

HETEROCYCLES, Vol. 79, 2009, pp. 583 - 598. © The Japan Institute of Heterocyclic Chemistry
Received, 9th September, 2008, Accepted, 28th January, 2009, Published online, 28th January, 2009.
DOI: 10.3987/COM-08-S(D)19

DIOICINE: A NOVEL PRENYLATED PURINE ALKALOID FROM *GYMNOCLADUS DIOICUS*¹

Richard W. Fitch,^{a*} Thomas F. Spande,^b H. Martin Garraffo,^b Rachael R. Chase,^{a,2} Mylaka A. Clinedinst,^{a,2} Derek A. Parkes,^{a,2} Richard Reed,^{a,2} Noel F. Whittaker,^b and John W. Daly^{b,3}

^aDepartment of Chemistry, Indiana State University, 600 Chestnut Street, Science S35E, Terre Haute, IN, 47809, USA ^bLaboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, 20892, USA

Abstract – The Kentucky coffeetree, *Gymnocladus dioicus* (L., K. Koch) is a leguminous tree reputed to be toxic to livestock and to contain the toxic nicotinic acetylcholine receptor agonist cytisine. Analysis of extracts of various tree parts by gas chromatography-mass spectrometry failed to reveal the presence of cytisine. Neither [³H]-epibatidine binding in rat cerebral cortex nor functional fluorescence assays in cultured cells expressing nicotinic receptor subunits indicated significant activity. However, a novel purine alkaloid, 3-((*E*)-3-methylbuta-1,3-dienyl)-1,7-dimethylisoguanine, which we have named dioicine, was identified as the major lipophilic alkaloid in the methanolic extract from leaves and seeds. Dioicine and its hydrolysis products are likely to be responsible for the historical use of the tree's seeds as a coffee substitute. Herein we describe the isolation, structure elucidation and preliminary biological characterization of dioicine.

INTRODUCTION

Nicotinic acetylcholine receptors (nAChR) are important targets for the development of therapeutics, owing to their involvement in cellular signaling in both the peripheral (PNS) and central (CNS) nervous systems.⁴ Our interest in these receptors stems from the ability of a number of naturally occurring alkaloids to function as nAChR ligands. These include the well known tobacco alkaloid nicotine, from

which the receptors derive their name, as well as the quinolizidine alkaloids cytisine and sparteine (**Figure 1**).^{4,5} Nicotine and cytisine function as nicotinic agonists, while sparteine is an antagonist. The latter two alkaloids are derived from leguminous plants (Fabaceae). Based on this observation and the well-documented occurrence of other quinolizidine alkaloids in this family,⁶ we have embarked on a survey of local legumes with the aim of identifying novel nicotinicly active alkaloids. As part of this study, we have examined the Kentucky coffeetree, *Gymnocladus dioica* (L., K. Koch).⁷ The common name for this tree derives from reports that early settlers used its seeds as a source of a coffee-like beverage purportedly produced by roasting and steeping them. However, there are also numerous published reports that the tree is toxic and contains the quinolizidine alkaloid cytisine as the active principle.⁸⁻¹⁰

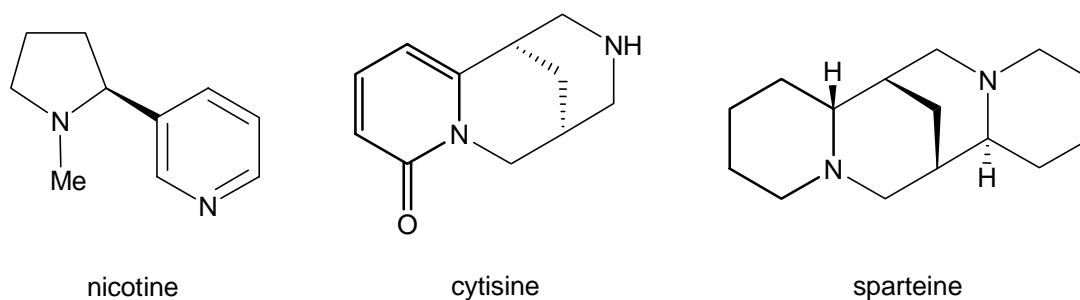
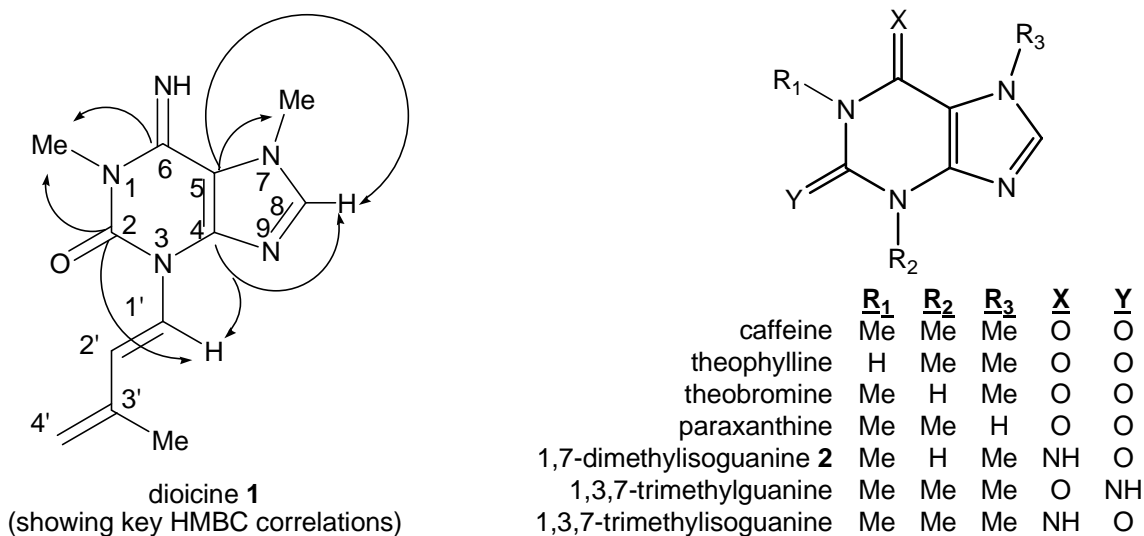


Figure 1

Reports of animal toxicity are documented for tree cuttings^{8a-d} and there is at least one anecdotal case of human poisoning, reportedly resulting from ingestion of the mucilage contained within the seed pods.^{9a-d} This poisoning was presumed to be due to confusion between this tree and the closely related Honey Locust *Gleditsia triacanthos*, whose mucilage is reported to be edible and sweet, hence the name.^{9d} These reports consistently cite cytisine as the toxic principle, but give no data or primary references to support this assertion.¹⁰ Nonetheless, the presence of cytisine would be consistent with the chemotaxonomy of this family⁶ and would also suggest the possibility of finding related bioactive alkaloids. Thus, we decided to examine the alkaloid content of this tree. To our knowledge, no reports of such an investigation exist.¹¹

We report herein the results of this analysis of this tree for alkaloids and the isolation and identification of a novel prenylated purine alkaloid, which we have named dioicine (**1**). Dioicine is a purine alkaloid, containing a 1,7-dimethylisoguanine moiety and bearing a significant resemblance to caffeine (**Figure 2**). The presence of methylpurines is more consistent with the historical reports of the use of this tree as a coffee substitute.^{7,8}



RESULTS AND DISCUSSION

Several specimens of *G. dioicus* were collected in the Terre Haute, IN area at three separate locations. We obtained leaf, stem, bark and seed pod samples over the course of several years. Extracts of the fresh or fresh-frozen parts were prepared by grinding the plant material under liquid N₂ and macerating with MeOH. The extracts were subsequently diluted with water and subjected to acid-base partitioning to obtain acid, base, and neutral fractions. Extracts and alkaloid fractions were prepared for the leaves, bark, green stems and seed pods/contents (husk, seed shells, seed pulp and surrounding mucilage). The crude extracts and base fractions were then evaluated by gas chromatography-mass spectrometry (GC-MS) to identify cytisine and any other volatile alkaloids that might be present. Cytisine is easily recognizable by the parent and primary fragments at *m/z* 190, 146 and 147, respectively in the EI spectrum or the protonated molecular ion at *m/z* 191 in the CI (NH₃) MS spectrum. Indeed 1.9 ng of cytisine (1 μL of a 10 μM solution in MeOH) was easily detected by either method. However, cytisine was not detected in any of the plant parts examined to a detection limit of 0.001% by weight. Instead, a single major lipophilic alkaloid was observed in the extracts. This alkaloid had a molecular weight of 245 by GC-EIMS, confirmed by CI-MS (NH₃, *m/z* 246). Major EI fragments were consistent with losses of methyl, formyl, formimino, acetamido, and a C₅ doubly-unsaturated hydrocarbon. LC-MS using an acetonitrile/water elution gradient with APCI ionization produced similar results to gas phase CI and indicated a single exchangeable proton when water was replaced with D₂O. HRMS in the EI mode gave a molecular formula of C₁₂H₁₅N₅O, which was inconsistent with known quinolizidine alkaloids.

A GC-FTIR spectrum indicated a carbonyl absorption at 1710 cm⁻¹ with alkene/arene absorptions at 1657 cm⁻¹ and 1531 cm⁻¹, but inconsistent with ordinary phenyl bands. A weak band at 3380 cm⁻¹ indicated a possible NH stretching absorption. Condensed phase spectra of a purified sample indicated significant

differences. While a sharp NH absorption band was present at 3332 cm^{-1} , no obvious carbonyl could be identified above 1700 cm^{-1} . Instead the major absorptions were seen at 1663 and 1634 cm^{-1} , possibly due to intermolecular interactions in the condensed phase.

The proton NMR spectrum in CDCl_3 showed three methyl singlets, two vinylic CH, and three possible aromatic CH signals. Though indicated by deuterium-exchange in LCMS and IR, no NH was observed in ^1H NMR in CDCl_3 . The carbon spectrum showed 12 signals with five quaternary carbons and four protonated carbons in the aromatic/alkene region between 105 and 150 ppm. Although the GC-FTIR indicated a carbonyl, no clearly assignable carbon for this group could be found initially. No signals were observed downfield of 150 ppm in CDCl_3 , although four quaternary candidates between 139 and 149 ppm were observed. HSQC-TOCSY spectra indicated five spin systems consisting of two isolated methyls, an allylic or *N*-methyl closely associated with a vinylic methylene, a coupled pair of vinyl/aryl CH, and an isolated vinyl/aryl CH. Given the sparse and isolated nature of these small systems and the presence of five nitrogens, a substituted purine became a working hypothesis. The correlation of the upfield methyl with the vinylic methylene suggested an isopropenyl moiety which was connected to the pair of vinylic CHs. This was further established by HMBC and NOESY correlations which clearly showed the allylic methyl related to both the terminal and internal vinylic systems (**Figure 2**). This allowed us to unambiguously establish an isopentadienyl system, likely as part of an enamine. The 7-methyl was cleanly established by COSY, HMBC and NOE correlations to H8 of the imidazole ring. The N1 methyl group was assigned by comparison with those of known methylxanthines and guanines/isoguanines in several solvents including CDCl_3 , $\text{DMSO-}d_6$, CD_3OD , and D_2O .¹²⁻¹⁸ The methyl pattern of the common methylpurines follows a trend in chemical shift where methyl on N7 (or N9) gives the greatest chemical shift, followed by N3 and N1, though the latter two are often close together. Our pattern was most consistent with that of 1,7-disubstitution. The UV maximum of 290 nm observed is also consistent with a number of purines,¹⁹ though one would expect a bathochromic shift with the extension of conjugation afforded by the isopentadienyl side chain. The position of the side chain was finally established by HMBC correlations, the key ones being H1' with both C2 and C4. The latter two atoms correlated in turn with the N1 methyl and H8 respectively (**Figure 2**). This also confirmed the placement of the N1 methyl group by correlation with both C2 and C6.

The greatest difficulty in the final structure assignment was the placement of the exocyclic nitrogen. From IR measurements it was not possible to correlate the two in the absence of reference material, as IR bands are often not reported in the literature. Attempts to corroborate NMR data from the literature¹²⁻¹⁸ were confounded by a number of factors, including the different solvents in which data were reported, the only very slight differences in chemical shift of the imine and carbonyl carbons, and the

unknown influence of the conjugated side-chain on chemical shifts through the intervening aromatic system.

In order to simplify matters, we decided to look at the hydrolysis products arising from the cleavage of the side chain at N3. Treatment of an NMR sample of **1** with excess trifluoroacetic acid (TFA) in CDCl_3 resulted cleanly in stoichiometric formation of prenal (3,3-dimethylacrolein), identical with an authentic sample. Also produced was a dimethylpurine (**2**), again having the methylation pattern consistent with paraxanthine by NMR. Unfortunately, the dimethylpurine could not be unambiguously assigned as the guanine or isoguanine from NMR data alone. We prepared the isomeric 1,7- and 1,9-dimethylguanines by treatment of guanine with methyl iodide, either in aqueous base with tetrabutylammonium chloride as a catalyst, or in aqueous MeOH made basic with sodium hydroxide, which provided isomeric dimethylguanines in low yield as a mixture, also containing unreacted guanine.²⁰⁻²² However, we were unable to match either dimethylguanine with our material. Although synthetic **2** is known, no NMR data was presented for the compound, but only a melting point, that we were unable to reproduce.²² The 1,3,7-trimethylguanine¹⁸ and isoguanine¹⁶ are also known and their NMR and mass spectra are fully assigned. Unfortunately NMR spectra were obtained in two different solvents. We prepared samples in both, but the spectra were not sufficiently distinct to allow unambiguous assignment.

Ultimately, we turned to mass spectrometry to answer the guanine/isoguanine question. The EI mass spectra of xanthenes and guanines/isoguanines are characterized by a retro-Diels-Alder extrusion of N1 and C2 along with any attached substituents.²⁴ In our case, the parent alkaloid **1** lost 58 to give m/z 187 (M-H - MeN=C=O) in accord with the same loss for the known 1,3,7-trimethylisoguanine,¹⁵ followed by consecutive losses of HCN yielding ions at m/z 160 and 133. Conversely, the known 1,3,7-trimethylguanine¹⁸ loses 55 and 56 mass units, consistent with the loss of methylcyanamide. The retro-Diels-Alder reaction also appears to be prevalent in chemical ionization spectra. CI-MS/MS of **1** (m/z 246) with ammonia as the reagent gas shows the losses of ammonia (m/z 229), MeN=C=O (m/z 189), the isopentadienyl side chain (m/z 180) and subsequent loss of MeOH (m/z 148).

GC-MS analysis of the TFA hydrolysate indicated prenal, as had NMR, identical in all respects with a commercial sample. The purine product gave the expected m/z 179 for a dimethylguanine/isoguanine. The loss of 57 amu (m/z 122) again supported the isoguanine structure **2** through loss of methyl isocyanate, which was confirmed by HRMS. The m/z 122 ion in EIMS was accompanied by the expected sequential losses of HCN (m/z 95 and 68).²⁴ Hydrolysis using aqueous HCl at rt also gave **2**, albeit at a much slower rate. In order to obtain complete hydrolysis, we treated the initial HCl hydrolysate with NaNO_2 , nitrosating the pendant primary amine and thereby facilitating hydrolysis.²¹ This gave paraxanthine which was identical in retention time and MS fragmentation with a commercial sample. Likewise, a sample allowed to stand in $\text{DCl}/\text{D}_2\text{O}$ for a few weeks showed the

dimethylisoguanine and a small amount of paraxanthine by ^1H NMR.¹³ The latter had identical chemical shifts compared with commercial material in the same solvent. This confirms the placement of the methyl groups at N1 and N7. Attempted hydrolysis in boiling water led to very little reaction and mostly recovery of starting dioicine.

Based on these data, we have assigned the isoguanine structure **1** to the alkaloid (**Figure 2**). We have named this compound dioicine, after the tree from which it was obtained. Dioicine is present in many of the parts of the coffetree, and its content varies across the plant based on GC-MS analysis. Higher amounts of **1** are found in the leaves and seeds, as well as immature seed pods, while lower amounts are found in the stem. Both new and old bark are essentially devoid of **1**, while content in leaves varies with season and altitude. Leaves collected in mid-July 2007 contained almost ten times more **1** than those collected in late August of the following year. This would be consistent with decreased need for deterrence of herbivory as senescence approaches.²⁵ Indeed, while immature seed pods contained relatively high levels of **1**, mature pods from the tree or ground contained very little.²⁶ Leaves on saplings at 1-2 meters elevation contained almost 10-fold more **1** than those from a limb at 4-5 meters, again consistent with deterrence of herbivory by browsers.²⁷ Also identified by GC-MS was an earlier-eluting isomer of **1**, that we ascribe to the *Z*-isomer of the dienamine side chain.²⁸ We initially ascribed its presence to isomerization in the injector, but variation of *Z* / *E* ratios in various samples suggested this was not the case. Substantial differences were found between samples, with greater proportions of the *Z* being present in the leaves collected later in the season or at higher points on the tree. Immature seed pods had a *Z* / *E* ratio of 1:4,²⁹ while the mature seed pods contained almost exclusively *Z* as did the leaf stems. We ascribe this to photoisomerization, with those parts having greater higher sun exposure containing higher proportions of *Z*-**1**. The seeds, which are sheltered from sunlight contain almost exclusively the *E*-isomer.

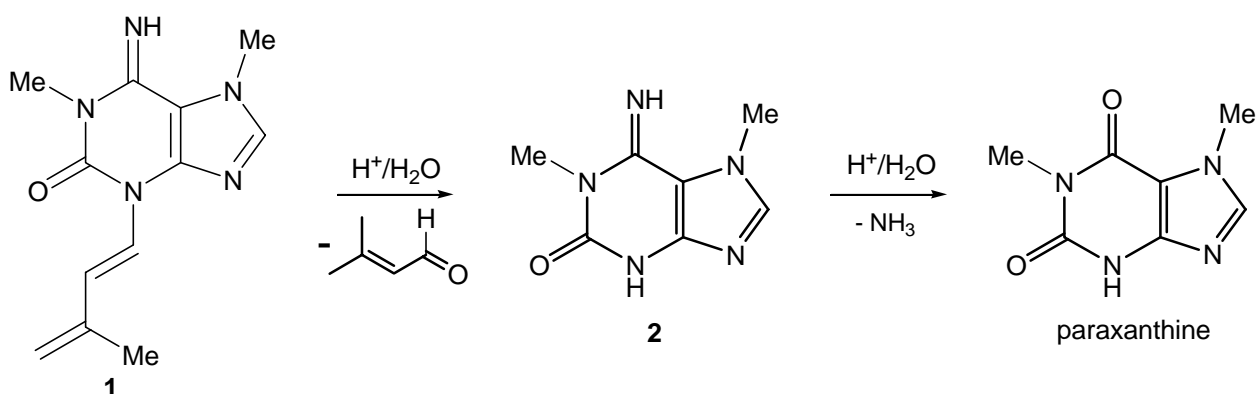
For comparison purposes, we also collected leaf samples from the related Honey locust, (*Gleditsia triacanthos* L.), Black locust (*Robinia pseudoacacia* L.) and Bristly locust (*Robinia hispida* L.) in the Terre Haute area (see **EXPERIMENTAL**). None were found to contain detectable levels of **1** by GC-MS.

The Kentucky coffetree gets its name from anecdotal reports of early American settlers roasting and steeping the seeds of this tree to produce a coffee-like beverage.⁷⁻¹⁰ Indeed, seeds of this tree were given to president Thomas Jefferson by General George Rogers Clark.³⁰ The president expressed significant interest in the tree and requested additional seeds in later correspondence.³¹ Reports suggest that the beverage was supplanted by true coffee as it became available.^{7b} While this may be true, the use of coffetree seeds to produce as a beverage would at first seem at odds with the reported toxicity of the tree, as well as the reputed presence of the potent nicotinic receptor agonist cytisine.⁸⁻¹⁰ While the presence

of cytosine would be consistent with the chemotaxonomy of this family,⁶ the use of a beverage containing cytosine as a stimulant similar to the use of coffee would seem unlikely, even though both are stimulants. The effects of oral cytosine/laburnum³² are quite different from that of oral caffeine.³³ Reports of coffeetree poisoning do indicate some similar effects, though reports are spotty in both animals⁸ and humans.⁹ Moreover, the reports describing cytosine as the toxic principle cite neither present nor past scientific studies in support.^{10,34} In the absence of other data, we ascribe the attributions of toxicity to cytosine to be based on possibly similar *in-vivo* pharmacological effects and physical properties.¹⁰

The presence of **1** and other isoguanine/xanthine alkaloids is certainly more consistent with a coffee-like beverage, given their structural similarity to caffeine (**Figure 2**). The structure of **1** supports the likely generation of **2** and/or paraxanthine from hydrolysis in such a preparation or following consumption (**Scheme 1**).³⁵

In retrospect, it is somewhat surprising that we obtained **1** intact from the acid-base workup in our isolation, given the ease of hydrolysis of enamines. We suspect that the conjugation of the dienamine side chain with the purine affords some measure of stability over the short period of the extraction and we have noted that hydrolysis in aqueous HCl is significantly slower than in nonaqueous TFA/ CDCl_3 . It may also be that the hydrolysis products are not easily extracted from the aqueous phase, such that **1** is obtained relatively free of other alkaloids. This would give the appearance that **1** is the major alkaloid, though we have not thoroughly examined this possibility.³⁵



Scheme 1. Dioicene hydrolysis

These observations provide new insights on the ethnopharmacological use of the Kentucky coffeetree and the pharmacological potential of **1**. Dioicene was found in high levels in the leaves and seeds of the coffeetree though it was likely only inefficiently extracted from the notoriously hard seed shells. Although we did not analyze roasted seeds for our experiments, one would anticipate that such treatment might facilitate extraction by degrading the seed coat while facilitating the hydrolysis of **1** *in situ* to **2**

and/or paraxanthine in the process. Likewise, its presence in the leaves at high levels would suggest that a tea³⁴ prepared from it could also have been used, though only one report notes such use, but not as a stimulating beverage.^{10a} It is uncertain whether **1** or its hydrolysis products are the primary pharmacological agents.³⁵ The dienamine side chain renders **1** substantially more lipophilic than paraxanthine, which could make it more easily absorbed.^{22,36} It is conceivable that **1** could act as a prodrug, being hydrolyzed in the plasma or other physiological compartment. Of the xanthine series, caffeine is much more lipophilic than the other methylxanthines. Paraxanthine is the major metabolite (>80%) of caffeine in humans and has a similar pharmacological profile.³³ Isoguanine **2** is bioisosteric with paraxanthine and may also have similar behavior. While the toxicology of paraxanthine has been much less studied in animals, it is possible that **1**, **2** or paraxanthine may be responsible for the reported toxicity.⁸ The animal toxicity of chocolate, for example is attributed to the decreased rate of metabolism and clearance of theobromine and caffeine.³⁷

In conclusion, the prior reports of the presence of cytisine in the Kentucky coffeetree, *Gymnocladus dioicus*, have not been confirmed by our GC-MS analysis. Instead, the tree has been found to contain a novel prenylated purine alkaloid, dioicine (**1**). Dioicine is related to the caffeine and its metabolite paraxanthine and may be responsible for the historical reports from which the tree derives its name. We are currently investigating the pharmacology and toxicology of **1** and its hydrolysis products. These results will be reported in due course.

EXPERIMENTAL

General Procedures. All solvents were reagent grade and used as received. Thin layer chromatography was performed using silica gel plates (EM Science 5715), visualized with UV light, and stained with KMnO₄, iodine, or phosphomolybdic acid. Melting points are uncorrected. NMR spectra were obtained on a Bruker Avance II 400 MHz spectrometer capable of pulsed field gradients on the z-axis. Chemical shifts are reported in ppm relative to internal TMS. Coupling constants (*J*) are reported in Hz and assignments of carbon multiplicity were determined on the basis of DEPT-edited experiments. Quaternary carbons were assigned on the basis of HMBC experiments. EI and CI-NH₃ mass spectra were obtained using a Thermo-Finnigan GCQ gas chromatograph-mass spectrometer or Hewlett-Packard 5890 gas chromatograph interfaced to a Hewlett-Packard 5971 quadrupole mass spectrometer. Each maintained a nominal ionization potential of 70 eV. GC-MS retention times are reported in minutes using splitless injection (1 μL, hold time 1 min) onto an RTX-5MS column (0.25 mm x 30 m) with a linear velocity of 40 cm/sec He and an oven program of 40 °C (hold 1 min), 10 °C/min to 280 °C (hold 10 min). Vapor Phase IR spectra were recorded by GC-FTIR using a Hewlett-Packard 5965B infrared detector. APCI-mass spectra were obtained by LC-MS using an Agilent 1100 binary

pump HPLC system equipped with a diode array detector and connected to a Thermo-Finnigan LCQ mass spectrometer. HPLC-MS retention times are reported in minutes on a Phenomenex Aqua C-18 column (125 Å, 4.6 mm x 250 mm, 5 µm particle size) using a gradient of 90% H₂O/MeCN to 50% over 40 min. High resolution mass spectra were collected in EI mode on a Waters-Micromass QToF spectrometer. Condensed phase IR spectra were obtained on a Midac M2000 FT instrument.

Collection. Specimens of *G. dioicus* were collected at three separate sites in the Terre Haute, IN area. Collection site 1 was located on the campus of St. Mary-of-the-Woods college in St. Mary-of-the-Woods, IN, west of Terre Haute at 39°30.816'N, 87°27.407'W. Site 2 was located at Deming Park in Terre Haute at 39°21.590'N, 87°27.905'W.²⁹ Site 3 was located near the entrance to Prairie Creek Nature Preserve in Sullivan, IN, south of Terre Haute, at 39°17.303'N, 87°27.902'W. At each site, a representative branch was pruned from the tree and separated into leaves, green stems, bark and seed pods (where available). As *G. dioicus* is dioicous, samples were taken from both male and female trees, though unambiguous identification could not be made for the Prairie Creek specimen, as it was a younger tree and may not yet have been capable of producing seed pods. Comparison samples of *Robinia pseudoacacia*, *R. hispida*, and *Gleditsia triacanthos* were collected from three locations on the campus of Indiana State University and Little Bluestem Nature Preserve.³⁸ Two separate cultivars of *G. triacanthos* were collected. The samples taken from ISU were of a thornless ornamental cultivar and those from Little Bluestem were of the normal thorned variety. Comparison samples of *Robinia pseudoacacia*, were collected from Little Bluestem Nature Preserve north of Terre Haute at 39°32.393'N, 87°23.950'W. A comparison sample of *Robinia hispida*, was collected near a residence in Terre Haute. Voucher specimens of each have been deposited at the Indiana State University Herbarium in the department of Biology.

Extraction and isolation of dioicine from leaves. A sample of *G. dioicus* leaves from Site 3 (365 g) was reduced to a coarse powder by placing in a household blender while maintaining the leaf material frozen with liquid N₂. The leaves were blended using a high speed setting and the top of the blender was covered with a standard coffee filter to allow N₂ gas to escape while retaining the plant material. Leaf material was added in portions with additional liquid N₂ as needed to maintain fluidity of the mixture. When complete, the coarse powder was transferred to a 1 L Erlenmeyer flask and the remaining N₂ allowed to evaporate. When evaporation was complete, MeOH (500 mL) was added and the macerated mixture stored for one month in a refrigerator at 4 °C. The crude extract was filtered through a cheesecloth and filter paper using a Büchner funnel and transferred to a separatory funnel. The solids were washed with additional MeOH (8 x 25 mL). The combined extract was diluted with water (700

mL) and extracted with CH₂Cl₂ (4 x 150 mL). The organic extract was subsequently extracted with 8 x 75 mL of 1N HCl. The combined HCl extracts were treated with concentrated NH₄OH to ~pH 14, saturated with NaCl, and extracted with CH₂Cl₂ (4 x 150 mL). The combined organic extracts were dried over Na₂SO₄ and the solvent removed by rotary evaporation to produce a green solid extract (157 mg, 0.043% wet weight), which consisted of >80% dioicine by NMR spectroscopy. A portion (60 mg) of the crude extract was sublimed under vacuum (145 °C at 0.15 mm Hg) to produce a white solid (46 mg) of virtually pure dioicine (mp. 123-131 °C). While pure dioicine could be obtained by flash chromatography on silica gel with a gradient of CHCl₃/MeOH (49:1 – 4:1), problems with compound stability and irreversible adsorption and/or decomposition on silica made the purified yields very low in several cases. A homogeneous sample was obtained by preparative thin-layer chromatography on silica using CHCl₃/MeOH (19:1) as the developing solvent.

Examination of plant parts and related species. Samples of plant parts of *G. dioicus* (~1 g) were cut into small pieces and ground under MeOH (3 mL). Similarly, samples of leaves from the related species *Robinia pseudoacacia*, *R. hispida*, and *Gleditsia triacanthos* (1 g each) were prepared. The mixtures were allowed to stand at rt for 1-4 days and analyzed by GC-MS. Extracted ion chromatograms using *m/z* 245 were used to detect dioicine. Dioicine-containing samples also indicated an isomer of dioicine (likely the *Z*-dieneamine),²⁸ which eluted at 17.9 min, whereas dioicine itself eluted at 20.1 min. Integration of the areas was used to estimate the relative amounts of each isomer as well as the total amount of dioicine. Background subtracted spectra for the peaks were used as confirmation of identity of each peak. Analysis of the related species indicated that none contained dioicine except *G. dioicus* and that the relative amounts of dioicine were as described in the text. Because of variations in the homogeneity, particle size and extraction efficiency from the leaf, bark, and especially seed shells, the results were only semiquantitative. Extracts from the leaves and seed shells contained the highest amounts of dioicine.³⁵ Leaf stems also contained dioicine, but in much lower amounts, while neither young nor older bark contained detectable dioicine. Mature seeded or sterile pods (outer husks) collected in the autumn contained very low amounts of dioicine.²⁶ Mucilage collected from these pods contained little or no dioicine. However, immature seed pods collected in midsummer at site 2 contained high amounts, approaching that of the leaves, though a direct comparison was not possible with this collection.²⁹ A 48 gram sample of pods was extracted as above, affording 65 mg (0.13%) of total bases which consisted of nearly pure dioicine by GC-MS and NMR.

1,6-Dihydro-6-imino-1,7-dimethyl-3-((*E*)-3-methylbuta-1,3-dienyl)-3*H*-purin-2(7*H*)-one (dioicine, 1).

TLC R_f 0.35 (19:1 CHCl₃/MeOH); ¹H NMR (400 MHz, CDCl₃) δ 7.49 (d, *J* = 14.9, 1H, H2'), 7.43 (s, 1H, H8), 7.15 (dd, *J* = 14.8, 0.5, 1H, H1'), 5.1 (d, *J* = 0.5, 1H, H4a), 5.0 (s, 1H, H4b), 4.03 (s, 3H, Me-N7), 3.42 (s, 3H, Me-N3), 1.95 (s, 3H, Me-C3'); ¹³C NMR (100 MHz, CDCl₃) δ 149.8 (C, C6), 148.5

(C, C2), 143.3 (C, C4), 140.6 (C, C3'), 139.5 (CH, C8), 124.2 (CH, C2'), 122.0 (CH, C1'), 116.9 (CH₂, C4'), 108.0 (C, C5), 34.3 (Me-N7), 28.5 (Me-N3), 18.5 (Me-C3'); ¹⁵N NMR (100 MHz, CDCl₃) δ 235 (N9), 150 (N7), 134 (N1), 131 (N3); ¹H NMR (400 MHz, CD₃OD) δ 7.73 (s, 1H, H8), 7.42 (d, *J* = 14.6, 1H, H2'), 7.18 (d, *J* = 14.6, 1H, H1'), 5.04 (m, *J* = 1.5, 1H, H4a), 5.01 (m, 1H, H4b), 4.01 (s, 3H, Me-N7), 3.40 (s, 3H, Me-N3), 1.94 (s, 3H, Me-C3'); ¹³C NMR (100 MHz, CD₃OD) δ 150.1 (C, C6), 151.3 (C, C2), 144.4 (C, C4), 141.7 (C, C3'), 141.9 (CH, C8), 125.8 (CH, C2'), 123.1 (CH, C1'), 117.2 (CH₂, C4'), 108.8 (C, C5), 34.6 (Me-N7), 29.5 (Me-N3), 18.5 (Me-C3'); ¹H NMR (400 MHz, DMSO-*d*₆) δ 7.89 (s, 1H, H8), 7.53 (d, *J* = 14.9, 1H, H2'), 7.17 (d, *J* = 14.9, 1H, H1'), 5.01 (m, *J* = 1.8, 1H, H4a), 4.98 (m, 1H, H4b), 3.96 (s, 3H, Me-N7), 3.30 (s, 3H, Me-N3), 1.87 (s, 3H, Me-C3'); ¹³C NMR (100 MHz, DMSO-*d*₆) δ 149.4 (C, C6), 149.4 (C, C2), 141.7 (C, C4), 140.6 (C, C3'), 140.3 (CH, C8), 122.5 (CH, C2'), 121.3 (CH, C1'), 116.0 (CH₂, C4'), 107.2 (C, C5), 33.8 (Me-N7), 28.7 (Me-N3), 18.1 (Me-C3'); ¹H NMR (400 MHz, D₂O) δ 7.50 (s, 1H, H8), 6.71 (d, *J* = 14.6, 1H, H2'), 6.41 (d, *J* = 14.6, 1H, H1'), 4.97 (m, *J* = 1.3, 1H, H4'a), 4.92 (m, 1H, H4'b), 3.74 (s, 3H, Me-N7), 3.10 (s, 3H, Me-N3), 1.72 (s, 3H, Me-C3'); ¹³C NMR (100 MHz, D₂O) δ 150.3 (C, C6), 150.2 (C, C2), 142.1 (C, C4), 141.0 (CH, C8), 140.1 (C, C3'), 127.6 (CH, C2'), 120.7 (CH, C1'), 118.2 (CH₂, C4'), 106.7 (C, C5), 34.0 (Me-N7), 29.3 (Me-N3), 17.3 (Me-C3'); UV λ_{max} (MeOH, nm) 205, 290; HPLC-UV λ_{max} (MeCN/H₂O, nm) 214, 288; Condensed-IR (KBr film) ν (cm⁻¹) 3332, 3111, 3073, 2921, 2851, 1663, 1634, 1536, 1465, 1423, 1291, 1251, 1216, 1172, 955, 735; Vapor-IR ν (cm⁻¹) 3380, 3089, 2956, 1710, 1657, 1579, 1531, 1463, 1416, 1357, 1286, 1224, 1164, 1043, 954, 883, 782.; GC-MS T_R 20.16 min. EIMS *m/z* 246.1332 (M+H, 13, calc. for C₁₂H₁₆N₅O 246.1355), 245.1284 (M+, 100, calc. for C₁₂H₁₅N₅O 245.1277), 230.1060 (M-Me, 20, calc. for C₁₁H₁₂N₅O 230.1042), 228 (20) 216.1220 (M-CHO, 10, calc. for C₁₁H₁₄N₅ 216.1249), 187.1003 (M - H - MeN=C=O, 11, calc. for C₁₀H₁₁N₄ 187.0984), 173.0842 (M - H - MeN=C=O, 33, calc. for C₉H₉N₄ 187.0984); CIMS (NH₃) *m/z* 246, MS/MS 229 (41), 189 (100), 180 (65) 148 (39); APCI-MS 246, (1 exchangeable proton with D₂O).

Hydrolysis experiments - 1,7-dimethylisoguanine (2). *Aqueous HCl* - A small sample (~5-7 mg) of sublimed dioicine in CDCl₃ (~200 μL) was placed in a microvial sealed with a Teflon-lined cap. To the solution was added 0.2 mL of 1 N HCl and the solution was stirred overnight at rt. The organic portion was separated and the aqueous portion extracted a second time with CH₂Cl₂. The combined organic extracts were dried over Na₂SO₄ and analyzed by GC-MS and found to contain prenal, identical in all respects with a commercial sample (Aldrich). The aqueous portion was neutralized with saturated aq. NaHCO₃ until bubbling ceased and then 1N NH₄OH was added to pH 9 (~ 3 drops). The solution was then extracted twice with CH₂Cl₂ (~200 μL) and the combined organic layers dried and analyzed by GC-MS as described above. The sample was found to contain unhydrolyzed **1** and small amounts of **2**

as described below. A similar experiment in ~2M DCl/D₂O produced a mixture of prenal and **2** as the DCl salt. ¹H NMR (400 MHz, TFA/CDCl₃) δ 8.05 (d, *J* = 0.6 Hz, 1H, H8), 3.95 (d, *J* = 0.6 Hz, 3H, N9Me), 3.43 (s, 3H, N1Me); ¹³C NMR (100 MHz, TFA/CDCl₃) δ 147.4 (C8), 103.6 (C5), 34.4 (N9Me), 30.5 (N1Me), quaternary carbons not identified. The mixture was evaporated to dryness and reconstituted in DMSO-*d*₆ to remove prenal leaving **2** as the DCl salt; ¹H NMR (400 MHz, DMSO-*d*₆) δ 9.27 (br s, 1H, NH), 8.09 (s, 1H, H8), 4.06 (s, 3H, N9Me), 3.58 (s, 3H, N1Me); ¹³C NMR (100 MHz, DMSO-*d*₆) δ 150.1 (C6), 149.8 (C2), 148.1 (C4), 147.7 (C8), 102.8 (C5), 34.5 (N9Me), 31.0 (N1Me)]. A second HCl hydrolysis experiment was conducted with the addition of sodium nitrite and allowed to stand overnight.²¹ The hydrolysate was neutralized and saturated with solid NaHCO₃ and extracted with EtOAc. Evaporation and analysis of the extract by GC-MS afforded mostly **2**, with a weak peak matching in retention time and mass spectrum with a commercial sample of paraxanthine (Aldrich).

Trifluoroacetic acid - To a sample of dioicine in CH₂Cl₂ prepared as described above was added 20 μL of CF₃CO₂H and the solution stirred overnight at rt. The solvent was evaporated under N₂ flow, the residue rubbed with CH₂Cl₂, and analyzed by GC-MS. A similar experiment in 50% CDCl₃/CF₃CO₂H produced an equimolar mixture of prenal and **2** as the trifluoroacetate; ¹H NMR (400 MHz, TFA/CDCl₃) δ 8.20 (s, 1H, H8), 4.22 (s, 3H, N9Me), 3.73 (s, 3H, N1Me); ¹³C NMR (100 MHz, TFA/CDCl₃) δ 150.3 (C2), 148.8 (C6), 148.1 (C4), 146.3 (C8), 103.6 (C5), 35.7 (N9Me), 31.2 (N1Me); . The mixture was evaporated to dryness and reconstituted in CD₃OD to remove prenal leaving **2** as the trifluoroacetate salt; ¹H NMR (400 MHz, CD₃OD) δ 8.09 (s, 1H, H8), 4.06 (s, 3H, N9Me), 3.58 (s, 3H, N1Me); ¹³C NMR (100 MHz, CD₃OD) δ 152.2 (C2), 152.1 (C6), 149.6 (C4), 149.3 (C8), 104.4 (C5), 34.9 (N7Me), 31.2 (N1Me)]. A sample was neutralized with aqueous ammonia, evaporated and purified by silica gel chromatography to give the free base; mp. 233-234.5 °C (decomp - sealed tube, lit., mp. 310-312 °C); R_f 0.35 (MeOH); ¹H NMR (400 MHz, CD₃OD) δ 7.77 (s, 1H, H8), 3.98 (s, 3H, N9Me), 3.44 (s, 3H, N1Me); ¹³C NMR (100 MHz, CD₃OD) δ 154.6 (C6/C2), 150.9 (C2/C6), 149.1 (C4), 145.3 (C8), 106.1 (C5), 34.3 (N7Me), 29.5 (N1Me)]; ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.0 (br s, 1H, NH), 7.78 (s, 1H, H8), 3.95 (s, 3H, N9Me), 3.28 (s, 3H, N1Me); ¹³C NMR (100 MHz, CD₃OD) δ 150.8 (C6/C2), 148.5 (C2/C6), (C4), 141.2 (C8), 105.6 (C5), 33.3 (N7Me), 27.5 (N1Me)]; Vapor-IR ν (cm⁻¹) 3476, 1734, 1656, 1552, 1472, 1269; UV λ_{max}(MeCN/H₂O, nm) 219, 235sh, 288 (9:6:10); EIMS 179.0807 (M⁺, 32, calc. for C₇H₉N₅O, 179.0807), 178.0738 (100, calc. for C₇H₈N₅O, 179.0729), 161.0496 (20, calc. for C₇H₅N₄O, 161.0463), 150.0783 (M-CHO, 18, calc. for C₆H₈N₅, 150.0780), 133.0491 (10, calc. for C₆H₅N₄, 133.0515), 122.0586 (9, calc. for C₅H₆N₄, 122.0594), 107.0440 (6, calc. for C₅H₅N₃, 107.0484), 95.0604 (8, calc. for C₅H₇N₂, 95.0610), 73.0481 (12, calc. for C₂H₅N₂O, 73.0402), 68.0384 (10, calc. for

C₄H₆N, 68.0501), 67.0308 (8, calc. for C₄H₅N, 67.0422), 53 (18). APCI-MS (MeCN/H₂O, *m/z*) 180 (M+H, 100), 163 (18), 149 (9), 123 (18).

Hot water. To a sample of **1** (~1 mg) was added deionized water (0.5 mL), and the solution heated with a water bath to 90 °C for six h with stirring, over which time most of the water evaporated. The residue was reconstituted in CH₂Cl₂, dried over Na₂SO₄, and analyzed by GC-MS, which indicated mostly starting **1** with ~2-3 % of **2** produced.

Pharmacology Nicotinic receptor pharmacology was determined in binding assays using [³H]-epibatidine in membranes of rat cerebral cortex³⁹ and functional fluorescence assays in HEK cells⁴⁰ transfected with nicotinic receptor α3 and β4 subunits.⁴¹ Both assays were performed essentially as described. Neither assay showed any significant activity for dioicine or the alkaloid fraction of *G. dioicus*.

ACKNOWLEDGEMENTS

The authors deeply mourn the passing of their friend and colleague, John Daly. RWF wishes to thank Dr. Marion T. Jackson, Dr. Peter E. Scott and Mr. Ian MacDonald for advice and assistance with taxonomy, collection, and preparation of herbarium specimens. Mr. Ryand J. Tucker² is also thanked for assistance with preparation of samples. A generous gift of HEK cells expressing various subunit combinations for nicotinic receptors^{41,42} by Drs. Kenneth J. Kellar and Yingxian Xiao of Georgetown University Medical School is appreciated. This work was performed with funds from Indiana State University including the Department of Chemistry and College of Arts and Sciences (startup funds for RWF), Office of the Provost (Promising Scholar Award to RWF), and Center for Public Service and Community Engagement (Faculty Fellow Award to RWF, Undergraduate Fellow Award to RRC), University Research Committee (UNR246), and the Lilly Endowment (CPSCE/Office of the Provost). A grant from the National Science Foundation (CHE0521075) for a 400 MHz NMR spectrometer at Indiana State University is also gratefully acknowledged. Research at NIH was funded by the intramural program of NIDDK.

REFERENCES AND NOTES

1. Presented at the 48th annual meeting of the American Society of Pharmacognosy, Portland, ME, July 14-18, 2007 (Poster P117).
2. Undergraduate student.
3. Deceased March 5, 2008.
4. S. P. Arneric, M. Holladay, and M. Williams, *Biochem. Pharmacol.*, 2007, **74**, 1092, and references cited therein.

5. T. Schmeller, M. Sauerwein, F. Sporer, M. Wink, and W. E. Muller, *J. Nat. Prod.*, 1994, **57**, 1316.
6. A. D. Kinghorn and M. F. Balandrin, in 'Alkaloids: Chemical and Biological Perspectives,' Vol. 2, ed. by S. W. Pelletier, Wiley and Sons, New York, 1984, pp. 105-148.
7. a) H. L. Keeler, 'Our Native Trees and How to Identify Them,' Charles Scriber's Sons, New York, 1900, pp.109-112. b) A. J. Downing, 'A Treatise on the Theory and Practice of Landscape Gardening, Adapted to North America; with a View to the Improvement of Country Residences,' A. O. Moore, New York, 1859, pp. 204-206.
8. a) R. Evers and R. Link, in 'Poisonous Plants of the Midwest and Their Effects on Livestock,' University of Illinois Press, Urbana, IL, 1972, pp. 57-59. b) K. Lampe and M. McCann, in 'AMA Handbook of Poisonous and Injurious Plants,' American Medical Association, Chicago, IL, 1985, pp. 86-87. c) G. Reynard and J. B. S. Norton, 'Poisonous Plants of Maryland In Relation to Livestock' Bulletin A 10, The University of Maryland Agricultural Experiment Station, College Park, MD, 1942, pp. 266-268. d) S. R. Hill, '100 Poisonous Plants of Maryland,' Bulletin 314, The University of Maryland Cooperative Extension Service, College Park, MD, 1985, pp. 20-21.
9. a) J. M. Kingsbury, 'Poisonous Plants of the United States and Canada,' Prentice-Hall, Englewood Cliffs, NJ, 1964, pp. 322-323. b) W. H. Lewis and P. F. Elvin-Lewis, 'Medical Botany: Plants Affecting Man's Health,' Wiley and Sons, New York, 1977, p. 43. c) J. W. Hardin and J. M. Arena, 'Human Poisoning from Native and Cultivated Plants,' Duke University Press, Durham, NC, 1974, pp. 84-85. d) V. K. Chestnut, 'Principal Poisonous Plants of the United States,' Bulletin 20, United States Department of Agriculture, 1898, p. 28.
10. a) C. F. Millspaugh, 'Medicinal Plants.', Yorston and Co., Philadelphia, PA, 1892, Vol. 1, plate 53 (3 pp). This work has been reproduced several times, most recently in 1974 by Dover Publications, New York. b) R. D. Dunglinson, 'Medical Lexicon: A Dictionary of Medical Science,' Blanchard and Lea, Philadelphia, 1860, p. 438. This is the earliest reference we identified which describes the tree (*Gymnocladus canadensis*, syn.) as containing cytisine in the leaves, while subsequent references list the seeds, fruit pulp and/or leaves.
11. A number of other compounds have been described including lipids, amino acids, and carbohydrates. A quite thorough compilation is found in 'Phytochemical Dictionary of the Leguminosae,' Vol. 1, ed. by F. A. Bisby, J. Buckingham, and J. B. Harborne, Chapman & Hall, London, UK, 1994, p. 348, and references cited therein.
12. A. F. Cook, R. T. Bartlett, R. P. Gregson, and R. J. Quinn, *J. Org. Chem.*, 1980, **45**, 4020.
13. S. S. Mitchell, A. B. Whitehill, H. G. Trapido-Rosenthal, and C. M Ireland, *J. Nat. Prod.*, 1997, **60**, 727.
14. a) F. Cafieri, E. Fattorusso, A. Mangoni, and O. Taglialatela-Scafati, *Tetrahedron Lett.*, 1995, **36**,

7893. b) X. Fu, P. L. Ng, F. J. Schmitz, M. B. Hossain, D. van der Helm, and M. Kelly-Borges, *J. Nat. Prod.*, 1996, **59**, 1104.
15. T. Itaya and T. Harada, *Chem. Pharm. Bull.*, 1990, **38**, 2971.
16. B. R. Copp, C. M. Wassvik, G. Lambert, and M. J. Page, *J. Nat. Prod.*, 2000, **63**, 1168.
17. D. Tazdemir, G. C. Mangalindan, G. P. Concepción, M. K. Harper, and C. M. Ireland, *Chem. Pharm. Bull.*, 2001, **49**, 1628.
18. N. B. Perry, J. W. Blunt, and M. H. G. Munro, *J. Nat. Prod.*, 1987, **50**, 307.
19. M. G. Harris and R. Stewart, *Can. J. Chem.*, 1977, **55**, 3807.
20. Z. L. Zhang, S. X. Zhou, X. W. Wang, H. T. Wang, Y. L. Chen, and J. Y. Liu, *Chem. Res. Chin. Univ.*, 2006, **22**, 451.
21. K. Yamauchi, T. Tanabe, and M. Kinoshita, *J. Org. Chem.*, 1976, **23**, 3691.
22. It should be noted that the less than fully substituted xanthenes and guanines/isoguanines have very poor solubility in most solvents other than DMSO or water (less at neutral pH). They are especially difficult to analyze by GC and HPLC giving broad GC peaks in many cases and sometimes precipitating on HPLC columns.
23. K. Grozinger, K. R. Freter, P. Farina, and A. Gladczyk, *Eur. J. Med. Chem. Chim. Ther.*, 1983, **18**, 221.
24. J. M. Rice and G. O. Dudek, *J. Am. Chem. Soc.*, 1967, **89**, 2719.
25. Seasonal variation in alkaloid content has been observed with other alkaloid bearing plants. For example, see M. H. Ralphs, D. R. Gardner, and J. A. Pfister, *J. Chem. Ecol.*, 2000, **26**, 1595.
26. The pods collected at site 1 in winter 2004 were small and almost exclusively sterile. Legumes and many other plants are known to abort seed production in response to water or nutrient stress as well as genetic factors. See A. G. Stephenson, *Ann. Rev. Ecol. Syst.*, 1981, **12**, 253.
27. Variation with age and elevation has been observed in other trees. For example, see D. McC. Newbery and H. de Foresta, *Biotropica*, 1985, **17**, 238.
28. While the *Z*-isomer was not isolated in pure form, it has an identical EI mass spectrum to dioicine. The isolated dioicine is determined to be the *E*-isomer based on the large ($J = 14$ Hz) coupling constant between the H1' and H2'.
29. Unfortunately, this tree was later damaged in a storm and removed from the site. The immature pods collected from this tree were collected several days after the tree was felled, precluding comparison with the leaves, which had already deteriorated. The degradation of the leaves allowed more sun exposure to the pods, which we believe contributed to the unexpectedly high *Z*/*E* ratio.
30. F. J. Turner, *Am. Hist. Rev.*, 1898, **3**, 672.
31. T. Jefferson, 'The Jeffersonian Cyclopaedia,' ed. by J. P. Foley, Funk and Wagnalls, New York, 1900,

- p. 45. Since the tree is desirable as an ornamental⁷ and as Jefferson's horticultural interest is well-documented, the request for seeds may have had little to do with the beverage.
32. A. Gramley and R. Goulding, *Brit. Med. J.*, 1981, **223**, 1220.
 33. a) J. W. Daly and B. B. Fredholm, in 'Coffee, Tea, Chocolate, and the Brain,' ed. by A. Nehlig, CRC Press, Boca Raton, FL, 2004, pp. 1-11. b) J. W. Daly and B. B. Fredholm, *Drug Alcohol Depend.*, 1998, **51**, 199.
 34. It is interesting to note that Chestnut^{9d} describes coffeetree poisoning, yet performed no apparent study on the alkaloids from this tree. However, he does perform analyses of *Cassina* (*Ilex vomicus*, Aquafoliaceae), the source of caffeine in a tea-like beverage called the black drink. See F. B. Power and V. K. Chestnut, *J. Am. Chem. Soc.*, 1919, **41**, 1307. See also: C. M. Hudson, 'Black Drink: A Native American Tea.' University of Georgia Press, Athens, GA, 1979.
 35. A number of other as yet incompletely characterized methylpurines are present in the extracts. Further compounds will be described as their structures are determined and a complete profile assembled.
 36. D. A. Smith, B. C. Jones, and D. K. Walker, *Med. Res. Rev.*, 1996, **16**, 243.
 37. G. E. Miller, L. L. Radulovic, R. H. Dewit, M. J. Brabec, S. M. Tarka, and H. H. Cornish, *Drug Metab. Disposition*, 1984, **12**, 154.
 38. a) R. A. Strait and M. T. Jackson, *Proc. Indiana Acad. Sci.*, 1986, **95**, 447. b) D. Dubis, R. A. Strait, M. T. Jackson, and J. O. Whitaker, Jr., *Nat. Areas J.*, 1988, **8**, 267.
 39. R. A. Houghtling, M. I. Davila-Garcia, and K. J. Kellar, *Mol. Pharmacol.*, 1995, **48**, 280.
 40. R. W. Fitch, Y. Xiao, K. J. Kellar, and J. W. Daly, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 4909.
 41. Y. Xiao, E. L. Meyer, J. M. Thompson, A. Surin, J. Wroblewski, and K. J. Kellar, *Mol. Pharmacol.*, 1998, **54**, 322.
 42. Y. Xiao and K. J. Kellar, *J. Pharmacol. Exp. Ther.*, 2004, **310**, 98.