Rapid HPLC Analytical Method Development for Herbal Medicine Formulae Based on Retention Rules Acquired from the Constituting Herbs

Hua YIN, Li ZOU, Yunjie SHENG, Xue BAI, Qiang LIU, and Binjun YAN†

College of Pharmaceutical Science, Zhejiang Chinese Medical University, No. 548, Binwen Road, Hangzhou, 310053, China

Herbal medicine (HM) formulae are the combinations of two or more types of constituting herbs. This study has proposed a novel approach to efficiently develop HPLC methods for HM formulae, which take advantage of the mutual retention rules between HM formulae and their constituting herbs. An HM formula composed of two herbs, Radix Salviae Miltiorrhizae and Rhizoma Chuanxiong, was taken as a case study. Based on design of experiments and stepwise multiple linear regression, models relating the analytical parameters to the chromatographic parameters were built (correlation coefficients >0.9870) for chemical compounds in the two herbs. These models representing the retention rules were utilized to predict the elution profile of the formula. The analytical parameters were numerically optimized to ensure adequate separation of the analytes. In validation experiments, satisfactory separations were achieved without any pre-experiments on the formula. The approach can significantly increase the HPLC method development efficiency for HM formulae.

Keywords HPLC method development, multivariate statistical modeling, quality by design, feedforward control, herbal medicine formula

(Received June 15, 2017; Accepted October 3, 2017; Published February 10, 2018)

An efficient approach to develop HPLC methods for compound matrices is needed. Since a compound matrix is the combination of its constituting ingredients, and HPLC methods for the individual ingredients have usually been developed before method development for the compound matrix, a question is raised: Are there any mutual retention rules between the compound matrix and its ingredients? If the answer is yes, then based on the retention rules acquired from the constituting ingredients, the impacts of the analytical parameters on the chromatographic separation of the compound matrix are predictable, which will enable rapid optimization of the analytical parameters for the compound matrix. When analyzing compound matrices and their constituting ingredients with the same HPLC analytical parameters, the same chemical compounds in these different samples have the same chromatographic parameters like retention time ($t_R$) and peak width. Therefore, the relationships between the analytical parameters and these chromatographic parameters can be defined as the mutual retention rules between the compound matrices and their constituting ingredients. By effectively acquiring and utilizing the retention rules, the HPLC method development efficiency can be increased.

In recent years, the analytical quality by design (AQbD) approach was proposed and has been successfully applied in the development of analytical methods. In the AQbD approach, the relationships between the analytical parameters and the quality of the analytical methods are systematically investigated using risk assessment and the design of experiments (DoE) methodologies, and are then quantitatively represented in the form of mathematical models using multivariate statistical modeling methods. The AQbD approach is introduced in this.
work to acquire retention rules, i.e., to construct models relating the analytical parameters to the chromatographic parameters.

AQbD originated from the concept of quality by design (QbD).\textsuperscript{18,19} According to the QbD initiative, pharmaceutical processes can be proactively adjusted within the design space, which was defined as "the multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide the assurance of quality".\textsuperscript{18} When the material attributes vary, the process parameters can be adjusted accordingly using a feedforward control strategy to ensure product quality.\textsuperscript{20–22} This work proposed that feedforward control can also be applied to the analytical process of compound matrices to efficiently develop HPLC methods. For example, from the three herbs of Radix Salviae Miltiorrhizae (RSM), Rhizoma Chuanxiong (RC) and Flos Carthami (FC), four different HM formulae for cardio-cerebrovascular diseases have been composed, namely, RSM-RC formula, RSM-FC formula, RC-FC formula and RSM-RC-FC formula. When developing three HPLC methods to determine the same compound (salvianolic acid B, SaB, a bioactive compound from RSM) in the three different HM formulae of RSM-RC, RSM-FC and RSM-RC-FC, different chemical compounds from RSM, RC and FC may coelute with SaB and obstruct the determination. The presence of different interferences in the three HM formulae can be regarded as variation in the material attributes. By utilizing the retention rules from RSM, RC and FC with the feedforward control strategy, the analytical parameters can be numerically optimized to achieve adequate chromatographic separation of SaB in each HM formula. Then, the three HPLC methods can be established with high efficiency (Fig. 1).

![Schematic diagram of the approach to systematically develop HPLC analytical methods for compound matrices based on the retention rules acquired from the constituting ingredients.](image)

**Fig. 1** Schematic diagram of the approach to systematically develop HPLC analytical methods for compound matrices based on the retention rules acquired from the constituting ingredients.

**Experimental**

**Materials and sample preparation**

HPLC-grade acetonitrile was purchased from Tedia Co., Inc. (Fairfield, OH, USA). Deionized water was produced with a UPH-III-5T water purification system (Chengdu Chaochun Technology Co., Ltd., Chengdu, China). Analytical-grade formic acid was purchased from Yixing Second Chemical Reagent Factory (Wuxi, China). Analytical-grade ethanol was purchased from Lingfeng Chemical Reagent Co., Ltd. (Shanghai, China). Standard substances of SaB, sodium danshensu, protocatechuic aldehyde, caffeic acid, rosmarinic acid (RA), lithospermic acid (LA), salvianolic acid A (SaA), ferulic acid (FA), senkyunolide I and ligustilide were purchased from Shanghai Winherb Medical Technology Co., Ltd. (Shanghai, China). RSM was purchased from the Affiliated HM clinic of Zhejiang Chinese Medical University (Hangzhou, China). RC was purchased from Huadong Medicine Co., Ltd. (Hangzhou, China).

The RSM sample and RC sample were prepared, respectively, as follows: 40 g of RSM or RC was cut into pieces and extracted with 400 mL of 50% (v/v) ethanol at reflux for 1 h. The extract solution was filtered, and the residue was further extracted at reflux with 400 mL of 50% (v/v) ethanol for 1 h. The two filtrates were combined and concentrated to 20 mL at 60°C with a vacuum evaporator. Then, 0.25 g of the concentrate was diluted to 100 mL with 50% (v/v) ethanol for HPLC analysis. The RSM-RC sample was prepared using 40 g each of RSM and RC following the same method.

**Selection of the key quality criteria for HPLC method development**

A key point concerning HPLC method development is finding
suitable analytical conditions to achieve adequate separations between all analytes and interferences within a reasonable time.

The resolution, R, is a commonly used indicator of chromatographic separation, calculated as $R = 2(t_{h,I} - t_{h,A})/(W_I + W_A)$, where $t_{h,A}$ and $t_{h,I}$ are the retention times of two adjacent peaks and $W_I$ and $W_A$ are their peak widths. However, for asymmetric peaks like tailing peaks, R is not an appropriate indicator of separation. Therefore, the beginning time ($t_{b}$) and the end time ($t_{e}$) of the peak, defined as the time points where the signal is equal to 5% of the peak height, were used in this work, and the separation criteria were as follows:

$$t_{b, A} < t_{b, I, B}$$  \hspace{1cm} (1)

$$t_{e, A} > t_{e, I, B}$$  \hspace{1cm} (2)

where $t_{b,A}$ and $t_{b,I}$ are the beginning time and the end time of the analyte peak, $t_{b,I,B}$ is the end time of the interference peak eluted in front of the analyte, and $t_{e,I,B}$ is the beginning time of the interference peak eluted behind the analyte. Equations (1) and (2) means that the interferences should either entirely elute before the analyte, or begin to elute after the analyte has entirely eluted. These two equations should be satisfied for all the analytes and interferences.

On the other hand, it is usually desirable for HPLC analysis to be completed in a short time. Since chromatographic analysis can be terminated soon after all of the analytes have been eluted, the largest $t_{e}$ of the analytes was taken as another key quality criterion, which should be minimized during HPLC method development.

**Design of experiments**

With the risk assessment tool of the Ishikawa diagram, the gradient, pH and flow rate of the mobile phase were identified as the key factors affecting the retention times and peak widths of the chemical compounds in RSM and RC, which could further affect the quality of the analytical method. A four-factor Box-Behnken design (Table 1) with five center points (DoE-25 to DoE-29) was performed, respectively, on the RSM sample and the RC sample to study their retention rules. In order to cover a wider range of mobile phase gradient setting, a two-stage linear gradient was used (Table 2), which was characterized by two analytical parameters, i.e., the time of gradient change ($T_C$) and the percentage of acetonitrile in the mobile phase ($P_A$) at gradient change from 0 to $T_C$, min, acetonitrile was increased linearly from 10% to $P_A$, and from $T_C$ to 80 min, acetonitrile was increased linearly from $P_A$ to 90%. With reasonable adjustment of $P_A$ and $T_C$, this kind of gradient on a C18 column can compose the analytical conditions suitable for a fair number of analytes. The gradient was all from 10 to 90% acetonitrile in all the DoE points, which can avoid the situation that some compounds elute in one DoE point but not elute in another DoE point. The third analytical parameter, $p_C$, was defined as the negative logarithm of the concentration (v/v) of the formic acid aqueous solution, which is related to the pH of the mobile phase B. When $p_C$ was 4.0, the concentration of formic acid was relatively low (0.01%) and the pH was 3.23. In contrast, at a small $p_C$ of 2.0 (i.e., a concentration of 1%), the pH was 2.23. Finally, the fourth analytical parameter, $F$, was the flow rate of the mobile phase.

Although the time for chromatographic separation was all fixed to 80 min in DoE (Table 2), the time of the optimized method for a specific determination would not be as long, because after all the analytes have been eluted, the remaining gradient does not affect the determination. This final stage of gradient can be replaced by a short column wash gradient (e.g., 90% acetonitrile) to elute the compounds remaining in the column. On the other hand, as shown below in the last paragraph of “Results and Discussion”, to further reduce the HPLC analysis time, the initial period of the gradient elution can be removed tentatively, if the mobile phase in this period has low eluting power to separate interferences from the analytes. When the initial period and/or final stage of the gradient elution is removed in developing the HPLC method for specific analytes, the two-stage gradient may become a single linear gradient, which is simpler and used more commonly.

**HPLC analysis**

The HPLC analyses were conducted on a Waters e2695 HPLC system equipped with a 2998 photodiode array (PDA) detector (Waters Corp., Milford, MA) using a C18 column (AkzoNobel Kromasil 100-5C18, 250 × 4.6 mm). Mobile phase A was acetonitrile, and mobile phase B was a formic acid aqueous solution with the formic acid concentrations shown in Table 1. The mobile phase gradient and the flow rate were set as shown in Tables 1 and 2. After 80 min of gradient elution, 5 min of column wash gradient (90% acetonitrile) was used to elute the compounds remaining in the column. The column temperature was set at 35°C, and the injection volume was 10 µL. UV spectra in the range of 250 – 400 nm with a resolution of 1.2 nm were collected by the PDA detector.

**Chromatographic peak matching**

Since the retention time of a compound changes under

---

**Table 1** Box-Behnken design of experiments

| Standard order | $P$, % | $T/\text{min}$ | $p_C$ | $F/\text{mL min}^{-1}$ |
|---------------|--------|----------------|-------|---------------------|
| DoE-1         | 30     | 30             | 3.0   | 1.0                 |
| DoE-2         | 50     | 30             | 3.0   | 1.0                 |
| DoE-3         | 30     | 60             | 3.0   | 1.0                 |
| DoE-4         | 50     | 60             | 3.0   | 1.0                 |
| DoE-5         | 40     | 45             | 2.0   | 0.8                 |
| DoE-6         | 40     | 45             | 4.0   | 0.8                 |
| DoE-7         | 40     | 45             | 2.0   | 1.2                 |
| DoE-8         | 40     | 45             | 4.0   | 1.2                 |
| DoE-9         | 30     | 45             | 3.0   | 0.8                 |
| DoE-10        | 50     | 45             | 3.0   | 0.8                 |
| DoE-11        | 30     | 45             | 3.0   | 1.2                 |
| DoE-12        | 50     | 45             | 3.0   | 1.2                 |
| DoE-13        | 40     | 30             | 2.0   | 1.0                 |
| DoE-14        | 40     | 60             | 2.0   | 1.0                 |
| DoE-15        | 40     | 30             | 4.0   | 1.0                 |
| DoE-16        | 40     | 60             | 4.0   | 1.0                 |
| DoE-17        | 30     | 45             | 2.0   | 1.0                 |
| DoE-18        | 50     | 45             | 4.0   | 1.0                 |
| DoE-19        | 30     | 45             | 4.0   | 1.0                 |
| DoE-20        | 50     | 45             | 4.0   | 1.0                 |
| DoE-21        | 40     | 30             | 3.0   | 0.8                 |
| DoE-22        | 40     | 60             | 3.0   | 0.8                 |
| DoE-23        | 40     | 30             | 3.0   | 1.2                 |
| DoE-24        | 40     | 60             | 3.0   | 1.2                 |
| DoE-25        | 40     | 45             | 3.0   | 1.0                 |
| DoE-26        | 40     | 45             | 3.0   | 1.0                 |
| DoE-27        | 40     | 45             | 3.0   | 1.0                 |
| DoE-28        | 40     | 45             | 3.0   | 1.0                 |
| DoE-29        | 40     | 45             | 3.0   | 1.0                 |

**Table 2** Two-stage linear gradient of the mobile phase

| Time/Min | Acetonitrile, % | Aqueous formic acid, % |
|----------|-----------------|------------------------|
| 0        | 10              | 90                     |
| 80       | 90              | 10                     |

---
different analytical conditions, peak matching was performed to identify the compound peaks in each chromatogram. The peaks of the known compounds were identified using their standard substances. For unknown compounds and compounds without standard substances, the chromatographic parameters were matched based on their peak areas and UV spectra. Specifically, two criteria were used to match peaks in different chromatograms: 1) the areas of the peaks should be close after being multiplied by the flow rate for normalization, because the same sample was used in the DoE and the concentration of the same compound was constant; 2) the correlation coefficient between the spectra of the peaks should be close to 1.

**Resolution of coeluting peaks**

In regions of coeluting peaks, the HPLC-PDA data \((X)\) were the superposition of the spectra of the coeluting compounds. Before peak matching, the resolution of the coeluting peaks was performed to acquire the spectrum as well as \(t_8\), \(t_9\), and \(t_{10}\) of each coeluting compound. With bilinearity, \(X (T \times J)\) of \(T\) time points and \(J\) wavelengths can be decomposed as \(X = CS + E\), where \(C (T \times N)\) is the elution profiles matrix of the \(N\) coeluting compounds, \(N (N \times J)\) is the spectra matrix of the \(N\) compounds, and \(E\) is the noise. One of the most popular methods, multivariate curve resolution-alternating least squares (MCR-ALS),\(^{23}\) was applied locally to the data regions of the coeluting peaks for resolution.

**Modeling of chromatographic parameters**

Models relating the analytical parameters to the chromatographic parameters were built in the form of a second-order polynomial model to mathematically represent the retention rules as follows:

\[
t = a_0 + a_1 t_8 + a_2 t_9 + a_3 pC_a + a_4 F + a_{11} P^2 + a_{12} T_8^2 + a_{13} P C_a^2 + a_{14} F^2 + a_{15} P T_8 + a_{16} P C_a + a_{17} P F + a_{18} T_8 pC_a + a_{19} t_9 F + a_{20} pC_a F,
\]

where \(t\) is the chromatographic parameter of \(t_8\), \(t_9\) or \(t_{10}\), and \(a_0 \ldots a_{20}\) are the estimated regression coefficients of the model. Independent models were built for each compound on the four analytical parameters \((P, T, pC_a, F)\) using stepwise multiple linear regression. The root-mean-square error of cross-validation (RMSECV) was calculated using leave-one-out cross-validation to estimate the uncertainties in each model. It is noted that the impacts of the analytical process uncertainties on the chromatographic parameters were also reflected by the RMSECVs.

**Optimization of the analytical parameters for HM formulae**

During the HPLC method development for HM formulae, an optimization model was built as Eqs. (4) – (7) to optimize the analytical parameters based on the feedforward control strategy:

\[
\text{minimize}_{P, T, pC_a, F} \max_{pC_a, F} \left( P, T, pC_a, F \right),
\]

s.t.

\[
t_{8,8} (P, T, pC_a, F) - \text{RMSECV}_{8,8} \geq t_{9,9} (P, T, pC_a, F) + \text{RMSECV}_{9,9},
\]

\[
t_{10,10} (P, T, pC_a, F) - \text{RMSECV}_{10,10} \geq t_{9,9} (P, T, pC_a, F) + \text{RMSECV}_{8,8},
\]

\[
P, T, pC_a, F \in \text{the ranges studied and modeled.}
\]

To minimize the HPLC analysis time, the objective function of Eq. (4) minimizes, the largest \(t_8\) of all the analytes, where \(t_8\) is a function of the four analytical parameters \((P, T, pC_a, F)\) as shown in Eq. (3). In consideration of the uncertainties in model prediction and to provide some extent of separation robustness to the analytical process uncertainties, Eqs. (5) and (6) were derived from Eqs. (1) and (2). In Eqs. (5) and (6), \(t_{8,8}, t_{8,9}, t_{9,9}, t_{9,10}, t_{10,10}\) are predicted by Eq. (3), and \(\text{RMSECV}_{8,8}, \text{RMSECV}_{8,9}, \text{RMSECV}_{9,9}, \text{RMSECV}_{9,10}, \text{RMSECV}_{10,10}\) are the RMSECVs of the models of \(t_{8,8}, t_{8,9}, t_{9,9}, t_{9,10}, t_{10,10}\), respectively. After decreasing \(t_8\) and increasing \(t_9\) by the RMSECVs, the peak ranges of the analytes and interferences are expanded in the constraints, which can ensure adequate chromatographic separation in a slightly biased prediction and compensate for the analytical process uncertainties. The constraints of Eqs. (5) and (6) should be satisfied for all analytes, and it is noted that the neighboring interferences eluted in front of and behind the analytes may be different compounds, depending on the analytical conditions and the compositions of the HM formulae. One constraint of Eq. (5) or (6) was set for one pair of analyte and interference. The number of these constraints set in the optimization model depended on the number of analytes and interferences considered. In the constraint of Eq. (7), the analytical parameters are limited to the ranges that have been studied and modeled to ensure reasonable predictions. This single-objective multi-constraints optimization model was solved using the grid search method.

**Software**

The Box-Behnken design was constructed using Design Expert 8.0 software (State-Ease Inc., Minneapolis, MN, USA). The MCR-ALS method was performed using the MCR-ALS toolbox 2.0.\(^{24}\) In-house programs were developed in Matlab 7.9 software (The Mathworks Inc., Natick, MA, USA) to perform stepwise multiple linear regression, cross-validation, and the optimization of the analytical parameters using the grid search method.

**Results and Discussion**

**Resolution of coeluting peaks and peak matching**

The DoE was conducted and 29 HPLC-PDA chromatograms of the RSM sample and RC sample were obtained respectively. Some of these chromatograms are shown in Fig. 2 with a wavelength of 286 nm. MCR-ALS was applied to the HPLC-PDA data to resolve coeluting peaks. In most cases, the coeluting peaks were slightly overlapped and MCR-ALS succeeded in acquiring the elution profiles and the spectra. However, in a few cases, the coeluting peaks were seriously overlapped, and the spectra of the coeluting compounds were too similar to acquire accurate elution profiles for each compound, which was similar to the situation reported in the literature.\(^{11}\) For these cases, the standard substances of the coeluting compounds were analyzed individually to acquire their elution profiles. It was found that in the elution profiles acquired by MCR-ALS, the biases in \(t_8\) were small and acceptable while the biases in \(t_9\) and \(t_{10}\) were relatively large, compared with the chromatograms of the standard solutions. Therefore, for coeluting compounds without standard substances, their \(t_8\) values from the MCR-ALS results were used for modeling.

Peak matching was conducted to identify specific compounds in each chromatogram. In total, 19 compounds in the RSM sample were matched, including SaB, sodium danshensu,
protocatechuic aldehyde, caffeic acid, RA, LA, SaA and 12 unknown compounds (UC-1, UC-2, ..., UC-12). In the RC sample, 29 compounds were matched, including FA, senkyunolide I, senkyunolide A, ligustilide, butylidene phthalide and 24 unknown compounds (UC-13, UC-14, ..., UC-36). The $t_R$ values of these compounds are listed in Tables S1 – S2 (Supporting Information).

Modeling of chromatographic parameters

The models of $t_R$, $t_B$ and $t_E$ were built using stepwise multiple linear regression, and the RMSECVs were calculated. All of the $t_R$ values were fitted successfully by the models, with correlation coefficients ($r$) greater than 0.9870 and RMSECVs less than 0.9 min. The regression coefficients, $r$ and RMSECVs of the $t_R$ models are listed in Tables S3 – S4 (Supporting Information). As shown in Fig. S1 (Supporting Information) with 6 compounds for examples, the $t_R$ values calculated from the models were very close to the experimental values acquired in the DoE. The impacts of the analytical parameters could be described quite well using second-order polynomial models, and transformation of $t_B$ was not required in this work.

As discussed above, the biases in the $t_B$ and $t_E$ of seriously overlapped peaks were relatively large. However, in this work, it was found that except for the tailing peaks like SaB (Fig. S2a in Supporting Information), the peak widths (roughly estimated as $t_E - t_B$) were all approximately 0.66 min (shown in Fig. S2b in Supporting Information taking senkyunolide I as an example), and the impacts of the analytical parameters on these peak widths were small and negligible, compared with the RMSECVs of the $t_R$ models. Therefore, from a practical point of view, the $t_B$ and $t_E$ models were necessary only for compounds whose peak widths were significantly affected by the analytical parameters. For other compounds, it was reasonable to estimate $t_R$ and $t_E$ as $t_R \pm 0.33$ min. When $t_B$ and $t_E$ were computed as $t_B \pm 0.33$ min, the RMSECV$_B$ and the RMSECV$_E$ (i.e., the RMSECVs of $t_B$ and $t_E$) were set equal to the RMSECV of $t_R$.

For the tailing peak of SaB, the $t_B$ and $t_E$ models were built, and Fig. S3 (Supporting Information) shows that the $t_B$ and $t_E$ values calculated from the models were very close to their experimental values in the DoE.

The retention rules acquired

With the regression models built, the quantitative relationships between the analytical parameters and the chromatographic parameters were constructed, which were the retention rules that could be used to predict the elution profile in the analysis of the RSM-RC formula. FA and SaB are discussed here as examples, while the impacts of the analytical parameters on the other compounds can be inferred from the regression coefficients listed in Tables S3 – S4 (Supporting Information). As shown in Fig. 3a, as $T_r$ decreased and $P$ increased, the $t_E$ of FA decreased significantly, because the percentage of acetonitrile in the mobile phase increased faster for the elution. These impacts of $T_r$ and $P$ are directly shown by comparing the $t_E$ of FA in Figs. 2a and 2b. In Fig. 3b, the $t_R$ of FA decreased with increasing flow rate, and the impact of $pC_a$ on the $t_R$ was small. As shown in Figs. 3c and 3d, the impacts of the analytical parameters on the $t_R$ of SaB were similar to those of FA. With the decrease of $P$ and increase of $T_r$, $t_R$ of SaB increased significantly, which can be directly seen by comparing Figs. 2c and 2d. The peak width of SaB was estimated as $t_E - t_B$, and the impacts of $pC_a$ and $T_r$ on the width are shown in Fig. 4. A higher
concentration of formic acid in the mobile phase inhibited the
tailing of SaB and decreased the peak width, which is directly
shown by comparing Figs. 2c and 2e. The impacts of the
analytical parameters on the $t_B$ and $t_E$ of SaB are shown in
Fig. 5. The models of the $t_B$ and $t_E$ were similar (Figs. 3c and
3d and Figs. 5a and 5b), while the models of the $t_R$ and $t_0$ were
quite different (Figs. 3c and 3d and Figs. 5c and 5d). The
tailing of the SaB peak at higher pH was serious, which
produced a large difference between $t_E$ and $t_R$.

**Optimization and validation of the analytical parameters for the
HM formula**

According to the analytes and interferences in the RSM-RC
formula, the specific optimization models were established
using the retention rules acquired from the herbs of RSM and
RC to optimize the analytical parameters.

When developing the HPLC method for the determination of
FA in the RSM-RC formula, the maximum absorption
wavelength of FA (321 nm) was used to ensure the sensitivity,
and the compounds that had absorption at 321 nm and had $t_R$
close to FA were taken as the interferences in Eqs. (5) and (6).
These interferences included UC-2, UC-3, UC-4 and RA from
RSM (Fig. 2c), in addition to UC-19 and UC-20 from RC
(Fig. 2a). Then, these related compounds were substituted into
the general optimization model of Eqs. (4) – (7), and a specific
optimization model (Eqs. (S1) – (S8) in Supporting Information)
was established. Equation (5) was used for the interferences
eluted in front of the analyte and Eq. (6) was used for the
interferences eluted behind the analyte. The reason for the form
of Eqs. (S3) and (S4) was that under different analytical
parameters, UC-3 and UC-4 would be possibly eluted either in
front or behind FA. The $t_R$, $t_0$ and $t_0$ of FA and its interferences
under different analytical parameters were all predicted by
Eq. (3), and finally, the analytical parameters were optimized to
$P = 41.7$, $T_s = 30.0$, $p_Ca = 2.0$ and $F = 1.19$ by the specific
optimization model. The predicted $t_R$ of FA was 13.2 min. In
the validation experiment, FA was well separated from the
interferences, and its actual $t_R$ was 13.8 min.

In the current Chinese Pharmacopoeia, single-compound
determination is still the most commonly used quality control
pattern, although it has been recognized that multi-compound determination is more suitable for the quality control of HMs. The analytical methods usually need to be redeveloped when upgrading quality standards from single-compound determination to multi-compound determination. When developing the HPLC method for the simultaneous determination of FA and SaB in the RSM-RC formula, senkyunolide I was easy to coelute with SaB, which is directly shown from the same $t_R$ values of SaB and senkyunolide I in Figs. 2a and 2c. With the analytical parameters in DoE-3, although the $t_R$ values of SaB and senkyunolide I were different (Figs. 2b and 2d), the large $t_R$ of SaB would result in a long HPLC analysis time. Using the proposed approach, the analytical parameters can be efficiently adjusted to achieve adequate separation of SaB within a relatively short analysis time. In the specific optimization model (Eqs. (S9) – (S25) in Supporting Information), SaB was also taken as the analyte in Eqs. (5) and (6), in addition to the analyte of FA in the previous case. The compounds had absorption at 286 nm (the maximum absorption wavelength of SaB) and $t_R$ close to SaB were taken as the interferences for SaB in Eqs. (5) and (6). These interferences included RA, LA, UC-5, SaA from RSM (Fig. 2c), and UC-19, UC-20, senkyunolide I, UC-21, UC-22 from RC (Fig. 2a). The $t_R$ of SaB was minimized by Eq. (4) because SaB had the larger $t_R$ of the two analytes. The analytical parameters were optimized to $P = 31.2$, $T_s = 45.3$, $p_{C_a} = 2.0$ and $F = 0.87$. Compared with the optimized analytical parameters of $P = 41.7$ and $T_s = 30.0$ for the determination of FA alone, the percentage of acetonitrile in the mobile phase was increased more slowly to ensure adequate separation of SaB. The predicted $t_R$ values of FA and SaB were 22.1 and 35.4 min, respectively. HPLC analyses were conducted on the RSM-RC sample in triplicate to validate the optimized analytical method. As shown in Fig. 6, FA and SaB were well separated from the interferences, especially senkyunolide I. The average $t_R$ of FA was 22.0 min, and the average $t_R$ of SaB was 35.0 min, very close to the predicted values. It should be noted that no pre-experiments were conducted on the RSM-RC sample to optimize the analytical parameters. The satisfactory separation result in Fig. 6 was achieved based only on the retention rules from RSM and RC. Besides, as shown in Fig. 6, since the analytes of FA and SaB have been eluted in the first 36 min of gradient elution, the remaining gradient can be removed because it does not affect the determination. Then, the two-stage gradient can be easily changed to a simple single linear gradient (10 – 26.8% acetonitrile in 0 – 36 min). The results demonstrated that the elution profiles of HM formulae could be predicted...
using the retention rules, and the proposed approach can significantly increase the HPLC methods development efficiency.

If the impacts of the analytical process uncertainties on $t_R$ were smaller and more accurate models were built, with smaller RMSECVs, the constraints of Eqs. (5) and (6) would be satisfied more easily, and $t_R$ could be further minimized. On the other hand, to reduce the HPLC analysis time, the 0 – 15 min gradient elution (10 – 17% acetonitrile) was tentatively removed, when it was assumed that the mobile phase in this period had low eluting power to separate the interferences from the analytes. The HPLC analysis was then conducted using the modified linear gradient, which was 17 – 31.2% acetonitrile in 0 – 30.3 min. In the result, FA and SaB were still well separated from the interferences, and the $t_R$ values of FA and SaB were 12.4 and 20.5 min, respectively. Based on these analytical parameters, other method validation experiments, including precision, linearity and limit of quantification, can be conducted to complete the method development.

Conclusions

In this work, a novel HPLC method development approach for compound matrices was proposed and successfully applied to the case study of an HM formula. The advantages of this approach include: 1) the data from the development of the analytical methods for the individual herbs can be effectively utilized; 2) the HPLC method development efficiency for HM formulae can be increased significantly; 3) the analytical methods can be efficiently adjusted for the determination of more analytes. Due to the diverse combinations of herbs in HM formulae, the number of frequently used formulae has far exceeded the number of their constituting herbs. The "herbs to formulate" approach proposed in this study provides a path to systematically develop HPLC methods for HM formulae with high efficiency. Moreover, this approach can also be widely applied to the development of other chromatographic methods, such as GC and LC-MS, in the future.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 81703462), the Public Welfare Technology Application Research Project of the Science Technology Department of Zhejiang Province (No. 2014C33216), the Research Project (No. 2015ZR04) and the Open Project (No. Yao2016008) of Zhejiang Chinese Medical University.

Supporting Information

The Supporting Information (Tables S1 - S4, Figs. S1 - S3, Eqs. (S1) - (S25)) is available free of charge on the Web at http://www.jsac.or.jp/analsci/.

References

1. D. C. Woollard, A. Bensch, H. Indyk, and A. McMahon, *Food Chem.*, 2016, 197, 457.
2. J. Li, R. Pawar, A. J. Krynitsky, and G. O. Noonan, *Planta Med.*, 2015, 81, 947.
3. L. Wang and H. Qu, *Anal. Bioanal. Chem.*, 2016, 408, 2133.
4. J. Xu, Q. Mao, H. Shen, L. Zhu, S. Li, and R. Yan, *J. Chromatogr. A*, 2013, 1304, 154.
5. J. Liang, W. Wu, G. Sun, D. Wang, J. Hou, W. Yang, B. Jiang, X. Liu, and D. Guo, *J. Chromatogr. A*, 2013, 1294, 58.
6. H. Chen, X. Chen, Q. Han, J. Wu, D. Tang, Q. Du, X. Yin, and D. Yang, *Anal. Bioanal. Chem.*, 2012, 404, 1851.
7. Z. Yan, X. Yang, J. Wu, H. Su, C. Chen, and Y. Chen, *Anal. Chim. Acta*, 2011, 691, 110.
8. H. Liao, Q. Li, R. Liu, J. Liu, and K. Bi, *Anal. Sci.*, 2014, 30, 1157.
9. M. Taraji, P. R. Haddad, R. I. J. Amos, M. Talebi, R. Szucs, J. W. Dolan, and C. A. Pohl, *Anal. Chem.*, 2017, 89, 1870.
10. B. Debrus, P. Lebrun, A. Ceccato, G. Caliari, E. Rozet, I. Nistor, R. Oprean, F. J. Ruperez, C. Barbaz, B. Boulanger, and P. Hubert, *Anal. Chim. Acta*, 2011, 691, 33.
11. B. Debrus, P. Lebrun, J. M. Kindenge, F. Lecomte, A. Ceccato, G. Caliari, J. M. T. Mbay, B. Boulanger, R. D. Marini, E. Rozet, and P. Hubert, *J. Chromatogr. A*, 2011, 1218, 5205.
12. P. J. Borman, M. J. Chatfield, I. Damjanov, and P. Jackson, *Anal. Chim. Acta*, 2011, 703, 101.
13. E. Rozet, P. Lebrun, B. Debrus, B. Boulanger, and P. Hubert, *TrAC, Trends Anal. Chem.*, 2013, 42, 157.
14. S. Furlanetto, S. Orlandini, B. Pasquini, M. Del Bubba, and S. Pinzauti, *Anal. Chim. Acta*, 2013, 802, 113.
15. C. Hubert, P. Lebrun, S. Houari, E. Ziemons, E. Rozet, and P. Hubert, *J. Pharm. Biomed. Anal.*, 2014, 88, 401.
16. S. Furlanetto, S. Orlandini, B. Pasquini, C. Caprini, P. Mura, and S. Pinzauti, *Anal. Bioanal. Chem.*, 2015, 407, 7637.
17. A. Tumpa, T. Miladinovic, T. Rakic, A. Stajic, and B. Jancic-Stojanovic, *Anal. Lett.*, 2016, 49, 445.
18. ICH Harmonised Tripartite Guideline: Pharmaceutical Development Q8(R2), Compiled by International Conference on Harmonisation, http://www. ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/ Quality/Q8_R1/Step4/Q8_R2_Guideline.pdf.
19. L. X. Yu, G. Amidon, M. A. Khan, S. W. Hoag, J. Polli, G. K. Raju, and J. Woodcock, *AAPS Journal*, 2014, 16, 771.
20. A. Burggraeve, A. F. T. Silva, T. Van Den Kerkhof, M. Hellings, C. Vervaet, J. Paul Remon, Y. Vander Heyden, and T. De Beer, *Talanta*, 2012, 100, 293.
21. K. Muteki, V. Swaminathan, S. S. Sekulic, and G. L. Reid, *AAPS PharmSciTech*, 2011, 12, 1324.
22. B. Yan, Y. Li, Z. Guo, and H. Qu, *Phytochem. Anal.*, 2014, 25, 59.
23. D. C. Montgomery, “Design and Analysis of Experiments”, 6th ed., 2004, Wiley, New York.
24. J. Jaumot, A. de Juan, and R. Tauler, *Chemom. Intell. Lab. Syst.*, 2015, 140, 1.
25. Chinese Pharmacopoeia Commission, “Chinese Pharmacopoeia”, 2015, China Medical Science Press, Beijing.
26. Y. Yang, H. Wang, J. Yang, A. H. Brantner, A. D. Lower-Nedza, N. Si, J. Song, B. Bai, H. Zhao, and B. Bian, *J. Chromatogr. A*, 2013, 1321, 88.
27. H. Wei, L. Sun, Z. Tai, S. Gao, W. Xu, and W. Chen, *Anal. Chim. Acta*, 2010, 662, 97.
28. Y. F. He, Q. Li, and K. S. Bi, *J. Sep. Sci.*, 2015, 38, 1090.
29. X. C. Gong, Y. Zhang, H. L. Chen, T. Chen, J. Y. Pan, X. Y. Wang, and H. B. Qu, *Anal. Methods*, 2016, 8, 2282.