Qualitative Analysis of Chemical Constituents in the Root of Codonopsis convolvulacea by Liquid Chromatography-Mass Spectrometry (Postprint)

Authors: Zhao Wenhui, Dou Zhiyang, Qiangba, Wu Jian, Baima Zhuoma, Liu Guangxue, Cai Shaoqing, Xu Feng

Date: 2024-12-15T00:00:00+00:00

Abstract

To elucidate the chemical constituents of the Tibetan medicine Jidan Shen [the root of Codonopsis convolvulacea var. pinifolia], qualitative analysis of its chemical components was performed using ultra-high performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry (UHPLC-Q-TOF-MS). A Waters ACQUITY UPLC BEH C18 column (1.7 m, $2.1\$\times 150mm)wasemployed, with 0.1mmol \cdot L\hat{}\{-1\}\$$ ammonium acetate aqueous solution (A)–acetonitrile (B) as the mobile phase, at a flow rate of 0.3 $mL \cdot min^{-1}$, injection volume of 3 L, and column temperature of 40 °C. Mass spectrometry was conducted using an ESI ion source in negative ion mode, with scanning ranges of m/z 100–1,800 (MS) and m/z 50–1,800 (MS/MS). The molecular formula was deduced based on the quasi-molecular ion in the MS^1 spectrum, while possible structural fragments and molecular structures were inferred from characteristic fragment ions in the MS^2 spectrum. Structural identification was further performed by comparison with reference standards and literature search. The results demonstrated: (1) Fifty-six compounds were identified from C. convolvulacea var. pinifolia for the first time, including six nitrogen-containing constituents, six phenylpropanoids, twenty-two lignans, two flavonoids, eight organic acids, and twelve glycosides or other constituents; among these, eleven components were unambiguously identified by comparison with reference standards. (2) Lignans and phenylpropanoids were identified as the major constituents in the root of C. convolvulacea var. pinifolia for the first time, and their fragmentation pathways were proposed. These findings indicate that the application of LC-MS technology enables rapid and efficient preliminary elucidation of the chemical constituents in the root of C. convolvulacea var. pinifolia, providing a chemical foundation for studies on quality standards, in vivo processes, and pharmacodynamic substances of the Tibetan medicine Jidan Shen.

Full Text

Qualitative Analysis of Chemical Constituents in the Roots of *Codonopsis convolvulacea* var. *pinifolia* by LC-MS

ZHAO Wenhui[1],[2], DOU Zhiyang[2], Qiangba[3], *Wujian*[3], *Baimazhuoma*[3], *LIU Guangxue*[2], *CAI Shaoqing*[2], *XU Feng*[1],[2]

[1] Medical College, Tibet University, Lhasa 850000, China
[2] School of Pharmaceutical Sciences, Peking University, Beijing 100191, China
[3] Tibet Qizheng Tibetan Medicine Co. Ltd., Nyingchi 860000, Tibet, China

Abstract: To elucidate the chemical constituents of the Tibetan medicine Codonopsis Convolvulaceae Radix (the roots of *Codonopsis convolvulacea* var. *pinifolia*), we performed qualitative analysis using ultra-high performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UHPLC-Q-TOF-MS). Chromatographic separation was achieved on a Waters ACQUITY UPLC BEH C18 column (1.7 μm, 2.1 × 150 mm) with a mobile phase consisting of 0.1 mmol · L^{-1} ammonium acetate aqueous solution (A) and acetonitrile (B) at a flow rate of 0.3 mL · min^{-1}. The injection volume was 3 μL and the column temperature was maintained at 40 °C. Mass spectrometry was conducted using an electrospray ionization (ESI) source in negative ion mode, with scanning ranges of m/z 100–1,800 (MS) and m/z 50–1,800 (MS/MS). Molecular formulas were deduced from quasi-molecular ions in the primary mass spectra, while structural fragments and possible molecular structures were elucidated based on characteristic fragment ions in the secondary mass spectra. Structures were ultimately identified by comparison with reference standards and literature searches. The results demonstrated: (1) Fifty-six compounds were identified from *C. convolvulacea* var. *pinifolia* for the first time, including 6 nitrogen-containing compounds, 6 phenylpropanoids, 22 lignans, 2 flavonoids, 8 organic acids, and 12 glucosides or other compounds. Among these, 11 compounds were unequivocally identified by comparison with reference standards. (2) Lignans and phenylpropanoids were identified as the major constituents in the roots of *C. convolvulacea* var. *pinifolia* for the first time, and their fragmentation pathways were deduced. This study demonstrates that UHPLC-Q-TOF-MS enables rapid and efficient preliminary elucidation of the chemical constituents in *C. convolvulacea* var. *pinifolia* roots, providing a chemical foundation for further investigations into quality standards, in vivo processes, and pharmacodynamic substances of Codonopsis Convolvulaceae Radix.

Keywords: *Codonopsis convolvulacea* var. *pinifolia*, UHPLC-Q-TOF-MS, chemical constituents, structural elucidation, fragmentation pathway

Introduction

According to authoritative texts including *Chinese Tibetan Medicine* (Qinghai Provincial Institute of Drug Control and Qinghai Institute of Tibetan Medicine, 1996), *Tibetan Medicine Records* (Northwest Institute of Plateau Biology, Chinese Academy of Sciences, 1991), *Chinese Materia Medica (Tibetan Volume)* (National Administration of Traditional Chinese Medicine Editorial Board, 2002), and *Sichuan Provincial Standards for Tibetan Medicinal Materials (2020 Edition)* (Sichuan Medical Products Administration, 2021), the Tibetan medicine Codonopsis Convolvulaceae Radix comprises the roots of six plant species: *Codonopsis convolvulacea* [currently *Pseudocodon convolvulaceus*], *Codonopsis convolvulacea* subsp. *vinciflora* [currently *Pseudocodon vinciflorus*], *Codonopsis convolvulacea* var. *pinifolia* [currently *Pseudocodon graminifolius*], *Codonopsis convolvulacea* var. *forrestii* [currently *Pseudocodon convolvulaceus* subsp. *forrestii*], *Codonopsis affinis*, and *Codonopsis macrocalyx* [currently *Codonopsis benthamii*]. The Tibetan transliteration is "Niwa" or "Niewa." The material used in this study was derived from the dried roots of *C. convolvulacea* var. *pinifolia*.

Codonopsis Convolvulaceae Radix possesses heat-clearing, qi-tonifying and blood-nourishing, lung-moistening and fluid-producing, spleen-strengthening and stomach-benefiting, and olfactory-enhancing effects. It is used to treat colds, chest pain, anemia, anorexia, cough due to lung yin deficiency, tonsillitis, and other conditions (Northwest Institute of Plateau Biology, 1991; Qinghai Provincial Institute of Drug Control and Qinghai Institute of Tibetan Medicine, 1996; National Administration of Traditional Chinese Medicine Editorial Board, 2002; Sichuan Medical Products Administration, 2021). Chemical investigations of this medicine remain in the preliminary stage. Beyond nutritional components including 17 amino acids, multiple vitamins, and inorganic elements (Zhong et al., 2002), only 27 compounds have been isolated and identified to date, primarily comprising anthraquinones, triterpenoids, steroids, aliphatic compounds, organic acids, lignans, and alkynes (Han et al., 2001; Wu, 2009; Sun et al., 2019).

No studies have reported qualitative analysis of *C. convolvulacea* var. *pinifolia* using UHPLC-Q-TOF-MS/MS. Given the technique's advantages of high sensitivity, resolution, mass accuracy, MS/MS capability, and rapid analysis, it plays a crucial role in natural product characterization (Zhang et al., 2017). This study employed UHPLC-Q-TOF-MS/MS for rapid detection and identification of chemical constituents in *C. convolvulacea* var. *pinifolia* roots to: (1) reveal the major chemical constituents and their structural types, and (2) elucidate the primary fragmentation patterns of these constituents during LC-MS analysis.

Materials and Methods

1.1 Instruments and Reagents The UHPLC-Q-TOF-MS/MS system consisted of a SCIEX Triple TOF 6660+ high-resolution time-of-flight mass spectrometer (ESI source, AB SCIEX, MA, USA) and an SCIEX Exion LC AD UHPLC system (AB SCIEX, MA, USA). The UHPLC system comprised two LC-20AD pumps, one LC-20AB pump, a DGU-20A3 degassing unit, an SIL-20AC autosampler with cooling, a CBM-20A system controller, a CTO-20A column oven, and an SPD-M20A diode array detector. Additional equipment included a KQ- (OHAUS, NJ, USA), a Milli-Q ultrapure water system (Millipore, USA), an SHZ-D( ) circulating water vacuum pump (Shanghai Lichenbang Instrument Technology Co., Ltd., China), and a Buchi R-20 rotary evaporator (Buchi, Switzerland). Ammonium acetate (Shanghai Aladdin Biochemical Technology Co., Ltd., China, batch No. G2222070, MS grade), methanol, formic acid, and acetonitrile (Fisher, USA) were all of MS grade.

The Tibetan medicine Codonopsis Convolvulaceae Radix was collected by senior Tibetan medicine engineer Qiangba from Tibet Qizheng Tibetan Medicine Co., Ltd. at Longkaikou Town, Heqing County, Dali, Yunnan Province, and identified by Associate Professor Xu Feng from Peking University School of Pharmaceutical Sciences as the roots of *Codonopsis convolvulacea* var. *pinifolia* (Campanulaceae) (Flora of China Editorial Committee, 1983). Reference standards of raffinose (L100210-5 g) and 1,2,3,4-tetrahydro-β-carboline-3-carboxylic acid (B066440-250 mg) were purchased from Beijing Hongbai Technology Co., Ltd.; L-phenylalanine (L115406-25 g), vanillin (BL017414-25 g), palmitic acid (B064315-100 g), and L-tryptophan (B074003-25 g) from Beijing Kaiguo Technology Co., Ltd.; (-)-secoisolariciresinol (MUST-24042513, 5 mg) from Chengdu Manster Biotechnology Co., Ltd.; and tangshenoside I (PSD240905-376, 5 mg), lobetyolin (PS012897, 10 mg), syringaresinol (PS011627, 10 mg), and wogonin (PS011541, 5 mg) from Chengdu Push Bio-Technology Co., Ltd. All reference standards had purities >98%.

1.2 Sample Preparation 1.2.1 Preparation of Test Sample Solution Approximately 2 g of coarse *C. convolvulacea* var. *pinifolia* root powder was accurately weighed and extracted with 60 mL of pure methanol by ultrasonication at 25 °C for 1 h. The mixture was filtered, and the residue was re-extracted under identical conditions. The combined filtrates were concentrated to dryness and reconstituted in 2 mL of pure methanol, then filtered through a 0.22 μm microporous membrane to obtain the test sample solution for LC-MS analysis.

1.2.2 Preparation of Reference Standard Solution
Each reference standard (1 mg) was accurately weighed and dissolved in 1 mL methanol to prepare 1 mg·mL^{-1} stock solutions. Aliquots (100 μL) of each standard were combined to prepare a mixed reference standard solution containing 11 compounds, which was filtered through a 0.22 μm microporous membrane prior to analysis.

1.3 LC-MS Analysis Conditions Chromatographic separation was performed on a Waters ACQUITY UPLC BEH C18 column (1.7 µm, 2.1 × 150 mm) maintained at 40 °C with an injection volume of 3 µL. The mobile phase consisted of 0.1 mmol · L^{-1} ammonium acetate aqueous solution (A) and acetonitrile (B) at a flow rate of 0.3 mL · min^{-1} using the following gradient: 0–5 min, 3% B; 5–45 min, 3–32% B; 45–52 min, 32–62% B; 52–57 min, 62–81% B; 57–62 min, 81–100% B; 62–67 min, 100% B.

Mass spectrometry parameters: ESI source in negative ion mode; Gas1 (nebulizer gas) 60 psi; Gas2 (heater gas) 60 psi; Gas3 (curtain gas) 35 psi; ion source temperature (TEM) 600 °C; ion spray voltage (IS) -4,500 V; declustering potential 60/-60 V; collision energy (35$±$15) eV; scan range m/z 100–1,800 (MS) and m/z 50–1,800 (MS/MS).

1.4 Data Processing and Structural Elucidation Mass spectrometric data were analyzed using PeakView v.1.2 software. For compounds with available reference standards, retention times, quasi-molecular ions, and fragment ion information were compared with those of the standards for confirmation. For compounds without reference standards, structures were tentatively identified based on retention time, quasi-molecular ions, and fragment ions, with reference to literature-reported mass spectrometric fragmentation patterns and SciFinder database searches.

Results and Analysis

2.1 Common Neutral Losses in Mass Spectrometry In this study, common neutral losses observed in the mass spectra included 18.01 Da (H_2O), 27.99 Da (CO), 15.02 Da (CH_3), 30.01 Da (CH_2O), 42.01 Da (C_2H_2O), 43.99 Da (CO_2), and 60.02 Da ($C_2H_4O_2$), indicating the presence of hydroxyl groups (lost as H_2O), carbonyl groups, methyl groups, formaldehyde groups, acetyl groups, carboxyl groups or lactones, and acetate groups in the molecular structures.

2.2 UHPLC-Q-TOF-MS/MS Analysis of Chemical Constituents in *C. convolvulacea* var. *pinifolia* Roots Analysis of the test sample solution under the conditions described in Section 1.2 yielded the base peak chromatogram (BPC) in negative ion mode shown in [Figure 3: see original paper], with 56 chemical constituents identified. Detailed information is presented in .

[Figure 3: see original paper] Base peak chromatogram of the roots of *Codonopsis convolvulacea* var. *pinifolia* in negative ion detection mode

Qualitative analysis results of 56 chemical constituents of the roots of *Codonopsis convolvulacea* var. *pinifolia* by UHPLC-Q-TOF-MS

2.3 Identification by Compound Class 2.3.1 Identification of Nitrogen-Containing Compounds

All nitrogen-containing compounds were categorized as one class, with six such compounds (A1–A6) preliminarily identified from *C. convolvulacea* var. *pinifolia*. Compounds A1, A3, and A5 were identified as L-phenylalanine, L-tryptophan, and 1,2,3,4-tetrahydro-β-carboline-3-carboxylic acid, respectively, by comparing their mass spectrometric data with purchased reference standards and SciFinder database searches.

In the negative-ion primary mass spectrum, compound A4 exhibited a quasi-molecular ion peak $[M-H]^-$ at m/z 406.1721, with a predicted molecular formula of $C_{17}H_{29}NO_{10}$. Its MS/MS spectrum showed fragment ions at m/z 244.1190, 226.1085, and 208.0979, corresponding to consecutive losses of 162.05 Da ($C_6H_{10}O_5$) and two 18.01 Da (H_2O) units from the quasi-molecular ion. Additional fragments at m/z 85.0295 and 59.0139 were observed from m/z 244.1190 after loss of 159.08 Da ($C_7H_{13}NO_3$) and 26.01 Da (C_2H_2). Fragment ion m/z 226.1085 lost 155.09 Da ($C_8H_{13}NO_2$) and 27.99 Da (CO) to produce m/z 71.0139 and 198.1136. Based on this fragmentation pattern and SciFinder database searches, compound A4 was tentatively identified as β-D-ribo-hexopyranoside, 2-propen-1-yl 2-(acetylamino)-2,3-dideoxy-4-O-β-D-galactopyranosyl or its isomer.

Compound A6 showed a quasi-molecular ion peak $[M-H]^-$ at m/z 485.1785, with a predicted molecular formula of $C_{21}H_{30}N_2O_{11}$. Its MS/MS spectrum revealed fragment ions at m/z 467.1671, 305.1143, 189.0676, and 161.0720, corresponding to sequential losses of 18.01 Da (H_2O), 162.05 Da ($C_6H_{10}O_5$), 116.10 Da ($C_5H_8O_3$), and 27.99 Da (CO). Additional fragments at m/z 323.1249, 203.0826, and 175.0877 resulted from consecutive losses of 162.05 Da ($C_6H_{10}O_5$), 120.04 Da ($C_4H_8O_4$), and 27.99 Da (CO). The proposed fragmentation pathway is illustrated in [Figure 4: see original paper]. Consequently, compound A6 was identified as tatarine C-4 -O-β-D-glucopyranoside or its isomer.

[Figure 4: see original paper] Proposed fragmentation pathways of tatarine C-4 -O-β-D-glucopyranoside (A6)

2.3.2 Identification of Phenylpropanoids

Phenylpropanoids are natural compounds consisting of a benzene ring connected to a three-carbon straight chain (C_6–C_3 unit), typically possessing phenolic structures. These compounds exhibit high response in negative ion detection mode, with characteristic losses of CO_2, CH_3, H_2O, and $C_6H_{10}O_5$ in their mass spectra. Six phenylpropanoids (A10–A15) were preliminarily identified from *C. convolvulacea* var. *pinifolia*.

Compound A10 displayed a quasi-molecular ion peak $[M-H]^-$ at m/z 677.2298, with a predicted molecular formula of $C_{29}H_{42}O_{18}$. Its MS/MS spectrum showed fragment ions at m/z 497.1664, 453.1766, and 291.1238, corresponding to consecutive losses of 180.06 Da ($C_6H_{12}O_6$), 43.99 Da (CO_2), and 162.05 Da ($C_6H_{10}O_5$). An additional fragment at m/z 323.0984 resulted from loss of 354.13 Da ($C_{17}H_{22}O_8$). Comparison with a reference standard confirmed compound A10 as tangshenoside I.

Compound A12 exhibited a quasi-molecular ion peak $[M-H]^-$ at m/z 371.1339 ($C_{17}H_{24}O_9$). Its MS/MS spectrum showed fragments at m/z 209.0829 ($[M-H-C_6H_{10}O_5]^-$), 194.0585 ($[M-H-CH_3]^-$), 179.0349 ($[M-H-CH_3-CH_3]^-$), and 161.0246 ($[M-H-CH_3-CH_3-CO]^-$). Based on SciFinder database searches and literature comparison (Liu et al., 2013), compound A12 was identified as tangshenoside II.

Compound A13 showed a quasi-molecular ion peak $[M-H]^-$ at m/z 469.1359 ($C_{21}H_{26}O_{12}$), with MS/MS fragments at m/z 325.0938 ($[M-H-C_6H_8O_4]^-$), 163.0406 ($[M-H-C_6H_8O_4-C_6H_{10}O_5]^-$), 119.0510 ($[M-H-C_6H_8O_4-C_6H_{10}O_5-CO_2]^-$), and 91.0553 ($[M-H-C_6H_8O_4-C_6H_{10}O_5-CO_2-CO]^-$). Compound A15 exhibited a quasi-molecular ion peak $[M-H]^-$ at m/z 823.2681 ($C_{38}H_{48}O_{20}$), with fragments at m/z 469.1351 ($[M-H-C_{17}H_{22}O_8]^-$), 265.0743 ($[M-H-C_{23}H_{30}O_{12}-C_2H_4O_2]^-$), and 235.0592 ($[M-H-C_{23}H_{30}O_{12}-C_2H_4O_2-CH_2O]^-$). Based on SciFinder database searches and literature comparison (Tang et al., 2023), compounds A13 and A15 were identified as tangshenoside V and tangshenoside VI, respectively. The detailed fragmentation pathway of tangshenoside V is shown in [Figure 5: see original paper].

[Figure 5: see original paper] Proposed fragmentation pathways of tangshenoside V (A13)

2.3.3 Identification of Lignans

Lignans are natural compounds formed by dimerization of two phenylpropanoid derivatives (C_6–C_3 monomers), some of which exist as glycosides in plant wood and resin. Lignans characteristically lose CH_2O, H_2O, CH_3, and CO in mass spectrometry. Twenty-two lignans (A7, A24–A29, A31, A34–A42, A55, A56) were preliminarily identified from *C. convolvulacea* var. *pinifolia*.

Compound A38 exhibited a quasi-molecular ion peak $[M-H]^-$ at m/z 417.1555 ($C_{22}H_{26}O_8$). Its MS/MS spectrum showed fragment ions at m/z 402.1320 and 387.1085, corresponding to consecutive losses of two 15.02 Da (CH_3) units. Additional fragments at m/z 191.0350, 223.0612, and 359.1136 resulted from losses of 196.07 Da ($C_{10}H_{12}O_4$), 164.04 Da ($C_9H_8O_3$), and 27.99 Da (CO) from m/z 387.1085, respectively. Comparison with a reference standard confirmed compound A38 as syringaresinol.

Compounds A26–A29 are isomers with a quasi-molecular ion peak $[M-H]^-$ at m/z 377.16 ($C_{20}H_{26}O_7$). Their MS/MS spectra showed fragment ions at m/z 359.15, 329.13, 283.09, and 121.02, corresponding to sequential losses of 18.01 Da (H_2O), 30.01 Da (CH_2O), 46.04 Da (C_2H_6O), and 162.06 Da ($C_{10}H_{10}O_2$). Additional fragments at m/z 195.06 ($[M-H-H_2O-C_{10}H_{12}O_2]^-$) and 180.04 ($[M-H-H_2O-C_{10}H_{12}O_2-CH_3]^-$) were also observed. Based on SciFinder database searches and literature comparison (Liu et al., 2016), compounds A26–A29 were identified as erythro-guaiacylglycerol-β-O-4 -dihydroconiferyl alcohol or its isomers.

Compound A34 showed a quasi-molecular ion peak $[M-H]^-$ at m/z 359.1497 ($C_{20}H_{24}O_6$), with MS/MS fragments at m/z 326.1163 ($[M-H-H_2O-CH_3]^-$),

311.0928 ([M−H−H_2O−CH_3−CH_3]$^-$), 299.0928 ([M−H−CH_2O−CH_3−CH_3]$^-$), and 269.0825 ([M−H−CH_2O−CH_3−CH_3−CH_2O]$^-$). Based on SciFinder database searches and literature comparison (Yin et al., 2023), compound A34 was identified as (+)-isolariciresinol or its isomer. Its detailed fragmentation pathway is shown in [Figure 6: see original paper].

Compounds A39–A42 are isomers with a quasi-molecular ion peak [M−H]$^-$ at m/z 525.19 ($C_{25}H_{34}O_{12}$). Their MS/MS spectra showed fragments at m/z 489.17 ([M−H−H_2O−H_2O]$^-$) and 311.12 ([M−H−H_2O−H_2O−$C_5H_{10}O_5$−CO−$C_9H_8O_2$]$^-$). Based on SciFinder database searches and literature comparison (Du et al., 2018), compounds A39–A42 were identified as 7R,8R-threo-4,7,9,9 -tetrahydroxy-3-methoxy-8-O-4 -neolignan-3 -O-β-D-glucoside or its isomers.

Compound A50 exhibited a quasi-molecular ion peak [M−H]$^-$ at m/z 523.2194 ($C_{26}H_{36}O_{11}$), with MS/MS fragments at m/z 361.1663 ([M−H−$C_4H_8O_4$−C_2H_2O]$^-$), 301.1094 ([M−H−$C_4H_8O_4$−C_2H_2O−CH_2O−CH_3−CH_3]$^-$), 223.0958 ([M−H−$C_4H_8O_4$−C_2H_2O−$C_8H_{10}O_2$]$^-$), and 179.0713 ([M−H−$C_4H_8O_4$−C_2H_2O−$C_8H_{10}O_2$−C_2H_4O]$^-$). Based on mass spectrometric data and SciFinder database searches, compounds A50–A52 were identified as icariside E3 or its isomers. The detailed fragmentation pathway is shown in [Figure 7: see original paper].

Compounds A55 and A56 are isomers with quasi-molecular ion peaks [M−H]$^-$ at m/z 419.17 ($C_{22}H_{28}O_8$). Their MS/MS spectra showed fragment ions at m/z 389.16, 181.05, 371.15, and 356.13, corresponding to sequential losses of 30.01 Da (CH_2O), 208.11 Da ($C_{12}H_{16}O_3$), 18.01 Da (H_2O), and 15.02 Da (CH_3). Additional fragments at m/z 341.13, 326.08, and 311.06 resulted from consecutive losses of three 15.02 Da (CH_3) units from m/z 356.13. No compounds matching these fragmentation patterns were found in the SciFinder database, suggesting they are isomers of the known compound 3-furanmethanol, tetrahydro-2-(4-hydroxy-3,5-dimethoxyphenyl)-4-[(4-hydroxy-3,5-dimethoxyphenyl)methyl]-, (2S,3R,4R) (CAS No.: 116498-58-9), representing two potential novel compounds. Their detailed fragmentation pathways are shown in [Figure 8: see original paper].

[Figure 6: see original paper] Proposed fragmentation pathways of (+)-isolariciresinol (A34)
[Figure 7: see original paper] Proposed fragmentation pathways of icariside E3 (A50–A52)
[Figure 8: see original paper] Proposed fragmentation pathways of two potential new compounds (A55 and A56)

2.3.4 Identification of Organic Acid Constituents

Organic acids typically fragment by losing neutral small molecules such as H_2O, CO_2, and C_2H_4. Eight organic acid compounds (A18–A23, A44, A45) were preliminarily identified from *C. convolvulacea* var. *pinifolia*.

Compound A18 exhibited a quasi-molecular ion peak [M−H]$^-$ at m/z 277.2668 ($C_{18}H_{30}O_2$), with MS/MS fragments at m/z 259.2067 ([M−H−H_2O]$^-$), 233.2275 ([M−H−CO_2]$^-$), 59.0139 ([M−H−$C_{16}H_{26}$]$^-$), 71.0139 ([M−H−$C_{15}H_{26}$]$^-$),

141.0921 ([M−H−$C_{10}H_{16}$]$^-$), and 127.0765 ([M−H−$C_{11}H_{18}$]$^-$). Based on SciFinder database searches and literature comparison (Olmo-García et al., 2018), compound A18 was identified as linolenic acid or its isomer. Its detailed fragmentation pathway is shown in [Figure 9: see original paper].

Compounds A19 and A20 are isomers with quasi-molecular ion peaks [M−H]$^-$ at m/z 327.21 ($C_{18}H_{32}O_5$), showing MS/MS fragments at m/z 229.14 ([M−H−$C_6H_{10}O$]$^-$), 221.13 ([M−H−$C_6H_{10}O$−H_2O]$^-$), 193.12 ([M−H−$C_6H_{10}O$−H_2O−H_2O]$^-$), and 185.11 ([M−H−$C_6H_{10}O$−CH_2O_2]$^-$). Based on SciFinder database searches and literature comparison (Ju et al., 2021), compounds A19 and A20 were identified as 9,12,13-trihydroxy-10,15-octadecadienoic acid or its isomers.

Compound A21 exhibited a quasi-molecular ion peak [M−H]$^-$ at m/z 329.2341 ($C_{18}H_{34}O_5$), with fragments at m/z 229.1461 ([M−H−$C_6H_{12}O$]$^-$), 211.1353 ([M−H−$C_6H_{12}O$−H_2O]$^-$), 193.1236 ([M−H−$C_6H_{12}O$−H_2O−H_2O]$^-$), and 171.1038 ([M−H−$C_6H_{12}O$−C_3H_6O]$^-$). Compound A22 showed [M−H]$^-$ at m/z 313.2392 ($C_{18}H_{34}O_4$), with fragments at m/z 295.2268 ([M−H−H_2O]$^-$), 277.2191 ([M−H−H_2O−H_2O]$^-$), 201.1139 ([M−H−C_8H_{10}]$^-$), and 183.1397 ([M−H−$C_7H_{12}O_2$]$^-$). Compound A44 exhibited [M−H]$^-$ at m/z 337.0932 ($C_{16}H_{18}O_8$), with fragments at m/z 191.0559 ([M−H−$C_9H_6O_2$]$^-$), 163.0405 ([M−H−$C_7H_{10}O_5$]$^-$), 127.0452 ([M−H−$C_9H_6O_2$−H_2O−HCOOH]$^-$), and 119.0499 ([M−H−$C_7H_{10}O_5$−CO_2]$^-$). Based on SciFinder database searches and literature comparison (Zhang et al., 2014), compound A21 was identified as 9,12,13-trihydroxyoctadecenoic acid or its isomer, compound A22 as 9,10-epoxy-18-hydroxyoctadecanoic acid or its isomer, and compound A44 as 3-p-coumaroylquinic acid or its isomer.

[Figure 9: see original paper] Proposed fragmentation pathways of linolenic acid (A18)

2.3.5 Identification of Other Constituents

Additional constituents identified included flavonoids, alkynes, and glycosides. Compound A16 exhibited a quasi-molecular ion peak [M−H]$^-$ at m/z 283.0622 ($C_{16}H_{12}O_5$), with MS/MS fragments at m/z 211.0401 ([M−H−$C_3H_4O_2$]$^-$) and 268.0377 ([M−H−CH_3]$^-$). Subsequent fragmentation of m/z 268.0377 yielded ions at m/z 239.0350 ([M−H−CH_3−CHO]$^-$), 163.0037 ([M−H−CH_3−CHO−C_6H_4]$^-$), and 135.0088 ([M−H−CH_3−CHO−C_6H_4−CO]$^-$). Comparison with a reference standard confirmed compound A16 as wogonin.

Compound A8 showed [M−H]$^-$ at m/z 395.1708 ($C_{20}H_{28}O_8$), with fragments at m/z 233.1183 ([M−H−$C_6H_{10}O_5$]$^-$), 215.1078 ([M−H−$C_6H_{10}O_5$−H_2O]$^-$), and 185.0972 ([M−H−$C_6H_{10}O_5$−H_2O−CH_2O]$^-$). Comparison with a reference standard identified compound A8 as lobetyolin, with compound A9 as its isomer. The detailed fragmentation pathway is shown in [Figure 10: see original paper].

Compound A32 exhibited [M−H]$^-$ at m/z 151.0408 ($C_8H_8O_3$), with fragments at m/z 136.0169, 108.0219, 95.0139, and 92.0274, identifying it as vanillin by comparison with a reference standard.

Compound A33 showed $[M-H]^-$ at m/z 503.1625 ($C_{18}H_{32}O_{16}$), with fragments at m/z 341.1088, 323.0985, 179.0555, and 113.0253, identifying it as raffinose by comparison with a reference standard.

Compound A43 exhibited $[M-H]^-$ at m/z 325.0930 ($C_{15}H_{18}O_8$), with fragments at m/z 161.0452 ($[M-H-C_9H_8O_3]^-$), 119.0500 ($[M-H-C_6H_{10}O_5-CO_2]^-$), 101.0380 ($[M-H-C_9H_8O_3-C_2H_4O_2]^-$), and 71.0161 ($[M-H-C_9H_8O_3-C_3H_6O_3]^-$). Based on SciFinder database searches and literature comparison (Sun et al., 2015), compound A43 was identified as p-coumaroylglucose or its isomer.

Compound A46 showed $[M-H]^-$ at m/z 293.1250 ($C_{12}H_{22}O_8$), with fragments at m/z 173.0774 ($[M-H-C_4H_8O_4]^-$), 143.0322 ($[M-H-C_4H_8O_4-C_2H_6]^-$), 131.0719 ($[M-H-C_6H_{10}O_5]^-$), and 101.0229 ($[M-H-C_4H_8O_4-C_2H_6-C_2H_2O]^-$). Based on its fragmentation pattern, compound A46 was identified as ethyl (3S)-3-(β-D-glucopyranosyloxy)butanoate or its isomer, with the detailed pathway shown in [Figure 11: see original paper].

Compounds A53 and A54 are isomers with $[M-H]^-$ at m/z 507.24 ($C_{23}H_{40}O_{12}$), showing fragments at m/z 447.23 ($[M-H-C_2H_4O_2]^-$), 315.18 ($[M-H-C_7H_{12}O_6]^-$), 161.04 ($[M-H-C_{17}H_{30}O_7]^-$), and 143.03 ($[M-H-C_{17}H_{30}O_7-H_2O]^-$). Based on this fragmentation pattern, compounds A53 and A54 were identified as 6-O-10-undecenoyltrehalose (CAS Registry No.: 151368-80-8) or its isomers, with the detailed pathway shown in [Figure 12: see original paper].

[Figure 10: see original paper] Proposed fragmentation pathways of lobetyolin (A8)
[Figure 11: see original paper] Proposed fragmentation pathways of ethyl (3S)-3-(β-D-glucopyranosyloxy)butanoate (A46)
[Figure 12: see original paper] Proposed fragmentation pathways of 6-O-10-undecenoyltrehalose (A53 and A54)

Discussion and Conclusion

Preliminary experiments revealed that chemical constituents in *C. convolvulacea* var. *pinifolia* roots exhibited higher ion intensities and more chromatographic peaks in negative ion mode than in positive ion mode. Therefore, mass spectrometric data obtained in negative ion mode were used for chemical constituent analysis.

Chemical investigations of Codonopsis Convolvulaceae Radix remain limited. Previous studies include: Chen et al. (2000) isolated and identified 9 compounds from the dried tuberous roots of *C. convolvulacea*; Han et al. (2001) identified 5 compounds from the dried roots using silica gel chromatography; Chen et al. (2001) isolated 4 phytosterol glycosides from Tibetan medicine "Niwa" (roots of *C. convolvulacea*); Sun et al. (2009) determined lobetyolin content in various *Codonopsis* species by HPLC; and Wu (2009) isolated 5 compounds from *C.*

macrocalyx roots. In total, 27 compounds have been identified from the six botanical sources of Codonopsis Convolvulaceae Radix. Additionally, Zhong et al. (2002) analyzed nutritional components, finding 17 amino acids, 5 vitamins, and 10 inorganic elements.

This study represents the first systematic qualitative analysis of *C. convolvulacea* var. *pinifolia*, one of the botanical sources of Codonopsis Convolvulaceae Radix. Using UHPLC-Q-TOF-MS/MS, we identified 56 constituents for the first time, with 11 confirmed by reference standards. All 56 compounds are newly reported from *C. convolvulacea* var. *pinifolia*. Except for L-phenylalanine, syringaresinol, and lobetyolin, the remaining 53 compounds are newly identified from all six botanical sources of Codonopsis Convolvulaceae Radix. Furthermore, two potential novel compounds (A55 and A56) were discovered. These findings substantially expand the chemical knowledge of *C. convolvulacea* var. *pinifolia*.

This study also provides, for the first time, detailed fragmentation pathways for multiple chemical constituents, offering an important reference for rapid chemical analysis of *Codonopsis* species. The results demonstrate that LC-MS combined with comprehensive data analysis enables efficient discovery and identification of constituents in Tibetan medicines with limited chemical research background. This work establishes a solid foundation for future studies on pharmacodynamic substances, quality evaluation, and in vivo processes of Codonopsis Convolvulaceae Radix.

References

CHEN QH, YANG PQ, LIU WJ, 2000. Studies on the chemical constituents of Convolvulate Asiabell (*Codonopsis convolvulacea*) [J]. *Chinese Traditional and Herbal Drugs*, 31(2): 84-86.

CHEN QH, YANG PQ, LIU WJ, 2001. Phytosterol glycosides from traditional Tibetan medicinal herb Niwa (*Condonopsis convolvulacea*) [J]. *West China Journal of Pharmaceutical Sciences*, 16(4): 245-247.

DU YE, LEE JS, KIM HM, et al., 2018. Chemical constituents of the roots of *Codonopsis lanceolate* [J]. *Archives of Pharmacal Research*, 41(11): 1082-1091.

Flora of China Editorial Committee in Chinese Academy of Sciences, 1983. *Flora Reipublicae Popularis Sinicae: Vol. 73 Book 2* [M]. Beijing: Science Press: 66-69.

HAN GX, GU L, YIN JS, et al., 2001. Studies on chemical constituents in roots of *Codonopsis convolvulacea* Kurz. [J]. *The Journal of Pharmaceutical Practice*, 19(3): 174-175.

JU LL, ZHANG JX, WANG FJ, et al., 2021. Chemical profiling of *Houttuynia cordata* Thunb. by UPLC-Q-TOF-MS and analysis of its antioxidant activity

in C2C12 cells [J]. *Journal of Pharmaceutical and Biomedical Analysis*, 204: 14271.

LIU MH, TONG X, WANG JX, et al., 2013. Rapid separation and identification of multiple constituents in traditional Chinese medicine formula Shenqi Fuzheng Injection by ultra-fast liquid chromatography combined with quadrupole-time-of-flight mass spectrometry [J]. *Journal of Pharmaceutical and Biomedical Analysis*, 74: 141-155.

LIU Y, YANG G, FENG F, 2016. Integrated chemical profiling of Zhi-Zi-Hou-Po decoction by liquid chromatography-diode array detector-time of flight mass analyzer and liquid chromatography-triple stage quadrupole mass analyzer combined with chemometrics [J]. *Analytical Methods-UK*, 8(23): 4689-4710.

National Administration of Traditional Chinese Medicine Editorial Board of Chinese Materia Medica, 2002. *Chinese Materia Medica: Vol. Tibetan Medicines* [M]. Shanghai: Shanghai Science and Technology Press: 203-204.

Northwest Institute of Plateau Biology of Chinese Academy of Sciences, 1991. *Tibetan medicine records* [M]. Xining: Qinghai People's Publishing House: 98-100.

OLMO-GARCÍA L, POLARI JJ, LI XQ, et al., 2018. Deep insight into the minor fraction of virgin olive oil by using LC-MS and GC-MS multi-class methodologies [J]. *Food Chemistry*, 261: 184-193.

Qinghai Provincial Institute of Drug Control, Qinghai Institute of Tibetan Medicine, 1996. *Chinese Tibetan medicine: Vol. 3* [M]. Shanghai: Shanghai Science and Technology Press: 103-104.

Sichuan Medical Products Administration, 2021. *Sichuan provincial standards for Tibetan medicinal materials (2020 edition)* [M]. Chengdu: Sichuan Science and Technology Press: 231-234.

SUN HY, LIU MX, LIN ZT, et al., 2015. Comprehensive identification of 125 multifarious constituents in Shuang-huang-lian powder injection by HPLC-DAD-ESI-IT-TOF-MS [J]. *Journal of Pharmaceutical and Biomedical Analysis*, 115(11): 86-106.

SUN QW, HUANG M, HE SZ, 2009. Determination of lobetyolin in 7 species plants of *Codonopsis* by HPLC [J]. *West China Journal of Pharmaceutical Sciences*, 24(3): 290-292.

SUN XL, YUAN F, JIN ZS, et al., 2019. Research progress on medicinal plant of *Codonopsis convolvulacea* Kurz. [J]. *Modern Agricultural Sciences and Technology*, (15): 69-71.

TANG XY, ZENG JX, WANG XX, et al., 2023. Chemical and metabolic profiling of *Codonopsis Radix* extract with an integrated strategy using ultra-high-performance liquid chromatography coupled with mass spectrometry [J]. *Journal of Separation Science*, 46(2): 2200723.

WU L, 2009. Studying on chemical constituents of root in *Codonopsis macrocalyx* Diels [D]. Chengdu: Southwest Jiaotong University: 1-58.

YIN JX, LI C, ZHANG J, et al., 2023. Comprehensive multicomponent characterization and quality assessment of Shuang-Huang-Lian powder injection using ultra-high-performance liquid chromatography-quadrupole time-of-flight-mass spectrometry ultra-high-performance liquid chromatography-quadrupole-Orbitrap-mass spectrometry [J]. *Rapid Communications in Mass Spectrometry*, 37(7): e9479.

ZHANG L, TU ZC, YUAN T, et al., 2014. Solvent optimization, antioxidant activity, and chemical characterization of extracts from *Artemisia selengnesis* Turcz. [J]. *Industrial Crops and Products*, 56: 223-230.

ZHANG Y, FENG BM, LU X, 2017. Research progress on application of UPLC/Q-TOF-MS in pharmaceutical analysis [J]. *Natural Product Research and Development*, 29(11): 1992-1996.

ZHONG HM, YUAN J, ZHANG SS, et al., 2002. Analysis of nutritional components of natural plant *S. indicum* Linn. and *Codonopsis convolvulacea* Kurz. [J]. *Journal of Yunnan University*, 24(6): 457-458.

Note: Figure translations are in progress. See original paper for figures.