JMS

JMS Letters

Dear Sir,

Structural elucidation of the wheat straw lignin polymer by atmospheric pressure chemical ionization tandem mass spectrometry and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

The structure of the cell wall of the lignocellulose fibers which constitute the skeletal substance of all terrestrial plants can be schematically viewed as bundles of cellulose microfibrils ordered in parallel in a matrix of amorphous hemicelluloses and lignins.[1,2] Obtaining cellulose fibers from vegetable matter using traditional industrial paper procedures consists of degrading large amounts of lignins and hemicelluloses and making them soluble in aqueous media.[3]

Lignin, one of the most abundant naturally occurring materials, was identified as a soluble residue of acidic hydrolysis of wood more than 150 years ago.[4] It is only in the last 50 years that the polymeric structure of lignin based on phenylpropenyl alcohol monomers (also called lignols) has been proposed. Knowledge about the lignin structure is still fragmentary. As shown in Fig. 1, it has been established that the monomeric lignol precursor in conifers is predominantly coniferyl alcohol (Fig. 1(A)), whereas in deciduous trees it is sinapyl alcohol (Fig. 1(B)) and in grasses and herblike dicotyledons it is *p*-cumaryl alcohol (Fig. 1(C)).[5]

In the last two decades, the biological heterogeneity of its molecular structure has been reported and it is well accepted that the chemical structure of the various lignin polymers depends on the botanical origin and chemical composition of the vegetable polymeric fibers.[6] It has been proposed that the native lignin polymer is produced through random polymerization processes. Although Freudenberg and Neish have postulated that the lignols polymerize to yield native lignin macromolecules by a radical mechanism, the exact biosynthetic pathway has not yet been definitely established.[7] The structure of the heterogenous lignin macromolecule remains unsolved and is a daunting task, especially as the extraction protocol appears to modify and degrade the lignin polymer. Thus, the initial step in lignin analysis will be to isolate lignin from the

*Correspondence to: J. H. Banoub, Fisheries and Oceans Canada, P.O. Box 5667, St. Johns, Newfoundland, A1C 5X1, Canada.
E-mail: banoubjo@dfo-mpo.gc.ca

pulp matrix without causing structural changes. Although many chemical and enzymatic methods have been developed for the isolation of lignin from vegetable matter fibers, none are without the risk of causing structural changes during isolation. To determine the constituents of lignin by chemical degradation is an extremely labor-intensive task, which, unfortunately, gives no clear answers as to its molecular makeup. Absolute molar mass values of lignins are difficult to determine because they have a strong tendency to form molecular aggregates in solution.[5,6] This is particularly true with the industrial chemical pulp manufacturing procedures used today, which favor either repolymerization and/or depolymerization of lignin molecules to afford erroneous molecular masses which may be higher or lower than the native lignin molecules. Nevertheless, attempts to determine the molecular weights of lignins have been made using gel permeation chromatography.[9]

Mass spectrometry, using electron ionization, has been used for the study and characterization of derivatized lignol monomer constituents which were released by either reductive cleavage or by pyrolysis. In this case, only monomeric, and to a small extent dimeric, products were identified by comparison of gas chromatographic retention times and mass spectra with authentic samples.[10–12]

We report here the structural characterization of the wheat straw lignin polymer using combined techniques such as atmospheric pressure chemical ionization mass spectrometry (APCI-MS), tandem mass spectrometry (MS/MS) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOFMS). The wheat straw lignin was extracted using the novel AVIDEL procedure, which selectively separates the cellulose, hemicellulose and lignin, and this allows the destructuring of the vegetable matter at atmospheric pressure by a catalyst–solvent system of formic acid–acetic acid.[13,14] Recently, a study on the structural elucidation by MALDI-TOFMS of lignins,[8] chestnut elligitannins and polyflavonoid tannins has been reported.[15,16]

The APCI-mass spectra of the lignin extracted from wheat straw in a mixture of formic acid, acetic acid and water were recorded in the negative and positive ion modes and are presented in Fig. 2(A) and (C). The lignin was dissolved in chloroform–methanol (2 : 1) and was injected directly through the APCI interface in to the mass spectrometer, which was set to scan for negative ions in the range m/z 100–650 (Fig. 2(A)). This APCI-mass spectrum indicates the presence of two major aromatic compounds tentatively assigned as a dimer **2**, $C_{18}H_{18}O_6$ (molecular mass = 330.11) and a trimer **3**, $C_{29}H_{30}O_8$ (molecular mass = 507.19), the structures of which are proposed in Fig. 1. The dimeric lignin fragment **2** gives a deprotonated molecular ion $[M - H]^-$ at m/z 329 and appears to consist of a modified coniferyl alcohol in which the *p*-hydroxyl group has been replaced by a carboxylic group that has condensed with another coniferyl alcohol via the production of a cyclic ether. The trimeric lignin fragment **3** gives a deprotonated molecular ion $[M - H]^-$ at m/z
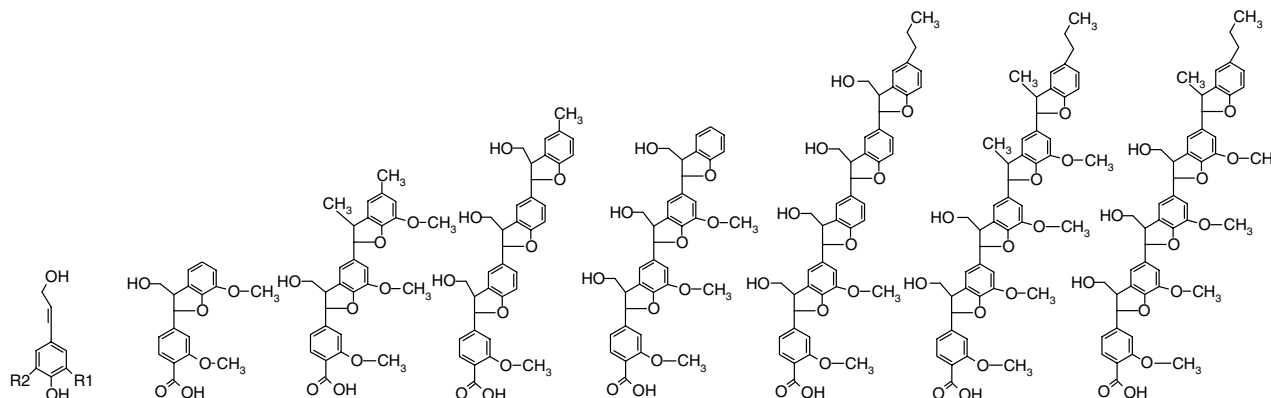
1A: R_1=OCH$_3$ R2=H	2: $C_{18}H_{18}O_6$	3: $C_{29}H_{30}O_8$	4: $C_{36}H_{34}O_9$	5: $C_{37}H_{36}O_{11}$	6: $C_{48}H_{48}O_{12}$	7: $C_{50}H_{52}O_{12}$	8: $C_{50}H_{52}O_{13}$
1B: R_1=R_2=OCH$_3$	Mr=330.11	Mr=506.19	Mr=610.22	Mr=656.22	Mr=816.31	Mr=844.34	Mr=860.34
1C: R_1=R_2=H							

Figure 1. Chemical structures of the various wheat straw lignin polymeric fragments.
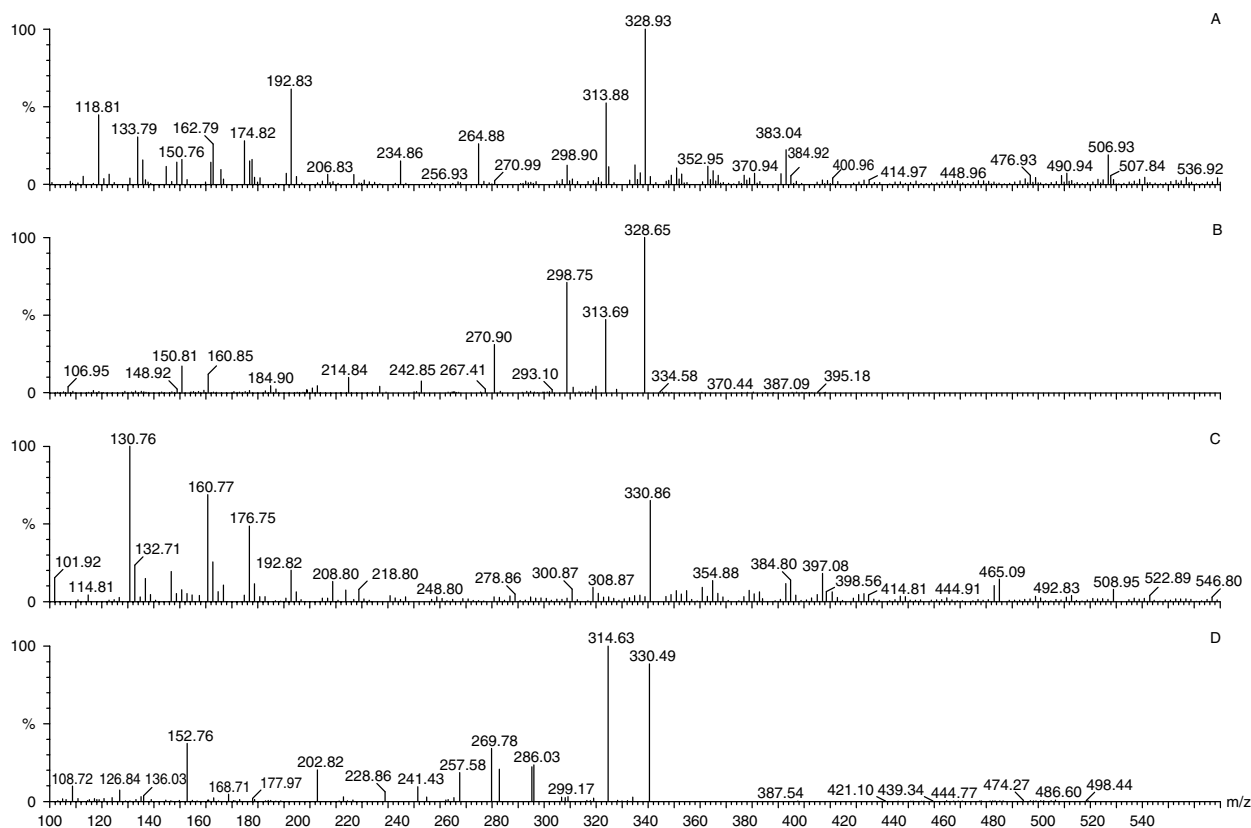
Figure 2. APCI mass spectra of the native wheat lignin polymer recorded in the negative ion mode (A) and positive ion mode (C) and product ion tandem mass spectra of the deprotonated molecular ion [M − H]⁻ at m/z 329 (B) and protonated molecular ion [M + H]⁺ at m/z 331 (D).

506 and appears to consist of three modified coniferyl alcohols assembled as previously described. The presence of one hydroxyl group within the dimeric and trimeric lignin fragments **3** and **4** was established by acetylation of the native lignin with acetic anhydride and pyridine. APCI-MS of the acetylated mixtures unambiguously indicated the presence of two newly acetylated lignin fragments, dimeric $C_{20}H_{20}O_7$ (molecular mass = 372.11) and trimeric $C_{31}H_{32}O_9$ (molecular mass = 550.19), which gave the deprotonated molecular ions [M − H]⁻ at m/z 371 and 549, respectively.

The presence of the *p*-carboxylic, hydroxyl and methoxyl groups was also established by measuring the product ion tandem mass spectra from the deprotonated molecular ions at m/z 329, 371, 506 and 549. The APCI product ion mass spectrum of the deprotonated molecular ion of the dimeric lignin fragment $[C_{18}H_{18}O_6 - H]^-$ at m/z 329 is presented in Fig. 2(B). From this product ion tandem mass spectrum, we can deduce that the deprotonated molecular ion cleaves into two diagnostic product ions at m/z 151 and 161 by fission between the aromatic ring containing the carboxylic group and the cyclic ether portion of the dimer. The product ions at m/z 215, 227, 243, 271, 299 and 314 are formed by consecutive losses of CO, CO_2, CH_3, CH_2O and CH_2 functional groups. The major fragmentation routes of the product ion tandem mass spectrum of [M − H]⁻ at m/z 329 are proposed in Fig. 3. The precursor ion tandem mass spectra of all the various fragment ions at m/z 151, 161, 215, 227, 242, 271, 299 and 314 were also measured and showed that all produced fragment ions originated only from the dimeric lignin deprotonated molecular ion [M − H]⁻ at m/z 329. These concerted losses indicate the various reactivities of the precursor ions (Figure 3).

The APCI mass spectrum of the wheat lignin was also recorded in the positive ion mode (Fig. 2(C)) and, once again, indicated the presence of the lignin dimeric fragment $C_{18}H_{18}O_6$, and trimeric fragment $C_{29}H_{39}O_8$, which gave their expected protonated molecular ions [M + H]⁺ at m/z 331 and 508, respectively. The APCI product ion tandem mass spectra of m/z 331 and 508 were also recorded and only that of m/z 331 is indicated in Fig. 3(D). The major fragmentation routes describing the formation of the product ion from the MS/MS of m/z 331 are tentatively proposed in Fig. 4. Once again, the formation of the various fragment ions produced

in this MS/MS suggests the proposed structure of the dimeric lignin fragment $C_{18}H_{18}O_6$. It should be mentioned that, in Fig. 4, the fragment ion at m/z 241 can be formed by concerted multiple neutral losses from the [M + H]⁺ ion at m/z 331, or by consecutive, stepwise dissociation involving the ion at m/z 315 and/or 285 as intermediates. In this context, note that 'consecutive' or 'concerted' losses of many such product ions in the MS/MS experiment simply means that they are both lost at the same time and within the same reaction region of the tandem mass spectrometer. It is very difficult to deduce the order of such elimination reactions under the MS/MS conditions and these were not investigated further.

Attempts to identify higher polymeric lignin fragment masses by APCI-MS in the negative and positive ion modes failed even when higher ion scans were set in the range m/z 600–1200. MALDI-MS appeared of interest for the structural characterization of the wheat lignin polymer because of the high sensitivity and the possibility of analyzing high molecular mass fragments. The possibility that perhaps the original wheat lignin polymer was indeed greater than a trimer did not escape our attention and we predicted that perhaps the chemical instability and fragility of the lignin polymer may have not been fully compatible with the atmospheric pressure ionization, breaking down selectively to both dimeric and trimeric fragment ions. It is noteworthy that similar results were obtained using electrospray ionization mass spectrometry (ESI-MS) and MS/MS.

MALDI-TOFMS of the wheat lignin polymer in the positive ion mode gave a series of prominent protonated molecules, *inter alia*, at m/z 861.33, 845.38, 657.44 and 331.67. These series of protonated molecules were tentatively assigned as pentameric lignin fragments $C_{50}H_{52}O_2$ (molecular mass = 844.39) assigned structures **8** and **7**, respectively (Fig. 1). The tetrameric lignin fragment $C_{37}H_{36}O_{11}$ (molecular mass = 656.22) was assigned structure **5** (Fig. 1). MALDI-TOFMS of the same wheat lignin polymer gave a series of deprotonated molecules in the negative ion mode at, *inter alia*, m/z 814.78, 609.88 and 328.68, assigned structures **6**, **4** and **3**, respectively (Figure 1).

A quick examination of structure **8** (Fig. 1) and other proposed structures **2–7** seems to indicate that these molecules are interrelated and are composed principally of a polycondensation of coniferyl
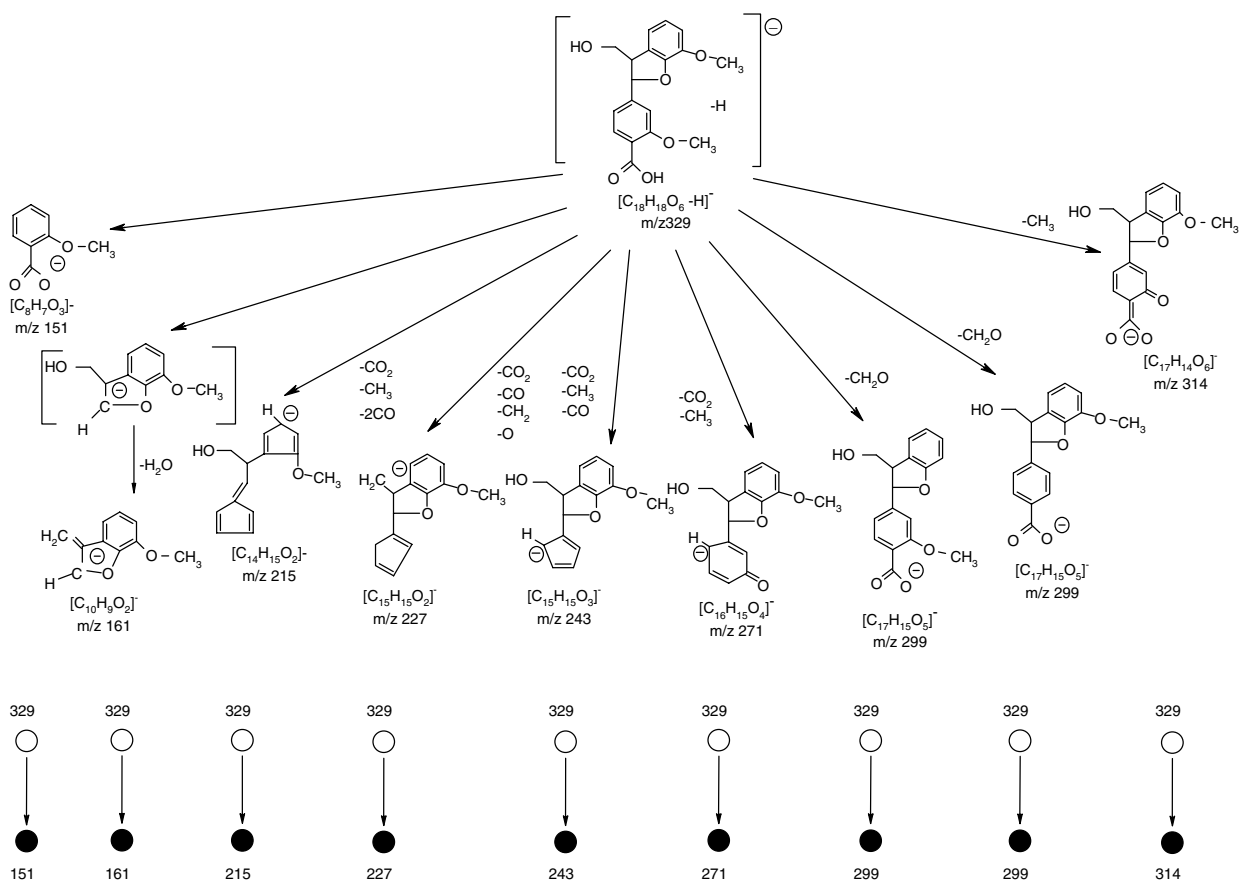
Figure 3. Proposed fragmentation routes of the product ion tandem mass spectra of the deprotonated molecular ion $[M - H]^-$ at m/z 329 and summaries of the precursor ion MS/MS scans of the various selected fragments.
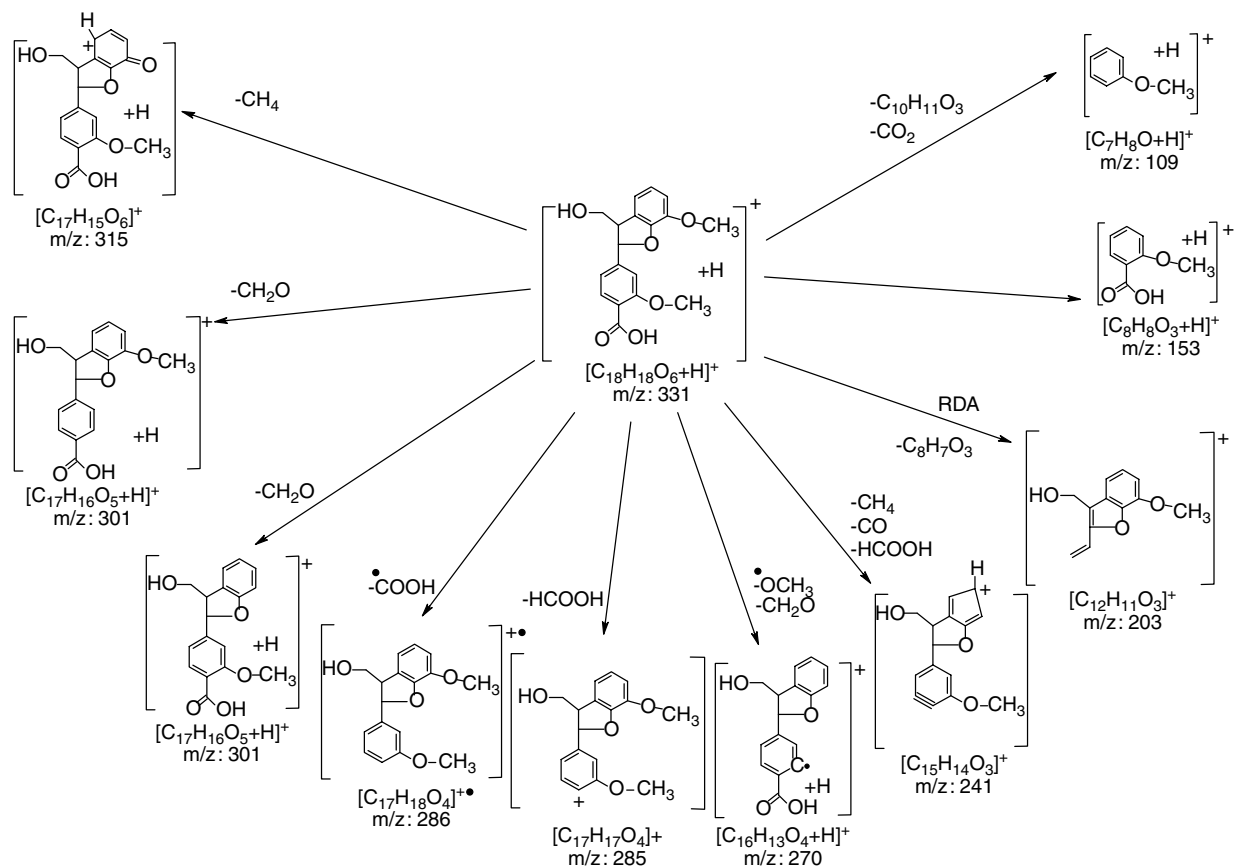
Figure 4. Proposed fragment routes of the product ion tandem mass spectrum of the protonated molecular ion $[M + H]^+$ at m/z 331.

alcohols (Fig. 1(A). We can observe a prevalent reduction of the terminal —CH=CH—CH$_2$OH end into the —CH$_2$CH$_2$CH$_3$ end group, in addition to losses of methoxyl groups and concurrent cleavages occurring at the cyclic ether locations. This indicates that the original wheat lignin is probably a longer polymer of polycondensed coniferyl alcohols in which the backbone is fixed in a cyclic ether repeating unit, one end being terminated by a carboxylic group and the other end with a hydroxyl group. These are attached by esterification/etherification to other cellulose or hemicellulose polymers. Such a native polymer is acid labile and breaks down to the series of compounds identified in Fig. 1.

To our knowledge, the method of extraction of lignin from wheat by the AVIDEL process seems to be one of the least destructive methods reported in the literature.[13,14] It has been suggested that, from the biological point of view, lignins appear to be synthesized during their polymerization by a non-enzymatic process occurring at the outside of the living cell (the cell wall). Thus, lignin is considered to be an infinite random three-dimensional network polymer, implying a uniform chemistry (or perhaps uniform destruction) of the pulping products.[5]

It appears to us that lignin perhaps is not fundamentally different from other biopolymers such as proteins, nucleic acids and polysaccharides in which the type of linkages and order of units are regulated. MALDI-MS of birch lignin indicates that the structure of this polymer is composed of repetitive units, with a molecular mass in the range 600–1800 Da.[8] In our hands, ESI-MS, ESI-MS/MS and MALDI-TOFMS of pine lignin polymer, isolated by the AVIDEL process, indicated the presence of a series of regular homopolymers with varying molecular masses.

The APCI mass spectra (negative and positive ion modes) were recorded with a Micromass Quattro quadrupole–hexapole–quadrupole mass spectrometer equipped with a APCI source and capable of analyzing ions up to m/z 4000. A personal computer (Compaq, PIII 500 MHz processor, running Windows NT 4, service pack 3) equipped with Masslynx 3.3 Mass Spectrometry Data System software was used for data acquisition and processing. The temperature of the APCI source was 100 °C and the APCI probe temperature was maintained at 400 °C. The operating voltage of the APCI corona was 1.80 kV and the high-voltage lens was set at 0.50 kV throughout the whole operation. APCI mass spectra were recorded with a cone voltage of 20 V. MS/MS experiments were conducted on the same instrument. Fragment ion and precursor ion spectra of mass selected ions were induced by collision with argon in the (r.f.-only) hexapole. MALDI-TOF mass spectra were acquired on a PerSeptive Biosystems Elite-STR instrument equipped with delayed extraction technology using α-cyano-4-hydroxycinnamic acid as matrix.

Acknowledgements

Yours,

JOSEPH H. BANOUB[1]* and MICHEL DELMAS[2]

[1] Department of Fisheries and Oceans, Science, Oceans and Environment Branch, St. John's, Newfoundland, A1C 5X1 and Biochemistry Department, Memorial University, St. John's, Newfoundland, Canada

[2] École Nationale Superieure des Ingénieurs en Arts Chimiques et Technologiques, Institut National Polytechnique de Toulouse, 118 route de Narbonne, 31077 Toulouse Cedex 4, France

References

1. Page DH. A note on the cell-wall structure of softwood tracheids. *Wood Fiber* 1976; **7**: 246.
2. Karlsson O, Pettersson B, Westermark U. The use of cellulases and hemicellulases to study lignin–cellulose as well as lignin–hemicellulose bonds in Kraft pulps. *J. Pulp Pap. Sci.* 2001; **27**: 196.
3. Salmén L, Olsson AM. Interaction between hemicelluloses, lignin and cellulose: structure–property relationships. *J. Pulp Pap. Sci.* 1998; **24**: 99.
4. Merdy P, Guillon E, Dumonceau J, Aplincourt M. Characterization of a wheat straw cell wall residue by various techniques. A comparative study with a synthetic and an extracted lignin. *Anal. Chim. Acta* 2002; **459**: 133.
5. Monties B. Novel structures and properties of lignins in relation to their natural and induced variability in ecotypes, mutants and transgenic plants. *Polym. Degrad. Stab.* 1998; **59**: 53.
6. Monties B. In *Methods in Plant Biochemistry*, Harbone JB (ed.). Academic Press: London, 1989; 113.
7. Freudenberg K, Neish AC. *Constitution and Biosynthesis of Lignin*. Springer: Berlin, 1968.
8. Metzger JO, Bicke C, Faix O, Fuszynski W, Angermann R, Karas M, Struptat K. Matrix-assisted laser desorption mass spectrometry of lignins. *Agnew Chem. Int. Ed. Engl.* 1992; **31**: 762.
9. Hortling B, Ranua M, Sundquist J. Investigation of the residual lignin in chemical pulps: part 1: enzymatic hydrolysis of the pulps and fractionation of the products. *Nordic Pulp Pap. Res. J.* 1990; **5**: 33.
10. Fachuang L, Ralph J. Derivitization followed by reductive cleavage (DFRC method), a new method for lignin analysis: protocol of analysis of DFRC monomers. *J. Agr. Food Chem.* 1997; **45**: 2590.
11. Faix O, Meier D, Fortmann L. Thermal degradation products of wood. Gas chromatographic separation and mass spectrometric characterization of monomeric lignin derived products. *Holz Roh Werkst.* 1990; **48**: 281.
12. Sonoda T, Ona T, Yokoï, Ishida Y, Ohtani H, Tsuge S. Quantitative analysis of detached lignin monomer composition by pyrolysis–gas chromatography combined with preliminary acetylation of samples. *Anal. Chem.* 2001; **73**: 5429.
13. Delmas M, Avignon G. Procédé de production de pâte à papier lignines sucres par fractionnement de matière végétale lignocellulosique en milieu acido formique/acetique. *Brevet Fr.* 1997; **97**: 13 658.
14. Lam HQ, Le Bigot Y, Delmas M, Avignon G. A new procedure for the destructuring of vegetable matter at atmospheric pressure by a catalyst/solvent system of formic acid/acetic acid. *Ind. Crops Prod.* 2001; **14**: 139.
15. Pasch H, Pizzi A. Considerations on the macromolecular structure of chestnut ellagitannins by matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *J. App. Polym. Sci.* 2002; **85**: 429.
16. Pasch H, Pizzi A, Rada K. MALDI-TOF mass spectrometry of polyflavanoid tannins. *Polymer* 2001; **42**: 7531.