Polyphenol Liquid-Liquid Extraction Process Development Using NRTL-SAC

Silva, Marcelo; García, Joan Cortada; Ottens, Marcel

DOI
10.1021/acs.iecr.8b00613

Publication date
2018

Document Version
Final published version

Published in
Industrial and Engineering Chemistry Research

Citation (APA)
Silva, M., García, J. C., & Ottens, M. (2018). Polyphenol Liquid-Liquid Extraction Process Development Using NRTL-SAC. *Industrial and Engineering Chemistry Research*, *57*(28), 9210-9221. https://doi.org/10.1021/acs.iecr.8b00613

I&EC research
Industrial & Engineering Chemistry Research

Article

Cite This: *Ind. Eng. Chem. Res.* 2018, 57, 9210−9221

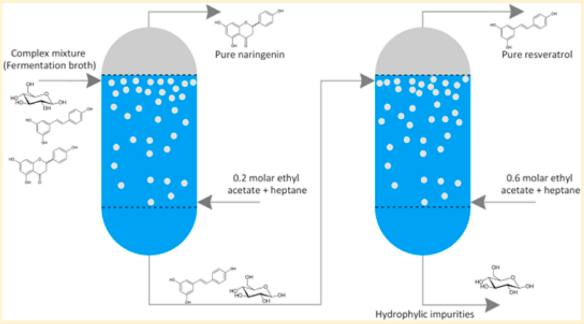
pubs.acs.org/IECR

Polyphenol Liquid−Liquid Extraction Process Development Using NRTL-SAC

Marcelo Silva, Joan Cortada García, and Marcel Ottens*

Delft University of Technology, Department of Biotechnology, Van der Maasweg 9, 2629 HZ Delft, The Netherlands

ABSTRACT: Liquid−liquid extraction (LLE) can be an effective strategy for the purification of polyphenols from a fermentation broth. However, solvents need to be chosen to ensure high extraction capacity and selectivity. For that purpose, a systematic study is here presented, where the partition of different polyphenols—naringin, naringenin, *p*-coumaric acid, and *trans*-resveratrol—was measured in different solvents and solvent mixtures and described using the semipredictive NRTL-SAC model. The minimum average absolute deviation obtained, based on predicted activity coefficients, was of 40%. With the exception of naringin, the NRTL-SAC molecular descriptors were estimated using solubility data already available in the literature. The obtained results made it possible to propose suitable LLE-based downstream process schemes for two possible purification scenarios: the recovery of *trans*-resveratrol and the purification of both naringenin and *trans*-resveratrol, two similar hydrophobic polyphenols, both from a fermentation broth containing hydrophilic impurities (e.g., sugars, proteins).



1. INTRODUCTION

Polyphenols are compounds that have interesting health benefits alongside other diverse (biotechnological) applications (e.g., colorants, nutraceuticals).[1] These molecules are secondary metabolites naturally produced by plants, which can act as radical scavengers due to the high stabilization provided by ring aromaticity.[2] Over the last years, research on their health properties has grown considerably,[3] with authors studying the properties of these molecules in the prevention of diseases such as Alzheimer and several types of cancer.[4]

Although these compounds are mostly obtained by plant extraction, there has been a growing interest in their production using fermentation processes, mainly due to sustainability issues.[5] The fact that the titers achieved for these molecules are usually low[5] demands that efficient downstream process strategies are developed in order to attain economic feasibility. One of the possible strategies that might combine simplicity and lower costs is liquid−liquid extraction (LLE).

LLE can be a suitable option for the recovery and purification of polyphenols. Not only can it offer high flexibility, but it is also adequate when the product of interest is heat-sensitive and when the feed stream is dilute.[6] One of the critical issues in the design of a LLE process is solvent selection, which has to provide—among other characteristics—extraction capacity, selectivity, and a high solute mass transfer rate.[7] For the particular case of polyphenols, there is no systematic study, to the best knowledge of the authors, concerning the adequacy of different solvent chemical classes for the recovery and purification of polyphenols. In order to perform that study, not only is a reliable thermodynamic model needed to describe their equilibrium properties, but this model also needs to have predictive capabilities.[8−10] Although some models have already been applied to describe the solubility of polyphenols in different solvents (e.g., NRTL-SAC, MPP-UNIFAC, etc.),[11,12] they were scarcely applied to liquid−liquid equilibrium data. One of the possibilities is using a model like COSMO-RS,[13] but some of the disadvantages, though, are that its predictive quality is still inferior to other models such as UNIFAC or NRTL-SAC.[14] On the other hand, group contribution models like UNIFAC are attractive, but a drawback is that group contribution does not take proximity effects into account,[15] which are likely very important in molecules with multiple strong functional groups, like polyphenols.[16] Moreover, some chemical groups present in polyphenols are still not accurately modeled, and the equilibrium data present in literature is not abundant.[12] Due to all the reasons mentioned above, the NRTL-SAC model[17] might be one of the best options available. It is an activity coefficient model that takes both excess enthalpy and excess entropy into account and that uses four molecular descriptors to characterize any molecule (X is a parameter related to hydrophobicity, Y^- and Y^+ to polarity, and Z to hydrophilicity) and only a relative small amount of equilibrium data is needed to regress those parameters.[18] This model has already been applied to describe solubility of polyphenols, but rarely to liquid−liquid equilibrium, despite being widely used for that purpose.[7]

Received: February 7, 2018
Revised: April 30, 2018
Accepted: June 19, 2018
Published: June 19, 2018

DOI: 10.1021/acs.iecr.8b00613
Ind. Eng. Chem. Res. 2018, 57, 9210−9221

In this work, four model polyphenols, naringenin (flavonoid), naringin (glycosylated flavonoid), *trans*-resveratrol (stilbene), and *p*-coumaric acid (phenolic acid) (Figure 1), were selected. Their partition coefficients in different solvents were predicted using NRTL-SAC and the results compared with obtained experimental data. The regressed molecular descriptors were then used to unveil which solvent properties were desirable (degree of hydrophobicity, polarity, and hydrogen bonding) for the purification of the studied polyphenols from a fermentation broth. The obtained results were ultimately used to suggest possible process configurations that can be applied to the downstream process of those compounds in two proposed situations: recovery of *trans*-resveratrol from a stream containing hydrophilic components (including *p*-coumaric acid) and recovery and purification of two similar hydrophobic polyphenols, *trans*-resveratrol and naringenin, from a fermentation broth.

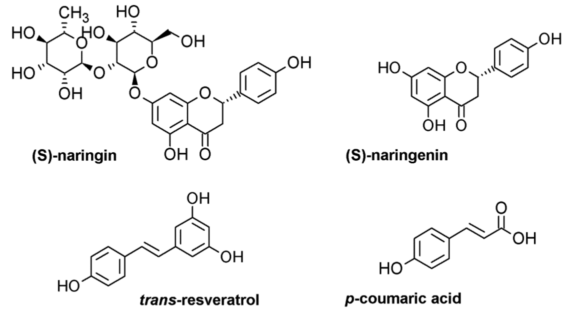


Figure 1. Chemical structure of the polyphenols studied in this work.

In the following section, a description of the used materials and methods is given, including a short description of the thermodynamic model employed. The main obtained results and their discussion are presented on section 3, and the conclusions are included in section 4.

2. MATERIALS AND METHODS

2.1. Chemicals. For the preparation of all the solutions, Milli-Q grade water was used. The polyphenol *trans*-resveratrol ≥98% was obtained from Olon S.P.A (Italy) for Evolva. Naringenin (natural (U.S.), 98%, lot #MKBW8466 V), naringin (≥95% (HPLC), lot #BCBM4171 V), and *p*-coumaric acid ≥98% (lot #BCBR8319 V) were purchased from Sigma-Aldrich. The used organic solvents are indicated in Table 1.

2.2. Partition Coefficient Determination. The partition coefficient of the used compounds was determined by the shake-flask method,[19] at the constant temperature of 25 °C. The experiments were performed in 15 mL Falcon tubes, using a liquid volume between 9 and 10 mL. The phase ratios were selected based on a preliminary guess of the partition coefficient, and, in most instances, a 1:1 ratio was used. The tubes were shaken for 90 min in a Sartorius Certomat BS-1 at 320 rpm (kinetic experiments were performed in order to check this was time enough to achieve equilibrium). Afterward, the tubes were centrifuged at 4000 rpm and 25 °C for 3 min. Samples were then taken from both phases and analyzed by UHPLC.

Table 1. List of Solvents Used for the Experiments Performed in This Work

solvent	supplier	purity
heptane	Sigma-Aldrich	anhydrous ≥99% (GC)
acetonitrile	Sigma-Aldrich	HPLC Plus ≥99.9%
isobutyl acetate	Fluka	Puriss ≈99%
butyl acetate	Fluka	Purum ≥98.5%
MTBE	Fluka	anhydrous 99.8%
1-octanol	Sigma-Aldrich	ACS reagent ≥99%
HCl 37% w/w	Sigma-Aldrich	ACS reagent 37%

For the experiments with *p*-coumaric acid, due to its acidic character which might influence partition, the aqueous phase consisted of a 10 mM HCl aqueous solution in order to ensure that the molecule would be in its neutral form.

2.3. Polyphenol Determination by UHPLC. The quantification of *p*-coumaric acid, *trans*-resveratrol, and naringenin was carried out by UHPLC (Ultimate 3000, Thermo Scientific, USA) in a C18 column (Acquity UPLC HSS column, 1.8 μm, 2.1 mm × 100 mm Waters, Milford, USA). Mobile phase A consisted of 10% formic acid in Milli-Q water and mobile phase B of 10% formic acid in acetonitrile. Every run was performed in isocratic mode, with the mobile phase containing 33.5% B and 66.5% A and flowing at 0.30 mL/min. The detection of *p*-coumaric acid was performed at 340 nm, *trans*-resveratrol at 304 nm, naringenin at 289 nm, and naringin at 283 nm.

2.4. The NRTL-SAC Model. The NRTL-SAC model is an excess Gibbs energy model, based on the polymer NRTL model. The activity coefficient of a given compound is taken to be the sum of a combinatorial (enthalpic) contribution and a residual (entropic) contribution:

$$\ln(\gamma_i) = \ln(\gamma_i^{\mathrm{C}}) + \ln(\gamma_i^{\mathrm{R}}) \tag{1}$$

The main difference from the original model is that instead of modeling the van der Waals interactions on a *per molecule* basis, the molecules are represented by four conceptual segments: X, Y^+, Y^-, and Z. Each of these segments represents the hydrophobic, polar, and hydrophilic character of each molecule, and the interaction energies between each segment are predefined.[17] The full mathematical treatment of this model can be found elsewhere.[17]

In this work, the four parameters (X, Y^+, Y^-, and Z) were regressed using solid–liquid equilibrium (SLE) data, available in the literature.[20−26] To relate the activity coefficient of the molecule in solution (γ_i) to its solubility, a simplified equation of the solid–liquid equilibrium relation was used:[27,28]

$$\ln(\gamma_i \cdot x_i) = \frac{\Delta H_{\mathrm{m}}}{RT_{\mathrm{m}}}\left(\frac{T_{\mathrm{m}}}{T} - 1\right) \tag{2}$$

Using the previous equation to describe solid–liquid equilibrium carries some approximations, among them the following:[29]

- The melting temperature of a given compound is close enough to its triple point temperature.
- The solvent is insoluble in the crystal lattice.
- The difference between the liquid and the solid heat capacities is considered negligible when compared to the fusion enthalpy.
- It is assumed that the solid–liquid transition occurs at a defined temperature point (the triple point temperature).

As most of the above-mentioned approximations are not far from reality, eq 2 has been applied before to the description of the solid−liquid equilibrium of polyphenols and found adequate.[28] For the case of naringin, however, one was confronted with a particular problem: some literature mentions that it is able to crystallize[30] (e.g., in water, as an octahydrate and with a melting temperature of 83 °C); another study suggests that it does not solidify in a crystalline structure and does not have a defined melting point.[31] In both cases, the approximations contained in eq 2 may no longer be valid (either the water of hydration is not taken into account or there is no defined melting point). Owing to that, two approaches were followed in this case: in one of them, an equation with the same structure as eq 2 was used, but leaving the melting temperature and enthalpy as regressing parameters; in the other one, the relation between the liquid−liquid partition coefficient and the activity coefficient of the molecule in the aqueous and organic phases was used instead:[19]

$$\frac{C_{i,\mathrm{org}}}{C_{i,\mathrm{aq}}}\cdot\frac{\overline{V}_{\mathrm{org}}}{\overline{V}_{\mathrm{aq}}} = P\cdot\frac{\overline{V}_{\mathrm{org}}}{\overline{V}_{aq}} = \frac{\gamma_{i,\mathrm{aq}}}{\gamma_{i,\mathrm{org}}} \tag{3}$$

where P is the volume-based partition coefficient, $\overline{V}_{\mathrm{org}}$ the molar volume of the organic phase, and $\overline{V}_{\mathrm{aq}}$ is the molar volume of the aqueous phase. Thus, the liquid−liquid equilibrium data obtained in this work for naringin was, in this last approach, not predicted using the NRTL-SAC model. Instead, it was only used for the purpose of parameter estimation.

3. RESULTS AND DISCUSSION

3.1. Model Validation Using SLE Data and Solvent Screening for LLE. The analyzed molecules in this work were selected from different polyphenol chemical classes. In that way, different downstream processes could be pictured, and a more general overview of the NRTL-SAC model prediction capabilities could be obtained. Both naringenin and naringin are flavonoids, although naringin has two additional sugar moieties, rendering it more hydrophilic. The molecule *trans*-resveratrol is a hydrophobic stilbenoid, while *p*-coumaric acid is a phenolic acid.

As previously mentioned, the data used for the regressed parameters for the NRTL-SAC model was obtained from different solid−liquid equilibrium data sets, which are available in the literature. The used references are described in Table A1 in the Appendix.

For the case of *p*-coumaric acid, being a weak acid, it is important to confirm that, in the solubility measurements, its main molecular form is the neutral form. Otherwise, two different molecules should be taken into account in the thermodynamic model. Given the pK_a of *p*-coumaric acid to be 4.6,[32] it was checked that for the measured solubility in water, the neutral species would account for 97% of the total concentration. Thus, it is considered that the estimated NRTL-SAC parameters are specified for the neutral molecule.

In order to apply eq 2, melting properties of the considered polyphenols are needed. This data is provided in Table 2. For naringin, and as previously mentioned, the melting enthalpy and melting temperature were added as regressing parameters, an approach suggested in previous work.[33]

Table 2. Chemical Properties of the Targeted Polyphenols[a]

polyphenol	M_w (g mol^{-1})	T_m (K)	ΔH_m (kJ mol^{-1})	ref
trans-resveratrol	228.25	541.3	30.6	20
p-coumaric acid	164.15	494.35	34.3	21
naringenin	272.26	523.15	39.8	25
naringin[b]	580.54	432.3	58.1	this work

[a]The melting temperature and the melting enthalpy were used to obtain the activity coefficient of each polyphenol in solution, using eq 2. [b]For this case, the melting temperature and enthalpy were not measured, but rather estimated from SLE data, using eq 2.

3.2. Parameter Determination for NRTL-SAC. For the parameter determination, the following objective function was used:[12]

$$\min_{X,Y^-,Y^+,Z} \sum_i [\ln(\gamma_i^{\mathrm{mod}}) - \ln(\gamma_i^{\mathrm{exp}})]^2 \tag{4}$$

In the previous equation, X, Y^+, Y^-, and Z are the molecular descriptors used by NRTL-SAC, γ_i^{mod} is the activity coefficient predicted by the model, and γ_i^{exp} the experimental activity coefficient, obtained using SLE data and by applying eq 2.

The needed parameters for the employed solvents were taken from the original NRTL-SAC paper.[17] However, whenever updated parameters were present in other sources,[34] those were used instead.

In order to check the applicability of the NRTL-SAC model for describing the measured solubility data, its predictions were compared with the experimental measurements. In Table 3, the regressed molecular descriptors are indicated, together with the percentage absolute average deviation (% AAD) (eq 5) associated with the determined activity coefficients:

Table 3. Regressed Molecular Descriptor Parameters for the NRTL-SAC Model[a]

polyphenol	X	Y^-	Y^+	Z	% AAD in γ (SLE data)	ref
trans-resveratrol	0.427	1.768	3.057	0.000	44.2	this work
p-coumaric acid	0.545	1.777	1.871	0.75	33.5	this work
naringenin	0.674	1.271	1.53	0.000	94.6	this work
naringin	0.190	0.000	2.016	0.748	49.3	this work

[a]It is also included the average absolute deviation between the experimentally determined activity coefficient and the one determined by NRTL-SAC.

$$\%\mathrm{AAD} = 100\cdot\frac{1}{N}\sum_{i=1}^{N}|\gamma_i^{\mathrm{mod}} - \gamma_i^{\mathrm{exp}}| \tag{5}$$

This previous parameter measures how distant, on average, the predicted activity coefficients are from the experimentally determined ones.

The correlation between experimental data and the predictions made by the NRTL-SAC model are also graphically represented in Figure 2.

As depicted in the Figure 2, and also corroborated by the obtained % AAD values, the NRTL-SAC was able to describe the observed data satisfactorily for most of the cases. The obtained average deviation values are also in line with similar studies.[35]

3.3. Liquid−Liquid Extraction Experiments. In order to prove that the predictions made by NRTL-SAC are concordant with the experimental values, liquid−liquid

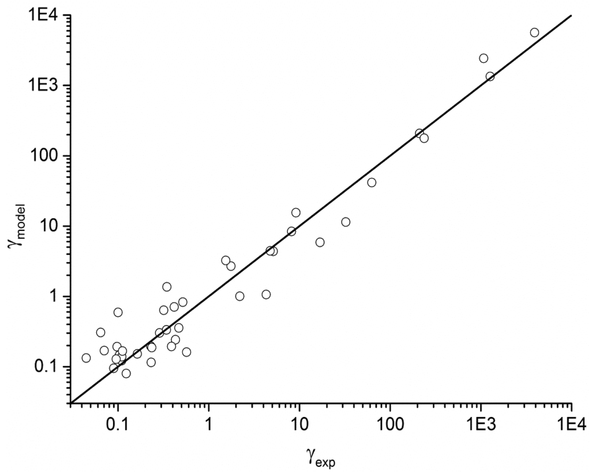


Figure 2. Comparison between experimental SLE data and predictions by NRTL-SAC, after having regressed the parameters for each polyphenol.

extraction experiments were performed and the obtained partition coefficients compared to those obtained from NRTL-SAC. From the analysis of the results obtained (see Figure 3), it was observed that for the case of naringin, although the model could describe the SLE data quite well (Table 3), the predictions for the LLE experiments were, in most of the cases, significantly higher than the experimental points. A possible explanation is related to the approximations contained in eq 2 to the specific case of naringin. In some literature, this compound is said to crystallize in water as an octahydrate and in other solvents as a dihydrate.[30] However, for the solid–liquid equilibrium relation in eq 2 to be valid, the energies of each dehydration step might have to be included.[36] In a different study, naringin has been observed not to have a defined melting point, possibly associated with having an amorphous rather than a crystalline form.[31] In this work, two approaches were followed in order to regress the molecular descriptor parameter for naringin: in the first one, the SLE data set available in the literature (indicated in Table A1) was used for the parameter estimation (first row in Table 4). In the other one, the same parameters were regressed, but using only the obtained liquid–liquid partition data in this work (second row in Table 4).

Table 4. Regressed Molecular Descriptor Parameters for Naringin Using the NRTL-SAC Model[a]

polyphenol	X	Y^-	Y^+	Z
naringin	0.190	0.000	2.016	0.748
naringin[b]	0.000	1.462	0.000	0.238

[a] Due to the difficulty associated with describing the solid–liquid equilibrium data for the molecule, a new set of parameters was estimated, using only liquid–liquid partition data. [b] The parameters for this molecule were estimated using only LLE data.

The overall performance of the NRTL-SAC model in describing the obtained partition data is indicated in Figure 3.

As shown in Figure 3, the naringin molecule was the one where the NRTL-SAC model had the worst performance, except when a different set of parameters was used, based on

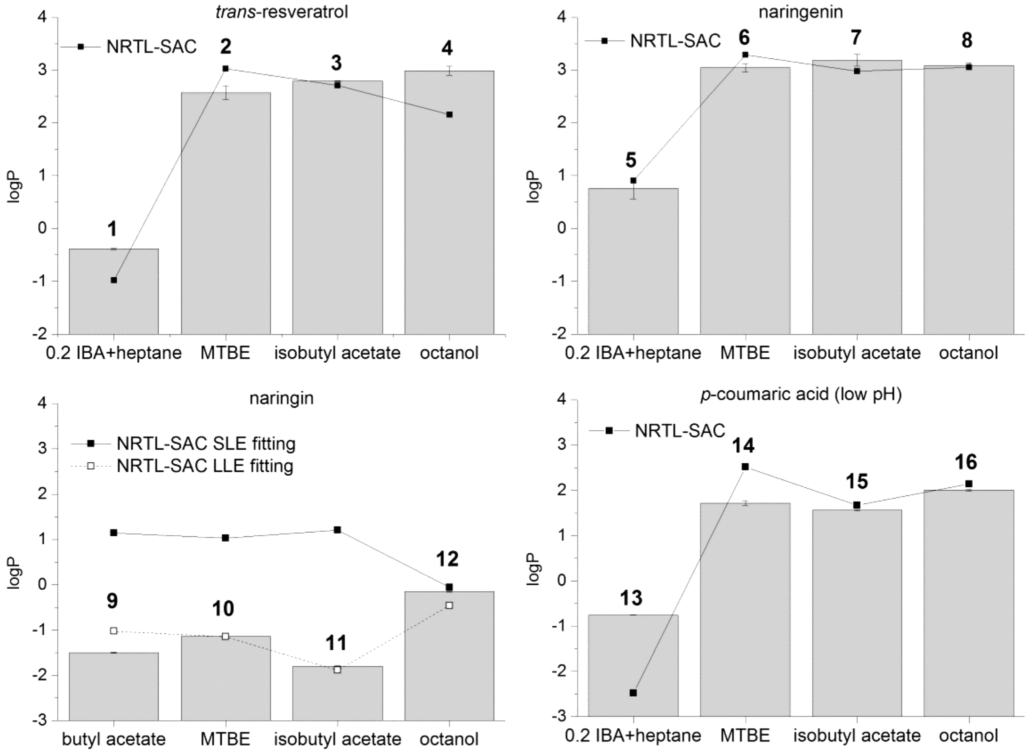


Figure 3. Determined partition coefficients of the four different polyphenols considered in this work. The partition was measured in four different solvents, for which the NRTL-SAC parameters were already determined. The vertical bars are the obtained experimental values, and the lines connected by squares are the NRTL-SAC model predictions. Numbers were included in the graphic in order to reference each experimental condition throughout this article.

DOI: 10.1021/acs.iecr.8b00613
Ind. Eng. Chem. Res. 2018, 57, 9210–9221

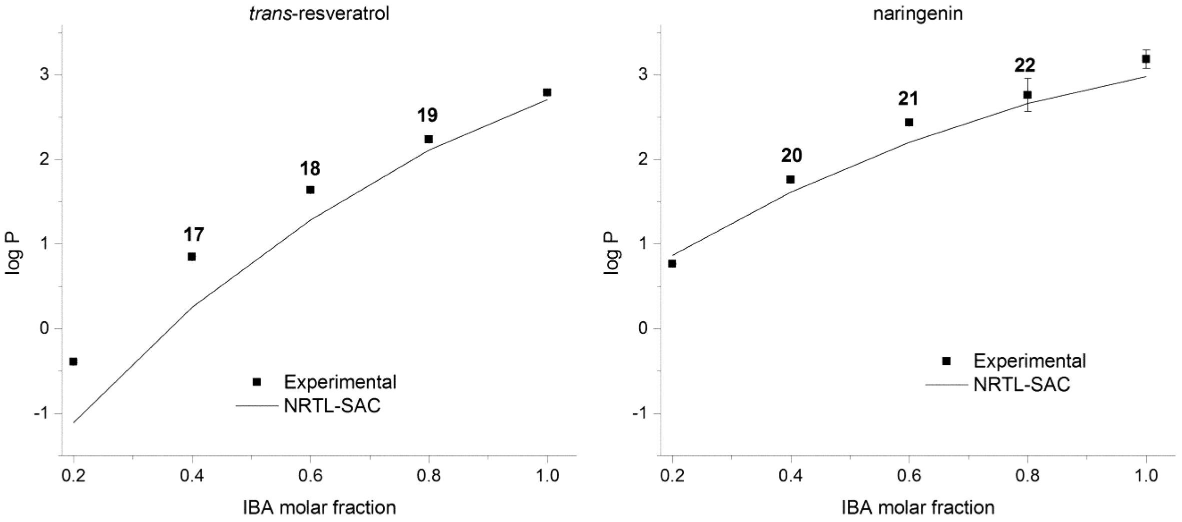


Figure 4. Partition coefficients of *trans*-resveratrol and naringenin in different mixtures of isobutyl acetate−heptane. The experimental values (solid squares) are compared with the predictions by NRTL-SAC (solid line). Numbers were included in the graphic in order to reference each experimental condition throughout this article.

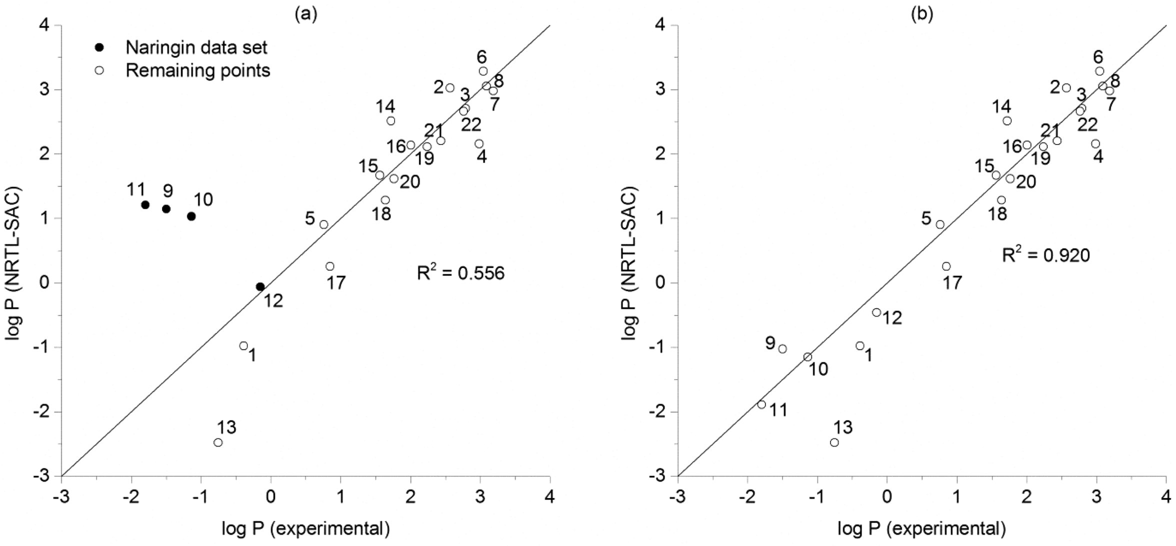


Figure 5. Comparison between experimental LLE data and the predictions made by NRTL-SAC. In graphic (a), all the points were predicted using the molecular descriptors of naringin, when regressed using SLE data. In graphic (b), the naringin data points were predicted by using the molecular descriptors when regressed with the LLE data. Numbers were included in the graphic in order to reference each experimental condition throughout this article.

the obtained LLE data. In all of the remaining situations, the predictions were in line with the experiments.

The partition coefficients of *trans*-resveratrol and naringenin were not only measured in pure solvents but also in different mixtures of heptane−isobutyl acetate (Figure 4). The goal with these experiments was to observe if by varying the amount of organic polar solvent (isobutyl acetate) to a hydrophobic solvent (heptane), the partition coefficient of two highly hydrophobic polyphenols could be fine-tuned. In this case as well, the NRTL-SAC is able to model the results almost quantitatively for every data point.

Like for the SLE predictions, the overall performance of the NRTL-SAC model was compared with the experimental liquid−liquid partition data. In Figure 5a, the results are shown when the parameters of naringin were obtained from SLE data. In Figure 5b, the parameters of naringin were obtained using only LLE data.

As indicated by the plots depicted in Figure 5, the NRTL-SAC predictions show a quite strong correlation with the experimental data, except for the case of naringin. However, the NRTL-SAC model itself is probably not the reason, but rather the possible simplified description of naringin solid−liquid equilibrium as previously mentioned. Either those details could be incorporated in a more complex thermodynamic model or the molecular descriptors of naringin (and possibly other glycosylated polyphenols) may be fitted to experimental LLE data, as it was performed in Figure 5b. The average relative error for the situation depicted in Figure 5a was 56% (35% when discarding naringin) and 40% when naringin parameters were regressed using LLE data (Figure 5b).

The information obtained up until this point, together with the newly regressed NRTL-SAC parameters, was used to propose possible scenarios for the recovery and purification of the considered polyphenols.

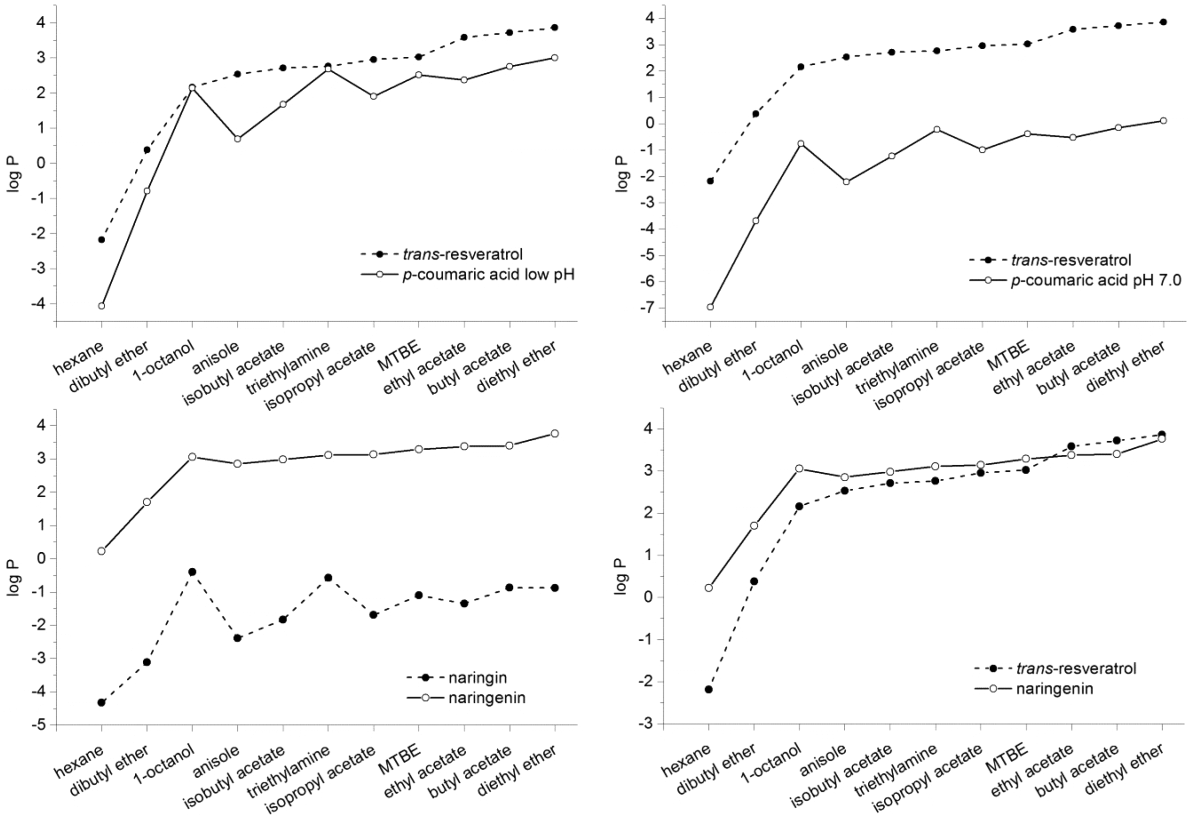


Figure 6. Predictions from NRTL-SAC for the partition coefficients of each polyphenol considered in this work. Each plot aims to compare the partition of two polyphenols that could be present in the same stream and that would have to be purified.

3.4. Process Design for Liquid–Liquid Extraction of Hydrophobic Polyphenols.

At this stage, it is important to determine the desired log *P* values for the different polyphenols, so that liquid–liquid extraction occurs as desired. For the recovery of hydrophobic polyphenols from a fermentation broth, the preferred log *P* values should be larger than 0. The reason is that, if the stream is going to be concentrated, the minimum solvent/aqueous feed ratio is 1:1. Moreover, the solvent stream should be able to extract the polyphenols (recovery depending on the number of stages) while leaving sugars, proteins, and organic acids behind, due to their low partition toward organic solvents. Because for those compounds the log *P* is lower than 0, purification is obtained if log $P \geq 0$ for the hydrophobic polyphenol. This statement is supported by obtained partition experimental data of glucose and proteins in a fermentation broth of *C. glutamicum* (data not shown) and on octanol/water partition data of organic acids present in the literature.[37]

If purification of closely related polyphenols is intended (e.g., *trans*-resveratrol and naringenin), the log *P* values have to meet more specific criteria. First of all, if the log *P* is too large for both of them, even after one stage of extraction, complete recovery of both compounds can be achieved as indicated by the Kremser equation.[7]

In the following equations, V_r stands for solvent/aqueous feed volume ratio, P_i is the partition coefficient (volume based) of compound *i*, *N* is the number of column stages, and Pur_i is the purity of compound *i*.

Assuming that the feed stream includes compounds 1 and 2 at 50% purity each and if both partition coefficients are so large that V_rP_1 and V_rP_2 become too large

$$\text{Pur}_1 = \frac{1 - \dfrac{V_rP_1}{(V_rP_1)^{N+1} - 1}}{2 - \dfrac{V_rP_1}{(V_rP_1)^{N+1} - 1} - \dfrac{V_rP_2}{(V_rP_2)^{N+1} - 1}} \tag{6}$$

$$\text{Pur}_1 \approx \frac{1 - (V_rP_1)^{-N}}{2 - (V_rP_1)^{-N} - (V_rP_2)^{-N}} \rightarrow \frac{1}{2}, \text{ for } N \text{ relatively small} \tag{7}$$

So, in the end, no purification occurs, as the purity of compound 1 remains at 50%. The ideal situation is when $V_rP_1 > 1$ and $V_rP_2 \approx 0$. Considering that maximum concentration factor is desired, that solubility of organic solvents in water is in the order of 50 g/L, and their density is on the order of 0.8 g/mL, V_r should be equal or larger than 0.0625. This is equivalent to saying that the feed stream should not be concentrated more than 16 times.

Assuming that the feed stream contains two compounds with 50% purity each and that the target is at least 70% purity, the log *P* of the impurity should be $0 < \log P \leq 0.83$, and the log *P* for the desired compound has to be at least 0.4 units higher. This information was obtained using the Kremser equation.

In Figure 6, and based on the regressed parameters for NRTL-SAC, the partition coefficients of the studied polyphenols were predicted for a wide list of solvents present in the NRTL-SAC model database. For the case of *p*-coumaric acid, being a weak acid ($\text{p}K_a = 4.6$), it is important to understand how the partition coefficient might vary with pH since, in some cases of interest, fermentations can occur at a pH considerably higher than its $\text{p}K_a$ (e.g., pH 7). Since the

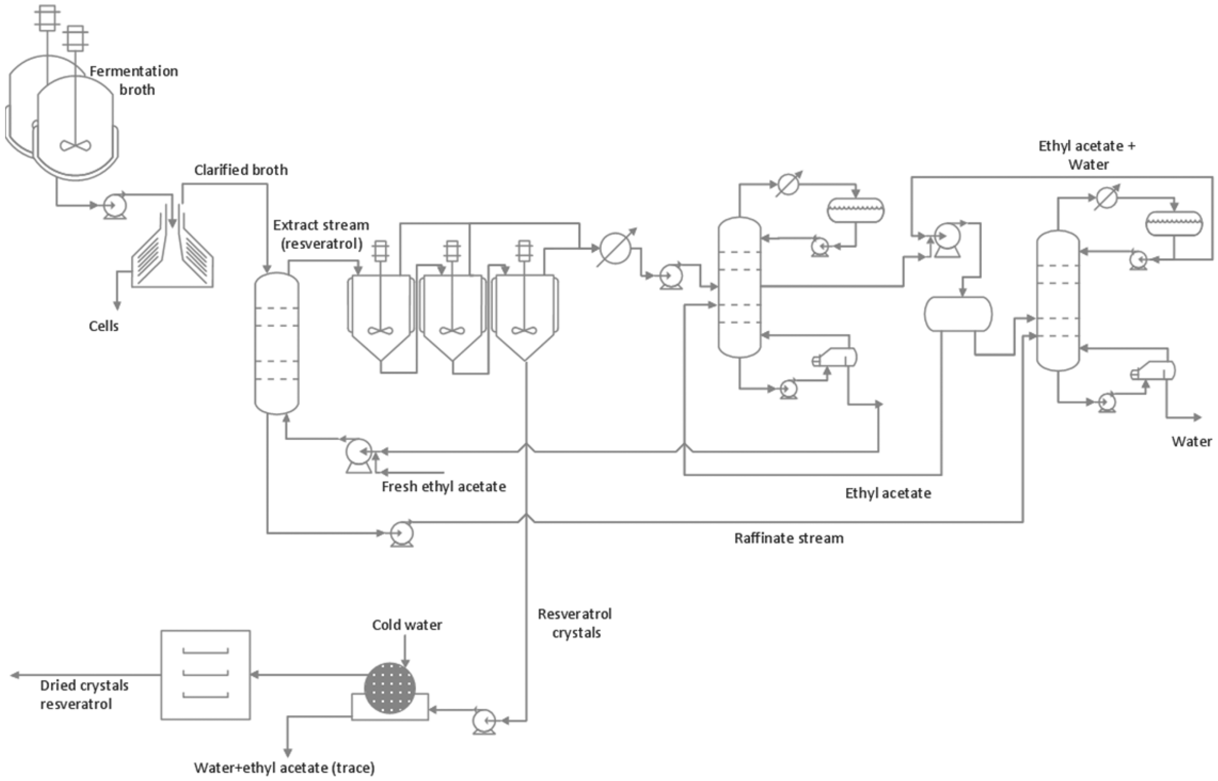


Figure 7. Conceptual downstream process train for the recovery of *trans*-resveratrol, a hydrophobic polyphenol, from a fermentation broth containing hydrophilic impurities. The clarified stream goes through a LLE step, where the polyphenol is preferentially extracted. Afterward, the desired compound is crystallized, filtered, washed, and dried to obtain the final formulation. The solvent used for the extraction, ethyl acetate, is distilled and recycled to the extraction column.

performed experiments for *p*-coumaric acid were performed at low pH, the partition coefficient at higher pH values (pH > pK_a + 2) was calculated by assuming that the deprotonated species *p*-coumarate does not partition to the organic phase:

$$\begin{aligned}\log D &= \log\left(\frac{[\mathrm{AH}]_{\mathrm{org}} + [A^{-}]_{\mathrm{org}}}{[\mathrm{AH}]_{\mathrm{aq}} + [A^{-}]_{\mathrm{aq}}}\right)\\ &\approx \log\left(\frac{[\mathrm{AH}]_{\mathrm{org}}}{[\mathrm{AH}]_{\mathrm{aq}} + K_{\mathrm{a}}[\mathrm{AH}]_{\mathrm{aq}}/[\mathrm{H}^{+}]_{\mathrm{aq}}}\right)\\ &= \log\left(\frac{[\mathrm{AH}]_{\mathrm{org}}}{[\mathrm{AH}]_{\mathrm{aq}}}\cdot\frac{1}{1 + K_{\mathrm{a}}/[\mathrm{H}^{+}]_{\mathrm{aq}}}\right)\\ &= \log P - \log(1 + 10^{\mathrm{pH}-\mathrm{p}K\mathrm{a}})\end{aligned} \tag{8}$$

In the previous equation, *D* is the partition coefficient of the weak acid *p*-coumaric acid that depends on pH. On the other hand, *P* is the partition coefficient of the neutral species.

Four hypothetical situations were considered. The first one considers a possible purification of *trans*-resveratrol and *p*-coumaric acid from a stream containing both compounds at low pH (approximately below 4, where most *p*-coumaric acid is in its protonated form). The second situation is equivalent but at pH 7, where *p*-coumaric acid should be mostly deprotonated. The remaining two cases represent a hypothetical separation between naringenin and naringin (a glycosylated and nonglycosylated polyphenol) and between naringenin and *trans*-resveratrol (two hydrophobic polyphenols), respectively.

Based on the results presented in Figure 6, one can expect the purification between naringenin and *trans*-resveratrol to be the most difficult one, as they have very similar partition profiles, the major difference being when heptane is used. Another important case to mention is naringin. Because it exhibits negative log *P* values for the considered solvents, it might be easy to purify naringin from naringenin but not to recover naringin from the remaining broth components such as sugars and proteins. As previously mentioned, a log *P* value lower than zero is not desired because that means that the flow rate of extract phase would need to be larger than the one for the aqueous stream. Since that would lead to dilution of a possibly already dilute feed stream (considering the current low titers in fermentation), other recovery methods, based on adsorption, for example, would probably be a better option.

Taking into account the previous discussion, two purification cases will be analyzed in more depth in the following section. In the first case, the recovery of a hydrophobic polyphenol recovery from a typical fermentation broth (containing hydrophilic compounds as impurities) will be considered. Next, the recovery and purification of two similar hydrophobic polyphenols—naringenin and *trans*-resveratrol—from a fermentation broth will also be tackled.

3.4.1. LLE Process Design for the Recovery of trans-Resveratrol.

In this section, three different scenarios will be given for the recovery of *trans*-resveratrol from a hypothetical fermentation broth containing other, hydrophilic molecules: sugars, proteins, and organic acids (p$K_a \approx 4.5$). As indicated in the previous section, the goal is to use a solvent to which the hydrophobic polyphenol can partition with log $P \geq 0$. For that, and as shown in Figure 6, almost any oxygenated organic solvent could meet the purpose.

3.4.1.1. First Scenario: Liquid Extraction with Organic Solvent. In the first scenario, a liquid–liquid extraction step is performed with ethyl acetate (Figure 7). The raffinate stream should contain the unextracted hydrophilic molecules, while the extract stream should carry the purified and concentrated *trans*-resveratrol. Because the solubility of this polyphenol in ethyl acetate is relatively high, it should be concentrated in a multiple effect evaporator, before obtaining enough supersaturation to be crystallized. While the obtained crystals are filtered, washed with cold water, and dried, the evaporated solvent is directed to a series of two distillation columns, where ethyl acetate is recovered and recycled to the extraction column.

3.4.1.2. Second Scenario: Using an Antisolvent for Precipitation. In the second proposed scenario, a similar strategy to the one depicted in Figure 7 is proposed. The hydrophobic polyphenol is first extracted using an organic solvent, and the solution is then concentrated near solubility limit by evaporation. The solvent evaporated in this step is also recycled back to the process. The difference, in this case, is that depending on the solvent used in the LLE step (if it can be made miscible with water or not), heptane or water can be added as an antisolvent to make the polyphenol precipitate. This could reduce the amount of energy spent in evaporating the organic solvent (ethyl acetate in the previous case). As in the previous case, the organic solvents would be recycled back to the process by distillation.

3.4.1.3. Third Scenario: Liquid Extraction with Switchable Solvent. The last, third scenario, is also a small modification to the first one. For that reason, the same Figure 7 can be taken as a reference. In this case, the extraction is performed with a "green" solvent such as an ionic liquid or a switchable solvent. Assuming that this solvent might have a high boiling point, the preliminary concentration step by evaporation is skipped. The polyphenol is instead precipitated by adding water as antisolvent. For the recovery, instead of using distillation the used solvent would first be made immiscible with water either by adding CO_2 in the case of the switchable solvent or by adding an additive or changing the temperature for the case of the ionic liquid. The two phases would then be separated, for example, by centrifugation, before recycling them back to the process.

3.4.2. LLE Process Design for the Recovery and Purification of trans-Resveratrol and Naringenin. As previously mentioned, for the goal of purification, the objective is to have a large selectivity (log P difference of minimum 0.4) but also to have a relatively low partition coefficient for one of the compounds ($0 < \log P \leq 0.83$).

One of the advantages of the NRTL-SAC model is that it represents each molecule to be composed of different segments: hydrophobic (X), polar (Y^{-} and Y^{+}), and hydrophilic (Z). Due to that, it was investigated which value combination of X, Y^{-}, Y^{+}, and Z for a hypothetical solvent would lead to an optimal purification scenario, where optimal is defined as meeting the following constraints:

$$0 < (\log P)_1 \leq 0.83 \wedge (\log P)_2 \geq 1.23 \quad (9)$$

Or the other way around:

$$0 < (\log P)_2 \leq 0.83 \wedge (\log P)_1 \geq 1.23 \quad (10)$$

For each of the molecular descriptors, their value was varied from 0 to 1 in 0.1 intervals, and the log P value evaluated. The result is present in Figure 8.

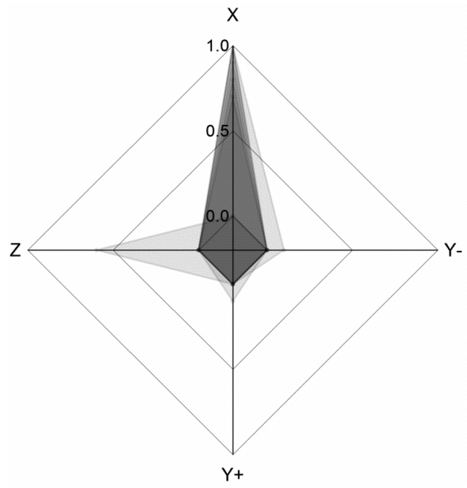


Figure 8. Prediction made by NRTL-SAC for the set of molecular descriptors that would make the optimal solvent for the purification of naringenin from *trans*-resveratrol.

As it is possible to verify, the predictions made by the NRTL-SAC model suggest using either a purely hydrophobic solvent (e.g., hexane) or a relatively hydrophilic solvent. The issue with the latter is that a solvent with high hydrophilic character will not be able to form two phases with a water-based fermentation broth. Thus, that hypothesis was disregarded. Because not always will a solvent with the desired hydrophobicity exist or be suitable (hexane, for example, has carcinogenic effects), one possible strategy for purification is to use a hydrophobic solvent such as heptane together with an organic polar solvent (e.g., octyl acetate, methyl ethyl ketone) in order to adjust the partition coefficient of the desired molecule. This was demonstrated in the experimental data shown previously, where the partition coefficients of both naringenin and *trans*-resveratrol were obtained for different heptane/isobutyl acetate mixtures.

Another way of fine-tuning the desired solvent characteristics for the purification of this polyphenols is to use ionic liquids. For that reason, the possible applicability of ionic liquids to the extraction of the polyphenols studied in this work was also examined. However, due to the relatively small database of NRTL-SAC parameters available in the literature,[38] only some ionic liquids were considered. Those were, from the indicated reference, the ones that NRTL-SAC predicted to form a biphasic mixture with water. Their associated molecular descriptors are indicated in Table A2 in the Appendix.

As it is possible to observe in Figure 9, the application of the examined ionic liquids seems to be only suitable for a possible purification scenario of *p*-coumaric acid and *trans*-resveratrol. For the other hypothetical purification steps, and as suggested by Figure 8, the applied solvents do not show sufficient hydrophobicity. For that reason, NRTL-SAC predicts hydrophobic polyphenols to have a log P value of equal to or larger than 1, which can make the purification task too challenging since purification may not be obtained (polyphenols are coextracted). Nonetheless, it is important to reinforce the idea

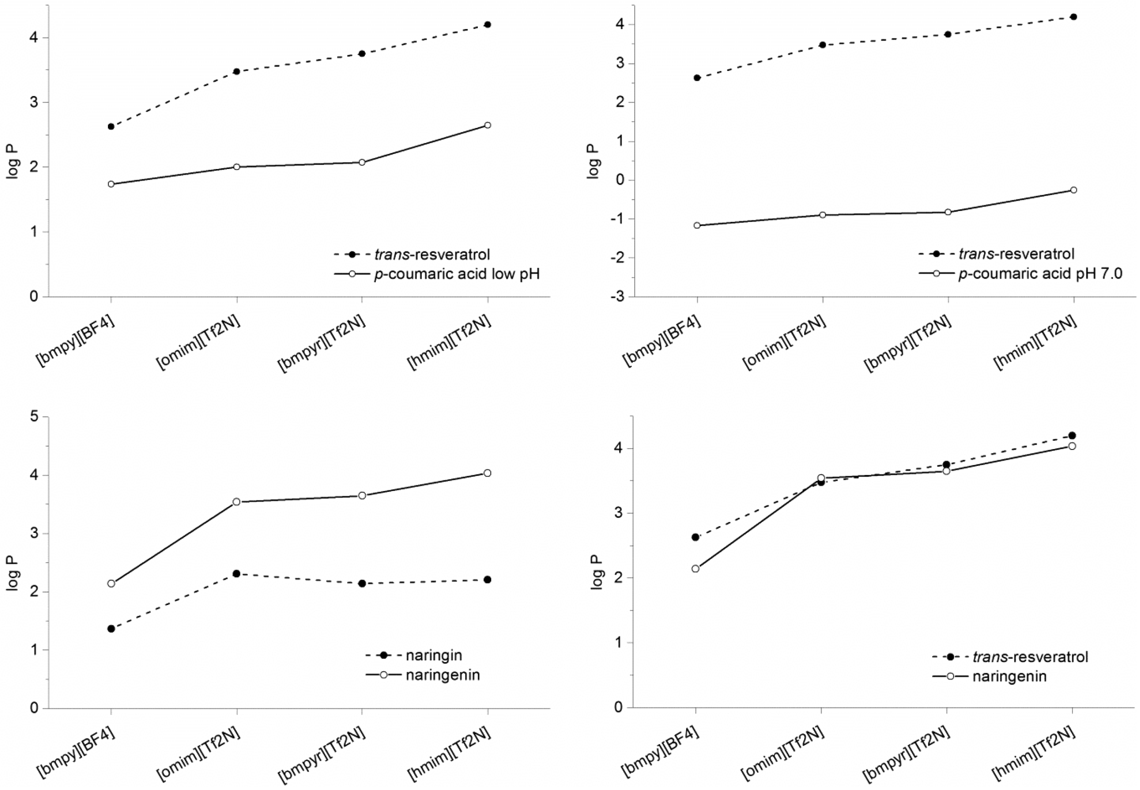


Figure 9. Predictions from NRTL-SAC for the partition coefficients of each polyphenol considered in this work. Each plot aims to compare the partition of two polyphenols that could be present in the same stream and that would have to be purified. The NRTL-SAC parameters for the considered ionic liquids are indicated in Table A2 in the Appendix.

that a very small set of ionic liquids were here evaluated due to the lack of sufficient parameters for the NRTL-SAC model.

Similarly to the previously proposed polyphenol recovery scenarios, a possible purification sequence for polyphenols such as *trans*-resveratrol and naringenin is indicated in Figure 10. The suggested downstream process can be made almost completely similar to the previously indicated one (Figure 7), the only modification being the addition of another liquid—liquid extraction column, where a second organic solvent is used. The idea behind this strategy is first to use a solvent that is able to extract the most hydrophobic polyphenol, naringenin (e.g., 0.2 molar fraction ethyl acetate in heptane), while leaving the less hydrophobic, *trans*-resveratrol, behind. The latter would be extracted using a more polar solvent mixture (e.g., 0.6 molar fraction ethyl acetate in heptane).

4. CONCLUSIONS

In this work, it was shown that the NRTL-SAC model could be applied to the description of liquid—liquid equilibrium data of complex molecules like polyphenols. Moreover, by only relying on pre-existent solubility data to regress the needed parameters, satisfactory predictions of the log *P* value were obtained (30% absolute average deviation), proving that the model is robust despite using a relatively small amount of data. The major exception occurred with naringin (56% AAD, when taking it into account), whose log *P* values were consistently overpredicted. The proposed explanation relies on the fact that the used solid-equilibrium equation is not applicable to amorphous substances but rather to compounds with a defined melting point. Moreover, different naringin solvates (different hydrates) might exist in the solid phase, depending on the solvent where solubility was measured.

The molecular descriptors (X, Y^{-}, Y^{+}, and Z) obtained for the four polyphenols considered in this study, together with the Kremser equation, allowed to define and find suitable solvents or solvent mixtures for applying liquid—liquid extraction to the recovery and purification of those molecules. Although the purification of naringin from a water-based fermentation broth was not studied in detail, the obtained liquid—liquid equilibrium results suggested that reverse-phase adsorption might be a more suitable alternative since its polarity is more similar to the remaining broth components than the other considered polyphenols.

Concerning the three remaining polyphenols (*trans*-resveratrol, naringin, and *p*-coumaric acid), two case studies were addressed in this study. The first one concerned the recovery of *trans*-resveratrol from a fermentation broth containing hydrophilic molecules, namely *p*-coumaric acid. By relying on NRTL-SAC, it was shown that, at neutral pH, almost any polar organic solvent would be able to purify *trans*-resveratrol while leaving *p*-coumaric acid behind (taken as the most challenging molecule to be removed). The second scenario dealt with the purification of both naringenin and *trans*-resveratrol from the same fermentation broth. In this case, the biggest challenged relied on the fact that both molecules are similar and highly hydrophobic. Thus, a binary solvent mixture which would combine a hydrophobic solvent (e.g., heptane) with a polar organic solvent (e.g., isobutyl acetate) was found to be one of the possible ways to fine-tune the partition value of each molecule in order to achieve separation.

DOI: 10.1021/acs.iecr.8b00613
Ind. Eng. Chem. Res. 2018, 57, 9210−9221

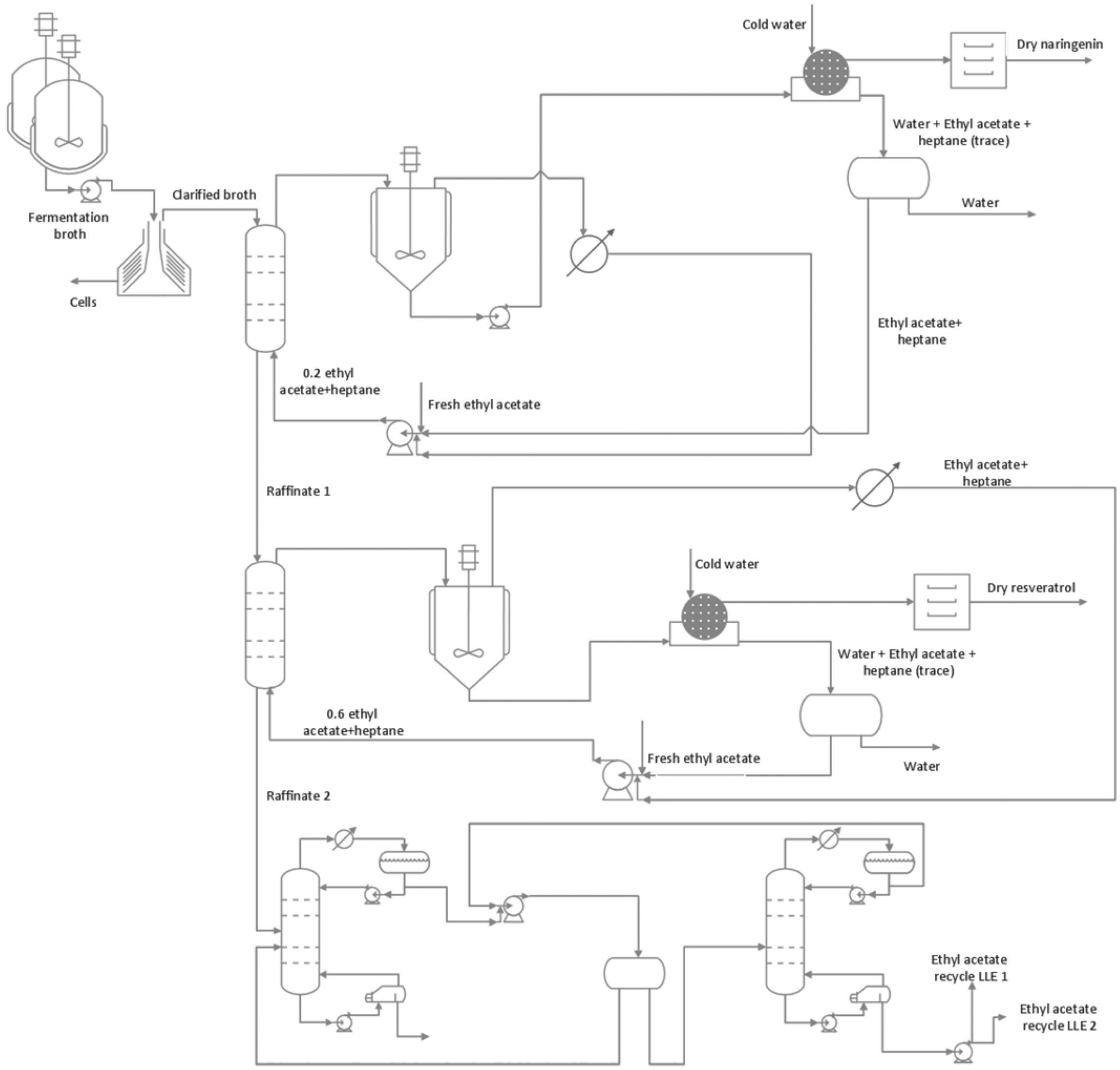


Figure 10. Conceptual downstream process train for the recovery and purification of two similar polyphenols (e.g., *trans*-resveratrol and naringenin) from a fermentation broth. The clarified stream goes through a first LLE step, where naringenin is preferentially extracted by a relatively hydrophobic stream (0.2 molar fraction ethyl acetate in heptane). Connected to it, there is another LLE step where *trans*-resveratrol, present in the "raffinate 1" stream, is recovered by using a more polar solvent mixture (0.6 molar fraction ethyl acetate in heptane). After liquid extraction is complete, both compounds are crystallized, filtered, washed, and dried to obtain the final formulation. The organic solvents are recovered by using distillation and then recycled to the different extraction columns.

The proposed conceptual downstream process designs for the two previously mentioned case studies had in consideration the possible usage of green solvents as ionic liquids or switchable solvents. Their advantages, like using CO_2 to promote phase-splitting in water or the fact that ionic liquids are considered "designer" solvents, are highly desired properties for the design of a liquid–liquid extraction process. Not only do they present such advantages—among others—but by using NRTL-SAC, the proper ionic liquid may now be developed with the right combination of hydrophobic–polar–hydrophilic character. Nonetheless, the lack of the needed parameters for NRTL-SAC in literature made it impossible not only to scan a wide database of solvents but also to find ionic liquids that would present better properties (regarding extraction capacity and selectivity) than the considered organic solvents.

■ APPENDIX

Literature data regarding the regressed parameters for the NRTL-SAC model are provided in Table A1. Associated molecular descriptors are indicated in Table A2.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: m.ottens@tudelft.nl. Telephone: +31 15 2782151.

ORCID

Marcelo Silva: 0000-0002-8606-5598

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the European Commission in the seventh Framework Programme (BacHBerry, Project No. FP7-613793).

Table A1. Description of the Solid−Liquid Equilibrium (SLE) Dataset Used for the NRTL-SAC Parameter Estimation for the Four Different Polyphenols Used in This Work

polyphenol	solvents	no. of data points	ref
trans-resveratrol	methanol, ethanol, 1-propanol, 2-propanol, 1-butanol, ethyl acetate, methyl acetate, acetone, and water	10	26
p-coumaric acid	methanol, ethanol, 1-propanol, 2-propanol, 1-butanol, acetone, ethyl acetate, methyl acetate, and water	9	21
naringenin	methanol, isopropyl alcohol, ethyl acetate, 1-butanol, hexane, water, and ethanol	11	22, 23
naringin	methanol, isopropyl alcohol, ethyl acetate, 1-butanol, hexane, water, and ethanol	11	23, 24

Table A2. Molecular Descriptors Parameters for Ionic Liquids Already Described in the Literature

polyphenol	X	Y^-	Y^+	Z	ref
[omim][Tf2N]	0.326	0.447	4.740	0.000	38
[bmpyr][Tf2N]	0.000	0.000	4.444	0.197	38
[hmim][Tf2N]	0.320	0.000	3.908	0.405	38
[bmpy][BF4]	0.000	0.062	1.224	0.158	38

■ ABBREVIATIONS

NRTL-SAC = nonrandom two-liquid segment activity coefficient
MPP-UNIFAC = modified polyphenol UNIFAC
LLE = liquid−liquid extraction
SLE = solid−liquid equilibrium

■ REFERENCES

(1) Transparency Market Research. Polyphenols Market by Product (Grape Seed, Green Tea, Apple, and Others), by Application (Functional Beverages, Functional Food, Dietary Supplements, and Others): Global Industry Analysis, Size, Share, Growth, Trends, and Forecast, 2012−2018; 2013.

(2) Alov, P.; Tsakovska, I.; Pajeva, I. Computational Studies of Free Radical-Scavenging Properties of Phenolic Compounds. *Curr. Top. Med. Chem.* **2015**, *15* (2), 85−104.

(3) Scalbert, A.; Johnson, I. T.; Saltmarsh, M. Polyphenols: Antioxidants and beyond. *Am. J. Clin. Nutr.* **2005**, *81* (1), 215S−217S.

(4) Pandey, K. B.; Rizvi, S. I. Plant polyphenols as dietary antioxidants in human health and disease. *Oxid. Med. Cell. Longevity* **2009**, *2* (5), 270−278.

(5) Dudnik, A. BacHBerry: BACterial Hosts for production of Bioactive phenolics from bERRY fruits. *Phytochem. Rev.* **2018**, *17*, 291.

(6) Goldberg, E. *Handbook of Downstream Processing*, 1st ed. ed.; Springer Netherlands: Amsterdam, 1997.

(7) Henley, E. J.; Seader, J. D.; Roper, D. K. *Separation process principles*; Wiley: New York, 2011.

(8) Gani, R.; Brignole, E. A. Molecular Design of Solvents for Liquid Extraction Based on Unifac. *Fluid Phase Equilib.* **1983**, *13* (Oct), 331−340.

(9) Weis, D. C.; Visco, D. P. Computer-aided molecular design using the Signature molecular descriptor: Application to solvent selection. *Comput. Chem. Eng.* **2010**, *34* (7), 1018−1029.

(10) Zhang, J. A.; Peng, D. L.; Song, Z.; Zhou, T.; Cheng, H. Y.; Chen, L. F.; Qi, Z. W. COSMO-descriptor based computer-aided ionic liquid design for separation processes. Part I: Modified group contribution methodology for predicting surface charge density profile of ionic liquids. *Chem. Eng. Sci.* **2017**, *162*, 355−363.

(11) Ferreira, O.; Pinho, S. P. Solubility of Flavonoids in Pure Solvents. *Ind. Eng. Chem. Res.* **2012**, *51* (18), 6586−6590.

(12) Sevillano, D. M.; van der Wielen, L. A. M.; Hooshyar, N.; Ottens, M. MPP-UNIFAC, a predictive activity coefficient model for polyphenols. *Fluid Phase Equilib.* **2014**, *384*, 82−88.

(13) Klamt, A. Conductor-Like Screening Model for Real Solvents - a New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99* (7), 2224−2235.

(14) Bouillot, B.; Teychené, S.; Biscans, B. An evaluation of thermodynamic models for the prediction of drug and drug-like molecule solubility in organic solvents. *Fluid Phase Equilib.* **2011**, *309* (1), 36−52.

(15) Dimian, A.; Bildea, C.; Kiss, A. *Integrated Design and Simulation of Chemical Processes*; 2nd ed.; Elsevier Science: Amsterdam, The Netherlands, 2014.

(16) Wu, H. S.; Sandler, S. I. Proximity effects on the predictions of the UNIFAC model: I. Ethers. *AIChE J.* **1989**, *35* (1), 168−172.

(17) Chen, C. C.; Song, Y. H. Solubility modeling with a nonrandom two-liquid segment activity coefficient model. *Ind. Eng. Chem. Res.* **2004**, *43* (26), 8354−8362.

(18) Chen, C.-C.; Crafts, P. A. Correlation and prediction of drug molecule solubility in mixed solvent systems with the nonrandom two-liquid segment activity coefficient (NRTL-SAC) model. *Ind. Eng. Chem. Res.* **2006**, *45* (13), 4816−4824.

(19) Sangster, J. *Octanol−Water Partition Coefficients: Fundamentals and Physical Chemistry*; John Wiley: New York, 1997.

(20) Sun, X. L.; Shao, Y. D.; Yan, W. D. Measurement and Correlation of Solubilities of trans-Resveratrol in Ethanol plus Water and Acetone plus Water Mixed Solvents at Different Temperatures. *J. Chem. Eng. Data* **2008**, *53* (11), 2562−2566.

(21) Ji, W.; Meng, Q.; Li, P.; Yang, B.; Wang, F.; Ding, L.; Wang, B. Measurement and Correlation of the Solubility of p-Coumaric Acid in Nine Pure and Water + Ethanol Mixed Solvents at Temperatures from 293.15 to 333.15 K. *J. Chem. Eng. Data* **2016**, *61* (10), 3457−3465.

(22) Zhang, P. P.; Lin, R.; Yang, G. D.; Zhang, J. Y.; Zhou, L.; Liu, T. T. Solubility of Naringenin in Ethanol and Water Mixtures. *J. Chem. Eng. Data* **2013**, *58* (9), 2402−2404.

(23) Zhang, L. Q.; Song, L.; Zhang, P. P.; Liu, T. T.; Zhou, L.; Yang, G. D.; Lin, R.; Zhang, J. Y. Solubilities of Naringin and Naringenin in Different Solvents and Dissociation Constants of Naringenin. *J. Chem. Eng. Data* **2015**, *60* (3), 932−940.

(24) Zhang, J. Y.; Zhang, P. P.; Liu, T. T.; Zhou, L.; Zhang, L. Q.; Lin, R.; Yang, G. D.; Wang, W. R.; Li, Y. P. Solubility of naringin in ethanol and water mixtures from 283.15 to 318.15 K. *J. Mol. Liq.* **2015**, *203*, 98−103.

(25) Chebil, L.; Humeau, C.; Anthoni, J.; Dehez, F.; Engasser, J. M.; Ghoul, M. Solubility of flavonoids in organic solvents. *J. Chem. Eng. Data* **2007**, *52* (5), 1552−1556.

(26) Sun, X.; Peng, B.; Yan, W. Measurement and correlation of solubility of trans-resveratrol in 11 solvents at T = (278.2, 288.2, 298.2, 308.2, and 318.2)K. *J. Chem. Thermodyn.* **2008**, *40* (4), 735−738.

(27) Chen, F. X.; Qi, Z. L.; Feng, L.; Miao, J. Y.; Ren, B. Z. Application of the NRTL method to correlate solubility of diosgenin. *J. Chem. Thermodyn.* **2014**, *71*, 231−235.

(28) Sevillano, D. M. n.; van der Wielen, L. A.; Trifunovic, O.; Ottens, M. Model Comparison for the Prediction of the Solubility of Green Tea Catechins in Ethanol/Water Mixtures. *Ind. Eng. Chem. Res.* **2013**, *52* (17), 6039−6048.

(29) Prausnitz, J. M.; Azevedo, E. G. d. *Molecular Thermodynamics of Fluid-Phase Equilibria*; 3rd ed.; Prentice Hall: Upper Saddle River, NJ, 1999.

(30) Hendrickson, R.; Kesterson, J. W. Purification of naringin. *Florida Agricultural Experiment Station J.* 1956.

(31) Cordenonsi, L. M.; Sponchiado, R. M.; Campanharo, S. C.; Garcia, C. V.; Raffin, P. R.; Schapoval, E. E. S. Study of Flavonoids present in Pomelo (Citrus maxima) by DSC, UV-VIS, IR, 1H and 13C NMR and MS. *Drug Analytical Res.* **2017**, *1*, 31−37.

(32) Blum, U. *Plant-Plant Allelopathic Interactions II*; Springer International Publishing: New York, 2014; p 322.

(33) von Stockar, U.; Van der Wielen, L. A. M. *Biothermodynamics: The Role of Thermodynamics in Biochemical Engineering*; CRC Press: Boca Raton, FL, **2013**.

(34) Haghtalab, A.; Yousefi Seyf, J. Vapor−Liquid and Solid−Liquid Modeling with a Universal Quasichemical Segment-Based Activity Coefficient Model. *Ind. Eng. Chem. Res.* **2015**, *54* (34), 8611−8623.

(35) Mota, F. L.; Carneiro, A. R.; Queimada, A. J.; Pinho, S. P.; Macedo, E. A. Temperature and solvent effects in the solubility of some pharmaceutical compounds: Measurements and modeling. *Eur. J. Pharm. Sci.* **2009**, *37* (3−4), 499−507.

(36) Franklin, S. J.; Younis, U. S.; Myrdal, P. B. Estimating the Aqueous Solubility of Pharmaceutical Hydrates. *J. Pharm. Sci.* **2016**, *105* (6), 1914−1919.

(37) Sangster, J. Octanol-Water Partition-Coefficients of Simple Organic-Compounds. *J. Phys. Chem. Ref. Data* **1989**, *18* (3), 1111−1229.

(38) Chen, C. C.; Simoni, L. D.; Brennecke, J. F.; Stadtherr, M. A. Correlation and prediction of phase behavior of organic compounds in ionic liquids using the nonrandom two-liquid segment activity coefficient model. *Ind. Eng. Chem. Res.* **2008**, *47* (18), 7081−7093.