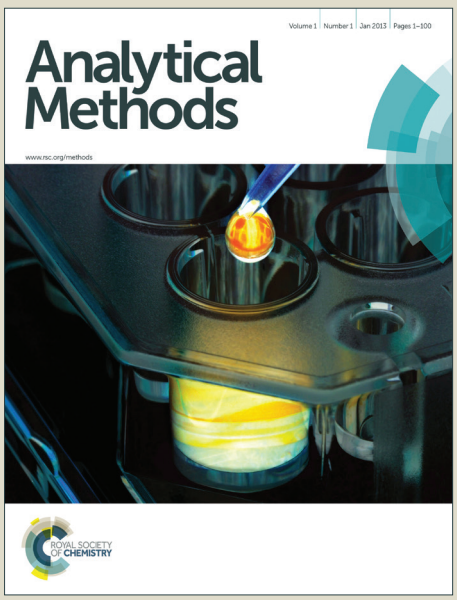
Analytical Methods

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

PAPER

Rapid classification and identification of complex chemical compositions in traditional Chinese medicine based on UPLC-Q-TOF/MS coupled with data processing technique using KuDieZi injection as an example

Lei Yuan[a,1], Zhenzhu Zhang[a,1], Zhiguo Hou[a,1], Bin Yang[a], Aizhu Li[a], Xuejun Guo[a], Yuming Wang[a] and Yubo Li[a]*

Abstract

KuDieZi (KDZ) injection is prepared by extracting and processing *Ixeris sonchifolia* [*Ixeris sonchifolia* Hance] belonging to *Lactuca* genus of the Asteraceae family. KDZ injection is a single-herb preparation in which components are analysed by liquid chromatography–mass spectrometry. However, the chemical compositions of this preparation are complex and diverse. Hence, data processing is complicated and time consuming; furthermore, data processing can not provide a systematic, accurate and repeatable method to rapidly classify and identify chemical constituents of the injection. In our study, the main components, particularly flavonoids, organic acids, amino acids and nucleosides in KDZ injection, were rapidly classified and identified by data processing technology based on UPLC-Q-TOF/MS. After we reviewed lots of studies and collected the fragments' information, then compared with the mass spectrometric analyses in standards, the rules of diagnostic fragments (DFs) and neutral losses (NLs) of the four substances were established and summarised. A rapid classification and identification method of the chemical compositions of KDZ injection was then constructed using DF filter (DFF) and NL filter (NLF). This method was applied to analyse the KDZ injection. A total of 31 chemical components, which included 8 flavonoids, 13 organic acids, 6 amino acids and 4 nucleosides, were obtained. DFF and NLF were used to rapidly classify and identify chemical substances in KDZ fingerprint. With this method, we effectively solved the technical difficulties in fingerprint resolution caused by complex components and low levels in traditional Chinese medicine (TCM). In addition, this study provided a novel approach for further studies on TCM.

1 Introduction

KuDieZi (KDZ) injection is prepared by extracting and processing *Asteraceae lactuca* genus *Ixeris sonchifolia* [*Ixeris sonchifolia Hance*]. It helps to improve microcirculation and anti-platelet aggregation, lowers blood pressure and treats coronary heart disease, angina, infarction and other clinical diseases.[1–2] Although KDZ injection is a single-herb preparation, its compositions are complicated; the main components of KDZ are nucleosides, organic acids, flavonoids, amino acids and other compounds.[3–4] Liquid chromatography–mass spectrometry (LC–MS) has been widely applied to investigate the components of single-herb and complex preparations with its advantages including high separation, high selectivity and good sensitivity.[5–6] LC–MS/MS or MS^n has also been applied to analyse substances in traditional Chinese medicine (TCM); however, this method displays drawbacks, including complicated and time-consuming data processing.[7–9] Therefore, a systematic and reliable method should be developed to determine active ingredients and elucidate action mechanisms of TCM.

In pharmaceutical analysis, the whole spectrum of drugs is scanned by UPLC-Q-TOF/MS; however, this method exhibits several disadvantages, such as complex spectrum and amount of information. Hence, chemometrics and data processing are necessary to mine and integrate original data information. With the development of data processing technology, diagnostic fragment filter (DFF, the characteristic fragment of a certain type of compound applied for screening and identification) and neutral loss filter (NLF, the neutral loss of a certain type of compound applied for screening and identification) show unique advantages of screening and identifying compounds.[8–13] Compounds with the same or similar skeleton can be cleaved into different fragments under energy bombardment; some of these fragments can be used to infer the type of cleavage and classify the subtype of substances; these diagnostic fragments (DFs) help screen target components and filter congeners.[14–20] Moreover, the discrepancy of *m*/*z* between a parent ion and a fragment ion (neutral loss, NF) in highly responsive parts is essential for identification.[21] Thus, researchers should develop a method in which DFF is combined with NLF to rapidly and accurately classify and identify complex components of KDZ injection.

The complexity of chemical substances and the unsystematic and unreliable analytical methods directly cause difficulty in classifying and identifying components of TCM during analysis; thus, the wide application of TCM injection in clinics is limited. Further developments related to global health have been hampered. Therefore, analytical methods should be developed to rapidly classify and identify complex TCM components. Based on combined relevant literature and previous experimental results, the rules of DFs and NLs regarding four categories, particularly flavonoids, organic acids, amino acids and nucleosides, were established and summarised in this study. Different DFs and NLs were used to rapidly classify and identify the chemical compositions of these categories in the KDZ injection fingerprint by UPLC-Q-TOF/MS. The identities of these compounds were then confirmed by comparing with those described in previous studies. Comprehensive analysis and systematic integration of the known components in the KDZ injection were conducted on the basis of commonalities and specificity in structure of the four categories. With the aid of data processing tools, such as DFF and NLF, we established a method to rapidly classify and identify

compounds, which exhibit important implications on the investigation of complex TCM injection. This study helped solve the key issue on rapid classification and identification of the complex chemical compositions of TCM injections to a certain degree; thus, this study provided a strong basis to assist in the development of rapid classification and identification methods for TCM.

2 Experimental

2.1 Standards and Reagents

Batches (no. 111201) of KDZ injections were provided by Jilin Tonghua China Pharmaceutical Co., Ltd. (Jilin, China). Luteolin-7-*o*-*β*-*D*-glucopyranoside, rutin, ferulic acid, proline, valine, uridine and adenine were obtained from Pharmaceutical and Biological Products (Beijing, China). Dicaffeoyl tartrate was purchased from Shanghai Yuanye Bio-Technology Co., Ltd. (Shanghai, China). Chlorogenic acid was obtained from Tianjin Yifang Technology Co., Ltd. (Tianjin, China). HPLC-grade methanol and acetonitrile were purchased from Oceanpak (Goteborg, Sweden). Distilled water was provided by Wahaha Co., Ltd. (Hangzhou, China). Standard compounds were dissolved in methanol for UPLC-Q-TOF/MS analysis.

2. 2 Preparation of samples

Approximately 1 mL of each of the 10 KDZ injections was mixed evenly and filtered using a microporous membrane (0.22 μm) for direct injection.

2. 3 UPLC-Q-TOF/MS conditions

A Waters Acquity UPLC Class I series equipped with a quat pump, an autosampler, a DAD detector and a column compartment was used for the analysis. The analytical column was a Waters ACQUITY UPLC BEH C18 column (2.1 mm × 50 mm, 1.7 μm) with a column temperature maintained at 35 ℃. The mobile phase was composed of eluent A (0.05% formic acid in water, *v*/*v*) and B (0.05% formic acid in acetonitrile, *v*/*v*), and a gradient elution was employed for the separation. The flow rate was maintained at 0.3 mL/min. The following elution conditions were applied with a linear gradient: 0–2 min, 2%–2% B; 2–5 min, 2%–9% B; 5–10 min, 9%–12% B; 10–16 min, 12%–12% B; 16–19 min, 12%–16%; 19–25 min, 16%–20% B; 25–30 min, 20%–25% B; 30–32 min, 25%–100% B; 32–34 min, 100%–100% B; 34–36 min, 100%–2% B; and 36–38 min, 2%–2% B. The injected sample volume was set at 5 μL.

UPLC was coupled to Q–TOF/MS equipped with electrospray ionisation in positive and negative ion modes. Ultra-high purity helium (He) was used as collision gas and high-purity nitrogen (N_2) was used as nebulising gas. The range of data acquisition was set from 50 Da to 1000 Da. Other operating parameters were listed as follows: capillary voltage, 3.0 kV; drying gas temperature, 325 ℃; desolvation gas flow rate, 600 L h^{-1}; and nebulising gas pressure, 350 psi. Leu–Enkephalin ion at *m*/*z* 556.2771 and 554.2615 were used to calibrate mass accuracy.

2.4 Data Analysis

After data were acquired, original data were obtained and processed by Markerlynx (Waters, UK) in Masslynx version 4.1 to detect and align the peaks. The parameters were set as follows: initial to final retention times of KDZ injection, 0 to 40; high and low mass, 1000 and 50, respectively; XIC window, 0.01 Da; and noise elimination level, 6.00. Data were processed and converted to an Excel format containing complete information of mass, retention time and peak area of the samples. Target compounds were obtained by processing the output data. The complete data were then screened using DF and NL information shown in Table 1 by following the procedure indicated in Figure. 2.

3 Results and discussion

3.1 Optimisation of UPLC-Q-TOF/MS conditions

The characteristics of the chemical compositions of TCM injections are used as basis to determine active ingredients and confirm the effect of drugs. In general, LC–UV detection is conducted to detect complex compounds in TCM;[22–23] however, the resulting fingerprint exhibits disadvantages, including low sensitivity, poor specificity and inaccurate structure information. In this study, UPLC-Q-TOF/MS was performed to obtain further material information of the fingerprint; the conditions of LC/MS were also optimised.

In this study, the physical structure and the chemical characteristics of the KDZ injection were integrated in four major categories (flavonoids, organic acids, amino acids and nucleotides) to acquire a wide range of material information; BEH C18 column was eventually used. Good resolution and high and narrow peaks were obtained at column temperature of 35 ℃, flow rate of 0.3 mL/min, injection volume of 7 μL and 0.05% added formic acid mobile phase. KDZ was comprehensively analysed and injected under positive and negative ion modes. The compounds in the KDZ injection exhibited peak behaviours in positive and negative ion modes in the total ion current (TIC); based on the mass behaviour described in previous studies, the two modes were used simultaneously. The typical TIC chromatograms of the substances in the KDZ injection under positive and negative ion modes are shown in Figure. 1.

3.2 Diagnostic fragment filtering and neutral loss filtering

Although KDZ injection is a single-herb preparation, chemical compositions are complex and diverse. In mass spectrometry collision-induction, compounds with similar or identical nucleus-skeletons generally show the same fracture behaviours and produce the same DFs; the strategy that uses DFs to screen compounds is called DFF. Therefore, this characteristic, which existed in the same category, was used to search for cleavage rules in MS. For example, quinine acids are characterised with DF at *m*/*z* 191 ($C_7H_{11}O_6^-$), which corresponds to [quinic acid–H]$^-$, and at *m*/*z* 173 ($C_7H_9O_5^-$), which corresponds to [quinic acid–H–H_2O]$^-$. The ion collection of quinine acids could be obtained after the two fragments were screened from the total data derived by Markerlynx and combined with the DFs of other subtypes to determine the type of compound. Neutral loss refers to the mass difference between level and tandem MS; substances substituted by a characterised group can be classified. In MS collision-induced dissociation, NLs often reveal the information related to the categories of compounds; the strategy that uses NLs to screen compounds is called NLF. Glycosides easily lose the neutral fragments of *m*/*z* 17 (NH_3), 43 (HNCO), 116 ($C_5H_8O_3$) and 132 ($C_5H_8O_4$); among these fragments, 43 (HNCO), 116 ($C_5H_8O_3$) and 132 ($C_5H_8O_4$) are characteristic NLs of nucleoside compounds. Thus, the complex components in KDZ injection can be rapidly and accurately classified and identified using a method in which DFF is combined with NLF. Based on the results of our study and those described in relevant and previous experiments, the rules of DFF and NLF regarding the four categories were established and summarised (Table 1). Indeed, substances in the KDZ injection were rapidly classified and identified using the proposed method in which DFF was combined with NLF (the procedure is indicated in Figure. 2).

3.2.1 Flavonoids

Based on chemical compositions, flavonoids in the KDZ injection could be divided into aglycone and glycoside. In MS, glycoside compounds were converted to aglycone as glycosyl was lost; thus, the characteristic fragments were analysed based on aglycone. After reviewed lots of relevant literature and reference experiment, we found different subclasses easily inform DFs with basic structure of aglycone. Although flavonoids easily produce the NL of glycosides, DF is better than NL in terms of specificity and operability. Therefore, DF was used to screen flavonoids. Flavonoids were classified on the basis of structure and divided into luteolin type (type A), apigenin type (type B), acacetin type (type C), quercetin type (type D) and isorhamnetin type (type E). The five subtypes exhibited the same structure of A ring at the aglycone part. C_1–C_3 of A ring easily underwent RDA reaction and produced the same fragment ion at *m*/*z* 153 $[C_7H_5O_4]^+$, which could be used as a DF of flavonoids. Types A–E contain different aglycone parts and cleav respectively to fragment ions at *m*/*z* 287 $[C_{15}H_{11}O_6]^+$, 271 $[C_{15}H_{11}O_5]^+$, 284 $[C_{16}H_{13}O_5]^+$, 303 $[C_{15}H_{11}O_7]^+$ and 315 $[C_{16}H_{12}O_7]^+$ in the positive ion mode.[24–29] In our study, data from these DF ions were integrated and applied to screen flavonoids and their subtypes. According to related literature and analysis results of standards in the positive ion mode, six DFs of *m*/*z* 153, 287, 271, 284, 303 and 315 were found and the specific information are shown in Table 1. In this study, the KDZ injection was analysed under LC–MS conditions, and the complete data were extracted by Makerlynx software. After the components were classified and identified, compounds 26–33 were obtained as flavonoids (the specific information are shown in Table 2). The specific cracking process for luteolin-7-*o*-glucoside is shown in Figure. 3.

Compound 27 exhibits a retention time of 14.41 and formula of $C_{21}H_{20}O_{11}$ (Table 2). In our experiment, fragment ions at *m*/*z* 449, 287, 153 and 135 were obtained. According to fragment rules, the fragment ion at *m*/*z* 153 $[C_7H_5O_4]^+$ was cleaved by C_1–C_3 of A ring via RDA–reaction; *m*/*z* 287 was the fragment ion at $[Y_0]^+$; thus, this compound can be inferred as a luteolin type. In addition, this compound presented $[M+H]^+$ ions with mass accuracy at *m*/*z* 449 and exhibited an NL of 162 Da between the parent ion (*m*/*z* 449) and the fragment ion (*m*/*z* 287); this result confirmed glycoside bond cleavage. The fragment ion at *m*/*z* 135 displayed a loss of 18 Da (H_2O) based on the fragment ion at *m*/*z* 153 and compared with that described in literature;[28] thus, compound 27 was inferred as luteolin-7-*o*-glucoside.

3.2.2 Organic acids

According to their chemical compositions, the organic acids in the KDZ injection could be classified as simple organic acids and acylated organic acids. Based on related literature and analysis of standards, simple organic compounds, which refer to unacylated single carboxylic acids, were found easily loss units of H_2O and CO_2 under the negative ion mode; thus, the substance could be identified according to its corresponding parent ion $[M+H]^+$ and NLs. Similar to flavonoids, acylated organic acids exhibited specific DFs, which favour screening and identification of the substances in the KDZ injection. Based on their structure, acylated organic acids were divided into six types: caffeoyltartaric acid (CTA, type A), dicaffeoyltartaric acid (DTA, type B), caffeoylquinic acid (CQA, type C), dicaffeoylquinic acid (DiCQA, type D), ferulylquinic acid (FQA, type E) and cinnamoylquinic acid (CiQA, type F).

Types A–B compounds contained the same unit of $[\text{tartaric acid}]^-$; thus, the fragment ion ($C_4H_5O_6^-$) at *m*/*z* 149 could be determined as their characteristic DF. For types C–F compounds, the common DFs were determined at *m*/*z* 191 ($C_7H_{11}O_6^-$) corresponding to $[\text{quinic acid–H}]^-$, and *m*/*z* 173 ($C_7H_9O_5^-$) corresponding to $[\text{quinic acid–H–H}_2\text{O}]^-$; these DFs could be used for preliminary screening. Types A–B compounds could be diagnosed as type A if the compound contained a fragment ion at *m*/*z* 311 only; with fragment ions at *m*/*z* 311 and 473 simultaneously, the compound could be confirmed as type B. For types C–D compounds, *m*/*z* 179 ($C_9H_7O_4^-$) corresponding to $[\text{caffeic acid–H}]^-$, and *m*/*z* 135 ($C_8H_7O_2^-$) corresponding to $[\text{caffeic acid–H–CO}_2]^-$, were determined as their additional DFs; the compound could be diagnosed as type C if this compound contained a fragment ion at *m*/*z* 353 ($C_{16}H_{17}O_9^-$) corresponding to $[\text{CQA–H}]^-$; the compound could be diagnosed as type D when this compound contained fragment ions at *m*/*z* 353 ($C_{16}H_{17}O_9^-$) corresponding to $[\text{CQA–H}]^-$, and *m*/*z* 515 ($C_{25}H_{23}O_{12}^-$) corresponding to $[\text{DiCQA–H}]^-$. Furthermore, this compound could be inferred as type E when this compound contained the fragment ions at *m*/*z* 193 ($C_{10}H_9O_4^-$) corresponding to $[\text{ferulic acid–H}]^-$, and *m*/*z* 367 ($C_{17}H_{19}O_9^-$). In the presence of *m*/*z* 163 ($C_9H_7O_3^-$) corresponding to $[\text{cinnamic acid}]^-$, and *m*/*z* 321 ($C_{16}H_{17}O_7^-$), the compound could be confirmed as type F. According to related literature and analysis results of standards in the negative ion mode, seven fragment ions at *m*/*z* 191, 173, 179, 135, 193, 163 and 149 were obtained,[17, 28–29] and the specific details are shown in Table 1. After classification and identification were performed, compounds 10–25 were obtained as organic acids (the specific details are shown in Table 2). The specific cracking process for 3-CQA is shown in Figure. 4.

Compounds 15–17 with retention times of 4.63–6.36 min and formula of $C_{16}H_{18}O_9$ showed fragment ions at *m*/*z* 353, 191, 179, 173and 135. According to the fragment rules, the fragment ions at *m*/*z* 191 for $[\text{quinic acid–H}]^-$, 173 for $[\text{quinic acid–H–H}_2\text{O}]^-$, 179 for $[\text{caffeic acid–H}]^-$, 135 for $[\text{caffeic acid–H–CO}_2]^-$ and 353 were their parent ions; thus, these compounds were inferred as type C, namely, CQA. Compounds 15–17 contained the same parent ion at *m*/*z* 353 and fragment ions at *m*/*z* 191, 179 and 135. Compounds 15-17 are difficult to identify based on character fragments; thus, the precise structure should be determined on the basis of the relative abundance of fragments. The base peak of 3-CQA and 5-CQA is generally the fragment ion at *m*/*z* 191; 3-CQA is larger than 5-CQA at *m*/*z* 179 fragment ion. Conversely, 4-CQA exhibits a base peak at *m*/*z* 173, and this finding is consistent with that in a previous study.[17] Therefore, compounds 15–17 can be inferred as 3-CQA, 5-CQA and 4-CQA, respectively.

3.2.3 Amino acids

Amino acids contain carboxyl and amino group; carboxyl and amino group can be easily cleaved in MS. On the basis of related literature and analysis of standards in MS, we found that amino acids can combine with protons and then form NLs of H_2O, HCOOH and NH_3 in the positive ion mode and the NLs of CO_2 and NH_3 in the negative ion mode (the specific details are shown in Table 1).[30] Thus, this characteristic of these compounds was considered in MS to classify and identify the amino acids. Compounds 1–5 were obtained as amino acids (the specific details are shown in Table 2).

Compound 3 with a retention time of 0.65 min and formula of $C_5H_9NO_2$ showed fragment ions at *m*/*z* 116, 98, 82 and 70. Compared with the precursor ion at *m*/*z* 116, the fragment ion at *m*/*z* 98 exhibited an NL of 18 Da for a H_2O molecule. The fragment ion at *m*/*z* 81 exhibited continuous losses of H_2O and NH_3 based on the parent ion and the fragment ion at *m*/*z* 70 showed a loss of 46 Da for a HCOOH molecule based on the parent ion at *m*/*z* 116. According to the NL rules, compound 3

can be inferred as an amino acid. In addition, the fragment of *m/z* 116 was a molecular ion $[M+H]^+$; compared with the fragment ions described in previous studies, compound 27 can be inferred as proline.[30]

3.2.4 Nucleosides

Nucleosides can be classified as aglycone and glycosides from their compositions. Based on references and standards in MS, aglycone of nucleoside exhibits NLs of *m/z* 17 (NH_3) and 43 (HNCO); glycosides are generally connected with ribose and cleaved in MS to form the specific NLs of *m/z* 116 ($C_5H_8O_3$) and 132 ($C_5H_8O_4$). However, flavonoids also produce neutral glycosyl debris. The glycosyl attached to flavonoids is a six-carbon glycoside and the glycosyl attached to nucleoside is a five-carbon glycoside. This finding showed that the NLs of the nucleoside components is specific for KDZ components. Otherwise, these substances do not display similar skeleton; therefore, NLF was used to screen nucleosides. Based on relevant literature and standard experiments in the positive ion mode, NLs of *m/z* 17 (NH_3), 43 (HNCO), 116 ($C_5H_8O_3$) and 132 ($C_5H_8O_4$) were obtained,[31–36] and the specific details are shown in Table 1. After classification and identification were performed, compounds 6–9 were obtained and identified as nucleosides (the specific details are shown in Table 2).

Compound 8 with a retention time of 1.44 min and formula of $C_{10}H_{13}N_5O_5$ showed fragment ions at *m/z* 268, 136, 119 and 93. The NL of 132 Da between the fragment ions *m/z* 268 and 136 confirmed the loss of a rib. Compared with the fragment ion at *m/z* 136, the fragment ion at *m/z* 119 exhibited an NL of 17 Da (NH_3); the fragment ion at *m/z* 93 displayed a loss of 43 Da for a HNCO molecule based on the fragment ion at *m/z* 136. According to NL rules, compound 8 can be inferred as a nucleoside. In addition, the fragment ion at *m/z* 268 was a molecular ion $[M+H]^+$; compared with the fragment ions described in previous studies,[31–35] compound 8 can be confirmed as adenosine.

In this study, after we integrated the fragment information and summarized the law from standards and existing references. These details were combined with data processing technologies DFF and NLF to construct a method that could be used to rapidly classify and identify 14 subclasses of 4 categories in the KDZ injection. Based on the UPLC-Q-TOF/MS platform, 31 kinds of compounds belonging to four categories were found: compounds 1–5 are amino acids, compounds 6–9 are nucleosides, compounds 10–23 are organic acids and compounds 24–31 are flavonoids. Compared with general data analytical methods, the designed method could be used to identify and classify the components during analysis, analyse active substances in the KDZ injection and provide guidance to evaluate medicinal components. Moreover, this study effectively combined DFF and NLF based on initial classification; thus, material information of different types was combined. With this method, unknown ingredients of TCM can be determined. Furthermore, complex workload in data processing can be reduced, impure ions can be removed effectively and compound identification procedures are simplified and clarified with this method. DFF and NLF also help classify and identify the KDZ components; thus, unknown compounds can be discovered. The proposed method also provides a basis to classify and identify other components in TCM.

4 Conclusions

In this study, the information of known substances in KDZ injection was integrated, and the rules of DFs and NLs were summarised. Using the UPLC-Q-TOF/MS platform, we established a method in which DFF was integrated with NLF; this method could be applied to rapidly screen and identify substances in KDZ injection. To a certain extent, this study solved technical problems related to parsing, classification and identification caused by the complex compositions of TCM injections; this study also contributed to the development of rapid classification and identification methods for TCM. Thus, TCM injections can be possibly applied in clinical practice. In addition, this study provided a new method to screen, classify and identify target components from complex samples in molecular biology and pharmacokinetics.

Acknowledgements

This project was supported by the National Science and Technology Major Project Subject (2011ZX09401-305-42) .

Notes and references

[a] *Tianjin State Key Laboratory of Modern Chinese Medicine, School of Traditional Chinese Materia Medica, Tianjin University of Traditional Chinese Medicine, 312 Anshan west Road, Tianjin 300193, China.*
Tel and Fax number: +86-22-59596223.
E-mail: tianjin_tcm001@sina.com, yuboli1@163.com

1. Nanjing University of Chinese Medicine, encyclopedia of Chinese Materia Medical, *Shanghai Science and Technology Press*, 2006: 1783 -1785.
2. L. J. Miao, N. Han and Z. H. Liu. *Asian Journal of Traditional Medicines*, 2011, **6,** 179-187.
3. X. Z. Feng, S. X. Xu and M. Dong, *J Chin Pharm Sci*, 2000, **9,** 134
4. J. Y. Ma, Z. T. Wang, L. S. Xu. *Phytochemistry*, 1998, **48,** 201
5. V. Tak, A. Purohit, D. Pardasani, D. Raghavender-Goud, R. Jain and D. K. Dubey, *Chromatography A*, 2014, **1370**, 80-92
6. M. H. Liu, X. Tong, J. X. Wang, W. Zhou, H. Gao and W. W. Su, *Pharmaceutical and Biomedical Analysis*, 2013, 74, 141-155
7. X. X. Zhang, J. R. Liang, Y. Zhao, J. Gao, W. J. Sun and Y. C. R. Ito, *Pharmaceutical and Biomedical Analysis*, 2014, **91,** 46-59
8. C. Huang, Q. M. Xu, C. Chen, C. W. Song, Y. Xu, Y. Xiang, Y. L .Feng, H. Ou-Yang, Y. Zhang and H. L. Jiang, *Chromatography A*, 2014, **1361,** 139-152
9. A. H. Zhang, D. Zou, G. L. Yan, Y. L. Tan, H. Sun and X. J. Wang, *J. Sep. Sci*, 2014, **37,** 1742–1747
10. T. Xie, Y. Liang, Jiye. A, H. P. Hao, L. S. Liu, X. Zheng, C. Dai, Y. Y. Zhou, T. Y. Guan, Y. N. Liu, L. Xie and G. J. Wang, *Chromatography B*, 2012, **905,** 43-53
11. M. Sandin, J. Teleman, J. Malmström and F. Levander, *Biophysica Acta*, 2014, **1844,** 29-41
12. C. Christin, R. Bischoff, P. Horvatovich, *Talanta*, 2011, **83,** 1209-1224
13. J. Guo, M. L. Zhang, C. S. Elmore and K. Vishwanathan, *Analytica Chimica Acta*, 2013, **780,** 55-64
14. K. M. Park, Y. J. Bae, J. H. Moon and M. S. Kim, *International Journal of Mass Spectrometry*, 2013, **352,** 70-76
15. L. W. Qi, H, Y. Wang, H. Zhang, C. Z. Wang, P. Li, C. S. Yuan, *Chromatography A*, 2012, **1230,** 93-99
16. J. Y. Zhang, Z. J. Wang, Q. Zhang, F. Wang, Q. Ma, Z. Z. Lin, J. Q. Lu and Y. J. Qiao, *Talanta*, 2014, **124,** 111-122
17. J.Y. Zhang, Q. Zhang, N. Li, Z. J. Wang, J. Q. Lu and Y. J. Qiao, *Talanta*, 2013,**104,** 1–9
18. A. Guerrero and C. B. Lebrilla, *International Journal of Mass Spectrometry*, 2013, **354-355,** 19-25
19. A. H.Zhang, H. Sun, G. L. Yan, P. Wang, Y. Han and X. J. Wang, *Cancer Letters*, 2014, **345,** 17-20
20. S. Y. Lee, Z. Y. Li, S. J. Valentine, S. M. Zucker, N. Webber, J. P. Reilly and D. E. Clemmer, *International Journal of Mass Spectrometry*, 2012, **309,** 154-160
21. J Qu, Q. L. Liang, G. A. Luo, and Y. M. Wang, *Anal. Chem*, 2004, **76,** 2239-2247

22. M. P. La, F. Zhang, S. H. Gao, X. W. Liu, Z. J. Wu, L. N. Sun, X. Tao and W. S. Chen, *Pharmaceutical and Biomedical Analysis*, 2015, **102,** 366-376
24. G. Rak, P. Fodor and L. Abrankó, *International Journal of Mass Spectrometry*, 2010, **290,** 32-38
25. M. Zehl, C. Braunberger, J. Conrad, M. Crnogorac, S. Krasteva, B. Vogler, U. Beifuss and L. Krenn, *Anal Bioanal Chem*, 2011, **400,** 2565-2576
26. A. L. S. Dias, E. Rozet, Y Larondelle, P. Hubert, H. Rogez and J. Quetin-Leclercq, *Anal Bioanal Chem*, 2013, **405,** 9235-9249
27. J. He, Y. Feng, H. Z. Ou-Yang, B. Yu, Y. X. Chang, G. X. Pan, G. Y. Dong, T. Wang, and X. M. Gao, *Pharmaceutical and Biomedical Analysis*, 2013, **84,** 189-195
28. S. Gouveia and P. C. Castilho, *Food Chemistry*, 2011, **129,** 333-344
29. H. J. An, H. Wang, Y. X. Lan, Y. K. Hashi, S. Z. Chen, *Pharmaceutical and Biomedical Analysis*, 2013, **85,** 295-304
30. I. Kıvrak, S. Kıvrak and M. Harmandar, *Food Chemistry*, 2014, **158,** 88-92
23. R. Canavesi, S. Aprile, E. Varese and G. Grosa, *Pharmaceutical and Biomedical Analysis*, 2014, **97,** 141-150
31. S. Guo, J. A. Duan, D. W. Qian, H. Q. Wang, Y. P. Tang, Y. F. Qian,D. W. Wu, S. L. Su and E. X. Shang, *Chromatography A*, 2013, **1301,** 147-155
32. R. Liu, J. Ji, L. C. Wang, S. Y. Chen, S. Guo and H. Wu, *Food Chemistry*, 2012, **135,** 548-554
33. F. T. Li, J. A. Duan, D. W. Qian, S. Guo, Y. H. Ding, X. H. Liu, Y. F. Qian, Y. R. Peng, Y. J. Ren and Y. Chen, *Pharmaceutical and Biomedical Analysis*, 2013, **83,** 10-18
34. W. Struck-Lewicka, R. Kaliszan and M. J. Markuszewski, *Pharmaceutical and Biomedical Analysis*, 2014, **xxx,** xxx–xxx
35. X. M. Wang, Y. Y.Xie, X. M. Hu, Y. K. Li, P. Hu, Y. M. Wang, Q. L. Liang and G. A. Luo, *Acta Pharmaceutica Sinica B*, 2013, **3,** 337-344
36. Y. P. Xie, J. Tian, P. Gao, G. W. Xu, X. Fu and Y. Wang, *Chin J Anal Chem*, 2013, **41,** 36–41.

Table 1. Diagnostic fragment and neutral loss information of chemical substances in KuDieZi injection

Compound classification	Subclass	Ion model	Diagnosed fragments	Neutral loss
Flavonoids		$[M+H]^+$	287$[C_{15}H_{11}O_6]^+$, 303$[C_{15}H_{11}O_7]^+$, 153$[C_7H_5O_4]^+$, 165$[C_8H_5O_4]^+$, 181$[C_8H_8O_5]^+$, 271$[C_{15}H_{11}O_5]^+$, 284$[C_{16}H_{13}O_5]^+$, 315$[C_{16}H_{12}O_7]^+$	
Organic acids	Simple	$[M-H]^-$		18$[H_2O]$, 44$[CO_2]$
	Acyl acid	$[M-H]^-$	191$[C_7H_{11}O_6]^-$,173$[C_7H_9O_5]^-$,179$[C_9H_7O_4]^-$,135$[C_8H_7O_2]^-$, 149$[C_4H_5O_6]^-$,163$[C_9H_7O_3]^-$,193$[C_{10}H_9O_4]^-$	
Nucleic acids	Bases	$[M+H]^+$		17$[NH_3]$, 43[HNCO]
	Nucleoside	$[M+H]^+$		17$[NH_3]$, 132$[C_5H_8O_4]$, 43[HNCO], 116$[C_5H_8O_3]$
Amino acids		$[M+H]^+$		17$[NH_3]$,18$[H_2O]$,46[HCOOH]
		$[M-H]^-$		17$[NH_3]$, 44$[CO_2]$

Table 2. Identification of the chemical constituents of KuDieZi injection by using UPLC–Q-TOF/MS in positive and negative ion mode.

	RT	m/z	Theoretical mass	Formula	Fragment ions	Chemical name
1	0.58	104.0346	105.0426	$C_3H_7NO_3$	104(42%),87(100)	Serine
2	0.59	146.0451	147.0531	$C_5H_9NO_4$	146(43%),129(52),102(3.2),	Glutamate
3	0.65	116.0710	115.0633	$C_5H_9NO_2$	116(100%),98(0.2),70(4.5)	Proline
4	0.81	118.0864	117.0790	$C_5H_{11}NO_2$	118(100%), 141(4.0), 101(0.2),72(22)	Valine
5	1.38	132.1023	131.0946	$C_6H_{13}NO_2$	132(100%), 115(0.1), 86(30)	Leucine
6	1.07	243.0613	244.0695	$C_9H_{12}N_2O_6$	243(65%), 244(5.8), 200(100),	Uridine
7	1.13	113.0344	112.0273	$C_4H_4N_2O_2$	113(100%), 96(11), 70(19)	Uracil
8	1.44	268.1043	267.0968	$C_{10}H_{13}N_5O_4$	268(100%), 136(40), 119(1.6), 93(0.1)	Adenosine
9	1.61	284.0993	283.0917	$C_{10}H_{13}N_5O_5$	284(19%), 152(100), 135(3.0)	Guanosine
10	0.94	111.0093	112.0160	$C_5H_4O_3$	111(100%), 93(0.1)	1- furancarboxylato
11	3.27	153.0187	154.0266	$C_7H_6O_4$	153(100%), 109(51)	3,4-dihydroxybenzoic acid
12	4.33	177.0560	178.0630	$C_{10}H_{10}O_3$	177(63%), 149(76), 133(100), 105(1.9)	4-methoxy cinnamic acid
13	4.34	311.0397	312.0481	$C_{13}H_{12}O_9$	311(22%), 179(100), 149(58), 135(2.4)	Caffeoyltartaric acid
14	4.64	179.0342	180.0423	$C_9H_8O_4$	179(61%), 135(100)	Caffeic acid
15	4.63	353.0874	354.0951	$C_{16}H_{18}O_9$	353(100%), 191(41), 179(26), 173(1.7),135(43)	3-Caffeoylquinic acid
16	5.94	353.0870	354.0951	$C_{16}H_{18}O_9$	353(38%), 191(100), 173(0.1)	5-Caffeoylquinic acid
17	6.36	353.0871	354.0951	$C_{16}H_{18}O_9$	353(100%), 179(28), 173(37)	4-Caffeoylquinic acid
18	6.03	167.0342	168.0423	$C_8H_8O_4$	167(100%), 121(48)	Vanilloid
19	6.34	367.1018	368.1107	$C_{17}H_{20}O_9$	367(1%), 193(100), 173(54)	Ferulylquinic acid
20	5.69	193.0503	194.0579	$C_{10}H_{10}O_4$	193(100%),178(16)	Ferulic acid
21	6.85	177.0555	178.0630	$C_{10}H_{10}O_3$	177(%), 149(0.2)	2-methoxy cinnamic acid
22	10.84	473.0725	474.0798	$C_{22}H_{18}O_2$	473(2.6%), 311(43), 179(35), 135(1.8)	Chicory acid
23	4.34	179.0341	180.0423	$C_9H_8O_4$	179(100%), 177(49), 135(2.4), 133(77)	2,5-hydroxy cinnamic acid
24	10.31	303.0495	302.0427	$C_{15}H_{10}O_7$	303(100%)	Quercetin
25	11.12	611.1611	610.1534	$C_{27}H_{30}O_{16}$	611(100%), 304(0.1),153(2.1)	Rutin
26	14.02	463.0870	462.0798	$C_{21}H_{18}O_{12}$	463(100%), 288(7.3), 287(57), 153(7.8), 135(2.6)	Luteolin-7-*o*-glucuronide
27	14.41	449.1076	448.1006	$C_{21}H_{20}O_{11}$	449(100%), 287(31), 153(9.7), 135(2.9)	Luteolin-7-*o*-glucoside
28	14.61	287.0553	286.0477	$C_{15}H_{10}O_6$	287(100%), 153(32), 135(9.5)	Luteolin
29	18.87	433.1119	432.1056	$C_{21}H_{20}O_{10}$	433(0.4%), 272(4.5), 271(100)	Apigenin-7-*o*-glucoside
30	19.41	447.0925	446.0849	$C_{21}H_{18}O_{11}$	447(100%), 271(58), 153(12)	Apigenin-7-*o*- glucuronide
31	20.07	271.0597	270.0528	$C_{15}H_{10}O_5$	271(6.9%), 153(45), 135(100)	Apigenin

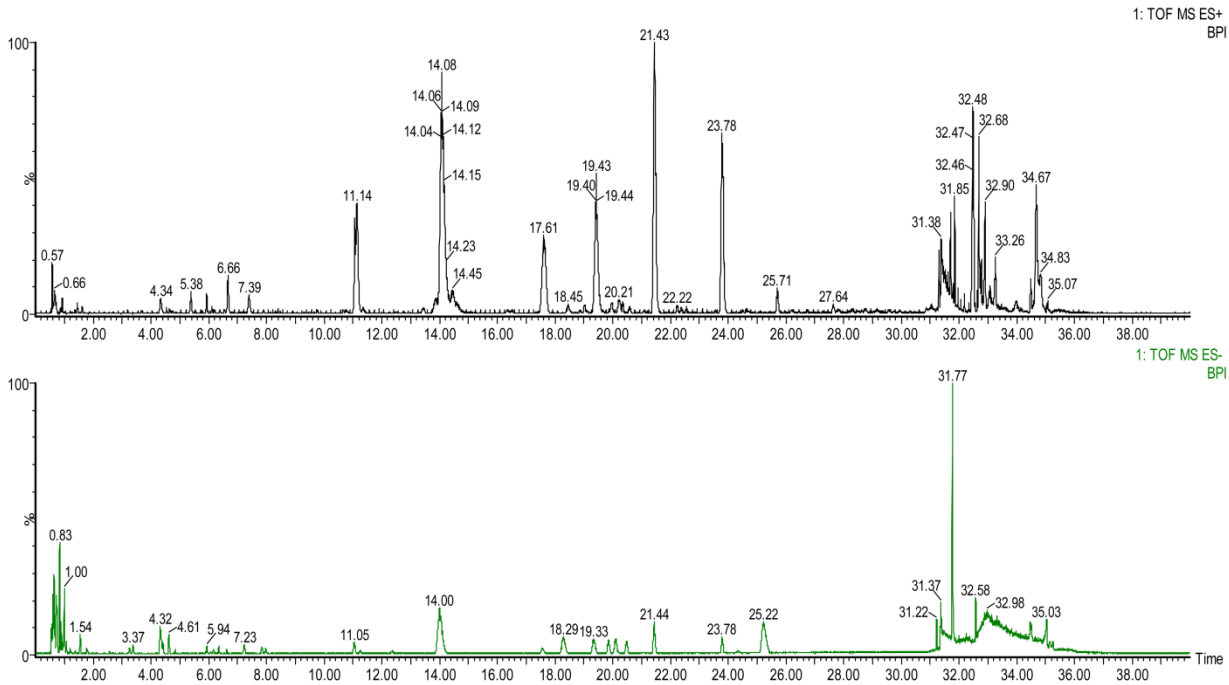


Figure.1 Typical total ion current (TIC) chromatograms of substances in KuDieZi injection, under positive and negative ion mode.

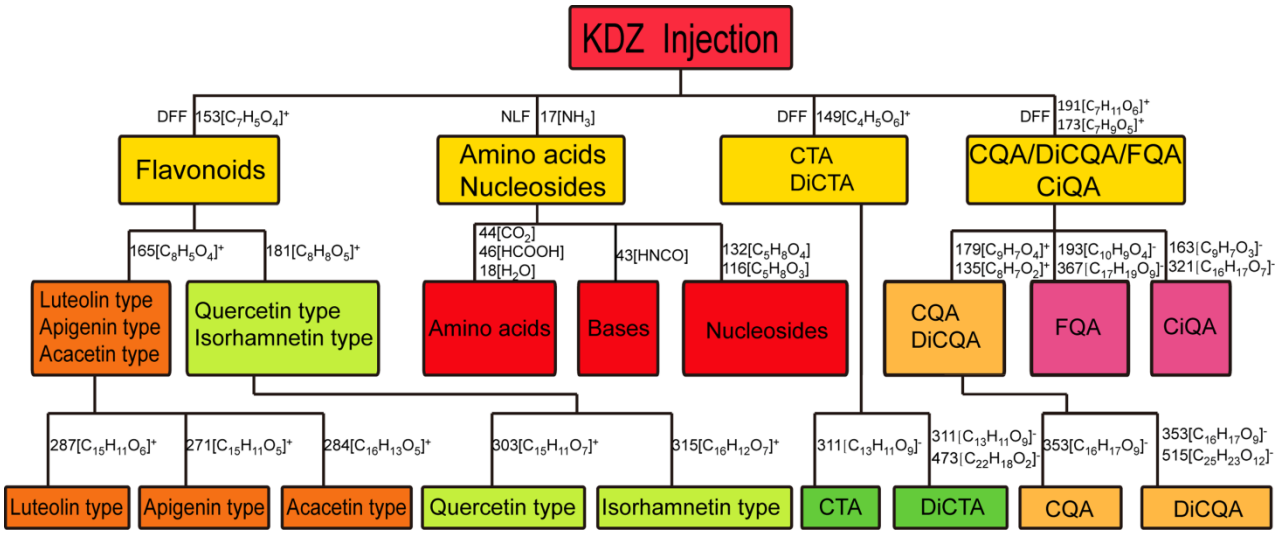


Figure.2 The screening process for the chemical substances of each subclass in KuDieZi (KDZ) injection. DFF: diagnostic fragments filter; NLF: neutral loss filter; CTA: Caffeoyltartaric acid; DiTA: Dicaffeoyltartaric acid; CQA: Caffeoylquinic acid; DiCQA: Dicaffeoylquinic acid; FQA: Ferulylquinic acid; CiQA: Cinnamoylquinic acid.

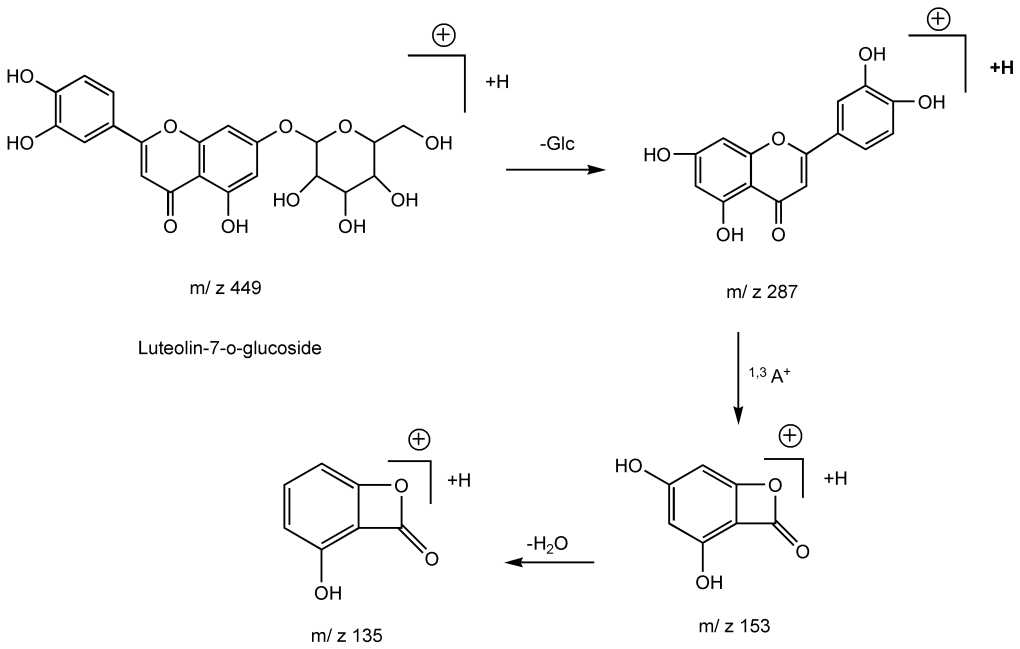


Figure.3 The proposed fragmentation pathways and common Diagnose Fragment Ions for flavonoids, taking Luteolin-7-o-glucoside for example. The ion at m/z 449 was molecular ion $[M+H]^+$, m/z 287 was the fragment ion of $[Y_0]^+$, the fragment ion of m/z 153 $[C_7H_5O_4]^+$ was cleaved by C1-C3 of A ring via RDA-reaction, while, the fragment ion (m/z 135) had a loss of 18Da(H_2O) on the basis of the fragment ion (m/z 153) .

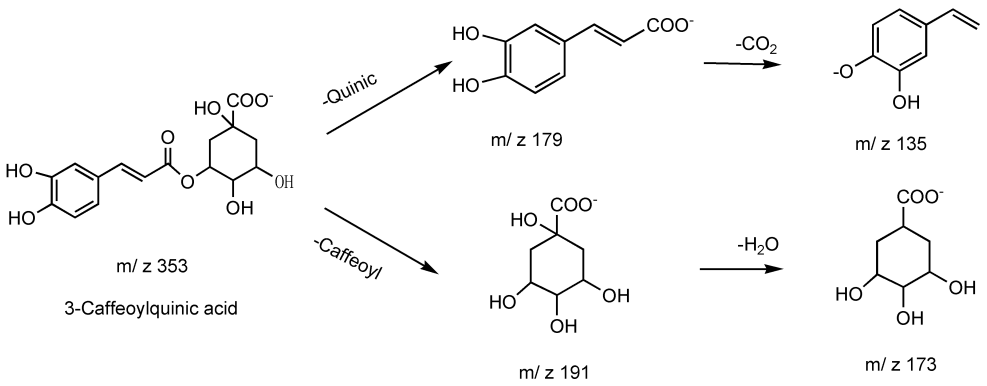


Figure.4 The structure and main way of fragmentation of Caffeoylquinic acid, as well as 3-Caffeoylquinic acid. The fragment ion at m/z 353 was parent ion, m/z 191 corresponding to [Quinic acid-H]$^-$, m/z 173 corresponding to [Quinic acid-H-H_2O]$^-$. Meanwhile, the fragment ion at m/z 179 corresponding to [Caffeic acid-H]$^-$, m/z 135 corresponding to [Caffeic acid-H-CO_2]$^-$.