ARTICLE

OBC
www.rsc.org/obc

The formation of β–β structures in lignin biosynthesis—are there two different pathways? †

Liming Zhang,* Gunnar Henriksson and Göran Gellerstedt
Department of Fibre and Polymer Technology, KTH, Royal Institute of Technology, SE-100 44 Stockholm, Sweden.
E-mail: liming@pmt.kth.se; ghenrik@pmt.kth.se; ggell@pmt.kth.se; Fax: +46-8-7906166; Tel: +46-8-7906165

Received 9th June 2003, Accepted 3rd September 2003
First published as an Advance Article on the web 17th September 2003

Based on results from 2D NMR studies, both pinoresinol and secoisolariciresinol structures were found to be present in native lignin from spruce wood as well as in spruce kraft lignin and residual kraft pulp lignin. These two structures constitute the major types of β–β inter-unit linkages present in spruce lignin, but their formation in the lignin polymer may follow different pathways leading to their different bonding patterns with the rest of the lignin polymer. The mechanisms involved are discussed.

Introduction

Lignin is a phenolic biopolymer formed through oxidative radical coupling of monolignols in the cell wall of vascular plants. Lignin macromolecules consist of phenylpropanoid units that are connected by various types of ether and carbon–carbon linkages, mainly including β–O–4, β–5, β–β, β–1, 5–5–O–4 and 4–O–5 types.[1–3] The formation of most of these well-recognized structures in lignin polymers can theoretically be explained by the mechanism of lignin biosynthesis, *i.e.* the oxidative radical coupling mechanism.

There are still controversies concerning the occurrence of the β–β structure in spruce lignin. The pinoresinol structure (**I**) (Fig. 1) has been well established through studies with NMR[4–6] as being part of spruce lignin. Nevertheless, degradation studies with β–O–4 cleavage methods, including thioacidolysis[7] and acid hydrolysis,[8] have failed to detect the existence of **I** or **II** (the expected product from **I** on thioacidolysis followed by reduction) in spruce lignin, but have always observed the presence of divanillyltetrahydrofuran (**IV**). On the other hand, pinoresinol-related structures have tentatively been found in trimeric and tetrameric thioacidolysis products from spruce lignin, *i.e.*, the pinoresinol structure seems always to occur together with a 5–5 or a 4–O–5 linked structure.[9]

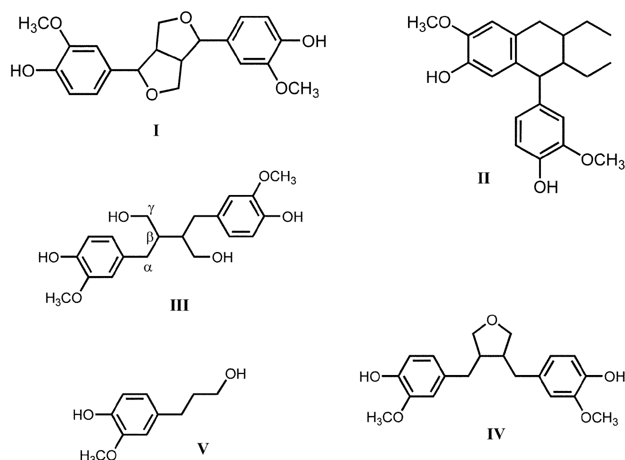


Fig. 1 Lignin structures mentioned in this work.

† Electronic supplementary information (ESI) available: NMR spectra. See http://www.rsc.org/suppdata/ob/b3/b306434d/

Based on studies with model compounds[8] and ^{1}H NMR analysis, the potential existence of the secoisolariciresinol structure (**III**) in spruce lignin has been suggested previously by Lundquist.[10–14] Model compound **III** was found to be converted into **IV** during acid hydrolysis.[8] The proton NMR signals observed on a spruce lignin sample[10–14] and assigned to **III** were, however, very ambiguous since they could easily be confused with the proton signals from the dihydroconiferyl alcohol structure (**V**). These NMR signals were finally attributed primarily to the latter.[12,13] Later, the presence of **V** in softwood lignin has been confirmed by mild hydrolysis[15] and by 2D NMR studies.[12,16,17] To date, the presence of the secoisolariciresinol structure (**III**) in intact softwood lignin has not been confirmed by diagnostic NMR experiments, and the origin of **IV** observed by thioacidolysis or by acid hydrolysis remains uncertain. The presence of secoisolariciresinol (**III**) in the mixture of softwood lignans is, however, well established[18] and, therefore, the structures **III** and **IV** have been suspected to be of lignan origin but present in lignin as "impurities".[13,19] Since the structure **IV**, isolated from lignin, was found to be optically inactive,[8,13] in contrast to the optically active **III** isolated as a softwood lignan,[18] this view has not found further support.

Even though results from thioacidolysis have clearly shown that pinoresinol degradation products (**II**) could hardly be released from softwood lignin through cleavage of β–O–4 bonds,[7] isolated β–β structures, produced by other β–O–4 cleavage methods, such as the DFRC method, have been automatically assigned to the presence of the pinoresinol structure.[20] Thus, wood chemists working in the area of lignin chemistry are generally unaware or unconvinced about the existence of **III** in lignin.

In the present paper, the occurrence of β–β structures in spruce lignins has been studied by 2D NMR techniques. Thereby, the presence of structures of type **I** as well as **III** could be unambigously assigned.

Results and discussion

A total of four lignin samples obtained from mature spruce wood were studied; two native lignins obtained after successive extractions of purified wood powder[21,22] and two lignins from a kraft cook of spruce.[23–25] The lignins were further purified by extraction with organic solvents in order to remove all types of low molecular weight material possibly present in the samples. The samples were subjected to NMR analysis employing the 2D NMR experiments HSQC, HSQC-DEPT and HSQC-

DOI: 10.1039/b306434d

TOCSY. Both unacetylated and acetylated samples were analysed. As indicated in Fig. 2, all lignin samples were found to contain both the pinoresinol (**I**) and the secoisolariciresinol structure (**III**). In addition, the dihydroconiferyl alcohol (**V**) structure was found.

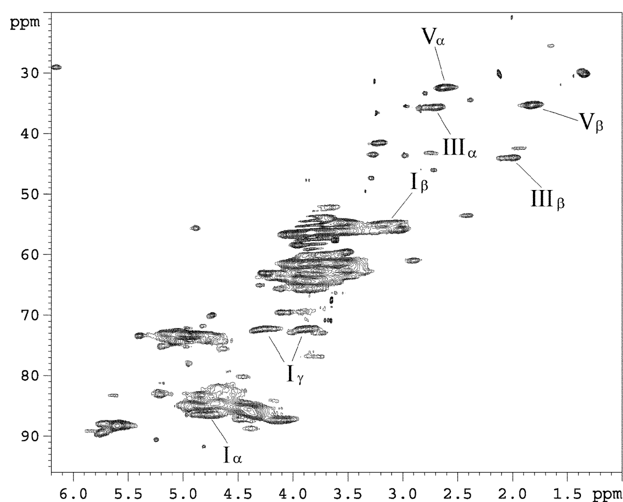


Fig. 2 HSQC spectrum of spruce milled wood lignin (unacetylated) and identification of correlation cross peaks originating from pinoresinol (**I**), secoisolariciresinol (**III**) and dihydroconiferyl alcohol (**V**) structures.

The NMR signals from **I**[5,6] and **V**[12,15,17] have been assigned before and the locations of these cross signals are well known. In the HSQC-DEPT spectra of the acetylated lignin samples, signals for the CH_2–α (2.64/35.2, 2.78/35.2 ppm) and the CH–β (2.20/41.0 ppm) groups of **III** were clearly observed (Fig. 3, Spectrum 1). By running a high-resolution HSQC experiment, signals for the CH_2–γ (4.01/64.8, 4.20/64.8 ppm) of **III** could also be unambiguously identified (Fig. 3, Spectrum 2). Furthermore, an HSQC-TOCSY experiment (Fig. 3, Spectrum 3), established the consecutive connections between the CH_2–α, CH–β and CH_2–γ groups of **III**. All the chemical shift data observed here are in agreement with previously published NMR data of **III**.[13,26] The observed NMR signals for **III** were found to have proton line widths as broad as the rest of the NMR signals from the same lignin sample, indicating that these secoisolariciresinol structures are part of the lignin macromolecule. The abundance of the secoisolariciresinol structure was estimated to be ~1.0 unit per 100 C9 in spruce lignin and 1.0–1.5 units per 100 C9 in kraft and kraft pulp residual lignin according to quantitative ^{13}C NMR analyses and quantitative HSQC NMR analysis.[27,28] In comparison, the abundance of the pinoresinol structure (**I**) was found to be ~2 units per 100 C9 in all of the lignin samples.

The above results show that both pinoresinol (**I**) and secoisolariciresinol (**III**) structures are part of the spruce lignin structure. They can, however, be linked to the lignin macromolecules through different types of linkages. **III** might be partly connected to the lignin polymer through two β–O–4 bonds, which can be cleaved during thioacidolysis[7,9] or during acid hydrolysis,[8] leading to the release of **III**, which, subsequently, can be converted into the vanillyltetrahydrofuran (**IV**) structure under acidic conditions.[8] Pinoresinol (**I**), on the other hand, must be connected to the lignin polymer through acid-resistant bonds, as has been suggested previously.[8,14] Such a type of linkage would make **I** inaccessible for both thioacidolysis and acid hydrolysis. Therefore, the degradation product **III**, observed in the DFRC analysis of lignin, may actually originate from β–O–4-linked **III** in the lignin polymer, rather than from the pinoresinol structure (**I**) as has been suggested.[20,26]

The assumption that β–β structures of the pinoresinol type might be connected to the lignin polymer through at least one acid-resistant linkage is in agreement with the mechanism of end-wise polymerization of lignin biosynthesis. A possible mechanism is shown in Scheme 1 and involves a growing polymer chain with a 4–O–5 linkage to the last phenylpropanoid unit (ring **A**). This may couple to a new monomer (ring **B**) through a β–β linkage, thus giving rise to a pinoresinol structure that is linked to the lignin polymer through an acid-resistant bond. As an alternative, the linkage between the lignin polymer and the **A**-ring (Scheme 1) can be of the 5–5 type, as has been suggested previously.[14]

The fact that the pinoresinol structure cannot be detected in spruce lignin through degradation methods involving the cleavage of the β–O–4 bonds demonstrates that dimerization of coniferyl alcohol radicals to pinoresinol structures is not an important reaction during lignin biosynthesis in native spruce wood. Such a reaction *in vitro*, initiated by peroxidase, has been shown to produce a significant amount of pinoresinol structures.[29,30] A portion of these would have been linked through β–O–4 linkages to the lignin polymer and consequently, they should have been detectable by acid hydrolysis or by thioacidolysis. The presence of pinoresinol structures (detectable by

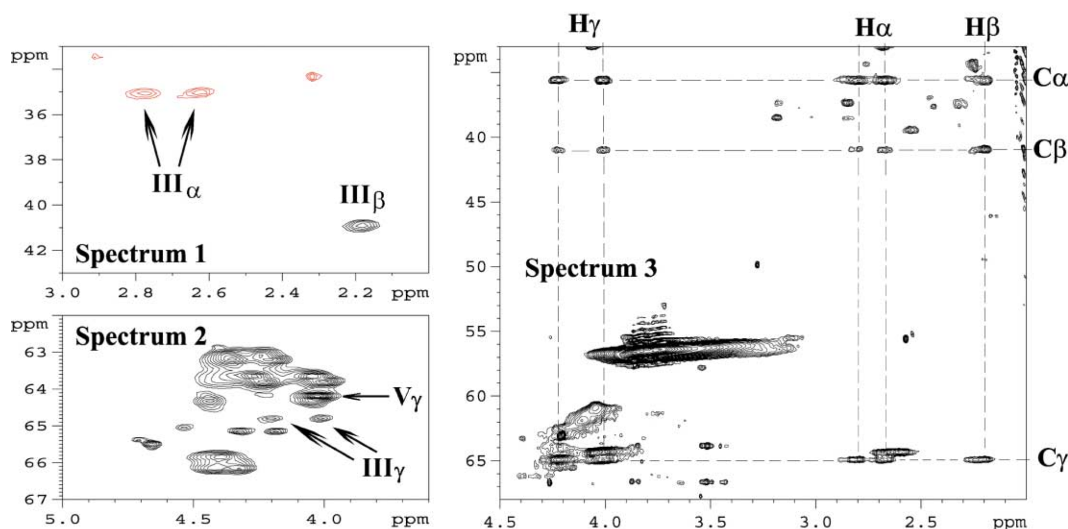
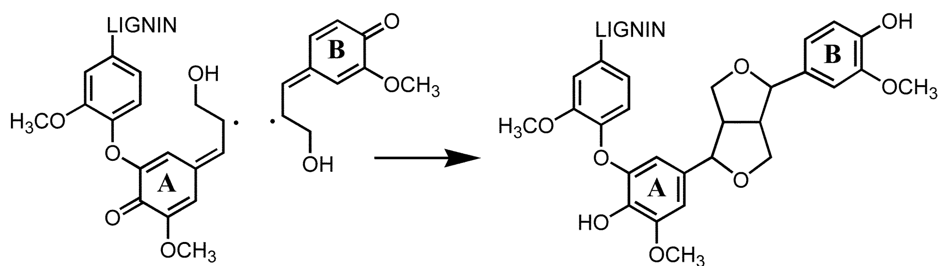
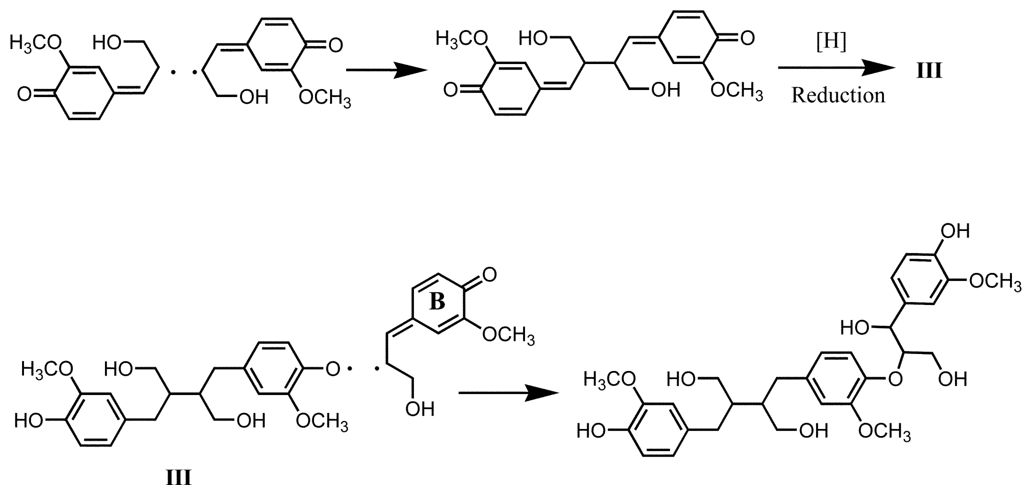


Fig. 3 Observation of the secoisolariciresinol (**III**) structure in acetylated lignin samples by HSQC, HSQC-DEPT and HSQC-TOCSY NMR experiments. Spectrum 1: HSQC-DEPT spectrum of milled wood lignin showing negative (red) signals for the CH_2–α and positive signals for the CH–β of **III**. Spectrum 2: High resolution HSQC spectrum of milled wood lignin showing signals for CH_2–γ of **III**. Spectrum 3: HSQC-TOCSY spectrum of kraft lignin showing the consecutive connections between α–, β– and γ–CH_n groups of **III**.



Scheme 1 A proposed mechanism for the formation of the pinoresinol (β–β) structure in spruce lignin.



Scheme 2 A proposed mechanism for the incorporation of the secoisolariciresinol (β–β) structure into spruce lignin.

thioacidolysis) in guaiacyl DHP[30] indicated that the initially formed pinoresinol dimers would indeed react with coniferyl alcohol radicals through β–O–4 coupling.

The presence of optically inactive secoisolaricinol (**III**) structures in lignin[8,13,20] is difficult to explain based only on the theory of oxidative radical coupling. Since **III** can be released from the lignin polymer through cleavage of the β–O–4 bonds, this structure is probably formed as a dimer initially through coupling of two coniferyl alcohol radicals followed by reduction (Scheme 2). On subsequent further radical coupling, the dimeric **III** should preferentially be incorporated into lignin polymer through β–O–4 linkages. One possibility might be that the quinonemethide intermediate formed in the radical coupling reaction is reduced. Similar quinone reductions have been observed to occur by several types of reductive enzymes,[31] and also possibly by other biomolecules. Thus, the nature of the β–β linkages in spruce lignin suggests that the lignin polymerization may be carried out in two modes, one of a dimerization type (monolignol coupling to monolignol) in a reductive environment during the initial stage of lignin formation, and one of an end-wise type (monlignol coupling to a phenol in the growing polymer) in an oxidative environment. The reductive environment should exist early during the cell wall synthesis, when the lignification initiating sites are formed. Although reduction seems to take place on a minor portion of the coniferyl alcohol units during lignin formation (*cf.* **V**), more information is needed before any firm conclusions about the role of reduction, its selectivity and mechanism can be made.

Experimental

Milled wood lignin (MWL) was prepared from pre-extracted wood meal of a mature (50–60 year old) Norway spruce (*Picea abies*) according to a standard method.[21] The yield of MWL was about 15% of the total lignin content in the wood material. After the extraction of the first MWL fraction, the residual wood powder was treated twice with 2 mL of Econase (a crude cellulase enzyme, Röhm Enzyme Finland Oy, Rajamäki, Finland) for two days, each time at 40 °C and at pH = 5, and subsequently washed and freeze-dried. The wood meal was again treated with steel ball milling for one day and extracted.[22] After purification, an "enzyme MWL" was isolated in 35% yield of the original lignin content in wood.

From a kraft cook of spruce, the corresponding dissolved lignin was isolated from the liquor.[23] Since this lignin has been found to contain significant amounts of low molecular weight material, successive extractions with methylene chloride, ethyl acetate and acetone were performed. In each extraction step, the kraft lignin (2 g) was suspended in the solvent (100 mL), stirred at room temperature overnight and recovered by filtration. The residual kraft pulp lignin was isolated from the pulp (pre-extracted with acetone) by acid hydrolysis in dioxane–water.[25]

In order to remove any traces of paramagnetic metal ions, each lignin sample (1 g) was dissolved in acetone–water (9 : 1, 10 mL) and treated overnight at room temperature with an ion exchange resin (Amberlite IRC-718, 1 g, pre-washed with acetone–water). The resin was filtered away and water (20 mL) added to the lignin solution. After removal of acetone by evaporation, the lignin precipitate was washed with water and freeze-dried. Acetylation of lignin was carried out as described in ref. 32. Acetylated lignin samples (~110 mg) were dissolved in acetone-d6 (600 μL) and unacetylated lignin samples (~100 mg) were dissolved in acetone-d6–D_2O (600 μL, 5 : 1). The lignin solutions were transferred into 5 mm NMR tubes. TMS was used as the reference for chemical shift values.

All NMR experiments were run on a Bruker Avance 400 MHz instrument. 2D NMR experiments were performed with a proton–carbon selective inverse probe equipped with a Z-gradient coil. Standard Bruker pulse programs were used in all experiments. Pulse lengths of the NMR probes were carefully calibrated and optimum tuning and matching as well as shimming of the NMR probe was performed with each run.

The HSQC-DEPT spectrum was acquired using the Bruker pulse program "inviedgstp", with a spectral window of 12.8 ppm in F2 and 150 ppm in F1 and with 2K × 1K increments.

An acquisition time 0.2 s was recorded. The spectral centre in the proton dimension was set at 6 ppm and in the ^{13}C dimension at 93 ppm. The delay time between the pulses was 0.5 seconds and 128 scans were acquired for each increment. HSQC spectra (pulse program "invieagssi") were generally acquired with a spectral window of 12.8 ppm in F2 and 150 ppm in F1 with increments 2 K × 0.5 K, giving an acquisition time of 0.2 s in F2. The spectral centre was set at 5.3 ppm in the ^{1}H and at 91 ppm in the ^{13}C dimension. Typical parameters for data acquisition included; a delay time of 2 s, an average coupling constant of 150 Hz, 32 scans per increment resulting in a total experimental time of 14 h. In the HSQC experiment, maximum sensitivity enhancement was set for the methine signals by setting the delay (D24) to 1/4J. The high resolution HSQC spectrum was acquired by using the Bruker pulse program "invieagssi", with a spectral window of 6.4 ppm in F2 and 60 ppm in F1 and with 1 K × 1 K increments. The spectral centre in the proton dimension was set at 4 ppm and in the ^{13}C dimension at 70 ppm. The delay time between pulses was 1 second, an acquisition time 0.2 s was recorded and 32 scans were acquired for each increment. The heteronuclear coupling constant value was set at 150 Hz. The acquired 2D NMR data sets were processed with 2 K × 1 K data points for the HSQC, HSQC-DEPT and HSQC-TOCSY spectra respectively and with 2 K × 2 K data points for the high resolution HSQC spectrum, using the $\pi/2$ shifted sine bell window function in both dimensions. The quantitative ^{13}C NMR experiments were performed with a 5 mm broadband probe and with an inverse gated proton-decoupling sequence. A pulse angle of 90° and a delay time of 12 s between pulses were applied during data acquisition.[33] Quantification of the HSQC spectra were carried out according to our recent developed method.[27,28]

Acknowledgements

This work was supported by the Commission of the European Communities, FAIR CT98 3460 – Bleachability of Alkaline Pulps to LZ, and by the Swedish Research Council for Environmental, Agricultural Sciences and Spatial Planning (FORMAS), 23.2/2002-1128 to GH and LZ.

References

1 E. Adler, *Wood Sci. Technol.*, 1977, **11**, 169–218.
2 P. Karhunen, P. Rummakko, J. Sipilä, G. Brunow and I. Kilpeläinen, *Tetrahedron Lett.*, 1995, **36**, 4501–4504.
3 L. Zhang and G. Gellerstedt, *Chem. Commun.*, 2001, 2744–2745.
4 H. H. Nimz and H.-D. Lüdemann, *Holzforschung*, 1976, **30**, 33–40.
5 R. M. Ede and G. Brunow, *J. Org. Chem.*, 1992, **57**, 1477–1480.
6 J. Ralph, J. Marita, S. A. Ralph, R. D. Hatfield, F. Lu, R. M. Ede, J. Peng, S. Quideau, R. F. Helm, J. H. Grabber, H. Kim, G. Jimenez-Monteon, Y. Zhang, H. J. G. Jung, L. L. Landucci, J. J. MacKay, R. R. Sederoff, C. Chapple, and A. M. Boudet, Solution-State NMR of Lignins, ed. Dimitris S. Argyropoulos, *Advances in Lignocellulosics Characterization*, TAPPI Press, Atlanta, GA, 1999, ch. 3, pp. 55–108.
7 C. Lapierre, B. Pollet, B. Monties and C. Rolando, *Holzforschung*, 1991, **45**, 61–68.
8 K. Lundquist, *Acta Chem. Scand.*, 1970, **24**, 889–907.
9 H. Önnerud, On the structure of native softwood and hardwood lignins, PhD thesis, Royal institute of Technology, KTH, Department of pulp and paper technology, 2002, Stockholm.
10 K. Lundquist, *Acta Chem. Scand.*, 1980, **B 34**, 21–26.
11 K. Lundquist, *Acta Chem. Scand.*, 1981, **B 35**, 497–501.
12 K. Lundquist and K. Stern, *Nordic Pulp Pap. Res. J.*, 1989, **4**, 210–213.
13 K. Lundquist and R. Stomberg, *Holzforschung*, 1988, **42**, 375–384.
14 K. Lundquist, *Nordic Pulp Pap. Res. J.*, 1992, **7**, 4–8.
15 G. Gellerstedt and L. Zhang, *Nordic Pulp Pap. Res. J.*, 1992, **7**, 75–80.
16 N. Fukagawa, G. Meshitsuka and A. Ishizu, *J. Wood Chem. Technol.*, 1991, **11**(3), 373–396.
17 J. Ralph, J. J. MacKay, R. D. Hatfield, D. M. O. Malley, R. W. Whetten and R. R. Sederoff, *Science*, 1997, **277**, 235–239.
18 A. Sakakibara, T. Sasaya, K. Miki and H. Takahashi, *Holzforschung*, 1987, **41**, 1–11.
19 Y. Z. Lai and K. V. Sarkanen, in *Lignins: Isolation and structural studies*, eds. K. V. Sarkanen and C. H. Ludwig, Wiley Interscience, New York, 1971, p 230.
20 J. Ralph, J. Peng, F. Lu, R. D. Hatfield and R. F. Helm, *J. Agric. Food Chem.*, 1999, **47**, 2991–2996.
21 A. Björkman, *Sven. Papperstidn.*, 1956, **59**, 477–485.
22 N. A. More, E. C. Johansson, I. D. Suckling and R. M. Ede, *Proceedings of 9th* International Symposium on Wood and Pulping Chemistry, p 73–1, vol. 1, Montreal, Canada, 1997.
23 G. Gellerstedt and E.-L. Lindfors, *Holzforschung*, 1984, **38**, 151–158.
24 G. Gellerstedt and E.-L. Lindfors, *Nordic Pulp Pap. Res. J.*, 1987, **2**(2), 71–75.
25 G. Gellerstedt, J. Pranda and E-L. Lindfors, *J. Wood Chem. Technol.*, 1994, **14**(4), 467–482.
26 J. Peng, F. Lu and J. Ralph, *J. Agric. Food Chem.*, 1998, **46**, 553–560.
27 L. Zhang and G. Gellerstedt, *Proceedings of 6th European Workshop on Lignocellulosics and Pulp*, Bordeaux, France, pp. 7–10, 2000.
28 L. Zhang, G. Gellerstedt, Manuscript in preparation.
29 K. Freudenberg, in *Constitution and biosynthesis of lignin*, eds K. Freudenberg and A. C. Neish, p. 86, Springer-Verlag, Berlin, 1968.
30 G. Jacquet, B. Pollet, C. Lapierre, C. Francesch, C. Rolando and O. Faix, *Holzforschung*, 1997, **51**, 349–354.
31 S. Ghisla and V. Massay, *Eur. J. Biochem.*, 1989, **181**, 1–17.
32 K. Lundquist, Proton (^{1}H) NMR spectroscopy, in *Methods in Lignin Chemistry*, eds. S. Y. Lin and C. W. Dence, Springer-Verlag, Berlin, Heidelberg, p. 242, 1992.
33 D. R. Robert and C.-L. Chen, *Holzforschung*, 1989, **43**, 323–332.