Available online at www.sciencedirect.com

ScienceDirect

*Chinese Journal of Natural Medicines* 2019, **17**(8): 0631–0640

Chinese Journal of Natural Medicines

# Profiling and identification of aqueous extract of *Cordyceps sinensis* by ultra-high performance liquid chromatography tandem quadrupole-orbitrap mass spectrometry

YAO Chang-Liang[1, 3], QIAN Zheng-Ming[2], TIAN Wen-Shuai[1], XU Xiao-Qian[1], YAN Yu[1], SHEN Yao[1], LU Song-Mao[1], LI Wen-Jia[2*], GUO De-An[1, 3*]

[1] *R&D Department, GenChim Testing (Shanghai) Co., Ltd., Shanghai 200131, China;*

[2] *Key Laboratory of State Administration of Traditional Chinese Medicine, Sunshine Lake Pharma Co. Ltd., Guangdong 523850, China;*

[3] *Shanghai Research Center for Modernization of Traditional Chinese Medicine, National Engineering Laboratory for TCM Standardization Technology, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China*

**[ABSTRACT]** Characterization of aqueous extract in traditional Chinese medicine (TCM) is challenging due to the poor retention of the analytes on conventional $C_{18}$ columns. This study presents a systematic characterization method based on a rapid chromatographic separation (8 min) on a polar-modified $C_{18}$ (Waters Cortecs T3) column of aqueous extract of *Cordyceps sinensis*. UHPLC-HRMS method was used to profile components in both untargeted and targeted manners by full MS/PIL/dd-MS$^2$ acquisition approach. The components were identified or tentatively identified by reference standards comparison, fragmentation rules elucidation and available databases search. A total of 91 components, including 10 nucleobases, 20 nucleosides, 39 dipeptides, 18 amino acids and derivatives and 4 other components, were characterized from the aqueous extract of *C. sinensis*. And this was the first time to systematically report the presence of nucleosides and dipeptides in *C. sinensis*, especially for modified nucleosides. The chemical basis inquiry of this work would be beneficial to mechanism exploration and quality control of *C. sinensis* and related products. Meanwhile, this work also provided an effective solution for characterization of aqueous extract in TCM.

**[KEY WORDS]** Aqueous extract; *Cordyceps sinensis*; Dipeptides; Nucleosides

**[CLC Number]** R917 **[Document code]** A **[Article ID]** 2095-6975(2019)08-0631-10

## Introduction

Full characterization of the chemical components in traditional Chinese medicines (TCMs) is a premise for the follow-up research, such as exploring the therapeutic basis and establishing feasible quality control standards. A lot of efforts have been made to successfully probe various compounds, including saponins [1], alkaloids [2], flavonoids [3], and terpenoids [4], in a single herb or Chinese patent medicines. High performance liquid chromatography tandem mass spectrometry (HPLC-MS) stands out among all the methods due to its high sensitivity, specificity and abundant structural information [5]. However, there are still some bottlenecks hindering the analysts: (1) poor retention of polar compounds on traditional reverse phase (RP) columns [6-7]; (2) lack in methods for characterization of trace components [2]; (3) laborious and manual interpretation of big MS data [1].

Aqueous extract was the main usage form of TCMs [8], and conventional $C_{18}$ columns are most widely used in TCM analysis. However, conventional $C_{18}$ columns often suffered from phase collapse in highly aqueous conditions. Hydrophilic interaction chromatography (HILIC) offered strong retention for polar analytes, but the wide use was restricted by its relatively low stability, poor resolution and the difficulty in characterizing the compounds due to its orthogonal selectivity to RPLC [9]. Recently, serially coupled

---

**[Received on]** 07-May-2019

**[Research funding]** This work was supported by National Key R&D Program of China (Nos. 2018YFC1707900 and 2018YFC1707300).

**[*Corresponding authors]** Tel: 86-21-50271516, E-mail: liwenjiapaper@163.com (LI Wen- Jia); E-mail: daguo@simm.ac.cn (GUO De-An)

These authors have no conflict of interest to declare.

RPLC-HILIC has been successfully developed to simultaneously separate both hydrophilic and hydrophobic components [6-7]. Another option was polar-modified $C_{18}$ columns, including polar-embedded and polar-endcapped $C_{18}$ columns, with the advantages of high stability under highly aqueous conditions, relatively similar selectivity compared to conventional $C_{18}$ columns, and improved peak shapes for basic compounds [10].

*Cordyceps sinensis* is a distinguished and reputable TCM, also known as “winter worm summer grass” (Dong Chong Xia Cao). *C. sinensis* has been traditionally used for centuries in China and showed a wide spectrum of biological functions including anti-aging, reparative, anti-tumor, immune-stimulation and antioxidant [11]. Nucleosides and nucleobases, polysaccharides, sterols, amino acids and polypeptides were reported from *C. sinensis* and contributed to the above biological effects [12]. Among which, nucleosides and nucleobases are believed to be major components and acted as chemical markers for quality control of *C. sinensis* [13-14]. However, up to date, only about ten nucleosides and nucleobases were isolated from *C. sinensis* [15]. The studies mainly focused on the determination of nucleosides and nucleobases [13-14, 16-17], and the chemical diversity of nucleosides in *C. sinensis* has not been well probed. Although the aqueous extract of *C. sinensis* appeared to be a promising antitumor and toxicity-reduced agent [18], the composition has not been deeply explored thus far.

Herein, we report a chemical profiling solution combing nontargeted and targeted identification methods for aqueous extract of *C. sinensis*. The nontargeted method was applied to analyze *C. sinensis* in 8 min, and the undetected minor nucleobases and nucleosides were further analyzed by the targeted identification method. Through these two methods, a total of 78 and additional 13 compounds were characterized by UHPLC/Q-orbitrap, respectively. By means of this approach, a considerable amount of minor modified nucleosides were sensitively detected and primarily identified.

## Materials and Methods

### *Chemicals and reagents*

Twenty-three compounds including adenosine, guanosine, inosine, adenine, hypoxanthine and arginine were purchased from the Institute for the Control of Pharmaceutical and Biological Products of China (Beijing, China). Guanine, adenosine-5'-monophosphate, phenylalanine, tyrosine, leucine, isoleucine, methionine, valine, proline, Leu-Gly, Leu-Ala and vitamin B2 were purchased from Anpel Inc (Shanghai, China). Ile-Leu, Leu-Leu, Ile-Ile and Leu-Ile were purchased from Nanjing Peptide Biotech Co., Ltd. (Nanjing, China), and uridine from Shanghai Standard Biotech Co., Ltd. (Shanghai, China) in this study. And the structures of the reference standards were displayed in Fig. 1.

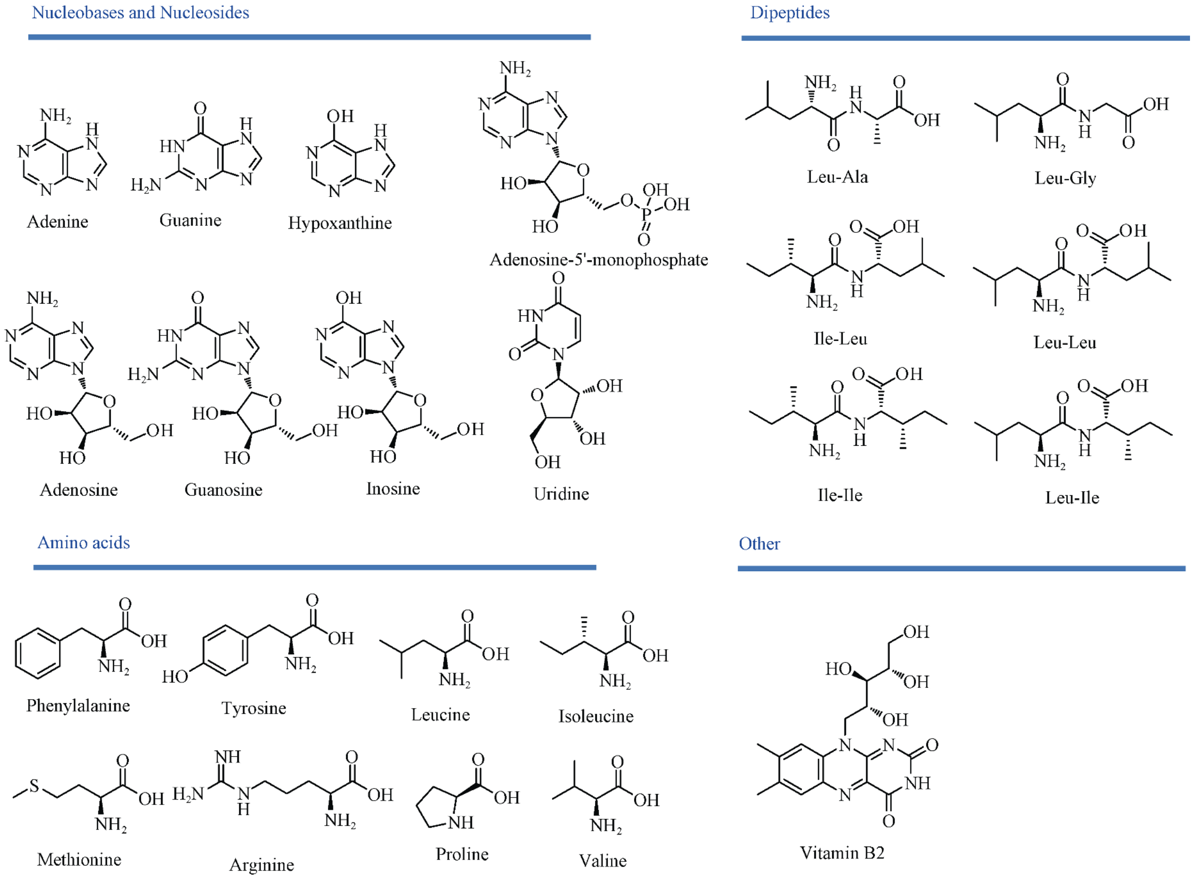


**Fig. 1 Chemical structures of 23 reference standards in this study**

Crude drug material of *C. sinensis* was artificially cultured and provided by Sunshine Lake Pharma Co., Ltd. (Dongguan, China) [19]. The sample was authenticated by Dr. MENG Qian-Wan [GenChim Testing (Shanghai) Co., Ltd.].

HPLC grade acetonitrile (Fisher scientific, Fair lawn, NJ, USA), LC/MS grade formic acid (Fisher scientific, Fair lawn, NJ, USA), and ultra-pure water in-house prepared by Milli-Q Advantage A10 water purification system (Millipore, Bedford, MA, USA) were applied in the mobile phases and sample preparation.

### *Sample preparation*

An aliquot of 0.2 g fine sample powder was accurately weighed and ultrasonically extracted in 10 mL ultra-pure water on a water bath (37 kHz) at room temperature for 60 min. The samples were centrifuged at 12 000 r·min$^{-1}$ for 10 min and the supernatant was stored at 4 °C prior to analysis.

### *UHPLC separation*

The separation of the extract was conducted on an Ultimate 3000 UHPLC system (Thermo Fisher Scientific, San Jose, CA, USA) equipped with a vacuum degasser, a binary pump, an autosampler, a diode array detector and a column compartment. An Waters CORTECS T3 column (2.1 mm × 100 mm, 1.6 μm) was used at 30 °C for chromatographic separation and eluted by the mobile phase consisting of acetonitrile (A) and water containing 0.1% formic acid (B) at a flow rate of 0.3 mL·min$^{-1}$ according to the following gradient elution program: 0−1 min, 1% A; 1−4 min, 1%−3% A; 4−6 min, 3%−30% A; 6−8 min 30% A. The DAD detector was set at 252 nm and between 190 and 400 nm. The sample injection volume was set at 1 μL. The total analysis time was 8 min for a single chromatographic run.

### *Q Exactive MS conditions*

A Q Exactive$^{\text{TM}}$ hybrid Q-orbitap mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) operated in positive mode was coupled to UHPLC system *via* a heated ESI (HESI) source for high-resolution data acquisition. HESI source parameters were set as follows: spray voltage, 3.5 kV; sheath gas pressure, 35 arb; aux gas pressure, 10 arb; sweep gas pressure, 0 arb, capillary temperature, 250 °C, aux gas heater temperature, 200 °C.

A full MS/data dependent (dd)-MS$^2$ method (injection 1) was firstly applied to untargeted profiling of various classes of compounds. The orbitrap scanned over *m/z* 100−500 at a resolution of 70 000 in full scan, and 17, 500 for MS$^2$ scan. AGC target values were set at $10^6$ and $10^5$ for full scan and MS$^2$ scan, respectively. Maximum injection times were 100 ms and 50 ms for full scan and MS$^2$ scan, respectively. The three most abundant precursors were selected to acquire MS/MS spectra fragmented by high-energy collision-induced dissociation (HCD) at normalized collision energy (NCE) 20%, 40% and 60%. Isolation window was set at 1.0 Da with an offset of 0.3 Da. An apex trigger of 2−4 s was defined to acquire the MS$^2$ fragments at high concentrations. Then a full MS/PIL/dd-MS$^2$ method (injection 2) was used to targeted profile the trace nucleosides and nucleobases. All the recorded modified bases (http: //modomics.genesilico.pl/) were imported into the precursor ion list (PIL) [20], and if idle "do not pick others" was on. The data was acquired and processed by Xcalibur 4.2 (Thermo Fisher Scientific, San Jose, CA, USA).

## Results and Discussion

### *Research strategy and advantages*

To analyze the high polar analytes efficiently in aqueous extract, a pure water tolerable core-shell column, Waters CORTECS T3 column was applied in this analysis. An 8 min gradient elution program was optimized and established. To get exact mass measurements and thereof successful elemental composition elucidations of precursor ions and product ions, the data was acquired by hybrid quadrupole-orbitrap mass spectrometer with high resolutions (70 000 for full scan and 17 500 for MS/MS). And to get abundant responses of the analytes, tune parameters including sheath gas pressure,

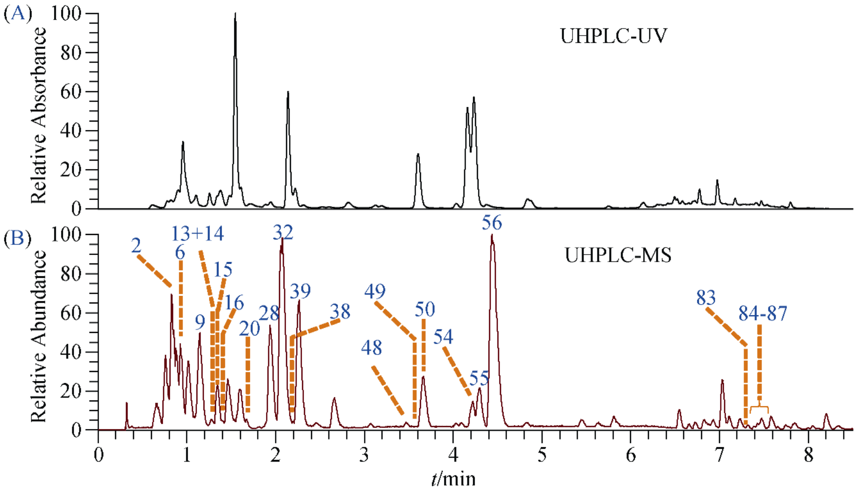


**Fig. 2 UHPLC-UV chromatogram acquired at 252 nm (A) and base peak chromatogram (B) of aqueous extract of *C. sinensis* obtained in positive mode on the UHPLC/Q Exactive-orbitrap instrument. The components identified by comparison with reference standards were annotated**

capillary temperature and aux gas heater temperature were optimized. Three different NCE (20%, 40% and 60%) were applied to simultaneously get a comprehensive MS/MS spectrum for various compounds ranging from the low mass to the high mass. A relatively narrow isolation window set at 1.0 Da was adopted to avoid potential interfering ions. Figs. 2A and 2B showed the UHPLC-UV chromatogram (252 nm) and base peak chromatogram of aqueous extract of *C. sinensis* obtained in the positive mode.

Recently, a strategy for simultaneously targeted and untargeted multicomponent characterization has been proposed and successfully applied to Erzhi pill based on the full MS/PIL/dd MS$^2$ function [21]. The strategy exhibited better performance in respect of identifying both the targeted components and untargeted ones. In this study, the aqueous extract of *C. sinensis* was firstly analyzed in an untargeted manner by full MS/dd-MS$^2$. Then, to get deeper knowledge of the main group of active compounds, targeted screening of nucleosides and nucleobases was conducted by PIL triggered MS$^2$ fragmentations, in which the reported modified nucleosides and nucleobases were included. To summarize, the detected compounds were identified by the following methods: (1) unambiguously identified by comparing with reference compounds; (2) characterized by fragmentation pathways and diagnostic product ions (DPIs) according to the related references; (3) tentatively identified by searching the databases, such as Human Metabolome Database (HMDB, http: //www.hmdb.ca), Modomics (http: //modomics.genesilico.pl) and MycompoundID (www.mycompoundid.org).

The superiority of the combined strategy was demonstrated by comparing the detected compounds. The nontargeted analysis (injection 1) offered multiple classes of compounds identified, including nucleosides, nucleobases, dipeptides, amino acids and others. While the targeted analysis (injection 2) only monitored nucleosides and nucleobases, an extra of 12 nucleosides (mainly methyl nucleosides) and 1 nucleobase were primarily characterized. Methyl nucleosides and dipeptides have been rarely reported in *C. sinensis*. Therefore, the combined strategy greatly revealed the chemical diversity of *C. sinensis*, with emphasis on the nucleosides and nucleobases.

Part of components ($t_R$: 0.5−1.8 min) were still not well retained, and serially coupled RPLC-HILIC may offer good separations for these compounds. RPLC-HILIC system showed a broader coverage of different compounds. However, modification of the instrument setup was necessary, and the compounds were eluted in more complex order. The polar-modified $C_{18}$ column provides a good choice for routine analysis of polar compounds.

### *Comprehensive characterization of multiple components*

Comparative analysis of the HCD features of 23 reference standards including three main structure subclasses (amino acid, dipeptide, nucleobase and nucleoside) was assisted further characterization of unknown components. By combing with database search method, 91 detected components in aqueous extract of *C. sinensis* were successfully characterized (Table 1).

**Table 1 Chemical components characterized from the aqueous extract of *C. sinensis***

| No | $t_R$ min | *m/z* | Molecular formula | Mass error (mDa) | MS/MS fragments | Identification |
|---|---|---|---|---|---|---|
| 1 | 0.81 | 189.1597 | $C_9H_{20}N_2O_2$ | –0.01 | 130.09, 84.08, 60.08 | Laminine |
| 2$^a$ | 0.83 | 175.1189 | $C_6H_{14}N_4O_2$ | –0.02 | 158.09, 130.10, 116.07, 70.07, 60.06 | Arginine |
| 3 | 0.83 | 156.0768 | $C_6H_9N_3O_2$ | 0.05 | 110.07$^c$, 95.06, 83.06 | Histidine |
| 4 | 0.87 | 147.0764 | $C_5H_{10}N_2O_3$ | 0.00 | 130.05, 84.05 | Glutamine |
| 5 | 0.88 | 118.0866 | $C_5H_{11}NO_2$ | 0.31 | 59.07, 58.07 | Betaine |
| 6$^a$ | 0.92 | 116.0709 | $C_5H_9NO_2$ | 0.34 | 70.07 | Proline |
| 7 | 0.92 | 266.1234 | $C_{10}H_{19}NO_7$ | –0.05 | 248.11, 230.10, 212.09, 194.08, 128.07, 116.07, 98.06 | GABA-*O*-Glucose |
| 8 | 1.06 | 306.0485 | $C_9H_{12}N_3O_7P$ | –0.07 | 178.06, 112.05 | Cytidine 2', 3'-cyclic monophosphate |
| 9$^a$ | 1.15 | 118.0865 | $C_5H_{11}NO_2$ | 0.28 | 72.08, 55.06 | Valine |
| 10 | 1.15 | 205.1184 | $C_8H_{16}N_2O_4$ | 0.13 | 72.08, 55.06 | Val-Ser |
| 11 | 1.27 | 280.1391 | $C_{11}H_{21}NO_7$ | –0.01 | 262.13, 244.12, 246.12, 198.11, 161.07, 130.09, 118.09, 84.08, 72.08 | Val-*O*-Glucose |
| 12 | 1.27 | 246.1449 | $C_{10}H_{19}N_3O_4$ | 0.05 | 229.12, 147.08, 130.05, 84.05, 72.08, 55.06 | Val-Gln |
| 13$^a$ | 1.3 | 136.0618 | $C_5H_5N_5$ | 0.04 | 119.04 | Adenine |
| 14$^a$ | 1.33 | 152.0567 | $C_5H_5ON_5$ | 0.04 | - | Guanine |
| 15$^a$ | 1.34 | 150.0584 | $C_5H_{11}NO_2S$ | 0.10 | 133.03, 104.05, 102.06, 87.03, 61.01, 56.05 | Methionine |
| 16$^a$ | 1.39 | 348.0705 | $C_{10}H_{14}N_5O_7P$ | 0.17 | 136.06, 97.03 | Adenosine-5'-monophosphate |
| 17 | 1.58 | 247.1288 | $C_{10}H_{18}N_2O_5$ | 0.00 | 148.06, 130.05, 102.06, 84.05, 72.08, 55.06 | Val-Glu |
| 18 | 1.59 | 169.0356 | $C_5H_4N_4O_3$ | 0.02 | 150.01, 141.04, 126.03, 98.04, 70.04 | Uric acid |

Continued

| No | $t_R$ min | *m/z* | Molecular formula | Mass error (mDa) | MS/MS fragments | Identification |
|---|---|---|---|---|---|---|
| 19 | 1.61 | 130.0500 | $C_5H_7NO_3$ | 0.10 | 84.05, 56.05 | Pyroglutamic acid |
| $20^a$ | 1.67 | 137.0459 | $C_5H_4N_4O$ | 0.08 | 119.04, 110.04 | Hypoxanthine |
| 21 | 1.7 | 246.1449 | $C_{10}H_{19}N_3O_4$ | 0.04 | 229.12, 133.06, 86.10 | Leu/Ile-Asn |
| 22 | 1.75 | 269.1609 | $C_{12}H_{20}N_4O_3$ | 0.01 | 251.15, 223.16, 178.13, 110.07, 93.05, 83.06 | His-Leu/Ile |
| $23^b$ | 1.8 | 258.1083 | $C_{10}H_{15}N_3O_5$ | –0.15 | 126.07 | 3-Methylcytidine |
| 24 | 1.81 | 205.1184 | $C_8H_{16}N_2O_4$ | 0.14 | 159.11, 118.09, 72.08, 60.05 | Ser-Val |
| 25 | 1.86 | 219.1339 | $C_9H_{18}N_2O_4$ | 0.00 | 86.10, 69.07 | Leu/Ile-Ser |
| 26 | 1.90 | 269.1609 | $C_{12}H_{20}N_4O_3$ | 0.04 | 251.15, 223.16, 178.13, 110.07, 93.05, 83.06 | His-Leu/Ile |
| 27 | 1.92 | 260.1969 | $C_{12}H_{25}N_3O_3$ | 0.03 | 243.17, 147.11, 129.10, 84.08 | Lys-Leu/Ile |
| $28^a$ | 1.94 | 132.1020 | $C_6H_{13}NO_2$ | 0.09 | 86.10, 69.07 | Isoleucine |
| 29 | 1.98 | 219.1340 | $C_9H_{18}N_2O_4$ | 0.09 | 86.10 | Leu/Ile-Ser |
| 30 | 2.00 | 153.0408 | $C_5H_4N_4O_2$ | 0.10 | 110.04 | Xanthine or isomer |
| $31^b$ | 2.06 | 282.1194 | $C_{11}H_{15}N_5O_4$ | –0.28 | 150.08 | 1-Methyladenosine |
| $32^a$ | 2.08 | 132.1020 | $C_6H_{13}NO_2$ | 0.14 | 86.10 | Leucine |
| 33 | 2.13 | 247.1288 | $C_{10}H_{18}N_2O_5$ | 0.00 | 86.10, 69.07 | Leu/Ile-Asp |
| 34 | 2.15 | 219.1340 | $C_9H_{18}N_2O_4$ | 0.09 | 201.12, 173.13, 155.12. 118.09, 74.06, 72.08, 56.05 | Thr-Val |
| 35 | 2.16 | 233.1497 | $C_{10}H_{20}N_2O_4$ | 0.08 | 215.14, 120.07, 86.10, 69.07 | Leu/Ile-Thr |
| 36 | 2.18 | 260.1969 | $C_{12}H_{25}N_3O_3$ | 0.03 | 243.17, 147.11, 129.10, 84.08 | Leu/Ile-Lys |
| 37 | 2.19 | 113.0349 | $C_4H_4N_2O_2$ | 0.35 | 96.01, 70.03 | Uracil isomer |
| $38^a$ | 2.19 | 245.0768 | $C_9H_{12}N_2O_6$ | 0.03 | 113.03, 96.01, 70.03 | Uridine |
| $39^a$ | 2.27 | 182.0812 | $C_9H_{11}NO_3$ | 0.05 | 165.05, 147.04, 136.08, 123.04, 119.05, 95.05, 91.05 | Tyrosine |
| 40 | 2.34 | 294.1546 | $C_{12}H_{23}NO_7$ | –0.13 | 276.14, 258.13, 248.15, 230.14, 212.13, 132.10, 86.10, 69.07 | Ile-*O*-Glucose |
| 41 | 2.43 | 233.1497 | $C_{10}H_{20}N_2O_4$ | 0.08 | 215.14, 120.07, 86.10, 74.06 | Leu/Ile-Thr |
| 42 | 2.45 | 294.1547 | $C_{12}H_{23}NO_7$ | –0.01 | 276.14, 258.13, 248.15, 230.14, 212.13, 132.10, 86.10, 69.07 | Leu-*O*-Glucose |
| 43 | 3.05 | 278.1249 | $C_{12}H_{15}N_5O_3$ | 0.11 | 260.11, 250.13, 208.11, 153.08, 112.05, 70.07, 60.06 | Qbase isomer |
| 44 | 3.07 | 203.1391 | $C_9H_{18}N_2O_3$ | 0.09 | 86.10, 69.07 | Leu/Ile-Ala |
| $45^b$ | 3.10 | 298.1148 | $C_{11}H_{15}N_5O_5$ | 0.03 | 166.07, 149.05 | 1-Methylguanosine |
| 46 | 3.25 | 252.0728 | $C_9H_9N_5O_4$ | 0.04 | 234.06, 206.07, 192.05, 188.06, 162.08, 148.06, 136.06 | Succinyladenine |
| $47^b$ | 3.38 | 258.1085 | $C_{10}H_{15}N_3O_5$ | 0.05 | 112.05 | 2'-*O*-methylcytidine |
| $48^a$ | 3.47 | 203.1391 | $C_9H_{18}N_2O_3$ | 0.09 | 86.10 | Leu-Ala |
| $49^a$ | 3.55 | 189.1234 | $C_8H_{16}N_2O_3$ | 0.06 | 86.10, 69.07 | Leu-Gly |
| $50^a$ | 3.66 | 268.1039 | $C_{10}H_{13}N_5O_4$ | –0.12 | 136.06 | Adenosine |
| 51 | 3.93 | 189.1234 | $C_8H_{16}N_2O_3$ | 0.06 | 86.10 | Leu/Ile-Gly |
| 52 | 4.03 | 217.1548 | $C_{10}H_{20}N_2O_3$ | 0.07 | 171.15, 118.09, 72.08, 55.06 | Val-Val |
| 53 | 4.09 | 252.1091 | $C_{10}H_{13}N_5O_3$ | 0.03 | 136.06, 117.05 | Cordycepin isomer |
| $54^a$ | 4.22 | 269.0880 | $C_{10}H_{12}N_4O_5$ | –0.10 | 137.05 | Inosine |
| $55^a$ | 4.3 | 284.0988 | $C_{10}H_{13}N_5O_5$ | –0.13 | 152.06, 135.03, 110.04 | Guanosine |
| $56^a$ | 4.44 | 166.0862 | $C_9H_{11}NO_2$ | –0.06 | 120.08, 103.05 | Phenylalanine |
| $57^b$ | 4.62 | 180.0881 | $C_7H_9N_5O$ | 0.11 | 110.04, 89.06 | PreQ1base isomer |
| 58 | 4.82 | 219.1340 | $C_9H_{18}N_2O_4$ | 0.09 | 201.12, 173.13, 132.10, 86.10, 60.05 | Ser-Leu/Ile |
| 59 | 4.89 | 268.1040 | $C_{10}H_{13}N_5O_4$ | 0.00 | 152.06, 135.03 | 2'-Deoxyaguanosine |
| 60 | 4.95 | 137.0459 | $C_5H_4N_4O$ | 0.08 | 119.04, 110.04, 94.04 | Hypoxanthine isomer |
| 61 | 5.10 | 189.1234 | $C_8H_{16}N_2O_3$ | 0.06 | 143.12, 132.10, 86.10, 75.06, 69.07 | Gly-Leu/Ile |

Continued

| No | $t_R$ min | *m/z* | Molecular formula | Mass error (mDa) | MS/MS fragments | Identification |
|---|---|---|---|---|---|---|
| 62 | 5.12 | 328.1392 | $C_{15}H_{21}NO_7$ | 0.08 | 310.13, 292.12, 264.12, 246.11, 178.09, 166.09, 132.08, 120.08 | Phe-*O*-Glucose |
| 63 | 5.17 | 233.1496 | $C_{10}H_{20}N_2O_4$ | 0.05 | 215.14, 187.14, 132.10, 86.10, 74.06, 56.05 | Thr-Leu/Ile |
| 64 | 5.44 | 203.1391 | $C_9H_{18}N_2O_3$ | 0.09 | 157.13, 132.10, 86.10 | Ala-Leu/Ile |
| 65 | 5.64 | 189.1234 | $C_8H_{16}N_2O_3$ | 0.06 | 143.12, 132.10, 86.10 | Gly-Leu/Ile |
| 66 | 5.84 | 233.1495 | $C_{10}H_{20}N_2O_4$ | −0.04 | 215.14, 187.14, 132.10, 86.10, 74.06, 56.05 | Thr-Leu/Ile |
| 67[b] | 5.98 | 282.1198 | $C_{11}H_{15}N_5O_4$ | 0.13 | 136.06 | 2'-*O*-methyladenosine |
| 68[b] | 6.3 | 298.1148 | $C_{11}H_{15}N_5O_5$ | 0.03 | 166.07 | *N*2-methylguanosine |
| 69[b] | 6.4 | 283.1039 | $C_{11}H_{14}N_4O_5$ | 0.20 | 137.05 | 2'-*O*-methylinosine |
| 70[b] | 6.4 | 298.1148 | $C_{11}H_{15}N_5O_5$ | 0.03 | 152.06 | 2'-*O*-methylguanosine |
| 71 | 6.49 | 198.0762 | $C_9H_{11}NO_4$ | 0.18 | 198.08, 181.05, 152.07, 139.04, 135.04, 111.04, 107.05 | Levodopa |
| 72[b] | 6.54 | 298.1149 | $C_{11}H_{15}N_5O_5$ | −0.01 | 166.07 | *N*6-hydroxymethyladenosine |
| 73 | 6.71 | 229.1550 | $C_{11}H_{20}N_2O_3$ | 0.36 | 70.07 | Pro-Leu/Ile |
| 74 | 6.72 | 231.1705 | $C_{11}H_{22}N_2O_3$ | 0.21 | 118.09, 86.10, 69.07 | Leu/Ile-Val |
| 75[b] | 6.81 | 312.1303 | $C_{12}H_{17}N_5O_5$ | 0.06 | 215.15, 180.09, 120.08, 110.07, 84.08, 70.07 | *N*2, *N*2-dimethylguanosine |
| 76 | 6.83 | 384.1153 | $C_{14}H_{18}N_5O_8$ | 0.34 | 252.07, 234.06, 206.07, 192.05, 162.08, 148.06, 136.06 | Succinyladenosine |
| 77 | 6.93 | 231.1706 | $C_{11}H_{22}N_2O_3$ | 0.28 | 118.09, 86.10, 72.08 | Val-Leu/Ile |
| 78 | 6.93 | 229.1551 | $C_{11}H_{20}N_2O_3$ | 0.41 | 116.07, 86.10, 70.07, 69.07 | Leu/Ile-Pro |
| 79[b] | 6.99 | 296.1356 | $C_{12}H_{17}N_5O_4$ | 0.23 | 164.09 | *N*6, *N*6-dimethyladenosine |
| 80 | 7.03 | 205.0972 | $C_{11}H_{12}N_2O_2$ | 0.11 | 188.07, 170.06, 159.09, 146.06, 118.07 | Tryptophan |
| 81 | 7.03 | 229.1549 | $C_{11}H_{20}N_2O_3$ | 0.24 | 116.07, 86.10, 70.07 | Leu/Ile-Pro |
| 82 | 7.11 | 231.1705 | $C_{11}H_{22}N_2O_3$ | 0.21 | 132.10, 86.10, 72.087, 55.06 | Val-Leu/Ile |
| 83[a] | 7.32 | 245.1862 | $C_{12}H_{24}N_2O_3$ | 0.22 | 199.18, 132.10, 86.10, 69.07 | Ile-Ile |
| 84[a] | 7.43 | 245.1862 | $C_{12}H_{24}N_2O_3$ | 0.22 | 199.18, 132.10, 86.10 | Leu-Ile |
| 85[a] | 7.48 | 245.1862 | $C_{12}H_{24}N_2O_3$ | 0.22 | 199.18, 132.10, 86.10, 69.07 | Ile-Leu |
| 86[a] | 7.52 | 377.1459 | $C_{17}H_{21}N_4O_6$ | 0.31 | 359.14, 243.09, 216.08, 198.07, 172.09 | Vitamin B2 |
| 87[a] | 7.58 | 245.1862 | $C_{12}H_{24}N_2O_3$ | 0.22 | 199.18, 132.10, 86.10 | Leu-Leu |
| 88 | 7.66 | 279.1707 | $C_{15}H_{22}N_2O_3$ | 0.33 | 233.17, 149.02, 132.10, 120.08, 103.05, 86.10 | Phe-Leu/Ile |
| 89 | 7.75 | 279.1705 | $C_{15}H_{22}N_2O_3$ | 0.18 | 166.09, 149.02, 120.08, 86.10, 69.07 | Leu/Ile-Phe |
| 90 | 7.85 | 279.1705 | $C_{15}H_{22}N_2O_3$ | 0.15 | 166.09, 149.02, 120.08, 103.05, 86.10 | Phe-Leu/Ile |
| 91[b] | 8 | 336.1666 | $C_{15}H_{21}N_5O_4$ | −0.03 | 279.17, 204.12, 177.10, 143.12, 136.06, 120.08, 86.10 | *N*6-isopentenyladenosine |

a: unambiguously identified by comparing with reference standards; b: data acquired by the targeted manner (injection 2);
c: the diagnostic product ions for characterizing the compound were underlined

*Characterization of nucleobases and nucleosides*

In total, 30 nucleobases and nucleosides were identified in the aqueous extract of *C. sinensis*, including 8 unambiguously identified by comparing with reference standards.

Firstly, the data acquired by the nontargeted analysis was analyzed. Nucleosides are composed of a nucleobase (purine base or pyrimidine base) and a five-carbon sugar (ribose or 2-deoxyribose). In positive mode, nucleosides and nucleobases were easily ionized into the protonated precursor ions ($[M + H]^+$). After elimination of sugar residues by dissociating glycosidic linkages, precursor ions of nucleosides produced DPIs associated with nucleobases, including *m/z* 112.05 for cytosine, 113.03 for uracil, 127.05 for thymine, 136.06 for adenine, 137.05 for hypoxanthine, and 152.06 for guanine. Neutral loss of 132.04 Da can be attributed to ribose, and 116.05 Da to 2-deoxyribose (**53** and **59**), respectively. In this way, most of the abundant nucleosides and nucleobases can be successfully characterized, except compounds **8**, **16**, **43**, **46** and **76**. Compound **16** (Fig. 3) was identified as adenosine-5'-monophosphate by comparing with the reference standard, DPI at *m/z* 136.06 and neutral loss of 212.01 Da (phosphatidyl ribose) were observed. Similarly, compound **8** was tentatively identified as cytidine 2', 3'-cyclic monophosphate or isomer, with DPI at *m/z* 112.05 (refer to cytosine) and neutral loss of 194.00 Da (2', 3'-cyclic phosphatidyl ribose). For compound **46** (Fig. 4), the molecular formula was determined as $C_9H_9N_5O_4$, and HCD of $[M + H]^+$ at *m/z* 252.07 generated product ions at *m/z* 206.07 ($[M + H - CH_2O_2]^+$), 192.05 ($[M + H - C_2H_4O_2]^+$), 162.08 ($[M + H - C_2H_2O_4]^+$) and 136.06 ($[M + H - C_4H_4O_4]^+$). Fragments at *m/z* 136.06 indicated an adenine residue, and the neutral losses of $C_4H_4O_4$, $C_2H_2O_4$, $C_2H_4O_2$ and $CH_2O_2$ all gave the evidences of a succinyl group. Therefore, compound **46** was tentatively identified as succinyladenine. For compound **76**, similar fragments were observed, and an extra of ribose group than compound **46** indicated the structure of succinyladenosine. Compound **43** was an isomer of Qbase, and the exact structure was not elucidated.

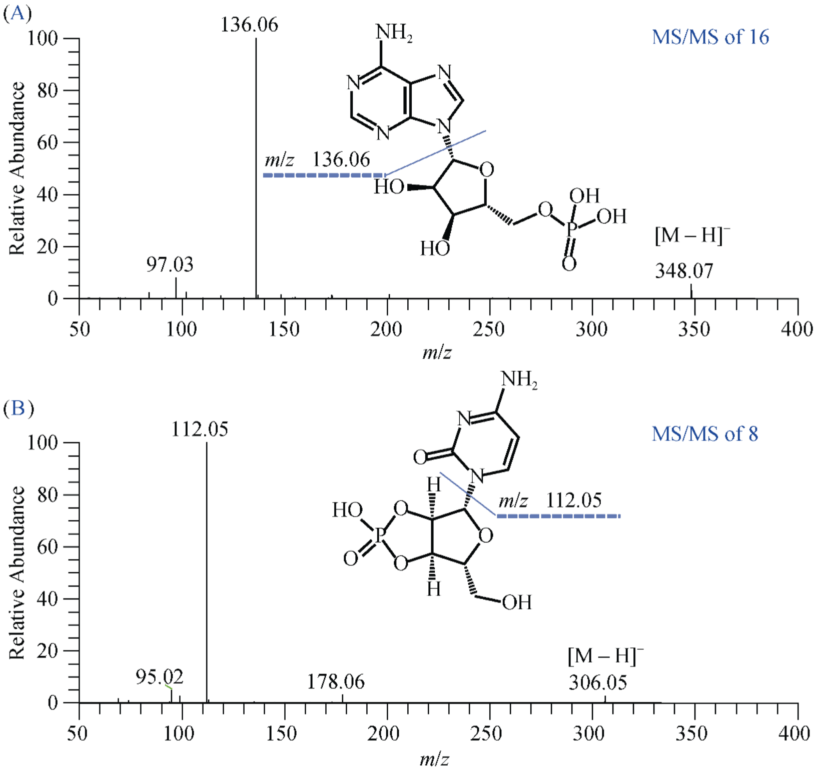


**Fig. 3 MS/MS spectra, proposed structures and fragmentation pathways for nucleoside phosphates: adenosine-5'-monophosphate (16, A) and cytidine 2', 3'-cyclic monophosphate (8, B)**

The modified nucleosides and nucleobases greatly enriched the chemical diversity of the aqueous extract of *C. sinensis*. Therefore, a targeted screening method was conducted by PIL function, and the reported modified nucleosides and nucleobases were imported into PIL. In this way, an extra of 1 nucleobase (**57**, preQ1base isomer) and 11 nucleosides were tentatively identified, including 10 methyl nucleosides (except compound **91**). Compound **91** was primarily identified as *N*6-isopentenyladenosine by searching the library, and the structure was further confirmed by the sequential neutral loss of 132.04 Da ($C_5H_8O_4$, ribose), 68.06 Da ($C_5H_8$, refer to isopentenyl) and DPI at *m/z* 136.06 (refer to adenosine). Compound **67** was an adenine glycoside with DPI at *m/z* 136.06, and the neutral loss of $C_6H_{10}O_4$ indicated a methyl ribose, and it was primarily identified 2'-*O*-methyladenosine by searching the databases. Compounds **45**, **68**, **70** and **72** are four isomers with the same precursor ions at *m/z* 298.11, and Modomics database searching feedback four compounds. They were attributed to 1-methylguanosine, *N*2-methylguanosine, 2'-*O*-methylguanosine and *N*6-hydroxymethyladenosine, respectively (Table 2), by DPI recognition and comparing *C*logP. Compound **70** (2'-*O*-methylguanosine) can be easily distinguished from others due to DPI at *m/z* 152.06. The other three compounds were primarily assigned by retention times and *C*logP [22], since compounds with larger *C*logP commonly eluted later in a RP separation system. In a similar way, two isomers (**23** and **47**) of methylcytidine were tentatively identified. Recently, HCD MS/MS spectra obtained at NCE of 80% were also successfully used to differentiate the nucleoside methylated positional isomers [23]. Compounds **75** and **80** were both dimethyl nucleosides, and an extra of two methyl groups than guanosine (*m/z* 180.09) and adenosine (*m/z* 164.09) were observed, respectively. Therefore compound **75** and **80** were separately tentatively identified as *N*2, *N*2-dimethylguanosine and *N*6, *N*6-dimethyladenosine, in accord with the fragments. In general, the targeted screening method greatly revealed the chemical diversity of modified nucleosides and nucleobases.

### *Characterization of amino acids and derivatives*

Amino acids are also main components of *C. sinensis* [24]. In total, 18 amino acids and derivatives were identified or tentatively identified. Eight amino acids were identified by

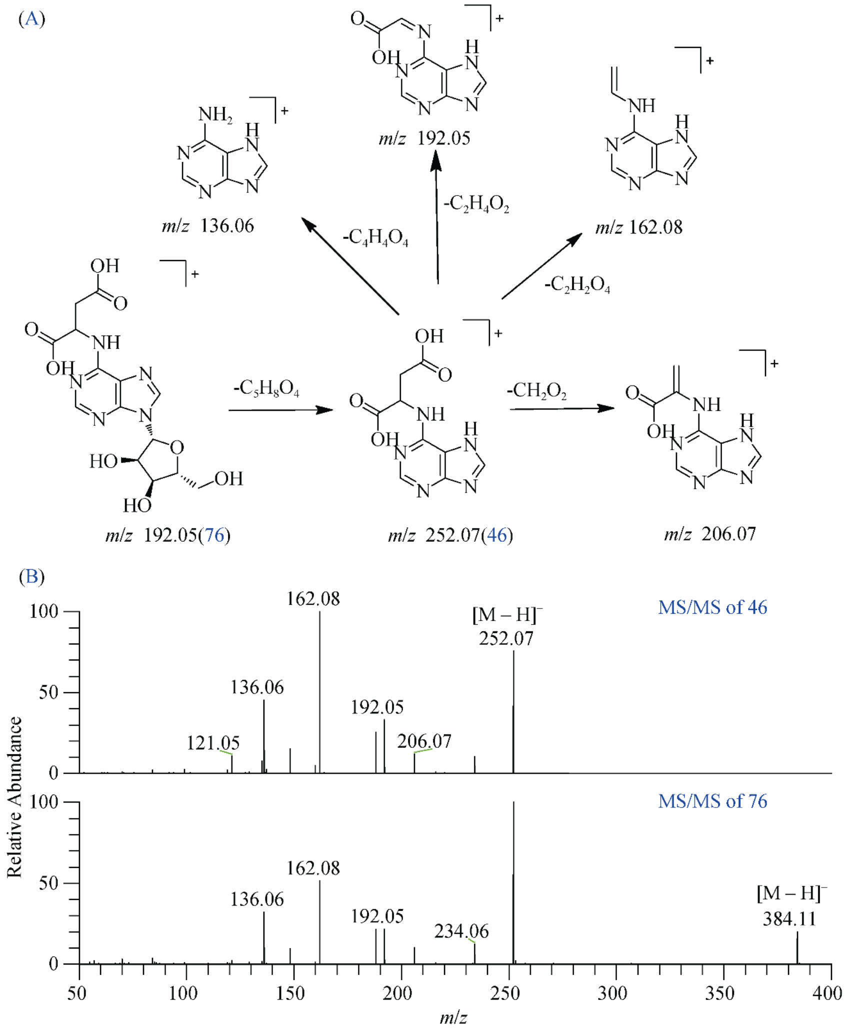
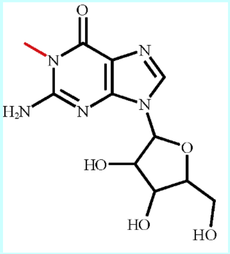
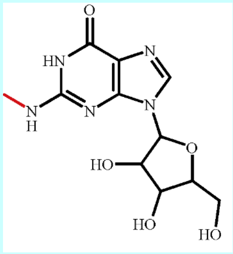
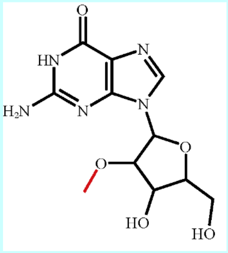
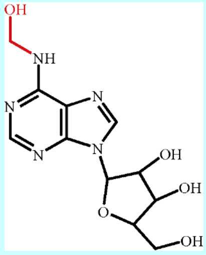


**Fig. 4 Proposed structures, fragmentation pathways (A) and MS/MS spectra (B) of compounds succinyladenine (46) and succinyladenosine (76)**

**Table 2 Structures, retention times, DPIs and *C*logP of four isomers**

| Compound | **45** | **68** | **70** | **72** |
|---|---|---|---|---|
| $t_R$/min | 3.10 | 6.30 | 6.40 | 6.72 |
| Structure |  |  |  |  |
| DPI | *m/z* 166.07 | *m/z* 166.07 | *m/z* 152.06 | *m/z* 166.07 |
| *C*logP | –3.01 | –2.32 | –2.34 | –2.19 |

comparing with reference standards. We take methionine (**15**) as an example, which has been confirmed by comparing with reference standards. HCD of precursor ions $[M + H]^+$ generated fragments at *m/z* 133.03, 104.05, 102.06, 87.03, 61.01 and 56.05 refer to neutral losses of $NH_3$, $CH_2O_2$, $CH_4S$, $CH_5NO_2$ and fragments of $C_2H_5S^+$ and $C_3H_6N^+$, respectively. Neutral losses of $NH_3$ and $CH_2O_2$ are common to the amino acids, except proline (an imino acid). Accordingly, two non-protein amino acids, pyroglutamic acid (**19**) and levodopa (**71**), were also detected and tentatively identified in the aqueous extract of *C. sinensis*. In addition, amino acid derivatives, five glucosides (**7**, **11**, **30**, **42** and **62**) were also detected and tentatively identified. For these compounds, characteristic neutral losses of a sugar, multiple $H_2O$ and CO were observed. Among which, **62** (phe-*O*-glucose) was one potential chemical marker to authenticate *C. sinensis* samples [25].

*Characterization of dipeptides*

Few reports on dipeptides of *C. sinensis* samples were available, and the reported dipeptides were also limited [26]. In positive mode, the MS/MS spectra of dipeptides showed product ions [amino acid + H]$^+$ and [amino acid + H − $CO_2H_2$]$^+$ of *C*-terminal amino acid, and [amino acid + H − $CO_2H_2$]$^+$ of *N*-terminal amino acid [27-28]. Therefore, [amino acid + H]$^+$ was characteristic to *C*-terminal residue. We take four isomers (**25**, **29**, **34** and **58**) with precursor ions [M + H]$^+$ at *m/z* 219.1339 for an example (Fig. 5). The molecular formulas were deduced as $C_9H_{18}N_2O_4$. For compound **58**, *m/z* 86.10 and 132.10 showed the existence of Leu/Ile at *C*-terminal, and *m/z* 60.05 indicated the existence of Ser at *N*-terminal, therefore compound **58** was tentatively identified as Ser-Leu/Ile. For compounds **25** and **29**, *m/z* 86.10 showed the existence of Leu/Ile at *N*-terminal and Ser was deduced at *C*-terminal according to the formula. And compounds **25** and **29** were both identified as Leu/Ile-Ser. Leu and Ile residues in oligopeptides can still not be differentiated. For compound **34**, *m/z* 72.08 and 118.09 showed the existence of Val at *C*-terminal, *m/z* 74.06 proved Thr at *N*-terminal, thus compound **34** was identified as Thr-Val. In this way, 39 dipeptides were identified in *C. sinensis*, among which 33 dipeptides contained Leu/Ile residues, accounting for the importance for differentiation between Leu and Ile residues.

*Characterization of other compounds*

Four other compounds were detected, including two *N*, *N*, *N*-trimethyl inner salt (laminine and betaine), uric acid, and vitamin B2 (also known as riboflavin). Vitamin B2 was confirmed by comparing with reference standard, and as it mainly existed in animal foods, it can be a potential marker to differentiate the fermented products [29].

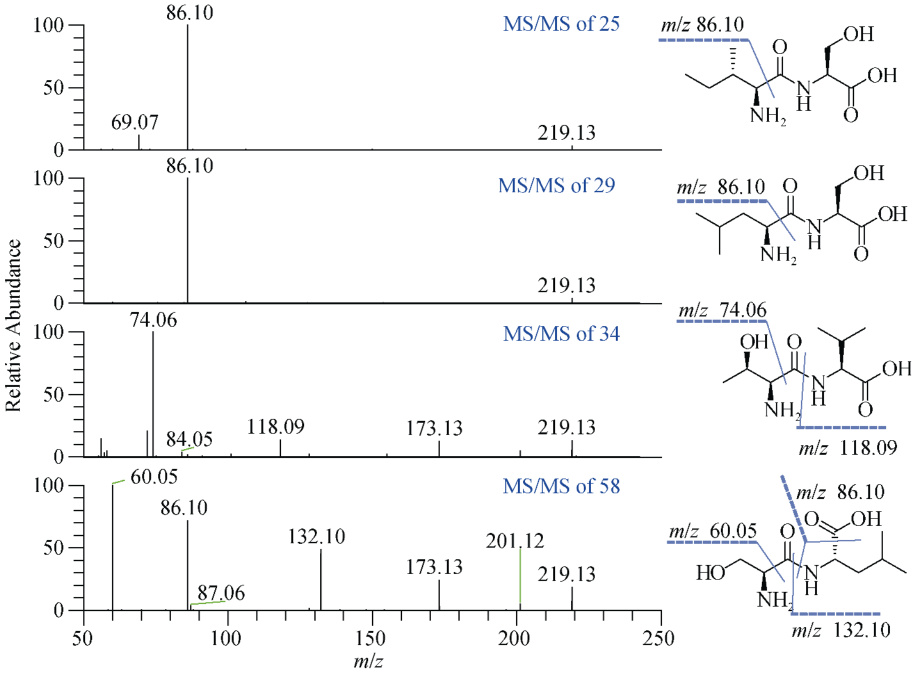


**Fig. 5 MS/MS spectra, proposed structures and fragmentation pathways for four dipeptide isomers Leu/Ile-Ser (25 and 29), Thr-Val (34) and Ser-Leu/Ile (58)**

## Conclusion

In the present work, a strategy was proposed to profile the components of aqueous extract of *C. sinensis*. A polar-modified $C_{18}$ column was adopted to characterize the components within 8 min, focusing on the polar compounds. The chemical basis was deeply explored in both targeted and untargeted manners. The targeted screening method offered 13 more trace nucleosides and nucleobases identified. The high resolution data of both precursor ions and product ions sharply augment the structural information. With all these efforts, 91 compounds were characterized and identified, including 30 nucleosides and nucleobases, 18 amino acids and derivatives, and 39 dipeptides. To the best of our knowledge, this was the first report unveiling the chemical basis of the aqueous extract of *C. sinensis*, which also laid a solid foundation for effect mechanism exploration and quality control of different sources of *C. sinensis* samples.

## References

[1] Yao CL, Pan HQ, Wang H, *et al.* Global profiling combined with predicted metabolites screening for discovery of natural

compounds: Characterization of ginsenosides in the leaves of *Panax notoginseng* as a case study [J]. *J Chromatogr A*, 2018, **1538**: 34-44.

[2] Pan HQ, Yao CL, Yang WZ, *et al*. An enhanced strategy integrating offline two-dimensional separation and step-wise precursor ion list-based raster-mass defect filter: Characterization of indole alkaloids in five botanical origins of *Uncariae Ramulus* Cum Unicis as an exemplary application [J]. *J Chromatogr A*, 2018, **1563**: 124-134.

[3] Zhang ML, Sun JH, Chen P. Development of a comprehensive flavonoid analysis computational tool for ultrahigh-performance liquid chromatography-diode array detection-high-resolution accurate mass-mass spectrometry data [J]. *Anal Chem*, 2017, **89**(14): 7388-7397.

[4] Wu HK, Mao YJ, Sun SS, *et al*. Leojaponic acids A and B, two new homologous terpenoids, isolated from *Leonurus japonicus* [J]. *Chin J Nat Med*, 2016, **14**(4): 303-307.

[5] Zhang ZX, Bo T, Bai Y, *et al*. Quadrupole time-of-flight mass spectrometry as a powerful tool for demystifying traditional Chinese medicine [J]. *Trac-Trends Anal Chem*, 2015, **72**: 169-180.

[6] Song YL, Zhang N, Shi SP, *et al*. Large-scale qualitative and quantitative characterization of components in Shenfu injection by integrating hydrophilic interaction chromatography, reversed phase liquid chromatography, and tandem mass spectrometry [J]. *J Chromatogr A*, 2015, **1407**: 106-118.

[7] Song QQ, Liu WJ, Chen XJ, *et al*. Serially coupled reversed phase-hydrophilic interaction liquid chromatography–tailored multiple reaction monitoring, a fit-for-purpose tool for large-scale targeted metabolomics of medicinal bile [J]. *Anal Chim Acta*, 2018, **1037**: 119-129.

[8] Qiao X, Wang Q, Wang S, *et al*. Compound to extract to formulation: a knowledge-transmitting approach for metabolites identification of Gegen-Qinlian decoction, a traditional Chinese medicine formula [J]. *Sci Rep*, 2016, **6**(1): 39534.

[9] Jin H, Liu Y, Guo Z, *et al*. Recent development in liquid chromatography stationary phases for separation of Traditional Chinese Medicine components [J]. *J Pharm Biomed Anal*, 2016, **130**: 336-346.

[10] Layne J. Characterization and comparison of the chromatographic performance of conventional, polar-embedded, and polar-endcapped reversed-phase liquid chromatography stationary phases [J]. *J Chromatogr A*, 2002, **957**(2): 149-164.

[11] Chen PX, Wang SA, Nie SP, *et al*. Properties of *Cordyceps Sinensis*: A review [J]. *J Funct Foods*, 2013, **5**(2): 550-569.

[12] Liu Y, Wang J, Wang W, *et al*. The chemical constituents and pharmacological actions of *Cordyceps sinensis* [J]. *Evid Based Complement Alternat Med*, 2015, **2015**: 575063.

[13] Zhao HQ, Wang X, Li HM, et al. Characterization of nucleosides and nucleobases in natural Cordyceps by HILIC-ESI/TOF/MS and HILIC-ESI/MS [J]. *Molecules*, 2013, **18**(8): 9755-9769.

[14] Zong SY, Han H, Wang B, *et al*. Fast simultaneous determination of 13 nucleosides and nucleobases in *Cordyceps sinensis* by UHPLC-ESI-MS/MS [J]. *Molecules*, 2015, **20**(12): 21816-21825.

[15] Yang ML, Kuo PC, Hwang TL, *et al*. Anti-inflammatory principles from *Cordyceps sinensis* [J]. *J Nat Prod*, 2011, **74**(9): 1996-2000.

[16] Yang FQ, Li DQ, Feng K, *et al*. Determination of nucleotides, nucleosides and their transformation products in Cordyceps by ion-pairing reversed-phase liquid chromatography-mass spectrometry [J]. *J Chromatogr A*, 2010, **1217**(34): 5501-5510.

[17] Yang FQ, Ge L, Yong JW, *et al*. Determination of nucleosides and nucleobases in different species of Cordyceps by capillary electrophoresis-mass spectrometry [J]. *J Pharm Biomed Anal*, 2009, **50**(3): 307-314.

[18] Huo XW, Liu CQ, Bai XL, *et al*. Aqueous extract of *Cordyceps sinensis* potentiates the antitumor effect of DDP and attenuates therapy-associated toxicity in non-small cell lung cancer *via* I kappa B alpha/NF kappa B and AKT/MMP2/ MMP9 pathways [J]. *Rsc Advances*, 2017, **7**(60): 37743-37754.

[19] Guo LN, Liu J, Huang Y, *et al*. Comparative study of DNA barcoding of cultivated and natural *Cordyceps sinensis* [J]. *Chin J Pharm Anal*, 2019, **39**(1): 147-155.

[20] Boccaletto P, Machnicka MA, Purta E, *et al*. MODOMICS: a database of RNA modification pathways, 2017 update [J]. *Nucleic Acids Research*, 2018, **46**(D1): D303-D307.

[21] Fu LL, Ding H, Han LF, *et al*. Simultaneously targeted and untargeted multicomponent characterization of Erzhi Pill by offline two-dimensional liquid chromatography/quadrupole-Orbitrap mass spectrometry [J]. *J Chromatogr A*, 2019, **1584**: 87-96.

[22] Zhang JY, Wang ZJ, Li Y, *et al*. A strategy for comprehensive identification of sequential constituents using ultra-high-performance liquid chromatography coupled with linear ion trap-Orbitrap mass spectrometer, application study on chlorogenic acids in Flos *Lonicerae Japonicae* [J]. *Talanta*, 2016, **147**(62): 16-27.

[23] Jora M, Burns AP, Ross RL, *et al*. Differentiating positional isomers of nucleoside modifications by higher-energy collisional dissociation mass spectrometry (HCD MS) [J]. *J Am Soc Mass Spectrom*, 2018, **29**(8): 1745-1756.

[24] Wang JQ, Kan LJ, Nie SP, *et al*. A comparison of chemical composition, bioactive components and antioxidant activity of natural and cultured *Cordyceps sinensis* [J]. *LWT-Food Sci Technol*, 2015, **63**(1): 2-7.

[25] Hu HK, Xiao L, Zheng BG, *et al*. Identification of chemical markers in *Cordyceps sinensis* by HPLC-MS/MS [J]. *Anal Bioanal Chem*, 2015, **407**(26): 8059-8066.

[26] Zhang P, Li SN, Li J, *et al*. Identification of *Ophiocordyceps sinensis* and its artificially cultured Ophiocordyceps mycelia by ultra-performance liquid chromatography/Orbitrap fusion mass spectrometry and chemometrics [J]. *Molecules*, 2018, **23**(5): 1013-1033.

[27] Andersen LT, Schlichtherle-Cerny H, Ardö Y. Hydrophilic di- and tripeptides are not a precondition for savoury flavour in mature Cheddar cheese [J]. *Dairy Sci Technol*, 2008, **88**(4-5): 467-475.

[28] Li XX, Fan PH, Zang MT, *et al*. Rapid determination of oligopeptides and amino acids in soybean protein hydrolysates using high-resolution mass spectrometry [J]. *Phytochem Anal*, 2015, **26**(1): 15-22.

[29] Zhang HY, Li YH, Mi JN, *et al*. GC-MS profiling of volatile components in different fermentation products of *Cordyceps Sinensis* Mycelia [J]. *Molecules*, 2017, **22**(10): 1800-1808.

**Cite this article as:** YAO Chang-Liang, QIAN Zheng-Ming, TIAN Wen-Shuai, XU Xiao-Qian[1], YAN Yu, SHEN Yao, LU Song-Mao, LI Wen-Jia, GUO De-An. Profiling and identification of aqueous extract of *Cordyceps sinensis* by ultra-high performance liquid chromatography tandem quadrupole-orbitrap mass spectrometry [J]. *Chin J Nat Med*, 2019, **17**(8): 631-640.