CHEMICAL ENGINEERING TRANSACTIONS
VOL. 56, 2017
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet
Guest Editors: Jiří Jaromír Klemeš, Peng Yen Liew, Wai Shin Ho, Jeng Shiun Lim

ISBN 978-88-95608-47-1; **ISSN** 2283-9216
DOI: 10.3303/CET1756134

Solubility Prediction of Flavonoids using New Developed UNIFAC-based Model

Mohd Shukri Mat Nor[a,b], Zainuddin Abdul Manan*[,a,b] , Azizul Azri Mustaffa[a,b], Suan Chua Lee[b,c]

[a]Process Systems Engineering Centre (PROSPECT), Research Institute of Sustainable Environment (RISE), Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[b]Faculty of Chemical and Energy Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[c]Institute of Bioproduct Development (IBD), Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
dr.zain@utm.my

Flavonoids are phytochemicals extensively used in the pharmaceutical, food, and pigment industries. They have many important biological properties including antioxidant, anti-inflammatory, antifungal, and anti-viral. The importance of flavonoids has motivated the development of many processes for the manufacture of flavonoids derivative products. The aim of this study is to develop a new set of Universal Functional Activity Coefficient (UNIFAC) parameters for solubility prediction of flavonoids in organic solvents. In this study, group interaction parameters of the UNIFAC have been regressed and improved from the solubility experimental data of flavonoids based on the activity coefficient model through the thermodynamic modelling of Solid–Liquid Equilibrium (SLE) relationship which involves an iterative step. The results showed that a more accurate prediction (lower prediction error) could be obtained using the new parameters. By using our developed parameter for flavonoids, better agreements were obtained between the experimental and the predicted values by the UNIFAC model with less than 5.57 % deviation. The results indicated that the newly developed UNIFAC-based model can adequately be used to represent the measured data with good accuracy and can be useful for the purpose of solubility estimation for flavonoids in various solvents.

1. Introduction

Flavonoids are one the most studied compounds for their pharmacological and biological activities among the numerous classes of phytochemicals that present in plants (Bravo, 1998). They are widely distributed in the leaves, seeds, bark, and flowers of plants, being known for comprising approximately 8,000 different compounds (Treml and Šmejkal, 2016). Flavonoids may be categorised into six major classes based on the differences in their molecular backbone structure. These classes of flavonoids comprise of an aromatic A-ring fused with a heterocyclic C ring and attached to an aromatic B-ring through a carbon-carbon bridge. The backbone structure of the flavonoids are given in Figure 1. It is known that flavonoids are important constituents of the human diet (Hertog et al., 1992), which is convinced to have a lot of therapeutic potentials and are able to reduce the risk of multiple diseases such as cancer (Mohammad Azmin et al., 2016).
The processes that are typically employed in the Industrial production of flavonoids (pharmaceutical, nutraceutical, food etc.) are extraction, formulation and crystallisation. It is important to have fundamental physicochemical properties data for the flavonoids in order to optimise these process designs. One of the important properties that plays an essential role in these processes is solubility. The measurement of their solubilities in solvents has been growing in the literature but is still insufficient to support the process design. In fact, solubility experimental measurement methods are time consuming, costly especially in raw material expenditure, and give rise to technical issues repeatedly due to the lack of experience in analytical skill.
One fascinating viewpoint is the application of the thermodynamic models of solid-liquid equilibrium, which are usually used for predicting the solubility through the relative importance of the melting properties (melting point and fusion enthalpy) and activity coefficient. These critical data are required to develop an effective process

Please cite this article as: Nor M.S.M., Manan Z.A., Mustaffa A.A., Lee S.C., 2017, Solubility prediction of flavonoids using new developed unifac-based model, Chemical Engineering Transactions, 56, 799-804 DOI:10.3303/CET1756134

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 is based on the UNIFAC (Universal Functional Activity Coefficient) correlation for estimating activity coefficients from group contributions. The UNIFAC is a well-known group contribution model for the molecule activity coefficients prediction in non-ideal mixtures. It showed that the Solid–Liquid Equilibrium (SLE) could be forecasted by thermodynamic modelling of the UNIFAC model. However, the results were precise only for a limited compounds and tend to overestimate the phytochemical solubility systems. For example, for cinnamic acid esters in organic solvents (Panteli et al., 2009) and phytochemicals from Orthosiphon Staminues (Mat Nor et al., 2015). The aim of this study is to predict the solubility of flavonoids using a newly developed UNIFAC-based model.

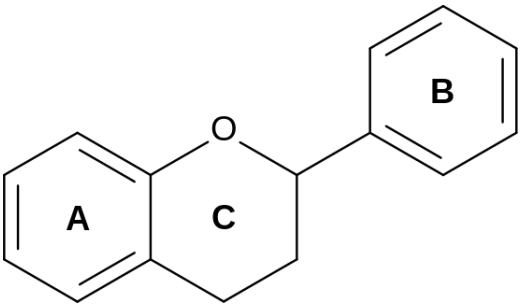

Figure 1: The general structure of flavonoid molecules

2. Methods and tools for solubility modeling

2.1 Database

The experimental solubilities of flavonoids were obtained from open literatures as tabulated in Table 1. There are 145 solubility data points considered as training set for the regression of interaction parameters. The studied temperature range is from 283.2 to 343.2 K with 6 different types of flavonoid compounds in a total of 12 pure solvents.

Table 1: The solubility data of Flavonoid compounds used for regression

Data Source	Flavonoid compounds	No. of data points	Temperature (K)	Solvents
Ferreira and Pinho (2012)	Hesperetin	8	298.2 - 313.2	acetonitrile, ethyl acetate
Liu and Chen (2008)		40	288.2 - 323.2	methanol, ethanol, 1-butanol, acetone, water
Zi et al. (2007)	Rutin	40	283.2 - 333.2	water, methanol, ethanol, 1-propanol, 2-propanol, 1-butanol, acetone, and ethyl acetate
Feizi et al. (2016)	Chrysin	2	298.2	dimethylformamide, tetrahydrofuran
Zhou et al. (2014)		14	293.2 - 323.2	Ethanol, water
Chebil et al. (2007)	Naringenin	3	323.2 - 343.2	Acetonitrile
Chebil et al. (2007)	Quercetin	2	323.2	2-methyl-2-Butanol, acetone
Chebil et al. (2007)		3	323.2 - 343.2	Acetonitrile
Razmara et al. (2010)		30	292.8 - 333.8	water, methanol, ethanol
Chebil et al. (2007)	Isoquercitrin	3	323.2 - 343.2	Acetonitrile

The experimental data of the enthalpy of fusion and melting temperature of flavonoids were collected from the references cited in Table 2.

2.2 Solid liquid equilibrium equation

The following standard thermodynamic correlation (Eq(1)) shows the relationship between solubility, activity coefficients and melting properties that is used to calculate the solubility of compound i in mole fraction, x_i.

Table 2: The enthalpy of fusion and melting temperature of Flavonoids under study (Chebil et al., 2007)

Flavonoid Compounds	Melting point (K)	Fusion enthalpy (kJ/mol)
Hesperetin	499.2	35.9
Rutin	450.2	82.3
Chrysin	558.2	39.2
Naringenin	523.2	39.8
Quercetin	595.2	41.5
Isoquercitrin	471.2	49.8

$$\ln(\gamma_i x_i) = -\frac{\Delta H_i^{fus}}{RT_{m,i}}\left(\frac{T_{m,i}}{T} - 1\right) - \frac{\Delta C_{p_i}^{fus}}{R}\left(\ln\left(\frac{T_{m,i}}{T}\right) - \frac{T_{m,i}}{T} + 1\right) \quad (1)$$

Gracin et al. (2002) found that ΔC_{pi}^{fus} has a small influence on the UNIFAC model. It is very difficult to measure ΔC_{pi}^{fus} especially for the case of decomposition, sublimation or parallel reactions occuring during melting (Bouillot et al., 2011). ΔC_{pi}^{fus} contribution is typically assumed as negligible. This Eq(1) can be further simplified as given by Eq(2):

$$x_i\gamma_i = \exp\left[\frac{\Delta H_i^{fus}}{RT_{m,i}}\ln\left(\frac{T}{T_{m,i}}\right)\right] \quad (2)$$

where γ_i is the activity coefficient, ΔH_i^{fus} is the enthalpy of fusion, $T_{m,i}$ is the melting temperature of the compound i, T is the temperature, R is the ideal gas constant and ΔCp_i^{fus} is the change of temperature at equlibrium for the solid and the liquid phase heat capacities for component i.

2.3 UNIFAC group contribution model

The procedure proposed here is based on the UNIFAC correlation for estimating activity coefficients from group contributions. The UNIFAC model is a combination of two parts of activity coefficient, which is combinatorial ($\ln \gamma^C$) and residual (γ^R) as stated in Eq(3):

$$\ln\gamma = \ln\gamma^c + \ln\gamma^R \quad (3)$$

$\ln \gamma^C$ relies on the mole fraction (x_i), area fraction (θ_i), segment fraction (ϕ_i), Van der Waals volumes (r_i), Van der Waals surface areas (q_i), volume/mole fraction ratios (V_i), relative volume/mole fraction ratios (V'_i) and surface area/mole fraction ratios (F_i). The expression of $\ln \gamma^C$ is given by Eq(4):

$$\ln\gamma_i^C = 1 - (V'_i) + 5q_i\left[1 - \frac{V_i}{F_i} + \ln\left(\frac{V_i}{F_i}\right)\right] \quad (4)$$

$$V'_i = \frac{r_i^{3/4}}{\sum_j x_j r_j^{3/4}};\quad V_i = \frac{r_i}{\sum_j x_j r_j};\quad F_i = \frac{q_i}{\sum_j x_j q_j};\quad r_i = \sum v_k^{(i)} R_k;\quad q_i = \sum v_k^{(i)} Q_k \quad (5)$$

γ^R represents the total of the functional group activity coefficients weighted by their number in solution. The equation for this part is expressed in Eq(6).

$$\ln\gamma^R = \sum_k v_k^{(i)}\left[\ln\Gamma_k - \ln\Gamma_k^{(i)}\right] \quad (6)$$

Γ_k and Γ_k^i are the residual activity coefficient of group k in the mixture and in a solution of pure compound i respectively. v_k and v_k^i are the number of groups of type k in the mixture and in compound i. They depend on the volume R_k and surface area Q_k of group k and adjustable binary interaction parameter a_{mn} that are usually regressed from VLE experimental data. The equations are expressed in Eqs(7) – (9):

$$\ln\Gamma_k = Q_k\left[1 - \ln\left(\sum_m \Theta_m\psi_{mk}\right) - \frac{\sum_m \Theta_m\psi_{km}}{\sum_n \Theta_n\psi_{nm}}\right] \quad (7)$$

$$\Theta_m = \frac{Q_m X_m}{\sum_n Q_n X_n};\qquad X_m = \frac{\sum_j v_m^{(j)} x_j}{\sum_j\sum_m v_n^{(j)} x_j} \quad (8)$$

$$\psi_{mn} = \exp\left(-\frac{a_{mn}}{T}\right) \quad (9)$$

Θ_m is the summation of the area fraction of group m, ψ_{mk} is the interaction energy parameter between group m and group k, X_m is the group m mole fraction, subscript n is referred to groups of type n and subscript or superscript j is represents compound j.

2.4 Mean absolute percent error

The mean absolute percent error (MAPE) was calculated in order to evaluate the performance of the model. The MAPE was determined by using Eq(10):

$$\text{MAPE} = \left[\frac{1}{n}\sum \frac{\left|\gamma_{\text{experiment}} - \gamma_{\text{predicted}}\right|}{\left|\gamma_{\text{experiment}}\right|}\right] \times 100\,\% \quad (10)$$

where n is the number of data points.

3. Results and discussion

3.1 Parameter regression

The UNIFAC model uses three types of parameters which are van der Waals volume (R_k), van der Waals area (Q_k) and interaction parameters (a_{mn}). The values of Rk and Qk were obtained from UNIFAC tables (Hansen et al., 1991) while the interaction parameters were regressed. Table 3 shows the resulting matrix of interaction parameters of this adjustment procedure. The interactions between groups have been assumed to be temperature independent due to the circumscribed amount of data set and the limitation of temperature ranges in this study. A matrix of the 13 possible interaction parameters is designed by considering the UNIFAC self-interaction parameters equal to zero. The total number of parameters is 156 obtained from (number of groups)^2 – number of groups.

3.2 Validation

A solubility dataset of puerarin in methanol (Wei and Zhang, 2013) not previously used for the regression was selected to validate the model described in this work. Figure 2 shows the newly developed UNIFAC model performance in comparison with experimental values. The prediction proclaimed by the model having a correct depiction of groups gives a mean absolute percent error (MAPE) of 5.57 %, showing that the predicted values are in good consensus with the experimental data.

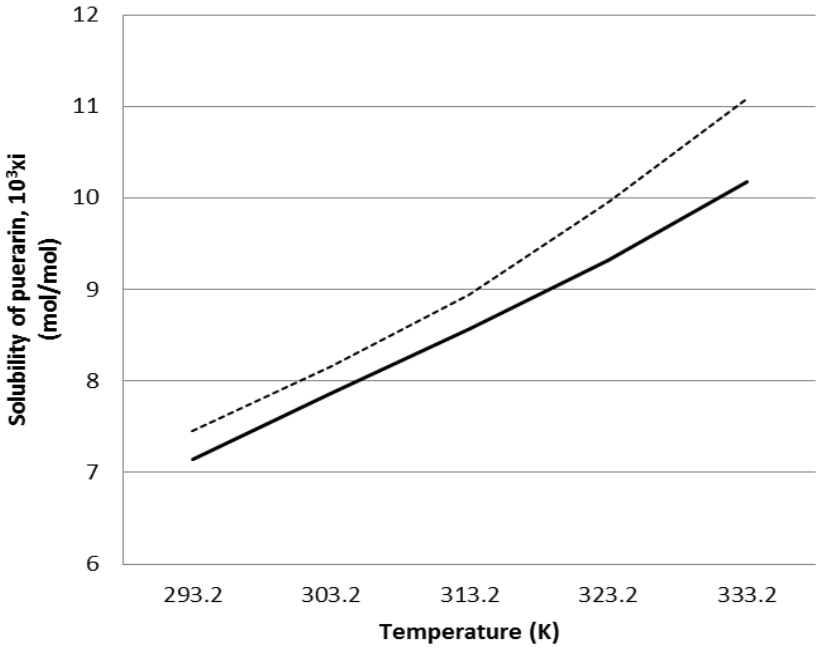

Figure 2: Solubility of puerarin in methanol in comparison between experimental (full lines) and prediction made by the newly developed UNIFAC model (dashed lines)

Table 3: Matrix of readjusted UNIFAC interaction parameters for flavonoids

	CH_2	C=C	ACH	$aCCH_2$	OH	CH_3OH	H_2O	ACOH	CH_2CO	CCOO	CH_2O	CCN	DMF
CH_2		86.0	61.1	76.5	986.5	697.2	1,318	1,333	476.4	232.1	251.5	597.0	485.3
C=C	-35.4		38.8	74.15	524.1	787.6	270.6	526.1	182.6	37.8	214.5	336.9	-70.45
ACH	-11.1	3.5		167.0	636.1	637.4	903.8	1,329	25.8	6.0	32.1	212.5	245.6
$aCCH_2$	-69.7	-113.6	-146.8		803.2	603.3	5,695	884.9	-52.1	5,688	213.1	6,096	5,629
OH	156.4	457.0	89.6	25.8		-137.1	353.5	-259.7	84.0	101.1	28.1	6.7	-143.9
CH_3OH	16.5	-12.5	-50.0	-44.5	249.1		-181.0	-101.7	23.4	-10.7	-128.6	53.3	-172.4
H_2O	300.0	496.1	362.3	377.6	-229.2	289.6		324.5	-195.4	72.9	540.5	112.6	319
ACOH	275.8	217.5	25.3	244.2	-451.6	-265.2	-601.8		-356.1	-449.4	-162.9	0	0
CH_2CO	26.76	42.9	140.1	365.8	164.5	108.7	472.5	-133.1		-213.7	-103.6	481.7	-61.7
CCOO	114.8	132.1	85.8	170.0	245.4	249.6	200.8	-36.7	372.2		-235.7	494.6	85.3
CH_2O	83.3	26.5	52.1	65.7	237.7	238.4	-314.7	-178.5	191.1	461.3		-18.5	254.8
CCN	24.8	-40.6	-23.0	-138.4	185.4	162.6	242.8	0.005	-287.5	-266.6	38.8		-151.5
DMF	-31.9	249.0	-133.9	-240.2	64.16	172.2	-287.1	0	97.0	-82.1	-158.2	150.6	

4. Conclusions

An extensive literature search to collect flavonoids solubility data in pure organic solvents and water has been accomplished. The new interaction parameters applicable for the case of flavonoids in a newly developed UNIFAC model was regressed from the data. This method was validated against puerarin in methanol experimental solubility dataset from literature showing better prediction with 5.57 % MAPE. The results presented in this study give a more accurate prediction of the flavonoid solubility systems. This newly developed UNIFAC model resulted in an improved prediction accuracy as shown by the better understanding between computed and experimental data, without losing the simplicity of the group contribution procedure.

Acknowledgments

The authors gratefully acknowledge the Ministry of Higher Education (Malaysia) and Universiti Teknologi Malaysia for the funding provided for this research under the Fundamental Research Grant Scheme (Reference: R.J130000.7844.4F270).

Reference

Bouillot B., Teychené S., 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.

Bravo L., 1998, Polyphenols: chemistry, dietary sources, metabolism, and nutritional significance, Nutrition reviews 56, 317-333.

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

Feizi S., Jabbari M., 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., Pinho S.O.P, 2012, Solubility of flavonoids in pure solvents, Industrial and Engineering Chemistry Research 51, 6586-6590.

Gracin S., Brinck T., Rasmuson Å.C. 2002, Prediction of solubility of solid organic compounds in solvents by UNIFAC, Industrial and Engineering Chemistry Research 41, 5114-5124.

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

Hertog M.G., Hollman P.C., 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.

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

Mat Nor M.S., Manan Z.A., Mustaffa A.A., Suan C.L., 2015, An Evaluation of Thermodynamic Models for the Prediction of Solubility of Phytochemicals from Orthosiphon Staminues in Ethanol, Computer Aided Chemical Engineering 37, 2087-2092.

Mohammad Azmin S.N.H., Abdul Manan Z., Wan Alwi S.R., Chua L.S., Mustaffa A.A.,Yunus N.A., 2016, Herbal Processing and Extraction Technologies, Separation and Purification Reviews 45, 305-320.
Panteli E., Saratsioti P., Stamatis H., Voutsas E., 2009, Solubilities of cinnamic acid esters in organic solvents. Journal of Chemical and Engineering Data 55, 745-749.
Razmara R.S., Daneshfar A., Sahraei R., 2010, Solubility of quercetin in water+ methanol and water+ ethanol from (292.8 to 333.8) K, Journal of Chemical and Engineering Data 55, 3934-3936.
Treml J., Šmejkal K., 2016, Flavonoids as Potent Scavengers of Hydroxyl Radicals, Comprehensive Reviews in Food Science and Food Safety 15 (4), 720-738.
Wei D., Zhang X., 2013, Solubility of puerarin in the binary system of methanol and acetic acid solvent mixtures, Fluid Phase Equilibria 339, 67-71.
Zhou L., Zhang P., Yang G., Lin R., Wang W., Liu T., Zhang L., Zhang J., 2014, Solubility of chrysin in ethanol and water mixtures, Journal of Chemical and Engineering Data 59, 2215-2220.
Zi J., Peng B., 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.