

## Developing an Analytical Method

In the presence of hydrogen peroxide,  $\text{H}_2\text{O}_2$ , and sulfuric acid,  $\text{H}_2\text{SO}_4$ , a solution containing vanadium ions forms a reddish-brown color. Although the exact chemistry of the reaction is uncertain, it allows for a simple qualitative “spot test” for vanadium—the formation of a reddish-brown color upon adding several drops of  $\text{H}_2\text{O}_2$  and  $\text{H}_2\text{SO}_4$  to a sample is a positive test for vanadium.

A spot test provides nothing more than a simple binary response: **YES**, the sample contains vanadium, or **NO**, the sample does not contain vanadium (at least at a concentration we can detect). Suppose we wish to adapt this qualitative test into a more quantitative method of analysis, one that allows us to report the concentration of vanadium in a sample. How might we accomplish this?

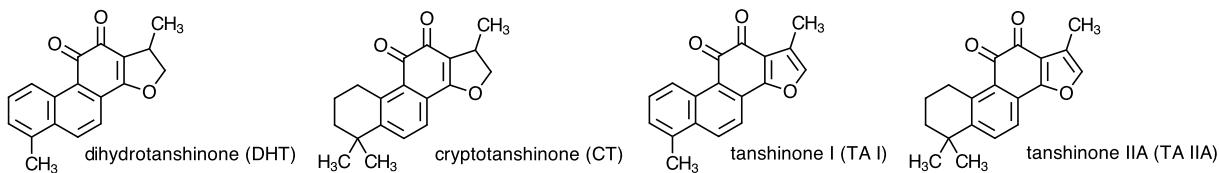
Given the reddish-brown color of a positive test, we might choose the solution’s absorbance at a wavelength of 450 nm as the analytical signal. In addition to the concentration of vanadium, the intensity of the solution’s color—and thus its absorbance—also depends on the amounts of  $\text{H}_2\text{O}_2$  and  $\text{H}_2\text{SO}_4$  added; in particular, a large excess of hydrogen peroxide decreases absorbance as the solution’s color changes from a reddish-brown color to a yellowish color. As well, we must ensure that the development of color is reproducible, and, of course, we want the method to yield accurate and precise results. We also need to determine if the method is susceptible to interferences and determine the smallest concentration of vanadium we can report with confidence. Finally, we want a method sufficiently rugged that different analysts will obtain similar results when analyzing the same sample. We call this process of optimizing and verifying a procedure method development.

This case study introduces you to method development within the context of the analysis of a medicinal plant using a combination of a microwave extraction to isolate the analytes from the plant’s roots and HPLC with UV detection to separate the analytes and to determine their concentrations. Interspersed within the case study’s narrative are a series of investigations, each of which asks you to stop and consider one or more important issues. Some of these investigations include data for you to analyze; you can copy and past this link (<http://bit.ly/YYgWL2>) into your browser to access interactive on-line versions of the data.

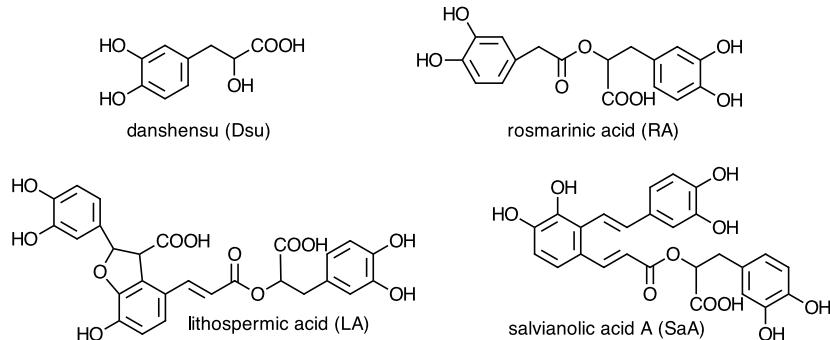
## Part I. Context of Analytical Problem

The dried root of *Salvia miltiorrhiza*—also known as red sage, Chinese sage, or Danshen, where “dan” and “shen” are Chinese for “red-colored” and “tonic herb,” respectively—is a traditional Chinese herbal medicine used to treat a variety of cardiovascular and cerebrovascular diseases, presumably because of its ability to prevent the formation of blood clots and its ability to dilate blood vessels.<sup>1</sup> Danshen is widely available throughout China, and is available, although to a lesser extent, in Europe and in the United States. The drug Dantonic®, a formulation that includes Danshen, is approved in 26 countries for the treatment of and prevention of angina; it currently is in phase III testing for use in the United States.<sup>2</sup>

As with any medicinal plant, the chemical composition of Danshen is complex with more than 70 constituent compounds identified in the literature. Early studies of Danshen’s chemical composition focused on lipophilic molecules, the four most important examples of which are:



Danshen also contains hydrophilic constituents, the four main examples of which are:



**Investigation 1.** What does it mean to characterize a molecule as hydrophilic or as lipophilic? How do they differ in terms of their chemical or physical properties?<sup>3</sup> Are there structural differences between these two groups of molecules that you can use to classify them as hydrophilic or as lipophilic? Consider the molecules below, both minor constituents of Danshen, and classify each molecule as lipophilic or hydrophilic.



<sup>1</sup> For a review of Danshen’s medicinal properties and uses, see “Danshen: An Overview of Its Chemistry, Pharmacology, Pharmacokinetics, and Clinical Uses,” the full reference for which is Zhou, L.; Zuo, Z.; Chow, M. S. S. *J. Clin. Pharmacol.* **2005**, *45*, 1345-1359 (DOI:10.1177/0091270005282630).

<sup>2</sup> You can view details regarding the phase III trial at <http://clinicaltrials.gov/show/NCT01659580>; the estimated completion date for the study is December 2015.

<sup>3</sup> A useful resource for exploring the chemical and physical properties of molecules is the Royal Society of Chemistry’s ChemSpider (<http://www.chemspider.com>) a free database that provides access to the properties of over 30 million compounds.

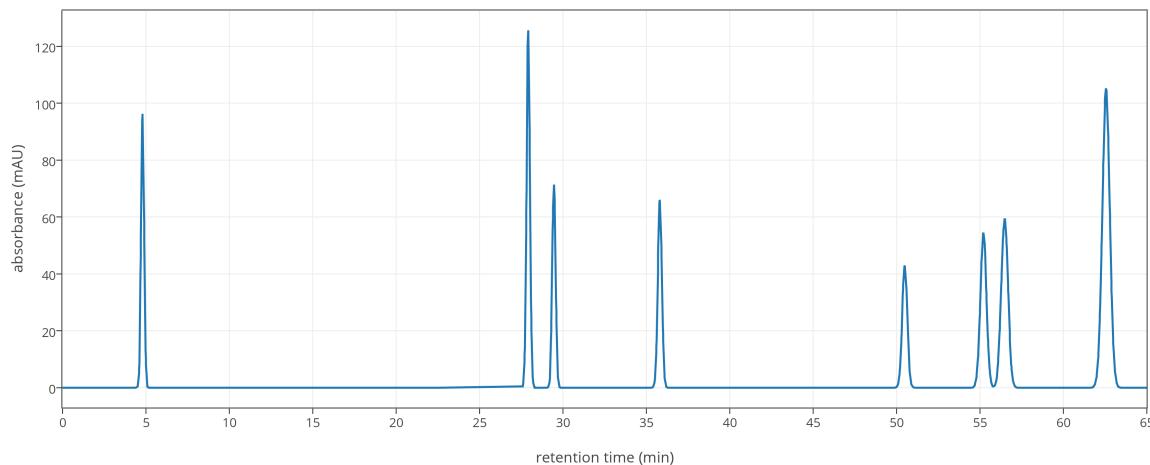
## Part II. Separating and Analyzing Mixtures Using HPLC with UV Detection

The analysis of a complex mixture usually is carried out using some form of chromatography, which allows us to separate the mixture's components prior to their individual detection. The sample is injected into a mobile phase that moves through a column containing a stationary phase. A component of the mixture that interacts more strongly with the stationary phase takes longer to pass through the column and reach the detector, eluting at a later time than a component that interacts less strongly with the stationary phase. The resulting chromatogram consists of a series of peaks, each characterized by a retention time ( $t_r$ ) and a peak height (or peak area).<sup>4</sup>

**Investigation 2.** For this study we will use a reverse-phase HPLC equipped with a UV detector to monitor absorbance. What is a reverse-phase separation and how is it different from a normal-phase separation? How does the choice between a reverse-phase separation and a normal-phase separation affect the order in which analytes elute from an HPLC?

Figure 1 shows a reverse-phase HPLC chromatogram for a standard mixture of the eight main components of Danshen. The column contains a non-polar C18 stationary phase. The mobile phase is a gradient of acetonitrile and an aqueous solution of 0.04% (v/v) phosphoric acid, beginning at 15% acetonitrile and ending at 80% acetonitrile, by volume. The flow rate is 0.80 mL/min. The elution order is danshensu, rosmarinic acid, lithospermic acid, salvianolic acid A, dihydrotanshinone, cryptotanshinone, tanshinone I, and tanshinone IIA.<sup>5</sup>

Figure 1: HPLC of Standard Dashen Sample



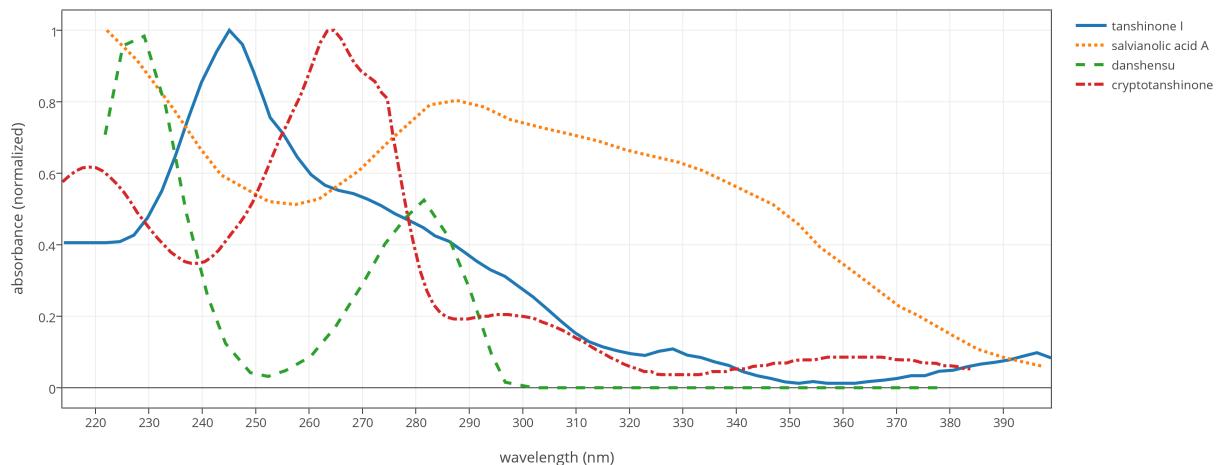
**Investigation 3.** Using the data in Figure 1 determine each analyte's retention time. Based on your answers to Investigation 1 and Investigation 2, does the relative order of elution order make sense? Why or why not?

<sup>4</sup> You can read more about chromatographic separations in general, and HPLC more specifically, in Chapter 12 of *Analytical Chemistry 2.0* (<http://bit.ly/1r3wJoz>).

<sup>5</sup> The data sets in this exercise are based loosely on work described in the paper “Simultaneous extraction of hydrosoluble phenolic acids and liposoluble tanshinones from *Salvia miltiorrhiza* radix by an optimized microwave-assisted extraction method,” the full reference for which is Fang, X.; Wang, J.; Zhang, S.; Zhao, Q.; Zheng, Z.; and Song, Z. *Sep. Purif. Technol.* **2012**, *86*, 149-156 (DOI:10.1016/j.seppur.2011.10.039). Although some data in this exercise are drawn directly from or extrapolated from data in the original paper, other data are drawn from additional sources or generated artificially. The original paper also includes data for the extraction and analysis of salvianolic acid B; because its concentration in Danshen is an order of magnitude greater than Danshen’s other constituents, it complicates the presentation of data and is not included in this exercise.

The chromatogram in Figure 1 was recorded using a UV detector. Figure 2 provides representative UV spectra from 220 nm to 380 nm for four of Danshen's constituents; you may assume these spectra are representative of Danshen's other hydrophilic and lipophilic compounds. Each spectrum is normalized so that its maximum absorbance is 1.00.

Figure 2: UV Spectra for Selected Constituents of Danshen



**Investigation 4.** Based on Figure 2, are there features in these UV spectra that distinguish Danshen's hydrophilic compounds from its lipophilic compounds? What wavelength should we choose if our interest is the hydrophilic compounds only? What wavelength should we choose if our interest is the lipophilic compounds only? What is the best wavelength for detecting all of Danshen's constituents?

Although separating the analytes from each other is essential to the analysis, our ultimate goal is to determine each analyte's concentration in samples of Danshen. The height of each peak in a chromatogram is proportional to the corresponding analyte's concentration in the sample as injected.

**Investigation 5.** For a UV detector, what is the expected relationship between peak height and the analyte's concentration in  $\mu\text{g/mL}$ ? For the results in Figure 1, can you assume the analyte with the smallest peak height is present at the lowest concentration? Why or why not?

The standard sample for the chromatogram in Figure 1 was prepared by diluting 1.00 mL of a stock standard solution to 10.00 mL in a volumetric flask. Table 1 details the stock standard's preparation.

**TABLE 1. PREPARATION OF STOCK STANDARD SOLUTION**

analyte	mg diluted to 10.00 mL	analyte	mg diluted to 10.00 mL
danshensu	6.00	dihydrotanshinone	1.51
rosmarinic acid	14.31	cryptotanshinone	2.89
lithospermic acid	13.31	tanshinone I	3.72
salvianolic acid A	4.17	tanshinone IIA	7.17

**Investigation 6.** Calculate the concentration, in  $\mu\text{g/mL}$ , for each analyte in the standard sample whose chromatogram is shown in Figure 1. Using this standard sample as a single-point external standard, calculate the proportionality constant for each analyte that relates its absorbance to its concentration in  $\mu\text{g/mL}$ . Do your results support your answer to Investigation 5? Why or why not?

### Part III. Extracting Analytes From Samples

The chromatogram in Figure 1 was obtained using samples of the eight analytes purchased from commercial sources. Because the analytes are available in pure form, there was no need to complete an extraction prior to injecting the standard sample into the HPLC; however, to analyze samples of Danshen, we first must extract the analytes from its roots using a suitable solvent.

**Investigation 7.** Brewing coffee is nothing more than a simple solvent extraction, which makes it a useful and a familiar model for considering how a solvent extraction works. There are a variety of methods for brewing coffee that differ in how the solvent and the coffee are brought together. Investigate at least five of the following methods for preparing coffee: Turkish, French Press, AeroPress, Chemex, Pour Over, Stovetop, Vacuum Pot, Espresso, and Cold Brew. In what ways are these methods similar to each other and in what ways are they different from each other? What variables in the extraction process are most important in terms of their ability to extract caffeine, essential oils, and fragrances from coffee?

The most common method for extracting analytes from a natural material—such as the roots, stems, and leaves of a medicinal plant—is to place a powdered sample in a suitable solvent and allow it to steep for 60 min at or near the solvent’s boiling point. After filtering, the solid residue is extracted a second time and the two extracts combined to give a final sample.<sup>6</sup>

**Investigation 8.** Why might a combination of high temperature, a lengthy extraction time, and the need for two extractions be undesirable when working with a medicinal plant such as Danshen?

Microwave-assisted solvent extractions are a promising method for addressing the limitations of a traditional solvent extraction because they use shorter extraction times and use smaller volumes of solvent. In this case study you will develop a method for the quantitative analysis in Danshen of the four lipophilic and the four hydrophilic compounds identified earlier. For this case study we will use a microwave-assisted solvent extraction that takes advantage of a microwave oven as a source of thermal energy.<sup>7</sup>

**Investigation 9.** What variables might we choose to control if we want to maximize the microwave extraction of Danshen’s constituent compounds? For each variable you identify, predict how a change in the variable’s value will affect the ability to extract from Danshen a hydrophilic compound, such as rosmarinic acid, and a lipophilic compound, such as tanshinone I.

<sup>6</sup> For a review of methods used for the quantitative analysis of Danshen, including different methods for extracting its active constituents, see “Advancement in Analysis of *Salviae miltiorrhiza* Radix et Rhizoma (Danshen),” the full reference for which is Li, Y.-G.; Song, L.; Liu, M.; Hu, Z.-B.; Wang, Z.-T. *J. Chromatogr. A*. **2009**, 1216, 1941-1953 (DOI: 10.1016/j.chroma.2008.12.032).

<sup>7</sup> For additional information on microwave extractions, see “Analytical-scale microwave-assisted extraction,” the full reference for which is Eskilsson, C. P.; Björklund, E. J. *J. Chromatogr. A* **2000**, 902, 227–250 (DOI:10.1016/S0021-9673(00)00921-3), and “Standardizing the World with Microwaves,” the full reference for which is Erickson, B. *Anal. Chem.* **1998**, 70, 467A–471A (DOI:10.1021/ac981908z). The microwave ovens used for solvent extractions essentially operate in the same manner as microwave ovens found in the home, although they are designed to allow more control over the microwave’s settings, and to handle better the harsher chemical environment found in a laboratory.

## Part IV. Selecting the Solvent, Temperature, and Microwave Power

There are a variety of methods to extract analytes from a solid sample, but the general principles are the same for most methods: find a suitable solvent and determine the experimental conditions—such as time, temperature, and solvent-to-solid ratio—that allow the solvent to extract completely the analytes from the sample. In this section of the case study you will select a solvent and optimize the extraction temperature and the microwave's power; in Part V you will optimize the solvent-to-solid ratio and the extraction time.

Because it is difficult to optimize simultaneously five different variables—in the lingo of method development we call a variable a factor, and we call the factor's value its level—you will complete three one-factor-at-a-time optimizations to identify a solvent, and to determine the temperature and the microwave power that maximizes the extraction of Danshen's constituents.<sup>8</sup>

In a one-factor-at-a-time optimization, the level for one factor is varied over a range of values while holding constant the levels of other factors. Each factor is optimized in turn, a process we repeat, if necessary, over multiple cycles until we find the set of factor levels that gives the best response; we call this set of factor levels and its response the system's global optimum.

**Investigation 10.** A one-factor-at-a-time optimization is an effective and an efficient algorithm when the factors behave independently, and an effective, although not necessarily an efficient, algorithm when the factors are dependent. What does it mean to say that two factors are independent or dependent? What does it mean to say that an optimization is efficient or effective? Why do dependent factors generally require that we optimize each factor more than once? Although the choice of solvent, temperature, and microwave power are dependent factors, for this case study you will optimize each factor once only. Explain why. For the analysis in this case study, is the order in which these three factors are optimized important? Why or why not?

*Selecting a Solvent.* To evaluate possible solvents, a 3.00-g sample of powdered Danshen is placed in a 100-mL flask and soaked for 20 min at room temperature using 60.0 mL of a suitable solvent. Following the initial soaking, the sample is transferred to a microwave oven and extracted for 5.00 min at a temperature of 60°C and a microwave power of 600 W.

**Investigation 11.** For the choice of solvent, consider ethanol, methanol, and water, as well as mixtures of water with ethanol or methanol, and predict how effective each is at extracting hydrophilic or lipophilic compounds. Why is a non-polar solvent, such as hexane, not a useful option for a microwave extraction? What limits, if any, might the choice of solvent place on the choice of temperature or microwave power?

The chromatograms obtained using methanol, ethanol, and water as the solvent are shown in Figures 3–5, respectively.

**Investigation 12.** Consider the data in Figures 3–5 and explain any trends you see in the relative extraction efficiencies of these three solvents. Are your results consistent with your predictions from Investigation 11? Why or why not? Which solvent is the best choice if you are interested in analyzing hydrophilic analytes only? Which solvent is the best choice if you are interested in analyzing lipophilic analytes only? Which solvent is the best choice if you are interested in analyzing both hydrophilic and lipophilic analytes?

<sup>8</sup> You can read more about optimization strategies in general, and one-factor-at-a-time optimizations more specifically, in Chapter 14 of *Analytical Chemistry 2.0* (<http://bit.ly/1r3wJoz>).

Figure 3: HPLC of Danshen Extracted with Methanol

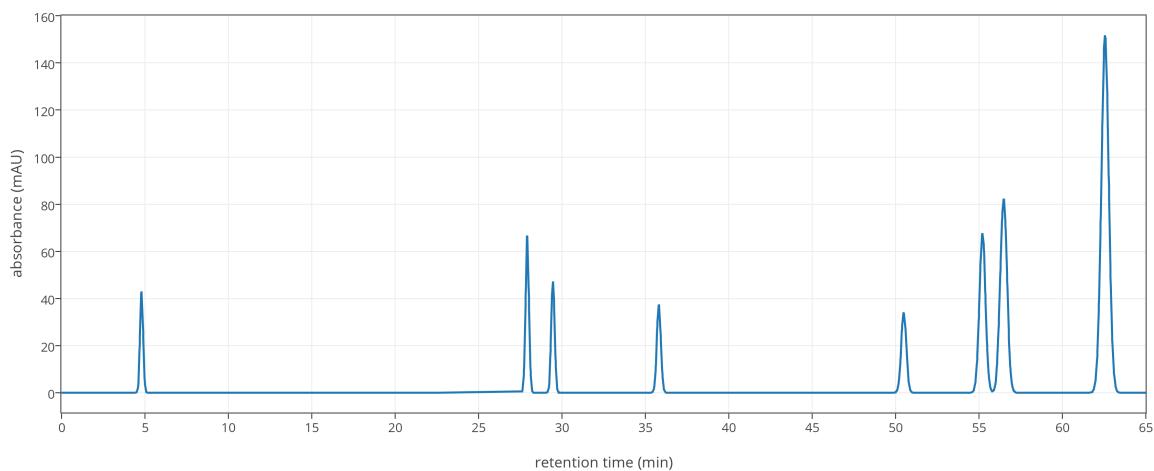


Figure 4: HPLC of Danshen Extracted with Ethanol

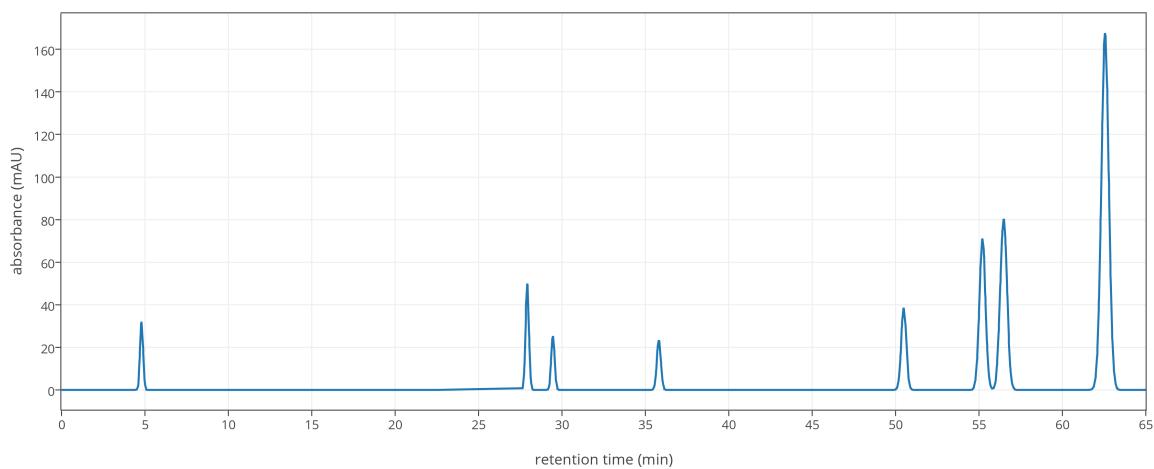
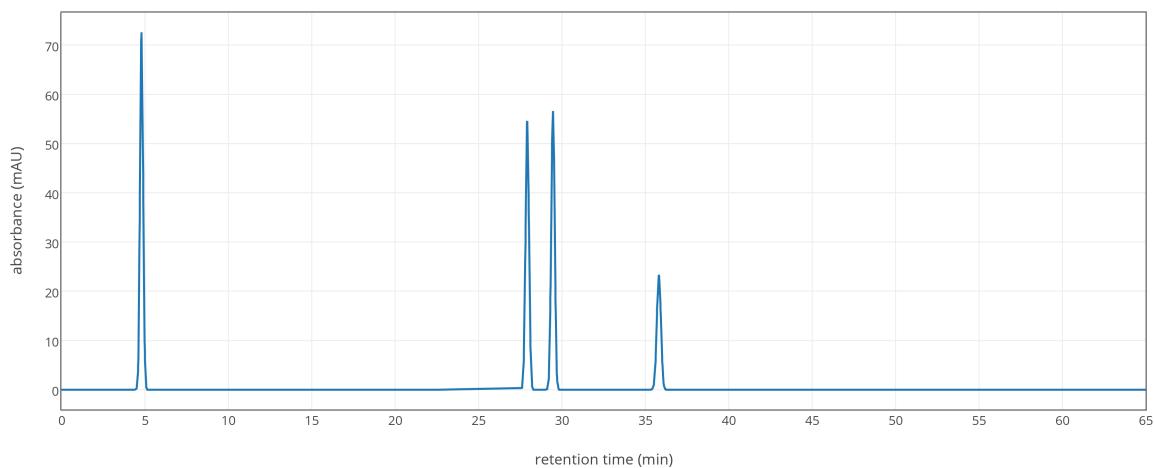


Figure 5: HPLC of Danshen Extracted with Water

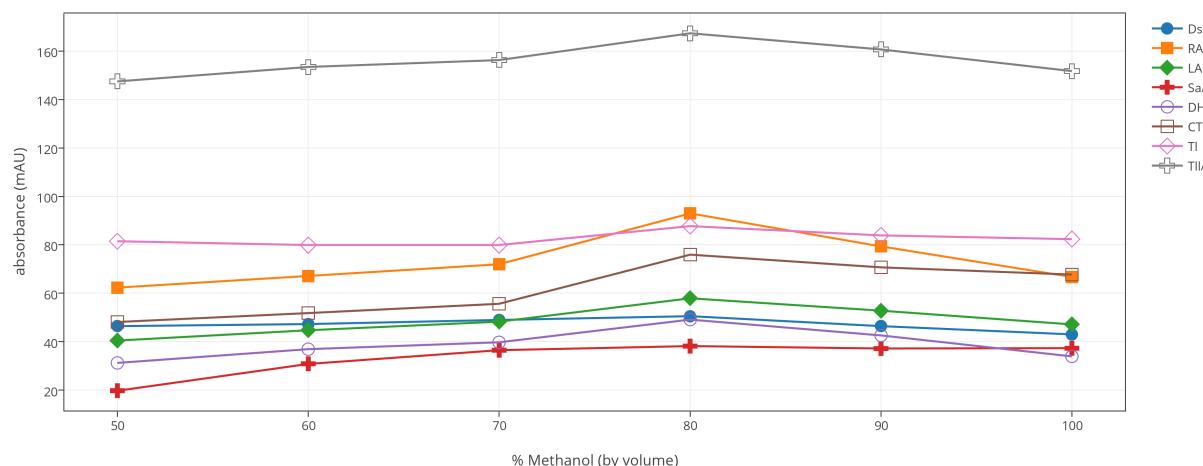


The temperature used to extract the samples in Figures 3–5 is limited by the boiling point of methanol, which has the lowest boiling point of the three solvents. Mixtures of methanol and water allow for higher boiling points, so it is worth exploring mixtures of these solvents. In addition, water and methanol have complimentary properties as solvents for microwave extractions: water is better than methanol at absorbing microwave radiation, but methanol is more efficient than water at converting absorbed microwave energy into heat.

**Investigation 13.** Propose a set of experiments that will effectively and efficiently allow you to determine the optimum mixture of methanol and water to use for this extraction. What range of methanol/water mixtures will you explore? How many samples will you run? Explain the reasons for the range of mixtures and the number of samples you selected. In describing the solvent mixtures, report values as percent methanol by volume (e.g. 55% methanol by volume).

Figure 6 shows results for the extraction of Danshen using a range of methanol-water mixtures from 50% methanol to 100% methanol by volume. Each extraction maintains the conditions used for the data in Figures 3–5: a solvent-to-solid ratio of 60.0 mL of solvent and 3.00 g of Danshen, an extraction time of 5.00 min, an extraction temperature of 60°C, and a microwave power of 600 W.

Figure 6: Effect of %Methanol:%Water (v/v) on Extraction of Danshen



**Investigation 14.** Consider the data in Figure 6 and explain any trends you see in the relative extraction efficiencies using different mixtures of methanol and water. What is the optimum mixture of methanol and water for extracting samples of Danshen? Are your results consistent with your predictions from Investigation 11 and the data from Investigation 12? Why or why not?

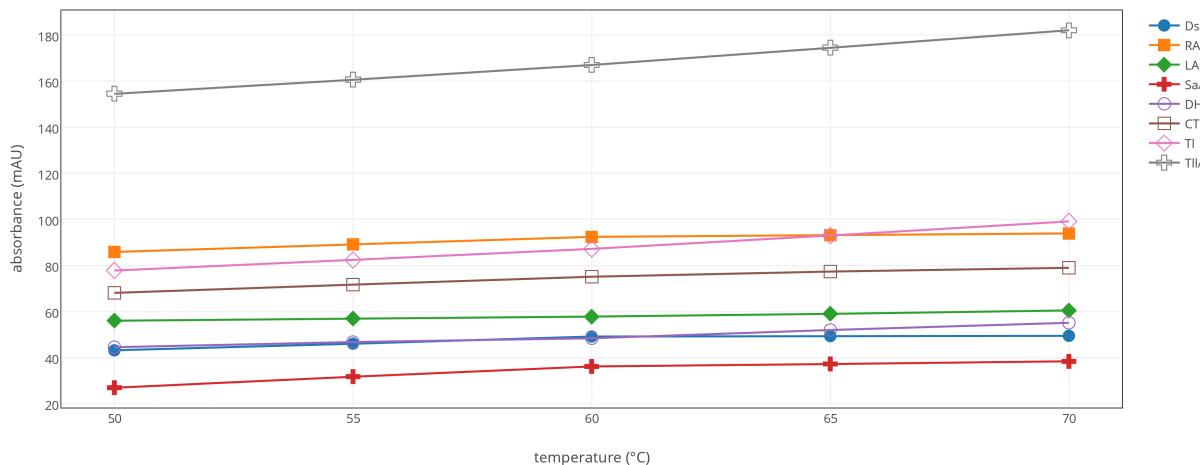
*Selecting a Temperature.* In general, extraction efficiency improves when using a higher temperature, although an excessively high temperature may decompose the sample and destroy some or all of its constituent compounds.

**Investigation 15.** Propose a set of experiments that will effectively and efficiently allow you to optimize the extraction temperature using the solvent selected in Investigation 14. What range of temperatures will you explore? How many samples will you run? Explain the reasons for the range of temperatures and the number of samples you selected.

Figure 7 shows results for the extraction of Danshen using a solvent of 80% methanol and 20% water (by volume) for temperatures from 50°C to 70°C. Each extraction maintains the remaining con-

ditions used for the data in Figures 3–5: a solvent-to-solid ratio of 60.0 mL of solvent and 3.00 g of Danshen, an extraction time of 5.00 min, and a microwave power of 600 W.

Figure 7: Effect of Temperature on Extraction of Danshen



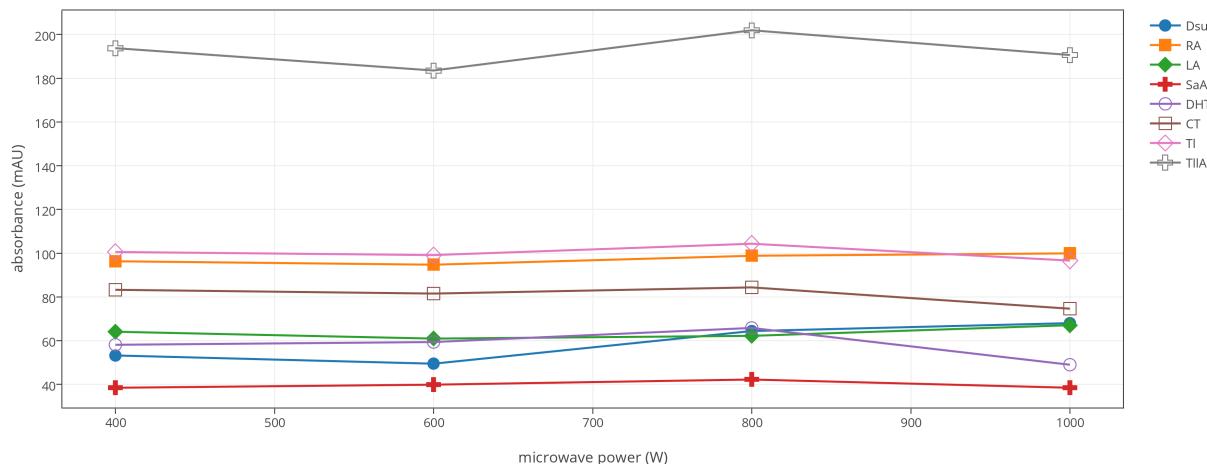
**Investigation 16.** Consider the data in Figure 7 and explain any trends you see in the relative extraction efficiencies as a function of temperature. What is the optimum temperature for extracting samples of Danshen? Are your results consistent with your expectations? Why or why not?

*Selecting the Microwave Power.* In a microwave extraction, temperature is controlled by applying short pulses of microwave radiation. The length of the applied pulse depends on the microwave's power, with a greater microwave power requiring shorter pulses to maintain the temperature. An increase in microwave power can improve extraction efficiency by increasing the breakdown of plant tissue, or it can decrease extraction efficiency by causing localized overheating of samples.

**Investigation 17.** Propose a set of experiments that will effectively and efficiently allow you to optimize the microwave power using the solvent and temperature selected in Investigation 16. What range of powers will you explore given that the microwave's power is adjustable between the limits of 0 W and 1000 W? How many samples will you run? Explain the reasons for the range of microwave powers and the number of samples you selected.

Figure 8 shows results for the extraction of Danshen using a solvent of 80% methanol and 20% water (by volume), a temperature of 70°C, and a range of microwave powers from 400 W to 1000 W. Each extraction maintains the remaining conditions used for the data in Figures 3–5: a solvent-to-solid ratio of 60.0 mL of solvent and 3.00 g of Danshen, and an extraction time of 5.00 min.

Figure 8: Effect of Microwave Power on Extraction of Danshen



**Investigation 18.** Consider the data in Figure 8 and explain any trends you see in the relative extraction efficiencies as a function of the microwave's power. What is the optimum power for extracting samples of Danshen using a solvent that is 80% methanol and 20% water by volume and an extraction temperature of 70°C?

## Part V. Optimizing the Solvent-to-Solid Ratio and the Extraction Time

Before continuing, let's review our progress in developing a method for extracting hydrophilic and lipophilic compounds from Danshen. In Part IV we completed a series of one-factor-at-a-time optimizations to determine the optimum solvent (80% methanol and 20% water, by volume), extraction temperature ( $70^{\circ}\text{C}$ ), and microwave power (800 W). For each of these optimizations we maintained a constant ratio of solvent-to-solid (60.0 mL of solvent and 3.00 g of Danshen) and a constant extraction time (5.00 min).

Now, in Part V, we turn our attention to optimizing the final two factors. First, however, we need to consider more carefully how we report the result of an extraction. In optimizing an extraction our goal is find a set of conditions that allow us to extract, or recover, all the analyte. Because we do not know how much analyte is in a sample, we seek, instead, to find the set of conditions that will recover the greatest amount of analyte, with results reported as mg analyte/g sample.

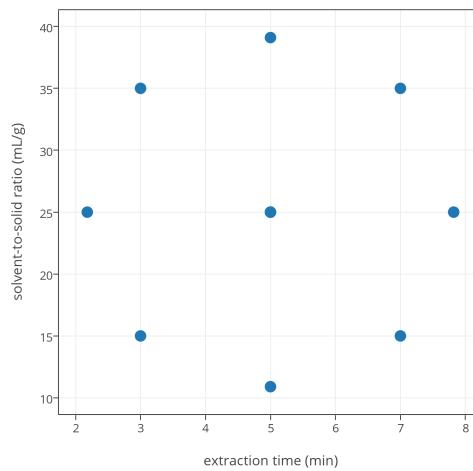
**Investigation 19.** When optimizing the choice of solvent, temperature, and microwave power, we used absorbance values taken directly from the HPLC analysis (see Figures 3–8) without first converting them into extraction yields reported in mg analyte/g sample. Why is it possible to use absorbance values for the optimizations in Part IV? Can you use absorbance values when optimizing the solvent-to-solid ratio or the extraction time? Why or why not? Using the optimum conditions from Figure 8 and your results from Investigation 6, report the extraction yield for each analyte as mg analyte/g sample.

Extraction time and the solvent-to-solid ratio are examples of dependent factors that may interact with each other in interesting and unpredictable ways. Although we can optimize both factors through a series of one-factor-at-a-time optimizations, a more efficient approach is to optimize them simultaneously using the experimental design shown in Figure 9. This experimental design, which is called a central-composite design, is efficient because it uses a small number of experiments—nine in this case, although replication of the center point is common—to explore a range of levels for each factor.

**Investigation 20.** We can divide the points in a central-composite design into three groups: a set of points that allow us to explore the effect on the extraction yield of extraction time only; a set of points that allow us to explore the effect on the extraction yield of the solvent-to-solid ratio only; and a set of points that allow us to explore the effect on the extraction yield of the interaction between extraction time and the solvent-to-solid ratio. Explain how each of these is accomplished in this experimental design.

Table 2 provides extraction yields for danshensu using the central-composite design in Figure 9. Note that the design's central point is run five times—which provides us with a measure of the reproducibility of extractions—and that the other points are run one time each.

Figure 9: Central-Composite Design



**TABLE 2. EXTRACTION YIELD FOR DANSHENSU USING A CENTRAL-COMPOSITE DESIGN**

extraction time (min)	solvent-to-solid ratio (mL/g)	extraction yield (mg/g)	extraction time (min)	solvent-to-solid ratio (mL/g)	extraction yield (mg/g)
5.00	10.9	0.721	5.00	25.0	0.785
5.00	25.0	0.790	5.00	39.1	0.784
3.00	15.0	0.743	7.00	35.0	0.805
2.18	25.0	0.742	5.00	25.0	0.801
3.00	35.0	0.754	5.00	25.0	0.773
5.00	25.0	0.813	7.82	25.0	0.820
7.00	15.0	0.785			

**Investigation 21.** Identify the five trials at the center of central-composite design and, for these trials, calculate the extraction yield's mean, standard deviation, relative standard deviation, variance, and 95% confidence interval about the mean.<sup>9</sup> What is the statistical meaning for each of these values? Transfer to Figure 9 the extraction yield for each experiment, using the mean extraction yield for the design's center point. What conclusions can you reach regarding the effect on danshensu's extraction yield of extraction time and solvent-to-solid ratio? Estimate the optimum conditions for maximizing danshensu's extraction yield and explain your reasoning?

Although the results in Table 2 are instructive in helping us understand how the extraction time and the solvent-to-solid ratio affect danshensu's extraction yield, a more quantitative model will provide us with a better ability to predict its extraction yield for any combination of factor levels. We can build an empirical model for danshensu's extraction yield using a second-order polynomial equation of the general form

$$EY = \beta_0 + \beta_a A + \beta_b B + \beta_{aa} A^2 + \beta_{bb} B^2 + \beta_{ab} AB$$

where  $EY$  is the extraction yield,  $A$  is the extraction time,  $B$  is the solvent-to-solid ratio, and  $\beta_0$ ,  $\beta_a$ ,  $\beta_b$ ,  $\beta_{aa}$ ,  $\beta_{bb}$ , and  $\beta_{ab}$  are the model's coefficients.

**Investigation 22.** What does it mean to describe a model as empirical instead of theoretical? What are the advantages and the disadvantages of using an empirical model? What is the significance for this empirical model of the coefficients  $\beta_0$ ,  $\beta_a$ ,  $\beta_b$ ,  $\beta_{aa}$ ,  $\beta_{bb}$ , and  $\beta_{ab}$ ? How does an empirical model that includes the coefficients  $\beta_{aa}$  and  $\beta_{bb}$  differ from a model that does not include these coefficients?

We can fit this empirical model to the data in Table 2 using a linear regression analysis.<sup>10</sup> The resulting empirical model of

$$EY = 0.575 + 0.0225A + 0.00905B - 0.00125A^2 - 0.000165B^2 + 0.000100AB$$

is significant at  $p = 0.0057$  with  $\beta_0$  significant at  $p < 0.001$ ,  $\beta_a$  significant at  $p < 0.01$ , and  $\beta_b$  and  $\beta_{bb}$  significant at  $p < 0.05$ .

<sup>9</sup> You can read more about characterizing data using means, standard deviations, variances, and confidence intervals in Chapter 4 of *Analytical Chemistry 2.0* (<http://bit.ly/1r3wJoz>).

<sup>10</sup> You can read more about linear regression in Chapter 5 of *Analytical Chemistry 2.0* (<http://bit.ly/1r3wJoz>). Although the context in this reference is fitting a straight-line to data with a single factor, the general approach, but not the specific equations, applies to fitting a full second-order polynomial to data with two factors. For a more detailed discussion of central-composite designs and linear regression, see Myers, R. H.; Montgomery, D. C. *Response Surface Methodology*, Wiley Series in Probability and Statistics, Wiley-Interscience:New York, 2002 or Brereton, R. G. *Chemometrics: Data Analysis for the Laboratory and Chemical Plant*, Wiley:Chichester, England, 2003.

**Investigation 23.** What does it mean to say that the regression analysis is significant at  $p = 0.0057$ ? Do the results of this regression analysis, as expressed in the model's coefficients, agree with your results from Investigation 21? Why or why not? What is the meaning of the intercept in this model and how does it affect your understanding of the empirical model's validity? Use the full regression model to calculate danshensu's predicted extraction yields for the central-composite design in Table 2. Organize your results in a table with columns for the factor levels, the experimental extraction yields, and the predicted extraction yields. Add a column showing the difference between the experimental extraction yields and predicted extraction yields. Calculate the mean, standard deviation, and the 95% confidence interval for these difference values and comment on your results.

The regression equation above describes the empirical model of danshensu's extraction yield for extraction times in the range 2.18–7.82 min and for solvent-to-solid ratios in the range 10.9–39.1 mL/g. Nevertheless, it is difficult to look at the equation and predict the extraction time and the solvent-to-solid ratio that maximizes danshensu's extraction yield; it is difficult, as well, to look at the regression equation and determine how sensitive is the optimum extraction yield to a small change in extraction time or solvent-to-solid ratio.

The factor levels giving the optimum extraction yield and the sensitivity of the extraction yield to a small change in factor levels are easier to visualize if we display the results as a three-dimensional plot with extraction yield on the  $z$ -axis and extraction time and the solvent-to-solid ratio on the  $x$ -axis and the  $y$ -axis, respectively. Figure 10 is one such a plot, which overlays a contour map of equivalent extraction yields on a heatmap that displays extraction yields using a variation in color. We call this type of plot a response surface.

**Investigation 24.** Does Figure 10 agree with your results from Investigations 21 and 23? Why or why not? Estimate the optimum conditions for maximizing danshensu's extraction yield and explain your reasoning. How sensitive is the optimum extraction yield to a small change in extraction time? How sensitive is the optimum extraction yield to a small change in the solvent-to-solid ratio?

Figures 11–15 show response surfaces for lithospermic acid, salvianolic acid A, cryptotanshinone, tanshinone I, and tanshinone IIA, each based on a regression analysis of data similar to that in Table 2 for danshensu. The regression models for rosmarinic acid and dihydrotanshinone are not significant, although the extraction of rosmarinic acid increases slightly for larger solvent-to-solid ratios and the extraction of dihydrotanshinone decreases slightly at longer extraction times; we will assume, however, that their extraction yield are independent of the extraction time and the solvent-to-solid ratio, with values of 2.317 mg/g for rosmarinic acid and 0.424 mg/g for dihydrotanshinone.

**Investigation 25.** Using Figures 11–15, determine the optimum extraction time and solvent-to-solid ratio for lithospermic acid, salvianolic acid A, cryptotanshinone, tanshinone I, and tanshinone IIA. How sensitive is the extraction of each analyte to a small change in the optimum extraction time and in the optimum solvent-to-solid ratio? Considering your responses here and to Investigation 24, are there combinations of extraction times and solvent-to-solid ratios that will optimize the extraction yield for all six of these analytes?

Figure 10: Response Surface for Extraction of Danshensu

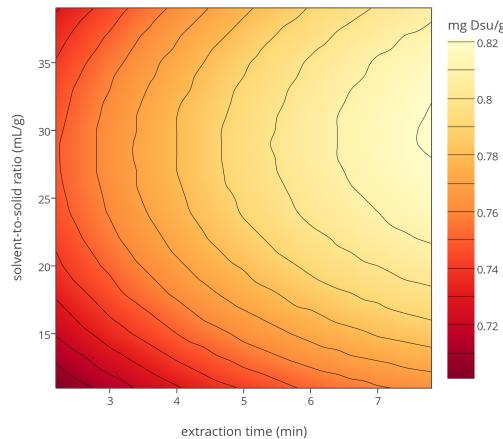


Figure 11: Response Surface for Extraction of Lithospermic Acid

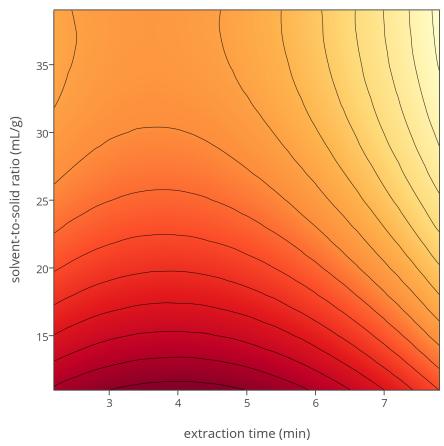


Figure 12: Response Surface for Extraction of Salvinolic Acid A

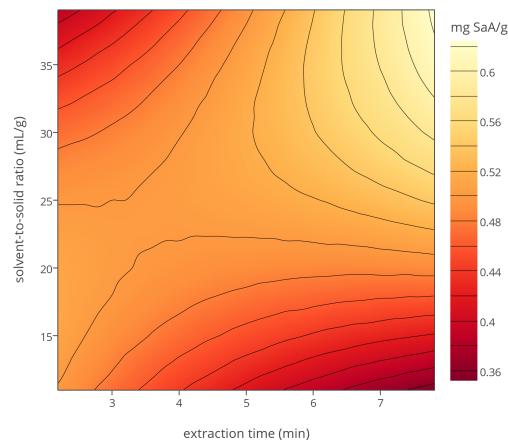


Figure 13: Response Surface for Extraction of Cryptotanshinone

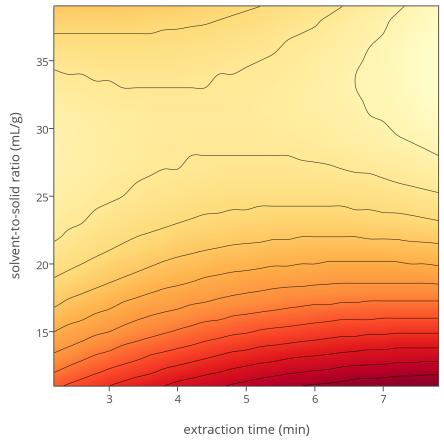


Figure 14: Response Surface for Extraction of Tanshinone I

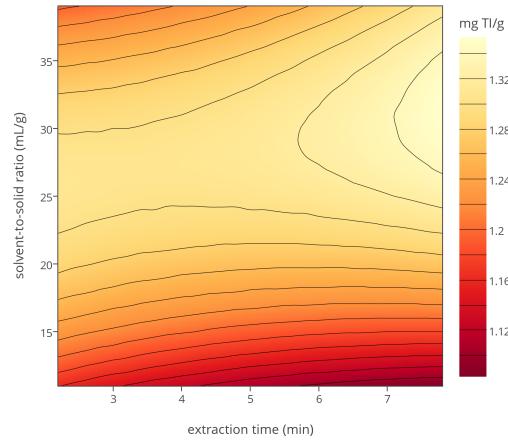
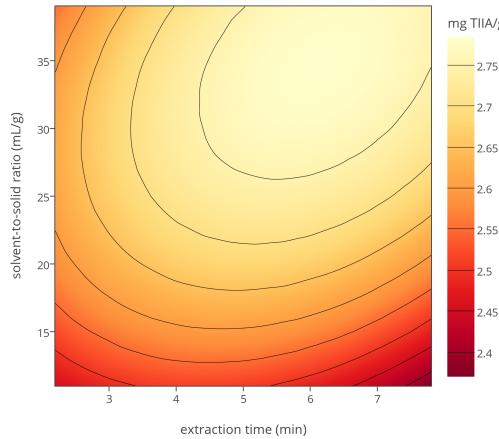


Figure 15: Response Surface for Tanshinone IIA



## Part VI. Finding the Global Optimum Across All Analytes

In Part V we determined that the individual extraction yields for danshensu, lithospermic acid, salvianolic acid A, cryptotanshinone, tanshinone I, and tanshinone IIA increase at longer extraction times and for larger solvent-to-solid ratios.<sup>11</sup> We also determined that the optimum extraction yield for tanshinone IIA is at a shorter extraction time than that for the other analytes, and that the optimum solvent-to-solid ratio for lithospermic acid and for salvianolic acid A are at larger solvent-to-solid ratios than that for the other analytes. Clearly selecting a single set of extraction conditions—what we call the global optimum—requires a compromise.

There are a variety of useful approaches to locating the global optimum when working with multiple analytes. For example, when working with a small number of analytes, typically two or three, it is possible to overlay contour plots and look for a set of factor levels where each analyte exceeds some threshold value. When working with more analytes, a more useful approach is to use Derringer's desirability function.<sup>12</sup>

The general form of the desirability function for  $n$  analytes is

$$D = (d_1^{r_1} \times d_2^{r_2} \times \cdots \times d_n^{r_n})^{1/\sum r_i}$$

where  $D$  is the global desirability,  $d_i$  is the individual desirability for the  $i^{\text{th}}$  analyte, and  $r_i$  is the relative importance for the  $i^{\text{th}}$  analyte, which allows us to weight the global desirability toward those analytes we deem more important. An analyte's individual desirability is determined by comparing its response,  $R_i$ , at a particular point on the response surface to an upper limit,  $U_i$ , and to a lower limit,  $L_i$ , of our choosing. If we wish to maximize the response, we set the individual desirabilities as

$$d_i = \begin{cases} 0 & \text{if } R_i < L_i \\ \left(\frac{R_i - L_i}{U_i - L_i}\right)^s & \text{if } L_i \leq R_i \leq U_i \\ 1.00 & \text{if } R_i > U_i \end{cases}$$

where the scaling factor,  $s$ , determines how slowly or quickly  $d_i$  approaches its maximum value of 1.

**Investigation 26.** To explore the effect of  $s$  on individual desirability, calculate  $d_i$  for responses from 0.0 to 1.0, in steps of 0.1, using an upper limit of 0.75 and a lower limit of 0.25, and values of 0.5, 1.0, and 5.0 for  $s$ . Examine your results and comment on any trends you see.

Although the desirability function seems complex, it is not hard to see how it works. As an example, let's consider how to calculate danshensu's individual desirability,  $d$ , for each combination of extraction time and solvent-to-solid ratio in Figure 10. First, we determine danshensu's maximum extraction yield and define the response,  $R$ , as the fraction of that maximum extraction yield. Next, we define the upper limit and the lower limit. Let's set the upper limit as 95% of danshensu's maximum extraction yield; thus,  $U$  is 0.95 and  $d = 1.00$  anytime the extraction yield exceeds 95% of its maximum value. If we define danshensu's lower limit as 90% of its maximum yield, then  $L$  is 0.90 and  $d = 0$  anytime the extraction yield is less than 90% of its maximum value. Between the upper limit

<sup>11</sup> As a reminder, the extraction yields for rosmarinic acid and for dihydrotanshinone do not show much variation, although the extraction of rosmarinic acid increases slightly with increasing solvent-to-solid ratios and the extraction of dihydrotanshinone decreases slightly with increasing extraction time. We will continue to assume that the extraction yields for these two analytes are relatively independent of extraction time and solvent-to-solid ratio.

<sup>12</sup> For additional information on Derringer's desirability function, see "Experimental design and multiple response optimization. Using the desirability function in analytical methods development," the full reference for which is Candioti, L. V.; De Zan, M. M.; Cámara, M. S.; Goicoechea, H. *Talanta*, **2014**, 124, 123–128 (DOI:10.1016/j.talanta.2014.01.034).

and the lower limit, we calculate  $d$  as defined above. Figure 16 shows danshensu's individual desirability function as a response surface using  $s = 1$ , which assumes a linear increase in the individual desirability between the upper limit and the lower limit.

**Investigation 27.** Compare the response surface for danshensu's individual desirability (Figure 16) to its response surface in terms of extraction yield (Figure 10). In what ways are these response surfaces similar and in what ways are they different?

An important feature of the global desirability function is that  $D$  is the product of each analyte's individual desirability function, which means the global desirability is zero for any combination of extraction time and solvent-to-solid ratio if at least one analyte's individual desirability function is zero. In addition, we can assign more weight to some analytes and less weight to other analytes by adjusting the value of  $r$  for each analyte.

**Investigation 28.** To explore the effect on the global desirability of weighting analytes, let's assume we have four analytes with individual desirabilities of 0.90, 0.80, 0.70, and 0.60. What is the global desirability if you (a) weight the factors evenly by assigning each an  $r$  of 1; (b) assign a weight of 3 to the first analyte and a weight of 1 to the other three analytes; (c) assign a weight of 5 to the first analyte and a weight of 1 to the other three analytes; (d) assign a weight of 3 to the last analyte and a weight of 1 to the other three analytes; and (e) assign a weight of 2 to the second and third analytes and a weight of 1 to the first and last analyte? Examine your results and discuss any trends you see.

The ability to adjust the upper limit,  $U$ , the lower limit,  $L$ , and the scaling factor,  $s$ , when calculating individual desirabilities, and to adjust the relative weighting,  $r$ , for each analyte when calculating the global desirability provides flexibility in identify the optimum conditions for extracting samples of Danshen. Figure 17 shows the global desirability function's response surface based on individual desirability functions for danshensu, lithospermic acid, salvianolic acid A, cryptotanshinone, tanshinone I, and tanshinone IIA. Each individual desirability function was calculated using an upper limit of 0.95, a lower limit of 0.90, and with  $s$  set to 1. All six analytes were weighted equally by setting their respective values of  $r$  to 1.

**Investigation 29.** A comparison of Figure 16 and Figure 17 shows that the global desirability function has a smaller range of maximum values than does the individual desirability function for danshensu. Which analytes limit the range of optimum values for the global desirability function? Based on Figure 17, what is the range of extraction times and range of solvent-to-solid ratios that result in an optimum global desirability? Given the range of possible values for the extraction time and the solvent-to-solid ratio, what values are the best option? Why?

Figure 16: Individual Desirability Function for Danshensu

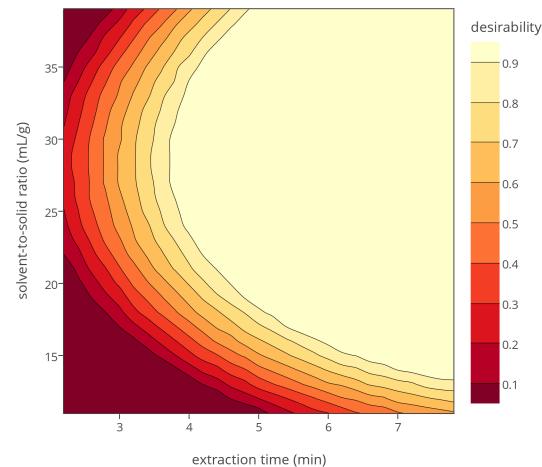
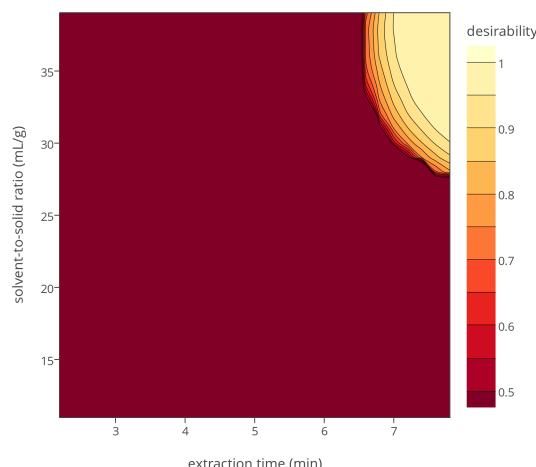


Figure 17: Response Surface for Global Desirability Function



## Part VII. Verifying the Analytical Method's Accuracy

Let's review our progress in developing a method for determining the concentration of hydrophilic and lipophilic compounds in Danshen. In Part IV we concluded that the optimum solvent is 80% methanol and 20% water (by volume), that the optimum extraction temperature is 70°C, and that the optimum microwave power is 800 W. In Part V we determined that the extraction yields for danshensu, lithospermic acid, salvianolic acid A, cryptotanshinone, tanshinone I, and tanshinone IIA increase at longer extraction times and for larger solvent-to-solid ratios, and that the extraction yields for rosmarinic acid and for dihydrotanshinone are not affected significantly by variations in extraction time and the solvent-to-solid ratio. We also learned in Part V that the analytes do not share a common optimum extraction time or solvent-to-solid ratio. Finally, in Part VI we used a global desirability function to show that an extraction time of 7.50 min, and a solvent-to-solid ratio of 35.0 mL/g allows for at least a 95% recovery of each analyte's optimum extraction yield.

Having optimized our method, we turn our attention to two additional steps in developing an analytical method: verifying that the analytical method works and applying the method to a range of different samples. Here, in Part VII, we examine ways to verify our analytical method; in Part VIII we apply the analytical method to samples of Danshen from different natural locations and samples from a controlled cultivation.

To verify our analytical method we need to show that the experimental extraction yields agree with the extraction yields predicted by the empirical models used to generate the response surfaces in Figures 10–15.

**Investigation 30.** In Part V we found that the empirical model for the extraction of danshensu is

$$EY = 0.575 + 0.0225A + 0.00905B - 0.00125A^2 - 0.000165B^2 + 0.000100AB$$

where  $EY$  is the extraction yield (in mg/g),  $A$  is the extraction time (in min), and  $B$  is the solvent-to-solid ratio (in mL/g). Using this model, calculate danshensu's predicted extraction yield for an extraction time of 7.50 min and a solvent-to-solid ratio of 35.0 mL/g. Is your predicted extraction yield consistent with the data in Table 2 and your response to Investigation 25?

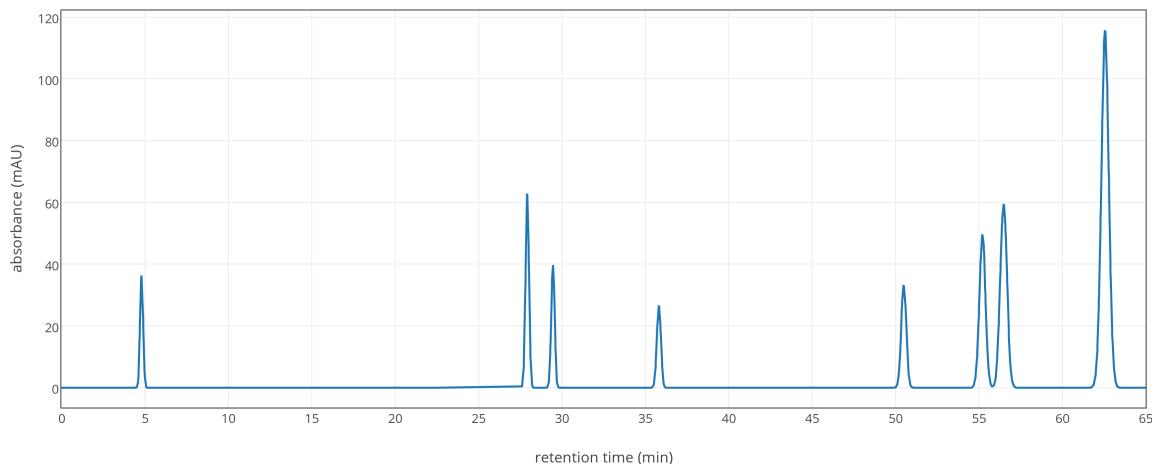
Table 3 summarizes the predicted extraction yields for the remaining compounds in Danshen. The predicted extraction yields for lithospermic acid, salvianolic acid A, cryptotanshinone, tanshinone I, and tanshinone IIA are from the empirical models used to construct the response surfaces in Figures 11–15; the predicted extraction yields for rosmarinic acid and for dihydrotanshinone are the average results of the 13 trials used in their respective central-composite designs.

**TABLE 3. PREDICTED EXTRACTION YIELDS FOR DANSHEN'S CONSTITUENTS**

analyte	extraction yield (mg/g)	analyte	extraction yield (mg/g)
danshensu	(see Investigation 30)	dihydrotanshinone	0.424
rosmarinic acid	2.317	cryptotanshinone	0.917
lithospermic acid	2.657	tanshinone I	1.336
salvianolic acid A	0.600	tanshinone IIA	2.762

**Investigation 31.** Figure 18 shows the chromatogram for a sample of Danshen extracted using the optimized conditions from Part VI. Using this chromatogram, calculate the actual extraction yield for each analyte and report its experimental extraction yield as a percentage of its predicted extraction yield from Table 3. Do your results provide confidence in our analytical method? Why or why not?

Figure 18: HPLC of Danshen Extracted Using Optimum Conditions



The motivation for developing this microwave-assisted extraction is the concern that conventional extraction methods require longer extraction times and that the extended application of a high temperature may result in the thermal degradation of Danshen's constituents. Table 4 provides extraction yields for the same sample of Danshen from Figure 18 using three conventional benchtop heat-reflux extraction (HRE) methods.

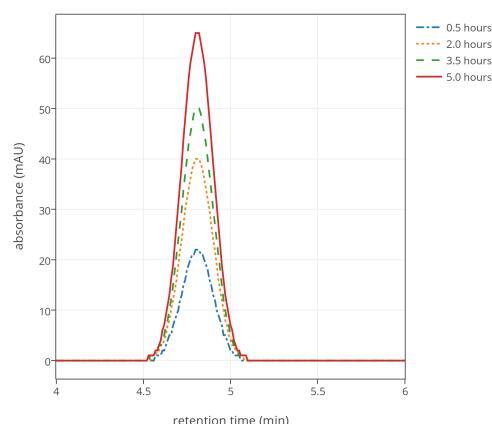
**TABLE 4. EXTRACTION YIELDS (MG/G) USING CONVENTIONAL METHODS**

analyte	HRE-1	HRE-2	HRE-3
danshensu	1.618	0.826	1.052
rosmarinic acid	2.032	2.016	1.619
lithospermic acid	2.675	1.785	2.265
salvianolic acid A	0.435	0.437	0.454
dihydrotanshinone	0.352	0.354	0.295
cryptotanshinone	0.571	0.599	0.543
tanshinone I	0.913	0.982	0.926
tanshinone IIA	1.952	2.280	1.738

The solvent for HRE-1 is 80% methanol and 20% water (by volume) with the extraction carried out twice for 60 min each. For HRE-2 the solvent is 100% methanol, with a single 60 min extraction. The solvent for HRE-3 is 75% methanol and 25% water (by volume), with a single 60 min extraction. The temperature for each extraction is the solvent's boiling point.

**Investigation 32.** Compare your results from Investigation 31 with the results reported in Table 4. Do these results support a concern that heat-reflux extractions may distort the apparent composition of Danshen? As you consider this question, you may wish to review the chemical structures of these compounds, which are shown in Part I, and the HPLC data in Figure 19 for samples drawn at different times during an extended heat-reflux extraction of Danshen.

Figure 19: HPLC of Danshen Using a Heat-Reflux Extraction

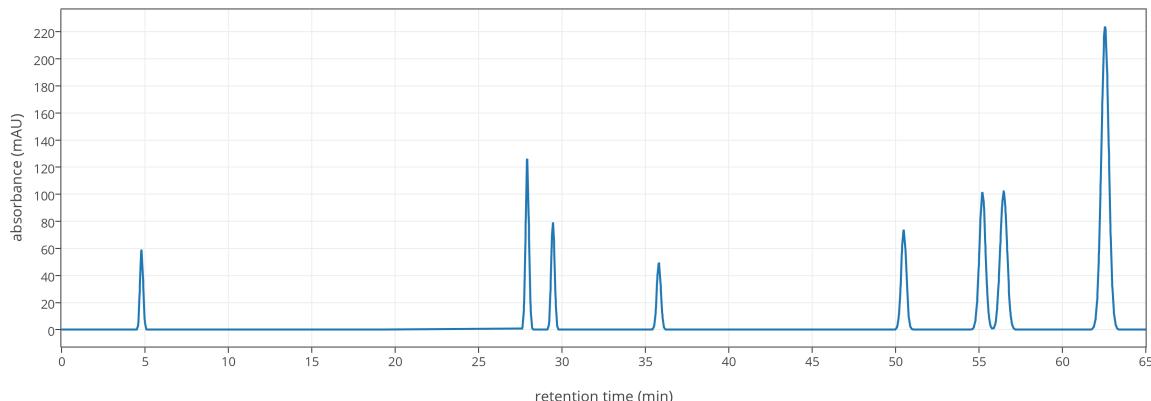


Although the results of Investigation 31 and Investigation 32 provide confidence in our analytical method, it is not the same as establishing that the concentrations we report represent accurately the actual concentrations of our analytes in Danshen. The possibility of thermal degradation during the microwave extraction, for example, is a potential concern not addressed in Investigation 32.

Establishing our method's accuracy is complicated because we do not know, *a priori*, the actual concentrations of the hydrophilic compounds and the lipophilic compounds in any particular sample of Danshen; indeed, their concentrations certainly vary from plant-to-plant and from field-to-field, particularly plants and fields in different geographic regions, and likely are affected by methods of cultivation. Nevertheless, we can establish our method's accuracy by analyzing a sample before and after adding a known amount of the analytes of interest.

**Investigation 33.** Explain why analyzing a sample before and after adding a known amount of an analyte allows you to evaluate a method's accuracy. Figure 20 shows the chromatogram for a sample of Danshen spiked prior to the microwave extraction with known amounts of each analyte, the concentrations of which are shown in Table 5. Using this data and your results for the unspiked sample in Investigation 31, how confident are you in the accuracy of our analytical method?

Figure 20: HPLC of Spiked Danshen Extracted Using Optimum Conditions



**TABLE 5. CONCENTRATIONS OF SPIKES ADDED TO DANSHEN**

analyte	$C_{\text{spiked}}$ (mg/g)	analyte	$C_{\text{spiked}}$ (mg/g)
danshensu	0.500	dihydrotanshinone	0.500
rosmarinic acid	2.500	cryptotanshinone	1.000
lithospermic acid	2.500	tanshinone I	1.000
salvianolic acid A	0.500	tanshinone IIA	2.500

## Part VIII. Applying the Analytical Method

With our analytical method optimized and its accuracy verified, we turn, at last, to applying our method to the analysis of samples of Danshen roots. Table 6 provides absorbance values (in mAU) for danshensu and for tanshinone I in wild plants harvested from five different cities in the province of Shandong, China, and in five plants harvested from a single cultivated field in which good agricultural practices that emphasize agricultural sustainability are used.

**TABLE 6. RESULTS FOR ANALYSIS OF DANSHEN SAMPLES**

Danshen Source	absorbance (mAU) for danshensu	absorbance (mAU) for tanshinone I
Wild Samples (Cities in Shandong Province)		
Sanshangou	21.6	123.8
Yuezhuang	10.3	55.3
Dazhangzhuang	11.8	67.6
Pingse	37.2	42.1
Mengyin	10.0	132.0
Cultivated Samples (Lot Number)		
020208	23.4	136.6
020209	23.7	137.1
020210	23.3	137.5
020211	22.8	148.0
020212	23.5	150.8

**Investigation 34.** Calculate the concentration of danshensu and the concentration of tanshinone I in each sample. For each set of samples—wild samples and cultivated samples—calculate the mean, the standard deviation, and the relative standard deviation for each analyte and comment on your results.

## **Part IX. Closing Thoughts**

The results for Investigation 34 are reported as the concentration, in mg/g, of danshensu and tanshinone I in samples of Danshen roots. Despite reporting the results this way, we cannot assume these values are the actual concentrations of danshensu and tanshinone I in these sample; they are, instead, the concentration of danshensu and tanshinone I extracted using 35.0 mL of a solvent that is 80% methanol and 20% water (by volume) per 1.000 g of sample, and using a microwave oven at 800 W to heat the solvent and sample for 7.50 min at 70°C. Different methods of extracting samples of Danshen yield different extraction yields, some of which recover smaller amounts of analytes (see Table 3 and Investigation 31), and some of which recover larger amounts of analytes (see Figures 3–5 and Investigation 12).

Although our analytical method reports the concentrations in Danshen of extractible hydrophilic and lipophilic compounds instead of their total concentrations, the analysis still has value because we ultimately are interested in the concentrations of these compounds that are easily recovered after harvesting the plants. In addition, and as suggested by Investigation 34, our analytical method provides us with a standard method for comparing the relative potency of different sources of Danshen and as a means of evaluating how changes in cultivation practices affect the relative potency of commercially grown Danshen. These are important and useful applications.