Structural Elucidation of New Dimeric Ellagitannins from *Quercus robur* L. Roburins A–E

Catherine L. M. Hervé du Penhoat,[a,*] Veronique M. F. Michon,[a] Shuyun Peng,[b] Carole Viriot,[b] Augustin Scalbert[b] and Douglas Gage[c]
[a] *Laboratoire de Chimie, Ecole Normale Supérieure, 24 rue Lhomond, Paris 75231, France*
[b] *Laboratoire de Chimie Biologique (INRA) I.N.A.–P.G., Thiverval-Grignon 78850, France*
[c] *Department of Biochemistry, Biochemistry Building, Michigan State University, East Lansing, Michigan 48824-1319, USA*

Eight ellagitannins from *Quercus robur* L. wood have been studied by high-resolution ^1H and ^{13}C NMR and FAB-MS spectroscopy. Three of these polyphenols are known compounds, namely, castalagin, vescalagin, and grandinin. Five of the ellagitannins are new oligomeric compounds containing vescalagin or castalagin moieties bonded to a pentose, lyxose or xylose. The inter-unit linkages are carbon–carbon bonds between C-1 of the glucosyl residue of one unit and either the C-2′ of the hexahydroxydiphenoyl (HHDP) group of a second unit or the C-1 of a pentosyl sugar.

Ellagitannins are a large group of polyphenolic compounds widely distributed in plants. An increasing interest in the role of these particular metabolites in the therapeutic action of traditional medicines of China and Japan[1,2] has led to a rapid growth of knowledge in this area. Correlation of the biological activity of tannins with their structure has been made possible due to their isolation by reversed-phase HPLC methods and structural determination with high-resolution NMR and FAB-MS techniques. Oligomeric hydrolysable tannins have been recently classified by Okuda *et al.*[3] according to both the types of monomeric structures present and the condensation modes between the monomers.

Heartwood of *Quercus robur* L. (pedunculate oak, English oak) is known to contain about 10% (w/w) of ellagitannins[4] which are responsible for the high durability of this wood.[5] These HHDP esters also contribute to the taste and the colour of brandies and wines aged in oak barrels.[6–8] Castalagin **1** and vescalagin **2**, the two main HHDP esters, were originally characterized by Mayer[9] but the presence of at least six other ellagitannins of unknown structure was reported recently.[4] These compounds have been purified by a combination of chromatography on Sephadex LH 20 and reversed-phase HPLC.[10] The physicochemical data of the fastest eluted component of these unknown tannins, roburin A **3**, indicated a dimeric compound composed of two vescalagin units but unambiguous evidence for the coupling mode was lacking.[11] The present paper describes the structural investigation of this group of ellagitannins based on high-resolution NMR and on FAB-MS techniques.[12]

Results and Discussion

The HPLC and FAB–MS data of the ellagitannins from *Quercus robur* L. are collected in Table 1. The new polyphenols, roburins A–E, are of three distinct structural types based on molecular-weight criteria. Dimeric tannins, roburin A and roburin D, formally result from the condensation of two 1,2,3,5-*o*-(nonahydroxytriphenoyl) 4,6-*o*-(hexahydroxydiphenoyl) glucosyl groups with the concomitant loss of a water molecule ($M = 2 \times 934 - 18$). The molecular weight of the largest molecules, roburins B and C, is higher than that of roburin A and roburin D by a mass increment of 132 ($M = 2 \times 934 - 18 + 132$). Finally, the molecular weight of the remaining two ellagitannins, roburin E and grandinin, is higher than that of monomers **1** and **2** by the same increment, 132 ($M = 934 + 132$). All the purified tannins, when submitted to degradation in methanol–hydrochloric acid,[13] gave one mol equivalent of ellagic acid, which indicates the presence of one HHDP unit in each molecule.

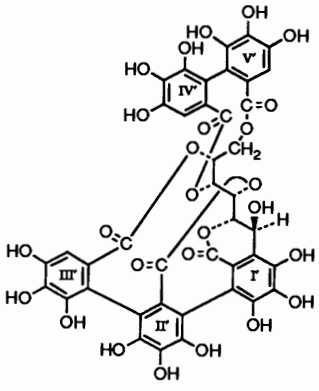
Table 1 FAB-MS, chemical and astringency data for *Quercus* ellagitannins

Polyphenol[a]	Molecular weight[b]	Number of HHDP units per molecule[c]	Astringency[d]
Roburin A **3**	1850	1	+
Roburin B **7**	1982	1	+
Roburin C **8**	1982	1	nt
Grandinin **5**	1066	1	(+)
Roburin D **4**	1850	1	+
Vescalagin **2**	934	1	(+)
Roburin E **6**	1066	1	(+)
Castalagin **1**	934	1	(+)

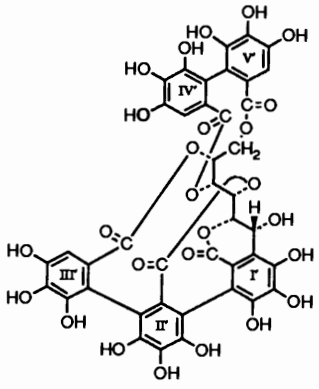
[a] In order of increasing retention time on a reversed-phase HPLC column.[10] [b] Based on the $[M - H]^-$ ions observed in the FAB-MS spectra in the negative mode with a glycerol matrix. [c] Identified as ellagic acid after acid degradation in MeOH–HCl. [d] (+), very weak; +, weak; in contrast, penta-*O*-galloyl-β-D-glucose showed strong astringency; nt, not tasted.

The NMR spectra of these ellagitannins were recorded in various deuteriated solvents {$[^2H_6]$acetone, $[^2H_4]$MeOH, and $(CD_3)_2SO$} and in the preliminary work $(CD_3)_2SO$ appeared to be the most satisfactory on the grounds of solubility. Furthermore, no overlapping solvent resonances were initially detected in the ^{13}C spectra. In all cases, fairly dilute solutions [< 60 mg/0.5 cm^3] were used as poorer spectral dispersion and an increase in linewidth were observed with higher concentrations. Although the heteronuclear experiments described here were carried out with classical carbon detection, sensitivity would be improved by reverse-detection.[14]

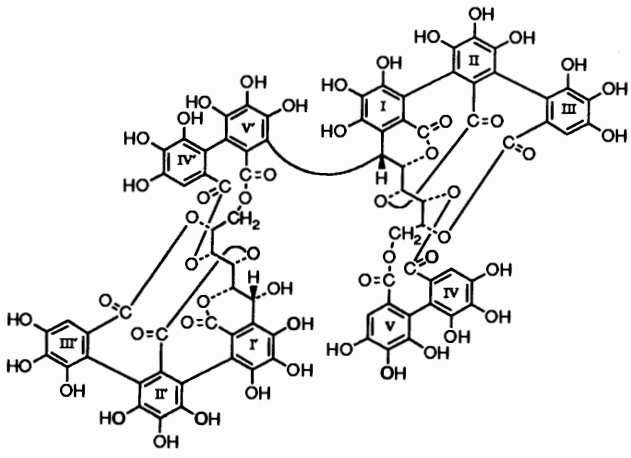
The NMR data of compounds **1** and **2** and of roburin A **3** in $(CD_3)_2SO$ were reported previously[11] and those of roburin D **4** are collected in the Experimental section. The latter two tannins each present proton signals for 2 glucosyl residues between δ 2.7 and 5.8. The glucose units with 5-H resonating at low and high field, respectively (Fig. 1), will be referred to as Glc1 and Glc2, and the numbering of the aromatic rings (I–V and I′–V′) is indicated in the formulae. The proton-coupling networks for the glucose units were extracted from (double-



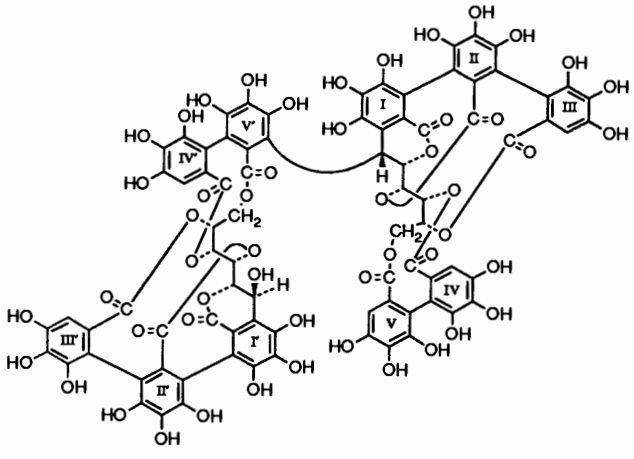
1



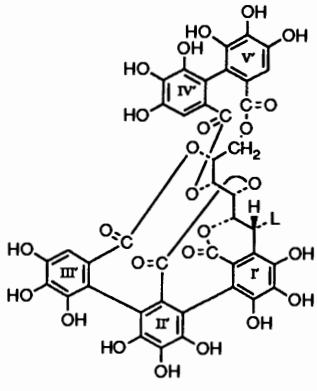
2



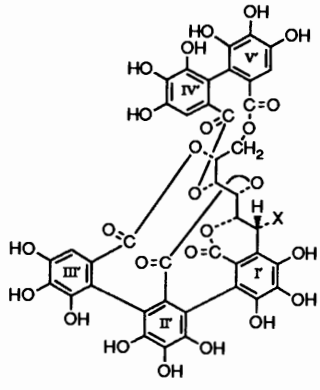
3 Roburin A



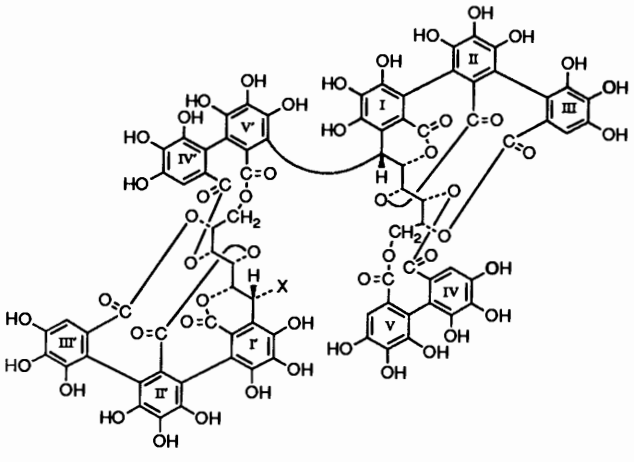
3 Roburin D



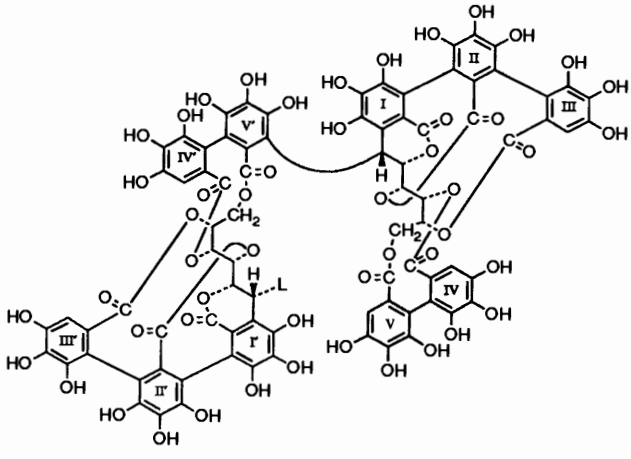
5



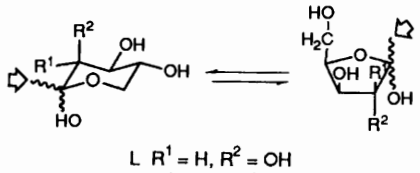
6 Roburin E



7 Roburin B



8 Roburin C



L R^1 = H, R^2 = OH
X R^1 = OH, R^3 = H

quantum-filtered COSY) DQF-COSY [δ–δ, ^{1}H–^{1}H] spectra. Both the ^{1}H chemical shifts and the $^3J_{H,H}$ coupling constants of the glycosyl residues of open-chain *C*-glucosidic ellagitannins are very different from those of compounds containing glucopyranose moieties in either the $^4C_1$[15] or $^1C_4$[16,17] conformations. Moreover, epimers **1** and **2** are readily distinguished by the values of their $^3J_{1\text{-H},\,2\text{-H}}$ coupling constants, <4.5 and 2 Hz, respectively. In the case of the glucosyl residues of roburin A **3**, the chemical shifts (with the exception of the primary hydroxy protons of Glc2) as well as the $^3J_{H,H}$ coupling constants were very similar to the data obtained for compound **2**, suggesting that this tannin was a dimer composed of two vescalagin subunits.[11] In contrast, the chemical shifts of Glc1 and Glc2 of roburin D are similar to those of compounds **2** and **1**, respectively, indicating that this dimer contains one vescalagin subunit and one castalagin subunit. However, while 6 aromatic protons are expected for the HHDP and nonahydroxytriphenoyl (NHTP) groups (3 per subunit of compound **1** or **2**), for compounds **3** and **4**, only 5 are detected. The highfield shifts of the signals of 6-H and 6-H′ of the Glc2 residue of **3** and **4**, δ -0.7 and -1.4, respectively, as compared with those of the monomers **1** and **2** are striking.

The broad-band ^{13}C NMR spectrum of compound **3**[11] and **4** (see Experimental section) contained 80–85 peaks, ostensibly twice the corresponding number in the spectra of monomeric compounds **1** and **2**. The resonances for C-2 appear at lowest field among the glucose signals of *C*-glucosidic tannins and are dependent on the configuration at C-1.[17] For roburin A **3** the signals at δ_C 76.24 and 76.84 could be attributed to the C-2 carbons of the glucose units and are analogous to the value observed for compound **2**. For roburin D **4** the signals at δ_C 73.07 and 76.31 can be attributed to C-2 of a castalagin and a vescalagin subunit, respectively. Both the number and chemical-shift range of the signals of the various groups of quaternary aromatic carbons were those expected for two 1,2,3,5-*o*-(nonahydroxytriphenoyl) 4,6-*o*-(hexahydroxydiphenoyl) glucosyl units (see Experimental section). However, unexpected signals of methine aromatic carbons and aliphatic carbons were observed, which suggested the presence of impurities. Therefore the exact number of a given type of carbon could not be unambiguously extracted from the ^{13}C NMR spectra.

In order to assign the signals of the aromatic protons, the carbonyls, and certain aromatic carbons, selective long-range heteronuclear chemical-shift correlation (INAPT[18]) experiments were performed. A 5 Hz filter was chosen to optimize polarization transfer from the glucosyl and aromatic protons to the carbonyl carbons through $^3J_{H,C}$ coupling. In the case of roburin A **3** the majority of the carbonyl resonances (all except those corresponding to carbons engaged in an ester linkage to a primary hydroxy group) and the signals of aromatic protons 2-H^{III}, 2-H^{IV}, 2-H$^{III'}$ and 2-H$^{IV'}$ were attributed. The low-field aromatic signals are those of 2-H^{IV} (Glc1) and 2-H$^{III'}$ (Glc2).[11] However, for roburin D **4**, owing to a lower signal-to-noise ratio only partial assignments could be made. As in the case of roburin A **3**, one of the low-field aromatic resonances is attributed to 2-H^{IV} (Glc1).

Recently, Nonaka *et al.*[19] published the structure of grandinin, **5**, a novel *C*-glycosidic ellagitannin in which a C_5-polyalcohol unit with a *lyxo* configuration was linked through a carbon–carbon bond to the C-1 position of the *C*-glucosyl moiety of a vescalagin unit. The $(M + H)^+$ peak in the positive-mode FAB-MS spectrum was 1067, indicating the same molecular weight as that of the low-weight ellagitannins from *Quercus robur* L. Hence, it appeared that the aforementioned mass increment of 132 might indicate the presence of a pentosyl residue. The ^{1}H NMR spectrum of grandinin **5** in [$^2H_6$]acetone was reported to be extremely complicated owing to the existence of an equilibrium mixture. It is known that the pyranose–furanose and anomeric equilibria of sugars show strong solvent dependence.[20] The equilibrium composition of solutions of the tannin **5** in common solvents was monitored by NMR spectroscopy and that of grandinin **5** in [$^2H_5$]pyridine (3 isomers; 53, 26 and 21%) (Fig. 2) was shown to contain a large proportion of the major isomer. The total ^{1}H coupling networks for the glucosyl and pentosyl residues of the major isomer of grandinin **5** (for the data of the minor isomers of compound **5** see the Experimental section) were extracted from the DQF-COSY spectrum and are collected in Table 2. The corresponding data for compound **2** were also determined. The chemical shifts and the coupling constants of the glucosyl protons of compounds **2** and **5** are very similar with the exception of the 1-H chemical shift. In the case of grandinin **5** this value is shifted by -1.2 ppm to high field. Such a shift would be expected upon replacement of the carbon-oxygen bond at C-1 of the glucose residue in compound **2** by the carbon–carbon bond in grandinin **5**.

The proton-coupling networks of the pentose moieties of the three isomers of compound **5** are the 5-spin systems expected for *C*-alkyl derivatives. For comparative purposes reported $^3J_{H,H}$-values for selected methyl α- and β-pentopyranoses in the $^4C_1$ conformation[21] and α- and β-pentofuranoses[22] are collected in Table 3. It was possible that a β-pyranose carrying a bulky vescalagin substituent might adopt the $^1C_4$ conformation instead of the $^4C_1$ conformation of methyl pentopyranose. The corresponding $^3J_{H,H}$-values were calculated for the methyl *C*-pentopyranoses according to Altona and Haasnoot[23] and are also collected in Table 3. Qualitatively, the relative values of $J_{2,3}$, $J_{3,4}$, $J_{4,5a}$ and $J_{4,5e}$ of the major isomer (small, large, large and small), which indicated that 2-OH was axial while 3- and 4-OH were equatorial, are compatible with a *lyxo*-pyranose-type structure. It is to be noted that, for the *lyxo* configuration, only the pyranose form in the $^4C_1$ conformation is expected to have a large $J_{3,4}$-value such as that (9.3 Hz) observed for the major isomer. The least abundant isomer (21%; 5-H resonating at δ 4.10, see Experimental section) presented very small $J_{4,5'}$- (<5 Hz) and $J_{4,5''}$- (<3 Hz) values compatible with the β-pyranose form in the $^1C_4$ conformation. Finally, the last isomer (26%) presented a tight AB spin system such as is often observed for primary hydroxy protons, suggesting the furanose form. The ^{13}C NMR spectrum of grandinin **5** is analogous to that of compound **2** with the exception of the glycosyl region. Partial data are collected in Table 4 and the characteristic signal of the hemiacetal carbon (major isomer) resonates at δ_C 103.42. Previously reported values[19] for the C-1 signals for an equilibrium mixture of grandinin **5** in [$^2H_6$]acetone were δ_C 102.5 and 103.7.

Roburin E **6**, the ellagitannin with the same molecular weight as that of grandinin but with a higher retention time on the reversed-phase column,[10] existed almost as a single isomer in [$^2H_5$]pyridine (Fig. 2). The ^{1}H NMR data of the glucosyl residue were similar to those of compounds **2** and **5**. However, the ^{1}H NMR chemical shifts of the pentosyl moiety of compound **6** were very different from those of grandinin **5** (Table 2). The coupling constants were compatible with the pyranose form but differed from those of grandinin **5**. In particular, the $^3J_{2\text{-H},3\text{-H}}$-value of 9.3 Hz for roburin E **6** instead of that of 3.4 Hz for grandinin **5** suggested that, for the former, 2-OH was equatorial and thus that compound **6** presented a *xylo* configuration. These $^3J_{H,H}$-values are analogous to those reported for methyl α- and β-*xylo*-pentopyranose[21] which are collected in Table 3. The ^{13}C NMR spectrum of compound **6** is similar to that of grandinin and the chemical shift of the hemiacetal carbon resonance is δ_C 101.86.

Steady-state NOE experiments were performed on samples of compounds **5** and **6** under conditions of incomplete saturation (300 ms) and the corresponding NOE-values which have not

Table 2 400 MHz ^{1}H NMR data for *Quercus* ellagitannins in [$^2H_5$]pyridine[a]

Protons	Compound 2	4	5[e]	6	7	8
Aromatic[b]						
H^{III}		7.321			7.299	7.327
H^{IV}		7.715			7.696	7.752
H^{V}		7.082			7.056	7.072
$H^{III'}$	7.286	8.407	7.265	7.265	8.291	8.369[g]
$H^{IV'}$	7.279	7.165	8.363	7.780	8.051	8.051[g]
$H^{V'}$	6.995		6.965	6.998		
Glc1[c]						
1-H		4.88 (br s)			4.89 (br s)	4.945 (br s)
2-H		5.57 (br s)			5.58 (br s)	5.547 (br s)
3-H		5.33 (d, 7.3)			5.32 (d, 7.8)	5.330 (d, 7.3)
4-H		6.19 (t, 7.1)			6.15 (m)[f]	6.163 (t, 7.3)
5-H		6.22 (d, 7.8)			6.223 (m)	6.245 (br d, 6.8)
6-H		5.24 (d, 11.2)			5.24 (br d, 11.2)	5.275 (br d, 11.2)
6-H′		4.18 (d, 11.7)			4.18 (d, 11.7)	4.17 (br d, 12.2)
Glc2[c]						
1-H	5.535 (d, 2)	5.82 (br s)	4.289 (br s)	3.987 (br s)	4.17 (br s)	3.887 (br s)
2-H	6.05 (~br t)	5.36 (br s)	6.390 (br s)	6.547 (br s)	6.10 (br s)	6.301 (br s)
3-H	5.417 (d, 7.3)	5.88 (d, 7.3)	5.42 (d, 7.3)	5.444 (d, 7.6)	5.20 (d, 6.3)	5.337 (d, 6.8)
4-H	6.119 (t, 7.1)	6.00 (t, 6.6)	6.198 (t, 6.8)	6.098 (t, 7.1)	5.90 (t, 6.4)	5.917 (t, 6.8)
5-H	6.409 (br d, 7.2)	5.86 (d, 7.3)	6.37 (d, 7)	6.290 (d, 7.8)	5.79 (br d, 6.8)	5.845 (br s, 6.8)
6-H	5.331 (dd, 12.7, 2.3)	3.96 (d, 11.2)	5.259 (dd, 12.5, 1.5)	5.137 (br d, 12.7)	3.79 (d, 11.2)	3.694 (br d, 12.2)
6-H′	3.986 (d, 12.7)	2.83 (d, 12.7)	3.93 (d, 12.2)	3.913 (d, 12.7)	2.64 (d, 12.7)	2.707 (br d, 12.7)
Pentose[d]						
2-H			5.112 (d, 3.4)	4.806 (d, 9.3)	5.086 (d, 3.4)	4.825 (d, 9.3)
3-H			4.909 (dd, 9.3, 3.4)	4.667 (t, 9)	4.971 (br dd, 8, 3)	4.649 (t, 9)
4-H			4.699 (td, 9.6, 3.7)	4.283 (m)	4.688 (m)	4.456 (complex m)
5-H[a]			4.471 (t, 10.8)	4.544 (t, 10.8)	4.477 (t, 10.7)	4.578 (br t, 11.2)
5-H[e]			4.210 (dd, 10.5, 5.6)	4.974 (dd, 10.5, 5.5)	4.226 (dd, 10.2, 5.8)	5.140 (complex m)

[a] δ_H 8.700 for the low-field pyridine signal. [b] From INAPT spectra. [c] From DQF-COSY spectrum. [d] From DQF-COSY and NOE difference spectra. [e] Also 2 minor isomers (see Experimental section). [f] Overlapping peaks. [g] Aromatic protons assigned by comparison with compound **6**.

been scaled up[24] are collected in Table 5. A longer pre-saturation time (5 s) led to a three-fold increase in the effect on the 4-H signal of the xylose residue of roburin E **6** upon irradiation of 2-H. However, in this case, indirect effects were observed for the glucose protons. The inversion at C-2 was corroborated by the relative values (small and large) of the NOE effects at 4-H upon irradiation of 2-H for compounds **5** and **6**, respectively. The inter-residue effects observed for 1-H and 2-H of glucose upon irradiation of 2-H of the pentopyranoses confirmed the branchpoints and the sequences of the isomers of grandinin **5**.

The molecular weight of roburins B and C (M = 1850 + 132) suggested that these compounds would be dimeric tannins with pentosyl substituents. The proton-coupling networks which were extracted from the DQF-COSY spectra are given in Table 2 while the ^{13}C NMR data are collected in Table 4. Both compounds contained two vescalagin subunits as evidenced by the glucosyl $^3J_{1\text{-H},2\text{-H}}$- and $\delta_{C\text{-}2}$-values. It is to be noted that, as in the case of roburins A and D, the primary hydroxy protons of Glc2 of roburins B and C were shifted to high field (δ −0.7 and −1.4). Inspection of the ^{1}H spectra taken in [$^2H_5$]pyridine (Fig. 2) indicated that these tannins were also structurally related to compounds **5** and **6**, respectively. The ^{1}H chemical shifts of the pentosyl moieties of compound **7** were almost identical with those of grandinin **5** while the shifts of compound **8** were very similar to those of roburin E **6**. In particular, the $^3J_{2\text{-H},3\text{-H}}$- and $\delta_{C\text{-}1}$-values of the pentosyl residues of compounds **7** and **8** were similar to those of compounds **5** and **6**, respectively.

Table 3 Reference and calculated $J_{H,H}$-values (Hz) for $^4C_1$ and $^1C_4$ pyranose forms and furanose forms of lyxose and xylose

Compound	Ref.	$J_{2,3}$	$J_{3,4}$	$J_{4,5}$	$J_{4,5'}$
Lyxose derivatives					
Methyl-D-pentosides					
α-Lyx*p*	21	3.8	4.0	9.0	4.8
β-Lyx*p*	21	3.8	7.5	7.5	4.0
α-Lyx*f*	22	4.8	4.3	6.7	4.4
β-Lyx*f*	22	5.0	4.6	7.6	4.5
$^4C_1$ *C*-alkyl-Lyx*p*	*a*	3.1	10.3	11.0	5.4
$^1C_4$ *C*-alkyl-Lyx*p*	*a*	3.1	3.6	0.8	1.3
Xylose derivatives					
Methyl-D-pentosides					
α-Xyl*p*	21	10.0		11.0	5.0
β-Xyl*p*	21	9.5	9.5	11.0	5.5
α-Xyl*f*	22	5.5	6.1	3.8	6.0
β-Xyl*f*	22	1.7	5.1	4.4	7.6
$^4C_1$ *C*-alkyl-Xyl*p*	*a*	9.8	9.8	10.5	5.4
$^1C_4$ *C*-alkyl-Xyl*p*	*a*	3.6	3.6	0.8	1.3

[a] Theoretical values for either α- or β-methyl-*C*-pentopyranose calculated according to ref. 23.

Table 4 Partial 100 MHz ^{13}C chemical shifts of ellagitannins **2**, **5**, **6**, **7** and **8** in [$^2H_5$]pyridine[a]

	Compound				
Carbon	**2**	**5**	**6**	**7**	**8**
Glycosyls					
C-1 of Glc1				42.03	
C-1 of Glc2		48.18	51.28	47.78	51.55
C-1 of pentose		103.42	101.86	103.37	101.85
Carbonyls[b]					
Glc2					
2-H to I′	165.63	166.07	166.13	166.05	165.04
3-H to II′	167.16	166.92	167.08	166.99	
4-H to IV′	167.79	167.69	167.90	167.28	
5-H to III′	168.11	168.06	168.02	168.30	
6-H to V′	170.27	170.41	170.10		
Glc1					
2-H to I				165.85	165.90
3-H to II				166.78	166.89
4-H to IV				168.16	167.94
5-H to III				168.10	167.96
6-H to V				170.44	170.29

[a] δ_C 124.2 for the central peak of the high-field pyridine signal. [b] Proton connectivity through $^3J_{C,H}$ correlation.

Evidence for the condensation modes for compounds **5**–**8**, obtained from the long-range heteronuclear chemical-shift correlation spectra, is collected in Table 6. Long-range coupling between 1-H of glucose and C-1 of the pentosyl moiety indicated C–C coupling at the anomeric carbons of these glycosyl units and in the case of tannins with two vescalagin subunits, the C_5-polyalcohol unit was always attached to C-1 of Glc2. The branchpoint between the subunits could be inferred from the existence of two long-range coupling constants between 1-H of Glc1 and two different C-2s of the aromatic rings concomitant with the absence of $H^{V'}$. This *C*-glycosylation of Glc1 by the V′ aromatic moiety of Glc2 leads to a characteristic downfield shift for the C-1 signal of Glc1 which resonates at δ_C 40. A corresponding high-field shift is not observed in the ^1H NMR spectra. For example, in the case of roburin A, the Glc1 and Glc2 1-H signals are both at δ 4.87. However, similar ^1H (δ 4.84 and 4.73) and ^{13}C (δ_C 37.9 and 38) chemical shifts have been reported[25] for 1-H and C-1 of acutissimin A and acutissimin B,

Table 5 Partial NOE data for the pentosyl protons of ellagitannins **5** and **6**

	% NOE Irradiated protons			
Ellagitannin and observed protons	Glc 1-H	Pentopyranose 2-H[a]	3-H[c]	4-H[e]
5				
Glc				
1-H		0.8		
2-H	4			
3-H	8			
Pentapyranose				
2-H	0.8		7	<1
3-H		7		3
4-H		<1	4	
5-H^{ax}			3	4
5-H^{eq}				3
6				
Glc				
1-H		1		
2-H	10	2		
3-H	5			
Pentopyranose				
2-H	1		6[d]	3
3-H		6		1
4-H		3[b]	2	
5-H^{ax}		1	5	3
5-H^{eq}		<2	<0.5	2

[a] 2-H of Lyx*p* of the minor isomer of grandinin **5** was also irradiated, resulting in 6 and 3% NOE at 3-H of Lyx*p* and 1-H of Glc, respectively. [b] 9% when the presaturation time was 5 s. [c] 6-H of Glc of Lyx*f* of grandinin **5** was also irradiated, resulting in 10% NOE at 6-H′ of Glc. [d] 2-H of Lyx*p* was undoubtedly partially saturated. [e] 5-H of Lyx*p* of the minor isomer of grandinin **5** was also irradiated, resulting in 14% NOE at 5-H′.

respectively, *C*-glycosylated ellagitannins bearing aromatic substituents. Unambiguous identification of the C-1 signals of Glc1 in the aliphatic regions of the ^{13}C spectra of compounds **3** and **4** is thwarted by the presence of signals from impurities. However, the characteristic downfield shifts of the signals of the primary hydroxy protons of Glc2 strongly suggest that the subunit linkage is the same as that of tannins **7** and **8**.

These downfield shifts of the primary hydroxy protons of Glc2 are clearly related to alkylation at C-2′ of ring V′. Examination of a Dreiding model of compound **2** suggests that the HHDP group can adopt several conformations, including a rigid, extended one with the primary hydroxy protons closest to ring V′ and a more mobile one with the primary hydroxy group situated below ring IV′ (as in formulae **3**–**8**). Molecular modelling will be necessary in order to determine the relative populations of these conformations.

During the preparation of this paper, an article[26] appeared suggesting that the configuration at the glucose C-1 position in all hitherto known *C*-glycosidic tannins should be inverted. As previously stated, the configuration at the C-1 position in these tannins has been assigned on the basis of the coupling constant between 1-H and 2-H, 5 and 2 Hz in compounds **1** and **2**, respectively, for dihedral angles of 50 and 70°. The structural revision is based on the presence (absence) of an NOE effect at 3-H upon irradiation of the 1-H glucosyl proton of vescalagin **2** (castalagin **1**). Chemical conversion of related compounds, casuarinin and stachyurin, into derivatives with the relevant structural features of compounds **1** and **2** followed by analysis of the NOE effects corroborated this contention. Our NOE

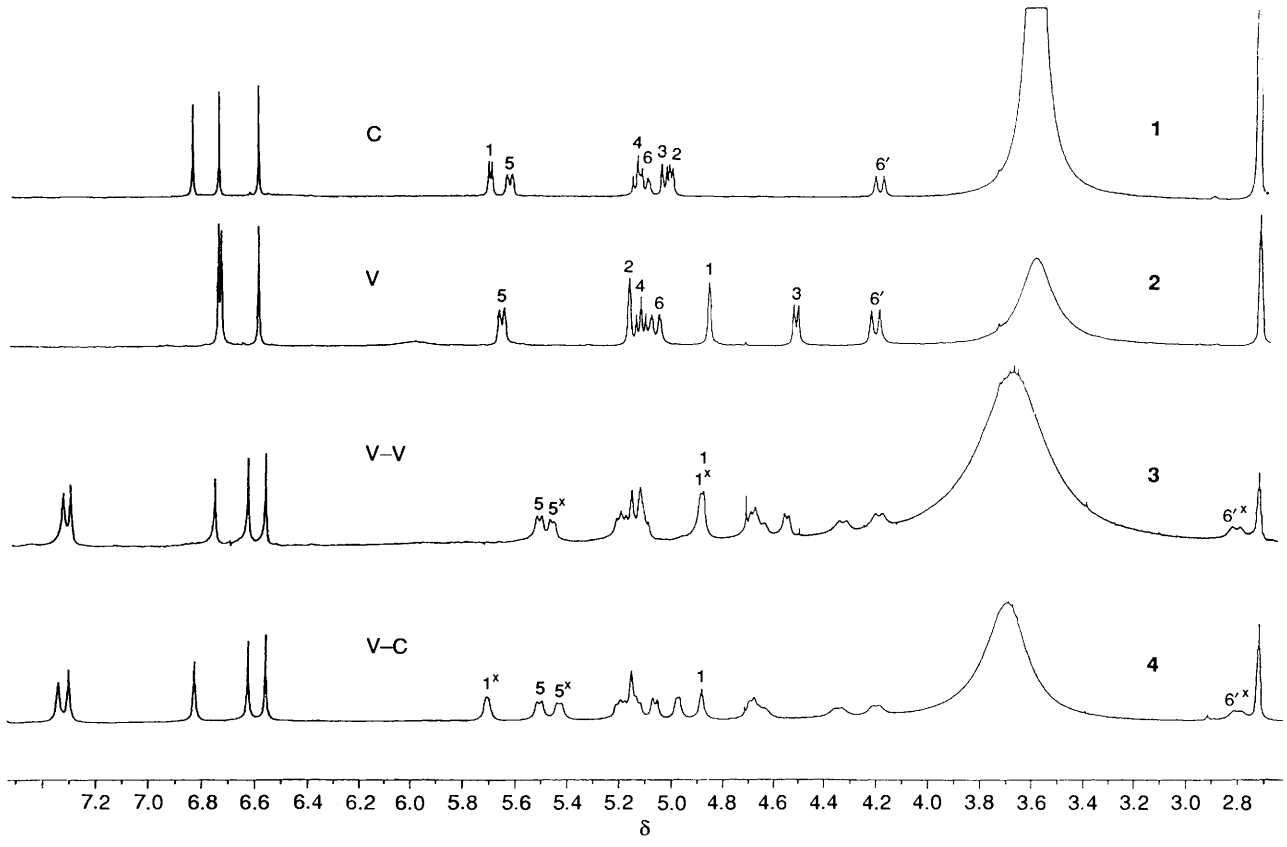


Fig. 1 400 MHz ^{1}H NMR spectra of castalagin **1**, vescalagin **2**, roburin A **3**, and roburin D **4** in $(CD_3)_2SO$ (δ_H 2.72). Key: x denotes signals of Glc2, C and V denote castalagin and vescalagin subunits respectively.

data for compounds **5** and **6**, Table 5, namely the important effect at 3-H upon irradiation of 1-H, are clearly in agreement with the revised structures although it is to be noted that only one epimer has been analysed. Furthermore, indirect effects are observed at 4-H (0.5–1%) upon irradiation of 1-H both in the present work and in the NOE difference spectrum of compound **2** that has just been published[26] suggesting that interpretation of the NOE data is not straightforward. Kinetic NOE measurements in conjunction with a molecular modelling study will be required both in order to interpret the NOE data and to describe the conformation(s) of the HHDP group in compounds **1**–**8**.

The eight ellagitannins from *Q. robur* L. wood whose structures have been determined correspond to the eight main peaks observed on HPLC chromatograms of the water-soluble fraction of wood extracts; these tannins which are not resolved by two-dimensional paper chromatography also correspond to the main spot observed by this technique.[4,10] As this water-soluble fraction exceeds 80% of the total extract dry weight,[10] we may therefore consider that the main soluble components of *Q. robur* L. heartwood have now been identified. Other soluble ellagitannins which appear as a streak on paper chromatograms[4] are also present within the wood. These ellagitannins, which cannot be resolved into any definite chemical entity by any chromatographic method tested, probably result from the non-specific oxidation of the identified ellagitannins, either during heartwood formation or heartwood ageing.[12] All the currently identified molecules belong to the *C*-glucosidic group of ellagitannins. The four dimers, roburins A, B, C and D, are the first dimers belonging to this group which have been fully described. Another dimer, alienanin B, which differs from roburin D by the lack of biphenyl linkages between aromatic residues **II** and **III**, has been previously isolated from *Quercus aliena*[27] but still awaits full description.

Table 6 Branchpoint long-range heteronuclear chemical-shift correlations [^{1}H–^{13}C] of *Quercus* ellagitannins

Ellagitannin	Long-range heteronuclear chemical-shift correlations
5	$^3J_{2\text{-H(Glc)},C\text{-}1(Lyx)}$; $^3J_{2\text{-H(Glc)},C\text{-}2'}$; $^2J_{1\text{-H(Glc)},C\text{-}1(Lyx)}$; $^3J_{1\text{-H(Glc)},C\text{-}2'}$
6	$^3J_{2\text{-H(Glc)},C\text{-}1(Xyl)}$; $^3J_{2\text{-H(Glc)},C\text{-}2'}$; $^2J_{1\text{-H(Glc)},C\text{-}1(Xyl)}$; $^3J_{1\text{-H(Glc)},C\text{-}2'}$
7	$^2J_{1\text{-H(Glc2)},C\text{-}1(Lyx)}$; two $^2J_{1\text{-H(Glc1)},C\text{-}2'}$
8	$^2J_{1\text{-H(Glc2)},C\text{-}1(Xyl)}$; two $^2J_{1\text{-H(Glc1)},C\text{-}2'}$

Key: C-2′ denotes the point of attachment of the glucose C-1 to the NHTP moiety (see Fig. 2).

In order to understand what may be the contribution of these tannins to the ageing of brandies in oak barrels, they were tasted. In comparison to gallotannins such as pentagalloylglucose, their astringency is very weak for monomers and weak for dimers (Table 1). It is possible that these tannins do not contribute directly to the taste of brandies, but rather indirectly through their complexing or reducing properties.

Experimental

NMR Spectroscopy.—The solvents [$(CD_3)_2SO$ was distilled over charcoal], internal standards and spectrometer frequencies for the NMR spectra are indicated in Figs. 1 and 2, and Tables 2 and 4. The digital resolution of the ^{1}H spectra was 0.5 Hz/point and the acquisition time was 2.01 s. *J*-Values are given in Hz. ^{13}C spectra were recorded with complete proton decoupling, an acquisition time of 1.11 s, digital resolution of 0.9 Hz/point, and a recycle time of 4.1 s. INAPT[18] spectra were acquired under similar conditions by using a 5 Hz filter for

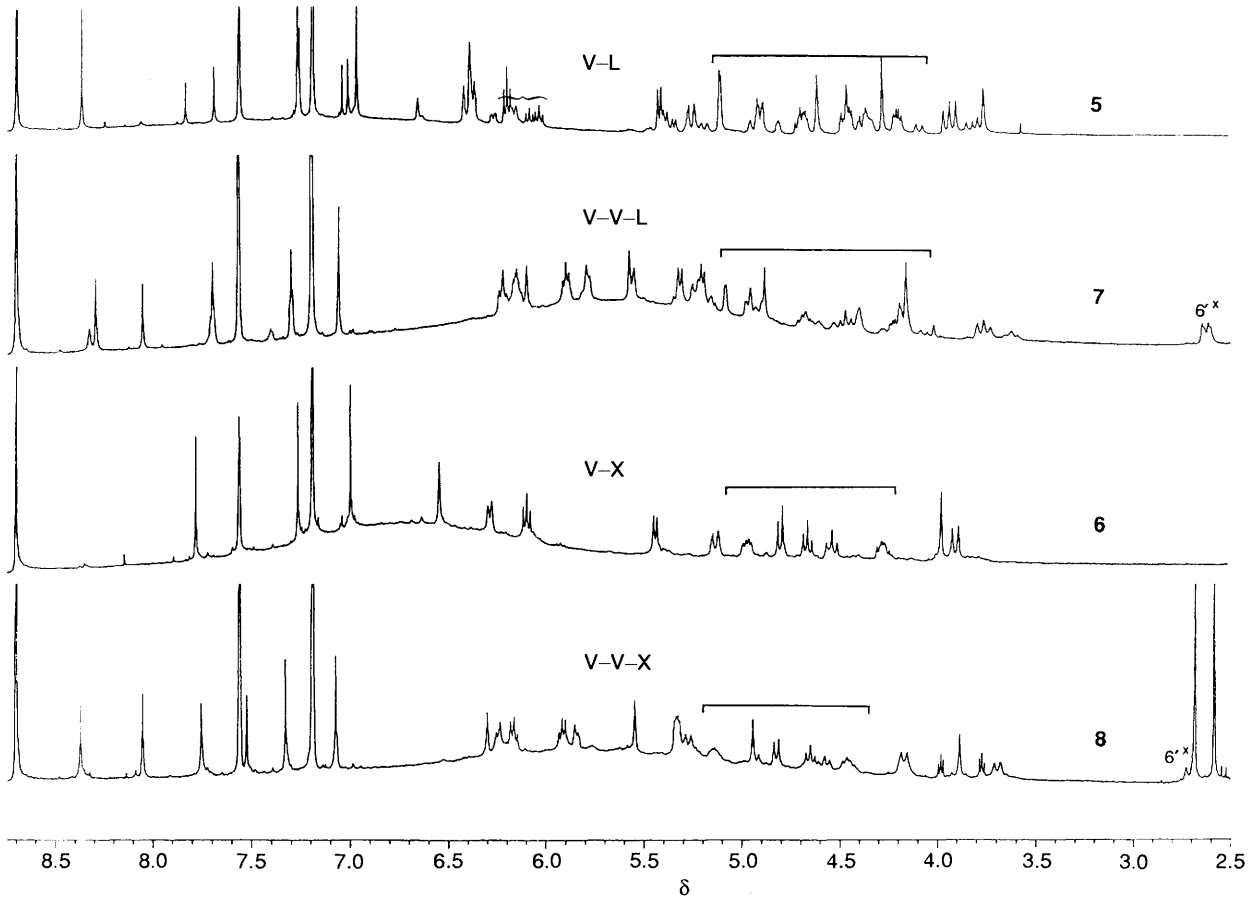


Fig. 2 400 MHz ^{1}H NMR spectra of roburins B **6**, C **7** and E **8** in [$^2H_5$]pyridine. The regions containing the signals of the pentosyl protons are indicated with a square bracket. The signals of the 4-H glucosyl protons for the three isomers of grandinin **5** at δ 6.15 are also bracketed, V, X and L denote vescalagin, xylose and lyxose subunits respectively.

polarization transfer. Double-quantum-filtered phase-sensitive COSY experiments[28] were performed using a (90°)–(t_1)–(90°)–(90°)–(FID, t_2) sequence. The spectral width in F1 and F2 was generally 2048 Hz; the number of data points in F2 was 1024, and 512 increments were recorded. The 90° pulse was 17 μs and the total acquisition time was ~16 h. Before Fourier transformation, the data were multiplied with a $\pi/2$-shifted squared sine bell. Zero filling was applied in F1. Steady-state NOE experiments were performed by applying low-power irradiation at the offset frequency of the saturated spin for 300 and 500 ms, for compounds **5** and **6** respectively, prior to acquisition. T_1-Values for compound **6** were 1–2 s for the methine protons and therefore a steady state was not attained in these experiments. Even under these conditions an indirect effect (0.5–1%) was observed for 4-H (Glc) upon irradiation of 1-H (Glc). Numerous indirect effects were observed at the glucose protons when longer saturation times were used.

FAB-MS Spectroscopy.—Spectra were acquired on a double-focussing magnetic sector instrument in the negative mode with a resolution of 3.000, an acceleration voltage of 10 kV and a glycerol matrix.

Extraction and Purification of Tannins.—Castalagin **1**, vescalagin **2**, grandinin **5**, and roburins A–E **3**, **4**, **6**, **7** and **8** were extracted by aq. methanol from *Q. robur* L. heartwood, and purified by Sephadex LH20 chromatography and reversed-phase high-performance liquid chromatography as previously described;[10] yields were respectively 1.1, 3.1, 3.0, 0.4, 1.3, 1.7, 0.3 and 0.4 g kg^{-1} wood. UV spectra were obtained with a diode array detector coupled to analytical HPLC apparatus. All compounds had identical spectra with no maximum between 240 and 400 nm but a shoulder at 280 nm.

Roburin A **3**. δ_C([$^2H_6$]acetone) 162.27–168.27 (10 C), 142–145 (> 15 C), 133–135 (~9 C), 122.5–126.75 (>9 C), 111.13–116.13 (> 10 C), 105.5–108.7 (~5 C), 63.8–76.6 (~11 C), 44.7 (1 C) and 38.9 (1 C).

Roburin D **4**. δ_H[$(CD_3)_2SO$] Glc1: 5.50 (d, *J* 6.8, 5-H), 5.19 (~t, *J* 6.8, 4-H), 5.15 [nd (not determined), 2-H], 4.88 (br s, 1-H), 4.68 (d, *J* 7.3, 3-H), 4.65 (br d, *J* 11.7, 6-H) and 4.20 (br d, *J* 10.3, 6-H′); Glc2: 5.70 (br d, *J* < 3, 1-H), 5.43 (d, *J* 6.4, 5-H), 5.13 (t, *J* 6.8, 4-H), 5.06 (d, *J* 7.3, 3-H), 4.97 (d, *J* 3.9, 2-H), 4.34 (d, *J* 10.3, 6-H) and 2.80 (d, *J* 10.7, 6-H′); δ_C[$(CD_3)_2SO$] 162.70–168.42 (10 C), 142.45–146.54 (~20 C), 132.94–137.25 (10 C), 121.04–126.56 (~10 C), 111.59–116.05 (~15 C), 104.77–108.18 (~5 C), 76.31 (C-2 of Glc1), 73.07 (C-2 of Glc2) and 64.01–70.61 (~9 C); δ_C([$^2H_5$]pyridine) 165.00–168.61 (10 C), 144.88–148.70 (~19 C), 137.08–140.40 (~9 C), 126.25–132.18 (~8 C), 114.55–118.33 (~13 C), 108.07–111.94 (5 C), 78.96 (C-2 of Glc1), 74.95 (C-2 of Glc2), 65–72.62 (~9 C), 42.14 (1 C) and 23.82–39.77 (5 C).

Grandinin **5**. δ_H([$^2H_5$]pyridine) isomer 2 (26%): for Glc: 6.42 (s, 2-H), 6.16 (br d, *J* 7.8, 5-H), 6.03 (t, *J* 7.3, 4-H), 5.39 (d, *J* 7.3, 3-H), 4.94 (d, *J* 11.5, 6-H), 3.79 (d, *J* 11, 6-H′) and 3.78 (s, 1-H); for Lyx: 4.62 (2 H, br s), 4.47 (1 H, nd) and 4.34 (2 H, nd). Isomer 3 (21%): for Glc: 6.66 (s, 2-H), 6.27 (br d, *J* 7.8, 5-H), 6.08 (t, *J* 7.3, 4-H), 5.35 (dd, J_1 6.4, J_2 1.5, 3-H), 5.19 (~dd, J_1 12.7, J_2 2, 6-H), 3.98 (s, 1-H) and 3.85 (d, *J* 12.7, 6-H′); for Lyx: 5.11 (nd, 2-H), 4.82 (br m, 3-H), 4.70 (~dd, J_1 12, J_2 < 5, 5-H), 4.38 (nd, 4-H) and 4.10 (~dd, J_1 12, J_2 < 3, 5-H′).

Estimation of Tannin Astringency.—Purified tannins in the

dry state (~1 mg) were tasted by two people not aware of the nature of the sample. Qualitative appreciation of astringency was expressed as very weak, weak, medium or strong.

Acknowledgements

We thank the Pierre and Marie Curie University (Paris VI) and the CNRS (UA 1110) for financial support.

References

1 E. Haslam, T. H. Lilley, Y. Cai, R. Martin and D. Magnolato, *Planta Med.*, 1989, **55**, 1.
2 T. Okuda, T. Yoshida and T. Hatano, *J. Nat. Prod.*, 1989, **52**, 1.
3 T. Okuda, T. Yoshida and T. Hatano, *Heterocycles*, 1990, **30**, 1195.
4 A. Scalbert, B. Monties and J.-M. Favre, *Phytochemistry*, 1988, **27**, 3483.
5 J. H. Hart and W. E. Hillis, *Phytopathology*, 1972, **62**, 620.
6 V. L. Singleton, *Adv. Chem. Ser.*, 1974, **137**, 254.
7 P. Ribereau-Gayon, *Vitis*, 1973, **12**, 119.
8 M. Moutounet, P. Rabier, J.-L. Puech, E. Verette and J.-M. Barillère, *Sci. Alim.*, 1989, **9**, 35.
9 W. Mayer, W. Gabler, A. Riester and H. Korger, *Justus Liebigs Ann. Chem.*, 1967, **707**, 177.
10 A. Scalbert, L. Duval, S. Peng, B. Monties and C. Hervé du Penhoat, *J. Chromatogr.*, 1989, **502**, 107.
11 C. L. Hervé du Penhoat, V. M. F. Michon, A. Ohassan, S. Peng, A. Scalbert and D. Gage, *Phytochemistry*, 1991, **30**, 329.
12 A preliminary account of this work was presented at the XVth International Conference of the Group Polyphenols (Strasbourg J.I.E.P., 1990): A. Scalbert, S. Peng, B. Monties, C. Hervé du Penhoat and D. Gage, *Bulletin Liason Groupe Polyphenols*, 1990, **15**, 203.
13 S. Peng, A. Scalbert and B. Monties, *Phytochemistry*, 1991, **30**, 775.
14 A. Bax and M. F. Summers, *J. Am. Chem. Soc.*, 1986, **108**, 2093.
15 E. A. Haddock, R. K. Gupta and E. Haslam, *J. Chem. Soc., Perkin Trans. 1*, 1982, 2535.
16 J. C. Jochims, G. Taigel and O. T. Schmidt, *Justus Liebigs Ann. Chem.*, 1968, **717**, 169.
17 T. Hatano, T. Yoshida, T. Shingu and T. Okuda, *Chem. Pharm. Bull.*, 1988, **36**, 3849.
18 A. Bax, J. A. Ferretti, N. Nashed and D. M. Jerina, *J. Org. Chem.*, 1985, **50**, 3029.
19 G. Nonaka, K. Ishimaru, R. Azuma, M. Ishimatsu and I. Nishioka, *Chem. Pharm. Bull.*, 1989, **37** 2071.
20 W. Mackie and A. S. Perlin, *Can. J. Chem.*, 1966, **44**, 2039.
21 K. Bock and H. Thogersen, *Annu. Rep. NMR Spectrosc.*, 1982, **13**, 40.
22 N. Cyr and A. S. Perlin, *Can. J. Chem.*, 1979, **57**, 2504.
23 C. Altona and C. A. G. Haasnoot, *Org. Magn. Reson.*, 1980, **13**, 417.
24 D. Neuhaus and M. Williamson in *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH, Cambridge, UK, 1989, pp. 95–96.
25 K. Ishimaru, G.-I. Nonaka and I. Nishioka, *Chem. Pharm. Bull.*, 1987, **35**, 602.
26 G. Nonaka, T. Sakai, T. Tanaka, K. Mihashi and I. Nishioka, *Chem. Pharm. Bull.*, 1990, **38**, 2151.
27 M. Nakayama, G. Nonaka and I. Nishioka, The 34th Annual Meeting of the Japanese Society of Pharmacognosy, Osaka, 1987, Abstract of Papers, p. 163.
28 U. Piantini, O. W. Sorenson and R. R. Ernst, *J. Am. Chem. Soc.*, 1982, **104**, 6800.

Paper 1/00097G
Received 8*th January* 1991
Accepted 18*th February* 1991