MPP-UNIFAC, a predictive activity coefficient model for polyphenols

David Méndez Sevillano [a], Luuk A.M. van der Wielen [a], Nasim Hooshyar [b], Marcel Ottens [a,*]

[a] Department of Biotechnology, Julianalaan 67, Delft 2628 BC, The Netherlands
[b] Unilever R&D Vlaardingen, Olivier van Noortlaan 120, Vlaardingen 3133 AT, The Netherlands

A B S T R A C T

Modified UNIFAC (Mod. UNIFAC) is a proven model for the prediction of activity coefficients of molecules in non-ideal mixtures. However, Mod. UNIFAC is often not accurate when hydrogen bonding or strong hydrophobic interactions are present. An interesting group of molecules called polyphenols presents both types of interactions and therefore, Mod. UNIFAC predictions are not always accurate. A polyphenol-specific UNIFAC-based method (MPP-UNIFAC) has been developed in this study for its application on polyphenol-like molecules. Interaction parameters were regressed from 410 solubility data points from polyphenols in literature showing a good fit. The average error of the fit is 0.094 log units, better than the average value for Mod. UNIFAC (0.5 log units). Moreover, the model was validated against three datasets that were not used for the regression giving more accurate predictions than the Mod. UNIFAC it was developed from.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Polyphenol is a vast concept that groups different types of molecules. These molecules consist of two or more phenyl rings with a number of hydroxyl groups attached to them. Besides these hydroxyl groups, they can have any type of group attached to it. The type of polyphenols that this work focuses on is flavonoids.

Flavonoids consist of two phenolic rings connected by a three-carbon and one-oxygen chain that forms a third ring fused to the first one. The name of this type of polyphenols comes from the similarities between their backbone to the molecule flavan (Fig. 1).

The non-aromatic ring can change by either the substitution of one of the alkane groups for a ketone group, the presence of double bonds or even the substitution of the entire ring by a saturated 2-carbon chain (i.e., stilbenoids). All flavonoids have a number of hydroxyl groups attached (either to the aromatic rings or the non-aromatic one). Other attached groups can be glucose molecules, methyl ethers, gallic esters, etc.

Flavonoids can be classified in different groups based on the conformation of the three rings that form the backbone of the molecule [1]. That classification can be found in Fig. 2.

Part of the importance of these molecules are the health benefits that some of them show. Flavanones such as naringenin or hesperidin show lowering of high cholesterol levels [2], some

studies relate isoflavones with reduction of hot flushes [3] or colon cancer [4] and catechins from green tea show a reduction of obesity and certain types of cancer [5].

Such polyphenols are usually present in food sources in very small concentrations. Therefore, their separation and purification entrails complex processes such as liquid–liquid extraction, solid–liquid extraction, crystallization, etc. In order to design and model these processes, specific physicochemical and thermodynamic properties of the polyphenols are needed which would require costly and lengthy experiments.

A large number of properties can be predicted a priori using activity coefficients. Activity coefficients have proven to be a powerful tool for the prediction of the behavior of a wide range of components in different kinds of mixtures. These activity coefficients can be either modeled with models such as NRTL-SAC [6] or predicted with models such as UNIFAC [7].

UNIFAC (UNIQUAC Functional-group Activity Coefficients) is a group contribution model for activity coefficients fully based on the UNIQUAC model [8]. The difference between these two models is based on the interaction parameters. In the case of UNIQUAC, they are regressed from experimental data while in the case of UNIFAC, they are calculated with the contribution of the different molecular subgroups. Even though the interaction parameters between those groups have been regressed from experimental data, they can be applied to molecules not considered in the original dataset. The strength of UNIFAC lies on its prediction capability without the need of previous experimental effort and therefore, it is thoroughly used in industry and found in the open

---

* Corresponding author. Tel.: +31 152782151.
  E-mail address: M.Ottens@tudelft.nl (M. Ottens).

Nomenclature

a	aromatic when applied to the definition of groups
a_{nm}	UNIFAC group interaction parameter between groups n and m (K)
b_{nm}	UNIFAC group interaction parameter between groups n and m
c_{nm}	UNIFAC group interaction parameter between groups n and m (K^{-1})
cyc	cyclical when applied to the definition of groups
F_i	auxiliary property for component i
q_i	relative van del Waals surface of the component i
Q_k	relative van der Waals surface area of the group k
r_i	relative van del Waals volume of the component i
R	ideal gas constant (J/mol K)
R_k	relative van der Waals volume of the group k
T	temperature (K)
$T_{m,i}$	temperature of melting of component i (K)
V_i	auxiliary property for component i
V_i'	empirically modified V_i-value
x_i	mole fraction in liquid phase
X_k	group mole fraction of group m in the liquid phase in Eqs. (12) and (13)

Greek letters

γ_i	activity coefficient (−)
$\Delta C_{p,i}^{\text{fus}}$	difference between the heat capacity of the solid phase and of the liquid phase at equilibrium for component i
ΔH_i^{fus}	enthalpy of solid–liquid transition of component i (J/mol)
Γ^k	group activity coefficient of group k in the mixture
$\Gamma^k_{(i)}$	group activity coefficient of group k in the pure substance i
θ_m	area fraction for group m in Eqs. (11) and (12)
$\upsilon_k^{(i)}$	number of structural groups of type k in molecule i.
Ψ_{nm}	UNIFAC group interaction parameter between groups n and m in Eqs. (11) and (14)

literature. Later on, the Modified UNIFAC model [9] (and its later revisions and extensions [10–13]) was developed to improve the performance of the original model.

Mod. UNIFAC has proven its usefulness in the prediction of properties for small and medium sized molecules [14,15]. However, it has shown rather inaccurate predictions when the studied molecules are either too large or present strong van der Waals forces, such as hydrogen bonding or weak ones, such as π–π interactions [16]. Méndez Sevillano et al. [17] showed how the prediction of Mod. UNIFAC was most inaccurate for smaller activity coefficient values, typically related to mixtures in which attractive forces being present, such as hydrogen bonds. These challenges have been overcome in the past by the extension and/or
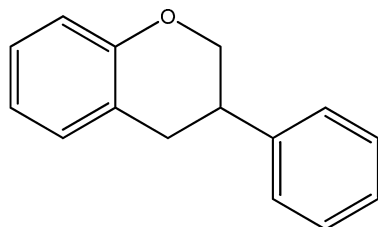

Fig. 1. Molecular structure of flavan.

redefinition of new and specific groups in cases like ionic liquids [18] or polymers [19].

The aim of this work is the development and validation of a mod-UNIFAC-based activity coefficient model, MPP (Modified Poly Phenol) UNIFAC, specific for polyphenols. A new set of 9 groups and 16 subgroups has been defined to its application to polyphenols and interaction parameters have been regressed from literature solubility data. Furthermore, the model has been validated by comparing its prediction to experimental datasets not used for the regression.

2. Theoretical framework

Activity coefficients in literature are calculated from different types of equilibrium properties (VLE, LLE, SLE, etc.). Due to the high melting point of polyphenols, they are mostly found in a solid state and solubility is the most commonly published equilibrium data. From solubility data, activity coefficients can be calculated based on Eq. (1). This equation is derived from basic well-known thermodynamic relations assuming a unique pure solid phase.

$$x_i\gamma_i^l = \exp\left\{\frac{\Delta H_i^{\text{fus}}}{RT_{m_i}}\left(\frac{T-T_{m_i}}{T}\right) + \frac{\Delta C_{p_i}^{\text{fus}}}{R}\left[\ln\left(\frac{T}{T_{m_i}}\right) - \left(\frac{T-T_{m_i}}{T}\right)\right]\right\} \tag{1}$$

This equation can be further simplified considering $\ln(T/T_{m,i})$ equal to $(T-T_{m,i})/T$ [20] leading to Eq. (2).

$$x_i\gamma_i^l = \exp\left[\frac{\Delta H_i^{\text{fus}}}{RT_{m_i}}\ln\left(\frac{T}{T_{m_i}}\right)\right] \tag{2}$$

Mod-UNIFAC calculates activity coefficients taking into account two contributions: the combinatorial and the residual. The first one accounts for the difference in size and shape of the groups while the second one accounts for the possible interactions between them.

$$\ln(\gamma_i) = \ln(\gamma_i^C) + \ln(\gamma_i^R) \tag{3}$$

Each of the contributions can be calculated with Eqs. (4) and (5)

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

$$\ln(\gamma_i^R) = \Sigma_k \upsilon_k^{(i)}\left[\ln(\Gamma_k) - \ln(\Gamma_k^{(i)})\right] \tag{5}$$

as a function of the following variables and the matrix $\upsilon_k^{(i)}$, a descriptor of the different groups that form the molecules present in the equilibrium.

$$V_i' = \frac{r_i^{3/4}}{\Sigma_j x_j r_j^{3/4}} \tag{6}$$

$$V_i = \frac{r_i}{\Sigma_j x_j r_j} \tag{7}$$

$$F_i = \frac{q_i}{\Sigma_j x_j q_j} \tag{8}$$

$$r_i = \Sigma \upsilon_k^{(i)} R_k \tag{9}$$