Pergamon

0031-9422(95)00148-4

Phytochemistry, Vol. 39, No. 5, pp. 1193–1199, 1995

0031-9422/95 $9.50 + 0.00

STRUCTURE SIMULATION OF TWO ELLAGITANNINS FROM *QUERCUS ROBUR* L.

NICOLAS VIVAS,* MICHEL LAGUERRE,† YVES GLORIES,‡ GUY BOURGEOIS§ and CHRISTIANE VITRY§

DEMPTOS cooperage posted to Institut d'Œnologie, Université de Bordeaux II, 351 cours de la libération, 33405 Talence, France; †Laboratoire de Chimie Analytique, CNRS-ER n°061, UFR des Sciences Pharmaceutiques, Université de Bordeaux II, 3 Place de la victoire, 33076 Bordeaux, France; ‡Laboratoire de Chimie Appliquée, Institut d'Œnologie, Université de Bordeaux II, 351 cours de la libération, 33405 Talence, France; §CESAMO-URA CNRS n°035, Université de Bordeaux I, 351 cours de la libération, 33405 Talence, France

(*Received in revised form* 27 *January* 1995)

Key Word Index—*Quercus robur*; Fagaceae; oak; heartwood; ellagitannins; vescalagin; castalagin; molecular mechanics.

Abstract—Two major ellagitannins from *Quercus robur*, vescalagin and castalagin, were studied by ^{1}H NMR and LSI-mass spectrometry, and molecular mechanics. A plausible 3D-structure is proposed for each of the two diastereoisomers. These structures were the most stable ones found after a conformational analysis study and their calculated NMR spectra are in good accordance with measured ones.

INTRODUCTION

Heartwood of *Quercus robur* contains *ca* 10% by weight of *C*-glucosidic ellagitannins [1–3]. These hexahydroxydiphenoyl (HHDP) esters, responsible for the durability of the wood [4], protect it against fungal or bacterial decay [5] and impart an astringent taste to the wood extract [6]. The study of ellagitannins is very important because this group of hydrolysable tannins is used in different industrial and agroalimentary applications [7–9] and is partially responsible for the colour of oak wood and contributes to its commercial quality [10]. They also greatly affect the quality and composition of wines [11, 12] or spirits [13] conditioned in oak barrels.

Ellagitannins are detected by a specific colorimetric reaction of the HHDP groups [14, 15]. The two principal HHDP esters of *Q. robur* heartwood are vescalagin (**1**) and castalagin (**2**); their structures were elucidated by NMR studies by Mayer *et al.* [16, 17]. Recently, a revised study concerning the configuration of the asymmetric C-1-glucosidic carbon was published by Nonaka *et al.* [18] and Yoshida *et al.* [19]. Very few studies on the conformation and stereochemistry of ellagitannins have been reported, nevertheless a first attempt was made by Spencer *et al.* [20].

Scarce data concerning the biosynthesis, particularly the conversion of β-penta-*O*-galloyl-D-glucose (**3**) into **1** and **2** [21] and the difficulty of producing suitable crystals for X-ray analysis, prompted us to make a complete reinvestigation of the problem. This paper describes a conformational study of **1** and **2** using molecular mechanics and computer simulation, and presents reasonable structures for them in aqueous solution. A tentative biosynthetic pathway for the conversion of penta-*O*-galloyl-D-glucose into vescalagin and castalagin is also proposed.

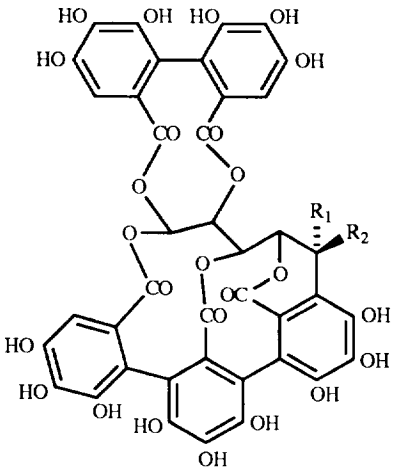
RESULTS AND DISCUSSION

Vescalagin (**1**) and castalagin (**2**) are diastereoisomers of 1, 2, 3, 4, 5, -*O*-(nonahydroxytriphenoyl)-4, 6-*O*-(hexahydroxydiphenoyl) glycosyl groups. The $[M - H]^-$ peaks in the negative mode LSI-mass spectrum were *m/z* 933 for **1** and **2** and *m/z* 939 for **3**. Okuda *et al.* [22] proposed reductive ring-cleavage and three C–C bondings starting from **3** to obtain **1** and **2**.

Several authors [20, 22, 23] consider that the biogenesis of ellagitannins, particularly the biogenetic conversion of **3** into **1** and **2**, is the result of a step-wise reaction involving a succession of intermediates, for which several formulae or structures have been postulated. However, when these compounds which are considered as the precursors of **1** and **2** are isolated, generally **1** and **2** cannot be found simultaneously. This makes the existence of these intermediates very questionable. This conclusion is confirmed in particular by the investigations of Tang *et al.* [8] and Ishimaru [24].

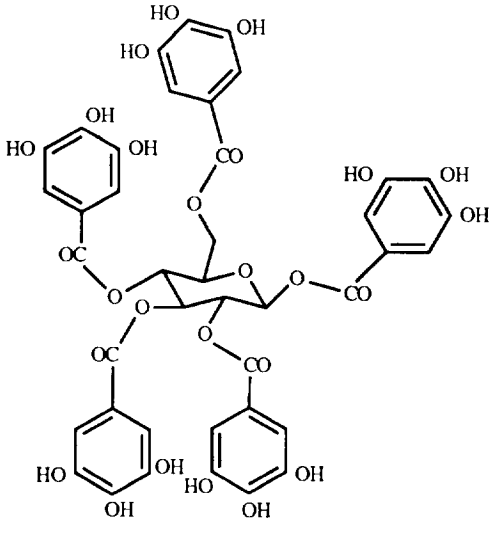
In order to build two satisfactory diastereoisomers of **1** and **2**, we found it necessary to follow a 'quasi-biosynthetic' stepwise pathway. In fact, all of our attempts to establish the model in a direct one-step procedure, starting from an open glucose moiety, failed owing to over-

*Author to whom correspondence should be addressed.



1, $R_1 = OH$, $R_2 = H$

2, $R_1 = H$, $R_2 = OH$



3

crowding and impossible spatial disposition of ring closures, so that minimization was no longer possible. Subsequently, a postulated 'biomimetic pathway' was followed. The starting structure was penta-*O*-galloylpyranoglucose (**3**) which is then opened (Fig. 1). One galloyl group is then transposed in order to let the aldehyde function appear, in a way reminiscent of the opening of the glucose molecule. Each intermediate was fully minimized as described in the Experimental before any further structural modification (MM2* force-field). In each case, the condensation step was performed between the closest carbon atoms, *i.e.* first between the galloyl groups in positions 2 and 3, then between positions 3 and 5. The remaining two steps, namely condensation between the positions 4 and 6 and position 1 *vs* the aldehyde can be conducted in any way resulting in the same global structure. It is interesting to note that during the whole process the galloyl groups in position 4 and 6 remain in a perfect parallel arrangement far from the three other groups. During condensation of the galloyl in position 2 and the aldehyde in position 1 there is a complete loss of chirality control; this process can then lead alternatively to vescalagin or castalagin (Fig. 2). These last fully minimized structures were the starting point of all calculations (see Experimental).

The best result was obtained after a Monte-Carlo search using the MM3* force-field. In this case, the phenolic groups were first omitted in order to avoid any disturbance due to misleading hydrogen-bonding. The best accordance between the calculated and the experimental coupling constants was, interestingly, obtained for the lowest energy conformer. Subsequent addition of the hydroxyls and minimization led to an even better accordance with measured data (Table 1).

It is interesting to note that, with all the possible hydrogen bonds conveniently set, castalagin is a little more stable than vescalagin (471.42 *vs* 472.98 kJ mol^{-1}). The difference arises only from the additional H-bond brought about by the enantiomeric alcoholic function in position 1; this is no longer possible in vescalagin.

Concerning the ^1H NMR spectra for **1** and **2**, there are two main differences when comparing their spectra in DMSO-d_6 (Fig. 3). In castalagin, H-1 is largely deshielded compared with its position in vescalagin (δ5.72 *vs* 4.85), as well as H-3 (δ5.05 *vs* 4.51). From a close examination of the proposed structures, it appears that in castalagin and vescalagin the H-1 protons are at the same distance from the centre of the aromatic ring I (3.50 Å), but in castalagin H-1 is only 0.53 Å above this plane (*i.e.* inside the deshielding cone), whereas in vescalagin it is 1.1 Å under the same plane (*i.e.* outside the deshielding cone). On the other hand, in castalagin, H-3 is in close vicinity to the hydroxyl in position 1 (D = 2.58 Å) capable of withstanding a strong anisotropic effect; this effect can explain the observed deshielding of this proton (Fig. 4).

During a comparison between measured NMR data (solution) and calculated coupling constants (vacuum), it was found important to check the stability of the proposed conformers in solution. They were submitted to reminimization with solvation made explicit (either using a continuum solvation model or solvent expressed, see Experimental). The results obtained are collected in Table 2.

After a complete conformational study and calculation of the NMR spectrum for each conformer, the two proposed structures are the only ones in good agreement with experimental data. It is interesting to note that if, on the one hand, they are the most stable conformers, on the other hand, they are only representative of a large family of conformers. In fact, these two ellagitannins are highly flexible molecules mainly around the glycosidic $-CH_2-$ and the large moiety containing the C-4 C-5 C-6 glycosidic linkage and the two aromatic rings IV and V. This part of the molecules can occupy a large conforma-

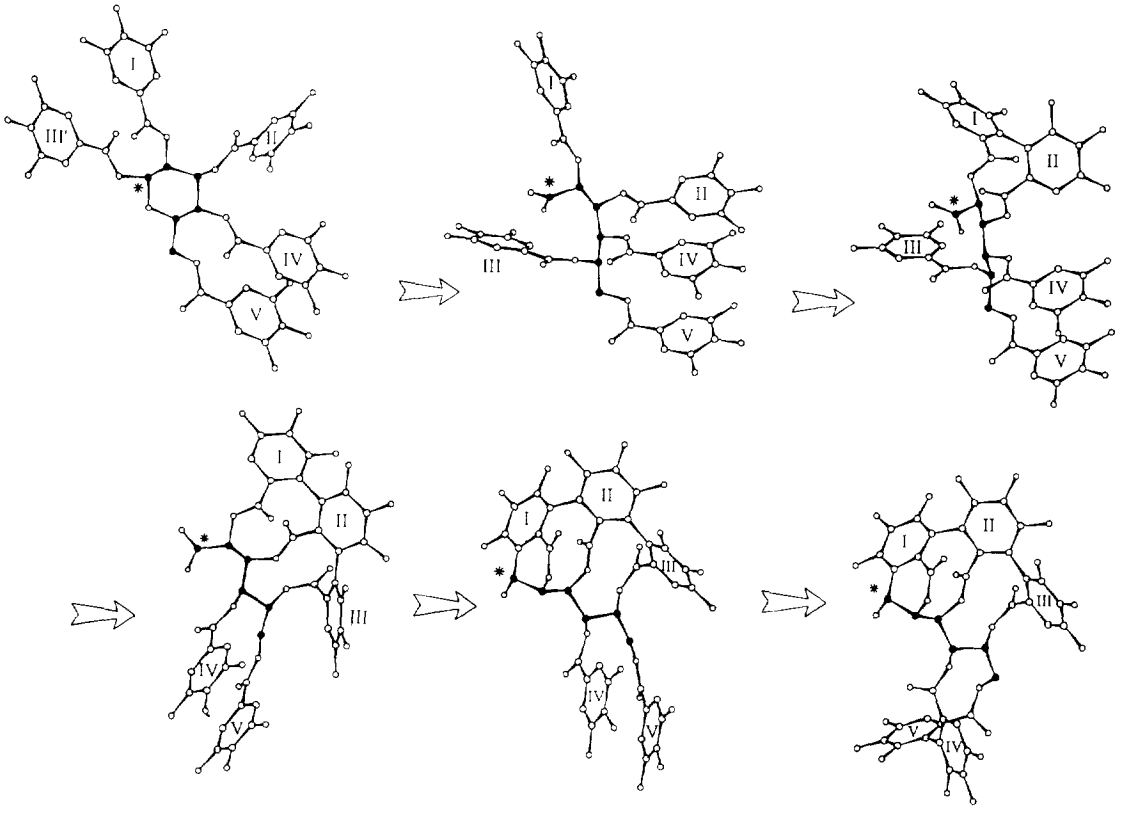


Fig. 1. Hypothetical 'biomimetic pathway' of β-pentagalloylglucose to castalagin (*C-1-glycosidic; •C-glycosidic groups).

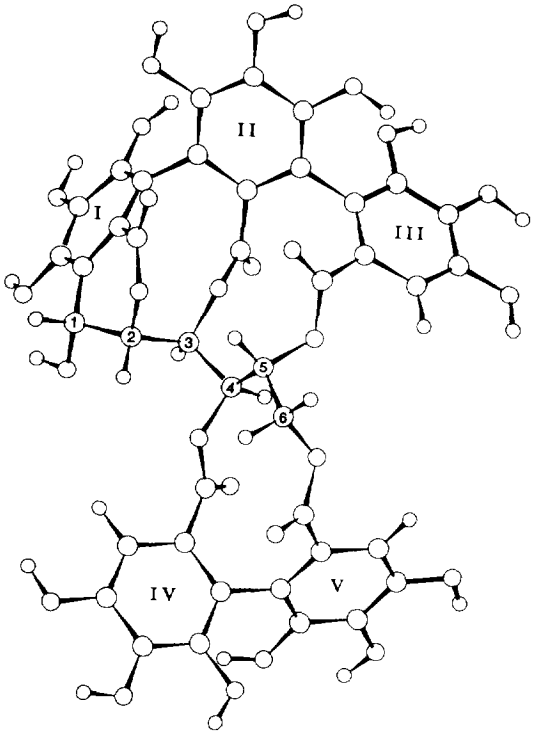


Fig. 2. Spatial structure of castalagin (I, II, III, IV, V phenyl groups; 1, 2, 3, 4, 5, 6, C-glycosidic groups).

tional space during a dynamics simulation run. Using NMR spectra, the two diastereoisomers are well characterized by their respective H-1 δ and H-3 δ ($\Delta_{\delta(2-1)}$: H-1 = 0.87 ppm; H-3 = 0.54 ppm); these differences are properly explained by the proposed 3D-structures. Even if it appears to be correct, the proposed 'biomimetic pathway' must be considered only as an assumption and must be confirmed by further studies. In particular, the presence of some of our postulated intermediates would have to be thoroughly checked during the first steps of an isolation procedure.

EXPERIMENTAL

General. Vescalagin (**1**) and castalagin (**2**), which were purified from *Q. robur* L., and β-pentagalloylglucose (**3**), were gifts from Augustin Scalbert (INRA, INA; Paris-Grignon, France). ^{1}H NMR spectra (400 MHz): Me_2CO-d_6–D_2O (7:3), pyridine-d_5 and DMSO-d_6; the digital resolution of the ^{1}H spectra was 0.5 Hz and the acquisition time, 2.01 sec. IR spectra: KBr discs. LSI-MS: Cs beam in the negative mode, acceleration voltage 35 kV (2 mA), temp. 40° and calibrating with Cs iodine (M_r 200–1500 Da) in a thioglycerol matrix.

Plant material. Samples of *Q. robur* L. were collected 1 year after 140-year-old trees were felled in the Limousin

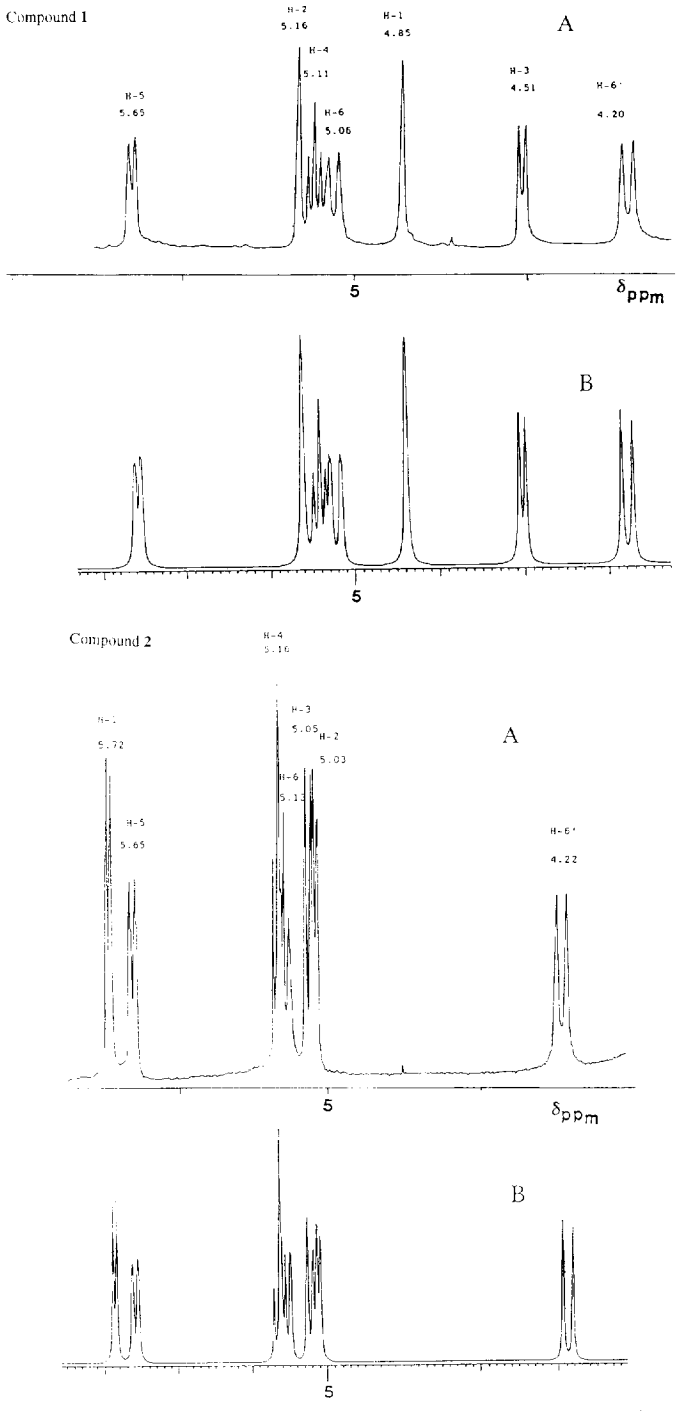


Fig. 3. ^{1}H NMR spectra measured (A) and simulated with calculated coupling constants (B) for compounds **1** and **2**.

area of France. Analytical samples were obtained by planing of heartwood.

Extraction. Shavings (100 g) were extd with Me_2CO–H_2O (7:3) at room temp. on a rotative table (135 rpm) during 48 hr. The extract was filtered, centrifuged (4500 rpm) and Me_2CO removed under red. pres. The aq. soln was freeze-dried and weighed (99.7 mg g^{-1} of shavings).

Isolation of ellagitannins. Freeze-dried extract (1 g dissolved in a min. vol. of H_2O) was loaded onto a Sephadex LH20–100 column (80 cm × 30 mm i.d.; 25–100 mm particule size). Elution was conducted with 300 ml H_2O (fr. I), 300 ml H_2O–MeOH (4:1) (fr. II) and 300 ml of MeOH (fr. III). Each fr. was dried on a rotary evaporator at 30°. Regeneration of the column was achieved by successive elutions with 500 ml HCl (10 mM), 500 ml of NaOH

Table 1. Coupling constants of compounds **1** and **2** measured (^{1}H NMR, 400 MHz) and calculated (MM3* force field) with and without phenolic hydroxyls

	Coupling constants					
	$J_{1,2}$	$J_{2,3}$	$J_{3,4}$	$J_{4,5}$	$J_{5,6}$	$J_{5,6'}$
			Vescalagin			
Measured	2.0	1.0	7.0	7.5	2.5	1.0
Without phenolic OH	2.1	1.4	6.6	8.2	2.3	1.7
With phenolic OH	2.1	1.3	6.5	8.1	2.4	1.5
			Castalagin			
Measured	4.8	1.0	7.0	7.5	2.5	1.0
Without phenolic OH	4.0	1.6	6.4	8.1	2.3	1.7
With phenolic OH	4.2	1.4	6.3	7.9	2.4	1.5

Table 2. Coupling constants of compounds **1** and **2** measured (^{1}H NMR, 400 MHz) and calculated using different solvatation models

	Coupling constants					
	$J_{1,2}$	$J_{2,3}$	$J_{3,4}$	$J_{4,5}$	$J_{5,6}$	$J_{5,6'}$
			Vescalagin			
Measured	2.0	1.0	7.0	7.5	2.5	1.0
8 Å layer of water	1.5	1.1	6.2	8.4	3.1	1.1
With GB/SA water	2.2	1.4	6.3	8.1	2.5	1.4
With GB/SA $CHCl_3$	2.1	1.4	6.6	8.2	2.4	1.5
			Castalagin			
Measured	4.8	1.0	7.0	7.5	2.5	1.0
8 Å layer of water	6.0	0.5	4.8	7.8	3.5	0.9
With GB/SA water	4.3	1.5	6.2	8.0	2.4	1.5
With GB/SA $CHCl_3$	4.2	1.5	6.4	7.9	2.4	1.5

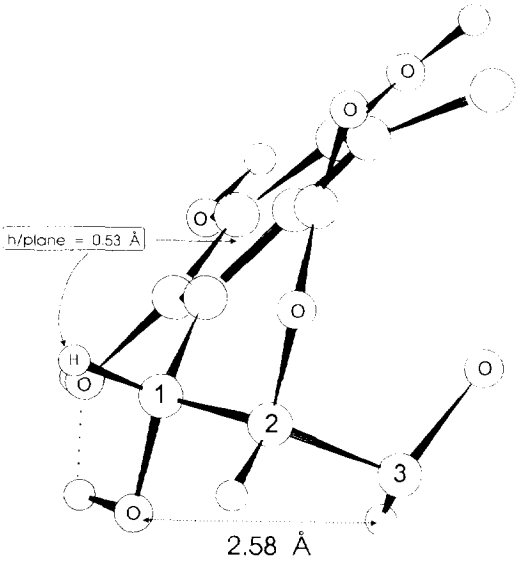


Fig. 4. Stereochemistry of C-1-glycosidic and phenyl-I-group of compound **2**.

(10 mM) and stabilized by 1 l of H_2O. Fr. II contained the main ellagitannin peaks observed by analytical HPLC [25].

Purification of castalagin and vescalagin. These compounds were purified from fr. II by prep. TLC on cellulose (WF254, Merck, 10 × 10 cm × 2 mm) and elution with H_2O–HOAc (47:3). Compounds were recovered from the adsorbent by elution with H_2O and the aq. filtrate freeze-dried. Spots corresponding to R_f 0.53 and 0.60 were collected. The R_f 0.53 product had the same R_t as castalagin by analytical HPLC (36.5 min.); the R_f 0.60 product had the same R_t as vescalagin (30 min).

Spectra. The 250–400 nm part of the UV spectra of **1** and **2** showed no maxima but a shoulder around 275 nm. Compound **1**: λ_{max}^{MeOH} 225 nm: (logε 3.8). Compound **2**: λ_{max}^{MeOH} 225 nm: (logε 3.62). *IR spectra*: vescalagin: 1455, 1510, 1620, 1740 cm^{-1}. Castalagin: 1445, 1500, 1615, 1740 cm^{-1}. ^{1}H NMR (400 MHz). Vescalagin (pyridine-d_5): δ Glc 5.53 (*d*, H-1), 6.05 (*br*, H-2), 5.41 (*d*, H-3), 6.11 (*t*, H-4), 6.40 (*br*, H-5), 5.33 (*dd*, H-6), 3.98 (*d*, H-6'). Vescalagin (DMSO-d_6), δ Glc 4.85 (*br s*, H-1), 5.15 (*br s*, H-2), 4.50 (*d*, H-3), 5.11 (*t*, H-4), 5.64 (H-5), 5.05 (*d*, H-6), 4.20 (*d*, H-6'). Castalagin (DMSO-d_6): δ Glc 5.69 (*d*, H-1), 5.0 (*d*, H-2), 5.02 (*d*, H-3), 5.12 (H-4), 5.61 (H-5), 5.09 (*d*, H-6), 4.18 (*d*, H-6'). The DMSO-d_6 spectra were simulated in order to obtain the exact values of the coupling constants, the spectrum of vescalagin in pyridine-d_5 being the best resolved one. Coupling constants were first measured in this solvent, then transposed and defined in DMSO. The coupling constants appeared to be essentially identical in these two solvents. The lowest ^3J constants (*ca* 1 Hz) are overestimated due to a general widening of peaks in DMSO. The recorded and

simulated spectra are shown in Fig. 3. Coupling constants used in final simulations are the 'measured' ones in Table 1.

Molecular mechanics. Calculations were performed on a SGI Indigo platform running Macromodel version 3.5 (Columbia University, New-York) or Insight II and Discover version 2.3.0 (Biosym technologies). Conformational minima were found using the modified MM2* (1987 parameters) or MM3* (1991 parameters) force-fields as implemented and completed in the MacroModel program. Build structures were minimized to a final RMS gradient $\leqslant 0.005$ kJ Å^{-1} mol^{-1} *via* the truncated Newton conjugate gradient (TNCG) method (1000 cycles). Coupling constant calculations were performed using ref. [26], as implemented in Macromodel.

Stochastic dynamics simulation. This variant of molecular dynamics (forces from the force field are augmented by frictional and random forces which simulate some of the properties of a solvent medium) is implemented in MacroModel [27]. The force field chosen was MM2* (1987 parameters). The kinetic energy was increased from 300 to 900 K with a bath constant of 0.2 psec. A total production time of 50 psec (time step 1 fsec) was chosen and one conformer was sampled each psec and minimized as following. Each conformer was minimized with the Truncated Newton Conjugate Gradient method (TNCG, 1000 cycles, RMS $\leqslant 0.005$ kJ Å^{-1} mol^{-1}) and all the unique (heavy atoms only) conformers within 50 kJ mol^{-1} energy range were reported and classed in order of ascending energy. The run used the first local minimum conformer as starting geometry. Each run was repeated starting from the lowest energy conformation found in the preceding run until no more lower energy conformations were found. A last run was then performed starting from the most unusual conformer, completely different from the lowest energy conformer. If the energy was still decreasing, then the entire procedure was repeated; if not, the procedure was stopped. In the present work, four runs were necessary for vescalagin and seven runs for castalagin.

Monte Carlo-style conformational search. This search is implemented in MacroModel [28, 29]. The automatic set-up has been selected, *i.e.* single bonds variable, chiral centres set and flexible ring opened and 1000 steps were made per input structure, in an energy range of 50 kJ mol^{-1}. Each conformer was fully minimized (1000 cycle. TNCG methods, RMS $\leqslant 0.005$ kJ Å^{-1} mol^{-1}, MM2* or MM3* force-field). The least-used structures were used as starting geometries only if their energies were within the energetic window (50 kJ mol^{-1} of the lowest energy structure yet found). All ^{1}H–^{1}H coupling constants were calculated for the resulting conformers for comparison with experimental data and the unique best-fit conformer for each enantiomer was extracted and stored for further calculations. In particular, re-minimizations were conducted using MM2* or MM3*, AMBER* and OPLS* force-fields with and without the GB/SA continuum solvation model of ref. [30] (solvents were H_2O and $CHCl_3$).

Minimization with solvent exprimed. This calculation was run using the Biosym Technologies softwares Insight and Discover, and the CVFF91 force-field. Conformers selected after the Monte-Carlo search were soaked in a water layer of 8 Å thickness (namely 371 molecules) and were submitted to a full minimization using 500 cycles of the Steepest Descent method followed by 2500 cycles of the Conjugate Gradient method until a R.M.S. of 0.001 kcal Å^{-1} mol^{-1} was reached. During the Steepest Descent minimization, heavy atoms of the solute molecule were tethered in order to correctly set the hydrogen bonds without distorting the starting geometry.

Acknowledgements—We sincerely thank Augustin Scalbert (INRA-INA Paris-Grignon, France) for providing different pure ellagitannins and for his continual support of our research. The authors are grateful to Mrs C. du Penhoat (Ecole Normale Supérieure, Paris, France) for revision of this manuscript and information about the structure of ellagitannins.

REFERENCES

1. Scalbert, A., Monties, B. and Janin, G. (1989) *J. Agric. Food Chem.* **37**, 1324.
2. Scalbert, A. (1992) in *Plant Polyphenols* (Hemingway, R. W. and Laks, P. E., eds), p. 259. Plenum Press, New York.
3. Klumpers, J., Scalbert, A. and Janin, G. (1994) *Phytochemistry* **36**, 1249.
4. Hart, J. M. and Hillies, W. E. (1972) *Phytopathology* **62**, 620.
5. Scalbert, A. (1991) *Phytochemistry* **20**, 2875.
6. Herve Du Penhoat, C. L. M., Michon, V. M. H., Peng, S., Viriot, L., Scalbert, A. and Gage, D. (1991) *J. Chem. Soc. Perkin Trans. I* 1653.
7. Davis, M. M., Montgomery, K. C. and Scroggie, J. L. (1985) *J. Soc. Leather Techn. Chem.* **69**, 130.
8. Tang, M. R., Hancock, R. A. and Covington, A. D. (1992) in *Plant Polyphenols* (R. W. Hemingway and P. E. Laks, eds), p. 259. Plenum Press, New York.
9. Vivas, N., Chauvet, S., Glories, Y. and Sudraud, P. (1993) *Ind. Agric. Alim.* **110**, 705.
10. Janin, G., Mazet, J. F., Flot, J. L. and Hofman, P. (1990) *Rev. For. Fr.* **XLII**, 134.
11. Quinn, M. K. and Singleton, V. L. (1985) *Am. J. Enol. Vitic.* **36**, 148.
12. Vivas, N. and Glories, Y. (1993) *Rev. Fr. OEnol.* **142**, 33.
13. Viriot, C., Scalbert, A., Lapierre, C. and Moutounnet, M. (1993) *J. Agric. Food Chem.* **41**, 1872.
14. Bate-Smith, E. C. (1972) *Phytochemistry* **11**, 1153.
15. Vivas, N., Chauvet, S., Glories, Y. and Sudraud, P. (1993) *Ann. Fals. Exp. Chim.* **86**, 215.
16. Mayer, W., Seitz, M., Jochims, J. C. (1969) *Liebigs Ann. Chem.* **721**, 186.
17. Mayer, W., Seitz, M., Jochims, J. C., Schauerte, K. S. and Shilling, N. G. (1971) *Liebigs Ann. Chem.* **751**, 60.

18. Nonaka, G., Sakai, T., Tanaka, T., Nihashi, K. and Nishioka, I. (1990) *Chem. Pharm. Bull.* **38**, 2151.
19. Yoshida, T., Ohbayashi, M., Ishimara, K., Ohwashi, W., Haba, K., Okano, Y., Shingu, T. and Okada, T. (1991) *Chem. Pharm. Bull.* **39**, 2233.
20. Spencer, C. M., Cai, Y., Martin, R., Lilley, T. H. and Haslam, E. (1990) *J. Chem. Soc., Perkin Trans.* **II** 651.
21. Viriot, C., Scalbert, A., Herve Du Penhoat, C. L. M. and Moutounnet, M. (1994) *Phytochemistry* **36**, 1253.
22. Okuda, T., Yoshida, T. and Hatano, T. (1989) *Planta Med.* **55**, 117.
23. Haslam, E. (1992) in *Plant Polyphenols* (Hemingway, R. W. and Laks, P. E., ed.), p. 169. Plenum Press, New York.
24. Ishimaru, K. (1987) PhD thesis. Kyushu University, Fukuoka.
25. Scalbert, A., Duval, L., Peng, S., Monties, B. and Herve Du Penhoat, C. L. M. (1990) *J. Chromatogr.* **502**, 107.
26. Altona, C., V. Haasnoot, C. A. G., and de Leew, F. A. A. M. (1980) *Tetrahedron* **36**, 2783.
27. Van Gunsteren, W. F. and Berendsen, H. J. C. (1988) *Molec. Simul.* **1**, 173.
28. Chang, G., Guida, W. C. and Still, W. C. (1989) *J. Am. Chem. Soc.* **111**, 4379.
29. Saunders, M., Houk, K. N., Wu, Y. D., Still, W. C., Lipton, M., Chang, G. and Guida, W. C. (1990) *J. Am. Chem. Soc.* **112**, 1419.
30. Still, W. C., Tempczyk, A., Hawley, R. C. and Hendrickson, T. (1990) *J. Am. Chem. Soc.* **112**, 6127.