

JOURNAL OF NATURAL PRODUCTS

© Copyright 1996 by the American Chemical Society and the American Society of Pharmacognosy

Volume 59, Number 7

July 1996

Full Papers

Anti-HIV Triterpene Acids from *Geum japonicum*

Hong-Xi Xu, Fa-Quan Zeng, Min Wan,[†] and Keng-Yeow Sim*

Departments of Chemistry and Biochemistry, National University of Singapore, 10 Kent Ridge Crescent, 119260 Singapore

Received December 4, 1995[©]

The methanol extract from the whole plant of *Geum japonicum* was found to inhibit the human immunodeficiency virus (HIV-1) protease. Through bioassay-directed fractionation of the extract, a new triterpene acid along with five known triterpene acids, ursolic acid, epimolic acid, maslinic acid, euscaphic acid, and tormentic acid, were isolated. The structure of the new compound was determined by spectral means including ¹H-¹H COSY, HMQC, HMBC, and NOE experiments to be 2 α ,19 α -dihydroxy-3-oxo-12-ursen-28-oic acid (**1**). Of these compounds, **1**, ursolic acid, and maslinic acid showed potent inhibitory activity against HIV-1 protease.

Introduction

The whole plant of *Geum japonicum* Thunb. (Rosaceae) has been used as a diuretic in traditional Chinese medicine.¹ From this plant, several hydrolyzable tannins and triterpenoids have been isolated.¹⁻³ In a previous study of *G. japonicum* aqueous extract we isolated geponin and gallic aldehyde, which showed antiviral activity.⁴

In the course of a continuing search for potential plant anti-HIV agents,⁵ the methanol extract of the whole plant of *G. japonicum* was found to show significant inhibitory activity against HIV-1 protease. In this paper, we describe the isolation and structure elucidation of a new triterpene acid, 2 α ,19 α -dihydroxy-3-oxo-12-ursen-28-oic acid (**1**), along with five known triterpene acids, ursolic acid (**2**), epimolic acid (**3**), maslinic acid (**4**), euscaphic acid (**5**), and tormentic acid (**6**) (Figure 1), and their inhibitory activity against HIV-1 protease.

Results and Discussion

A methanol extract of the whole plant of *G. japonicum* was successively partitioned into hexane- and ethyl

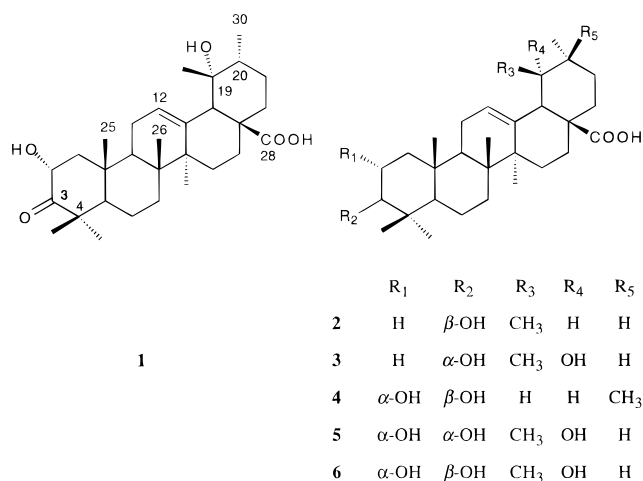


Figure 1. Structures of compounds 1-6.

acetate-soluble fractions. Biological evaluation of the extract and fractions indicated that most of the activity was in the ethyl acetate-soluble fraction. The active EtOAc-soluble fraction was then subjected to silica gel column chromatography to give a new triterpene acid (**1**) together with five known triterpene acids, ursolic acid (**2**), epimolic acid (**3**), maslinic acid (**4**), euscaphic acid (**5**), and tormentic acid (**6**). The structures of these

* To whom correspondence should be addressed.

[†] Department of Biochemistry.

[©] Abstract published in *Advance ACS Abstracts*, May 15, 1996.

compounds were elucidated on the basis of various spectroscopic means.

Compound (**1**) was obtained as a colorless amorphous powder. The infrared (IR) spectrum indicated the presence of hydroxyl group (3600–3200 cm^{-1}), carboxylic acid group (1684 cm^{-1}), and double bond (1650 cm^{-1}). Its molecular formula was determined to be $\text{C}_{30}\text{H}_{46}\text{O}_5$ by high-resolution mass spectroscopy (M^+ 486.3373).

The ^{13}C NMR together with the DEPT spectrum revealed 30 carbon signals including characteristic signals due to a trisubstituted double bond [δ 127.55 (d) and 140.07 (s)], a carboxylic acid group (δ 180.65), a ketone carbonyl group (δ 216.41), and two alcoholic carbons [δ 69.73 (d) and 72.67 (s)]. The ^1H NMR spectrum showed the presence of a hydroxy methine proton at δ 4.80 (1H, dd, $J = 12.5, 6.5$ Hz), an olefinic proton at δ 5.55 (1H, br s), six singlets at δ 1.01 (3H), 1.09 (3H), 1.15 (3H), 1.22 (3H), 1.42 (3H), 1.64 (3H) for six tertiary methyl group, and a doublet at δ 1.12 (3H, d, $J = 6.6$ Hz) for a secondary methyl group. The secondary methyl signal on ring E provides a most useful indicator for the presence of an urs-12-ene skeleton.⁶ In addition, the signals in its ^{13}C NMR at δ 127.55 (d) and 140.07 (s) are characteristic for a C-12/C-13 double bond in the ursene-type structure.⁷ The presence of the ursene skeleton was further verified using 2D NMR spectroscopy including HMBC measurement. The ^1H – ^{13}C long-range COSY and ^1H – ^1H COSY spectrum of this compound gave very good information for establishing the assignment of the partial structure of C-1–C-3 carbons. There were long-range correlations between the H-2 proton at δ 4.80 and the methylene carbon at δ 50.10 (C-1) in the HMBC spectrum and a cross-peak between the signals of the H-2 proton and the H-1 proton in the ^1H – ^1H COSY spectrum, which indicated that the ketone carbonyl group must be either at position C-1 or C-3. Furthermore, the observation of the correlation between the ketone carbonyl group and two methyl carbons (C-23 at δ 25.28 and C-24 at δ 21.77) suggested the ketone must be at the C-3 position. The secondary hydroxyl group was assigned to C-2 due to the correlation between the C-1 methylene carbon at δ 50.10 and the C-25 methyl group. Splitting patterns for the hydroxy methine proton (δ 4.80, dd, $J = 12.5, 6.5$ Hz) showed that the hydroxyl group is equatorial. The stereochemistry of the hydroxyl group was further confirmed by NOE experiments. In particular, on irradiation of the signal of the H-2 proton at δ 4.80, the NOE effect was observed on the H-24 at δ 1.01 and H-25 at δ 1.15 methyl protons. Furthermore, irradiation of the signal of the H-25 methyl protons gave the NOE enhancements for H-26 methyl protons at δ 1.09, H-24 methyl protons, and H-2 proton. These data clearly suggested the equatorial configuration of the C-2 hydroxyl group, as indicated in structure **1**. On the other hand, a signal at δ 3.03 (1H, s, H-18) assigned to the bridgehead H-18 suggested the presence of a 19-O-substituted urs-12-ene skeleton.⁸ The hydroxyl group at C-19 was assigned to the α position by analogy with the ^{13}C NMR chemical shifts of similarly substituted triterpene acids.⁹ The carboxylic acid group was assigned to C-17 on the basis of a HMBC correlation to H-18. Furthermore, the structure of compound **1** was supported by analysis of ion peaks in the mass spectrum. The peaks at m/z 246 [$264 - \text{H}_2\text{O}$]⁺, 220 [$264 -$

Table 1. ^{13}C NMR Spectral Data^a of Compound **1** (δ Relative to TMS in Pyridine- d_5 , 125 MHz)

C	δ	C	δ	C	δ
1	50.10 t ^b	11	24.12 t	21	26.92 t
2	69.73 d	12	127.55 d	22	38.45 t
3	216.41 s	13	140.07 s	23	25.28 q
4	48.10 s	14	42.16 s	24	21.77 q
5	57.67 d	15	29.26 t	25	17.30 q
6	19.55 t	16	26.31 t	26	16.77 q
7	33.14 t	17	48.25 s	27	24.66 q
8	40.37 s	18	54.56 d	28	180.65 s
9	47.33 d	19	72.67 s	29	27.06 q
10	37.90 s	20	42.35 d	30	15.95 q

^a δ in ppm. ^b Multiplicities determined by DEPT sequences.

Table 2. ^1H – ^{13}C Long-Range Correlation Detected in the HMBC Spectrum of Compound **1** (δ Relative to TMS in Pyridine- d_5)

proton ^a	correlated C
1.01 (H-24)	216.41 (C-3), 48.10 (C-4), 57.67 (C-5)
1.09 (H-26)	33.14 (C-7), 40.37 (C-8), 47.33 (C-9), 42.16 (C-14)
1.12 (H-30)	72.67 (C-19), 26.92 (C-21)
1.15 (H-25)	50.10 (C-1), 57.67 (C-5), 47.33 (C-9)
1.22 (H-23)	216.41 (C-3), 48.10 (C-4), 57.67 (C-5)
1.42 (H-29)	54.56 (C-18), 72.67 (C-19), 42.35 (C-20)
1.64 (H-27)	40.37 (C-8), 140.07 (C-13), 42.16 (C-14), 29.26 (C-15)
1.87 (H-9)	50.10 (C-1), 40.37 (C-8), 37.90 (C-10), 17.30 (C-25), 16.77 (C-26)
2.48 (H-1)	69.73 (C-2), 216.41 (C-3), 37.90 (C-10), 17.30 (C-25)
3.03 (H-18)	127.55 (C-12), 140.07 (C-13), 42.16 (C-14), 26.31 (C-16), 48.25 (C-17), 72.67 (C-19), 42.35 (C-20), 180.65 (C-28)
4.80 (H-2)	50.10 (C-1)
5.55 (H-12)	47.33 (C-9), 42.16 (C-14), 54.56 (C-18)

^a δ in ppm.

CO_2^+ , 219 [$264 - \text{CO}_2\text{H}$]⁺, and 218 [$264 - \text{HCO}_2\text{H}$]⁺ corresponded to the typical retro-Diels–Alder (RDA) cleavage fragments commonly found in the spectra of urs-12-ene derivatives possessing a carboxyl group on C-17 and a hydroxyl substituent on ring D or E. Besides, there was a base peak at m/z 146 apparently due to the secondary retro-Diels–Alder cleavage of the m/z 218 fragment ion. These fragmentations resemble very closely those of tormentic acid.¹⁰ Other important ions were at m/z 438, 366, and 220, which indicated that a secondary hydroxyl group and a ketone group were located in the A/B ring portion. The complete assignment of all the carbons was determined (Table 1) with the help of DEPT, ^1H – ^1H COSY, ^1H – ^{13}C COSY, and NOE and followed by analysis of the HMBC spectral data (Table 2). On the basis of all of the above information, compound **1** was found to be a new triterpene acid, and the structure of the natural product was formulated as $2\alpha,19\alpha$ -dihydroxy-3-oxo-12-ursen-28-*oic* acid.

The other known compounds, ursolic acid (**2**),¹¹ epimolic acid (**3**),¹² maslinic acid (**4**),^{11,13} euscaphic acid (**5**),¹¹ and tormentic acid (**6**),¹⁴ were identified by comparing their MS and ^1H and ^{13}C NMR data with those reported in the literature. They were further confirmed by 2D NMR spectra. Among these known triterpene acids, maslinic acid and tormentic acid have previously been reported³ to be present in *G. japonicum*.

All the compounds isolated from the EtOAc-soluble fraction were tested for their anti-HIV-1 protease activity by HPLC assay.¹⁵ $2\alpha,19\alpha$ -Dihydroxy-3-oxo-12-ursen-

28-oic acid (**1**) (72%), ursolic acid (**2**) (85%), and maslinic acid (**4**) (100%) exhibited stronger activity than the others [epipomolic acid (**3**) (42%) and tormetic acid (**6**) (49%), respectively] at the concentration of 17.9 $\mu\text{g/mL}$. Maslinic acid (**4**) was the most active, while eusaphic acid (**5**) (0%) did not show any activity at the tested concentration. This is the first report of anti-HIV-1 activity of this plant. Further studies of structure-activity relationships are currently in progress in our laboratory, but the results of the experiments obtained in the present study indicate that compounds **1**, **2**, and **4** may contribute toward the anti-HIV-1 protease activity of the EtOAc-soluble fraction of *G. japonicum*.

Experimental Section

General Experimental Procedures. Optical rotations were obtained on a Perkin automatic polarimeter. IR spectra were recorded on a Perkin-Elmer 1600 Series FTIR spectrometer. EIMS were recorded on a MICRO-MASS 7035E mass spectrometer at 70 eV. All NMR spectra (^1H , ^{13}C , COSY, NOE, HMQC and HMBC) were recorded on a Bruker AMX 500 spectrometer (500 MHz for ^1H and 125 MHz for ^{13}C) in pyridine- d_5 , the chemical shifts are reported in ppm with TMS as an internal standard, and coupling constants (J) are given in Hz. TLC was done on 0.25 mm silica gel (60 F₂₅₄, Merck) plates.

Plant Material. The whole plant of *G. japonicum* used in this experiment was collected from China in Aug 1994. The material was identified as *G. japonicum* Thunb. by Dr. Dao-Feng Chen, Department of Pharmacology, Shanghai Medical University, China. A voucher specimen is deposited in Department of Pharmacology, Shanghai University of Traditional Chinese Medicine, China.

Extraction and Isolation. Dried whole plants (2.6 kg) were chopped into small pieces and extracted three times with MeOH (20 L) at room temperature for 7 days, and the extract was evaporated in *vacuo* to yield MeOH extract (350 g). The MeOH extract was suspended in distilled water (1 L) and successively extracted with hexane (2 L \times 5) and EtOAc (2 L \times 5). The EtOAc-soluble fraction was filtered, and the filtrate was evaporated under reduced pressure to give a brown oil (120 g).

The EtOAc extract (100 g) was subjected to silica gel column chromatography using a linear gradient CHCl_3 -MeOH system. The fractions with CHCl_3 -MeOH 9:1-7:3 were further purified by repeated silica gel column chromatography and/or preparative TLC to give $2\alpha,19\alpha$ -dihydroxy-3-oxo-12-ursen-28-oic acid (**1**) (44 mg), 3β -hydroxy-12-ursen-28-oic acid (**2**) (70 mg), $3\alpha,19\alpha$ -dihydroxy-12-ursen-28-oic acid (**3**) (23 mg), $2\alpha,3\beta$ -dihydroxy-12-oleanen-28-oic acid (**4**) (30 mg), $2\alpha,3\alpha,19\alpha$ -trihydroxy-12-ursen-28-oic acid (**5**) (50 mg), and $2\alpha,3\beta,19\alpha$ -trihydroxy-12-ursen-28-oic acid (**6**) (20 mg).

$2\alpha,19\alpha$ -Dihydroxy-3-oxo-12-ursen-28-oic acid (1**):** colorless amorphous powder, $[\alpha]_D^{25} +1.7^\circ$ (c 1.0, EtOH); mp 215-217 $^\circ\text{C}$; HREIMS 486.3373 (calcd for $\text{C}_{30}\text{H}_{46}\text{O}_5$, 486.3345); EIMS m/z 486 (M^+ , 0.8), 438 (5.8), 366 (34.6), 246 (10.7), 219 (9.9), 185 (32.5), 146 (100); IR KBr (cm^{-1}) 3600-3200, 2928, 2869, 1684, 1650, 1384, 1105; ^1H NMR (pyridine- d_5 , 500 MHz) δ 1.01 (3H, s, 24- H_3), 1.09 (3H, s, 26- H_3), 1.12 (3H, d, $J=6.6$ Hz, 30- H_3), 1.15 (3H, s, 25- H_3), 1.22 (3H, s, 23- H_3), 1.25 (1H, m, 5-H), 1.37 (1H, m, 1- H_a), 1.42 (3H, s, 29- H_3), 1.50 (1H,

m, 20-H), 1.64 (3H, s, 27- H_3), 1.87 (1H, t, $J=8.9$ Hz, 9-H), 2.08 (1H, m, 21- H_a), 2.11 (1H, m, 22- H_a), 2.13 (1H, m, 21- H_b), 2.15 (1H, m, 22- H_b), 2.48 (1H, dd, $J=12.5$, 6.5 Hz, 1- H_b), 3.03 (1H, s, 18-H), 4.80 (1H, dd, $J=12.5$, 6.5 Hz, 2-H), 5.55 (1H, br s, 12-H); ^{13}C NMR (pyridine- d_5 , 125 MHz) see Table 1.

Biological Assay. Recombinant HIV-1 protease was obtained from the expression vector, PGEX-PR 107, in *Escherichia coli* DH5 α and was purified according to the method established in our laboratory.¹⁵

The proteolytic activity of HIV-1 PR was measured using the synthetic heptapeptide Ser-Gln-Asn-Tyr-Pro-Ile-Val (SQNYPIV), corresponding to the p24-p17 cleavage site in the natural *gag* precursor, as substrate. The products of cleavage were analyzed by reversed-phase high-performance liquid chromatography (RP-hplc, Ultrasphere ODS, 5 μm , 4.6 mm \times 15 cm, Beckman, CA) with a 7-33% acetonitrile gradient in 0.1% trifluoroacetic acid (TFA). Product peak areas were integrated by a data system (Chromatopac C-R3A, Shimadzu, Japan) and compared with the standard peptide SQNY, which is expected to be generated from heptapeptide SQNYPIV after HIV-1 protease digestion. The extinction coefficient of the peptide SQNY was used to estimate the rate of proteolysis of the heptapeptide substrate. The assay was performed in a volume of 140 μL containing 90 pmol of substrate, 2 U of HIV-1 protease, and compound solution in 0.1 M NaOAc buffer, pH 5.5, with 1 M $(\text{NH}_4)_2\text{SO}_4$, at 37 $^\circ\text{C}$. The reaction, which proceeded for 2 h, was terminated by the addition of 20 μL of aqueous 10% trifluoroacetic acid. The supernatant obtained by centrifugation at 12 000 rpm for 3 min was then analyzed by the HPLC method. The unit of enzyme activity was defined as the amount of enzyme which yields 1 nmol of tetrapeptide SQNY under the above conditions in 1 min.

Proteolytic activity of HIV-1 protease inhibited by the compound was calculated as the activity of the control subtracting that of the sample and then divided by itself.

Acknowledgment. We are grateful to the National University of Singapore for a grant in support of this research.

References and Notes

- Yoshida, T.; Okuda, T.; Memon, M. U.; Shingu, T. *J. Chem. Soc., Perkin Trans. 1* **1985**, 315.
- Yoshida, T.; Okuda, T.; Memon, M. U.; Shingu, T. *Phytochemistry* **1985**, *24*, 1041.
- Shigenaga, S.; Kouno, I.; Kawano, N. *Phytochemistry* **1985**, *24*, 115.
- Xu, H. X.; Kadota, S.; Wang, H.; Kurokawa, M.; Shiraki, K.; Masumoto, T.; Namba, T. *Heterocycles* **1994**, *38*, 167.
- Xu, H. X.; Wan, M.; Loh, B. N.; Kon, O. L.; Chow, P. W.; Sim, K. Y. *Phytother. Res.* **1996**, in press.
- Johns, S. R.; Lamberton, J. A.; Morton, T. C.; Soares, H.; Willing, R. I. *Aust. J. Chem.* **1983**, *36*, 2537.
- Durham, D. G.; Liu, X.; Richards, R. M. E. *Phytochemistry* **1994**, *36*, 1469.
- Yoshioka, I.; Sugawara, T.; Kitagawa, I. *Chem. Pharm. Bull.* **1971**, *19*, 1700.
- Seto, T.; Tanaka, T.; Tanaka, O.; Naruhashi, N. *Phytochemistry* **1984**, *23*, 2829.
- Takahashi, K.; Ogura, M.; Tanabe, Y. *Chem. Pharm. Bull.* **1969**, *17*, 2223.
- Seo, S.; Tomita, Y.; Tori, K. *J. Chem. Soc., Chem. Commun.* **1975**, 270, 954.
- Rogelio, P. M.; Mariano, G. F. *J. Nat. Prod.* **1988**, *51*, 996.
- Kojima, H.; Ogura, H. *Phytochemistry* **1989**, *28*, 1703.
- Ojinaka, C. M.; Okogun, J. I.; Okorie, D. A. *Phytochemistry* **1980**, *19*, 2482.
- Wan, M.; Loh, B. N. *Biochem. Mol. Int.* **1995**, *35*, 899.