Tannins and Related Compounds. CII. 1) Structures of Terchebulin, an Ellagitannin Having a Novel Tetraphenylcarboxylic Acid (Terchebulic Acid) Moiety, and Biogenetically Related Tannins from Terminalia chebula Retz.

Ta-chen Lin, a Gen-ichiro Nonaka, a Isuo Nishoka a, c and Feng-chi Ho b

Faculty of Pharmaceutical Sciences, Kyushu University 62, a 3-1-1 Maidashi, Higashi-ku, Fukuoka 812, Japan and Taiwan Forestry Research Institute, Heng-chun Branch, b Heng-chun, Taiwan, Republic of China. Received June 1, 1990.

A chemical examination of myrobolans (the fruits of Terminalia chebula Retz., Combretaceae) has led to the isolation and characterization of puniclagin (1), terflavin A (2) and a new ellagitannin named terchebulin (3), which possesses a novel tetraphenylcarboxylic acid (terchebulic acid) moiety. Furthermore, from the leaves of T. chebula, a series of biogenetically related hydroxyizable tannins, terflavins B (7), C (9) and D (10), punicalagin (1) and punicalin (8), have been isolated and structurally elucidated. The concomitant isolation of terflavins A (2) and B (7) provides biogenetic evidence that the terchebulic acid moiety is derived by an oxidative carbon-oxygen coupling of adjacent flavagallonic acid and gallic acid esters.

Keywords myrobolan; Terminalia chebula; Combretaceae; terchebulin; terchebulic acid; flavagallonic acid; terflavin C; terflavin D; tannin; oxidative phenol coupling

Recent chemical work has shown that one of the major metabolic patterns of gallic acid esters is oxidative carbon-carbon and/or carbon-oxgen coupling(s) of vicinal aromatic rings leading to esters of higher-molecular-weight phenolcarboxylic acids, 2) and that as well as hexahydroxydiphenic acid esters (ellagitannins), esters of so-called gallic acid trimers, such as valoneic acid, flavagallonic acid, etc., occur widely in the plant kingdom. 3) To date, however, only one gallic acid tetramer, gallocatechin, has been found as a constituent of pomegranate (Punica granatum L.) 4) and Indian almond (Terminalia catappa L.) 5) tannins. In the course of our chemical work on tannins in myrobolans, the fruits of Terminalia chebula Retz. (Combretaceae), which are important not only for medicinal use but also as a commercial source of vegetable tannins, we have isolated a hydroxyizable tannin [terchebulin (3)] having a new gallic acid tetramer (terchebulic acid) and biogenetically related tannins [punicalagin (1) 9) and terflavin A (2) 10), besides the tannins [chebulagic acid, 6) chebulinic acid 6, 7) and corilagin 8) which were isolated previously from this material. Furthermore, examination of the leaf extract has resulted in the isolation of a series of related tannins, terflavins B (7), 9) C (9) and D (10), punicalagin (1) and punicalin (8). 9) This paper deals with the isolation and structure elucidation of these compounds.

Initial fractionation of the aqueous acetone extract of commercial myrobolans was achieved by Sephadex LH-20 chromatography with water containing increasing proportions of methanol. 9) Each fraction was repeatedly chromatographed over a variety of reversed-phase gels to yield compounds 1—3, among which two compounds were found to be identical with punicalagin (1) 9) and terflavin A (2) 9) by comparisons of their physical and proton-nuclear magnetic resonance (1H-NMR) spectral data with those of samples isolated from T. catappa. The new tannin, terchebulin (3) was obtained as tan crystals, but the 1H-NMR spectrum was duplicated, showing that in solution 3 exists as an equilibrium mixture of α- and β-forms [δ 5.32 (2/3H), d, J=4 Hz, α-anomeric H; δ 5.01 (1/3H), d, J=8 Hz, β-anomeric H]. The fact that all the aromatic signals correspond to five protons in total appear as sharp singlets (see Experimental) indicated the presence of five penta-substituted aromatic rings. Furthermore, lowfield shift (δ 7.58) of one of the aromatic signals implied the occurrence of a depside-like linkage. 9) The carbon-13 nuclear magnetic resonance (13C-NMR) spectrum more clearly indicated the presence of α- and β-anomers (δ 90.8, α-anomeric C; δ 94.1, β-anomeric C) and two aromatic δ-lactone rings (δ 157.9, 160.3). 9) The observation of four carboxyl carbon signals with normal
chemical shifts (δ 166.8, 167.6, 169.1 and 169.9) is consistent with the 2,3,4,6-tetra-acylation pattern of the hexopyranose moiety.

Partial acid hydrolysis of 3 yielded a hydrolysate (4), together with ellagic acid. The negative fast atom bombardment mass spectrum (FAB-MS)⁶ of 4 exhibited a prominent (M−H)⁻ peak at m/z 781, which was 302 mass units less than that of 3, in agreement with the des-hexahydroxydiphenol structure. In the ¹H-NMR spectrum of 4, the aliphatic signal pattern, although complicated, was similar to that of punicalin (8),⁴¹ suggesting the 4,6-substitution of the glucopyranose ring.

Ordinary phenol methylation of 3 afforded the hexadecamethyl ether [field-desorption mass spectrum (FD-MS): m/z 1308 (M⁺)], which was methanalysed with sodium methoxide in methanol to give dimethyl-(S)-hexa-methoxydiphenolate (6)¹⁰ [(δ)₂³⁵ − 29.4° (CHCl₃)] and a new phenolcarboxylic acid (terchebulic acid) methylate (5). The electron-impact mass spectrum (EI-MS) of 5 exhibited, together with a prominent M⁺ peak at m/z 792, characteristic fragment peaks at m/z 581 and 239 (Chart 3) resulting from the cleavage of the phenyl-phenyl bond.⁴,¹¹ The ¹H-NMR spectrum showed signals due to eleven methoxyl groups and three aromatic one-proton singlets (δ 7.21, 7.45 and 7.74). The presence of two δ-lactone rings in 5 was confirmed by ¹³C-NMR spectroscopy, which showed relatively upfield carboxyl carbon signals at δ 156.6 and 159.0. Furthermore, the two-dimensional nuclear Overhauser effect (NOESY) spectrum of 5 displayed cross peaks between two (δ 7.21, 7.45) of the above three aromatic signals and methoxyl signals, whereas there was no cross...
peak between the aromatic singlet at $\delta$ 7.74 and the methoxyl signals, thus indicating clearly that the aromatic proton and the methoxyl group are not located adjacent. As for the atropisomerism of the phenyl-phenyl bond, the circular dichroism (CD) spectral comparison of 5 with the $S$-flavagallonic acid methylate (11)\textsuperscript{53} (Fig. 1) established it to be in the $S$-series.

Although there is no unequivocal evidence for the orientation of the 4,6-substituted terchebulic acid ester in terchebulin, the fact that terchebulin co-occurs with terflavin A (2) led us to conclude its structure to be as shown by the formula 3.

Our previous work demonstrated that the compositions of tannins almost invariably differ quite markedly in each plant material, particularly in leaf and bark.\textsuperscript{50,40} Thus, we examined the leaves of \textit{T. chebula} to isolate biogenetically related compounds.

The aqueous acetone extract of the leaves collected in Taiwan was separated in a similar manner to that used for the fruits to afford terflavins C (9) and D (10), together with the known tannins, terflavin B (7),\textsuperscript{5} punicalagin (1)\textsuperscript{40} and punicalin (8).\textsuperscript{4} The $^1$H- and $^{13}$C-NMR spectra of 9 and 10 were also duplicated owing to the presence of $\alpha$- and $\beta$-anomers. The relatively large coupling constants of the $^1$H-NMR signals in 9 and 10 suggested the presence of a glucopyranose ring with a $^{4}$C$_1$ conformation. One of the characteristic features in the $^{13}$C-NMR spectra of these compounds was the observation of two $\beta$-lactone carbonyl carbon signals ($\delta$ 159.0, 161.0 in 9; $\delta$ 159.7, 161.0 in 10). The appearance of aromatic singlets at $\delta$ 6.39, 6.40 (1H in total) and 6.50 (1H) in the $^1$H-NMR spectrum of 9 suggested the presence of a hexahydroxydiphenol ester group, whereas 10 showed no such peaks.

Complete acid hydrolysis of 9 yielded glucose, ellagic acid and flavagallonic acid.\textsuperscript{51} On the other hand, partial acid hydrolysis of 9 gave 10, together with ellagic acid. The location of each acyl group was unequivocally established by conversion of terflavin A (2) into 9 on partial enzymic hydrolysis with tannase. Based on these chemical and spectroscopic findings, the structures of terflavins C and D were determined to be 9 and 10, respectively.

The co-occurrence of 4-O-flavagollonyl-6-O-galloyl-$D$-glucose derivatives [terflavins A (2) and B (7)] with terchebulin (3) and punicalagin (1) provides support for an earlier scheme of biogenesis of component phenolicarboxylic acids.\textsuperscript{52} Finally, it should be noted that the molecular weights of leaf tannins are, on the whole, less than those of the fruit tannins, that is, the leaf contains compounds of an earlier biosynthetic stage.

### Experimental

Melting points were determined on a Yanagimoto micro-melting point apparatus and are uncorrected. Optical rotations were measured with a Jasco DIP-4 digital polarimeter. $^1$H- and $^{13}$C-NMR spectra were taken with a JEOL FX-100 spectrometer, with tetramethylsilane as an internal standard; chemical shifts are given on a $\delta$ (ppm) scale. FAB- and FD-MS were recorded on JEOL JMS DX-300 and D-300 spectrometers. Column chromatography was carried out with Sephadex LH-20 (25–100 $\mu$), Pharmacia Fine Chemical Co., Ltd.), MCI-gel CHP 20P (75–150 $\mu$), Mitsubishi Chemical Industries, Ltd.), Fuji-gel ODS G-3 (43–65 $\mu$, Fuji Gel Hanbui Co., Ltd.), Bondapak C$_{18}$/Porasil B (37–75 mesh, Waters Associates, Inc.), Prep-pak 500C$_{18}$ (Waters Associates, Inc.), silica gel 60 (70–230 mesh, Merck) and Avicel cellulose (Funakoshi). Thin-layer chromatography (TLC) was performed on precoated Kieselgel 60 F$_{254}$ plates (0.2 mm thick, Merck) with solvent systems of benzene-ethyl formate-formic acid (1:7:1, 2:10:3), and spots were located by ultraviolet
illumination (Manulight, 2536 Å) and by spraying 1% ferric chloride reagent or 10% sulfuric acid, followed by heating.

Isolation of Tannins a) From Myobalanus: The commercial myobalanus (2.9 g) were extracted four times with 60% aqueous aceton at room temperature. Concentration of the extracts under reduced pressure afforded dark brown precipitates, which were removed by filtration. The filtrate was applied to a Sephadex LH-20 column. Elution with H2O containing increasing amounts of MeOH (20° to 100°) and finally with 100° MeOH (1:1) yielded four fractions: fr. 1 (550 g), fr. 2 (195 g), fr. 3 (178 g) and fr. 4 (117 g). Fraction 1 was recrystallized over MCI-gel CHP 20P with H2O-MeOH (1:0–1:1) and then over Sephadex LH-20 with H2O-MeOH (1:0–0:1) to give gallic acid (12.5 g), ellagic acid (2 g), chelidonic acid (25 g), corosolic acid (25 g) and quercetin 3-O-glucuronide (575 mg). Fraction 2 was recrystallized over MCI-gel CHP 20P, Fuji-gel ODS G-3, Bondapak C18/Porasil B and Prep-pak C18 with H2O-MeOH (1:0–1:1), and Sephadex LH-20 with EtOH to yield triflavin A (2) (48 mg) and purpurogalin (1) (3.8 g). Repeated chromatography of fraction 4 as for fraction 3 gave terechinol (3) (158 mg).

b) From Leaves: The dried leaves (2.1 kg) of *T. chebula*, collected at Ping-tung, Taiwan, O.R.C., were extracted with 70% aqueous acetic acid at room temperature. After concentration of the extract under reduced pressure, the resulting precipitates were removed by filtration. The aqueous solution was applied to a column of Sephadex LH-20, and elution with H2O containing increasing amounts of MeOH yielded three fractions: fr. 1 (550 g) (16%); fr. 2 (347 g) (44%); fr. 3 was recrystallized over MCI-gel CHP 20P and Fuji-gel ODS G-3 with H2O-MeOH (1:0–1:1) and Sephadex LH-20 with H2O-MeOH (1:0–0:1) to give gallic acid (631 mg) and 2,3,5-(S)-hexahydroxydiphenyl-ω-glucopyranose (640 mg). Fraction 2 was twice recrystallized over Sephadex LH-20 with EIOH and 60% aqueous MeOH and MCI-gel CHP 20P, Fuji-gel ODS G-3 and Prep-pak C18 with H2O-MeOH (1:0–1:1) to give ellagic acid (795 mg), 2,3-di-O-galloyl-β-glucose (29 mg), corilagin (616 mg), terechinol (7) (54 mg), purpurogalin (8) (246 mg), triflavin B (7) (180 mg) and triflavin D (7) (38 mg). Fraction 3 was chromatographed over MCI-gel CHP 20P and Fuji-gel ODS G-3 with H2O-MeOH (1:0–1:1) and Sephadex LH-20 with H2O and 80% aqueous MeOH to yield terechinol (3) (108 mg) and purpurogalin (1) (105 mg).

Terechinol (3) A tan crystalline powder (H2O, mp 222°–224° C, [α]25° +136.2° (c=1.5, MeOH). Anal. Calc: C12H10O3 3H2O, C 46.90; H 3.58. Found: C 47.08; H 3.51. 1H-NMR (acetone-d6+D2O): 3.10 (1H, d, J=12 Hz, glc-H6), 4.88 (2H, d, J=3 Hz, 3-H, glc-H2), 5.01 (1H, d, J=18 Hz, glc-H1), 5.32 (2H, s, H2-17), 5.31 (2H, s, H2-17), 6.37 (2H, s, aromatic H), 6.37 (3H, 1H, s, aromatic H), 6.44, 6.63 (each 1H, s, aromatic H), 6.73 (1H, s, aromatic H), 6.83 (2H, s, aromatic H), 7.18 (1H, s, aromatic H). 13C-NMR (acetone-d6+D2O): 64.4 (C-6), 69.6, 70.1, 74.6, 75.1, 77.5, 77.5 (C-10), 90.8 (C-16), 94.1 (C-17), 157.9, 161.7 (2×), 195.9 (C-30) (CDCl3).

Partial Acid Hydrolysis of 3 A solution of 3 (55 mg) in 1H2SO4 (3 ml) was heated at 95° C for 4 h. After cooling, the reaction mixture was chromatographed over MCI-gel CHP 20P (H2O-MeOH) and then over Sephadex LH-20 (EIOH) to yield ellagic acid (9 mg) and the hydrolysate (4) (13 mg) as a tan amorphous powder, [α]28° +195° (c=1.0, MeOH). Negative FAB-MS m/z: 781 (M–H). Anal. Calc: C12H8O6, C 44.06; H 4.10. Found: C 44.41; H 3.88. 1H-NMR (acetone-d6+D2O): 3.0–4.5 (glc-H5), 5.02 (d, J=4 Hz, glc-H3), 6.54, 6.55 (1H, aromatic H), 6.83 (each 1H, s, aromatic H), 7.12, 7.60 (each 1H, t, s, aromatic H). 13C-NMR (acetone-d6+D2O): 61.7 (C-6), 92.8 (C-2), 97.4 (C-1), 157.9, 161.0 (2×) (CDCl3) (COOH).

Partial Acid Hydrolysis of 3 A solution of 3 (50 mg) in 2% H2SO4 (10 ml) was heated at 95° C for 5 h. After cooling, the reaction mixture was concentrated under reduced pressure and the residue was recrystallized from a column of Sephadex LH-20 with EIOH to give triflavin D (10) (13 mg).

Terechinol (10) A pale yellow crystalline powder (H2O-acetone), mp 215°–228° C (dec)., [α]28° +112° (c=1.0, acetone-H2O, 1:1). Anal. Calc: C12H8O6, 3H2O, C 44.88; H 4.16. Found: C 44.57; H 3.50. Negative FAB-MS m/z: 631 (M–H). 1H-NMR (acetone-d6+D2O): 3.0–4.2 (glc-H), 4.58 (t, J=8 Hz, glc-H4), 5.00 (d, J=4 Hz, glc-H3), 6.31, 7.50 (each 1H, s, flavoglyconyl H). 13C-NMR (acetone-d6+D2O): 61.7 (C-6), 92.8 (C-2), 97.4 (C-1), 157.9, 161.0 (2×) (CDCl3) (COOH).

Acid Hydrolysis of 3 A solution of 3 (10 mg) in 3N HCL (3 ml) was heated at 190° C for 20 h. After cooling, the reaction mixture was concentrated under reduced pressure and the residue was recrystallized from a column containing a mixture of H2SO4 and H2O. After neutralization with BaCO3, the solution was analyzed by TLC on cellulose [n-ButOH-pyridine–H2O (6:4:3)] which showed a spot with Rf (0.4) corresponding to glucose. Further elution with 60% aqueous MeOH gave flavaglucic acid (7 mg).

**Terechinol (7)** A solution of 7 (43 mg) in H2O was shaken with tannase (kindly provided by Dr. M. Kanazawa, Tokyo Med. Coll., Japan) at room temperature for 20 h. The reaction mixture was directly subjected to Sephadex LH-20 chromatography with EIOH to yield gallic acid and terflavin C (9) (10 mg).

Acknowledgement The authors are grateful to Mr. Y. Tanaka, Miss
References


