Tannins of Rosaceous Medicinal Plants. Part 2.[1] Gemins A, B, and C, New Dimeric Ellagitannins from *Geum japonicum*

Takashi Yoshida and Takuo Okuda *
Faculty of Pharmaceutical Sciences, Okayama University, Tsushima, Okayama 700, Japan
M. Usman Memon and Tetsuro Shingu
Faculty of Pharmaceutical Sciences, Kobe-Gakuin University, Ikawadani, Nishi-ku, Kobe 673, Japan

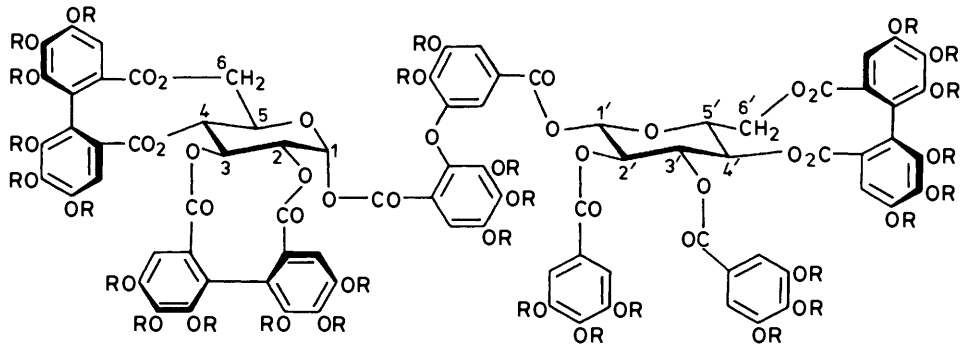
Three novel dimeric hydrolysable tannins, gemins A, B, and C, have been isolated from the leaves of *Geum japonicum* (Thunb.) and their structures, which have α- and β-glucose cores linking through a dehydrodigalloyl group, have been elucidated from spectral and chemical evidence.

Some plants of the *Geum* species (Rosaceae) have been used as diuretics and astringents in Japan and China.[2] Gein was previously found as one of the chemical constituents in *G. urbanum* L.,[3] *G. japonicum* (Thunb.), and *G. macrophyllum* (Willd) *var. sachalinense* (Koidz.) Hara.[4] The plants of this genus are also known to be rich in tannins.[5] In our continuing investigation of the tannins of medicinal plants of the Rosaceae, we have examined the tannins of *G. japonicum* Thunb., and here report the isolation and structure determination of three new dimeric hydrolysable tannins, named gemins A (**1**), B (**16**), and C (**17**).[6]

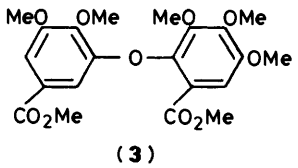
Results and Discussion

The aqueous acetone homogenate of the leaves was successively extracted with ether, ethyl acetate, and butan-1-ol. The ethyl acetate extract was fractionated by droplet counter-current chromatography by the ascending method using the butan-1-ol–propan-1-ol–water (2:1:3, v/v) system. The crude tannins thus obtained were further purified by Sephadex LH-20 column chromatography to yield potentillin (**5**),[7] tellimagrandin I (**10**),[8] tellimagrandin II (**11**),[8] and 1,2,3-tri-*O*-galloyl-β-D-glucose (**14**),[9] which were identified by direct comparison of their spectra with those of authentic samples. Besides them, a tannin whose ^{1}H n.m.r. spectrum showed the presence of a galloyl group (δ 7.21) and a hexahydroxydiphenoyl group (δ 6.73 and 6.48) was also isolated, and was identified as 1-*O*-galloyl-2,3-*O*-[(*S*)-hexahydroxydiphenoyl]-α-D-glucose (sanguiin H-4)[10] (**15**) by comparsion of spectral data. The butan-1-ol-soluble portion was also fractionated in a way similar to the ethyl acetate extract, leading to isolation of gemins A (**1**), B (**16**), and C (**17**) along with pedunculagin (**6**).[8] The three new tannins (**1**), (**16**), and (**17**) were shown to be ellagitannins by their diagnostic colour reaction[11] upon being sprayed with the $NaNO_2$–AcOH reagent on cellulose t.l.c. plates.

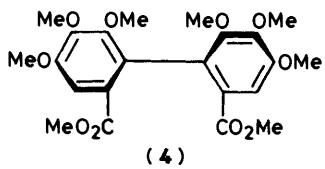
Gemin A (**1**), the main tannin, $[\alpha]_D$ +156° was isolated as a chromatographically homogeneous, amorphous powder. High-performance gel-permeation chromatography of (**1**) showed a retention volume close to that of agrimoniin, a dimeric ellagitannin recently isolated from *Agrimonia japonica* (Miq.) Koidz. and *Potentilla kleiniana* Wight et Arn.[7] The dimeric nature of compound (**1**) was further substantiated by the peak at *m/z* 1 873 $(M + H)^+$ in the fast-atom bombardment mass spectrum (F.A.B.-M.S.), leading to the formula $C_{82}H_{56}O_{52}$. The ^{1}H n.m.r. spectrum (200 MHz) showed seven 1-H singlets at δ 7.35, 6.70, 6.67, 6.65, 6.51, 6.48, and 6.38, two *meta*-coupled doublets at δ 7.31 and 6.88, and two 2-H singlets at δ 7.05 and 7.00 in the aromatic region, suggesting the presence of a dehydrodigalloyl, three hexahydroxydiphenoyl, and two galloyl groups in the molecule.



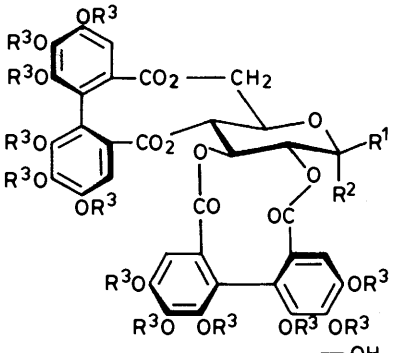
(**1**) R = H
(**2**) R = Me



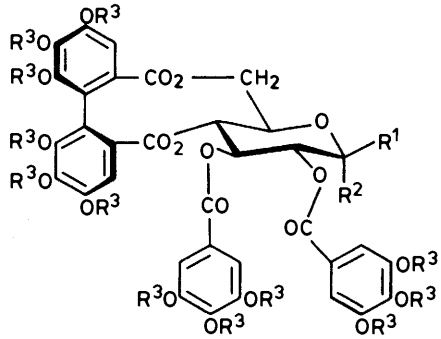
(**3**)



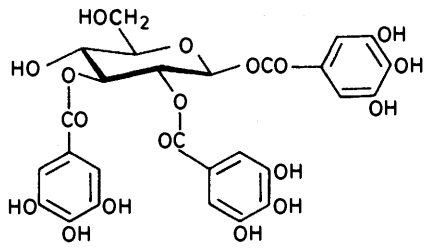
(**4**)



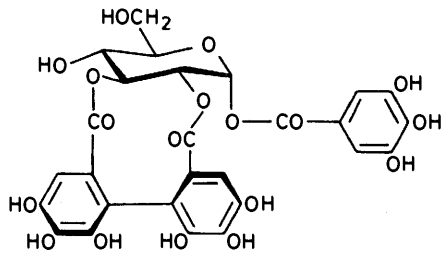
(5) $R^1 = R^3 = H$, R^2 = -OCO- (3,4,5-trihydroxybenzoyl)
(6) R^1, R^2 = H, OH, R^3 = H
(7) R^1 = H, R^2 = OMe, R^3 = Me
(8) R^1 = OMe, R^2 = H, R^3 = Me
(9) R^1 = H, R^2 = -OCO- (3,4,5-trimethoxybenzoyl)-O- (2-methoxy-3,4,5-trimethoxycarbonylphenyl)
R^3 = Me



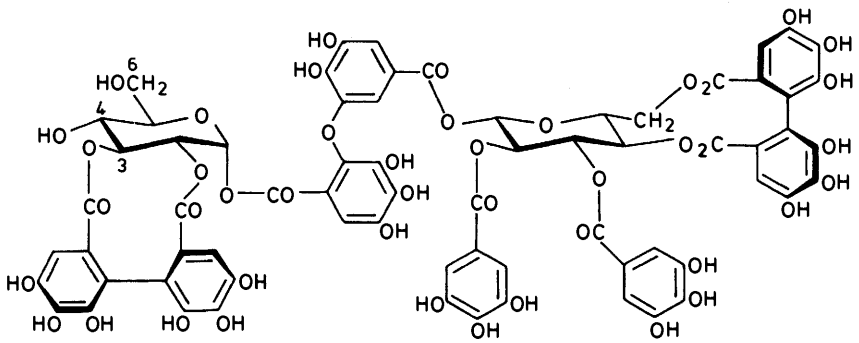
(10) R^1, R^2 = H, OH, R^3 = H
(11) R^1 = -OCO- (3,4,5-trihydroxybenzoyl), $R^2 = R^3$ = H
(12) R^1 = H, R^2 = OMe, R^3 = Me
(13) R^1 = -OCO- (3,4,5-trimethoxybenzoyl)-O- (2-methoxy-3,4,5-trimethoxycarbonylphenyl), R^2 = H, R^3 = Me



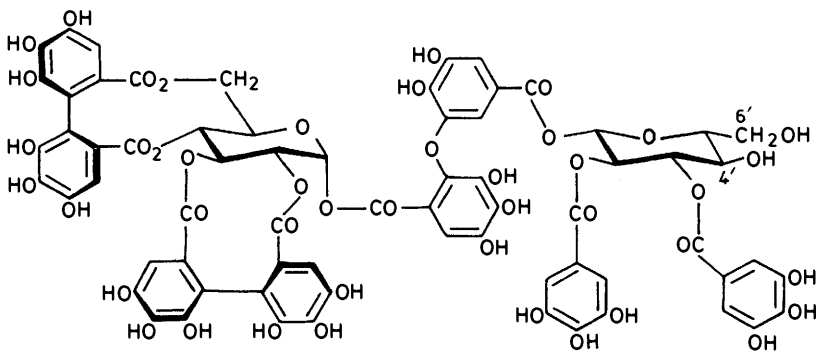
(14)



(15)



(16)



(17)

Table 1. ^{1}H N.m.r. spectral data of gemins A (**1**), B (**16**), and C (**17**): 200 MHz; δ values from $SiMe_4$ in [$^2H_6$]acetone; *J* values in Hz

	Gemin A (**1**)	Gemin B (**16**)[a]	Gemin C (**17**)
Galloyl	7.05, 7.00	7.03, 6.99	7.05, 7.00
DHDG[b]	7.35 (s)	7.25 (s)	7.30 (s)
	7.31 (d, *J* 2)	7.28 (d, *J* 2)	7.24 (d, *J* 2)
	6.88 (d, *J* 2)	6.85 (d, *J* 2)	6.81 (d, *J* 2)
HHDP[c]	6.70, 6.67, 6.65, 6.51, 6.48, 6.38	6.73, 6.65, 6.51, 6.50,	6.66, 6.61, 6.52, 6.35
α-Glucose core			
1-H	6.56 (d, *J* 4)	6.42 (d, *J* 3.5)	6.53 (d, *J* 3.5)
2-H	5.38 (dd, *J* 4, 10)	5.13 (dd, *J* 3.5 and 10)	5.35 (dd, *J* 3.5 and 10)
3-H	5.54 (t, *J* 10)	5.44 (t, *J* 10)	5.49 (t, *J* 10)
4-H	5.19 (t, *J* 10)	4.03 (t, *J* 10)	5.14 (t, *J* 10)
5-H	4.52 (m)	} 3.80[e]	4.45 (dd, *J* 6 and 10)
6-H	5.31 (dd, *J* 7 and 14),[d] 3.69 (d, *J* 14)		5.18 (dd, *J* 6 and 13), 3.61 (d, *J* 13)
β-Glucose core			
1′-H	6.17 (d, *J* 8)	6.15 (d, *J* 8)	6.01 (d, *J* 8)
2′-H	5.59 (dd, *J* 8 and 10)	5.60 (dd, *J* 8 and 9)	5.35 (dd, *J* 8 and 10)
3′-H	5.85 (t, *J* 10)	5.83 (t, *J* 9)	5.56 (t, *J* 10)
4′-H	5.22 (t, *J* 10)	5.21 (t, *J* 9)	3.98 (t, *J* 10)
5′-H	4.52 (m)	4.52 (m)	} 3.80 (m)
6′-H	5.24 (dd, *J* 6 and 14),[d] 3.79 (d, *J* 14)	5.28 (dd, *J* 6 and 13), 3.80[e]	

[a] Measured in [$^2H_6$]acetone–D_2O. [b] Dehydrodigalloyl. [c] Hexahydroxydiphenoyl. [d] Assignments for 6-H and 6′-H may be reversed. [e] Overlapped by the signal of DHO.

Table 2. ^{13}C Chemical shifts of glucose carbons of dimers, (**1**), (**16**), and (**17**), and related monomers; values in p.p.m. from $SiMe_4$ in [$^2H_6$]acetone

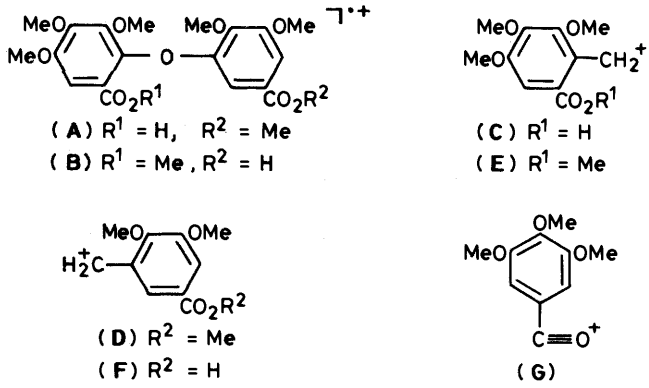
	α-Glucose						β-Glucose					
Compound	C-1	C-2	C-3	C-4	C-5	C-6	C-1′	C-2′	C-3′	C-4′	C-5′	C-6′
(**1**)	91.3	75.1	76.7	69.4	71.5	63.5	94.3	72.2	74.1	71.3	73.6	63.8
(**16**)[a]	91.3	74.5	79.1	67.7	76.3	61.5	94.4	72.2	74.0	71.3	73.6	63.7
(**17**)	90.8	74.3	76.0	69.1	71.0	63.2	93.6	71.8	76.1	69.1	78.5	61.7
(**5**)	90.7	74.1	76.0	69.1	71.0	63.0						
(**11**)							93.8	71.8	73.3	70.8	73.1	63.1
(**14**)							94.2	72.5	76.7	69.8	79.2	62.4
(**15**)	91.4	74.5	79.1	68.0	76.7	61.8						

[a] Measured in [$^2H_4$]methanol.

Upon methylation with diazomethane, gemin A (**1**) formed nonacosa-*O*-methylgemin A (**2**), whose ^{1}H n.m.r. spectrum exhibited signals corresponding to 29 methoxy groups in the region δ 3.60—4.00. Methanolysis of the ether (**2**) afforded methyl tri-*O*-methylgallate, dimethyl penta-*O*-methylde-hydrodigallate (**3**),[12,13] dimethyl (*S*)-4,4′,5,5′,6,6′-hexameth-oxydiphenate (**4**), and glucose which was identified by g.l.c. of the trimethylsilyl ether. The presence of two glucose cores in (**1**) was indicated by two sets of the sequentially coupled seven-spin system at δ 6.56—3.69, which were fully assigned by spin–spin decoupling experiments which started with irradiation of the anomeric proton signals, at δ 6.56 and 6.17, of different coupling constants (4 and 8 Hz, respectively) (Table 1). These data indicate the $^4C_1$ conformation with α- and β-glucosidic linkages for the two glucose cores in (**1**). The ^{13}C n.m.r. spectrum of (**1**) showed 12 distinctive glucose carbon signals including those of the α-and β-anomeric carbons at δ_C 91.3 and 94.3 p.p.m., which are comparable with those of potentillin (**5**) and tellimagrandin II (**11**) (Table 2).

Based on these data, combined with the occurrence of 10 ester carbonyl carbon resonances over the range δ_C 164.0—170.9 p.p.m. in the ^{13}C n.m.r. spectrum of (**1**), gemin A is presumed to be a dimeric tannin possessing fully acylated α-and β-glucose cores which are linked to each other through the dehydro-digalloyl group in a way analogous to agrimoniin.[7] Allocation of each phenolic ester group on the glucopyranose ring was based on the following evidence. Methylation of (**1**) with dimethyl sulphate and potassium carbonate in refluxing acetone for 30 h provided five products of partial degradation, (**7**)—(**9**), (**12**), and (**13**), in addition to (**2**)—(**4**) and methyl tri-*O*-methylgallate, which were separated by repeated p.l.c. (prepar-ative-layer chromatography) with multi-development. Three of the five products, (**7**), (**8**), and (**12**), were identified as the α- and β-anomer of trideca-*O*-methylpedunculagin and the α-anomer of trideca-*O*-methyltellimagrandin I, respectively, by direct comparisons with authentic specimens. The ^{1}H n.m.r. spectrum of the fourth product (**9**) showed two *meta*-coupled 1-H doublets (*J* 2 Hz) at δ 7.34 and 6.88, and a singlet at δ 7.40, which are due to the dehydrodigalloyl group, and also the signals due to two hexahydroxydiphenoyl groups at δ 6.98, 6.92, 6.77, and 6.61 (1 H each). The anomeric proton signal is shown at δ 6.62 as a doublet of small coupling constant (*J* 3.5 Hz). These data and the molecular ion peak at *m/z* 1 356 ($C_{66}H_{68}O_{31}$) in the mass spectrum, and also the isolation of (**7**) and (**8**) upon the partial hydrolysis described above, lead to the structure (**9**) for this compound. The determination of the mode of attachment of the dehydrodigalloyl group to the anomeric centre as in structure (**9**) was based on the fragmentation peaks in the mass spectrum:

Table 3. Rearrangement product peaks of (**9**) and (**13**)[a]



	(C) *m/z* 225	(D) *m/z* 209	(E) *m/z* 239	(F) *m/z* 195
(**9**)	5.0%	5.0%		
(**13**)			4.0%	19%

[a] Recorded at ionizing potential 20 eV, and the base peaks in (**9**) and (**13**) are *m/z* 404 and 212, respectively.

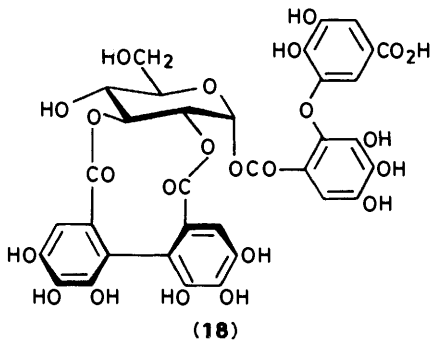
the peaks at *m/z* 225 [fragment (**C**)] and 209 [fragment (**D**)] arise on cleavage of the ether linkage of the monomethyl ester ion (**A**), accompanied by rearrangements of a methyl group and a hydrogen[14] (Table 3). The possibility of formation of fragment (**C**) from the hexamethoxydiphenoyl group by an analogous breakdown of the biphenyl bond was excluded by the absence of a peak at *m/z* 225 in the mass spectra of (**7**) and (**8**). Structure (**9**) was confirmed by identification with one of the partial hydrolysates produced upon similar methylation of agrimoniin.[1]

The structure, (**13**), of the fifth product was similarly deduced based on the ^{1}H n.m.r. spectrum which showed the presence of a dehydrodigalloyl group [δ 7.37 and 6.83 (1 H each, d, *J* 2 Hz) and 7.35 (1 H, s)], a hexahydroxydiphenoyl group [δ 6.77 and 6.95 (1 H each, s)], and two galloyl groups [δ 7.21 and 7.22 (2 H each, s)], and also on the mass spectral peak at *m/z* 239 which is ascribable to the fragment (**E**) produced from the ion (**B**). The *m/z* 195 peak which corresponds to fragment (**F**) was also exhibited, but we prefer to assign this to fragment (**G**) arising from the galloyl group, since the intensity (19%) is stronger than that expected for the rearranged fragment (**F**).

The chirality of the hexahydroxydiphenoyl groups in (**1**) was assigned as *S* on the basis of production of (**4**), (**7**), (**8**), and (**12**) having (*S*)-hexahydroxydiphenoyl groups.

Based on these findings, the structure of gemin A was unambigously established as (**1**).

Gemin B (**16**), $[\alpha]_D$ +67°, $C_{68}H_{50}O_{44}\cdot 7H_2O$, and gemin C (**17**), $[\alpha]_D$ +133°, $C_{68}H_{50}O_{44}\cdot 8H_2O$, were obtained as amorphous powders. These tannins are indistinguishable from each other on normal-phase h.p.l.c., but different retention times are observed on reversed-phase h.p.l.c. The presence of a dehydrodigalloyl, two hexahydroxydiphenoyl, and two galloyl groups, and α- and β-glucose residues, in both (**16**) and (**17**) is shown by their ^{1}H (Table 1) and ^{13}C (Table 2) n.m.r. spectra, and also by identification of the products from their acid hydrolysis. The ^{1}H and ^{13}C n.m.r. spectra of (**16**) and (**17**) are very similar to those of gemin A (**1**) except that the signals attributable to a hexahydroxydiphenoyl group are missing. These data indicate that gemins B and C are dimeric ellagitannins closely related to gemin A, and that the former two are isomers differing in the location of a hexahydroxydiphenoyl



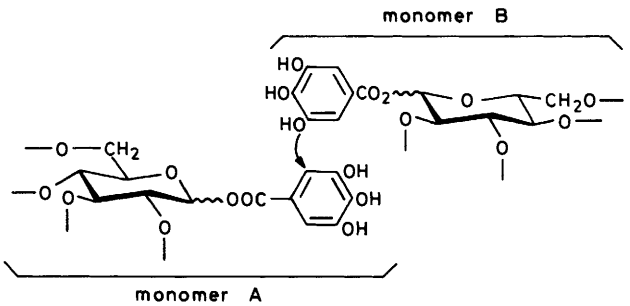
group. The glucose proton signals of (**16**) and (**17**) were assigned by exhaustive spin–spin decoupling experiments in a manner similar to that of gemin A (**1**) (Table 1). The protons on the β-glucose core in (**16**) show essentially the same chemical shifts as those of (**1**), while the signals due to 4-H and 6-H on the α-glucose core are shifted upfield from those of (**1**), indicating the presence of free hydroxy groups at C-4 and C-6 in gemin B (**16**). Analogously, the presence of free hydroxy groups at C-4′ and C-6′ in gemin C (**17**) was revealed by the upfield shifts of 4′-H and 6′-H from those of (**1**). The location of one of the two hexahydroxydiphenoyl groups in these tannins is regarded as being 4′-O ~ 6′-O for gemin B (**16**), and 4-O ~ 6-O for gemin C (**17**) on the basis of the substantial upfield shift of one of the methylene protons at C-6 (6′) which is analogous to that[16] generally exhibited by the ellagitannins having the hexahydroxydiphenoyl ester linkage on the 4,6-positions of the D-glucopyranose ring. Upon enzymatic hydrolysis with tannase, both gemins B (**16**) and C (**17**) gave a partial hydrolysate (**18**), which is identical with a product of the analogous treatment of gemin A (**1**). The ^{1}H n.m.r. spectrum of (**18**) shows peaks characteristic of a dehydrodigalloyl group (singlet at δ 7.27 and two meta-coupled doublets at δ 6.85 and 7.27) and of a hexahydroxydiphenoyl group (two 1-H singlets at δ 6.69 and 6.49). A proton on the acylated anomeric centre is exhibited at δ 6.41 as a doublet (*J* 3.5 Hz), and the 2-H and 3-H signals were observed at δ 5.10 (dd, *J* 3.5 and 10 Hz) and 5.41 (t, *J* 10 Hz), respectively, while the signals due to 4-, 5-, 6-, 6′-H appeared at higher field. Therefore, this hydrolysate was formulated as (**18**), which established the location of the second hexahydroxydiphenoyl group at 2-O and 3-O and a dehydrodigalloyl group at 1-O on the glucose residue in gemins B and C.

The ^{13}C resonances of the glucose moieties in gemins B and C were reasonably assigned, on consideration of acylation shifts,[16] by comparison with those of gemin A (**1**) which were unequivocally assigned by $\{^1H\}$–^{13}C selective decoupling experiments.[16] These assignments were also substantiated by correlation of the resonances with the corresponding peaks of their monomeric units [(**15**) and (**11**) for (**16**), and (**5**) and (**14**) for (**17**)], which is analogous to the spectral relationships among (**1**), (**5**), and (**11**).

The c.d. spectra of gemins A (**1**), B (**16**), and C (**17**) exhibited a strong positive Cotton effect at 234—236 nm, indicating that the hexahydroxydiphenoyl groups in gemins B and C have the (*S*)-configuration which is identical with that established for these groups in gemin A. This conclusion is also in accord with the empirical rule[17] in which the sign and amplitude of the Cotton effect around 235 nm reflect the chirality and the number of the hexahydroxydiphenoyl groups in the molecule of ellagitannins.

Interpretation of these chemical and spectral features for gemins B and C yielded the structures (**16**) and (**17**).

Biogenetically the dimers found in this study are presumed to



Dimer	Monomer A	Monomer B
(1)	(5)	(11)
(16)	(15)	(11)
(17)	(5)	(14)
Agrimoniin	(5)	(5)

Scheme.

be formed by an intermolecular C–O oxidative coupling between the galloyl groups at C-1 of two monomeric tannins (Scheme) as proposed for biogenesis of a dehydrodigallic acid.[18] This assumption is supported by the coexistence of these dimers and their conceivable monomeric precursors, (**5**), (**11**), (**14**), and (**15**) in the same plant. It is also noteworthy that in *Agrimonia japonica* and *Potentilla kleiniana*, in which agrimoniin is the sole dimer, only potentillin (**5**) has been found as the monomeric tannin possessing the galloyl group at C-1.[1,7] Although several dimeric hydrolysable tannins have been found recently,[19–21] all of them are regarded as dimers in which monomers link through a valoneoyl or sanguisorboyl[20] group formed by an analogous oxidative coupling between the galloyl group at C-1 of one monomer and the hexahydroxydiphenoyl group at C-4 and C-6 of the other. Gemins A, B, and C, in addition to agrimoniin, are therefore dimers of a novel type.

Experimental

Optical rotations were measured on a JASCO DIP-4 polarimeter at room temperature. I.r. spectra were recorded on a JASCO A-102 spectrophotometer and u.v. spectra on a Hitachi 200-10 spectrophotometer. Mass spectra were recorded on a Shimadzu LKB-9000 GC-MS spectrometer. ^{1}H N.m.r. and ^{13}C n.m.r. spectra were measured on either a Hitachi R-22 FTS or a JEOL FX-200 spectrometer, with $SiMe_4$ as internal standard, and the chemical shifts are given in δ-values. Normal-phase h.p.l.c. was performed on a column of Nomura Develosil 60-5, 4 × 150 mm, with hexane–MeOH–tetrahydrofuran–formic acid (55:33:11:1, v/v), containing oxalic acid (450 mg l^{-1}), as eluant. Reversed-phase h.p.l.c. was run on a column of Merck LiChrosorb RP-18 (10 μ), 4 × 300 mm, with 0.1M H_3PO_4–0.1M KH_2PO_4–EtOH–EtOAc (50:50:2:5, v/v) as developer. Detection was effected by u.v. absorption at 280 nm. T.l.c. was performed on cellulose (Funakoshi) plates (0.3 mm) in the solvent systems (A), 7% aqueous acetic acid, and (B), butan-1-ol–AcOH–water (4:1:5; upper layer). Kieselgel PF_{254} (Merck) was used for analytical and preparative t.l.c. (0.25 and 0.7 mm, respectively) (solvent system C: benzene–acetone 6:1, v/v). The spots were visualized by u.v. light (254 nm), or with iron(III) chloride or $NaNO_2$–AcOH spray reagents. Sephadex LH-20 (100 μ) (Pharmacia Fine Chemicals) and Avicel microcrystalline cellulose (Funakoshi) were used for column chromatography. Solutions were evaporated under reduced pressure below 40 °C. Light petroleum refers to that fraction boiling in the range 85—120 °C.

Isolation of Tannins.—Fresh leaves (1.6 kg) of *G. japonicum* collected in August was homogenized in acetone–water (1:1, v/v) (3 × 4.3 l) and filtered through Celite-545. The homogenate was concentrated to *ca.* 1.0 l, and extracted with ether (4 × 500 ml), and then with ethyl acetate (16 × 600 ml). The ethyl acetate layer was evaporated to give a dark brown residue (13 g). The aqueous layer was further extracted with butan-1-ol saturated with water (14 × 600 ml). Removal of the solvent yielded a hygroscopic brown residue (20 g). A portion (4 g) of the ethyl acetate extract was subjected to droplet counter-current chromatography (100 glass tubes, 3.2 mm × 120 cm) by the ascending method using butan-1-ol–propan-1-ol–water (4:1:5), v/v) as the solvent system. Every fifth fraction (12 g each) was monitored by u.v. absorption at 280 nm, t.l.c. (cellulose; solvent A), and h.p.l.c. to give three batches of crude tannin fractions; A, fractions 20—40 (450 mg), B, fractions 41—65 (346 mg), and C, fractions 66—85 (227 mg).

Fraction A was purified by Sephadex LH-20 column chromatography (3 × 46 cm) with ethanol as eluant to afford tellimagrandin II (**11**) (27 mg). Fraction B was similarly chromatographed over Sephadex LH-20 (ethanol) to give tellimagrandin I (**10**) (10 mg), potentillin (**5**) (10 mg), and 1,2,3-tri-*O*-galloyl-β-D-glucose (**14**) (10 mg). Purification of fraction C by Sephadex LH-20 chromatography gave 1-*O*-galloyl-2,3-*O*-[(*S*)-hexahydroxydiphenoyl]-α-D-glucose (**15**) (12 mg), $[\alpha]_D$ + 99.4° (*c* 1.6 in acetone); $\lambda_{max.}$(MeOH) 222 (log ε 4.66) and 265 nm (4.29); δ_H (90 MHz; CD_3COCD_3–D_2O) 7.21 (2 H, s, ArH), 6.73 and 6.48 (1 H each, s, ArH), 6.52 (d, *J* 3.5 Hz, 1-H), 5.48 (t, *J* 9 Hz, 3-H), 5.18 (dd, *J* 3.5 and 9 Hz, 2-H), and 4.05—3.83 (4 H, m, 4-H, 5-H, and 6-H_2). These data were coincident with those of sanguiin H-4.[12] A portion (3 g) of the butan-1-ol-soluble fraction was similarly subjected to droplet counter-current chromatography (ascending method; butan-1-ol–propan-1-ol–water, 4:1:5, v/v). Fractions 56—80 contained a mixture of *gemins A* and *C*, which was separated by Sephadex LH-20 column chromatography (2.2 × 14 cm) with gradient elution with ethanol→water to give pure gemin A (**1**) (270 mg) and gemin C (**17**) (35 mg). The crude tannin from fractions 91—140 was also rechromatographed over Sephadex LH-20 to give *gemin B* (**16**) (39 mg) and an additional crop of gemin A (200 mg). Fractions 156—200 yielded pedunculagin (**6**) (190 mg).

Gemin A (**1**).—This tannin was obtained as an amorphous powder (Found: C, 49.2; H, 3.5. $C_{82}H_{56}O_{52}\cdot 7H_2O$ requires C, 49.3; H, 3.5%); $[\alpha]_D$ +156° (*c* 1.6 in ethanol); $\lambda_{max.}$(MeOH) 222 (log ε 5.26) and 272 nm (4.92); $\nu_{max.}$(KBr) 3 425, 1 730, 1 615, 1 510, 1 445, 1 340—1 240, 1 180, 1 040, and 1 010 cm^{-1}; c.d. (MeOH) $[\theta]_{235}$ +33.5 × 10^4, $[\theta]_{262}$ −6.8 × 10^4, and $[\theta]_{283}$ + 6.5 × 10^4; δ_C (50.1 MHz; CD_3COCD_3) 170.9, 169.8, 169.6 (2 C), 169.1 (2 C), 167.6, 166.7, 166.1, and 164.7 (ester carbonyl), and glucose carbons, see Table 2; t.l.c. (cellulose) R_F 0.31 (solvent A), 0.30 (B).

Methylation of Gemin A (**1**) *with Diazomethane.*—A solution of gemin A (**1**) (50 mg) in acetone (2 ml) was treated with excess of ethereal diazomethane for 3 h at room temperature. The crude methyl ether was purified by p.l.c. (solvent C) to give *nonacosa-O-methylgemin A* (**2**) (3.7 mg) as a white amorphous solid (Found: C, 57.8; H, 5.3. $C_{111}H_{114}O_{52}\cdot H_2O$ requires C, 58.0; H, 5.1%); $[\alpha]_D$ +46.5° (*c* 1.3 in acetone); $\lambda_{max.}$(MeOH) 215 (log ε 5.15), 250 (4.74), and 300 nm (4.13); $\nu_{max.}$(KBr) 1 745, 1 585, 1 480, 1 455, 1 410, 1 390, 1 330, 1 280, 1 200, 1 150, 1 120, and 1 100 cm^{-1}; δ_H (200 MHz; CD_3COCD_3) [ArH: 7.22 and 7.20 (2 H each, s), 7.41 and 6.90 (1 H each, d, *J* 2 Hz), 7.45, 6.98, 6.93, 6.92, and 6.63 (1 H each, s)], [α-glucose core: 6.57 (1 H, d, *J* 3.5 Hz, 1-H), 5.66 (1 H, t, *J* 9 Hz, 3-H), 5.38 (1 H, dd, *J* 3.5 and 9 Hz, 2-H), 5.07 (1 H, t, *J* 9 Hz, 4-H), 4.90 (1 H, dd, *J* 6 and 13 Hz, 6-H), 4.31 (1 H, dd, *J* 9 Hz, 5-H), *ca.* 4.1 (6′-H, overlapped by

OMe signal)], [β-glucose core: 6.24 (1 H, d, *J* 8 Hz, 1′-H), 5.88 (1 H, t, *J* 9 Hz, 3′-H), 5.61 (1 H, dd, *J* 8 and 9 Hz, 2′-H), 5.35 (1 H, dd, *J* 7 and 14 Hz, 6′-H), 5.27 (1 H, t, *J* 9 Hz, 4′H), 4.57 (1 H, dd, *J* 7 and 9 Hz, 5′-H), *ca.* 4.1 (6′-H, overlapped by OMe signal)]; t.l.c. (silica gel) R_F 0.35 (solvent C).

*Methanolysis of Nonacosa-*O*-methylgemin A* (**2**).—To a solution of compound (**2**) (29 mg) in absolute methanol (1 ml) was added 1% aqueous sodium methoxide (0.2 ml). After 12 h at room temperature, the reaction mixture was neutralized with Amberlite IR-120 resin (H-form) and evaporated to dryness after filtration. The residue was treated with diazomethane for 30 min and evaporated to give a syrupy residue which was partitioned between dichloromethane and water. Evaporation of dichloromethane phase and purification of the residue by p.l.c. (solvent C) yielded methyl tri-*O*-methylgallate (4.6 mg), m.p. 80—81 °C; *m/z* 226 (M^+); dimethyl penta-*O*-methyldehydrodigallate (**3**) (4.5 mg), m.p. 113—114 °C; *m/z* 436 (M^+); and dimethyl (*S*)-4,4′,5,5′,6,6′-hexamethoxydiphenate (**4**) (7.6 mg), $[\alpha]_D$ −38° (*c* 1.3 in ethanol); *m/z* 450 (M^+). These methanolysates were identified with authentic samples by $[\alpha]_D$, m.s., and ^{1}H n.m.r. spectra. The aqueous layer gave glucose which was identified by g.l.c. and g.c.-m.s. of its trimethylsilyl ether.

Methylation of Gemin A (**1**) *with Dimethyl Sulphate.*—To a solution of gemin A (**1**) (110 mg) in dry acetone (7 ml) were added anhydrous potassium carbonate (550 mg) and dimethyl sulphate (0.4 ml). The reaction mixture was stirred for 12 h at room temperature, and then refluxed for 30 h. Removal of the solvent after filtration gave a syrupy residue which was separated into 7 zones by p.l.c. (benzene–acetone, 14:1; 5 developments). From zones 1—3 (zone 1: highest R_F), methyl tri-*O*-methylgallate (5.1 mg), dimethyl penta-*O*-methyldehydrodigallate (**3**) (2.5 mg), and dimethyl (*S*)-4,4′,5,5′,6,6′-hexamethoxydiphenate (**4**) (7.6 mg) were obtained, and identified by comparison with authentic specimens. Zone 4 contained a mixture of trideca-*O*-methyl-α-tellimagrandin I (**12**) (2 mg), and trideca-*O*-methyl-α-pedunculagin (**7**) (3 mg), which were separated by further p.l.c. (light petroleum–acetone, 2:1; double development). Zone 5 was further purified by p.l.c. (benzene—acetone, 5:1) to give trideca-*O*-methyl-β-pedunculagin (**8**) (5 mg) and octadecamethyl derivative (**9**) (6 mg). Zones 6 and 7 gave octadecamethyl derivative (**13**) (13 mg) and nonacosa-*O*-methylgemin A (**2**) (25 mg), respectively.

Octadecamethyl derivative (**9**). White amorphous solid (Found: C, 57.2; H, 5.2. $C_{66}H_{68}O_{31}\cdot 2H_2O$ requires C, 56.9; H, 5.2%); $[\alpha]_D$ +46° (*c* 1.3 in acetone); m.s. (70 eV) *m/z* 1 356 (M^+); $\lambda_{max.}$(MeOH) 222 (log ε 5.3), 252sh (4.92), and 295 nm (4.13); δ_H (200 MHz; CD_3COCD_3) 7.34 and 6.88 (1 H each, d, *J* 2 Hz, ArH), 7.39, 6.97, 6.91, 6.79, and 6.60 (1 H each, s, ArH), 6.62 (1 H, d, *J* 3.5 Hz, 1-H), and 4.02—3.65 (18 × OMe); t.l.c. (silica gel) R_F 0.61 (solvent C).

Octadecamethyl derivative (**13**). White amorphous powder (Found: C, 58.2; H, 5.5. $C_{66}H_{70}O_{31}\cdot H_2O$ requires C, 58.3; H, 5.2%); $[\alpha]_D$ +24° (*c* 1.3 in acetone); m.s. (70 eV) *m/z* 1 358 (M^+); $\lambda_{max.}$(MeOH) 215 (log ε 5.09), 260 (4.6), and 300: sh nm (4.11); δ_H (200 MHz; CD_3COCD_3) 7.22 and 7.21 (2 H each, s, ArH), 7.35, 6.95, and 6.77 (1 H each, s, ArH), 7.37 and 6.83 (1 H each, d, *J* 2 Hz, ArH), 6.28 (1 H, d, *J* 8 Hz, 1-H), 5.89 (1 H, t, *J* 10 Hz, 3-H), 5.62 (1 H, dd, *J* 8 and 10 Hz, 2-H), 5.31 (1 H, dd, *J* 6 and 14 Hz, 6-H), 5.29 (1 H, t, *J* 10 Hz, 4-H), and 4.62 (1 H, dd, *J* 6 and 10 Hz, 5-H), 6′H is hidden by methoxy-group signals at 3.64—3.97 (18 × OMe); t.l.c. (silica gel) R_F 0.48 (solvent C).

Gemin B (**16**).—This compound was isolated as an off-white amorphous powder (Found: C, 48.4; H, 4.2. $C_{68}H_{50}O_{44}\cdot 7H_2O$ requires C, 48.1; H, 3.8%); $[\alpha]_D$ +67° (*c* 1.4 in ethanol); $\lambda_{max.}$(MeOH) 221 (log ε 5.01) and 275 nm (4.68); $\nu_{max.}$(KBr) 3 425, 1 725, 1 610, 1 510, 1 445, 1 220, 1 190, and 1 030 cm^{-1}; c.d. (MeOH) $[\theta]_{236}$ +25 × 10^4, $[\theta]_{265}$ −6.4 × 10^4, and $[\theta]_{286}$ +3.2 × 10^4; δ_C (50.1 MHz; CD_3OD) 171.5, 170.1, 169.5, 169.1, 167.5, 166.7, 166.4, and 164.1 (ester carbonyl); glucose carbons, see Table 2; t.l.c. (cellulose) R_F 0.36 (solvent A), 0.33 (B).

Gemin C (**17**).—Gemin C was obtained as an off-white amorphous powder (Found: C, 47.4; H, 3.7. $C_{68}H_{50}O_{44}\cdot 8H_2O$ requires C, 47.6; H, 3.5%); $[\alpha]_D$ +133° (*c* 1.1 in ethanol); $\lambda_{max.}$(MeOH) 221 (log ε 5.05) and 276 nm (4.73); $\nu_{max.}$(KBr) 3 450, 1 720, 1 610, 1 440, 1 360—1 300, 1 210, 1 180, and 1 020 cm^{-1}; c.d. (MeOH) $[\theta]_{235}$ +22 × 10^4, $[\theta]_{260}$ −3.8 × 10^4, and $[\theta]_{281}$ +6.3 × 10^4; δ_C (50.1 MHz; CD_3COCD_3) 169.6, 168.4 (2 C), 168.0, 166.7, 166.1, 165.4, and 163.2 (ester carbonyl); glucose carbons, see Table 2; t.l.c. (cellulose) R_F 0.29 (solvent A).

Acid Hydrolysis of Gemins B (**16**) *and C* (**17**).—A solution of gemin B (**16**) (5 mg) in 5% sulphuric acid (2 ml) was refluxed for 4 h. After neutralization with Amberlite IR-45 ion-exchange resin (OH-form), the mixture was filtered and extracted with ethyl acetate (3 × 2 ml). The ethyl acetate extract, after evaporation, was treated with ethereal diazomethane for 1 h at room temperature to yield methyl tri-*O*-methylgallate, dimethyl penta-*O*-methyldehydrodigallate (**3**), and tetra-*O*-methylellagic acid which were separated by p.l.c. (silica gel; solvent C) and identified by comparison of t.l.c. and mass spectra of authentic samples. The aqueous layer gave glucose which was identified by g.l.c. after trimethylsilylation.

Gemin C (**17**) was similarly hydrolysed to give the same products.

Hydrolysis of Gemins A (**1**), *B* (**16**), *and C* (**17**) *with Tannase.*—A solution of gemin A (**1**) (100 mg) in water (10 ml) was treated with tannase at 37 °C for 5 days. The reaction mixture was chromatographed over Sephadex LH-20 (1.1 × 30 cm), eluted with water mixed with increasing amount of ethanol, and 18-g portions of eluant were collected. Elution with water–ethanol (75:25) yielded 2,3-*O*-[(*S*)-hexahydroxydiphenoyl]-D-glucose (4 mg) and gallic acid (15 mg). Elution with water–ethanol (1:1) afforded hydrolysate (**18**) (17 mg), $[\alpha]_D$ +56° (*c* 1.0 in ethanol); $\lambda_{max.}$(MeOH) 222 (log ε 4.68) and 265 nm (4.36); δ_H (200 MHz; CD_3COCD_3) 7.27 and 6.85 (1 H each, d, *J* 2 Hz, ArH), 7.27, 6.69, and 6.49 (1 H each, s, ArH), 6.41 (1 H, d, *J* 3.5 Hz,1-H), 5.41 (1 H, t, *J* 10 Hz, 3-H), 5.10 (1 H, dd, *J* 3.5 and 10 Hz, 2-H), 4.00 (1 H, t, *J* 10 Hz, 4-H), and 3.70 (3 H, 5-, 6-, and 6′-H); t.l.c. (cellulose) R_F 0.65 (solvent A); h.p.l.c. R_t 3.53 min (normal phase) and 5.90 min (reversed phase).

Gemins B (**16**) and C (**17**) (1 mg each) were similarly treated with tannase, and the reaction mixture was analysed by t.l.c. (cellulose; solvent A) and h.p.l.c. (normal and reversed phases) to show the presence of the same products, including the main product (**18**), as those from gemin A (**1**).

Acknowledgements.
We thank Professor T. Koga and Miss N. Toh of Daiichi College of Pharmaceutical Sciences for measuring the c.d. spectra, and Dr. K. Fukushima of Research Institute of Biochemodynamics, Chiba University for measurement of the fast-atom bombardment mass spectra. One of us (M. U. M.) is grateful to the Ministry of Education, Science and Culture of Japan for a foreign student scholarship.

References

1 Part 1, T. Okuda, T. Yoshida, M. Kuwahara, M. U. Memon, and T. Shingu, *Chem. Pharm. Bull.*, 1984, **32**, 2165.

2 L. M. Perry, 'Medicinal Plants of East and Southeast Asia,' MIT Press, Cambridge, Massachusetts, 1980, p. 343.
3 H. Herissey and J. Cheymol, *Compt. Rend.*, 1925, **180**, 384.
4 M. Takahashi and T. Fujita, *Yakugaku Zasshi*, 1955, **75**, 1567.
5 R. Hegnauer, 'Chemotaxonomie der Pflanzen,' Birkhäuser Verlag, Basel and Stuttgart, 1973, vol. **6**, p. 108.
6 This study has been outlined in preliminary communications: T. Yoshida, T. Okuda, M. U. Memon, and T. Shingu, *J. Chem. Soc., Chem. Commun.*, 1982, 351; T. Yoshida, Y. Maruyama, T. Okuda, M. U. Memon, and T. Shingu, *Chem. Pharm. Bull.*, 1982, **30**, 4245.
7 T. Okuda, T. Yoshida, M. Kuwahara, M. U. Memon, and T. Shingu, *J. Chem. Soc., Chem. Commun.*, 1982, 163.
8 T. Okuda, T. Yoshida, M. Ashida, and K. Yazaki, *J. Chem. Soc., Perkin Trans. 1*, 1983, 1765.
9 T. Okuda, T. Hatano, N. Ogawa, and K. Yazaki, *Chem. Pharm. Bull.*, 1982, **30**, 4230.
10 T. Tanaka, G. Nonaka, and I. Nishioka, Abstract Papers, 102nd Annual Meeting of the Pharmaceutical Society of Japan, 1982, April.
11 E. C. Bate-Smith, *Phytochemistry*, 1972, **11**, 1153.
12 W. Mayer, *Justus Liebigs Ann. Chem.*, 1954, **578**, 34.
13 T. Ozawa, K. Haga, D. Kobayashi, T. Kamiyama, and Y. Takano, *Agric. Biol. Chem.*, 1977, **41**, 1249.
14 Ch. Wünsche, A. Sachs, A. Eiwiller, and W. Mayer, *Tetrahedron*, 1968, **24**, 3407.
15 J. C. Jochims, G. Taigel, and O. T. Schmidt, *Justus Liebigs Ann. Chem.*, 1968, **717**, 169.
16 T. Yoshida, T. Hatano, T. Okuda, M. U. Memon, T. Shingu, and K. Inoue, *Chem. Pharm. Bull.*, 1984, **32**, 1790.
17 T. Okuda, T. Yoshida, T. Hatano, T. Koga, N. Toh, and K. Kuriyama, *Tetrahedron Lett.*, 1982, **23**, 3937.
18 O. T. Schmidt and W. Mayer, *Angew. Chem.*, 1956, **68**, 103.
19 T. Okuda, T. Hatano, and N. Ogawa, *Chem. Pharm. Bull.*, 1982, **30**, 4234.
20 G. Nonaka, T. Tanaka, and I. Nishioka, *Chem. Pharm. Bull.*, 1982, **30**, 2255 and references cited therein.
21 R. K. Gupta, S. M. K. Al-Shafi, K. Layden, and E. Haslam, *J. Chem. Soc., Perkin Trans. 1*, 1982, 2525.

Received 30*th May* 1984; *Paper* 4/866