.







Py-MS: Curie-point pyrolysis mass spectrometry,
ER: ethanol residue,
AR: amylase residue,
OAR: blank run on AR, omitting the enzyme,
KR: residue after 0.1 N KOH extraction,
KLR: Klason lignin residue.

Table 4 (cont'd.).
Compilation of the chemical compounds found in the ER, KR and KLR fractions of tobacco U.S.A. 4.

Peak	$M^{+\bullet}$	Scan	Elution temperature (deg. C)	Compound name	Residue after 80% ethanol extraction [ER]	Residue after 0.1 N KOH extraction [KR]	Klason lignin residue [KLR]
49	172	6549	209	6,8-dimethylnaphthol			×
50	156	6615	211	3,3'-bipyridyl			×
51	180	6713	213	4-vinyl-2,6-dimethoxyphenol	×	×	×
52	180	6816	216	4-propanal-2-methoxyphenol			×
53	172	6841	217	7,8-dimethylnaphthol			×
54	194	6906	219	4-(prop-1-enyl)-2,6-dimethoxyphenol	×	×	×
55	194	7146	225	4-(prop-*cis*-2-enyl)-2,6-dimethoxyphenol		×	×
56	186	7292	229	2-acetyl-1-naphthol			×
57	194	7383	231	4-(prop-*trans*-2-enyl)-2,6-dimethoxyphenol	×	×	×
58	196	7496	233	4-acetyl-2,6-dimethoxyphenol			×
59	180	7560	236	4-(prop-2-enol)-2,6-dimethoxyphenol	×		
60	266	7660	240	3,7,11,15-tetramethylhexadec-1-ene			×
61	278	8130	252	3-methylene-7,11,15-trimethyl-hexadec-1-ene			×
62	256	8560	258	*n*-hexadecanoic acid		×	×

phenol (peak No. 51), 4-(prop-*trans*-2-enyl)-2,6-dimethoxyphenol (peak No. 57), and 4-(prop-2-enol)-2,6-dimethoxyphenol (peak No. 59). These compounds are also present in the KLR fraction, but the relative abundances are different.

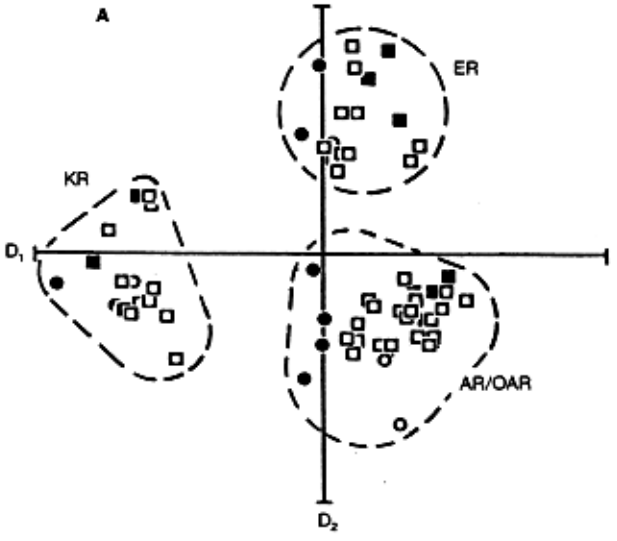
Comparative Studies of Py-MS Data using Multivariate Analysis

Large datasets can be compared using multivariate analysis. This technique offers the possibility of determining the mass-spectrometric differences in the Py-MS fingerprints of the tobacco samples which are analyzed quantitatively. These differences can be related to differences in chemical composition between the samples. In Figure 5 the factor scores of the first principal component of the BOKELMAN fractions of all tobaccos are given. It should be stressed here that the presence of replicate analyses in the dataset is not known to the computer performing the factor analysis; the mathematical analysis is not supervised.

The replicates lie close together in Figure 5, as indicated by the solid lines (error bars), and this reflects the reproducibility of each triplicate analysis. The Euclidean distances between the samples indicate the different distribution of mass peaks in the fingerprints and thus differences in chemical composition. The Euclidean distances between the ER, AR, and OAR fractions are relatively small compared to the KR and KLR fractions. The KLR fractions exhibit a relatively wide range of compositions. A subdivision within each fraction related to the various types of tobacco is not clear. The mass peaks responsible for the separate positions of the KLR and KR fractions relative to the ER, AR, and OAR can be found in the reconstructed mass spectra of the first principal component (Figures 5B and 5C). The KLR fractions reveal markers for lignin (*m/z* 124, 138, 150, 164, 180, 194 and 210 (Fig. 5B)), whereas the KR fractions are relatively enriched in polysaccharides (*m/z* 57, 60, 73, 114 and 126 (Fig. 5C)) and residual pectins (*m/z* 85, 32 and 31). The spread in the factor scores of the first principal component of the KLR fractions points to substantial chemical differences in the samples, which appear to be introduced by the Klason lignin procedure. The variations within the KLR fraction overshadow possible variations in the other fractions. In order to analyze the latter variations in more detail we performed a discriminant analysis on all but the KLR fractions. Using this mathematical procedure the variation between the triplicates of a sample is minimized in relation to the variation between the samples, but at the same time discriminant analysis maximizes the differences between the Py-MS fingerprints of the tobacco samples.

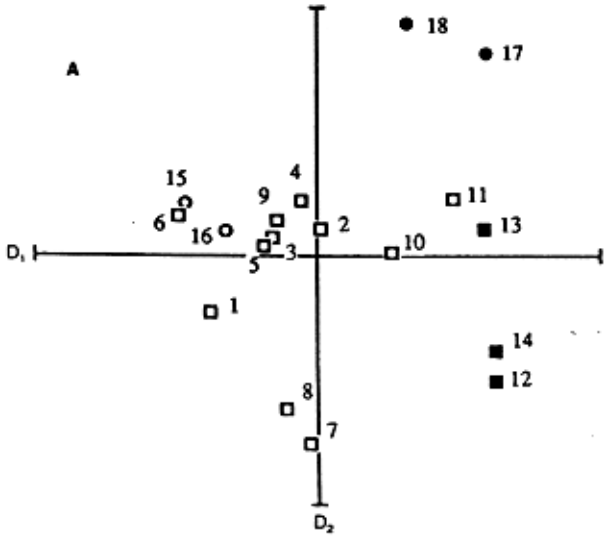
The result of this analysis is expressed as a D_1D_2 discriminant map (Fig. 6). Only the group centroids are plotted here, which is done routinely through Figures 6 to 11. A subdivision into three clusters is found: the KR fraction, ER fraction, and the combined AR/OAR fractions. The mass peaks responsible for the separate position of the KR fractions relative to the ER, AR, and OAR fractions are shown in the reconstructed mass spectra of the first discriminant function (Figs. 6B and 6C). The KR fraction is relatively enriched in polysaccharides as deduced from the mass peaks 144, 126 (hex-

Figure 6.
A: D_1D_2 map resulting from discriminant analysis of the Py-MS fingerprints of the combined ER, AR, OAR and KR fractions of all the tobaccos analyzed. Figures B and C show the reconstructed mass spectra of the first discriminant function, revealing characteristics for the ER, AR and OAR (see B) **and the KR** (see C) **fractions.**

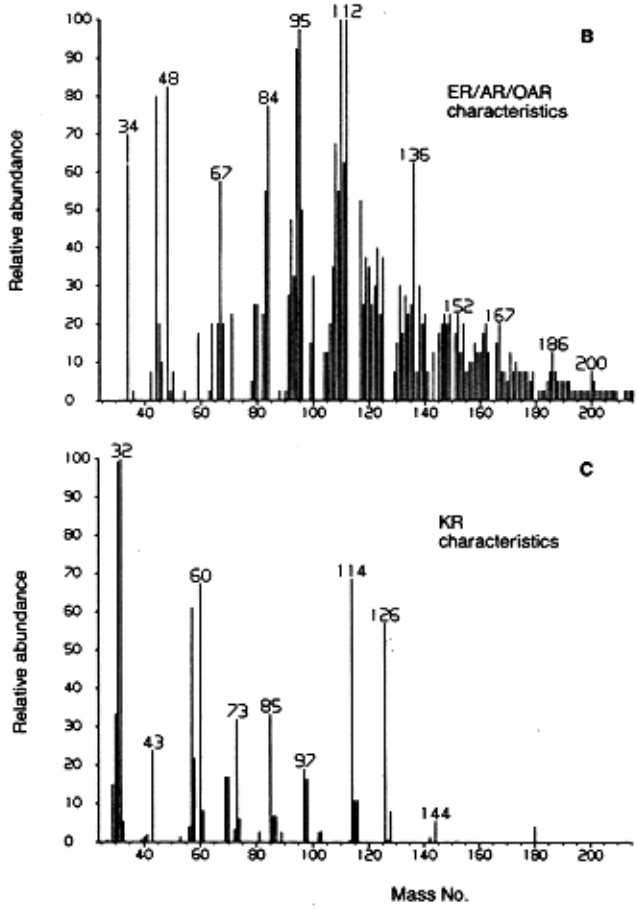


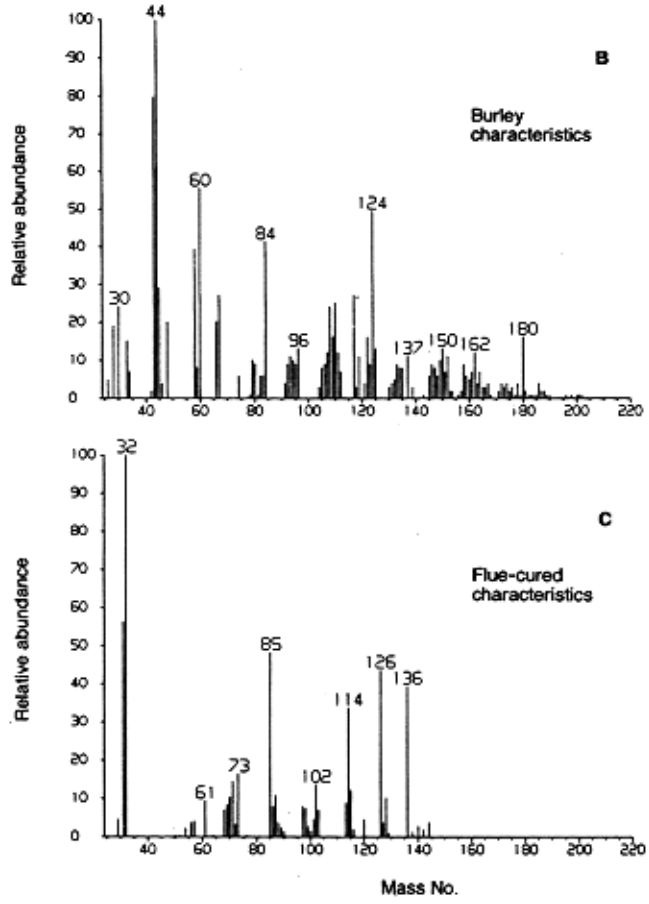
□: flue-cured tobacco, ○: Oriental tobacco,
■: Burley tobacco, ●: stem material.

Figure 7.
A: D_1D_2 map resulting from discriminant analysis of the Py-MS fingerprints of the ER fractions of all the tobaccos analyzed. Only the group centroids are plotted here. The numbers correspond to the tobacco sample Nos. listed in Table 1. **Figures B and C show the reconstructed mass spectra of the first discriminant function, revealing characteristics for the Burley** (see B) **and flue-cured** (see C) **type tobacco.**



□: flue-cured tobacco, ○: Oriental tobacco,
■: Burley tobacco, ●: stem material.





Py-MS: Curie-point pyrolysis mass spectrometry, AR: amylase residue,
ER: ethanol residue. OAR: blank run on AR, omitting the enzyme, KR: residue after 0.1 N KOH extraction.

oses), 114 (pentoses), 73, 60, 57 (anhydro sugars), and 85, 32, 31 (residual pectins). Figure 6B contains characteristics for the ER, AR, and OAR fractions, but the information is too dense for the chemical significance of the data to be determined.

Discrimination on the level of the tobacco varieties is not obvious from these data. Thus, the chemical differences induced by the fractionation procedure are larger than the differences between the various tobaccos or tobacco types. Separate evaluation of the Py-MS data of the residues from separate fractionation steps can serve as an alternative for investigating possible chemical differences between tobacco types, because in this way the dominating effects of the extraction procedures in the Py-MS data files are mainly overruled.

ER Fractions of the Tobaccos

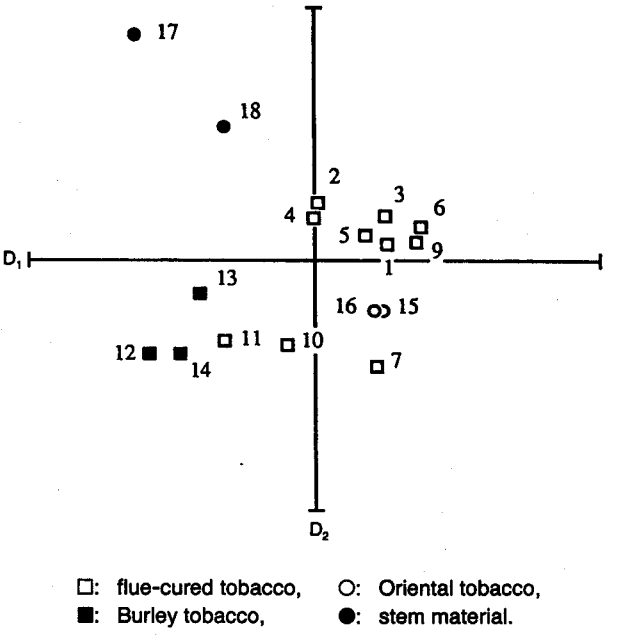
Figure 7 shows the graphical presentation of the first and second discriminant functions resulting from discriminant analysis of the combined ER fractions of all the tobaccos analyzed. The stem material (samples 17 and 18) and Burley tobacco (12 to 14) with positive D_1 scores are clearly distinguished from the flue-cured (1 to 11) and Oriental (15 and 16) tobaccos with negative D_1 scores. The Oriental tobaccos are separated from the flue-cured tobaccos by the D_3 function, plotting above the D_1D_2 plane. The spread in the Euclidean distances of the flue-cured tobaccos is rather wide, pointing to a wider range in chemical composition. Consistent differences between samples 7 and 8, 10 and 11, and the other flue-cured tobaccos are observed. The Euclidean distance between samples 7 and 8, both originating from Taiwan, is very small, pointing to a high degree of resemblance. The Malaysian flue-cured tobaccos 10 and 11 have positive D_1 scores and are classified as Burley tobaccos rather than flue-cured tobaccos. The main cluster of flue-cured tobaccos originates from Brazil, Canada, and the U.S.A. Thus, among the flue-cured tobaccos some phenotypic variations are present due to differences in climate and cultural management.

The mass peaks which account for the separation of the Burley tobaccos from the flue-cured and Oriental tobaccos are found in the reconstructed mass spectra of the first discriminant function (Figures 7B and 7C). The ER fractions of the Burley tobaccos are enriched in nicotine (*m/z* 84) and proteins (*m/z* 92, 117) compared to the flue-cured tobaccos. The stem material is distinguished from all the other samples by mass peaks which are indicative of lignin.

OAR Fractions of the Tobaccos

The differences in chemical composition of the tobacco samples on amylase treatment are not only due to the action of amylase but also to the hot water involved. In order to investigate this parameter a blank run without the addition of the enzyme was performed (OAR fractions), and the residues were analyzed with Py-MS. The D_1D_2 map resulting from discriminant analysis of the Py-MS fingerprints obtained is shown in Figure 8. Several clusters of tobacco samples can be observed. The Burley tobaccos (12 to 14), and stem materials (17 and 18) are clearly distinguished from the flue-cured tobaccos. Within the latter group the Euclidean distances are very large, indicating a greater variety in chemical composition. As observed before, sample 11 and, to a lesser degree, sample 10, both from Malaysia, are classified as Burley tobaccos because of their negative D_1 scores. Sample 7 from Taiwan is found to be an anomalous flue-cured tobacco (the OAR fraction of sample 8, also from Taiwan, was not available). The Oriental tobaccos (samples 15 and 16) are distinguished from the flue-cured tobaccos in D_3. On closer examination the relative distribution of the tobacco sample Py-MS fingerprints in the D_1D_2 map of the discriminant analysis of the OAR fractions in fact reflects the distribution in the D_1D_2 map resulting from the discriminant analysis of the ER fractions, rotated and inverted by the second discriminant function. Hot-water extraction evidently does not affect the mass-peak ratios in the Py-MS fingerprints of the tobacco samples.

Figure 8.
D_1D_2 map resulting from discriminant analysis of the Py-MS fingerprints of the OAR fractions of all the tobaccos analyzed. Only the group centroids are plotted here. The numbers correspond to the tobacco sample Nos. listed in Table 1.

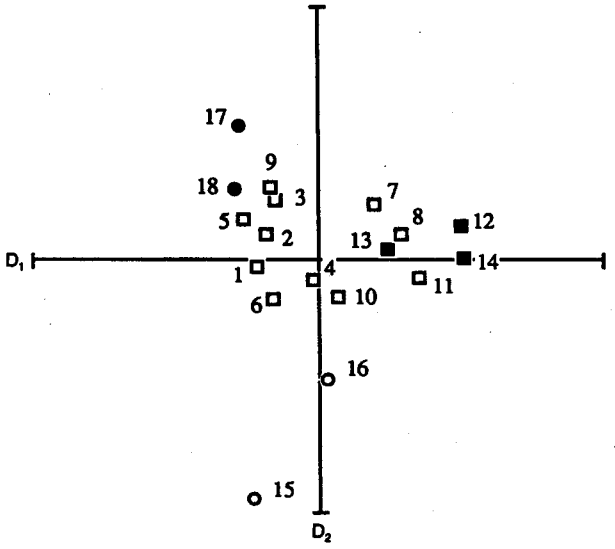


Py-MS: Curie-point pyrolysis mass spectrometry,
OAR: blank run on amylase residue, omitting the enzyme.

AR Fractions of the Tobaccos

The Euclidean distances between the Burley type and flue-cured type tobacco and the stem material de-

Figure 9.
D_1D_2 map resulting from discriminant analysis of the Py-MS fingerprints of the AR fractions of all the tobaccos analyzed. Only the group centroids are plotted here. The numbers correspond to the tobacco sample Nos. listed in Table 1.



□: flue-cured tobacco, ○: Oriental tobacco,
■: Burley tobacco, ●: stem material.

Py-MS: Curie-point pyrolysis mass spectrometry,
AR: amylase residue.

creased after the amylase treatment, as observed in the D_1D_2 plot of the discriminant analysis of the AR fractions (Fig. 9). These samples appear to have lost some of the characteristics responsible for their distinction from the other flue-cured tobaccos in the ER and OAR fractions. The Malaysian tobaccos (samples 10 and 11) are still plotted among the Burley tobaccos, whereas the Oriental tobaccos show a divergence after amylase treatment. Sample 15 is distinguished from sample 16 in D_2, something which is not observed in the OAR fractions. This cannot be explained by their different starch contents (Table 1). The stem material is distinguished from all the other samples in D_3.

Discriminant analysis of the Py-MS fingerprints of the combined AR and OAR fractions offers the possibility of investigating the net result of the amylase treatment in detail. The calculated D scores are given in Table 5. The flue-cured tobaccos respond relatively uniformly to amylase treatment; the D scores of these samples decreased, except for both the Malaysian tobaccos, 1 and 2. The D scores of the Burley tobaccos do not change significantly. This is related to their low starch content. Again, these results indicate that the Malaysian tobaccos must be regarded as Burley tobaccos rather than flue-cured tobaccos. The differences between the flue-cured and Burley-type tobaccos decreased on amylase treatment, which can be related to the different starch contents of the individual tobacco samples, but this does not appear to explain all the changes in composition.

Table 5.
Compilation of the D scores resulting from discriminant analysis of the Py-MS spectra of the combined AR and OAR fractions of all the tobaccos analyzed.

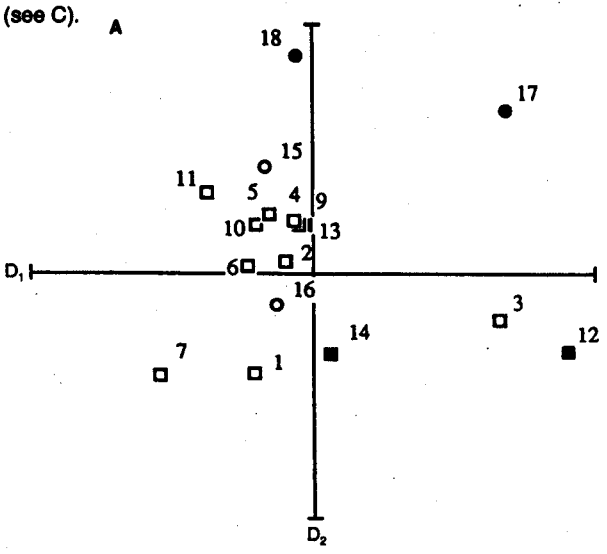
Sample No.	Tobacco sample	D scores	
		OAR fractions	AR fractions
Flue-cured tobacco:			
1	U.S.A. 1	0.0209	–1.0942
2	U.S.A. 2	0.4855	–0.3308
3	U.S.A. 3	0.1108	–0.8977
4	Brazil 1	0.5162	–0.6296
5	Brazil 2	–0.4479	–1.1551
6	Canada 1	–0.2719	–1.3017
7	Taiwan 1	0.3129	–0.5249
8	Taiwan 2	–	0.5852
9	China 1	–0.1209	–1.0567
10	Malaysia 1	0.0080	0.3794
11	Malaysia 2	1.2909	1.1196
Burley tobacco:			
12	South Korea 1	1.5489	1.3112
13	U.S.A. 4	0.5707	1.0302
14	Brazil 3	0.9054	1.1512
Oriental tobacco:			
15	Turkey 1	–2.4657	–0.8985
16	Turkey 2	–0.3666	–1.0801
Stem material:			
17	U.S.A. 3S	0.6736	1.1832
18	China 1S	0.5059	0.4723

KR Fractions of the Tobaccos

After extraction with 0.1 N KOH there is less distinction between the flue-cured and Burley tobaccos (Figure 10). The flue-cured tobaccos 4 and 9, and Burley tobacco 13 plot on top of each other. The total variance in the sample set decreased, and consequently the Euclidean distances between the replicates of a sample increased (not shown). The stem material (samples 17 and 18) can easily be distinguished from the leaf material. Only sample 12, a Burley tobacco from South Korea, still clearly differs from all the other tobaccos. Interestingly, sample 3, a flue-cured tobacco grown in the U.S.A., plots away from the main cluster of flue-cured tobaccos. The reconstructed mass spectra of the first discriminant function (Figures 10B and 10C) reveal that samples 3, 12 and 17 are relatively enriched in hexoses (*m/z* 144), pentoses (*m/z* 114), and residual pectins (*m/z* 85, 32 and 31), which could be due to differences in the extractability of polysaccharides in these samples.

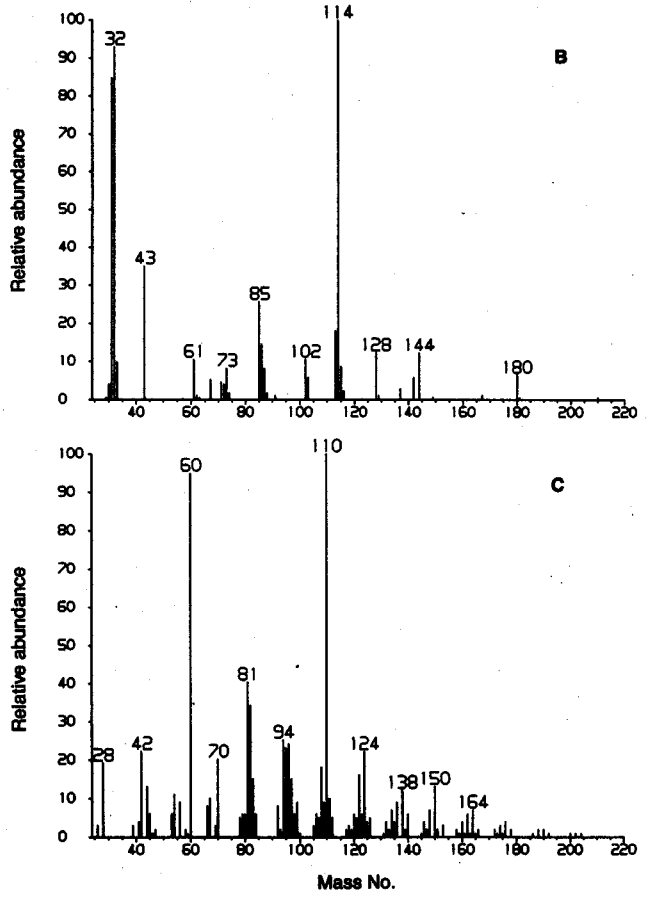
Figure 10.

A: D_1D_2 map resulting from discriminant analysis of the Py-MS fingerprints of the KR fractions of all the tobaccos analyzed. Only the group centroids are plotted here. The numbers correspond to the tobacco sample Nos. listed in Table 1. **Figures B and C represent the reconstructed mass spectra of the first discriminant function, revealing characteristics for samples 3, 12 and 17 (see B) and the remaining samples** (see C).



□: flue-cured tobacco, ○: Oriental tobacco,
■: Burley tobacco, ●: stem material.

Py-MS: Curie-point pyrolysis mass spectrometry,
KR: residue after 0.1 N KOH extraction.



KLR Fractions of the Tobaccos

The clustering behaviour between the flue-cured and Burley tobaccos also changed in the KLR fractions (see Fig. 11A) as a result of the treatment with 72% sulfuric acid. The leaf material (flue-cured, Burley, and Oriental tobaccos; samples 1 to 16) is clearly distinguished from the stem material (samples 17 and 18). Under an angle of 130° in the D_1D_2 map we observe an elongated cluster of Burley, flue-cured, and Oriental tobaccos. By graphical rotation of the D_1 and D_2 axis (37), spectra were obtained along the main axis of this cluster at angles of 40° (Fig. 11C), 130° (Fig. 11B) and 310° (Figure 11D). Several samples in the leaf-material cluster (flue-cured 3, 4, 6, 9; Oriental 15) show higher relative abundances of anhydro sugars in their pyrolyzates, pointing to residual polysaccharides (Fig. 11D). The Burley-type tobaccos reveal markers for residual alkaloids (*m/z* 84, 145, 146, 157, 158 and 162 (Fig. 11B)), which is confirmed with Py-GC-MS (Fig. 4C). A forthcoming paper will give more details on this subject (30).

The relatively large Euclidean distance between both Oriental tobaccos in the D_1D_2 plot of Figure 11 is also observed here as it was noticed earlier in the AR fractions. The stem material (sample Nos. 17 and 18) and both Malaysian tobaccos (samples 10 and 11) are clearly differentiated from the other tobaccos by their higher abundances of lignin markers in the Py-MS fingerprints (Fig. 11C).

Treatment with sulfuric acid results in a modification of the biopolymer systems present in tobacco, as observed in the Py-GC-MS trace of U.S.A. 4 (Fig. 4C). As a result, the spread in factor scores has increased (Fig. 5A). Thus the Klason lignin procedure promotes artifacts in the chemical composition of the residues, which do not reflect the refractory polymers in tobacco. The presence of polysaccharides and alkaloids in several samples can be explained assuming that these compounds are trapped within a modified lignin matrix, which prevents their breakdown by sulfuric acid.

DISCUSSION

The fractionation procedure investigated in this paper is designed to serve as an alternative for a serial-extraction procedure with the object of roughly estimating cell content and cell-wall materials in tobacco (3). This method for the systematic fractionation and characterization of cell-wall biopolymers in tobacco is based on procedures used with other plant material (28, 36). Errors in serial-extraction procedures are attributed to the fact that many distinct types of biopolymers have overlapping ranges of solubility (3). It is believed that ethanol extraction removes reducing sugars and oligosaccharides, water-soluble salts, lipids, waxes, resins, pigments, small acids, phenolic components, and alkaloids (3). The amylase removes starch, but this treatment

must be accompanied by a parallel experiment without the enzyme, in order to evaluate the extraction power of the hot water used. Dilute-alkali extraction is thought to remove protein, soluble hemicelluloses, and pectins. This should result in a residue which consists of cellulose and xyloglucans. The Klason procedure is thought to leave a residue enriched in lignin and condensed protein (29).

In order to discriminate between the various tobacco types multivariate analysis was performed on the Py-MS fingerprints of the samples listed in Table 1. Factor analysis on the Py-MS data of all polymer fractions reveals that the differences in chemical composition within each fraction are relatively small compared to the differences between the fractions. The multivarious origin of the samples analyzed does not give rise to a discrepancy in the scores of the first factor, except for the KLR fractions. The lignin markers in the pyrolyzate of the KLR fraction (Fig. 4C) are a number of mono and dimethoxyphenol derivatives with different substituents on position 4 of the aromatic nucleus. The monomethoxyphenols are in general more predominant than their dimethoxy analogues in the pyrolyzate. Several markers for residual alkaloids are observed among these guaiacyl and syringyl derivatives, such as 1-methylpyrrolidine, myosmine, 3-methyl-5-phenylpyrazole, and nicotine. All the nitrogen-containing compounds present in this KLR fraction are not known as pyrolysis products of proteins, except for indole (scan No. 5249) and methylindole (scan No. 5803). As a consequence the protein content of this fraction is overestimated by BOKELMAN and co-workers (3). This effect is most striking in the Burley tobaccos. Consequently the organic nitrogen content of the KLR fraction of a tobacco cannot be used to calculate the protein content of this fraction. These results imply that the maximum value for the lignin content of a tobacco is estimated incorrectly (3). Py-GC-MS analysis reveals that, besides alkaloids, fatty acids (up to $C_{24:0}$), aliphatic hydrocarbons (up to $C_{34}H_{70}$), and terpenoic hydrocarbons are observed when the film thickness of the stationary phase is decreased to 0.4 µm. Thus, in the KLR fractions differences in the response to the chemical processing are found, which can be related to the individual tobacco types. These differences result in characteristic refractory polymers which are present in the KLR fractions, e.g. proteins in Burley tobacco and polysaccharides in flue-cured tobacco. Perhaps an additional extraction step of the KLR fractions is needed to obtain more unambiguous results.

The high abundance of 1,6-anhydro-β-D-glucopyranose, 2-hydroxymethyl-5-hydroxy-2,3-dihydro-4*H*-pyran-4-one, levoglucosenone, pyranone and furan derivatives in the Py-GC-MS data of the KR fraction reveals the polysaccharide nature of this fraction. Among these saccharide markers several guaiacyl and syringyl derivatives can be observed. The D_1D_2 plot resulting from discriminant analysis of the Py-MS data on the KR fractions of all tobaccos analyzed (Fig. 10) indicates that the variability within this fraction has greatly diminished compared to the other polymer fractions. Characteristics for the individual tobacco types (flue-cured, Burley, and Oriental) are not detected in the KR fractions, which is clearly a result of the chemical processing of the tobacco samples. These observations support the existence of a conservative xyloglucan/cellulose/lignin framework in tobacco, which is less dependent on its geographical origin. The pectin, which is not completely removed by the KOH treatment, is thought to represent a polymer which is relatively tightly attached to the surface of the cell-wall system mentioned. The Euclidean distances between samples 3 and 12, and the rest of the sample set can be explained if incorrect treatment of the samples during the base extraction is assumed. Incomplete removal of KOH influences the pyrolytic breakdown of biopolymer systems (21). This is known to promote chain-fission reactions through reverse aldol condensation (32, 33). Py-GC analysis of these KR fractions with simultaneous FID and NPD revealed a minor contribution of nitrogen-containing compounds; only 1*H*-pyrrole and pyridine were found (not shown). The presence of proteins in this fraction as stated by BOKELMAN et al. (3) cannot be confirmed.

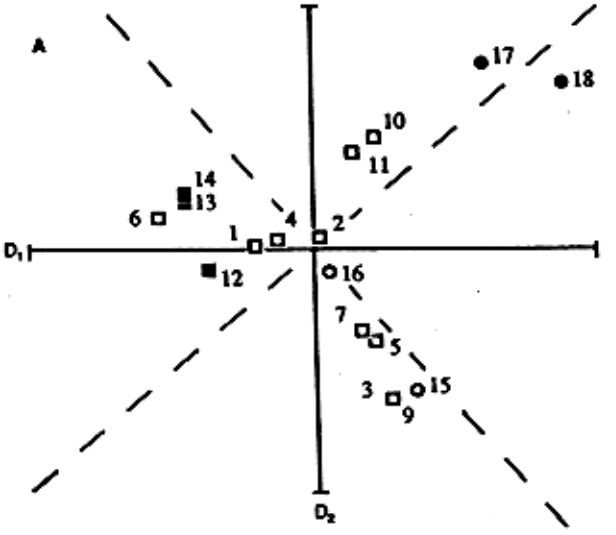
The main peak in the Py-GC-MS trace of the ER fraction of U.S.A. 4, a typical Burley tobacco, represents nicotine. In these data markers are found for alkaloids, polysaccharides, proteins, and lignin. The ethanol extraction is far from complete with respect to the removal of alkaloids and, to a minor degree, lipids and waxes. Discriminant analysis of the Py-MS data of the separate ER and AR fractions of the tobacco samples analyzed reveal characteristics for the flue-cured and Burley-type tobaccos. The flue-cured tobaccos are enriched in carbohydrates, whereas the Burley tobaccos contain relatively higher amounts of alkaloids and protein. These properties appear to be of cytosolic nature reflecting phenotypic variations and differences in cultural management. Discriminant analysis of the Py-MS fingerprints of the separate ER and AR fractions is excellent for classifying the tobaccos or tobacco types.

Factor analysis of a total file containing the Py-MS data of the ER, AR, OAR, KR, and KLR fractions of all the tobaccos analyzed, revealed that the chemical differences induced by the subsequent extraction steps are larger than the chemical differences related to the individual tobacco types. Thus, comparison of the weight data of polymer fractions derived from tobacco is not suitable for determining and classifying the tobacco type used. The chemical interpretation of the weight of fractions resulting from classical wet-chemical methods needs reconsideration. A great problem in any chemical fractionation approach is the assumption that the solvents, acids, and bases used have the same effect with different plant materials, resulting in comparable fractions of a uniform chemical nature. Our data clearly show that widely-accepted abundances for chemical-compound classes, e.g. proteins and lignin in the Klason lignin residues, are estimated incorrectly. As a consequence we have to state that new parameters are re-

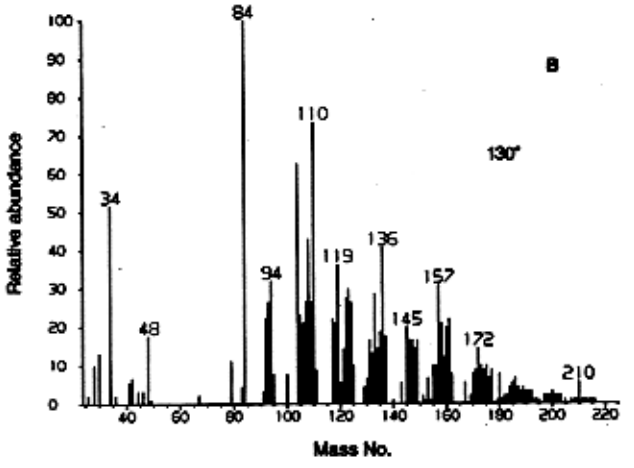
Figure 11.

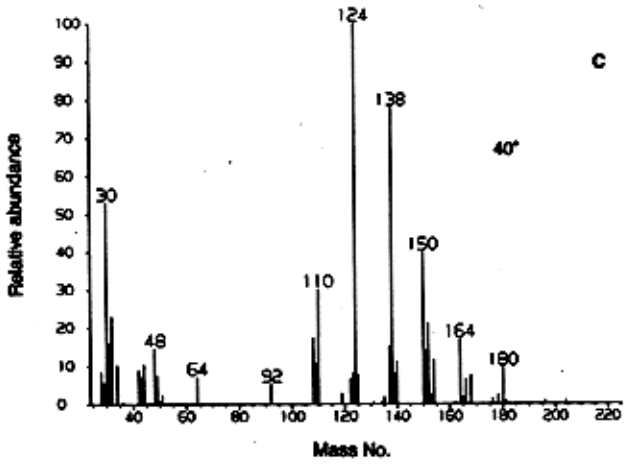
D_1D_2 map resulting from discriminant analysis of the Py-MS fingerprints of the KLR fractions of all the tobaccos analyzed. Only the group centroids are plotted here.
The numbers correspond to the tobacco sample Nos. listed in Table 1.
Figures B, C and D show the reconstructed mass spectra of the first discriminant function, after rotation by 130°, 40° and 310°, respectively.

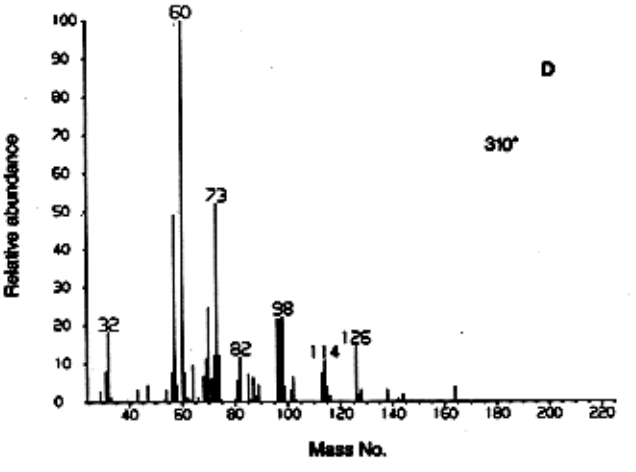
Py-MS: Curie-point pyrolysis mass spectrometry,
KLR: Klason lignin residue.



□: flue-cured tobacco, ○: Oriental tobacco,
■: Burley tobacco, ●: stem material.







quired to describe characteristics for the individual tobacco types.

This survey shows that the chemical nature and origin of several Py-MS markers and associated pyrolysis products can be correlated with various classes of chemical compounds and known biopolymer systems. At the present state of investigations the impurity of the tobacco-derived polymer fractions, especially the ER and AR fractions, causes severe interpretation problems and precludes a more detailed understanding of the polymer systems present in tobacco. With analytical pyrolysis these impurities could be analyzed and investigated in detail. An improved knowledge of the structure of pyrolysis products and their correlation with biopolymer systems will have to be developed and should result in a tool-box for the rapid mass-spectrometric analysis of plant material. Further development of a fractionation strategy for plant material either by chemical or biological means, e.g. through the use of enzymes and subsequent LC purification, should include analytical pyrolysis techniques at an early stage, as applied in this paper.

CONCLUSIONS

Py-MS analyses of different tobaccos and polymer fractions derived therefrom (ER, AR, OAR, KR, and KLR) can be performed reproducibly for the meaningful multivariate analysis of the data obtained. Multivariate analysis of the Py-MS fingerprints is an excellent technique for comparing large numbers of tobacco samples, and hence suitable for industrial applications.

Py-MS data on the distinct types of tobacco (flue-cured, Burley, and Oriental) show that differences in the chemical composition of the fractions obtained are relatively small compared to changes induced by the fractionation procedure itself. Thus, the fractionation scheme proposed by BOKELMAN and co-workers (3) is a meaningful tool for investigating the chemical composition of tobacco, but not suitable for determining and classifying the separate types of tobacco, unless discriminant analysis of each polymer fraction is performed individually. In this way characteristics for the flue-cured and Burley-type tobacco, and stem material are extracted in the ER, OAR, and AR fractions. These characteristics reflect phenotypic variations and differences in cultural management. On alkali extraction a relatively uniform cell-wall skeleton is revealed. The Klason lignin procedure used to estimate refractory compound classes, e.g. proteins, produces artifacts in the chemical composition of the residues, which are related to the tobacco types used.

Py-GC-MS shows that several classes of chemical constituents, e.g. proteins, alkaloids, lipids, and terpenoids, are not completely extracted, as is widely assumed. The extraction steps in the fractionation procedure, especially the ER, AR, and KLR fraction, require further investigation to obtain well-defined residues in which interference from refractory polymers is minimized.

Analytical pyrolysis techniques clearly show that widely-accepted abundances for several compound classes in these fractions are determined incorrectly, which makes characterization of a tobacco more difficult. In the future the significance of these compound classes for identifying and classifying tobaccos has to be reconsidered in order to obtain a more solid foundation for distinguishing between tobaccos or tobacco types.

REFERENCES

1. Aries, R. E., C. S. Gutteridge and W. A. Laurie: A pyrolysis mass spectrometry investigation of pectin methylation; Personal communication.
2. Bokelman, G. H., and W. S. Ryan: Analyses of Bright and Burley tobacco laminae and stems; Beitr. Tabakforsch. Int. 13 (1985) 29—36.
3. Bokelman, G. H., W. S. Ryan and E. T. Oakley: Fractionation of Bright tobacco; J. Agric. Food Chem. 31 (1983) 897—901.
4. Boon, J. J., B. Brandt - de Boer, W. Genuit, J. Dallinga and E. Turkstra: Characterization of particulate organic matter from sediments in the estuary of the Rhine and from offshore dump sites of dredging soils; *in* Organic marine geochemistry, edited by M. L. Sohn, ACS (Am. Chem. Soc.) Symp. Ser. 305, 1985, 76—90.
5. Boon, J. J., L. Dupont and J. W. de Leeuw: Characterization of a peat bog profile by Curie-point pyrolysis - mass spectrometry combined with multivariate analysis and by pyrolysis gas chromatography - mass spectrometry; *in* Peat and water, edited by C. H. Fuchsman, Elsevier Science Publishers, Amsterdam, 1986, pp. 215—239.
6. Boon, J. J., G. B. Eijkel and E. P. H. Best: Study of the seasonal change in composition in roots and shoots of *Phragmites australis* by pyrolysis mass spectrometry combined with multivariate data analysis; Proc. Int. Symp. on Aquatic macrophytes (Fac. Sci., Cath. Univ. Nijmegen, The Netherlands), 1983, 1—5.
7. Boon, J. J., and J. Haverkamp: Pyrolysis mass spectrometry of intact and decomposed leaves of *Nuphar variegatum* and *Zostera marina*, and some archeological eelgrass samples; Hydrobiol. Bull. 16 (1982) 71—82.
8. Boon, J. J., A. D. Pouwels and G. B. Eijkel: Pyrolysis high-resolution gas chromatography - mass spectrometry studies on beech wood — Capillary high-resolution mass spectrometry of a beechwood lignin fraction; Biochem. Soc. Trans. 15 (1986) 170—174.
9. Boon, J. J., A. Tom, B. Brandt - de Boer, G. B. Eijkel, P. G. Kistemaker, F. J. W. Notten and F. H. M.

Mix: Mass spectrometric and factor discriminant analysis of complex organic matter from bacterial culture environment of *Bacteroides gingivalis;* Anal. Chim. Acta 163 (1984) 193—205.

10. Boon, J. J., W. Windig, R. G. Wetzel and G. L. Godshalk: The analytical pyrolysis of particulate residues of decomposing *Myriophyllum heterophyllum;* Aquat. Bot. 15 (1983) 307—320.
11. Brock, T. C. M., J. J. Boon and B. G. F. Paffen: The effects of the season and of water chemistry on the decomposition of *Nymphea alba* L. — Weight loss and pyrolysis mass spectrometry of the particulate matter; Aquat. Bot. 22 (1988) 197—229.
12. Cartwright, W., A. Cornell and V. Olender: Preliminary studies on micro-pyrolysis gas chromatography profiling of tobacco components; Abstr. 34th Tobacco Chemists' Research Conference (1980), Richmond, Va., U.S.A.
13. Dube, M. F., and C. R. Green: Methods of collection of smoke for analytical purposes; Recent Adv. Tob. Sci. 8 (1982) 42—102.
14. Enzell, C. R., and I. Wahlberg: Leaf composition in relation to smoking quality and aroma; Recent Adv. Tob. Sci. 6 (1980) 64—122.
15. Finster, P., F. Tödter and F. Seehofer: Automatisierte Simultanbestimmung von Tabakinhaltsstoffen (Teil I); LaborPraxis 6 (1982) 1362—1364.
16. Genuit, W., J. J. Boon and O. Faix: Characterization of beech-milled wood lignin by pyrolysis - gas chromatography - photoionization mass spectrometry; Anal. Chem. 59 (1987) 508—512.
17. Halket, J. M., and H. R. Schulten: Rapid characterization of tobacco by combined direct pyrolysis - field ionization mass spectrometry and pyrolysis - gas chromatography - mass spectrometry; J. Anal. Appl. Pyrolysis 8 (1985) 547—560.
18. Hartley, R. D., and J. Haverkamp: Pyrolysis - mass spectrometry of the phenolic constituents of plant cell walls; J. Sci. Food Agric. 35 (1984) 14—20.
19. Hoogerbrugge, R., S. J. Willig and P. G. Kistemaker: Discriminant analysis by double-stage principal component analysis; Anal. Chem. 55 (1983) 1711—1712.
20. Kaaden, A. van der, J. J. Boon and J. Haverkamp: The analytical pyrolysis of carbohydrates, 2. Differentiation of homopolyhexoses according to their linkage type, by pyrolysis - mass spectrometry and pyrolysis - gas chromatography / mass spectrometry; Biomed. Mass Spectrom. 11 (1984) 486—492.
21. Kaaden, A. van der, J. Haverkamp, J. J. Boon and J. W. de Leeuw: Analytical pyrolysis of carbohydrates, 1. Chemical interpretation of matrix influences on pyrolysis - mass spectra of amylose using pyrolysis gas chromatography - mass spectrometry; J. Anal. Appl. Pyrolysis 5 (1983) 199—220.
22. McComb, E., and R. M. McCready: Colometric determination of pectic substances; Anal. Chem. 29 (1957) 819—821.
23. Meuzelaar, H. L. C., J. Haverkamp and F. D. Hilleman: Pyrolysis mass spectrometry of recent and fossil biomaterials; Elsevier Science Publishers, Amsterdam, 1982.
24. Oakley, E. T.: Enzymatic determination of starch in fresh green, lyophilized green, and cured tobacco; J. Agric. Food Chem. 31 (1983) 902—905.
25. Pomogyi, P., E. P. H. Best, J. H. A. Dassen and J. J. Boon: On the relation between age, plant composition and nutrient release from living and killed *Ceratophyllum* plants; Aquat. Bot. 19 (1984) 243—250.
26. Pouwels, A. D., A. Tom, G. B. Eijkel and J. J. Boon: Characterization of beech wood and its holocellulose and xylan fractions by pyrolysis - gas chromatography - mass spectrometry; J. Anal. Appl. Pyrolysis 11 (1987) 417—436.
27. Rosa, N.: Pyrolysis - gas chromatographic estimation of tobacco alkaloids and neophytadiene; J. Chromatogr. 171 (1979) 419—423.
28. Salomonsson, A.-C., O. Theander and P. Aman: Composition of normal and high-lysine barleys; Swed. J. Agric. Res. 10 (1980) 11—16.
29. Sarkanen, K. V., and C. H. Ludwig: Lignins — Occurrence, formation, structure and reactions; John Wiley & Sons Inc., New York, N.Y., 1971.
30. Scheijen, M. A., J. J. Boon, W. Hass and V. Heemann: Characterization of tobacco lignin preparations by Curie-point pyrolysis - mass spectrometry and Curie-point pyrolysis - high resolution - gas chromatography - mass spectrometry; J. Anal. Appl. Pyrolysis 15 (1989) 97—120.
31. Schlotzhauer, W. S., and O. T. Chortyk: Recent advances in studies on the pyrosynthesis of cigarette smoke constituents; J. Anal. Appl. Pyrolysis 12 (1987) 193—222.
32. Shafizadeh, F., and Y. Z. Lai: Thermal degradation of 1,6-anhydro-β-D-glucopyranose; J. Org. Chem. 37 (1972) 278—284.
33. Shafizadeh, F., C. W. Philpot and N. Ostojic: Thermal degradation of 1,6-anhydro-β-D-glucopyranose; Carbohydr. Res. 16 (1971) 279—287.
34. Stedman, R. L.: The chemical composition of tobacco and tobacco smoke; Chem. Rev. 68 (1968) 153—207.
35. Thome, F. A., G. W. Young and W. C. Hamlin, Jr.: Applications of pyrolysis - gas chromatography - mass spectrometry to tobacco chemistry; presented at the 1982 CORESTA Symposium held in Winston-Salem, North Carolina, U.S.A., CORESTA Information Bulletin 1982, Special Issue, pp. 22—23.
36. Soest, P. J. van, R. H. Wine and L. A. Moore: Estimation of the true digestibility of forages by the *in vitro* digestion of cell walls; Proc. X. Int. Grassl. Congr., Helsinki, Finland, 1966, 438—441.

37. Windig, W., P. G. Kistemaker and J. Haverkamp: Chemical interpretation of differences in pyrolysis - mass spectra of simulated mixtures of biopolymers by factor analysis with graphical rotation; J. Anal. Appl. Pyrolysis 3 (1982) 199—212.

Acknowledgements

The authors gratefully acknowledge the technical assistance of Mr. Gert B. Eijkel and Mrs. G. Spremberg.

This work is part of the research programme of the Foundation for Fundamental Research on Matter (FOM).

Addresses of authors:

Martin A. Scheijen,
Unit for Mass Spectrometry of Macromolecular Systems,
FOM Institute for Atomic and Molecular Physics,
Kruislaan 407,
NL-1098 SJ Amsterdam,
The Netherlands.

Werner Hass,
B·A·T Cigarettenfabriken GmbH,
Forschung und Entwicklung,
Bahrenfelder Chaussee 139,
D-2000 Hamburg 50,
Federal Republic of Germany.