Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1057 (2004) 141–149
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

High-performance liquid chromatographic method with quantitative comparisons of whole chromatograms of raw and steamed *Panax notoginseng*

Aik-Jiang Lau[a], Bee-Hong Seo[b], Soo-On Woo[c], Hwee-Ling Koh[a,*]

[a] *Department of Pharmacy, Faculty of Science, National University of Singapore, 18 Science Drive 4, Singapore 117543, Singapore*
[b] *Waters Asia Ltd, 396 Alexandra Road, #04-06 BP Tower, Singapore 119954, Singapore*
[c] *Universiti Tunku Abdul Rahman, Faculty of Engineering and Science, 13 Jalan 13/6, 46200 Petaling Jaya, Selangor Darul Ehsan, Malaysia*

Received 14 April 2004; received in revised form 15 September 2004; accepted 21 September 2004

Abstract

A high-performance liquid chromatographic (HPLC) method coupled with chromatographic pattern matching was developed to differentiate whole chromatograms of raw and steamed *Panax notoginseng* objectively and quantitatively. The major peaks differentiating chromatograms of raw and steamed samples were also identified for the first time in this herb. The raw and steamed *P. notoginseng* roots and its products were successfully differentiated. The quantitative differences between the chromatograms were correlated to the duration of steaming. Chromatographic pattern matching allows rapid, simple, automated, and quantitative comparisons of complex chromatograms. It is a useful tool in ensuring safety and quality of herbal products.

Keywords: Chromatographic pattern matching; *Panax notoginseng*; Medicinal herbs; Ginsenosides; Chemical fingerprinting; Alignment algorithm

1. Introduction

Medicinal herbs have gained popularity in many countries. This raises many concerns with regards to their quality control, which is still a common problem today. The quality and chemical content of herbs vary greatly due to many factors such as species variation, geographical source, cultivation, harvest, storage, and processing [1]. Better scientific methodologies are still needed to evaluate and assess medicinal herbs and their products.

Unlike synthetic drugs of high purity, medicinal herbs and their products have a very complex mixture of chemical components whose identity is only partially known. Often, a few chemical markers in the chromatograms are selected and employed in evaluating the quality and authenticity of herbs. This method does not evaluate the entire chromatographic profile and large amounts of data in the chromatograms are discarded. Furthermore, similar integration results may not be achievable if baseline resolution is not achieved in complex mixtures. Selection of suitable markers to correctly identify the herb is also difficult and subjective. Therefore, this approach is neither sufficient nor satisfactory for quality control of herbs.

In recent years, the use of chromatographic chemical fingerprinting for the identification and quality control of medicinal herbs has attracted a lot of interest [2–5]. Analysis of chromatographic profiles, generally with the goal of making a classification, is known as 'fingerprinting'. Fingerprinting using chromatographic methods is also one of the requirements proposed by US Food and Drug Administration (FDA) for botanicals [6] and The European Agency for the Evaluation of Medicinal Products for herbal

* Corresponding author. Tel.: +65 68747962; fax: +65 67791554.

0021-9673/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2004.09.069

preparations [7]. It is of importance because the chromatographic fingerprints are unique and represent powerful tools for the comparison, classification, identification, and evaluation of samples. However, due to the complex fingerprints of herbal samples and chromatographic variations, accurate analysis and interpretation of the chromatograms in chemical fingerprinting still pose a great challenge to analysts.

One method to compare complex fingerprints is by visual comparison. This traditional method of visual chromatographic comparison is simple, but it is very subjective and non-quantitative. For complex chromatograms with incomplete separation of peaks, visual comparisons can be difficult and may miss subtle differences. Moreover, chromatography always varies from run-to-run due to pump, temperature, sample injection variations as well as changes in mobile phase and column chemistries. The resulting run-to-run chromatographic variations such as retention time drift and baseline drift, make the visual comparison method more ambiguous. In some cases, these variations also make the methods analyzing simple difference or variance of the chromatographic response non-applicable. Therefore, there is a need for a simple, valuable tool to objectively compare the entire chromatograms, detect real sample differences between them and measure the degree of differences quantitatively.

With the rapidly increasing computational power and rapid development of the field of chemometrics in the last two decades, it is now possible to use complex mathematical algorithms to analyze the whole chromatogram quantitatively, and automatically. Such large data sets, which are previously considered impractical, can now be handled. This approach has been applied to medicinal herbs [3–5], pharmaceuticals [8], and food [9,10], etc. Typically, most chromatographic analysis techniques involve developing a method of identifying some peaks to be compared, developing a method of comparing the various aspects of the peaks (such as peak area, height), and then actually performing the comparisons. However, the inevitable variations of chromatographic peak parameters such as retention time have always been a major impediment against accurate data processing in most chemometric analysis techniques. Multivariate chemometric analysis with entire chromatographic profiles as input data, is very sensitive to even minute variations [3,11–13]. Therefore, many new approaches for retention time adjustments/peak alignment [3,11–13] and extracting information from the fingerprints [4] have also been studied to address this challenging problem for chemometric analysis. Malmquist and Danielsson's alignment algorithm [12] involves four rounds of iterative shifting to optimize sample-to-target correlation. Gong et al. [3] recently used a combination of chemometric resolution with cubic spline data interpolation to select marker compounds, correct the retention time shifts and reconstruct the chromatographic fingerprints with correction. Torgrip et al. [13] suggest that the alignment methods can be divided roughly into two major categories, coarse and fine. The 'coarse' methods such as dynamic time warping, correlation optimized warping (COW), shift the peaks up to the same resolution as the scanning instrument. The 'fine' methods shift peaks beyond the instrument's spectral resolution. The authors also developed a new search algorithm, the breadth first search (BFS), which is shown to be favorable in terms of computational speed. After alignment of chromatograms and extracting the relevant information from the fingerprints, samples can then be classified based on several available multivariate data analysis methods such as principal component analysis (PCA), soft independent modeling of class analogy (SIMCA), K-nearest neighbors (KNN), and linear discriminant analysis (LDA). Artificial neural networks (ANN) are also rapidly emerging in the field of analytical chemistry as a powerful tool for pattern searching, mapping, and fingerprinting. There are also a number of studies [8,9] using ANN and the comparison [8] of several ANN architectures with standard classifiers such as KNN and SIMCA.

Panax notoginseng (Burk.) F.H. Chen or Sanqi is a highly valued and important Chinese medicinal herb, belonging to the same genus as Chinese and Korean ginseng (*Panax ginseng*) and American ginseng (*Panax quinquefolium*). *P. notoginseng* is available in two different forms—the raw and steamed forms. Traditionally, the raw form is widely used in Chinese medicine for its hemostatic and cardiovascular properties [14,15], while the steamed form has been claimed to be a tonic used to "nourish" blood and to increase production of various blood cells in anaemic conditions [16]. Due to their different pharmacological actions and clinical indications, using the wrong form of herb may lead to undesirable results. Quality control of this herb is hence important. A previous study [17] showed that there were visually distinct differences between the chromatograms of raw and steamed samples and the concentration of some saponins were changed.

This study aims to use a high-performance liquid chromatographic (HPLC) pattern matching method as a new approach to objectively and quantitatively differentiate between the raw and steamed *P. notoginseng*. To date, this pattern matching analysis tool has not been studied or applied to complex samples such as medicinal herbs. The tool will take into account 5 parameters of chromatographic variations to align the chromatograms, and it does not require the characterization of peaks or other chromatographic features that are required by other known techniques. After chromatographic alignment and comparison, a quantitative value showing the relative difference between the raw and steamed samples will be generated and it is used as a criterion to differentiate between raw and steamed samples. Major peaks in the chromatograms of steamed *P. notoginseng*, which can serve to differentiate raw and steamed forms, are also identified for the first time. This method will then be applied on 11 pairs of raw and steamed *P. notoginseng* proprietary products to measure their degree of similarity or differences between each other.

2. Experimental

2.1. Materials

The water used was treated with a Milli-Q water purification system (Millipore, Molsheim, France). HPLC-grade solvents were used for the analysis. The raw *P. notoginseng* root was obtained from a Chinese medical shop in Singapore. Eleven pairs of raw and steamed *P. notoginseng* Chinese Proprietary Medicines (CPMs) were also obtained from various local Chinese medical shops (Table 1).

Ginsenosides Rb1, Rc, Rd, Re, and Rg1 standards were purchased from Indofine Chemical Company (Somerville, NJ, USA). Notoginsenoside R1 was obtained from National Institute for the Control of Pharmaceutical and Biological Products (Beijing, China). 20R-ginsenosides Rg3, 20S-ginsenoside Rg3, mixture of ginsenoside Rg5 and Rk1 were generous gifts from Professor J.H. Park (College of Pharmacy, Seoul National University, Korea). 20RS-ginsenoside Rh1 were obtained from Delta Information Centre for Natural Organic Compounds, China.

2.2. Sample preparation

A 10 ml volume of 70% (v/v) methanol was added to 1 g of the powdered sample. The suspension was ultrasonically (230 V, Branson model 5510, Danbury, CT, USA) extracted for 20 min and filtered. This extraction was repeated two additional times. The combined filtrate was evaporated to dryness in vacuo. The residue was then dissolved in 5 ml of 70% (v/v) methanol and filtered through a 0.45 μm nylon filter membrane prior to HPLC analysis.

2.3. Steaming of raw P. notoginseng herb

Samples of the powdered raw *P. notoginseng* root were steamed at 120 °C using an autoclave (Hirayama, Japan) for 2, 6, and 9 h. The powder was then dried in a vacuum oven at about 80 °C until constant weight and extracted using ultrasonication as described above.

2.4. Isolation and identification of peaks in steamed P. notoginseng

Two hundred and thirty grams of steamed *P. notoginseng* (9 h) was extracted with methanol (2 l) by ultrasonication for 2 h and filtered. This step was repeated for a total of five times. The combined filtrates were evaporated in vacuo. The residue (34 g) was dissolved in 140 ml water, washed with 200 ml *n*-hexane (three times), and extracted with 300 ml water saturated *n*-butanol (five times). The major differentiating peaks were then isolated from the butanol extract using normal phase open-column chromatography (silica gel 60, 63–200 μm, 500 mm × 30 mm, gradient elution using dichloromethane and methanol), followed by semi-preparative HPLC (Agilent Zorbax SB-C_{18}, 250 mm × 9.4 mm i.d., 5 μm) to obtain the pure compounds. Their identities were determined by comparing their ^{13}C NMR data, MS data and melting points with the data obtained from literature [18–21] as well as by comparing their retention times with those of the standards, if available.

2.5. HPLC with chromatographic pattern matching analysis

HPLC analysis was performed on a Waters Alliance liquid chromatograph (Milford, MA, USA) equipped with Alliance separation module 2695 and photodiode array detector 2996. The reversed-phase column used was Waters Symmetry C_{18} (250 mm × 4.6 mm i.d., 5 μm). The binary gradient elution system consisted of (A) water and (B) acetonitrile. Separation was achieved using the following gradient: 0–30 min, 20% B; 30–60 min, 20–45% B; 60–78 min, 45–75% B; 78–80 min, 75–100% B. The column temperature was kept constant at 35 °C. The flow rate was 1 ml/min and the injection volume was 5 μl. The UV detection wavelength was set at 203 nm. Sample analysis was processed by Waters Empower software 2002 with chromatographic pattern matching tool.

The chromatographic pattern match processing method parameters used for the comparisons of samples were optimized as follows. Replicate injections of identical samples were carried out and used to develop method parameters, namely, scan start and stop times, peak width, alignment interval, retention time search limit, detection threshold, response value, and percent peak height. The scan start and stop times were 10 and 78 min. The detection threshold was set at 350. The response value and percent peak height were 0.0005 and 4.3%, respectively. After the parameters were optimized, the same pattern match processing method was then applied to all the samples compared.

Statistical data analysis in this study was performed using the unpaired Student's *t*-test with a minimum of six replicates. Differences were considered to be significant when p values were <0.05.

3. Results and discussion

3.1. Identification of major peaks in steamed samples

Eight major potential markers (20S-ginsenoside Rh1, 20R-ginsenoside Rh1, 20S-ginsenoside Rg3, 20R-ginsenoside Rg3, ginsenosides Rk3, Rh4, Rk1, Rg5) that play key roles in differentiating the chromatograms of raw and steamed *P. notoginseng* were identified (Fig. 1). This is the first report of the detection and isolation of ginsenoside Rk3 from *P. notoginseng* herb. These peaks were not detected or were present in very low amounts in raw samples. The steaming process has caused chemical degradation and conversion of some saponins to new compounds and their

Table 1
Pattern match standard deviations (PMSD) of raw and steamed root, and the 11 pairs of products (raw form is taken as the reference)

CPM pair no.	Samples	Brand name	Pattern match standard deviation (AU)
	Raw *P. notoginseng* root (replicate injections)		0.0004
	Steamed *P. notoginseng* root (replicate injections)		0.0004
	Steamed (2 h) vs. raw *P. notoginseng* root		0.0024
	Steamed (6 h) vs. raw *P. notoginseng* root		0.0096
	Steamed (9 h) vs. raw *P. notoginseng* root		0.0120
1	Pure raw pseudoginseng powder Pure steamed pseudoginseng powder	Meihua	0.0041
2	Raw tienchi ginseng tablet Steamed tienchi ginseng tablet	Meihua	0.0021
3	Yunnan tienchi powder (raw) Yunnan tienchi powder (steamed)	Nature's Green	0.0012
4	Yunnan tienchi tablets (raw) Yunnan tienchi tablets (steamed)	Nature's Green	0.0009
5	Tienchi powder (raw) Tienchi powder (steamed)	Yunfeng	0.0050
6	Tienchi tablets (raw) Tienchi tablets (steamed)	Yunfeng	0.0028
7	Tienchi powder (raw) Tienchi powder (steamed)	Camellia	0.0026
8	Tienchi tablets (raw) Tienchi tablets (steamed)	Camellia	0.0025
9	Tienchi tablet (raw) Tienchi tablet (steamed)	Yulin	0.0005
10	Yunnan tian qi powder (raw) Yunnan tian qi powder (steamed)	Kiat Ling	0.0004
11	Chinese yunnan tien chi tablet (raw) Chinese yunnan tien chi tablet (steamed)	Luen Shing	0.0004

structures are shown in Fig. 2. Therefore, the standardization of the steaming process is important in ensuring consistent quality of steamed products. This is also the first report of 14 well-resolved saponins in a single HPLC chromatogram of *P. notoginseng*.

3.2. Chromatographic pattern matching software

The chromatographic pattern matching software [22,23] treats the entire chromatogram as a pattern and compares chromatograms in pairs. One is specified as reference (typically a known standard) against which the software compares the other sample chromatogram. In this application, the raw *P. notoginseng* samples were specified as reference. It uses a chromatographic alignment algorithm to align corresponding retention intervals from two chromatograms.

The alignment algorithm is based on a two-step procedure. First, the algorithm fixes the responses in the sample chromatogram and mathematically adjusts the reference responses by least-squares optimization to produce the best fit between the two chromatograms. The algorithm is based on the assumption that the normal chromatographic variations can be described by five parameters (concentration/response ratio, baseline offset, baseline drift, retention time offset, retention time scale). It measures and applies these alignment parameters to the reference chromatogram. The response ratio refers to the factor by which the reference chromatogram is multiplied along the *y*-axis to best align with the sample chromatogram. The retention time offset refers to the amount of time by which the peak apices of reference chromatogram is adjusted. Retention time scale refers to the amount by which the retention time scale (*x*-axis) of the reference chromatogram is stretched or compressed to best align with the sample. Baseline offset refers to the amount by which the reference chromatogram was adjusted in the y-axis while baseline drift refers to the slope applied to the reference chromatogram baseline. In measuring all these values, only the raw data is used and no peak integration is performed. The pattern matching algorithm simulates, or models a possible range of responses for the reference chromatogram and finds the values for the five parameters by finding those values that minimize the sum of squared differences between the two chromatograms.

Secondly, after alignment, the algorithm calculates the degree of differences (in terms of standard deviations) between the sample chromatogram and the parameter-adjusted reference chromatogram. The standard deviation is the square root

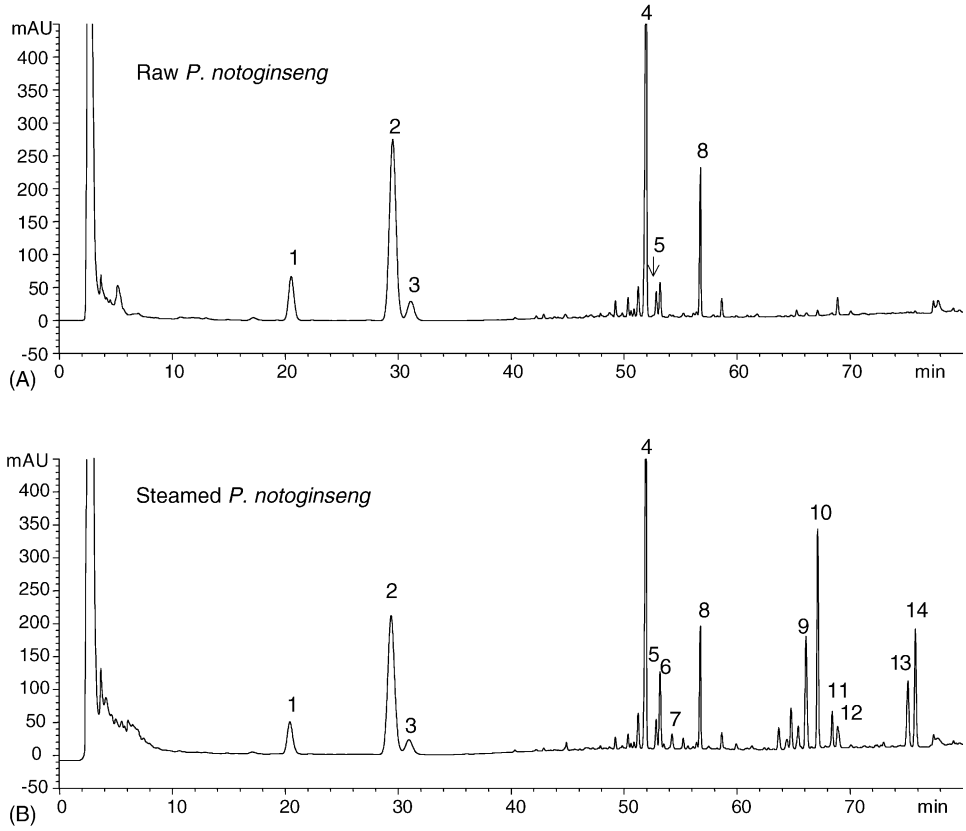


Fig. 1. HPLC chromatograms of (A) raw and (B) steamed *P. notoginseng* (2 h). (1) R1, (2) Rg1, (3) Re, (4) Rb1, (5) Rc, (6) 20S-Rh1, (7) 20R-Rh1, (8) Rd, (9) Rk3, (10) Rh4, (11) 20S-Rg3, (12) 20R-Rg3, (13) Rk1, (14) Rg5.

of the average of a squared difference. The formula is:

$$\text{Standard deviation} = \sqrt{\frac{\sum_{i=1}^{N}(f_i - a_i)^2}{(N-5)}}$$

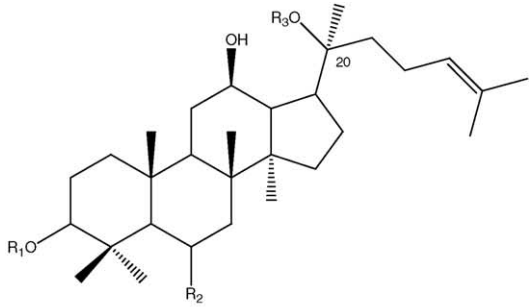
where f_i are the response values in the fixed, sample chromatogram and a_i are the adjusted response values from the reference chromatogram. The index i ranges over the N time samples in the comparison interval. The value for N is reduced by 5 to take into account the effect of the five-parameter alignment on the magnitude of differences.

During a scan of the entire chromatogram, the algorithm centers an alignment interval of fixed width (typically 2-peak widths) on each point over the entire chromatogram, regardless of the presence of peaks. It performs the alignment for each interval and computes the five alignment parameters as well as the standard deviations for each interval within the scan region (from start to stop time). From the whole scan, the pattern match standard deviation (PMSD) value is computed. It is the root-mean-square value of all the individual standard deviations measured within the entire scan region. A Matlab implementation of this alignment algorithm is available in the reference [22]. The term "standard deviation" although being statistically meaningless for a pair of samples, is used in this paper and means a value of the pairwise differences calculated by the corresponding formula.

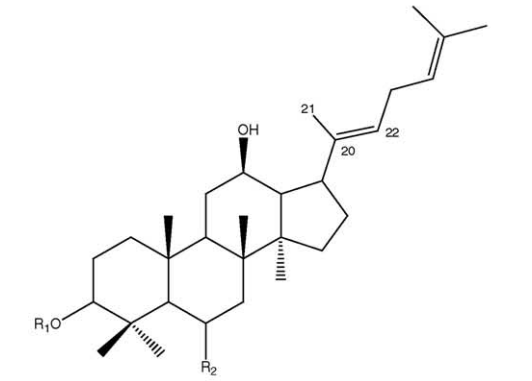
In this method, the variations due to normal chromatographic variations will not affect the comparisons, revealing only changes due to the samples. This alignment algorithm has the advantage of adjusting for five underlying chromatographic variations simultaneously. Furthermore, no selection of peaks, internal standards, or traditional peak integration is needed, compared to known techniques. The algorithm detects peaks using the second derivative of the chromatogram. The apex of the inverted second derivative identifies the apex of a peak. Besides showing how similar or different the samples are, it can also rapidly identify which specific retention time interval of the chromatograms have significant pattern differences or new peaks. This may help to identify and isolate the differentiating peaks of interest. Other than chromatograms from LC separation, it can also be applied to GC, CE separation or other imported data.

3.3. Optimization of pattern match processing method

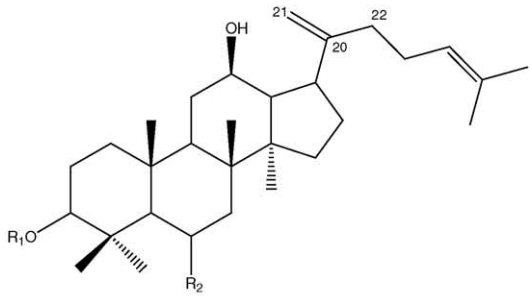
The pattern match processing method was developed with repeated ($n \geq 6$) replicate injections of the raw samples, as well as replicate injections of steamed (2 h) samples, to detect true differences between samples. The scan region of 10–78 min was chosen to include the region of



Saponins	R_1	R_2	R_3
Rb1	-Glc-Glc	-H	-Glc-Glc
Rc	-Glc-Glc	-H	-Glc-Ara(f)
Rd	-Glc-Glc	-H	-Glc
Re	-H	-O-Glc^2-Rha	-Glc
Rg1	-H	-O-Glc	-Glc
R1	-H	-O-Glc^2-Xyl	-Glc
20S Rg3	***-Glc-Glc***	***-H***	***-H***
20R Rg3	***-Glc-Glc***	***-H***	***-H***
20S Rh1	***-H***	***-O-Glc***	***-H***
20R Rh1	***-H***	***-O-Glc***	***-H***



Saponins	R_1	R_2
Rg5	***-Glc-Glc***	***-H***
Rh4	***-H***	***-O-Glc***



Saponins	R_1	R_2
Rk1	***-Glc-Glc***	***-H***
Rk3	***-H***	***-O-Glc***

Fig. 2. Structures of 14 saponins, including those saponins (in bold) that are characteristic for the steamed samples. Abbreviations: Glc, glucose; Ara(f), arabinose in furanose form; Rha, rhamnose.

interest and to exclude regions of void volume, injection artifacts at the beginning and the re-equilibration region at the end of the chromatogram. In the preliminary pattern matching method, the scan time was selected while the software automatically calculated the rest of the parameters (alignment interval, retention time search limit, peak width, detection threshold, response value, and percent peak height) for the pattern matching process. In the optimization routine, the autocalculated values can be further optimized to suit the particular set of samples and applications of the study.

The retention time search limit affects the alignment results. It was further optimized to ensure the retention time offset encompasses the largest possible retention time offsets between the chromatograms. Retention time search limit of 15 s specified the retention time range over which the reference chromatogram was offset in search for the best alignment. The detection threshold affects the interpretation of the results. It was set to determine which peak apices were detected and plotted on the chromatogram. Detection threshold was set at 350 for optimal number of detected peaks in this case and to prevent over-clustering of peak apex markers in the plots.

To objectively determine if a difference is genuine (not due to baseline noise), the individual standard deviations should be compared against a threshold. The threshold values do not affect the alignment or the standard deviation values. It only helps in the interpretation of the results. This standard deviation threshold is obtained from comparing similar chromatograms (replicate injections). This threshold is calculated automatically by the software from response value and percent peak height. These two parameters were further optimized to correspond to the highest standard deviation values that an interval could have when the compared samples were the same. It ensured that the standard deviations of the replicate injections (identical samples) were detected below the threshold line. Therefore, it is a kind of limit-of-detection test, which should be sensitive to detect true differences between different samples and not similar samples. The formula for the standard deviation threshold is:

$$\text{Standard deviation threshold} = (\text{response} + \text{percent peak height}/100) \times (\text{maximum peak height in the interval})$$

The response is the minimum value for the standard deviation threshold. The percent peak height adds a value to the response that is proportional to the peak height within the compared interval.

3.4. Chromatographic pattern matching results of raw and steamed P. notoginseng roots

Before comparing the raw and steamed samples, pairs ($n \geq 6$ pairs) of identical raw samples (replicate injections)

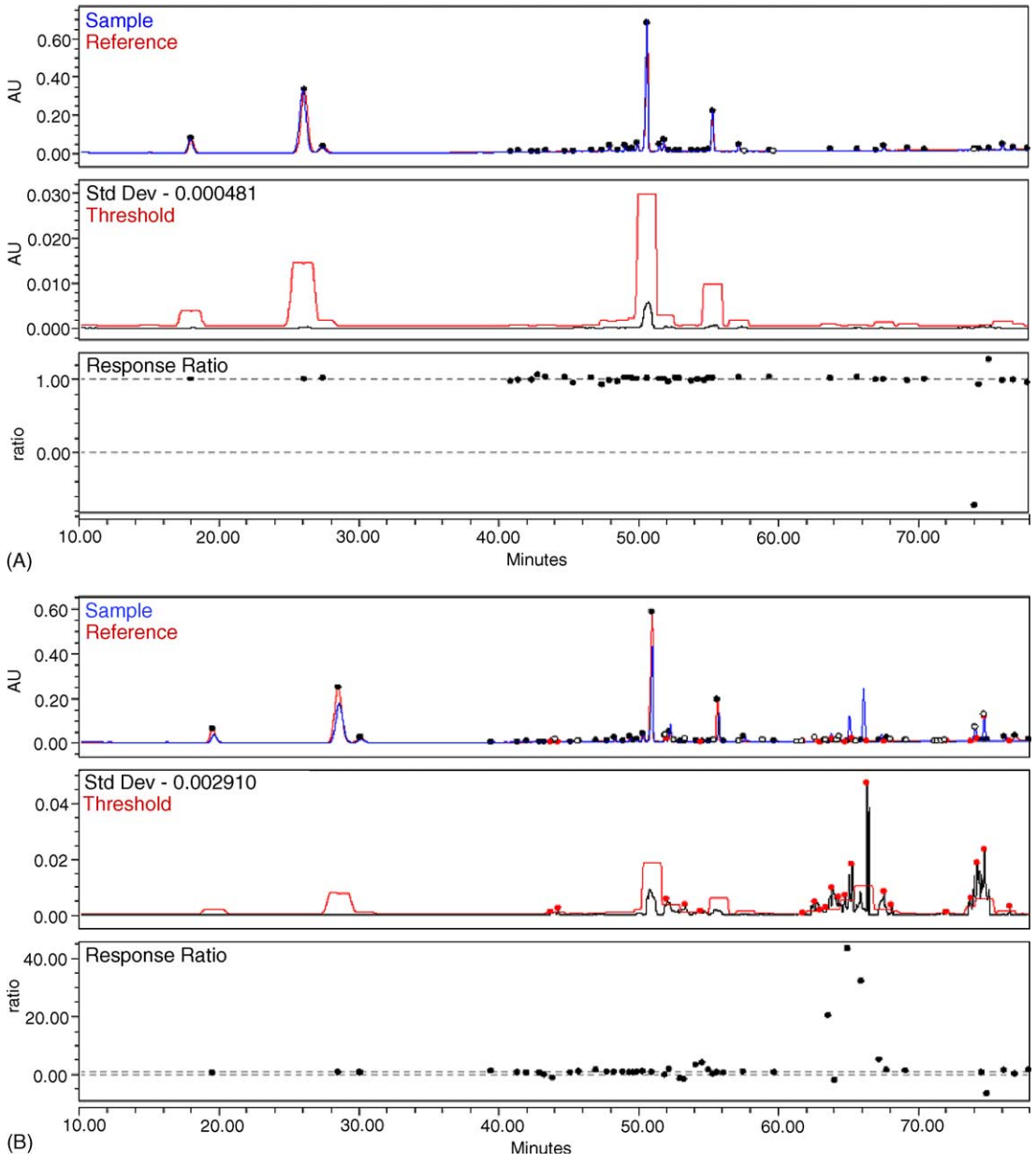


Fig. 3. Typical results from chromatographic pattern matching for (A) replicate injections of raw *P. notoginseng* herb and (B) raw and steamed (2 h) *P. notoginseng*. Each of the top plots shows an overlay of the chromatograms, with black markers on peak apices. Each of the middle plots shows their corresponding standard deviations for all points in the scan region. Each of the bottom plots shows response ratios (sample/reference) of all points within scan region.

were compared. Replicate injections of identical steamed samples ($n \geq 6$ pairs) were also performed and they gave similar results as replicate injections of raw samples. The typical pattern matching results of a pair of replicate injections of raw samples are shown in Fig. 3(A). The top plot shows the overlay of a pair of chromatograms in the scan region between 10 and 78 min. The middle plot shows their individual standard deviations of all points within the scan region. The standard deviation is a measure of the magnitude of point-to-point differences between the two chromatograms after alignment, as shown by the formula above. For similar samples (such as these replicate injections), the standard deviations were small and below the threshold line and they represented the background noise present. For replicate injections of raw *P. notoginseng* and replicate injections of steamed (2 h) *P. notoginseng*, the pattern match standard deviation values were found to be below 0.0006. Response ratios (sample/reference) of the peaks in both chromatograms were close to 1 for identical samples, as shown in the bottom plot. The response ratios measure the approximate ratio of concentrations within the compared intervals, if the alignment is good.

Using pattern matching analysis, the raw and steamed samples were successfully differentiated. The top plot of

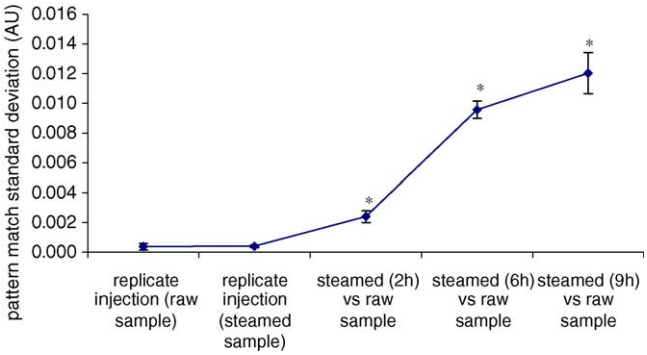


Fig. 4. Pattern match standard deviation values of replicate injections and *P. notoginseng* herbs that were steamed for 2, 6, and 9 h. Values were means ± S.D., $n \geq 6$. For the steamed samples, the pattern match standard deviation values were obtained from the pattern matching comparisons with the corresponding raw sample (before steaming). The asterisk (*) denotes statistically significant differences between the PMSD values of the steamed samples and replicate injections at $p < 0.05$.

Fig. 3(B) shows the typical overlaid chromatograms of raw and steamed (2 h) *P. notoginseng*. The main differences were in the region 63–76 min. The steamed form showed numerous peaks eluting out in this region. These peaks were not distinct in the chromatogram of the raw samples. In addition, the standard deviations increased above threshold line at various points in the scan region, especially in the region 63–76 min, as shown in the middle plot of Fig. 3(B). This indicated that the samples were not similar. Response ratio also increased to a value of 45 at one point (65th minute). The pattern match standard deviation value was also significantly higher than that obtained from replicate injections of samples. As the duration of steaming increased, the differences between the chromatograms also increased. This is reflected in the pattern match plot. The average pattern match standard deviation values increased with duration of steaming (Fig. 4). The values (Table 1) increased from 0.0024 to 0.0120 for 2 and 9 h of steaming, respectively. These were about 3–17 times the PMSD values for replicate injections of identical samples and the PMSD values were statistically higher ($p < 0.05$, Student's *t*-test) than that obtained for replicate injections. Therefore, these values gave useful indications of the similarity of the sample pairs and may serve as a similarity or match index.

3.5. Chromatographic pattern matching of raw and steamed P. notoginseng products

After establishing the method and differentiating the known raw and steamed samples, the chromatographic pattern matching method was then applied to the comparisons of 11 pairs of raw and steamed *P. notoginseng* proprietary herbal products. Their pattern match standard deviation values were summarized in Table 1. Eight pairs (pairs 1–8) of raw and steamed CPMs showed distinctive differences from the pattern match plots. The pattern match standard deviation values of these eight pairs ranged from 0.0009 to 0.0050. Values above 0.0008 were found to be statistically different ($p < 0.05$) from the replicate injections of identical samples. Interestingly, three raw and steamed pairs (pairs 9–11) were found to have similar chromatographic patterns from their pattern match plots. For pairs 9 and 10, the products labeled as 'steamed' were found to have chromatograms resembling a raw sample. Whereas, for pair 11, the product labeled as 'raw' have distinctive peaks in the region 63–76 min, resembling a steamed product. The pattern match standard deviation values for the three pairs were also close to that obtained for replicate injections. The standard deviations were below threshold line and their response ratios were close to 1, which were similar to the typical results of the replicate injections. This implies that these pairs of sample were similar, although they were labeled as 'raw' and 'steamed'. Thus, there may be possible mislabeling, insufficient steaming of the 'steamed' CPMs, or high temperatures during processing/harvesting may have changed the 'raw' CPMs. Our previous paper (Lau et al. 2003) reported the clear differentiation of chromatographic fingerprints of extracts of raw and steamed *P. notoginseng.* The current method in this paper provided quantitative comparisons of pairs of such products. Visual chromatographic comparisons of the extracts of three pairs out of 11 pairs of such products showed that the differentiation between the pairs was not clear and was inconsistent with the label claims. Indeed, upon quantitative comparisons of the chromatograms, we have demonstrated that the method is able to detect such discrepancies. The lack of quality control and standardizations of herbal products and samples is an important issue to address. In the preparation of raw samples, the raw root may have been subjected to excessively high temperatures during harvesting or drying, and this may have caused some chemical components to be changed. Thus, these raw products may have chromatographic profiles that resemble those of steamed products. On the other hand, if the steamed products are not steamed/processed sufficiently, they may yield profiles similar to those of the raw products. Currently, among the different manufacturers, there is no standardization of the processing conditions used to steam the samples. Such information is also not available and is kept confidential by the manufacturers. As quality control and standardization of the processing methods are currently not present, the wide variations of results for different pairs of CPMs are not unexpected.

From the values of the pattern match standard deviation, additional useful information regarding the degree in which the raw herb was steamed or changed by the process may also be obtained. Among the CPMs, pair 10 has the lowest PMSD value, while pair 5 has the highest value. The greater the differences, the greater the degree in which the herbs were steamed or the components were changed by the processing method, with respect to the corresponding raw sample. This is useful in standardization of the

steamed herb. A 'blind' test was also carried out whereby the identities of four pairs of raw and steamed products were not known to the analyst. Using pattern matching analysis, accurate identification of the raw and steamed products was obtained, confirming the usefulness of the method. This method may also be applied to identify herbs, check the quality or source by comparing an unknown with a known or standard herb. Using the pattern match standard deviation values, it can determine which samples are most similar to the given reference and provide the closest match.

4. Conclusion

The new HPLC method combined with chromatographic pattern matching analysis, allows rapid, simple, automated, and quantitative comparisons of complex chromatograms. The method is successful in providing quantitative differentiation of raw and steamed *P. notoginseng* roots and its products, and in rapidly detecting inconsistencies in the products (e.g. in the labeling, degree of steaming in "steamed" products). The degree in which the raw herb has been changed by the steaming process may also be deduced from the pattern matching analysis. This is also the first report for the detection of ginsenoside Rk3 in *P. notoginseng* and its isolation from this herb. Fourteen well-resolved saponins have been identified in a single HPLC chromatogram of an extract of *P. notoginseng*. This method can potentially differentiate herbs or complex samples based on their chromatographic profiles. It can also determine the similarity between samples. Therefore, it is a potentially useful tool in ensuring the quality and safety of herbal products.

Acknowledgements

The authors gratefully acknowledge the financial support from the National University of Singapore Academic Research Fund and a research scholarship (A.-J.L.). The authors will also like to thank Waters Asia Ltd (Singapore) for the help in this work and Professor J.H. Park (College of Pharmacy, Seoul National University, South Korea) for providing some of the ginsenoside standards.

References

[1] G.B. Mahady, H.H.S. Fong, N.R. Farnsworth, Botanical Dietary Supplements: Quality, Safety, and Efficacy, Swets and Zeitlinger, The Netherlands, 2001.
[2] B.T. Schaneberg, S. Crockett, E. Bedir, I.A. Khan, Phytochemistry 62 (2003) 911.
[3] F. Gong, Y.Z. Liang, Y.S. Fung, F.T. Chau, J. Chromatogr. A 1029 (2004) 173.
[4] Y.Y. Cheng, M.J. Chen, W.J. Welsh, J. Chem. Inf. Comput. Sci. 43 (2003) 1959.
[5] F. Gong, Y.Z. Liang, P.S. Xie, F.T. Chau, J. Chromatogr. A 1002 (2003) 25.
[6] Guidance for Industry, Botanical Drug Products, Draft Guidance, CDER/FDA, 2000.
[7] CPMP/CVMP, Note for guidance on specifications: test procedures and acceptance criterion for herbal drugs, herbal drug preparations, and herbal medicinal and herbal medicinal products, EMEA, 2000.
[8] W.J. Welsh, W. Lin, S.H. Tersigni, E. Collantes, R. Duta, M.S. Carey, W.L. Zielinski, J. Brower, J.A. Spencer, T.P. Layloff, Anal. Chem. 68 (1996) 3473.
[9] C.B. Cordella, J.S. Militao, M.C. Clement, D. Cabrol-Bass, J. Agric. Food Chem. 51 (2003) 3234.
[10] J. Quesada Granados, J.J. Merelo Guervos, M.J. Oliveras Lopez, J. Gonzalez Penalver, M. Olalla Herrera, R. Blanca Herrera, M.C. Lopez Martinez, J. Agric. Food Chem. 50 (2002) 1470.
[11] K.J. Johnson, B.W. Wright, K.H. Jarman, R.E. Synovec, J. Chromatogr. A 996 (2003) 141.
[12] G. Malmquist, R. Danielsson, J. Chromatogr. A 687 (1994) 71.
[13] R.J.O. Torgrip, M. Aberg, B. Karlberg, S.P. Jacobsson, J. Chemometrics 17 (2003) 573.
[14] Y.P. Zhu, Chinese Material Medica, Harwood Medicine Association, Australia, 1998.
[15] The State Pharmacopoeia Commission of PR China, Pharmacopoeia of the People's Republic of China, vol. 1, Chemical Industry Press, Beijing, 2000.
[16] State Administration of Traditional Chinese Medicine, Zhong Hua Ben Cao Jin Xuan Ben, vol. 1, Shanghai Science and Technology Publishers, Shanghai, 1996.
[17] A.J. Lau, S.O. Woo, H.L. Koh, J. Chromatogr. A 1011 (2003) 77.
[18] I.H. Park, N.Y. Kim, S.B. Han, J.M. Kim, S.W. Kwon, H.J. Kim, M.K. Park, J.H. Park, Arch. Pharm. Res. 25 (2002) 428.
[19] S.I. Kim, J.H. Park, J.H. Ryu, J.D. Park, Y.H. Lee, J.H. Park, T.H. Kim, J.M. Kim, N.I. Baek, Arch. Pharm. Res. 19 (1996) 551.
[20] R.W. Tang, H.Z. Li, J.T. Chen, D.Z. Wang, Y.N. He, C.R. Yang, Magn. Reson. Chem. 40 (2002) 483.
[21] V.U. Ahmad, A. Basha, Spectroscopic Data of Saponins: The Triterpenoid Glycosides, CRC Press, Boca Raton, FL, 2000.
[22] M.V. Gorenstein, US Pat. No. 5969228 (1999); http://www.uspto.gov/, accessed 2 July 2004.
[23] Waters Corporation. Empower chromatographic pattern matching software getting started guide, USA, 2002; http://www.waters.com/WatersDivision/pdfs/71500031903rA.pdf, accessed 2 July 2004.