Response surface modeling and optimizing conditions for anthocyanins extraction from *Hibiscus sabdariffa* L. (Roselle) grown in Lam Dong, Vietnam

Tri Nhut Pham$^{1,2}$, Thuong Nhan Phu Nguyen$^1$, Lam Tri Duc$^{1,3}$, Minh Tien Nguyen$^1$, Tran Quoc Toan$^4$, Le Thi Hong Nhan$^5$, Dai Viet N-Vo$^6$, Thanh Sang Vo$^1$, Le Minh Bui$^{1,*}$

$^1$NTT Hi-Tech Institute, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam
$^2$Faculty of Chemical Engineering and Food Technology, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam
$^3$Graduate School of Biotechnology, Kyung Hee University, South Korea
$^4$Institute of Natural Products Chemistry, Vietnam Academy of Science and Technology, Ha Noi, Vietnam
$^5$Department of Chemical Engineering, HCMC University of Technology, VNU-HCM, Ho Chi Minh City, Vietnam
$^6$Faculty of Chemical and Natural Resources Engineering, Universiti Malaysia Pahang, Malaysia

*Corresponding author: blminh@ntt.edu.vn, labams2013@gmail.com

**Abstract.** *Hibiscus sabdariffa* L. (Hs, roselle; Malvaceae) is a common ingredient used in the food and pharmaceutical industries for color and health benefits. Anthocyanins along with protocatechuic acid and quercetin have been recognized as bioactive compounds in *Hibiscus sabdariffa* L. aqueous extracts. In this work, Response Surface Methodology was implemented together with Box-Behnken experimental design to investigate the impacts of process variables on anthocyanin extraction yields from *Hibiscus sabdariffa* L. (Roselle) grown in Lam Dong, Vietnam. The effects of three independent variables; namely the extraction temperatures of solvent ethanol (X1); the duration of extraction (X2); solid-liquid ratios (X3) on the total anthocyanins content (Y1) were explored. The optimization model showed that with the solid liquid ratio 7.7:1, extraction time 33.29 min, and temperature 56.9°C, the highest anthocyanin concentration of 186.006 mg/L was obtained. *Hibiscus sabdariffa* L. anthocyanin yield detailed significant correlation with high F values, low P values (<0.0001), and desirable determination coefficient (R$^2$ = 0.9770). The resulting quadratic model could be used to predict yields with given conditions. Actual experimental verification and residual testing revealed that the actual anthocyanin yields could be accurately predicted by the model.
1. Introduction

Hibiscus sabdariffa L., commonly known as roselle or red sorrel, is one of the most commonly cultivated plants globally. The plant has tropical origin and belongs to the Malvaceae family. The plant is a common ingredient in traditional medicine and extracts from the plant has been shown to exhibit a wide spectrum of therapeutic effects regarding various diseases such as cardiovascular disease, liver disease, fever and hypertension [1-3]. In addition, due to its colour and unique flavor, Hibiscus sabdariffa L. are utilizable in culinary as a cooking ingredient and colorant. A prominent feature of the Hibiscus sabdariffa L extract is its high content of antioxidants where anthocyanins are major components. Anthocyanins are water-soluble glycosides and acylglycosides of anthocyanidins, and they are found in the form of polyhydroxylated and or methoxylated heterosides which derive from the flavylum ion or 2-phenylbenzopyrilium in nature. Anthocyanins are compounds of interest because they are suitable replacements for chemical dyes. The compounds is characterized polyphenolic family, commonly existing in the plant kingdom [4-6]. However, a suitable extraction technique is important to maintain the antioxidant activity of the extract from the plant and to obtain the highest yield of anthocyanins. Although parameters including solvent acidity, solid-to-solvent ratio, time or temperature of extraction were individually studied in different studies, a study which include all important parameters is lacking in the literature. The aforementioned issue could be addressed by a technique that accommodates variables in an interactive manner. Optimization, in most cases, is performed with response surface methodology (RSM). The technique is a powerful and widely used technique devised for testing of multiple-process variables and optimizing experimental responses. The advantage of the technique is numerous ranging from simplicity to high efficiency and consistency with comprehensive theories. In addition, it also allows fewer experimental trials in comparison with studies of one variable at a time and is applicable in various fields, including construction field [7-15]. In order to achieve the above purpose, RSM combined with Box–Behnken design was employed in this study to identify the set of conditions which yields the highest anthocyanin yields from Hibiscus sabdariffa L. (Roselle) grown in Lam Dong, Vietnam. Following parameters were optimized: solid-liquid ratios, extraction temperatures and duration of extraction. The established model will serve as reasonable relationship between the independent variables and the response.

2. Material and methods

2.1. Sample preparation

Hibiscus sabdariffa L. dried calyces were used, provided by farmers of Da Lat city, Lam Dong Province, Viet Nam. Dried petals of roselle were ground using a commercial grinder and were weighed to 10g, put in the two neck round bottom flask and was extracted by Ethanol 50° solutions. Ethanol (C₂H₅OH) is purchased from Sigma Aldrich (US). The different particle diameter ground samples were stored in polyethylene vacuum bags at room temperature until extraction was carried out. First of all, each experimental condition will be tested individually to determine the optimum condition. For extraction parameter study, 10g of roselle powder was placed in the two neck round bottom flask and was extracted by ethanol with concentration at 50. The liquid/solid ratio in this experiment ranges from 2:1 to 10:1 (mL/g) and. The extraction temperature is adjusted from 40 to 80 (°C) and time varies from 15 to 35 (min). Then, centrifugation took place at 4500 rpm for 15 min by high speed centrifuge Model LACE16 (from COLO lab expert). The supernatant was collected and the extract, after being filtered with filter paper, was transferred into plastic bottle to estimate anthocyanin content. The pH scanning of supernatant ranges from λvis_max to 700 nm were recognized for anthocyanin content. The total anthocyanin content was determined by using the pH-differential method [16]

2.2. Experimental design using RSM

Following the estimation of the statistical model, analysis of variance (ANOVA) was performed to evaluate the significance of the experimental variables of consideration. These variables include solid/liquid ratio (X₁), extraction temperature (X₂), and time extraction (X₃). The whole optimization process was administered in the Design Expert 7.0 software. Based on the results obtained from the single factor investigation, 3-level factors (Table 1) were produced and attempted in actual
experiments to generate data (Table 2) for estimating the second order polynomial model, which is described as follows:

\[ Y = \beta_0 + \sum_{i=1}^{3} \beta_i x_i + \sum_{i=1}^{2} \sum_{j=i+1}^{2} \beta_{ij} x_i x_j \]

where, \( Y \) is a response variable, \( \beta_0 \), \( \beta_i \), \( \beta_{ij} \) and \( x_i \) and \( x_j \) are the regression coefficients of the intercept, linear, quadratic and interaction terms, respectively.

The difference between actual values and predicted values was not substantial indicating that the model is significant as demonstrated by the F-value of 47.18. The probability of the model to obtain such value owing to noise is minimal (0.79%). All model terms are also statistically significant as indicated by low p-values. Insignificance of the lack of fit shows that the model fits well with the real data. Overall, the model is significant as demonstrated by the F-value of 47.18. The probability of the model to obtain such value owing to noise is minimal (0.79%). All model terms are also statistically significant as indicated by low p-values. Insignificance of the lack of fit shows that the model fits well with the data, thus no further specification of the model is required. After estimating model coefficients, the following model is obtained:

\[ Y = 186.6 - 2.82X_1 - 1.31X_2 + 1.36X_3 - 1.37X_1X_2 - 2.57X_2X_3 - 2.16X_1X_3 - 5.01X_2^2 - 3.00X_2^2 - 3.89X_3^2 \]

The predictive model calculation (second order polynomial equation model) leads to the optimal theoretical calculation of extraction in order to achieve the maximum value of 187.006 mg/L anthocyanin content, using 50% EtOH, at 56.9 °C, and 7:1 (mL/g) solid/liquid ratio.

| Table 1. Coded and levels of the three independent variables for full factorial design of RSM model |

| Variables | Coded factor | Level |
|-----------|-------------|-------|
| Solid-liquid ratios | \( X_1 \) (mL/g) | Low (-1) 7 : 1 | Medium (0) 8 : 1 | High (+1) 9 : 1 |
| Extraction temperatures | \( X_2 \) (°C) | 60 | 65 | 70 |
| Duration of extraction | \( X_3 \) (min) | 35 | 40 | 45 |

3. Result and discussion

3.1 Process optimisation and statistical analysis of experimental design

The experiments were designed according to the design method of complex CCD center. Experimental results (20 experiments) and predicted values generated by Design-Expert 11 are shown in Table 2. To be specific, variables of interest include concentration of ethanol 50%, liquid/solid ratio ranged from 4:1 to 6:1 (mL/g), temperature ranged from 60 to 70 °C and reaction time ranged from 35-45 (min). No.1–14 were the factorial experiments, and No.15–20 were the central experiments. The resulting anthocyanin content ranges from 163.86 mg/L to 186.77 mg/L. The results indicated that all factors showed statistical significance, solid/liquid ratio at the p ≤ 0.05 level and both temperature and time at a ≤ 0.01 level. Additionally, there was a high statistical significance (p ≤ 0.01) in the interaction between solid/liquid, temperature and time; This occurs when one of the independent variables exerts an effect on another of the independent factors of the experimental design. The results implied that the anthocyanin content would be severely altered as survey parameters change (ratio, temperature, time).

The difference between actual values and predicted values was not substantial indicating that the results of the experimental experiments have high accuracy. As indicated by Table 3, determination coefficient \( R^2 = 0.9770 \) shows a high goodness-of-fit of the predicted values to the real data. Overall, the model is significant as demonstrated by the F-value of 47.18. The probability of the model to obtain such value owing to noise is minimal (0.79%). All model terms are also statistically significant as indicated by low p-values. Insignificance of the lack of fit shows that the model fits well with the data, thus no further specification of the model is required. After estimating model coefficients, the following model is obtained:

\[ Y = 186.6 - 2.82X_1 - 1.31X_2 + 1.36X_3 - 1.37X_1X_2 - 2.57X_2X_3 - 2.16X_1X_3 - 5.01X_2^2 - 3.00X_2^2 - 3.89X_3^2 \]

Table 2. The matrix of observed and predicted values for RSM model

| No | \( X_1 \) | \( X_2 \) | \( X_3 \) | Actual | Predicted | Y (mg/L) | No | \( X_1 \) | \( X_2 \) | \( X_3 \) | Actual | Predicted | Y (mg/L) |
|----|-----|-----|-----|-------|------|-------|    | ----|-----|-----|-----|-------|------|-------|
| 1  | 7   | 55  | 25  | 173.35| 171.37| 11    |    | 8   | 51.59| 30  | 180.13| 180.32|
| 2  | 9   | 55  | 25  | 173.47| 173.62| 12    |    | 8   | 68.41| 30  | 176.15| 175.91|
| 3  | 7   | 65  | 25  | 176.19| 175.80| 13    |    | 8   | 60  | 21.59| 171.34| 173.31|
| 4  | 9   | 65  | 25  | 173.69| 172.56| 14    |    | 8   | 60  | 38.41| 179.87| 177.8 |
| 5  | 7   | 55  | 35  | 182.44| 183.55| 15    |    | 8   | 60  | 30  | 186.01| 186.60|
| 6  | 9   | 55  | 35  | 175.13| 175.51| 16    |    | 8   | 60  | 30  | 186.73| 186.60|
| 7  | 7   | 65  | 35  | 179.52| 179.36| 17    |    | 8   | 60  | 30  | 186.66| 186.60|
Table 3 ANOVA data for removal models

| Source | Sum of Squares | Degree of freedom | Mean Square | F-value | Prob. > F | Comment |
|--------|----------------|-------------------|-------------|---------|-----------|---------|
| Model  | 862.56         | 9                 | 95.84       | 47.18   | < 0.0001  | significant SD = 1.43 |
| A-A    | 108.64         | 1                 | 108.64      | 53.48   | < 0.0001  | Mean = 178.48 |
| B-B    | 23.50          | 1                 | 23.50       | 11.57   | 0.0068    | CV(%) = 0.7986 |
| C-C    | 25.32          | 1                 | 25.32       | 12.46   | 0.0054    | R² = 0.9770 |
| AB     | 15.04          | 1                 | 15.04       | 7.40    | 0.0215    | AP = 20.60% |
| AC     | 52.99          | 1                 | 52.99       | 26.09   | 0.0005    |         |
| BC     | 37.20          | 1                 | 37.20       | 18.31   | 0.0016    |         |
| A²     | 361.94         | 1                 | 361.94      | 178.16  | < 0.0001  |         |
| B²     | 129.60         | 1                 | 129.60      | 63.79   | < 0.0001  |         |
| C²     | 217.76         | 1                 | 217.76      | 107.19  | < 0.0001  |         |
| Residual | 20.32        | 10                | 2.03        |         |           |         |
| Lack of Fit | 19.89   | 5                 | 3.98        | 46.37   | 0.1503    | not significant |
| Pure Error | 0.4289   | 5                 | 0.0858      |         |           |         |

3.2 Analysis of response surface

The coinciding relationship of various variables and the response could be described by the surface plot. Based on the model estimated by the software, three response surface plots of the anthocyanin content, corresponding to different combinations of the solid-liquid ratios (X₁), extraction temperatures (X₂) and duration of extraction (X₃) were plotted as in Fig.1. The contour plots represented impact intensity of the interaction variables on the content. According to the model results, all three interaction terms representing combinations of three factors are significant (P<0.05). From the response surface analysis, optimum process of anthocyanin extraction are identified with the following parameters: temperature 56.9 (°C), time 33.29 (min), and liquid-solid ratio 7.7 :1 (mL/g). The maximum predicted anthocyanin content was 187.006 mg/L.

Figure1. 3D model about the influence and interaction of Y with (a) the solid-liquid ratios (X₁) and extraction temperatures (X₂); (b) the solid-liquid ratios (X₁) and duration of extraction (X₃); (c) extraction temperatures (X₂) and duration of extraction (X₃)

3.3 Testing the Model Precision

The yield of anthocyanin extraction could be predicted using the above model. To validate the model, residuals of 20 runs and yields of oil were plotted in the Figure 2. A normal plot of residuals was generated (Figure 2) and it can be observed that most data points are grouped reasonably near to, and
on either side of the straight 45° line.

**Figure 2.** Normal plot of residuals of anthocyanin content

4. **Conclusion**

In this work, a RSM procedure was applied to study the effect of three variables (solid-liquid ratios, extraction temperatures, duration of extraction) on anthocyanin extraction yield from Vietnamese Hibiscus sabdariffa L. (*Roselle*). The 20 experiments were planned by CCD to generate the data for estimation of the model. The best combination of factors yielding the highest anthocyanin content (187.006 mg/L) was as follows: 33.29 (min) temperature and 7.7:1 (mL/g) solid/liquid ratio.

The ANOVA results indicated that the solid-liquid ratios ($X_1$), extraction temperatures ($X_2$) and duration of extraction ($X_3$) had significant effects on the responses. The obtained results suggested that RSM is a reasonable tool for optimization not only for extraction processes, but also for other applications involving multiple experimental conditions.

**Acknowledgments**

This work was supported by grants from Nguyen Tat Thanh University, Ho Chi Minh City, Viet Nam.

**Reference**

1. G Riaz, R Chopra 2018 *Biomedicine & Pharmacotherapy* **102** 575-586
2. I DaCosta Rocha, B Bonlaender, H Sievers, I Pischel, M Heinrich, L 2014 *Food Chemistry* **165** 424-443
3. Z Aziz, SY Wong, NJ Chong 2013 *Journal of Ethnopharmacology* **150** 442-450
4. ARC Braga, D Murador, VV Rosso 2018 *Journal of Food Composition and Analysis* **68** 31-40
5. P Ongkowijoyo, DA Luna-Vital, E. Mejia 2018 *Food Chemistry* **250** 113-126
6. L Xie, H Su, C Sun, X Zheng, W Chen 2018 *Trends in Food Science & Technology* **72** 13-24
7. TH Tran, HHHm Nguyen, DC Nguyen, TQ Nguyen, H Tan, LTH Nhan, DH Nguyen, LD Tran, ST Do, and TD Nguyen 2018 *Processes* **6** 206
8. TT Hien, NPTNhan, DT Nguyen, VTT Ho, LGBach 2018 *Solid State Phenomena* **279** 217-221.
9. NP Nhan, T Hien, LTH Nhan, PNQ Anh, LT Huy, NTC Trinh, DT Nguyen, L.G. Bach 2018 *Solid State Phenomena* **279** 235-239
10. TV Tran, TP Bui, T Nguyen, NTH Le, LG Bach 2017 *Adsorption Science & Technology*, vol. **35**, 2017, pp. 72–85
11. VT Tran, DT Nguyen, VTT Ho, PQH Hoang, PQ Bui, LGBach 2017 *Journal of materials and Environmental Sciences* **8** 426-437
12. TV Tran, QTP Bui, TD Nguyen, VTT Ho, LG Bach, 2017 *Water Science & Technology*, **75** 2047-2055.
13. TV Thuan, VTT Ho, ND Trinh, NT Thuong, BTP Quynh, LG Bach 2016 *Carbon – Science and Technology* **863**-69
14. T Van Thuan, BTP Quynh, TD Nguyen, LG Bach, 2017 *Surfaces and Interfaces*, **6** 209-217
15. LG Bach, T Tran, TD Nguyen, TV Pham, ST Do 2018 *Research on Chemical Intermediates* **44** 1661-1687
16. J Lee, RW Durst, RE Wrolstad 2005 *Journal of AOAC International* **88** 1269-1278