LWT - Food Science and Technology 55 (2014) 96–103

Contents lists available at ScienceDirect

LWT - Food Science and Technology

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

# Artificial intelligence in predicting extraction of anti-cancer compounds

B.L. Shrestha, O.D. Baik*

*Department of Chemical and Biological Engineering, University of Saskatchewan, 57 Campus Dr., Saskatoon, SK S7N 5A9, Canada*

ARTICLE INFO

*Article history:*
Received 11 April 2013
Received in revised form
1 September 2013
Accepted 2 September 2013

*Keywords:*
Neural networks
Diffusional model
Saponins
*Saponaria vaccaria* L.

ABSTRACT

Saponins from the particulate *Saponaria vaccaria* L. seeds (0.29–0.84 mm, 15.35–61.40 g $H_2O$/100 g dry mass) were extracted for methanol concentrations (MeOH) of 30, 50, 70, and 90 mL/100 mL $H_2O$ and temperatures ($T$) of, 30, 45, and 60 °C at ten extraction intervals ($t$) between 1 and 180 min. A calibration equation was developed from the liquid-chromatogram–mass spectroscopy peaks to quantify the extract yields (mg $mL^{-1}$) for various types of saponins. An artificial neural network (ANN) with three inputs, MeOH, $T$, and $t$ predicted the extraction kinetics and the yields with less than ca. 12% error. The ANN model not only slightly outperformed the numerical diffusional model, but it also made the prediction simple and faster eliminating the use of the partition coefficient and the effective diffusivity. Therefore an ANN model can be a right approach to predict the yields of saponins and similar products.



## 1. Introduction

Saponins are the natural compounds with anti-cancer and anti-tumor functions (Kwon, Jacqueline, Bélanger, Paré, & Yaylayan, 2003; Oleszek & Marston, 2000). They are currently in high demand for pharmaceutical, nutraceutical, and other health-care applications. Triterpene bisdesmosidic saponins such as segetoside and vaccaroside from the seed of *Saponaria vaccaria* L. have potentials for the treatments of tumor and cancer in human prostate, lungs, breasts, and colon due to their ability to induce apoptosis *in vitro* in human cancer cell lines at low concentrations (Ramirez-Erosa, 2008). Synthesis of active biomaterials contained in plants for pharmaceutical and nutraceutical use has been in place for last few decades. The synthesized products are expensive to produce and also possess some side effects to human health. According to the World Health Organization (WHO), about 80% of the world's population still relies on natural medicine (Chen & Han, 2011) and 41% of the approved drugs between 1983 and 1994 contain natural products as their source (Farnsworth, Akerele, & Bingel, 1985). For anti-cancer and anti-infective compounds the percentage of drugs with natural products as their source has increased to over 60% (Cragg, Newman, & Snader, 1997). In a recent WHO event held in Beijing, health representatives from more than 70 countries agreed up on an idea of using natural products as alternatives or supplements to modern health care. In Canada and Germany, according to WHO, more than seven in ten people have tried the natural products. Sales revenue from traditional medicine in China reached approximately US$ 21 billion in 2007, and was forecasted to reach US$ 28 billion by 2010.

The *S. vaccaria* L. (Prairie Carnation) is a novelty medicinal plant native to North America, and its seeds, particularly the germs (embryo) are the major source of saponins. They contain a large proportion of triterpene saponins including quillaic acid and gypsogenin bisdesmosides, and gypsogenic acid monodesmosides in addition to several phenolics and small cyclic peptides (Balsevich, Bishop, & Ramirez-Erosa, 2006; Jia, Koike, Sahu, & Nikaido, 2002; Sang et al., 2003; Shrestha & Baik, 2012). Use of warm solvent of water and alcohol such as ethanol or methanol is a common practice for the extractions of saponins from their resource materials. But, sample preparation and assay of chromatograms involve several steps that demand expert knowledge and skilled operator, and are very time consuming requiring several hours for a single extraction, and importantly, the operator exposes to highly toxic material such as methanol. On the other side, use of mathematical models such as DM with proper boundary conditions to predict yields and extraction kinetics is computationally intensive. ANNs have been known for successful modeling of complex non-linear systems. A well trained ANN can implicitly map inputs to outputs accurately and rapidly. With an outstanding commercial success of the adaptive channel equalizer, a device built on a single neuron network used for stabilizing voice signals in long distance telephone systems in 1984, ANN has been used in aerospace, banking, defense, manufacturing, financial, oil and gas, telecommunications, etc. In medicine, ANN has been used to diagnose severe diseases

* Corresponding author.
*E-mail address:* oon-doo.baik@usask.ca (O.D. Baik).

0023-6438/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.lwt.2013.09.001

such as breast cancer (Janghel, Shukla, Tiwariand, & Kala, 2010; Verma, 2008), coronary artery stenosis utilizing coronary arteriography (Mobley, Schechter, Moore, McKee, & Eichner, 2005), and liver cancer (Kondo, Kondo, Takao, & Ueno, 2010). A review of ANN applications to the clinical functions of diagnosis, prognosis, and survival analysis in the medical sectors of oncology, critical care and cardiovascular medicine can be found in Lisboa (2002). There is a growing research interests in application of ANN in food science including modeling of microbial growth and prediction of food safety, interpreting spectroscopic data, and predicting physical, chemical, functional and sensory properties of various food products during processing and distribution (Huang, Kangas, & Rasco, 2007).

Application of ANN in the field of extraction, however, is very limited. Camara et al. (2010) predicted lycopene and β-carotene in fruits and vegetables. Pan and Luo (2010) determined a model parameter in supercritical fluid extraction for soybean isoflavones using ANN, and Wei, Liu, and Ni (2006) simulated extraction kinetics for rhubarb anthraquinones obtained using supercritical $CO_2$ fluid extraction with ANN. No research has yet been conducted to predict the yield, and to model solid–liquid extraction kinetics of anti-cancer and -tumor bioactive materials using ANN.

There is a multitude of benefits with a good extraction model, but the phenomenon to be modeled could be very complex, and mathematically intensive, or are not even possible to formulate explicitly due to lack of a complete knowledge of underlying principles. In this context, ANN can be a simple, reliable and an effective model to be used.

The objectives of this research were 1) to quantify saponins yields at different extraction conditions, 2) to develop ANN model for accurate and rapid prediction of saponins yields, and 3) to compare saponins yields, and extraction kinetics obtained using ANN and DMs.

## 2. Materials and methods

Sections 2.1 through 2.5 have been briefly explained as detailed explanation can be found in Shrestha and Baik (2012) and Izadifar and Baik (2008).

### 2.1. Sample preparation

Only the fraction of *S. vaccaria* L. seed particles (germs) enriched with saponins was used. Since the smaller particles are extremely difficult to filter out, and might pass through all the stages of the extraction process showing up as a contamination in the final product of saponins, only the particles falling between 0.297 mm and 0.84 mm in diameter were used. The average particle thickness based up on 200 particles was 124 μm, and the particle porosity was 60.72 ± 0.76% (mean ± s.d.). The initial moisture content of the particles was 15.35 g $H_2O$/100 g dry mass. The samples were stored in the cold room at 4 °C.

### 2.2. Extracts preparation

A 50 mL of the aqueous methanol at desired temperature was slowly poured into the tube containing 0.500 ± 0.001 g of the germs stamping the time as soon as the methanol came into contact with the particles. The solution was constantly stirred and heated for a specified interval of time in the closed tube. A volume of 2.5 mL of the solution was extracted and centrifuged for 2 min at 1006 × g. A volume of approximately 1.5 mL of the solution was filtered through a nylon syringe filter (0.45 μm) into a HPLC vial to obtain the extract of *S. vaccaria* L. particles. A total of 120 extracts were prepared using four concentrations of methanol (30, 50, 70, and 90 mL/100 mL $H_2O$) using 99.99% pure methanol and HPLC grade water at three temperatures (30 °C, 45 °C and 60 °C) and 10 extraction intervals, 1, 2, 4, 6, 8, 30, 60, 100, 150 and 180 min.

### 2.3. HPLC and solvent system

Hewlett Packard 1100 HPLC with solvent A, 0.12 mL acetic acid/100 mL $H_2O$ in 10 mL acetonitrile/100 mL $H_2O$ and solvent B, 0.12 mL acetic acid/100 mL $H_2O$ in 100 mL acetonitrile/100 $H_2O$ was used as eluants in gradient mode over the time of 58 min at the flow rate of 0.2 mL $min^{-1}$. The column temperature, the injection volume, and the wavelength were 35 °C, 4 μL, and 210 nm respectively.

### 2.4. Mass transfer Biot number

The mass transfer Biot number, $Bi_m = h_m L_c / D_{eff}$, where, $h_m$: mass transfer coefficient (m $s^{-1}$), $L_c$: characteristic length (m), and $D_{eff}$: effective mass diffusivity ($m^2$ $s^{-1}$), was used to investigate the mass transfer resistances within the particles and at particle–solvent interface. The $Bi_m$ greater than 30 implies well mixing of the solvent and the solute from the particles.

### 2.5. Calibration equation to estimate saponins yield

#### 2.5.1. Preparation of saponin-rich sample from the particles of *S. vaccaria L.*

The saponins rich sample was prepared from the particles of *S. vaccaria* L. by following a long procedure consisting of de-fatting, and repeated washing with methanol/water solution, air and vacuum drying, and elution through a solid-phase extraction cartridge. The saponin-rich eluates were dried in vacuum at 25 °C, and the residue, the reference saponins were used as an external standard to develop calibration curve for the quantitative estimation of saponins in *S. vaccaria* L. particles. Güçlü-Üstündag and Mazza (2007) used glycyrrhizic acid ammonium salt as an external standard for total saponin estimation.

#### 2.5.2. Formation of a calibration equation

HPLC chromatograms were obtained in duplicates for the standard solution varying in concentrations, and the mean total peak areas of saponins peaks were plotted against known concentrations to obtain a calibration equation, $y_{\text{saponin (mg mL}^{-1})} = (x_{\text{total area}} + 236.5)/1673$, $R = 0.99$. All bisdesmosidic saponins identified in the reference saponins following Balsevich, Bishop, and Deibert (2008) are shown in Fig. 1, and listed in Table 1. All saponins peaks appeared between 12 and 27 min of retention time, and were close to that reported by Güçlü-Üstündag and Mazza (2007).

### 2.6. Diffusion model

A DM was used with the experimental extraction data to estimate partition coefficient and effective diffusivity, and the extraction kinetics of saponins. This model was obtained by conversion of each term in the general mass balance equation incorporating the flat particle shapes only with the appropriate initial and boundary conditions. Since only flat particles were present in the sample, a single partial differential equation was discredited using finite difference method, and was solved using Gauss–Seidel method.

### 2.7. ANN models

Fig. 2 shows a three-layer back propagation ANN with an input column vector **x**, weight matrices **u**, **v**, and **w**, and the intermediate