Partition coefficients of Phenolic Compounds in Biphasic Organic Solvent Systems

Isabella Fernanda Weber Cordova

Final dissertation report submitted to ***Escola Superior de Tecnologia e Gestão*** of ***Instituto Politécnico de Bragança*** to obtain the **Master's Degree in Chemical Engineering** in the scope of the double diploma with *Universidade Tecnológica Federal do Paraná – Campus Ponta Grossa*

Supervisors

Maria Olga de Amorim e Sá Ferreira

Simão Pedro de Almeida Pinho

Priscilla dos Santos Gaschi Leite

João Manuel Costa Araújo Pereira Coutinho

Bragança

July 2021

"We are star stuff which has taken its destiny into its own hands."

Carl Sagan

Acknowledgments

First of all, I would like to express my sincere gratitude to the supervisors from IPB, professors Olga Ferreira and Simão Pinho, whose guidance, dedication and support has been so meaningful throughout this work.

I also want to thank the professors Priscilla Leite and Juliana Pietrobelli for support and encouragement for participating in the Double Degree Program. I am very grateful for having provided me with this incredible experience.

I would like to thank all my friends and colleagues that I made during this time, especially to Sérgio Vilas-Boas, for all the help, guidance and patience during this work.

I am also very grateful for my family and friends, for all the encouragement, especially for my parents Carlos and Julieta, for all the love, support and education.

Finally, many thanks to everyone that took part in the study and enabled this dissertation to be possible.

This work was developed within the scope of the project AllNat – "Using natural deep eutectic solvents for the extraction of bioactive compounds from plant material" (reference POCI-01-0145-FEDER-030463, PTDC/EQU-EPQ/30463/2017), funded by FEDER funds through COMPETE2020 – Programa Operacional Competitividade e Internacionalização (POCI), Portugal 2020 and by national funds through Foundation for Science and Technology (FCT/MCTES). I.W. Cordova also thanks project AllNat - POCI-01-0145-FEDER-030463 for her scholarship.

Abstract

Centrifugal Partition Chromatography (CPC) can be used to fractionate and purify complex mixtures, presenting several advantages over traditional separation methods, such as one-step separation of natural compounds from crude extracts and the potential to be scaled up to industrial levels.

In this work, quercetin, vanillin, rutin, ferulic acid and hesperetin were chosen as representative of a few important families of phenolic compounds that can be found in natural extracts (mostly flavonoids and phenolic acids). One of the most important and time consuming steps in the design of a separation process using CPC is the selection of the most effective biphasic solvent system due to the huge number of possible combinations. Following a literature review, the following biphasic solvent systems were selected: the conventional Arizona quaternary mixture N (equal volumes of heptane, ethyl acetate, methanol and water), and two modified versions containing alternative green solvents (version 1: replacing heptane by limonene, and methanol by ethanol; version 2: replacing only methanol by ethanol) and four binary solvent systems (water + ethyl acetate, water + 1-butanol, water + 1-octanol and heptane + methanol). The partition coefficients of the model solutes were measured at (298.2 ± 0.5) K by the shake-flask method, generally showing consistent results in comparison with existent literature data.

Regarding the thermodynamic modelling of the partition data, the UNIFAC (Universal Quasichemical Functional Group Activity Coefficient), the NRTL-SAC (Nonrandom Two-Liquid Segment Activity Coefficient Model) and the Abraham solvation model were already suggested. The UNIFAC model is totally predictive, however, it did not present consistent results in comparison with the other methods. In addition, quercetin and rutin could not be represented by the available groups. On the other hand, the Abraham solvation model was able to describe the partition coefficients data, with a global RMSD (root-mean-square deviation) of 0.57. The results obtained by the Abraham solvation model are comparable to the NRTL-SAC model, with a global RMSD of 0.3. The former has the advantage of its mathematical simplicity, represented by two linear free energy relationships, but it is more difficult to obtain the parameters for biphasic systems composed of more than two solvents.

Keywords: Partition coefficient, solubility, UNIFAC, Abraham, NRTL-SAC, chromatography.

Resumo

A Cromatografia de Partição Centrífuga (CPC) pode ser usada para fracionar e purificar misturas complexas, apresentando várias vantagens em relação aos métodos de separação tradicionais, como a separação em uma etapa de extratos brutos e poder ser projetada para escala industrial.

Uma das etapas mais importantes e demoradas no projeto de um processo de separação usando CPC é a seleção do sistema de solvente bifásico mais eficaz devido ao grande número de combinações possíveis. Após uma revisão da literatura, os seguintes sistemas de solventes bifásicos foram selecionados: o sistema quaternário Arizona N convencional (volumes iguais de heptano, acetato de etila, metanol e água), e duas versões modificadas contendo solventes verdes alternativos (versão 1: substituição de heptano por limoneno, e metanol por etanol; versão 2: substituindo apenas metanol por etanol) e quatro sistemas de solventes binários (água + acetato de etila, água + 1-butanol, água + 1-octanol e heptano + metanol). Os coeficientes de partição dos solutos modelo foram medidos a (298,2 ± 0,5) K pelo método do *shake-flask*, geralmente apresentando resultados consistentes em comparação com dados existentes na literatura.

Em relação à modelagem termodinâmica dos dados de partição, já foram sugeridos o UNIFAC (*Universal Quasichemical Functional Group Activity Coefficient*), o NRTL-SAC (*Nonrandom Two-Liquid Segment Activity Coefficient Model*) e o modelo de solvatação de Abraham. O modelo UNIFAC é totalmente preditivo, porém não apresentou resultados consistentes em comparação com os outros métodos. Além disso, quercetina e rutina não puderam ser representadas pelos grupos disponíveis. Por outro lado, o modelo de solvatação de Abraham foi capaz de descrever os dados dos coeficientes de partição, com um RMSD global (raiz quadrada do erro-médio) de 0,57. Os resultados obtidos pelo modelo de solvatação de Abraham são comparáveis ao modelo NRTL-SAC, com um RMSD global de 0,3. O primeiro tem a vantagem de sua simplicidade matemática, representada por duas relações lineares de energia livre, mas é mais difícil de obter os parâmetros para sistemas bifásicos compostos por mais de dois solventes.

Palavras-chave: Coeficiente de partição, solubilidade, UNIFAC, NRTL-SAC, Abraham, cromatografia.

Table of Contents

List of Symbols and Acronyms

List of symbols

K_D°	Partition constant
$a_{X,n}$	Ratio of activities of compound X in phase n
K	Distribution constant
D	Distribution ratio
P	Partition coefficient
α	*Separation factor*
S	Solubility
E	Liquid excess molar refraction at 20 ºC
S	Dipolarity/polarizability
A	Overall or effective hydrogen bond acidity
B	Overall or effective hydrogen bond basicity
V	McGowan characteristic volume

List of acronyms

CPC	Centrifugal Partition Chromatography
LLE	Liquid-liquid extraction
CCC	Countercurrent chromatography
UNIFAC	Universal Quasichemical Functional Group Activity Coefficient
NRTL-SAC	Nonrandom Two-Liquid Segment Activity Coefficient Model
RMSD	Root-mean-square deviation

List of Figures

List of Tables

Chapter 1. Introduction

1.1. Scope and objectives

Centrifugal Partition Chromatography (CPC) is a technology that allows the fractionation and purification of complex mixtures that can also be scaled up to industrial levels. The separation principle of this equipment is based on the partition of a sample between two liquid phases of a solvent mixture, presenting some advantages over the more traditional solid-liquid separation methods: the possibility of one-step purification from crude extracts, lower solvent usage, no irreversible adsorption on solid matrices, total recovery of the injected sample, low risk of sample denaturation.

The most important and time consuming task for operating CPC is the selection of the optimum biphasic solvent system. Thus, a preliminary screening of the mixture of solvents that provides a suitable distribution ratio of the target solutes is needed. The optimum solvent system can be obtained experimentally. However, due to the wide range of possible solvent systems, it is advisable to complement the experimental study with thermodynamic modelling tools, such as the UNIFAC (Universal Quasichemical Functional Group Activity Coefficient) group contribution model, the NRTL-SAC (Nonrandom Two-Liquid Segment Activity Coefficient Model) model or the Abraham solvation model.

The main objective of this work is the experimental measurement and thermodynamic modelling of the partitions coefficients, using the UNIFAC and Abraham solvation models, of several model compounds of bioactive phytochemicals (quercetin, vanillin, rutin, ferulic acid and hesperetin acid), to support the design of fractionation processes of natural extracts and oils derived from biomass, using centrifugal partition chromatography.

1.2. Contents and outline

Chapter 2 starts with an introduction about liquid-liquid extraction, counter current chromatography, their advantages, applications and the selection of the solvent system, describing the measurement of the distribution of a solute of interest in a solvent system, an important concept in liquid-liquid extraction processes. Chapter 2 also includes a description of the chemical, biological and pharmaceutical properties, as well as possible applications, of the five phenolic compounds selected in this work. Finally, it also contains a literature review of their solubility and partition coefficient data which is key information to plan the experimental and modelling work.

The experimental results obtained in this work are described and critically compared to literature data in Chapter 3. Chapter 4 presents the thermodynamic modelling, by applying the UNIFAC and Abraham Solvation models, and using as database the partition and solubility data from the literature review and experimentally obtained here. Lastly, Chapter 5 presents the main conclusions and future studies related to this work.

Chapter 2. Partitioning of phenolic compounds

2.1. Liquid-liquid extraction

Liquid-liquid extraction (LLE), also referred as partitioning [1], is a separation process that consists of transferring a solute present in a particular solvent by bringing it into contact with another immiscible or partially miscible solvent, creating a two phase solvent system [1,2]. Often, one of the phases of the solvent system is an aqueous mixture and the other is an organic mixture [1].

Traditional methods used to purify molecules can require several steps to achieve purification, like dialysis, ionic and affinity chromatography, among others. Nonetheless, liquid-liquid extraction is extensively used to purify and isolate biomolecules due to its simplicity, flexibility, low cost and facility to large-scale production compared to traditional methods [2,3]. Thus, LLE can be applied in the separation and purification of several organic molecules, including polyphenols [3]. LLE is also useful for the isolation of heat-sensitive compounds and for diluted feed streams [3]. LLE include two steps: mixing followed by phase separation. The first step promotes contact and solute transfer between two immiscible phases. Usually, mixing increases the solute transfer between the phases, however, it can also difficult phase separation if emulsion formation occurs. Hence, when selecting a solvent system, it is important to consider both steps of LLE extraction [1].

2.2. Countercurrent chromatography

Countercurrent chromatography (CCC) is an all-liquid separation technique based on the partition of a solute or several solutes of interest between two immiscible phases, a stationary phase and a mobile phase [4,5]. This liquid-liquid separation has some advantages in comparison with traditional solid-liquid separation, such as the absence of irreversible adsorption of the solute in a

solid phase, low solvent consumption, therefore, less expensive [4], versatility, and good sample recovery without loss of its chemical properties and structure [6].

Countercurrent chromatography has the potential to become an environmental friendly approach due to the reduced solvent usage in comparison with traditional separation methods. Furthermore, considering the wide range of possible biphasic solvent systems used in a CCC, it is possible and desirable to select and study solvents with less environmental impact [7]. On a large scale, countercurrent chromatography can be used in continuous processes with the recycling of the biphasic systems, being more sustainable in comparison with batch processes [8].

2.3. Partition constant, distribution constant and distribution ratio

During the solute transfer between the immiscible phases, as stated by Berk (2018), "equilibrium is reached when the chemical potential of the extractable solute is the same in the two phases" [1]. Mathematically, for a compound A that is present in the same chemical form in the two immiscible phases, during the equilibrium, the ratio in both phases is obtained by the partition constant K_D°, [9]:

$$K_D^{\circ}(A) = \frac{a_{A,1}}{a_{A,2}} \quad (1)$$

Where $a_{A,1}$and $a_{A,2}$ are the ratio of activities of compound *A* in phase 1 and phase 2 when equilibrium is reached [9]. In a similar concept, the partition coefficient P is defined as "the concentration of the compound in the upper phase divided by the concentration of the compound in the lower phase" [10].

The ratio of activity of a compound depends on temperature and pressure. However, "distribution isotherms are generally linear over a reasonable concentration range" [9], so it is possible to consider concentrations instead of activity ratios. When the linear distribution isotherm is taken into consideration, the partition constants are referred to as the distribution constant K [9]. Considering counter current chromatography, the expression of the distribution constant K is the concentration of the solute in the stationary phase divided by the concentration of the solute in the mobile phase [9]. Thus, Equation 1 is rewritten as follows:

$$K(A) = \frac{C_{A,S}}{C_{A,M}} \quad (2)$$

In a CCC solvent system, the ideal value of K for a particular solute is close to 1. Large K values can contribute to an improper sample broadening and excessive run times, and small K values can lead to a loss of peak resolution [10].

For solutes that can be found in more than one chemical form in one or both of the liquid phases and it is considerably affecting the solvent-solute system, the distribution constant is referred to as the distribution ratio, D. The distribution ratio is described as the total concentration of all chemical forms of a solute in the stationary phase divided by the total concentration in the mobile phase [5].

In normal phase elution in counter current chromatography, where the upper phase is the mobile phase and the lower phase is the stationary phase, the relationship between the partition coefficient P and the distribution constant K is defined by the relation $K = 1/P$, according to Friesen and Pauli [10].

In reverse phase elution, where the upper phase is stationary and the lower phase is mobile, the relationship between the partition coefficient P and the distribution constant K can be defined as $K = P$ [10].

2.4. Selection of biphasic solvent system

Choosing the most appropriate two-phase liquid solvent system to promote the separation of a sample by countercurrent chromatography is a complex task. According to Costa and Leitão (2010), there are some basic requirements for a biphasic solvent system in a CCC separation: the settling time, that is, the necessary time for the separation of two phases after being mixed, should be ideally shorter than 30 seconds or rarely exceeding 60 seconds; the separation factor α, where $\alpha = K_2/K_1$, $K_2 > K_1$, needs to be greater than 1.5 [6]. It is also important that the solvent system allows satisfactory retention of the stationary phase [11]. Theoretically, the ideal value for P in countercurrent chromatography is, like K, close to 1 [10]. Experimentally, K should be between 0.4 and 2.5 [5,10], since "small K values usually result in a loss of peak resolution, while large K values tend to produce excessive sample band broadening" [10]. It is also important that the two-

phase solvent systems produce equal or similar volumes for both phases in order to avoid wasting solvents and consequent environmental issues [6].

Woźniak and Garrard (2015) analyzed over 4000 papers related to CCC or CPC, classifying 2322 isocratic solvent systems according to the water-immiscible solvent in the system [12]. Evaluating the data, they conclude that water, ethyl acetate methanol and hexane were the most commonly used solvents, that compose the HEMWat system [12]. This quaternary solvent system features a range of polarities based on the proportion of each component in the HEMWat system, being very versatile in the separation of several chemical compounds. Figure 1 illustrates the range of solvent proportions (named from -7 to +8) proposed by Friesen and Pauli (2007) for the HEMWat system [10].

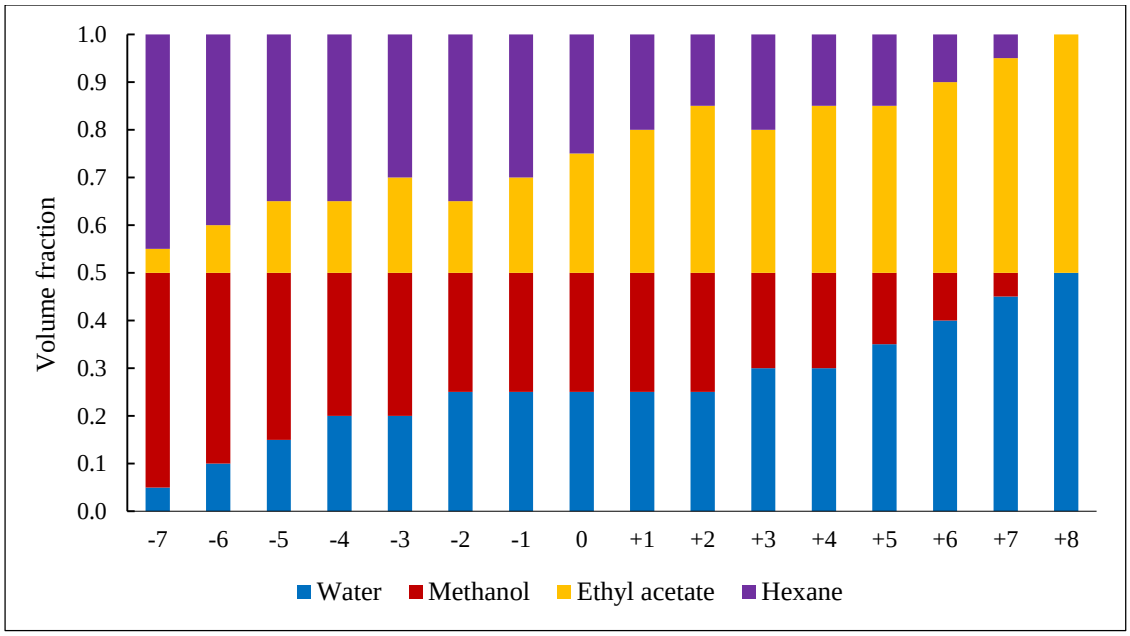

Figure 1. HEMWat global volume fractions of hexane, ethyl acetate, methanol and water [10].

Hopmann *et al.* (2010) determined, by gas chromatography, the composition of the upper and lower phases of the HEMWat solvent system [13]. The distribution of the solvents presented in Figure 2 indicates that water is mostly located in the lower phase, and hexane in the upper phase. Ethyl acetate and methanol are then partitioned between both phases, depending on the global phase composition [14].

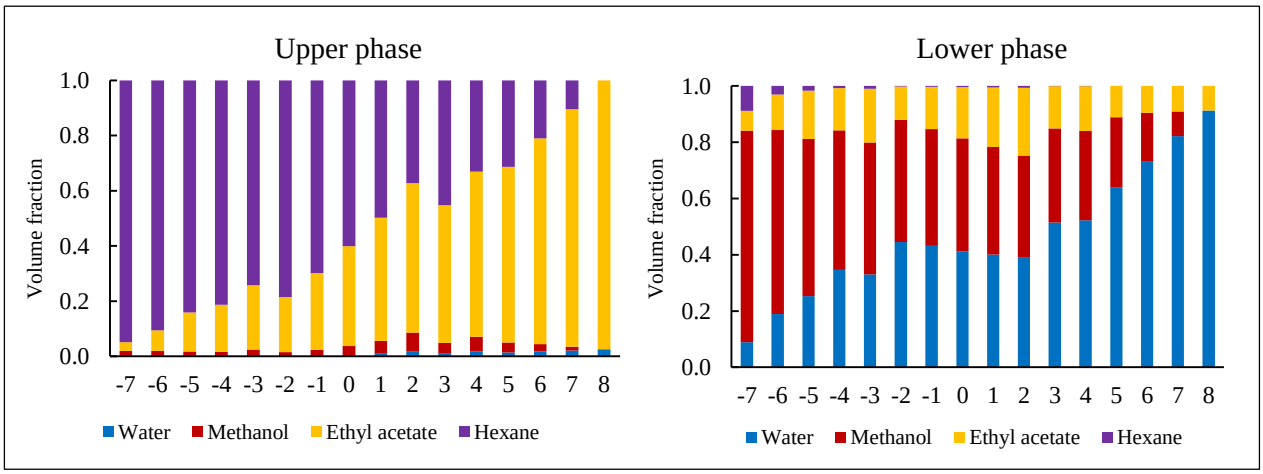

Figure 2. HEMWat volume fractions of hexane, ethyl acetate, methanol and water in the upper and lower phases [13].

The Arizona scale system can also represent an array of different proportions and polarities by setting a range of compositions for the heptane/ethyl acetate/methanol/water mixture, obeying the condition that the methanol/water volume ratio must be the same as heptane/ethyl acetate volume ratio [7]. There are 23 compositions named from A to Z, with the exclusion of the letters E, I and O [7], as illustrated in Figure 6. The name Arizona comes from the fact that the array of compositions starts in A, the most polar system with 0/1/0/1 v/v heptane/ethyl acetate/methanol/water, ending in Z, the least polar system with 1/0/1/0 v/v heptane/acetate/methanol/water [7].

The distributions of the solvents in the Arizona system vary due to their polarity and density differences. The general tendencies are: as in the HEMWat solvent system, the alkane is concentrated in the upper phase and water in the lower phases; going from A to Z, the upper phase increases in alkane and decreases in ethyl acetate content; similarly, the lower phase increases in methanol and decreases in water content [14]. In summary, the polarity of the system increases with the presence of water and ethyl acetate in the lower and upper phases, respectively.

Similar to the ARIZONA solvent system, the n-heptane/ethyl acetate/1-butanol/methanol/water system covers a considerable range in polarity and hydrophobicity regularly from heptane/methanol to 1-butanol/water, as shown in Figure 3 [15]. The addition of butanol in the ARIZONA system enhances the separation of moderately to highly polar solutes, since their low solubility in apolar organic solvents can be troubling [5,15].

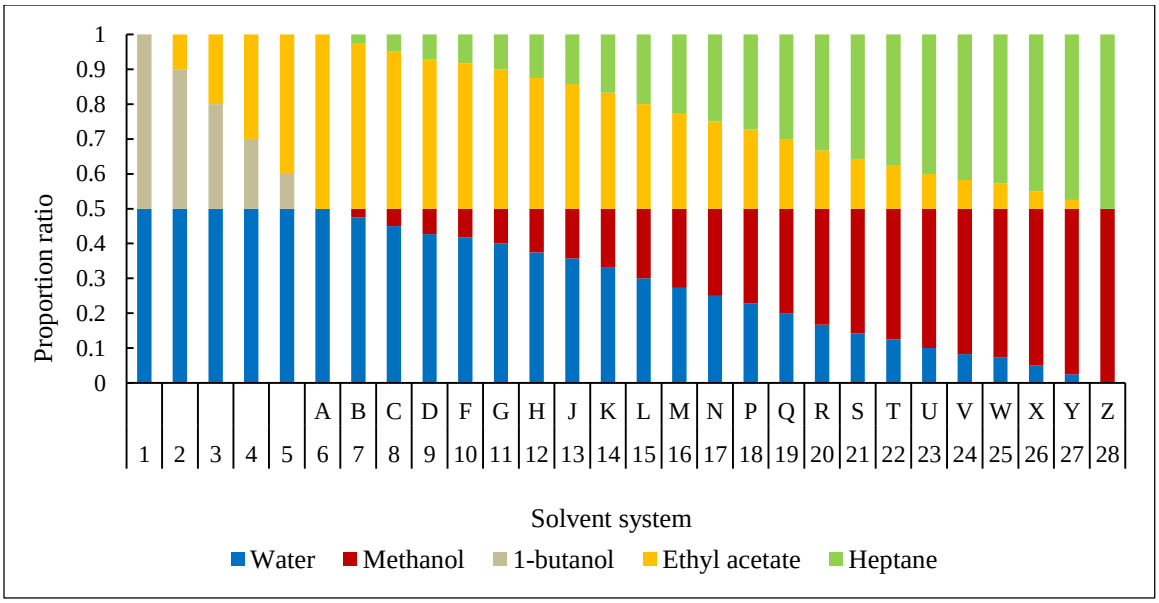

Figure 3. Global volume fraction of heptane, ethyl acetate, 1-butanol, methanol and water solvent systems [16].

As can be seen in Figure 3, only 1 to 5 solvent systems contain 1-butanol. Therefore, the compositions of systems 6 to 28 will be equal to the ARIZONA solvent systems [13]. The distribution of all components in the upper and lower phases are described in Figure 4 and Figure 5, respectively.

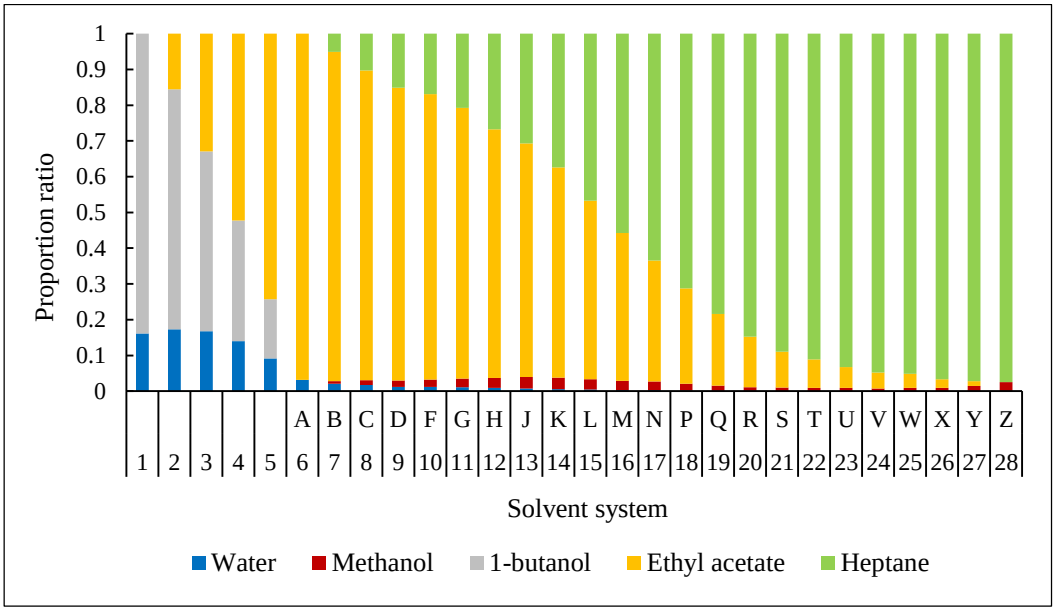

Figure 4. Upper phase volume fractions of heptane, ethyl acetate, 1-butanol, methanol and water solvent systems [16].

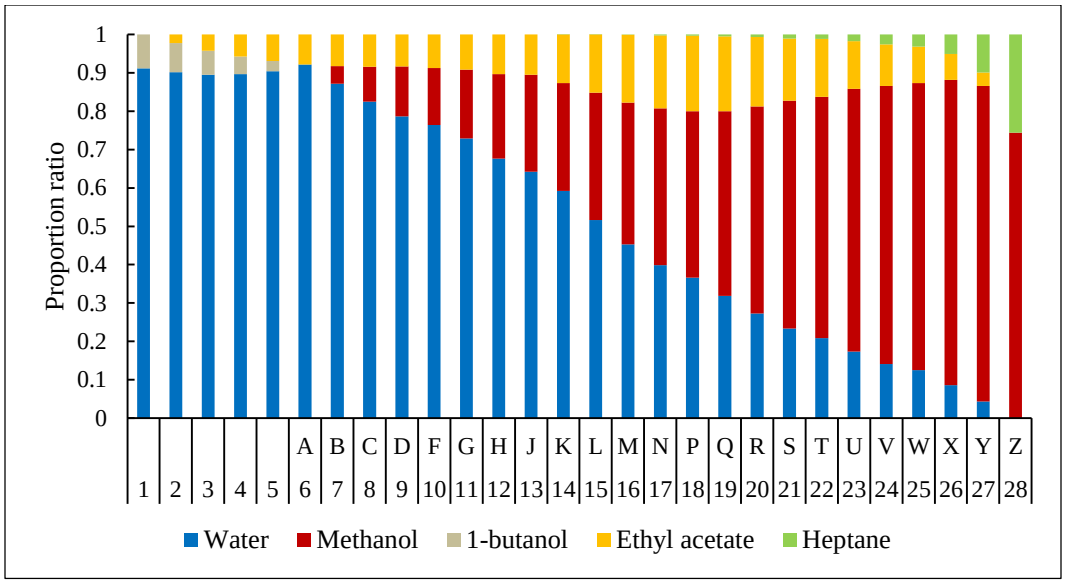

Figure 5. Lower phase volume fractions of heptane, ethyl acetate, 1-butanol, methanol and water solvent systems [16].

Regarding possible variations in the presented systems, the HEMWat system can include solvent substitutions; for example, replacing methanol with ethanol, propanol or butanol, or replacing ethyl acetate with chloroform ($CHCl_3$) or dichloromethane (CH_2Cl_2) [6].

Berthod *et al.* (2005) studied the effect of replacing heptane by other alkanes (pentane, hexane, cyclohexane, isooctane) in the composition of the upper and lower phases in the Arizona system [14]. Since the alkane is mostly present in the upper phase, changing the alkane used in the biphasic solvent system does not significantly change the lower phase composition [14]. Nonetheless, the upper phase composition alters significantly when the alkane is substituted. As the alkane chain increases, the methanol molar fraction decreases in the upper phase composition [14]. As a rule, as the alkane becomes larger, the solubility in water consequently decreases, with the only exception being cyclohexane. The study indicates that substituting heptane with a lighter alkane, the upper phase will be slightly more polar, and vice-versa [14]. Analysing the polarity, the heptane-ethyl acetate ratio is responsible for increasing the upper layer polarity and the methanol-water ratio predominantly decreases the lower layer polarity [16]. The replacement of heptane by cyclohexane, that presents higher density, will difficult the stationary phase retention, since the difference between upper and lower phase densities will be lower [14].

2.5. Green solvent selection

A green solvent can be described by two assessments: environmental, health and safety (EHS), and energy demand [17]. It is possible to calculate the net cumulative energy demand (CED) by obtaining the energy demand required during the production of a solvent and the options available to recover part of that energy, such as incineration, purification and recycling [17]. Thus, a green solvent presents possibilities of CED reduction during production and recovery, also having beneficial environmental, health and safety impact.

The HEMWat solvent system enables solvent substitutions, such as replacing ethyl acetate with chloroform or DCM [6]. According to the World Health Organization, chloroform and dichloromethane are likely carcinogenic to humans, hence, this substitution is not advisable [17].

According to Byrne et al. (2016), Slater and Saveski of Rowan University developed an approach to evaluate and compare environmental, health and safety parameters of several solvents, showing a good differentiation between hydrocarbon solvents [17]. This approach concludes that cyclohexane is a greener substitution to n-heptane, and n-pentane and n-hexane are more hazardous solvents than n-heptane [17].

Besides energy demand and EHS, the origin of each solvent should be considered, since "renewable feedstocks will need to be embraced to secure the sustainability of the chemical industry" [17]. In this regard, ethanol can be considered a green solvent, since its origin is bio-based and, also, ethyl acetate can be obtained from renewable sources, although biomass is not a primary feedstock [17]. Therefore, ethanol can be a green substitute for methanol, since both components have similar polarities and ethanol has much lower toxicity [18].

The substitution of methanol for ethanol, a "renewable polar solvent less toxic than methanol" and heptane for limonene, a natural renewable solvent, in the Arizona solvent system, was already proposed by Faure *et al.* (2014), demonstrating that liquid chromatography can present an environmental friendly approach in separation and purification processes [7]. The volume fractions of each component in this Green Arizona system is described in Figure 6.

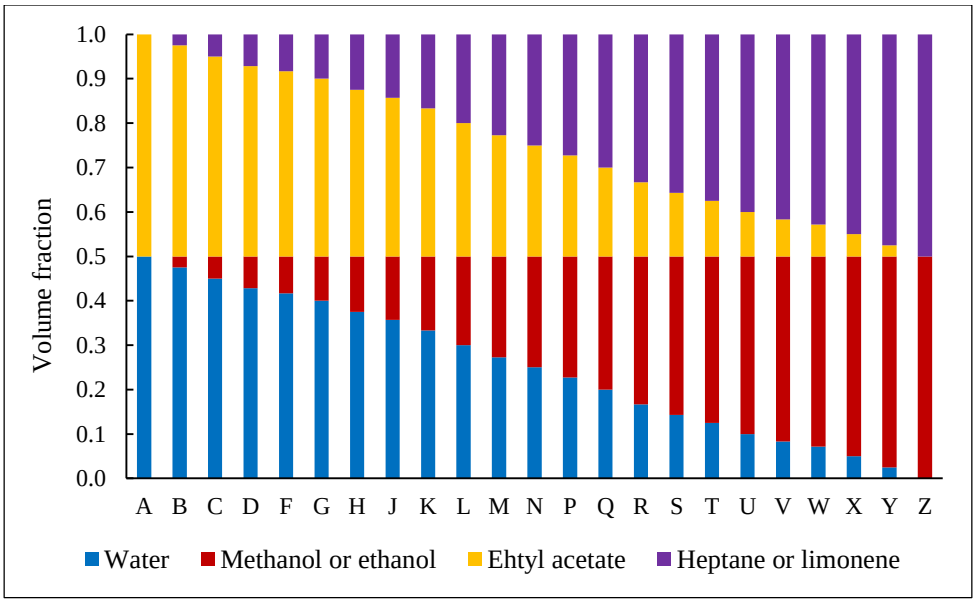

Figure 6. Green Arizona global volume fractions of heptane or limonene, ethyl acetate, methanol or ethanol and water [7].

The substitution resulted in significant differences in the distribution of ethanol and ethyl acetate between the upper and lower phases [7]. The upper phases contain two to six times more ethanol in comparison with the traditional Arizona system; as result, more limonene was present in the lower phase. Due to the high density of limonene, the difference between upper and lower phase densities was smaller, which can negatively impact the retention of the stationary phase during CCC separation [7]. However, a low-density difference can be beneficial in hydrostatic CCC columns, also called centrifugal partition chromatography (CPC) [18]. The difference of density between phases is important to the mobile phase driving pressure, and working with high centrifugal fields can retain the stationary phase and promote a satisfactory partition using the green limonene/ethyl acetate/ethanol/water system [18].

The possibility to scale-up countercurrent chromatography and the consequential increase of solvent usage enhances the importance of developing approaches to use green alternatives and minimize the necessary solvent volume [16]. Garrard *et al.* (2007) analyzed two strategies to minimize solvent usage: recycling solvents by removing the solute through evaporation and formulating the two-phase solvent system separately, since the mobile phase volume needed is higher than the stationary phase [16]. To re-use the recycled mobile phase, the solvent was condensed and a calculated volume of pure solvent was added to the previous condensed solvent in order to return to the original composition before re-use in next runs [16]. According to the

authors [14], for solvent systems containing heptane, ethyl acetate, methanol and butanol, “no difference was detected between the runs initially performed with fresh solvent and subsequent runs performed with the recycled solvent reconstituted into the mobile phase” [16].

2.6. Polyphenols: rutin, ferulic acid, quercetin, vanillin and hesperetin analysis of separation solvent systems

Polyphenols are secondary products from natural metabolic processes in plants that presents “interesting health benefits alongside other diverse (biotechnological) applications (e.g., colorants, nutraceuticals)” [3]. These compounds can be obtained by plant extraction, but can also be produced using fermentation processes, a more sustainable technique [3]. The chemical structures of the phenolic compounds studied in this work are represented in Figure 7, together with vanillin.

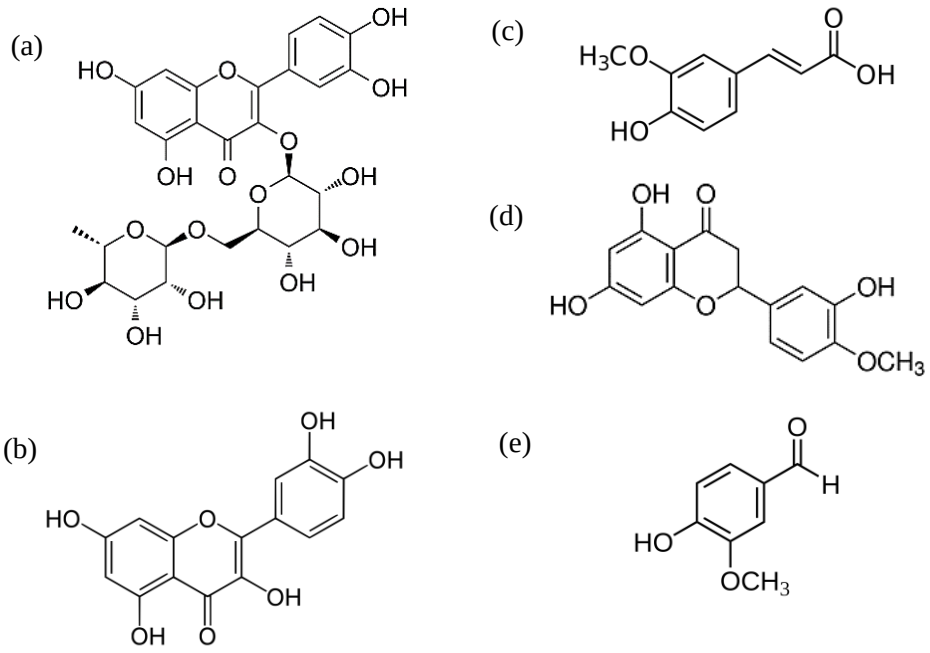

Figure 7. Chemical structures of (a) rutin, (b) quercetin, (c) ferulic acid, (d) hesperetin and (e) vanillin.

Rutin (quercetin-3-O-rutinoside), also known as vitamin P, is one of the most common quercetin glycosides found in a variety of fruits and vegetables, especially grapes and buckwheat, synthesized as a protectant against UV radiation [19]. Rutin is a flavonol glycoside, that is, the phenolic group is linked to a sugar moiety, which increases the solubility of the molecule in water

[19]. Rutin has a medical application in the treatment of vascular disorders associated to capillary permeability and fragility [19].

Ferulic acid (4-hydroxy-3-methoxycinnamic acid) is a phenolic acid with strong antioxidant properties [20]. It can be widely found in the cell walls of grains, fruits, and vegetables, usually conjugated with polysaccharides and other compounds, having the capacity of interrupting UV-peroxidative chain reactions in phospholipid membranes [20].

Quercetin is a flavonoid present in a variable of natural sources, such as onions, garlic, and apples, also being present in tea and wine [21]. Like other flavonoids, quercetin has antioxidant proprieties, providing "some protection against osteoporosis, pulmonary and cardiovascular diseases, and chronic degenerative diseases, including cancers"[21].

Vanillin is a phenolic aldehyde present in the pod or bean of *Vanilla orchid* [22]. Vanillin can be used as a food flavouring, fragrance agent in food and cosmetics, and as a food preservative because of its antioxidant and antimicrobial proprieties due to presence of a phenolic group [22,23].

Hesperetin is a natural flavanone, a derivative of hesperidin [24]. This compound can be found in plants and fruits, especially in citrus fruits, and presents several pharmaceutical functions [24–26]. Hesperetin has also been shown neuroprotective effects protect neurons and brain cells against oxidative damage [24,27].

Friesen and Pauli (2005) partitioned the powdered methanolic extract of *Valeria officinallis* roots employing several compositions of HEMWat solvent system [10]. Searching for partition coefficients between 0.4 and 2.5, with optimal value of 1, the separation of several chemical components was achieved [10]. For quercetin and ferulic acid, the optimum value was obtained with a hexane/ethyl acetate/methanol/water mixture with proportions of 5/5/5/5 v/v and 3/7/4/6, respectively [10]. Vanillin achieved optimum partition with a HEMWat system with proportions of 3/7/5/5 v/v [10].

According to Woźniak and Garrard (2015) that analysed 2322 isocratic solvent systems, alkane-based solvent systems are the most commonly used for the separation of flavonoids, whereas this is true for almost all of the solutes in the database, such as terpenoids, alkaloids and polyphenols [12]. Alkane-based solvent systems are also the most common used solvent systems used to separate phenolic acids and derivatives, followed by ether-based, chlorinated-based and water-immiscible alcohols solvent systems [12].

In comparison with free flavonoids, "the presence of a sugar moiety on the flavonoid structure enhances its polarity" [6]. Thus, the HEMWat solvent system must be used in higher polarity ratios. Better partitioning of the solute of interest can be achieved with the addition of butanol in the hexane/ethyl acetate/methanol/water system [6]. Even so, for high polarity glycosylated flavonoid solutes, it may be necessary to use more polar solvent systems than HEMWat [6]. In this case, Costa and Leitão (2010) suggested employing ethyl acetate/butanol/water or chloroform/methanol/water solvent system [6].

2.7. Experimental solubility data available in the literature

Before starting the experimental work and the thermodynamic modelling of the partition coefficients, it is necessary to collect experimental information in the literature about the solubility of the model compounds in water and organic solvents as well as their partition in different solvent systems. That experimental database is reported in this section for the solubility data and in Section 2.8 for the partition data.

2.7.1. Vanillin

Shakeel *et al.* (2015) measured the solubility of vanillin in ten solvents, in the temperature range of 298 K to 318 K, using an isothermal method [22]. Noubigh *et al.* (2008) also measured the solubility of vanillin in water at 298.15 K [28]. Figure 8 presents a summary of the information collected at 298.2 K.

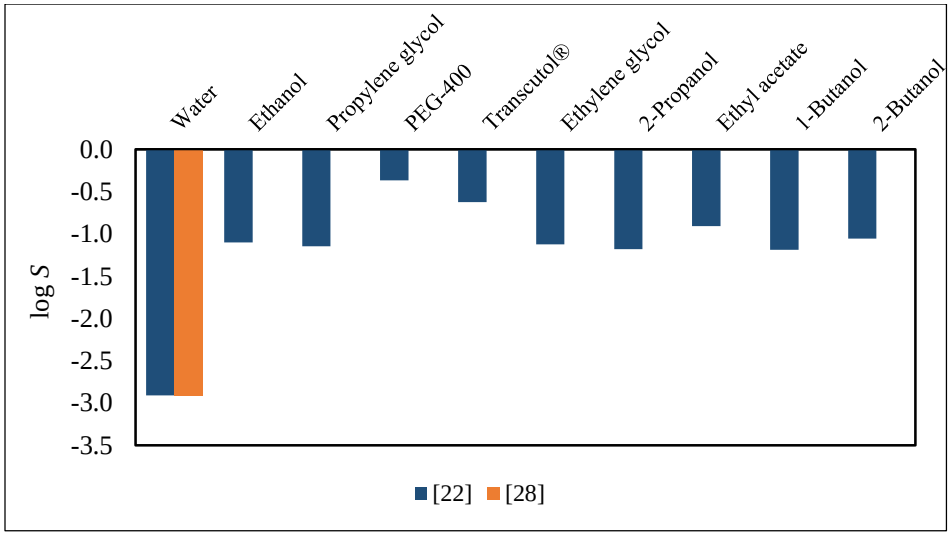

Figure 8. Solubility (in mole fraction) of vanillin in several solvents at 298.2 K [22,28].

According to Figure 8, the experimental measurements of the solubility of vanillin in water are in agreement. Comparing the solubility of this solute among the analysed solvents, the solubility is lower in water in comparison with other solvents. Vanillin presents higher solubility in polyethylene glycol-400 (PEG-400), also showing similar solubility value in Transcutol®. Vanillin shows comparable solubility in alcohols, ethyl acetate, propylene glycol and ethylene glycol.

2.7.2. Rutin

Peng *et al.* (2009) measured the solubility of rutin in several ethanol + water mixtures, between 273 K and 323.15 K [29]. The mole fraction solubility of rutin in pure ethanol, $S = 4.27 \cdot 10^{-3}$, was also determined by Peng *et al.* (2009) [29]. The results obtained are shown in **Erro! Fonte de referência não encontrada.** for several ethanol + water mixtures.

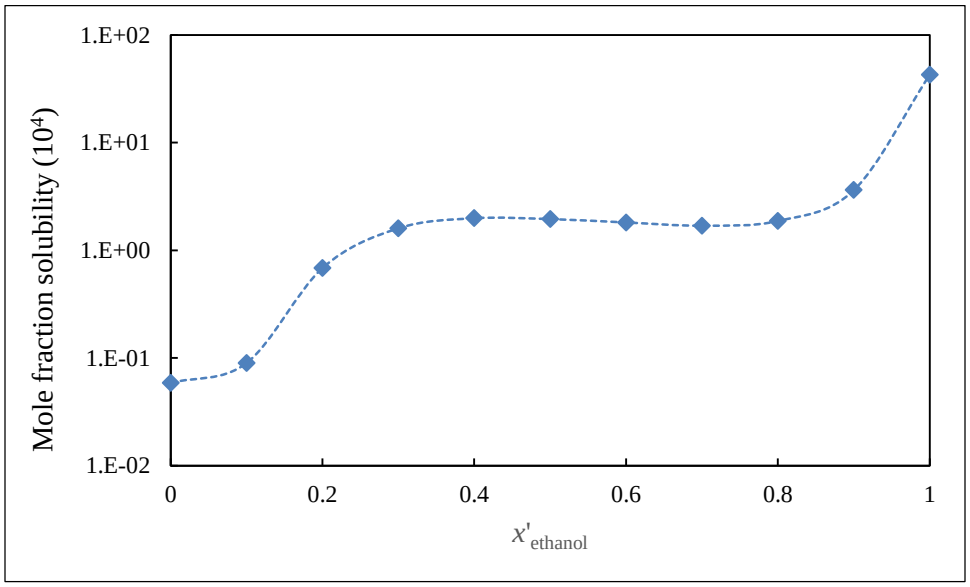

Figure 9. Solubility of rutin in several ethanol + water mixtures at 293.15 K [29].

The solubility of rutin increases from $x'_{ethanol}$ = 0.0 to 0.4 and slowly decreases until $x'_{ethanol}$ = 0.7. The solubility increases rapidly $x'_{ethanol}$ = 0.8 to $x'_{ethanol}$ = 1, being much higher in pure ethanol than in any other mixture [29]. Moreover, the solubility increases with temperature [29].

Zi *et al.* (2007) measured the solubilities of rutin at different temperatures in water and 7 organic solvents [30]. **Erro! Fonte de referência não encontrada.** presents the results obtained at 298.2 K.

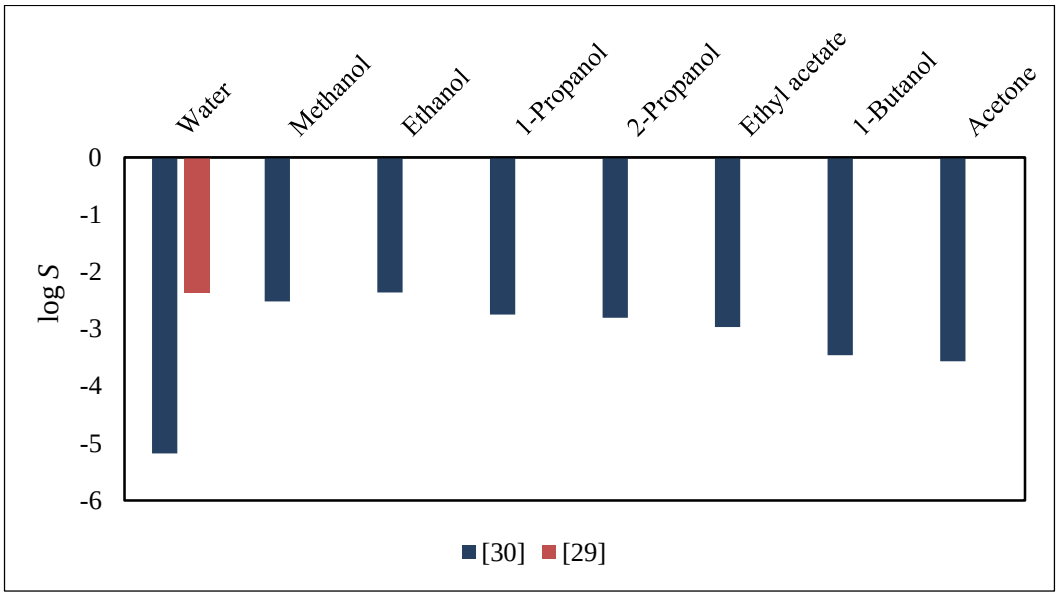

Figure 10. Solubility (in mole fraction) of rutin in several solvents at 298.15 K [29][30].

For all solvents presented in this study [30], except ethyl acetate, the solubility of rutin increases with temperature. Comparing the solubility of rutin among the analysed solvents, the solubility is higher in ethanol followed by methanol. Rutin also shows comparable solubility values in 1-propanol, 2-propanol and ethyl acetate. The solubility in water is significantly lower.

2.7.3. Ferulic acid

Shakeel *et al.* (2017) measured the solubility of ferulic acid, between 298.2 K and 318.2 K, using a static equilibrium method, in several solvents: water, ethanol, methanol, ethylene glycol, propylene glycol, isopropyl alcohol, 1-butanol, 2-butanol, ethyl acetate, dimethyl sulfoxide (DMSO), polyethylene glycol-400 (PEG-400) and Transcutol® [31]. Similarly, Vilas-Boas *et al.* (2020) measured the solubility of ferulic acid in methanol, ethanol, 1-propanol, 2-propanol, 2-butanone, ethyl acetate, acetonitrile and water at 298.2 K and 313.2 K by the isothermal shake-flask method [32]. Zhou *et al.* (2007) measured the solubility of ferulic acid in methanol, ethanol, dichloromethane, chloroform, methyl acetate, ethyl acetate and butyl acetate using the synthetic

method [33]. This method consists of detecting the temperatures of crystal disappearance of the solute of interest [34]. The collected information is compared in Figure 11.

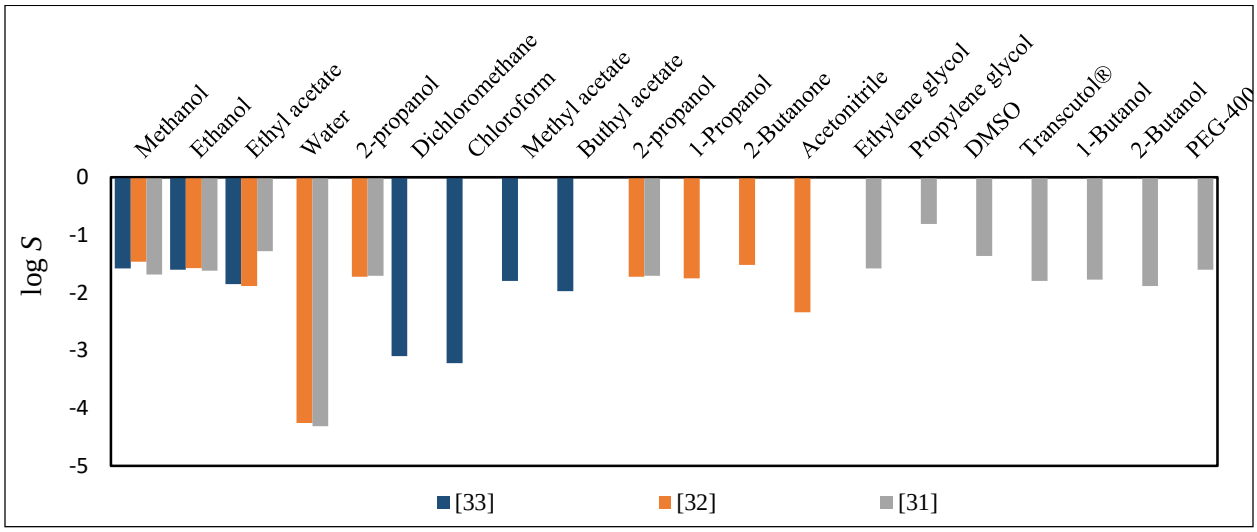

Figure 11. Solubility (in mole fraction) of ferulic acid in several solvents at 298.2 K [31–33].

For all solvents analysed by Shakeel *et al.* (2017), the solubility of ferulic acid increases with temperature [31]. Comparing the solubility of ferulic acid among the solvents in this study [31], the solubility assumes considerably higher values in propylene glycol. Ferulic acid has its lowest value of solubility in water in comparison with other solvents in this study [31].

According to Vilas-Boas *et al.* (2020), the solubility is higher at 313.2 K than at 298.2 K for all solvents analysed [32]. The solubility of ferulic acid is also "larger in alcohols and 2-butanone and considerably smaller in water" [32].

The experimental measurements of the solubility of ferulic acid are in agreement for all five solvents, with the exception of ethyl acetate. Haq *et al.* (2017) measured the solubility of ferulic acid in various isopropanol + water mixtures at different temperatures using a static equilibrium method [35], as presented in Figure 12.

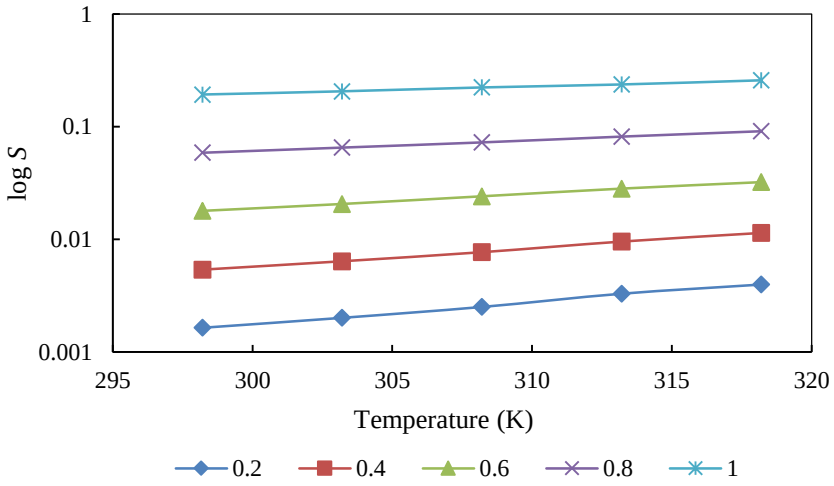

Figure 12. Solubility data (in mole fraction) of ferulic acid in isopropanol/water mixtures (mass fraction of 0.2, 0.4 0.6, 0.8 and 1 of isopropanol) as a function of temperature [35].

For all mixtures, the solubility also increases with temperature. The solubility of ferulic acid assumes higher values with the increase of isopropanol in the mixture, having its maximum solubility in pure isopropanol [35].

2.7.4. Hesperetin

Ferreira and Pinho (2012) experimentally obtained, for several solvents, the solubility of hesperetin between 298.15 K and 313.15 K [36]. In a similar study, Liu and Chen (2008) measured the solubility of hesperetin in methanol, ethanol, 1-butanol, acetone and water between 288.2 K and 323.2 K [37]. Finally, Zhang *et al.* (2017) determined the solubility of hesperetin in water and 1-octanol between 298.15 K and 318.15 K [38]. Figure 13 presents an overview of the literature data.

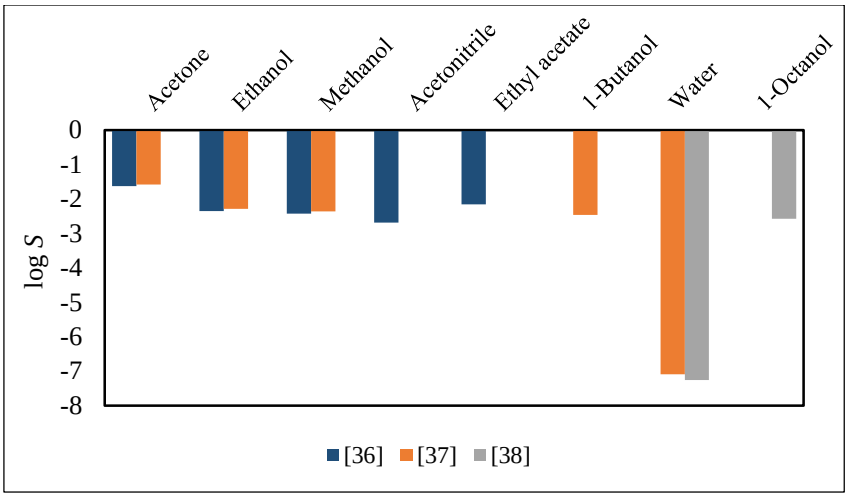

Figure 13. Solubility of hesperetin (in mole fraction) in several solvents at 298.2 K [36–38].

According to Figure 13, the experimental measurements of solubility from different studies are in agreement in acetone, ethanol, methanol and water. The solubility of hesperetin assumes lower values in water and the highest value in acetone. The solubility also increases with temperature in all studied solvents [36,37].

Ferreira *et al.* (2013) measured the solubility of hesperetin in the mixed solvents water + ethanol, water + methanol and water + acetone at 298.15 K [39]. The solubilities were obtained using the analytical isothermal shake-flask method [39]. Figure 14 summarizes the solubility data described in this study [39].

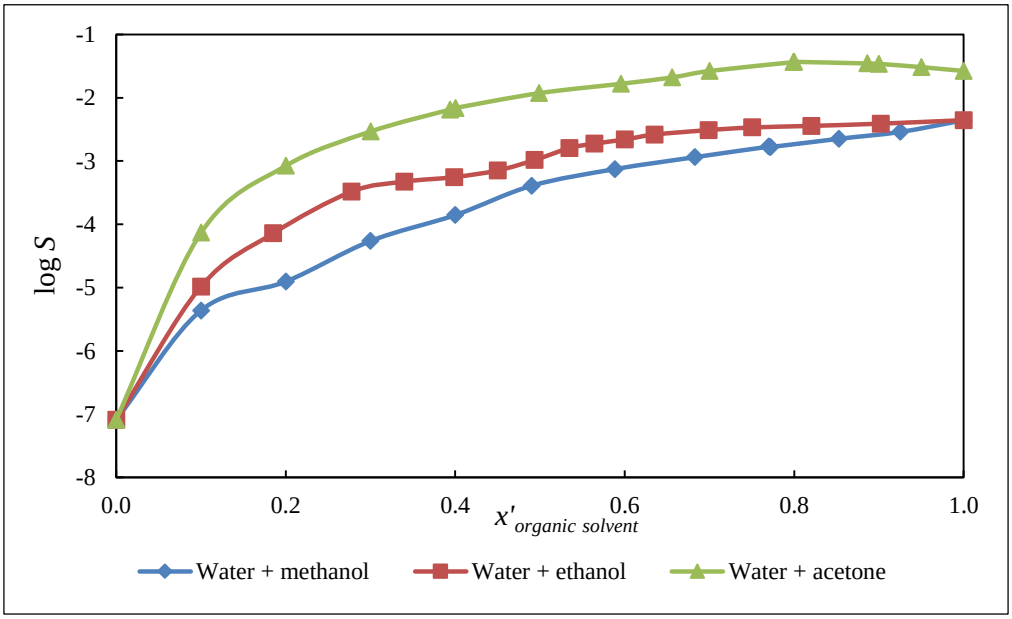

Figure 14. Solubility of hesperetin (in log value) in mixed solvents [39].

The solubility of hesperetin in water + methanol slowly increases from $x'_{methanol}$ = 0.0 to $x'_{methanol}$ = 0.4. After that, mole fraction solubility rapidly increases. For water + ethanol, the solubility also slowly increases from $x'_{ethanol}$ = 0.0 to $x'_{ethanol}$ = 0.4, although this increase is higher than the observed in water + methanol. For this mixed solvent, "the sharpest increase in the solubility curve is obtained between $x'_{ethanol}$ = 0.50 and $x'_{ethanol}$ = 0.60" [39]. The water + acetone system present the highest solubilities values than the other two mixed solvents, with the maximum mole fraction solubility being around $x'_{acetone}$ = 0.8, decreasing from $x'_{acetone}$ = 0.8 to $x'_{acetone}$ = 1.0.

2.7.5. Quercetin

Dománska *et al.* (2018) measured the solubility of quercetin in 1-propanol, 1-butanol, ethyl acetate and butyl acetate by the synthetic method in the temperature range between 295.1 K to 300.4 K [34]. Althans *et al.* (2014) and Razmara *et al.* (2010) also experimentally obtained the solubility of quercetin at 298 K in ethanol and methanol, respectively [40,41]. Figure 15 summarizes the solubility data described in these studies at 298 K.

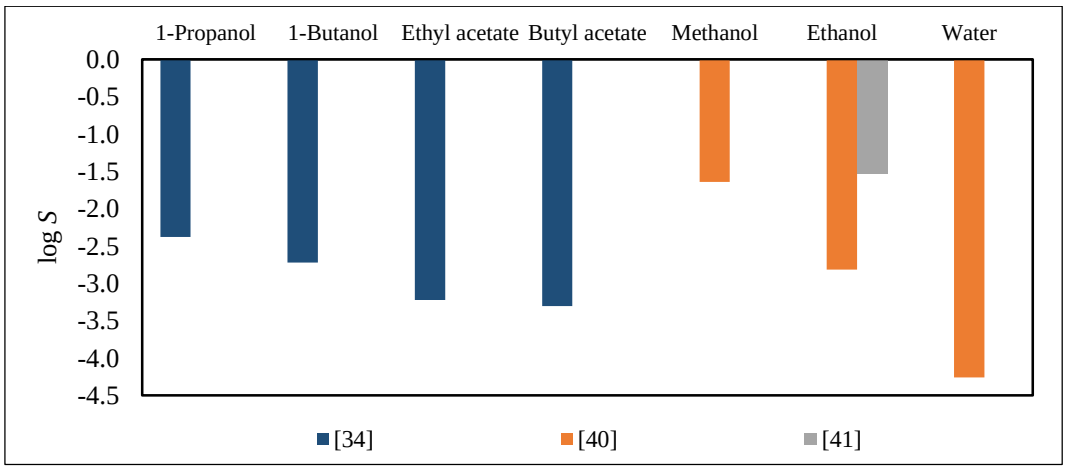

Figure 15. Mole fraction solubility of quercetin at 298 K [34,40,41].

According to Figure 15, quercetin presents greater solubility in ethanol and methanol, and lower solubility in the ester solvents and water.

Razmara *et al.* (2010) also measured the solubility of quercetin dehydrate in several water + methanol and water + ethanol mixtures by the isothermal shake-flask method, between 292.8 K and 298.6 K [40]. The results obtained at 298.6 K are represented in Figure 16.

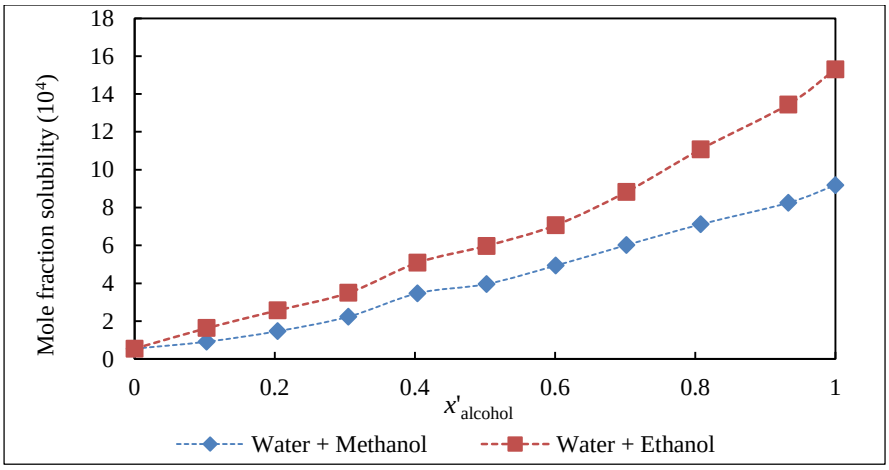

Figure 16. Solubility of quercetin in several water + methanol and water + ethanol mixtures at 298.6 K [40].

The solubility of quercetin constantly increases from $x'_{alcohol}$ = 0.0 to $x'_{alcohol}$ = 1.0. As in agreement with other study [40], the solubility in ethanol is greater than the solubility in methanol.

2.8. Experimental partition data available in the literature

2.8.1. Arizona and HEMWat solvent systems

Marlot (2018) determined the partition coefficient of several solutes, highly diluted to avoid saturation, in 21 solvent systems using the shake-flask method [42]. The analysis was made with an aliquot with 2 mL of each phase diluted by a factor 10 in methanol and it was tested with a considerable number of solutes, among them, ferulic acid, vanillin and quercetin [42].

Berthod *et al.* (2009) and Lu *et al.* (2009) determined several partition coefficients in the Arizona solvent system. Lu *et al.* (2009) obtained different partition coefficients by screening the natural extracts of *Piper longum L.* and *Polygonum cuspidatum* by CCC and HPLC [43]. Berthod *et al.* (2009) determined the partition coefficients by countercurrent chromatography of several solutes, including vanillin, quercetin and ferulic acid [44]. Figure 17 summarizes the partition coefficient data of vanillin, quercetin and ferulic acid in several Arizona systems.

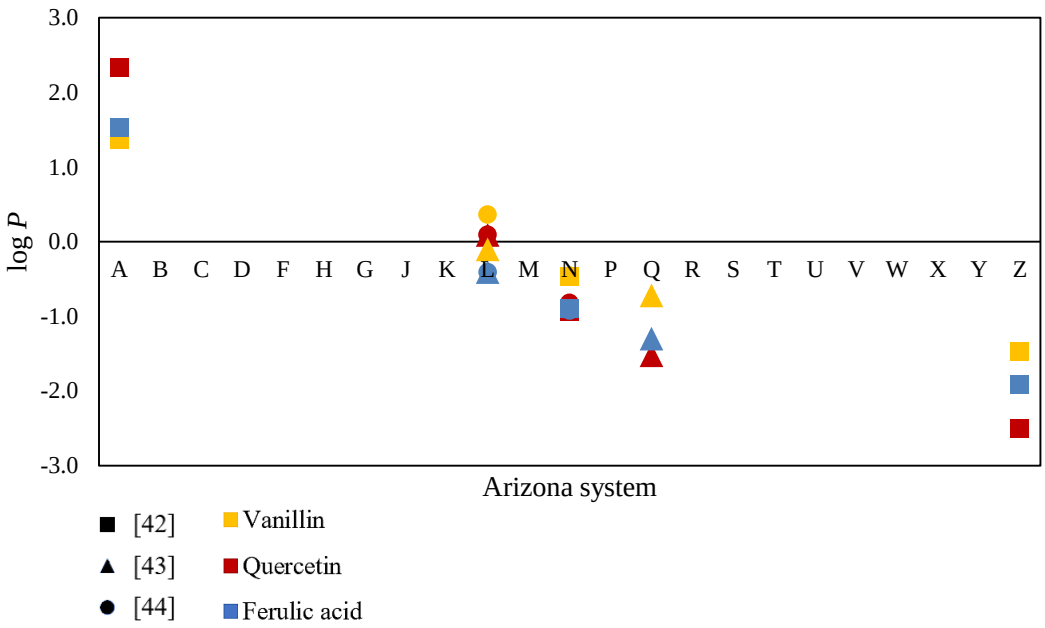

Figure 17. Partition coefficient data of vanillin, quercetin and ferulic acid in several Arizona systems [42–44].

Friesen and Pauli (2007) measured the partition coefficient for several solutes in the HEMWat solvent system, among other biphasic systems [10]. The compound was added to a given solvent system in a test tube, shaken two minutes and centrifuged [10]. The UV analysis made the quantification of each solute from the bottom phase obtained after centrifugation [10]. Friesen *et al.* (2014) also measured the partition of vanillin in five HEMWat solvent systems by high-speed countercurrent chromatography (HSCCC) [45]. Berthod and Faure (2015) obtained the partition coefficient in HEMWat +3 for ferulic acid and quercetin using CCC [46]. Yang *et al.* (2016) also measured by CCC the partition of quercetin in HEMWat 0, +3 and +5 [47]. Figure 18 summarizes the data for three solutes: ferulic acid, vanillin and quercetin.

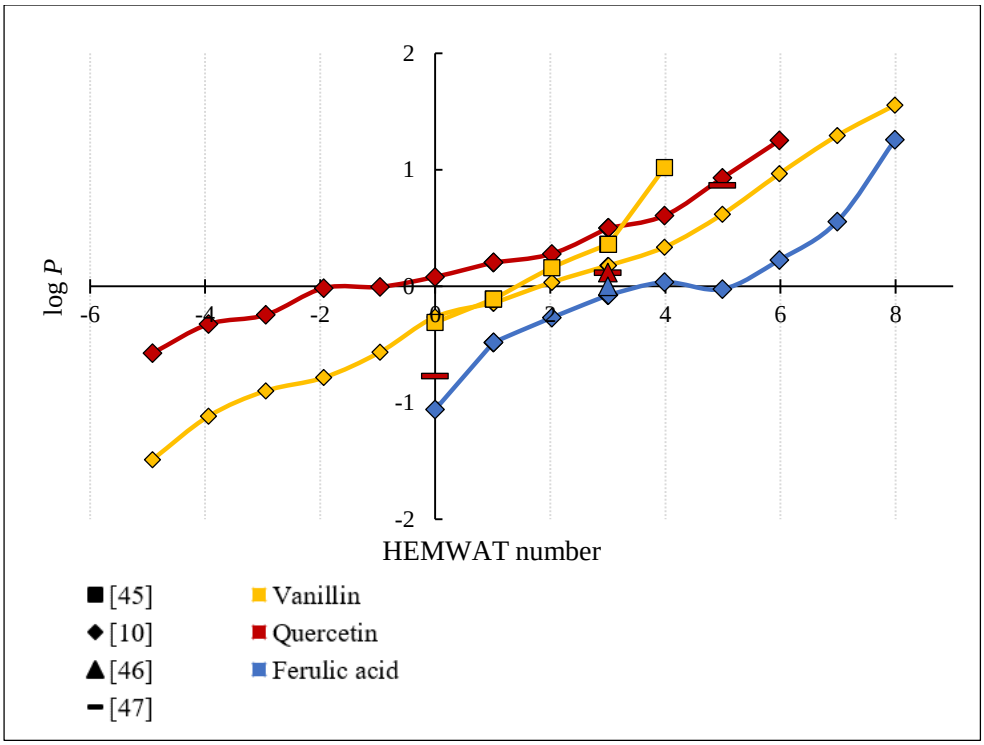

Figure 18. Partition coefficient data of vanillin, quercetin and ferulic acid in several HEMWat systems [10,45–47].

The optimal value for partition coefficient is 1, or 0 when analysing its log value [10]. Hence, the ideal HEMWat number described in Figure 18 for quercetin and ferulic acid is 0 and +4, respectively [10]. For vanillin, the optimum solvent systems was +2 [10]. The partition data for quercetin was obtained by Yang *et al.* (2016) and Friesen and Friesen and Pauli (2007) using the shake flask and CCC, respectively [10,47]. Observing Figure 18, the results for quercetin diverges for HEMWat 0, and are similar for HEMWat +3 and +5. In addition, considering the solvent system HEMWat number +4, the partition coefficient value of vanillin measured by Friesen and Pauli (2007) and Friesen *et al.* (2014) slightly diverge.

2.8.2. Other solvent systems

2.8.2.1. Vanillin

Marlot (2018) determined the partition coefficient of several solutes including ferulic acid, vanillin and quercetin studied in this work [42]. Lamprecht and Blochberger (2009) separated vanillin from ice cream and yoghurt samples, obtaining partition coefficients by reversed-phase HPLC [48]. Kaygorodov *et al.* (2010), also obtained the partition coefficient of vanillin for several

solvent systems [23,49]. Likewise, Marlot (2018) determined the partition coefficient of vanillin for several solvent systems. The results of these studies are summarized in Table 1.

Table 1. Partition coefficient data for vanillin in some solvent systems.

Solvent system	T (K)	Log P_{exp}	Reference
Heptane/acetonitrile	298.15	-1.90	[42]
Methyl terc-butyl ether/water	298.15	0.99	
Methyl isobutyl ketone/water	Room temperature[a]	1.47	[48]
1-pentanol/water		1.35	
Chloroform/water		1.31	
Dichloromethane/water		1.28	
1,2-Dichloroethane/water		1.12	
Diethyl ether/water		0.80	
Toluene/water		0.61	
Benzene/water		0.57	
Hexane/water		-0.70	[49]
Chloroform/water		1.42	
Butyl acetate/water		1.45	

[a]Temperature of the experiment not specified.

2.8.2.2. Rutin

Pedriali *et al.* (2008) obtained the partition coefficient of rutin and some derivatives by dissolving the components in 1-octanol/water (pH 7.4) system and analysing the fractions using UV spectroscopy [50]. Das and Kalita (2014) determined the partition coefficient by the shake flask method [51]. The results of both studies are summarized in Table 2.

Table 2. Partition coefficient data for rutin in some solvent systems.

Solvent System	Partition coefficient	Reference
octanol/water	0.85 ± 0.07	[50]
octanol/water	0.85 ± 0.02	[51]
octanol/phosphate buffer (pH 6.8)	0.80± 0.02	[51]
octanol/acetate buffer (pH 4.5)	0.91 ± 0.02	[51]

Chapter 3. Experimental work

3.1. Materials

All the solutes and solvents used to obtain the experimental partition coefficients are listed in Table 3. They were stored at room temperature and kept in desiccators. The water used in the experiments had ultrapure quality (resistivity of 18.2 MΩ·cm, free particles ≥ 0.22 μm and total organic carbon < 5 $\mu g \cdot dm^{-3}$). The water content of ferulic acid, hesperetin and quercetin was determined by the Karl-Fisher method.

Table 3. List of organic compounds used in this work.

Compound	Mass purity (%)[a]	CAS number	Source	Water Content (%)
***trans*-ferulic acid**	≥ 99.9	537-73-5	Alfa Aesar	0.1
(S)-hesperetin	≥ 98.0	520-33-2	Cayman Chemicals	1.1
quercetin	≥ 95.0	117-39-5	Sigma-Aldrich	6.2
vanillin	≥ 99.0	121-33-5	Sigma-Aldrich	-
methanol	≥ 99.9	67-56-1	J. T. Baker	-
ethanol	≥ 99.8	64-17-5	Fisher	-
1-butanol	≥ 99,5	71-36-3	Sigma-aldrich	-
1-octanol	≥ 99.0	111-87-5	Carlo Erba	-
ethyl acetate	≥ 99.9	141-78-6	Carlo Erba	-
heptane	≥ 99.0	142-82-5	Honeywell	-
(R)-(+)-limonene	≥ 97.0	5989-27-5	Sigma-Aldrich	-
rutin	≥ 94.0	207671-50-9	Sigma-Aldrich	6

3.2. Methodology

In this work, the partition coefficients were measured by the analytical shake-flask method, which has been conventionally applied to obtain the distribution of biomolecules in biphasic systems [52,53]. First, the solid solute was dissolved in the solvent in which it has higher solubility, considering the maximum concentration of 0.5 mg of solute/mL of solvent system as infinite dilution. After, the solution was transferred to a glass tube along with predefined volumetric

proportions of the other solvents present in the selected solvent system, resulting in a total volume of 14 ml. Then, the tubes were manually flipped upside down for five minutes, followed by continuous agitation at 300 rpm in an Eppendorf ThermoMixer C equipment, at (298.2 ± 0.5) K for at least 5 hours. After the agitation, the system is left to settle for approximately 15 hours. The experiments were performed twice, in two independent tubes. For each tube, at least two samples of the upper and lower phases were taken and the concentration of the solutes in each phase was measured by UV-Vis spectrophotometry. For each calibration curve, at least seven standard solutions were prepared ($R^2 > 0.999$), and the absorbance read at the wavelength of maximum absorbance (Table 4). Depending on the composition of the upper or lower phases, and to avoid immiscibility issues, the calibration curves were prepared using pure ethanol or a mixture of water and ethanol (wt 50:50%) as solvents.

Table 4. Wavelengths of maximum absorbance found for each solute, when the solvent was pure ethanol or a mixture of ethanol + water.

Solute	Maximum wavelength in pure ethanol	Maximum wavelength in ethanol-water (w_t 50:50%)
ferulic acid	322	313
hesperetin	289	289
quercetin	370	374
vanillin	308	313
rutin	258	258

3.3. Results and discussion

The experimental partition coefficients of *trans*-ferulic acid, (*S*)-hesperetin, quercetin, rutin and vanillin in 4 binary biphasic systems (methanol + heptane, 1-butanol + water, 1-octanol + water, ethyl acetate + water), and in the three quaternary biphasic systems "Arizona N" (water:methanol:ethyl acetate:heptane, in the volumetric proportion 1:1:1:1) and two modified versions containing alternative green solvents (version 1: replacing heptane by limonene, and methanol by ethanol; version 2: replacing only methanol by ethanol) are presented and compared to literature data in Table 5. An indication of the experimental methods used for the measurements is also provided in the footnotes of the table.

Table 5. Experimental log *P* values measured in this work and collected from literature.

Solvent system	ferulic acid	hesperetin	quercetin	vanillin	rutin
Methanol + 1-heptane	-2.12 ± 0.01[a] -1.91[b] [54]	-2.20 ± 0.01[a]	-3.01± 0.03[a] -2.51[b] [54]	-1.48 ± 0.02[a] -1.47[b] [54]	-3.18 ± 0.01[a]
1-butanol + water	1.48 ± 0.02[a] 1.45[b] [54]	2.27 ± 0.01[a]	2.42 ± 0.10[a] 1.29[b] [54]	1.19 ± 0.01[a] 1.15[b] [54] 1.15[c] [48]	0.73 ± 0.01[a]
1-octanol + water	1.32 ± 0.01[a] 1.08[b] [54] 1.77[b] [56] 1.69 ± 0.01[c] [56]	2.39[a] ± 0.01[a] 2.15 ± 0.43[b] [55] 2.26 ± 0.09[b] [57] 1.50[b] [59]	2.21 ± 0.10[a] 0.99[b] [54] 1.82 ± 0.32[b] [58] 2.45[b] [56] 3.15 ± 0.01[c] [58]	1.29 ± 0.01[a] 0.92[b] [54] 1.19 ± 0.06[b] [41] 1.17[b] [60] 1.22[d] [61]	-0.01 ± 0.01[a]
ethyl acetate + water	1.57 ± 0.02[a] 1.53[b] [54]	2.34 ± 0.01[a]	2.33 ± 0.01[a] 2.33[b] [54]	1.43 ± 0.02[a] 1.36[b] [54] 1.34[c] [48]	-0.76 ± 0.01[a]
ARIZONA N	-0.96 ± 0.02[a] -0.92[b] [54] -0.90[e] [44]	-0.40[a] ± 0.04[a]	-0.89 ± 0.05[a] -0.93[b] [54] -0.82[e] [44] -0.89[e] [62]	-0.56 ± 0.03[a] -0.47[b] [54] -0.46[e] [44]	-2.53 ± 0.01[a]
Green ARIZONA N (v1)	-0.42 ± 0.01[a]	0.15 ± 0.01[a]	-0.23 ± 0.02[a]	-0.18 ± 0.03[a]	-0.30 ± 0.02[a]
Green ARIZONA N (v2)	-0.83 ± 0.01[a]	-0.45 ± 0.02[a]	-0.70 ± 0.01[a]	-0.40 ±0.01[a]	-2.27 ± 0.01[a]

[a]Experimentally obtained in this work. Temperature and pressure standard uncertainties are u(T) = 0.50 K and ur(p) = 0.05, respectively. Standard deviations are placed after the minus-plus sign.
[b]Obtained by the shake-flask method.
[c]Experimentally determined by reversed phase HPLC analysis
[d]Experimentally measured by a slow equilibrium method through an octanol drop apparatus.
[e]Obtained by counter-current chromatography (CCC).

The consistency of the measurements can be confirmed by the low standard deviations presented in Table 5, resulting in 2.8% average coefficient of variation. The experimental values are in good agreement with the log *P* values found in literature, except for quercetin in 1-butanol/water and 1-heptane/methanol.

Large deviations are noticed between the partition coefficients reported the literature for 1-octanol + water systems, although the data obtained in this work is comparable to most of the data

available in literature for vanillin [41,60,61] and hesperetin [55,57,59]. For the 1-octanol/water systems, the pre-saturation of both phases is highly recommended [63] to avoid the formation of emulsions during the shake-flask method. Therefore, this procedure was adopted in this work.

In the second version of the Green Arizona system, the replacement of methanol by ethanol increases the polarity in the upper phase, causing an increment in partition coefficient value when compared with the conventional Arizona [18]. In the first version of the green Arizona system, the additional substitution of heptane by limonene also increases the value of the partition coefficient for all five solutes, since the replacement of heptane for limonene rises the polarity of the solvent system in comparison with the conventional Arizona [18].

For 1-butanol + water, 1-octanol/water and ethyl acetate/water, the solutes present more affinity towards the organic upper phases, with the exception of rutin. For the conventional Arizona N system, the partition coefficients are much closer to zero, presenting intermediate polarity between the most polar ethyl acetate + water biphasic system (ARIZONA A) and the least polar heptane + methanol (ARIZONA Z). To the best of our knowledge, the log *P* of all the studied solutes in the green versions of the Arizona N system were reported here for the first time.

Chapter 4. Thermodynamic modelling

4.1. The UNIFAC model

The Universal Quasichemical Functional Group Activity Coefficient model, or UNIFAC, is a solution-of-groups model that allows the estimation of activity coefficients in nonelectrolyte liquid mixtures in non-critical conditions [64,65]. The activity coefficients of the UNIFAC model are mainly specified by the geometry (shape and size) of the groups, the structure of the molecule, that is, the distribution of the chemical groups and subgroups, and the energy related to the interaction between the groups [64]. The UNIFAC equations are presented in Appendix A.

According to Galanakis *et al.* (2013), the UNIFAC model has some limitations that need to be considered: "the pressure should be less than 5 bar, the temperature should be less than 150°C, and calculations are only applicable to condensable non electrolytes, while components should not contain more than ten functional groups" [64].

In the present study, the calculation of the distribution constant K_i can be obtained from the thermodynamic equilibrium according to Equation 3 [13]:

$$K_i = \frac{x_i^S}{x_i^M} = \frac{\gamma_i^M}{\gamma_i^S};\ x_i \to 0 \tag{3}$$

Where K_i *is* the partition coefficient of solute i, x_i is the molar fraction, S represents the stationary phase and M the mobile phase. To obtain the distribution constant based on molar concentrations, it is necessary to multiply the value of K_i by the ratio of the weighted sum of molar volumes (v_0) of all pure compounds j in the mobile and stationary phases [13]:

$$K_i^m = \frac{C_i^S}{C_i^M} = K_i . \frac{\sum x_j^M v_{0j}}{\sum x_j^S v_{0j}} \tag{4}$$

Since $x_i \to 0$, the molar volume of the solute can be neglected.

According to Friesen and Pauli (2007), the partition coefficient P_i for a solute i is defined by the following equation [10] :

$$P_i = \frac{C_i^{upper}}{C_i^{lower}} \tag{5}$$

In order to calculate the partition coefficients considering CPC, the aqueous lower phase was defined as the stationary phase. In reverse phase, when the upper phase is the stationary phase, the relation is defined by $K_i^m = P_i$ [10].

To obtain the activity coefficients in the solvent system, the free software XLUNIFAC version 1.0 was used (Copyright © 2000 Preben Randhol and Hilde K Engelien). The UNIFAC model predicts activity coefficient values only requiring information about the phase composition of the solvent system and the chemical structure of the molecules of solutes and solvents [66]. The upper and lower phase compositions used in the UNIFAC predictions for methanol/ethyl acetate/heptane/water/butanol, ethanol/ethyl acetate/limonene/water/butanol and HEMWat solvent systems can be found in Appendix B.

However, this model is unable to obtain activity coefficient values for complex molecules, because either the functional group is not defined or the additivity rule becomes invalid [66]. For this reason, it was only possible to describe the ferulic acid, hesperetin and vanillin using the UNIFAC model. The hesperetin prediction data was not contemplated in this discussion due to the lack of experimental data for comparison, with the exception of the partition coefficient in Arizona N obtained in this work. The hesperetin prediction data using the UNIFAC model can be found in Appendix A.

As previously mentioned, Friesen and Pauli (2007) measured the partition coefficient for several solutes in the HEMWat solvent system, including ferulic acid and vanillin [10]. Figure 19 compares those experimental results with the predictions of the UNIFAC method.

It is possible to conclude that the UNIFAC model qualitatively predicts the same increasing trend. The behaviour of the curves are similar, since as the lower phase becomes more polar (richer in water), the organic compounds tend to travel towards the upper phase, resulting in higher *P* [10]. Moreover, for ferulic acid very good quantitative predictions are obtained in comparison with vanillin results.

Marlot (2018) measured the partition coefficient for several solutes in the Arizona solvent system, including vanillin and ferulic acid [42]. Also, Lu *et al.* (2009) obtained the partition coefficients of ferulic acid and vanillin for solvent system numbers 15 and 19, equivalent to Arizona L and Q, respectively [43]. Berthod *et al.* (2009) determined by CCC the partition coefficients of ferulic acid and vanillin for solvent system Arizona L and N [44]. Figure 20 and Figure 21 compares the results predicted by the UNIFAC method and the experimental data available in the literature, and measured in this work.

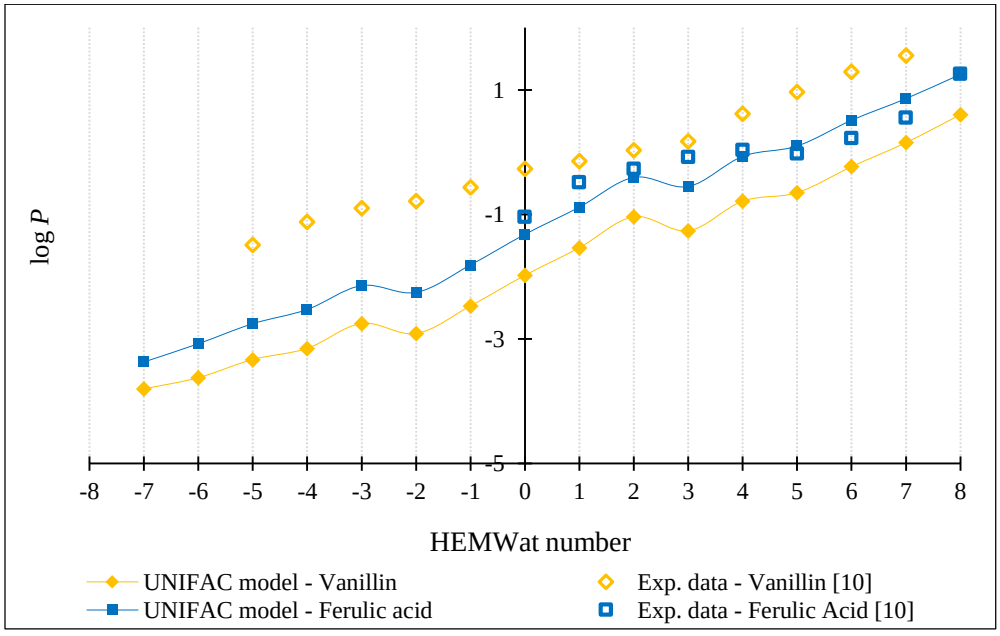

Figure 19. Partition coefficients of ferulic acid and vanillin in different HEMWat solvent systems: experimental data [10].

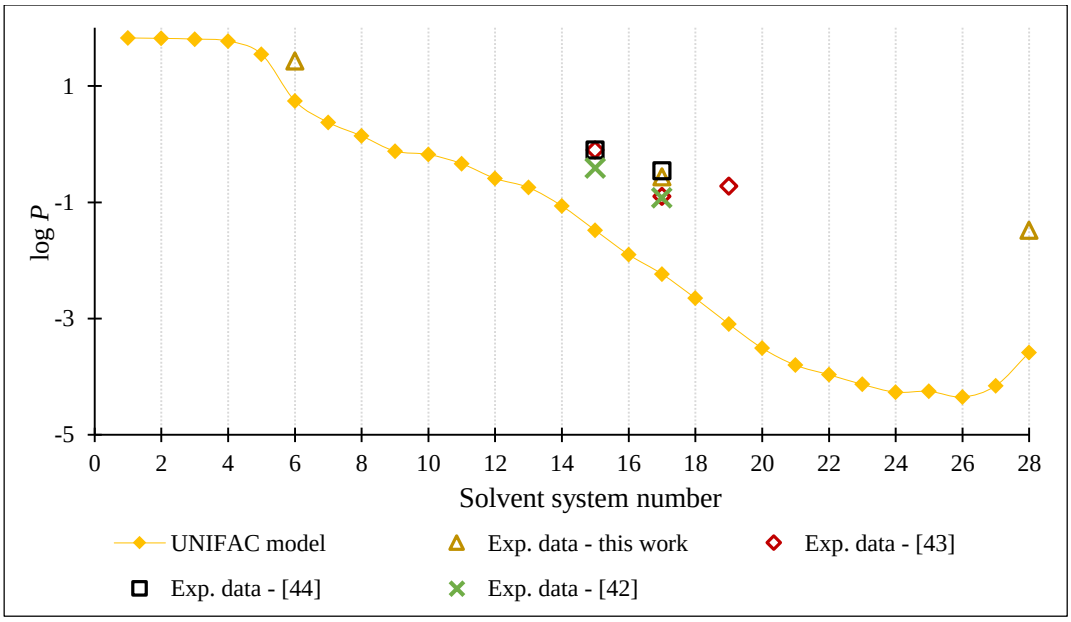

Figure 20. Partition coefficients of vanillin in different n-heptane/ethyl acetate/1-butanol/methanol/water solvent systems: experimental data obtained in this work and from the literature [42–44].

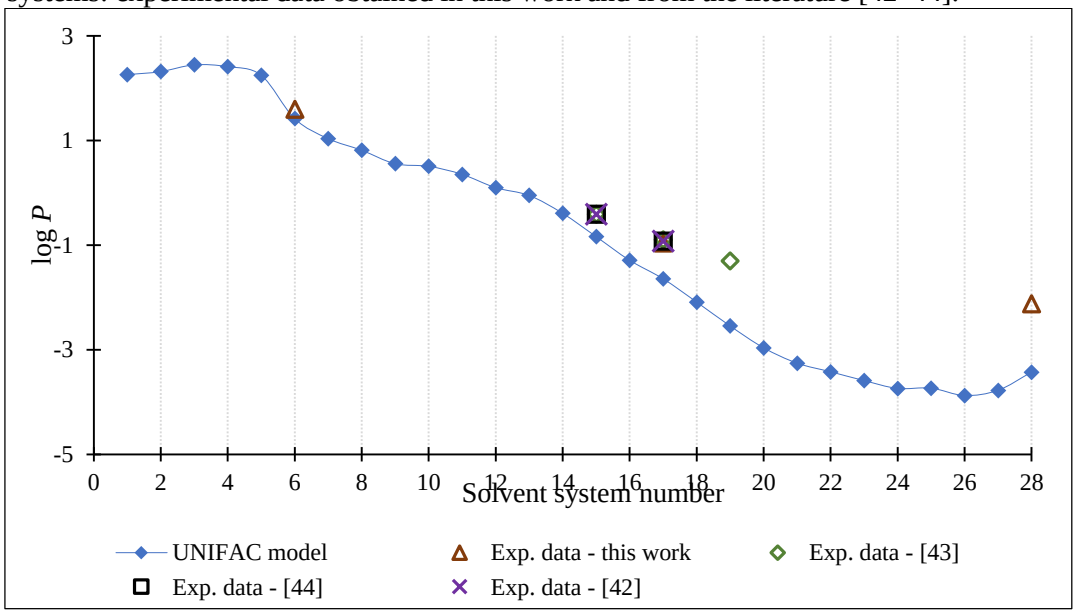

Figure 21. Partition coefficients of ferulic acid in different n-heptane/ethyl acetate/1-butanol/methanol/water solvent systems: experimental data obtained in this work and from the literature [42–44].

The UNIFAC predictions of the partition coefficients are in qualitative agreement with the experimental results, for the Arizona solvent system. In the case of ferulic acid, lower deviations were obtained. In this case, the organic solutes concentration rises in the lower phase as it becomes less aqueous, corresponding to an increase in the solvent system number.

Table 6 presents the comparison between UNIFAC predictions and experimental log *P*, for the three solutes partitioning in the version 1 of the Green Arizona N solvent system.

Table 6. Comparison between UNIFAC prediction and experimental log *P* values for vanillin, ferulic acid and hesperetin for Arizona N and green Arizona N (version 1).

Solute	Conventional Arizona N		Green Arizona N (version 1)	
	UNIFAC prediction	Experimental	UNIFAC prediction	Experimental
Vanillin	-2.23	-0.56	-4.55	-0.18
Ferulic acid	-1.64	-0.96	-2.64	-0.42
Hesperetin	-4.55	-0.40	-6.07	0.15

The UNIFAC predictions in limonene/ethyl acetate/1-butanol/ethanol/water solvent systems present similar behaviour in comparison with the conventional Arizona system. As the solvent system number increases, the polarity decreases, also decreasing the concentration of the solute in the upper phase, lowering the partition coefficient value. The substitution of heptane by limonene and methanol by ethanol increases the polarity of the upper phase, also increasing the partition coefficient. However, this behavior is not noticed in the predictions made by the UNIFAC model, since the predictions of the partition coefficients in the green Arizona are generally lower in comparison with the predictions in the conventional Arizona solvent system. Besides, according to Table 6, the partition values differs significantly from the experimental values, showing that the UNIFAC model was not able to satisfactory predict the partition values for the conventional and green Arizona N (version 1).

4.2. The NRTL-SAC model

The UNIFAC model predicts activity coefficient values only requiring information about chemical structure of the molecules of solutes and solvents [66]. However, this model is unable to obtain activity coefficient values for complex molecules, because either the functional group is not defined or the additivity rule becomes invalid [66].

A possible alternative that is able to overcome UNIFAC limitations regarding complex molecules calculations is the Nonrandom Two-Liquid Segment Activity Coefficient Model, or NRTL-SAC, a thermodynamic model that can predict the solubility of a solute in pure and mixed solvent systems [66,67]. This model uses experimental data to obtain the molecular parameters of the solute and, then, the model can be used to predict the phase equilibria in other solvents, as long

as the molecular descriptors of the latter are also available [66]. Nonetheless, the NRTL-SAC model presents limitations because it requires a large databank of experimental equilibrium data for the determination of the molecular descriptors. In addition, the melting temperatures and enthalpies are needed for the solubility data and the phase composition for all solvent systems for the partition coefficient calculations [66].

To obtain the necessary parameters for the solubility calculations, the NRTL-SAC outlines molecules in predefined conceptual segments, also known as molecular descriptors, according to their interaction and behavior in solutions [66]. There are mainly three molecular descriptors defined in this model: hydrophobic segment (X), hydrophilic segment (Z) and polar segments (+Y and –Y) [66,68]. Unlike the UNIFAC group model, the NRTL-SAC molecular descriptors are not determined based on the molecular structure, but from the interaction behavior of the molecules of the solute in solution based on experimental equilibrium data [66]. The hydrophobic segment corresponds to molecular surfaces that presents adversity to form hydrogen bond; the hydrophilic segment amounts molecular surface areas that show tendency to form hydrogen bonds and the polar segment represents polar molecular surface regions that can "donate or accept electrons" [66]. The binary interaction parameters express the characteristics of the segment-segment interaction [66]. Thereby, the activity coefficient and, as consequence, the partition coefficient are a function of the X, +Y, -Y and Z molecular parameters and of the two liquid phase equilibrium composition (PE) [68].

In relation with the NRTL-SAC model, it is possible to characterize organic groups regarding the three molecular segments previously described. Alkanes are mainly hydrophobic, therefore can be characterized only with hydrophobicity [66]. Differently, alcohols are well described by X and Z, since they carry both hydrophobic and hydrophilic segments [66]. Ketones, esters and ethers are described as polar molecules with more or less hydrophobic segments. They are well represented by X and Y's [66]. Lastly, acids exhibit hydrophilicity, polarity, and hydrophobicity [66].

The solubility and partition coefficients of vanillin, ferulic acid, quercetin and hesperetin were already described in the literature [69]. In that work, solubility and partition in binary solvent systems were selected in order to fit the molecular parameters. The molecular descriptors of 62 common solvents were the ones reported by Chen and Crafts (2006) [67]. Those results will be compared here with ones obtained by the Abraham solvation model.

4.3. Abraham solvation model

The Abraham solvation model describes the solubility or the partition of a solute of interest in a linear correlation of a relatively small number of descriptors, as described in Equation 6 [70]:

$$\log(P_S) = c + eE + sS + aA + bB + vV \quad (6)$$

Equation 6 calculates the partition coefficient between two liquid phases [32]. The lowercase coefficients characterize the condensed phase in solubility predictions or the solvent system under consideration in partition coefficients predictions [32,70,71]. The uppercase descriptors are the Abraham descriptors that describe the solute properties [32,70]. *E* is the liquid excess molar refraction at 20 ºC, and can be calculated as described by Zissimos *et al.* [70]. *S* describes the dipolarity/polarizability of the solute [70]. *A* and *B* are the overall or effective hydrogen bond acidity and basicity, respectively [70]. *S*, *A* and *B* are calculated by regression of several biphasic partition coefficients, and the descriptor *S* can also be obtained from gas liquid chromatographic measurements [70]. *V* is the McGowan characteristic volume that can be calculated from bond and atom contributions of the molecular structure [32,70].

The Abraham solvation model predicts solubilities based on the assumption that the partition of a solute (P_S) can be given by the solubility in the solvent of interest (S_S) divided by the solute solubility in water (S_w), as expressed in the Equation 7 [71]:

$$P_S = \frac{S_S}{S_W} \quad (7)$$

Equation 7 is valid if the solute is present in the same form in both phases, that is, no polymorphism, dissociation, solvate or hydrate formation and the solute is not too soluble in water or in the solvent [71].

The main advantage of the partition coefficient predictions using the Abraham solvation model is its simplicity; however, this method is highly dependent on the availability of the solute and solvents descriptors and of solubility and partition data [71].

In this work, the solute descriptors already available in the literature were first used to predict the solubilities and partition coefficients of vanillin, ferulic acid, quercetin, hesperetin and rutin. If necessary, the molecular descriptors were recalculated considering also the partition

coefficients measured in this work in binary mixtures. For solubilities, the data reported in literature were also used to estimate the correspondent partition coefficient as in Equation 7. The descriptors *E* and *V* were calculated from the molecular structure and refractive index of the solute, respectively. The latter was estimated by the software ACD/ChemSketch Freeware. Then, the descriptors *S*, *A* and *B* were optimized, considering the experimental partition coefficient and solubility database. All the solvent descriptors used in this work are presented in Appendix C.

In this section, the results obtained with the Abraham solvation model were compared with the results calculated by the NRTL-SAC model [69]. In this work, the NRTL-SAC calculations were carried out for all solutes, with the exception of rutin [69]. The correlation step included a database composed of partition coefficients in binary solvent systems (measured in this work) and solubilities. The descriptors obtained in the correlation step were then used to predict a set of partition coefficients in other binary and quaternary solvent systems [69].

The quality of the predictions obtained by the NRTL-SAC and Abraham solvation model is evaluated by the root-mean square deviation (RMSD), described in Equation 8:

$$\mathrm{RMSD} = \sqrt{\left[\frac{\sum_i \left(log\ P_{s,i}{}^{exp} - log\ P_{s,i}{}^{calc} \right)^2}{n}\right]} \qquad (8)$$

Where $P_{s,i}$ represents the partition coefficient of a solute in the solvent system i, exp represents the experimental values and $calc$ the calculated values, n is the total number of data points.

4.3.1. Vanillin

Abraham *et al.* (2010) reported the 5 descriptors for vanillin using log(*P*) values from the Medicinal Chemistry Database [71], as indicated in Table 7. In that work, the authors also measured the solubility of vanillin in water, 1-propanol, toluene and butanone by diluting and continuously stirring the compounds and solvents for 24 h [71]. The octanol-water partition was determined by the shake-flask method [71].

Table 7. Abraham descriptors of vanillin published in the literature [71].

E	***S***	***A***	***B***	***V***
0.99	1.30	0.31	0.68	1.131

For vanillin, the partition coefficients were calculated for a dataset composed by 14 partition coefficient values and 6 solubilities values using the Abraham descriptors in Table 7. Experimental values of partition in 1-octanol/water, 1-butanol/water, ethyl acetate/water, heptane/methanol were determined by the shake-flask method in this work. Other values in the dataset were reported in literature, as described in Section 2.7. The results and comparison with literature values are explicitly shown in Appendix C. Figure 22 summarizes the calculated partition coefficients of vanillin using the Abraham solvation model.

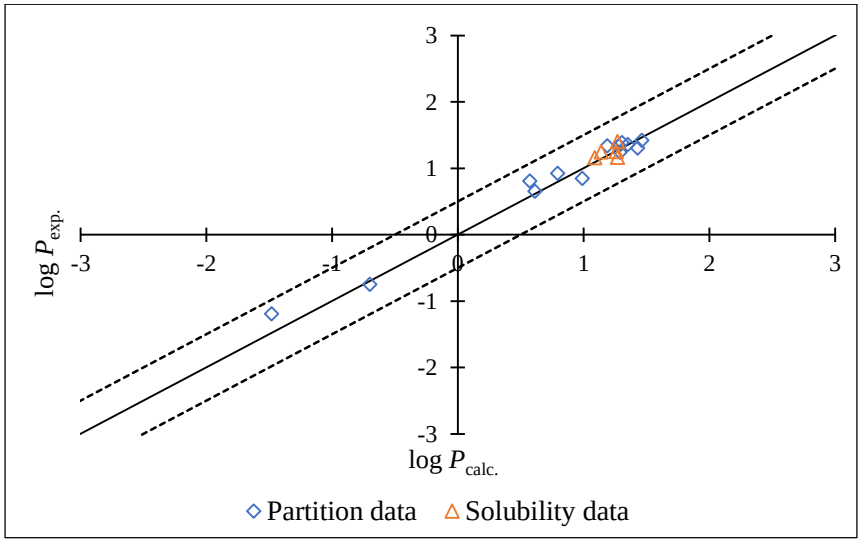

Figure 22. Comparison between experimental and calculated data by the Abraham solvation model for vanillin. The dashed lines correspond to the log P ± 0.5 log unit.

As observed in Figure 22, the calculated values of partition using the Abraham solvation model are in good agreement with the experimental data. In this dataset, it was observed a global RMSD equal to 0.12. The model was able to predict satisfactorily both partition coefficients and solubilities in distinct solvents. That was also the case of the NRTL-SAC model, that presented a RMSD of 0.10 [69]. For this model, in the prediction step, a RMSD of 0.18 was obtained [69]. Vanillin showed the lowest RMSD in the correlation and prediction steps in comparison with other solutes. The excellent agreement of the calculated data from NRTL-SAC and the Abraham solvation model with the experimental data can be due to the simplicity of the molecular structure of vanillin, and the availability of experimental information that allowed to estimate robust parameters.

4.3.2. Rutin

Abraham and Acree Jr. (2011) used the descriptors of rutin (Table 8) for the prediction of the solubility in several ethanol + water mixtures obtained by Peng *et al.* (2009), as detailed in 2.7.2. Rutin [29,72].

Table 8. Abraham descriptors of rutin published in the literature [72].

E	***S***	***A***	***B***	***V***
4.41	2.50	2.50	3.73	3.965

These descriptors were used to predict the partition coefficients measured in this work (1-octanol/water, 1-butanol/water, ethyl acetate/water and 1-heptane/methanol) and solubilities in pure solvents and water + ethanol mixtures as described in Section 2.7.2. Figure 23 presents a comparison of the experimental and calculated values of partition coefficients and solubilities using the descriptors of rutin obtained by Abraham and Acree Jr. (2011). The detailed results are presented in Appendix C.

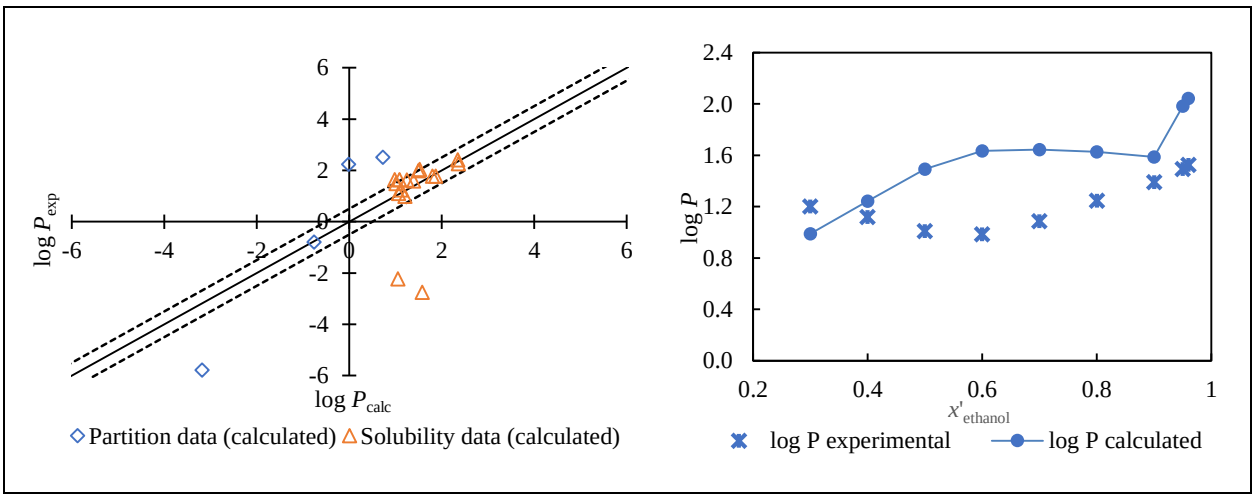

Figure 23. Comparison between experimental and calculated data by the Abraham solvation model: partition data in binary biphasic systems and solubilities (in the left) and solubilities in mixed solvents of water and ethanol (in the right), for rutin. The dashed lines correspond to the log P ± 0.5 log unit.

As observed in Figure 23, the absolute difference in log P between experimental and calculated partition coefficient data is higher than 0.5. In addition, the trend observed in the correlated solubility in mixed solvents is not well captured by the model. For the solubility in alcohols, the absolute difference in log P is lower than 0.5 in comparison with experimental and calculated values, while ethyl acetate and acetone presented an absolute difference in log P of 4.32

and 3.28, respectively. Therefore, the rutin descriptors *S*, *A* and *B* were reestimated. The optimization database included the partition coefficients in 1-octanol/water, 1-butanol/water, ethyl acetate/water and heptane/methanol, obtained in this work, and the solubility of rutin in pure solvents determined by Zi *et al.* (2007) and Peng *et al.* (2009) [29,30].

Table 9. Abraham descriptors of rutin obtained in this work.

E	***S***	***A***	***B***	***V***
5.93	3.95	0.32	3.38	3.965

Figure 24 summarizes the partition coefficients and solubility values calculated using the optimized descriptors from Table 9, in comparison with the experimental values. The detailed results are present in Appendix C.

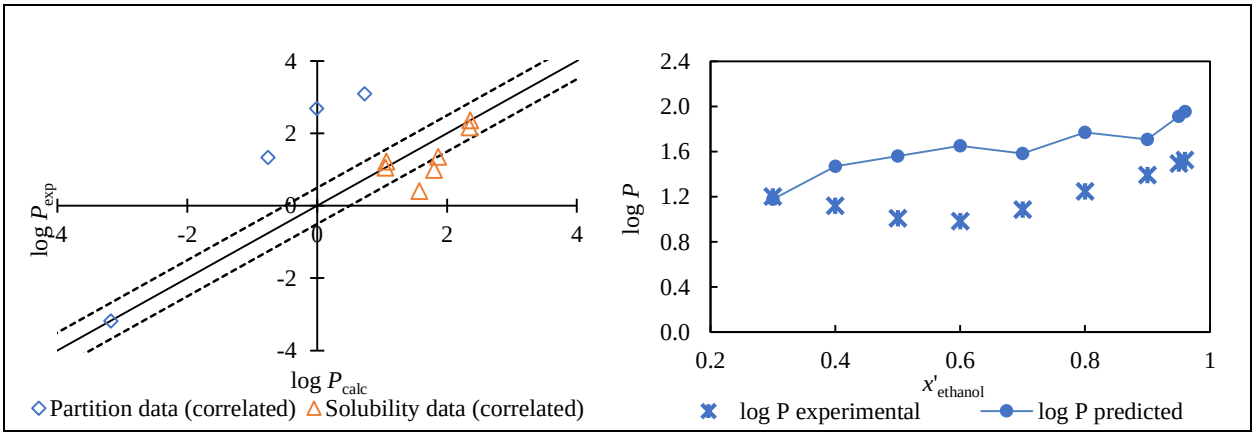

Figure 24. Comparison between experimental partition and solubility in pure solvents and correlated data for rutin (in the left) and comparison between experimental solubilities in mixed solvents and predicted data (in the right) using optimized descriptors by the Abraham solvation model. The dashed lines correspond to the log P ± 0.5 log unit.

Observing Figure 24, the optimized descriptors were not able to improve the description of the partition data and the prediction of the solubility of rutin in the mixed solvents, presenting a RMSD of 1.14 for correlation and 0.45 for prediction. The high value RMSD in the correlation step was mostly due to the inability of the model to correlate the partition data. The comparison between the solubility data in mixed solvents and predicted data shows that the Abraham model was not able to predict the changes in solubility from $x'_{ethanol}$ = 0.3 to $x'_{ethanol}$ = 0.96.

Krewson and Naghski (1952) reported that rutin forms different solvates in several organic solvents, including ethanol and water [73]. The formation of rutin alcoholate, that is, rutin with ethanol of crystallization, or rutin hydrate interferes drastically in the solubility of rutin in water

[73]. Peng et al. [29] also reported SEM analysis that demonstrate that the crystal forms of rutin changed in morphology at different compositions of the mixed solvent. The formation of different solid phases of rutin that present distinct solubilities can interfere in the quality of the experimental results. Therefore, a consistent database of solubilities and/or partition coefficients of rutin in different solvents is necessary in order to obtain suitable descriptors for this solute. In the case of the solubilities, the characterization of the solid phase is mandatory. As mentioned before, Equation 7 is only valid if the solute is present in the same form in both phases [71].

4.3.3. Ferulic acid

Vilas-Boas *et al.* (2020) estimated the Abraham descriptors of ferulic acid with a dataset that included the solubility in methanol, ethanol, 1-propanol, 2-propanol, 2-butanone, ethyl acetate, DMSO and acetonitrile at 298.2 K [32]. The descriptors are presented in Table 10.

Table 10. Abraham descriptors of ferulic acid published in the literature [32].

E	***S***	***A***	***B***	***V***
1.472	1.138	0.290	0.877	1.429

Using those descriptors, the values of several partition coefficients and solubilities were calculated. The complete dataset of partition coefficients and solubilities, as well as the correspondent calculated values for ferulic acid, are presented in Appendix C. Figure 25 presents the calculated values in comparison with the experimental values.

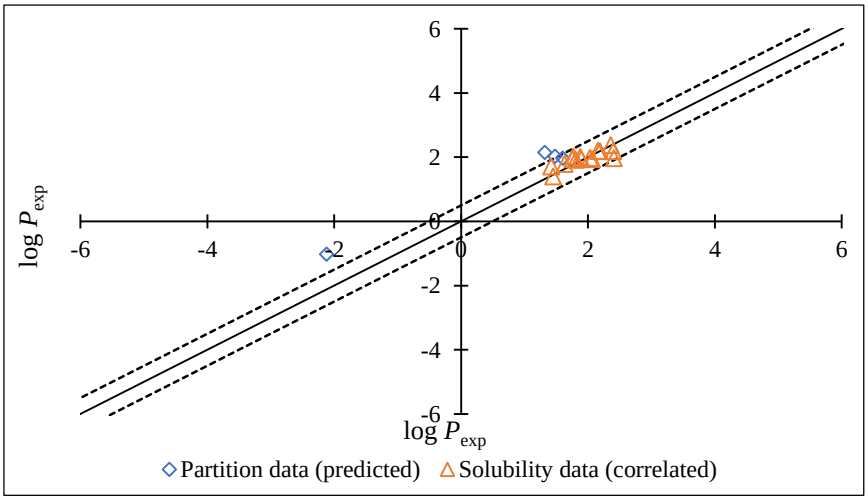

Figure 25. Comparison between experimental and calculated data by the Abraham solvation model for ferulic acid. The dashed lines correspond to the log P ± 0.5 log unit.

As observed in Figure 25, the estimated descriptors can satisfactorily describe the solubility data, since the dataset used to estimate these Abraham descriptors is composed only by solubility values. However, the model was not able to predict the partition values measured in this work, with three solvent systems presenting an absolute difference in log greater than 0.5, especially for heptane/methanol, with an absolute difference of 1.10. Therefore, an optimization with a dataset that includes those partition coefficients besides the solubility data measured in literature [32] was carried out. Table 11 presents the optimized Abraham descriptors.

Table 11. Abraham descriptors of ferulic acid obtained in this work.

E	***S***	***A***	***B***	***V***
1.472	1.440	0.280	0.892	1.429

The parameters obtained were also used to predict the solubility in a set of solvents not included in the correlation step. Figure 26 summarizes all the results obtained. The detailed results are in Appendix C.

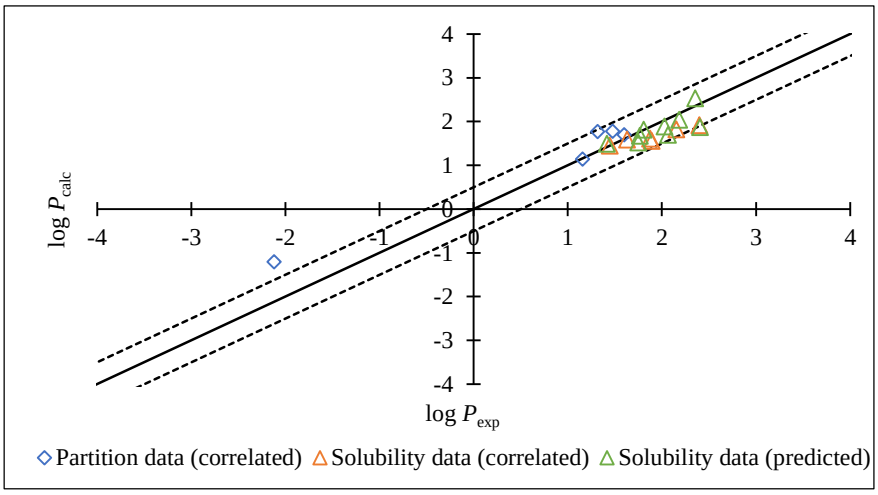

Figure 26. Comparison between experimental and calculated data using optimized descriptors by the Abraham solvation model for ferulic acid. The dashed lines correspond to the log P ± 0.5 log unit.

Using the optimized descriptors, the model was able to satisfactory correlate both partition coefficient and solubility values, with exception of heptane/methanol partition coefficient, which presents an absolute difference in log of 0.91. In this model, ferulic acid presented a RMSD of 0.39 and 0.25 in the correlation and prediction step, respectively.

In the NRTL-SAC model, the systems containing ferulic acid presented larger deviations than the ones containing vanillin, as in the Abraham solvation model. RMSD of 0.37 and 0.26

were obtained in the correlation and prediction steps, respectively [69]. In general, both models provide a good description of the experimental data. It should be mentioned that the quality of the solubility data calculated by the NRTL-SAC model largely depends on the accuracy of the available melting properties of the solutes, which can be complex to be measured for some phenolic compounds [69].

4.3.4. Quercetin

Abraham and Acree Jr. (2014) estimated the descriptors of quercetin with a dataset composed by solubility in pure methanol and ethanol, and partition data in 1-octanol/water and gas/water [74]. Due to the large disparity of literature solubility values of quercetin in water [40,75], Abraham and Acree Jr. (2014) calculated the solubility of quercetin in water, obtaining log C_w = -3.90 at 298 K, with units of C_w in mol/L [74]. Table 12 shows the estimated descriptors for quercetin.

Table 12. Abraham descriptors of quercetin published in the literature [74].

E	***S***	***A***	***B***	***V***
2.68	2.20	1.92	1.40	1.9632

The descriptors in Table 12 were used to predict the partition coefficients measured in this work and solubility data in 1-propanol, 1-butanol, ethyl acetate, ethanol, methanol, acetonitrile and butyl acetate, as described in Section 2.7.5. Figure 27 presents the calculated solubility and partition values in comparison with the experimental values. The complete results are presented in Appendix C.

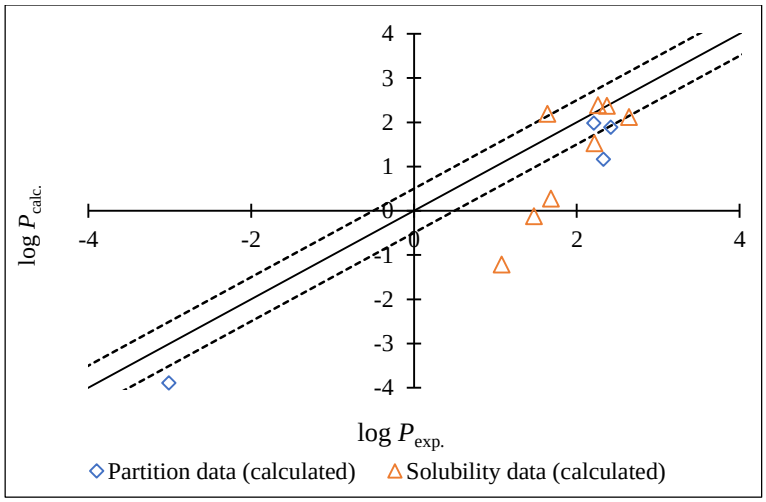

Figure 27. Comparison between experimental and calculated data for quercetin. The dashed lines correspond to the log P ± 0.5 log unit.

With the descriptors available in literature, the model was not able predict both partition coefficient and solubility values. Therefore, the optimization of the Abraham solute descriptors was made with a dataset composed of partition coefficients values and solubilities in pure solvents (1-propanol, 1-butanol, ethanol and ethyl acetate) to predict solubilities in methanol, butyl acetate and acetonitrile.

Table 13 presents the optimized Abraham descriptors.

Table 13. Abraham descriptors of quercetin obtained in this work.

E	***S***	***A***	***B***	***V***
3.534	2.568	1.027	1.232	1.963

Figure 28 shows the comparison between experimental and calculated data using the optimized descriptors:

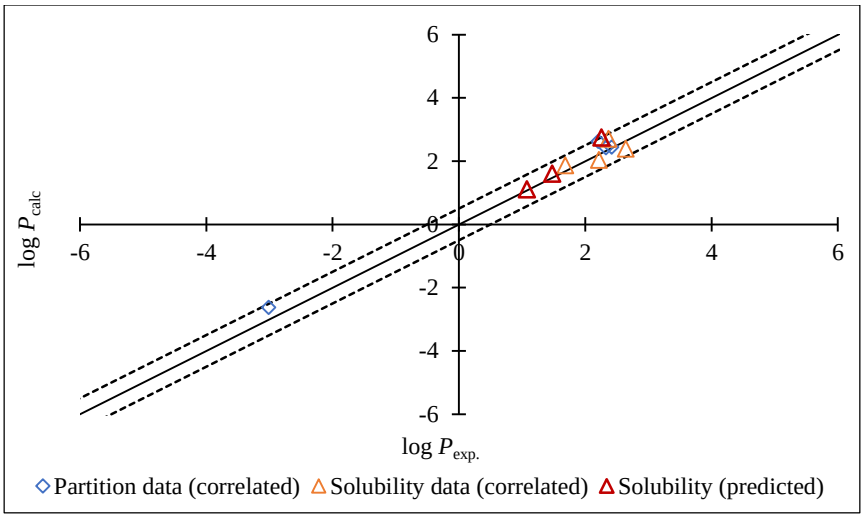

Figure 28. Comparison between experimental and calculated data for quercetin using optimized descriptors by the Abraham solvation model. The dashed lines correspond to the log P ± 0.5 log unit.

Using the optimized descriptors, the Abraham model was able to better correlate and predict both partition and solubility data. Observing Figure 28, the correlated solubilities in pure solvents are better fitted in comparison with the values obtained using the descriptors found in literature [74], with a RMSD of 0.27 for correlation and 0.29 for prediction.

In the NRTL-SAC model, quercetin presented larger deviations than the results obtained by the Abraham solvation model, with a RMSD of 0.41 and 0.63 for the prediction step [69]. In general, the values predicted in the NRTL-SAC were also satisfactory. However, the Abraham solvation model also presented better results with a smaller dataset in the optimization step.

4.3.5. Hesperetin

Talebi *et al.* (2020) calculated the descriptors of hesperetin using a set of 28 log P_s of solubility data available in literature [36–38] in mixed solvents (methanol + water and ethanol + water), and also in pure solvents (1-octanol, methanol, 1-butanol, ethanol, acetonitrile, ethyl acetate and acetone) [25].

Table 14 shows the estimated descriptors for hesperetin.

Table 14. Abraham descriptors of hesperetin published in the literature [25].

E	***S***	***A***	***B***	***V***
2.357	2.795	0.891	0.906	2.088

Table 14 were used to predict the partition data measured in this work. Those predictions and the correlated solubilities in pure and mixed solvents are compared with the experimental values in Figure 29. The detailed results can be found in Appendix C.

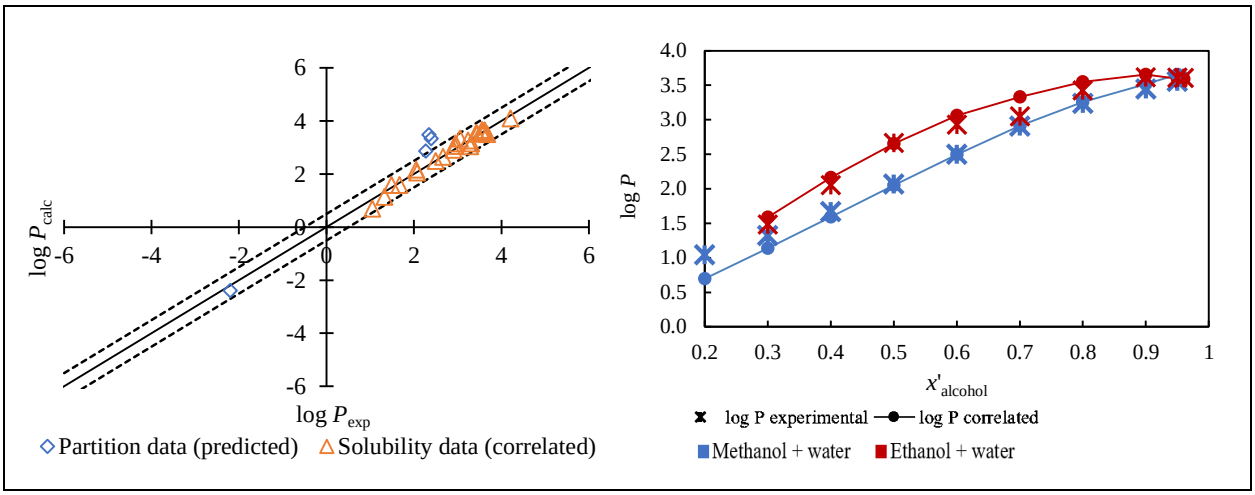

Figure 29. Comparison between experimental partition and solubility data and calculated data for hesperetin (in the left) and comparison between experimental solubilities in mixed solvents and calculated data (in the right). The dashed lines correspond to the log P ± 0.5 log unit.

With the descriptors available in literature, the model was able to correlate satisfactory the solubility in pure and mixed solvents, with a RMSD of 0.14 in the correlation step. However, it could not predict the partition values (RMSD = 0.72), since the dataset used to obtain the Abraham descriptors was only composed by solubility data.

The optimization of the Abraham solute descriptors was made with a dataset composed of partition coefficients and solubilities in pure solvents (acetone, ethanol, ethyl acetate, methanol, 1-octanol, 1-butanol and acetonitrile) to predict solubilities in the methanol + water and ethanol + water. Table 15 presents the optimized Abraham descriptors.

Table 15. Abraham descriptors of hesperetin obtained in this work.

E	***S***	***A***	***B***	***V***
2.358	3.047	0.785	0.929	2.088

Figure 30 shows the comparison between experimental and calculated data using the descriptors optimized in this work.

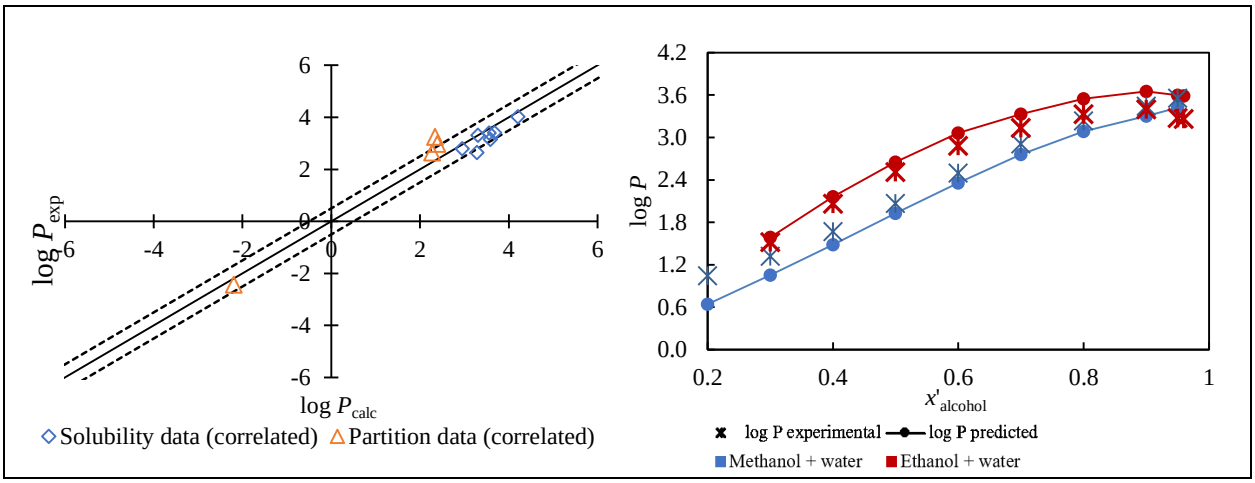

Figure 30. Comparison between experimental partition and solubility data and correlated data for hesperetin (in the left) and comparison between experimental solubilities in mixed solvents and predicted data (in the right) using optimized descriptors by the Abraham solvation model. The dashed lines correspond to the log P ± 0.5 log unit.

Observing Figure 30, it is possible to conclude that the results obtained using the optimized descriptors did not significantly differ from the results using the descriptors from literature [25] for solubility results, presenting a RMSD of 0.44 in the correlation and 0.20 in the prediction step. The correlation of the partition coefficient values presented a RMSD of 0.58, a better result in comparison with the predictions obtained with the descriptors from literature (RMSD = 0.72).

In the NRTL-SAC model, hesperetin presented a RMSD of 0.40 in the correlation step [69]. In the prediction step, the only data included in the dataset was the partition for the ARIZONA N system measured in this work, with an absolute log difference 0.29 between the experimental and predicted value [69]. Overall, the values calculated in the NRTL-SAC were satisfactory. However, due to the small dataset used in the prediction step in the NRTL-SAC model, it is not possible to correlate the adequacy of this model in comparison with the Abraham solvation model.

5. Conclusions and future work

In this work, the partition coefficient of vanillin, rutin, ferulic acid, hesperetin and quercetin were measured in several biphasic solvent systems at (298.2 ± 0.5) K, by the shake-flask method. The data obtained are consistent to most of the data available in literature, although larger deviations for the 1-octanol/water system are observed. For this particular biphasic system, the possible formation of emulsions during the shake-flask method was already reported and can be the cause of error. So, to avoid those issues, in this work, the solvents were pre-saturated before the preparation of the shake-flasks.

Besides the four biphasic systems composed of the binary mixtures water + ethyl acetate, water + 1-butanol, water + 1-octanol and heptane + methanol, three biphasic systems composed of quaternary mixtures were studied. Thus, the partition coefficients were also measured in the conventional Arizona N system, and in two modified versions containing alternative green solvents (version 1: replacing heptane by limonene, and methanol by ethanol; version 2: replacing only methanol by ethanol). Comparing with the conventional Arizona system, substitution of methanol by ethanol in the green Arizona version 1 and the additional replacement of heptane by limonene in the green Arizona version 2 increases the polarity in the upper phase. To the best of our knowledge, the log(K) of all the studied solutes in the Arizona N green systems are reported here for the first time. Those systems can be considered viable alternatives, as small increments are obtained in the partition coefficients when using the green versions compared to the conventional one, while providing a more environmental-friendly approach, with beneficial health and safety impact.

Obtaining the experimental values of the partition coefficients of a given solute to select the most favourable biphasic solvent system can be a time consuming task. Under certain conditions, thermodynamic tools such as the UNIFAC model and the Abraham solvation model can be applied in the preliminary screening of the solvents systems. In the studied systems, the UNIFAC method was unable to accurately predict the partition coefficients of vanillin and hesperetin, presenting an absolute log difference of 1.68 and 4.15, respectively, between the conventional Arizona N calculated partitions and obtained experimentally in this work. For ferulic

acid, the absolute log difference between the conventional Arizona N calculated partition and experimental was 0.68, a satisfactory result in comparison with the other solutes. The global RDSM for ferulic acid and vanillin were 0.84 and 1.42, respectively, considering the experimental results of 1-butanol/water, ethyl acetate/water, heptane/methanol and conventional Arizona N. For the green Arizona N version 1, the absolute log difference between the calculated and experimental partition coefficient are 4.37, 2.22 and 6.22 for vanillin, ferulic acid and hesperetin, respectively. In conclusion, the UNIFAC model was not able to predict the partition for the Arizona N version 1 for all solutes. For quercetin and rutin, the UNIFAC model could not be applied due to the lack of some parameterized groups.

On the other hand, the Abraham solvation model needs experimental data input, preferably partition coefficients and/or solubility data in solvent systems of diverse polarity, to allow estimating the solute descriptors. In this work, new sets of parameters were proposed for vanillin, ferulic acid, quercetin and hesperetin, aiming to improve the description of the partition coefficients measured in this work. The Abraham solvation model provided a good description of the selected partition coefficients and solubilities, with a global RMSD of 0.57.

The NRTL-SAC model applied in the literature to the same solutes also needed experimental data input, including a dataset of partitions and solubilities, in mixtures with a diverse polarity range. For this model, a global deviation in log units of 0.3 was obtained, achieving a global RMSD of 0.3 for the phenolic compounds under study. The overall results are very similar to the Abraham solvation model. However, the latter has advantages such as its simplicity and requiring less experimental data input, but it is more difficult to obtain the parameters for biphasic systems composed of more than two solvents.

In order to select a suitable solvent system to promote the separation of a given solute, the experimental results reported in this work are of great importance. Nonetheless, the information regarding the solubility and partition of these solutes is still scarce in the literature, with the exception of vanillin. For future work, additional experimental measurements and thermodynamic modelling of partition coefficients and solubilities in several solvent systems should be carried out, extending those studies to alternative green solvents. It is also suggested to obtain the parameters for limonene in the NRTL-SAC model and the Abraham descriptors for the Arizona systems (conventional and green versions) as well as the experimental measurement of the phase compositions of version 2. Finally, the application of the fully predictive Conductor-like Screening

Model for Real Solvents (COSMO-RS model) is suggested for preliminary solvent screening, particularly when no experimental data are available for the solvents and/or solutes understudy.

References

[1] Z. Berk, Food Process Engineering and Technology, in: Z. Berk (Ed.), Food Process Eng. Technol., 3rd ed., Academic Press, 2018: pp. 289–310. https://doi.org/10.1016/B978-0-12-812018-7.00011-7.

[2] P.G. Mazzola, A.M. Lopes, F.A. Hasmann, A.F. Jozala, T.C. Penna, C.O. Magalhaes, Perola O Rangel-Yagui, A. Pessoa Jr, Liquid–liquid extraction of biomolecules: an overview and update of the main techniques, J. Chem. Technol. Biotechnol. 83 (2008) 113–157. https://doi.org/10.1002/jctb.

[3] M. Silva, J.C. García, M. Ottens, Polyphenol Liquid-Liquid Extraction Process Development Using NRTL-SAC, Ind. Eng. Chem. Res. 57 (2018) 9210–9221. https://doi.org/10.1021/acs.iecr.8b00613.

[4] K. Hostettnann, A. Marston, M. Hostettnann, K. Hostettmann · A. Marston · M. Hostettmann Preparative Chromatography Techniques, 2nd ed., Springer-Verlag Berlin Heidelberg New York, New York, 1998. https://doi.org/10.1007/978-3-662-03631-0.

[5] W. Vetter, M. Müller, M. Englert, S. Hammann, Countercurrent chromatography-when liquid-liquid extraction meets chromatography, in: Liq. Extr., 2019: pp. 289–325. https://doi.org/10.1016/B978-0-12-816911-7.00010-4.

[6] F. Das Neves Costa, G.G. Leitão, Strategies of solvent system selection for the isolation of flavonoids by countercurrent chromatography, J. Sep. Sci. 33 (2010) 336–347. https://doi.org/10.1002/jssc.200900632.

[7] K. Faure, E. Bouju, P. Suchet, A. Berthod, Limonene in Arizona liquid systems used in countercurrent chromatography. i physicochemical properties, Anal. Bioanal. Chem. 406 (2014) 5909–5917. https://doi.org/10.1007/s00216-014-8005-3.

[8] L. Lorántfy, R. Örkenyi, D. Rutterschmid, D. Bakonyi, J. Faragó, G. Dargó, Á. Könczöl, Continuous industrial-scale centrifugal partition chromatography with automatic solvent system handling: Concept and instrumentation, Org. Process Res. Dev. 24 (2020) 2676–2688. https://doi.org/10.1021/acs.oprd.0c00338.

[9] C.F. Poole, Milestones in the development of liquid-phase extraction techniques, in: Liq. Extr., Elsevier Inc., 2019: pp. 1–44. https://doi.org/10.1016/B978-0-12-816911-7.00001-3.

[10] J.B. Friesen, G.F. Pauli, G.U.E.S.S. - A generally useful estimate of solvent systems in CCC, J. Liq. Chromatogr. Relat. Technol. 28 (2005) 2777–2806. https://doi.org/10.1080/10826070500225234.

[11] H.T. Lu, Y. Jiang, F. Chen, Application of preparative high-speed counter-current chromatography for separation of chlorogenic acid from Flos Lonicerae, J. Chromatogr. A. 1026 (2004) 185–190. https://doi.org/10.1016/j.chroma.2003.11.004.

[12] K. Skalicka-Woźniak, I. Garrard, A comprehensive classification of solvent systems used for natural product purifications in countercurrent and centrifugal partition chromatography, Nat. Prod. Rep. 32 (2015) 1556–1561. https://doi.org/10.1039/c5np00061k.

[13] E. Hopmann, W. Arlt, M. Minceva, Solvent system selection in counter-current chromatography using conductor-like screening model for real solvents, J. Chromatogr. A.

1218 (2011) 242–250. https://doi.org/10.1016/j.chroma.2010.11.018.
[14] A. Berthod, M. Hassoun, M.J. Ruiz-Angel, Alkane effect in the Arizona liquid systems used in countercurrent chromatography, Anal. Bioanal. Chem. 383 (2005) 327–340. https://doi.org/10.1007/s00216-005-0016-7.
[15] Y. Liu, P. Kuang, S. Guo, Q. Sun, T. Xue, H. Li, An overview of recent progress in solvent systems, additives and modifiers of counter current chromatography, New J. Chem. 42 (2018) 6584–6600. https://doi.org/10.1039/c7nj04747a.
[16] I.J. Garrard, L. Janaway, D. Fisher, Minimising solvent usage in high speed, high loading, and high resolution isocratic dynamic extraction, J. Liq. Chromatogr. Relat. Technol. 30 (2007) 151–163. https://doi.org/10.1080/10826070601077252.
[17] F.P. Byrne, S. Jin, G. Paggiola, T.H.M. Petchey, J.H. Clark, T.J. Farmer, A.J. Hunt, C. Robert McElroy, J. Sherwood, Tools and techniques for solvent selection: green solvent selection guides, Sustain. Chem. Process. 4 (2016). https://doi.org/10.1186/s40508-016-0051-z.
[18] K. Faure, E. Bouju, J. Doby, A. Berthod, Limonene in Arizona liquid systems used in countercurrent chromatography. II Polarity and stationary-phase retention, Anal. Bioanal. Chem. 406 (2014) 5919–5926. https://doi.org/10.1007/s00216-014-8030-2.
[19] M.J. Frutos, L. Rincón-Frutos, E. Valero-Cases, Rutin, in: Nonvitamin Nonmineral Nutr. Suppl., Elsevier Inc., Alicante, 2018: pp. 111–117. https://doi.org/10.1016/B978-0-12-812491-8.00015-1.
[20] K.E. Burke, Prevention and Treatment of Aging Skin with Topical Antioxidants, William Andrew Inc., 2009. https://doi.org/10.1016/B978-0-8155-1584-5.50012-0.
[21] A. Jalan, R.W. Ashcraft, R.H. West, W.H. Green, Predicting solvation energies for kinetic modeling, Annu. Reports Prog. Chem. - Sect. C. 106 (2010) 211–258. https://doi.org/10.1039/b811056p.
[22] F. Shakeel, N. Haq, N.A. Siddiqui, Solubility and thermodynamic function of vanillin in ten different environmentally benign solvents, Food Chem. 180 (2015) 244–248. https://doi.org/10.1016/j.foodchem.2015.01.102.
[23] M.I.F. Mota, P.C.R. Pinto, J.M. Loureiro, A.E. Rodrigues, Recovery of Vanillin and Syringaldehyde from Lignin Oxidation: A Review of Separation and Purification Processes, Sep. Purif. Rev. 45 (2016) 227–259. https://doi.org/10.1080/15422119.2015.1070178.
[24] T. Muhammad, M. Ikram, R. Ullah, S.U. Rehman, M.O. Kim, Hesperetin, a citrus flavonoid, attenuates LPS-induced neuroinflammation, apoptosis and memory impairments by modulating TLR4/NF-κB signaling, Nutrients. 11 (2019) 1–20. https://doi.org/10.3390/nu11030648.
[25] J. Talebi, N. Saadatjou, A. Farajtabar, Hesperetin Solubility in Aqueous Co-solvent Mixtures of Methanol and Ethanol: Solute Descriptors, Solvent Effect and Preferential Solvation Analysis, J. Solution Chem. 49 (2020) 179–194. https://doi.org/10.1007/s10953-020-00948-6.
[26] R. Jaffe, J. Mani, Polyphenolics evoke healing responses: Clinical evidence and role of predictive biomarkers, 2nd ed., Elsevier Inc., 2018. https://doi.org/10.1016/B978-0-12-813006-3.00029-5.
[27] S.L. Hwang, P.H. Shih, G.C. Yen, Citrus Flavonoids and Effects in Dementia and Age-Related Cognitive Decline, Elsevier Inc., 2015. https://doi.org/10.1016/B978-0-12-407824-6.00080-X.

[28] A. Noubigh, M. Cherif, E. Provost, M. Abderrabba, Solubility of gallic acid, vanillin, syringic acid, and protocatechuic acid in aqueous sulfate solutions from (293.15 to 318.15) K, J. Chem. Eng. Data. 53 (2008) 1675–1678. https://doi.org/10.1021/je800205e.
[29] B. Peng, R. Li, W. Yan, Solubility of rutin in ethanol + water at (273.15 to 323.15) K, J. Chem. Eng. Data. 54 (2009) 1378–1381. https://doi.org/10.1021/je800816f.
[30] J. Zi, B. Peng, W. Yan, Solubilities of rutin in eight solvents at T = 283.15, 298.15, 313.15, 323.15, and 333.15 K, Fluid Phase Equilib. 261 (2007) 111–114. https://doi.org/10.1016/j.fluid.2007.07.030.
[31] F. Shakeel, M.M. Salem-Bekhit, N. Haq, N.A. Siddiqui, Solubility and thermodynamics of ferulic acid in different neat solvents: Measurement, correlation and molecular interactions, J. Mol. Liq. 236 (2017) 144–150. https://doi.org/10.1016/j.molliq.2017.04.014.
[32] S.M. Vilas-Boas, R.S. Alves, P. Brandão, L.M.A. Campos, J.A.P. Coutinho, S.P. Pinho, O. Ferreira, Solid-liquid phase equilibrium of trans-cinnamic acid, p-coumaric acid and ferulic acid in water and organic solvents: Experimental and modelling studies, Fluid Phase Equilib. 521 (2020) 1–10. https://doi.org/10.1016/j.fluid.2020.112747.
[33] C. Zhou, X. Shi, H. Wang, N. An, Measurement and correlation of solubilities of trans-ferulic acid in solvents, Chinese J. Chem. Eng. 58 (2007) 2705–2709. https://doi.org/10.1016/S1004-9541(08)60233-5.
[34] U. Domańska, A. Wiśniewska, Z. Dąbrowski, M. Więckowski, Ionic Liquids as Efficient Extractants for Quercetin from Red Onion (Allium cepa L.), J. Appl. Solut. Chem. Model. 7 (2018) 21–38. https://doi.org/10.6000/1929-5030.2018.07.04.
[35] N. Haq, N.A. Siddiqui, F. Shakeel, Solubility and molecular interactions of ferulic acid in various (isopropanol + water) mixtures, J. Pharm. Pharmacol. 69 (2017) 1485–1494. https://doi.org/10.1111/jphp.12786.
[36] O. Ferreira, S.P. Pinho, Solubility of flavonoids in pure solvents, Ind. Eng. Chem. Res. 51 (2012) 6586–6590. https://doi.org/10.1021/ie300211e.
[37] L. Liu, J. Chen, Solubility of Hesperetin in Various Solvents from (288.2 to 323.2) K, J. Chem. Eng. Data. 53 (2008) 1649–1650. https://doi.org/10.1021/je800078j.
[38] H. Zhang, M. Wang, L. Chen, Y. Liu, H. Liu, H. Huo, L. Sun, X. Ren, Y. Deng, A. Qi, Structure-solubility relationships and thermodynamic aspects of solubility of some flavonoids in the solvents modeling biological media, J. Mol. Liq. 225 (2017) 439–445. https://doi.org/10.1016/j.molliq.2016.11.036.
[39] O. Ferreira, B. Schröder, S.P. Pinho, Solubility of hesperetin in mixed solvents, J. Chem. Eng. Data. 58 (2013) 2616–2621. https://doi.org/10.1021/je400513s.
[40] R.S. Razmara, A. Daneshdar, R. Sahraei, Solubility of Quercetin in Water + Methanol and Water + Ethanol from (292.8 to 333.8) K, J. Chem. Eng. Data,. 5 (2010) 3934–3936. https://doi.org/10.1021/je9010757.
[41] A. Noubigh, M. Abderrabba, E. Provost, Salt addition effect on partition coefficient of some phenolic compounds constituents of olive mill wastewater in 1-octanol-water system at 298.15 K, J. Iran. Chem. Soc. 6 (2009) 168–176. https://doi.org/10.1007/BF03246517.
[42] L. Marlot, Développement de méthodes bidimensionnelles préparatives CPCxLC : application à la purification de molécules d ’ intérêt issues de matrices végétales, Université de Lyon, 2018. https://tel.archives-ouvertes.fr/tel-02069116v2.
[43] Y. Lu, A. Berthod, R. Hu, W. Ma, Y. Pan, Screening of complex natural extracts by countercurrent chromatography using a parallel protocol, Anal. Chem. 81 (2009) 4048–

4059. https://doi.org/10.1021/ac9002547.
[44] A. Berthod, S. Ignatova, I.A. Sutherland, Advantages of a small-volume counter-current chromatography column, J. Chromatogr. A. 1216 (2009) 4169–4175. https://doi.org/10.1016/j.chroma.2008.11.013.
[45] J.B. Friesen, S. Ahmed, G.F. Pauli, Qualitative and quantitative evaluation of solvent systems for countercurrent separation, J. Chromatogr. A. 1377 (2015) 55–63. https://doi.org/10.1016/j.chroma.2014.11.085.
[46] A. Berthod, K. Faure, Separations with a Liquid Stationary Phase: Countercurrent Chromatography or Centrifugal Partition Chromatography, Anal. Sep. Sci. (2015) 1177–1206. https://doi.org/10.1002/9783527678129.assep046.
[47] Z. Yang, Y. Wu, S. Wu, A combination strategy for extraction and isolation of multi-component natural products by systematic two-phase solvent extraction-13C nuclear magnetic resonance pattern recognition and following conical counter-current chromatography separation: Podophyl, J. Chromatogr. A. 1431 (2016) 184–196. https://doi.org/10.1016/j.chroma.2015.12.074.
[48] G. Lamprecht, K. Blochberger, Protocol for isolation of vanillin from ice cream and yoghurt to confirm the vanilla beans origin by 13C-EA-IRMS, Food Chem. 114 (2009) 1130–1134. https://doi.org/10.1016/j.foodchem.2008.10.071.
[49] K.L. Kaygorodov, Y. V Chelbina, V.E. Tarabanko, N. V Tarabanko, Extraction of vanillin by aliphatic alcohols, Журнал Сибирского Федерального Университета. Химия. 3 (2010) 228–233.
[50] C.A. Pedriali, A.U. Fernandes, L.D.C. Bernusso, B. Polakiewicz, The synthesis of a water-soluble derivative of rutin as an antiradical agent, Quim. Nova. 31 (2008) 2147–2151. https://doi.org/10.1590/S0100-40422008000800039.
[51] M.K. Das, B. Kalita, Design and evaluation of phyto-phospholipid complexes (phytosomes) of Rutin for transdermal application, J. Appl. Pharm. Sci. 4 (2014) 51–57. https://doi.org/10.7324/JAPS.2014.40110.
[52] J. Sangster, Octanol water partition coefficients of simple organic compounds, J. Phys. Chem. Ref. Data. 18 (1989) 1111–1229.
[53] L. Marlot, M. Batteau, K. Faure, Classification of biphasic solvent systems according to Abraham descriptors for countercurrent chromatography, J. Chromatogr. A. 1617 (2020). https://doi.org/10.1016/j.chroma.2019.460820.
[54] L. Marlot, Développement de méthodes bidimensionnelles préparatives CPCxLC : application à la purification de molécules d ’ intérêt issues de matrices végétales, (2019).
[55] Q.Q. Wang, J.B. Shi, C. Chen, C. Huang, W.J. Tang, J. Li, Hesperetin derivatives: Synthesis and anti-inflammatory activity, Bioorganic Med. Chem. Lett. 26 (2016) 1460–1465. https://doi.org/10.1016/j.bmcl.2016.01.058.
[56] Z. Shi, J. He, T. Yao, W. Chang, RP-HPLC Determination of Octanol-Water Partition Coefficients for Bioactive Compounds from Chinese Herbal Medicines, J. Liq. Chromatogr. Relat. Technol. 27 (2004) 465–479.
[57] Y.H. Tsai, K.F. Lee, Y. Bin Huang, C. Te Huang, P.C. Wu, In vitro permeation and in vivo whitening effect of topical hesperetin microemulsion delivery system, Int. J. Pharm. 388 (2010) 257–262. https://doi.org/10.1016/j.ijpharm.2009.12.051.
[58] J.A. Rothwell, A.J. Day, M.R.A. Morgan, Experimental determination of octanol-water partition coefficients of quercetin and related flavonoids, J. Agric. Food Chem. 53 (2005) 4355–4360. https://doi.org/10.1021/jf0483669.

[59] I.A. Sima, A. Kot-Wasik, A. Wasik, J. Namiésnik, C. Sârbu, Assessment of lipophilicity indices derived from retention behavior of antioxidant compounds in RP-HPLC, Molecules. 22 (2017). https://doi.org/10.3390/molecules22040550.
[60] L.J. Jin, Z. Wei, J.Y. Dai, P. Guo, L.S. Wang, Prediction of partitioning properties for benzaldehydes by various molecular descriptors, Bull. Environ. Contam. Toxicol. 61 (1998) 1–7. https://doi.org/10.1007/s001289900721.
[61] D.M. Miller, Evidence that interfacial transport is rate-limiting during passive cell membrane permeation, BBA - Biomembr. 1065 (1991) 75–81.
[62] Z. Yang, P. Guo, R. Han, D. Wu, J.M. Gao, S. Wu, Methanol linear gradient counter-current chromatography for the separation of natural products: Sinopodophyllum hexandrum as samples, J. Chromatogr. A. 1603 (2019) 251–261.
[63] J.C. Dearden, G.M. Bresnen, The measurement of partition coefficients, Quant. Struct. Relationships. 7 (1988) 133–144.
[64] C.M. Galanakis, V. Goulas, S. Tsakona, G.A. Manganaris, V. Gekas, A knowledge base for the recovery of natural phenols with different solvents, Int. J. Food Prop. 16 (2013) 382–396. https://doi.org/10.1080/10942912.2010.522750.
[65] B.E. Poling, J.M. Prausnitz, J.P. O'Connell, The Properties of Gases and Liquids, 5th ed., McGraw-Hill, 2001. https://doi.org/10.1063/1.3060771.
[66] C.C. Chen, Y. Song, Solubility modeling with a nonrandom two-liquid segment activity coefficient model, Ind. Eng. Chem. Res. 43 (2004) 8354–8362. https://doi.org/10.1021/ie049463u.
[67] C.C. Chen, P.A. Crafts, 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. 45 (2006) 4816–4824. https://doi.org/10.1021/ie051326p.
[68] D.B. Ren, Z.H. Yang, Y.Z. Liang, Q. Ding, C. Chen, M.L. Ouyang, Correlation and prediction of partition coefficient using nonrandom two-liquid segment activity coefficient model for solvent system selection in counter-current chromatography separation, J. Chromatogr. A. 1301 (2013) 10–18. https://doi.org/10.1016/j.chroma.2013.05.029.
[69] S.M. Vilas-Boas, I.W. Cordova, K.A. Kurnia, H. Almeida, P.S. Gaschi, J.A.P. Coutinho, S.P. Pinho, O. Ferreira, Distribution of phenolic compounds in biphasic solvent systems: experimental and thermodynamic modelling studies, in: 2021: p. 29. https://iupac.org/event/issp-19/.
[70] A.M. Zissimos, M.H. Abraham, M.C. Barker, K.J. Box, K.Y. Tam, Calculation of Abraham descriptors from solvent–water partition coefficients in four different systems; evaluation of different methods of calculation, J. Chem. Soc. Perkin Trans. 2. 2 (2002) 470–477. https://doi.org/10.1039/b110143a.
[71] M.H. Abraham, R.E. Smith, R. Luchtefeld, A.J. Boorem, R. Luo, W.E. Acree, Prediction of Solubility of Drugs and Other Compounds in Organic Solvents, J. Pharm. Sci. 99 (2010) 1500–1515. https://doi.org/10.1002/jps.21922.
[72] M.H. Abraham, W.E. Acree, Partition coefficients and solubilities of compounds in thewater-Ethanol solvent system, J. Solution Chem. 40 (2011) 1279–1290. https://doi.org/10.1007/s10953-011-9719-x.
[73] C.F. KREWSON, J. NAGHSKI, Some physical properties of rutin., J. Am. Pharm. Assoc. Am. Pharm. Assoc. (Baltim). 41 (1952) 582–587. https://doi.org/10.1002/jps.3030411106.
[74] M.H. Abraham, W.E. Acree, On the solubility of quercetin, J. Mol. Liq. 197 (2014) 157–

159. https://doi.org/10.1016/j.molliq.2014.05.006.
[75] D. Althans, P. Schrader, S. Enders, Solubilisation of quercetin: Comparison of hyperbranched polymer and hydrogel, J. Mol. Liq. 196 (2014) 86–93. https://doi.org/10.1016/j.molliq.2014.03.028.
[76] M.H. Abraham, Scales of Solute Hydrogen-bonding: Their Construction and Application to P hysicoc hem ica I and B ioc hem ica I Processes Michael, Chem. Soc. Rev. 096 (1993) 73–83.
[77] S. Endo, K.U. Goss, Applications of polyparameter linear free energy relationships in environmental chemistry, Environ. Sci. Technol. 48 (2014) 12477–12491. https://doi.org/10.1021/es503369t.
[78] L.M. Sprunger, A. Proctor, W.E. Acree, M.H. Abraham, N. Benjelloun-Dakhama, Correlation and prediction of partition coefficient between the gas phase and water, and the solvents dry methyl acetate, dry and wet ethyl acetate, and dry and wet butyl acetate, Fluid Phase Equilib. 270 (2008) 30–44. https://doi.org/10.1016/j.fluid.2008.06.001.

Appendix A: UNIFAC model equations and information data

The activity coefficient $\gamma_i$of a component i is obtained by the following Equation A.1 [65]:

$$\ln \gamma_i = \ln \gamma_i^{combinatorial} + \ln \gamma_i^{residual} \quad (A.1)$$

The combinatorial part is related to the size and shape of the groups, whereas the residual part is essentially due to energy and surface areas interactions [65]. Equation A.2 expresses the combinatorial and residuals parts of the Equation (1) [65]:

$$\ln\gamma_i^{combinatorial} = \ln\frac{\Phi_i}{x_i} + \frac{z}{2} q_i \frac{\theta_i}{\Phi_i} + \ l_i + \frac{\Phi_i}{x_i} x_j l_j \quad (A.2)$$

Where x_i is the mole fraction of the component i, j is the main group according to UNIFAC group specifications tables, Φ_i is the segment fraction, concept similar to volume fraction, and θ_i is the area fraction of component i [65]. The following equations describe the remaining terms of Equation A.1 [65]:

$$l_i = \frac{z}{2}(r_i - q_i) - (r_i - 1) \quad (A.3)$$

$$\theta_i = \frac{q_i x_i}{\sum_j q_j x_j} \quad (A.4)$$

$$\Phi_i = \frac{r_i x_i}{\sum_j r x_j} \quad (A.5)$$

$$z = 10 \quad (A.6)$$

The summations in Equations A.4 and A.5 are over all components, including component i [65]. The pure component parameters r_i and q_i are obtained in the UNIFAC group specifications tables according to Equations A.7 and A.8 [65]:

$$r_i = \sum_k v_k^{(i)} R_k \quad (A.7)$$

$$q_i = \sum_k v_k^{(i)} Q_k \quad (A.8)$$

$$R_k = \frac{V_{wk}}{15.47} \quad (A.9)$$

$$Q_k = \frac{A_{wk}}{2.5 \times 10^9} \tag{A.10}$$

Where $v_k^{(i)}$ is the number of groups of a type k in molecule i, V_{wk} and A_{wk} are the Van der Waals group volume and surface area, respectively [65].

In the UNIFAC model, the residual part is defined by the following equation:

$$\ln\gamma_i^{residual} = \sum_k v_k^{(i)} (ln\Gamma_k - ln\Gamma_k^{(i)}) \tag{A.11}$$

Where "Γ_k is the group residual activity coefficient and $\Gamma_k^{(i)}$ is the residual activity coefficient of group k in a reference solution containing only molecules of type i" [65].

Equation A.12 defines the group residual activity coefficient for Γ_k and $\Gamma_k^{(i)}$ [65]:

$$\ln\Gamma_k = Q_k \left[1 - \ln\left(\sum_m \theta_m \psi_{mk} \right) - \sum_m \frac{\theta_m \psi_{km}}{\theta_n \psi_{nm}} \right] \tag{A.12}$$

θ_m is the area fraction of a group m and the sums in the Equation A.12 are over all groups. Similar to θ_i, expressed by the Equation A.6, θ_m is calculated accordingly to Equation A.13, where X_m is the mole fraction of group m presents in the mixture [65]:

$$\theta_m = \frac{Q_m X_m}{\sum_n Q_n X_n} \tag{A.13}$$

In Equation A.14, ψ is defined as the group-interaction parameter between two groups [65]:

$$\psi_{mn} = \exp\left(-\frac{a_{mn} - a_{nm}}{RT} \right) \tag{A.14}$$

Where a_{mn} is the interaction energy between groups m and n [65]. It is important to notice that a_{mn} is different than a_{nm}.

Table A.1. Logarithmic partition coefficients calculated using the UNIFAC model of vanillin, ferulic acid and hesperetin logarithmic at 22ºC for the HEMWat system [13].

Solvent system number	Vanillin	Ferulic acid	Hesperetin
-7	-3.80	-3.37	-7.61
-6	-3.62	-3.08	-7.24
-5	-3.33	-2.75	-6.64
-4	-3.15	-2.53	-6.34
-3	-2.75	-2.14	-5.48
-2	-2.91	-2.25	-5.93
-1	-2.47	-1.81	-4.97
0	-1.98	-1.32	-3.91
1	-1.53	-0.88	-2.97
2	-1.04	-0.40	-1.93
3	-1.26	-0.55	-2.48
4	-0.78	-0.07	-1.48
5	-0.65	0.10	-1.31
6	-0.23	0.51	-0.54
7	0.16	0.86	0.15
8	0.60	1.25	0.93

Table A.2. Logarithmic partition coefficients calculated using the UNIFAC model at 22ºC for n-heptane/ethyl acetate/1-butanol/methanol/water and limonene/ethyl acetate/1-butanol/ethanol/water solvent systems [16].

Solvent system number	Arizona system	n-heptane/ethyl acetate/1-butanol/methanol/water			limonene/ethyl acetate/1-butanol/ethanol/water		
		Vanillin	Ferulic acid	Hesperetin	Vanillin	Ferulic acid	Hesperetin
1		1.83	2.26	3.31	1.83	2.26	3.31
2		1.82	2.32	3.35	1.82	2.32	3.35
3		1.81	2.45	3.37	1.81	2.45	3.37
4		1.77	2.41	3.34	1.77	2.41	3.34
5		1.55	2.25	2.90	1.55	2.25	2.90
6	**A**	0.74	1.42	1.23	0.74	1.42	1.23
7	**B**	0.37	1.04	0.50	0.31	0.89	0.30
8	**C**	0.14	0.82	0.06	0.01	0.52	-0.33
9	**D**	-0.12	0.56	-0.46	-0.30	0.15	-1.00
10	**F**	-0.18	0.51	-0.55	-0.38	0.04	-1.16
11	**G**	-0.34	0.35	-0.85	-0.58	-0.20	-1.58
12	**H**	-0.59	0.10	-1.33	-0.89	-0.57	-2.23
13	**J**	-0.75	-0.05	-1.62	-1.08	-0.79	-2.63
14	**K**	-1.06	-0.39	-2.26	-1.42	-1.20	-3.36
15	**L**	-1.48	-0.84	-3.06	-1.90	-1.74	-4.36
16	**M**	-1.90	-1.29	-3.89	-2.35	-2.25	-5.31
17	**N**	-2.23	-1.64	-4.55	-2.71	-2.65	-6.08
18	**P**	-2.65	-2.09	-5.42	-3.10	-3.06	-6.93
19	**Q**	-3.09	-2.55	-6.31	-3.53	-3.50	-7.83
20	**R**	-3.51	-2.96	-7.15	-3.91	-3.88	-8.64
21	**S**	-3.80	-3.26	-7.74	-4.15	-4.11	-9.15
22	**T**	-3.96	-3.42	-8.07	-4.28	-4.23	-9.42
23	**U**	-4.13	-3.59	-8.39	-4.41	-4.32	-9.66
24	**V**	-4.26	-3.74	-8.65	-4.51	-4.39	-9.85
25	**W**	-4.25	-3.73	-8.61	-4.47	-4.34	-9.75
26	**X**	-4.35	-3.88	-8.81	-4.46	-4.25	-9.67
27	**Y**	-4.15	-3.78	-8.41	-4.09	-3.80	-8.83
28	**Z**	-3.58	-3.43	-7.34	-3.10	-2.68	-6.65

Appendix B: Compositions of the upper and lower phases of methanol/ethyl acetate/heptane/water/butanol, ethanol/ethyl acetate/limonene/water/butanol and HEMWat solvent systems

Table B.1. Upper and lower phase composition in mole fraction of HEMWat solvent system family [13].

Solvent system	Upper				Lower			
	Hexane	**Ethyl acetate**	**Methanol**	**Water**	**Hexane**	**Ethyl acetate**	**Methanol**	**Water**
-7	0.8989	0.04	0.0604	0.0007	0.0271	0.0289	0.7467	0.1973
-6	0.8446	0.0941	0.0579	0.0034	0.0081	0.0458	0.5752	0.3709
-5	0.77	0.1733	0.052	0.0047	0.0043	0.0593	0.4642	0.4722
-4	0.7388	0.2095	0.0464	0.0053	0.0018	0.0463	0.371	0.5809
-3	0.6488	0.2754	0.065	0.0108	0.0023	0.0612	0.3639	0.5726
-2	0.7093	0.2435	0.0413	0.0059	0.0003	0.0333	0.2927	0.6737
-1	0.6034	0.3242	0.0605	0.0119	0.0005	0.0433	0.2868	0.6694
0	0.4883	0.3971	0.0858	0.0288	0.0009	0.054	0.2863	0.6588
1	0.3776	0.457	0.1142	0.0512	0.0011	0.064	0.2798	0.6551
2	0.2557	0.4999	0.1521	0.0923	0.0015	0.0746	0.2712	0.6527
3	0.342	0.5086	0.0962	0.0532	0.0002	0.04	0.216	0.7438
4	0.2292	0.5601	0.1206	0.0901	0.0002	0.0422	0.2043	0.7533
5	0.2252	0.6161	0.0831	0.0756	0	0.0265	0.145	0.8285
6	0.1465	0.7031	0.059	0.0914	0	0.0211	0.0927	0.8862
7	0.0718	0.793	0.0297	0.1055	0	0.019	0.0451	0.9359
8	0	0.8741	0	0.1259	0	0.0173	0	0.9827

Table B.2. Upper and lower phase composition in mole fraction of n-heptane/ethyl acetate/1-butanol/methanol/water solvent system family [16]

Solvent system	Upper					Lower				
	Methanol	**Ethyl acetate**	**Heptane**	**Water**	**Butanol**	**Methanol**	**Ethyl acetate**	**Heptane**	**Water**	**Butanol**
1	0	0	0	0.4939	0.5061	0	0	0	0.9811	0.0189
2	0	0.086	0	0.5167	0.3973	0	0.0046	0	0.9791	0.0163
3	0	0.1859	0	0.5107	0.3034	0	0.0086	0	0.978	0.0134
4	0	0.3186	0	0.4614	0.22	0	0.0117	0	0.9784	0.0099
5	0	0.5249	0	0.3509	0.1242	0	0.0138	0	0.9803	0.0059
6	0	0.8508	0	0.1492	0	0	0.0155	0	0.9845	0
7	0.0168	0.8482	0.0313	0.1037	0	0.0226	0.0169	0	0.9605	0
8	0.0307	0.8181	0.0648	0.0864	0	0.0461	0.0177	0	0.9362	0
9	0.0426	0.7974	0.0979	0.0621	0	0.0676	0.018	0	0.9144	0
10	0.0481	0.7779	0.1097	0.0643	0	0.0781	0.0192	0	0.9027	0
11	0.0574	0.7484	0.1367	0.0575	0	0.0967	0.0206	0	0.8827	0
12	0.0683	0.7038	0.18	0.0479	0	0.1242	0.0241	0	0.8517	0
13	0.0778	0.6701	0.2097	0.0424	0	0.1455	0.0252	0	0.8293	0
14	0.0836	0.6223	0.2635	0.0306	0	0.1694	0.0317	0	0.7989	0
15	0.0782	0.5522	0.3443	0.0253	0	0.2136	0.0406	0.0001	0.7457	0
16	0.0716	0.4789	0.4295	0.02	0	0.2534	0.0503	0.0002	0.6961	0
17	0.0695	0.4054	0.5064	0.0187	0	0.2963	0.057	0.0005	0.6462	0
18	0.0574	0.3352	0.5955	0.0119	0	0.3243	0.0611	0.0007	0.6139	0
19	0.0452	0.2624	0.6852	0.0072	0	0.3769	0.0635	0.0011	0.5585	0
20	0.0342	0.1922	0.769	0.0046	0	0.4396	0.0609	0.0016	0.4979	0
21	0.0324	0.139	0.8256	0.003	0	0.4997	0.0566	0.0025	0.4412	0
22	0.0302	0.1125	0.855	0.0023	0	0.5414	0.0535	0.0029	0.4022	0
23	0.0296	0.0834	0.8855	0.0015	0	0.6059	0.0454	0.0044	0.3443	0
24	0.024	0.0644	0.9101	0.0015	0	0.6635	0.041	0.0067	0.2888	0
25	0.0297	0.0569	0.9119	0.0015	0	0.695	0.0367	0.0081	0.2602	0
26	0.031	0.0352	0.933	0.0008	0	0.7724	0.027	0.0137	0.1869	0
27	0.0506	0.019	0.9296	0.0008	0	0.8555	0.0149	0.0286	0.101	0
28	0.0845	0	0.9155	0	0	0.9133	0	0.0867	0	0

Table B.3. Upper and lower phase composition in mole fraction of limonene/ethyl acetate/1-butanol/ethanol/water solvent system family [16]

Solvent system	Upper					Lower				
	Ethanol	**Ethyl acetate**	**Limonene**	**Water**	**Butanol**	**Ethanol**	**Ethyl acetate**	**Limonene**	**Water**	**Butanol**
1	0	0	0	0.4939	0.5061	0	0	0	0.9811	0.0189
2	0	0.086	0	0.5167	0.3973	0	0.0046	0	0.9791	0.0163
3	0	0.1859	0	0.5107	0.3034	0	0.0086	0	0.978	0.0134
4	0	0.3186	0	0.4614	0.22	0	0.0117	0	0.9784	0.0099
5	0	0.5249	0	0.3509	0.1242	0	0.0138	0	0.9803	0.0059
6	0	0.8508	0	0.1492	0	0	0.0155	0	0.9845	0
7	0.0117	0.8552	0.0286	0.1045	0	0.0158	0.017	0	0.9672	0
8	0.0216	0.831	0.0596	0.0878	0	0.0324	0.0179	0	0.9497	0
9	0.0302	0.8157	0.0906	0.0635	0	0.0478	0.0184	0	0.9338	0
10	0.0342	0.7981	0.1018	0.0659	0	0.0554	0.0197	0	0.9249	0
11	0.041	0.7721	0.1276	0.0593	0	0.0691	0.0213	0	0.9096	0
12	0.0492	0.7317	0.1693	0.0498	0	0.0895	0.025	0	0.8855	0
13	0.0564	0.7009	0.1983	0.0444	0	0.1055	0.0264	0	0.8681	0
14	0.061	0.6557	0.2511	0.0322	0	0.1238	0.0335	0	0.8427	0
15	0.0574	0.5856	0.3302	0.0268	0	0.1584	0.0434	0.0001	0.7981	0
16	0.053	0.511	0.4146	0.0214	0	0.1904	0.0546	0.0002	0.7548	0
17	0.0517	0.4358	0.4924	0.0201	0	0.2259	0.0626	0.0005	0.711	0
18	0.043	0.3622	0.582	0.0128	0	0.2495	0.0679	0.0007	0.6819	0
19	0.034	0.285	0.6731	0.0079	0	0.2954	0.0718	0.0011	0.6317	0
20	0.0258	0.2099	0.7593	0.005	0	0.3522	0.0704	0.0017	0.5757	0
21	0.0246	0.1525	0.8196	0.0033	0	0.4091	0.0668	0.0027	0.5214	0
22	0.023	0.1239	0.8506	0.0025	0	0.45	0.0642	0.0032	0.4826	0
23	0.0227	0.092	0.8836	0.0017	0	0.5161	0.0557	0.0049	0.4233	0
24	0.0183	0.0711	0.9089	0.0017	0	0.5777	0.0516	0.0076	0.3631	0
25	0.0228	0.0629	0.9126	0.0017	0	0.6128	0.0467	0.0094	0.3311	0
26	0.0238	0.0391	0.9362	0.0009	0	0.7028	0.0354	0.0163	0.2455	0
27	0.0391	0.0212	0.9388	0.0009	0	0.807	0.0203	0.0352	0.1375	0
28	0.066	0	0.934	0	0	0.8898	0	0.1102	0	0

Appendix C: Abraham solvation model information data

Table C.1. Experimental and calculated values from partitions (*P*) and solubilities (*S*) of vanillin as log *P* using descriptors from literature [71].

Type of data	Solvent	Reference		log *P*	
		Experimental data	Solvent descriptors	Exp.	[71]
P	**1-octanol/water**	This work	[76]	1.29	1.25
P	**1-Butanol/water**		[77]	1.19	1.34
P	**Ethyl acetate/water**		[21]	1.43	1.31
P	**Heptane/methanol**		[78]	-1.48	-1.19
P	**Butyl acetate/water**	[23]	[71]	1.45	1.51
P	**Chloroform/water**			1.31	1.39
P	**Hexane/water**			-0.70	-0.75
P	**Diethyl ether/water**	[48]		0.80	0.93
P	**Benzene/water**			0.57	0.81
P	**1-pentanol/water**			1.35	1.36
P	**Dichloromethane/ Water**			1.28	1.34
P	**Toluene/ water**			0.61	0.66
P	**Methyl isobutyl ketone/water**			1.47	1.42
P	**Methyl *tert*-butyl ether/Water**	[42]		0.99	0.85
S	**Ethanol**	[22]		1.27	1.40
S	**Ethylene glycol**			1.27	1.17
S	**2-propanol**			1.09	1.16
S	**Ethyl acetate**			1.26	1.26
S	**1-Butanol**			1.01	1.08
S	**2-Butanol**			1.14	1.25

Table C.2. Experimental and calculated values for partitions (*P*) and solubilities (*S*) of rutin as log *P* using descriptors reported in the literature [72] and obtained in this work.

	Solvent	Reference		log *P*			
		Experimental data	Solvent descriptors	Exp.	[72]	This work	
P	**1-Octanol/water**	This work	[76]	-0.01	2.23	0.70	Correlated
P	**1-Butanol/water**		[77]	0.73	2.51	1.81	
P	**Ethyl acetate/water**		[21]	-0.76	-0.79	-0.80	
P	**Heptane/methanol**		[78]	-3.18	-5.78	-3.58	
S	**Ethanol**	[29]	[74]	2.35	2.08	0.06	
S	**Methanol**	[30]		2.36	2.26	2.36	
S	**1-Propanol**		[71]	1.87	1.78	1.36	
S	**2-Propanol**			1.80	1.78	0.99	
S	**Ethyl acetate**			1.57	-2.75	0.41	
S	**1-Butanol**			1.07	1.11	1.22	
S	**Acetone**			1.05	-2.23	1.05	
S	**30% (v/v) Ethanol**	[29]	[25]	1.20	0.99	0.51	Predicted
S	**40% (v/v) Ethanol**			1.12	1.24	0.55	
S	**50% (v/v) Ethanol**			1.01	1.49	0.38	
S	**60% (v/v) Ethanol**			0.98	1.64	0.24	
S	**70% (v/v) Ethanol**			1.09	1.64	0.04	
S	**80% (v/v) Ethanol**			1.25	1.63	0.07	
S	**90% (v/v)Ethanol**			1.39	1.59	-0.14	
S	**95% (v/v) Ethanol**			1.49	1.98	-0.05	
S	**96% (v/v) Ethanol**			1.52	2.04	-0.04	

Table C.3. Experimental and calculated values for partitions (*P*) and solubilities (*S*) of ferulic acid as log *P* using descriptors reported in the literature [31] and obtained in this work.

	Solvent	Reference		log *P*			
		Experimental data	Solvent descriptors	Exp.	[32]	This work	
P	**1-octanol/water**	This work	[76]	1.32	2.14	1.77	Correlated
P	**1-Butanol/water**		[77]	1.48	2.02	1.77	
P	**Ethyl acetate/water**		[21]	1.6	1.96	1.70	
P	**Heptane/methanol**		[78]	-2.12	-1.02	-1.21	
S	**Methanol**	[32]	[74]	2.40	2.19	1.92	
S	**Ethanol**			2.15	2.19	1.82	
S	**Isopropanol**		[71]	1.89	1.94	1.56	
S	**1-Propanol**			1.88	2.01	1.61	
S	**Ethyl acetate**			2.03	1.98	1.89	
S	**Acetonitrile**			1.63	1.79	1.59	
S	**2-Butanone**			1.45	1.40	1.45	
S	**Ethylene glycol**	[31]		2.07	1.94	1.69	Predicted
S	**1-butanol**			1.75	1.94	1.52	
S	**2-butanol**			1.77	2.02	1.67	
S	**dimethyl sulfoxide (DMSO)**			2.36	2.37	2.53	
S	**Dichloromethane**	[33]		2.41	1.96	1.87	
S	**Chloroform**			2.19	2.18	2.04	
S	**Methyl acetate**			1.81	1.91	1.81	
S	**Butyl acetate**			1.42	1.71	1.49	

Table C.4. Experimental and calculated values for partitions (*P*) and solubilities (*S*) of quercetin acid as log *P* using descriptors reported in the literature [74] and obtained in this work.

	Solvent	Reference		log *P*			
		Experimental data	Solvent descriptors	Exp.	[74]	This work	
P	**1-octanol/water**	This work	[76]	2.21	1.98	2.48	Correlated
P	**1-Butanol/water**		[77]	2.42	1.89	2.42	
P	**Ethyl acetate/water**		[21]	2.33	1.16	2.29	
P	**Heptane/methanol**		[78]	-3.01	-3.89	-3.01	
S	**1-propanol**	[34]	[71]	2.64	2.12	2.11	
S	**1-butanol**			2.22	1.53	1.73	
S	**Ethyl acetate**			1.68	0.27	1.76	
S	**Ethanol**	[74]	[74]	2.37	2.37	2.50	
S	**Methanol**			2.26	2.39	2.66	Predicted
S	**Butyl acetate**			1.08	-1.21	1.40	

Table C.5. Experimental and calculated values for partitions (*P*) and solubilities (*S*) of hesperetin acid as log *P* using descriptors reported in the literature [25] and obtained in this work.

	Solvent	Reference		log *P*			
		Experimental data	Solvent descriptors	Exp.	[25]	This work	
P	**1-octanol/water**	This work	[76]	2.39	3.34	2.98	Correlated
P	**1-Butanol/water**		[77]	2.27	2.87	2.64	
P	**Ethyl acetate/water**		[21]	2.34	3.48	3.24	
P	**Heptane/methanol**		[78]	-2.2	-2.40	-2.44	
S	**acetone**	[36]	[71]	4.20	4.20	4.02	
S	**acetonitrile**			3.30	3.30	3.31	
S	**ethanol**		[74]	3.59	3.59	3.15	
S	**ethyl acetate**		[71]	3.56	3.56	3.40	
S	**methanol**		[74]	3.68	3.68	3.41	
S	**1-octanol**	[38]		2.95	2.95	2.79	
S	**1-Butanol**	[37]	[71]	3.28	3.28	2.64	
S	**20% (v/v) Methanol**	[25]	[25]	1.04	1.04	0.64	Predicted
S	**30% (v/v) Methanol**			1.32	1.32	1.05	
S	**40% (v/v) Methanol**			1.67	1.67	1.48	
S	**50% (v/v)Methanol**			2.07	2.07	1.93	
S	**60% (v/v) Methanol**			2.50	2.50	2.36	
S	**70% (v/v) Methanol**			2.91	2.91	2.76	
S	**80% (v/v) Methanol**			3.24	3.24	3.08	
S	**90% (v/v) Methanol**			3.44	3.44	3.30	
S	**95% (v/v) Methanol**			3.56	3.56	3.42	
S	**30% (v/v) Ethanol**			1.48	1.48	1.52	
S	**40% (v/v) Ethanol**			2.05	2.05	2.06	
S	**50% (v/v) Ethanol**			2.66	2.66	2.51	
S	**60% (v/v) Ethanol**			2.93	2.93	2.88	
S	**70% (v/v) Ethanol**			3.05	3.05	3.14	
S	**80% (v/v) Ethanol**			3.42	3.42	3.33	
S	**90% (v/v)Ethanol**			3.61	3.61	3.40	
S	**95% (v/v) Ethanol**			3.61	3.61	3.28	
S	**96% (v/v) Ethanol**			3.60	3.60	3.26	

Table C.6. Abraham coefficients for wet and dry solvents used in this work.

	Solvent	c	e	s	a	b	v	Reference
d	**1-butanol**	0.152	0.438	-1.177	0.096	-3.919	4.122	[71]
d	**1-octanol**	-0.034	0.489	-1.044	-0.024	-4.235	4.218	[71]
d	**1-propanol**	0.148	0.436	-1.098	0.389	-3.893	4.036	[71]
d	**2-butanol**	0.194	0.383	-0.956	0.134	-3.606	3.829	[71]
d	**2-Butanone**	0.246	0.256	-0.08	-0.767	-4.855	4.148	[71]
d	**2-propanol**	0.099	0.344	-1.049	0.406	-3.827	4.033	[71]
d	**20% (v/v) Methanol**	0.022	0.142	-0.138	0.088	-0.574	0.559	[25]
d	**30% (v/v) Methanol**	0.016	0.187	-0.172	0.165	-0.953	0.898	[25]
d	**40% (v/v) Methanol**	0.02	0.222	-0.205	0.218	-1.329	1.259	[25]
d	**50% (v/v)Methanol**	0.023	0.223	-0.222	0.264	-1.747	1.662	[25]
d	**60% (v/v) Methanol**	0.053	0.207	-0.238	0.272	-2.157	2.073	[25]
d	**70% (v/v) Methanol**	0.098	0.192	-0.26	0.266	-2.558	2.474	[25]
d	**80% (v/v) Methanol**	0.172	0.197	-0.319	0.241	-2.912	2.842	[25]
d	**90% (v/v) Methanol**	0.258	0.25	-0.452	0.229	-3.206	3.175	[25]
d	**95% (v/v) Methanol**	0.27	0.278	-0.52	0.23	-3.368	3.365	[25]
d	**30% (v/v) Ethanol**	-0.269	0.107	-0.098	0.133	-1.316	1.414	[25]
d	**40% (v/v) Ethanol**	-0.221	0.131	-0.159	0.171	-1.809	1.918	[25]
d	**50% (v/v) Ethanol**	-0.142	0.124	-0.252	0.251	-2.275	2.415	[25]
d	**60% (v/v) Ethanol**	-0.04	0.138	-0.335	0.293	-2.675	2.812	[25]
d	**70% (v/v) Ethanol**	0.063	0.085	-0.368	0.311	-2.936	3.102	[25]
d	**80% (v/v) Ethanol**	0.172	0.175	-0.465	0.26	-3.212	3.323	[25]
d	**90% (v/v)Ethanol**	0.243	0.213	-0.575	0.262	-3.45	3.545	[25]
d	**95% (v/v) Ethanol**	0.239	0.328	-0.795	0.294	-3.514	3.697	[25]
d	**96% (v/v) Ethanol**	0.238	0.353	-0.833	0.297	-3.533	3.724	[71]
d	**Acetonitrile**	0.413	0.077	0.326	-1.566	-4.391	3.364	[71]
d	**Butyl acetate**	0.248	0.356	-0.501	-0.867	-4.973	4.281	[71]
d	**Chloroform**	0.191	0.105	-0.403	-3.112	-3.514	4.395	[71]

	Solvent	c	e	s	a	b	v	Reference
d	**Dimethyl sulfoxide (DMSO)**	-0.194	0.327	0.791	1.26	-4.54	3.361	[71]
d	**Ethanol**	0.222	0.471	-1.035	0.326	-3.596	3.857	[74]
d	**Ethyl acetate**	0.328	0.369	-0.446	-0.7	-4.904	4.15	[71]
d	**Ethylene glycol**	-0.27	0.578	-0.511	0.715	-2.619	2.729	[71]
d	**Isopropanol**	0.099	0.344	-1.049	0.406	-3.827	4.033	[71]
d	**Methanol**	0.276	0.334	-0.714	0.243	-3.32	3.549	[74]
d	**Methyl acetate**	0.351	0.223	-0.15	-1.035	-4.527	3.972	[71]
w	**1-Butanol/water**	0.369	0.426	-0.719	-0.091	-2.246	2.689	[77]
w	**1-octanol/water**	0.088	0.562	-1.054	0.034	-3.46	3.814	[76]
w	**1-pentanol/water**	0.185	0.367	-0.732	0.105	-3.1	3.395	[71]
w	**Benzene/water**	0.142	0.464	-0.588	-3.099	-4.625	4.491	[71]
w	**Butyl acetate/water**	-0.475	0.428	-0.094	-0.241	-4.151	4.046	[71]
w	**Chroloform/water**	0.191	0.105	-0.403	-3.112	-3.514	4.395	[71]
w	**Diethyl eter/water**	0.248	0.561	-1.016	-0.226	-4.553	4.075	[71]
w	**Ethyl acetate/water**	0.441	0.591	-0.669	-0.325	-4.261	3.668	[71]
w	**Heptane/methanol**	-0.056	0.164	-0.62	-1.337	-0.957	0.507	[78]
w	**Hexane/water**	0.361	0.579	-1.723	-3.599	-4.764	4.344	[71]
w	**Methyl isobutyl ketone/water**	0.383	0.801	-0.831	-0.121	-4.441	3.876	[71]
w	**Toluene/ water**	0.143	0.527	-0.72	-3.01	-4.824	4.545	[71]
w/d	**Methyl t butyl ether**	0.376	0.264	-0.788	-1.078	-5.03	4.41	[71]
w/d	**Dichloromethane**	0.319	0.102	-0.187	-3.058	-4.09	4.324	[71]
w/d	**Methyl ter-butyl ether/Water**	0.376	0.264	-0.788	-1.078	-5.03	4.41	[71]

Appendix D: Calibration curves in pure ethanol and ethanol/water 50/50 (m/m) in UV-Vis spectroscopy for vanillin, rutin, ferulic acid, quercetin and hesperetin

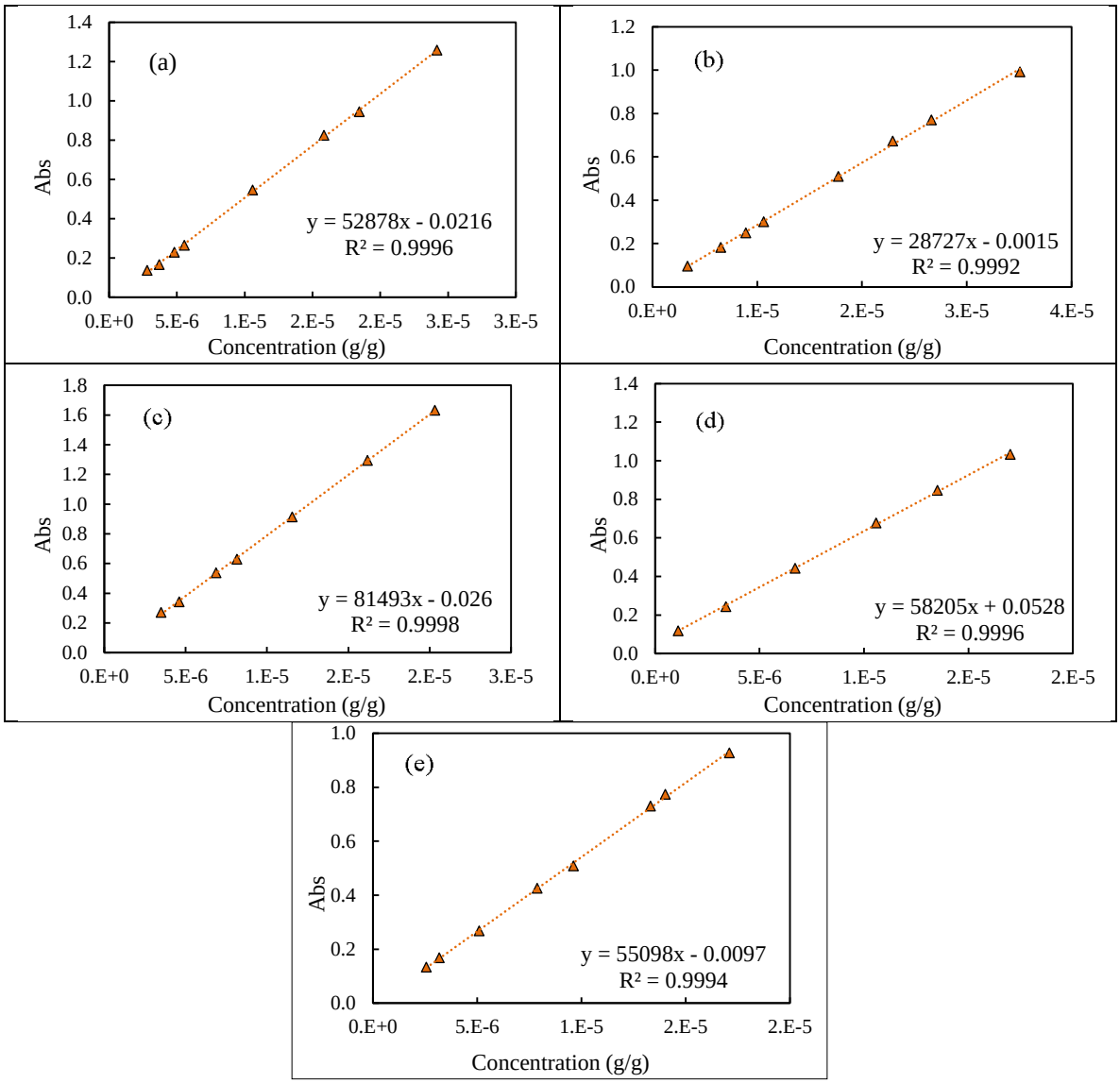

Figure D.1. Calibration curves in pure ethanol in UV-Vis spectroscopy for (a) vanillin, (b) rutin, (c) ferulic acid, (d) quercetin and (e) hesperetin. Concentration in g of solute/ g of solution.

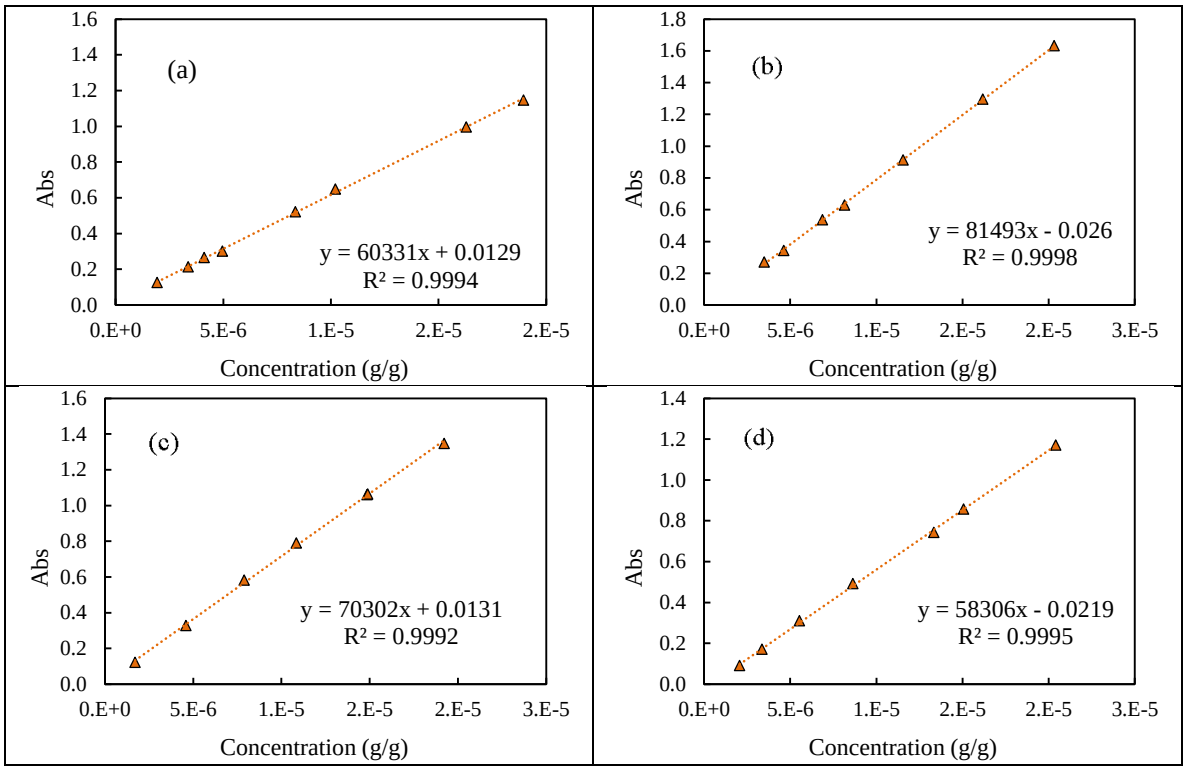

Figure D.2. Calibration curves in ethanol/water 50/50 (m/m) in UV-Vis spectroscopy for (a) vanillin, (b) ferulic acid, (c) quercetin and (d) hesperetin. Concentration in g of solute/ g of solution.