JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

New Sweet-tasting C21-pregnane Glycosides from Myriopteron extensum (Wight) K. Schum.

Guo Sun, Qin Dai, Hongxia Zhang, ZhiJian Li, and ZhiZhi Du

J. Agric. Food Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jafc.6b04121 • Publication Date (Web): 19 Nov 2016

Just Accepted

ACS Publications

Journal of Agricultural and Food Chemistry is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

New Sweet-Tasting C21-Pregnane Glycosides from *Myriopteron extensum* (Wight) K. Schum.

Guo Sun[†,‡,▽], Qin Dai[†,‡,▽], Hong-Xia Zhang[†], Zhi-Jian Li[†], Zhi-Zhi Du*[†]

[†] Key Laboratory of Economic Plants and Biotechnology, Yunnan Key Laboratory for Wild Plant Resources, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, Yunnan, China

[‡] University of Chinese Academy of Sciences, Beijing 100049, China

Corresponding Author

* Phone: +86-871-65223224; Fax: +86-871-65216335

E-mail address: duzhizhi@mail.kib.ac.cn

ABSTRACT

Ten novel C21 pregnane glycosides, extensumside C–L (**1**–**10**), were isolated as highly sweet-tasting substances from the edible pericarps of *Myriopteron extensum* (Wight) K. Schum by sensory-guided fractionation and purification. Their structures were determined through 1D and 2D NMR, such as HSQC, HMBC, ^{1}H-^{1}H COSY, HSQC-TOCSY and ROESY, as well as other spectroscopic analysis combined with chemical evidences. These compounds shared the same aglycone, 3β,16α-dihydroxy-preg-5-en-20-one, and contained the deoxysugar chain and the glucose chain which were linked to C-3 and C-16 of the aglycone, respectively. The sweetness potency was evaluated by a human sensory panel test and preliminary structure-taste relationship was discussed. The sweetness intensities of these compounds are between 50 to 400 times greater than that of sucrose. Furthermore, quantitation analyses of compounds **1**, **3**, **4**, and **6** in different parts of *M. extensum* indicated that the concentrations of these sweet components in the pericarps are obviously higher than those in stems and roots.

KEYWORDS

C21-pregnane glycosides, highly sweet natural products, sensory evaluation, sweet intensity

INTRODUCTION

In our daily life, we can often access to all kinds of delicacy. The color, smell and taste of the food may directly have an important impact on our appetite. Human have five basic taste sensations at least, including sweet, umami, bitter, salty and sour taste,[1] of which sweet, umami and bitter can affect the acceptance of people to some food to a great extent.[2] Sweet agents are important substances in our daily life, and they include sugar and non-sugar. It is useful for us to consume some sugar substances appropriately, while it can harm our body if overdosed. Nowadays, more and more people have some diseases,[3] such as dental caries, hypertension, hyperglycemia, cardiovascular diseases and obesity, which are more or less related to excessive intake of sugar.[4-6] According to the data published in *New England Journal of Medicine* in 2010, 92.4 million adults suffer from diabetes in China, which has the most people suffering from diabetes in the world.[7] To avoid the health problem caused by excessive intake of sugar substances, people turn their attention to developing non-sugar sweeteners and low-calorie sweet agents. Most of the current available low-calorie sweeteners in the world market are artificial synthetic sweeteners, such as acesulfame-K, alitame, aspartame, cyclamate, saccharin, and sucralose.[8] Some of these artificial sweeteners are approved for use as food additives in the USA and the European Union, but debates in terms of their safety, stability, and quality of taste have arisen in recent years.[9] These impel people to turn their attention to developing natural non-sugar sweeteners. By far, all the known natural sweet-tasting substances are derived from green plant, as opposed to lower plants, microbial, or marine

organism.[10-12] Although many natural products derived from plants have a sweet taste, only few of them have either been commercialized as sweeteners or flavors, or are under development, such as rebaudioside A, stevioside, steviol glycosides, mogrol glycosides (Luo Han Guo sweetener), morgroside V, glycyrrhizic acid, thaumatin, brazzein, and monatin.[9,13] It is still very necessary to search and discover naturally occurring non-sugar and high-potency molecules.

In our ethnobotanical investigation and sensory evaluation of characteristic edible plants in the Yunnan province of China, a plant called "shankugua" was found to exhibit obvious sweet taste. The Yao people in the Xinpin county of Yunnan province utilize the fruits of this plant to make pickles. This plant was identified taxonomically as *Myriopteron extensum* (Wight) K. Schum, a single species of Asclepiadaceae. It is grown in thickets and open woods of 600-1600 meters above sea level and widely distributed in Guangxi, Guizhou, Yunnan of China as well as in India, Indonesia, Laos, Myanmar, Thailand, Vietnam.[14] The roots of *M. extensum* can be used as medicine for the treatment of pulmonary tuberculosis and cough.[14] Only few researches on the chemistry and pharmacological activity of *M. extensum* are reported up to date. Extensumsides A and B,[15] a C21 pregnane glycoside and a cardiac glycoside, with some triterpenoids and several known steroids were isolated and identified from the stems of *M. extensum*, and the cytotoxic activity of EtOAc extract and n-Butanol extract were also tested.[16] A novel flavanoid, Lupinifolin, was identified from the stems of *M. extensum.*[17] Three new dinormonoterpenoid glucosides were reported from the pericarps of *M. extensum*.[18] However, there are no reports on the

sweet-tasting components from *M. extensum*.

To clarify the main chemical constituents responsible for the sweet taste of the pericarps of *M. extensum*, we did the phytochemical investigations on the plant. Ten novel C21 pregnane glycosides were isolated and identified as highly sweet compounds and they were proposed to the main sweet-tasting constituents from the pericarps of *M. extensum*

MATERIALS AND METHODS

Chemicals. The following materials were used: AR Grade Methanol (Huada, Guangzhou, China), L-cysteine methyl hydrochloride (Sigma-aldrich, Shanghai, China), N-trimethylsilylimidazole (Sangon Biotech, Shanghai, China), n-hexane (Damao, Tianjin, China), D-(+)-glucose (J & K Scientific, Guangzhou, China), HPLC acetonitrile (Merck, Shanghai, China), Ethyl acetate (Jige, Tianjin, China), Acetic anhydride (Damao, Tianjin, China), Sulphuric acid (Xilong Chemical Co. Ltd, Guangdong, China), Hydrochloric acid (Xilong Chemical Co. Ltd), Ferric chloride (Damao, Tianjin, China), Chloroform (Rionlon, Tianjin, China), Dioxane (Sinopharm chemical reagent Co. Ltd., Shanghai, China), Sodium dicarbonate (Damao, Tianjin, China).

General Experimental Procedures. Optical rotations were measured with a Horiba Sepa-300 polarimeter (Horiba, Tokyo, Japan). UV spectra were obtained using a Shimadzu UV-2401A spectrophotometer (Shimadzu, Tokyo, Japan). A Bruker tensor 27 spectrophotometer (Bruker, Bremen, Germany) was used for scanning IR

spectroscopy using KBr pellets. 1D and 2D NMR spectra were recorded on Bruker Avance III 600 spectrometer (Bruker, Bremen, Germany) at 298 K. Unless otherwise specified, chemical shifts (δ) were expressed in ppm with reference to the solvent signals. ESIMS were obtained on a Bruker HTC/Esquire spectrometer (Bruker, Bremen, Germany). HRESIMS were recorded on an Agilent G6230 TOF MS spectrometer (Agilent Technologies, Santa Clara, America).

Column chromatography (CC) was done using silica gel (200–300 mesh, Qingdao Marine Chemical Ltd.Co., China), RP-18 silica gel (40-63μm, Merck, Germany), Diaion HP-20 (Mitsubishi Chemical Corporation, Japan) and Sephadex gel LH-20 (GE healthcare bio-sciences AB, Sweden). TLC was performed on silica gel GF254 (Qingdao Marine Chemical Ltd. Co., China), and spots were visualized by heating silica gel plates sprayed with 10% H_2SO_4 in ethanol. HPLC analyses and separations were performed on an Agilent HP-1100 system (Agilent Technologies, Santa Clara, America) equipped with a quaternary gradient pump (G1311A) and a multiwavelength detector (G1314A VWD). The lyophilizer (Virtis Benchtop K, America) was used to dry the samples and eliminate the residual solvents.

Plant Material. The fruits of *Myriopteron extensum* (Wight) K. Schum were collected from Xinpin county of Yunnan province, southwest of China, and identified by Professor Lisong Wang. A voucher specimen (KUN 0309000) was deposited in Herbarium of Kunming Institute of Botany, Chinese Academy of Sciences.

Extraction of Plant Material. The air-dried and powdered pericarps of *M. extensum* (1.4 kg) were extracted with methanol then 60% aqueous methanol three

111 times at room temperature (15 L × 3 days × 3 times), respectively. The filtrates were
112 combined and concentrated under reduced pressure until elimination of methanol to
113 obtain crude extracts. Crude extracts were made up with water to 1 L, and then
114 extracted with petroleum ether (1 L × 5) to yield petroleum ether portion (A, 24.5 g).
115 The emulsifying layer produced in the extraction process and the residual aqueous
116 layer were concentrated to eliminate solvent and freeze-dried to get the emulsified
117 portion (B, 45.2 g) and aqueous portion (C, about 300 g), respectively. Extract C was
118 identified as the sweet-tasting portion by sensory experiments.

119 **Bioassay-Guided Fractionation and Purification of the Sweet-Tasting**
120 **Compounds.** Extract C was subjected to macro-porous resins column
121 chromatography (Diaion HP-20, 2 kg) and eluted with water, 30%, 50%, 70%
122 aqueous methanol and methanol (6 L each). Every elutant was evaporated to remove
123 organic solvent in reduced pressure and lyophilized to give fraction C1 (about 250 g,
124 eluted by water), fraction C2 (21 g, eluted by 30%, 50%, 70% aqueous methanol and
125 combined), fraction C3 (22 g, eluted by methanol). Fraction C3 was identified as the
126 sweet-tasting fraction according to sensory evaluation results.

127 Fraction C3 was separated by column chromatography on silica gel with
128 chloroform-methanol solvent system as eluent. After gradient elution with C (2000
129 mL), C/M (50:1, 2500 mL), C/M (20:1, 2500 mL), C/M (10:1 + 0.5% W, 2500 mL),
130 C/M (5:1 + 1% W, 2500 mL), C/M (2:1 + 2% W, 2500 mL), M (2500 mL) (C:
131 chloroform, M: methanol, W: water), 9 fractions were obtained on the basis of TLC
132 analysis and each was freeze-dried to use for sensory experiments. According to

sensory evaluation, fractions C3-1 to C3-3 were non-sweet fractions, while fractions C3-4 to C3-8 were sweet-tasting fractions.

Fraction C3-4 (2.20 g) was chromatographed on RP-18 MPLC system (BUCHI, φ 4.0 × 25 cm; 220 g RP-18 silica gel, Merck, Germany) with MeOH-H_2O as eluent in a gradient mode (50–100% aqueous methanol, 2000 mL each eluent) to give five subfractions, of which subfraction C3-4-3 (350 mg) was screened as sweet fraction through sensory evaluation. The fraction was subjected to Sephadex gel LH-20 column chromatography, followed by semi-preparative HPLC on a ZORBAX SB-C18 column (9.4 × 250 mm, flow rate 2.5 mL/min) with acetonitrile-water (45:55) to yield **1** (115 mg, t_R = 13.28 min) and **2** (123 mg, t_R = 16.13 min). Fraction C3-5 (2.70 g) was chromatographed on RP-18 (MPLC system, BUCHI), Sephadex gel LH-20, respectively, and then purified by semi-preparative HPLC on a ZORBAX SB-C18 column (9.4 × 250 mm, flow rate 2.5 mL/min) with acetonitrile-water (45:55) to yield **3** (142 mg, t_R = 10.89 min). Fraction C3-6 (3.10 g) was chromatographed on RP-18 (MPLC system, BUCHI), Sephadex gel LH-20, respectively, followed by semi-preparative HPLC on a ZORBAX SB-C18 column (9.4 × 250 mm, flow rate 2.5 mL/min) with acetonitrile-water (45:55) to afford **4** (138 mg, t_R = 8.65 min) and **5** (139 mg, t_R = 11.34 min). Fraction C3-7 (3.40 g) was chromatographed on RP-18 (MPLC system, BUCHI), Sephadex LH-20 gel, respectively, and then isolated by semi-preparative HPLC on a ZORBAX SB-C18 column (9.4 × 250 mm, flow rate 2.5 mL/min) with acetonitrile-water (40:60) to afford **6** (266 mg, t_R = 10.35 min) and **7** (115 mg, t_R = 13.11 min). Fraction C3-8 (2.50 g) was chromatographed on RP-18

(MPLC system, BUCHI), Sephadex LH-20 gel, respectively, and then isolated by semi-preparative HPLC on a ZORBAX SB-C18 column (9.4 × 250 mm, flow rate 2.5 mL/min) with acetonitrile-water (40:60) to yield **8** (125 mg, t_R = 10.67 min), **9** (118 mg, t_R = 12.51 min) and **10** (113 mg, t_R = 15.60 min).

Extensumside C (**1**): white amorphous powder; $[\alpha]_D^{24}$ +0.6 (c = 0.10, MeOH); UV (MeOH) λ_{max} (nm) (log ε): 217 (3.98); IR (KBr) ν_{max} cm^{-1}: 3443, 2930, 1694, 1633, 1384, 1060; ^1H and ^{13}C NMR data, see table 1; HREI-MS *m/z* 1186.6190 (calcd. for $C_{59}H_{94}O_{24,}$ 1186.6135); ESI-MS/MS *m/z* 1209 $[M + Na]^+$, 1047 $[M + Na – Glc]^+$, 983 $[M + Na – Cym – mBe]^+$, 839 $[M + Na – Cym – mBe – Cym]^+$, 821 $[M + Na – Cym – mBe – Glc]^+$, 677 $[M + Na – Cym – mBe – Cym – Glc]^+$.

Extensumside D (**2**): white amorphous powder; $[\alpha]_D^{24}$ –39.0 (c = 0.10, MeOH); UV (MeOH) λ_{max} (nm) (log ε): 217 (3.91); IR (KBr) ν_{max} cm^{-1}: 3445, 2934, 1721, 1642, 1382, 1079; ^1H and ^{13}C NMR data, see table S1 to S4; HRESI-MS *m/z* 1209.6018 $[M + Na]^+$ (calcd. for $C_{59}H_{94}O_{24}Na$, 1209.6027). ESI-MS/MS *m/z* 1209 $[M + Na]^+$, 983 $[M + Na – Ole – mBe]^+$, 839 $[M + Na – Ole – mBe – Cym]^+$, 821 $[M + Na – Ole – mBe – Glc]^+$, 677 $[M + Na – Ole – mBe – Cym – Glc]^+$.

Extensumside E (**3**): white amorphous powder; $[\alpha]_D^{24}$ –6.9 (c = 0.10, MeOH); UV (MeOH) λ_{max} (nm) (log ε): 217 (3.96); IR (KBr) ν_{max} cm^{-1}: 3424, 2933, 1702, 1641, 1384, 1079; ^1H and ^{13}C NMR data, see table S1 to S4; HRESI-MS *m/z* 1371.6569 $[M + Na]^+$ (calcd. for $C_{65}H_{104}O_{29}Na$, 1371.6561). ESI-MS/MS *m/z* 1371 $[M + Na]^+$, 1145 $[M + Na – Cym – mBe]^+$, 1001 $[M + Na – Cym – mBe – Cym]^+$, 1209 $[M + Na – Glc]^+$, 1047 $[M + Na – Glc – Glc]^+$, 983 $[M + Na – Cym – mBe – Glc]^+$, 839 $[M + Na$

– Cym – mBe – Cym – Glc]$^+$, 677 [M + Na – Cym – mBe – Cym – Glc – Glc]$^+$.

Extensumside F (**4**): white amorphous powder; $[\alpha]_D^{24}$ –7.1 (c = 0.10, MeOH); UV (MeOH) λ_{max} (nm) (log ε): 217 (3.95); IR (KBr) ν_{max} cm^{-1}: 3425, 2933, 1701, 1637, 1384, 1060; ^{1}H and ^{13}C NMR data, see table S1 to S4; HRESI-MS *m/z* 1371.6565 [M + Na]$^+$ (calcd. for $C_{65}H_{104}O_{29}Na$, 1371.6561). ESI-MS/MS *m/z* 1371 [M + Na]$^+$, 1145 [M + Na – Cym – mBe]$^+$, 1001 [M + Na – Cym – mBe – Cym]$^+$, 1209 [M + Na – Glc]$^+$, 1047 [M + Na – Glc – Glc]$^+$, 839 [M + Na – Cym – mBe – Cym – Glc]$^+$.

Extensumside G (**5**): white amorphous powder; $[\alpha]_D^{24}$ –34.2 (c = 0.11, MeOH); UV (MeOH) λ_{max} (nm) (log ε): 218 (3.87); IR (KBr) ν_{max} cm^{-1}: 3445, 2932, 1642, 1383, 1061; ^{1}H and ^{13}C NMR data, see table S1 to S4; HRESI-MS *m/z* 1371.6483 [M + Na]$^+$ (calcd. for $C_{65}H_{104}O_{29}Na$, 1371.6555). ESI-MS/MS *m/z* 1371 [M + Na]$^+$, 1145 [M + Na – Ole – mBe]$^+$, 1001 [M + Na – Ole – mBe – Cym]$^+$, 1209 [M + Na – Glc]$^+$, 1047 [M + Na – Glc – Glc]$^+$, 839 [M + Na – Ole – mBe – Cym – Glc]$^+$.

Extensumside H (**6**): white amorphous powder; $[\alpha]_D^{24}$ –9.5 (c = 0.10, MeOH); UV (MeOH) λ_{max} (nm) (log ε): 217 (4.06); IR (KBr) ν_{max} cm^{-1}: 3440, 2933, 1700, 1633, 1383, 1061; ^{1}H and ^{13}C NMR data, see table S1 to S4; HRESI-MS *m/z* 1533.7057 [M + Na]$^+$ (calcd. for $C_{71}H_{114}O_{34}Na$, 1533.7084). ESI-MS/MS *m/z* 1533 [M + Na]$^+$, 1307 [M + Na – Cym – mBe]$^+$, 1163 [M + Na – Cym – mBe – Cym]$^+$, 1371 [M + Na – Glc]$^+$, 1209 [M + Na – Glc – Glc]$^+$, 1047 [M + Na – Glc – Glc – Glc]$^+$, 1145 [M + Na – Cym – mBe – Glc]$^+$, 1001 [M + Na – Cym – mBe – Cym – Glc]$^+$, 839 [M + Na – Cym – mBe – Cym – Glc – Glc]$^+$, 677 [M + Na – Cym – mBe – Cym – Glc – Glc – Glc]$^+$.

Extensumside I (**7**): white amorphous powder; $[\alpha]_D^{24}$ –12.1 (c = 0.10, MeOH); UV (MeOH) λ_{max} (nm) (log ε): 216 (3.92); IR (KBr) ν_{max} cm^{-1}: 3443, 2927, 1721, 1641, 1384, 1079; ^1H and ^{13}C NMR data, see table S1 to S4; HRESI-MS *m/z* 1533.7049 [M + Na]$^+$ (calcd. for $C_{71}H_{114}O_{34}Na$, 1533.7084). ESI-MS/MS *m/z* 1533 [M + Na]$^+$, 1307 [M + Na – Ole – mBe]$^+$, 1163 [M + Na – Ole – mBe – Cym]$^+$, 1371 [M + Na – Glc]$^+$, 1209 [M + Na – Glc – Glc]$^+$, 1047 [M + Na – Glc – Glc – Glc]$^+$, 1001 [M + Na – Ole – mBe – Cym – Glc]$^+$, 839 [M + Na – Ole – mBe – Cym – Glc – Glc]$^+$, 677 [M + Na – Ole – mBe – Cym – Glc – Glc – Glc]$^+$.

Extensumside J (**8**): white amorphous powder; $[\alpha]_D^{24}$ –32.2 (c = 0.11, MeOH); UV (MeOH) λ_{max} (nm) (log ε): 218 (3.92); IR (KBr) ν_{max} cm^{-1}: 3445, 2930, 1642, 1382, 1064; ^1H and ^{13}C NMR data, see table S1 to S4; HRESI-MS *m/z* 1519.6913 [M + Na]$^+$ (calcd. for $C_{70}H_{112}O_{34}Na$, 1519.6927). ESI-MS/MS *m/z* 1519 [M + Na]$^+$, 1293 [M + Na – Cym – mBe]$^+$, 1163 [M + Na – Cym – mBe – Dig]$^+$, 1357 [M + Na – Glc]$^+$, 1195 [M + Na – Glc – Glc]$^+$, 1033 [M + Na – Glc – Glc – Glc]$^+$, 969 [M + Na – Cym – mBe – Glc – Glc]$^+$, 1001 [M + Na – Cym – mBe – Dig – Glc]$^+$, 839 [M + Na – Cym – mBe – Dig – Glc – Glc]$^+$, 677 [M + Na – Cym – mBe – Dig –Glc – Glc – Glc]$^+$, 679 [M + Na – Cym – mBe – Dig – The – Glc – Glc]$^+$.

Extensumside K (**9**): white amorphous powder; $[\alpha]_D^{24}$ –16.7 (c = 0.10, MeOH); UV (MeOH) λ_{max} (nm) (log ε): 217 (3.98); IR (KBr) ν_{max} cm^{-1}: 3443, 2930, 1633, 1384, 1060; ^1H and ^{13}C NMR data, see table S1 to S4; HRESI-MS *m/z* 1519.6910 [M + Na]$^+$ (calcd. for $C_{70}H_{112}O_{34}Na$, 1519.6927). ESI-MS/MS *m/z* 1519 [M + Na]$^+$, 1293 [M + Na – Ole – mBe]$^+$, 1163 [M + Na – Ole – mBe – Dig]$^+$, 1357 [M + Na – Glc]$^+$,

221 1195 [M + Na – Glc – Glc]$^+$, 1033 [M + Na – Glc – Glc – Glc]$^+$, 969 [M + Na – Ole –
222 mBe – Glc – Glc]$^+$, 1001 [M + Na – Ole – mBe – Dig – Glc]$^+$, 839 [M + Na – Ole –
223 mBe – Dig – Glc – Glc]$^+$, 677 [M + Na – Ole – mBe – Dig – Glc – Glc – Glc]$^+$, 679
224 [M + Na – Ole – mBe – Dig – The – Glc – Glc]$^+$.

225 Extensumside L (**10**): white amorphous powder; $[\alpha]_D^{24}$ –12.5 (*c* = 0.10, MeOH); UV
226 (MeOH) λ_{max} (nm) (log ε): 217 (3.98); IR (KBr) ν_{max} cm^{-1}: 3443, 2930, 1633, 1384,
227 1060; ^{1}H and ^{13}C NMR data, see table S1 to S4; HRESI-MS *m/z* 1519.6917 [M +
228 Na]$^+$ (calcd. for $C_{70}H_{112}O_{34}Na$, 1519.6927).

229 **Acid Hydrolysis of Total Steroidal Glycosides and Compounds 1-10.** 2 g of
230 fraction C3 was refluxed with 1M HCl (dioxane-H_2O, 1:1, 100 mL) at 95 ℃ for 2
231 hours. After cooling and evaporated to dryness in a vacuum, the residue was
232 suspended in water and extracted with EtOAc for three times. The aqueous layer was
233 neutralized with $NaHCO_3$ and evaporated to dryness under reduced pressure to give
234 the sugar residue.[19] The sugar residue was chromatographed on a silica gel column
235 with $CHCl_3$-MeOH-H_2O (7:1:0.1 to 7:3:0.5) as gradient eluents to obtain D-Cymarose,
236 D-Digitoxose, D-Oleandrose, D-Thevetose. Their structures were assigned to be the
237 D-series on the basis of ^{1}H and ^{13}C NMR data as well as their optical rotation values[20]:
238 D-Cymarose, $[\alpha]_D^{24}$ +55.3 (*c* 0.12, H_2O); D-Digitoxose, $[\alpha]_D^{24}$ +44.9 (*c* 0.13, H_2O);
239 D-Oleandrose, $[\alpha]_D^{24}$ –23.1 (*c* 0.11, H_2O); D-Thevetose, $[\alpha]_D^{24}$ +38.2 (*c* 0.10, H_2O).

240 2 mg of each compound (compounds **1**–**10**) was refluxed with 1 M HCl
241 (dioxane-H_2O, 1:1, 1 mL) at 95 ℃ for 2 hours. After cooling and evaporated to
242 dryness in a vacuum, the residue was suspended in water and extracted with EtOAc

for three times. The aqueous layer was neutralized with $NaHCO_3$ and evaporated to dryness under reduced pressure to give the sugar residue.[19] All sugars of compounds **1**–**10** were identified by co-chromatography of the sugar solutions with the authentic samples on the TLC plate using $CHCl_3$-MeOH-H_2O (7:2:1, bottom layer) and EtOAc-MeOH-H_2O (9:1:0.1, top layer) as developing solvents, respectively.[20]

GC Analysis for Determination of Glucose Absolute Configuration. A portion of each sugar residue above was dissolved in anhydrous pyridine (1 mL), to which 2 mg of L-cysteine methyl hydrochloride was added. The mixture was stirred at 60 °C for 2 h, and after evaporation in vacuo to dryness, 0.2 mL of N-trimethylsilylimidazole was added; the mixture was kept at 60 °C for another 2 h.[21] The reaction mixture was partitioned between n-hexane and H_2O (2 mL each), and the n-hexane extract was analyzed by Agilent 7890A gas chromatography with a flame ionization detector (FID) to identify the derivative of D-(+)-glucose. A HP-5 capillary column (5% phenyl methyl siloxan, 50 m × 0.32 mm, 0.52 μm film thickness) was used with helium as carrier gas (1.28 mL/min). Oven program: initial temperature 160 °C, then raised to 280 °C at 5 °C/min. The injector and detector temperature were held constant at 250 °C. The retention time of D-(+)-glucose standard is 21.7 min.

Sensory Evaluation of Subfractions of Extract and Compounds 1-10. More than 30 employees and graduate students from Kunming Institute of Botany (KIB) were invited for screening of sensory sensitivity on sweetness using Givaudan's panelist selection procedure (taste intensity ranking test). A total of 10 people were selected to form a taste panel and trained following ISO norms.[22] A 0.2% (w/v) water

solution of each elutant obtained from chromatographic isolation was tested by taste panel to find the sweet subfractions. The sweetness intensities relative to sucrose of pure compounds were evaluated by the sensory panel consisting of seven sweet sensitive tasters (four women and three men, age from 24 to 45, Chinese only) as previously described[23] and modified as needed. All glycosides were dissolved in water in graduated concentrations from 0.02 to 0.002% (w/v), while sucrose solutions were made at concentrations of 1%, 2% and 4%. The relative sweetness of compounds compared to a 1–4% solution (w/v) of sucrose was determined by tasting its solutions at different concentrations and selecting the concentration at which the taste was approximately closest to that of the sucrose solution.

Samples were coded and randomly presented to panelists of 10 mL in each cup and the total cups less than 20 at ambient temperature. The panelists were asked to rinse their mouths with water in between samples and rest for some time after tasting several cups of samples. The assays were performed at least in triplicate on separate occasions.

Each sample was dissolved in water to make a stock solution. Dilution solutions of the samples were presented in an order of ascending concentrations (0.010, 0.015, 0.020, 0.025, 0.033, 0.050, 0.100 mg/mL). The panelists were asked to taste the sample solutions until they cannot percept sweet taste from the sample solution. The concentration of the second last tested sample solution was taken as the individual recognition threshold of this compound.

Quantitation by Analytical HPLC-DAD.

An Agilent HP-1100 HPLC system (Agilent, America) with quaternary gradient unit (G1311A), with degasser (G1322A), autosampler (G1313A), column oven (G1316A), and multiwavelength detector (G1314A VWD) was used for quantitation of C21 pregnane glycosides. A series of external standard solutions were prepared using isolated authentic references (**1**, **3**, **4**, and **6**) to make working curves. Chromatography separation was performed on a stainless steel column (ZORBAX SB-C18, 250 × 4.6 mm I.D., 5 μm; Agilent, America) using a flow rate of 1.0 mL/min with acetonitrile-water as mobile phase. All samplers were analyzed using a gradient system with the detector wavelength of 217 nm and oven temperature at 30 ℃. The gradient program was as follows: the samples were injected at 20% acetonitrile-water (A/W). The gradient was changed to 30% A/W within 5 min, and then to 40% A/W from 5 min to 15 min, to 50% A/W from 15 min to 35 min, to 100% A from 35 min to 40 min, respectively. Within 5 min the proportion of the eluent was changed to starting conditions and equilibrated for 10 min.

Three different parts (pericarps, stems, and roots) of *M. extensum* were dried and powdered to use for subsequent experiments. The powdered materials of the three parts were weighed 10.0 g precisely to the conical flasks (250 mL) respectively, to which 100 mL methanol was added. The samples were extracted for 30 min under the ultrasonic conditions and the supernatants were filtered through a 0.45 μm filter membrane to make tested solutions used for quantitation of the sweet C21 pregnane glycosides in the botanical materials.

RESULTS AND DISCUSSION

Isolation and Elucidation of C21 Pregnane Glycosides. The MeOH and 60% MeOH extract of dried pericarps of *M. extensum* was partitioned between H_2O and petroleum ether. The H_2O fraction containing a delicate sweet taste was submitted to repeated chromatography by macroporous resin, normal-phase silica gel, reverse-phase C18 silica gel, and Sephadex gel LH-20 respectively, and finally purified by semi-preparative and preparative HPLC to yield compounds **1-10** (Figure 1).

Extensumside C (**1**) was obtained as a white amorphous powder. The molecular formula of **1** was deduced as $C_{59}H_{94}O_{24}$, with thirteen degrees of unsaturation, on the basis of HR-EI-MS experiment with a quasi-molecular ion peak at *m/z* 1186.6190 (calcd. for $C_{59}H_{94}O_{24}$, 1186.6135). The typical absorption peaks at 3445, 1694, 1633 cm^{-1} in the IR spectrum, suggested the existence of hydroxyl group, carbonyl group, and olefinic group. Compound **1** showed positive Liebermann-Burchard[24] and Keller-Kiliani reactions,[25] suggesting that it belonged to steroidal 2-deoxyglycoside. Detailed analysis of the carbon signals for the aglycone moiety of **1** in the ^{13}C NMR spectrum revealed that it was in good agreement with those of stelmatocryptonoside D (**1a**) (Figure 3),[26] which was isolated from *Stelmatocrypton khasianum*, except that the carbon signal of C-3 at 71.3 in stelmatocryptonoside D (**1a**) was shifted downfield to δ 78.1 in **1** due to the glycosidation of 3-OH. This was ultimately confirmed by 1D and 2D NMR spectra. In HMBC spectrum, the correlations between H19 (0.80, s) and C1(37.4), C9 (50.1), C10 (36.9), C5 (140.9), H18 (0.58, s) and C12 (38.6), C13 (44.9),

C17 (72.0), C14 (54.4), and H21 (2.36, s) and C20 (208.5), C17 (72.0), C16 (80.5) provided solid evidence for the existence of 18-CH_3, 19-CH_3, Δ^5, 20-oxo, 21-CH_3 and 16-OH. In addition, the correlations between H3 (3.83, m) and H2a (1.67, m), H2b (2.06, m), H4a (2.4, m), H4b (2.64, dd, 13.3, 2.4), and H16 (5.23, t, 7.8) and H17 (2.93, d, 6.2), H15a (1.58, m), H15b (1.89, m) in ^{1}H-^{1}H COSY spectrum proved the presence of 3-OH and 16-OH unambiguously as well. The stereochemistry of the aglycone was determined to be 3β,16α-dihydroxy-pregn-5-en-20-one through ROESY experiment (Figure 4) combined with the biogenesis of steroids.

The anomeric protons at 4.84 (d, 7.8), 5.31 (d, 9.6), 5.11 (dd, 9.6, 1.6), 4.85 (d, 8.0), 5.13 (d, 7.8) in the ^{1}H NMR spectrum of **1** as well as the corresponding carbon resonances at 102.1, 98.9, 100.4, 104.7, 105.3 in the ^{13}C NMR spectrum indicated the presence of five sugar units. The assignments of all the ^{1}H and ^{13}C NMR signals for sugar moiety were accomplished through 1D and 2D NMR techniques, especially by HSQC-TOCSY and HMBC spectra, which revealed that the five saccharide units are one D-thevetopyranosyl unit, two D-cymaropyranosyl units, and two D-glucopyranosyl units. In addition to the signals for the aglycone and five sugar residues, the correlations from $H2_{mBe}$ (5.83, s) to $C4_{mBe}$ (20.2), $C5_{mBe}$ (27.2), $C3_{mBe}$ (158.4), $C1_{mBe}$ (165.9), from $H4_{mBe}$ (2.2, s) to $C5_{mBe}$ (27.2), $C2_{mBe}$ (116.1), $C3_{mBe}$ (158.4), $C1_{mBe}$ (165.9), and from $H5_{mBe}$ (1.73, s) to $C4_{mBe}$ (20.2), $C2_{mBe}$ (116.1), $C3_{mBe}$ (158.4) in the HMBC spectrum, as well as the correlations between $H2_{mBe}$ (5.83, s) and $H5_{mBe}$ (1.73, s), $H4_{mBe}$ (2.20, s) in the ^{1}H-^{1}H COSY spectrum, indicated the presence of 3-methyl-2-Butenoyl group (mBe). The position of all the sugar residues and

353 3-methyl-2-Butenoyl group were determined through HMBC correlations (Figure 2).
354 The correlation signals from $H1_{Thev}$ (4.84, d, 7.8) to C3 (78.1), from $H1_{CymI}$ (5.31, d,
355 9.6) to $C4_{Thev}$ (82.8), from $H1_{CymII}$ (5.11, dd, 9.6, 1.6) to $C4_{CymI}$ (83.3), from $H4_{CymII}$
356 (4.87, dd, 9.9, 2.9) to $C1_{mBe}$ (165.9), from $H1_{GlcI}$ (4.85, d, 8.0) to C16 (80.5), and from
357 $H1_{GlcII}$ (5.13, d, 7.8) to $C6_{GlcI}$ (69.7) in the HMBC spectrum suggested that
358 3-methyl-2-Butenoyl group was linked to C-4 of *β*-D-Cym_{II}, *β*-D-Cym_{II} to C-4 of
359 *β*-D-Cym_I, *β*-D-Cym_I to C4 of *β*-D-Thev, *β*-D-Thev to C-3 of the aglycone, *β*-D-Glc_{II} to
360 C-6 of *β*-D-Glc_I, and *β*-D-Glc_I to C-16 of the aglycone, respectively. The ESI-MS/MS
361 data were also in accordance with those deductions. The absolute configuration of the
362 deoxysugars was determined to the D-seris by co-chromatography on TLC plate with
363 authentic samples, which was obtained through the acid hydrolysis of the crude
364 pregnane glycoside mixture and the isolation of the hydrolysate by column
365 chromatography on silica gel. The absolute configuration of Glc was established to be
366 D-form through GC analysis of derivatives of the hydrolysate and the D-Glc standard
367 and comparison of their retention time. Thus, compound **1** was deduced as
368 3*β*,16*α*-dihydroxy-preg-5-en-20-one-3-*O*-*β*-[4-*O*-(3-methyl-2-butenoyl)]-D-cymaropy
369 ranosyl-(1→4)-*β*-D-cymaropyranosyl-(1→4)-*β*-D-thevetopyranosyl-16-*O*-*β*-D-glucopy
370 ranosyl-(1→6)-*β*-D-glucopyranoside, named extensumside C.
371 Extensumside C–L (**1**–**10**) were obtained as a white amorphous powder and showed
372 positive Liebermann-Burchard[24] and Keller-Kiliani reactions,[25] suggesting that they
373 belonged to steroidal 2-deoxyglycoside. The NMR data of **2**–**10** indicated that they
374 were 3- and 16-*O*-glycosides of 3*β*,16*α*-dihydroxy-pregn-5-en-20-one, and each of

them differed from the others in the sugar moiety at C-3 or C-16.

The molecular formula of extensumside D (**2**) was determined as $C_{59}H_{94}O_{24}$ (calcd. for $C_{59}H_{94}O_{24}Na$, 1209.6027) by a quasi-molecular ion peak at *m/z* 1209.6018 in the HR-ESI-MS spectrum. Detailed analysis of the ^{1}H and ^{13}C NMR data (Table S1 and S2) for the aglycone of **1** and **2** revealed that most of the resonances in **2** were the same as those of **1**, and the significant differences between them were that the *β*-D-Cym linked to C-4 of *β*-D-Cym_I in **1** was replaced by another 3-methoxy-6-deoxy-sugar in **2**. The assignment of the ^{1}H and ^{13}C NMR signals through HSQC-TOCSY spectrum combined with other 2D NMR techniques including HMBC, ^{1}H-^{1}H COSY, ROESY spectra identified this sugar residue as *β*-D-Ole. Therefore, the structure of **2** was elucidated as 3*β*,16*α*-dihydroxy-preg-5-en-20-one-3-*O*-*β*-[4-*O*-(3-methylbut-2-enoyl)]-D-oleandropyranosyl-(1→4)-*β*-D-cymaropyranosyl-(1→4)-*β*-D-thevetopyranosyl-16-*O*-*β*-D-glucopyranosyl-(1→6)-*β*-D-glucopyranoside, named extensumside D.

The molecular formula of extensumside E (**3**) was determined as $C_{65}H_{104}O_{29}$ (calcd. for $C_{65}H_{104}O_{29}Na$, 1371.6561) by a quasi-molecular ion peak at *m/z* 1371.6569 in the HR-ESI-MS spectrum. Compared the molecular weight of **3** with that of **1**, it could be found that the molecular weight of **3** was 162 more than that of **1**, corresponding to the residue of $C_6H_{10}O_5$. Through further detailed analysis of ^{1}H, ^{13}C NMR and DEPT spectra of **3**, it was found that **3** and **1** shared the same aglycone and C-3 saccharide linkage, and the main difference was that the 2-OH of *β*-D-Glc_{II} in **1** was glycosylated by another sugar in **3**, which could be inferred from the glycosidation shifts of C-2

from δ 75.4 to δ 85.1 and C-3 from δ 78.4 to δ 78.1. According to the rest several carbon resonances, it could be deduced that it was D-Glc which was linked to C-4 of *β*-D-Glc_{II} in **3**. The assignment of sugar signals was mainly through HSQC-TOCSY combined with HMBC, ^{1}H-^{1}H COSY spectra. Thus, the structure of **3** was deduced as 3*β*,16*α*-dihydroxy-preg-5-en-20-one-3-*O*-*β*-[4-*O*-(3-methylbut-2-enoyl)]-D-cymaropyranosyl-(1→4)-*β*-D-cymaropyranosyl-(1→4)-*β*-D-thevetopyranosyl-16-*O*-*β*-D-glucopyranosyl-(1→2)-*β*-D-glucopyranosyl-(1→6)-*β*-D-glucopyranoside, named extensumside E.

The molecular formula of extensumside F (**4**) was identified as $C_{65}H_{104}O_{29}$ on the basis of the quasi-molecular ion peak at *m/z* 1371.6565 in the HR-ESI-MS spectrum, which is the same as **3**. The carbon resonances at 98.9 ($C1_{CymI}$), 100.4 ($C1_{CymII}$), 102.1 ($C1_{Thev}$), 104.9 ($C1_{GlcI}$), 105.0 ($C1_{GlcII}$), 105.6 ($C1_{GlcIII}$) and the corresponding proton signals at 5.31 ($H1_{CymI}$), 5.11 ($H1_{CymII}$), 4.83 ($H1_{Thev}$), 4.85 ($H1_{GlcI}$), 5.02 ($H1_{GlcII}$), 5.07 ($H1_{GlcIII}$) in the HSQC spectrum suggested that **4** contained six sugar residues. Compared the ^{1}H and ^{13}C NMR spectra of **4** with those of **3**, it was found that **3** and **4** shared the same aglycone and C-3 sugar linkage, and the significant difference between them was that the *β*-D-Glc linked to C-2 of *β*-D-Glc_{II} in **3** was connected to C-6 of *β*-D-Glc_{II} in **4**, which could be inferred from the glycosidation shifts of C-6 from 62.4 to 69.8 and C-5 from 78.2 to 76.9. This was consistent with the ^{1}H and ^{13}C NMR data of stelmatocryptonoside D (**1a**).[26] All the signals of sugars were assigned through HSQC-TOCSY spectrum combined with HMBC, ^{1}H-^{1}H COSY spectra. Therefore, the structure of **4** was identified as

3*β*,16*α*-dihydroxy-preg-5-en-20-one-3-*O*-*β*-[4-*O*-(3-methylbut-2-enoyl)]-D-cymaropyranosyl-(1→4)-*β*-D-cymaropyranosyl-(1→4)-*β*-D-thevetopyranosyl-16-*O*-*α*-D-glucopyranosyl-(1→6)-*β*-D-glucopyranosyl-(1→6)-*β*-D-glucopyranoside, named extensumside F.

The molecular formula of extensumside G (**5**) was determined as $C_{65}H_{104}O_{29}$ (calcd. for $C_{65}H_{104}O_{29}Na$, 1371.6555) by quasi-molecular ion peak at *m/z* 1371.6483 in the HRESI-MS spectrum, which is the same as **3** and **4**. Compared the ^{1}H and ^{13}C NMR spectra of **5** with those of **4**, it was found that **5** and **4** shared the same aglycone and C-16 sugar linkage, and the significant difference between them was that the sugar linked to C-4 of *β*-D-Cym_I in **4** was replaced by another deoxysugar in **5**. In addition, further comparison of ^{1}H and ^{13}C NMR data of **5** (Table S1 to S4) with those of **2**, it was found that **5** shared the same aglycone and C-3 oligosaccharide chain as **2**. Combined analysis above, the structure of **5** was identified as 3*β*,16*α*-dihydroxy-preg-5-en-20-one-3-*O*-*β*-[4-*O*-(3-methylbut-2-enoyl)]-D-oleandropyranosyl-(1→4)-*β*-D-cymaropyranosyl-(1→4)-*β*-D-thevetopyranosyl-16-*O*-*β*-D-glucopyranosyl-(1→6)-*β*-D-glucopyranosyl-(1→6)-*β*-D-glucopyranoside, named extensumside G.

The molecular formula of extensumside H (**6**) was determined as $C_{71}H_{114}O_{34}$ (calcd. for $C_{71}H_{114}O_{34}Na$, 1533.7084) by a quasi-molecular ion peak at *m/z* 1533.7057 in the HR-ESI-MS spectrum. The proton resonances at 5.31 ($H1_{CymI}$), 5.13 ($H1_{CymII}$), 4.82 ($H1_{Thev}$), 5.07 ($H1_{GlcII}$), 4.89 ($H1_{GlcI}$), 5.01 ($H1_{GlcIV}$), 5.15 ($H1_{GlcIII}$) and the corresponding carbon signals at 98.9 ($C1_{CymI}$), 100.4 ($C1_{CymII}$), 102.2 ($C1_{Thev}$), 102.6

441 ($C1_{GlcII}$), 104.9 ($C1_{GlcI}$), 105.4 ($C1_{GlcIV}$), 106.6 ($C1_{GlcIII}$) in the HSQC spectrum
442 indicated the presence of seven sugar residues and β-orientation of all the sugars. The
443 molecular weight of **6** was 162 more than that of **3**, which was corresponding to the
444 residue ($C_6H_{10}O_5$) of D-Glc. Compared the ^{13}C NMR and DEPT spectra of **6** with
445 those of **3**, it could be found that **6** and **3** shared the same aglycone and C-3 saccharide
446 linkage, and the obvious distinction between them was that C-6 of *β*-D-Glc_{II} in **6** was
447 substituted by another D-Glc, which could be inferred from the glycosidation shifts of
448 C-6 of *β*-D-Glc_{II} in **6** from 62.4 to 69.4. All the signals of sugars were assigned
449 through HSQC-TOCSY spectrum combined with HMBC, ^1H-^1H COSY spectra. Thus,
450 the structure of **6** was deduced as
451 3*β*,16*α*-dihydroxy-preg-5-en-20-one-3-*O*-*β*-[4-*O*-(3-methylbut-2-enoyl)]-D-cymaropy
452 ranosyl-(1→4)-*β*-D-cymaropyranosyl-(1→4)-*β*-D-thevetopyranosyl-16-*O*-bis-*β*-D-gluc
453 opyranosyl-(1→2 and 1→6)-*β*-D-glucopyranosyl-(1→6)-*β*-D-glucopyranoside, named
454 extensumside H.
455 Extensumside I (**7**) had the molecular formula $C_{71}H_{114}O_{34}$ (calcd. for $C_{71}H_{114}O_{34}Na$,
456 1533.7084) based on the quasi-molecular ion peak at *m/z* 1533.7049 in the
457 HR-ESI-MS spectrum, which is the same as **6**. Compared the ^{13}C NMR and DEPT
458 spectra of **7** with those of **6**, it was found that they shared the same aglycone and C-16
459 oligosaccharide chain, and the main difference between them was that *β*-D-Cym_{II} in **6**
460 was replaced by another sugar unit in **7**. Further comparison of the ^{13}C NMR and
461 DEPT spectra of **7** with those of **2**, revealed that the C-3 oligosaccharide chain of **7**
462 was the same as that of **2**. The deduction was supported by the HMBC correlations

463 and the ESI-MS/MS data. The assignment of all the sugar signals was accomplished
464 through HSQC-TOCSY spectrum combined with HMBC, ^{1}H-^{1}H COSY spectra.
465 Therefore, the structure of **7** was identified as
466 3*β*,16*α*-dihydroxy-preg-5-en-20-one-3-*O*-*β*-[4-*O*-(3-methylbut-2-enoyl)]-D-oleandrop
467 yranosyl-(1→4)-*β*-D-cymaropyranosyl-(1→4)-*β*-D-thevetopyranosyl-16-*O*-bis-*β*-D-gl
468 ucopyranosyl-(1→2 and 1→6)-*β*-D-glucopyranosyl-(1→6)-*β*-D-glucopyranoside,
469 named extensumside I.
470 Extensumside J (**8**) had the molecular formula $C_{70}H_{112}O_{34}$ (calcd. for $C_{70}H_{112}O_{34}Na$,
471 1519.6927) based on the quasi-molecular ion peak at *m/z* 1519.6913 in HR-ESI-MS
472 spectrum. The proton signals at 5.53 ($H1_{Dig}$), 5.16 ($H1_{Cym}$), 4.82 ($H1_{Thev}$), 5.06
473 ($H1_{GlcII}$), 4.89 ($H1_{GlcI}$), 5.01 ($H1_{GlcIV}$), and 5.15 ($H1_{GlcIII}$) as well as the corresponding
474 carbon resonances at 99.0 ($C1_{Dig}$), 100.0 ($C1_{Cym}$), 102.2 ($C1_{Thev}$), 102.6 ($C1_{GlcII}$),
475 105.0 ($C1_{GlcI}$), 105.5 ($C1_{GlcIV}$), and 106.7 ($C1_{GlcIII}$) in the HSQC spectrum, indicated
476 the existence of seven monosaccharide units with C1-OH of all the sugar units to be
477 *β*-orientation. Comparison of the ^{13}C NMR and DEPT spectra of **8** with those of **6**
478 revealed that **8** shared the same aglycone and C-16 oligosaccharide chain as **6**, and the
479 significant difference between them was that the *β*-D-Cym connected to C-4 of
480 *β*-D-Thev in **6** was replaced by another sugar unit in **8**. In the HSQC-TOCSY
481 spectrum, the signals at 5.53 ($H1_{Dig}$), 4.66 ($H3_{Dig}$), 2.44 ($H2b_{Dig}$), 2.01 ($H2a_{Dig}$)
482 correlated with $C1_{Dig}$ (99.0), and $H3_{Dig}$ (4.66, m) correlated with $C1_{Dig}$ (99.0), $C4_{Dig}$
483 (83.4), $C5_{Dig}$ (68.8), $C3_{Dig}$ (67.8), $C2_{Dig}$ (39.2), $C6_{Dig}$ (18.5). Combined with analysis
484 of HMBC, ^{1}H-^{1}H COSY and ROESY spectra, this sugar unit was identified as

485 *β*-D-digitoxopyranose (*β*-D-Dig). Therefore, the structure of **8** was deduced as
486 3*β*,16*α*-dihydroxy-preg-5-en-20-one-3-*O*-*β*-[4-*O*-(3-methylbut-2-enoyl)]-D-cymaropy
487 ranosyl-(1→4)-*β*-D-digitoxopyranosyl-(1→4)-*β*-D-thevetopyranosyl-16-*O*-bis-*β*-D-glu
488 copyranosyl-(1→2 and 1→6)-*β*-D-glucopyranosyl-(1→6)-*β*-D-glucopyranoside,
489 named extensumside J.

490 Extensumside K (**9**) had the molecular formula of $C_{70}H_{112}O_{34}$ (calcd. for
491 $C_{70}H_{112}O_{34}Na$, 1519.6927) based on the quasi-molecular ion peak at *m/z* 1519.6910 in
492 the HR-ESI-MS spectrum. Compared the ^{13}C NMR and DEPT spectra of **9** with those
493 of **7**, it was found that **9** shared the same aglycone and C-16 oligosaccharide chain
494 with **7**, while the sugar linked to C-4 of *β*-D-Thev in **7** was different in **9**. Detailed
495 analysis of the HSQC-TOCSY spectrum indicated that $H1_{Dig}$ (5.56), $H4_{Dig}$ (3.53),
496 $H2a_{Dig}$ (2.03), $H2b_{Dig}$ (2.46) correlated with $C1_{Dig}$ (99.0), and $H4_{Dig}$ (3.53) correlated
497 with $C1_{Dig}$ (99.0), $C4_{Dig}$ (83.6), $C3_{Dig}$ (67.8), $C5_{Dig}$ (68.8), $C2_{Dig}$ (39.3), $C6_{Dig}$ (18.6)
498 as well. Combined with the HMBC, ^1H-^1H COSY, ROESY spectra, it could be
499 deduced that the sugar unit connected to C-4 of *β*-D-Thev was D-digitoxopyranose
500 (*β*-D-Dig). All the signal assignment of the sugars was mainly through the
501 HSQC-TOCSY spectrum combined with the HMBC, ^1H-^1H COSY, ROESY spectra.
502 Thus, the structure of **9** was elucidated as
503 3*β*,16*α*-dihydroxy-preg-5-en-20-one-3-*O*-*β*-[4-*O*-(3-methylbut-2-enoyl)]-D-oleandrop
504 yranosyl-(1→4)-*β*-D-digitoxopyranosyl-(1→4)-*β*-D-thevetopyranosyl-16-*O*-bis-*β*-D-gl
505 ucopyranosyl-(1→2 and 1→6)-*β*-D-glucopyranosyl-(1→6)-*β*-D-glucopyranoside,
506 named extensumside K.

Extensumside L (**10**) had the molecular formula $C_{70}H_{112}O_{34}$ based on a quasi-molecular ion peak at *m/z* 1519.6917 (calcd. for $C_{70}H_{112}O_{34}Na$, 1519.6927) in the HR-ESI-MS spectrum. Compared the ^{13}C NMR and DEPT spectra of **10** with those of **6**, it was found that **10** shared the same aglycone and C-16 oligosaccharide chain as **6**, while the sugar linked to C-4 of β-D-Cym_{II} in **7** was different in **10.** From the HMBC correlations from $H1_{Dig}$ (5.42) to $C2_{Dig}$ (39.6), from $H4_{Dig}$ (4.91) to $C5_{Dig}$ (67.8), $C6_{Dig}$ (18.4), from $H6_{Dig}$ (1.40) to $C5_{Dig(}$ 67.8), $C4_{Dig}$ (75.6), from $H2a_{Dig}$ (2.02) to $C1_{Dig}$ (100.7), and from $H2b_{Dig}$ (2.40) to $C1_{Dig}$ (100.7), together with ^{1}H-^{1}H COSY correlations between $H1_{Dig}$ (5.42, d, 9.6, 1.2) and $H2a_{Dig}$ (2.02, m), $H4_{Dig}$ (4.91, m) and $H5_{Dig}$ (4.54, m), $H3_{Dig}$ (4.65, m) and $H2a_{Dig}$ (2.02, m), and $H5_{Dig}$ (4.54, m) and $H6_{Dig}$ (1.40, d, 6.2) and combined with the ROESY spectrum, it could be deduced that the sugar unit connected to C-4 of *β*-D-Cym was D-digitoxopyranose (*β*-D-Dig). Therefore, the structure of **10** was elucidated as 3*β*,16*α*-dihydroxy-preg-5-en-20-one-3-*O*-*β*-[4-*O*-(3-methylbut-2-enoyl)]-D-digitoxopyranosyl-(1→4)-*β*-D-cymaropyranosyl-(1→4)-*β*-D-thevetopyranosyl-16-*O*-bis-*β*-D-glucopyranosyl-(1→2 and 1→6)-*β*-D-glucopyranosyl-(1→6)-*β*-D-glucopyranoside, named extensumside L.

Sweetness Intensities and Structure-Activity Relationship of C21 Pregnane Glycosides. The sweetness intensities of **1**–**10** were evaluated by seven sweet sensitive panelists and the results are reported in table 2. All these C21 pregnane glycosides are natural sweet compounds with a high sweetness potency and low sweet-tasting threshold. Their sweetness intensities are between 50 to 400 times

greater than that of sucrose. These C21 pregnane glycosides represent the main sweet components in the pericarps of *M. extensum* and provide a reasonable explanation for the ethnobotanical findings that Yao people in Xinpin county of Yunnan province prefer to utilize the pericarps of *M. extensum* to make pickles.

Based on our research on the sensory evaluation of the C21 pregnane glycosides, the sweetness intensities of these compounds are closely related to the aglycone (to be reported elsewhere) and the saccharide moieties. As for the saccharide moieties, the sorts and number of the sugar units and the linkage of them may be critical to the sweetness intensities of this kind of compounds. The sweetness intensities decrease rapidly with the number of glucose in the glucose chain linked to the C16 of the aglycone increasing, suggesting that the glucose chain affect the sweetness intensities largely. The deoxysugar chain linked to the C3 of the aglycone also has an unnegligible influence on the sweetness. However, the detailed structure-taste relationship, and the mechanism of action of these compounds to be sweet-tasting are still unknown, which may be the subject of our further investigations.

Distribution of Sweet C21 Pregnane Glycosides in Different Parts of *M. extensum*. The occurrence and quantitation of compounds **1**, **3**, **4**, and **6** in different tissues (pericarps, stems, and roots) of *M. extensum* were analyzed by means of HPLC-DAD. The results are shown in table 3. In general, the highest concentration of these compounds was found in the pericarps, followed by in the stems, whereas the roots of this plant had the lowest contents of these compounds. On the other hand, the relative concentrations of these compounds were also unevenly distributed in different

551 tissues. From the aspect of enrichment of botanical secondary metabolites, the
552 growing years of roots are generally longer than that of pericarps, and accordingly the
553 enrichment of metabolites in the roots is higher than that in the pericarps.
554 Nevertheless, what is surprised to us is that the contents of these sweet components in
555 the pericarps are obviously higher than those in the roots with longer growing years.
556 This implies that pericarps of *M. extensum* may be a good and sustainable resource to
557 use for the research and development of natural high-potency sweeteners.
558

559 ## ASSOCIATED CONTENT

560 **Supporting Information**

561 The ^{1}H NMR, ^{13}C NMR data of extensumside D–L (**2**-**10**); The scheme for the
562 isolation of compounds **1-10** from the pericarps of *M. extensum*

563 The Supporting information is available free of charge on the ACS Publications
564 website at DOI:

565 ## AUTHOR INFORMATION

566 **Corresponding Author**

567 * Phone: +86-871-65223224; E-mail address: duzhizhi@mail.kib.ac.cn

568 **Author Contributions**

569 $^{\nabla}$ Guo Sun and Qin Dai equally contributed to this work.

570 **Funding**

571 This research was supported by grants from the Natural Science Foundation of the
572 Yunnan Province (2013FB065), the 45th Scientific Research Foundation for the

Returned Overseas Chinese Scholars from State Education Ministry and National S&T Basic Work Program of China (2012FY110300).

Note

The authors declare no competing financial interests.

ACKNOWLEDGMENT

We thank Professor Lisong Wang for identification of the plant. We also thank associate professor Yuehu Wang for the help in elucidating the structures of the compounds and the supporting of HPLC. We are very grateful to all the panelists participating in the sensory tests, too.

583 ## REFERENCES

584 (1) Lindemann, B. Receptors and transduction in taste. *Nature* **2001**, *413*, 219-225.

585 (2) Temussi, P. A. Sweet, bitter and umami receptors: a complex relationship. *Trends*

586 *in Biochem. Sci.* **2009**, *34*, 296-302.

587 (3) Rastogi, S. S.; Singh, R. B.; Singh, N. K. et al. Prevention of cardiovascular

588 disease and diabetes mellitus in low and middle income countries. *The Open*

589 *Nutraceuticals Journal* **2011**, *4*, 97-106.

590 (4) Anderson, A. S. Sugars and health – risk assessment to risk management. *Public*

591 *Health Nutr.* **2014**, *17*, 2148–2150.

592 (5) Vos, M. B.; Kaar, J. L.; Welsh, J. A. Added sugars and cardiovascular disease

593 risk in children: a scientific statement from the American Heart Association.

594 *Circulation* **2016***, 134*, 1-18.

595 (6) Yang, Q. H.; Zhang, Z. F.; Gregg, E. W. et al. Added sugar intake and

596 cardiovascular diseases mortality among US adults. *JAMA Intern. Med*. **2014**, *174*,

597 516-524.

598 (7) Yang, W. Y.; et al. Prevalence of diabetes among men and women in China. *N.*

599 *Engl. J. Med.* **2010**, *362*, 1090-1101.

600 (8) Duffy, V. B.; Anderson, G. H. Position of the American dietetic association: use of

601 nutritive and nonnutritive sweeteners. *J. Am. Diet. Assoc.***1998**, *98*, 580-587.

602 (9) Kim, N.-C.; Kinghorn, A. D. Highly sweet compounds of plant origin. *Arch.*

603 *Pharm. Res.* **2002**, *25*, 725-746.

604 (10) Kurihara, Y. Characterization of antisweet substance, sweet proteins, and

605 sweetness-inducing proteins. *Crit. Rev. Food Sci. Nutr.* **1992**, *32*, 231-252.

606 (11) Kinghorn, A. D.; Kennelly, E. J. Discovery of highly sweet compounds from

607 natural sources. *J. Chem. Educ.* **1995**, *72*, 676-680.

608 (12) Suttisri, R.; Lee, I. S.; Kinghorn, A. D. Plant-derived triterpenoid sweetness

609 inhibitors. *J. Ethnopharmacol*. **1995**, *47*, 9-26.

610 (13) DuBois, G. E.; Prakash, I. Non-caloric sweeteners, sweetness modulators, and

611 sweetener enhancers. *Annu. Rev. Food Sci. Technol.* **2012**, *3*, 353-380.

612 (14) Editorial Committee of Flora of China, Chinese Academy of Sciences. Flora of

613 China. Science Press: Beijing, China, **1977**; *63*, pp 270.

614 (15) Yang, M. F.; Li, Y. Y.; Zhang, G. L. Steroidal saponins from *Myriopteron*

615 *extensum* and their cytotoxic activity. *Planta Med.* **2004**, *70*, 556-560.

616 (16) Li, Y. Y.; Zhang, G. L. Chemical constituents and cytotoxic activity of

617 *Myriopteron extensum*. *Nat. Prod. Res. Dev.* **2003**, *15*, 113-115.

618 (17) Noppamas Soonthornchareonnon. Lupinifolin, a bioactive flavanone from

619 *Myriopteron extensum* (Wight) K. Schum. Stem. *Thai J. Phytopharmacy* **2004**, *2*,

620 19-28.

621 (18) Dai, Q.; Wang, Y. H.; Zhang, H. X.; Sun, G.; Du, Z. Z. Three new

622 dinormonoterpenoid glucosides from pericarps of *Myriopteron extensum. Phytochem.*

623 *Lett.* **2015**, *12*, 164-167.

624 (19) Wang, J. S.; Yang, X. W.; Di, Y. T.; Wang, Y. H.; Shen, Y. M.; Hao, X. J.

625 Isoflavone diglycosides from *Glycosmis pentaphylla*. *J. Nat. Prod.* **2006**, *69*, 778-782.

626 (20) Abe, F.; Mori, Y.; Okabe, H.; Yamauchi, T. Steroidal constituents from the roots

and stems of *Asclepias fruticosa*. *Chem. Pharm. Bull.* **1994**, *42*, 1777-1783.

(21) Liang, D.; Hao, Z. Y.; Zhang, G. J.; Chen, R. Y.; Yu, D. Q. Cytotoxic triterpenoid saponins from *Lysimachia clethroides*. *J. Nat. Prod.* **2011**, *74*, 2128-2136.

(22) ISO 8586: 2012. *Sensory Analysis—General guidelines for the selection, training and monitoring of selected assessors and expert sensory assessors*, 2012.

(23) Jia, Z. H.; Yang, X. G. A minor, sweet cucurbitane glycoside from *Siraitia grosvenorii*. *Nat. Prod. Commun.* **2009**, *4*, 769-772.

(24) Xiong, Q.; Wilson, W. K.; Pang, J. The Liebermann-Burchard reaction: sulfonation, desaturation, and rearrangment of cholesterol in acid. *Lipids* **2007,** *42*, 87-96.

(25) Halilu, M. E.; Abubakar, A.; Garba, M. K.; Isah, A. A. Antimicrobial and Preliminary Phytochemical studies of Methanol Extract of Root Bark of Crossopteryx febrifuga (Rubiaceae). *J. Appl. Pharm. Sci.* **2012,** *2*, 66-70.

(26) Zhang, Q. Z.; Zhao, Y. Y.; Wang, B.; Feng, R.; Liu, X. H.; Cheng, T. M. New pregnane glycosides from *Stelmatocrypton khasianum*. *Steroids* **2002**, *67*, 347-351.

643 **Figure captions**

644 **Figure 1.** Structures of compounds **1-10** isolated from the pericarps of *M. extensum*

645 **Figure 2.** Key HMBC and COSY correlations of **1** (HMBC →, COSY —)

646 **Figure 3.** The structure of compound **1a**

647 **Figure 4.** The ROESY correlations of the aglycone moiety

648 **Figure 5.** The analytical HPLC chromatogram of sweet compounds (**1**, **3**, **4** and **6**) in

649 the pericarps, stems, roots of *M. extensum*.

650

651

Table 1. The ^{1}H NMR (600 MHz) and ^{13}C NMR (150 MHz) data of compound 1

	^{13}C	^{1}H		^{13}C	^{1}H
1	37.4	1.65 (m) 0.96 (m)	2	37.0	2.37 (d, 10.5) 1.87 (dd, 10.5, 2.6)
2	30.2	2.06 (m) 1.67 (m)	3	78.0	4.06 (m)
			4	83.3	3.50 (dd, 9.6, 2.6)
3	78.1	3.83 (m)	5	69.2	4.21 (m)
4	39.2	2.64 (dd, 13.3, 2.4) 2.40 (m)	6	18.4	1.36 (d, 6.2)
			3-OCH_3	58.8	3.58 (s)
5	140.9	—			D-Cym_{II}
6	121.6	5.26 (m)	1	100.4	5.11 (dd, 9.6, 1.6)
7	31.9	1.70 (m) 1.40 (m)	2	36.0	2.33 (m) 1.82 (m)
8	31.4	1.20 (m)	3	75.8	3.90 (m)
9	50.1	0.83 (td, 10.8, 4.4)	4	74.6	4.87 (dd, 9.9, 2.9)
10	36.9	—	5	68.3	4.28 (m)
11	20.9	1.37 (m) 1.22 (m)	6	18.4	1.33 (d, 6.3)
			3-OCH_3	58.4	3.40 (s)
12	38.6	1.79 (d, 2.8) 1.24 (m)			mBe
			1	165.9	—
13	44.9	—	2	116.1	5.83 (s)
14	54.4	1.30 (m)	3	158.4	—
15	33.4	1.89 (m) 1.58 (m)	4	20.2	2.20 (s)
			5	27.2	1.73 (s)
16	80.5	5.23 (t, 7.8)			D-Glc_I
17	72.0	2.93 (d, 6.2)	1	104.7	4.85 (d, 8.0)
18	14.6	0.58 (s)	2	75.2	3.99 (t, 8.4)
19	19.3	0.80 (s)	3	78.5	4.19 (t, 8.6)
20	208.5	—	4	71.3	4.23 (m)
21	32.4	2.36 (s)	5	77.2	3.98 (m)
		D-Thev	6	69.7	4.80 (d, 11.3) 4.32 (dd, 11.3, 4.4)
1	102.1	4.84 (d, 7.8)			
2	74.6	3.96 (d, 8.5)			D-Glc_{II}
3	85.8	3.72 (m)	1	105.3	5.13 (d,7.8)
4	82.8	3.70 (m)	2	75.4	4.09 (m)
5	71.6	3.68 (m)	3	78.4	4.30 (m)
6	18.7	1.45 (d, 5.5)	4	71.6	4.29 (m)
3-OCH_3	60.5	3.94 (s)	5	78.4	4.01 (m)
		D-Cym_I	6	62.7	4.53 (dd, 11.7, 1.7) 4.41 (d, 11.7)
1	98.9	5.31 (d, 9.6)			

[a] All spectra were recorded in pyridine-d_5 solvent; δ in ppm, J in Hz; hydrogen and carbon signal assignments were verified by HSQC, HMBC, ^{1}H-^{1}H COSY, HSQC-TOCSY and DEPT measurements.

Table 2. The sensory evaluation of compounds 1-10

Comp.	R_1	R_2	Sweet potency*	Threshold** (mg/mL)
1	-*β*-D-Thev^4-*β*-D-Cym^4-*β*-D-Cym^4-mBe	-*β*-D-Glc^6-*β*-D-Glc	400	0.02
2	-*β*-D-Thev^4-*β*-D-Cym^4-*β*-D-Ole^4-mBe	-*β*-D-Glc^6-*β*-D-Glc	300	0.02
3	-*β*-D-Thev^4-*β*-D-Cym^4-*β*-D-Cym^4-mBe	S_1	200	0.025
4	-*β*-D-Thev^4-*β*-D-Cym^4-*β*-D-Cym^4-mBe	S_2	200	0.025
5	-*β*-D-Thev^4-*β*-D-Cym^4-*β*-D-Ole^4-mBe	S_2	100	0.033
6	-*β*-D-Thev^4-*β*-D-Cym^4-*β*-D-Cym^4-mBe	S_3	200	0.025
7	-*β*-D-Thev^4-*β*-D-Cym^4-*β*-D-Ole^4-mBe	S_3	150	0.033
8	-*β*-D-Thev^4-*β*-D-Dig^4-*β*-D-Cym^4-mBe	S_3	100	0.05
9	-*β*-D-Thev^4-*β*-D-Dig^4-*β*-D-Ole^4-mBe	S_3	50	0.1
10	-*β*-D-Thev^4-*β*-D-Cym^4-*β*-D-Dig^4-mBe	S_3	50	0.1

* sweet intensity relative to the sucrose solution (1%, w/w);

** minimum concentration that human can percept;

Table 3. The amount of sweet compounds (1, 3, 4 and 6) in the pericarps, stems, and roots of *M. extensum*[a]

Comp.	Fruits (mg/g)	Stems (mg/g)	Roots (mg/g)
1	0.262 ± 0.012	0.050 ± 0.006	0.031 ± 0.003
3	0.115 ± 0.004	0.116 ± 0.007	0.046 ± 0.002
4	0.183 ± 0.010	0.060 ± 0.004	0.032 ± 0.005
6	0.664 ± 0.021	0.251 ± 0.011	0.031 ±0.002

[a] concentrations are means (± SD) of duplicate analyses, and calculated in dry weight of the plant materials.

Figure 1

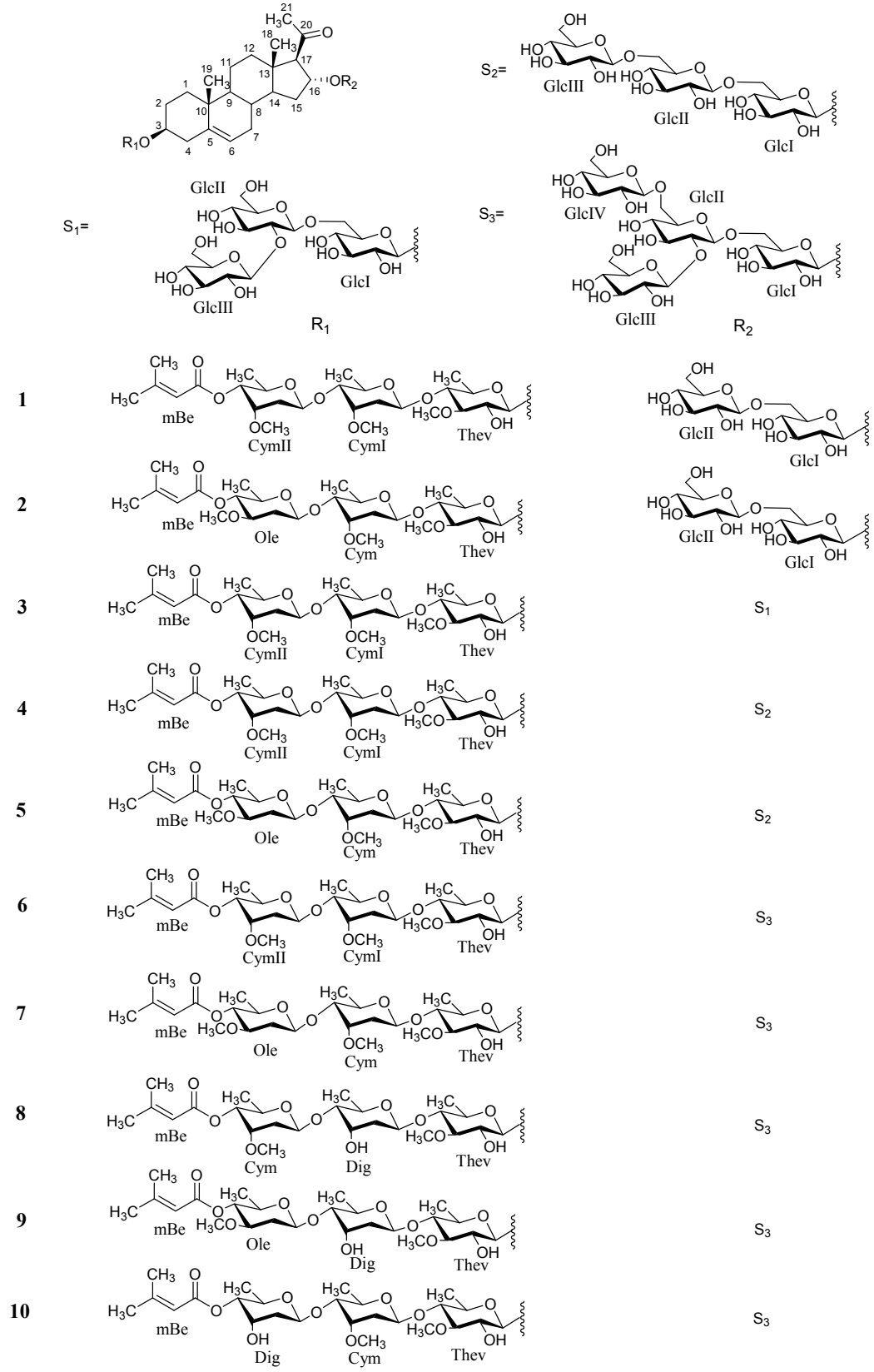

S1=
S2=
S3=
R1
R2
GlcI
GlcII
GlcIII
GlcIV
1
2
3
4
5
6
7
8
9
10
mBe
Ole
OCH3
Cym
CymII
CymI
Dig
Thev
OH
S1
S2
S3

Figure 2

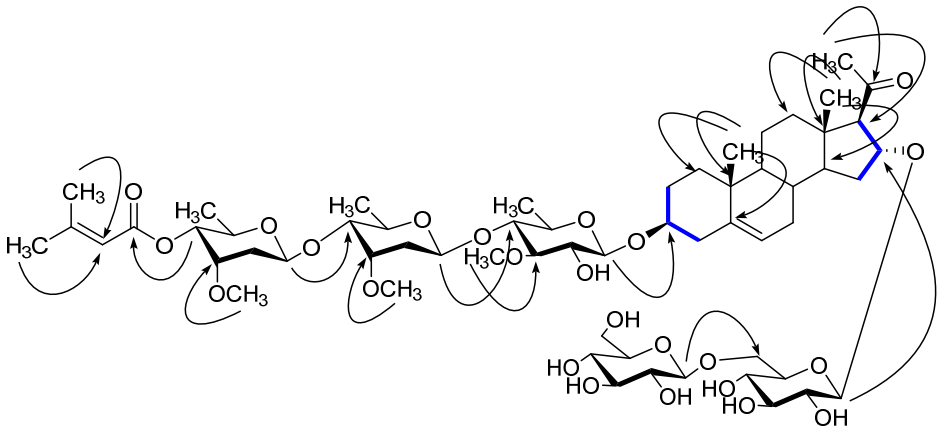

CH3
O
H3C
H3C
O
O
H3C
O
O
H3C
O
O
H3C
O
O
OCH3
OCH3
H3CO
OH
CH3
H3C
O
CH3
O
OH
HO
HO
O
OH
O
HO
HO
O
OH

Figure 3

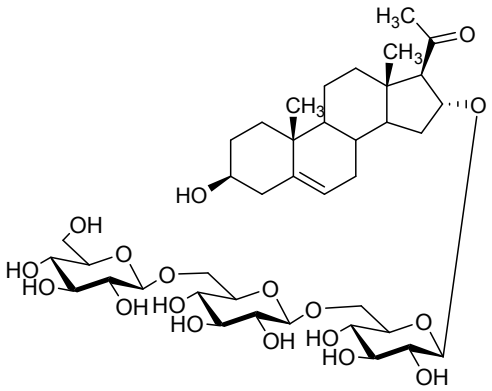


Figure 4

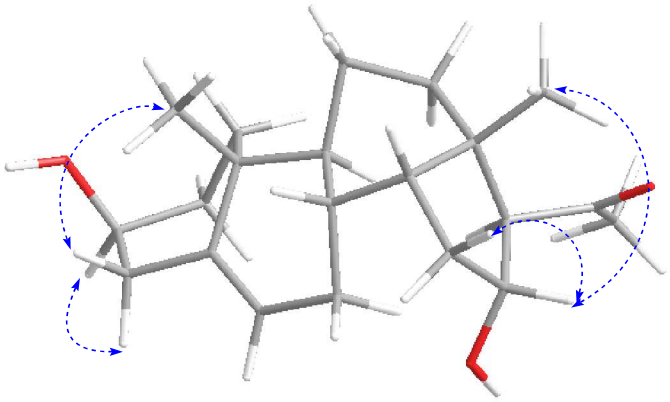
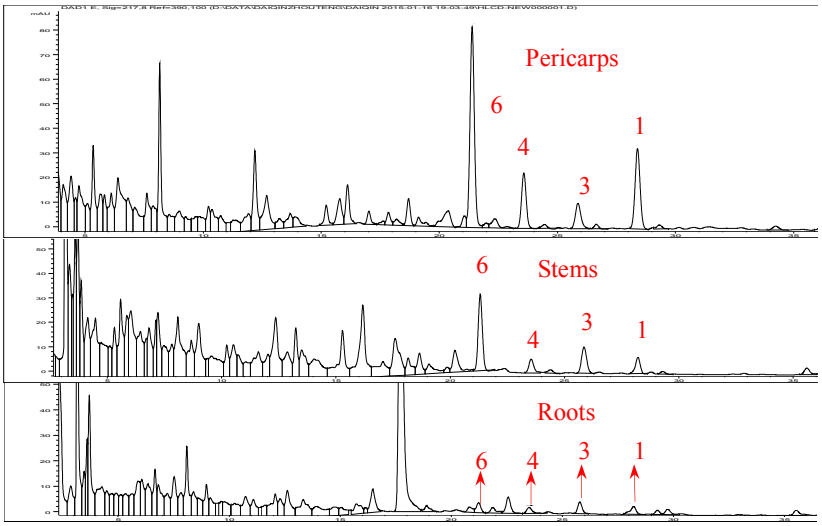
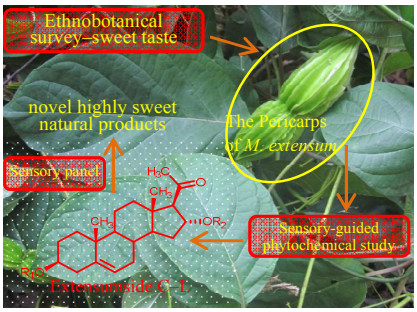


Figure 5.



TOC graphic


Ethnobotanical survey–sweet taste
novel highly sweet natural products
The Pericarps of M. extensum
Sensory panel
Sensory-guided phytochemical study
Extensumoside C–I