JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Molecular Modeling of Syringyl and *p*-Hydroxyphenyl *β*-*O*-4 Dimers. Comparative Study of the Computed and Experimental Conformational Properties of Lignin *β*-*O*-4 Model Compounds

STÉPHANE BESOMBES, DANIELLE ROBERT, JEAN-PIERRE UTILLE, FRANÇOIS R. TARAVEL, AND KARIM MAZEAU*

Centre de Recherches sur les Macromolécules Végétales, CNRS (associated with University Joseph Fourier), BP 53, 38041 Grenoble Cedex 9, France

As a new approach for the study of the ultrastructure of lignin, the conformational preferences of lignin *β*-*O*-4 model compounds have been investigated by molecular modeling. The computed results have been compared with the experimental data (X-ray crystal structures and $^3J_{H\alpha H\beta}$ NMR coupling constant values) reported in the literature. This comparison has led to an improved understanding of the influence of the structure, stereochemistry, and intramolecular H-bonding upon the conformational properties of the *β*-*O*-4 dimers. A large number of low-energy conformations have been predicted for the structures. It has also appeared that the conformational features are predominantly governed by local steric interactions rather than by differences in the H-bonding interactions. The threo and erythro forms differ significantly in their conformational features, with a preferential extended overall shape for the threo form in which the bulky aromatic groups are distant from each other.

KEYWORDS: Lignin; *β*-*O*-4 structure; molecular modeling; conformational study; hydrogen bonding; X-ray crystal structure; NMR coupling constant

INTRODUCTION

Lignin constitutes the second most abundant biopolymer in nature after cellulose. However, its native molecular structure and the nature of its association with the cellulosic components of plant cell walls still remain incompletely understood (*1*). The selective extraction of lignin is a major problem for the pulp and paper industry, especially the removal of the residual lignin from the pulp at the bleaching stage. Lignin also constitutes a drawback in the animal feed industry because the high content of lignin phenolic acids in forage is thought to reduce the digestibility of carbohydrates by ruminants.

The ultrastructural characterization of lignin is a considerable challenge for experimental methods, due to the lack of a regular and ordered structure. Up to now, structural data have essentially been collected by techniques that require extraction and isolation stages during which the native structure of lignin is more or less altered. Moreover, the amorphous character of lignin prevents the use of X-ray diffraction techniques. Although such studies have been useful for the elucidation of the chemical composition and the primary structure of lignin, they do not give information on its three-dimensional characteristics. Experimental advances in the study of the three-dimensional structure of lignin have mainly been collected for small model compounds, through X-ray diffraction studies and measurements of proton−proton NMR coupling constants in solution.

Among alternative experimental techniques, computational methods hold great promise for studying lignin. Whereas molecular modeling using empirical force fields has proven to be very efficient in predicting the conformational characteristics and establishing the structure−property relationships of all the major biopolymers, very few studies have concerned lignin or lignocellulosic structures (*2*−*14*). Thus, given the challenges associated with the study of lignin, there is potential for molecular modeling methods to provide valuable insights into the depiction of its native structure. However, in regard of the complexity of the lignin network, it appears necessary to adopt a step by step approach for its modeling. For this reason, we chose to begin with the study of *β*-*O*-4 model compounds (**Figure 1**), as the *β*-*O*-4 linkage represents the predominant interunit linkage in lignin. Softwood lignins contain predominantly guaiacyl units, whereas hardwood lignins exhibit a broad range of guaiacyl/syringyl ratios. The *p*-hydroxyphenyl unit is much less present in wood lignin. However, its study is of interest in regard to its large presence in nonwoody plants used as fodder.

A first attempt to model a *p*-hydroxyphenyl *β*-*O*-4 dimer using the MM3 force field has previously been reported (*8*). However, no critical comparison of the computed results with the structural data obtained from experimental techniques was performed. To the best of our knowledge, no molecular modeling study has been performed for a syringyl structure. The present study has been undertaken in order to reconcile and rationalize experimental and computational results. In a

* To whom correspondence should be addressed. Telephone: 33-476-03-76-39. Fax: 33-476-54-72-03. E-mail: karim.mazeau@cermav.cnrs.fr.

10.1021/jf0206668 CCC: $25.00
Published on Web 11/23/2002

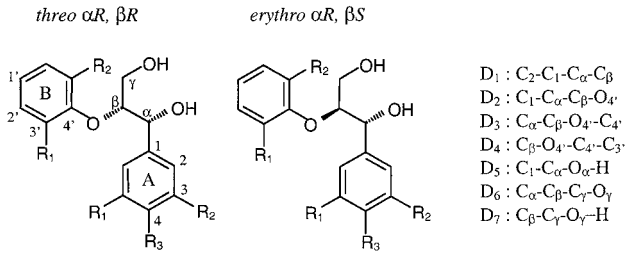

Figure 1. Structure and stereochemistry of the threo αR,βR and erythro αR,βS β-O-4 model compounds, together with the labeling of the atoms and torsion angles of interest used for the conformational searching. *p*-Hydroxyphenyl $R_1 = R_2 = R_3 = H$; guaiacyl $R_1 = OMe$, $R_2 = H$, $R_3 = OH$; syringyl $R_1 = R_2 = OMe$, $R_3 = OH$.

previous work, we have reported the conformational preferences of the threo and erythro diastereomeric forms of a guaiacyl β-*O*-4 dimer using molecular mechanics calculations with the CHARMM force field (*15*). On the basis of the satisfactory agreement obtained between the predicted conformational characteristics and the corresponding experimental data from X-ray crystal structures and $^3J_{H\alpha H\beta}$ NMR coupling constant values reported in the literature, the molecular modeling protocol has been validated. The influence of intramolecular hydrogen bonding (H-bonding) has also been extensively discussed. A consistent structural model has emerged indicating that the conformational preferences of guaiacyl β-*O*-4 dimers are predominantly governed by steric interactions rather than by differences in the H-bonding pattern. Calculations have also predicted a large variety of low-energy conformations.

In this study, we have investigated the computed conformational preferences for the threo and erythro forms of syringyl and *p*-hydroxyphenyl β-*O*-4 dimers in relation to the corresponding conformational data reported in the literature (*16*−*22*). Finally, a global comparison of the predicted and experimental conformational properties gathered for the guaiacyl, syringyl, and *p*-hydroxyphenyl β-*O*-4 structures has been performed.

MATERIALS AND METHODS

Quanta 97 software (Accelrys) was used for the building and analysis of all the structures. Energy evaluations were performed with the CHARMM 22 force field (*23*) using the PARM22 parameter file, which has extensively been validated for organic molecules. Charges were assigned to atoms by using the charge template method implemented in Quanta as the default option. All calculations were carried out on Silicon Graphics workstations running the IRIX 5.3 operating system.

Conformational searching was carried out by using a stochastic protocol, the RIS random sampling search method, operating in internal coordinates. Stochastic search procedures allow sampling of a larger conformational space for a given time when compared to molecular dynamics or systematic searching (*24*−*26*) and offer a reasonable solution to the problem of finding the global minimum and all the relevant local minima, provided the number of generated structures is large enough. For the generation of different starting geometries, geometrical parameters associated with high-energy constraints (bond lengths and bond angles) were kept constant, whereas less-constrained energy terms (torsion angles) were randomly assigned in one of the three staggered positions at values of 60°, −60°, and 180°, corresponding to gauche+, gauche−, and trans orientations, respectively. The torsion angles used for the generation of starting geometries in the conformational searching are presented in **Figure 1**. For each β-*O*-4 diastereomer, 6000 starting geometries were randomly generated and optimized. Energy minimization was achieved by using the Adopted Basis Newton Raphson algorithm with a RMS convergence criterion of 0.01 kcal/mol/Å^2. Both the nonbonded pair list and the H-bond list were periodically updated (every 10 steps) during minimization. Computations were performed by using an explicit term for H-bond interactions.

Modeling conditions correspond to the isolated state (vacuum) of the β-*O*-4 structures. Thus, no intermolecular interactions either from solvent or other molecules are considered. In the energy calculations, a distance-independent dielectric constant set to 4 was used, which is an appropriate value to represent a crystal environment for further comparison with X-ray crystal structures. This low value is also appropriate to simulate the dielectric characteristics of a $CDCl_3$ solvent environment ($\epsilon = 4.8$) in which the $^3J_{H\alpha H\beta}$ NMR coupling constants reported in the literature were measured (*20*−*22*).

In a preliminary step, the preferential orientations of the phenolic and the aromatic methoxy groups were studied for the syringyl structure. One largely favored orientation was characterized, corresponding to the methoxy in plane with the aromatic ring and, for the A aromatic ring, with the phenolic hydrogen directed toward one methoxy oxygen in a H-bond interaction. Only this orientation was considered for generating the starting geometries, leading to a simplification of the torsions set used in the conformational searching. In the same manner, the OH phenolic function was removed from the *p*-hydroxyphenyl structure as it is too far from the β-*O*-4 linkage or any other H-bond donor or acceptor group to exert any significant conformational role.

For analysis, fully minimized conformations that were within a 4 kcal/mol range above the global minimum were grouped into families based on similar values for the main torsions angles. Only the lowest energy structure of each conformational grouping has been considered in the analysis part of this paper. The relative populations of the conformers were calculated according to a Boltzmann distribution for a temperature of 298 K. Averaged theoretical $^3J_{H\alpha H\beta}$ values were calculated from the Boltzmann distribution of conformers according to the improved Karplus relationship proposed by Haasnoot et al. (*27*). This relationship allows an accurate correction for the influence of electronegativity and orientation of substituents on $^3J_{H\alpha H\beta}$ than the original Karplus equation (*28*).

RESULTS AND DISCUSSION

In this study, we present and discuss the computed conformational preferences for the threo αR,βR and erythro αR,βS diastereomeric forms of syringyl and *p*-hydroxyphenyl β-*O*-4 dimers (**Figure 1**). Energetical and geometrical characteristics of the low-energy conformers are given in **Tables 1** (syringyl) and **2** (*p*-hydroxyphenyl). Selected syringyl low-energy conformers are presented in **Figure 2**, showing the most representative conformational shapes that the syringyl structure can adopt. The results and conclusions presented in this paper also apply to the corresponding enantiomeric forms, i.e., threo αS,βS and erythro αS,βR, apart from the opposite sign for the torsion angles values.

For each structure, the small relative energy differences between the low-energy conformers indicate that a large number of conformations are accessible. It is not recommended to discard for good any conformer on the sole basis of its low relative Boltzmann population, because the energies calculated here correspond to the isolated state of the β-*O*-4 dimer and might change in the presence of other molecules. As a consequence, we must consider that all the conformers are relatively significant. Nevertheless, for each diastereomer, these conformations can easily be grouped in three major families, differing by the value of the D_2 torsion angle about the $C_\alpha C_\beta$ bond (i.e., $C_1-C_\alpha-C_\beta-O_{4'}$), trans, gauche−, or gauche+ (see **Figure 1** for the labeling of the atoms). Indeed, D_2 appears to be the major torsion angle as it predominantly governs the conformational extension of β-*O*-4 structures. These major families have been subsequently divided into sub-families according to the orientation of the second major torsion angle,

Table 1. Energetical and Geometrical Characteristics for the Threo αR, βR and Erythro αR, βS Syringyl Low-Energy Conformers

	conformer[a]	ΔE[b]	Boltzmann %[c]	torsion angles,[d] deg				$d_{C_1C_4'}$,[e] Å	donor[g]	H-bond[f]	
				D_1	D_2	D_3	D_4			d_{O-O}, Å	angle,[h] deg
						Threo					
1	**T1 (2)**	0.00	52.6	64	–162	135	116	4.67	γ-OH	2.94	176
									α-OH	2.88	164
2	**T5 (2)**	0.43	25.3	78	–178	101	–115	4.56	α-OH	2.93	177
									γ-OH	2.90	165
3	**T3′ (2)**	0.88	11.8	148	–80	92	–110	3.37	α-OH	2.97	166
									γ-OH	3.04	165
4	**T3′**	1.68	3.1	142	–78	80	–112	3.22	α-OH	2.93	170
5	**T1′ (2)**	1.95	1.9	151	–143	130	115	4.52	γ-OH	2.93	177
									α-OH	2.86	159
6	**T1**	2.06	1.6	72	–166	136	117	4.69	γ-OH	2.95	177
7	**T4**	2.20	1.3	80	–67	146	116	3.92	γ-OH	2.95	173
8	**T7**	2.25	1.2	90	–68	160	112	3.99	γ-OH	2.91	165
9	**T5**	2.78	0.5	78	–178	101	–116	4.55	α-OH	2.93	177
10	**T2**	3.04	0.3	73	–174	177	115	4.89	γ-OH	2.92	172
11	**T10**	3.46	0.2	71	173	96	71	4.55	γ-OH	2.95	167
12	**T9**	3.91	0.1	87	54	130	114	4.20	γ-OH	2.93	176
13	**T6**	4.09	0.1	93	65	–172	119	4.14	γ-OH	3.00	177
						Erythro					
1	**E3′ (2)**	0.00	70.8	54	–68	–96	113	4.19	α-OH	2.93	173
									γ-OH	2.92	167
2	**E2 (2)**	1.06	12.0	77	–74	–138	–114	4.42	γ-OH	2.94	172
									α-OH	2.89	167
3	**E3″ (2)**	1.21	9.2	–177	–55	–94	108	4.04	α-OH	2.93	169
									γ-OH	2.93	167
4	**E1**	1.70	4.0	75	–64	173	–118	4.18	γ-OH	2.98	177
5	**E2**	2.57	0.9	77	–60	–139	–115	4.31	γ-OH	2.94	174
6	**E5**	2.66	0.8	83	61	–150	–116	3.90	γ-OH	2.95	172
7	**E10 (2)**	2.85	0.6	96	–161	–91	–69	4.59	γ-OH	2.99	165
									α-OH	2.92	163
8	**E3′**	3.16	0.3	59	–68	–100	114	4.20	α-OH	2.94	175
9	**E2′**	3.18	0.3	59	–69	–138	–114	4.43	γ-OH	2.94	174
10	**E6**	3.18	0.3	71	50	–171	–115	4.01	γ-OH	2.91	172
11	**E8′**	3.36	0.2	54	60	–104	–58	3.34	γ-OH	2.90	163
12	**E7**	3.47	0.2	88	168	–129	–118	4.66	γ-OH	2.95	178
13	**E4**	3.53	0.2	98	–175	–58	116	4.33	α-OH	2.93	173
14	**E9**	3.94	0.1	102	177	–179	–115	4.90	γ-OH	2.92	173
15	**E6′**	3.97	0.1	42	48	–176	–117	4.10	γ-OH	2.92	174

[a] The number 2 in parentheses indicates the presence of two intramolecular H-bonds. [b] Relative energy (kcal/mol). [c] Boltzmann population percentage calculated at 298 K. [d] Atom labeling listed in **Figure 1.** [e] Distance between atoms C_1 (A ring) and $C_{4'}$ (B ring). [f] H-bonds between the α- or γ-hydroxyl hydrogen and the oxygen of the methoxy groups borne by the B aromatic ring. [g] Hydroxyl group corresponding to the H-bond donor hydrogen. [h] H-bond geometrical characteristics: donor–acceptor distance (hydroxyl oxygen–methoxy oxygen) and angle (hydroxyl oxygen–hydroxyl hydrogen–methoxy oxygen).

$C_\alpha-C_\beta-O_4-C_{4'}$, referred to as D_3. For instance, threo syringyl conformations **T1**, **T2**, and **T5** are all characterized by a trans orientation for D_2 and thus exhibit a rather similar global shape. However, they differ by their D_3 values (respectively 135°, 177°, and 105°), leading to different conformational extension (as pictured by the $C_1-C_{4'}$ distance listed in **Table 1**) but also leading to different H-bonding interactions. Conformers having an orientation for the A aromatic ring differing from the most frequently observed (i.e., D_1 approximately in the range 70°−90°) are noted as prime (′) for $D_1 \sim \pm 150°$ or as second (″) for $D_1 \sim 180°$.

Hydrogen Bonding. Different intramolecular H-bonds can coexist in syringyl and *p*-hydroxyphenyl *β-O*-4 dimers. A first type of intramolecular H-bond, involving only the α and γ hydroxyl groups, can be detected in some conformational families, but its existence depends on the orientation about the $C_\alpha-C_\beta$ bond. Indeed, for stereochemical reasons, conformations having a trans (threo) or a gauche+ (erythro) orientation for D_2 cannot accommodate the presence of such a H-bond. The conformational significance of this H-bond has been extensively discussed for the guaiacyl structure (*15*). Given the modeling assumptions (i.e., the isolated state) and on the basis of a critical comparison with the X-ray crystal structures, it has appeared necessary not to take this H-bond into consideration if one intends to establish a realistic distribution of energy of the conformers. Thus, the syringyl and *p*-hydroxyphenyl conformers considered in this paper do not contain any such H-bond.

A second kind of H-bond, between the α- or γ-hydroxyl hydrogen and the *β*-ether oxygen, is also present in each syringyl and *p*-hydroxyhenyl low-energy conformer. However, it is poorly directional (∼105°) and has previously been found to be too weak to influence the conformational properties of guaiacyl dimers in a significant manner. In the same way, we think that this H-bond is not a major stabilizing interaction for the syringyl and *p*-hydroxyphenyl structures.

Syringyl Structure. The energetic and geometric characteristics of the syringyl low-energy conformers are given in **Table 1**. A highly directional intramolecular H-bond, between the α- or γ-hydroxyl hydrogen and the B aromatic ring methoxy oxygen (noted OH_α---OMe or OH_γ---OMe, respectively), is present in all the low-energy conformers. With optimal geometrical features (i.e., donor−acceptor distances below 3.0 Å and angles close to 180°) and a stabilizing energy calculated to be ∼4.0 kcal/mol, this H-bond can be considered as strong. This H-bond systematically locks the B aromatic ring in an anticlinal orientation ($D_4 = \pm 120°$) rather than in the antiperiplanar

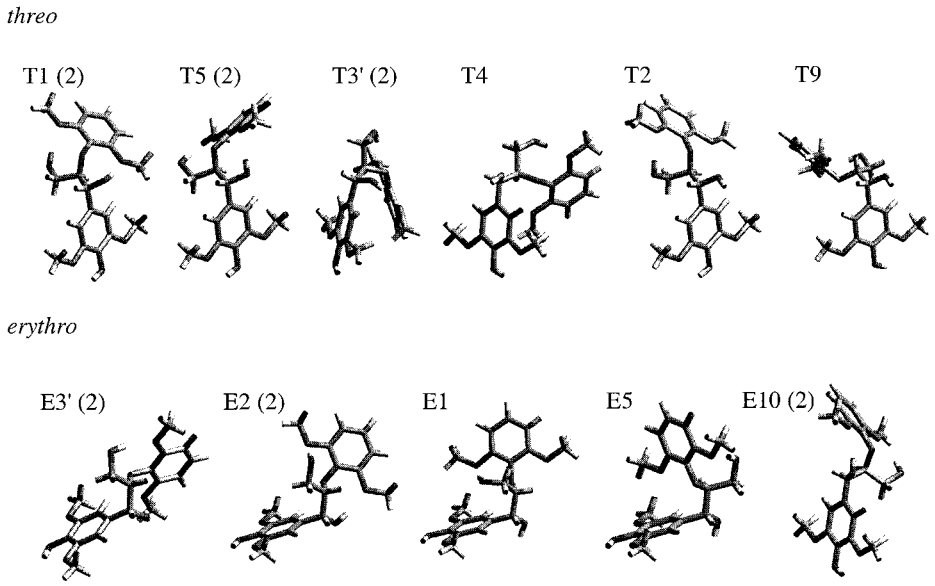

Figure 2. Selected low-energy conformers for the threo $\alpha R,\beta R$ and erythro $\alpha R,\beta S$ syringyl β-*O*-4 structure (relative energies and selected torsion angles values are given in Table 1).

Table 2. Energetical and Geometrical Characteristics for the Threo $\alpha R,\beta R$ and Erythro $\alpha R,\beta S$ *p*-Hydroxyphenyl Low-Energy Conformers

conformer		ΔE[a]	Boltzmann %[b]	torsion angles,[c] deg D_1	D_2	D_3	D_s	$d_{C_1C_4}$,[d] Å
				Threo				
1	**T5**	0.00	33.8	71	179	90	178	4.56
2	**T2**	0.35	18.8	72	–176	169	170	4.92
3	**T1**	0.44	16.2	71	–178	150	–178	4.88
4	**T11**	0.90	7.4	85	54	72	–168	3.66
5	**T7**	1.14	4.9	87	–70	167	169	4.14
6	**T10′**	1.23	4.3	48	175	79	–160	4.52
7	**T4**	1.26	4.0	80	–67	152	178	4.00
8	**T6**	1.48	2.8	90	64	169	163	4.29
9	**T1′**	1.48	2.8	47	–180	152	–179	4.92
10	**T2′**	1.65	2.1	48	–177	168	170	4.95
11	**T12**	2.09	1.0	89	–72	93	–174	3.36
12	**T4′**	2.13	0.9	44	–82	155	174	4.22
13	**T3**	2.13	0.9	63	–80	76	–140	3.18
				Erythro				
1	**E3**	0.00	36.9	80	–55	–81	173	3.82
2	**E1**	0.02	35.4	77	–64	–169	–163	4.31
3	**E4**	0.96	7.3	88	–179	–74	170	4.45
4	**E3′**	1.12	5.5	57	–66	–79	170	4.00
5	**E5**	1.26	4.4	72	56	–150	176	3.86
6	**E6**	1.37	3.6	72	52	–167	–170	4.00
7	**E8**	1.51	2.9	69	57	–95	135	3.38
8	**E9**	1.97	1.3	104	177	–167	–168	4.93
9	**E2**	1.97	1.3	77	–60	–132	–118	4.27
10	**E5′**	2.29	0.8	45	53	–150	175	3.88
11	**E6′**	2.63	0.4	45	50	–167	–170	4.04

[a] Relative energy (kcal/mol). [b] Boltzmann population percentage calculated at 298 K. [c] Atom labeling listed in **Figure 1.** [d] Distance between atoms C_1 (A ring) and $C_{4'}$ (B ring).

orientation ($D_4 = \pm 180°$) present in higher energy conformers lacking this H-bond. Some conformations (**T1**, **T5**, **T3′**, **E2**, **E3′**, and **E10**) contain an additional strong intramolecular H-bond involving the other methoxy group of the B aromatic ring. The existence of this additional H-bond has been experimentally revealed by the X-ray diffraction study of an erythro syringyl compound (*19*).

The Boltzmann populations summed up for the three distinct major families are given in **Table 3**. It shows that the threo form is characterized by a large excess of conformers having a trans orientation for D_2 (82.5%) while the erythro form predominantly exists in a gauche-D_2 orientation (97.5%). Although trans conformations are largely predominant for the threo form, gauche conformations remain significant enough.

Table 3. Computed Boltzmann Populations (%) Summed for the Three D_2 Staggered Orientations (Trans, Gauche–, Gauche+)

	threo			erythro		
	trans	gauche–	gauche+	trans	gauche–	gauche+
p-hydroxyphenyl	77.9	11.9	10.2	8.7	79.2	12.1
guaiacyl[a]	71.1	25.3	3.6	5.6	85.3	9.1
syringyl	82.5	17.4	0.1	1.0	97.5	1.5

[a] Values taken from ref 15.

***p*-Hydroxyphenyl Structure.** The energetic and geometric characteristics of the *p*-hydroxyphenyl low-energy conformers are given in **Table 2**. Due to the absence of any methoxy group on the aromatic rings, no strong H-bond is detected in the *p*-hydroxyphenyl structure. Finally, all the low-energy conformers have the B aromatic ring locked in an antiperiplanar orientation ($D_4 = \pm 180°$). The Boltzmann populations summed up for the three distinct major families (**Table 3**) indicate that the threo form predominantly exists in a trans D_2 orientation (78%), whereas the erythro form is characterized by an excess of conformers having a gauche orientation for D_2 (79%).

Comparison with the X-Ray Crystal Structures. Crystal structures that have been reported for the threo syringyl and the erythro *p*-hydroxyphenyl compounds (*16*, *18*) slightly differ from the modeled structures. They have a carboxylic acid group instead of a hydroxymethyl group at the γ-position of the propane chain. A hydroxymethyl group is also present at the C-1′ carbon of the B aromatic ring for one of the two erythro syringyl compounds (*17*). Nevertheless, despite these small structural differences, comparison with the minimized structures is quite relevant. Visual comparisons between the computed and the crystalline conformations are presented in **Figure 3**. The main geometrical parameters are compared in **Table 4**. To the best of our knowledge, no crystal structure is available for a threo *p*-hydroxyphenyl compound.

The threo syringyl crystalline structure (**TCS**) adopts a trans orientation for D_2. It has also similar torsion angle values and

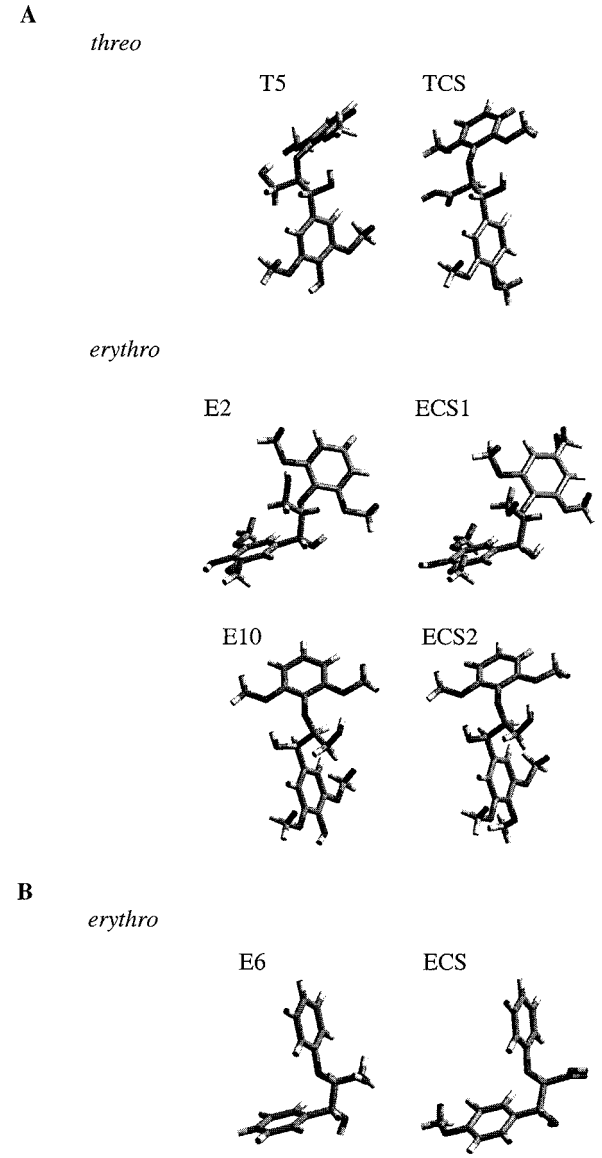

Figure 3. Comparison of the computed conformers and the crystal structures (**TCS**, **ECS**): (A) syringyl structure; (B) *p*-hydroxyphenyl structure. Torsion angles values and distances are listed in **Table 4**. RMS values (Å): (A) threo, **T5/TCS**, 0.436; erythro, **E2/ECS1**, 0.219; **E10/ECS2**, 0.325. (B) erythro: **E6/ECS**, 0.312. RMS values obtained by least-squares fit of the aromatic carbons, the propane chain carbons, and the ether oxygen.

the same intramolecular H-bonding pattern than the two crystal structures reported for the guaiacyl (*30*, *31*). The T5 conformer compares favorably with the crystal structure. The slight differences between the two conformations are similar to those previously observed and explained in the study of the guaiacyl structure (*15*). The erythro syringyl can adopt two orientations for D_2, gauche− or trans, rather than the preferred gauche+ orientation reported for the guaiacyl and *p*-hydroxyphenyl structures (*18*, *29*). As seen in **Figure 3**, there is a satisfactory similarity between the crystal structures (**ECS1** and **ECS2**) and the computed conformers (**E2** and **E10**), also confirmed by the low calculated RMS values (0.219 and 0.325 Å, respectively). It is important to notice that **ECS1** and **ECS2** conformations are fundamentally different. Indeed, **ECS2** adopts a trans orientation for D_2 rather than the gauche− one in **ECS1**. Nevertheless, both conformations are characterized by a similar conformational extension (pictured by the $C_1-C_{4'}$ distance), in other respects much larger than the ones found for the guaiacyl and the *p*-hydroxyphenyl compounds. This suggests that the repulsion between the bulky aromatic groups plays an important role in the conformational preferences of the syringyl structure. The fact that the erythro syringyl structure can exist in two different conformations in the solid state also leads to two other conclusions. First, it is an indication of the flexibility of *β-O-4* linked structures expected from the numerous low-energy conformers established by molecular modeling. Second, it indicates that intramolecular H-bonding is not the determining factor that exclusively governs the conformational preferences of the erythro syringyl structure. Indeed, H-bond interactions are not strong enough to lock the erythro syringyl in a single preferential conformation. As strong H-bond interactions having similar optimal geometries are present in the two crystal structures, almost the same stabilization in energy is provided to the two conformers. Hence, these two conformations appear to be low in energy mainly because of having low steric strain associated with their dihedral angles and with the removal of the two bulky aromatic rings.

The *p*-hydroxyphenyl crystal structure (*18*) compares favorably with the **E6** minimized conformer (**Figure 3**). In particular, the antiperiplanar orientation for the B aromatic ring is observed, as correctly predicted by molecular modeling. It is encouraging to notice that this erythro conformation with a gauche+ orientation for D_2 is advantageously predicted as a low-energy structure by the CHARMM force field, whereas it was not by the MM3 force field (*8*).

The satisfactory comparison with the crystal structures suggests that the CHARMM force field is able to correctly reproduce the geometry of *β-O-4* structures. The slight divergences observed between the predicted and the experimental conformations can easily be explained by the limits of the modeling assumptions (isolated state), which cannot account for the influence of intermolecular interactions due to packing constraints in the crystal lattice. Finally, it must be emphasized that, as for the guaiacyl compound, the syringyl and *p*-hydroxyphenyl crystal structures correspond to low energy computed conformers.

Comparison of Predicted and Experimental Conformational Properties of the Guaiacyl, Syringyl and *p*-Hydroxyphenyl *β-O-4* Structures. A global comparison of the predicted and experimental conformational data established for the three structures can help to rationalize the conformational properties of *β-O-4* compounds, particularly the influence of the methoxy groups and intramolecular H-bonding. For a proper analysis, one must distinguish their influence upon the geometry and upon the distribution of energy of the conformers.

Influence of the Methoxy Groups. The predicted geometrical characteristics of each conformational family remain globally unchanged along with the number of methoxy groups borne by the aromatic groups. The most noticeable variation concerns the stabilization of the B aromatic ring in an antiperiplanar orientation ($D_4 = \pm 180°$) in the *p*-hydroxyphenyl rather than in an anticlinal orientation ($D_4 = \pm 120°$) as in the guaiacyl and the syringyl. Experimentally, this point is confirmed by the crystal structure of the erythro *p*-hydroxyphenyl compound (*18*). Evolution of the geometry for the threo **T1** and erythro **E1** low-energy conformers is presented in Figure 4. Except for the B aromatic ring orientation, the geometrical similarity is confirmed by the low calculated RMS values.

By contrast, the predicted distribution of energy of the conformers is more markedly influenced by the presence of methoxy groups. The relative energetic ranking of the syringyl

Table 4. Comparison of Selected Geometrical Characteristics between the Computed Conformers and the Crystal Structures (**CS**) for the Syringyl and the *p*-Hydroxyphenyl Model Compounds

	torsion angles,[a] deg					angle between A	H-bond[c]	
conformer	D_1	D_2	D_3	D_4	$d_{C_1C_{4'}}$,[b] Å	and B rings, deg	angle, deg	d_{O-O}, Å
					Syringyl			
threo								
T5	78	–178	101	–115	4.56	66	177/165	2.93/2.90
TCS	62	178	139	–102	4.82	94	162	2.91
erythro								
E2	77	–74	–138	–114	4.42	68	167/172	2.89/2.94
ECS1	**–114**	–71	–150	–104	4.43	63	158	3.06
E10	96	–161	–91	–69	4.59	64	165/163	2.99/2.92
ECS2	107	179	–89	–85	4.50	76	157/125	2.99/3.19
					p-Hydroxyphenyl			
erythro								
E6	72	52	–167	–170	4.00	88		
ECS	40	57	–173	–159	4.11	89		

[a] Atom labeling given in **Figure 1.** [b] Distance between atoms C_1 (A ring) and $C_{4'}$ (B ring). [c] H-bond donor (oxygen)–acceptor (oxygen) distance and angle.

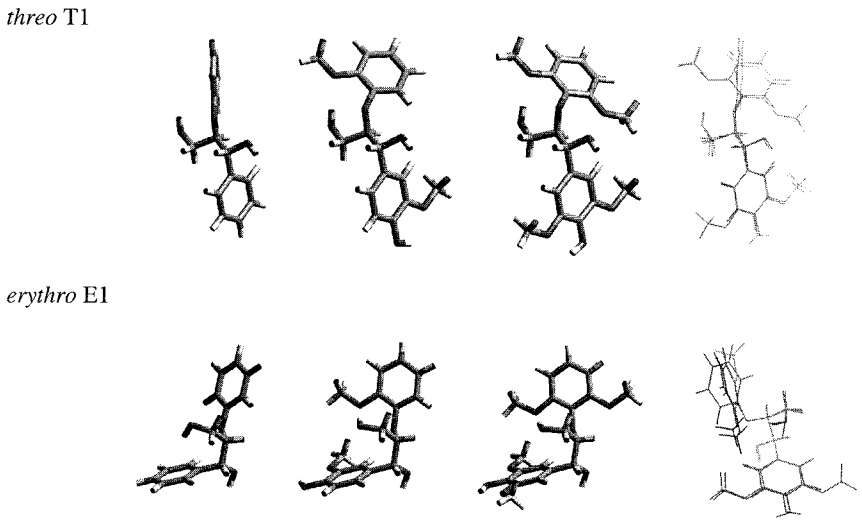

Figure 4. Comparison and superimposing of the three β-*O*-4 structures for the threo T1 and erythro E1 computed conformers. RMS values are obtained by least-squares fit of the aromatic carbons, the propane chain carbons and the ether oxygen. RMS values (Å): **T1**, H/G, 0.720; H/S, 0.695; G/S, 0.162; overall RMS H/G/S, 0.522. **E1**, H/G, 0.304; H/S, 0.398; G/S, 0.163; overall RMS H/G/S, 0.243.

conformers slightly differs from the guaiacyl one. This is mainly due to the inability for all the syringyl conformers to accommodate the presence of a second strong intramolecular H-bond to the other aromatic methoxy group. As a consequence, some low-energy guaiacyl conformations, such as threo **T2** or erythro **E1**, become less favorable for the syringyl structure. However, the syringyl globally adopts conformational preferences similar to those predicted for the guaiacyl. Indeed, as shown by the summed Boltzmann populations given in **Table 3**, the threo form is still characterized by a large excess of conformers having a trans orientation for D_2 and the erythro form also predominantly exists in a gauche− D_2 orientation. Finally, it appears that the threo gauche+ and the erythro trans conformations are much less favored than for the guaiacyl structure. The distribution of the threo *p*-hydroxyphenyl low-energy conformers is perceptibly changed in comparison with those of the guaiacyl and syringyl. Indeed, the relative ranking of the threo conformers inside each major family is inversed; i.e., the more favored conformers for the guaiacyl become the less favored ones for the *p*-hydroxyphenyl. The threo gauche+ conformers also become more favored (**Table 3**). However, the threo form still retains a predominant trans D_2 orientation and the erythro also exhibits a large amount of gauche− conformations, as for the guaiacyl and syringyl structures.

Table 5. Experimental and Computed Values of $^3J_{H\alpha H\beta}$ (Hz)

	threo			erythro		
	computed[a]	NMR[b]	CS[c]	computed[a]	NMR[b]	CS[c]
p-hydroxyphenyl	7.7	7.5		3.5	5.7	9.3
guaiacyl	7.0[d]	8.0	9.2	3.3[d]	4.8	9.1
syringyl	8.0	8.8	9.2	3.2	3.7	3.5

[a] Value calculated from the Boltzmann distribution, using the Karplus relationship reported by Haasnoot et al. (*27*). [b] Measured in $CDCl_3$. [c] Value obtained from the $H_\alpha H_\beta$ orientation in the crystal structure, using the Karplus relationship reported by Haasnoot et al. (*27*). [d] Values taken from ref 15. $^3J_{H\alpha H\beta}$ values corresponding to the three staggered orientations of D_2. Threo: 9.3 Hz (180°), 0.7 Hz (–60°), 3.9 Hz (+60°). Erythro: 2.2 Hz (180°), 2.4 Hz (–60°), 9.3 Hz (+60°).

Experimentally, the influence of the number of methoxy groups is directly depicted by the evolution of the $^3J_{H\alpha H\beta}$ value (**Table 5**). For a proper understanding of the use of $^3J_{H\alpha H\beta}$, one has to recall the connection between the $C_1-C_\alpha-C_\beta-O_{4'}$ (D_2) and $H_\alpha-C_\alpha-C_\beta-H_\beta$ ($H_\alpha H_\beta$) torsion angles. D_2 and $H_\alpha H_\beta$ orientations are identical for the threo form (e.g., a trans orientation of D_2 corresponds to a trans orientation of $H_\alpha H_\beta$), whereas they are shifted by +120° for the erythro form (e.g., a gauche+ orientation of D_2 corresponds to a trans orientation of $H_\alpha H_\beta$). As a consequence, an extended conformation (i.e., a

trans D_2 orientation) leads to a large $^3J_{H\alpha H\beta}$ value for the threo form (around 9 Hz) and to a small value for the erythro form (around 2 Hz). The values of $^3J_{H\alpha H\beta}$ corresponding to the three staggered orientations of D_2 (trans, gauche−, and gauche+) are given in **Table 5**.

The evolution of $^3J_{H\alpha H\beta}$ obtained by NMR studies in solution (*20*−*22*) clearly indicates an increase in the amount of extended conformations from the *p*-hydroxyphenyl to the syringyl structure, both for the threo and the erythro forms. This means that the presence of methoxy groups promotes extended conformations with the bulky aromatic groups distant from each other. Molecular modeling calculations and X-ray diffraction studies previously established that the favored conformations of guaiacyl *β-O*-4 dimers are not merely governed by repulsion between the bulky aromatic groups (*15*, *29*). It is clear that this conclusion also applies to the *p*-hydroxyphenyl structure, as both the predicted and experimental data presented here indicate a significant amount of compact low-energy conformers. In contrast, for the syringyl structure, the steric repulsion between the aromatic groups appears to largely influence the conformational preferences by promoting extended conformers. The influence of the methoxy groups upon the $^3J_{H\alpha H\beta}$ value measured by NMR is less pronounced for the threo form than for the erythro form. Indeed, for the threo form, extended conformations represent the favored conformational state for the three structures. As a consequence, the evolution of $^3J_{H\alpha H\beta}$ is moderate. In contrast, this influence is most noticeable for the erythro form for which compact conformations, predominant for the *p*-hydroxyphenyl structure, become less favored for the guaiacyl and much less for the syringyl.

As shown in **Table 5**, NMR data are, on the whole, in agreement with the conformers reported by X-ray diffraction studies. Indeed, the crystalline conformation systematically corresponds to the predominant D_2 conformational state expected from the $^3J_{H\alpha H\beta}$ value measured by NMR. For the erythro form, the progressive decrease of $^3J_{H\alpha H\beta}$ is also well-correlated with the change of the crystalline conformation when going from the guaiacyl to the syringyl form. Thus, a global trend emerges which suggests that the conformational preferences found in the solid state are significantly preserved in solution.

This influence of the methoxy groups is on the whole correctly predicted by molecular modeling. There is a satisfactory agreement between the experimental and computed $^3J_{H\alpha H\beta}$ values for the threo form. Indeed, the high $^3J_{H\alpha H\beta}$ values measured by NMR for the three structures indicate a predominance of a trans $H_\alpha H_\beta$ orientation, in accordance with the conformational preferences established by molecular modeling. In particular, molecular modeling correctly predicts the higher $^3J_{H\alpha H\beta}$ value for the syringyl than for the guaiacyl and the *p*-hydroxyphenyl. However, molecular modeling tends to slightly overestimate the relative population of extended conformers for the *p*-hydroxyphenyl structure. Finally, comparison of the experimental $^3J_{H\alpha H\beta}$ value measured by NMR with the value extracted from the X-ray crystal structures confirms the coexistence in solution of gauche conformations for the three structures, in accordance with the computed distribution of conformers. This predicted significance of gauche− orientations for D_2 was also experimentally confirmed by the complementary measurement of the heteronuclear long-range coupling constant $^3J_{C\gamma H\alpha}$ for a guaiacyl compound (*15*). Comparison of the experimental and computed $^3J_{H\alpha H\beta}$ values for the erythro form leads to a poorer agreement. Molecular modeling tends to underestimate the percentage of trans $H_\alpha H_\beta$ orientation for the *p*-hydroxyphenyl and the guaiacyl structures and thus predicts similar $^3J_{H\alpha H\beta}$ values for the three structures. Nevertheless, the global trend suggesting a higher $^3J_{H\alpha H\beta}$ value for the threo form in comparison with the erythro form is quite well predicted by molecular modeling. The differences observed between calculated and measured $^3J_{H\alpha H\beta}$ values can be easily explained by the fact that solvation effects might change the relative energies of conformers that have been established for the isolated state. Besides, the use of the generic and empirical Karplus relationship allows the prediction of the coupling constant with an estimated accuracy of approximately 0.4 Hz.

Finally, molecular mechanics calculations predict rather restricted distributions of energy of the low-energy conformers, suggesting that both the threo and the erythro forms can adopt a large variety of conformations. This result suggests a significant flexibility of *β-O*-4 linked structures. Experimentally, this trend is confirmed by the $^3J_{H\alpha H\beta}$ NMR coupling constant, whose values systematically indicate the coexistence in solution of conformers having at least two possible orientations for D_2, and also by the X-ray diffraction studies, with the two possible conformations found for the erythro syringyl structure.

Influence of Intramolecular H-Bonding. The same H-bonding patterns with identical optimal geometries (donor−acceptor distances and angles) are present in the guaiacyl and syringyl structures. As a consequence, the same anticlinal orientation of the B aromatic ring is retained. The additional strong intramolecular H-bond (i.e., the one involving the second methoxy group of the B aromatic ring), which is present in some syringyl conformers, is systematically less directional (∼165°) than the first one (∼175°) that is already present in the guaiacyl structure. As can be seen in **Table 2**, no significant adjustment of the torsion angles in these syringyl conformers occurs to adapt the presence of the second H-bond, so that the first H-bond always remains at its maximum strength. This suggests that this first H-bond constitutes the predominant stabilizing interaction.

For the *p*-hydroxyphenyl structure, the B aromatic ring is preferentially locked in an antiperiplanar orientation rather than in an anticlinal orientation. This can be mainly attributed to the lack of the strong H-bond involving the aromatic methoxy oxygen, rather than resulting from a simple steric effect due to the absence of the methoxy group, as such an antiperiplanar orientation for D_4 has also been identified for guaiacyl (*15*) and syringyl conformers. The absence of this strong H-bond also affects the other bond of the *β-O*-4 linkage, namely, D_3. The D_3 value for the *p*-hydroxyphenyl structure noticeably differs from the value observed in the corresponding guaiacyl conformation. The D_3 value in the *p*-hydroxyphenyl is identical to the value found for the higher energy guaiacyl conformers that lack the H-bond to the aromatic methoxy oxygen. This confirms that the torsions of the *β-O*-4 linkage are perceptibly influenced by the presence or not of a strong H-bond to the B aromatic ring methoxy oxygen.

Computed and experimental results clearly indicate that H-bonding is not the determining factor that governs the conformational preferences of *β-O*-4 dimers. Molecular modeling predicts that the energetic ranking of the low-energy conformers is predominantly governed by steric interactions rather than by differences in the H-bonding pattern. Indeed, since all the low-energy conformers contain intramolecular H-bonds with similar geometries, almost the same stabilization in energy is provided to all those conformers. Experimentally, this trend is confirmed by the X-ray studies performed on guaiacyl and syringyl compounds. The conformations of erythro and threo guaiacyl compounds were found to be independent of the

presence or not of a H-bond to the methoxy group (*29*, *30*). In the same way, the erythro syringyl has been found to exist in two possible conformations having different H-bonding patterns, although having identical optimal geometries (*17*, *19*).

In conclusion, the computations performed in this study have led to an improved rationalization of the conformational data of β-*O*-4 dimers collected by experimental techniques. From the satisfactory agreement between the predicted and experimental results, the CHARMM force field is validated for the study of β-*O*-4 structures. However, the comparison with the NMR data obtained in solution is limited by the modeling conditions chosen in this study (i.e., the isolated state). An improved prediction would require taking into account the molecular environment.

It is obvious that one cannot draw definitive conclusion about the three-dimensional structure of native lignin in the plant cell wall from the study of small model compounds. Although this further step remains still far away, our study constitutes an encouraging first step in that direction and a good basis for the building and study of larger lignocellulosic structures for which conformational data remain experimentally unattainable. To this end, modeling simulations performed in the presence of a cellulosic matrix and water molecules are under progress.

ACKNOWLEDGMENT

We are very grateful to Prof. K. Lundquist and Prof. T. J. Elder for fruitful discussions. We thank Dr J.P. Simon and Prof. K. E. L. Eriksson for useful information.

Supporting Information Available: Molecular conformations of the most representative low-energy structures for the threo and erythro *p*-hydroxyphenyl structure are presented. This material is available free of charge via the Internet at the http://www.acs.org.

LITERATURE CITED

(1) Argyropoulos, D. S.; Menachem, S. B. Lignin. In *Biopolymers from renewable resources*; Springer: New York, 1998; pp 292−322.

(2) Elder, T. J.; McKee, M. L.; Worley, S. D. The application of molecular orbital calculations to wood chemistry. The formation and reactivity of quinone methide intermediates. *Holzforschung* **1988**, *42*, 233−240.

(3) Elder, T. J. Application of computational methods to the chemistry of lignins. In *Lignin. Properties and materials*; Glasser, W. G., Sarkanen, S. Eds.; ACS Symposium Series; American Chemical Society: Washington, DC, 1989; Vol. 397, pp 262−271.

(4) Bízik, F.; Tvaroska, I.; Remko, M. Conformational analysis of ester and ether linkages in lignin-arabinoxylan complexes. *Carbohydr. Res.* **1994**, *261*, 91−102.

(5) Faulon, J. L.; Hatcher, P. G. Is there any order in the structure of lignin? *Energy Fuels* **1994**, *8*, 402−407.

(6) Faulon, J. L.; Carlson, G. A.; Hatcher, P. G. A three-dimensional model for lignocellulose from gymnospermous wood. *Org. Geochem.* **1994**, *21* (12); 1169−1179.

(7) Houtman, C. J.; Atalla, R. H. Cellulose−lignin interactions. A computational study. *Plant Physiol.* **1995**, *107*, 977−984.

(8) Simon, J. P.; Eriksson, K. E. L. A molecular mechanics investigation of lignin structure. Conformational analysis of 1-phenyl-2-phenoxy-1,3-propanediol using MM3. *Holzforschung* **1995**, *49*, 429−438.

(9) Garver, T. M.; Maa, K. J.; Marat, K. Conformational analysis and 2D NMR assignment strategies for lignin model compounds. The structure of acetoguaiacyl-dehydro-diisoeugenol methyl ether. *Can. J. Chem.* **1996**, *74*, 173−184.

(10) Simon, J. P.; Eriksson, K. E. L. The significance of intramolecular hydrogen bonding in the β-*O*-4 linkage of lignin. *J. Mol. Struct. (THEOCHEM)* **1996**, *384*, 1−7.

(11) Shevchenko, S. M.; Bailey, G. W. The mystery of the lignin-carbohydrate complex: a computational approach. *J. Mol. Struct. (THEOCHEM)* **1996**, *364*, 197−208.

(12) Simon, J. P.; Eriksson, K. E. L. Computational studies of the three-dimensional structure of guaiacyl β-*O*-4 lignin models. *Holzforschung* **1998**, *52*, 287−296.

(13) Houtman, C. J. What factors control dimerization of coniferyl alcohol? *Holzforschung* **1999**, *53*, 585−589.

(14) Russell, W. R.; Forrester, A. R.; Chesson, A. Predicting the macromolecular structure and properties of lignin and comparison with synthetically produced polymers. *Holzforschung* **2000**, *54*, 505−510.

(15) Besombes, S.; Robert, D.; Utille, J. P.; Taravel, F. R.; Mazeau, K. *Holzforschung*, accepted for publication.

(16) Lundquist, K.; Stomberg, R.; von Unge, S. Stereochemical assignment of the threo and erythro forms of 2-(2,6-dimethoxyphenoxy)-1-(3,4-dimethoxyphenoxyl)-1,3-propanediol from X-ray analyses of the synthetic intermediates (*Z*)-2-(2,6-dimethoxyphenoxy)-3-(3,4-dimethoxyphenyl)-2-propenoic acid and *threo*-2-(2,6-dimethoxyphenoxy)-3-(3,4-dimethoxyphenoxyl)-3-hydroxypropanoic acid. *Acta Chem. Scand.* **1987**, *B41*, 499−510.

(17) Stomberg, R.; Lundquist, K. On the stereochemistry of lignin model compounds of the arylglycerol-β-aryl ether type: crystal structure of *erythro*-1-(4-hydroxy-3,5-dimethoxyphenyl)-2-(4-hydroxymethyl-2,6-dimethoxyphenoxy)-1,3-propanediol, $C_{20}H_{26}O_9$. *J. Crystallogr. Spectrosc. Res.* **1989**, *19* (2), 331−339.

(18) Johansson, A.; Lundquist, K.; Stomberg, R. Stereochemistry of arylglycerol β-aryl ethers. Crystal structure of *erythro*-3-hydroxy-3-(4-methoxyphenyl)-2-phenoxypropanoic acid. *Acta Chem. Scand.* **1992**, *46*, 901−905.

(19) Langer, V.; Lundquist, K. *erythro*-2-(2,6-Dimethoxyphenoxy)-1-(3,4,5-trimethoxyphenyl)-1,3-propanediol. *Acta Cryst E* **2001**, *57*, o1219-o1221.

(20) Sipilä, J.; Syrjänen, K. Synthesis and ^{13}C NMR spectroscopic characterisation of six dimeric arylglycerol-β-aryl ether model compounds representative of syringyl and *p*-hydroxyphenyl structures in lignins. On the aldol reaction in β-ether preparation. *Holzforschung* **1995**, *49*, 325−331.

(21) Kawai, S.; Okita, K.; Sugishita, K.; Tanaka, A.; Ohashi, H. Simple method for synthesizing phenolic β-O-4 dilignols. *J. Wood Sci.* **1999**, *45*, 440−443.

(22) Ralph, S.; Landucci, L.; Ralph, J. NMR database of lignin and cell wall model compounds. http://www.dfrc.wisc.edu/.

(23) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.

(24) Saunders, M. Stochastic exploration of molecular mechanics energy surfaces. Hunting for the global minimum. *J. Am. Chem. Soc.* **1987**, *109*, 3150−3152.

(25) Saunders, M.; Houk, K. N.; Wu, Y. D.; Still, W. C.; Lipton, M.; Chang, G.; Guida, W. C. Conformations of cycloheptadecane. A comparison of methods for conformational searching. *J. Am. Chem. Soc.* **1990**, *112*, 1419−1427.

(26) Chang, G.; Guida, W. C.; Still, W. C.; An internal coordinate Monte Carlo method for searching conformational space. *J. Am. Chem. Soc.* **1989**, *111*, 4379−4386.

(27) Haasnoot, C. A. G.; de Leeuw, F. A. A. M.; Altona, C. The relationship between proton−proton NMR coupling constants and substituent electronegativities. I. *Tetrahedron* **1980**, *36*, 2783−2792.

(28) Karplus, M. Vicinal proton coupling in NMR. *J. Am. Chem. Soc.* **1963**, *85*, 2870−2871.

(29) Stomberg, R.; Lundquist, K. Stereochemistry of lignin structures of the β-O-4 type. Crystal structures of model compounds. *Nord. Pulp Pap. Res. J.* **1994**, *1*, 37−43.
(30) Lundquist, K.; Li, S.; Stomberg, R. Stereochemistry of lignin structures of the arylglycerol β-aryl ether type. Crystal structures of the threo form of guaiacylglycerol β-guaiacyl ether and its triacetate. *Nord. Pulp Pap. Res. J.* **1996**, *1*, 43−47.
(31) Stomberg, R.; Hauteville, M.; Lundquist, K. Studies on lignin model compounds of the β-*O*-4 type: crystal structures of *threo*-1-(4-hydroxy-3,5-dimethoxyphenyl)-2-(2-methoxyphenoxy)1,3-propanediol and 3-hydroxy-1-(4-hydroxy-3,5-dimethoxyphenyl)2-(2-methoxyphenoxy)-1-propanone-methanol (1/1). *Acta Chem. Scand*. **1988**, *B42*, 697−707.

Received for review June 14, 2002. Revised manuscript received September 19, 2002. Accepted September 20, 2002. S.B. is supported by a grant from the French Ministère de l'Enseignement Supérieur et de la Recherche.

JF0206668