Pergamon

Tetrahedron Vol. 51, No. 45, pp. 12203-12210, 1995

0040-4020/95 $9.50+0.00

0040-4020(95)00794-6

Characterization of Anti-HIV Lignans from *Larrea tridentata*

John Gnabre,[a] Ru Chih C. Huang,[a]* Robert B. Bates,[b] Jennifer J. Burns,[b] Sriyani Caldera,[b] Mark E. Malcomson[b] and Kelly J. McClure[b]

[a]Department of Biology, Johns Hopkins University, Baltimore, MD 21218-2685

[b]Department of Chemistry, University of Arizona, Tucson, AZ 85721

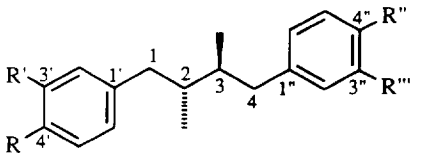
Abstract: Fractions from *Larrea tridentata* with anti-HIV-1 activity (specifically, inhibition of HIV Tat transactivation) were analyzed by GC/MS and NMR and found to contain lignans **1a-i** and **2a-d**. Assay-guided purification by countercurrent chromatography established **1g** (mal.4) to be especially active. Compounds **1b-f,h,i** and **2d** are new.

Extracts of the leaves of the creosote bush, *Larrea tridentata* (DC.) Cov. (Zygophyllaceae), with chloroform:methanol were found to possess anti-HIV-1 activity, especially inhibition of HIV Tat transactivation.[1] Assay-guided separation by countercurrent chromatography (CCC) of the active extracts gave a series of mostly new lignans whose characterization we report herein.

Previous investigations of *L. tridentata* resulted in the isolation of the lignans **1a** (nordihydroguaiaretic acid, NDGA),[2] **1g** (first called "partially-demethylated dihydroguaiaretic acid",[3] more recently mal.4[4]), **1j** (dihydroguaiaretic acid),[2] **2a** (3'-demethoxyisoguaiacin),[3] **2b** (norisoguaiacin),[3] **2c** (3'-demethoxy-6-*O*-demethylisoguaiacin),[4] and **2e** (6,3'-di-*O*-demethylisoguaiacin).[4] Some of these and other lignans have been

identified in other *Larrea* species.[5]

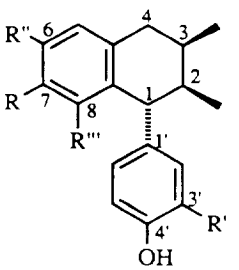
We wish to report the characterization of lignans **1a-i**, **2a-d** during the examination of the fractions which inhibited Tat transactivation. The most active compound, mal.4 (**1g**), suppresses HIV basal transcription, HIV Tat-dependent and Tat-independent transactivation; protects human lymphoblastoid CEM-SS cells against HIV-1 infection; and inhibits HIV replication in mitogen-stimulated peripheral blood mononuclear cells from acute AIDS patients.[6] We also isolated the flavone ayanin (quercetin 3,7,4′-trimethyl ether, **3**), devoid of anti-Tat transactivation activity, found in several plants[7,8] but apparently not before in *Larrea* species.



1

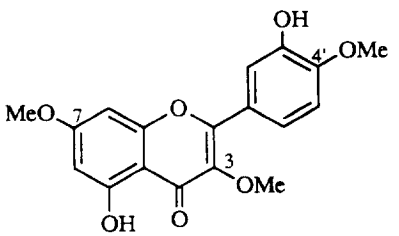
			R	R	R"	R"
a	NDGA		OH	OH	OH	OH
b	G_1		OMe	OMe	OH	OMe
c	G_2		OMe	OMe	OMe	OH
d	G_3		OAc*	OMe*	OH	OMe
e	G_4		OAc*	OMe*	OMe	OH
f	L_1		OH	OH	OH	H
g	L_2	(mal.4)	OH	OMe	OH	OH
h	L_3		OAc*	OMe*	OH	OH
i	L_4		OMe	OH	OH	OH
j			OH	OMe	OH	OMe
k			OH	OMe	OMe	OH

*R and R′ may be reversed



2

		R	R	R"	R"
a		OMe	H	OH	H
b		OH	OMe	OH	H
c		OH	H	OH	H
d		OH	H	H	OH
e		OH	OH	OH	H



3

Active fractions *Lo* and *Gr* (less polar) from a pilot CCC study using hexane:ethyl acetate:methanol:0.5%NaCl (6:4:5:5 and later 7:3:5:5) were selected for further fractionation.[1] Contrary to earlier findings,[3] the lignans in these fractions could be analyzed by GC/MS without derivatization. This analysis showed *Lo* to be a mixture of lignans **1f** (L_1, 6%), **1g** (L_2, 75%), **1h** (L_3, 10%) and **1i** (L_4, 9%). Their GC retention times and main mass spectral peaks are given in Table I. The most characteristic MS peaks other than the molecular ion peaks are from cleavage of the benzylic bonds, and show the distribution of mass on each aromatic ring.

Table I. GC Retention Times and MS Peaks for **1b-i** and **2a-b.**

	GC ret. time, min.	M/z of MS Ions (% of Base in Parentheses) Mol. Ion	Benzyl Ions	Other Ions
1b (G_1)[a]	14.92	344(51)	151(100),137(60)	
1c (G_2)	15.19	344(40)	151(100),137(48)	
1d (G_3)	15.87	372(18)	137(100)	330(51),43(9)
1e (G_4)	16.23	372(21)	137(100)	330(52),43(13)
1f (L_1)	15.26	286(57)	123(9?),107(10)	
1g (L_2)	16.10	316(37)	137(100),123(37)	
1h (L_3)	16.17	358(25)	179(8),137(100), 123(26)	316(58),43(26)
1i (L_4)	16.30	316(42)	137(100),123(41)	
2a	17.16	298(100)		269(30),297(17)
2b	17.49	314(100)		313(15),271(15)

[a] Designation in ref. 1.

Lignan **1g** was reported previously in *L. tridentata*,[3] but may not have been separated from its isomer **1i**. In the present case, **1g** and **1i** were separated from the *Lo* mixture by further CCC.[1] Their mass spectra (Table I) are nearly identical, but their NMR parameters (Table II) show useful differences in the 2′ and 5′ proton absorptions. The protons *ortho* to OMe

(2′ for **1g**, 5′ for **1i**) absorb farther upfield than the corresponding protons in the other isomer; this is because the effect of being *meta* to OH vs OMe outweighs the effect of being *ortho* to OH vs OMe. The methylene protons on C-1 in **1i** absorb slightly farther upfield than those in **1g**. The structural assignments for **1g** vs **1i** were firmly established by difference nOe measurements in which the methoxyl protons were irradiated: In the case of **1g**, the doublet at δ6.61 (J = 1.9 Hz) was enhanced, and in the case of **1i**, a doublet at δ6.77 (J = 8.1 Hz).

Table II. ^{1}H NMR Chemical Shifts (δ) and Coupling Constants (J, in Hz, in Parentheses) for **1f**, **1g** and **1i**.

proton	**1f**	**1g**	**1i**
1	2.23dd(13.3,10.0)	2.25dd(13.4,9.3)	2.21dd(13.5,9.3)
	2.69dd(13.3,5.1)	2.71dd(13.4,4.7)	2.69dd(13.5,4.9)
2,3	1.72m	1.72m	1.72m
4	2.28dd(13.3,9.1)	2.25dd(13.5,9.3)	2.24dd(13.5,9.2)
	2.71dd(13.3,5.2)	2.68dd(13.5,5.2)	2.68dd(13.5,4.9)
5,6	0.81d(6.6)	0.82,0.83d(6.7)	0.81,0.82d(6.7)
2′	6.66d(1.9)	6.61d(1.9)	6.71d(2.1)
5′	6.77d(8.0)	6.82d(8.0)	6.77d(8.1)
6′	6.59dd(8.0,1.9)	6.64dd(8.0,1.9)	6.63dd(8.1,2.1)
2"	7.01d(8.4)	6.67d(2.0)	6.67d(2.0)
3"	6.75d(8.4)		
5"		6.77d(8.0)	6.77d(8.0)
6"		6.58dd(8.0,2.0)	6.58dd(8.0,2.0)
OMe		3.86s	3.87s
OH	4.59,4.93.5.07s	5.00,5.15,5.45s	5.08,5.37,5.60s

Compound **1f** was identified in the *Lo* mixture from its MS. That the OH on the mono-oxygenated ring was in the *para* position was indicated by the AA'BB' pattern for aryl protons at δ6.75 and 7.01. Compound **1h** showed an acetate group through its molecular ion, loss of 42 (ketene), and peak at 43 (Ac'). The relative positions of Ac and OMe in **1h** were not determined.

Fraction *Gr* was shown by GC/MS to be a complex mixture containing many fatty acid derivatives, but lignans **1b** (G_1, 23%), **1c** (G_2, 20%), **1d** (G_3, 7%) and **1e** (G_4, 9%) were identified along with what from their masses and very strong molecular ion peaks are presumably the previously characterized[3] 3'-demethoxyisoguaiacin (**2a**, 12%) and norisoguaiacin (**2b**, 6%). Which is which between **1b** and **1c** was assigned from GC retention times by analogy with **1g** and **1i**. The relative positions of OH and OMe in **1d** and **1e** were assigned on the same basis.

Concerning the stereochemistry of the NDGA derivatives **1**, Gisvold and Thaker[3] showed the mixture of lignans **1** from *L. divaricata* to afford *meso*-NDGA (shown by X-ray to be meso[9]) upon removal of all alkyls (and presumably acyls) from oxygens. On this basis, **1b-j** are presumed to also have opposite configurations at the chiral centers. The asymmetric members of the series (**1b-i**) with sufficient pure material for optical rotation measurement are chiral (**1g**, $[\alpha]_D^{25}$ +63 , c 0.19 in MeOH; **1i**, $[\alpha]_D^{25}$ +76°, c 0.17 in MeOH), but their absolute configurations are not known.

Lignans **1** can adopt various conformations, including the extended conformation found in an X-ray study of a dibromo derivative,[9] and nearly extended conformations with the largest groups about the central bond gauche calculated using molecular mechanics to be most stable.[10] The NMR evidence, however, which shows the protons on C1 to have chemical shifts different by 0.4-0.5 ppm and the upfield C1 proton to have stronger coupling to the C2 proton, requires contributions from more compact conformations in which the upfield C1 proton lies over the plane of the farther ring.

Further CCC of the more polar fraction *Lo* with chloroform:methanol:water (8:2:1-lower phase) gave pure samples of five lignans (FB_1-FB_5),[1] each containing three phenolic hydroxyls including two in a catechol unit. FB_1, FB_2 and FB_4 were readily identified as lignans **1g**, **1i** and **1f**, respectively. Unlike the other new lignans, **1f** crystallized (mp 125-127 °C).

From their NMR (Table III) and mass spectra (mol. ions 284), FB_3 and FB_5 are isomeric lignans **2d** and **2c**, respectively. Compound **2c** has been found in *L. tridentata* previously,[4] but **2d** is new.

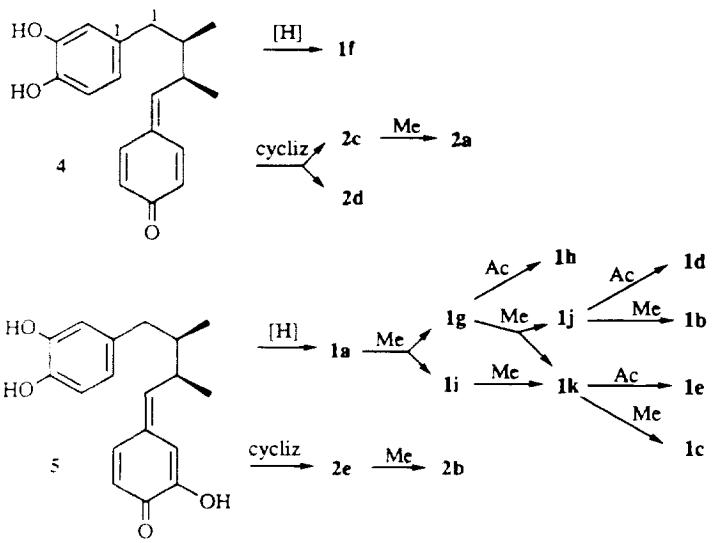
Table III. ^{1}H NMR Chemical Shifts (δ) and Coupling Constants (J, in Hz, in Parentheses) for **2a**, **2c** and **2d**.

proton	2a	2c	2d
1	3.61d(6.2)	3.59d(6.3)	3.82d(3.8)
2,3	1.90,2.02m	1.90,2.01m	1.88,1.95m
4α	2.87dd(16.2,4.7)	2.84dd(16.3,5.3)	2.78dd(16.9,4.9)
4ß	2.45dd(16.2,7.2)	2.42dd(16.3,7.2)	2.47dd(16.9,9.8)
5	6.57s	6.62s	6.79d(8.7)
6			6.67d(8.7)
8	6.39s	6.31s	
2′	6.87d(8.5)	6.89d(8.5)	6.99d(8.5)
3′	6.70d(8.5)	6.72d(8.5)	6.76d(8.5)
CMe	0.88,0.89d(6.9)	0.87,0.88d(6.9)	0.89,0.95d(6.9)
OMe	3.86s		
OH	4.78,5.35s	very broad	4.44,4.70,5.11s

The aromatic protons on the fused aromatic ring in **2d** are *ortho* to one another from the 8.7 Hz coupling between them. The choice of C7 and C8 for the hydroxyl groups is based partly on biogenetic grounds, namely that **2d** may be biosynthesized from the same precursor **4** which is reduced to **1f** and cyclized to 2c (→ **2a**), by 180° rotation about the 1-1′ bond in **4** before cyclization. Structure **2d** is supported by the relatively large changes in the chemical shifts of the protons at C1 and C2′, presumably due to the presence of a hydroxyl group at C8. Indeed, many features of the NMR spectrum of **2d** as compared to **2a** and 2c can be explained in terms of a larger proportion in **2d** of the half-chair with the bulky 1-substituent axial (to avoid the 8-hydroxyl), e.g., the downfield shifts of H1 and one of the C-methyls (the one on C3), the smaller coupling constant between H1 and H2, and the larger coupling constant between H3 and H4ß. The stereochemistry shown for **2d** is consistent with its NMR spectra,[4] proposed biosynthesis, and observed optical rotation ($[\alpha]_D^{25}$ +118°, c 0.09 in MeOH).

An X-ray study on the triacetate of **2c** put the relative configurations of this group on firm ground.[11]

A possible scheme for the biosynthesis of all of the types **1** and **2** lignans found in *L. tridentata* from the hypothetical intermediates **4** and **5** is shown below. Compound **1k**, a likely intermediate in the formation of **1c** and **1e**, has not yet been found in the plant.



The antioxidant NDGA (**1a**) and some of its derivatives have been shown before to possess numerous biological properties,[12] but not previously to be inhibitors of Tat transactivation and replication.[8] While the overall anti-HIV of NDGA (**1a**) is weak, its methyl-substituted congener mal.4 (**1g**) has very promising activity.[8] Further biological studies of these natural products and synthetic analogs are underway, as this class of compounds may spur clinical interest.

EXPERIMENTAL

General. GC/MS analyses were performed on a Hewlett-Packard system consisting of a Model 5890 gas chromatograph, a Model 5970 mass selective detector and an RTE-6 data system. The GC column was an HP-5 fused silica capillary with a 5% phenyl methylsilicone stationary phase, a film thickness of 0.33 μm, a length of 25 m and an internal diameter of 0.2 mm. The carrier gas was helium with a column head pressure of 20 psig. The GC oven temperature program was used as follows: 70°C initial, 1 min hold,

increased at 20°/min to 300°C, 6 min hold. The samples were dissolved in CH_2Cl_2 and a split injection technique was used. Direct insertion probe analyses were used with some of the lignans **2** on a Hewlett-Packard model 5988A mass spectrometer temperature programmed from 30-325°C at 30°/min. NMRs were measured in $CDCl_3$/TMS at 250 or 500 MHz on Bruker AM-250 and AM-500 spectrometers. Molecular mechanics calculations were carried out using PCMODEL 4.0, Serena Software, Bloomington, IN 47407-3076.

Ayanin. Some ^1H NMR parameters of ayanin (**3**) have been reported in DMSO-d_6[6] and acetone-d_6.[7] We measured parameters in $CDCl_3$: δ3.87s, 3.88s, 3.99s, 5.73br s, 6.36d(J = 2.2 Hz), 6.45d(J = 2.2 Hz), 6.97d(J = 8.5 Hz), 7.70d(J = 2.1 Hz), 7.73dd(J = 8.5,2.1 Hz) and in C_6D_6: δ3.06s, 3.12s, 3.71s, 5.42br s, 6.17d(J = 2.2 Hz), 6.39d(J = 8.6 Hz), 6.45d(J = 2.2 Hz), 7.69dd(J = 8.6,2.2 Hz), 7.88d(J = 2.2 Hz).

Acknowledgements - This work was supported in part by National Institutes of Health Grant AI 32301 (R.C.H.), U.S. Army Medical Research Grant DAND 17-93-C3122 (R.C.H.) and by a Dimitri V. d'Arbelloff Fellowship awarded to J.N.G. by the Millipore Foundation.

REFERENCES

1. Gnabre, J.N.; Ito, Y.; Ma, Y.; Huang, R., accepted for *J. Chromatog. A.*
2. Waller, C.W., Gisvold, O. *J. Am. Pharm. Asn., Sci. Ed.* **1945**, *34*, 78-80.
3. Gisvold, O.; Thaker, E. *J. Pharm. Sci.* **1974**, 63, 1905-1907.
4. Konno, C.; Xue, H.-Z.; Lu, Z.-Z.; Ma, B.-X.; Erdelmeier, C. A. J.; Che, C.-T.; Cordell, G. A.; Soejarto, D. D.; Waller, D. P.; Fong, H. H. S. *J. Nat. Prod.* **1989**, *52*, 1113-1117.
5. Mabry, T. J.; DiFeo, D. R. Jr.; Sakakibara, M.; Bohnstedt, C. F. Jr.; Seigler, D. in *Creosote Bush: Biology and Chemistry of Larrea in New World Deserts.* Mabry, T. J.; Hunziker, J. H.; DiFeo, D. R. Jr. (Eds). Dowden, Hutchinson & Ross, Inc., Stroudsburg, PA, **1977**, pp. 115-134.
6. Gnabre, J. N.; Brady, J. N.; Clanton, D. J.; Ito, Y.; Dittmer, J.; Bates, R. B.; Huang, R. C. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, in press.
7. Rao, K. V.; Owoyale, J. A. *J. Het. Chem.* **1976**, *13*, 1293-1295.
8. Malan, E.; Roux, D. G. *J. Chem. Soc. Pl* **1979**, 2696-2703.
9. McKechnie, J. S.; Paul, I. C. *J. Chem. Soc. B* **1969**, 699-702.
10. Sakurai, H.; Nikaido, T.; Ohmoto, T.; Ikeya, Y.; Mitsuhashi, H. *Chem. Pharm. Bull.* **1992**, *40*, 1191-1195.
11. Fronczek, F. R.; Caballero, P.; Fischer, N. H.; Fernandez, S.; Hernandez, E.; Hurtado, L. M. *J. Nat. Prod.* **1987**, *50*, 497-499.
12. MacRae, W. D.; Towers, G. H. N. *Phytochemistry* **1984**, *23*, 1207-1220.

(*Received in USA* 29 *April* 1995; *revised* 14 *September* 1995; *accepted* 15 *September* 1995)