Available online at www.sciencedirect.com
ScienceDirect
Journal of Pharmaceutical and Biomedical Analysis 47 (2008) 516–525
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
www.elsevier.com/locate/jpba

Characterization of phenolic compounds in the Chinese herbal drug *Artemisia annua* by liquid chromatography coupled to electrospray ionization mass spectrometry

Jian Han, Min Ye, Xue Qiao, Man Xu, Bao-rong Wang, De-An Guo *

The State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Xueyuan Road 38, Beijing 100083, China

Received 26 October 2007; received in revised form 31 January 2008; accepted 6 February 2008
Available online 10 March 2008

Abstract

A simple and rapid method has been established for the screening of the main phenolic compounds in *Artemisia annua* by LC–DAD–ESI-MS^n. A total of 40 phenolic compounds were identified or tentatively characterized in the methanol extract of *A. annua*, including 8 *C*-glycosyl flavonoids, 5 *O*-glycosyl flavonoids, 3 flavonoid aglycones, 21 quinic acid derivatives, 2 benzoicacid glucosides and 1 coumarin. The *C*-glycosyl flavonoids were reported from *A. annua* for the first time and they were found to be a new type of main constituents, and might be responsible for its antioxidant and antiviral activity. Quinic acid derivatives were also found to be the major constituents of *A. annua*.
© 2008 Elsevier B.V. All rights reserved.

Keywords: *Artemisia annua*; LC–DAD–ESI-MS^n; Phenolic compounds

1. Introduction

Qinghao is a traditional Chinese medicine prepared from the aerial parts of *Artemisia annua* L. which belongs to *Compositae* family. It is used as an anti-parasitic and fever-relieving agent. Artemisinin was isolated from Qinghao as a major bioactive constituent and has been used to treat malaria. Besides artemisinin, the chemical constituents of Qinghao include volatile oil, terpenes, flavonoids and coumarins [1]. Previous reports mainly focused on terpenes since they are the active fractions of the anti-parastic, while little attention was paid to the phenolic compounds. Recently, phenolic compounds have gained more and more attention because of their antioxidant and anticancer activities. It has been reported that the flowers and leaves of *Compositae* plants also show anti-bacterial, anti-fungal, anti-viral and anti-inflammatory activities [2]. As far as we know, no study has provided comprehensive information of the constituents in *A. annua* until Lai et al. [3] reported the identification of major flavonoids and caffeoylquinic acids by LC–APCI–MS/MS. In our recent studies, we found that most phenolic compounds could be more efficiently ionized in negative ion mode ESI-MS than positive mode [4–6]. And the chromatographic fingerprints have become a useful tool in the quality control of complex herbal plants. Thus we tried to develop an LC–DAD–ESI-MS^n method for the comprehensive and accurate identification of phenolic compounds in *A. annua* to provide some evidence for the study of other functional aspects of this important traditional herbal medicine.

Liquid chromatography coupled with mass spectrometry (LC/MS) is a powerful tool for the rapid identification of chemical constituents in plant extracts. The high sensitivity of MS as an LC detector facilitates to discover new minor constituents, which are difficult to obtain by conventional means. The tandem mass spectrometric fragmentation behavior of flavonoids has been investigated extensively, which allows the characterization of unknown compounds even without the reference standards [7–11]. In the present study, HPLC–DAD–ESI-MS^n was used to analyse phenolic compounds globally in the methanol extracts of *A. annua* for the first time.

* Corresponding author. Tel.: +86 10 8280 1516; fax: +86 10 8280 2700.
E-mail address: gda@bjmu.edu.cn (D.-A. Guo).

0731-7085/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2008.02.013

2. Materials and methods

2.1. Chemicals and materials

Standard of chlorogenic acid was purchased from the National Institute for the Control of Pharmaceutical and Biological Products (NICPBP) (Beijing, China).

HPLC grade acetonitrile (CH_3CN) (Fisher, Fair Lawn, NJ, USA) and ultra-pure water were used for all analyses. The methanol used for extraction of sample was AR grade, purchased from Beijing Chemical Corporation (Beijing, China).

2.2. Sample preparation

The samples of *A. annua* were purchased from different drug stores in Beijing and identified by the authors. The materials were pulverized and an aliquot of 1 g was extracted with 100 mL of methanol by ultrasonication for 40 min and then filtered through a 0.45-μm membrane prior to use. Aliquots of 10 μL was injected into the LC–MS analysis.

2.3. HPLC conditions

An Agilent series 1100 HPLC instrument (Agilent, Waldbronn, Germany) equipped with a quaternary pump, a diode-array detector, an autosampler and a column compartment was used for analyses. The sample was separated on a Zorbax XDB-C_{18} column (5 μm, φ4.6 mm × 250 mm, Agilent). The mobile phase was (A) CH_3CN and (B) water–acetic acid (100:0.3, v/v). A gradient program was used as follows: 12–20% A at 0–10 min, hold for 5 min, linearly gradient to 46% A in 30 min, then linearly gradient to 80% A at 60 min. The mobile phase flow rate was 0.8 mL/min; the chromatogram was recorded at 280 nm and spectral data for all peaks were accumulated in the range of 190–400 nm. Column temperature was controlled at 25 °C.

2.4. Mass spectrometric conditions

A Finnigan LCQ advantage ion trap mass spectrometer (ThermoFinnigan, San Jose, CA) was connected to the Agilent 1100 HPLC instrument via ESI interface for HPLC/MS^n analysis. Ultrahigh pure helium (He) was used as the collision gas and high-purity nitrogen (N_2) as the nebulizing gas. The optimized parameters in the negative ion mode were as follows: ion spray voltage, 4.5 kV; sheath gas, 55 arbitrary units; auxiliary gas, 5 U; capillary temperature, 350 °C; capillary voltage, −4 V; tube lens offset voltage, −40 V. For full scan MS analysis, the spectra were recorded in the range of *m/z* 100–1000. The isolation width of precursor ions was 3.0 U. MS^n data were acquired in the automatic data-dependent mode.

3. Results and discussion

Samples from different drug stores were analysed, respectively, and the LC–ESI TIC profiles were shown in Fig. 1S. It showed that the repeatability was acceptable. Unexpectedly, the phenolic constituents of *A. annua* were found to be much more complex than hitherto known, and some *C*-glycoside flavonoids were identified for the first time. A typical HPLC/UV and MS chromatograms of the methanol extracts are shown in Fig. 1. Chemical structures of the identified compounds are shown in Fig. 2 and their MS^n fragmentations were shown in Table 1.

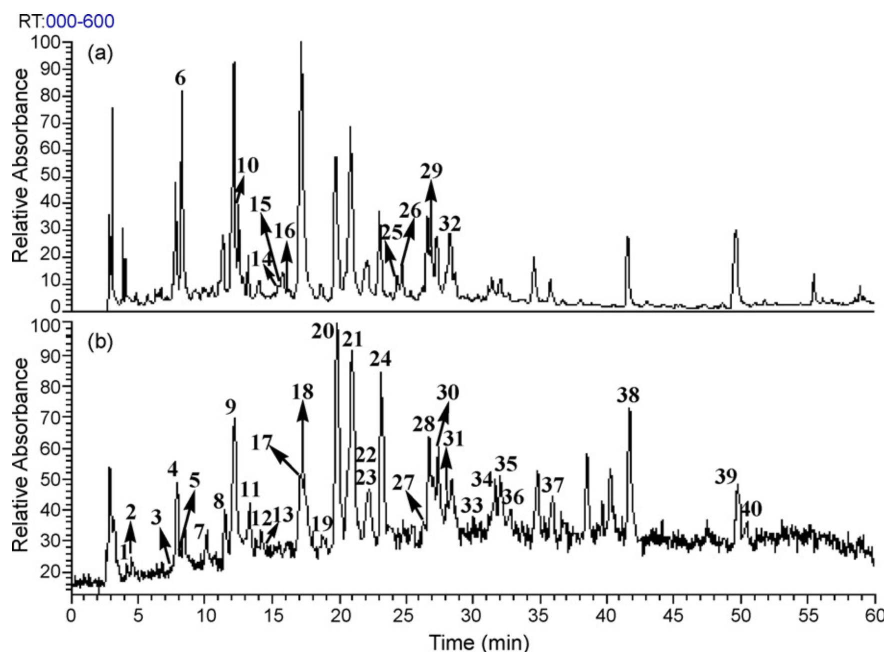
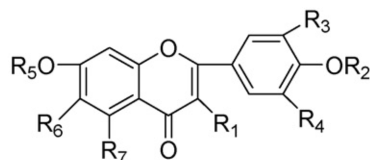
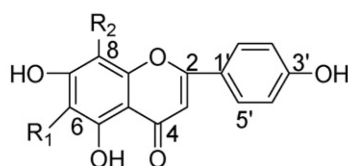
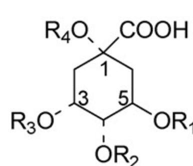


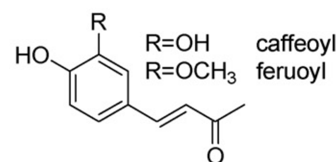
Fig. 1. HPLC–DAD–ESI-MS^n analysis of the methanolic extract of *Artemisia annua* L. (a) HPLC–UV chromatogram monitored at 280 nm and (b) LC–negative ion ESI-MS total ion current (TIC) profile.



Compound	R_1	R_2	R_3	R_4	R_5	R_6	R_7
3	O-di-glu	CH_3	OH	OH	H	H	OH
14	O-rha-glu	H	OH	H	H	H	OH
17	O-glu	H	OH	H	H	H	OH
18	O-glu	CH_3	OH	OH	H	H	OH
22	O-glu	H	OCH_3	H	H	H	OH
38	OH	CH_3	H	H	CH_3	OCH_3	OH
39	OH	CH_3	H	H	CH_3	OCH_3	H
40	OH	CH_3	H	OCH_3	CH_3	OCH_3	OH

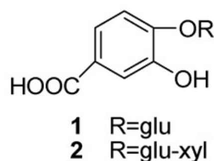






Compound	R_1	R_2
apigenin	H	H
7	glu	glu
8	ara	glu
9	glu	ara
10	pent	glu
11	glu	pent
12	rha	glu
15	H	glu
16	glu	rha

Compound	R_1	R_2	R_3	R_4
4	caffeoyl	H	H	H
5	H	H	caffeoyl	H
13	feruoyl	H	H	H
19	H	caffeoyl	caffeoyl	H
20	caffeoyl	H	caffeoyl	H
24	caffeoyl	caffeoyl	H	H
25	H	caffeoyl	feruoyl	H
26	caffeoyl	feruoyl	H	H
27	caffeoyl	H	H	feruoyl
28	feruoyl	H	H	caffeoyl
29	caffeoyl	H	feruoyl	H
30	feruoyl	H	caffeoyl	H
31	feruoyl	caffeoyl	H	H
32	H	feruoyl	caffeoyl	H
33	H	caffeoyl	caffeoyl	H
34	caffeoyl	H	caffeoyl	H
35	caffeoyl	caffeoyl	caffeoyl	H
36	feruoyl	feruoyl	H	H



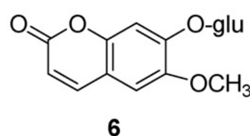


Fig. 2. Chemical structures of phenolic compounds identified in *A. annua*. glu: glucosyl; ara: arabinosyl; rha: rhamnosyl; xyl: xylosyl.

3.1. Identification of flavone glycosides

Thirteen flavone glycosides were plausibly identified according to their online UV spectra and MS^n fragmentations. Five of them were *O*-glycosyl flavonoids previously reported from *A. annua*. The other eight compounds were *C*-glycosyl flavonoids reported from *A. annua* for the first time and were tentatively identified on the basis of their tandem mass spectra.

3.1.1. O-Glycosides

Five *O*-glycosides were identified and all of them were previously reported from *A. annua* [3]. Compounds **3** and **18** exhibited [M−H]$^-$ ion at *m/z* 655 and 493, respectively, and both of them gave the [aglycone−H]$^-$ (abbreviated as [A−H]$^-$) ion at *m/z* 331. Then the collision-induced dissociation (CID) of the *m/z* 331 yielded a [A−H−15]$^-$ fragment at *m/z* 316, consistent with mearnsetin. The MS^n spectra of the compound **3** gave ions at *m/z* 493 and 331, originating from successive losses of 162 Da, suggesting the presence of two glucose residues, while the MS^n information of compound **18** indicated the presence of one glucose residue. Thus these two compounds were identified as mearnsetin-*O*-diglucoside and mearnsetin-*O*-glucoside, respectively. And the fragmentation pathway of compound 3 was proposed in Scheme 1.

CID of compound **14** yielded a [A−H]$^-$ ion at *m/z* 301, resulting from the concomitant losses of a rhamnose (146 Da) and a hexose (162 Da). Thus it was preliminarily characterized as quercetin-rhamnosylglucoside.

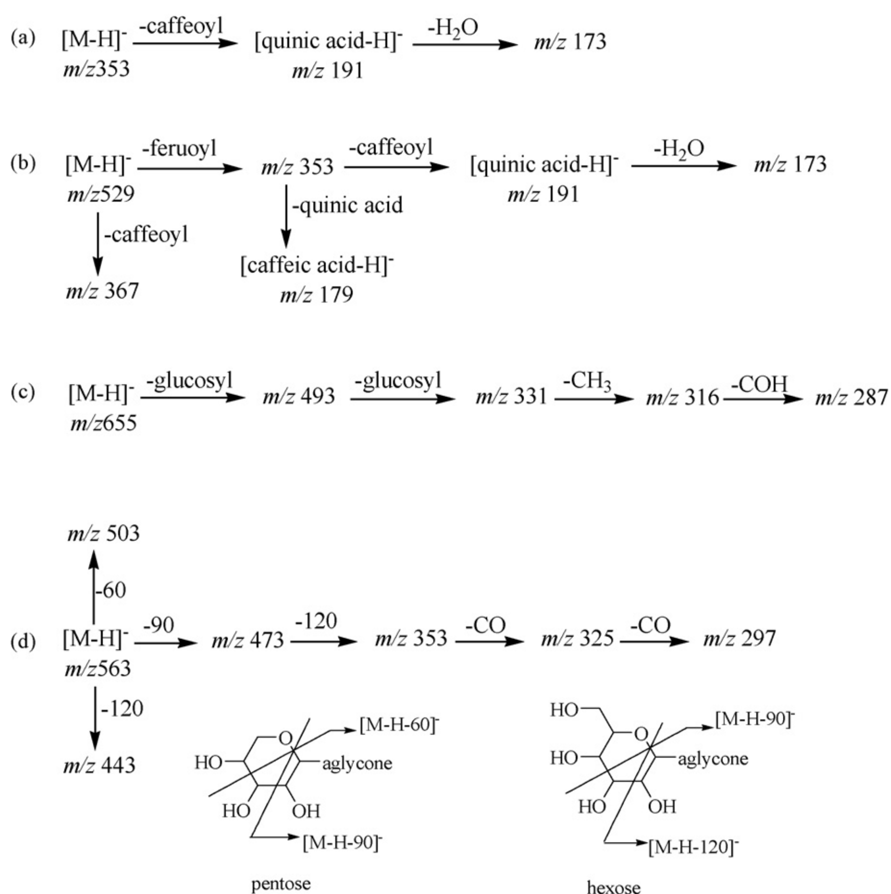
Compound **17** also yielded a [A−H]$^-$ ion at *m/z* 301, resulting from the loss of a glucose (162 Da). The CID spectrum of the *m/z* 301 ion was consistent with quercetin. Thus it was preliminarily characterized as quercetin-*O*-glucoside.

Compound **22** gave [M−H]$^-$ ion at *m/z* 477 and [A−H]$^-$ ion at *m/z* 315, suggesting the existence of a glucose residue. And the MS^3 spectrum of *m/z* 315 was very similar to that of isorhamnetin [4]. This compound was thus tentatively identified as isorhamnetin-*O*-glucoside.

3.1.2. C-Glycosides

The UV spectra of the eight compounds (peaks **7**, **8**, **9**, **10**, **11**, **12**, **15** and **16** in Fig. 1) in the HPLC chromatogram showed maximum absorption bands at about 224, 272 and 314 (sh) nm, consistent with the characteristic absorption of flavones, however giving different fragmentation pattern from those of the *O*-glucosides. The MS/MS spectra of them exhibited ions of *m/z* [M−H−60]$^-$, [M−H−90]$^-$ and [M−H−120]$^-$ at different relative abundances, which were demonstrated as characteristic ions of *C*-glycosides [12–14]. It had been also reported that C-6 and C-8 flavonoid glycosides could be distinguished by the relative abundance of ions [M−H−60]$^-$ and [M−H−120]$^-$ [15,16]. Similar difference was observed for MS/MS spectra of these compounds (Table 2). And the fragmentation pathway of compound **8** was proposed in Scheme 1.

Compound **7** gave an [M−H]$^-$ ion at *m/z* 593, and its MS/MS spectrum produced ions at *m/z* 503 ([M−H−90]$^-$), 473 ([M−H−120]$^-$) and 383 ([M−H−210]$^-$), while the ion of [M−H−60]$^-$ was absent, all of which coincided with those reported previously [16], and suggested the presence of apigenin



Scheme 1. Major fragmentations for the molecular adduct ions of typical compounds identified in *Artemisia annua* L. (a) Compound **4** (chlorogenic acid); (b) compound **30** (3-caffeoyl-5-feruoyl-quinic acid); (c) compound **3** (mearnsetin-di-glucoside); (d) compound **8** (apigenin-6-*C*-arabinoside-8-*C*-glucoside).

Table 1
Identification of phenolic compounds from *Artemisia annua*

Peak no.	Retention time, t_R (min)	$[M-H]^-$	UV λ_{max} (nm)	LC–ESI-MS^n m/z (% base peak)	Identification
1	4.33	315	208, 216, 252	MS^2[315]: 153(100), 152(31), 109(10), 108(13)	3,5-Dihydroxybenzoicacid-*O*-glu
2	4.76	447	214, 254	MS^2[447]: 315(100), 153(18), 152(12), 151(3); MS^3[315]: 153(100)	3,5-Dihydroxybenzoicacid-*O*-glu-xyl
3	7.55	655	–	MS^2[655]: 535(3), 493(100), 492(31), 491(2), 331(28), 329(10), 316(1), 301(1), 217(1); MS^3[493]: 478(3), 477(2), 373(7), 331(100), 329(47), 315(39), 314(2), 301(2); MS^4[331]: 316(100), 287(4), 270(3); MS^5[316]: 287(100), 271(33), 270(27)	Mearnsetin-di-glu
4	7.90	353	218, 326	MS^2[353]: 191(100), 179(4), 161(1), 135(1); MS^3[191]: 173(100), 111(63), 85(80)	Chlorogenic acid
5	8.25	353	226, 338	MS^2[353]: 191(100), 179(24), 173(33)	3-Caffeoylquinic acid
6	8.38	191	226, 290, 338	MS^2[191]: 176(100)	Scopoletin
7	9.87	593	220, 256, 328	MS^2[593]: 503(31), 485(4), 473(100), 455(2), 443(1), 438(1), 425(3), 419(1), 413(3), 397(1), 395(3), 383(42), 367(1), 353(68), 341(2), 329(2), 325(4), 321(2), 297(1), 197(1); MS^3[473]: 395(3), 383(19), 353(100), 325(4), 311(3), 297(4); MS^4[353]: 325(48), 297(100), 282(36)	Apigenin-6,8-di-*C*-glu (vicenin-2)
8	11.47	563	228, 274, 328	MS^2[563]: 545(18), 503(67), 485(7), 473(100), 443(59), 425(21), 413(14), 383(76), 353(94), 325(3); MS^3[473]: 413(3), 383(15), 353(100), 325(3); MS^4[353]: 325(100), 323(4), 297(42); MS^5[325]: 297(100)	Apigenin-6-*C*-ara-8-*C*-glu (isoschaftoside)
9	12.09	563	220, 274, 328	MS^2[563]: 545(7), 503(13), 485(5), 473(77), 455(6), 443(100), 425(8), 413(8), 395(6), 383(55), 365(2), 353(69), 325(2), 297(2); MS^3[443]: 407(2), 397(2), 395(3), 383(27), 353(100), 325(4); MS^4[353]: 325(100), 323(10), 297(48), 189(18), 117(13); MS^5[325]: 297(100)	Apigenin-6-*C*-glu-8-*C*-ara (schaftoside)
10	12.54	563	224, 274, 328	MS^2[563]: 545(17), 503(50), 473(96), 443(86), 425(20), 413(12), 383(97), 353(100); MS^3[353]: 325(100), 308(7), 297(60); MS^4[325]: 297(100)	Apigenin-6-*C*-pent-8-*C*-glu
11	13.33	563	236, 272, 328	MS^2[563]: 545(6), 527(2), 515(3), 503(6), 488(2), 485(7), 473(52), 455(8), 443(100), 437(2), 425(2), 413(5), 395(2), 383(30), 365(2), 353(48); MS^3[443]: 383(15), 353(100); MS^4[353]: 325(100), 307(7), 297(44), 283(3), 233(5), 175(3); MS^5[325]: 297(100)	Apigenin-6-*C*-glu-8-*C*-pent
12	14.04	577	236, 294, 318	MS^2[577]: 559(10), 533(5), 503(37), 487(8), 473(77), 457(92), 439(14), 437(3), 413(15), 395(12), 383(70), 365(5), 353(100); MS^3[353]: 325(100), 297(31); MS^4[325]: 297(100)	Apigenin-6-*C*-rha-8-*C*-glu
13	14.15	367	234, 292, 318	MS^2[367]: 193(4), 191(100)	5-Feruoylquinic acid
14	15.24	609	236, 274, 326	MS^2[609]: 300(100), 301(70), 273(3), 271(10), 255(11), 254(2); MS^3[300]: 271(99), 255(100), 239(13), 193(13), 179(20), 151(50); MS^4[271]: 229(100)	Quercetin-rha-glu
15	15.60	431	–	MS^2[431]: 341(15), 311(100), 283(3); MS^3[311]: 283(100), 268(1), 253(2), 191(2), 147(1); MS^4[283]: 211(32), 209(21), 163(100), 145(52)	Apigenin-8-*C*-glu (vitexin)
16	16.10	577	240, 324	MS^2[577]: 559(7), 487(49), 473(5), 457(100), 383(22), 353(47); MS^3[457]: 383(10), 365(3), 353(100), 325(7), 297(2); MS^4[353]: 325(100), 323(9), 297(11), 282(30), 266(23); MS^5[325]: 297(100)	Apigenin-6-*C*-glu-8-*C*-rha
17	16.90	463	258, 348	MS^2[463]: 445(2), 427(1), 343(2), 301(100), 300(98), 283(3), 271(3), 255(3), 254(1), 231(1), 214(1), 179(3), 151(2); MS^3[301]: 273(6), 271(100), 255(82), 229(3), 211(9), 179(82), 169(15), 164(3), 151(16), 134(9); MS^4[271]: 227(100)	Quercetin-glu
18	17.22	493	228, 286, 336	MS^2[493]: 478(2), 331(100), 330(51), 316(17), 315(8) MS^3[331]: 316(100), 315(37), 287(2), 272(3), 163(1); MS^4[316]: 298(4), 287(100), 273(6), 271(22), 270(22), 259(6), 255(8), 245(3), 243(18), 231(2), 229(3), 215(3), 186(1), 179(6), 166(20), 165(2), 138(2)	Mearnsetin-glu
19	18.63	515	244, 324	MS^2[515]: 353(100), 341(1), 337(1), 335(12), 317(3), 299(2), 255(6), 227(2), 211(2), 203(2), 191(3), 179(8), 175(1), 173(14); MS^3[353]: 191(47), 179(100), 173(58), 135(12)	3,4-Dicaffeoylquinic acid

Table 1 (*Continued*)

Peak no.	Retention time, t_R (min)	$[M-H]^-$	UV λ_{max} (nm)	LC–ESI-MS^n *m/z* (% base peak)	Identification
20	19.81	515	242, 328	MS^2[515]: 353(100), 352(2), 335(4), 191(18), 179(3); MS^3[353]: 191(100), 179(2), 145(2), 135(3)	3,5-Dicaffeoylquinic acid
21	20.92	515	240, 326	MS^2[515]: 353(100), 352(1), 191(6), 179(4); MS^3[353]: 191(100), 179(53), 173(2), 137(2), 135(5)	Quinic acid derivate
22	21.84	477	240, 328	MS^2[477]: 462(7), 315(100), 314(26), 313(3), 300(33), 299(16), 272(3), 271(3), 255(1); MS^3[315]: 300(100), 299(23), 272(6), 271(3), 254(2), 153(3); MS^4[300]: 271(100), 255(38), 254(21), 226(6); MS^5[271]: 255(100)	Isorhamnetin-glu
23	22.17	515	226, 324	MS^2[515]: 353(100), 335(1), 191(5), 179(5), 173(1); MS^3[353]: 191(100), 179(66), 173(21), 135(15)	Quinic acid derivate
24	23.14	515	240, 326	MS^2[515]: 353(100), 335(2), 317(2), 299(8), 291(1), 255(6), 227(1), 203(7), 191(2), 179(6), 173(3); MS^3[353]: 191(69), 179(84), 173(100), 135(34), 109(3)	4,5-Dicaffeoylquinic acid
25	24.71	529	236, 284, 318 (sh)	MS^2[529]: 485(6), 447(6), 412(6), 387(8), 367(92), 353(100), 349(40), 335(73), 324(3), 299(4), 287(3), 275(5), 255(22), 240(6), 193(69), 181(4), 179(11), 175(4), 173(48); MS^3[353]: 178(20), 173(100)	3-Feruoyl-4-caffeoyl-quinic acid
26	25.51	529	244, 288, 324	MS^2[529]: 367(100), 366(2), 353(1), 349(3), 335(9), 313(1), 299(1), 277(1), 255(1), 193(3), 175(1), 173(22), 161(1); MS^3[367]: 193(52), 173(100), 155(7), 111(7), 108(10); MS^4[173]: 111(100)	4-Feruoyl-5-caffeoyl-quinic acid
27	26.36	529	244, 328	MS^2[529]: 367(36), 353(100), 344(3), 257(2), 255(4), 191(23); MS^3[353]: 191(100), 179(17), 109(34)	1-Feruoyl-5-caffeoyl-quinic acid
28	26.74	529	244, 328	MS^2[529]: 367(100), 353(30), 291(2), 191(6), 179(4); MS^3[367]: 191(100), 173(4)	1-Caffeoyl-5-feruoyl-quinic acid
29	27.03	529	242, 326	MS^2[529]: 367(100), 366(3), 353(9), 193(10), 191(4), 173(5), 161(2); MS^3[367]: 193(100), 191(36), 134(2), 117(4)	3-Feruoyl-5-caffeoyl-quinic acid
30	27.40	529	242, 328	MS^2[529]: 367(63), 353(100), 352(1), 193(2), 191(12), 179(5), 173(2); MS^3[353]: 191(90), 179(100), 173(19), 135(21)	3-Caffeoyl-5-feruoyl-quinic acid
31	28.01	529	242, 302	MS^2[529]: 367(100), 366(4), 353(42), 349(2), 193(3), 191(10), 173(4), 161(2); MS^3[367]: 193(71), 191(100), 173(83)	4-Caffeoyl-5-feruoyl-quinic acid
32	28.39	529	244, 286, 322	MS^2[529]: 367(43), 353(100), 255(5), 203(4), 191(6), 179(7), 173(5); MS^3[353]: 191(38), 179(100), 173(87), 161(5), 155(10), 135(35); MS^4[179]: 135(100)	3-Caffeoyl-4-feruoyl-quinic acid
33	29.70	543	–	MS^2[543]: 367(10), 349(100), 348(1), 269(3), 193(2), 188(1), 175(5), 173(10); MS^3[349]: 331(8), 287(33), 243(11), 228(14), 193(20), 175(100), 173(31), 155(10), 134(15)	3,4-Diferuoylquinic acid
34	31.53	543	246, 328	MS^2[543]: 367(100), 349(15), 348(5), 261(7), 191(15); MS^3[367]: 191(100)	3,5-Diferuoylquinic acid
35	32.04	677	242, 326	MS^2[677]: 515(100), 514(1), 497(2), 469(1), 353(8), 299(1); MS^3[515]: 353(100), 335(3), 299(2), 255(5), 179(12), 173(7); MS^4[353]: 179(100), 173(82)	3,4,5-Tricaffeoylquinic acid
36	32.87	543	246, 328	MS^2[543]: 367(100), 366(2), 349(28), 193(15), 173(7); MS^3[367]: 193(100), 173(58), 154(23); MS^4[193]: 134(100)	4,5-Diferuoylquinic acid
37	35.69	691	246, 326	MS^2[691]: 529(100), 515(96), 511(4), 497(2), 483(1), 367(2), 353(25), 352(2), 349(1), 335(2), 317(2), 289(2); MS^3[529]: 367(75), 353(100), 299(52), 161(69)	Dicaffeoylferuoylquinic acid
38	41.68	359	258, 352	MS^2[359]: 344(100), 329(4); MS^3[344]: 329(100), 314(4), 301(6), 286(1); MS^4[329]: 314(100), 311(1), 301(21), 296(2), 286(63), 285(5), 283(2), 270(1), 258(7), 242(1), 230(2), 202(1), 175(2); MS^5[314]: 286(100), 269(1), 268(1), 258(20), 257(1), 230(12)	3,5-Dihydroxy-6,7,4′-trimethoxyflavone
39	48.82	343	248, 336	MS^2[343]: 328(100), 327(1), 313(3), 261(1); MS^3[328]: 313(100), 298(2), 285(1); MS^4[313]: 298(100), 285(5), 270(34); MS^5[298]: 270(100), 242(6), 187(6)	3-Hydroxy-6,7, 4′-trimethoxyflavone

Table 1 (*Continued*)

Peak no.	Retention time, t_R (min)	[M−H]$^-$	UV λ_{max} (nm)	LC–ESI-MS^n *m/z* (% base peak)	Identification
40	49.74	373	256, 348	MS^2[373]: 358(100), 343(9); MS^3[358]: 343(100), 342(1), 328(2), 315(2), 300(1), 299(1); MS^4[343]: 328(100), 315(14), 313(11), 300(75), 299(14), 297(1), 287(5), 285(13), 284(10), 272(15), 271(3), 269(5), 257(3), 241(1), 240(1), 213(1); MS^5[328]: 313(6), 311(2), 300(100), 299(4), 285(29), 272(8), 269(1), 256(2), 227(2), 213(2)	3,5-Dihydroxy-6,7,3′,4′-tetramethoxyflavone

(270) + glucose + glucose. Thus it was tentatively identified as 6,8-di-*C*-glucosyl apigenin, also named vicenin-2.

Four compounds (**8–11**) exhibited [M−H]$^-$ ions at *m/z* 563. Similar to compound **7**, CID of them gave product ions of [M−H−90]$^-$, [M−H−120]$^-$ and [M−H−210]$^-$. However different from compound **7**, all of these four compounds yielded the ions of [M−H−60]$^-$ in their MS/MS spectra, suggesting the presence of a pentose substitution. It was proved that the [M−H−60]$^-$ ion had much more relative abundance in the spectrum of C-6 glycoside than that of C-8 glycoside. In view of the retention behavior on the RP-HPLC column, the 6-*C*-pentosyl-8-*C*-glucosyl eluted before the isomeric 6-*C*-glucosyl-8-*C*-pentosyl [16]. According to all of the above information, these four compounds were tentatively characterized as 6-*C*-arabinosyl-8-*C*-glucosyl apigenin, 6-*C*-glucosyl-8-*C*-arabinosyl apigenin, 6-*C*-pentosyl-8-*C*-glucosyl apigenin and 6-*C*-glucosyl-8-*C*-pentosyl apigenin, respectively. These fragments are in accordance with previous data [16].

Compounds **12** and **16** both gave the [M−H]$^-$ ions at *m/z* 577. In their MS/MS spectra, ions of *m/z* 559 ([M−H−H_2O]$^-$), 503 ([M−H−74]$^-$) (observed in compound **12**, not in compound **16**), 473 ([M−H−104]$^-$), 487 ([M−H−90]$^-$), 457 ([M−H−120]$^-$) were observed, suggesting the presence of a methylpentose moiety and a hexose substitute. Thus they were plausibly characterized as 6-*C*-rhamnosyl-8-*C*-glucosyl apigenin and 6-*C*-glucosyl-8-*C*-rhamnosyl apigenin, respectively.

Compound **15** gave the [M−H]$^-$ ion at *m/z* 431, and its MS/MS spectrum yielded the ions at *m/z* 341 ([M−H−90]$^-$) and 311 ([M−H−120]$^-$), thus it was identified as 8-*C*-glucosyl apigenin, also named vitexin.

Table 2
The relative abundances of ions obtained from the respective pseudomolecular ions [M−H]$^-$ of identified *C*-glycosyl flavones in MS^2

Compounds	Relative abundances (%) of ions from [M−H]$^-$					
	−18	−60	−74[a]	−90	−104[a]	−120
7	–	–	–	31	–	100
8	18	67	–	100	–	59
9	7	13	–	77	–	100
10	17	50	–	96	–	86
11	6	6	–	52	–	100
12	10	–	37	8	77	92
15	–	–	–	15	–	100
16	7	–	–	49	5	100

[a] Methylpentose substitute.

C-glycosylflavonoids are common constituents in the flowering plants [17]. Currently, more than 100 *C*-glycosylflavonoids have been isolated from plants. Well-known mono-*C*-glycosylflavonoids include vitexin and isovitexin, which are the apigenin-8-*C*-glycoside and apigenin-6-*C*-glycoside, respectively. Most of the aglycone of the *C*-glycosylflavonoids is apigenin or luteolin. The typical di-*C*-glycosylflavonoid is schaftoside (apigenin-6-*C*-glu-8-*C*-ara), which is also relatively common. In some families the *C*-glycosylflavonoids are the major class of flavonoids present, e.g. in *Passifloraceae* and *Poaceae*, whereas in others they occur only as minor constituents in specific taxa, e.g. *Scutellaria baicalensis* in *Lamiaceae*. Previously no *C*-glycosylflavonoid has been identified in *A. annua*, though some *C*-glycosylflavonoids did have been isolated from other plants in *Compositae* like *Achillea species* [18–20], *Centaurea* species [21–24]. In present study, a total of 8 *C*-glycosylflavonoids were tentatively identified and they seemed to be one of the predominant constituents.

3.2. Identification of flavone aglycone

Three aglycones were identified or tentatively characterized, all of which were known constituents previously identified in *A. annua*.

Compound **38** yielded [M−H]$^-$ at *m/z* 359. Its MS^n fragmentations indicated the existence of three –CH_3 groups. And the ion at *m/z* 258 resulted from the loss of two –CO (−56 Da) groups from the ion at *m/z* 314, which was only observed in flavones, reported by Faber et al. [7]. By searching the known flavones in *A. annua*, it was tentatively characterized as 3,5-dihydroxy-6,7,4′-trimethoxyflavone. Compound **39** exhibited [M−H]$^-$ at *m/z* 343, 16 Da less than compound **38**. And its MS^n fragmentations also indicated the existence of three –CH_3 groups. Thus it was tentatively identified as 3-hydroxy-6,7,4′-trimethoxyflavone. Compound **40** gave [M−H]$^-$ at *m/z* 373, and its CID gave peaks at *m/z* 358, 343, 328 and 313, indicating the presence of four methoxyl groups. Thus it was identified as 3,5-dihydroxy-6,7,3′,4′-tetramethoxyflavone, which was reported from *A. annua* previously [25].

Flavonoids are a class of natural products widely distributed in the plants. It has been reported that the major bioactive constituents of many medicinal herbs were flavonoids for their antioxidant, anti-inflammatory, and anti-carcinogenic activities. Among them *C*-glycosylflavonoids also have various activities, such as hepatoprotective effect [26,27], antioxidant [28], antithyroid effects [29] and antiviral activity [30]. In present

study, a total of 16 flavonids, including 8 *C*-glycosylflavonoids, have been identified and they were one type of the major constituents of *A. annua*. These findings could facilitate the discovery of essential components that take the therapeutic effects. And it would provide some new basis for successful quality control of this medicinal herb.

3.3. Identification of quinic acid derivatives

A total of 21 quinic acid derivatives were identified in present study. Compound **4** was confirmed as chlorogenic acid (5-caffeoylquinic acid) by comparing with the standard. Its fragmentation pathway is proposed in Scheme 1. It was demonstrated that the linkage position of acyl groups on monoacyl quinic acid could be determined according to the MS/MS fragmentation behavior of the $[M-H]^-$ ions [31,32]. Briefly, when the acyl group was linked to 3-OH or 5-OH, the $[\text{quinic acid}-H]^-$ ion at *m*/*z* 191 was the base peak, and the $[\text{caffeic acid}-H]^-$ ion at *m*/*z* 179 was more significant for 3-*O*-caffeoylquinic acids. While the $[\text{quinic acid}-H_2O-H]^-$ ion at *m*/*z* 173 was the prominent peak when the acyl group connected to 4-OH.

Compound **5** displayed $[M–H]^-$ ion at *m*/*z* 353, the same as chlorogenic acid. But it exhibited different MS^2 spectra. In the MS^2 spectrum, the *m*/*z* 191 was the base peak. By referring to literature data [4], it was identified as 3-caffeoylquinic acid.

Compound **13** gave $[M-H]^-$ ion at *m*/*z* 367, and was identified as 5-feruoylquinic acid. Its MS/MS spectrum contained a base peak at *m*/*z* 191 corresponding to the loss of 176 Da (feruoyl group), consistent with the literature data [4].

Compounds **19**, **20** and **24** all gave the $[M-H]^-$ ions at *m*/*z* 515 and the $[M-H-162]^-$ ions at *m*/*z* 353. However, their MS^3 spectra were significantly different. Compound **20** produced base peak ion at *m*/*z* 191, as reviewed above, it was identified as 3,5-dicaffeoylquinic acid, which was consistent with the previous report [4]. Other two compounds produced base peak at *m*/*z* 179 or 173, hence they were identified as 4-substituted quinic acids. According to the literature [33], 3,5-dicaffeoylquinic acid was more easily eluted from the reverse-phase column when compared with 4,5-dicaffeoylquinic acid. Thus compound **24** was identified as 4,5-dicaffeoylquinic acid, and the compound **19** was identified as 3,4-dicaffeoylquinic acid. Compounds **21** and **23** also displayed $[M-H]^-$ ions at *m*/*z* 515 and possessed similar fragmentations with the above three diacyl quinic acids, thus they were tentatively identified as quinic acid derivatives.

At least eight compounds (compounds **25**–**32**) with $[M-H]^-$ ions at *m*/*z* 529 were detected in our study, and the selected ion monitoring (SIM) at it was shown in Fig. 3. All of these eight compounds showed MS^2 base peaks at either *m*/*z* 353 ($[\text{caffeoylquinic acid}-H]^-$) or *m*/*z* 367 ($[\text{feruoylquinic acid}-H]^-$). And MS^3 spectra of six of them (compounds **25**, **26**, **29**–**32**) significantly yielded 179 ($[\text{caffeoyl}-H]^-$) or 193 ($[\text{feruoyl}-H]^-$). Based on these MS^n fragmentations, they were characterized as caffeoylferuoylquinic acids (CFQA) isomers [34]. Previously six CFQAs have been identified from coffee and the hierarchical key established by LC–MS^n permitted specific structures to be assigned to each of the six isomers. In present paper, six of the compounds with *m*/*z* 529 were tentatively identified by referring to the hierarchical key. The fragmentation pathway of compound **30** is proposed in Scheme 1. Other two compounds (compounds **27** and **28**) showed a little different MS^n fragmentation and relative abundance from the other six compounds. Both of their MS^3 spectrum displayed *m*/*z* 191 as the base peak, thus they were 3- or 5-substitutes. If they were 3-substitutes, the peak abundance at *m*/*z* 179 should generally be above 50% [34], they

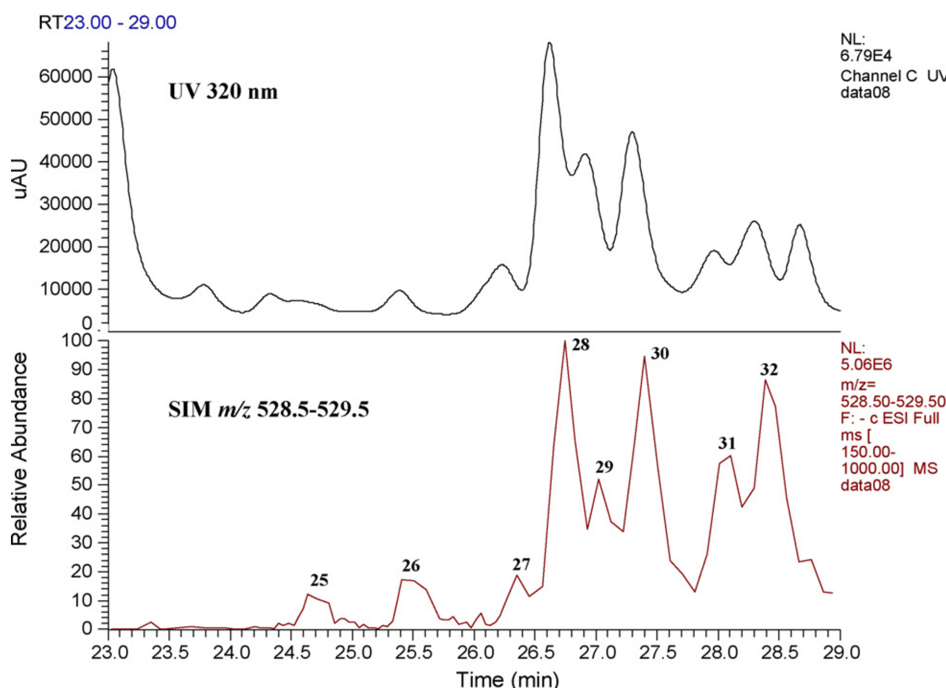


Fig. 3. Selected ion monitoring at *m*/*z* 529.

were thus plausible identified as 1-feruoyl-5-caffeoylquinic acid and 1-caffeoyl-5-feruoylquinic acid, respectively.

Three compounds (compounds **33**, **34**, **36**) exhibited $[M-H]^-$ ions at *m/z* 543. CID of them yielded product ions at *m/z* 367 ($[M-H-\text{ferulic acid}]^-$) or 349 ($[M-H-\text{ferulic acid}-H_2O]^-$) and 191 ($[M-H-2\text{ferulic acid}]^-$) or 193 ($[\text{ferulic acid}-H]^-$) corresponding to the identification of diferuoylquinic acids. The comparison of their fragmentation pattern and retention time with the dicaffeoylquinic acid, they were tentatively assigned as 3,4-diferuoylquinic acid, 3,5-diferuoylquinic acid, and 4,5-diferuoylquinic acid, respectively.

Compound **35** displayed an $[M-H]^-$ ion at *m/z* 677 and the CID of it yielded an $[M-H-162]^-$ ion at *m/z* 515 and an $[M-H-324]^-$ ion at *m/z* 353, indicating the successive losses of two caffeoyl groups. The MS^4 spectrum of it produced ion at *m/z* 179 ($[\text{caffeoyl acid}-H]^-$). It was thus identified as 3,4,5-tricaffeoylquinic acid.

CID of compound **37** led to competitive losses of 162 and 176 Da, representing caffeic acid and feruoyl acid, respectively. It was plausibly identified as dicaffeoylferuoylquinic acid.

Previously only chlorogenic acid and quinic acid have been identified from *A. annua* [3]. It has been reported that quinic acids have antioxidant activity [35,36] and hepatocyte protective activity [37]. In our present study, varieties of quinic acid derivatives were identified and they were also the relative abundant constituents in it. Hence we assume that Qinghao might possess these activities.

3.4. Identification of other compounds

Compound **6** gave $[M-H]^-$ ion at *m/z* 191, and the MS^2 spectra yielded the product ion at *m/z* 176, corresponding to the loss of one $-CH_3$ group. By searching the known constituents in *A. annua*, it was tentatively identified as scopoletin, a coumarin.

Two benzoicacid glucosides, 3,5-dihydroxybenzoicacid-glucoside (compound **1**, $[M-H]^-$ *m/z* 315) and 3,5-dihydroxy-benzoicacid-glucopylxyloside (compound **2**, $[M-H]^-$ *m/z* 447), were detected in *A. annua*. Their structures were established based on the characteristic MS^n fragmentations.

4. Conclusion

A simple and rapid method has been established for the screening of main phenolic compounds in *A. annua* by LC–DAD–ESI-MS^n. The *C*-glycosylflavones were found to be a new type of main constituents, and might suggest its antioxidant and antiviral activity. Quinic acid derivatives were also found to be the major constituents of the *A. annua*. It should be further investigated if these types of compounds are related to its antiparastic effect or they possess other bioactivities. On the other hand, quality control of traditional medicine is essentially different from chemical drugs for the complexity of multi-component feature. And it has been recognized that the therapeutic effects of traditional Chinese medicine are due to the synergistic contribution of multi-components, not just by the major constituents. Thus it is valuable to establish method covering most of the chemical constituents for the quality control of herbal medicines. Our present study provided scientific evidence for the establishment of global quality control method for *A. annua* and its related preparations.

Acknowledgements

We thank the Program for Changjiang Scholar and Innovative Team in University (985-2-063-112), the Cultivation Fund of the Key Scientific and Technical Innovation Project (no. 104218), Ministry of Education of China for financial support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jpba.2008.02.013.

References

[1] Y. Yang, J. Zhang, Y.T. Chen, The Extraction and Isolation Handbook of Natural Product Constituents, China Press of Traditional Chinese Medicine, 2003, p. 439.
[2] H. Matsuda, T. Morikawa, I. Toguchida, S. Harima, M. Yoshikawa, Chem. Pharm. Bull. 50 (2002) 972–975.
[3] J.P. Lai, Y.H. Lim, J. Su, H.M. Shen, C.N. Ong, J. Chromatogr. B 848 (2007) 215–225.
[4] M. Ye, Y.N. Yan, D.A. Guo, Rapid Commun. Mass Spectrom. 19 (2005) 1469–1484.
[5] M. Ye, J. Han, H.B. Chen, J.H. Zheng, D.A. Guo, J. Am. Soc. Mass Spectrom. 18 (2007) 82–91.
[6] R.X. Liu, M. Ye, H.Z. Guo, K.S. Bi, D.A. Guo, Rapid Commun. Mass Spectrom. 19 (2005) 1557–1565.
[7] N. Fabre, I. Rustan, E. de Hoffman, J. Quetin-Leclercq, J. Am. Soc. Mass Spectrom. 12 (2001) 707–715.
[8] E. de Rijke, H. Zappey, F. Ariese, C. Gooijer, U.A.T. Brinkman, J. Chromatogr. A 984 (2003) 45–58.
[9] Y.L. Ma, I. Vedernikova, H. van den Heuvel, M. Claeys, J. Am. Soc. Mass Spectrom. 11 (2000) 136–144.
[10] R.J. Hughes, T.R. Croley, C.D. Metcalfe, R.E. March, Int. J. Mass Spectrom. 210 (2001) 371–385.
[11] J.F. Stevens, M. Ivancic, V.L. Hsu, M.L. Deinzer, Phytochemistry 44 (1997) 1575–1585.
[12] P. Waridel, J.L. Wolfender, K. Ndjoko, K.R. Hobby, H.J. Major, K. Hostettmann, J. Chromatogr. A 926 (2001) 29–41.
[13] C.A.M. Pereira, J.H. Yariwake, M. McCullagh, Phytochem. Anal. 16 (2005) 295–301.
[14] Q.M. Li, H. van den Heuvel, O. Delorenzo, J. Corthout, L.A.C. Pieters, A.J. Vlietinck, M. Claeys, J. Chromatogr. 562 (1991) 435–446.
[15] W. Wu, F. Song, Z. Lu, S. Lu, Chem. J. Chin. Univ. 26 (2005) 27–30.
[16] F. Ferreres, B.M. Silva, P.B. Andrade, R.M. Seabia, N.A. Ferreira, Phytochem. Anal. 14 (2003) 352–359.
[17] F. Gerhard, G. Michael, Planta Med. 47 (1983) 131–140.
[18] A.A. Ahmed, N.H. El-Sayed, J. Tom, Rev. Lat. Am. Quim. 20 (1989) 5.
[19] K. Valant-Vetschera, Studies in Organic Chemistry, Amsterdam, 11 (1982), 213–220.
[20] K.M. Valant-Vetschera, Sci. Pharm. 62 (1994) 323–330.
[21] J.F. Gonnet, Biochem. Syst. Ecol. 21 (1993) 389–396.
[22] J. Reynaud, A. Couble, J. Raynaud, Pharmazie 47 (1992) 51–52.
[23] A.M. Zaghloul, O.M. Salama, A.F. Halim, J. Pharm. Sci. 6 (1990) 61–68.
[24] K. Kamanzi, J. Raynaud, B. Voirin, Pharmazie 38 (1983) 494–495.
[25] R.S. Frank, N.S. Lacey, T. George, L. Kim, Planta Med. 68 (2002) 1140–1141.
[26] O.D. Deliorman, M. Aslan, G. Aktay, Life Sci. 72 (2003) 2273–2283.
[27] K.H. Bohm, H. Lotter, O. Seligmann, Planta Med. 58 (1992) 544–548.

[28] J. Budzianowski, G. Pakulski, J. Robak, Pol. J. Pharmacol. Pharm. 43 (1991) 395–401.
[29] E. Gaitan, R.C. Cooksey, J. Legan, J. Clin. Endocrinol. Metab. 80 (1995) 1144–1147.
[30] Y.L. Li, S.C. Ma, Y.T. Yang, J. Ethnopharmacol. 79 (2002) 365–368.
[31] R.E. March, X.S. Miao, C.D. Metcalfe, Rapid Commun. Mass Spectrom. 18 (2004) 931–934.
[32] K. Schüta, D.R. Kammerer, R. Carle, A. Schieber, Rapid Commun. Mass Spectrom. 19 (2005) 179–186.
[33] C. Löffler, A. Sahm, V. Wray, F.C. Czygan, P. Proksch, Biochem. Syst. Ecol. 23 (1995) 121–128.
[34] N.C. Michael, L.J. Kelly, K. Susan, K. Nikolai, J. Agric. Food Chem. 51 (2003) 2900–2911.
[35] H.J. Kim, Y.S. Lee, Planta Med. 71 (2005) 871–876.
[36] T.M. Huang, M.K. Na, P.T. Thuong, N.D. Su, D.E. Sok, K.S. Song, Y.H. Seong, K.H. Bae, J. Ethnopharmacol. 108 (2006) 188–192.
[37] T. Xiang, Q.B. Xiong, A. Ketut, Y. Tezuka, T. Nagaoka, L.J. Wu, S. Kadota, Planta Med. 67 (2001) 322–325.