Analytical Methods

Rapid identification of polyphenols in Kudiezi injection with a practical technique of mass defect filter based on high performance liquid chromatography coupled with linear ion trap/Orbitrap mass spectrometry

Journal:	*Analytical Methods*
Manuscript ID:	AY-TEC-01-2014-000213.R1
Article Type:	Technical Note
Date Submitted by the Author:	28-Feb-2014

Complete List of Authors:	Wang, Fang; School of Chinese Pharmacy, Beijing University of Chinese Medicine, lu, jianqiu; Center of Scientific Experiment, Beijing University of Chinese Medicine, Zhang, Jia Yu; State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University Health Science Center, Yin, Pei hua; Center of Scientific Experiment, Beijing University of Chinese Medicine, Wang, Zi jian; Center of Scientific Experiment, Beijing University of Chinese Medicine, Dong, Lu yan; Beijing University of Chinese Medicine, School of Chinese Pharmacy,

Rapid identification of polyphenols in Kudiezi injection with a practical technique of mass defect filter based on high performance liquid chromatography coupled with linear ion trap/Orbitrap mass spectrometry

Fang Wang[a, b], Jiayu Zhang[b], PeihuaYin[a], Zijian Wang[b], Luyan Dong[a] and Jianqiu Lu[b*]

[a]*School of Chinese Pharmacy, Beijing University of Chinese Medicine, 6 Wangjing Zhonghuan South Road, Chao yang Dsitrict, Beijing 100102, China. E-mail: wangfangblx@gmial.com; Tel: +86 18810481565*

[b]*Center of Scientific Experiment, Beijing University of Chinese Medicine, 11 East Rd of North 3rd Ring, Chaoyang District, Beijing 100029, China. E-mail: lujq@vip.sina.com; Tel./Fax: +86 010 64286203*

Abstract

In the present study, a practical approach of mass defect filter (MDF), a data-mining technique, was developed and evaluated for the rapid classification of complicated peaks into well-known chemical families based on the exact mass acquired by high resolution mass spectrometry. The full-scan mass data of Kudiezi injection was acquired by high performance liquid chromatography coupled with linear ion trap-Orbitrap mass spectrometer system (HPLC-LTQ-Orbitrap) that features high resolution, mass accuracy and sensitivity. To screen the polyphenols including chlorogenic acids (CGAs), flavonoids in the injection, MDF approach was employed to rapidly screening them form the complex system. First, two filtering templates and several filters were set to remove the interference ions of complex matrix by MetWorks 1.3 Software. Then the target peaks filtered were characterized according to their accuracy mass data and MS/MS fragment ions. Utilizing the proposed approach, 14 CGAs and 16 flavonoids could be screened and identified. The results of rapid screening and detection showed that the developed MDF approach based on high-resolution mass spectrometry would be adaptable to the analysis of complex system of traditional Chinese medicines.

Keywords: Mass defect filter; LTQ-Orbitrap; Kudiezi injection; Polyphenols

Introduction

Since electrospray mass spectrometry has emerged as a major analytical tool for rapid identification of multiple compounds in traditional Chinese medicines (TCMs),[1-3] scientists have made great efforts to develop various LC/MS techniques and approaches for sensitive and selective detection of constituents in complex matrix. Reliable accurate mass measurements contributes to the right elemental formulae determination and structural elucidation of constituents hugely.[4] To date, high-resolution mass spectrometers (HRMS) have made a huge impact in a number of analytical fields, such as protein identification, protein modification, metabolomics, biomarker discovery, pesticide residue, drug screening, poison detection as well.[5] For example, the hybrid linear ion trap-Orbitrap mass spectrometer (LTQ-Orbitrap) has been introduced extensively,[6-7] which combines high trapping capacity and MS^n scanning function of the linear ion trap along with accurate mass measurements within 5 ppm and a resolving power of up to 100,000 over a wider dynamic range than that is achievable with many other mass spectrometers.[4] Particularly, Orbitrap facilitates fast data-dependent acquisition of accurate MS^n spectra on an LC timescale. Undoubtedly, these advantages could be used to increasing the throughput and identification efficiency of compounds. The combined use of LTQ and Orbitrap would be a better choice for identifying complicated components in TCMs.

However, it can be imaged that the application of LTQ-Orbitrap on TCMs would generate a large quantity of high-resolution information content which will lead to a new challenge of useful information processing. Therefore, some computational tools for data acquired on high resolution mass spectrometers have been reported to solve this problem. Mass defect filter (MDF) as a data-mining technology for finding drug metabolites was first proposed in 2003.[8] It is based on the realization that mass defect values (MD, the exact mass difference of a compound from a given nominal mass; for example, ^{16}O=15.9949, MD=-5.1 mDa; ^{14}N=14.0031, MD=3.1 mDa) of phase I and phase II metabolites typically fall within a defined narrow window (≤ 50 mDa, for example, mass defect shifts -5 mDa by hydroxylation, -23 mDa by demethylation, +32 mDa by glucuronidation, -43 mDa by sulfation) related to the parent drugs or core substructures determined. With mass range and mass defect range set, a significant number of ions outside the window can be removed.[9] Furthermore, in order to achieve more common or uncommon metabolites, multiple templates were set, according to the parent drugs and their metabolic

pathways, meanwhile algorithm was optimized.[10-12] In the past few years, MDF has evolved into a comprehensive methodology through a number of wonderful applications to drug metabolism, drug discovery, food field, natural organic matter, and natural products.[13-19] TCMs usually contain hundreds of secondary metabolites of plants undergoing some biosynthetic pathways such as shikimic acid pathway, acetate-malonate pathway, mevalonic acid pathway, which lead to owning similar structure skeletons. These series of processes of biosynthesis are similar to drug metabolism *in vivo*. Therefore, MDF technology is suitable for screening of compounds in TCMs. Although extract ion chromatogram (EIC) process is highly effective in the detection of common compounds with predictive molecular weights, screening global categorized compounds from TCMs by EIC is labour-intensive, and especially difficult to distinguish uncertain and trace components from complex chemical background in full-scan mass chromatograms rapidly.[20] As a post-acquisition data processing, MDF can overcome these problems to a certain extent, and EIC can only be adopted as the complementary approach of MDF to target certain and uncertain compounds.[21]

This study was thus designed to develop a practical methodology for rapid screening and identification of serial components in TCMs based on the combination of LTQ-Orbitrap and MDF technology. In order to examine the feasibility and reliability of the present approach, Kudiezi injection, a TCM preparation extracted and purified from the whole herb of *Ixeris sonchifolia* (Bge.) Hance was taken as an example. Kudiezi injection has been playing an important role in treatment of cardiovascular and cerebrovascular diseases. However, it is not clear about its effective constituents. As the main active components, chlorogenic acid (CGAs) and flavonoids were chosen to be illuminated using the established methodology. Significantly, this methodology could be envisioned to a wide application for the identification of categorized compounds or micro constituents *in vitro*.

Experimental

Materials and chemicals

The reference standards were obtained from the National Institutes for Food and Drug Control (Beijing, China), including luteolin, luteolin-7-O-β-D-glucoside and apigenin. The reference standards of neochlorogenic acid (3-CQA), chlorogenic acid (5-CQA), crypt chlorogenic acid (4-CQA), isochlorogenic acid A (3, 5-DiCQA), isochlorogenic acid B (3, 4-DiCQA) and

isochlorogenic acid C (4, 5-DiCQA) were purchased from Chengdu Biopurify Phytochemcials Co, Ltd (Sichuan, China). Apigenin-7-O-β-D-glucoside was purchased from Chengdu Deitian Creature Science Co. Ltd (Sichuan, China). Their purities were determined to be no less than 98% by HPLC-UV. The commercial products of Kudiezi injection, which were produced by Tonghua Huaxia Pharmaceutical Co, Ltd (Jilin, China), were purchased by prescription from hospital.

Acetonitrile, methanol and formic acid (HPLC grade) were purchased from Merck (Darmstadt, Germany). Ultra-pure water used throughout the experiment was produced by a Milli-Q purification system (Millipore, Bedford, USA).

HPLC conditions

Chromatography was performed on an Accela HPLC system equipped with a binary pump and an autosampler (Thermo Scientific, Bremen, Germany). Separation of the compounds was achieved on a Thermo Hypersil BDS C_{18} column (250 × 4.6 mm i.d., 5 μm) at room temperature. The mobile phase was composed of 0.1% formic acid (A) and acetonitrile (B), with the following gradient elution: 0-18 min, 2-8% B; 18-36 min, 8-12% B; 36-55 min, 12-18% B; 55-70 min, 18-25% B; 70-80 min, 25-30% B; 80-85 min, 30-40% B. The flow rate was at 1.0 $mL \cdot min^{-1}$.

Mass spectrometric conditions

A hybrid LTQ-Orbitrap XL mass spectrometer (Thermo Scientific, Bremen, Germany) was connected to the LC system via an electrospray ionization (ESI) interface in a post-column splitting ratio of 1: 3. High-resolution MS and MS/MS analysis was operated in negative ion mode with a mass range of *m/z* 100-1 200. Accurate mass analysis were calibrated according to the manufacturer's guidelines using a standard solution mixture of caffeine, sodium dodecyl sulfate, sodium taurocholate, the tetrapeptide MRFA acetate salt and Ultramark (Sigma Aldrich, St. Louis, MO, USA). The resolution of the Orbitrap mass analyzer was set at 30 000. Data-dependent MS^n scanning was performed to minimize total analysis time as it can trigger fragmentation spectra of target ions. Nitrogen was used as sheath and auxiliary gas. Helium served as collision gas. The isolation width was 2 amu, and the normalized collision energy was 35% for all compounds. Collision-induced dissociation (CID) was conducted in LTQ with an activation q of 0.25 and activation time of 30 ms. The key optimized ESI parameters were listed as follows: capillary temperature of 350 °C; sheath gas flow of 30 arb; auxiliary gas flow of 10 arb; source voltage of 4.0 kV; capillary voltage of -35 V; tube lens voltage of -110 V. MS scan functions and HPLC

solvent gradients were controlled by the Xcalibur data system (Thermo Scientific), while the data was collected and analyzed with Xcalibur 2.1 (Thermo Scientific).

Sample and standards preparation

Stock solutions of reference standards were prepared in methanol. All working solutions were prepared by diluting the stock solutions of CGAs and flavonoids, respectively. Kudiezi injection was filtered through a 0.22 μm nylon microporous membrane filter and an aliquot of 10 μL of the successive filtrate was injected into the HPLC-MS system for analysis.

Mass defect filter approach

The accurate mass full-scan raw data was processed by MDF using MetWorks 1.3 (Thermo Scientific) which facilitates the screening of characteristic components. The structural analogues in herbs usually shared similar core substructure, and characteristic compounds were generated via various substituents. Each substituent generated relatively changes in the mass defect of the core substructure. Therefore, parameter settings were related to the core substructure and combination of different substituents essentially. According to the above points, the first step was to establish MDF templates which is based on all structures of compounds published, and the second was to limit the mass defect range and mass range according to the substituents of various constitutes. Meanwhile, the number of filters was also essential to obtain satisfactory filtering chromatograms in which characteristic ions remain visible.

Results and discussion

Establishment of MDF approach to detect the polyphenols

Polyphenols in Kudiezi injection included CGAs and flavonoids, therefore two filtering templates were defined according to the two categories of compounds. CGAs are formed between quinic acid and one to four residues of certain cinnamic acids, including caffeic, *p*-coumaric, ferulic and sinapic commonly,[22] as illustrated in Fig. 1C. The distinctive characteristic of CGAs is that they have many isomers owing to the different substituted positions of cinnamic acids on quinic acid. Hence, the filtering template of CGAs was determined as quinic acid, and the four cinnamic acids introduced above were set as the substituents of CGAs predominantly. According to the summary of mass defects of the various substituents (Table II), the calculated mass defect of caffeoyl (MD: 31.2 mDa) was the smallest among them, and sinapoyl (57.9 mDa) was the largest which had the maximum value of mass weight (MW: 206 Da), while *p*-coumaroyl had the minimum value (146

Da). Owing to the number of substituents which usually came from one to three, the isomers had undergone relatively major changes in their molecular formulae which lead to wide mass defect and mass range. Therefore, three filters were set related to the substituted number (Table III). Filer 1 was set for screening the single esters. Caffoylquinic acid was assigned to obtain the minimum value of mass defect (86.7 mDa), and *p*-coumaroylquinic acid was assigned to obtain the minimum change of formula corresponding to $C_{16}H_{17}O_8$, while sinapoylquinic had the maximum value of mass defect and mass weight corresponding to $C_{18}H_{21}O_{10}$ (MW: 397 Da, MD: 112.9 mDa). The calculated mass defect range was from 86 to 113 mDa over the mass range of 337-398 Da. Similarly, mass defect ranges from 118 to 171 mDa and 150 to 229 mDa with mass ranges of 483-604 Da and 629-810 Da of filter 2 and filter 3 respectively were set for screening diesters and triesters. The generated chromatograms after filtrations were displayed in Fig. 3. It was anticipated that a total of 14 CGAs were detected without triesters observed.

Flavonoids in Kudiezi injection had the common skeleton of 5, 7-Dihydroxyflavone, therefore the filtering template was determined as 5, 7-dihydroxyflavone ($C_{15}H_9O_4$, MD: 50.6 mDa), as illustrated in Fig. 1B. The substituents of flavonoids were predominantly hydroxyl, methoxyl, formyl, glucuronide, hexoses, deoxyheoses. According to the summary of the mass defects of the various substituents, both hydroxyl and formyl produced the minimum decrease of mass defects; while others produced increase among which glucuronide was the least and deoxyheose the most (Table II). The sugar moieties resulting that the formulae of isomers had great changes made the mass range too wide, hence one filter would not obtain satisfactory results. Therefore three filers were set up according to the number of conjugated sugars (Table IV). Filter 1 was set for screening flavonoids without sugar moieties. The maximum number of three hydroxyls was assigned to obtain the minimum value of mass defect corresponding to $C_{15}H_9O_7$ (MD: 34.3 mDa) and the maximum was obtained by two methoxyls substituted. However, an assignment of one hydroxyl and two methoxyls was to produce the maximum value corresponding to $C_{17}H_{13}O_7$ (mass weight, MW: 329 Da; MD: 65.6mDa). Then the calculated mass defect range was from 34 to 71 mDa over the mass range of 253-330 Da. On the basis of above that, single sugar group linked, glucuronide and deoxyheose produced the minimum and maximum values of mass defect (66.4 mDa and 128.6 mDa) respectively, however deoxyheose and glucuronide possessed the minimum and maximum elemental compositions. Meanwhile there were three substituted

positions left, one hydroxyl and two methoxyls were contributed to the greatest change of formulae. Therefore the mass range was from 399-506 Da with 63-129 mDa of mass defect. Similarly, Filter 3 for screening flavonoids with two sugar groups conjugated set was based on the principles interpreted above. Then, the calculated mass defect range from 124 to 177 mDa over the mass range of 561-654 Da was applied to detect the flavonoid-disaccharides. As shown in Fig. 4, 16 potential flavonoids were filtered (Fig. 4C).

Identified CGAs of Kudiezi injection

By comparing with the reference standards, compound 2, 4, 6, 12, 13, 14 were 3-caffeoylquinic acid (3-CQA), 5-CQA, 4-CQA, respectively. For compound 1, its deprotonated molecular ion $[M - H]^-$ at *m/z* 353.0875 (molecular formula $C_{16}H_{17}O_9$) generated fragment ions at *m/z* 191, 173 and 335 corresponding to $[\text{quinic acid} - H]^-$, $[\text{quinic acid} - H - H_2O]^-$, $[M - H - H_2O]^-$, respectively. According to the elution order on the reversed-phase column, compound 1 was tentatively identified to be 1-CQA. Compounds 3, 7 and 8 all gave $[M - H]^-$ ion at *m/z* 337.0918 ($C_{16}H_{17}O_8$) corresponding to *p*-coumaroylquinic acid (*p*-CoQA). In their MS^2 spectra, the base peaks were different significantly. 3-*p*CoQA and 5-*p*CoQA yielded their respective MS^2 base peak at *m/z* 163 $[\text{coumaric acid} - H]^-$ and *m/z* 191 $[\text{quinic acid} - H]^-$, while 4-*p*CoQA generated MS^2 base peak at *m/z* 173 $[\text{quinic acid} - H - H_2O]^-$. Both 1-*p*CoQA and 5-*p*CoQA would generate same MS^2 base peak at *m/z* 191 $[\text{quinic acid} - H]^-$.[22] However, the polarity of 5-*p*CoQA is weaker than that of 3-*p*CoQA. Hence, compound 3, 7 and 8 were characterized to be 3-*p*CoQA, 5-*p*CoQA and 4-*p*CoQA, respectively. Furthermore, three feruloylquinic acids (FQA) were detected, including 3-FQA, 4-FQA and 5-FQA. In the previously reports,[23-24] 3-FQA, 4-FQA and 5-FQA generated MS^2 base peak at *m/z* 193 $[\text{ferulic acid} - H]^-$, *m/z* 173 $[\text{quinic acid} - H - H_2O]^-$ and *m/z* 191 $[\text{quinic acid} - H]^-$, respectively. Their ESI-MS^n information was shown in Table I.

In the same experiment, four dicaffeoylquinic acids (DiCQA) were observed in Fig. 3B. By comparing with the reference standards, compounds 12, 13 and 14 were assigned as 3, 4-DiCQA, 3, 5-DiCQA and 4, 5-DiCQA, respectively. For peak 11 owing same deprotonated molecular ion $[M - H]^-$ at *m/z* 515.1193 ($C_{25}H_{23}O_{12}$) produced the predominant fragment ion at *m/z* 353 in MS^2 spectrum and *m/z* 191 in MS^3 spectra. According to the reports in the literature,[25-26] 1-, 3-, or 5-substituted positions of the compound would be substituted. On RP-ODS column, 1, 3-DiCQA was remarkably in advance of 3, 4-DiCQA eluted.[22] Therefore, compound 11 was identified as 1,

3-DiCQA.

Identified flavonoids of Kudiezi injection

Compounds 22, 27 and 29 were identified as luteolin-7-*O*-β-D-glucoside, apigenin-7-*O*-β-D-glucoside and luteolin by comparing their retention times and high-resolution mass spectra with those of reference standards.

In Fig. 3A, compound 30 produced its $[M - H]^-$ ion at *m/z* 283.0608 ($C_{16}H_{11}O_5$). Further fragmentation of the ion resulted in $[M - H - CH_3\cdot]^-$ ion at *m/z* 268 and $[M - H - CH_3\cdot - CO]^-$ ion at *m/z* 240, consistent with acacetin.[27] Therefore, compound 30 was tentatively identified as acacetin.

Flavonoid-*O*-saccharide: Both of the deprotonated molecular ions of compound 21 and 28 eliminated a glucuronic acid residue to produce $[aglycone - H]^-$ ions at *m/z* 285 and *m/z* 269, respectively. Furthermore, the fragment ions of the *m/z* 285 and *m/z* 269 ions were in accordance with those fragmentation pathways of luteolin and apigenin. Compound 21 and 28 were therefore assigned as luteolin-7-*O*-*β*-D-glucuronide and apigenin-7-*O*-*β*-D-glucuronide, respectively.

Flavonoid-*O*-disaccharide: Compound 15 in Fig. 4C with the deprotonated molecular ion $[M - H]^-$ at *m/z* 625.1407 ($C_{27}H_{29}O_{17}$), yielded Y_1^- $[M - H - 162]^-$, Y_0^- $[M - H - 324]^-$ and $[Y_1 - H]^-$ (*m/z* 462) ions named by the structured naming rules proposed by Demon and Costello. The mass difference of 162 and 324 Da indicated the loss of two glucosyls. The $[M - H - 162]^-$ ion yielded Y_0^- $[M - H - 324]^-$ as the base peak, accompanied by the ions of $[Y_0 - H]^-$ and $[Y_0 - 2H]^-$, indicating that the compound was belong to flavonol, and glycosyls were linked to two different hydroxyl positions.[28] Hence, compound 15 was deduced as quercetin-di-*O*-glycosides. Meanwhile, the ion at *m/z* 301 $[M - H - 324]^-$ was the base peak of compound 17 with $[Y_0 - H]^-$, $[M - H - 162]^-$ and $[M - H - 180]^-$ ions not observed. So compound 17 was plausibly identified as quercetin-7-*O*-gentiobioside.

Compound 16 produced $[M - H - 162]^-$ ion at *m/z* 477 (100%) and $[M - H - 324]^-$ ion at *m/z* 315 (15.7%) in its MS^2 spectrum corresponding to two glycosyls linked. As 1→2 substituted glycoside eliminated glycoside easily one by one resulting in *m/z* 315 ion at lower relatively intensity, its disaccharide moiety was 1→2 linkage. Y_0^- and $[Y_0 - H]^-$ ions (*m/z* 315 and 314) of aglycone residue were detected simultaneously, and the fragment ions of aglycone were consistent with those of isorhamnetin. Hence, compound 16 was identified as isorhamnetin-3-*O*-sophorosides.[29]

In MS^2 spectrum, the deprotonated molecular ion of compound 18 produced $[M - H - 162]^-$ ion

at *m/z* 447 (100%), and [M - H - 162 - 162]$^-$ ion at *m/z* 285 (100%) in its MS^3 spectrum, corresponding to the loss of a disaccharide moiety. By ananlyzing the relative intensities of fragment ions, its disaccharide moiety was 1→2 linkage. Owing to the fragment ions of aglycone consistent with those of luteolin, compound 18 was plausibly identified as luteolin-7-*O*-*β*-D-sophoroside. Compound 19 had the same deprotonated molecular ion, and yielded prominent [M - H - 324]$^-$ ion at *m/z* 285. In addition, the *m/z* 447 [M - H - 162]$^-$ ion was found at very low relative abundance (3.4%) and *m/z* 429 [M - H - 180]$^-$ ion was not detected.[30] Hence, compound 19 was deduced as luteolin-7-*O*-*β*-D-gentiobioside.[31] In the same way, compound 23 was identified as apigenin-7-*O*-*β*-D-gentiobioside. The aglycone of compound 24 was assigned as *O*-methylated luteolin (diosmetin or chrysoeriol) due to the appearance of the fragment [Y_0 - CH_3·]$^-$ of Y_0^- in its MS^2 spectrum.[32-33] Nevertheless, methylated position could not be determined owing to the similar properties and the lack of standards. Thus, compound 24 was tentatively characterized as methylated luteolin-*O*-gentiobioside.

Compound 20 and 25 produced [M - H - 308]$^-$ ion at *m/z* 285 in their MS^2 spectra, without [M - H - 120]$^-$ ion, suggesting the loss of a rutinose.[28] Therefore, compound 20 and 25 were deduced as luteolin-*O*-rutinosides. In MS^2 and MS^3 spectra of compound 26, ions at *m/z* 431 [M - H - 146]$^-$ and *m/z* 285 [M - H - 292]$^-$, suggesting that two rhamnose residues eliminated, could be observed. Therefore, compound 26 was tentatively deduced as luteolin-di-*O*-rhamnoses.

Discussion

In this experiment, owing to CGAs and flavonoids which belong to two different chemical families, two mass defect filter templates were set for screening. Meanwhile, considering the structural characteristics, one filter window set of which the mass defect range and mass range were too wide, lead to weak changes between the filtered and original chromatograms. The background interference ions fell within the ranges of filtering window, and their intensity was stronger than the target. As a result, the target ions were still covered and difficult to identification. Hence, appropriate number of filter was necessary.

Conclusions

This report took the advantage of the LTQ-Orbitrap mass spectrometry system to establish an post-acquired data processing method with MDF in order to perform rapid and global detection of structural analogues (CGAs and flavonoids). Compared with the conventional manual inspection,

the MDF approach enabled original data to be analyzed in a much faster time frame and the compounds in chromatograms to be displayed clearly by reducing the potential interferences of matrix ions. Additionally, various filtering templates and filter number would be beneficial to obtain satisfactory results and classify homologous families, especially when the mass weight and mass defect ranges of categorized formulae are too wide. Suitable parameters are extremely crucial for global screening of homologous compounds in TCMs. Significantly, this methodology could be extended to other analysis fields, such as natural organic matter in natural waters and soils or sediments, nutrients in food, pesticide residues in fruit and vegetables, new drug screening or poison detection and so on.

Acknowledgments

The authors greatly appreciate the financial support from the National Basic Research and Development Program of China (No. 2012CB518406), and the Post-graduate Automatic Research of Beijing University of Chinese Medicine (No. 2013-JYBZZ-XS-062).

References

1. J. Y. Zhang, Q. Zhang, H. X. Zhang, Q. Ma, J. Q. Lu and Y. J. Qiao, *J. Agric. Food Chem.*, 2012, **60**, 9023-9034.
2. J. Y. Zhang, N. Li, Y. Y. Che, Y. Zhang, Y. Jiang and P. F. Tu, *J. Pharm. Biomed. Anal.*, 2011, **56**, 950-961.
3. J. Y. Zhang, J. Q. Lu, Q. Zhang, L. Dai, Y. Liu, P. F. Tu and Y. J. Qiao, *J. Chromatogr. Sci.*, 2013, doi:10.1093/chromasci/bms253.
4. Y. Liang, G. J. Wang, L. Xie and L. S. Sheng, *Curr. Drug Metabol.*, 2011, **12**, 329-344.
5. R. M. Markus and H. M. Hans, *Anal. Bioanal. Chem.*, 2012, **403**, 1221-1231.
6. A. Makarov, E. Denisov, A. Kholomeev, W. Balschun, O. Lange, K. Strupat and S. Horning, *Anal. Chem.*, 2006, **78**, 2113-2120.
7. A. Makarov, E. Denisov, O. Lange and S. Horning, *J. Am. Soc. Mass Spectrom.*, 2006, **17**, 977-982.
8. H. Y. Zhang, D. L. Zhang and K. Ray., *J. Mass Spectrom.*, 2003, **38**, 1110-1112.
9. M. Zhu, H. Y. Zhang and W. G. Humphreys, *J. Biol. Chem.*, 2011, **286**: 25419-25425.
10. J. E. Barbara and J. M. Castro Perez, *Rapid Commun. Mass Spectrom.*, 2011, **25**, 3029-3040.
11. M. Zhu, L. Ma, D. L. Zhang, K. Ray, W. P. Zhao, W. G. Humphreys, G. Skiles, M. Sanders and H. Zhang, *Drug Metab. Dispos.*, 2006, **34**, 1722-1733.
12. F. Cuyckens, R. Hurkmans, J. M. Castro Perez, L. Leclercq and R. J. Mortishire Smith, *Rapid Commun. Mass Spectrom.*, 2009, **23**, 327-332.
13. J. X. Tian, C. Peng, L. Xu, Y. Tian and Z. J. Zhang, *Biomed. Chromatogr.*, 2013, **27**, 775-783.
14. F. Du, Q. Ruan, M. Zhu and J. Xing, *J. Mass Spectrom.*, 2013, **48**, 413-422.
15. M. Zhu, H. Y. Zhang and W. G. Humphreys, *J. Biol. Chem.*, 2011, **286**, 25419-25425.
16. F. L. Chu, L. Pirastru, R. Popovic, and L. Sleno, *J. Agric. Food Chem.*, 2011, **59**, 3004-3013.
17. K. Nagy, L. Sandoz, B. D. Craft, and F. Destaillats, *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess*, 2011, **28**, 1492-1500.
18. H. Zhang, M. Zhu, K. L. Ray, L. Ma, and D. Zhang, *Rapid Commun. Mass Spectrom.*, 2008,

22, 2082-2088.

19. Z. M. Gu, L. Q. Wang, and J. Wu, *Natural Product Journal*, 2011, **1**, 135-145.
20. G. L. Yan, H. Sun, W. J. Sun, L. Zhao, X. C. Meng and X. J. Wang, *J. Pharm. Biomed. Anal.*, 2010, **53**, 421-431.
21. Q. Ruan, S. Peterman, M. A. Szewc, L. Ma, D. Cui, W. G. Humphreys and M. Zhu, *J. Mass Spectrom.* 2008, **43**, 251–261.
22. J. Y. Zhang, Q. Zhang, N. Li, Z. J. Wang , J. Q. Lu and Y. J. Qiao, *Tanlata*, 2013, **104**, 1-9.
23. I. Parveen, M. D. Threadgill, B. Hauck, I. Donnison and A. Winters, *Phytochemistry*, 2011, **72**, 2376-2384.
24. D. Perrone, A. Farah, C. M. Donangelo, P. Tomas de and P. R. Martin, *J. Food Chem.*, 2008, **106**, 859-867.
25. M. N. Clifford, S. Knight and N. Kuhnert, *J. Agric. Food Chem.*, 2005, 53, 3821-3832.
26. R. Jaiswal, J. Kiprotich and N. Kuhnert, *Phytochemistry*, 2011, **72**, 781-790.
27. F. Kuhn, M. Oehme, F. Romero, E. Abou-Mansour and R. Tabacchi, *Rapid Commun. Mass Spectrom.*, 2003, **17**, 1941-1949.
28. F. Ferreres, R. Llorach and A. Gil-Izquierdo, *J. Mass Spectrom.*, 2004, **39**, 312-321.
29. L. Z. Lin and J. M. Harnly, *Food Chem.*, 2010, **120**, 319-326.
30. W. Wu, Z. Liu, F. Song and S. Liu, *Anal. Sci.*, 2004, **20**, 1103-1105.
31. P. Y. Shi, Y. F. Zhang, H. B. Qu and X. H. Fan, *Phytochem. Anal.*, 2011, **22**, 66-73.
32. P. Y. Shi, Q. He, Y. Song, H. B. Qu and Y. Y. Cheng, *Anal. Chim. Acta,* 2007, **598**, 110-118.
33. G. Gattuso, C. Caristi, C. Gargiulli, Bellocco, E., G. Toscano and U. Leuzzi, *J. Agric. Food Chem.*, 2006, **54**, 3929-3935.

Table I Characterization of polyphenols in Kudiezi injection by HPLC-HR-ESI-MS^n

No.	t_R(min)	Formula $[M-H]^-$	Theoretical Mass *m/z*	Experimental Mass *m/z*	Mass error (ppm)	MS^n (*m/z*) P-ion (%)[b]	Identification
1	9.92	$C_{16}H_{17}O_9$	353.0867	353.0875	2.2	MS^2[353]: 191 (100), 173 (30.1), 335 (18.4)	1-CQA
2	17.29	$C_{16}H_{17}O_9$	353.0867	353.0876	2.5	MS^2[353]: 191 (100), 179 (44.7), 135 (6.8), 173 (4.0)	3-CQA
3	22.44	$C_{16}H_{17}O_8$	337.0918	337.0927	2.6	MS^2[337]: 163 (100)	3-*p*-CoQA
4	25.63	$C_{16}H_{17}O_9$	353.0867	353.0877	2.7	MS^2[353]: 191 (100), 179 (3.2)	5-CQA
5	26.23	$C_{17}H_{19}O_9$	367.1024	367.1030	1.8	MS^2[367]: 193 (100), 134 (10.8)	3-FQA
6	27.69	$C_{16}H_{17}O_9$	353.0867	353.0876	2.6	MS^2[353]: 173 (100), 179 (53.6), 191 (13.0), 135 (5.9)	4-CQA
7	34.29	$C_{16}H_{17}O_8$	337.0918	337.0926	2.4	MS^2[337]: 191 (100), 163 (13.3)	5-*p*-CoQA
8	35.30	$C_{16}H_{17}O_8$	337.0918	337.0927	2.5	MS^2[337]: 173 (100), 163 (6.2)	4-*p*-CoQA
9	38.87	$C_{17}H_{19}O_9$	367.1024	367.1032	2.2	MS^2[367]: 191 (100), 193 (10.2)	5-FQA
10	39.44	$C_{17}H_{19}O_9$	367.1024	367.1032	2.3	MS^2[367]: 173 (100), 193 (13.4), 191 (2.6)	4-FQA
11	38.79	$C_{25}H_{23}O_{12}$	515.1184	515.1193	1.8	MS^2[515]: 353 (100), 179 (19.1), 335 (9.2) MS^2[353]: 191 (100), 179 (38.2), 135 (6.5)	1,3-CQA
12	58.78	$C_{25}H_{23}O_{12}$	515.1184	515.1203	3.7	MS^2[515]: 353 (100), 173 (16.3), 335 (15.3) MS^2[353]: 173 (100), 179 (72.3), 191 (46.6), 135 (8.1)	3,4-CQA
13	60.17	$C_{25}H_{23}O_{12}$	515.1184	515.1201	3.3	MS^2[515]: 353 (100), 471 (3.5), 179 (3.3) MS^3[353]: 191 (100), 179 (47.5),173 (5.9),135 (5.4)	3,5-CQA
14	64.73	$C_{25}H_{23}O_{12}$	515.1184	515.1199	3.0	MS^2[515]: 353 (100), 173 (6.7) MS^3[353]: 173 (100), 179 (56.9), 191 (26.1), 135 (4.0)	4,5-CQA
15	32.57	$C_{27}H_{29}O_{17}$	625.1399	625.1407	1.2	MS^2[625]: 463 (100), 301 (32.0), 462 (17.8) MS^3[463]: 301 (100), 300 (39.8), 299 (12.0) MS^4[301]: 151 (100), 179 (61.3)	Quercetin-di-*O*-glycoside

16	40.78	$C_{28}H_{31}O_{17}$	639.1556	639.1564	1.2	MS^2[639]: 477 (100), 315 (15.7) MS^3[477]: 314 (100), 315 (44.6) MS^3[314]: 271 (100), 285 (99.7), 300 (59.7)	Isorhamnetin-3-*O*-sophorosides
17	44.36	$C_{27}H_{29}O_{17}$	625.1399	625.1403	0.5	MS^2[625]: 301 (100) MS^3[301]: 251 (100), 283 (88.5), 257 (83.4)	Quercetin-7-*O*-gentiobioside
18	44.65	$C_{27}H_{29}O_{16}$	609.1450	609.1461	1.8	MS^2[609]: 447 (100) MS^3[447]: 285 (100), 327 (1.7) MS^4[285]: 199 (100), 241 (92.8), 175 (84.4), 217 (72.1), 243 (55.4), 151 (31.4), 201 (22.0), 213 (21.7)	Luteolin-7-*O*-*β*-D-sophorosides
19	49.13	$C_{27}H_{29}O_{16}$	609.1450	609.1459	1.5	MS^2[609]: 285 (100), 286 (7.3), 447 (3.4) MS^3[285]: 285 (100), 199 (13.5), 241 (12.9), 217 (10.1), 243 (8.4), 151 (5.7)	Luteolin-7-*O*-*β*-D-gentiobioside
20	54.42	$C_{27}H_{29}O_{15}$	593.1501	593.1509	1.4	MS^2[593]: 285 (100), 286 (7.8), 447 (1.2) MS^3[285]: 199 (100), 241 (98.5), 175 (97.1), 217 (83.3), 243 (65.4), 151 (38.4)	Luteolin-*O*-rutinosides.
21	55.10	$C_{21}H_{17}O_{12}$	461.0715	461.0725	2.3	MS^2[461]: 285 (100) MS^3[285]: 241 (100), 175 (92.4), 199 (91.4), 217 (73.5), 243 (63.8), 151 (37.6)	Luteolin-7-*O*-*β*-D-glucuronide
22	55.57	$C_{21}H_{19}O_{11}$	447.0922	447.0933	2.5	MS^2[447]: 285 (100) MS^3[285]: 199 (100), 241 (93.1), 175 (91.4), 217 (76.1), 243 (63.6), 151 (39.3), 213 (27.0)	Luteolin-7-*O*-β-D-glucoside
23	55.57	$C_{27}H_{29}O_{15}$	593.1501	593.1509	-0.3	MS^2[593]: 269 (100) MS^3[269]: 225 (100), 197 (32.8), 227 (30.2), 183 (27.3), 201 (26.5), 149 (25.8), 181 (24.5)	Apigenin-7-*O*-*β*-D-gentiobioside
24	57.14	$C_{28}H_{31}O_{16}$	623.1607	623.1619	1.9	MS^2[623]: 299 (100), 284 (24.5)	Methylated luteolin-*O*-gentiobio

							side
25	57.99	$C_{27}H_{29}O_{15}$	593.1501	593.1514	2.2	MS^2[593]: 285 (100), 447 (0.6) MS^3[285]: 199 (100), 241 (99.6), 175 (88.9), 217 (82.8), 243 (46.32)	Luteolin-*O*-rutinosides
26	59.03	$C_{27}H_{29}O_{14}$	577.1552	577.1564	2.1	MS^2[577]:431 (100) MS^3[577]:285 (100)	Luteolin-di-*O*-rhamnoses
27	62.40	$C_{21}H_{20}O_{10}$	431.0973	431.0986	3.0	MS^2[431]: 269 (100) MS^3[269]: 225 (100), 197 (39.5), 149 (26.2), 201 (23.1), 227 (22.8), 151 (16.3)	Apigenin-7-*O*-β-D-glucoside
28	62.65	$C_{21}H_{17}O_{11}$	445.0765	445.0780	3.2	MS^2[445]: 269 (100) MS^3[269]: 225 (100), 149 (37.3), 201 (29.9), 151 (24.4), 183 (22.8), 227 (19.8) MS^4[225]: 181 (100), 197 (58.2), 183 (28.3), 196 (27.3)	Apigenin-7-*O*-*β*-D-glucuronide
29	74.59	$C_{15}H_9O_6$	285.0394	285.0403	3.3	MS^2[285]: 241 (100), 199 (80.7), 175 (79.8), 217 (69.1), 243 (63.5), 151 (33.4), 213 (23.7)	Luteolin
30	75.21	$C_{16}H_{11}O_5$	283.0601	283.0608	2.4	MS^2[283]: 268 (100) MS^3[268]: 268 (100), 240 (43.0), 239 (14.8)	Acacetin

Table II The substituents of CGAs and flavonoids

Substituent	Formula Change(mass)	Mass Weight Change (Da)	Mass Defect Shift (mDa)
p-Coumaroyl	+ $C_9H_6O_2$	146.0362	+ 36.2
Caffeoyl	+ $C_9H_6O_3$	162.0312	+ 31.1
Feruloyl	+ $C_{10}H_8O_3$	176.0468	+ 46.8
Sinapoyl	+ $C_{11}H_{10}O_4$	206.0579	+ 57.9
Formyl	+ CO	15.9949	- 5.1
Hydroxyl	+ O	15.9949	- 5.1
Methoxyl	+ OCH_2	30.0106	+ 10.6
Hexose (Glc)	+ $C_6H_{10}O_5$	162.0528	+ 52.8
Deoxyheose (Rha/ Fuc)	+ $C_6H_{10}O_4$	146.0679	+ 57.9
Glucuronide	+ $C_6H_8O_6$	176.0320	+ 32.1

"__": the maximum mass weight change and mass defect shift;

"...": the minimum mass weight change and mass defect shift.

Table III The filter settings of CGAs

Filters	Mass Change (Da)		Mass Defect Shift (mDa)	
	Min	Max	Min	Max
1	+ *p*-Coumaroyl	+ Sinapoyl	+ Caffeoyl	+ Sinapoyl
	$C_{16}H_{17}O_8$ (337)	$C_{18}H_{21}O_{10}$ (398)	$C_{16}H_{17}O_9$ (353.0867, MD: 86)	$C_{18}H_{21}O_{10}$ (397.1129, MD: 113)
2	+ 2*p*-Coumaroyl	+ 2Sinapoyl	+ 2Caffeoyl	+ 2Sinapoyl
	$C_{25}H_{23}O_{10}$ (483)	$C_{29}H_{31}O_{14}$ (604)	$C_{25}H_{23}O_{12}$ (515.1184, MD: 118)	$C_{29}H_{31}O_{14}$ (603.1708, MD: 171)
3	+ 3*p*-Coumaroyl	+ 3Sinapoyl	+ 3Caffeoyl	+ 3Sinapoyl
	$C_{34}H_{29}O_{12}$ (629)	$C_{40}H_{41}O_{18}$ (810)	$C_{34}H_{29}O_{15}$ (677.1501, MD: 150)	$C_{40}H_{41}O_{18}$ (809.2287, MD: 229)

Table IV The filter settings of flavonoids

Filters	Mass Change (Da)		Mass Defect Shift (mDa)	
	Min	Max (+ 2-OCH_2 + -OH)	Min (+ 3-OH)	Max (+ 2-OCH_2)
1 (Aglycone)	$C_{15}H_9O_4$ (253)	$C_{17}H_{13}O_7$ (330)	$C_{15}H_9O_7$ (301.0343, MD: 34)	$C_{17}H_{13}O_6$ (313.0707, MD: 71)
2 (Flavonoid-saccharide)	+ Deoxyheose	+ Glucuronide	+ Glucuronide	+ Deoxyheose
	$C_{21}H_{19}O_8$ (399)	$C_{23}H_{21}O_{13}$ (506)	$C_{21}H_{17}O_{13}$ (477.0664, MD: 66)	+ $C_{23}H_{23}O_{10}$ (459.1286, MD: 129)
3 (Flavonoid-disaccharide)	+ -OH + 2Deoxyheose	+ -OH + 2Hexose	+ 2Glc	+ Deoxyheose + Glc
	$C_{21}H_{19}O_8$ (561)	$C_{29}H_{33}O_{17}$ (654)	$C_{27}H_{27}O_{17}$ (623.1243, MD: 124)	$C_{29}H_{33}O_{16}$ (637.1763, MD: 177)