Journal of Pharmaceutical and Biomedical Analysis 53 (2010) 279–286

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

# A novel approach to rapidly explore analytical markers for quality control of Radix Salviae Miltiorrhizae extract granules by robust principal component analysis with ultra-high performance liquid chromatography–ultraviolet–quadrupole time-of-flight mass spectrometry

Jing-Zheng Song[a], Song-Lin Li[a], Yan Zhou[a], Chun-Feng Qiao[a], Shi-Lin Chen[b], Hong-Xi Xu[a,*]

[a] Chinese Medicine Laboratory, Hong Kong Jockey Club Institute of Chinese Medicine, Shatin, N. T., Hong Kong, China
[b] Institute of Medicinal Plant Development, Chinese Academy of Medicinal Sciences, Beijing, China

ARTICLE INFO

*Article history:*
Received 22 December 2009
Received in revised form 18 March 2010
Accepted 20 March 2010
Available online 27 March 2010

*Keywords:*
Chemical marker
Robust principal component analysis
Radix Salviae Miltiorrhizae
UHPLC
Q-Tof
Mass spectrometry

ABSTRACT

In a well-controlled experiment, outliers discriminated by robust principal component analysis (RPCA) represent contents in samples which are of particular quality distinguishable from the rest of the others, therefore chemical constituents in a natural product causing discrimination between outliers and the majority of samples could be considered as analytical markers for quality control. Based on this strategy, a novel approach for rapidly exploring characteristic analytical markers was proposed for the quality control of extract granules of Radix Salviae Miltiorrhizae (EGRSM). In this study, large sizes of samples were analyzed via high-throughput ultra-high performance liquid chromatography–ultraviolet–quadrupole time-of-flight mass spectrometry (UHPLC–UV–Q-Tof MS). RPCA was first performed on the three groups of samples: RSM (the raw material), the in-house prepared aqueous extract of Radix Salviae Miltiorrhizae (AERSM) and commercial product of EGRSM, to determine the variation of specific constituents between raw material and the final products as well as the effect of manufacturing process on the overall quality. Then RPCA was performed on the commercial products of EGRSM to explore the applicability of identified characteristic markers for the quality control of EGRSM. Candidate markers were extracted by RPCA, and their molecular formulae were determined by high resolution electrospray ionization-mass spectrometric (ESI-MS) analysis. The suitability of identified markers was then evaluated by determining the relationship between quantities of the identified markers with their antioxidant activities biologically, and further confirmed in a variety of samples. In conclusion, the combination of RPCA with UHPLC–UV–Q-Tof MS is a reliable means to identify chemical markers for evaluating quality of herbal medicines.



## 1. Introduction

Chemical markers are frequently used for assuring quality consistency of natural products derived from botanical sources. An ideal chemical marker for a natural product should be not only a characteristic constituent but also the therapeutic constituent. In many cases, however, it is not well known which components in a botanical product are responsible for these activities. As suggested by the European Medicines Agency (EMEA), four categories of constituents or groups of constituents were defined as chemical markers including principle (constituents that have known clinical activities), active (constituents that have some known pharmacological activities), negative (constituents that may have allergenic or toxic properties) and analytical (constituents that are chosen as markers for identification and quantitative determination) markers [1]. Recently, we have suggested a new classification of chemical markers into eight categories including therapeutic, bioactive, characteristic, main, synergistic, correlative, toxic and general components [2]. Irrespective of the definitions, selection of chemical markers for a particular product is usually affected by subjective assessments, empirical evidences or the commercial availabilities of reference standards, and generally one or two constituents with the highest contents were considered and selected as markers. Since the natural products consist of complex mixtures of phytochemical constituents, a mere of chemical markers will not be sufficient and specific enough for quality control, quality assurance and stability assessment of a particular natural product. Moreover, because of losses due to chemical conversion or degradation, markers that are specific for particular herbs may not be suitable for the quality control of herbal preparations, such as aqueous extract

* Corresponding author. Tel.: +852 34062875; fax: +852 35517333.
*E-mail address:* xuhongxi@hkjcicm.org (H.-X. Xu).

0731-7085/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2010.03.022

granules of herbal medicines. Therefore, the determining criteria and strategies to select adequate chemical markers for the quality control of aqueous extract granules of herbal medicines would be required.

Aqueous extract granules of herbal medicines, a novel form of processed medicinal herb in which a single herb (or several herbs) is (are) extracted with hot water and then dried by spray drying or fluidized bed drying techniques, has been developed and marketed in China, Southeast Asia and North America. The advantages for such extract granules are ease for storage and handling, and convenient for oral administration with minimized consumption volume (compared to water decoction or raw materials). However, quality of the aqueous extract granules is being affected by numerous factors, such as the quality of raw herbal materials, differences in manufacturing processes and chemical stabilities of various constituents. The most critical issue is that no official method or standard for quality control of extract granules is currently available making this task very challenging.

Radix Salviae Miltiorrhizae (RSM, Chinese name Danshen), the dried root of *Salviae Miltiorrhizae* Bunge, is a commonly used natural medicine for treating coronary heart disease, cerebrovascular disease, hepatitis, hepatocirrhosis, chronic renal failure, dysmenorrheal and neurosasthenic insomnia in China and Southeast Asia [3–5]. High quality aqueous extract granule of Radix Salviae Miltiorrhizae (EGRSM) should be a rich source of phenolic compounds with strong antioxidant activity, and it is also used as dietary supplement in healthy food products. Up till now, the major constituents such as salvianolic acid B and tanshinone IIA were selected as markers for the quality control of RSM according to the official quality standard [6]. Despite of the debate that the quality of RSM products could not be well monitored using only a few markers, a combination of several chemical constituents as markers for quality assessment of RSM and its products based on various analytical methods such as HPLC–UV, LC–MS and capillary electrophoresis methods were reported [7–12], where the reasons why those constituents were chosen have not been clearly explained. With multiple influencing factors such as quality of RSM, the manufacturing processes and the instabilities of phenolic components in RSM, the resulting chemical compositions of EGRSM are complex leading to difficulty in ensuring the quality consistency of EGRSM, and in determining the criteria for selecting suitable markers for quality control of EGRSM.

From statistical point of view, outliers are observations which deviate markedly from other members of the sample set. Outliers of the dataset can be attributed to either experimental errors or their unique characteristics making them recognizably (statistically) different from the designated sample population. In a well-controlled experiment, the outliers could mainly be attributed to the unique features of that particular dataset rather than experimental errors. Therefore, once the ways for unraveling "unique features" are identified, they can be a useful approach to select chemical markers for the quality control of natural medicines.

Principal component analysis (PCA) is a very popular data exploration and reduction technique for extracting the most important information from the large, and confusing data sets. Since its algorithm is based on the sample mean data and the sample covariance matrix of the data with the assumption that all data have not been spoiled by outliers; once the dataset contain outliers, it will be significantly skewed to the outliers. Thus, the classical PCA may wrongly delineate the shape of the majority of the data [13]. On contrary to the classical PCA, robust principal components analysis (RPCA) utilizes the median instead the mean of data, therefore, it can minimize the negative impact of outliers to obtain reliable evaluation, and can be used to generate a diagnostic plot to locate and classify the outliers [13,14].

For the quality control of EGRSM, high quality EGRSM should have a constituent profile which falls into an acceptable, well-defined and narrow range of contents, which will then exert biological activities in a consistent and reproducible manner. A constituent in a sample does not meet the requirement of quality standard can be treated as an outlier. Thus, the chemical constituents causing significant quality variations between failed samples (outliners) and passed samples, should be considered as characteristic analytical markers reflecting the quality of samples. Based on this strategy, a novel approach for rapid and appropriate analytical markers identification was proposed for the quality control of natural medicines: firstly, large amount of samples were analyzed through high-throughput ultra-high performance liquid chromatography ultraviolet quadruple time-of-flight mass spectrometer (UHPLC–UV–Q-Tof MS); then robust principal components analysis (RPCA) was performed to differentiate outliers from the majority of samples. Finally, characteristic analytical markers were identified from the loadings of RPCA.

In this study, EGRSM was used as an example to demonstrate the applicability of this strategy to identify characteristic analytical markers for overall quality assessment of EGRSM. RPCA was first performed on the three groups of samples: RSM (the raw material), the in-house prepared aqueous extract of Radix Salviae Miltiorrhizae (AERSM) and commercial product of EGRSM, to understand the differences in specific constituents between raw material and in the final products, as well as to delineate these variations during manufacturing process. Then, RPCA was performed on the commercial products of EGRSM to explore the characteristic markers for the quality control of EGRSM. Candidate markers were extracted by RPCA, and their molecular formulae were identified by high resolution electrospray ionization-mass spectrometric (ESI-MS) analysis. The applicability of identified markers was then evaluated by determining the relationship between their contents and antioxidant activities, and then further confirmed in various samples.

## 2. Materials and methods

### 2.1. Materials

Thirty-nine samples of RSM were purchased from Chinese medicine retailers in Shenzhen, Guangzhou, Shanghai, Nanjing, Zhengzhou and Hong Kong of China. Thirty-eight batches of commercial EGRSM produced by 13 manufacturers in mainland China, Hong Kong and Taiwan were purchased from their distributors in Hong Kong or directly from the manufacturers.

### 2.2. Chemicals and reagents

Acetonitrile of HPLC and LC–MS grades were purchased from Tedia (OH, USA) and Fisher Scientific (Loughborough, UK), respectively. Methanol and ethanol of analytical grade were bought from Tedia (OH, USA). Formic acid of MS grade was from Fluka (Steinheim, Germany). The 2,4,6-tri(2-pyridyl)-s-triazine (TPTZ, ≥99.0%, TLC) was purchased from Fluka (Sigma–Aldrich, Steinheim, Switzerland). The 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox, 97%) was purchased from Aldrich (Sigma–Aldrich, St. Louis, USA). Reference standards of danshensu (**1**), protocatechuic aldehyde (**2**), rosmarinic acid (**3**), lithospermic acid (**5**), salvianolic acid B (**6**), salvianolic acid A (**8**), dihydrotanshinone I (**10**), cryptotanshinone (**11**), tanshinone I (**12**) and tanshinone IIA (**13**) were provided by Hong Kong Jockey Club Institute of Chinese Medicine (Shatin, Hong Kong, China). The purities of all these reference standards were ≥98.0% (HPLC), their chemical structures were shown in Fig. 1. Ultra-pure water was produced by Milli-Q system (Millipore, Bedford, USA) and other chemicals were of analytical grades.

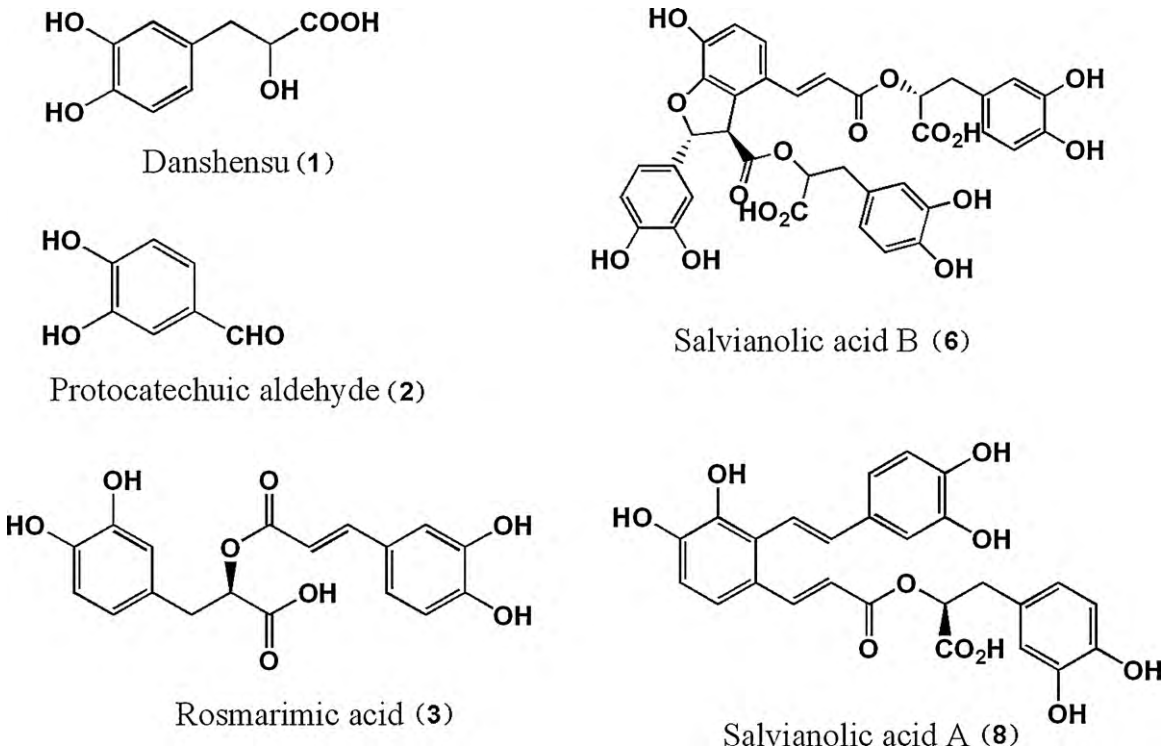


**Fig. 1.** Chemical structures of identified markers.

### 2.3. Instrumentation

All data were acquired from Waters Acquity UPLC Q-Tof Premier MS system which consisted of a binary solvent manager, a sample manager, a photo diode array (PDA) detector (Waters Co., Milford, USA) and a Waters Q-Tof Premier mass spectrometer equipped with an ESI source (Micromass MS Technologies, Manchester, UK). The sample was separated on a Waters Acquity UPLC BEH $C_{18}$ (2.1 × 100 mm, 1.7 μm, Ireland) column at 30 °C. A binary gradient elution with mobile phase of (A) 10% 1,4-dioxane and 0.1% formic acid in water and (B) 10% 1,4-dioxane and 0.1% formic acid in acetonitrile: 0–4 min, B 0–20%; 4–7 min, B 20–50%; 7–14 min, A 50–100%. The temperature of autosampler was maintained at 10 °C and injection volume was 2.0 μL. The chromatogram was recorded at 280 nm. A high resolution electrospray ionization-mass spectrometric (ESI-MS) experiment was performed in positive and negative ion modes, respectively. The source temperature was set at 100 °C. The capillary voltage and cone voltage were set at 2800 V and 40 V, respectively. The nebulization gas was set at 600 L/h and 300 °C, the cone gas was set at 50 L/h. Data were centroided during acquisition using independent reference lock-mass ions via the LockSpray™ interface to ensure accuracy and precision. Leucine-enkephalin at *m/z* 556.2771 was used as the lock mass in positive mode and that at *m/z* 554.2615 was used in negative mode at a concentration of 50 pg/μL with an infusion flow rate of 10 μL/min. Dynamic range enhancement (DRE™) was applied throughout the MS experiment to ensure accurate mass measurement over a wider dynamic range.

### 2.4. Sample preparation

#### 2.4.1. Aqueous extract of Radix Salvia Miltiorrhizae (AERSM)

Five grams of RSM was boiled twice with 100 mL of water for 2 h. Aliquot was filtered and evaporated to dryness with rotary vacuum evaporator. Residues were collected, weighed and stored in a desiccators at 0 °C for further analysis.

#### 2.4.2. Standard solutions

Stock solutions of standards were prepared in the concentration ranging from 1.2 to 3.0 mg/mL in methanol and stored at −20 °C until use. Standard working solutions were prepared by diluting stock solutions to desired concentrations. Six-point calibration curves were established for quantitation and triplicates of each standard were measured.

#### 2.4.3. Time-course study on heating of AERSM

Powdered AERSM of 0.1 g was dissolved in 10 mL of water and then heated in boiling water. Solution of 0.2 mL was aspirated at 0, 1, 2, 4, 8, 12, 16 and 20 h, respectively. All the solutions were stored at 0–4 °C refrigerator prior to analysis.

#### 2.4.4. Sample solutions

For RSM and EGRSM, powdered samples of 0.2 g was accurately weighed and ultrasonicated with 10 mL of 70% methanol for 20 min. For AERSM, powdered sample of 0.1 g was accurately weighed and ultrasonicated with 10 mL of 70% methanol for 20 min. After cooling down to the room temperature, supernatant was collected and filtered through a membrane filter (0.2 μm) for further analysis. All the solutions were stored at 0–4 °C refrigerator prior to analysis.

### 2.5. Chemonetric data analysis

All chromatographic data were calculated with MassLynx application manager version 4.1 (Waters Co., Milford, USA). Peak areas at UV 280 nm were integrated with "ApexTrack Peak Integration" mode. The retention time ($t_R$) was ranged from 0 to 15 min with tolerance of 0.01 min and the peak integration parameters were set as automatic. Then, the detected peak areas together with their respective retention time were exported as Excel files. The peak areas were normalized (area of each peak at 280 nm was divided by its sample weight, μV/mg) to obtain the datasets. Results from both classical PCA and RPCA were processed with TOMCAT, a graphical interface for robust multivariate calibration technique [15]

**Table 1**
The on-line chromatograms and high resolution mass data of constituents in RSM and EGRSM which were subjected to RPCA.

| Peak no. | Retention time (min) | HR-mass (*m*/*z*) | Tolerance (ppm) | Formula | Compound |
|---|---|---|---|---|---|
| **1** | 1.45 | 197.0444 $[M-H]^-$ | −3.0 | $C_9H_{10}O_5$ | Danshensu |
| **2** | 2.30 | 137.0234 $[M-H]^-$ | −3.6 | $C_7H_6O_3$ | Protocatechuic aldehyde |
| **3** | 6.97 | 359.0758 $[M-H]^-$ | −2.5 | $C_{18}H_{16}O_8$ | Rosmarinic acid |
| **4** | 7.27 | 417.0818 $[M-H]^-$ | −1.0 | $C_{20}H_{18}O_{10}$ | Salvianolic acid D |
| **5** | 7.47 | 537.1022 $[M-H]^-$ | −2.0 | $C_{27}H_{22}O_{12}$ | Lithospermic acid |
| **6** | 8.13 | 717.1438 $[M-H]^-$ | −2.5 | $C_{36}H_{30}O_{16}$ | Salvianolic acid B |
| **7** | 8.23 | 717.1458 $[M-H]^-$ | 0.3 | $C_{36}H_{30}O_{16}$ | Salvianolic acid E/L |
| **8** | 8.69 | 493.1139 $[M-H]^-$ | 0.8 | $C_{26}H_{22}O_{10}$ | Salvianolic acid A |
| **7′** | 8.81 | 717.1454 $[M-H]^-$ | −0.3 | $C_{36}H_{30}O_{16}$ | Salvianolic acid E/L |
| **9** | 9.23 | 491.0970 $[M-H]^-$ | −1.6 | $C_{26}H_{20}O_{10}$ | Salvianolic acid C/Isosalvianolic acid C |
| **9′** | 9.89 | 491.0967 $[M-H]^-$ | −2.2 | $C_{26}H_{20}O_{10}$ | Salvianolic acid C/Isosalvianolic acid C |
| **10** | 12.96 | 279.1015 $[M+H]^+$ | −2.1 | $C_{18}H_{14}O_3$ | Dihydrotanshinone I |
| **11** | 13.16 | 297.1484 $[M+H]^+$ | −2.4 | $C_{19}H_{20}O_3$ | Cryptotanshinone |
| **12** | 13.34 | 277.0857 $[M+H]^+$ | −2.9 | $C_{18}H_{12}O_3$ | Tanshinone I |
| **13** | 13.88 | 295.1326 $[M+H]^+$ | −2.7 | $C_{19}H_{18}O_3$ | Tanshinone IIA |

in Matlab 7.0 (The Mathworks, Inc.). After centering the data by *Qn*-autoscaling, RPCA was applied and the outputs of a matrix containing the robust principal components and a matrix containing robust loadings were transferred to Matlab to construct three-dimensional scatter plots (score plots and loading plots for the first three principal components). All data with three replicates were reported as mean ± standard deviation.

### 2.6. Antioxidant activity assay

The antioxidant activity was determined using a modified ferric reducing ability of plasma (FRAP) assay [16]. The stock solutions including 200 mM acetate buffer (pH 3.6), 10 mM TPTZ in 40 mM HCl and 20 mM ferric chloride solution were prepared. The fresh working solution (FRAP solution) was prepared by mixing acetate buffer, TPTZ solution and ferric chloride solution in the ratio of 10:1:1. The reaction was performed in a 96-well microplate and extracts of RSM, AERSM and EGRSM (5 μL) were allowed to react with 245 μL of the FRAP solution for 30 min in the dark. Then, readings at the absorption maximum (593 nm) were measured using a Biotek ELX 808 microplate reader (Biotek Instruments Inc., Winooski, USA). Different concentrations of Trolox solution were used to establish the calibration curve and the concentration of sample was expressed as Trolox equivalent antioxidant capacity (TEAC) in μM per gram of dried sample.

## 3. Results and discussion

### 3.1. Optimization of UHPLC separation and validation of methodology

Since chemical compositions of RSM and EGRSM are complex with isomeric constituents, the UHPLC condition was optimized to provide sufficient separation within a short running time for high throughout analysis. In our previously study, it was found that using 1,4-dioxane as the mobile phase modifier could significantly improve the separation of phenolic compounds, especially for several isomeric components, peaks **6**, **7** and **7′** with the same *m*/*z* of 717.1456, peaks **9** and **9′** with the same *m*/*z* of 491.0978 (Table 1 and Fig. 2). The separation condition was then modified and improved. An optimum mobile phase consisting of 0.2% formic acid, 10% of 1,4-dioxane and acetonitrile was selected which showed satisfactory separation on contents of RSM and EGRSM.

The repeatability was evaluated by measuring the variation of quintuple sample solutions, respectively. The relative standard deviation (RSD) of retention times ($t_R$) and peak areas (PA) were 0.02–0.05% and 1.34–3.26%, respectively. The stability study was performed at 10 °C within 48 h and at 0–4 °C for 1 week. The RSD for stability study at 0–4 °C and 10 °C were 2.1% and 3.2%, respectively. The results indicated that the optimized method was valid.

### 3.2. Extraction of specific chemical markers by robust principal components analysis

In the present study, the PCA and RPCA (via Projection Pursuit PP approach based on the *Qn*-estimator scale [14]) were applied to the three groups of samples including 39 RSM, 20 AERSM and 38 EGRSM in order to determine the variation of specific constituents between RSM and EGRSM as well as the effects of manufactur-

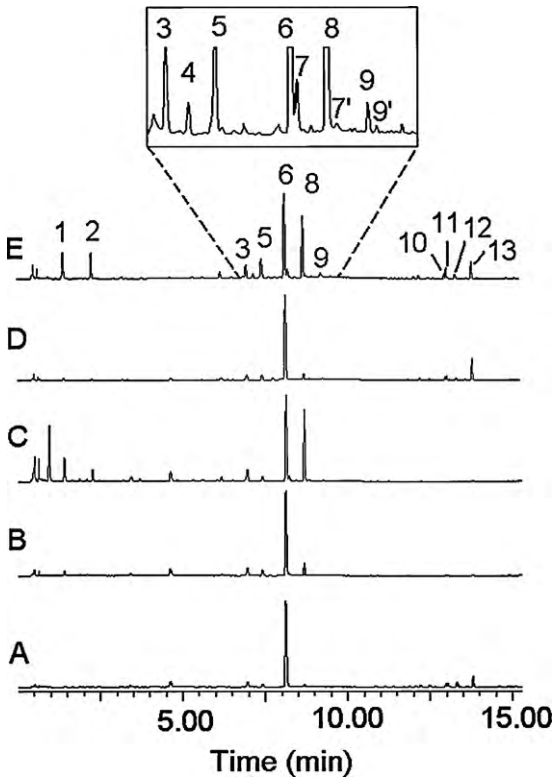


**Fig. 2.** Typical chromatograms of RSM (A); AERSM (B); EGRSM sample #67 (C) showed high content of **1**, **2**, **7** and **8** representing manufacturing of aqueous extract under high temperature or prolonged heating; EGRSM sample #95 (D) contained both phenolic compounds and tanshinones with very low content of **1**, **2** and **8** representing manufacturing of alcohol extract under low temperature or short heating; and EGRSM sample #73 (E) contained both phenolic compounds and tanshinones with high content of **1**, **2**, **7** and **8** representing manufacturing of alcohol extract under high temperature or prolonged heating. Peak identification was illustrated in Table 1.

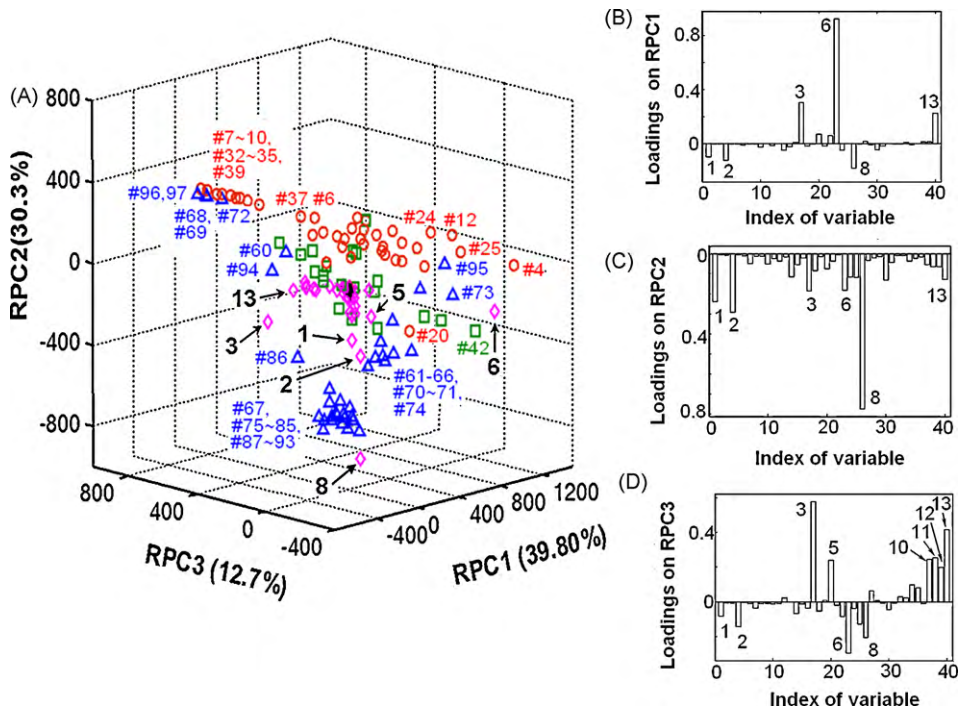


**Fig. 3.** Spatial plot of eigenvectors and eigenvalues from RPCA in RSM (samples #1–#39), AERSM and EGRSM (samples #60–#97, panel A); and the loadings of individual variables (detected components in the samples) in RPC1 (B), RPC2 (C) and RPC3 (D); Symbols in (A), (B) and (C): red circles represented RSM samples, green squares represented AERSM samples and blue triangles represented EGRSM samples. Pink diamonds showed the loadings of variables (loadings × 20 in A, loadings × 200 in B and loadings × 100 in C). The numbers of **1**–**13** represent the extracted markers which were identified in Table 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

ing process on the overall quality of EGRSM the effects RSM on the quality of EGRSM. The classical PCA study showed that the scores data were scattered over all directions where no clear pattern for these data could be observed (data not shown). Thus, it is uncertain whether these results reveal the main features of the dataset or simply reflect the presence of outliners. Also, it is unsure whether all outliers could be detected using classical PCA method which could possibly lead to incorrect illustration. In contrast, these samples were clearly clustered according to the differences in contents of individual compositions as revealed by RPCA. To investigate

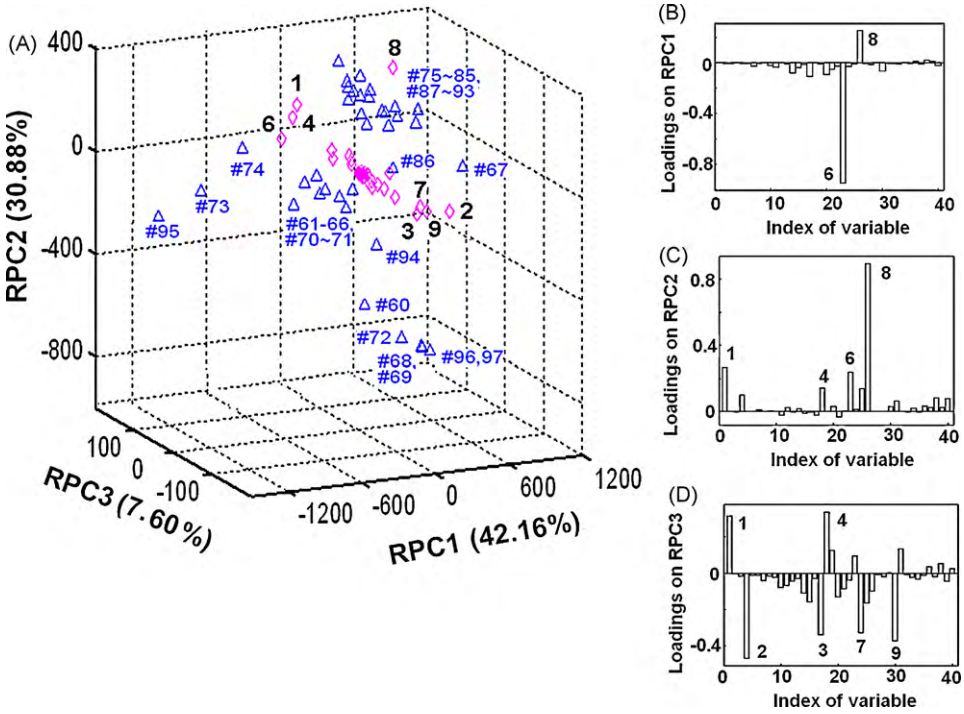


**Fig. 4.** Spatial plot of robust eigenvectors and eigenvalues from RPCA in EGRSM (samples #60–#97, panel A), and the loadings of individual variables (detected components in the samples) in RPC1 (B), RPC2 (C) and RPC3 (D).

the correlation between the loadings (representing constituents in samples) and the scores (representing quality of samples), the superimposed three-dimensional scatter plots on projection of the scores and loadings for the first three RPCs (RPC1, RPC2 and RPC3, respectively) were illustrated in Fig. 3A. It was shown that the first three RPCs accounted for more than 82.8% of the total variance. Samples were clearly clustered according to their differences in chemical compositions, such as one group of data with high RPC1 and RPC2 values (observation #4, #12, #24–25, #42, #73, #95) which corresponded to samples with high content of compounds **3**, **6** and **13** but low content of compound **8**, and another group of data with high RPC2 and RPC3 values (#7–10, #32–35, #68–69, #72, #96–97) corresponded to samples with low contents of all constituents. It was also found that data in the scores plots were ranked according to their degree of variations in chemical compositions, for example, the highest scores rank order for RPC1 was associated with the content of **6** while the order for RPC2 was associated with the content of **1** and **8**. Therefore, RPCA could discriminate samples of different qualities.

When comparing AERSM with RSM, the scores of AERSM showed similar trends to that of RSM with respect to RPC1 and RPC3 values. The RPC2 value for an AERSM sample was lower than that of RSM sample. Since the RPC2 value correlated well to the content of compound **8** (Fig. 3C) which is a degradation product of the major constituent of RSM compound **6**, it could be used as an indicator to monitor the manufacturing process. This was further proven by a time-course study where an AERSM sample was heated for 20 h as mentioned in Section 2.4.3, the content of compound **6** decreased while that of compound **8** increased. With further heating to 4 h, compound **8** was degraded to **1** and **2**.

By investigating the relationship between scores and loadings, RPC1 (39.8%) showed that two phenolic compounds **3**, **6**, and one of the tanshinones compound **13** were positively correlated to RSM but negatively correlated to AERSM and EGRSM (Fig. 3B). This profile is highly indicative of the major constituents in RSM. Since **3** and **6** are the major phenolic components and **13** is the major tanshinone in RSM, they were frequently used as markers for quality evaluation of RSM as stated in scientific publications and official quality standard. The RPC2 (30.3%) was highly correlated to the three phenolic compounds, **1**, **2** and **8** (Fig. 3C) which are degraded from **6**, suggesting them as characteristic markers for quality control of AERSM and EGRSM. The RPC3 (12.7%) was correlated to the phenolic components, **3**, **5** and tanshinones **10**–**13** (Fig. 3D), which revealed the contribution of low-content constituents to overall quality. Therefore, RPC1 and RPC3 could be used to unravel the differences in contents between EGRSM with its raw material, RSM, while RPC2 could be used to reveal the effect of manufacturing process on the quality of EGRSM.

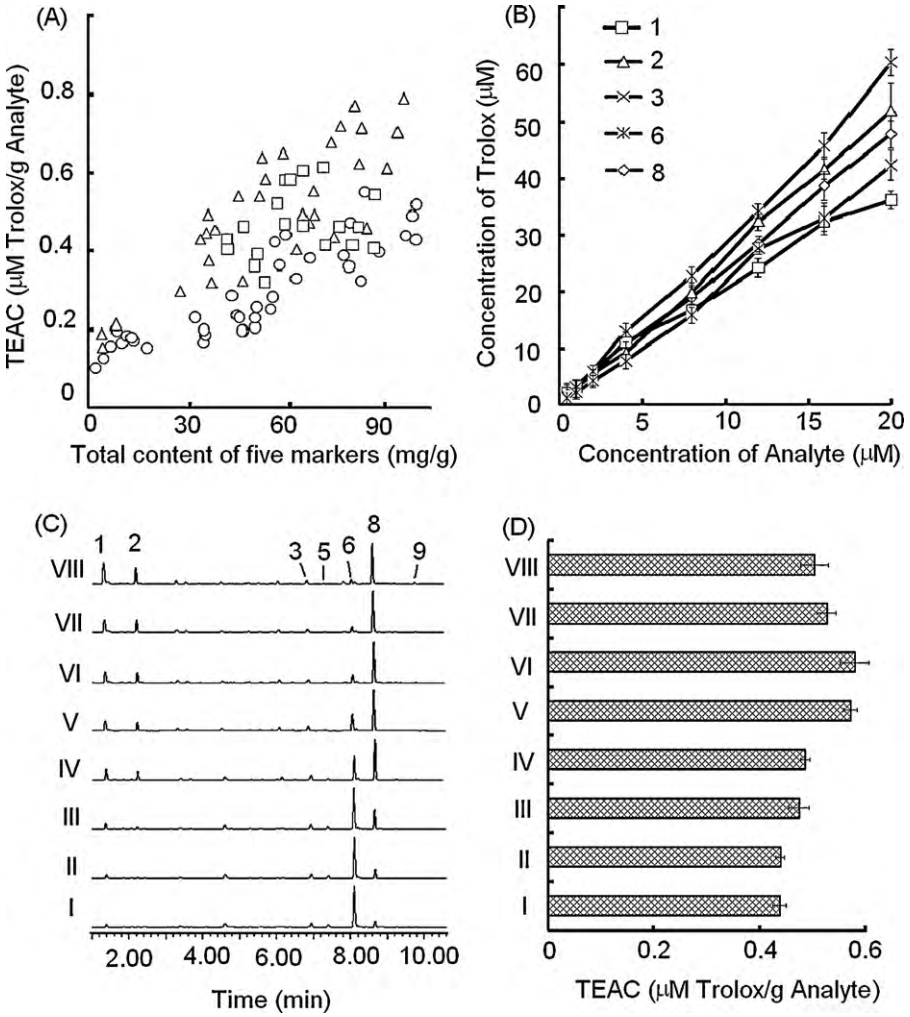


**Fig. 5.** Chromatographic profiles of AERSM at 100 °C in the time-course study (panel A) and their corresponding antioxidant activities (panel B); relationship between Trolox equivalent antioxidant capacity (TEAC) and total contents of five identified markers (**1**, **2**, **3**, **6** and **8**; panel C); the concentration-course relationship between the five identified markers and antioxidant activities (panel D); Symbols: I–VIII in panel (A) and (B) represented an AERSM heated for 0, 1, 2, 4, 8, 12, 16 and 20 h, respectively. Peaks in panel (C) were the same as those showed in Table 1; circles, squares and triangles in panel (C) represented samples of RSM, AERSM and EGRSM, respectively.

Furthermore, RPCA was performed on the EGRSM samples to identify specific markers for quality control of EGRSM. The superimposed three-dimensional scatter plots for the first three RPCs were illustrated in Fig. 4A and loadings of the first three RPCs were individually illustrated in Fig. 4B–D. It was shown that RPC1 (42.16%) was positively correlated to **8** but negatively correlated to **6** (Fig. 4B). RPC2 (30.88%) was highly correlated to **1**, **4** (salvianolic acid D), **6** and **8** (Fig. 4C) and RPC3 (7.60%) was correlated to **1**, **2**, **3**, **4**, **7** (salvianolic acid E or salvianolic acid L) and **9** (salvianolic acid C) (Fig. 4D). The RPC3 represented only 7.6% of the total variation suggesting that it was less specific for quality differentiation. Therefore, from plots of RPCA, the most relevant variables which can be used as characteristic marker compounds were identified as **1**, **2**, **3**, **4**, **6** and **8**.

### 3.3. Identification of characteristic peaks

More than 40 peaks were detected from the EGRSM samples, in which 15 of them were found to have contributions to results of RPCA (Table 1). The on-line UHPLC combined with high resolution ESI-Q-Tof MS was used for identifying structures of the characteristic markers. ESI operating in both negative and positive modes were attempted to detect phenolic compounds and tanshinones. Results indicated that the negative mode was more sensitive to phenolic compounds while the positive mode was more sensitive to tanshinones. Therefore, both negative and positive modes were adopted.

With negative ionization mode, most of the collected *m*/*z* data were of $[M-H]^-$ or $[2M-H]^-$ ions. Since the chemical structures of most polyphenolic acids contained danshensu or caffeic acid moiety, the presence of $[M-H-198]^-$ and $[M-H-180]^-$ fragments indicated that there were losses of danshensu and caffeic acid, respectively. With positive ionization mode, most of the *m*/*z* were of $[M+H]^+$ or $[M+Na]^+$ ions and the major fragments were $[M+H-18]^+$, $[M+H-28]^+$ and $[M+H-15]^+$ corresponding to the loss of $H_2O$, CO and $CH_3$, respectively. The structures of **1**–**3**, **5**, **6**, **8**, **11**–**13** were identified by comparing their respective retention time and accurate *m*/*z* data with the corresponding reference standards. The peaks **4**, **7**, **7′**, **9**, **9′** and **10** were tentatively identified by calculating their elemental compositions from accurate *m*/*z* of parent ions and comparing their fragments with those published data in literatures [8,17].

It was noted that although the peak area of compound **4** was smaller than those of other markers, its contribution to RPC2 was equivalent to that of compound **2** which is a frequently used marker for quality evaluation of RSM products. Compound **4** could only be found in AERSM and EGRSM, instead of the RSM samples suggesting that it is one of the characteristic analytical markers for quality control of AERSM and EGRSM. Since no reference standard of compound **4** is available, its chemical structure needs to be further confirmed and the development of its commercial reference standard is also required.

### 3.4. Evaluation of selected markers by antioxidant activity assay

The major components in RSM are phenolic compounds which have high antioxidant activities and associated with many biological effects [18–20]. To verify whether the selected markers were suitable for the quality control, the relationship between selected markers and their antioxidant activities as determined by FRAP-based Trolox equivalent antioxidant capacity (TEAC) was investigated. The results were shown in Fig. 5A–D. It was indicated that antioxidant activity was positively correlated to the total contents of five selected phenolic components (**1**, **2**, **3**, **6**, **8**) and the antioxidant activities were ranked as follow EGRSM > AERSM > RSM (Fig. 5A). The antioxidant activities were positively correlated to the five independent markers, respectively. As shown in Fig. 5B, **3** exhibited the highest antioxidant activity, followed by **2**, **6** and **8** while **1** showed the lowest activity. This result suggested that **6** and its two major decomposed compounds, **2** and **8**, contributed much to the antioxidant activities exerted by AERSM and EGRSM.

Furthermore, a sample of AERSM was boiled in water for 1, 2, 4, 8, 12, 16 and 20 h respectively to evaluate the suitability of the selected markers for quality control during manufacturing process. The UHPLC profile and the corresponding antioxidant activities for each time point were shown in Fig. 5C and D, respectively. When the heating time was increased to 4 h, peak area of **6** was decreased while those of **2** and **8** were significantly increased (Fig. 5C, I–IV). Further increasing the heating time, peak areas of **6** and **8** were decreased while those of **1** and **2** were increased (Fig. 5D, V–VIII). In contrast, peak of **3** was decreased with the increasing heating time. The antioxidant activities among samples at different heating time intervals and variations in chemical compositions could be well-correlated, for example, when content of **8** was increased during heating for 12 h and the cor-

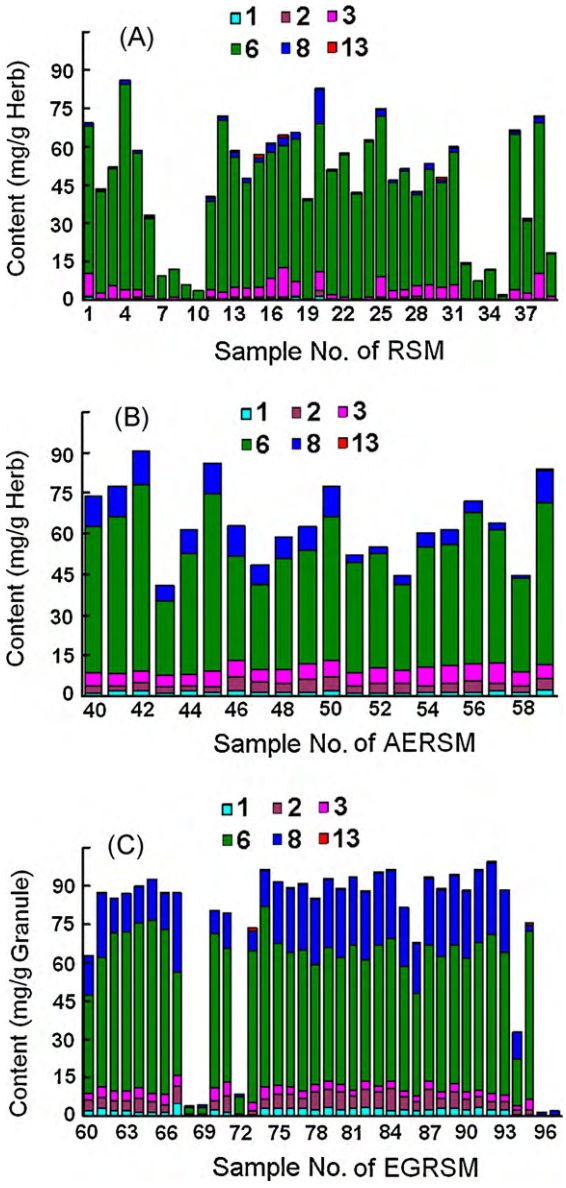


**Fig. 6.** Contents of five identified markers in RSM (A), AERSM (B), commercial products of EGRSM (C). Results were average of triplicate determinations, and value of RSD was ranged from 0.47 to 3.15%.

responding antioxidant activity increased accordingly. Therefore, the selected marker could not only reflect the final quality of EGRSM, but also be useful for quality control during manufacturing process.

### 3.5. Quality assessment of EGRSM by UHPLC analysis

Chromatograms recorded at UV 280 nm were used for the quality assessment. Totally 38 batches of EGRSM from 13 manufacturers were determined. In parallel, qualities of 39 RSM and 20 AERSM were also determined (Fig. 6A–C). Chromatographic profiles showed different fingerprint patterns. Typical chromatograms with representative patterns were illustrated in Fig. 2. Contents of the five identified markers (**1**, **2**, **3**, **6** and **8**) in EGRSM were calculated with reference to their respective calibration curves and the results were shown in Fig. 6C. It was shown that content of **6**, the dominated compound in RSM was lower in EGRSM and AERSM. Contents of **1**, **2** and **8** were much higher in EGRSM and AERSM than those in RSM which was attributed to the long heating time as illustrated in Fig. 5C. In addition, the quality of granules varied between manufacturers, high consistency was found in two of the manufacturers (Sample No. 2–7 and No. 16–34, respectively), indicated that monitoring the manufacturing process is one of the most important sections for product quality assurance. Therefore, the quality of commercial products of EGRSM can be well evaluated and interpreted using markers identified by the newly developed approach.

## 4. Conclusion

In this study, robust principal component analysis combined with UHPLC–UV–Q-Tof MS provides a rapid approach to explore characteristic analytical markers for quality evaluation of AERSM. With high-throughput UHPLC–UV–Q-Tof MS, large amount of samples can be rapidly analyzed and then RPCA can discriminate samples with tremendous quality variations as outliers, which are beneficial for the exploration of characteristic analytical markers for quality control. The identified markers were positively correlated to the bioactivities, suggested that the selected analytical markers are practical and reliable for quality evaluation. In summary, the present approach is rapid and reliable for analytical marker selection and can be applicable to quality evaluation of herbal medicines with undefined therapeutic active constituents.

## Acknowledgements

This research is funded by the Hong Kong Jockey Club Charities Trust.

## References

[1] The European Medicines Agency: reflection paper on markers used for quantitative and qualitative analysis of herbal medicinal, http://www.emea.europa.eu/pdfs/human/hmpc/25362907en.pdf.
[2] S.L. Li, Q.B. Han, C.F. Qiao, J.Z. Song, C.C. Lung, H.X. Xu, Chemical markers for the quality control of herbal medicines: an overview, Chin. Med. 3 (2008), doi:10.1186/1749-8546-3-7.
[3] A. Fugh-Berman, Herbs and dietary supplements in the prevention and treatment of cardiovascular disease, Prev. Cardiol. 3 (2000) 24–32.
[4] M. Imanshahidi, H. Hosseinzadeh, The pharmacological effects of *Salvia* species on the central nervous system, Phytother. Res. 20 (2006) 427–437.
[5] X. Wang, S.L. Morris-Natschke, K.H. Lee, New developments in the chemistry and biology of the bioactive constituents of tanshen, Med. Res. Rev. (2007) 133–148.
[6] Chinese Pharmacopoeia Commission, Pharmacopoeia of the People's Republic of China, vol. I, People's Medical Publishing House, 2005, p. 213–214.
[7] Y.G. Li, L. Song, M. Liu, B.H. Zhi, Z.T. Wang, Advancement in analysis of *Salviae miltiorrhizae* Radix et Rhizoma (Danshen), J. Chromatogr. A 1216 (2009) 1941–1953.
[8] M. Liu, Y.G. Li, F. Zhang, L. Yang, G.X. Chou, Z.T. Wang, Z.B. Hu, Chromatographic fingerprinting analysis of Danshen root (*Salvia miltiorrhiza Radix et Rhizoma*) and its preparations using high performance liquid chromatography with diode array detection and electrospray mass spectrometry, J. Sep. Sci. 30 (2007) 2256–2267.
[9] M. Liu, Y. Li, G. Chou, X. Cheng, M. Zhang, Z. Wang, Extraction and ultra-performance liquid chromatography of hydrophilic and lipophilic bioactive components in a Chinese herb Radix Salviae Miltiorrhizae, J. Chromatogr. A 1157 (2007) 51–55.
[10] J.L. Zhang, M. Cui, Y. He, H.L. Yu, D.A. Guo, Chemical fingerprint and metabolic fingerprint analysis of Danshen injection by HPLC–UV and HPLC–MS methods, J. Pharm. Biomed. Anal. 36 (2005) 1029–1035.
[11] Y. Zhao, B. Chen, S. Yao, Simultaneous determination of abietane-type diterpenes, flavonolignans and phenolic compounds in compound preparations of *Silybum marianum* and *Salvia miltiorrhiza* by HPLC-DAD-ESI MS, J. Pharm. Biomed. Anal. 38 (2005) 564–570.
[12] Z. Shi, J. He, T. Yao, W. Chang, M. Zhao, Simultaneous determination of cryptotanshinone, tanshinone I and tanshinone IIA in traditional Chinese medicinal preparations containing *Radix salvia miltiorrhiza* by HPLC, J. Pharm. Biomed. Anal. 37 (2005) 481–486.
[13] C. Croux, A. Ruiz-Gazen, High breakdown estimators for principal components: the projection-pursuit approach revisited, J. Multivar. Anal. 95 (2005) 206–226.
[14] M. Daszykowski, K. Kaczmarek, Y.V. Heyden, B. Walczak, Robust statistics in data analysis—a review: basic concepts, Chemometrics Intell. Lab. Syst. 85 (2007) 203–219.
[15] M. Daszykowski, S. Serneels, K. Kaczmarek, P. Van Espen, C. Croux, B. Walczak, TOMCAT: a MATLAB toolbox for multivariate calibration techniques, Chemometrics Intell. Lab. Syst. 85 (2007) 269–277.
[16] Y.Q. Ma, X.Q. Ye, Z.X. Fang, J.C. Chen, G.H. Xu, D.H. Liu, Phenolic compounds and antioxidant activity of extracts from ultrasonic treatment of Satsuma Mandarin (Citrus unshiu Marc.) Peels, J. Agric. Food Chem. 56 (2008) 5682–5690.
[17] A.H. Liu, H. Guo, M. Ye, Y.H. Lin, J.H. Sun, M. Xu, D.A. Guo, Detection, characterization and identification of phenolic acids in Danshen using high-performance liquid chromatography with diode array detection and electrospray ionization mass spectrometry, J. Chromatogr. A 1161 (2007) 170–182.
[18] T.Y. Lee, H.H. Chang, G.J. Wang, J.H. Chiu, Y.Y. Yang, H.C. Lin, Water-soluble extract of *Salvia miltiorrhiza* ameliorates carbon tetrachloride-mediated hepatic apoptosis in rats, J. Pharm. Pharmacol. 58 (2006) 659–665.
[19] J. Sun, S.H. Huang, B.K. Tan, M. Whiteman, Y.C. Zhu, Y.J. Wu, Y. Ng, W. Duan, Y.Z. Zhu, Effects of purified herbal extract of *Salvia miltiorrhiza* on ischemic rat myocardium after acute myocardial infarction, Life Sci. 76 (2005) 2849–2860.
[20] J. Liu, C.F. Yang, S. Wasser, H.M. Shen, C.E. Tan, C.N. Ong, Protection of *Salvia Miltiorrhiza* against aflatoxin-$B_1$-induced hepatocarcinogenesis in Fischer 344 rats: dual mechanisms involved, Life Sci. 69 (2001) 309–326.