COMPUTER-AIDED MODELLING FOR FLAVONOID SOLUBILITY PREDICTION USING COMBINED COSMO-RS AND UNIFAC

MOHD SHUKRI BIN MAT NOR

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (*Chemical Engineering*)

Faculty of Chemical and Energy Engineering
Universiti Teknologi Malaysia

JULY 2018

Specially dedicated to my beloved wife, daughters, son and family

Thanks for their endless love, support and encouragement

ACKNOWLEDGEMENT

In the name of Allah, the Most Gracious and the Most Merciful

All praises to Allah for bestowing the health, strengths and opportunity for me to experience this valuable path in completing this thesis. The knowledge and experiences gained throughout the process is a precious treasure that will not be forgotten. This thesis would not have been possible without the guidance, assistance and encouragements of kind-hearted people around me.

First and foremost, I would like to express my deepest gratitude to Prof. Ir. Dr. Zainuddin bin Abdul Manan, my main supervisor, who is also the Deputy Vice Chancellor (Academic and International) for UTM, for his dedication in guiding me throughout my PhD studies and the writing the thesis, despite his heavily occupied administrative work. I would like to express my heartiest thanks to my other two co-supervisors, Dr. Azizul Azri Mustaffa and Assoc. Prof. Dr. Chua Lee Suan, for their patience and guidance to share their knowledge with me, throughout these few years. Special thanks to Assoc. Prof. Dr. Mohamad Azmi Bustam for his support and facilities provide especially the usage of the COSMO-RS software during the research attachment program at UTP. In addition, I would like to thank to the Ministry of Higher Education of Malaysia for providing the financial support (R.J130000.7844.4F270) to complete this research.

Last but not least, I would like to express my boundless appreciation to my dearest father, Mat Nor Abdullah and my wife, Siti Nuurul Huda for their continuous support and encouragement. Also, to my daugthers and sons, Nur ‘Aisyatul Ibtisam, Nur Aleesyatuz Zyafeerah; Muhammad Ayyash Arrayan, and Muhammad Areeysh Azzyafran, thank you for the warm embraces and the joyful moments shared.

ABSTRACT

Flavonoids are groups of molecules with a broad spectrum of pharmacological activities such as antioxidant, anti-bacterial, anti-carcinogenic and anti-inflammatory properties. From the pharmaceutical point of view, the effectiveness of these compounds is largely controlled by their solubility to obtain acceptable bioavailability with minimal side effects in effective therapeutic dosages. The study of flavonoid solubility in solvent is important for effective extraction and better understanding of their physiochemical properties. The experimental works for flavonoids solubility measurement is laborious, time-consuming and costly. As a result, there is limited data on the solubility of flavonoids for the processing of flavonoid-based products. The prediction of solubility via solid-liquid equilibrium thermodynamic model is the method of choice to overcome these drawbacks. Therefore, the main objective of this study was to develop a new UNIFAC-based model assisted by COSMO-RS for predicting the solubility of flavonoids in solvents. The methodology of this study can be summarised into four main stages, namely, (1) data collection and database development of pure components (fusion enthalpy and melting temperature) and mixture (solubility and activity coefficient) properties, (2) UNIFAC-based model development, (3) model validation, and (4) model application (case studies). The missing data were determined using modeling approach after the accuracy of the model has been verified. Melting temperature was determined using improved Marrero and Gani model using stepwise and simultaneous regression methods, and fusion enthalpy data were calculated using original Marrero and Gani model, while solubility was computed using COSMO-RS computer-aided tool. The solubility data were regressed to determine new UNIFAC interaction parameters applicable for the case of flavonoids. This solubility model was validated against four datasets, two datasets from experimental work involving baicalein and kaempferol in methanol, ethanol and 1-propanol at various temperatures between 298.15 to 373.15 K and another two compounds from the literature (luteolin and apigenin). The validation results showed better predictions for all four datasets with confidence level higher than 94 %. The model was applied to two case studies involving solvent selections for flavonoids extraction and crystallisation. From the results of these case studies, the model shows reasonable accuracy and predictive capability with high confidence level in estimating the solubilities of flavonoids. As a conclusion, this study has proven that the proposed combination of COSMO-RS computer-aided and UNIFAC approaches can offer a new and reliable model for solubility prediction of flavonoids, thereby time saving and cost effective for product design and development.

ABSTRAK

Flavonoid merupakan sebatian molekul yang mengandungi aktiviti farmakologi dan tindak balas fisiologi yang menyeluruh melibatkan anti oksidan, anti bakteria, anti karsinogenik dan anti keradangan. Dalam bidang farmaseutikal, keberkesanan flavonoid untuk bioperolehan yang memuaskan, kesan sampingan yang minimum dan dos ubat yang efektif lazimnya dipengaruhi oleh kebolehlarutan sebatian. Kajian kebolehlarutan flavonoid sangat penting untuk memastikan proses pengekstrakan yang lebih efektif dan memahami sifat biofarmaseutikalnya.Walau bagaimanapun, eksperimen untuk menentukan kebolehlarutan flavonoid biasanya memerlukan masa, kos dan tenaga kerja yang tinggi. Ini mengakibatkan kurangnya data kebolehlarutan sebatian ini sedangkan ianya penting dalam pemprosesan produk berasaskan flavonoid. Salah satu cara yang digunakan untuk mengatasi masalah ini adalah melalui kaedah ramalan data kebolehlarutan menggunakan model termodinamik keseimbangan pepejal-cecair. Matlamat utama kajian ini adalah untuk menambahbaik model UNIFAC berbantukan perisian COSMO-RS bagi meramal data kebolehlarutan flavonoid dalam sesuatu pelarut. Kajian ini diringkaskan kepada empat peringkat, (1) pengumpulan data dan pembangunan pengkalan data melibatkan sifat komponen tulen (entalpi pelakuran dan suhu takat lebur) dan campuran komponen (kebolehlarutan dan pekali aktiviti), (2) pembangunan model berasaskan UNIFAC, (3) pengesahan model, dan (4) aplikasi model melalui kajian kes. Kelompongan data yang diperlukan boleh dikira menggunakan model yang telah dipastikan kejituannya. Data suhu takat lebur dikira menggunakan model Marrero dan Gani yang telah ditambahbaik melalui kaedah regresi langkah demi langkah dan regresi serentak, entalpi pelakuran dikira menggunakan kaedah asal Marrero dan Gani, data kebolehlarutan pula dikira menggunakan COSMO-RS iaitu kaedah berbantukan komputer. Data kebolehlarutan dikira berulangkali bagi mendapatkan parameter interaksi baharu yang terbaik, sesuai untuk flavonoid. Model kebolehlarutan ini disahkan menggunakan empat set data; dua set data diambil daripada eksperimen yang dijalankan sendiri melibatkan kebolehlarutan *baicalein* dan *kaempferol* di dalam metanol, etanol dan 1-propanol pada suhu di antara 298.15 K ke 373.15 K manakala dua set data lagi diambil daripada literatur (*luteolin* dan *apigenin*). Keputusan pengesahan model bagi keempat-empat set data menunjukkan nilai ramalan yang baik dengan paras keyakinan lebih daripada 94%. Model ini diaplikasikan dalam dua kajian kes melibatkan pemilihan pelarut bagi pengekstrakan dan pengkristalan flavonoid. Melalui kedua-dua kajian kes ini, kemampuan model dalam membuat ramalan data yang jitu bagi kebolehlarutan flavonoid telah dibuktikan dengan paras keyakinan yang tinggi. Kesimpulannya, kajian ini membuktikan bahawa model baharu yang dikemukakan untuk meramal data kebolehlarutan flavonoid melalui pendekatan gabungan antara COSMO-RS dan UNIFAC mampu menjimatkan masa dan kos terutamanya berkaitan pembangunan dan rekabentuk produk.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

COSMO-RS	-	Conductor-like screening model for a real solvent
DFT	-	Density functional theory
DMSO	-	Dimetyl sulfoxide
DSC	-	Differential scanning calorimetry
GSE	-	General solubility equation
HPLC	-	High performance liquid chromatography
ICAS	-	Integrated Computer Aided System
JR	-	Joback and Reid
LLE	-	Liquid-liquid equilibrium
MG	-	Marrero and Gani
MSDS	-	Material safety data sheet
NIST	-	National Institute of Standards and Technology
NRTL	-	Non-Random Two Liquid
PXRD	-	Powder X-ray diffraction
QSAR	-	Quantitative structure-activity relationship
QSPR	-	Quantitative structure-property relationship
SLE	-	Solid-liquid equilibrium
SMILES	-	Simplified molecular-input line-entry system
UNIFAC	-	Universal Functional Activity Coefficient
VLE	-	Vapour-liquid equilibrium

LIST OF SYMBOLS

a_{mn}	-	Adjustable binary interaction parameters between groups m and n
A_i	-	Group contributions of compound i
Aj	-	Group contributions of compound j
A_H	-	Slope of each empirical correlation for enthalpic term
A_S	-	Slope of each empirical correlation for entropic term
B_H	-	Intercept of each empirical correlation for enthalpic term
B_S	-	Intercept of each empirical correlation for entropic term
H_{fus0}	-	Universal constant of fusion enthalpy
H_{fus1i}	-	Fusion enthalpy contribution value of first order groups occurring in the compound
H_{fus2j}	-	Fusion enthalpy contribution value of second order groups occurring in the compound
H_{fus3k}	-	Fusion enthalpy contribution value of third order groups occurring in the compound
n	-	Number of consecutive repeat units in the series.
N	-	Total number of groups
N_i	-	Number of first order groups in the compound
N_i	-	Mumber of groups in which carbon element forms the centre of the group
N_j	-	Number of groups in which non-carbon element forms the center
m_i	-	Contribution of group i related to the enthalpy of melting
M_j	-	Number of second order groups in the compound
M_W	-	Molecular weight
n_i	-	Number of times a group i appears in a compound
O_k	-	Number of third order groups in the compound

P_k	-	Characterizes position factor
q_i	-	Van der Waals molecular volume of component i
q_j	-	Van der Waals molecular volume of component j
Q_k	-	Van der Waals surface area parameters
R	-	Ideal gas constant
r_i	-	Van der Waals molecular area of component i
r_j	-	Van der Waals molecular area of component j
R_k	-	Van der Waals volume parameters
R^2	-	Coefficient of determination
T	-	Temperature
$T_{m,i}$	-	Melting temperature of component i
T_{mo}	-	Universal constant of meting temperature
T_{m1i}	-	Melting temperature contribution value of first order groups occurring in the compound
T_{m2j}	-	Melting temperature contribution value of second order groups occurring in the compound
T_{m3k}	-	Melting temperature contribution value of third order groups occurring in the compound
$v_k^{(i)}$	-	Number of groups of k in component i
x_i	-	Solubility or mole fraction of component i
X_j^{exp}	-	Property experimental data
X_j^{pred}	-	Property predicted value
$\Delta C_{p,i}^{fus}$	-	Heat capacity of component i
ΔH_i^{fus}	-	Fusion enthalpy of component i
ΔS_i^{fus}	-	Total fusion entropy of component i
$\Delta S_{tot,i0}^{fus}$	-	Total fusion entropy of component i at 0 K
α	-	Confidence level
σ	-	Rotational symmetry number of a compound
ϕ_i	-	Segment fraction of component i
θ_i	-	Area fraction of component i

Φ	-	Molecular flexibility number of a compound
Θ_m	-	Area fraction of group m
Θ_n	-	Area fraction of group n
Γ_k	-	Residual activity coefficient of group k in the mixture
$\Gamma_k^{(i)}$	-	Residual activity coefficient of group k in a solution of pure component i.
γ_i	-	Activity coefficient of component i
γ_i	-	Activity coefficient of component i
γ_i^C	-	Combinatorial term of activity coefficient of component i
γ_i^R	-	Residual term of activity coefficient of component i
Ψ_{mk}	-	Boltzmann factor between groups m and k
Ψ_{km}	-	Boltzmann factor between groups k and m
Ψ_{mn}	-	Boltzmann factor between groups m and n

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Flavonoids are widely used phytochemicals in the pharmaceutical, food, and pigment industries. This chapter provides an outlook of the flavonoid compounds followed by an introduction of the research background and problem statement. Next, the research objectives and the scope of this work related to the development of new UNIFAC-based model to predict the flavonoids solubility are described. Finally, this chapter highlights three (3) key contributions of this thesis that are relevant to the research field, and the potential application of the developed method in the pharmaceutical and bioproduct industries.

1.2 Research Background

Phytochemicals are chemical compounds extracted from plants. “Phyto” is the prefix derived from a Greek word, meaning plant. Phytochemicals are used to protect plants against bacteria, viruses, fungi, and other climate stimuli as well as to provide colour, aroma and flavour. Researchers have demonstrated that phytochemicals can be used to protect humans against diseases (Rao, 2003). Historically, phytochemicals was first consumed as a traditional medicine for many centuries and across civilisations, from the west to the east in the form of Greek medicine, Indian Ayurvedic medicine, Japanese Kampo medicine, traditional Chinese medicine

(represented by acupuncture and Chinese herbal medicine), and Unani medicine in the South Asia and Middle East (Mosihuzzaman and Choudhary, 2008). Evidence of herbal remedies being used by Neanderthal man around sixty thousand years ago was found in an uncovered burial site of a cave in northern Iraq in 1960 (Kunle *et al*., 2012).

Phytochemicals usage for the prevention and treatment of various illnesses is drawing worldwide attention. The synthetic allopathic drugs may also cause unavoidable side-effects. Islam *et al*. (2015) reported that 121 active compounds currently being used are derived from plants. This makes up about 25 % from all drugs prescribed worldwide. It includes several well-known allopathic medicines such as morphine, digitoxin, quinidine, taxol, quinine, colchicine, digoxin, and L-hyoscyamine. It is believed that the therapeutic effects can be attributed to the bioactive compounds that exist naturally in plants. The process of tracing these bioactive compounds is known as phytochemical screening. The screening has revealed the presence of four phytochemical classes which are phenolics, terpernoids (Saxena *et al*., 2013), nitrogen containing compounds (Wink, 2004) and organosulfurs (Vazquez-Prieto and Miatello, 2010) as the major group of phytochemicals in the successive extracts of different plants. A summary of the phytochemicals classification is shown in Figure 1.1. The classification may also be derived from their biosynthetic origin, such as terpenoids resulting from the condensation of a varying number of isoprene units formed through the mevalonate pathway. As a result, most phytochemical classification schemes are based on chemical structure definitions.

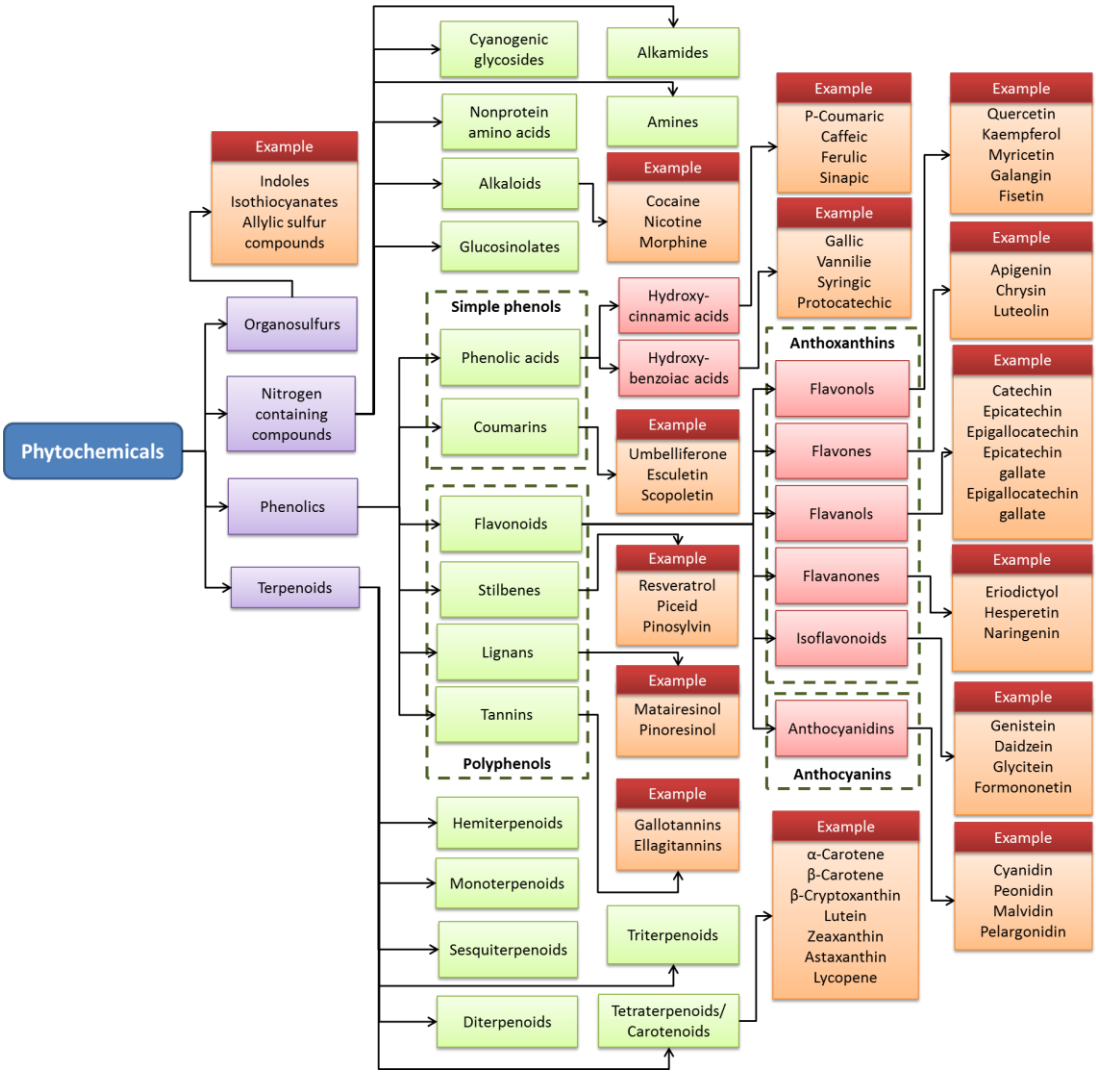

Figure 1.1 Classification of phytochemicals

Flavonoids are one the most studied compounds among the numerous classes of phytochemicals that are present in plants (Chaaban *et al*., 2017) and known as one of the important plant constituents that are consumed in diet (Hertog *et al.*, 1992) and applied in medicine. They are believed to have many therapeutic capabilities and potential to reduce the risk of multiple diseases. They have been utilised to enhance human health via their multiple biological activities including antimicrobial (Cushnie and Lamb, 2005), anti-inflammatory (Yamamoto and Gaynor, 2001), antioxidant (Shahidi *et al.*, 1992), the prevention of osteoporosis (Migliaccio and Anderson, 2003) and anticancer functions (Jose *et al*., 2014). Their potential role as anticancer agents have been intensely explored in the medication of breast, ovarian, cervical,

prostate cancer, and pancreatic (Liu *et al.*, 2010). Due to these vast potentials, flavonoids have driven the development of many processes for the manufacture of flavonoid derived products. The processes that are typically employed in the industrial production of flavonoids (pharmaceutical, nutraceutical, food etc.) are extraction, formulation and crystallisation. According to the new report published by Grand View Research, global demand for flavonoids market was valued at over USD 410.1 million in 2015 and is forecasted to reach above USD 1.06 billion by 2025 (Haven and Ranjan, 2017).

It is important to have fundamental physicochemical property data for the flavonoids in order to optimise these process designs. One of the most important properties that play an essential role in these processes is solubility. It is crucial to know the flavonoids solubility in various solvents to select the proper solvents and design an optimised production process. The solubility can be determined experimentally via solubility measurement techniques such as batch method (Jouyban *et al*., 2001), flow column method (Wasik *et al*., 1981), the potentiometric, the micro solubility self-calibrating direct UV, miniaturized shake-flask microtitre plate, the micro dissolution methods (Dressman and Reppas, 2007), gravimetric shake flask method (Liu *et al*., 2014) and synthetic method (Wang *et al*., 2006). The work on measurement of the flavonoid solubilities in solvents has been growing in the literature but the technical data is still insufficient to support the process design. As a matter of fact, experimental solubility measurement methods are expensive especially raw material cost, laborious, and often technically problems due to the lack of experience in analytical skills. For instance, in the shake flask method, saturated solutions are prepared at a certain temperature with a surplus of solid solute. The suspension is centrifuged or filtrated to ensure the solution is fully free from undissolved solute. Centrifugation or filtration steps are often conducted at a lower temperature than the saturation experiments, leading to the variance in the degree of precision and accuracy of obtained values.

Thus, one of the attractive directions is to employ the solid-liquid equilibrium (SLE) thermodynamic model. This model is commonly used for predicting solubility

through the relative importance between the melting properties (melting temperature and fusion enthalpy), heat capacity and activity coefficient. These critical data are needed to develop an efficient process model that can be integrated into process design methods similar to those that have been established for the petrochemical and palm-oil-based oleo-chemical industries (Mohammad Azmin *et al*., 2016). The thermodynamic framework for solubility modelling suggested here is based on the Universal Functional-group Activity Coefficients (UNIFAC) correlation for estimating activity coefficients from group contributions. UNIFAC is a well-known model for the estimation of molecule activity coefficients in non-ideal mixtures developed by Fredenslund *et al*. (1975).

To date, research to establish solubility of flavonoids in solvents has been very scarce due to the lack of physicochemical property data. From the literature search, there is very little reported publication related to the flavonoids modeling and experimental works. Previously, Guo *et al*. (2007) predicted the flavonoids solubility in a broad variety of ionic liquids with over 1800 available structures which were investigated based on the conductor-like screening model for a real solvent (COSMO-RS) computation, while Chebil *et al*. (2007) used the Austin Model 1 method carried out in HyperChem 7 molecular simulation and visualization program for the prediction of flavonoids (quercetin, chrysin, rutin, naringenin, isoquercitrin and hesperetin) solubility. Hansen solubility parameters have also been applied for optimizing the solubility in extraction of flavonoids from grape and berry substrates (Srinivas *et al*., 2009). In addition, Ferreira and Pinho (2012) applied NRTL-SAC to predict flavonoids solubility for apigenin, genistein, hesperetin and luteolin in pure solvent. Sevillano *et al*. (2014) proposed MPP-UNIFAC for estimating polyphenols using flavonoids and stilbenoids as a regression dataset. In fact, most of the flavonoids solubility predictions were simply correlated using non-predictive empirical models such as the Apelblat or Modified Apelblat equation such as for quercetin and quercetin dehydrate hesperetin (Liu and Chen, 2008), luteolin (Peng and Yan, 2009), (Srinivas *et al*., 2010), genistein (Wu *et al*., 2010), apigenin (Xiao *et al*., 2010) and chrysin (Zhou *et al*., 2014).

In this work, the flavonoids solubility data was generated using COSMO-RS related to quantum chemistry calculations since there are very limited experimental data in literature for parameter regression. Previously, the generated COSMO-RS data was compared with available literature data in preliminary study and the results were very promising for alcohol solvents. The produced COSMO-RS data was used for UNIFAC parameter regression. This combination of modeling between UNIFAC and COSMO-RS approaches is expected to contribute towards the establishment of the vital solubility database and prediction thermodynamic model of flavonoids.

Throughout this research, the questions that will be answered are as follows:

i. Is there a systematic and effective framework to model the solubility of components such as flavonoids that are typically complex and have scarce physicochemical property data?
ii. Can a thermodynamic SLE model approach be used to guide designers to find the optimum flavonoids solubility in extraction, formulation and crystallisation processes?
iii. What are the relevant property relationships between a solvent and flavonoids to estimate their solubilities?
iv. Can UNIFAC model combined with COSMO-RS be used to predict flavonoids solubility with less resources and effort?
v. How accurate would be these models and predictive methods?

1.3 Problem Statement

The solubility of the solutes in solvents plays a vital role in the operation and development of extraction, formulation and crystallisation process. Therefore, knowing the product solubility is a necessary method in optimising and designing those processes properly. However, there are currently very limited data available on the solubility of flavonoids that are vital for the design and optimisation of industrial processes. The work on flavonoid solubilities measurement has been growing in the

literature but the data is still inadequate to support the process design. It is also well-known that the flavonoid standards (raw materials) are quite expensive because of the typically small quantity of flavonoids that are recoverable from the appropriate sources. As a result, process design and optimisation tasks are currently performed on trial-and-error basis based on researchers' experiences through case studies (Aziz *et al*., 2003). The disadvantage of the trial-and-error method is that the number of experimental repetitions is unknown and might require a lot of work and time. This will lead to high cost of raw materials, utilities, time, and labour.

Since solubility experiments are very expensive because of the raw materials cost, SLE thermodynamic model which allows for solubility predictions in solvents is of great importance for increasing extraction efficiency. A comprehensive database with solubilities, activity coefficient and pure component data (heat capacity, melting temperature and fusion enthalpy) is required for model development. Sometimes, error is quite high for the prediction of the melting temperature, especially for complex molecules. To improve the model, a regression on melting temperature data needs to be conducted specifically for flavonoids. Until now, no study has examined the potential of MG group contribution model for predicting melting temperature, particularly for flavonoids. Numerous models for the prediction of solubility have been developed such as the Hansen solubility parameter (Hansen, 1967), UNIFAC, COSMO-RS, - Non-Random-Two-Liquid (NRTL) (Renon and Prausnitz, 1969), the general solubility equation (GSE) (Ran *et al*., 2001) and many more. One of the most commonly used and well-established methods is UNIFAC. Nevertheless, UNIFAC accuracy is depending on the contribution parameters determined by regression of experimental data collected in large databases.

One of the main issues which can be improved in UNIFAC prediction of flavonoids solubility is the regression data size of interaction parameters. A common point in the case of flavonoid molecules is that the experimental data are needed for the model parameters regression but the measurement is often expensive, time-consuming and may even be infeasible in some cases. COSMO-RS has been introduced as a full predictive model-based approach to produce the data for

UNIFAC interaction parameters regression. The main advantage of the COSMO-RS is that it uses only data from quantum chemical calculations, thus enabling predictions when there are no available experimental data. The model uses the sigma profile (a molecule-specific distribution of the charge density) as input for the calculation of solubility. The most time-consuming aspect of the method is the generation of sigma profiles, and also required quite expensive license to perform the software. To date, no study has used COSMO-RS to generate flavonoids solubility data for UNIFAC parameter regression. In UNIFAC, the parameters are determined with the contribution of the diverse functional groups in flavonoids. Even though the interaction parameters between those groups have been regressed from COSMO-RS data, they still can be applied to other flavonoids and solvents not evaluated in the initial dataset. UNIFAC strength lies on its predictive ability without the need of previous experimental effort.

Hence, there is a need to develop a combination of UNIFAC and COSMO-RS approach that could predict the solubility of flavonoids in solvents in high accuracy. This will indirectly contribute towards the development of high value flavonoids based product. Following is the problem statement of this research:

"Given flavonoids with a set of solubility data, and each flavonoid solubility differs based on the range of temperature and solvent used, it is desired to study the relationship between solvent and solute, and develop UNIFAC-based model that can predict the solubility of flavonoids in solvents at different temperatures. In addition, the model developed must be fast, reliable, and cost-effective method to determine the solubility at any temperature for extraction, formulation and crystallisation of flavonoid components. This UNIFAC based model uses a computer aided approach (COSMO-RS) in generating the solubility data. The method involves development of COSMO-RS solubility for flavonoids, prediction of SLE pure properties and regression of new UNIFAC parameters. This UNIFAC based model then is verified with experimental work in order to validate the developed model in terms of accuracy and hence, reliability".

1.4 Objective of the Study

The primary aim of this research is to establish new UNIFAC-based parameters and model in predicting the solubility of flavonoids in common solvents. The main objective is supported by four specific sub-objectives:

i. To predict flavonoids solubility via solid-liquid equilibrium.
ii. To develop a UNIFAC-based model to predict the activity coefficient that is a vital parameter to be used in solubility prediction of flavonoids.
iii. To generate flavonoids solubility and activity coefficient data using COSMO-RS tool that is important in regression of UNIFAC interaction paramenters.
iv. To validate the developed UNIFAC-based solubility model through experimental study, as well as experimental data from the literature.
v. To apply the developed UNIFAC-based model to related case studies.

1.5 Scope of the Study

The scope of study contains:

i. State-of-the-art review and analyses of flavonoids physicochemical property models (pure and mixture properties) and their connections in the SLE thermodynamic model.
ii. Application of MG method in predicting the fusion enthalpy (vital property in SLE calculation) for flavonoids.
iii. Improvement of MG melting temperature (vital property in SLE calculation) model for flavonoids by introducing new group contibutions.
iv. Development of fusion enthalpy and melting temperature (pure component properties) database, collected either from literature/ open database or calculated using MG models for unavailable published data.
v. Development of solubility and activity coefficient (mixture properties) database, collected from literature for various solvents and computed using COSMO-RS for common alcoholic solvents (methanol, ethanol and 1-propanol) at temperature range from 298.15 to 373.15 K.

vi. Regression of new UNIFAC interaction parameters using developed property database for flavonoids solubility

vii. Solubility measurements for baicalein and kaempferol (in methanol, ethanol and 1-propanol at 298.15 to 373.15 K) using shake flask method for the purpose of validation of the UNIFAC based model.

viii. Improvement of processes in extraction and crystallisation through solvent selection as case studies.

1.6 Significance of the Study

Due to the time and cost intensiveness of performing temperature-dependent measurements for many different binary systems, the availability of a reliable method to predict this property is of foremost importance. Therefore a fully predictive model, able to well characterise the behaviour of such solubility systems, would be a smart alternative.

This study would be of significant interest to scholars, as well as practitioners, particularly in the efficient design of flavonoid products and processes to produce high value products in Malaysia. In addition, the understanding of the concept of solubility provides promising commercial benefits for industry. From a practice standpoint, this study is relevant and timely, especially for the pharmaceuticals, cosmetics and food sector.

1.7 Contributions of the Study

The key specific contributions of this study are compiled as follows:

i. Creation of flavonoids physicochemical property database: Flavonoid databases including pure component (melting temperature (from open

REFERENCES

Abildskov, J.; Constantinou, L. and Gani, R. (1996). Towards the development of a second-order approximation in activity coefficient models based on group contributions. *Fluid phase equilibria,* 118**,** 1-12.

Abramov, Y. A. (2015). Major source of error in QSPR prediction of intrinsic thermodynamic solubility of drugs: solid vs nonsolid state contributions? *Molecular pharmaceutics,* 12**,** 2126-2141.

Abramowitz, R. and Yalkowsky, S. H. (1990). Melting point, boiling point, and symmetry. *Pharmaceutical research,* 7**,** 942-947.

Abrams, D. S. and Prausnitz, J. M. (1975). Statistical thermodynamics of liquid mixtures: a new expression for the excess Gibbs energy of partly or completely miscible systems. *AIChE Journal,* 21**,** 116-128.

Agarwal, R.; Katiyar, S. K.; Zaidi, S. I. and Mukhtar, H. (1992). Inhibition of skin tumor promoter-caused induction of epidermal ornithine decarboxylase in SENCAR mice by polyphenolic fraction isolated from green tea and its individual epicatechin derivatives. *Cancer Research,* 52**,** 3582-3588.

Aguda, R. and Chen, C.-C. (2016). Solubility of Nutraceutical Compounds in Generally Recognized as Safe Solvents at 298 K. *International Journal of Chemical Engineering and Applications,* 7**,** 289.

Alevizou, E. I. and Voutsas, E. C. (2013). Solubilities of p-coumaric and caffeic acid in ionic liquids and organic solvents. *The Journal of Chemical Thermodynamics,* 62**,** 69-78.

Alminger, M.; Aura, A. M.; Bohn, T.; Dufour, C.; El, S.; Gomes, A.; Karakaya, S.; Martínez-Cuesta, M.; Mcdougall, G. and Requena, T. (2014). In vitro models for studying secondary plant metabolite digestion and bioaccessibility. *Comprehensive Reviews in Food Science and Food Safety,* 13**,** 413-436.

Augustijns, P. and Brewster, M. E. (2007). *Solvent systems and their selection in pharmaceutics and biopharmaceutics.* Springer.

Austin, J. B. (1930). A relation between the molecular weights and melting points of organic compounds. *Journal of the American Chemical Society,* 52**,** 1049-1053.

Avdeef, A. (2001). Physicochemical profiling (solubility, permeability and charge state). *Current topics in medicinal chemistry,* 1**,** 277-351.

Aziz, R. A.; Sarmidi, M. R.; Kumaresan, S.; Taher, Z. M. and Foo, D. (2003). Phytochemical Processing: The Next Emerging Field in Chemical Engineering–Aspects and Opportunities. *Jurnal Kejuruteraan Kimia Malaysia,* 3**,** 45-60.

Baka, E.; Comer, J. E. and Takács-Novák, K. (2008). Study of equilibrium solubility measurement by saturation shake-flask method using hydrochlorothiazide as model compound. *Journal of pharmaceutical and biomedical analysis,* 46**,** 335-341.

Bard, B. (2008). *Development of new experimental tools for fast determination of solubility and lipophilicity.* University of Geneva.

Bard, B.; Martel, S. and Carrupt, P.-A. (2008). High throughput UV method for the estimation of thermodynamic solubility and the determination of the solubility in biorelevant media. *european journal of pharmaceutical sciences,* 33**,** 230-240.

Batra, P. and Sharma, A. K. (2013). Anti-cancer potential of flavonoids: recent trends and future perspectives. *3 Biotech,* 3**,** 439-459.

Bevan, C. D. and Lloyd, R. S. (2000). A high-throughput screening method for the determination of aqueous drug solubility using laser nephelometry in microtiter plates. *Analytical chemistry,* 72**,** 1781-1787.

Bhagwat, S.; Haytowitz, D. B. and Holden, J. M. (2014). USDA database for the flavonoid content of selected foods, Release 3.1. *US Department of Agriculture: Beltsville, MD, USA.*

Bondi, A. A. (1968). Physical properties of molecular crystals liquids, and glasses.

Bouillot, B.; Teychené, S. and Biscans, B. (2011). An evaluation of thermodynamic models for the prediction of drug and drug-like molecule solubility in organic solvents. *Fluid Phase Equilibria,* 309**,** 36-52.

Box, K. J.; Völgyi, G.; Baka, E.; Stuart, M.; Takács-Novák, K. and Comer, J. E. (2006). Equilibrium versus kinetic measurements of aqueous solubility, and the ability of compounds to supersaturate in solution—a validation study. *Journal of pharmaceutical sciences,* 95**,** 1298-1307.

Brickman, A. M.; Khan, U. A.; Provenzano, F. A.; Yeung, L.-K.; Suzuki, W.; Schroeter, H.; Wall, M.; Sloan, R. P. and Small, S. A. (2014). Enhancing dentate gyrus function with dietary flavanols improves cognition in older adults. *Nature neuroscience,* 17**,** 1798-1803.

Cassidy, A.; O'reilly, É. J.; Kay, C.; Sampson, L.; Franz, M.; Forman, J.; Curhan, G. and Rimm, E. B. (2011). Habitual intake of flavonoid subclasses and incident hypertension in adults. *The American journal of clinical nutrition,* 93**,** 338-347.

Chaaban, H.; Ioannou, I.; Chebil, L.; Slimane, M.; Gérardin, C.; Paris, C.; Charbonnel, C.; Chekir, L. and Ghoul, M. (2017). Effect of heat processing on thermal stability and antioxidant activity of six flavonoids. *Journal of Food Processing and Preservation*.

Charton, M. and Charton, B. (1994). Quantitative description of structural effects on melting points of substituted alkanes. *Journal of Physical Organic Chemistry,* 7**,** 196-206.

Chebil, L.; Humeau, C.; Anthoni, J.; Dehez, F.; Engasser, J.-M. and Ghoul, M. (2007). Solubility of flavonoids in organic solvents. *Journal of Chemical & Engineering Data,* 52**,** 1552-1556.

Chen, C.-C. and Song, Y. (2004). Solubility modeling with a nonrandom two-liquid segment activity coefficient model. *Industrial & engineering chemistry research,* 43**,** 8354-8362.

Chen, M.; Xie, C. and Liu, L. (2010). Solubility of Andrographolide in Various Solvents from (288.2 to 323.2) K. *Journal of Chemical & Engineering Data,* 55**,** 5297-5298.

Chen, T.; Shen, H. and Zhu, C. (2002). Evaluation of a method for high throughput solubility determination using a multi-wavelength UV plate reader. *Combinatorial chemistry & high throughput screening,* 5**,** 575-581.

Chickos, J. S. and Acree, W. E. (2009). Total phase change entropies and enthalpies. An update on fusion enthalpies and their estimation. *Thermochimica Acta,* 495**,** 5-13.

Chickos, J. S. and Nichols, G. (2001). Simple relationships for the estimation of melting temperatures of homologous series. *Journal of Chemical & Engineering Data*, 46**,** 562-573.

Constantinou, L. and Gani, R. (1994). New group contribution method for estimating properties of pure compounds. *AIChE Journal*, 40**,** 1697-1710.

Cushnie, T. and Lamb, A. J. (2005). Antimicrobial activity of flavonoids. *International journal of antimicrobial agents*, 26**,** 343-356.

De Lemos, M. L. (2001). Effects of soy phytoestrogens genistein and daidzein on breast cancer growth. *Ann. Pharmacother*, 35**,** 1118–1121.

Dearden, J. and Rahman, M. (1988). QSAR approach to the prediction of melting points of substituted anilines. *Mathematical and Computer Modelling*, 11**,** 843-846.

Di, L. and Kerns, E. H. (2015). *Drug-like properties: concepts, structure design and methods from ADME to toxicity optimization.* Academic Press.

Diedrichs, A. and Gmehling, J. R. (2010). Solubility calculation of active pharmaceutical ingredients in alkanes, alcohols, water and their mixtures using various activity coefficient models. *Industrial & Engineering Chemistry Research*, 50**,** 1757-1769.

Dressman, J. and Reppas, C. (2007). Drug solubility: how to measure it, how to improve it. *Advanced Drug Delivery Reviews*, 59**,** 531-532.

E. Mullins; Y. Liu; A. Ghaderi and S.D. Fast. (2008). Sigma profile database for predicting solid solubility in pure and mixed solvent mixtures for organic pharmacological compounds with COSMO-based thermodynamic methods. *Industrial & Engineering Chemistry Research*, 47**,** 1707–1725.

Fang, J.; Zhang, M.; Zhu, P.; Ouyang, J.; Gong, J.; Chen, W. and Xu, F. (2015). Solubility and solution thermodynamics of sorbic acid in eight pure organic solvents. *The Journal of Chemical Thermodynamics*, 85**,** 202-209.

Feizi, S.; Jabbari, M. and Farajtabar, A. (2016). A systematic study on solubility and solvation of bioactive compound chrysin in some water+ cosolvent mixtures. *Journal of Molecular Liquids*, 220**,** 478-483.

Ferreira, O. and Pinho, S. O. P. (2012). Solubility of flavonoids in pure solvents. *Industrial & Engineering Chemistry Research*, 51**,** 6586-6590.

Fredenslund, A.; Jones, R. L. and Prausnitz, J. M. (1975). Group Contribution Estimation of Activity Coefficients in Non-ideal Liquid Mixtures. *AlChE Journal,* 21**,** 1086 - 1099.

Gabas, N. and Laguérie, C. (1993). Predictions with UNIFAC of liquid-solid phase diagrams: application to water-sucrose-glucose, water-sucrose-fructose and water-xylose-mannose. *Journal of crystal growth,* 128**,** 1245-1249.

Gani, R. (2004). Computer-aided methods and tools for chemical product design. *Chemical Engineering Research and Design,* 82**,** 1494-1504.

Gharagheizi, F. and Salehi, G. R. (2011). Prediction of enthalpy of fusion of pure compounds using an artificial neural network-group contribution method. *Thermochimica acta,* 521**,** 37-40.

Gmehling, J.; Tiegs, D. and Knipp, U. (1990). A comparison of the predictive capability of different group contribution methods. *Fluid Phase Equilibria,* 54**,** 147-165.

Gmehling, J. G.; Anderson, T. F. and Prausnitz, J. M. (1978). Solid-Liquid Equilibria Using UNIFAC. *Industrial & Engineering Chemistry Fundamentals,* 17**,** 269-273.

Goodarzi, M.; Chen, T. and Freitas, M. P. (2010). QSPR predictions of heat of fusion of organic compounds using Bayesian regularized artificial neural networks. *Chemometrics and Intelligent Laboratory Systems,* 104**,** 260-264.

Gracin, S.; Brinck, T. and Rasmuson, Å. C. (2002). Prediction of solubility of solid organic compounds in solvents by UNIFAC. *Industrial & engineering chemistry research,* 41**,** 5114-5124.

Gross, J. and Sadowski, G. (2001). Perturbed-chain SAFT: An equation of state based on a perturbation theory for chain molecules. *Industrial & engineering chemistry research,* 40**,** 1244-1260.

Guo, Z.; Lue, B.-M.; Thomasen, K.; Meyer, A. S. and Xu, A. X. (2007). Predictions of flavonoid solubility in ionic liquids by COSMO-RS: experimental verification, structural elucidation, and solvation characterization. *Green Chem.,* 9**,** 1362-1373.

Gupta, R. B. and Heidemann, R. A. (1990). Solubility models for amino acids and antibiotics. *AIChE Journal,* 36**,** 333-341.

Han, R. M.; Tian, Y. X.; Liu, Y.; Chen, C. H.; Ai, X. C.; Zhang, J. P. and Skibsted, L. H. (2009). Comparison of flavonoids and isoflavonoids as antioxidants. *J. Agric. Food Chem,* 57**,** 3780–3785.

Hansen, C. M. (1967). The three-dimensional solubility parameter-key to paint component affinities: solvents, plasticizers, polymers, and resins. II. Dyes, emulsifiers, mutual solubility and compatibility, and pigments. III. Independent cal-culation of the parameter components. *Journal of paint technology,* 39**,** 505-510.

Hansen, H. K.; Fredenslund, P.; Schiller, M. and Gmehling, J. (1991a). Vapor-Liquid Equilibria by UNIFAC Group Contribution. 5. Revision and Extension. *Industrial & Engineering Chemistry Research,* 30**,** 2352 - 2355.

Hansen, H. K.; Rasmussen, P.; Fredenslund, A.; Schiller, M. and Gmehling, J. (1991b). Vapor-liquid equilibria by UNIFAC group contribution. 5. Revision and extension. *Industrial & Engineering Chemistry Research,* 30**,** 2352-2355.

Hao, H.-X.; Hou, B.-H.; Wang, J.-K. and Zhang, M.-J. (2005). Solubility of erythritol in different solvents. *Journal of Chemical & Engineering Data,* 50**,** 1454-1456.

Harjo, B.; Wibowo, C. and Ng, K. M. (2004). Development of natural product manufacturing processes: Phytochemicals. *Chemical Engineering Research and Design,* 82**,** 1010-1028.

Haven, B. and Ranjan, B. (2017). *Flavonoids Market Size To Reach $1.06 Billion By 2025* [Online]. Available: https://www.grandviewresearch.com/press-release/global-flavonoids-market [Accessed 1 Jan. 2017.

Hertog, M.; Hollman, P. and Vandeputte, B. (1993). Content of potentially anticarcinogenic flavonoids of tea infusions, wines, and fruit juices. *J Agr Food Chem,* 41**,** 1242–1246.

Hertog, M. G.; Hollman, P. C. and Venema, D. P. (1992). Optimization of a quantitative HPLC determination of potentially anticarcinogenic flavonoids in vegetables and fruits. *Journal of Agricultural and Food Chemistry,* 40**,** 1591-1598.

Higuchi, T. and Connors, A. (1965). Phase-solubility techniques.

Hukkerikar, A. S.; Sarup, B.; Kate, A. T.; Abildskov, J.; Sin, G. and Gani, R. (2012a). Group-contribution+ (GC+) based estimation of properties of pure

components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilibria,* 321**,** 25 - 43.

Hukkerikar, A. S.; Sarup, B.; Ten Kate, A.; Abildskov, J.; Sin, G. and Gani, R. (2012b). Group-contribution+(GC+) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilibria,* 321**,** 25-43.

Indofine Chemical, C. (2016). *The Flavonoid Company* [Online]. Available: indofinechemical.com/product-details.aspx?Sku=021127S [Accessed 22 March 2016.

Islam, M. T.; Guha, B.; Uddin, M. A.; Mutsuddy, R.; Khastagir, A.; Junaid, M. S. A.; De Carvalho Melo-Cavalcante, A. A. and De Freitas, R. M. (2015). Requesting a complete biosensor system in phyto-source drug discovery and development. *IOSR J Pharm,* 5**,** 32-41.

Jain, A. and Yalkowsky, S. H. (2006). Estimation of melting points of organic compounds-II. *Journal of pharmaceutical sciences,* 95**,** 2562-2618.

Joback, K. G. and Reid, R. (1987). Estimation of pure-component properties from group-contributions. *Chem. Eng. Commun.,* 57**,** 233-234.

Jose, J.; Sudheesh, S.; Sumesh Kumar, T.; Sony, J. and Jayadevi Variyar, E. (2014). A comparative evaluation of anticancer activities of flavonoids isolated from Mimosa pudica, Aloe vera, Phyllanthus niruri against human breast carcinoma cell line (MCF-7) using MTT assay. *Int J Pharm Pharm Sci,* 6**,** 319-22.

Jouyban-Gharamaleki, V.; Jouyban-Gharamaleki, K.; Shayanfar, A.; Khoubnasabjafari, M. and Jouyban, A. (2015). An Automated System for Determining Drug Solubility Based on Laser Monitoring Technique. *Journal of laboratory automation,* 20**,** 3-9.

Jouyban, A. (2008). Review of the cosolvency models for predicting solubility of drugs in water-cosolvent mixtures. *Journal of Pharmacy and Pharmaceutical Sciences,* 11**,** 32-58.

Jouyban, A. (2009). *Handbook of solubility data for pharmaceuticals.* CRC Press.

Jouyban, G.; York, P.; Hanna, M. and Clark, B. (2001). Solubility prediction of salmeterol xinafoate in water–dioxane mixtures. *International journal of pharmaceutics,* 216**,** 33-41.

Kan, A. and Tomson, M. (1996). UNIFAC prediction of aqueous and nonaqueous solubilities of chemicals with environmental interest. *Environmental science & technology,* 30**,** 1369-1376.

Kang, J. W.; Abildskov, J. and Gani, R. (2002). Estimation of Mixture Properties from First- and Second-Order Group Contributions with the UNIFAC Model. *Industrial & Engineering Chemistry Research,* 41**,** 3260 - 3273.

Kang, J. W.; Diky, V.; Chirico, R. D.; Magee, J. W.; Muzny, C. D.; Abdulagatov, I.; Kazakov, A. F. and Frenkel, M. (2011). A new method for evaluation of UNIFAC interaction parameters. *Fluid Phase Equilibria,* 309**,** 68 - 75.

Karthikeyan, M.; Glen, R. C. and Bender, A. (2005). General melting point prediction based on a diverse compound data set and artificial neural networks. *Journal of chemical information and modeling,* 45**,** 581-590.

Katritzky, A. R. and Gordeeva, E. V. (1993). Traditional topological indexes vs electronic, geometrical, and combined molecular descriptors in QSAR/QSPR research. *Journal of chemical information and computer sciences,* 33**,** 835-857.

Katritzky, A. R.; Lomaka, A.; Petrukhin, R.; Jain, R.; Karelson, M.; Visser, A. E. and Rogers, R. D. (2002). QSPR correlation of the melting point for pyridinium bromides, potential ionic liquids. *Journal of chemical information and computer sciences,* 42**,** 71-74.

Katritzky, A. R.; Maran, U.; Karelson, M. and Lobanov, V. S. (1997). Prediction of melting points for the substituted benzenes: A QSPR approach. *Journal of chemical information and computer sciences,* 37**,** 913-919.

Kaufman, P. B.; Duke, J. A.; Brielmann, H. and Boik, J. H., J. E. (1997). A comparative survey of leguminous plants as sources of the isoflavones, genistein and daidzein: implications for human nutrition and health. *J. Altern. Complement. Med,* 3**,** 7-12.

Kawakami, K.; Miyoshi, K. and Ida, Y. (2005). Impact of the amount of excess solids on apparent solubility. *Pharmaceutical research,* 22**,** 1537-1543.

Kerns, E. H.; Di, L. and Carter, G. T. (2008). In vitro solubility assays in drug discovery. *Current drug metabolism,* 9**,** 879-885.

Kier, L. (2012). *Molecular connectivity in chemistry and drug research.* Elsevier.

Klamt, A. (1995). Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. *The Journal of Physical Chemistry,* 99**,** 2224-2235.

Klamt, A. and Schüürmann, G. (1993). COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of the Chemical Society, Perkin Transactions 2***,** 799-805.

Kolár, P.; Shen, J.-W.; Tsuboi, A. and Ishikawa, T. (2002). Solvent selection for pharmaceuticals. *Fluid Phase Equilibria***,** 194–197.

Kontogeorgis, G. M. and Gani, R. (2004). *Computer Aided Property Estimation for Process and Product Design: Computers Aided Chemical Engineering.* Elsevier.

Krzyzaniak, J. F.; Myrdal, P. B.; Simamora, P. and Yakowsky, S. H. (1996). Boiling Point and Melting Point Prediction for Aliphatic , Non-Hydrogen-Bonding Compounds. *Ind. Eng. Chem. Res.,* 34**,** 2530-2535.

Kumari, A.; Sandeepa, K.; Prathap Kumar, T. and Satyavathi, B. (2015). Solubility, thermodynamic properties, and derived excess properties of Benzoic Acid in (Acetic Acid+ Water) and (Acetic Acid+ Toluene) binary mixtures. *Journal of Chemical & Engineering Data,* 61**,** 67-77.

Kunle, O. F.; Egharevba, H. O. and Ahmadu, P. O. (2012). Standardization of herbal medicines-A review. *International Journal of Biodiversity and Conservation,* 4**,** 101-112.

Kuramochi, H.; Noritomi, H.; Hoshino, D. and Nagahama, K. (1996). Measurements of solubilities of two amino acids in water and prediction by the UNIFAC model. *Biotechnology progress,* 12**,** 371-379.

Larsen, B. L.; Rasmussen, P. and Fredenslund, A. (1987a). A Modified UNIFAC Group-Contribution Model for Prediction of Phase Equilibria and Heats of Mixing. *Industrial & Engineering Chemistry Research,* 26**,** 2274 - 2286.

Larsen, B. L.; Rasmussen, P. and Fredenslund, A. (1987b). A modified UNIFAC group-contribution model for prediction of phase equilibria and heats of mixing. *Industrial & engineering chemistry research,* 26**,** 2274-2286.

Le Marchand, L.; Murphy, S. P.; Hankin, J. H.; Wilkens, L. R. and Kolonel, L. N. (2000). Intake of flavonoids and lung cancer. *Journal of the National Cancer Institute,* 92**,** 154-160.

Lee, S. E.; Yang, H.; Son, G. W.; Park, H. R.; Park, C.-S.; Jin, Y.-H. and Park, Y. S. (2015). Eriodictyol Protects Endothelial Cells against Oxidative Stress-Induced Cell Death through Modulating ERK/Nrf2/ARE-Dependent Heme Oxygenase-1 Expression. *Int. J. Mol. Sci,* 16**,** 14526-14539.

Li, R.-S.; Wang, X.-B.; Hu, X.-J. and Kong, L.-Y. (2013). Design, synthesis and evaluation of flavonoid derivatives as potential multifunctional acetylcholinesterase inhibitors against Alzheimer's disease. *Bioorganic & medicinal chemistry letters,* 23**,** 2636-2641.

Li, X.; Yin, Q.; Chen, W. and Wang, J. (2006). Solubility of hydroquinone in different solvents from 276.65 K to 345.10 K. *Journal of Chemical & Engineering Data,* 51**,** 127-129.

Lin, S.-T. and Sandler, S. I. (2002). A priori phase equilibrium prediction from a segment contribution solvation model. *Industrial & engineering chemistry research,* 41**,** 899-913.

Lipinski, C. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev,* 23**,** 3-25.

Liu, C.; Dang, L.; Bai, W.; Wang, R. and Wei, H. (2014). Solid–Liquid Equilibrium of Theophylline in Solvent Mixtures. *Journal of Chemical & Engineering Data,* 59**,** 263-268.

Liu, H. L.; Jiang, W. B. and Xie, M. X. (2010). Flavonoids: recent advances as anticancer drugs. *Recent patents on anti-cancer drug discovery,* 5**,** 152-164.

Liu, L. and Chen, J. (2008). Solubility of Hesperetin in Various Solvents from (288.2 to 323.2) K. *Journal of Chemical & Engineering Data,* 53**,** 1649-1650.

Magnussen, T.; Rasmussen, P. and Fredenslund, A. (1981). UNIFAC Parameter Table for Prediction of Liquid-Liquid Equilibria. *Industrial & Engineering Chemistry Process Design and Development,* 20**,** 331 - 339.

Marrero, J. and Gani, R. (2001). Group-contribution based estimation of pure component properties. *Fluid Phase Equilibria,* 183**,** 183-208.

Matsuda, H.; Kaburagi, K.; Kurihara, K.; Tochigi, K. and Tomono, K. (2010). Prediction of solubilities of pharmaceutical compounds in water+ co-solvent systems using an activity coefficient model. *Fluid Phase Equilibria,* 290**,** 153-157.

Mccullough, M. L.; Peterson, J. J.; Patel, R.; Jacques, P. F.; Shah, R. and Dwyer, J. T. (2012). Flavonoid intake and cardiovascular disease mortality in a prospective cohort of US adults. *The American journal of clinical nutrition,* 95**,** 454-464.

Messina, M. J.; Persky, V.; Setchell, K. D. and Barnes, S. (1994). Soy intake and cancer risk: a review of the in vitro and in vivo data. *Nutrition and cancer,* 21**,** 113-131.

Migliaccio, S. and Anderson, J. (2003). Isoflavones and skeletal health: are these molecules ready for clinical application? *Osteoporosis International,* 14**,** 361-368.

Miura, T.; Yuan, L.; Sun, B.; Fujii, H.; Yoshida, M.; Wakame, K. and Kosuna, K. I. (2002). Isoflavone aglycon produced by culture of soybean extracts with basidiomycetes and its anti-angiogenic acitivity. *Biosci. Biotechnol. Biochem,* 66**,** 2626–2631.

Mohammad Azmin, S. N. H.; Abdul Manan, Z.; Wan Alwi, S. R.; Chua, L. S.; Mustaffa, A. A. and Yunus, N. A. (2016). Herbal Processing and Extraction Technologies. *Separation & Purification Reviews,* 45**,** 305-320.

Mosihuzzaman, M. and Choudhary, M. I. (2008). Protocols on safety, efficacy, standardization, and documentation of herbal medicine (IUPAC Technical Report). *Pure and Applied Chemistry,* 80**,** 2195-2230.

Mulvihill, E. E.; Burke, A. C. and Huff, M. W. (2016). Citrus flavonoids as regulators of lipoprotein metabolism and atherosclerosis. *Annual review of nutrition,* 36**,** 275-299.

Nan, G.; Shi, J.; Huang, Y.; Sun, J.; Lv, J.; Yang, G. and Li, Y. (2014). Dissociation constants and solubilities of daidzein and genistein in different solvents. *Journal of Chemical & Engineering Data,* 59**,** 1304-1311.

Ncube, N.; Afolayan, A. and Okoh, A. (2008). Assessment techniques of antimicrobial properties of natural compounds of plant origin: current methods and future trends. *African Journal of Biotechnology,* 7.

Ong, K. and Khoo, H. (1997). Biological effects of myricetin. *Gen Pharmacol-Vasc S,* 29**,** 121–126.

Pan, L.; Ho, Q.; Tsutsui, K. and Takahashi, L. (2001). Comparison of chromatographic and spectroscopic methods used to rank compounds for aqueous solubility. *Journal of Pharmaceutical sciences,* 90**,** 521-529.

Pandey, K. B. and Rizvi, S. I. (2009). Plant polyphenols as dietary antioxidants in human health and disease. *Oxidative medicine and cellular longevity,* 2**,** 270-278.

Peng, B. and Yan, W. (2009). Solubility of luteolin in ethanol+ water mixed solvents at different temperatures. *Journal of Chemical & Engineering Data,* 55**,** 583-585.

Peng, B.; Zi, J. and Yan, W. (2006). Measurement and correlation of solubilities of luteolin in organic solvents at different temperatures. *Journal of Chemical & Engineering Data,* 51**,** 2038-2040.

Pick, A.; Müller, H.; Mayer, R.; Haenisch, B.; Pajeva, I. K.; Weigt, M.; Bönisch, H.; Müller, C. E. and Wiese, M. (2011). Structure–activity relationships of flavonoids as inhibitors of breast cancer resistance protein (BCRP). *Bioorganic & medicinal chemistry,* 19**,** 2090-2102.

Pinho, S. P.; Silva, C. M. and Macedo, E. A. (1994). Solubility of amino acids: a group-contribution model involving phase and chemical equilibria. *Industrial & engineering chemistry research,* 33**,** 1341-1347.

Pulley, G. N. (1936). Solubility of naringin in water. *Industrial & Engineering Chemistry Analytical Edition,* 8**,** 360-360.

Puri, S.; Chickos, J. S. and Welsh, W. J. (2002). Three-dimensional quantitative structure− property relationship (3d-qspr) models for prediction of thermodynamic properties of polychlorinated biphenyls (PCBs): enthalpy of sublimation. *Journal of chemical information and computer sciences,* 42**,** 109-116.

Queimada, A. J.; Mota, F. L.; Pinho, S. P. and Macedo, E. A. (2009). Solubilities of Biologically Active Phenolic Compounds: Measurements and Modeling. *The Journal of Physical Chemistry B,* 113**,** 3469-3476.

Ran, Y.; Jain, N. and Yalkowsky, S. H. (2001). Prediction of aqueous solubility of organic compounds by the general solubility equation (GSE). *Journal of chemical information and computer sciences,* 41**,** 1208-1217.

Rao, B. N. (2003). Bioactive phytochemicals in Indian foods and their potential in health promotion and disease prevention. *Asia Pacific journal of clinical nutrition,* 12**,** 9-22.

Rashed, K.; Ciric, A.; Glamoclija, J. and Al., E. (2014). Antibacterial and antifungal activities of methanol extract and phenolic compounds from Diospyros virginiana L. *Ind Crop Prod,* 59**,** 210-215.

Razmara, R. S.; Daneshfar, A. and Sahraei, R. (2010). Solubility of quercetin in water+ methanol and water+ ethanol from (292.8 to 333.8) K. *Journal of Chemical & Engineering Data,* 55**,** 3934-3936.

Renon, H. and Prausnitz, J. (1969). Estimation of parameters for the NRTL equation for excess Gibbs energies of strongly nonideal liquid mixtures. *Industrial & Engineering Chemistry Process Design and Development,* 8**,** 413-419.

Renon, H. and Prausnitz, J. M. (1968). Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE journal,* 14**,** 135-144.

Savjani, K. T.; Gajjar, A. K. and Savjani, J. K. (2012). Drug solubility: importance and enhancement techniques. *ISRN pharmaceutics*, 2012.

Saxena, M.; Saxena, J.; Nema, R.; Singh, D. and Gupta, A. (2013). Phytochemistry of medicinal plants. *Journal of Pharmacognosy and Phytochemistry*, 1.

Sevillano, D. M.; Van Der Wielen, L. A.; Hooshyar, N. and Ottens, M. (2014). MPP-UNIFAC, a predictive activity coefficient model for polyphenols. *Fluid Phase Equilibria,* 384**,** 82-88.

Sevillano, D. M.; Van Der Wielen, L. a. M.; Trifunovic, O. and Ottens, M. (2013). Model Comparison for the Prediction of the Solubility of Green Tea Catechins in Ethanol/Water Mixtures. *Industrial & Engineering Chemistry Research,* 52**,** 6039-6048.

Shahidi, F.; Janitha, P. and Wanasundara, P. (1992). Phenolic antioxidants. *Critical Reviews in Food Science & Nutrition,* 32**,** 67-103.

Shen, Y.; Zhang, J.; Sheng, R.; Dong, X.; He, Q.; Yang, B. and Hu, Y. (2009). Synthesis and biological evaluation of novel flavonoid derivatives as dual binding acetylcholinesterase inhibitors. *Journal of enzyme inhibition and medicinal chemistry,* 24**,** 372-380.

Shimoyama, Y. and Iwai, Y. (2009). Development of activity coefficient model based on COSMO method for prediction of solubilities of solid solutes in supercritical carbon dioxide. *The Journal of Supercritical Fluids,* 50**,** 210-217.

Si, H. and Liu, D. (2007). Phytochemical genistein in the regulation of vascular function: new insights. *Curr. Med. Chem,* 14**,** 2581–2589.

Simamora, P. and Yalkowsky, S. H. (1994). Group Contribution Methods for Predicting the Melting Points and Boiling Points of Aromatic Compounds. *Industrial & Engineering Chemistry Research,* 33**,** 1405-1409.

Sousa, C. C. S.; Matos, M. a. R. and Morais, V. M. F. (2009). Energetics of flavone and flavonone. *Journal of Chemical Thermodynamics,* 41**,** 1408 - 1412.

Spiliotis, N. and Tassios, D. (2000). A UNIFAC model for phase equilibrium calculations in aqueous and nonaqueous sugar solutions. *Fluid Phase Equilibria,* 173**,** 39-55.

Srinivas, K.; King, J.; Monrad, J.; Howard, L. and Hansen, C. (2009). Optimization of subcritical fluid extraction of bioactive compounds using Hansen solubility parameters. *Journal of food science,* 74**,** E342-E354.

Srinivas, K.; King, J. W.; Howard, L. R. and Monrad, J. K. (2010). Solubility and solution thermodynamic properties of quercetin and quercetin dihydrate in subcritical water. *Journal of food engineering,* 100**,** 208-218.

Towler, G. and Sinnott, R. K. (2012). *Chemical engineering design: principles, practice and economics of plant and process design.* Elsevier.

Treml, J. and Šmejkal, K. (2016). Flavonoids as Potent Scavengers of Hydroxyl Radicals. *Comprehensive Reviews in Food Science and Food Safety*.

Tsavas, P.; Voutsas, E.; Magoulas, K. and Tassios, D. (2004). Phase equilibrium calculations in aqueous and nonaqueous mixtures of sugars and sugar derivatives with a group-contribution model. *Industrial & engineering chemistry research,* 43**,** 8391-8399.

Vazquez-Prieto, M. A. and Miatello, R. M. (2010). Organosulfur compounds and cardiovascular disease. *Molecular aspects of medicine,* 31**,** 540-545.

Walas, S. M. (2013). *Phase equilibria in chemical engineering.* Butterworth-Heinemann.

Wang, J. and Hou, T. (2011). Recent advances on aqueous solubility prediction. *Combinatorial chemistry & high throughput screening,* 14**,** 328-338.

Wang, L.-C. and Wang, F.-A. (2004). Solubility of Niacin in 3-Picoline+ Water from (287.65 to 359.15) K. *Journal of Chemical & Engineering Data,* 49**,** 155-156.

Wang, L.-H. and Cheng, Y.-Y. (2005). Solubility of Puerarin in Water, Ethanol, and Acetone from (288.2 to 328.2) K. *Journal of Chemical & Engineering Data,* 50**,** 1375-1376.

Wang, Q.; Ma, P. and Neng, S. (2009). Position Group Contribution Method for Estimation of Melting Point of Organic Compounds. *Chinese Journal of Chemical Engineering,* 17**,** 468-472.

Wang, Z.; Wang, J.; Zhang, M. and Dang, L. (2006). Solubility of erythromycin A dihydrate in different pure solvents and acetone+ water binary mixtures between 293 K and 323 K. *Journal of Chemical & Engineering Data,* 51**,** 1062-1065.

Wasik, S.; Tewari, Y.; Miller, M. and Martire, D. (1981). Octanol/water partition coefficients and aqueous solubilities of organic compounds.

Wedlich, U. and Gmehling, J. (1987). A Modified UNIFAC Model. 1. Prediction of VLE, enthalpy of excess, and infinite dilution activity coefficient. *Industrial & Engineering Chemistry Research,* 26**,** 1372 - 1381.

Wei, D. and Zhang, X. (2013). Solubility of puerarin in the binary system of methanol and acetic acid solvent mixtures. *Fluid Phase Equilibria,* 339**,** 67-71.

Wei, H. and Löbenberg, R. (2006). Biorelevant dissolution media as a predictive tool for glyburide a class II drug. *european journal of pharmaceutical sciences,* 29**,** 45-52.

Weidlich, U. and Gmehling, J. (1987). A modified UNIFAC model. 1. Prediction of VLE, hE, and. gamma.. infin. *Industrial & Engineering Chemistry Research,* 26**,** 1372-1381.

Welch, C. R.; Wu, Q. and Simon, J. E. (2008). Recent advances in anthocyanin analysis and characterization. *Current analytical chemistry,* 4**,** 75-101.

Wen, X. and Qiang, Y. (2004). a New Group Contribution Method for Estimating Melting and Boiling Points of Organic Compounds. *Chemical Engineering Communications,* 191**,** 75-86.

Wink, M. (2004). Encyclopedia of Plant and Crop Science. *Phytochemical Diversity of Secondary Metabolites.* New York: Marcel Dekker, Inc.

Winnike, R. (2005). Solubility assessment in pharmaceutical development, practical considerations for solubility profiling and solubilization techniques. *AAPS Short Course on Fundamentals of Preformulation in Pharmaceutical Product Development, Nashville, TN*.

Wu, J.-G.; Ge, J.; Zhang, Y.-P.; Yu, Y. and Zhang, X.-Y. (2010). Solubility of genistein in water, methanol, ethanol, propan-2-ol, 1-butanol, and ethyl

acetate from (280 to 333) K. *Journal of Chemical & Engineering Data,* 55**,** 5286-5288.

Xiao, M.; Shao, Y.; Yan, W. and Zhang, Z. (2011). Measurement and correlation of solubilities of apigenin and apigenin 7-O-rhamnosylglucoside in seven solvents at different temperatures. *The Journal of Chemical Thermodynamics,* 43**,** 240-243.

Xiao, M.; Yan, W. and Zhang, Z. (2010). Solubilities of apigenin in ethanol+ water at different temperatures. *Journal of Chemical & Engineering Data,* 55**,** 3346-3348.

Yalkowsky, S. H. and Banerjee, S. (1992). *Aqueous solubility: Methods of estimation for organic compounds.* Marcel Dekker New York.

Yamamoto, Y. and Gaynor, R. B. (2001). Therapeutic potential of inhibition of the NF-κB pathway in the treatment of inflammation and cancer. *Journal of Clinical Investigation,* 107**,** 135-142.

Yang, Z.-S.; Zeng, Z.-X.; Xue, W.-L. and Zhang, Y. (2008). Solubility of bis (benzoxazolyl-2-methyl) sulfide in different pure solvents and ethanol+ water binary mixtures between (273.25 and 325.25) K. *Journal of Chemical & Engineering Data,* 53**,** 2692-2695.

Yu, M.; Lee, J.; Lee, J. and Al., E. (2012). Identification of myricetin and scutellarein as novel chemical inhibitors of the SARS coronavirus helicase, nsP13. *Bioorg Med Chem Lett,* 22**,** 4049–4054.

Zhou, L.; Zhang, P.; Yang, G.; Lin, R.; Wang, W.; Liu, T.; Zhang, L. and Zhang, J. (2014). Solubility of chrysin in ethanol and water mixtures. *Journal of Chemical & Engineering Data,* 59**,** 2215-2220.

Zi, J.; Peng, B. and Yan, W. (2007). Solubilities of rutin in eight solvents at T= 283.15, 298.15, 313.15, 323.15, and 333.15 K. *Fluid Phase Equilibria,* 261**,** 111-114.