Vibrational Structure of Selected Compounds Derived from Biomass: Lignin Dimers, Selected Aldopentoses and Aldohexoses

Izabela Czekaj and Natalia Sobuś

Institute of Organic Chemistry and Technology, Faculty of Chemical Engineering and Technology, Cracow University of Technology, Warszawska 24, Kraków 31-155, Poland

Abstract: The production of chemicals from biomass is a very challenging process due to its diverse chemical composition. Lignin, cellulose and hemicellulose are the three main biopolymers of wood biomass, with cell walls of plant origin. Lignin has been chosen for the present studies due to its range of different linkages and structures. The present work involved a computational study of the most dominant lignin dimers and their vibrational structures, based on the Density Functional Theory method. Full geometry optimization of the compartments used the StoBe code with cluster model and non-local functional (RPBE) approach. The calculations of the vibrational frequencies were performed with harmonic approximations as well as an anharmonicity fit in the Morse potential function, as implemented in the StoBe code. In the case of lignin, the calculations included three different precursors based on: coumaryl alcohol, coniferyl alcohol and sinapyl alcohol. To represent the cellulose and hemicellulose derivatives, selected aldopentoses and aldohexoses (arabinose, xylose, glucose, galactose, and mannose) were considered. Presented here are the theoretical investigations for a variety of biomass derived compounds, to give the possibility of obtaining a theoretical VBD (Vibrations Basis Database) for experimental spectra interpretation. Such a database could be further used in the preliminary composition assessment of biomass derived substrates, which will be discussed here in more detail.

Key words: Wood biomass, lignin, cellulose, vibrational structure, DFT (density functional theory).

1. Introduction

A major challenge for society in research is how to build fuels and chemical intermediates from available and renewable materials that do not compete with food crops in terms of water or fertilizer, such materials including wood biomass or agricultural waste. The urgent need to achieve the more sustainable production of chemicals from renewable feedstock, including biomass, has resulted in intense research. However, the production of chemicals from biomass is a very challenging process due to the diversity in chemical composition. Biomass from woody plant materials (e.g. corn and wood wastes, grass etc.) is a promising source of biorenewable feedstock [1]. However, the use of biomass for feedstock production requires a sequence of chemical conversions, which means overcoming the primary issue: carbohydrates. Recent global efforts have focused on plant cell walls (collectively known as lignocellulosic biomass) in order to create biofuels for transportation needs [2-4]. The first stage is to break down the lignocellulose nanostructure of the biomass. Lignocellulose is built from cellulose crystals embedded in hemicellulose sugars and lignin polyphenols [5]. Fig. 1 shows the main compounds derived from cellulose, hemicellulose and lignin. Cellulose is the most abundant and important constituent in plant cell walls. Its crystal structure and complex network shared with other carbohydrate polymers are key factors determining the mechanical strength and degradability of plant cell walls [6, 7]. Another constituent is lignin, which is commonly removed from wood meal by the paper

Corresponding author: Izabela Czekaj, D.Sc., Ph.D., research field: catalysis, computational chemistry, chemical technology.
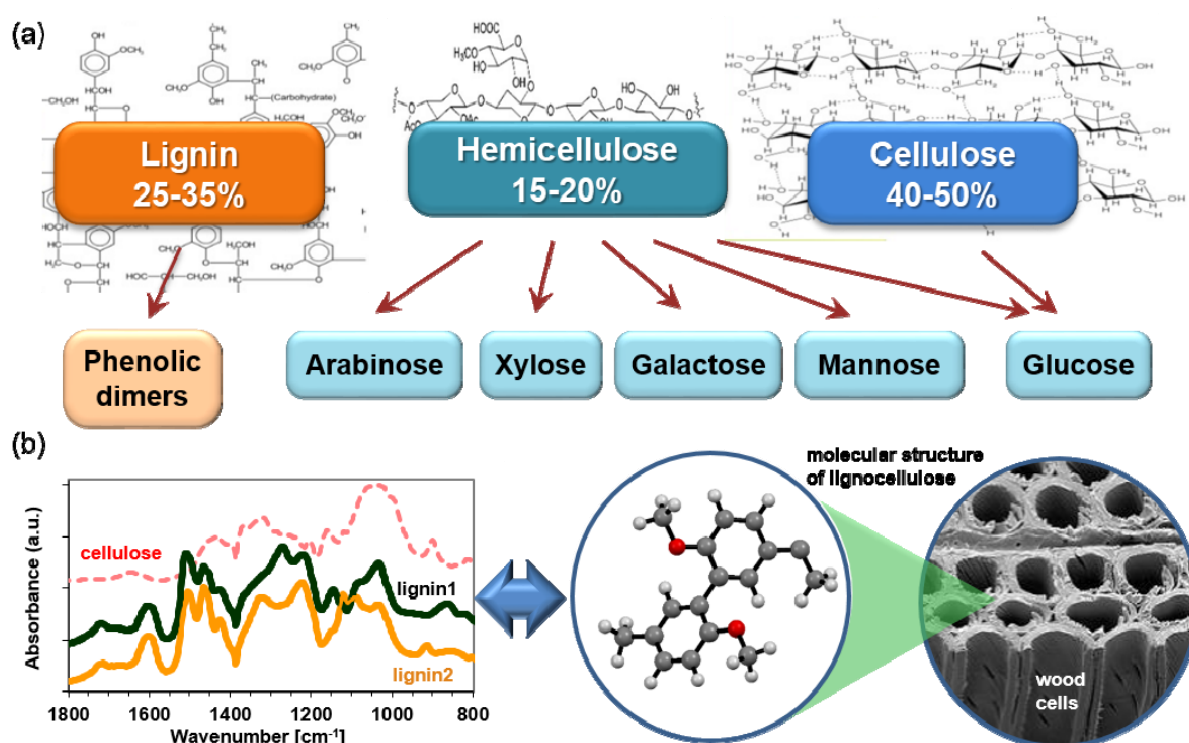
Fig. 1 A schematic representation of research on lignocellulose: (a) the main compounds derived from cellulose, hemicellulose and lignin, (b) link between experimental and theoretical methods.

industry. Lignin is used as a base material to manufacture biopolymers by conventional polymer techniques. Another field of application is agriculture; lignin products can form the basis of delayed-action fertilizers, which release nutrients slowly or serve as support for the production of humus in the soil. In view of its chemical composition, lignin is a raw material rich in aromatic (phenolic) compounds with a high added value. In 2004, the pulp and paper industry alone extracted 50 million tons of lignin, although only approximately 2% of the lignin available is used commercially, with the remainder used as a low-value fuel [9]. An especially difficult task is to extract lignin from wood, and there are a number of explanations for this [10]: (i) strong covalent bonds exist between lignin and carbohydrates, (ii) lignin has a high molecular weight and may form a three-dimensional network [11], and (iii) hydrogen bonds and physical phenomena such as a solid solution may be involved in the retention of lignin in the wood-fiber wall. There have been several

electronic structure investigations into lignin and lignocellulose [12-14]. Cellulose builds up the structural substance of unlignified cell walls and consists of fibrous macromolecules based on anhydroused *d*-glucose units. Single polysaccharide chains are hydrogen bonded and build microfibrils, which are resistant to hydrolysis. The high degree of polymerization causes the orientation, elasticity and large tensile strength of this carbohydrate. Hemicellulose is also a polysaccharide (such as arabinose, glucose, galactose, mannose and xylose) and consists of different monomers, which together create a random, amorphous structure [15-17].

More detailed and integrated studies into the differences between the various vibration structures of biomass-derived compounds, as well as comparisons between experimental and theoretical vibration studies of lignocellulose, should help us to achieve new understanding in the use of biomass for certain chemical processes (Fig. 1b) [18].

This study examines the design of a new theoretical approach to simulating the vibrational structure of lignin dimers as well as selected aldopentoses and aldohexoses from cellulose/hemicellulose, and an analysis of their vibrational spectra. At the present moment we concentrate on comparing only the calculated spectra with each other and the development of a vibrational basis set structure.

2. Computational Method

The electronic structure of the all lignin fragments was calculated by *ab initio* DFT (density functional theory) methods (StoBe code [19]). The generalized gradient corrected functionals according to Perdew, Burke, and Ernzerhof (RPBE) were used in order to account for electron exchange and correlation [20, 21]. All Kohn-Sham orbitals were represented by LCAOs (linear combinations of atomic orbitals) using extended contracted-Gaussian basis sets for the atoms [22, 23]. In the calculations for all reaction steps, the low spin states were found to be energetically favorable.

A DZVP (double zeta valence polarization) type was used for the orbital basis sets of O, C (621/41/1), and H (41). Auxiliary basis sets, such as (4.3:4.3) for O, C, N, and (41) for H, were applied to fit the electron density and the exchange-correlation potential.

The vibration frequencies of the adsorbed molecules were calculated by single point energy calculations for the optimized geometries. The calculations of the vibrational frequencies were performed with harmonic approximations as well as an anharmonicity fit in the Morse potential function, as implemented into the StoBe code [24]. Theoretical vibrational spectra, including realistic intensities, were obtained by convolution of the vibrational spectra of the individual adsorbates, applying Gaussian line-shapes. The frequencies were reported as obtained from the calculations, without scaling. The intensity of adsorption, A, of the peak in the IR spectrum was calculated as proportional to the square of the first derivative of the electric dipole moment, μ, of the molecule with respect to the normal-mode vibrational coordinates ξ (a combination of nuclear displacement coordinates from equilibrium geometry): $A = C|d\mu/d\xi_{\xi=0}|^2 = C|d\mu^{\xi}|^2$ [33].

2.1 Geometrical Model

The lignin structure consists of the three most typical alcohols: coniferyl, coumaryl, and sinapyl. First we built six types of dimers based on coumaryl alcohol, and Fig. 2 shows these dimer structures of lignin [12]: with ether linkage (arylglycerol β-aryl ether, β-O-4 and phenylpropane a-aryl ether, α-O-4), with a carbon-carbon linkage (1,2-diaryl propane, β-1, phenylcoumaran, β-5, and bifhenyl/dibenzodioxocin, 5-5) and resinol β-β-linked structure.

According to the literature [25], lignin mostly contains arylglycerol β-aryl ether (45-50%), bifhenyl/dibenzodioxocin (18-25%), 1,2-diaryl propane (9-12%) and phenylcoumaran (7-10%). Additionally, in the case of 1,2-diaryl propane, β-1 and phenylcoumaran, β-5, other forms were considered [26]: H-, G- and S-lignin (the names are directly connected with the subunits present in the lignin structure: guaiacylpropane (G), syringylpropane (S), phydroxyphenylpropane (H)). Fig. 3 shows all three considered types of lignin. In comparison, the coumaryl structure of certain methyl groups was exchanged into -O-CH$_3$. The arrows indicate the changes in the functional groups compared to the previous structure model (coumaryl in the case of G-lignin and coniferyl in the case of S-lignin).

Finally, selected aldopentoses and aldohexoses as derivatives from cellulose and hemicellulose were considered (Fig. 4): arabinose, D- and L-xylose, glucose, D- and L-galactose, as well as D-and L-mannose.

3. Results and Discussion

The lignin vibrational structure strongly depends on the source of wood, which is sketched in the FTIR spectra of lignin (Fig. 1, left part). The most interesting
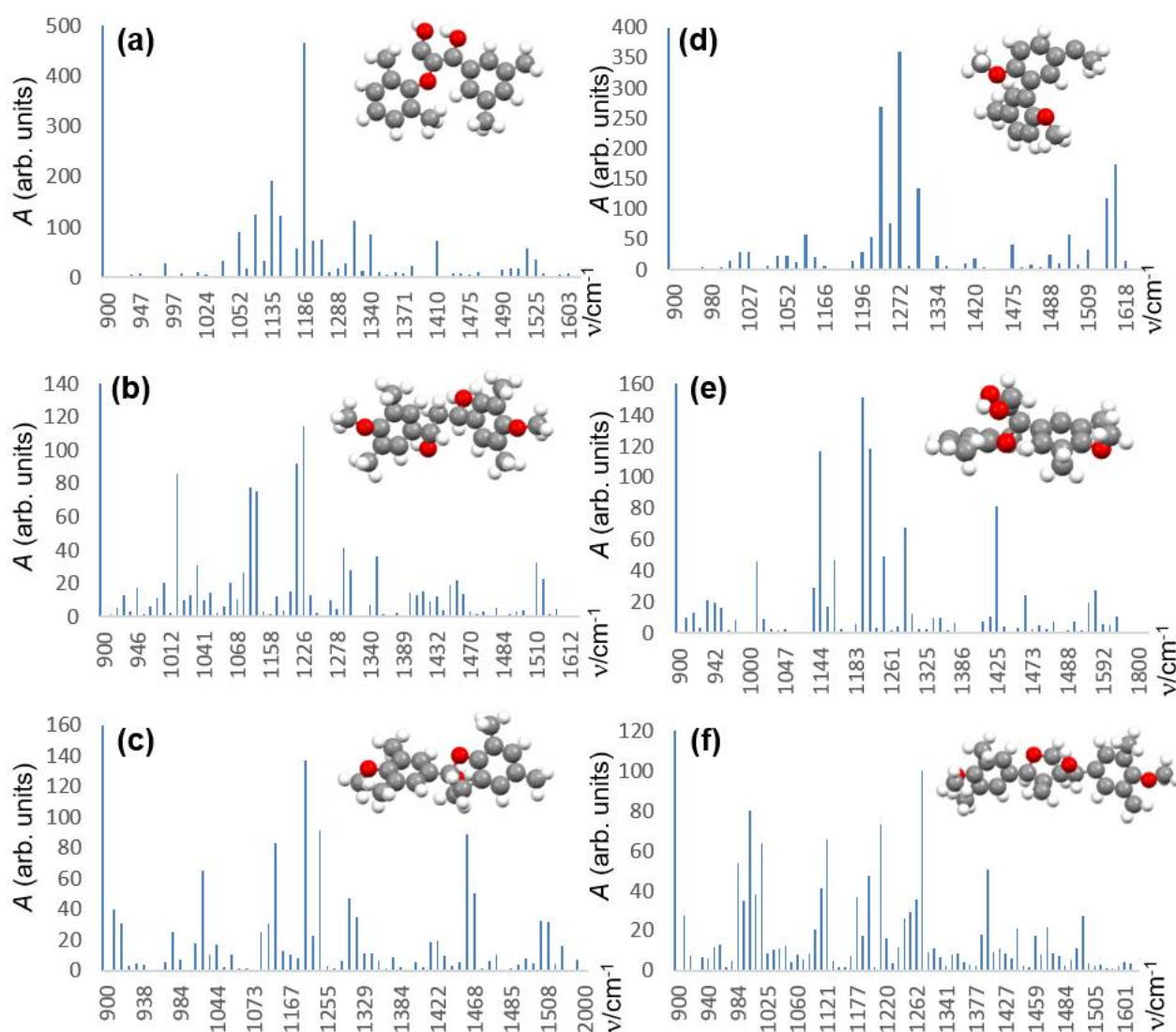
Fig. 2 The calculated frequencies and intensities of normal vibrational modes of lignin dimer models—coumaryl representation: (a) arylglycerol-aryl ether β-O-4; (b) 1,2-diaryl propane-1, β-1; (c) phenylcoumaran, β-5; (d) bifhenyl/dibenzodioxocin, 5-5, (e) phenylpropane-aryl ether, α-O-4; (f) β-β-linked structure.

and variable region is that visible between 900 and 2,000 cm^{-1} [15]. All the observed peaks formed part of the fingerprint region of the various functional groups present in the lignin structure and hence were very difficult to match and their definite assignment difficult to achieve.

Therefore, the theoretical spectra were considered to help to assign vibrations of certain groups. Fig. 2 shows the calculated frequencies and intensities of normal vibrational modes of lignin dimer models in the frequency range 900-2,000 cm^{-1}. The spectra were very diverse and therefore difficult to analyze and compare. However, we ascertained by way of systematic analysis of the frequencies, together with the changes in their intensity, for different dimers. Other than a number of bands that were similar and intense in all dimers (mostly in the region of 1,180-1,280 cm^{-1}), each dimer structure had fingerprint vibrations, with a high intensity specific for each dimer. This allowed us to distinguish specific dimer structures by using the calculated spectra from a basis set. The lignin vibration structure in the region between 900 and

Vibrational Structure of Selected Compounds Derived from Biomass: Lignin Dimers,
Selected Aldopentoses and Aldohexoses

15

2,000 cm^{-1} contains several bands which can be assigned to the vibrations of certain groups, according to the literature [27-29]. For example, the region between 1,730-1,760 cm^{-1} is assigned to C=O stretch vibrations (ν(C=O)). However, in the considered models these vibrations were not observed. The aromatic skeletal vibrations were responsible for the band in the range 1,500-1,550 cm^{-1}. The C-H deformation (δ(CH) symmetric and asymmetric) was observed in the range 1,370-1,450 cm^{-1}. The wagging vibrations, ω(CH$_2$), were suggested in the frequency range 1,300-1,350 cm^{-1}. Our calculations show that this region has a contribution from bending vibrations of O-H and the rocking vibration of the C-H groups.

The vibrations of the aromatic ring were observed around 1,300 cm^{-1}. The ether stretching vibrations, ν_{as}(C-O-C), in the range 1,150-1,180 cm^{-1} overlap the hydroxyl group stretching vibrations, ν(O-H), in the range 1,100-1,130 cm^{-1} as well as the guaiacyl C-H + syringyl C-H vibrations at 1,110-1,130 cm^{-1}. The ether group vibrations were observed in the case of β-O-4 dimer at 1,186 cm^{-1} and for α-O-4 dimer at 1204 cm^{-1}, which were slightly higher frequencies than the experimentally assigned ranges for the vibrations of these functional groups. These vibrations also have high intensities (Figs. 4a and 4e). The lower frequencies (1,134 cm^{-1} for β-O-4 and 1,144 cm^{-1} for α-O-4) were assigned to the C-H vibrations of the whole dimer, including the perypherial methyl groups.

Differences in the intensities of the calculated spectra for the H-, G- and S-lignin models were also observed. Fig. 3 presents the calculated frequencies and intensities of the normal vibrational modes of the lignin dimer models of 1,2-diaryl propane β-1 (left column) and phenylcoumaran, β-5 (right column) modified according to: coumaryl (H-lignin), coniferyl (G-lignin), and sinapyl (S-lignin) representation.

Increasing the number of C-O-CH$_3$ bonds resulted in a significant change in the intensities of certain bands and the presence of new frequencies. Therefore, it is expected that a more complex polymeric model of these three types of lignin will give a distinctive band structure.

The calculated IR bands of glucose, galactose and mannose (Figs. 4a, 4c and 4d) show similarities, due to the fact that galactose and mannose are isomers of glucose with different orientations of the C-4 and C-2 hydroxyl groups, respectively. Their IR spectra have similar regions in the ranges: 1,000-1,080 cm^{-1} including stretching and in-plane bending vibrations (ν(CC), ν(CO), β(COH)), 1,080-1,150 cm^{-1} dominated by stretching vibrations (ν(CO), ν(CC)) and 1,300-1,350 cm^{-1} including CH$_2$ wagging (ω(CH$_2$)). The most dominant bands for arabinose were found at the frequency with stretching (ν) and in-plane bending vibrations (β): 979 cm^{-1} (vas(COC), ν(CC), ν(CO), β(CCH)), 1,046 and 1,062 cm^{-1} (ν(CO), ν(CC), β(COH)) and 1,112 cm^{-1} (ν(CO), ν(CC)). The characteristic IR bands of L- and D-xylose are found at 972 (β(CCH)), 1,049 and 1,072/1,079 (ν(CO), ν(CC), β(COH)), 1,399 cm^{-1} (ω(CH$_2$)) cm^{-1}.

To summarize, comparing the calculated vibrational spectra of lignin dimers suggests a good potential to form basis sets of vibrations for the assignment of experimental spectra. The comparison with real experimental spectra of lignin has yet to be performed. First, the vibration basis set of lignin dimers is prepared. Second, the more complicated polymeric lignin model is formed, based on various configurations of previously investigated dimers. Third, the compared calculated spectra of polymeric lignin is then compared with the experimental spectra. In our opinion, such a scenario should guarantee a more accurate comparison between the calculated and experimental vibrational spectra.
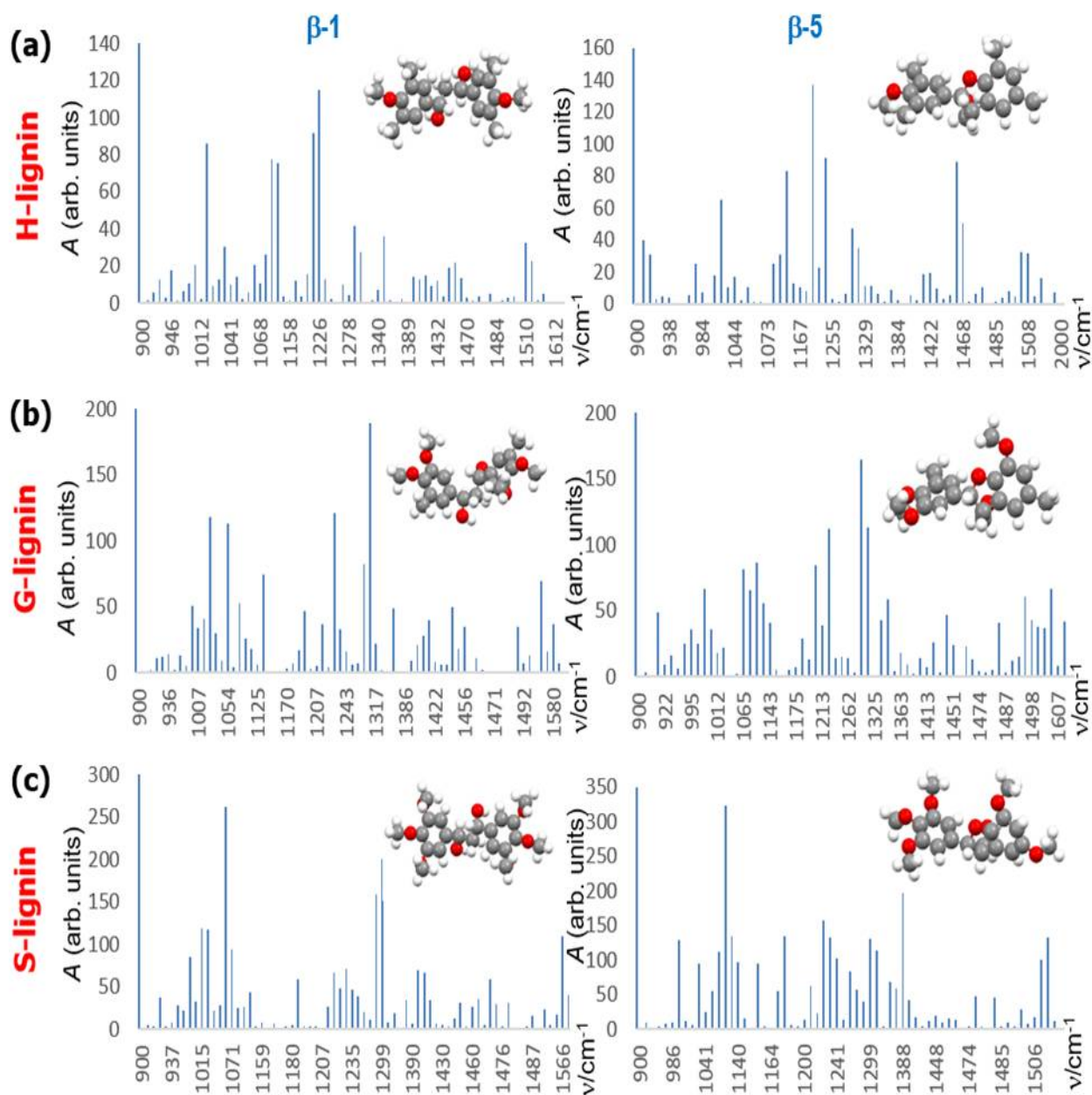
Fig. 3   The calculated frequencies and intensities of normal vibrational modes of lignin dimer models of 1,2-diaryl propane β-1 (left column) and phenylcoumaran, β-5 (right column) modified according to: (a) coumaryl (H-lignin), (b) coniferyl (G-lignin), and (c) sinapyl (S-lignin) representation.
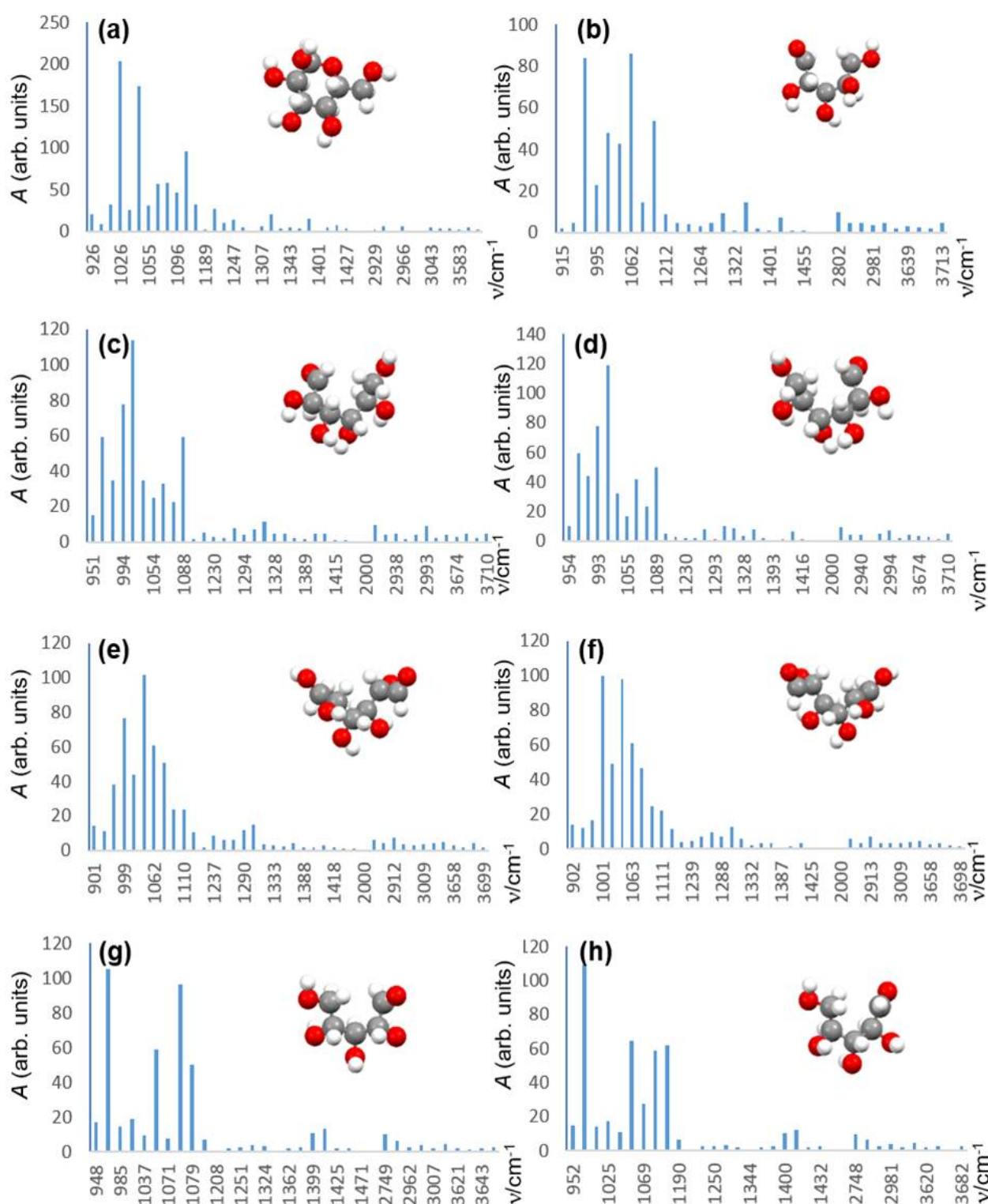
Fig. 4 **The calculated frequencies and intensities of normal vibrational modes of sugars: (a) glucose, (b) arabinose, (c) D-galactose, (d) L-galactose, (e) D-mannose, (f) L-mannose, (g) D-xylose, and (h) L-xylose.**

4. Conclusions

The assignment of lignin vibrational bands observed experimentally is extremely difficult. Therefore the use of a basis set of simulated vibrational spectra could significantly help to describe lignin from different sources. The systematic analysis of frequencies, together with the changes in their intensity for different dimers representing lignin, shows that each dimer structure has fingerprint vibrations with a high intensity level specific for each dimer. This creates the opportunity to distinguish specific dimer structure/functional groups by using calculated spectra from a basis set.

The vibration structure of the dimers was presented for three different types of lignin: H-lignin (coumaryl structure), G-lignin (coniferyl structure) and S-lignin (sinapyl structure). The addition of CH_3-O- parts significantly changed the vibrational structure of the investigated β-1 and β-5 dimers.

In the case of the investigated aldopentoses and aldohexoses the situation is much more complex due to the fact that the spectra are similar to each other. Therefore, more complicated models, such as polysaccharides, should be considered as a next step in modelling cellulose and hemicellulose.

The next step should be establishing a more complete model of the vibrational structure of lignin, taking into consideration the content of certain dimers and the optimization of the geometrical structure of the lignin multidimensional polymer.

Acknowledgment

References

[1] Huber, G. W., Iborra, S., and Corma, A. 2006. "Synthesis of Transportation Fuels from Biomass: Chemistry, Catalysts, and Engineering." *Chemical Reviews* 106: 4044-98.

[2] Langan, P., Gnanakaran, S., Rector, K. D., Pawley, N., Fox, D. T., Cho, D. W., and Hammel, K. E. 2011. "Exploring New Strategies for Cellulosic Biofuels Production." *Energy and Environmental Science* 4: 3820-33.

[3] da Costa Sousa, L., Chundawat, S. P. S., Balan, V., and Dale, B. E. 2009. "'Cradle-to-grave' Assessment of Existing Lignocellulose Pretreatment Technologies." *Current Opinion in Biotechnology* 20: 339-47.

[4] Himmel, M. E., Ding, S. Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W. and Foust, T. D. 2007. "Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuel Production." *Science* 315: 804-7.

[5] Ragauskas, A. J., Williams, C. K., Davison, B. H., Britovsek, G., Cairney, J., Eckert, C.A., Frederick, W. J. Jr., Hallett, J. P., Leak, D. J., and Liotta, C. L.. 2006. "The Path Forward for Biofuels and Biomaterials." *Science* 311: 484–9.

[6] Jarvis M. C., and McCann M. C. 2000. "Macromolecular Biophysics of the Plant Cell Wall: Concepts and Methodology." *Plant Physiology and Biochemistry* 38: 1-13.

[7] Ding, S. Y., Liu, Y. S., Zeng, Y., Himmel, M. E., Baker, J. O., and Bayer, E. A. 2012. "How does Plant Cell Wall Nanoscale Architecture Correlate with Enzymatic Digestibility?" *Science* 338: 1055-60.

[8] Dorrestijn, E., Laarhoven, L. J. J., Arends, I. W. C. E., and Mulder, P. 2000. "The Occurrence and Reactivity of Phenoxyl Linkages in Lignin and Low Rank Coal." *Journal of Analytical and Applied Pyrolysis* 54: 153-92.

[9] Gosselink, R. J. A., de Jong, E., Guran, B., and Abächerli, A. 2004. "Co-ordination Network for Lignin Standardisation, Production and Applications Adapted to Market Requirements (EUROLIGNIN)." *Industrial Crops and Products* 20: 121-9.

[10] Björkman, A. 1957. "Lignin and Lignin-Carbohydrate Complexes." Industrial and Engineering Chemistry 49: 1395-8.

[11] Dorrestijn, E., Laarhoven, L. J. J., Arends, I. W. C. E., and. Mulder, P. 2000. "The Occurrence and Reactivity of Phenoxyl Linkages in Lignin and Low Rank Coal." *Journal of Analytical and Applied Pyrolysis* 54: 153-92.

[12] Parthasarathi, R., Romero, R. A., Redondo, A., and Gnanakaran, S. 2011. "Theoretical Study of the Remarkably Diverse Linkages in Lignin." *Journal of*

Physical Chemistry Letters 2: 2660-6.

[13] Martinez, C., Rivera, J. L., Herrera, R., Rico, J. L., Flores, N., Rutiaga, J. G., and López, P. 2007. "Evaluation of the Chemical Reactivity in Lignin Precursors using the Fukui Function." *Journal of Molecular Modeling* 14: 77-81.

[14] Haensel, T., Reinmoeller, M., Lorenz, P., Beenken, W. J. D., Krischok, S., and Imad-Uddin Ahmed, S. 2012. *Cellulose* 19: 1005.

[15] Wiercigroch, E., Szafraniec, E., Czamara, K., Pacia, M. Z., Majzner, K., Kochan, K., Kaczor, A., Baranska, M., and Malek, K. 2017. "Raman and Infrared Spectroscopy of Carbohydrates: A review." *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 185: 317-35.

[16] Brauer, B., Pincu, M., Buch, V., Bar, I., Simons, J. P., and Gerber, R. B. 2011. "Vibrational Spectra of α-glucose, β-glucose, and Sucrose: Anharmonic Calculations and Experiment." *Journal of Physical Chemistry A* 115: 5859-72.

[17] Zheng, R., Wei, W., and Shi, Q. 2009. "Density Functional Theory Study on Sum-Frequency Vibrational Spectroscopy of Arabinose Chiral Solutions." *Journal of Physical Chemistry A* 113: 157-64.

[18] Konopka, C. 2011. "Influence of the Preparation Method of Lignin and Comparison to Commercial Products." Bachelor Thesis, Faculty of Materials Engineering of Georg Simon Ohm University of Applied Sciences (Nuremberg, Germany).

[19] Hermann, K., Pettersson, L. G. M., Casida, M. E., Daul, C., Goursot, A., Koester, A., Proynov, E., St-Amant, A., Salahub, D. R., Carravetta, V., Duarte, A., Godbout, N., Guan, J., Jamorski, C., Leboeuf, M., Leetmaa, M., Nyberg, M., Pedocchi, L., Sim, F., Triguero, L., and Vela, A. 2005. StoBe-deMon, deMon Software: Stockholm, Berlin.

[20] Perdew, J. P., Burke, K., and Ernzerhof, M. 1996. "Generalized Gradient Approximation Made Simple." *Physical Review Letters* 77: 3865-8.

[21] Hammer, B., Hansen, L. B., and Nørskov, J. K. 1999. "Improved Adsorption Energetics within Density-Functional Theory using Revised Perdew-Burke-Ernzerhof Functionals." *Physical Review B* 59: 7413-21.

[22] Labanowski, J. K., and Anzelm, J. W. 1991. Eds., Density Functional Methods in Chemistry, Springer, New York.

[23] Broclawik, E., and Salahub, D. R. 1993. "Density Functional Theory and Quantum Chemistry: Metals and Metal Oxides." *Journal of Molecular Catalysis* 82: 117-29.

[24] Friedrich, C. 2004. "Geometrische, Elektronische und Vibronische Eigenschaften der Reinen und Defektbehafteten V2O5(010)-Oberfläche und Deren Wechselwirkung mit Adsorbaten: Theoretische Untersuchungen." Ph.D. Thesis, Free University of Berlin.

[25] Santos, R., Hart, P. W., Jameel, H., and Chang, H. M. 2013. "Wood Based Lignin Reactions Important to the Biorefinery and Pulp and Paper Industries." *BioResources* 8: 1456-77.

[26] Chen, H. 2014. "Biotechnology of Lignocellulose: Theory and Practice, Chapter 2: Chemical Composition and Structure of Natural Lignocellulose." Chemical Industry Press, Beijing, and Springer Science + Business Media Dordrecht.

[27] Schneider, B., Stokr, J., Schmidt, P., Mihailov, M., Dirlikov, S., and Peeva, N. 1979. "Stretching and Deformation Vibrations of CH_2, $C(CH_3)$ and $O(CH_3)$ Groups of Poly(methyl methacrylate)." *Polymer* 20: 705-12.

[28] van den Broek, M. A. F. H., Nienhuys, H. K., and Bakker, H. J. 2001. "Vibrational Dynamics of the C–O Stretch Vibration in Alcohols." *The Journal of Chemical Physics* 114: 3182-6.

[29] Silverstein, R. M., Bassler, G. C., and Morrill, T. C. 1981. "Spectrometric Identification of Organic Compounds." 4th ed. New York: John Wiley and Sons, 1981.