Phytochemistry 71 (2010) 453–462

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Macromolecular replication during lignin biosynthesis

Yi-ru Chen, Simo Sarkanen *

Department of Bioproducts and Biosystems Engineering, University of Minnesota, St. Paul, MN 55108, USA

ARTICLE INFO

Article history:
Received 12 December 2008
Received in revised form 24 October 2009
Available online 4 January 2010

Keywords:
Lignin primary structure
Replication
Template polymerization
Density functional theory
Dynamical electron correlation

ABSTRACT

Lignins play a crucial role in the cell-wall architecture of all vascular plants. They are composed of *p*-hydroxyphenylpropanoid units interconnected through covalent bonds formed during lignol radical coupling between six different pairs of atomic centers. For 50 years, the primary structures of lignins have been thought to be random, but for a number of reasons such an assumption is not tenable. For example, it has been reported that, by simple physicochemical means, the rather recalcitrant lignins in spruce wood can be decisively separated into two fractions containing quite dissimilar biopolymer chains. Thus, a paradigm shift should be imminent, and a detailed working hypothesis for the mechanism of lignin biosynthesis would be invaluable in delineating how the process of macromolecular lignin assembly can be properly investigated. In conjunction with an earlier experimental result, an explicit model for a template dehydropolymerization process has been developed that describes how lignin primary structure is replicated. The strengths of the powerful noncovalent interactions have been calculated that control the transient placement of lignol radicals about to undergo coupling on a double-stranded lignin template. These elementary steps engender, in the growing daughter chain, a primary structure identical to that of the distal template strand. The interactions are governed by dynamical electron correlation in the π-orbitals of each immobilized lignol radical and the complementary aromatic ring in the antiparallel proximal strand. The resulting noncovalent forces are computed to be stronger than those stabilizing GC/CG base pairs in DNA double-helices, but the mechanism of replication is fundamentally different from that of any other biopolymer.

1. Introduction

"Natural lignin and lignin formed *in vitro* are identical" (Freudenberg, 1959). Karl Freudenberg made this statement over 50 years ago; it represented his assessment of a pioneering series of investigations into lignin structure carried out during the 1950s. He was alluding to a comparison between milled-wood lignin extracted from spruce and the dehydropolymerisates formed from the monolignol, coniferyl alcohol, through single-electron oxidation by laccase, peroxidase or cupric ion. The conclusion was based upon similarities that had been observed in the infrared spectra and analytical degradation products obtained from these natural and synthetic lignin preparations.

1.1. Randomness in lignin primary structure

As was soon pointed out, a corollary of this working hypothesis would be that the "sequence of the individual units in lignin is fortuitous, for they are not moulded like proteins on a template" (Freudenberg, 1968). It was understandably difficult to conceive of a reason why the substructures resulting *in vitro* from the coupling of coniferyl alcohol radicals to the growing end of a lignin-like chain should be arranged in any particular order. Thus, the primary structures of lignins formed *in vitro* and *in vivo* were assumed to be random.

A very substantial body of opinion in lignin biochemistry continues to hold that the essence of this 50-year-old working hypothesis is fundamentally correct (Ralph et al., 2008). There have certainly been significant changes in the identities and/or proportions of accepted substructures in lignin macromolecules. Nevertheless, monolignol radicals are still generally thought to couple with the oxidized end of a lignin chain through atomic centers that are selected by their unpaired electron densities and other simple physicochemical constraints which account for the probabilities of formation of the possible alternative covalent bonds. The impact of the adjoining noncovalent interactions has not been fully recognized, however, and thus the coupling process has remained random conceptually (although it has more recently been designated "combinatorial" in reference to the different proportions of substructures created (Ralph et al., 2004)).

As cell-wall constituents of all vascular plants, lignins represent the second most abundant group of biopolymers. They are composed of *p*-hydroxyphenylpropanoid units linked through β–*O*–4

* Corresponding author. Tel.: +1 612 624 6227; fax: +1 612 625 6286.
E-mail address: sarka001@umn.edu (S. Sarkanen).

0031-9422/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2009.11.012

(primarily), β–5, β–β, 5–5, β–1 and 4–*O*–5 bonds designated according to the atomic centers in the radicals that are coupled together during the final step of lignin biosynthesis. The equivalent substructures in the biopolymer consist of alkyl aryl ethers, phenylcoumarans, resinols, dibenzodioxocins and biphenyls, tetrahydrofuran-*spiro*-cyclohexadienones, and diaryl ethers (Fig. 1).

1.2. Order in the physicochemical properties of lignins

An important discovery began to attract attention over the past 4 years (Gellerstedt, 2007; Lawoko et al., 2005). By means of consecutive treatments with aqueous 8 M urea, alkaline borate and/or barium hydroxide, it has proven possible to separate about 90% of the lignin in endoglucanase-treated spruce wood meal into two homogeneous fractions possessing quite different macromolecular configurations. In both, the chains are covalently linked to hemicelluloses (55% average content) through a small proportion (about 3% (Balakshin et al., 2007)) of the lignin monomer residues. Those that are bound primarily to xylans are composed almost exclusively of β–*O*–4 alkyl aryl ether substructures; the others, which are predominantly attached to glucomannans, are thought to embody (in appropriately adjusted proportions) all of the substructures found in softwood lignins (Fig. 1).

Such a striking result is clearly inconsistent with a random (or "combinatorial") distribution of inter-unit linkages in lignin macromolecules. This finding was anticipated in a preliminary way by the observation (first disclosed in 1970) that β–β-linked pinoresinol moieties are connected to macromolecular lignin chains predominantly (if not obligatorily) through at least one of their aromatic C-5 atoms. Thus, neither acid-catalyzed hydrolysis (Lundquist, 1970) nor thioacidolysis (Lapierre et al., 1991) of gymnosperm lignins yields any pinoresinol-related dimers; however, the release of such substructures is readily detected during thioacidolysis of synthetic dehydropolymerisates formed from coniferyl alcohol *in vitro* (Terashima et al., 1996).

The underlying concept behind the macromolecular scheme in Fig. 1, where combinations of different but well-defined chain configurations contribute to the overall structure of lignin, is consistent with comprehensive degradative analyses that are rigorously quantitative. Thus, the monomeric and dimeric thioacidolytic cleavage products derived from β–*O*–4, β–5, 5–5 (3–3), β–1 and 4–*O*–5 lignin substructures in a variety of alfalfa plant lines have

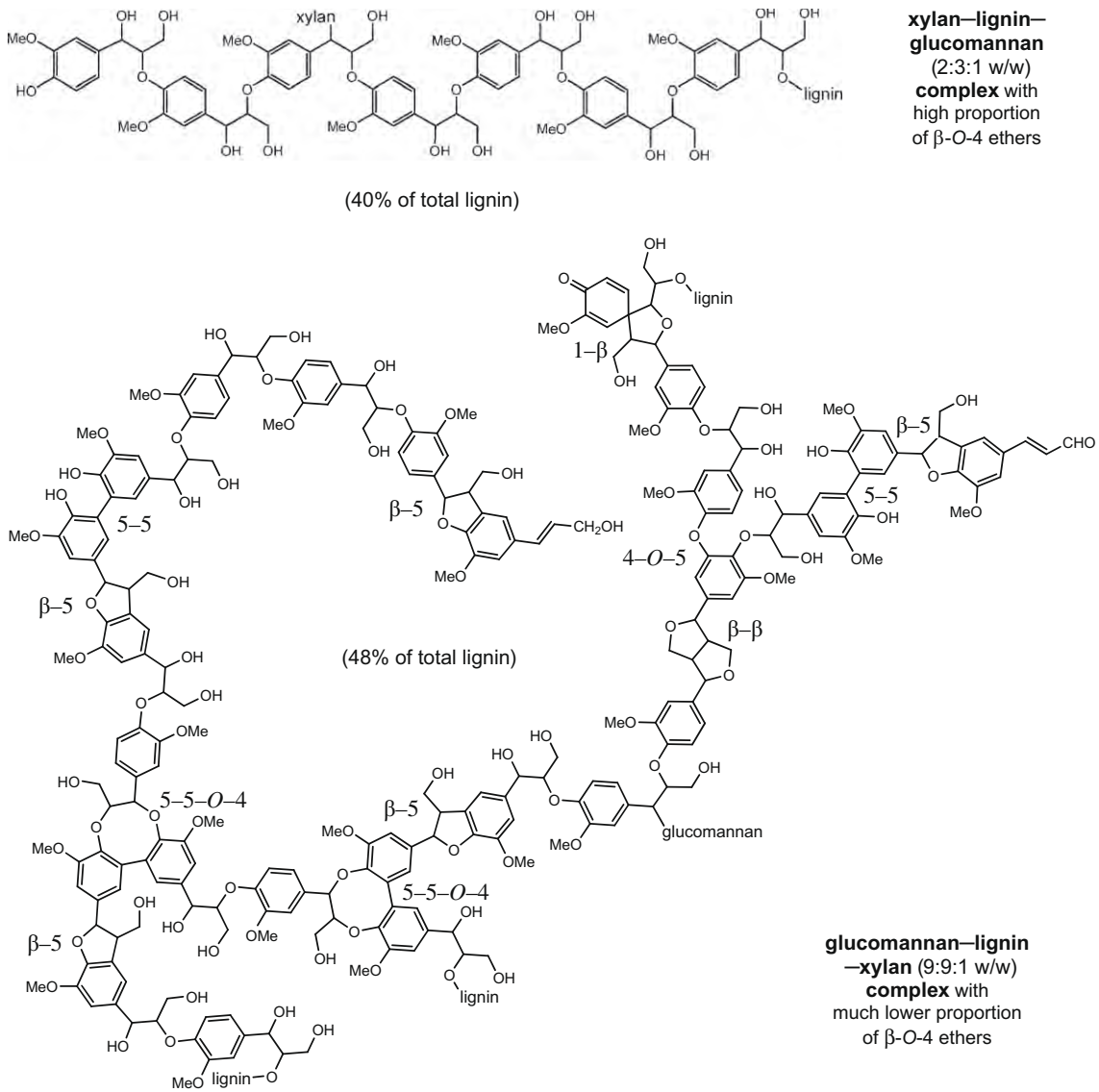

Fig. 1. Structures suggested for gymnosperm lignin–carbohydrate complexes isolated by successive fractionation of endoglucanase-treated spruce wood meal with aqueous 8 M urea, alkaline borate and/or barium hydroxide (adapted from Gellerstedt, 2007).

been found to increase linearly with stem growth and development until maturation (Davin et al., 2008a,b). The key point is that these correlations do not depend on whether there is a methoxyl substituent at the aromatic C-5 atom, one of the positions through which a monolignol may undergo coupling when becoming linked to the growing end of a lignin chain. Such a result implies that lignin macromolecules encompass specific sets of primary structures, the relative proportions of which vary linearly with overall lignin content.

Even kraft lignins, which are formed through cleavage and further transformation of the native biopolymer under strenuous alkaline conditions, retain a memory of the original macromolecular structure. These derivatives consist mostly of large supramacromolecular complexes composed of many individual components; yet the largest of these associated entities are remarkably few in number even though they are interconvertible (Chen and Sarkanen, 2006). Early in the process of their formation, the collective behavior of the associated kraft lignin complexes is consistent with there being a different locus on each of the species that is complementary to only one of the individual components present (distinguished by molecular weight) (Garver et al., 1989).

In keeping with this high degree of selectivity, the dissociation of softwood kraft lignin complexes in aqueous 0.10 M NaOH is a first-order process characterized by a rate coefficient of $3.9 \times 10^{-3}\ h^{-1}$ at 25.0 °C (Garver et al., 1989). Such a value is remarkably similar in magnitude to the corresponding parameter ($7.0 \times 10^{-3}\ h^{-1}$) describing the dissociation of Norway spruce lignin in 0.1 M NaOH at room temperature (Contreras et al., 2008). Since the rates of dissociation of individual components from associated (kraft or native) lignin complexes should depend quite strongly on molecular weight, dissociative processes that are first-order overall would not be expected to manifest themselves. However, overall first-order behavior would prevail under circumstances where the individual lignin components are released from each associated complex in a particular order through consecutive dissociative events controlled by the same rate-limiting step (Garver et al., 1989).

1.3. Lignin replication by a template polymerization process

The question then arises as to how different primary structures (such as those depicted in Fig. 1) could be replicated during lignin biosynthesis in plant cell walls. When the first chain in any lignin domain has been assembled, it will not be able to be displaced from the locus that, by whatever means, encodes its primary structure: its mobility will be highly restricted because of its insolubility. Hence, it is most likely that lignin configuration is preserved through a direct template polymerization mechanism (Guan et al., 1997; Sarkanen, 1998) where each daughter strand is topologically complementary to the parent chain from which it is derived.

The foregoing working hypothesis has a number of consequences. Since the topology of each macromolecular lignin chain is dependent on the conformations of the respective substructures that constitute its inter-unit linkages, fidelity of replication requires that the template be relatively rigid. Thus, it has been proposed that the functional template at the leading edge of a lignifying domain is comprised of two polymeric strands interacting to form a relatively rigid complex (Chen and Sarkanen, 2003).

Just before coupling, a monolignol radical is transiently immobilized on the double-stranded template in a position adjacent to the radical at the growing end of the daughter lignin chain. Formation of a new covalent bond between these two open-shell species leads to a substructure that is complementary to the dimeric moiety in the proximal template strand with which the radicals were interacting immediately prior to the event. It may be supposed that the assembly of each macromolecular segment in a lignin domain proceeds to completion before biosynthesis of the next chain can be initiated in the reverse direction. Consequently, an antiparallel relationship should prevail between each daughter and parental lignin strand.

The double-stranded template at the developing front of each lignifying domain perpetuates itself in a dynamical state. Conformational analysis reveals that the overall structure is puckered when it is fully extended (Chen and Sarkanen, 2003). As a monolignol radical couples to the end of a growing lignin chain, the new substructure formed tends to associate in a head-to-tail fashion with the corresponding dimeric *motif* in the proximal template strand. Because of the puckering, however, the same geometric relationship between the complementary dimeric substructures cannot be preserved across all three interacting lignin chains. Therefore, the growth of the new lignin chain must be followed by a replication fork that is caused by incremental dissociation of the distal strand from the proximal template strand (Chen and Sarkanen, 2003). This process occurs in an orderly manner when the nonbonded attraction between the respective lignol radicals and the proximal dimeric substructure is stronger than the intermolecular forces between the closed-shell aromatic moieties in the double-stranded template. Consequently, arrays of lignin chains are unable to accommodate the existence of crystalline domains in either plant cell walls or in isolated lignin (derivative) preparations.

It has been reported that, without participating covalently, methylated lignin macromolecules are capable of promoting the formation of high molecular weight dehydropolymerisate components during oxidative monolignol coupling in open solution (Guan et al., 1997). The likelihood that this was actually the first preliminary demonstration of monolignol template dehydropolymerization *in vitro* cannot be ignored. However, the function of the methylated lignin macromolecule in this experiment has been dismissed by some workers as "likely nothing more than an assembly surface" onto which monolignols may become adsorbed before reaction (Ralph et al., 2004). In principle, the matter could be resolved by comparing the primary structures of the dehydropolymerisates with those of the putative lignin template strands, but this is not yet feasible.

On the other hand, effective template dehydropolymerization of monolignols on a pre-existing lignin chain will require strong intermolecular attraction between the aromatic moieties in the parent substructures and the lignol radicals about to undergo coupling. Moreover, for the orderly expansion of a lignin domain, a head-to-tail orientation should, in most cases, prevail between a monolignol radical and the closed-shell residue with which it interacts in the proximal template strand (*vide supra*). Under these circumstances, the degeneracy of the replicative process (whereby the primary structure of the distal template strand is transcribed into the daughter chain) will be determined by the potential well that characterizes the intermolecular interactions between each lignol radical and the corresponding aromatic moiety in the proximal lignin strand. The present work has sought to examine the first two of these conditions upon which a template polymerization mechanism in lignin biosynthesis would depend.

2. Results and discussion

2.1. Stabilization energies of lignol radical complexes

The question arises as to how the noncovalent attraction between aromatic residues in lignin substructures may be reliably calculated. If the interaction energy is dominated by contributions from electrostatic and dispersion forces, when charge transfer is

insignificant there could be a continuum of favorable configurations from edge-on to cofacially offset placements of the rings (Hunter and Sanders, 1990). Hartree–Fock wave function theory fails to locate nonbonded complexes between closed-shell lignin monomer units that are separated by distances within which π–π interactions are significant (Sarkanen and Chen, 2005). The second-order Møller-Plesset perturbation (MP2) technique affords one of the least expensive ways of improving on traditional Hartree–Fock wave function theory in regard to correcting for electron correlation. Unfortunately, MP2 calculations overestimate the attraction in edge-on and parallel benzene dimers quite substantially (Tsuzuki et al., 2002). Thus, it is to be expected that the stabilization energies previously estimated at the MP2/6-31G(d) level of wave function theory (*vide infra*) for complexes of cofacially offset veratryl alcohol molecules (representing monomeric lignin substructures) would be too high (Sarkanen and Chen, 2005).

On the other hand, Kohn–Sham density functional theory affords advantages in the study of biological systems owing to its computational efficiency. Indeed, geometry optimizations for six dilignols have been performed at the B3LYP/6-31G(d,p) level of theory, and intramolecular hydrogen bonding was found to stabilize the lowest energy conformations of several of them (Durbeej and Eriksson, 2003). However, as recently as 2005, it had not been possible to locate stacked complexes of nucleobases or aromatic amino acids with density functional theory. Early in 2006, a dramatic improvement in the situation occurred with the publication of the M05-2X hybrid meta exchange–correlation functional (Zhao et al., 2006) that is parameterized to provide reliable results for noncovalent interactions, *inter alia*. For example, using the 6-31+G(d,p) basis set, this method yielded a mean unsigned error of just 1.2 kcal/mol relative to the best estimates for nucleobase π···π stacking energies (Zhao et al., 2006).

Nevertheless, the results of these calculations are afflicted by basis set superposition error, which arises from the opportunity that each molecule in a complex has to exploit Gaussian basis functions of the other component(s) in describing its own intramolecular structure more accurately (Boys and Bernardi, 1970). The counterpoise approach developed to deal with this effect is widely thought to overestimate the requisite correction, which is generally smaller for density functional theory than for correlated wave function theories (Zhao and Truhlar, 2005). Some investigators have adopted half of the counterpoise correction as a reasonable reflection of basis set superposition error (Kim et al., 2000; Zhao et al., 2006) and such a practice has been employed in the present work. Otherwise, frequency calculations have been routinely employed to identify saddle points on the potential energy surfaces.

2.2. Lignol radical complexes with lignin template substructures

Apart from their stabilization energies, the configurations of complexes between lignol radicals and the aromatic moieties in lignin chains are of direct relevance to the feasibility of any prospective mechanism for macromolecular lignin replication. The first question that arises is whether they embody edge-on or cofacially offset arrangements of the aromatic rings. These alternative configurations for complexes between the coniferyl alcohol (monolignol) radical and veratryl alcohol (representing a guaiacyl unit in a lignin chain) are depicted in Fig. 2a and b. At the M05-2X/6-31+G(d,p) level of density functional theory, the stabilization energies are calculated to be 5.1 kcal/mol for the edge-on (head-to-head) placement of the rings and 8.9 kcal/mol for the cofacially offset (head-to-tail) configuration.

The ~3.3-Å separation between the aromatic rings in the (more stable) cofacial complex is within a distance that could allow significant π-orbital overlap, but according to the standard Mulliken population analysis (not shown), less than 1% of the unpaired electron in the coniferyl alcohol radical is delocalized over the atomic centers in the veratryl alcohol. Concomitantly, the unpaired electron density may vary by as much as 3% (of the total) among the individual atomic centers of the coniferyl alcohol radical itself.

At the same time, there is no intermolecular hydrogen bonding in the complexes with the configurations depicted. Therefore, the intermolecular attraction arises primarily from dynamical electron correlation among the interacting π-orbitals. Consequences of this have been observable, for example, during the reversible dissociation of softwood kraft lignin complexes in aqueous 0.10 M NaOH: the accompanying ultraviolet–visible spectral changes were small but well-defined in maintaining isosbestic points at 305 and 455 nm (Garver et al., 1989).

At the MP2/6-31G(d) level of wave function theory, the configurations of the edge-on and cofacially offset complexes located for coniferyl alcohol radical and veratryl alcohol (not shown) are quite comparable (although not identical) to those in Fig. 2a and b, and their stabilization energies are estimated to be 4.9 and 11.2 kcal/mol, respectively. This comparison between the results from density functional theory and wave function theory strengthens the likelihood that the complexes between monolignol radicals and closed-shell aromatic moieties in lignin chains are cofacial.

A similar situation appears in the evaluation of edge-on and cofacially offset complexes between the vanillyl alcohol radical (representing the growing end of a lignin chain) and veratryl alcohol (Fig. 2c and d). Their stabilization energies are calculated to be 4.0 and 9.7 kcal/mol, respectively, at the M05-2X/6-31+G(d,p) level of density functional theory (Sarkanen and Chen, 2007). Again, there is no hydrogen bonding between the vanillyl alcohol radical and veratryl alcohol in the complexes shown, and thus dynamical electron correlation is largely responsible for the prevailing intermolecular attractive interactions.

The stabilization energies of the cofacial complexes between lignol radicals and veratryl alcohol (Fig. 2b and d) are probably lower limits to the actual values in aqueous solution: here contributions from intermolecular hydrogen bonding are attenuated because of the high dielectric constant of the medium and competition from solvent molecules. Nevertheless, electron correlation alone in lignol radical complexes engenders a level of stability greater than that arising from the three hydrogen bonds and aromatic stacking interactions which stabilize GC/CG base pairs in DNA double-helical structures in aqueous solution (Every and Russu, 2007).

The question of how much intermolecular hydrogen bonding could contribute to the interactions between a monolignol radical and veratryl alcohol can be established by comparing the stabilization energies of corresponding hydrogen-bonded and non-hydrogen-bonded complexes that possess very similar configurations. Appropriate examples are depicted in Fig. 3a and b for a coniferyl alcohol radical where the Cα–Cβ double bond is *syn* (rather than *anti*) to the aromatic methoxyl group. The difference between the gas-phase stabilization energies of the two complexes is about 6 kcal/mol, but the actual contribution from the αOH···O4 hydrogen bond in aqueous solution will be considerably smaller (less than 2 kcal/mol).

Whether or not intermolecular hydrogen bonding is present, the configuration of the complex between a monolignol radical and a closed-shell lignin substructure strongly affects the stabilization energy. Two such instances are depicted in Fig. 3c and d. The (non-hydrogen-bonded) head-to-head complex between the coniferyl alcohol radical and veratryl alcohol (Fig. 3c) has a stabilization energy of 6.2 kcal/mol, almost 3 kcal/mol less than that for a corresponding head-to-tail complex (Figs. 2b and 3a). On the other hand, the (non-hydrogen-bonded) head-to-tail complex between the coniferyl alcohol radical and β–1 tetrahydrofuran-*spiro*-cyclohexadienone (Fig. 3d) has the largest stabilization

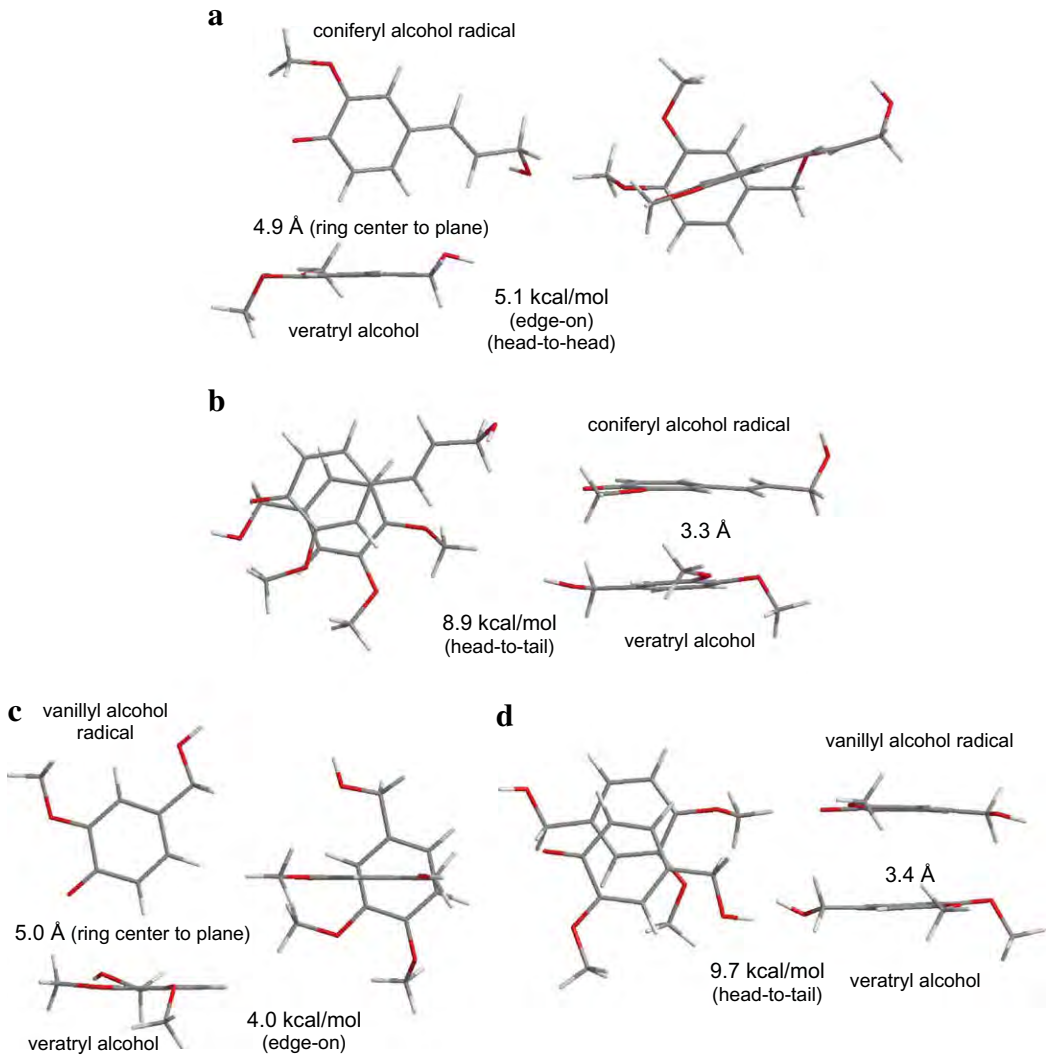

Fig. 2. Interactions of lignol radicals with closed-shell lignin substructures in conformations that preclude intermolecular hydrogen bonds: perpendicular Chem3D depictions of (a) edge-on and (b) cofacially offset coniferyl alcohol radical complexes with veratryl alcohol; (c) edge-on and (d) cofacially offset complexes of vanillyl alcohol radical (representing oxidized lignin chain-end) with veratryl alcohol.

energy (11.0 kcal/mol) among all the monolignol radical complexes hitherto located. Yet this noncovalent complex possesses a 3.8 Å average separation between the rings with a 25° interplanar angle that brings the two methoxyl O-atoms to a 2.9 Å distance of one another.

All of the geometry-optimized structures and stabilization energies of the complexes in Figs. 2 and 3 were obtained from gas-phase density functional theoretical calculations. Owing to the physical origins of electron correlation, the condensed-phase structures and energies of these complexes may not be very different from those in the gas phase. The situation in aqueous solution has been explored by solvating the coniferyl alcohol radical with eight water molecules, four around the O4 atom, one coordinated to the methoxy group, and three around the γ-hydroxyl group. After geometry optimization, the solvated coniferyl alcohol radical was cofacially positioned 3.3 Å away from a veratryl alcohol molecule in the same relative orientation as in Fig. 2b. Geometry optimization of the resulting complex was allowed to proceed towards convergence to a point before significant migration of the water molecules had begun (Fig. 4a). Even though the resulting configuration has not been fully optimized (owing to water molecule mobility), the stabilization energy of the partially solvated complex in this state (Fig. 4a) is only 0.7 kcal/mol lower than that of the corresponding gas-phase entity (Fig. 2b). Nevertheless, the first water molecules in the immediate vicinity of the complex will have the greatest effect on its stabilization energy. Therefore, the presence of water would not be expected to have a significantly adverse effect on the attractive interactions arising from electron correlation in complexes between lignol radicals and aromatic rings in lignin macromolecules.

The lignin domains themselves are hydrophobic, and thus are not solvated in the usual manner. The question of whether neighboring lignin units may influence the configurations and stabilization energies of complexes between lignol radicals and closed-shell lignin substructures is examined in Fig. 4b–d. A geometry-optimized ternary complex between coniferyl alcohol radical, veratryl methyl ether and veratryl alcohol (Fig. 4b) is compared with the corresponding geometry-optimized binary complexes between the monolignol radical and veratryl methyl ether (Fig. 4c) and veratryl methyl ether and veratryl alcohol (Fig. 4d), respectively. In relative terms, each of the individual molecular entities is very similarly positioned and oriented in these binary and ternary complexes. Moreover, the sum of the stabilization energies for the two binary complexes (17.6 kcal/mol) is very similar to the overall

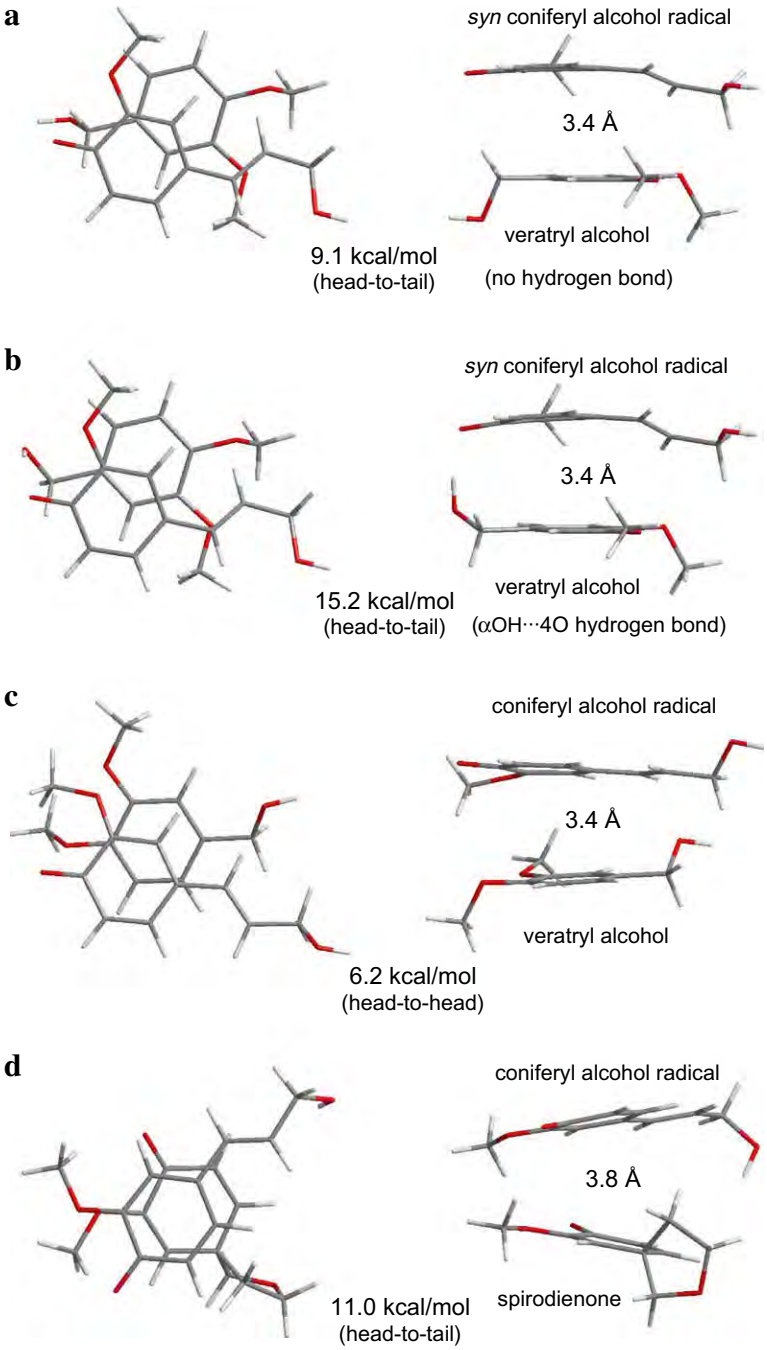

Fig. 3. Dependence of stabilization energy on configuration of cofacial monolignol radical complex: perpendicular Chem3D depictions of *syn* coniferyl alcohol radical complexes with veratryl alcohol in (a) absence and (b) presence of intermolecular hydrogen bond; *anti* coniferyl alcohol radical complexes with (c) head-to-head veratryl alcohol and (d) head-to-tail tetrahydrofuran-*spiro*-cyclohexadienone, both without intermolecular hydrogen bonds.

energy of the ternary complex (18.0 kcal/mol) with respect to those of the individual molecules. Thus, the nonbonded attraction caused by electron correlation in complexes between open- and closed-shell lignin species is not significantly influenced by the presence of the other substructures in lignin domains.

2.3. Closed-shell/closed-shell and open-shell/open-shell interactions

When each lignol-radical-coupling event creates a new inter-unit linkage at the growing end of a lignin macromolecule, the resulting closed-shell substructures will interact quite strongly with suitably positioned aromatic rings in the proximal lignin template strand. As illustrated in Fig. 4d, powerful interactions may persist even if the relative placement of the aromatic rings changes significantly. The effect is quite general; it is exemplified by complexes between two dibenzodioxocins and two tetrahydrofuran-*spiro*-cyclohexadienones, respectively, in Fig. 5a and b. In the spirodienone complex, even though the average separation (~4.1 Å) and angle (28°) between the ring planes are the greatest encountered in any complex so far, the stabilization energy (11.0 kcal/mol) is also equal to the largest despite the rather short distance (2.9 Å) between the methoxyl O-atoms. Hence, the attractive forces between macromolecular chains in lignin domains are very strong even though intermolecular orbital overlap contributes

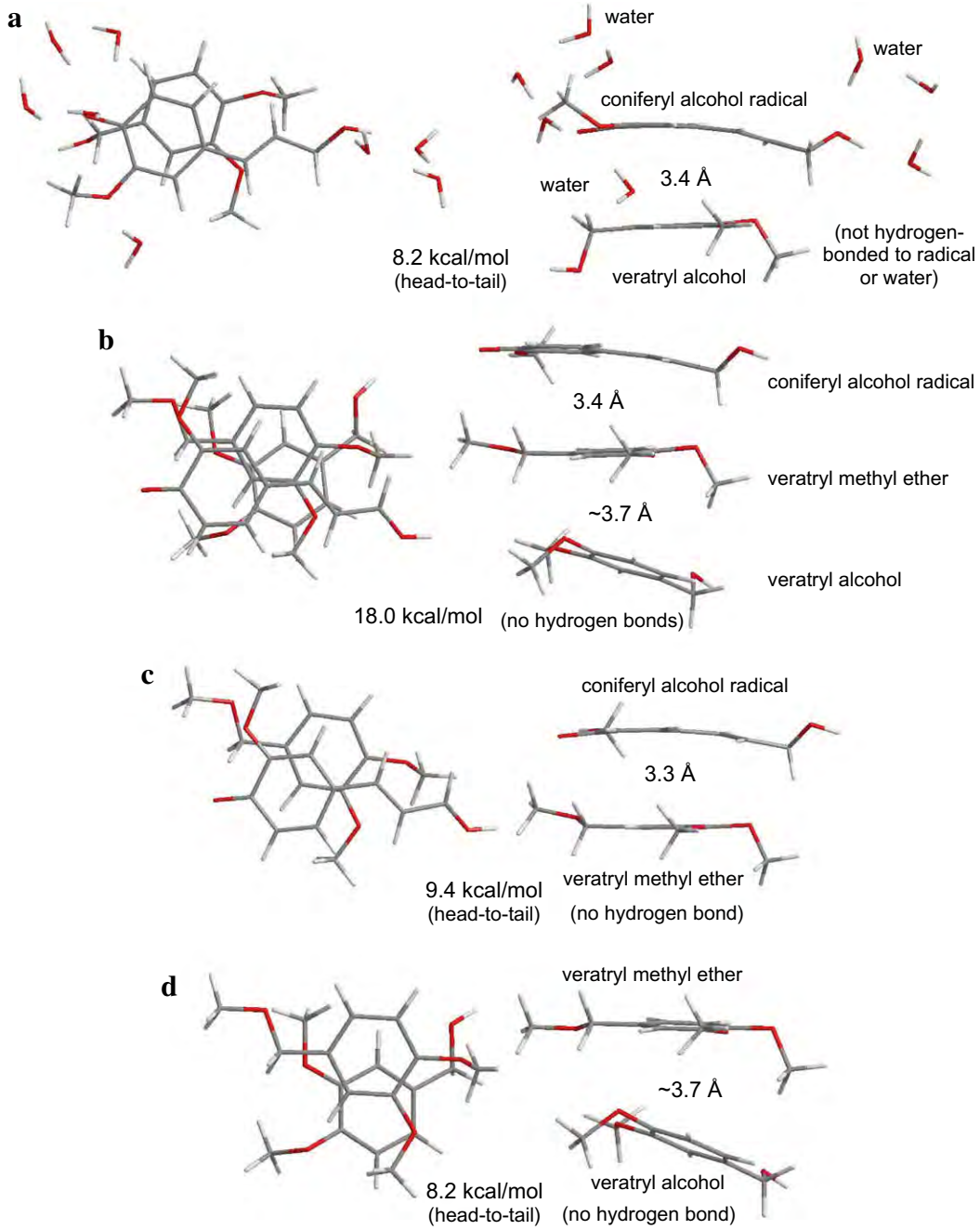

Fig. 4. Stabilization energies of cofacial monolignol radical complexes in condensed phases: perpendicular Chem3D depictions of (a) complex of coniferyl alcohol radical solvated by eight water molecules with non-hydrogen-bonded veratryl alcohol; complexes without hydrogen bonds of coniferyl alcohol radical with (b) veratryl methyl ether noncovalently bound to veratryl alcohol and (c) veratryl methyl ether alone; (d) closed-shell cofacially offset complex without hydrogen bond between veratryl methyl ether and veratryl alcohol.

little in their favor. Indeed, the interactions would be enhanced considerably by intermolecular hydrogen bonding within these hydrophobic regions from which water molecules are excluded.

It was found 40 years ago that oxidative β–β coupling of isoeugenol stereospecifically yields only the *threo*-coupling product, a bisquinone methide which, upon nucleophilic addition of water, undergoes cyclization to form the corresponding tetrahydrofuranoid compound. It was proposed that the total absence of the *erythro*-coupling product arises from the transient formation of a head-to-head π-complex between the two isoeugenol radicals (Sarkanen, 1971; Sarkanen and Wallis, 1973). Indeed, at the M05-2X/6-31+G(d,p) level of theory, the geometry-optimized complexes leading to *threo*-coupling (*e.g.*, Fig. 5c) have the shortest distances between the Cβ-atoms among the various alternative configurations.

2.4. Overall replication of lignin chains

In view of what has come to light, it would be quite surprising if the final step in lignin biosynthesis were not to take place through a template polymerization mechanism. The striking way in which lignin macromolecules noncovalently promote the dehydropolymerization of coniferyl alcohol to high molecular weight components (Guan et al., 1997) is fully corroborated by the strengths of

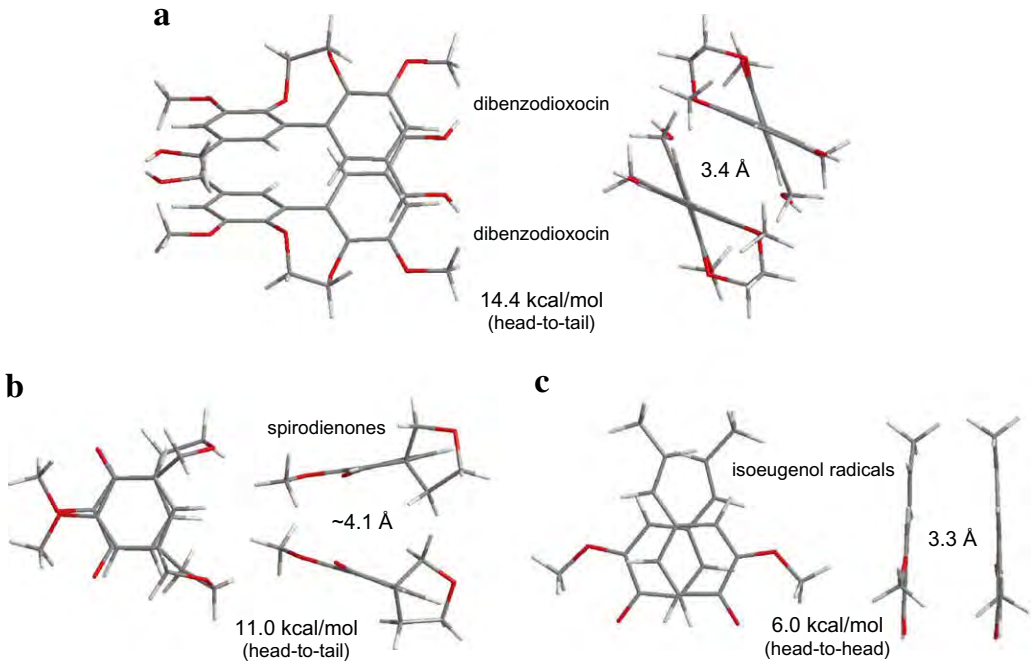

Fig. 5. Closed-shell/closed-shell and open-shell/open-shell interactions between macromolecular lignin substructures and monolignol radicals, respectively: perpendicular Chem3D depictions of cofacial bimolecular complexes between (a) dibenzodioxocin and (b) tetrahydrofuran-*spiro*-cyclohexadienone molecules, and (c) isoeugenol radicals, all without intermolecular hydrogen bonds.

the nonbonded attractive interactions between lignol radicals and closed-shell lignin substructures. The fidelity of replication (*i.e.*, the degeneracy of the process) of course rests on the contours of the potential wells that accommodate the lignol radicals as they are about to undergo coupling on the lignin template strand.

There are two remaining matters that call for clarification. First, how might the orderly assembly of successive lignin chains be achieved within a lignin domain? The most straightforward possibility calls for each macromolecular lignin chain segment to be bounded, at least at one end, by β–β-linked resinol moieties (Chen and Sarkanen, 2003); these dilignol precursors are biosynthesized regio- and stereospecifically elsewhere from monolignols through the action of a dirigent protein in concert with a suitable (per)oxidase (Davin et al., 1997). The locus that is complementary to this dimeric substructure in a (proximal) lignin template strand should interact particularly strongly (*cf.* Fig. 5a) with an incoming β–β-linked resinol molecule that is about to act as an initiation site for the assembly of a new antiparallel lignin chain (Chen and Sarkanen, 2003).

Finally, there is the question of how lignin primary structures may be encoded. Spatial and temporal correlations of lignins with proline-rich protein epitopes have been documented in the developing cell walls of the *Zea mays* coleoptile (Müsel et al., 1997) and in secondary walls of differentiating protoxylem elements in the *Glycine max* hypocotyl (Ryser et al., 1997). Moreover, a 35:32:14 lignin-protein-carbohydrate complex released by *Picea abies* suspension cultures in response to a *Rhizosphaera kalkhoffii* elicitor has been reported to possess a 41 mol% combined Hyp-Pro content with respect to the amino acids present (Lange et al., 1995). Such a Hyp-Pro composition is consistent with the participation of proline-rich proteins in these elicitor-induced stress lignin complexes. Therefore, lignin primary structure is most probably encoded in proline-rich cell-wall proteins (Sarkanen, 1998). The conformations of such progenitorial templates for lignin biosynthesis would present contiguous arrays of sites analogous to those individually present in the dirigent proteins that govern the regio- and stereospecific coupling of monolignol radicals to form lignans (Davin et al., 1997).

3. Concluding remarks

The detailed working hypothesis substantiated here has a twofold purpose. It seeks to demonstrate that a template polymerization process which replicates lignin primary structure in plant cell walls is eminently reasonable from a mechanistic point of view. It also points to some intriguing experiments that would explicitly confirm and extend many of the concepts advanced. Thus, the present study should attest to the importance of computational chemistry in its ability to steer an implacable field onto a productive new avenue of investigation.

4. Experimental

All M05-2X/6-31+G(d,p) calculations were carried out with a locally modified Gaussian 03 suite of programs on an IBM Power4 machine in the University of Minnesota Supercomputing Institute. The work has focused primarily on the interactions between lignol radicals and aromatic moieties in lignin chains, but some complexes involving closed-shell lignin residues alone were probed. The monolignol radical examined was the one derived from coniferyl alcohol, while the vanillyl alcohol radical was chosen to depict the growing end of a macromolecular lignin chain. Monomer residues in lignin chains were represented by veratryl alcohol or veratryl methyl ether, as shown in Figs. 2–5. The essential aspects of the β–1 "spirodienone" lignin substructure (Figs. 3d and 5d) were embodied in methyl tetrahydrofuran-3-*spiro*-4′-cyclohexadienon-2′-yl ether, while the most important characteristics of the 5–5′–O–4″ dibenzodioxocin *motif* (Fig. 5a) were preserved in 4,9-dimethoxy-6,7-dihydro-5,8-dioxa-dibenzo[*a*,*c*]cyclooctene.

Spin-unrestricted and spin-restricted models were employed for open-shell and closed-shell complexes with their respective

spin multiplicities of 2 and 1. The only triplet complexes considered were seven involving two isoeugenol radicals (of which the one with the shortest nonbonded β–β′ distance is depicted in Fig. 5c).

The complexes are considered to be cofacial or edge-on, depending on whether the angle between the planes of the interacting aromatic rings is less or greater than 45°. In general, the complexes portrayed in Figs. 2, 3 and 5 are the most stable (*i.e.*, lowest in energy) of each particular type located at the M05-2X/6-31+G(d,p) level of theory. The stabilization energy is (the absolute magnitude of) the difference between the energy of the complex and those of the individual (isolated) molecular species. This is usually smaller than the interaction energy, which is the corresponding quantity calculated relative to the energies of individual radicals or molecules fixed in the same conformations as found in the complex.

In the quest for complete geometry optimization (and thus a global minimum in potential energy), the number of (systematically different) initial structures considered for a particular kind of complex varied with the impact that changes in its configuration would have on the replicative scheme being developed. The individual radical or molecular constituents were always geometry-optimized prior to incorporation into these initial structures, except where alternative rotamers had to be employed to avoid early hydrogen-bond formation. Otherwise, the number of initial structures varied from 15 cofacial and 10 edge-on complexes between the coniferyl alcohol radical and veratryl alcohol to seven complexes each involving cofacially disposed spirodienones and isoeugenol radicals, respectively. The sole exception appears in Fig. 5a, where the availability of computational resources dictated that only one dibenzodioxocin complex could be geometry-optimized.

The intermolecular separations given in Figs. 2–5 between the constituents of the geometry-optimized complexes are the average distances between aromatic ring planes in the cofacial configurations or perpendicular distances between ring plane and ring center in the edge-on configurations.

After geometry optimization, harmonic vibrational frequencies were calculated (at the same level of theory) to establish whether a minimum or first-order saddle point had been located. An imaginary frequency of $-25\ cm^{-1}$ was adopted as a limiting criterion for an effective minimum in the potential energy surface for each complex, since vibrations of such magnitude would vanish in the face of very small (energetically inconsequential) conformational changes. No imaginary frequencies greater than $-50\ cm^{-1}$ in absolute magnitude were encountered for any of the complexes. Half of the counterpoise correction was applied to each stabilization energy as a remedy for possible basis set superposition error (Kim et al., 2000; Zhao et al., 2006). The standard Mulliken population analysis was used to determine spin densities on the atomic centers of open-shell structures.

Acknowledgements

The authors express their gratitude for the computational resources provided by the University of Minnesota Supercomputing Institute. Hatch funds and a subcontract from the BioEnergy Science Center, which is a US Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science, are gratefully acknowledged.

References

Balakshin, M.Y., Capanema, E.A., Chang, H.-M., 2007. MWL fraction with a high concentration of lignin-carbohydrate linkages: isolation and 2D NMR spectroscopic analysis. Holzforschung 61, 1–7.

Boys, S.F., Bernardi, F., 1970. The calculation of small molecular interactions by the differences of separate total energies—some procedures with reduced errors. Mol. Phys. 19, 553–566.

Chen, Y.-r., Sarkanen, S., 2003. Macromolecular lignin replication—a mechanistic working hypothesis. Phytochem. Rev. 2, 235–255 (©2004).

Chen, Y.-r., Sarkanen, S., 2006. From the macromolecular behavior of lignin components to the mechanical properties of lignin-based plastics. Cellulose Chem. Technol. 40, 149–163 (and references therein).

Contreras, S., Gaspar, A.R., Guerra, A., Lucia, L.A., Argyropoulos, D.S., 2008. Propensity of lignin to associate: light scattering photometry study with native lignins. Biomacromolecules 9, 3362–3369.

Davin, L.B., Wang, H.-B., Crowell, A.L., Bedgar, D.L., Martin, D.M., Sarkanen, S., Lewis, N.G., 1997. Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. Science 275, 362–366.

Davin, L.B., Jourdes, M., Patten, A.M., Kim, K.-W., Vassão, D.G., Lewis, N.G., 2008a. Dissection of lignin macromolecular configuration and assembly: comparison to related biochemical processes in allyl/propenyl phenol and lignan biosynthesis. Nat. Prod. Rep. 25, 1015–1090.

Davin, L.B., Patten, A.M., Jourdes, M., Lewis, N.G., 2008b. Lignins: a twenty-first century challenge. In: Himmel, M.E. (Ed.), Biomass Recalcitrance—Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, Oxford, pp. 213–305.

Durbeej, B., Eriksson, L.A., 2003. A density functional theory study of coniferyl alcohol intermonomeric cross linkages in lignin—three-dimensional structures, stabilities and the thermodynamic control hypothesis. Holzforschung 57, 150–164.

Every, A.E., Russu, I.M., 2007. Probing the role of hydrogen bonds in the stability of base pairs in double-helical DNA. Biopolymers 87, 165–173.

Freudenberg, K., 1959. Biosynthesis and constitution of lignin. Nature 183, 1152–1155.

Freudenberg, K., 1968. The constitution and biosynthesis of lignin. In: Freudenberg, K., Neish, A.C. (Eds.), Constitution and Biosynthesis of Lignin. Springer-Verlag, New York, pp. 45–122.

Garver, T.M., Jr., Iwen, M.L., Sarkanen, S., 1989. The kinetics of macromolecular kraft lignin complex dissociation. In: Fifth International Symposium on Wood and Pulping Chemistry, vol. I. Tappi Proceedings, Tappi Press, Atlanta, pp. 113–119.

Gellerstedt, G., 2007. Lignin complexity: fundamental and applied issues. <http://rfparois.free.fr/LIG2G/Seminaire%20LIG2G-WEB-vs-tout-public.htm>.

Guan, S.-Y., Mlynár, J., Sarkanen, S., 1997. Dehydrogenative polymerization of coniferyl alcohol on macromolecular lignin templates. Phytochemistry 45, 911–918.

Hunter, C.A., Sanders, J.K.M., 1990. The nature of π–π interactions. J. Am. Chem. Soc. 112, 5525–5534.

Kim, K.S., Tarakeshwar, P., Lee, J.Y., 2000. Molecular clusters of π-systems—theoretical studies of structures, spectra, and origin of interaction energies. Chem. Rev. 100, 4145–4185.

Lange, B.M., Lapierre, C., Sandermann Jr., H., 1995. Elicitor-induced spruce stress lignin—structural similarity to early developmental lignins. Plant Physiol. 108, 1277–1287.

Lapierre, C., Pollet, B., Monties, B., Rolando, C., 1991. Thioacidolysis of spruce lignin: GC–MS analysis of the main dimers recovered after Raney nickel desulphuration. Holzforschung 45, 61–68.

Lawoko, M., Henriksson, G., Gellerstedt, G., 2005. Structural differences between the lignin–carbohydrate complexes present in wood and in chemical pulps. Biomacromolecules 6, 3467–3473.

Lundquist, K., 1970. Acid degradation of lignin. II. Separation and identification of low molecular weight phenols. Acta Chem. Scand. 24, 889–907.

Müsel, G., Schindler, T., Bergfeld, R., Ruel, K., Jacquet, G., Lapierre, C., Speth, V., Schopfer, P., 1997. Structure and distribution of lignin in primary and secondary cell walls of maize coleoptiles analyzed by chemical and immunological probes. Planta 201, 146–159.

Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P.F., Marita, J.M., Hatfield, R.D., Ralph, S.A., Christensen, J.H., Boerjan, W., 2004. Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochem. Rev. 3, 29–60.

Ralph, J., Brunow, G., Harris, P.J., Dixon, R.A., Schatz, P.F., Boerjan, W., 2008. Lignification: are lignins biosynthesized via simple combinatorial chemistry or via proteinaceous control and template replication? In: Daayf, F., Lattanzio, V. (Eds.), Recent Advances in Polyphenol Research, vol. 1. Wiley-Blackwell, Oxford, pp. 36–66.

Ryser, U., Schorderet, M., Zhao, G.-F., Studer, D., Ruel, K., Hauf, G., Keller, B., 1997. Structural cell-wall proteins in protoxylem development: evidence for a repair process mediated by a glycine-rich protein. Plant J. 12, 97–111.

Sarkanen, K.V., 1971. Precursors and their polymerization. In: Sarkanen, K.V., Ludwig, C.H. (Eds.), Lignins—Occurrence, Formation, Structure and Reactions. Wiley-Interscience, New York, pp. 95–163.

Sarkanen, S., 1998. Template polymerization in lignin biosynthesis. In: Lewis, N.G., Sarkanen, S. (Eds.), Lignin and Lignan Biosynthesis, ACS Symposium Series 697. American Chemical Society, Washington, DC, pp. 194–208.

Sarkanen, S., Chen, Y.-r., 2005. Towards a mechanism for macromolecular lignin replication. 59th Appita Proceedings 2, 407–414.

Sarkanen, S., Chen, Y.-r., 2007. Noncovalent interactions between substructures control macromolecular replication during lignin biosynthesis. 14th International Symposium on Wood and Fiber Pulping Chemistry 152, 1–6.

Sarkanen, K.V., Wallis, A.F.A., 1973. Oxidative dimerizations of (*E*)- and (*Z*)-isoeugenol (2-methoxy-4-propenylphenol) and (*E*)- and (*Z*)-2,6-dimethoxy-4-propenylphenol. J. Chem. Soc. Perkin I 1973, 1869–1878.

Terashima, N., Atalla, R.H., Ralph, S.A., Landucci, L.L., Lapierre, C., Monties, B., 1996. New preparations of lignin polymer models under conditions that approximate cell wall lignification. II. Structural characterization of the models by thioacidolysis. Holzforschung 50, 9–14.

Tsuzuki, S., Honda, K., Uchimaru, T., Mikami, M., Tanabe, K., 2002. Origin of attraction and directionality of the π/π interaction: model chemistry calculations of benzene dimer interaction. J. Am. Chem. Soc. 124, 104–112.

Zhao, Y., Truhlar, D.G., 2005. Benchmark databases for nonbonded interactions and their use to test density functional theory. J. Chem. Theory Comput. 1, 415–432.

Zhao, Y., Schultz, N.E., Truhlar, D.G., 2006. Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. J. Chem. Theory Comput. 2, 364–382.