

The structural and conformational analyses and antioxidant activities of chebulinic acid and its thrice-hydrolyzed derivative, 2,4-chebuloyl- β -D-glucopyranoside, isolated from the fruit of *Terminalia chebula*

Karel D. Klika,^a Ammar Saleem,^b Jari Sinkkonen,^a Marja Kähkönen,^c Jyrki Loponen,^b Petri Tähtinen,^a and Kalevi Pihlaja*^{a,b}

^a Structural Chemistry Group, Department of Chemistry, University of Turku, Vatselankatu 2, FIN-20014 Turku, Finland.

^b Environmental Chemistry Group, Department of Chemistry, University of Turku, Vatselankatu 2, FIN-20014 Turku, Finland.

^c Division of Food Chemistry, Department of Applied Chemistry and Microbiology, University of Helsinki, D-Talo, Viikki, FIN-00014 Helsinki, Finland

E-mail: kalevi.pihlaja@utu.fi

Dedicated to Professor Sándor Antus on the occasion of his 60th birthday

(received 02 Nov 03; accepted 29 Jan 04; published on the web 01 Feb 04)

Abstract

1,3,6-Tri-*O*-galloyl-2,4-chebuloyl- β -D-glucopyranoside (chebulinic acid, **1**) and its novel thrice-hydrolyzed derivative, 2,4-chebuloyl- β -D-glucopyranoside (galloyl-free chebulinic acid, **2**), together with ellagic (**3**) and gallic acids (**4**), ethyl gallate (**5**), and luteolin (**6**), were isolated from the dried fruit of *Terminalia chebula* by bioactivity-guided fractionation of the extract. The compounds were variously identified on the basis of UV, MS, and NMR data. The latter proved that the glucose unit in **1** adopts a well defined ¹C₄ chair conformation with all substituents in axial positions. However, the absolute configurations of the three stereocenters in the chebuloyl group have been reassigned. For **2**, a solvent-dependent mixture of both open-chain and cyclic pyranose forms was observed, the latter with a β configuration for the anomeric C-1 sugar carbon being heavily preferred in aqueous solution. In contrast to **1**, a dynamic equilibrium consisting of several possible skew conformers (⁵S₁, ¹S₅, ¹S₃, and ²S₀) was designated for the sugar ring of **2**. Antioxidant activities of the isolated compounds were assessed by measuring their ability to scavenge 1,1-diphenyl-2-picrylhydrazyl (DPPH) radicals and to inhibit the autoxidation of methyl linoleate *in vitro*. Of the isolated compounds, **1** showed the highest radical scavenging activity in the DPPH assay. For the methyl linoleate assay, **2** and **4–6** all exhibited strong antioxidant activities whereas the activities of **1** and **3** were only moderate. The fruit extract itself was highly effective in both tests.

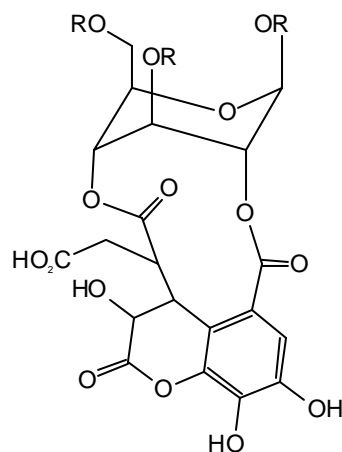
Keywords: *Terminalia chebula*, conformational analysis, NMR spectroscopy, antioxidant activities

Introduction

The hydrolyzable tannins¹ of dicotyledonous plants are useful compounds that find wide application as tanning agents, but they may also be important components of traditional medical concoctions.^{2,3} Structurally they incorporate the basic components of a glucose moiety with various numbers of galloyl ester substituents, but despite this basic template the group comprises of a vast array of interesting structures incorporating all possible permutations; for the ellagitannins alone there are some 500-odd representatives that have been structurally characterized.^{3,4} For these reasons, they are an enticing class of compounds attracting as much attention recently for their bioactive properties^{2,3} as they did originally for their tanning properties.³ However, the excitement that this class of compounds is potentially able to generate is tempered by the confusion and structural inconsistencies that cloud some of its members. In particular, the gallotannin 1,3,6-tri-*O*-galloyl-2,4-chebuloyl- β -D-glucopyranoside (**1**, chebulinic acid) has had a colorful structural history. Chebulinic acid (**1**) can be found in several plant species, such as *Phyllanthus emblica*,² *Terminalia arborea*,⁵ and *T. chebula*,²



The fruit of *Terminalia chebula*.



Chebulinic acid (**1**)
(R = galloyl)

the latter from whence it derives its name and a plant which is widely distributed throughout the Indian subcontinent and South-East Asian regions. Moreover, the dried fruit of *T. chebula* is regarded as the "king of medicines" by Tibetans and second-to-none by Ayurvedic apothecaries and is also held in high esteem by other folk medicinal practitioners.⁶ Its uses span the full gamut of medical applications and include use as an astringent, purgative, supplement, *etc.*, and to retard the effects of aging and impart longevity as well as for boosting the immune system. It is also used for the treatment of numerous maladies including diarrhea, fever, cough, asthma, rheumatism, chronic ulcers, leucorrhoea, pyorrhea, dermal fungal infections, *etc.* to name but a few of the many ailments that this purported panacea can cure.⁶

Although hydrolyzable tannins have been intensively studied^{2,3,7} by NMR, it was only recently that the full NMR assignments of **1**, despite being a particularly well known, albeit of limited access, tannin have been reported⁸ Very often, the NMR data for many chebuloyl-containing compounds have not been fully reported with only the protons of the glucose moiety being described⁹ and the aliphatic protons of the chebuloyl group being omitted rendering identification a more burdensome task as well as precluding the observation of structural similarities. This is also true of **1** where the conformation of the sugar ring was based on ¹H NMR studies.⁹⁻¹¹ The deficiencies in the NMR data for these particular compounds are slowly being addressed, *e.g.*, Ding, Liu *et al.*^{8,12} have reported the NMR assignments of the twice-hydrolyzed derivative of **1**, 1-*O*-galloyl-2,4-chebuloyl- β -D-glucopyranoside (known either as chebulanin or terminalic acid), and amariinic acid¹³ has also been described.¹⁴

The colorful structural history of **1**, though, seems set to continue unabated, for example, recent depictions have represented the relative stereochemistry for H-3 and H-4 in the chebuloyl moiety as either *cis*² or *trans*.³ It was not possible to ascertain the stage at which the allocations of the various chebuloyl group stereochemistries were made, and it seems that often the stereochemistry is simply assumed, or based on other analogous structures which in themselves may be wrong or irrelevant. Implicit assumptions based on biosynthetic arguments,¹⁴ analogous structures,¹⁴ or chebulic acid (the hydrolyzed chebuloyl moiety) also encounter the problem of the potential configurational lability of the stereocenters at 3 and 2'. The lack of comprehensive NMR data certainly only exacerbates this problem. Indeed, in addition to the structural irregularities persisting for **1**, even now the conformational behavior of the glucose moiety, seemingly well established, can be misrepresented.² Conformational evaluation is of prime importance given that it can play a determinate role in the biological activity of a compound,¹⁵ in particular for sugars and their associated enzyme affinities.¹⁶

To address this issue and to further augment the characterization of this class of compounds, we herein report the full NMR assignment and structural characterization of chebulinic acid (**1**) and its novel thrice-hydrolyzed derivative, 2,4-chebuloyl- β -D-glucopyranoside (galloyl-free chebulinic acid, **2**, see Chart 1 overleaf). Both **1** and **2** were isolated for this work as part of a comprehensive phytochemical investigation^{6,17-19} into the traditional medicinal plants of Pakistan which included *T. chebula*. An aqueous methanol extract of the dried fruit of *T. chebula* Retz. (Combretaceae) subjected to antioxidant¹⁷ and cancer-cell growth inhibition¹⁸ activity-guided fractionation furnished both **1** and **2**, together with the familiar compounds ellagic acid (**3**), 3,4,5-trihydroxy benzoic acid (gallic acid, **4**), ethyl 3,4,5-trihydroxy benzoate (ethyl gallate, **5**), and 5,7,3',4'-tetrahydroxyflavone (luteolin, **6**). Compounds **1-6** were identified variously on the basis of UV, MS, and NMR data in comparison, where applicable, to literature data and/or authentic samples. The antioxidant potentials of **1-6** were assessed by measuring their ability to scavenge 1,1-diphenyl-2-picrylhydrazyl (DPPH) radicals and to inhibit the autoxidation of methyl linoleate *in vitro*. Given the structure of **1**, it seemed pertinent, in addition to the standard reference substances Trolox and α -tocopherol, to also make a comparison to tannic acid—a commercial mixture consisting of various gallotannins.

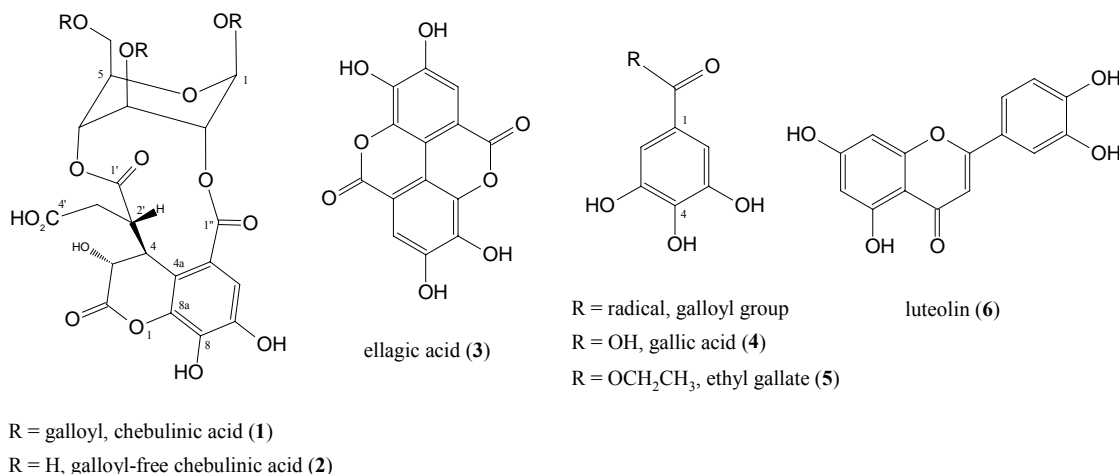


Chart 1. The structures of the compounds **1-6** isolated from *Terminalia chibula*. The stereochemistries of chebulinic acid (**1**) and galloyl-free acid (**2**) are as indicated, including that of the anomeric carbon C-1. The conformation portrayed for the glucose ring, an inverted chair conformation, is the conformation adopted by **1**, the conformation of the glucose ring **2** is discussed separately. The numbering system in use is also indicated for **1** and **2** and their attendant chebuloyl and galloyl groups.

Results and Discussion

Isolation, characterization, and antioxidant activities of *T. chebula* compounds

The extraction and isolation procedures for the compounds **1-6** have been described in detail previously,¹⁸ but it was the water partition of a dried aqueous methanol extract of *T. chebula* fruit originally showing promising antioxidant¹⁷ and cancer-cell growth inhibition¹⁸ properties that furnished, after SEC fractionation and consequent HPLC, chebuloyl-free chebulinic acid (**2**). After similar processing, but from the 1-butanol partition, chebulinic acid (**1**) was obtained. The six compounds in total (compounds **1-6**) so isolated by these procedures were subsequently subjected to further bioassays, *viz.* antioxidation assays (*vide infra*). Although the possible presence of chebulagic acid was also noted in the extract by HPLC-ESI-MS, its low concentration in these particular samplings precluded the isolation of sufficient quantities for further investigative work. Three of the compounds isolated were readily identified as gallic acid (**4**), ethyl gallate (**5**), and luteolin (**6**) on the basis of retention times, UV and MS data, and by comparison to authentic samples. Ellagic acid (**3**), the lactonized dimer of gallic acid, was also readily identified by comparison of its ¹H and ¹³C NMR to literature values.²⁰

Compounds **1** and **2**, however, were only identified after extensive NMR analysis in conjunction with MS (which furnished ions of *m/z* 955 and 499 amu for the [M - H]⁻ ion under ESI⁻ conditions for **1** and **2**, respectively, and for which accurate mass analysis provided the compositions C₄₁H₃₁O₂₇ and C₂₀H₁₉O₁₅, respectively). The gross structural identification of both **1** and **2** relied on the standard application of COSY, HMQC, and HMBC experiments to obtain the structural framework and to effect assignment of the spins. The identification of **1** was not

immediately apparent despite being a known compound because of the limited, or limited access to,⁸ the literature NMR data. Intriguingly, despite the greater simplicity of **2** and the apparent structural similarities of **1** and **2**, the identification of **2** was much more problematic and the approaches to the solutions of the two structures in effect, diverged. For this reason they are discussed separately in the ensuing sections. Common for both structures was the utilization of the anomeric proton, H-1, or the anomeric carbon itself, C-1, as convenient starting points for mapping the ¹H–¹³C spin network of the sugar ring (and the realization of the presence of a hexose itself). Analogously for the chebuloyl group, H-3' provided a convenient starting point for mapping the ¹H–¹³C spin network and the determination of the structural framework of this unit, which proceeded in a similar manner for both compounds given that little variation in the structure of this moiety was observed between the compounds. Needless to say, it proceeded at an accelerated rate for the latter examination.

In both the DPPH radical scavenging and the inhibition of the autoxidation of methyl linoleate assays, the fruit extract of *T. chebula* was found to exhibit strong antioxidant potential (see Table 1). All of the six compounds (**1**–**6**) isolated from the extract were also found to be highly active in one or the other or both of these tests.

Table 1. *In vitro* antioxidant activities of *T. chebula* fruit extract and its constituent phenolics

Sample	Dry weight content, ppt	DPPH radical scavenging, % of radicals scavenged ^a	Methyl linoleate autoxidation, % of inhibition ^a
<i>T. chebula</i> fruit extract	–	90 ± 1	97 ± 0
chebulinic acid (1)	1.46	93 ± 2	34 ± 4
galloyl-free chebulinic acid (2)	0.88	38 ± 1	87 ± 3
ellagic acid (3)	0.88	41 ± 1	50 ± 1
gallic acid (4)	38 × 10 ⁻³	46 ± 3	97 ± 1
ethyl gallate (5)	88 × 10 ⁻³	45 ± 1	98 ± 1
luteolin (6)	55 × 10 ⁻³	13 ± 0	98 ± 0
tannic acid ^b	–	89 ± 0	80 ± 1
Trolox ^b	–	36 ± 1	98 ± 0
α-tocopherol ^b	–	not assayed	98 ± 0

Results are mean values ±SD, *n* = 3. ^a Individual compounds were tested in the DPPH and methyl linoleate assays at concentrations of 16.7 μM and 0.1 mM, respectively, whilst *T. chebula* fruit extract and tannic acid were tested at a concentration of 1 mg mL⁻¹ in the DPPH assay and 50 ppm in the methyl linoleate assay. ^b Reference mixture/compounds.

For example, in the DPPH assay **1** was by far the most efficient whilst the other compounds, including **2**, displayed more moderate scavenging powers by comparison, although they performed equitably or slightly better than the reference compound Trolox. The one exception was **6** which only performed weakly as a radical scavenger. Under the conditions of this test, tannic acid was found to be comparable to both the fruit extract itself and chebulinic acid (**1**). These results are in line with the known²¹ ability of galloyl glycosides to scavenge DPPH radicals and an increasing number of hydroxyl groups capable of donating hydrogen to the radical should enhance the capability of a substrate to scavenge the radical. Tannic acid and **1** are well endowed with hydroxyl groups (a total of 13 for the latter) and as such provided high scavenging rates. Compounds **2–5**, possessing from 3–7 hydroxyl groups each, all performed comparably to each other and significantly below tannic acid and **1**. The exception to this trend, of course, was **6**, which, with 4 hydroxyl groups, performed well below compounds **2–5**.

In the lipid-containing model (the inhibition of the autoxidation of methyl linoleate), the majority of the compounds exhibited significant antioxidant activities. However, despite its high DPPH radical scavenging capacity, **1** possessed only moderate antioxidant activity in bulk methyl linoleate, whereas **2**, virtually conversely, was a good oxidation inhibitor. The contrasting results of these two tests might be due to the different mechanisms of oxidation involved (artificial DPPH radical or lipid radical) and the different media in which the tests were conducted—dissolution in methanol or dissolution/dispersion in bulk oil resulting in starkly contrasting, and even non-uniform, reaction environments. Whilst in the DPPH model the antioxidant efficiency of phenolic compounds is mainly dependent on the number of hydroxyl groups capable of donating hydrogen to the radical, in the methyl linoleate assay, other factors also play a role, such as the distribution of the antioxidants between the oil and the oil–air interfaces, the metal chelation properties of the compounds, *etc.*, in addition to their radical scavenging capacities.²² As the fruit extract proved to be both a powerful DPPH radical scavenger and a strong antioxidant in the methyl linoleate assay, it is likely that this efficiency probably arises from the presence of several active components as nearly all of the isolated compounds showed moderate or strong action in the two tests. Given the low concentration (in the ppm range) of **4–6**, however, the antioxidant activities of the fruit extract can probably be attributed to the more mass prevalent constituents **1–3** which are present in the ppt range.

Conformation and configuration of chebulinic acid (**1**)

The basic structural analysis of **1** provided evidence for the presence of a sugar moiety—explicitly a hexose—as a structural entity which formed a core template from which various groups were tethered. The presence of three distinct galloyl groups was evidenced by the three two-proton singlets in the aromatic region of the ¹H NMR spectrum, complemented by the sets of corresponding aromatic signals in the ¹³C NMR spectrum. HMBC was important for not only supporting the existence of these distinct galloyl groups and for rendering the assignment of the quaternary carbons within each of these units, it also established the site of connection of the galloyl groups to the sugar ring *via* the correlations to the carbonyl carbons of the galloyl groups from both the respective H-2s of the galloyl units and the sugar ring protons at the site of attachment (H-1, H-3, and H-6). Similarly for the chebuloyl group, the skeletal framework was

constructed from the basic structural analysis and the sites of attachment to the sugar ring, sites 2 and 4, realized in a like-wise manner to the galloyl groups. The HMBC correlations observed for **1** are presented in Figure 1 and the ^1H and ^{13}C NMR data of **1** in Tables 2 and 3, respectively. Due to the limited access to the literature data, recourse was made to comparison with the spectral data for the five-membered analog of chebulic acid^{23,24} (where C-3 of the lactone ring is lacking) for confirmation of these assignments and the final realization that the compound was indeed the known compound chebulinic acid (**1**).

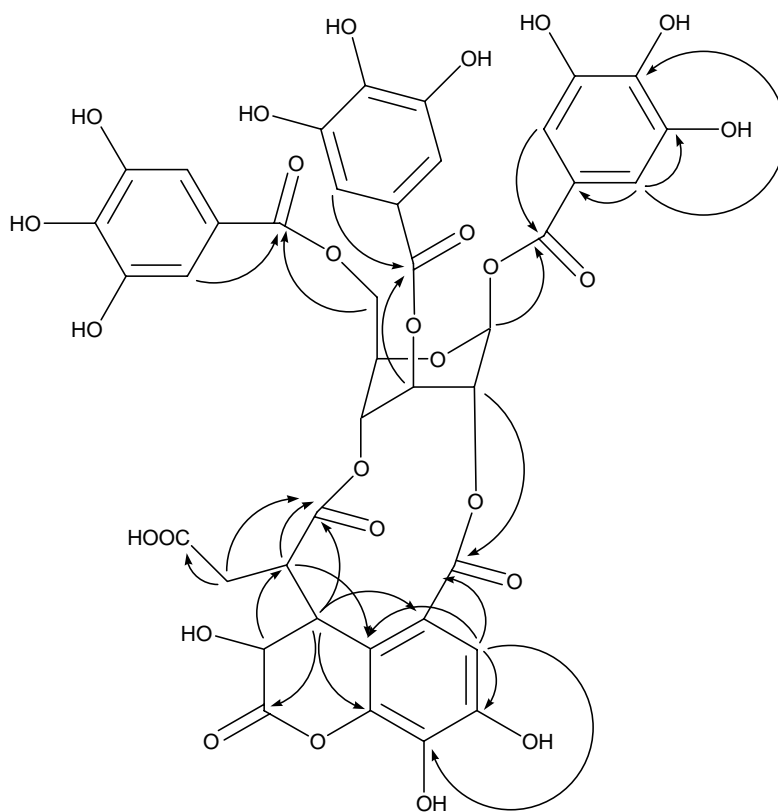


Figure 1. Stereostructure and selected HMBC correlations for chebulinic acid (**1**). Similar HMBC correlations were also observed for the other galloyl groups as indicated for the galloyl-1 group and, in addition, virtually all possible two- and three-bond correlations for the glucose moiety were also observed; these additional correlations have been omitted for the sake of clarity.

Although the conformation of the sugar ring in **1** has been discussed in the literature on a number of occasions, because of the unconventional nature of the adopted conformation of the sugar ring (evident by the magnitude of the vicinal ^1H - ^1H coupling constants for the ring protons), the identity of the hexose was unequivocally confirmed as glucose after hydrolysis of all the attendant groups. Analysis by ^1H and ^{13}C NMR of the hydrolyzed sample provided sets of signals consistent with an equilibrium solution of α - and β -glucose by comparison to an authentic solution of glucose, which was further confirmed by mixed analysis resulting in the

enhancement of the glucose signals in the ^1H NMR spectrum. The earliest report for the correct gross structure of **1**, albeit with hydrolysis of the lactone ring, appears to be by Haslam¹⁰ where an inverted chair conformation, $^1\text{C}_4$, was assigned as the conformation of the glucose moiety, later reinforced by Haddock.⁹ Other conformations such as $^3\text{O}B$, or a conformational equilibrium of $^3\text{O}B$ and $B_{1,4}$ have also been proposed to account for the observed coupling constants in other solvents (DMSO- d_6).¹¹ The prime reason, of course, for contemplating a departure from a conventional $^4\text{C}_1$ chair conformation is to account for the vicinal coupling constants of the sugar ring protons which are all very small and thus imply equatorial orientations for these protons, *i.e.* a $^1\text{C}_4$ conformation. Furthermore, considerably large four-bond w -couplings were also observed in this study for the sugar ring protons lending further support to their equatorial disposition. The NOE results obtained in this work with respect to the sugar ring conformation were relatively uninformative, but implied a similar geometry as NOE contacts between sugar ring protons were only observed between vicinal protons. Although inverted chair conformations are well established for this class of compounds,^{2,3,9-11} and indeed so for this particular compound (**1**),^{3,9,10} they are not always necessarily apparent. Therefore it was considered warranted to re-examine the conformational preference of chebulinic acid (**1**) by considering whether any of the skew conformers could accommodate all of the vicinal proton couplings of the sugar ring whilst alleviating not only the associated strain and steric hindrance that an inverted chair might inflict, but also addressing the geometric requirements of the macrocycle.

Of the possible six skew forms (see Chart 2), the $^3\text{S}_1$ conformer appears as the best candidate for $J_{\text{H}2,\text{H}3}$, $J_{\text{H}3,\text{H}4}$, and $J_{\text{H}4,\text{H}5}$ are all likely to be small (see Table A1, Appendix for a guide on the evaluation of dihedral angle *vs* the magnitude of J).

However, irrespective of the anomeric configuration at C-1, $J_{\text{H}1,\text{H}2}$ should be sizeable, thus discounting this conformer as the predominant structure or a significant contribution of it to a dynamic equilibrium. Furthermore, for the three long-range couplings observed, $^4J_{\text{H}1,\text{H}3}$, $^4J_{\text{H}2,\text{H}4}$, and $^4J_{\text{H}3,\text{H}5}$, for $^4J_{\text{H}1,\text{H}3}$ to be maintained, then the anomeric configuration must be α which is not the case given an observed NOE between H-1 of the glucose moiety and H-4 of the chebuloyl moiety.

Table 2. ^1H NMR data (δ in ppm, J in Hz) for chebulinic acid (**1**) and galloyl-free chebulinic acid (**2**) at 25 °C

site	$1^{a,b}$ in acetone- d_6	$2^{c,e}$ pyranose form in D_2O	$2^{c,e}$ pyranose form in acetone- $d_6 + \text{D}_2\text{O}$	$2^{c,e}$ open-chain form in acetone- $d_6 + \text{D}_2\text{O}$
glucose unit				
H-1	6.51, dd, $J_{\text{H}2}$ 2.8, $J_{\text{H}5}$ 0.8	5.22, apparent s, $J_{\text{H}2}$ 7.2	5.14, d, $J_{\text{H}2}$ 6.8	9.55, s
H-2	5.47, ~dt, $J_{\text{H}1}$ 2.8, $J_{\text{H}2}$ 1.9, $J_{\text{H}4}$ 1.4	5.24, apparent s, $J_{\text{H}1}$ 7.2, $J_{\text{H}2}$ 2.1	5.02, dd, $J_{\text{H}1}$ 6.8, $J_{\text{H}2}$ 2.0	5.37, d, $J_{\text{H}2}$ 1.8
H-3	6.33, ddt, $J_{\text{H}4}$ 3.6, $J_{\text{H}2}$ 1.9, $J_{\text{H}5}$ 1.0, $J_{\text{H}1}$ 0.8	4.66, br. s, $J_{\text{H}2}$ 2.1, $J_{\text{H}4}$ 0.4	4.67, br. d, $J_{\text{H}2}$ 2.0, $J_{\text{H}4}$ < 1	4.95, br. d, $J_{\text{H}2}$ 1.8, $J_{\text{H}4}$ < 1
H-4	5.08, dt, $J_{\text{H}3}$ 3.6, $J_{\text{H}2}$ 1.4, $J_{\text{H}5}$ 1.0	4.98, d, $J_{\text{H}5}$ 9.4, $J_{\text{H}3}$ 0.4	4.82, br. D, $J_{\text{H}5}$ 9.5, $J_{\text{H}3}$ < 1	5.01, br. D, $J_{\text{H}5}$ 9.5, $J_{\text{H}3}$ < 1
H-5	4.71, m, $J_{\text{H}6}$ 8.0, $J_{\text{H}4}$ 6.4, $J_{\text{H}1}$ 1.0, $J_{\text{H}2}$ 1.0	3.95, ddd, $J_{\text{H}4}$ 9.4, $J_{\text{H}6}$ 6.1, $J_{\text{H}2}$ 2.8	3.85, ddd, $J_{\text{H}4}$ 9.5, $J_{\text{H}6}$ 6.7, $J_{\text{H}2}$ 2.7	3.90, ddd, $J_{\text{H}4}$ 9.5, $J_{\text{H}6}$ 6.4, $J_{\text{H}2}$ 2.8
H-6a	4.73, m, $J_{\text{H}6}$ -11.5, $J_{\text{H}5}$ 6.4	3.55, d(AB)d, $J_{\text{H}6}$ -12.1, $J_{\text{H}5}$ 6.1	3.36, dd, $J_{\text{H}6}$ -11.9, $J_{\text{H}5}$ 6.7	3.41, dd, $J_{\text{H}6}$ -11.9, $J_{\text{H}5}$ 6.4
H-6b	4.85, d(AB)d, $J_{\text{H}6}$ -11.5, $J_{\text{H}5}$ 8.0	3.75, d(AB)d, $J_{\text{H}6}$ -12.1, $J_{\text{H}5}$ 2.8	3.61, dd, $J_{\text{H}6}$ -11.9, $J_{\text{H}5}$ 2.7	3.66, dd, $J_{\text{H}6}$ -11.9, $J_{\text{H}5}$ 2.8
chebuloyl group				
3	4.94, d, $J_{\text{H}4}$ 7.2	5.00, br. D, $J_{\text{H}4}$ 7.6, $J_{\text{H}2}$ -0.4	4.81, overlapped	4.80, overlapped
4	5.17, dd, $J_{\text{H}2}$ 7.2, $J_{\text{H}2}$ 1.6	4.85, dd, $J_{\text{H}2}$ 7.6, $J_{\text{H}2}$ 1.8	4.85, overlapped	4.84, overlapped
6	7.54, s	7.44, s	7.33, s	7.46, s
2'	3.94, ddd, $J_{\text{H}6}$ 12.2, $J_{\text{H}4}$ 3.0, $J_{\text{H}4}$ 1.6	3.55, ddd, $J_{\text{H}4}$ 12.5, $J_{\text{H}6}$ 3.8, $J_{\text{H}4}$ 1.8, $J_{\text{H}5}$ -0.4	3.55, ddd, $J_{\text{H}4}$ 11.9, $J_{\text{H}6}$ 3.7, $J_{\text{H}4}$ 1.5	3.60, ddd, $J_{\text{H}4}$ 11.0, $J_{\text{H}6}$ 4.6, $J_{\text{H}4}$ 1.2
3'a	2.24, m, $J_{\text{H}6}$ -17.0, $J_{\text{H}2}$ 3.0	1.98, d(AB)d, $J_{\text{H}6}$ -16.6, $J_{\text{H}2}$ 12.5	2.0-2.1, overlapped with solvent	2.0-2.1, overlapped with solvent
3'b	2.27, m, $J_{\text{H}4}$ -17.0, $J_{\text{H}2}$ 12.2	2.11, d(AB)d, $J_{\text{H}4}$ -16.6, $J_{\text{H}2}$ 3.8	2.0-2.1, overlapped with solvent	2.0-2.1, overlapped with solvent

Acetone-containing samples are referenced internally to the solvent signal (2.04 ppm) and the D_2O sample to TSP (0.00 ppm). ^a Extraction of the ^1H NMR parameters for **1** and for the pyranose form of **2** in D_2O were made using PERCH iteration software. The OH protons were not observed in any of the samples. ^b Galloyl protons: H-2 in O-1 galloyl, 7.20, s; H-2 in O-3 galloyl, 7.27, s; H-2 in O-6 galloyl, 7.05, s. ^c A strongly second-order system was normally observed for H-1 and H-2 due to their close chemical shifts, but the shifts of both spins are concentration dependent. In dilute solutions, H-1 is upfield of H-2 and both move downfield with increasing concentration but H-1 moves at a faster rate than H-2 such that in concentrated solutions it is downfield of H-2. Quite obviously, at intermediate concentrations the two resonances are overlapped.

Table 3. ^{13}C NMR data (δ in ppm, J in Hz) for chebulinic acid (**1**) and galloyl-free chebulinic acid (**2**) at 25 °C

structure	glucose unit														
	C-1	$^1J_{\text{HC}}$	C-2	$^1J_{\text{HC}}$	C-3	$^1J_{\text{HC}}$	C-4	$^1J_{\text{HC}}$	C-5	$^1J_{\text{HC}}$	C-6	$^1J_{\text{HC}}$			
1 in acetone- d_6	92.28	173.3	71.19	157.4	62.18	157.9	69.10	157.3	75.78	153.5	64.93	151.2			
2 , pyranose form in D_2O	91.51	167.1	82.06	147.3	67.82	142.0	76.54	146.2	71.85	144.8	64.85	142.8			
2 , pyranose form in acetone- d_6 + D_2O	89.14	–	79.49	–	65.32	–	74.09	–	69.52	–	62.74	–			
2 , open-chain form in acetone- d_6 + D_2O	197.69	–	83.96	–	64.50	–	73.74	–	69.26	–	62.64	–			
α -D-glucose in D_2O	94.91	169.6	74.30	<i>ca.</i> 143	75.59	142.9	72.47	143.6	74.26	<i>ca.</i> 143	63.42	144.0			
β -D-glucose in D_2O	98.73	161.8	76.96	144.0	<i>78.76</i>	<i>142.3</i>	72.42	<i>ca.</i> 143	<i>78.58</i>	<i>143.5</i>	63.58	142.9			
	chebuloyl group														
	C-2	C-3	C-4	C-4a	C-5	C-6	C-7	C-8	C-8a	C-1'	C-2'	C-3'	C-4'	C-1''	
1 in acetone- d_6	169.53	66.64	41.26	116.18	119.13	117.10	146.49	139.50	141.11	173.80	39.49	30.19	172.69	165.17	
2 , pyranose form in D_2O	172.86	68.16	42.03	115.93	122.81	119.84	147.84	140.34	142.61	177.43	41.48	31.95	177.88	170.74	
2 , pyranose form in acetone- d_6 + D_2O	169.67	65.71	39.59	113.43	120.30	116.93	145.70	137.99	140.06	<i>174.06</i>	38.70	29.44	<i>174.15</i>	167.55	
2 , open-chain form in acetone- d_6 + D_2O	169.61	65.88	40.06	115.02	117.99	117.19	145.91	139.06	140.27	<i>174.17</i>	38.34	30.83	<i>174.22</i>	166.42	
	galloyl groups														
	O-1 galloyl group				O-3 galloyl group				O-6 galloyl group						
	C=O	C-1	C-2	C-3	C-4	C=O	C-1	C-2	C-3	C-4	C=O	C-1	C-2	C-3	C-4
1 in acetone- d_6	165.06	120.21	110.42	146.02	139.71	164.95	120.01	110.46	146.11	139.77	166.55	121.01	109.94	145.90	138.99

Acetone-containing samples are referenced internally to the solvent signal (29.80 ppm) and the D_2O sample to TSP (0.00 ppm). For reference purposes, α -D-glucose and β -D-glucose are also included. Assignments italicized are interchangeable, as appropriate, within each row.

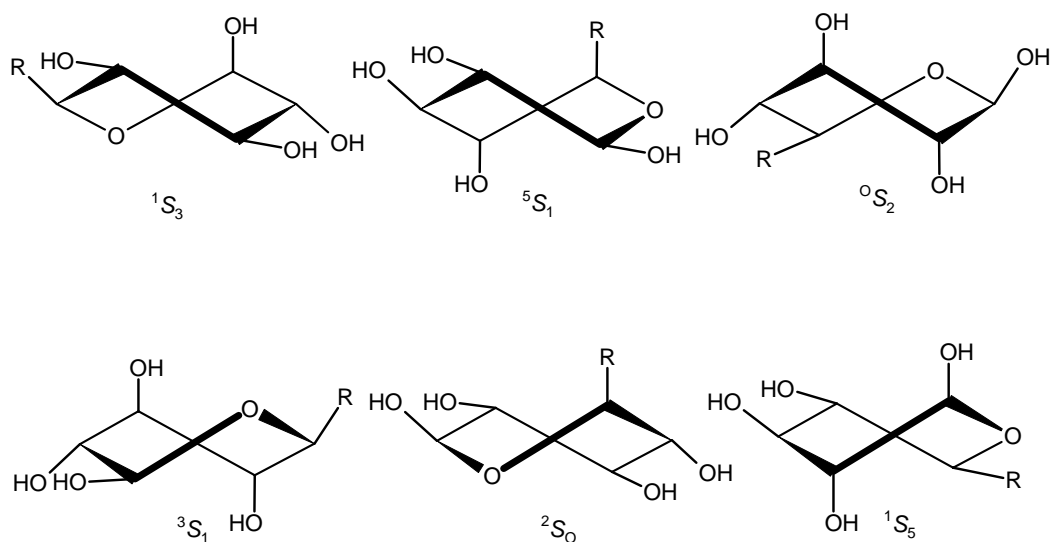


Chart 2. The six skew conformers of β -D-glucose.

For the boat conformations (see Chart 3), only one of the possible six conformations, $^{3,0}B$, is able to provide a set of dihedral angles that could conceivably give rise to a set of small vicinal couplings as observed for the sugar ring protons (see Table A2, Appendix for a guide on the evaluation of dihedral angle vs the magnitude of J).

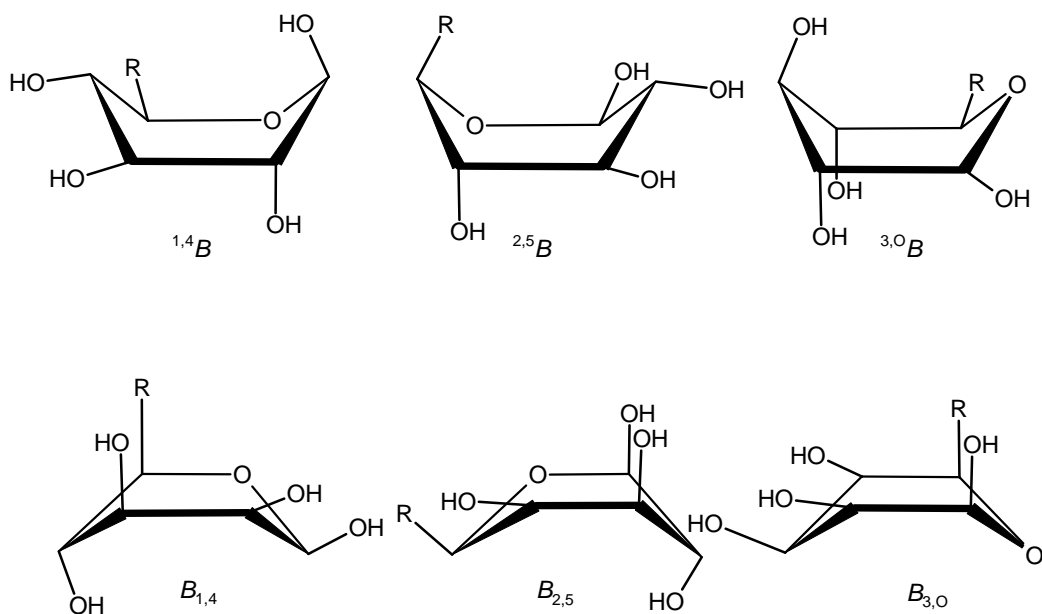


Chart 3. The six skew conformers of β -D-glucose.

Although the linkages to the chebuloyl group are both axially orientated in this conformer, based on chemical intuition it is considered that not all of the vicinal ring couplings should be so small given the dihedral angles. Furthermore, H-1 is axially orientated in this conformer and therefore the magnitude of $J_{H1,C1}$ should be small if this was a significant contributing conformer

on the basis of the experimental and theoretical study of $J_{H1,C1}$ for α - and β -D-glucose by Callam *et al.*²⁵ For glucose, a Perlin effect for the anomeric position of -7.8 Hz can be apportioned as -2 Hz when the hydrogen is trans-diaxial to an atom (*cf.* the Perlin effect for cyclohexane, -4 Hz) and -5.8 Hz when the hydrogen is trans-diaxial to an oxygen lone pair. The observed value of 173.3 Hz for $J_{H1,C1}$ (see Table 3) is therefore more suggestive of an equatorially-orientated H-1 in comparison to α - (169.6 Hz) and β -D-glucose (161.8 Hz). Also for the 3S_1 conformer, an equatorially-orientated H-1 is only possible for the α anomer, which is excluded due to the observed NOE between H-1 of the sugar ring and H-4 of the chebuloyl group thereby demanding that the anomeric configuration be β . However, inspection of the other ${}^1J_{H,C}$ values for the sugar moiety of **1** in comparison to both α - and β -D-glucose reveals that all of the values for ${}^1J_{H,C}$ for the ring protons increase dramatically well beyond that expected for Perlin-type effects. If the axial/equatorial disposition was the only factor influencing the magnitude of ${}^1J_{H,C}$, then $J_{H2,C2}$, $J_{H3,C3}$, and $J_{H4,C4}$ should each increase by only 4 Hz whilst $J_{H1,C1}$ and $J_{H5,C5}$ should increase by *ca.* 8 Hz. Clearly the effect of strain and steric hindrance experienced by this structure renders the magnitude of ${}^1J_{H,C}$ somewhat unreliable for determining the disposition of the hydrogens. Nevertheless, the value of $J_{H1,C1}$ is consistent with a 1C_4 conformation, where for the β anomer, H-1 is equatorially orientated.

Even the ${}^3J_{H,C}$ values, accessible through the correlations present in the HMBC spectra, lend support for a 1C_4 conformation. Of the eight possible three-bond couplings involving the sugar carbons and the ring protons, both ${}^3J_{H2,C4}$ and ${}^3J_{H3,C1}$ would be unexpected to, and ${}^3J_{H5,C1}$ would be very unlikely to, give rise to observable correlations for the 3S_1 conformer in a HMBC spectrum optimized on a long-range coupling of 8 Hz due to their unfavorable dihedral angles in this conformation. These three correlations were in fact observed and of the eight possible three-bond couplings, it was only the ${}^3J_{H5,C3}$ correlation that was lacking. Hence, the original 1C_4 conformation is confirmed as this is the only conformer thoroughly consistent with all observed ${}^3J_{H,H}$, ${}^4J_{H,H}$, ${}^1J_{H,C}$, and ${}^3J_{H,C}$ values.

For the chebuloyl moiety of **1**, it is quite evident from the sizeable magnitude of the observed vicinal coupling (7.2 Hz) between H-3 and H-4 that these two protons are not in a cis relationship to one another. For a fused-ring system in a half-chair conformation (or more precisely because the structure is a lactone, a sofa conformation with C-3 as the head), a cis configuration for these two protons (axial-equatorial relationship for either of two conformations) would only be expected to furnish a considerably smaller coupling constant (4 – 5 Hz). Similarly, a trans-diequatorial relationship would also be expected to furnish a small coupling constant. However, an anomeric conformation with the two protons in a trans-diaxial relationship would be expected to provide a much larger coupling constant than the one observed. Thus a dynamic equilibrium consisting of two interconverting half-chair conformations where H-3 and H-4 have a trans configuration are held to account for the observed, intermediately-sized coupling constant as contributions arise from both diequatorial and diaxial relationships.²⁶ This trans relationship between these two protons is a point that seems to be occasionally misrepresented in the literature,² and is perhaps persistently so because of the frugal reporting of the NMR data for this otherwise well known compound. With the relative stereochemistry between these centers now established, examination of a Dreiding™

model revealed that of the two possible configurations, only the 3*R*,4*S* configuration could be manipulated to bring H-4 of the chebuloyl group into reasonable proximity to H-1 from the glucose moiety for observation of a NOE contact. It simply was not possible due to the constraints of the macrocycle to do so for the 3*S*,4*R* configuration. This then allowed assessment of the final stereocenter at position 2'. The observed coupling between H-2' and H-4 was 1.6 Hz, suggestive of a dihedral angle of *ca.* $\pm 63^\circ$ or $\pm 114^\circ$. Examination of the Dreiding™ model revealed that for the 2'*R*3*R*4*S* configuration, the geometric constraints of the macrocycle permitted only an angle of $+114^\circ$, whilst for the 2'*S*3*R*4*S* configuration angles of -63° and -114° were permissible. For the 2'*R*3*R*4*S* configuration with an angle of $+114^\circ$, the internuclei distances $r_{\text{H}2',\text{H}3}$ and $r_{\text{H}2',\text{H}4}$ suggested that the NOE should be larger between H-2' and H-3 in comparison to the NOE between H-2' and H-4. By contrast, the 2'*S*3*R*4*S* configuration with an angle of either -63° or -114° should result in only a comparatively small NOE between H-2' and H-3 in relation to H-2' and H-4. Experimentally, 0.67 and 3.09% were measured for $\eta_{\text{H}2',\text{H}3}$ and $\eta_{\text{H}2',\text{H}4}$, respectively. Thus, the 2'*S*3*R*4*S* configuration is concluded to be the absolute configuration of these three stereocenters. This is different in absolute terms even to those cases where a trans configuration has been reported³ for H-3 and H-4, although we note that the relative stereochemistry of these three centers in that reference is in accord with our determination. Needless to say, the observation of the NOE between H-1 of the glucose moiety and H-4 of the chebuloyl moiety was pivotal in fully describing the structure of chebulinic acid.

Structure, dynamic equilibria, and conformation of galloyl-free chebulinic acid (2)

Immediately apparent for **2** was the lack of the three galloyl groups present in **1** based on ¹H and ¹³C NMR. This, together with the MS, which accounted fully for the loss of these three units from chebulinic acid (**1**), led to the presumption that this compound was simply chebulinic acid (**1**) whereby the three galloyl groups had been hydrolyzed off as the presence of the chebuloyl group was still apparent from the ¹H and ¹³C NMR and the mapping the ¹H–¹³C spin network of the sugar ring again realized a hexose moiety based on the standard application of HMQC and COSY spectra. (The exception was the connectivity of sites 1 and 2 of the sugar ring which was based on HMBC due to the small value of $J_{\text{H}1,\text{H}2}$.) But, although bearing close similarity to **1** in that a chebuloyl group was bound to a sugar moiety, either the nature of the adopted sugar ring conformation or the sugar itself was clearly different based on the observed vicinal ¹H–¹H coupling constants of the sugar ring and thus the identity of the hexose was also required to be unequivocally confirmed as glucose by way of hydrolysis. Analysis by ¹H and ¹³C NMR of the hydrolyzed sample again yielded sets of signals consistent with an equilibrium solution of α - and β -glucose by comparison to an authentic solution of glucose, further confirmed by admixture with authentic glucose resulting in the enhancement of the glucose signals in both the ¹H and ¹³C NMR spectra. Thus the identification of **2** as a thrice-hydrolyzed derivative of **1**—hence the term galloyl-free chebulinic acid—was clear.

However, this apparent structural simplification was countered by several complications. For example, with the freeing of the anomeric carbon not only were α - and β anomers possible, but also other forms. Indeed the substantial presence of the open-chain form was clearly discernable in acetone-*d*₆ solution (together with a few drops of D₂O to facilitate dissolution) based on a

second set of signals observed in both the ^1H and ^{13}C NMR spectra and, most notably, the observed presence of an aldehydic proton and aldehydic carbon in their respective spectra. The open-chain form in the acetone- d_6 solution constituted as much as 40% of the sample, but with increasing amounts of D_2O the equilibrium shifted towards the cyclic form and in D_2O alone the pyranose form essentially predominated ($> 10:1$). In DMSO solution, the equilibrium was also shifted towards the open-chain form with approximately equal amounts of both forms, but the system was a complex mixture with a number of unidentified forms present and, hence, was difficult to unravel. Removal of the DMSO and redissolution in D_2O provided spectra congruent with the original sample in D_2O proving that the changes were not irreversible in addition to limited signs of interconversion by EXSY spectra. In the mixed solvent system D_2O – CD_3OD (1:1), both epimeric methyl hemi-acetals were formed in essentially equal amounts with only trace amounts of the free aldehyde ($< 2\%$). However, D_2O was the preferred solvent for analysis as, not only did it simplify the ^1H NMR spectrum, but it also eliminated overlap, *e.g.* between the residual acetone- d_5 signal and some of the analyte signals which precluded assessment of those particular coupling constants. The ^1H and ^{13}C NMR data are presented in Tables 2 and 3, respectively, for the pyranose form in D_2O and for both forms in acetone- d_6 – D_2O .

In fact, the absence of the galloyl groups opens up a bevy of structural possibilities. For example, septanose ring formation in addition to a pyranose ring, is conceivable. However, septanoses are generally unstable and are usually not observed unless the anomeric carbon is tied²⁷ and furthermore, suitable evidence was not forthcoming, *e.g.* by the observation of suitable HMBC correlations (see Figure 2 for the observed HMBC correlations) for this type of structure.

By the same token though, no direct evidence (*i.e.* appropriate HMBC correlations) for a pyranose ring could be elicited either. The 2,4 site attachment of the chebuloyl group though, was confirmed in both the cyclic and open-chain forms thereby eliminating both furanose ring formation and formation of a hemi-acetal ester, (*via* chebuloyl attachment to position 1 of the glucose), the latter of which is a quite stable structural unit.²⁸ The reason for consideration of structures other than the expected pyranose ring formation was the unusual set of vicinal couplings observed for the sugar ring protons of **2** ($J_{\text{H}_2,\text{H}_3}$ and $J_{\text{H}_3,\text{H}_4}$ are both small (1.8 and 0.4 Hz, respectively), as for **1**, but both $J_{\text{H}_1,\text{H}_2}$ and $J_{\text{H}_4,\text{H}_5}$ are considerable (7.8 and 9.4 Hz, respectively)) and not just that they were distinct to the corresponding couplings observed for **1**.

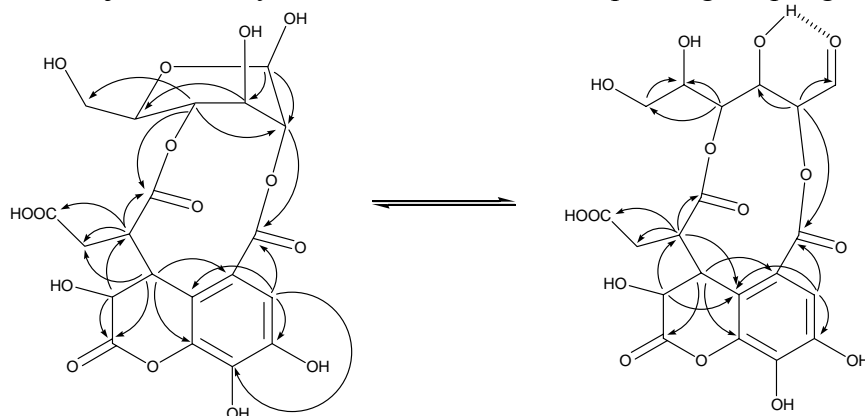


Figure 2. Selected HMBC correlations for the cyclic pyranose and open-chain forms of galloyl-free chebulinic acid (**2**).

It thus makes it apparent though, that the conformation of the sugar for the cyclic form of **2** differs substantially from that of **1** if a pyranose form is present for **1**, which is deemed to be the case lacking evidence for any other form. It is somewhat surprising that the conformation of the pyranose form of **2** is different to that found in **1** given that with removal of the galloyl groups a reduction in steric hindrance would be anticipated and that the inverted chair conformation would be maintained to accommodate the macrocycle by having the O-2 and O-4 linkages axially orientated. But it must be the case given the observations and, needless to say, other factors must also be playing a role in determining the adopted conformation. For the spectrum of **2** obtained in acetone- d_6 -D₂O solution, the vicinal couplings of the sugar ring are all of comparable size to their counterparts in the spectrum of **2** obtained in D₂O solution alone, inferring that the conformation of the pyranose form of **2** essentially does not change with the solvent. Also peculiar, except for J_{H_1,H_2} , the set of couplings corresponding to the ring vicinal couplings observed for the open-chain form were extremely close in magnitude to their cyclic form counterparts. The surprisingly small value observed for J_{H_1,H_2} in the linear form of **2** may be due to a stable structure where the dihedral angle between H-1 and H-2 is near 90°, which might arise from intramolecular hydrogen bonding (as indicated in Figure 2) restraining the free rotation about the C₁-C₂ bond which would otherwise yield a time-averaged dihedral angle between the two protons and result in a coupling of reasonable magnitude between H-1 and H-2. The formation of this hydrogen bond would presumably be energetically favorable, but with increasing competition from water the equilibrium would be expected, as observed, to shift towards the pyranose form. Essentially the same vicinal coupling constants were observed for both epimeric methyl hemi-acetals, and the magnitudes of J_{H_2,H_3} , J_{H_3,H_4} , and J_{H_4,H_5} were very similar to their counterparts in both the pyranose and open-chain (aldehyde) forms, implying a similar geometry also for these derivatives. The observed value of 7.2 Hz, though, for J_{H_1,H_2} in the methyl hemi-acetals is indicative of free rotation about the C₁-C₂ bond and thus substantiates the hydrogen bonding postulate for the open-chain form of **2**.

Given that pyranose was the likely structure of the cyclic form of **2** and that clearly the conformation of the glucose ring was not predominantly in either of the two possible chair conformations, an examination of all possible skew and boat conformations was conducted for **2** to rationalize the observed vicinal ring couplings. Additionally, another consideration that needed to be taken into account was that the observed results could be a time-average of a rapidly interconverting (on the NMR-timescale) equilibrium of two or more skew, boat, or chair conformations. To eliminate unlikely contributing conformers, three selection criteria were taken into account: extreme coupling values (either very small or very large values for a particular vicinal ring coupling in comparison to an observed value for that coupling which was conversely large or small, respectively); the observed NOEs; and the constraint of the chebuloyl group restricting at least one of the glucose linkages to an axial orientation. These selection criteria rule out both chair forms as significant contributors to the dynamic equilibrium: the ⁴C₁ conformation is ruled out on the basis of both chebuloyl linkages being equatorial and the large values expected for J_{H_2,H_3} and J_{H_3,H_4} in comparison to the observed small values; the ¹C₄ conformation is ruled out on the basis of the small values expected for J_{H_1,H_2} (the anomeric configuration was

proven to be β by an NOE between H-1 and H-3) and $J_{H4,H5}$ in comparison to the observed large values.

From an examination of the six possible skew conformers (see Chart 2) and an assessment of the likely vicinal couplings for the glucose ring protons that would result for each skew conformer, conformers 1S_3 and 2S_0 could both be possibly eliminated as significant contributors based on both the extreme divergence of three couplings (either very small or very large) from that observed coupled with the fact that for both of these conformers, tethering of the chebuloyl group would be equatorially orientated for one link and isoclinally orientated for the other (though in a cis relationship as governed by the configuration of glucose). For the other four skew conformations the combinations for the chebuloyl group tethering would either be axial and isoclinal or axial and equatorial. Various combinations and permutations of these four conformers could give rise to the observed time-averaged couplings, however, though NOE measurements yielded enhancements between H-1 and H-3 and between H-3 and H-5, none were forthcoming between H-1 and H-5. Both the 3S_1 and 0S_2 conformers could potentially yield an NOE between H-1 and H-5. Elimination of these two conformers would leave only the 5S_1 and 1S_5 conformers; the former of which should permit an observable NOE between H-1 and H-3 and the latter of which should permit an observable NOE between H-3 and H-5, and in both cases, exclusive of the other. A near proportionate contribution of these two conformers, qualitatively at least, would yield vicinal couplings for the ring protons consistent with those observed. However, there is also an NOE between H-4 and H-2 which needs to be accounted for. The 4C_1 conformer would readily provide this NOE contact, but would also provide an NOE between H-1 and H-5 which was not observed (meaningful in the presence of $\eta_{H1,H3}$ and $\eta_{H3,H5}$). Both linkages would also be equatorially orientated for this conformer. Returning to the 1S_3 and 2S_0 conformers, both of these can provide $\eta_{H2,H4}$ but not $\eta_{H1,H5}$ (the former will also provide $\eta_{H3,H5}$ whilst the latter will also provide $\eta_{H1,H3}$). So it is probable that one or other or both of these conformers are contributing also to the conformational equilibrium.

For the boat conformations (see Chart 3), $B_{3,0}$ could be eliminated as the chebuloyl tethering would be equatorial for both linkages whilst ${}^{1,4}B$ and ${}^{2,5}B$ conformations have three likely couplings that are too divergent. Interestingly, whilst a ${}^{3,0}B$ conformation would have both chebuloyl linkages axial, and all couplings could potentially be in accordance with those observed, there should also be a strong NOE between H-1 and H-5, thus eliminating it. Boat conformations $B_{1,4}$ and $B_{2,5}$ are similar to the skew conformations 5S_1 and 1S_5 in that whilst neither of them individually would be expected to give rise to the observed couplings, a conformational equilibrium could be envisaged that did. Although neither of these boat conformations ($B_{1,4}$ and $B_{2,5}$) would give rise to NOEs in contradiction of those observed, it is probably unlikely that they, as boat conformations, would be present to any great degree considering that boats are only observed in exceptional circumstances. It is worth noting though, that the boat conformations $B_{1,4}$ and $B_{2,5}$ are related to the skew conformations 1S_5 and 5S_1 , respectively, by a single twisting motion. However, it is postulated that the two skew conformers, 5S_1 and 1S_5 , and one or other or both of the 1S_3 and 2S_0 conformers probably represent the conformations contributing to the dynamic equilibrium of the sugar ring in **2**.

Comparison of the sugar ring $^1J_{H,C}$ values to those of α - and β -D-glucose reveal that with the reduction in strain and steric hindrance, normality has returned with respect to the magnitude of the $^1J_{H,C}$ values but, with the likely presence of a number of conformers, it is difficult to base a clear interpretation on them. Consideration of the dihedral angles suggests that in the 5S_1 conformer, J_{H_1,C_1} should be only slightly smaller (*ca.* 2 Hz) than the value observed in α -glucose as H-1 is trans-diaxial to only one atom. For the 1S_5 , 1S_3 , and 2S_0 conformers, the value of J_{H_1,C_1} should be similar to that of α -glucose. In other words, an equilibrium mixture of these four conformers should give rise to a slightly smaller value for J_{H_1,C_1} in comparison to α -glucose, and this is consistent with the observed value of 167.1 Hz. The abundance of both long-range $J_{H,HS}$ and $J_{H,CS}$ that were present for the glucose moiety in **1** was noticeably diminished in **2**.

The gross structural elucidation and signal assignment of the chebuloyl group that was present in **2** proceeded rapidly given previous experience and comparison with **1**. However, the change in the glucose ring conformation did not permit the same interactions between the glucose and chebuloyl moieties, but given the similarity of the NMR parameters for the chebuloyl moiety of **2** in comparison to that of **1**, despite the different solvents in use, the same arguments apply regarding the relative stereochemistry of the three stereocenters within the chebuloyl moiety. These can therefore be taken as 2'S3R4S. It can be assumed further that in all likelihood the absolute stereochemistry also remains invariant.

Finally, although the presence of **2** as an artifact of the extraction and isolation procedure virtually has been discounted by the careful raw extraction of the dried fruit under mild conditions and immediate analysis by HPLC-ESI-MS – thereby confirming its presence in the dried fruit– it remains to be determined whether it is a biosynthetic intermediate, either catabolic or anabolic, of the producing plant or, indeed, if it is an intentional biosynthetic target *per se*.

Experimental Section

General Procedures. NMR spectra were acquired on a JEOL Alpha 500 NMR (or JEOL Lambda 400) spectrometer {11.75 (9.40) T} equipped with either a 5 mm normal-configuration tunable ($^{13}C\{^1H\}$) probe or a 5 mm inverse z-axis field-gradient ($^1H\{X\}$) probe operating at 500.16 (399.78) MHz for 1H and 125.78 (100.54) MHz for ^{13}C . Spectra were generally acquired at 25 °C and variously at other temperatures as indicated using acetone-*d*₆ and/or D₂O, DMSO-*d*₆, or CD₃OD as solvent. 1H and ^{13}C spectra for acetone-containing samples were referenced internally to the solvent signal using values of 2.04 and 29.80 ppm, respectively; for D₂O (DMSO-*d*₆) samples the spectra were referenced internally to TSP (TMS) using 0.00 (0.00) ppm for both nuclei. Spectral widths of 2-D spectra were appropriately selected from the 1-D spectra and acquired with an adequate level of resolution. All experiments were performed using standard, vendor-supplied pulse sequences.

1-D 1H spectra were acquired with single-pulse excitation, 45° flip angle, pulse recycle time of 9.5 s, and with spectral widths of 7 kHz consisting of 64 k data points (digital resolution 0.11 Hz/pt), zero-filled to 128 k prior to Fourier transformation. Spin analysis was performed using Perch²⁹ iteration software for the extraction of 1H chemical shifts and $J_{H,H}$ coupling constants.

Since the reliable extraction of small couplings approaching the linewidth is heavily dependent on whether they are to a degree resolvable on at least one spin for Perch to reliably extract them, only those couplings reliably extracted by Perch are reported whilst couplings buried in the linewidth on both interacting spins are not reported (*i.e.*, not extracted) even if their likely presence is probable or is evident from homodecoupling experiments or COSY experiments. NOE difference measurements were acquired using saturation times of 7–12 s at a reduced level of resolution (3.9 Hz/pt); 1 Hz of exponential weighting was usually applied prior to Fourier transformation and signal enhancement was integrated relative to the intensity of the irradiated signal set to –100%. Prior to NOE measurements, samples were deoxygenated by effusion with nitrogen. Both double quantum-filtered COSY (including the fg version) and field-gradient EXSY spectra were acquired in phase-sensitive mode and processed with zero-filling ($\times 2$, $\times 4$) and exponential weighting (1–3 Hz, 3–5 Hz) applied in both dimensions prior to Fourier transformation. For EXSY spectra, various mixing times of up to 8 s were utilized.

1-D ^{13}C spectra were acquired with single-pulse excitation, broad-band ^1H decoupling, 45° flip angle, pulse recycle time of 3.5 s and with spectral widths of 30 kHz consisting of 64 k data points (digital resolution 0.46 Hz/pt), zero-filled to 128 k and with 1 Hz exponential weighting applied prior to Fourier transformation. DEPT 135° spectra were acquired with similar spectral windows and processing as utilized for the 1-D carbon spectra but with a post acquisition delay time of 3 s. Field-gradient HMQC, field-gradient HMBC, and HMBC-BIRD experiments were all acquired in magnitude mode and processed with zero-filling ($\times 2$, $\times 4$), a $\pi/(3-8)$ -shifted sinebell function, and exponential weighting (3–5 Hz, 5–25 Hz) applied in both dimensions prior to Fourier transformation. For HMBC-BIRD experiments, the length of the BIRD relaxation delay (typically 400 ms) was optimized on the incoming FID. For all ^1H – ^{13}C correlation experiments, $^1J_{\text{H,C}}$ was optimized on a value of 145 Hz whilst the HMBC correlations were optimized for a long-range $^nJ_{\text{H,C}}$ coupling of 3 or 8 Hz.

Analytical HPLC-DAD analyses were performed using a Spherisorb 5 ODS-2 column (250 \times 4.6 mm). The HPLC system consisted of a L-7100 pump, a L-7455 diode array detector, and a programmable auto sampler L-7250). The flow rate was maintained at 1.0 mL min^{-1} with a column back pressure of 70–136 bar. Two solvents were used for elution, acetonitrile (A) and 0.02–5% formic acid (B). The elution profile was 0–35 min, 0–30% A in B (linear gradient). The injection volume was 20 μL .

Analytical HPLC-ESI-MS analysis was performed using the same column as above and a Perkin-Elmer Sciex API-365 triple quadrupole mass spectrometer equipped with a pneumatically-assisted ion spray interface and operated in negative ion mode. The HPLC system consisted of a Perkin-Elmer Series 200 HPLC unit equipped with a UV-vis detector (detection wavelength 280 nm). Prior to entering into the MS, the HPLC eluent was split 5:1 from a constant flow rate of 1 mL min^{-1} . Two solvents were used for elution, acetonitrile (A) and 1% formic acid (B). The elution profile was 0–35 min, 0–30% A in B (linear gradient). The injection volume was 20 μL . For the MS, the needle voltage was set at –4,000 V, the orifice plate voltage at –35 V and the ring voltage at –220 V. The nebulizer gas (purified air) flow was set at position 9 and that for the curtain gas (N_2) at position 12. The heated nitrogen gas temperature was set at 300 $^\circ\text{C}$ and the gas flow rate at 7 L min^{-1} . Masses were scanned from m/z 100 to 1,100 in 0.30

amu steps. HRMS were performed using a VG Analytical ZabSpec instrument under ESI⁻ conditions using a peak matching technique with PFK as a reference substance at a resolution of 8,000–10,000 (at 10% peak height).

Plant material. Air-dried *T. chebula* fruit was purchased from a local herb market in Lahore, Pakistan and identified by Prof. Dr. Zahoor Ahmed from the Department of Botany, Punjab University, Lahore, Pakistan. A voucher specimen has been deposited in the Turku University Herbarium (voucher no. TUR 354114).

Extraction and isolation of compounds 1–6. Detailed extraction and isolation procedures for 1–6 have been described previously.¹⁸ Briefly, the dried fruits (200 g) were extracted with 70% aqueous methanol (4 × 1 L with 1 h of continuous stirring for each extraction) after which the methanol was removed under reduced pressure at 35 °C followed by freeze-drying to yield 62 g of extract. The extract (30 g) was then redissolved in 750 mL of deionized water and successively partitioned into *n*-hexane, chloroform, and 1-butanol phases (3 × 750 mL each) using a separatory funnel. After removal of the solvents, the extracts from the water and 1-butanol phases were fractionated using size-exclusion chromatography (column, 40 × 3.8 cm; stationary phase, Sephadex LH-20; particle size, 18–111 μm) by consecutive elution with water, 30, 50, and 80% aqueous methanol, and 30, 50, and 80% aqueous acetone. Separate 30% aqueous methanol fractions of the 1-butanol-phase partition furnished gallic acid (**4**; 3.7 mg; UV, λ_{max}, 269.0 nm) and ethyl gallate (**5**; 8.5 mg; UV, λ_{max}, 270.0 nm) as precipitants at ambient temperature. The 80% aqueous acetone fraction of the water-phase partition furnished ellagic acid (**3**) as a precipitant upon cooling to 4 °C. The 80% acetone fraction of the 1-butanol-phase partition was redissolved in 60% aqueous methanol and subjected to preparative RP-HPLC (LiChroprep RP-C18 column, 44 × 3.7 cm, 40–63 μm). The elution profile consisted of a linear gradient, 0–25% acetonitrile in 5% aqueous formic acid over 100 min (flow rate, 4 mL/min; UV detection at 280 nm). Fractions containing chebulinic acid (**1**) were pooled. The 30% aqueous methanol fraction of the water-phase partition was subjected to semi-preparative RP-HPLC (LiChrospher RP-C18 column, 25 × 1.0 cm, 10 μm). The elution profile consisted of a linear gradient, 0–25% acetonitrile in 0.1% aqueous formic acid over 40 min (flow rate, 3 mL min⁻¹; UV detection at 280 nm). Fractions containing galloyl-free chebulinic acid (**2**) were pooled. Fractions containing luteolin (**6**) were pooled and the solvent removed yielding 5.3 mg of **6** (UV, λ_{max}, 254, 350 nm). The identification of 4–6 was based on comparison of their retention times, ESI⁻-MS, and λ_{max}s with standard samples.

1,3,6-Tri-*O*-galloyl-2,4-chebuloyl-β-D-glucopyranoside (1, chebulinic acid) was obtained as a white powder after removal of the solvent (141 mg). UV (λ_{max}, nm) 246.4 (shoulder), 273.3; ¹H NMR, see Table 2; ¹³C NMR, see Table 3; ESI⁻-MS *m/z* 955 [M – H]; HRESI-MS *m/z* 955.1066 [M – H] (calcd for C₄₁H₃₁O₂₇, 955.1053).

2,4-Chebuloyl-β-D-glucopyranoside (2, galloyl-free chebulinic acid) was obtained as orange, amorphous solid after removal of the solvent (85 mg). UV (λ_{max}, nm): 243.3, 271.0; ¹H NMR, see Table 2; ¹³C NMR, see Table 3; ESI⁻-MS *m/z* 499 [M – H], 999 [2M – H]; HRESI-MS *m/z* 499.0756 [M – H] (calcd for C₂₀H₁₉O₁₅, 499.0724).

Ellagic acid (3) was obtained as a white solid (85 mg). UV (λ_{\max} , nm): 249.7, 366.9; ^1H NMR (δ ppm, DMSO- d_6 , lit value²⁰ in parentheses): 7.44 (s, 7.47); ^{13}C NMR (δ ppm, DMSO- d_6 , lit values²⁰ in parentheses) 107.39 (s, 107.6), 109.99 (d, 110.2), 112.07 (s, 112.2), 136.14 (s, 136.3), 139.34 (s, 139.4), 147.86 (s, 148.0), 158.86 (s, 159.0); ESI⁻-MS m/z 301 [M – H], 603 [2M – H].

Hydrolysis of 1 and 2. Aqueous solutions of **1** and **2** were separately treated with dilute ammonia solution and left to stand at room temperature overnight. Following evaporation of the solvent, the residues were taken up in D₂O and analyzed by ^1H and ^{13}C NMR yielding sets of signals consistent with an equilibrium solution of α - and β -D-glucose by comparison to a solution of authentic glucose at similar concentration. Further confirmation of the presence of glucose in each case was effected by mixed analysis at appropriate rates with the authentic glucose solution resulting in the enhancement of the glucose signals in all spectra.

DPPH radical scavenging assay. The assay was carried out essentially according to the procedure of Kähkönen and Heinonen.³⁰ In a cuvette, 2.950 mL of 0.1 mM methanolic DPPH solution was mixed with 50 μL of a methanolic solution of the sample to yield final cuvette concentrations of 16.7 μM for the individual compounds and 1 mg mL⁻¹ for the extract and tannic acid. The absorption was monitored at intervals of 15 s for 5 min. The results were expressed as the percentage of radicals scavenged after 4 min of reaction. The values expressed were calculated according to the following equation:³⁰

$$\% \text{ of radicals scavenged} = 100 \times (A1 - A2) / A1$$

where A1 = absorbance immediately after the addition of the test sample ($t = 0$) and A2 = absorbance after 4 min. Trolox was used as a reference.

Inhibition of the autoxidation of methyl linoleate assay. The assay was carried out essentially according to the procedure of Kähkönen and Heinonen.³⁰ Test samples were dissolved in methanol and added to methyl linoleate (0.2 g) followed by evaporation of the methanol under nitrogen. The extract and tannic acid were tested at a concentration of 50 ppm whilst pure compounds were tested at a concentration of 0.1 mM. The oxidation of methyl linoleate was performed in the dark at 40 °C and the formation of hydroperoxides was monitored by measuring the formation of conjugated diene hydroperoxides spectrophotometrically at 234 nm.³¹ The antioxidant activity was expressed as the percentage inhibition of the formation of methyl linoleate hydroperoxides after 70 h of oxidation. α -Tocopherol and Trolox were used as references.

Appendix

Table A1. Skew (twist-boat) conformers

vicinal relationship	dihedral angle	magnitude of J
eq–eq	60	medium to small
ax–ax	180	large
ax–eq	60	medium to small
ax–iso <i>cis</i>	30	medium to large
ax–iso <i>trans</i>	120	medium to large
eq–iso <i>cis</i>	30	medium to large
eq–iso <i>trans</i>	90	very small

Table A2. Boat conformers

vicinal relationship	dihedral angle	magnitude of J
eq–eq	0	large
ax–ax	0	large
ax–eq	120	medium to large
ax–bowsprit	60	medium to small
ax–flagpole	180	large
eq–bowsprit	60	medium to small
eq–flagpole	60	medium to small

Acknowledgments

Financial support from the Academy of Finland, grant no. 4284 (K.P.), is gratefully acknowledged.

References and Notes

1. The term “hydrolyzable tannins” is one that is open-ended in definition and can encompass various sets of compounds. In essence,² it refers to the ellagitannins and gallotannins which are compounds where one or more gallic acid units are attached to a glucose core *via* ester bonds formed between the hydroxyls of the sugar and the acid functionalities of the gallic acids. The distinction between ellagitannins and gallotannins is that in the former one or more pairs of galloyl groups are linked together by a C₂–C₂ bond, thus forming a hexahydroxydiphenoyl unit (HHDP) spanning positions C-1/C-6, C-3/C-6, C-4/C-6, C-2/C-3, C-3/C-4, or C-2/C-4.⁴ In the latter category the galloyl groups are not linked.
2. *Phytochemical Dictionary. A Handbook of Bioactive Compounds from Plants*; Harborne, J. B.; Baxter, H.; Moss, G. P., Eds; Taylor & Francis: London, 1999; 2nd Edn, Ch. 45, pp 570–574.
3. Quideau, S.; Feldman, K. S. *Chem. Rev.* **1996**, *96*, 475.
4. Feldman, K. S.; Malliga, R. I.; Liu, Y. *J. Org. Chem.* **2003**, *68*, 7433.
5. Lin, T.-C.; Hsu, F.-L. *Chin. Pharm. J.* **1996**, *48*, 167.
6. Saleem, A. Biological activity of plant phenolics with an emphasis on Pakistani plants used in traditional health care. Ph.D. Thesis, University of Turku. ISBN 951-29-2188-9, Painosalama Oy. Turku, Finland, **2002**.
7. Feldman, K. S.; Ensel, S. M. *J. Am. Chem. Soc.* **1994**, *116*, 3357.
8. Ding, G.; Liu, Y.; Wang, L.; Ji, C.; Sheng, L. *Zhongguo Yaoke Daxue Xuebao* **2001**, *32*, 91.
9. Haddock, E. A.; Gupta, R. K.; Haslam, E. *J. Chem. Soc., Perkin Trans. 1* **1982**, 2535.
10. Haslam, E.; Uddin, M. *J. Chem. Soc. (C)* **1967**, 2381.
11. Jochims, J. C.; Taigel, G.; Schmidt, O. Th. *Liebigs Ann. Chem.* **1968**, 717, 169.
12. Liu, Y.; Yu, B.; Ding, G.; Wu, Y. *Chin. Chem. Lett.* **1998**, *9*, 827.
13. In amariinic acid, a hydroxy group replaces H-2' of chebulagic acid, the corresponding ellagitannin of chebulinic acid (**1**).
14. Foo, L. Y. *Phytochemistry* **1995**, *39*, 217.
15. James, L. C.; Roversi, P.; Tawfik, D. S. *Science* **2003**, *299*, 1362.
16. Abe, I.; Kashiwagi, Y.; Noguchi, H.; Tanaka, T.; Ikeshiro, Y.; Kashiwada, Y. *J. Nat. Prod.* **2001**, *64*, 1010.
17. Saleem, A.; Ahotupa, M.; Pihlaja, K. *Z. Naturforsch.* **2001**, *56c*, 973.
18. Saleem, A.; Husheem, M.; Härkönen, P.; Pihlaja, K. *J. Ethnopharmacol.* **2002**, *81*, 327.
19. Saleem, A.; Engström, M.; Wurster, S.; Savola, J.-M.; Pihlaja, K. *Z. Naturforsch.* **2002**, *57c*, 332.
20. Li, X.-C.; Elsohly, H. N.; Hufford, C. D.; Clark, A. M. *Magn. Reson. Chem.* **1999**, *37*, 856.
21. Chevalley, I.; Marston, A.; Hostettmann, K. *Phytochemistry* **1999**, *50*, 151.
22. Frankel, E. N. *Lipid Oxidation*; The Oily Press: Scotland, 1998.
23. Huang, Y.-L.; Chen, C.-C.; Hsu, F.-L.; Chen, C.-F. *J. Nat. Prod.* **1998**, *61*, 1194.
24. Latté, K. P.; Kolodziej, H. *Phytochemistry* **2000**, *54*, 701.
25. Callam, C. S.; Gadikota, R. R.; Lowary, T. L. *J. Org. Chem.* **2001**, *66*, 4549.

26. Pihlaja, K.; Tähtinen, P.; Klika, K. D.; Jokela, T.; Salakka, A.; Wähälä, K. *J. Org. Chem.* **2003**, *68*, 6864.
27. Tran, T. Q.; Stevens, J. D. *Aust. J. Chem.* **2002**, *55*, 171.
28. Satoh, T.; Sugiyama, S.; Ota, H. *Tetrahedron Lett.* **2002**, *43*, 3033.
29. See for example, Laatikainen, R.; Niemitz, M.; Weber, U.; Sundelin, J.; Hassinen, T.; Vepsäläinen, J. *J. Magn. Reson., Ser. A* **1996**, *120*, 1, or the program website at <http://www.perchsolutions.com>
30. Kähkönen, M. P.; Heinonen, M. *J. Agr Food Chem.* **2003**, *51*, 628.
31. Kähkönen, M. P.; Hopia, A. I.; Vuorela, H. J. ; Rauha, J.-P.; Pihlaja, K.; Kujala, T. S.; Heinonen, M. *J. Agr. Food Chem.* **1999**, *47*, 3954.