INSTITUTO POLITÉCNICO DE BRAGANÇA
Escola Superior de Tecnologia e Gestão

Studies on the Solubility of Phenolic Compounds

Sérgio Antonio Mendes Vilas Boas

Dissertation presented to

Escola Superior de Tecnologia e Gestão

Instituto Politécnico de Bragança

In order to obtain Master's Degree in

Chemical Engineering

This work was supervised by

Simão Pedro Almeida Pinho

Maria Olga Amorim e Sá Ferreira

Luciano Fernandes

July, 2017

“Simplicity is the ultimate sophistication.”

Clare Boothe Luce

Acknowledgments

I have never imagined that I would grow so much since I arrived in Portugal. From professional learnings to personal improvement, I only have to thank this country and all the people that I have met so far and those who have supported me during the time of my stay.

First of all, I would like to give a special thanks to my supervisors from IPB, professors Simão Pinho and Olga Ferreira, for the opportunity to perform this work and for all their dedication, patience, encouragement, guidance, open discussions and friendship. Sometimes they have been more than supervisors, but truly friends.

I also express my gratitude to professor Luciano Fernandes, my supervisor from UTFPR, for his support, availability, enlightening discussions and friendship. His support really encouraged me to come to Portugal to develop this project.

I cannot let to thank professors Juliana Peitrobelli, Priscila Leite and Erica Lovo for offering me the opportunity to attend the Double Degree Program formed by the partnership of UTFPR and IPB. I also thank all the professors, from UTFPR and IPB, who have been part of my journey helping me to finish this work.

Special thanks to Laboratory of Separation and Reaction Engineering (LSRE) for providing me all the facilities and equipment to reach the outcomes of this work.

I am immensely grateful to my parents, Antonio Sergio and Elaine, and to my brother, Carlos Eduardo, for all their understanding, care, incentive and confidence since I decided to come to Portugal. I cannot let to mention that without their support, I couldn't have become who I am today. Thank you for believing in me!

I also thank to my friends from Brazil, who always supported to fight for my dreams, even being far away, and to the friends that I made in Portugal. From the bottom of my heart, thank you for contributing for who I am and for all the things I have accomplished until now.

To everyone who have contributed directly or indirectly to accomplish this dissertation, thank you very much!

Professors Simão Pinho and Olga Ferreira, supervisors of this Master thesis, are members of the Associate Laboratory LSRE-LCM financed by Project POCI-01-0145-FEDER-006984 funded by FEDER through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI) – and by national funds through FCT - Fundação para a Ciência e a Tecnologia.

This work is a result of project "AIProcMat@N2020 - Advanced Industrial Processes and Materials for a Sustainable Northern Region of Portugal 2020", with the reference NORTE-01-0145-FEDER-000006, supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the Portugal 2020 Partnership Agreement, through the European Regional Development Fund (ERDF) and of Project POCI-01-0145-FEDER-006984 – Associate Laboratory LSRE-LCM funded by ERDF through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI) – and by national funds through FCT - Fundação para a Ciência e a Tecnologia.

Abstract

Phenolic compounds generally act as antioxidant and free radical scavengers, having several practical applications in the pharmaceutical, food, oil and chemical industrial processes. Among those compounds, phenolic acids represent a group that is widely present in some natural products, showing interesting properties, such as preventers of some degenerative diseases, with application in the pharmaceutical industry. Furthermore, solubility studies play a key role to obtain a significant yield and a representative product, being an important parameter for the development of new drugs as well as the optimization of already existent processes. In this context, the main objective of this work is to measure the solubility of gallic, protocatechuic, gentisic and α-resorcylic acids in water and organic solvents (methanol, ethanol, 1-propanol, isopropanol, 2-butanone, ethyl acetate, dimethylformamide and acetonitrile) at 298.15 and 313.15 K and to employ the NRTL-SAC thermodynamic model coupled to the Reference Solvent Approach (RSA) to describe the solubility data.

The experimental methodology was the shake-flask method coupled to the gravimetric method and, in general, the results obtained were satisfactorily consistent with the information available in literature for gallic and protocatechuic acids. For gentisic and α-resorcylic acids, no solubility studies were found at the analyzed temperatures until now. Melting points and enthalpies of fusion of the selected phenolic acids were also measured via Differential Scanning Calorimetry (DSC).

Finally, the NRTL-SAC segment descriptors were obtained by fitting the solubility data in seven solvents, obtaining average relative deviations (ARD) between 25 and 34%. The model was then applied to predict the solubility in 1-propanol and dimethylformamide and the ARD% were 70 and 78%, respectively. Those values are satisfactory for semi-predictive models, using a limited set of solvents, showing that the NRTL-SAC is adequate to model binary systems containing the selected phenolic acids.

Keywords: solubility, phenolic acids, NRTL-SAC.

Resumo

Compostos fenólicos geralmente agem como antioxidantes e sequestradores de radicais livres, possuindo diversas aplicações práticas, tais como em processos farmacêuticos, alimentícios, na indústria de petróleo e na indústria química. Dentre esses compostos, os ácidos fenólicos representam um grupo amplamente presente in alguns produtos naturais, apresentando propriedades interessantes, como preventivos de algumas doenças degenerativas, o que os torna amplamente utilizados na indústria farmacêutica. Além disso, estudos de solubilidade desempenham um papel chave para a obtenção de rendimentos significativos e produtos representativos na indústria farmacêutica, sendo a solubilidade um parâmetro primordial no desenvolvimento de novos medicamentos bem como na otimização de processos já implementados. Nesse contexto, o principal objetivo deste trabalho é a medição de solubilidade dos ácidos gálico, protocatechuico, gentísico e α-resorcílico em água e em solventes orgânicos (metanol, etanol, 1-propanol, 2-propanol, 2-butanona, acetato de etilo, acetonitrilo e dimetilformamida) a 298,15 e 313,15 K e aplicar o modelo termodinâmico NRTL-SAC combinado com a abordagem do Solvente Referência para descrever os dados de solubilidade.

A metodologia experimental utilizada foi o método do frasco agitado combinado com o método gravimétrico e, em geral, os resultados obtidos foram consistentes com a informação disponível na literatura para os ácidos gálico e protocatechuico. Em relação aos ácidos gentísico e α-resorcílico, nenhum estudo de solubilidade foi encontrado nas temperaturas analisadas até o momento. Pontos e entalpias de fusão dos compostos selecionados também foram medidos por Calorimetria Diferencial de Varrimento (DSC).

Finalmente, os descritores de segmentos NRTL-SAC forram obtidos através de ajuste de dados de solubilidade em sete solventes, obtendo-se um erro relativo médio (ARD) entre 25 e 34%. O modelo foi então aplicado na previsão da solubilidade em 1-propanol e em dimetilformamida e os ARD% foram de 70 e 78%, respectivamente. Esses valores são satisfatórios para modelos semipreditivos, com base em um pequeno conjunto de solventes, o que indica o modelo NRTL-SAC como adequado para modelar sistemas binários contendo os ácidos fenólicos selecionados.

Palavras-chave: solubilidade, ácidos fenólicos, NRTL-SAC.

Table of Contents

List of Symbols and Acronyms

List of Symbols

LogP	Partial Coefficient
M	Molar Mass
pH	Potential of Hydrogen
R	Ideal Gas Constant
Ref	Referent Solvent
S	Mass Fraction Solubility
T	Absolute Temperature
T_m	Melting Point Temperature (K)
x_i^{calc}	Calculated Mole Fraction Solubility
x_i^{exp}	Experimental Mole Fraction Solubility
α	Randomness Factor
γ_I	Activity Coefficient of species *I*
ΔH_{fus}	Enthalpy of Fusion (kJ/mol)
ΔS_{fus}	Entropy of Fusion (J/mol K)

List of Acronyms

ARD	Average Relative Deviation
CAS	Chemical Abstracts Number
DMF	Dimethylformamide
DSC	Differential Scanning Colorimetry

HPLC	High-Performance Liquid Chromatography
IPB	Polytechnic Institute of Bragança
GA	Gallic Acid
GEA	Gentisic Acid
LLE	Liquid-Liquid Equilibrium
LQA	Laboratory of Analytic Chemistry
MATLAB	Matrix Laboratory Computing Environment
MOSCED	Modified Separation of Cohesive Energy Density Model
NRTL	Nonrandom Two Liquid
NRTL-SAC	Nonrandom Two-Liquid Segment Activity Coefficient
PCA	Protocatechuic Acid
RA	α-Resorcylic Acid
RSA	Reference Solvent Approach
TOC	Total Organic Carbon
UTFPR	Federal Technological University of Parana
UNIFAC	Universal Functional-Group Activity Coefficients
UV-Vis	Ultraviolet Visible Radiation
VLE	Vapor-Liquid Equilibrium

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Importance and Objectives

Solubility is a fundamental physical property for the design of processes to extract, separate, concentrate, and purify a given target species. In particular, the solubility of phenolic compounds in water and organic solvents plays an important role in the design of separation processes such as extraction, precipitation or crystallization in the food, pharmaceutical, and cosmetic industries.

Previous work, carried out in our research group, was focused on the solubility of some natural phenolic compounds in water (Mota et al. 2008) and the solubility of flavonoids in pure organic solvents (Ferreira & Pinho 2012) or mixed solvents (Ferreira et al. 2013). In this context, the main objective of this master thesis is to extend those studies to a group of phenolic acids (gallic acid, protocatechuic acid, gentisic acid and α-resorcylic acid) by establishing an experimental work plan to measure their solubility in water and organic solvents (methanol, ethanol, 1-propanol, 2-propanol, butanone, ethyl acetate, dimethylformamide and acetonitrile) at 298.15 and 313.15 K and by applying the NRTL-SAC thermodynamic model to describe the experimental data.

1.2 Contents

Chapter 2 starts with a brief description of the chemical and biological properties of the phenolic compounds selected in this work, considering also their current scientific and industrial applications. The most common experimental methods to measure the solubility of solids in liquids are also described. Special attention was given to the traditional shake-flask method, which was applied to perform the experimental measurements. Moreover, a literature review

focusing on solubility measurements of the four compounds evaluated in this work was also performed. This chapter finishes with the presentation of the main thermodynamic models generally used to describe the low pressure solid-liquid equilibria, their applications, range and limitations with particular emphasis given to the NRTL-SAC model.

The experimental materials and methods are described in Chapter 3. This chapter also contains the experimental solubility results as well as the melting properties measurements. Chapter 4 is dedicated to the thermodynamic modeling of the solubility results presented in Chapter 3, by applying the NRTL-SAC model. Finally, in Chapter 5, the main conclusions are summarized and some suggestions for future work are pointed out.

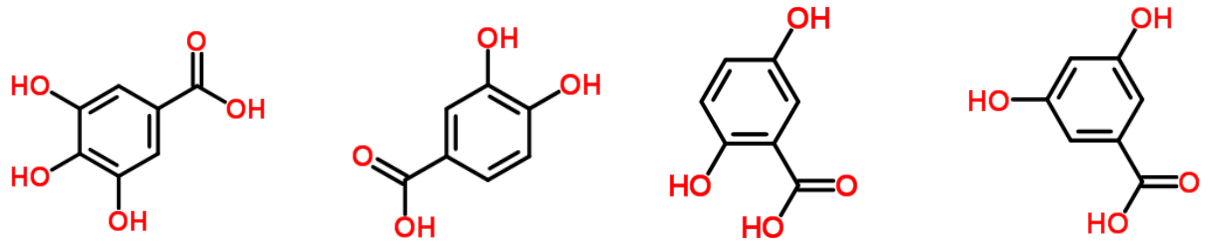
Chapter 2 State of Art

2.1 Phenolic Compounds

Phenolic compounds are a chemical family whose members have one or more hydroxyl groups directly attached to an aromatic ring. They are abundant in fruit, aromatic herbs and vegetables and are well known to have the capacity of scavenging free radicals and oxidizing compounds, which is related to their hydrogen-bonding ability and aromaticity (De Oliveira & Bastos 2011; Vermerris & Nicholson 2009). Those substances have several applications, from anti-antioxidants in the food and oil industry, up to antiviral and anti-inflammatory activity that can be exploited by the pharmaceutical industry. In another perspective, some are considered toxic to a large number of bacteria, indicating that phenolic compounds could also be used in the wastewater common treatment (Noubigh et al. 2013; De Oliveira & Bastos 2011).

Other studies report that some phenolic compounds may be applied in medicine as cardioprotective agents and retardants in cancer cell multiplication (Ferguson et al. 2005; Obied et al. 2005). In 2004, Kampa *et al.* studied the inhibitory effect of some phenolic acids, including the sinapic acid and the protocatechuic acid, on the human breast cancer T47D cells growth, *in vitro*. In addition, the low incidence of coronary diseases and atherosclerosis presented by people who consume olive oil regularly was related to its high content of phenolic compounds (de Lorgeril et al. 1999).

As mentioned before, four phenolic acids, will be studied here in more detail: gallic acid, protocatechuic acid, gentisic acid and α-resorcylic acid. Their chemical structures are presented in Figure 2.1.



(a) (b) (c) (d)

Figure 2.1: Chemical Structures of: (a) gallic acid; (b) protocatechuic acid; (c) gentisic acid and (d) α-resorcylic acid.

Gallic acid (GA, 3,4,5-trihydroxybenzoic acid) is widely distributed in fruits and plants and has many industrial applications, such as antioxidant in food and oil companies, source material for the manufacturing of inks and colors, anti-cancer and antimicrobial agent for the drug industry and raw material for the chemical synthesis of propyl gallate and trimetropim (Ow & Stupans 2003; Mota et al. 2010).

Furthermore, Liu et al. (2013) and Wang et al. (2009) reported that gallic acid is a very strong inhibitor of kappa-casein (k-CN), a milk protein associated to the formation of amyloid, a substance related to the development of several human diseases, such as Alzheimer's, Parkinson's and Huntington's diseases.

Protocatechuic acid (PCA, 3,4-dihydroxybenzoic acid) is also largely found in some fruits and vegetables, such as in acai, mango, grapes, green propolis and yellow and red onions. In addition, high quantities of protocatechuic acid were identified in some plants, such as *Indigofera hirsute*, *Camelina sativa* seeds and *Scutellaria barbata*, widely used in the Chinese medicine (Paula et al. 2016; Vermerris & Nicholson 2009). This compound, due to its strong antioxidant activity, has the property of preventing the germination of onion smudge fungus, *Colletotrichum circinans*, which means that the protocatechuic acid may be used as a barrier for protecting onion's crops (Vermerris & Nicholson 2009). It can also bring benefits to the human body, such as anti-inflammatory and anti-diabetic effects. Another study shows that PCA enhances the activity of superoxide dismutase (SOD), an enzyme that is related to prevention of some neurodegenerative diseases (Hatzipanayioti & Petropouleas 2010).

Gentisic acid (GEA, 2,5-dihydroxybenzoic acid) is also a phenolic compound that has similar biological characteristics to gallic acid and protocatechuic acid, such as antioxidant, anti-inflammatory and antimutagenic properties (Nafees et al. 2012). Other studies indicate

alternative uses for this compound. Vrsalović et al. (2010) published a work concluding that gentisic acid acts as a very good inhibitor of the corrosion of aluminum-magnesium alloys (Vrsalović et al. 2010). Also, gentisic acid presents laboratory application as sample matrix for laser desorption-ionization (LDI) and has shown acceptable results to detect peptides (Strupat, K.; Karas, M.; Hillenkamp 1991).

Finally, α-resorcylic acid (RA, 3,5-dihydroxybenzoic acid) is a phenolic acid found in human urine. In the few studies encountered about this compound, this acid is considered a good inhibitor for lipolysis in adipocyte (Liu et al. 2012). In addition, α-resorcylic acid may be used to condense hyperbranched polyesters, substances that are attracting attention as novel optical, electronic and magnetic materials (Gao & Yan 2004; Mansour et al. 2005).

2.2 Importance of Solubility Measurements

Solubility may be described as the property that measures the ability of one substance (solute) to dissolve within another (solvent) in chemical equilibrium, describing whether they mix up easily or not. Therefore, it is a quantitative term that plays an important role in the behavior of systems containing chemical substances (Martins et al. 2013). Moreover, solubility studies in different solvents provide essential information for the design of separation process, such as precipitation, crystallization and superficial fluid extraction in the food, pharmaceutical, cosmetic and chemical areas (Letcher et al. 2007; Noubigh et al. 2013).

The solubility of an organic compound is directly related to its molecular structure and the polarity of the molecular bonds of solute and solvent. Usually, polar solutes tend to dissolve better in polar solvents, whereas apolar or weakly polar substances are more likely to be dissolved in less polar systems. Actually, the solubility of solids or liquids in another liquid will only occur if the interaction between the solute and the solvent is sufficiently high to promote the rupture of the solute-solute and solvent-solvent interactions. Also, the entropy change, which is related to the system's temperature, is a factor that should be considered to evaluate whether a substance dissolves easily or not in a solvent (Martins et al. 2013).

The solubility has particular relevance in the pharmaceutical industry for which one of the most challenging aims is the discovery of new drugs and formulations that have to be routinely tested. To perform those tests and optimize the drugs' formulation, a large amount of water solubility data are required as this property is directly related to a drug's pharmacokinetic

properties, and consequently, its effects in human organism (Mota et al. 2010; Martins et al. 2013; Baka et al. 2008).

The solubility data in organic solvents, among other properties, are also important to the process and product design in the pharmaceutical industry. The chemical species responsible for the desired activity, usually called active ingredient, is generally isolated via crystallization, requiring solubility data to design the process (Mota et al. 2010).

2.3 Experimental Work

In this section, a literature review is presented regarding experimental methods to measure solubility of solids in liquids, as well as a database containing the solubility of the four selected compounds (gallic acid, protocatechuic acid, gentisic acid and α-resorcylic acid) in water and organic solvents already reported by other authors.

2.3.1 Experimental Methods

The solubility of solids in liquids can be measured by several direct methods, which are usually classified as analytical or synthetic methods. While the former requires the chemical analysis of the liquid and solid phases in equilibrium to determine the solubility of the solid, in the synthetic methods, the solubility is measured by varying a thermodynamic property of the system, such as temperature, pressure or composition, avoiding any chemical analysis (Hefter & Tomkins 2003).

2.3.2 Analytical Methods – Shake-Flask Technique

The analytical methods are considered the most classical approach and are usually based in the saturated shake-flask methodology, proposed more than 50 years ago and still offering satisfactory reliability to measure the solubility of several systems. The basic idea of this method consists on adding an excess amount of solute to the solvent, where a saturated solution should be formed, and the solubility is measured under isothermal-isobaric conditions (Hefter & Tomkins 2003; Baka et al. 2008; Shefter & Higuchi 1963). A sample prepared at saturated conditions is thermostatized and kept under agitation until the system reaches the equilibrium, which may vary between 12 hours and 7 days, depending on the solute and the solvent natures, agitation employed, the amount of material used and the equilibrium method applied (Apley et al. 2015). When the equilibrium is achieved, the remaining solid, also called residue, is removed

from the supernatant (mother solution) by filtration or centrifugation, and then the concentration of the solute in the solution can be determined. Various analytical techniques may be used to identify the solubility of the solid in the mother solution, such as gravimetry, UV-Vis spectroscopy, HPLC and X-ray diffraction (Hefter & Tomkins 2003; Mota et al. 2010).

For poorly soluble systems, the time required to reach the equilibrium is normally higher than for systems containing soluble solutes. One way to speed up the process is to increase the surface available area for dissolution, which can be achieved by either vortexing or sonicating the samples during the process. Other challenges in determining the solubility of poorly soluble solids are their tendency to float (Apley et al. 2015).

In order to achieve reliable measurements, it is essential to ensure that the equilibrium state is reached, which can be obtained by studying the solute's dissolution profile in the system. The shortest time required to obtain a constant solute concentration can be considered as a suitable equilibrium time, which can be easily obtained by isothermal gravimetry. In this method, a super-saturated mother solution is prepared and maintained stirred and under isothermal-isobaric conditions while several samples are collected at different times. The supernatant is then removed from the samples and the remaining solid part is weighted. When there is no considerable variation in the solid solubility, the equilibrium is reached (Hefter & Tomkins 2003).

Although requiring longer times of experimental work, the gravimetric method of analysis can be considered quite accurate and reproducible to perform solubility measurements in pure and mixed solvents. It may present some loss issues when applied to systems containing lipophilic insoluble compounds, as well as some limitations due to the retention of solvent in the solid inner interstices (Mota et al. 2010).

UV-Vis spectroscopy allows solubility measurements of several systems due to the large wavelength range that can be applied and correlated to a calibration curve previously built. They also provide satisfactory reproducibility and speed of analysis and impurities can be easily identified. However, when the UV absorption decreases, the uncertainty of the solubility results considerably increases (Mota et al. 2010).

High performance liquid chromatography (HPLC) is a powerful analytical technique that can be coupled to a saturated solution generation column to measure aqueous solubilities. It can reduce colloidal dispersions, solute adsorption in the material walls, minimize sample loss by

evaporation and the use of organic solvents (Mota et al. 2010). On the other hand, HPLC is considered a time consuming technique because it requires long runs and a calibration curve to be correlated to the desired parameter (Lin et al. 2009).

Even though the shake-flask method is considered one of the simplest procedures to determine equilibrium solubilities, it is time consuming and requires lots of manual work. The reliability of the results depends on a rigorous control of external variables, such as temperature, pressure, sedimentation time, stirring time and technique applied to separate the solid and the liquid phases (Baka et al. 2008)

In order to minimize the experimental error involved in the shake-flask method, Baka et al. (2008) recommend some procedures that should be observed while performing the experimental work:

- The measurements must be carried out at controlled, standard temperature;
- The amount of solid in excess present in the solution should be around 1 - 2 mg/ml of solution, to avoid difficulties in sampling;
- Equilibrium time must be checked for each compound studied. However, a minimum time to reach the equilibrium should be around 24 hours, summing 6 hours of stirring and 18 hours for sedimentation.

According to these authors, when the procedures above mentioned are strictly followed, the experimental error of the solubility measurements can be reduced to about 4 % (Baka et al. 2008).

2.3.3 Experimental Database

For comparison purposes, a literature review of the experimental methodology employed by other authors in similar analysis was performed. From the four selected benzoic acids, only studies related to gallic acid and protocatechuic acid were found. The experimental methodology and its specifications are described in Table 2.1.

Table 2.1: Methodologies employed by different authors to measure the solubility of gallic acid and protocatechuic acid .

System	Reference	Range of Temperature (K)	Experimental Methodology	Shaking Time (h)	Settling Time (h)
Gallic Acid plus Water	(Mota et al. 2008)	288 – 323	Shake-Flask coupled to UV-VIS Spectroscopy and to Gravimetric Methods	64-117	7-26
Gallic Acid plus Methanol, Ethanol, Water, and Ethyl Acetate	(Daneshfar et al. 2008)	298.2 – 333.2	Shake-Flask coupled to UV-VIS Spectroscopy	4[a]	1
Gallic Acid plus Water	(Lu & Lu 2007)	273.2 – 363.2	Shake-Flask Method coupled to HPLC	2[b]	6
Gallic Acid plus Methanol and Water	(Noubigh et al. 2013)	293.15 – 318.15	Shake-Flask coupled to UV-VIS Spectroscopy	3	NA[c]
Gallic Acid plus Water 1-Propanol, 2- Propanol and Acetonitrile	(Dali et al. 2016)	293.15 – 318.15	Shake-Flask coupled to UV-VIS Spectroscopy	3	NA[c]
Protocatechuic Acid plus Water	(Queimada et al. 2009)	288.2 – 323.2	Shake-Flask coupled to UV-VIS Spectroscopy and to Gravimetric Method	120 - 140	24 - 40
Protocatechuic Acid plus Methanol, Ethanol, Methyl Acetate and Ethyl Acetate	(Noubigh et al. 2015)	293.15 – 318.15	Shake-Flask coupled to UV-VIS Spectroscopy	3	NA[c]

[a]Samples were stirred at 400 – 500 rpm. [b]Samples were agitated at 200 rpm using an electronic stirrer. [c]Information not available.

The solubility data compiled in Table 2.1 are graphically presented in Figures 2.2. to 2.5. (values obtained from Tables A.1 to A.4 of Appendix A) and Table 2.2.

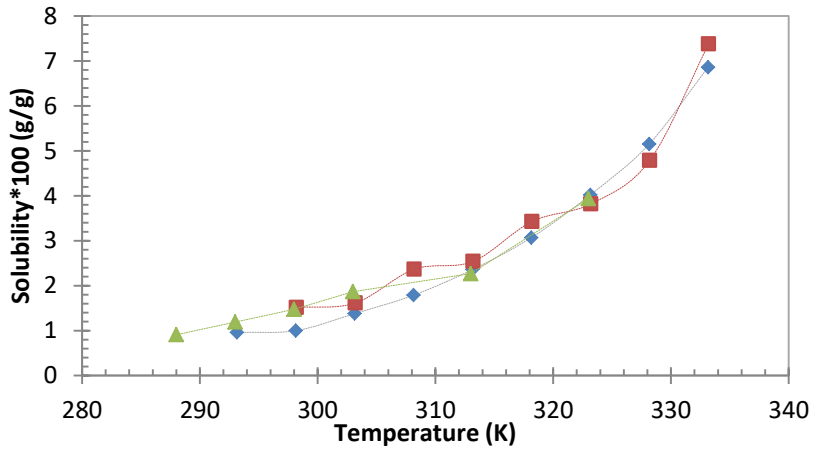


Figure 2.2: Solubility in weight fraction of gallic acid in water available in literature: (Lu & Lu 2007) (◆), (Daneshfar et al. 2008) (■) and (Mota et al. 2008) (▲).

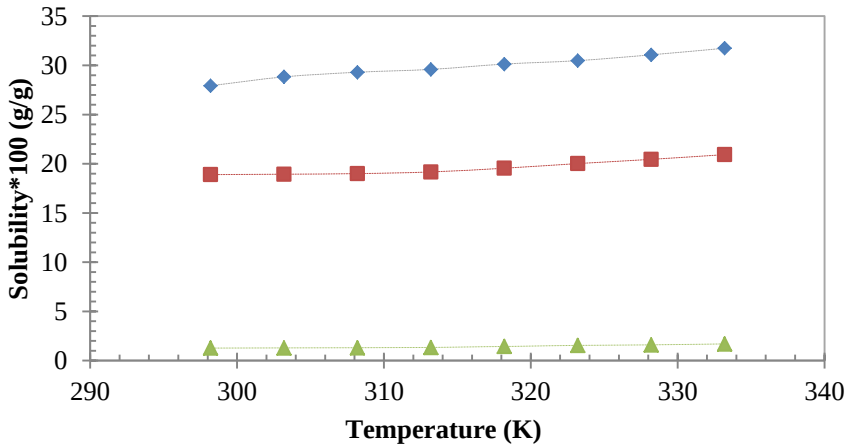


Figure 2.3: Solubility in weight fraction of gallic acid in different organic solvents as a function of temperature: methanol (◆), ethanol (■) and ethyl acetate (▲) (Daneshfar et al. 2008).

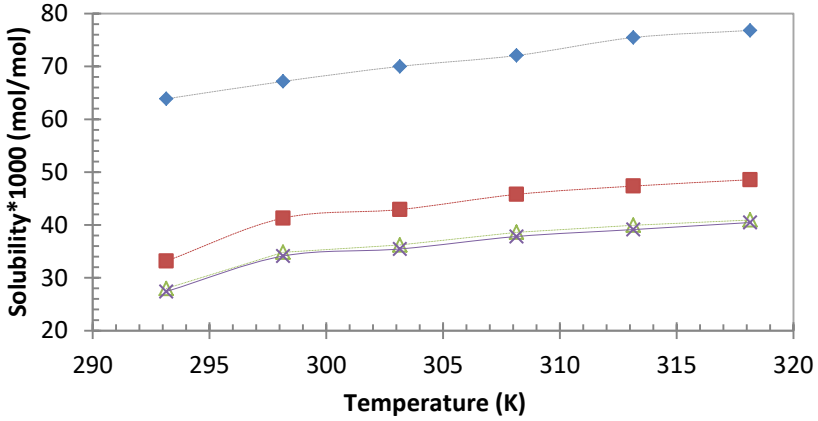


Figure 2.4: Solubility in mole fraction of gallic acid in different solvents as a function of temperature: methanol (◆) (Noubigh et al. 2015), 1-propanol (■), 2-propanol (▲) and acetonitrile (×) (Dali et al. 2016).

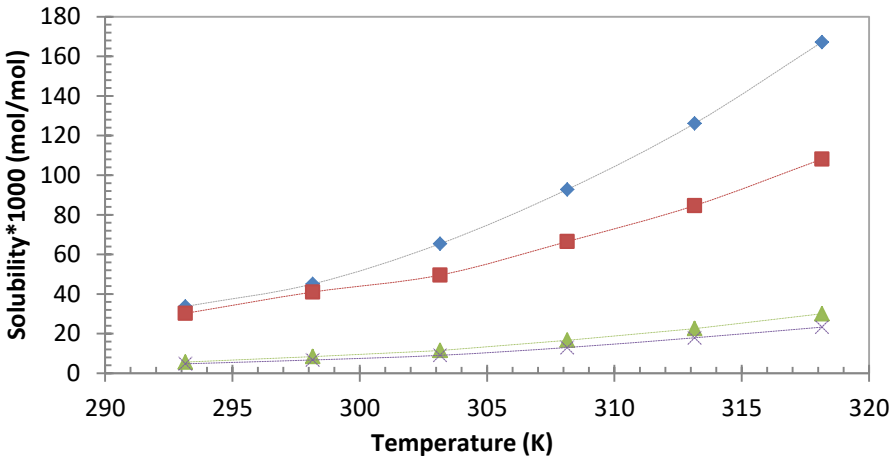


Figure 2.5 Solubility in mole fraction of protocatechuic acid as a function of temperature in different solvents methanol (◆), ethanol (■), methyl acetate (▲) and ethyl acetate (×) (Noubigh et al. 2015).

Table 2.2: Solubility in (g/L) of protocatechuic and gentisic acids in water found in literature

Solute	Temperature range (K)	Solubility (g/L)	Reference
Gentisic acid	298.15	22	(Herzog & Swarbrick 1971)
Protocatechuic acid	287.15	18.2	(Yalkowsky et al. 2010)

The wide application of gallic acid and protocatechuic acid in the pharmaceutical, food and chemical industries is probably the reason why more information is available for these compounds.

Although solubility measurements of gallic acid were performed by several authors, the values do not totally agree. For instance, the solubility of GA in water provided by Lu and Lu (2007) are generally lower than the results reported by Mota et al. (2008) and Daneshfar et al. (2008). Those differences may be due to the different experimental methodologies including the analytical techniques employed. Lu & Lu (2007) applied the shake-flask methodology coupled with HPLC, stirring the samples during 2 hours, what may not be enough time to reach the equilibrium state. Daneshfar et al. (2008) employed UV-Vis spectroscopy technique to quantify the solubility and 3 hours for the stirring time.

On the other hand, few solubility data were found for gentisic acid (see Table 2.2.) in line with the scarce applications described before. Similarly, no data were found for α-resorcylic acid.

2.4 Thermodynamic Modeling

Although solubility data as a function temperature are fundamental to design several industrial processes, there is still a lack of information about many solid-liquid and liquid-liquid systems, especially involving organic solvents. Even with the constant improvement of the analytical equipment, experimental works focused on solubility measurements usually take substantial time to be performed and must be carried out very rigorously to achieve reliable results (Mota et al. 2010).

In many cases, solubility data are unavailable due to time restrictions and limited amount of samples. In addition, given the complexity of most drug molecules and the large diversity of their interactions, solubility may be measured very easily in simple systems, but it can be very complex task, for instance, when the system contain multicomponent solvents (Mota et al. 2012)

In order to model, complement and support the experimental measurements, many thermodynamic methods have been proposed. Among those tools, there are some theoretical models that rely on information about the molecules under study, and others that are based on mathematical correlations of the experimental data. Some semi-empirical models may be applied to several non-ideal binary or multicomponent systems and describe their thermodynamic behavior. (Chen & Song 2004; Mota et al. 2011)

2.4.1 Review of Models to Calculate Phase Equilibria

Some of the most commonly thermodynamic models used to predict the equilibrium of drugs in the pharmaceutical industry are the models of Wilson, UNIQUAC, NRTL, Hansen, UNIFAC, NRTL-SAC, among others (Letcher et al. 2007; Prausnitz et al. 1999).

Wilson's model, one of the first thermodynamic methods proposed to determine a non-ideal equilibrium, is based on molecular considerations for binary and miscible systems. The model has two adjustable parameters, Λ_{12} and Λ_{21}, which are related to the pure-component molar volumes and to the interaction energies, and can be obtained, for binary systems from experimental data. Generally, in practical applications, the systems of interest are multicomponent or multiphase systems (Farajnezhad et al. 2016). For those cases, the multicomponent system may be considered composed by several binary systems, which generates more Wilson's interaction parameters that usually are not available in literature. An

alternative way to determine those parameters requires the knowledge of several experimental solubility data to be fitted and applicable for further calculations (Farajnezhad et al. 2016).

Although Wilson's method was one of the first applied models, it has some restrictions. The first disadvantage is that those equations are not useful for systems where the logarithms of the activity coefficients reach maxima or minima. Another limitation is that the Wilson's model shou1d only be used to predict solubilities involving completely miscible liquid systems or, else, for those limited regions where just one liquid phase is present (Rowlinson 1970).

To overcome some of the limitations of Wilson's method, some other models were developed boosted by the necessity of measuring solubilities of pharmaceuticals and polymers. Hansen model is based on the Hansen solubility parameters, which are obtained from mathematical regression of experimental solubility data (Srinivas et al. 2009). It was formulated based on the fact that solubility parameters have shown great industrial application to aid in solvent selection, being considered a correlative model (Hansen 2013).

A well-known category of thermodynamic predictive models follows a group contribution methodology, which is based on the concept that the properties of a molecule can be derived from the functional groups that compose it (Nouar et al. 2016). The most successful method based on the functional group concept is the Universal Functional-Group Activity Coefficients (UNIFAC) model, which has been constantly applied to predict vapor-liquid, liquid-liquid and solid-liquid equilibria (Nouar et al. 2016). By using chemical structure information from the molecules that compose the studied system and some binary interaction coefficients, the model calculates activity coefficients of the components present in solution, and consequently, determines the system equilibrium (Chen & Song 2004; Nouar et al. 2016).

Although both UNIFAC and Hansen models fit well many systems, they have some limitations. For instance, they are inadequate to estimate the solubility of either large molar weigh molecules (above 200 g/mol) or systems containing electrolyte solutes (Chen & Song 2004). In spite of the fact that Hansen's model is based on the simple assumption that relates solubility parameters to experimental data, the method has limited practical applications regarding drug solubility (Mota et al. 2010).

The UNIFAC model is not applicable to predict solubilities from systems containing isomers or at high pressure, above 10 atmospheres (Pistikopoulos et al. 2010). Furthermore, in

some cases, the parameters for certain functional groups and binary interactions are not available making the UNIFAC method unsuitable (Pistikopoulos et al. 2010; Valavi et al. 2016).

Similarly to the Wilson's model, Renon and Prausnitz (1968) also considered the local composition concept to derivate the Non-Random Two-Liquid (NRTL) model, which is one of the most successful thermodynamic models in the chemical industry to provide precise presentation of nonideal VLE and LLE systems (Chen & Song 2004). The following equation describes the NRTL activity for a multicomponent system:

$$ln\,\gamma_I = x_j^2\left[\tau_{ji}\left(\frac{G_{ji}}{x_i+x_jG_{ji}}\right)^2 + \frac{\tau_{ij}G_{ij}}{\left(x_j+x_iG_{ij}\right)^2}\right] \quad (1)$$

with G_{ij} and τ_{ij} defined as follows:

$$G_{ij} = \exp\left(-\alpha_{ij}\tau_{ij}\right)\tau_{ji} \quad (2)$$

The parameter G_{ij} also can be calculated through the following expression:

$$\exp\left(-\alpha_{ij}\tau_{ij}\right)\tau_{ji} = \frac{g_{ij}-g_{ji}}{RT} = \frac{a_{ij}}{RT} \quad (3)$$

where g_{ij} is the energy interaction between *i* and *j* molecules, α is the non-randomness factor, *T* is absolute temperature of the system and *R* is the ideal gas constant (Renon & Prausnitz 1968). The model requires three binary interaction parameters that are determined by regression of experimental data to a specific system: a_{12}, a_{21} and α_{12}. The reduction of experimental data indicates that α_{12} varies from 0.2 to 0.47 for a large number of binary systems, which suggests that this parameter can be fixed when the experimental data are scarce (Rowlinson 1970).

For moderate nonideal systems, the NTRL model provides no advantages over the simpler Wilson's models. However, this model predicts more accurately the solubility of very nonideal systems, composed by partially immiscible phases (Rowlinson 1970). The model has shown to be very precise to determine some equilibrium properties; for instance it has been used to correlate the solubility of niflumic acid, flufenamic acid and diclofenac sodium in different solvents, reporting a deviation of 2% from the experimental solubility measurements (Valavi et al. 2016).

2.4.2 Non-Random Two-Liquid Segment Activity Coefficient Model (NRTL-SAC)

In 2004, Chen and Song, considering the successful range of the NRTL industrial applications, especially in polymer industry, and the mentioned limitations of the group contribution models, proposed an innovative variant method to describe liquid-liquid equilibrium systems: the NRTL segment activity coefficient (NTRL-SAC) model (Chen & Song 2004; Mota et al. 2010). In this method, the liquid non-idealities are defined in terms of three distinct conceptual molecules' segments: hydrophilic, polar and hydrophobic. Those concept segments, or molecules descriptors, are represented respectively by hexane, acetonitrile and water, and represent the possible surface interactions that a solute and a solvent may have in a binary system. (Chen & Song 2004; Letcher et al. 2007)

By resorting to an extensive VLE and LLE database of 62 common solvents commonly used in the pharmaceutical industry and assuming that the non-idealities may be described in terms of four molecules descriptors, Chen and Song (2004) estimated the number of segments required in each solvent and their values (Letcher et al. 2007). Considering those values and a few selected experimental solubility data of the target solute, it's possible to predict the solute's number of segments readily and, consequently, use them to estimate its solubility in other solvents and systems (Chen & Song 2004; Chen & Crafts 2006).

One advantage of this method is to require less experimental work to predict a solute's solubility in several solvents (Fakhraian et al. 2016). In addition, NRTL-SAC can be a very convenient tool to design a crystallization process due its capability of identifying both solvent and anti-solvent candidates (Chen & Crafts 2006).

2.4.2.1 NRTL-SAC Model Equations

NTRL-SAC model revealed to be a very consistent thermodynamic tool to qualitatively correlate and predict drug solubility of pure and multicomponent systems, based only in a small initial set of experimental solubility data (Chen & Crafts 2006). In this model, the activity coefficient for a component *I* present in solution is the sum of a combinatorial term, γ^{C}_{I} and a residual term, γ^{R}_{I}.

$$ln\,\gamma_I = \ln\gamma_I^C + \ln\gamma_I^R \tag{4}$$

The combinatorial term, $\gamma^{c}_{I,}$ is calculated from the Flory-Huggins equation for the combinatorial entropy mixing, as follows:

$$\ln \gamma_I^C = \ln \frac{\phi_I}{x_I} + 1 - r_I \sum_J \frac{\phi_I}{r_J} \quad (5)$$

where *I* and *J* are component indices, $\phi_I$is the segment mole fraction of component *I*, r_1 and r_J are the total segment number in components *I* and *J* and x_I is the molar fraction of component *I*.

. Those terms are calculated by the following expressions:

$$r_I = \sum_i r_{i,I} \quad (6)$$

$$\phi_I = \frac{r_I x_I}{\sum_J r_J x_J} \quad (7)$$

where *i* the segment based-species indices, x_I and x_J are the molar fractions of components *I* and *J*, $r_{i,I}$ is the number of segment species *i* contained in component *I*.

The residual term, $\gamma^R{}_I$, is based on the NRTL's model local composition interaction contribution, $\gamma^{lc}{}_{I,}$ which is represented by the following equation:

$$\ln \gamma_I^R = \ln \gamma_I^{lc} = \sum_m r_{m,I} [\ln I_m^{lc} - \ln I_m^{lc,I}] \quad (8)$$

where $r_{m,I}$ is the number of segments *m* contained in species *I*, I_m^{lc} is the activity coefficient of segment species *m* and $\ln I_m^{lc,I}$ is the activity coefficient of segment species restricted only in component *I*. Those terms can be computed from NRTL model's equation:

$$\ln I_m^{lc} = \frac{\sum_j x_j G_{jm} \tau_{jm}}{\sum_k x_k G_{km}} + \sum_{m'} \frac{x_{m'} G_{mm'}}{\sum_k x_k G_{km'}} \left(\tau_{mm'} - \frac{\sum_j x_j G_{jm'} \tau_{jm'}}{\sum_k x_k G_{km'}} \right) \quad (9)$$

$$\ln I_m^{lc,I} = \frac{\sum_j x_{j,I} G_{jm} \tau_{jm}}{\sum_k x_{k,I} G_{km}} + \sum_{m'} \frac{x_{m,I'} G_{mm'}}{\sum_k x_{k,I} G_{km'}} \left(\tau_{mm'} - \frac{\sum_j x_{j,I} G_{jm'} \tau_{jm'}}{\sum_k x_{k,I} G_{km'}} \right) \quad (10)$$

where *i*, *j*, *k*, *m* and m' are the segment-based species indices, *I* and *J* are the component indices and x_j is the segment-based mole fraction of the species *j*. As in the NRTL model, the NRTL-SAC has two parameters, *G* and τ, which are related to each other by the non-randomness factor α:

$$G = e^{-(\alpha.\tau)} \quad (11)$$

In the NRTL-SAC model, the behavior of the mixtures is determined by the segment compositions of the molecules and their pairwise segment-segment interactions, which are

represented by characteristics of hydrophobicity (*X*), polarity (*Y*) and hydrophilicity (*Z*). Those parameters are obtained not from the molecular structures, but from the interaction characteristics of the molecules in solution expressed in terms of their experimental equilibrium data (Chen & Song 2004). The local binary quantities and the non-randomness factor were determined to be constants and are shown in appendix B.

As mentioned before, the solvents chosen to represent the hydrophobic, polar and hydrophilic behavior are hexane, acetonitrile and water, respectively. The binary parameters for hydrophobic (*X*) and hydrophilic (*Z)* interactions are obtained from liquid-liquid equilibrium (LLE) data of hexane-water mixture. The non-randomness factor α was fixed at 0.2 for these interactions because this is the ordinary value for systems that present liquid-liquid separation. Likewise, the binary parameters for hydrophobic segment (*X*) – polar (*Y*) segment and for polar segment (*Y*) – hydrophilic segment (*Z*) were obtained from available data of hexane-acetonitrile and acetonitrile-water, respectively, which lead the authors to fix α at 0.2 for both cases (Chen & Song 2004). On the other hand, the binary parameters for the polar segment (*Y*) – hydrophilic (*Z*) segment were obtained from vapor-liquid equilibrium (VLE) data, fixing the non-randomness factor at 0.3. Furthermore, the polar segment was subdivided in *Y*- and *Y*+ and τ_{12} was established to vary between -2 and +2 to reflect the fact that interaction between polar segment and the hydrophilic segment may be positive or negative. It's also assumed in the model that there is no interaction among segments of the same nature, i.e, polar-polar segments (Chen & Song 2004).

In addition to the conceptual segment parameters, the NRTL-SAC also requires molecular parameters for the solvents present in the systems that will be analyzed. Those parameters can be determined from regression of the available LLE and VLE data for binary systems of the intended solvent and the reference molecule descriptors (hexane, acetonitrile and water). The molecular parameters identified by Chen and Song (2004) for 62 solvents commonly used in the pharmaceutical industry are presented in appendix B.

Although three conceptual segments were generally defined, in many cases only one or two molecular parameters are necessary to describe a solvent`s behavior. Therefore, considering the parameters expressed in Appendix B and the NRTL-SAC equations, it is possible to estimate the activity coefficient of the solute in a binary system, which is directly related to the solute's solubility.

Assuming that the solubility of an organic nonelectrolyte solid can be described by the expression:

$$\ln x_I^{SAT} = \frac{\Delta_{fus}S}{R}\left(1 - \frac{T_m}{T}\right) - \ln \gamma_I^{SAT} \quad (12)$$

for $T \leq T_m$

$$\Delta_{fus}S = \frac{\Delta_{fus}H}{T_m} \quad (13)$$

where x_I^{SAT} is the mole fraction of the solute *I* dissolved in the solvent phase at saturation, $\Delta_{fus}S$ is the entropy of fusion of the solute, *R* is the ideal gas constant, *T* is the absolute temperature (measured in Kelvin), T_m is the melting point of the solute, $\Delta_{fus}H$ is the enthalpy of fusion of the solute and γ_I^{SAT} is the activity coefficient of the solute at saturation (Frank et al. 1999). The terms $\Delta_{fus}S$, $\Delta_{fus}H$ and T_m are thermodynamic properties that vary among polymorphic forms of solute. Therefore, considering a polymorph at a specific temperature, the solute solubility is only function of its activity coefficient, which can be obtained by the NTRL-SAC model (Chen & Song 2004; Chen & Crafts 2006).

2.4.2.2 NRTL-SAC Applications

In this section, a few applications of the NRTL-SAC model, relevant to this work, will be discussed. Very recently the NRTL-SAC model has shown to be very useful of the in the description of of a set of pharmaceutical compounds in aqueous solution with a reported average error of 38%, which may be considered lower compared to other thermodynamic approaches (Valavi et al. 2016). It is also relevant to mention that NRTL-SAC, unlike other thermodynamic models, ignores temperature dependence of the activity coefficient, which means the influence of temperature is only considered on the activity of the solid (Valavi et al. 2016; Chen & Song 2004).

Despite its empirical nature, this model has shown to be relatively advantageous compared to other thermodynamic models due its robustness and accuracy (Valavi et al. 2016). The simplicity and the wide applicability of the model, either for organic non-electrolytic and organic electrolytic molecules, are also points that should be accounted while evaluating the NRTL-SAC features (Chen & Crafts 2006). Furthermore, Valavi et al. (2016) concluded that

NRTL-SAC, when predicting the solubility of pharmaceutical compounds in organic solvents, presents more accurate results than the group contribution UNIFAC model.

Mota et al. (2010) presented a review of the successful applications of the NRTL-SAC model until 2008. The authors identified that the model has been applied to estimate the solubility of complex chemicals, such as acetylsalicylic acid, benzoic acid, testosterone, theophylline, estriol, hydrocortisone, among others, obtaining acceptable deviations between experimental and predicted values.

In order to better evaluate the performance of the NTRL-SAC model, a literature review of the works published between 2006 and 2016 is compiled in Table 2.3., updating the review made by Mota et al. (2008).

Table 2.3: Literature review of the NRTL-SAC successful applications from 2008 to 2016.

Works	**Author**
NRTL-SAC presented excellent solubility predictions of paracetamol, sulfadiazine, cimetidine and sulfamerazine in mixed solvents at 293, 298 and 303 K	(Chen & Crafts 2006)
Prediction of infinite-dilution activity coefficient data of 22 ionic liquids in 35 solvents.	(Chen et al. 2008)
The symmetric Nonrandom Two-Liquid Segment Activity Model (*e*-NRTS-AC) was proposed for electrolytes and the prediction of salt solubilities in a few representative solvents was carried out. They also estimated the solubility of sodium acetate and sodium salicylate in water-ethanol mixed solvent.	(Song & Chen 2009)
Prediction of the solubility of paracetamol, budesonide, allopurinol and furosemide in water, acetone/water, ethanol/water and ethanol/ethyl acetate.	(Mota et al. 2009)
The authors designed and compared the crystallization of acetaminophen in ethanol by applying the Van Laar equation, Wilson's model, NRTL, NTRL-SAC and UNIFAC models. The most precise method found at predicting the equilibrium solubility and the crystal size was NRTL-SAC.	(Widenski et al. 2010)
NRTL-SAC, UNIFAC, MOSCED and Jouyban-Acree methods have been used to model and optimize an isothermal anti-solvent crystallization of acetaminophen in acetone/water systems. NRTL-SAC and Jouyban-Acree showed to be the most accurate methods in the analyzed situation.	(Widenski et al. 2011)
The authors combined an original optimization procedure with NRTL-SAC to screen, among 62 solvents, the best option to perform the crystallization of seven pharmaceutical molecules (lovastatin, valsartan, paracetamol, budesonide, allopurinol, furosemide and sulfadiazine).	(Sheikholeslamzadeh et al. 2012)
Solubility prediction of drug-like molecules, such as salicylic acid, benzoic acid,	(Mota et al. 2012)

acetylsalicylic acid, ibuprofen, hydroquinone, estirol and estradiol in systems containing ethanol, 1-butanol, 1-pentanol, 1-octanol.	
The solubility of three compounds (3-pentadecylphenol, lovastatin, and valsartan) in different solvents and solvent mixtures was studied. The prediction results showed a better performance of the NRTL-SAC model compared to the UNIFAC model.	(Sheikholeslamzadeh & Rohani 2012)
Prediction of partition coefficients and selection of suitable solvents employed in counter-current chromatography systems. Several solutes were tested in heptane/methanol/water, heptane/ethyl acetate/methanol/water (Arizona and hexane/ethyl acetate/methanol/water systems to validate the method).	(Ren et al. 2013)
Solubility predictions of epicatchin, epigallocatechin, epicatchin gallate and epigallocatechin gallate in water/ethanol mixtures at 293 K and 303 K were performed using UNIFAC and NRTL-SAC methods. The NRTL-SAC model was found to be the most accurate model.	(Sevillano et al. 2013)
Prediction of androstenedione solubility in binary mixtures of methanol + water and ethanol + water at temperatures from 275 to 325 K.	(Tang et al. 2014)
Solubility estimation of phosphoryl chloride and trimethylamine in several solvents, such as dichloromethane, acetic acid, ethyl acetate, acetone, n- hexane, 1-butanol, 2-propanol, isopropyl ether at temperatures from 283.15 to 323.15 K	(Feng et al. 2014)
The authors employed NTRL-SAC model to screen a suitable biphasic liquid system, between four possibilities, to be applied on a phenolic extraction process by high speed counter-current chromatography. The extracted phenolic compounds were 3,4-dihydroxyphenylethanol, vanillic acid, orientin, vitexin, veratric acid, 2"-O-(3"', 4"-dimethoxybenzoyl) orientin, 2"-O-feruloylorientin, 2"-O-feruloylvitexin, 2"-O-(2" methylbutyryl) vitexin, 2"-0-(2"'-methylbutyryl) isoswertiajaponin, 2"-O-(2"'-methylbutyryl) isoswertisin and the solvent systems evaluated were composed by different compositions of hexane/ethyl/acetate/ethanol/methanol/water	(Qin et al. 2015)
By using eleven solvent system families containing 33 biphasic liquid systems, the authors, based on NRTL-SAC model, proposed a systematic and practical solvent system selection strategy to predict partition coefficients of eleven more solvent families containing partially or totally different solvents.	(Ren et al. 2015)
Development of a temperature-dependent NTRL-SAC model applied to systems containing risperidone, fenofibrat, fenoxycarb, tolbutamide, meglumine, butyl paraben, butamben, salicylamide in organic solvents, such as methanol, toluene, ethanol, 1-propanol, 2-propanol, 1-butanol, acetone and ethyl acetate	(Valavi et al. 2016)
NRTL-SAC was applied to model ternary phase diagram for chiral medetomidine salts in alcohols. The systems analyzed were composed by medetominide hydrochloride and 2-propanol, medetominide hydrobromide and 2-propanol, and medetominide oxalate and ethanol	(Fakhraian et al. 2016)

Table 2.3 shows that the NRTL-SAC model has been widely studied since 2006. Some studies point the model as a very good tool to assist the design of several separation processes, such as cooling crystallization, isothermal anti-solvent crystallization and counter-current chromatography. In addition, many works have compared NRTL-SAC with other models, like UNIFAC, pointing that the segment methodology usually provided the most reliable results.

Besides the application pointed in Table 2.3, NTRL-SAC was also included in the thermodynamic library of the commercial software *Aspen Properties* and *Aspen Polymers Plus*. As mentioned by the ASPEN Technology "NRTL-SAC can be used for fast, qualitative estimation of the solubility of complex organic compounds in common solvents" (Aspen Technology 2010).

To our knowledge, the NRTL-SAC was not used to predict the solubility of gallic acid, protocatechuic acid, gentisic acid and α-resorcylic acid. However, several authors applied it to drug molecules, with complex structures, which lead us to believe that NRTL-SAC may be used to determine molecule descriptors of the compounds addressed in this work, and consequently calculate its solubility in water and organic solvents.

2.4.3 Reference Solvent Approach (RSA)

Despite the robustness and accuracy of some thermodynamic models, prediction of the solubility of a solid in organic liquids may find some hindrances. Group-contribution approaches, such as UNIFAC, require the availability of all the group parameters that compose the molecules in the analyzed systems. Although the values can be measured in some cases or be obtained from similar structures, in many cases the available database is insufficient to estimate the missing parameters (Abildskov & O'Connell 2003).

The NRTL segment activity coefficient approach, on the contrary, requires only four parameters for each component in the analyzed system, what is a considerable advantage over the UNIFCAC model when the systems are composed by complex molecules. On the other hand, in order to predict reliable solubility values, the knowledge of the solute's melting point and enthalpy of fusion is required. However, those values are not always available and sometimes they are not simple to be determined.

Considering the limitations mentioned above, Abildskov & O'Connell (2003) proposed an alternative approach that may be incorporated in the thermodynamic approaches to maximize

their realibility and accuracy. The Reference Solvent Approach (RSA) employs one of the solected solvents as reference to predict the activity coefficients and solubilities of the same solute in different systems. The selection of the reference solvent is related to minimization of the errors between the experimental and predicted solubilites, as shown in the next topic.

2.4.3.1 Thermodynamic Framework

The solubility of a solid solute in a liquid solvent at equilibrium, when both molecules have sufficiently different sizes and shapes that no solid solutions are formed, can be calculated by:

$$\ln x_s^L = \ln x_s^{id} - \ln \gamma_s^L \tag{14}$$

where x_s^L is the solute mole fraction solubility and γ_s^L is the solute activity coefficient using the Lewis/Randall standard state (pure component as a liquid at the system temperature *T*). For nonideal solutions, x_s^L needs to be calculated iteratively by equation (14) and thermodynamic methods such as UNIFAC and NRTL-SAC may be used to determine the activity coefficient iteratively. The ideal solubility is obtained approximately from the ratio of standard-state fugacities, generally approximated with the use of the melting properties T_{mS} and ΔH_{mS}, as shown in the following expression:

$$\ln x_I^{id} = \frac{\Delta H_{mS}}{R T_{mS}} \left(1 - \frac{T_{mS}}{T}\right) \tag{15}$$

Therefore, the solubility of the solute can be assumed as function of the melting parameters and activity coefficient. However, corrections to equation (14) are very important, especially if the system presents solid phase transitions between the system temperatures T and T_{mS}. In several cases, the solid form of the solute can vary with the solvent, either by crystalline lattice, solute' solvation or compound formation (Shefter & Higuchi 1963; Abildskov & O'Connell 2003). Those aspects cannot be taken in account in the methodology presented above and accurate melting data are required to perform reliable predictions of solubility in solid-liquid systems.

Considering the discussion pointed above and disregarding any variations of the solid in the system, the term represented by equation (15) can be considered constant and the solubility of the solid *S* in a solvent *i* can be calculated as a function of the reference solvent *j* as shown below:

$$\ln x_{Si} = \ln x_{Sj} + \ln \gamma_{Sj}\left(T, \{x_S\}_j\right) - \ln \gamma_{Si}(T, \{x_S\}_i) \quad (16)$$

Where the reference terms in equation (15) can be gathered and written as:

$$Ref_{Sj}\left(T, \{x_S\}_j\right) = \ln x_{Sj} + \ln \gamma_{Sj}\left(T, \{x_S\}_j\right) \quad (17)$$

Assuming that the solute solubility in the reference solvent *j* is obtained experimentally, the solute activity coefficient and, consequently, the parameter *Ref* in equation (17) can be obtained interactively through robust thermodynamic methods such as UNIFAC, NRTL and NRTL-SAC in different temperatures. Then, a set of values for equation (17) in different times can be combined to one of the mentioned thermodynamic models to predict iteratively the solute' solubility in solvent *i* through equation (15). The present methodology is very useful when the pure-solute properties are either unknown or very difficult to measure (Abildskov & O'Connell 2003).

2.4.3.2 Reference Solvent Approach

Despite RSA cancels errors of measurements or assumption of pure-solute properties, it requires the selection of the reference solvent. In order to choose the best solvent option, Abildskov & O'Connell (2003) proposed the evaluation of the residual term represented by the following expression:

$$\delta \ln x_{S,ij} = \ln x_{Si} + \ln \gamma_{Si}(T, \{x_S\}_i) - Ref_{Sj}\left(T, \{x_S\}_j\right) \quad (18)$$

The residual term represents the error obtained by assuming the solvent *j* as reference, in other words, it describes how much the approximation assumed by equation (16) diverge from reality, especially due to experimental errors or solid phase transitions (Abildskov & O'Connell 2003). In order to obtain the reference solvent, a minimization of the sum of *N* available residual terms should be performed, as shown in the following equation:

$$min \left| Ref_{Sj}\left(T, \{x_S\}_j\right) - \sum_{i=data} \frac{\ln x_{Si} + \ln \gamma_{Si}(T, \{x_S\}_i)}{NP} \right| \quad (19)$$

where *NP* is the total number of data points for the different *i* solvents in the database (Mota et al. 2012; Abildskov & O'Connell 2003). This strategy was developed to aid in the selection of

the solvent that adjusts the system better in terms of minimum error. However, it is important to evaluate the scattering of the data adjusted by the RSA technique.

Chapter 3 Solubility Measurements

3.1. Experimental Methodology for the Solubility Measurements

3.1.1. Compounds

All the solutes and solvents were used as received, without further purification. Ultrapure water (resistivity of 18.2 MΩ.cm, free particles ≥ 0.22 μm and total organic carbon (TOC) < 5 μd·dm^{-3}) was obtained at the laboratory LQA (at IPB) through a reverse osmosis process using a Direct-Q® Water Purification system. The identification, source and purity of the remaining components are described in Table 3.1. The normal boiling point of the solvents is also provided.

Table 3.1: CAS, Molar Mass, Assay (Purity %), provider and boiling points at atmospheric pressure of each of the components employed in this work.

Component	CAS	Molar Mass (M)	Mass Purity (%)	Source	Normal Boiling Point (ºC)
Gallic Acid	149-91-7	170.12	≥ 98	Merck KGaA	---------
Protocatechuic Acid	99-50-3	154.12	≥ 96	Merck KGaA	---------
Gentisic Acid	490-79-9	154.12	≥ 99	Merck KGaA	---------
α-Resorcylic Acid	99-10-5	154.12	≥ 98	Merck KGaA	---------
Methanol	67-56-1	32.04	≥ 99.9	Carlo Erba	64.7[a]
Ethanol	64-17-5	46.07	≥ 99.9	Carlo Erba	78.4[a]
Isopropanol	67-63.0	60.10	≥ 99.8	Honeywell	82.6 [a]
1-Propanol	71-23-8	60.10	≥ 99.5	Carlo Erba	97.0 [a]
2-Butanone	78-93-3	72.11	≥ 99.5	Sigma Aldrich	79.6 [a]
Ethyl Acetate	141-78-6	88.11	≥ 99.7	Chromaslv®	77.1 [a]
Acetonitrile	75-05-8	41.05	≥ 99.9	Sigma Aldrich	82.0 [a]
Dimethylformamide	68-12-2	73.09	≥ 99.9	Carlo Erba	153.0 [a]

[a]Data obtained from (David R. Lide 2003).

3.1.2. Experimental Procedure

As summarized in Table 2.1, all the authors employed the well-known shake-flask methodology in order to reach the solid-liquid equilibrium. This technique presents very good results for systems that aren't poorly soluble (Apley et al. 2015) and was selected in this work.

Moreover, different analytical techniques can be employed to evaluate the solubility in the solution after the equilibrium state is achieved. Despite the robustness of UV-Vis spectroscopy and HPLC approaches, the gravimetric method was selected. Although this technique takes longer to achieve the final results, they are quite accurate in systems that don't present insoluble solutes (Mota et al. 2010).

3.1.2.1. Isothermal Shake-Flask

The saturated solutions were prepared by mixing a small amount of solid in excess to the Erlenmeyer flasks containing between 70 and 80 ml of solvent and a magnetic stirrer bar. The flasks were placed on a plate stirrer inside a thermostatic bath (Lauda Instruments, model E20, Ecoline 025) operating with distilled water, as shown in Figure 3.1.

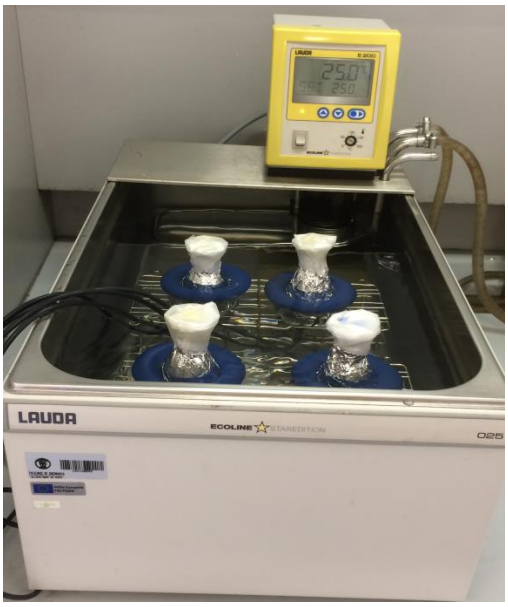


Figure 3.1: Experimental setup of the shake-flask methodology.

All the flasks were covered with aluminum foil to protect the solutions from possible light degradation.

Considering the wide range of shaking and settling times reported on Table 2.3, previous experiments were made to determine those conditions for the systems under study. It was found that 24 hours and 8 hours are the minimum times to shake and settle the solutions, respectively. The actual shaking and settling times employed for each solubility assay performed in this work are described in Table C.1 of Appendix C.

From the available literature data, the solubility of the solutes in water was expected to be lower than the solubility in organic solvents. Therefore, to be sure aqueous systems reached the equilibrium state in 24 hours, the flasks were first placed in an ultrasonic bath (Ultrasons-H, JP Selecta S.A.), for one hour, at the same temperature employed on the thermostatic bath (298.15 and 313.15 K).

In order to confirm that the solutions in the bath were saturated, the flasks were checked periodically and solid was added when necessary during the stirring process. For the systems with no data available in literature (systems with 2-butanone and DMF and all the systems involving gentisic and α-resorcylic acids), preliminary experiments were performed at ambient temperature (around 293 K) by placing the flasks directly over the plate stirrer (Magnetic Stirrer MSH-300N, BOECO Germany). After reaching a saturated solution with a small quantity of solid in excess, the flasks were placed in the bath. This procedure helped to achieve the equilibrium state more easily and minimized experimental errors.

After the settling period, three samples with volume varying between 1.5 and 5 ml were taken from the supernatant solution, using plastic syringes with metallic needles and placed in a previously weighted glass flask (± 10^{-4} g). The third sample was collected and filtered with polypropylene filters of 0,45 μm pore diameter. However, for some flasks containing a large amount of solid in excess, filters were used in the three samples, to prevent the transfer of any suspended particle. Figure 3.2 shows the material mentioned above to collect the samples from the solutions.

Figure 3.2: Syringes and filters used to collect the samples from the solutions.

Each week, the set of first experiments was performed at 298.15 K. Immediately after collecting three samples of the Erlenmeyer flask, additional solid was added to each system and the flasks returned to the thermostatic bath at 313.15 K.

3.1.2.2. Gravimetric Method

After the samples were taken from the solution and placed into small flasks, these were immediately covered with a screw cap. This procedure should be particularly fast for systems containing volatile solvents. Then, the flasks were weighted and put in a hood until all the visible solvent evaporated. Afterwards, the samples were transferred to a drying oven operating at 343.15 K, for at least 7 days. Then, the samples were taken from the drying oven and placed into a desiccator for 2 hours until their masses were registered again. This procedure was repeated for each sample until a constant mass was reached. The average required time to obtain completely dried samples was usually 20 days. However, dimethylformamide has a significantly higher boiling point than the other solvents and at least 30 days were necessary to achieve complete dryness.

In the gravimetric method, the solubility in weight fraction $S_{solution}$can be calculated by the following equation:

$$S_{solution} = \frac{m_{F+S} - m_F}{m_{F+Sol+C} - m_{F+C}} \quad (20)$$

where m_{F+S} is the mass of the flask plus the dry solid, m_F is the mass of the flask, m_{F+S+C} is the mass of the flask and cap plus the amount of collected solution and m_{F+C} is the mass of the flask and cap. The solubility in grams of solute per 100 grams of solvent can be calculated as follows:

$$S = \frac{S_{solution}}{1 - S_{solution}} * 100 \quad (21)$$

3.2. Results and Discussion

3.2.1. Melting temperature and enthalpy

In addition to the solubility studies, Differential Scanning Calorimetry (DSC) was also employed in this work to determine the melting point and the enthalpy of fusion of gallic acid, protocatechuic acid, gentisic acid and α- resorcylic acid. The analyses were performed at the University of Aveiro and the average results (3 samples per component) are summarized in Table 3.2 (the original DSC thermograms are shown in Appendix D).
Those parameters are essential to evaluate the solubility as shown in equation 11.

Table 3.2: Melting temperature and enthalpy determined experimentally via DSC, in this work.

Substance	***Tm* (K)**	**ΔH_{fus} (kJ/mol)**	**Observations**
Gallic Acid	524.2	$74.3 \leq \Delta H_{fus} \leq 79.44$	Very high value. It seems to be from the enthalpy of sublimation instead of fusion
Protocatechuic Acid	475.9	33.4 ± 0.7	It appears to have solid-solid transitions before reaching the melting point
Gentisic Acid	471.5 478.9	2.3 28.15 ± 1.3	Solid-solid transition
α- resorcylic acid	510.5	37.0 ± 1.5	

For comparison purposes, a literature review about the melting temperature and enthalpy of the selected solids was performed. The results found are shown in Table 3.3.

Table 3.3 Meting points, enthalpies of fusion and methodologies employed to obtain those parameters for gallic acid, protocatechuic acid, gentisic acid and α-resorcylic acid.

Compound	Tm (K)	ΔH_{fus} (kJ/mol)	Methodology	Reference
Gallic Acid	499	38.77	Marrero and Gani Group Contribution Model	(Mota et al. 2008)
	535	ND**	DSC	(Mota et al. 2008)
	524.2	62.38	DSC	(Jr et al. 2016)
Protocatechuic Acid	472.3 ± 1.6	31.2 ± 1.6	DSC	(Queimada et al. 2009)
	469.3	34.2	Marrero and Gani Group Contribution Model	(Queimada et al. 2009)
	474.8	33.5	DSC	(Vecchio 2013)
	474.9	34.0	DSC	(Vecchio & Brunetti 2011)
	474.9	NM*	DSC	(Price et al. 1999)
Gentisic Acid	476.2 ± 0.2	20.8 ± 1.7***	DSC	(Monte et al. 2010)
	478.9	NM*	DSC	(Price et al. 1999)
α- Resorcylic Acid	509.9	29.3	DSC	(Sarma et al. 2010)
	508.3 ± 0.2	38.3 ± 0.4	DSC	(Monte et al. 2010)
	508.9	NM*	DSC	(Price et al. 1999)

*Not measured; ** Not determined due to decomposition upon melting; *** Authors indicate that a phase transition seems to occur immediately followed by fusion.

In general, the measured melting temperatures of protocatechuic, gentisic and α-resorcylic acids are in close agreement with those found in literature. For gallic acid, Jr et al. (2016) report a similar value to the one found in this work.

Regarding the enthalpy of fusion, the literature values are more uncertain and less consistent with each other. The results found in this work for protocatechuic and α- resorcylic acids are close to those found by Vecchio (2013) and Monte et al. (2010), respectively.

The major divergence occurred with gallic acid, which started to decompose before reaching the melting point. Mota et al. (2008) tried to measure the enthalpy of fusion of gallic acid via DSC and also observed some degradation in the process. Other authors reported values that are considerably divergent from each other, which may indicate that DSC technique is unfeasible to measure the enthalpy of fusion of gallic acid.

Despite some of the measured enthalpies of fusion being close to those found in literature, the thermograms shown in Appendix G indicate very high values, which may include either enthalpies of sublimation or other parameters related to solid-solid transitions. Also, in

some cases, other peaks were detected, making it difficult to identify the enthalpy of fusion for the studied compounds.

3.2.2 Solubility in water and organic solvents

The experimental solubilities obtained in this work at 298.15 K and 313.15 K are shown in Table 3.4 and Table 3.5, respectively. More detailed information about those data is included in Tables E.1 to E.8. of Appendix E, namely, number of collected samples average solubility, standard deviation and coefficient of variation. The number of collected samples varied between 3 and 4 (depending on the amount of solid in excess in the Erlenmeyer flasks).

Table 3.4: Experimental solubilities (g of solute/100 g of solvent) of gallic acid, protocatechuic acid, gentisic acid and α-resorcylic acid in water and organic solvents at 298.15 K.

Solvent	Gallic Acid	Protocatechuic Acid	Gentisic Acid	α-Resorcylic Acid
Water	1.072 ± 0.001	1.293 ± 0.001	2.196 ± 0.001	10.176 ± 0.002
Methanol	38.623 ± 0.002	79.193 ± 0.007	67.565 ± 0.001	43.376 ± 0.001
Ethanol	23.732 ± 0.001	55.577 ± 0.002	45.503 ± 0.001	13.068 ± 0.001
Isopropanol	13.007 ± 0.001	45.146 ± 0.002	33.156 ± 0.002	12.823 ± 0.001
1-Propanol	10.585 ± 0.001	40.904 ± 0.001	35.277 ± 0.001	34.512 ± 0.001
2-Butanone	6.196 ± 0.001	49.272 ± 0.001	36.163 ± 0.001	3.621 ± 0.001
Ethyl Acetate	0.996 ± 0.001	7.894 ± 0.001	11.222 ± 0.001	3.317 ± 0.001
Acetonitrile	0.492 ± 0.001	5.910 ± 0.001	7.680 ± 0.001	3.271 ± 0.001
DMF	44.514 ± 0.009	60.728 ± 0.002	73.025 ± 0.007	47.774 ± 0.004

Table 3.5: Experimental solubilities (g of solute/100 g of solvent) of gallic acid, protocatechuic acid, gentisic acid and α-resorcylic acid in water and organic solvents at 313.15 K.

Solvent	Gallic Acid	Protocatechuic Acid	Gentisic Acid	α-Resorcylic Acid
Water	2.417 ± 0.001	3.046 ± 0.001	5.137 ± 0.001	22.452 ± 0.001
Methanol	41.472 ± 0.001	92.404 ± 0.003	78.613 ± 0.004	52.207 ± 0.001
Ethanol	24.522 ± 0.001	57.988 ± 0.001	51.607 ± 0.003	16.485 ± 0.001
Isopropanol	14.422 ± 0.001	50.261 ± 0.001	44.943 ± 0.001	16.083 ± 0.001
1-Propanol	11.697 ± 0.001	43.987 ± 0.001	40.499 ± 0.001	37.816 ± 0.001
2-Butanone	6.027 ± 0.001	50.444 ± 0.001	40.252 ± 0.001	4.395 ± 0.001

Ethyl Acetate	1.096 ± 0.001	12.991 ± 0.001	18.261 ± 0.001	5.064 ± 0.001
Acetonitrile	0.699 ± 0.001	10.785 ± 0.001	10.989 ± 0.001	5.344 ± 0.001
DMF	49.403 ± 0.002	67.401 ± 0.004	78.101 ± 0.010	53.318 ± 0.003

In order to better compare the solubility values shown in Tables 3.4 and 3.5, bubble graphics were built and are displayed in Figure 3.3.

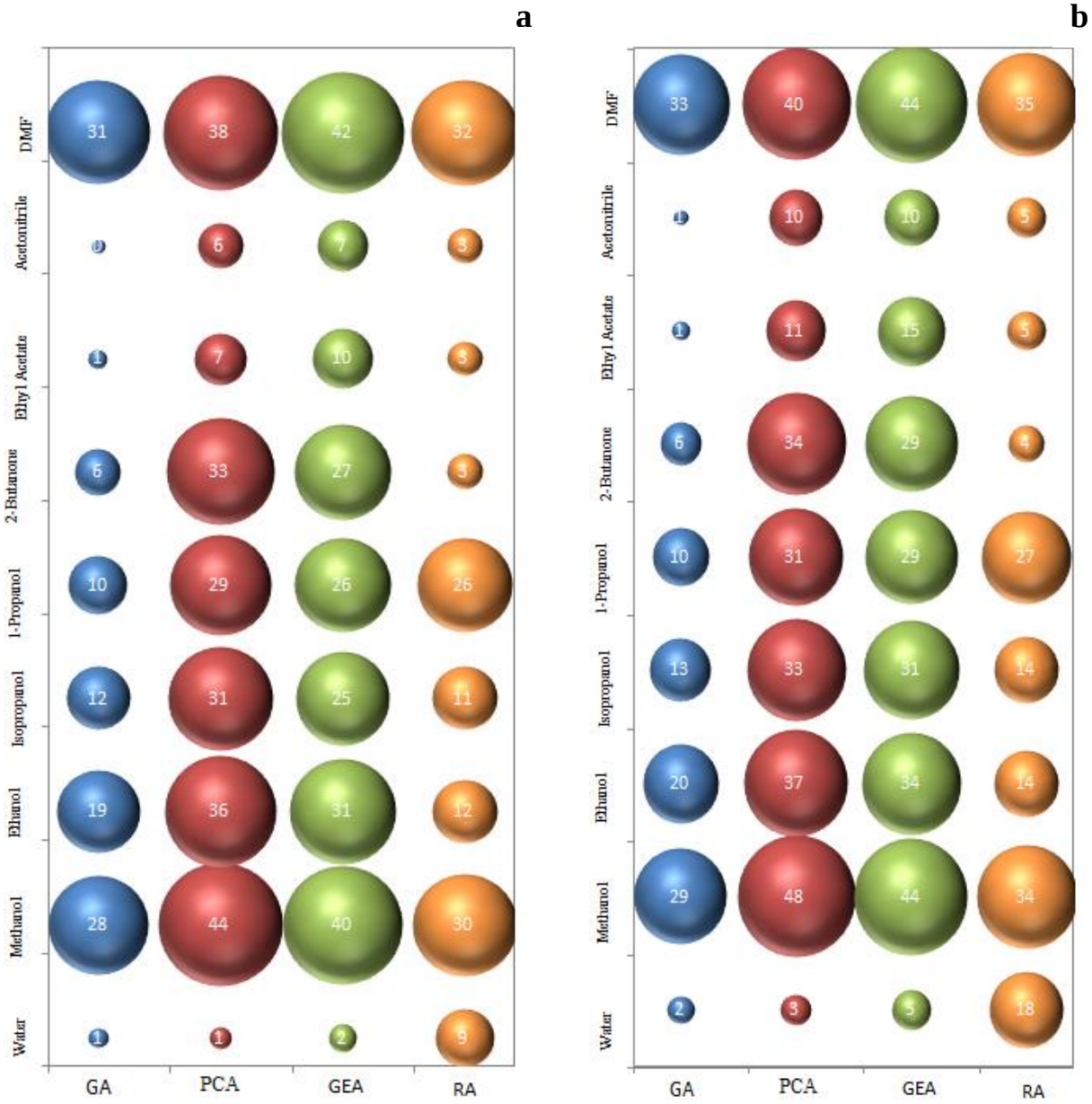


Figure 3.3: Experimental solubilities of gallic acid, protocatechuic acid, gentisic acid and α-resorcylic acid in different solvents at 278.2 K (a) and 313.2 K (b).

The binary systems containing protocatechuic and gentisic acids usually present the highest solubilities in the organic solvents. In the case of water, α-resorcylic acid is the most soluble solute, followed by gentisic, protocatechuic and gallic acids.

As expected, the solubility of these phenolic acids in short-chain alcohols is high, not only due the dispersion forces but also to the large number of hydrogen bonds formed between solute and solvent (the polarity parameters of all the selected compounds are shown in Tables F.1. and F.2. of Appendix F). In general, the solubility decreases with the increase of the alkyl chain of the alcohol with one exception: the solubility of α-resorcylic acid in 1-propanol is unexpectedly high and much larger than the solubility in isopropanol.

A peculiar behavior was also observed in the solubilities of the systems containing 2-butanone, acetonitrile and ethyl acetate. In those cases, solubilities of protocatechuic and gentisic acids are much higher than the solubilities of gallic and α-resorcylic acids.

The coefficients of variation of the experimental data are considerably low, being 1.99% and 1.60% the maximum values at 298.15 and 313.15 K, respectively (data shown in Tables E.1. and E.2.). This coefficient of variation is a statistical parameter that helps to evaluate the precision and the analytical method employed, which means that the results obtained by the shake-flask coupled to gravimetric method were quite accurate.

The percentages of solubility increase between 298.15 K and 313.15 K are shown in Table 3.6.

Table 3.6: Percentage of solubility increase from 298.15 K to 313.15 K for each binary system.

Solvent	Gallic Acid	Protocatechuic Acid	Gentisic Acid	α-Resorcylic Acid
Water	125.5	135.6	133.4	120.6
Methanol	7.4	16.7	16.4	20.4
Ethanol	3.3	4.3	13.4	26.2
Isopropanol	10.9	11.3	35.6	25.5
1-Propanol	10.5	7.5	14.8	9.6
2-Butanone	-2.7	2.4	11.3	21.4
Ethyl Acetate	10.0	64.6	62.7	52.7
Acetonitrile	42.1	82.5	43.1	63.4
DMF	0.11	0.11	0.07	0.12

In general, as temperature increases the solubility also increases. This variation is much stronger in the aqueous systems than in the organic solvents. The exception is the system formed by gallic acid and 2-butanone, for which a small decrease of 2.7 % in the solubility was detected, which is uncommon, but possible. However, the solubility's percentages of increase in 2-butanone systems are not very high (except for α-resorcylic acid), which may have led to oversaturation in the system containing gallic acid. Further studies should be performed to check the results obtained for this system.

On the other hand, although solubility values in dimethylformamide were high for the four addressed solutes, their percentages of increase from 298.15 K to 313.15 K were very small. The highest percentage of increase was obtained for protocatechuic acid in water (135.6 %), and the lowest occurred in the system composed by gentisic acid and DMF (0.07 %).

3.2.3. Comparison of the experimental solubilities to literature data

In order to better evaluate the experimental results obtained in this work, a comparison between them and the available literature data is made in Figures 3.3 to 3.5.

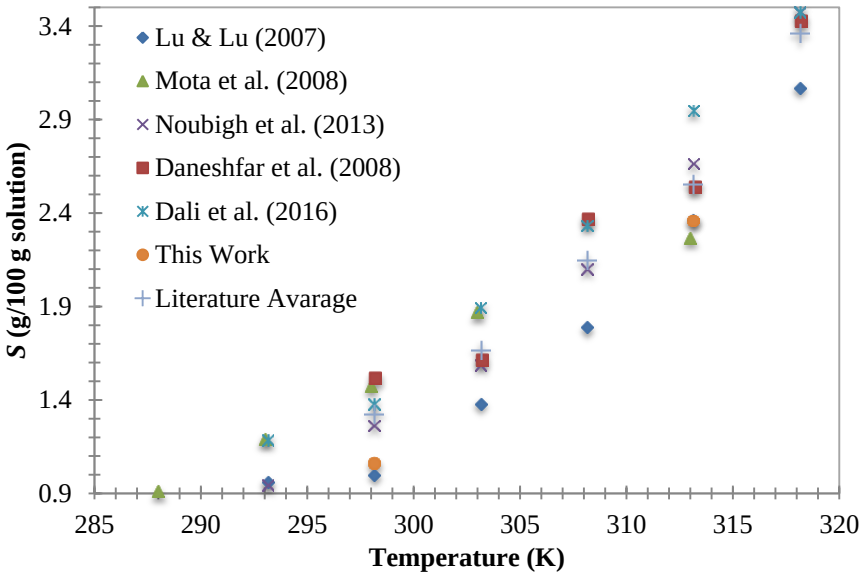


Figure 3.4: Comparison between experimental and literature data of gallic acid solubility in water.

As can be seen in Figure 3.3, the solubility data obtained in this work is lower than the average solubility data collected from literature, being close to those obtained by Lu and Lu (2007).

All the other authors but Mota et al. (2008) considered shaking and settling times lower than those assumed in this work, as shown in Table 2.1. In addition, the solubility data reported by Daneshfar et al. (2008) present a high scattering with temperature, which may indicate possible experimental errors.

Figures 3.4 and 3.5 compare the solubility of gallic acid and protocatechuic acid in organic solvents, respectively.

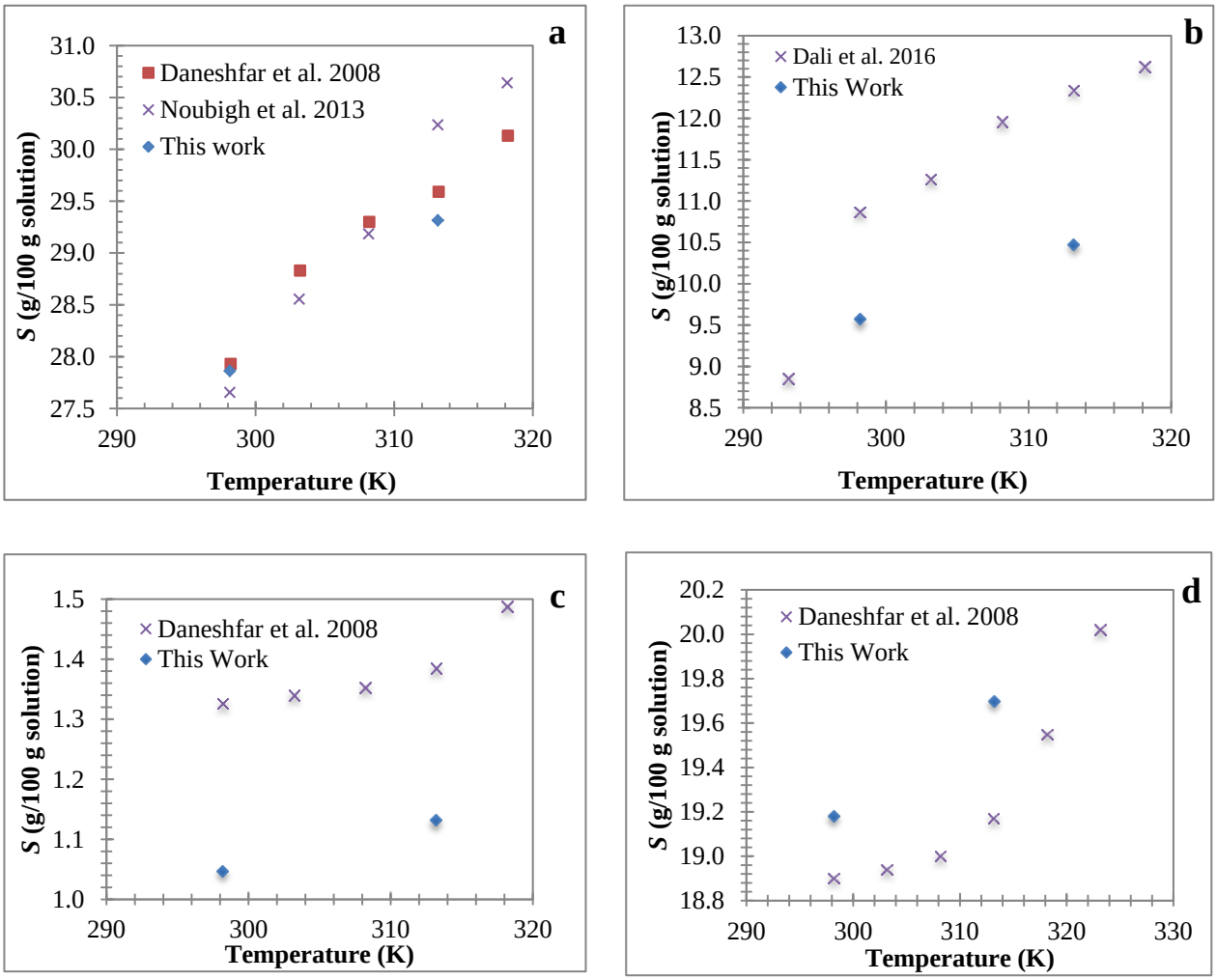


Figure 3.5: Comparison between experimental and literature data of gallic acid solubility in methanol (a), 1-propanol (b), ethyl acetate (c) and ethanol (d).

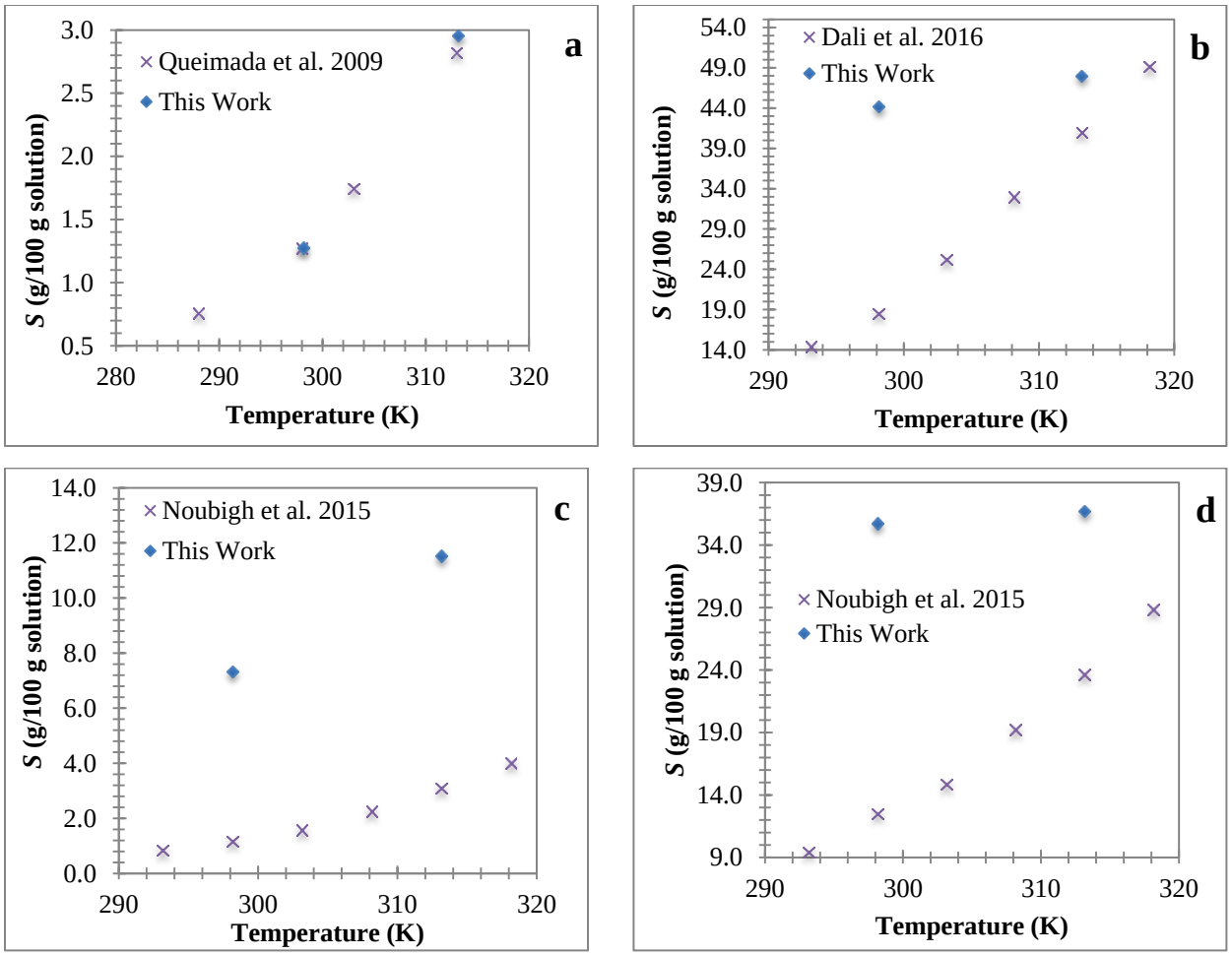


Figure 3.6: Comparison between experimental and literature data of protocatechuic acid solubility in water (a), methanol (b), ethyl acetate (c) and ethanol (d).

Figure 3.4 shows that the solubility of gallic acid in methanol, ethyl acetate and ethanol doesn't vary more than ± 0.6 g/100 g from the results obtained by Daneshfar et al. (2008).

For the binary systems formed by gallic acid and 1-propanol, the solubility values published by Dali et al. (2016) are higher than the values measured in this work. However, the literature data present a higher leap between the values at 293.15 and 298.15 K, which is not observed at higher temperatures. This behavior may be related to some oversaturation during their analysis.

On the other hand, for the systems containing protocatechuic acid, the reported literature results and the experimental measurements obtained in this work are close only for aqueous systems. The methodology employed by Queimada et al. (2009) considered high shaking and settling times (120-140 h and 24-48 h, respectively), which means that the equilibrium state was

certainly reached. On the other hand, the works published by Dali et al. (2016) and Noubigh et al. (2015) assumed shaking and settling times much lower than those employed in this work (shown in Table 2.1.). In those cases, the equilibrium state may not have been reached, which could explain the reason for the solubility values obtained by the authors in the binary systems containing protocatechuic acid and methanol, ethyl acetate and ethanol be inferior to the values obtained in this work, especially at 298.15 K. In addition, differences in the solid phase may also explain the different solubility values.

Chapter 4 Solubility Modeling

4.1. NRTL-SAC Programming

4.1.1. Methodology and Simulations Conditions

The NRTL-SAC model and the optimization of the parameters for the four selected compounds were implemented using the software MATLAB version R2013a. The selected optimization algorithm was the MATLAB routine "Isqnonlin", which is based on nonlinear least-squares curve fitting of the objective function (absolute value of the difference between the experimental and the calculated solubility data). This algorithm is based on the minimization of the following objective function:

$$min_x \|f(x)\| = min_x(f_1(x)^2 + f_2(x)^2 \ldots + f_n(x)^2) \quad (22)$$

where $f(x)$ is the relative difference between the experimental and the calculated solubility at each temperature, obtained by:

$$f(x) = \frac{x_{exp} - x_{cal}}{x_{exp}} \quad (23)$$

In order to reduce the number of calculations, bound constraints were considered. The minimum and maximum values of the NRTL-SAC parameters for all the simulations were set at 0.000 and 3.000, respectively. Those values were fixed considering that the NRTL-SAC molecular descriptors cannot be negative and are seldomly higher than 3.000.

The Reference Solvent Approach (RSA) proposed by Abildskov & O'Connell (2003) was also coupled to the NRTL-SAC method, as the available temperature and enthalpy of fusion are, in some cases, highly uncertain.

The main goal of the first set of simulations is to determine the four NRTL-SAC conceptual molecules' segments (X, Y^+, Y^-, Z) for gallic acid, protocatechuic acid, gentisic acid

and α-resorcylic acid using part of the solubility data measured here. After, those parameters will be used to predict the solubility in a different set of solvents.

4.2. Results and Discussion

4.2.1. Correlation

In order to evaluate the accuracy of the results obtained, the average relative deviations (ARD %) were calculated for each binary system as follows:

$$ARD(\%) = \frac{1}{NP}\sum_i \frac{\left|{x_i}^{exp} - {x_i}^{calc}\right|}{{x_i}^{exp}} * 100 \quad (24)$$

where *NP* is the number of data points, and ${x_i}^{exp}$ and ${x_i}^{calc}$ are the experimental and calculated solubility in mole fraction, respectively.

The first optimization approach involved the application of Equation 15, using the values of the melting properties presented in Table 3.2. While performing the simulations, it wasn't possible to converge to a set of parameters that could correlate the data. This is probably due to the high values found for the enthalpy of fusion, which may represent other phenomena instead of the melting point. Therefore, the Reference Solvent Approach was adopted as a second strategy to describe the solid-liquid equilibria.

For this correlation step, water, methanol, ethanol, isopropanol, 2-butanone, acetonitrile and ethyl acetate were selected. After, the parameters found were used to predict the solubility in dimethylformamide and isopropanol.

In the gallic acid simulations, 2-butanone was discarded due to the decrease of the solubility between 298.15 K and 313.15 K.

Table 4.1 shows the optimized segment parameters, the selected reference solvent, the number of solvents employed, the general ARD% and the system outlier for each solute studied in this work. More detailed information about the ARD% found per binary system as well as predicted solubility and activity coefficient data are shown in Table G.1 from Appendix G.

Table 4.1: NRTL-SAC parameters, RSA, system Outlier, number of solvents and ARD (%) for each simulation.

Compound	X	Y-	Y+	Z	RSA	Outlier	NS	ARD (%)
Gallic Acid	0.496	0.430	0.000	2.290	Acetonitrile	Ethyl Acetate	6	29
Protocatechuic Acid	0.579	1.080	0.000	0.726	Acetonitrile	2-Butanone	7	29
Gentisic Acid	1.525	0.037	0.530	1.838	2-Propanol	Water	7	25
α-Resorcylic Acid	0.188	0.139	0.000	1.044	Acetonitrile	Ethyl Acetate	7	34

From all the solvents evaluated, acetonitrile was employed as the reference solvent three times, for gallic, protocatechuic and α-resorcylic acids. For the simulations performed with gentisic acid, 2-propanol was the reference solvent that presented the minimum ARD.

The correlation results lead to conclusion that NRTL-SAC is an adequate model to estimate solubility of the studied compounds, with minimum and maximum ARD values of 25 and 34 % for gentisic and α-resorcylic acids, respectively. Queimada et al. (2009) and Mota et al. (2012) also employed NRTL-SAC to predict solubility of drug molecules, such as salicylic acid, benzoic acid, paracetamol and furosemide, and reported ARD values of 67% and 29%, respectively, which are similar to those found in this work.

In general, the models adjusted the solubility data better in systems containing alcohols compared to those containing ethyl acetate and 2-butanone. Good correlation results were also obtained for the solubility in water of all solutes with the exception of gentisic acid.

Figures 4.1 to 4.4 compare the experimental and calculated solubility data for each binary system as a function of temperature.

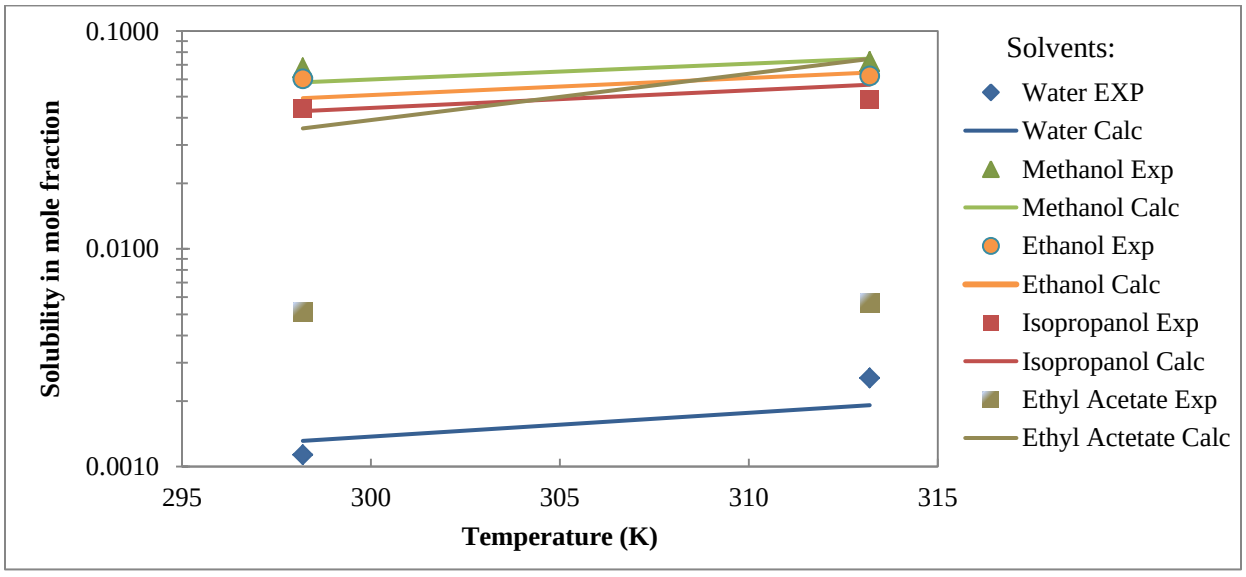


Figure 4.1: NRTL-SAC prediction results for gallic acid.

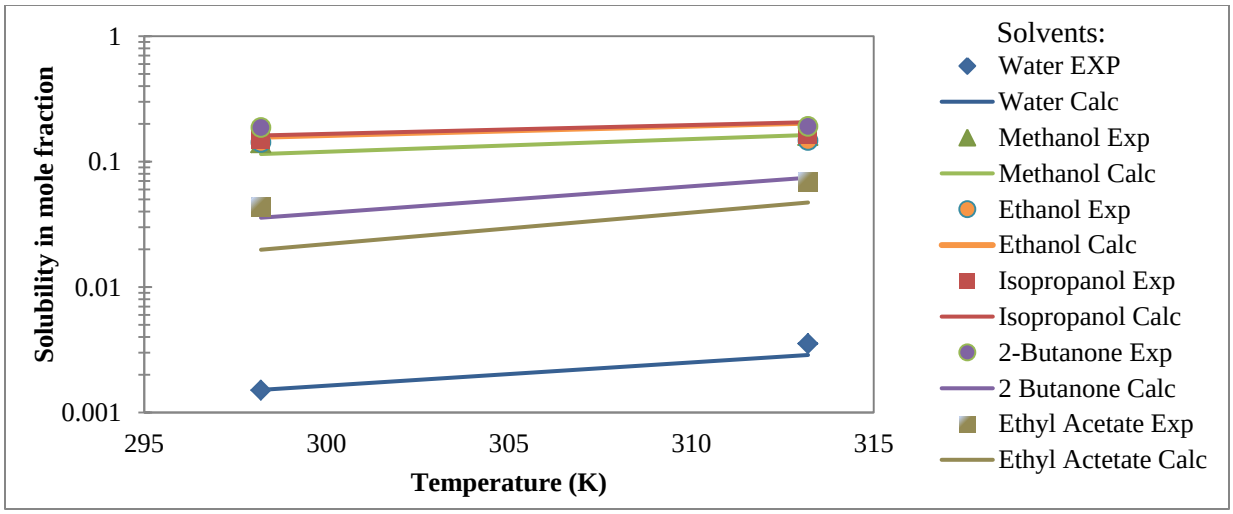


Figure 4.2: NRTL-SAC prediction results for protocatechuic acid.

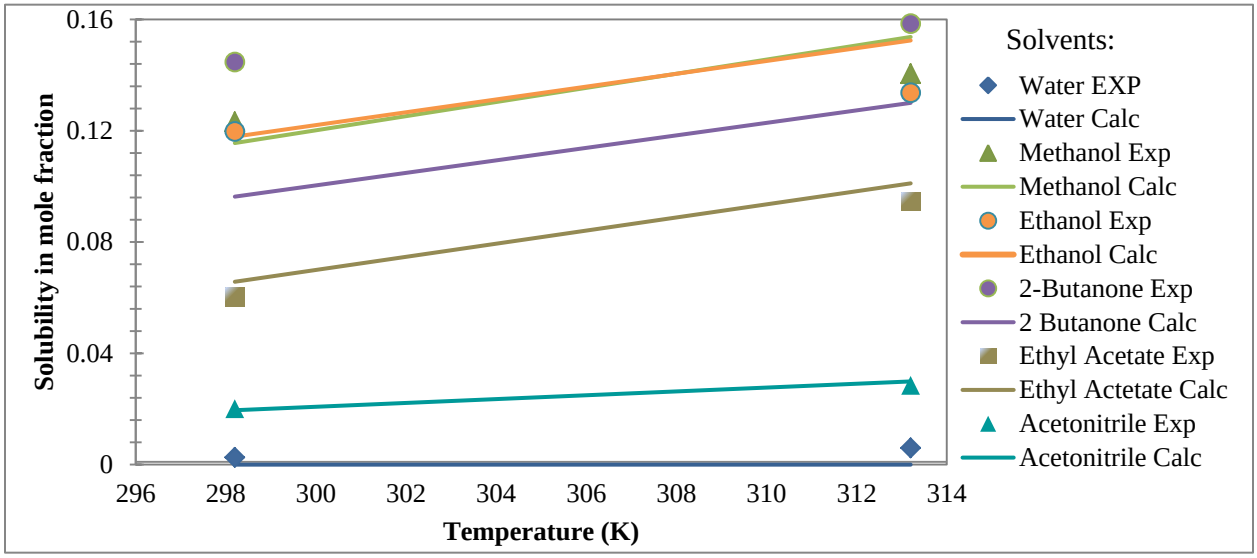


Figure 4.3: NRTL-SAC prediction results for gentisic acid.

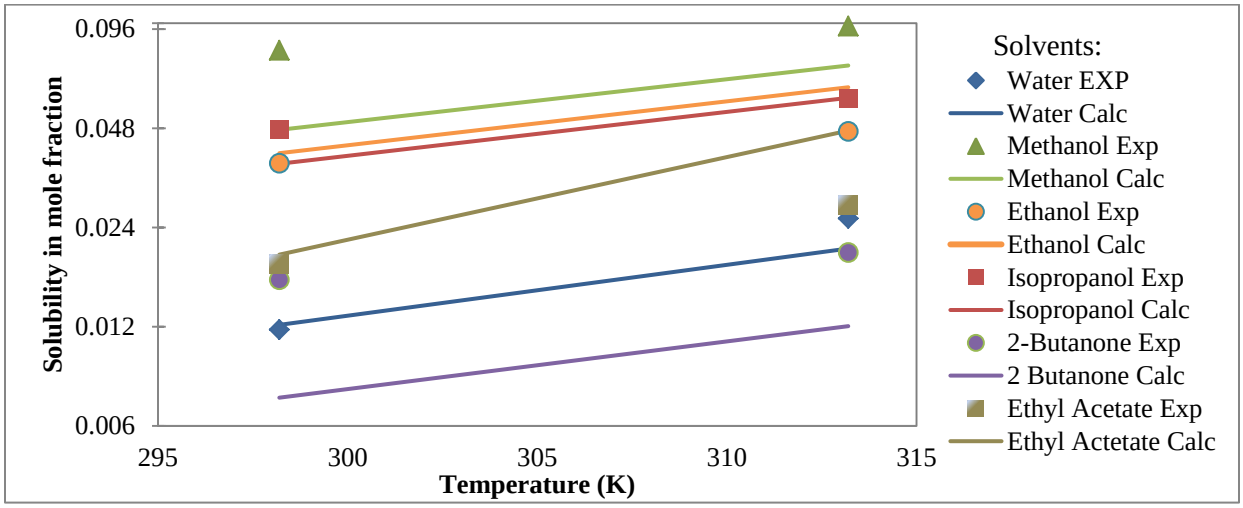


Figure 4.4: NRTL-SAC prediction results for α-resorcylic acid.

For the gallic acid systems, the best fitted solvent was ethanol, with a maximum deviation of 9% and all the solvents but ethyl acetate presented ARD lower than 21%. However, the outlier presented a very high deviation of 95 %.

For the simulations performed with protocatechuic acid, 2-butanone was the outlier, with 71% of maximum deviation. In this case, water and methanol presented the best calculated results with 10% of ARD.

The simulations performed with gentisic acids exhibited peculiar results. In this case, all the solubility data were adjusted very well in systems composed by organic solvents (all ARD were inferior to 25%). However, the system presented a maximum deviation of 99.99 % for the solubility calculated in aqueous system, predicting a value close to zero.

For α-resorcylic acid, ethyl acetate was the outlier, presenting a deviation of 80% and water and 2-propanol were the solvents best fitted in terms of the calculated solubility, showing each an ARD of 11%.

The predicted activity coefficients, shown in Table G.1, are higher than 1 for systems that present low solubilities, such as those containing water, ethyl acetate and acetonitrile. For those systems, the interactions between the solute and the solvent are weak, causing lower solubilities. On the other hand, systems containing alcohols usually presented predicted activity coefficients lower than 1, which means in these systems, the solute-solvent interactions are strong provoking an increase in the solid solubility.

4.2.2. Prediction

Once the NRTL-SAC segment parameters were obtained, the model can be used to estimate the solubility of the same solutes in different solvents. Figure 4.5 shows the predicted solubility of the four solutes in 1-propanol and dimethylformamide (DMF).

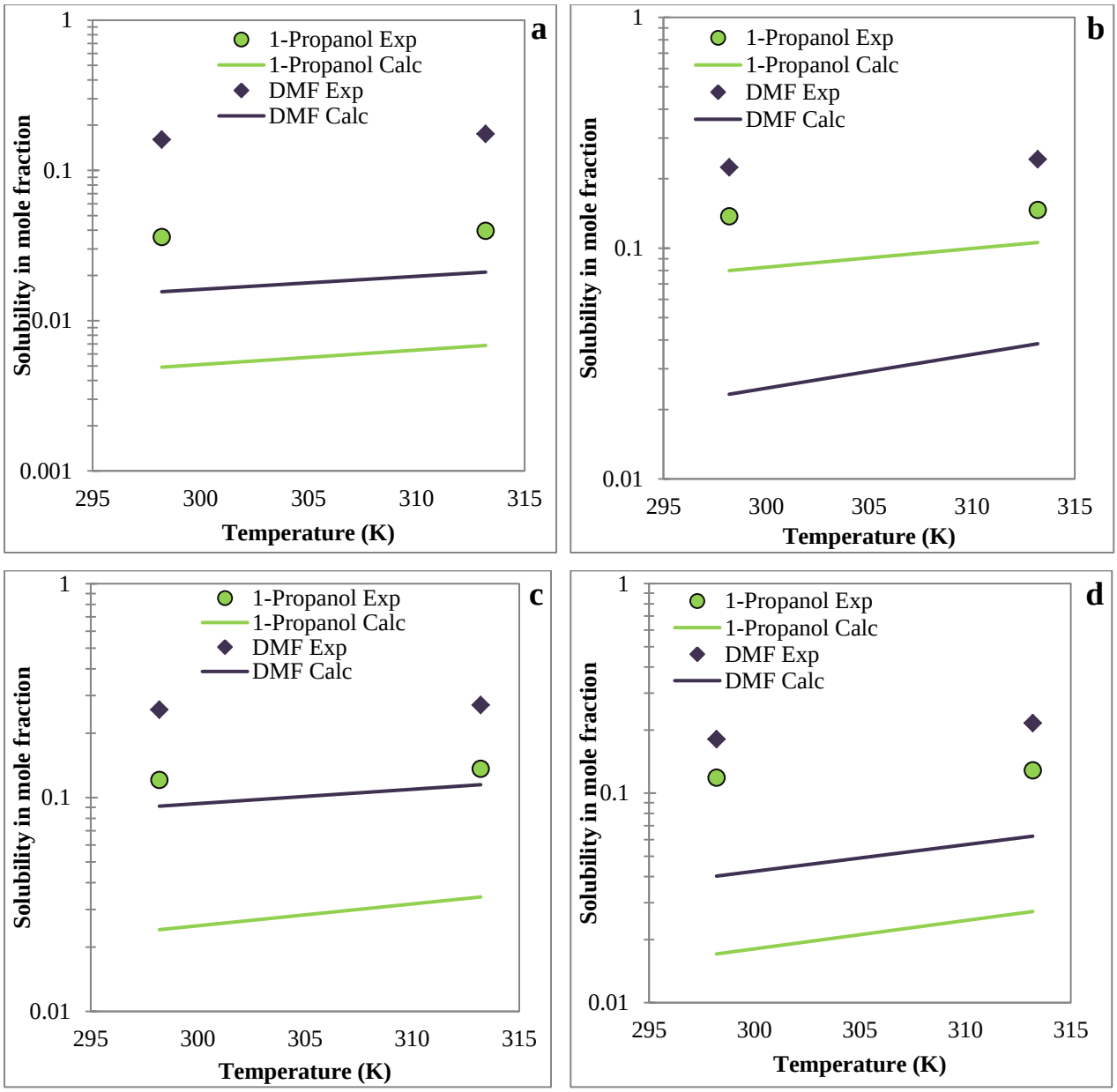


Figure 4.5: Predicted solubility data obtained through NRTL-SAC for gallic acid (a), protocatechuic acid (b), gentisic acid (c), and α-resorcylic acid (d).

The ARD% found for 1-propanol and DMF were 70 and 78%, respectively. More detailed information is displayed in Table G.2 of Appendix G. Although Figure 4.5 shows that the predicted solubility data were not as good as the calculated solubility data obtained in the optimizations of the NRTL-SAC parameters, Queimada et al. (2009) and Mota et al. (2010) also presented results containing ARD% higher than 70% for the predicted solubilities of binary systems containing drug molecules, such as allopurinol, ibuprofen and estradiol.

In the case of the systems containing 1-propanol, there might have been some oversaturation for gallic acid and α-resorcylic acids (as mentioned before), which may be the cause of the high difference observed between the experimental and the estimated solubilities. In the future, further experimental assays will be performed with this solvent to verify the experimental values pointed in this work.

The systems containing DMF presented very high experimental solubility values for all the analyzed solutes. In addition, the crystals formed during the crystallization processes presented different coloration from those observed in other systems, which may indicate either that the solutes reacted with the solvents or the formation of distinct solid phases. X-ray analysis will be performed in the future in order to characterize the solids.

Finally, Figure 4.6 shows the big picture by presenting the calculated solubility using the NRTL-SAC model as a function of the experimental solubility data (the predicted solubility data for the system formed by gentisic acid and water were disregarded due to their low log values). All the solubility information used to plot the graphics in Figure 4.6 are displayed in Tables G.1 and G.2 of Appendix G.

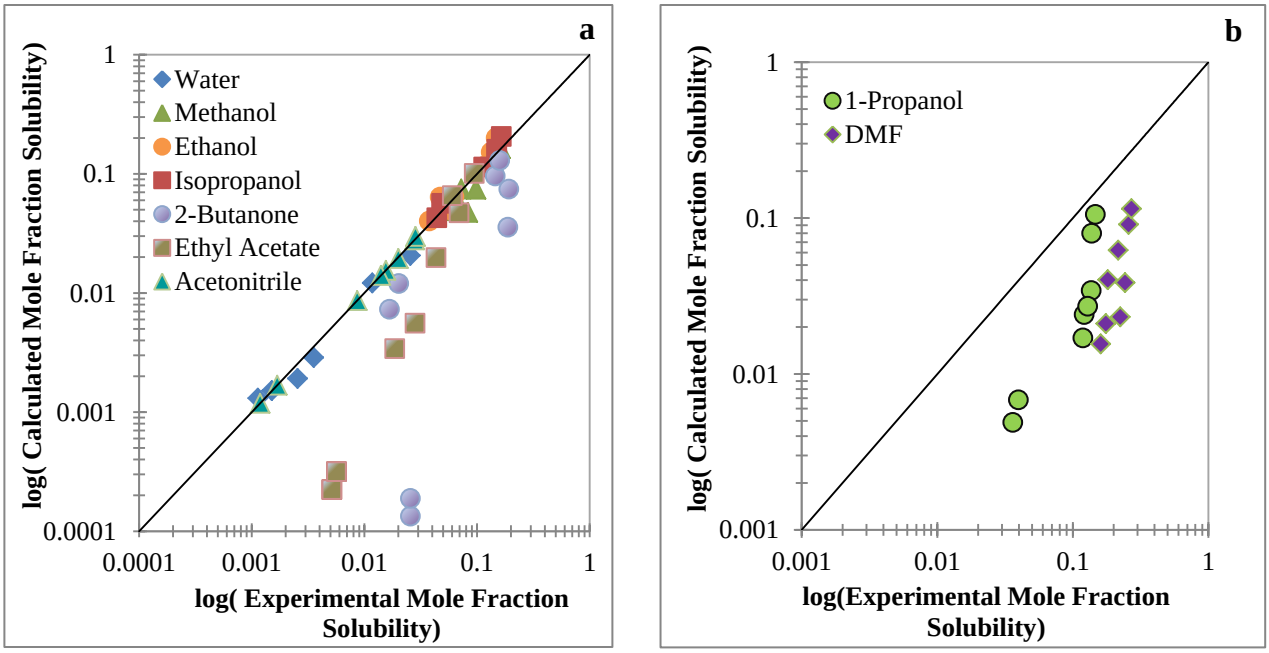


Figure 4.6: Comparison between experimental and predicted solubility for solvents used in the determination of the NRTL-SAC segment descriptors (a) and other organic solvents (b).

As can be seen in Figure 4.6, in general, the calculated solubilities that presented the highest deviations were inferior to the experimental solubility data. From the seven binary systems employed to optimize the NRTL-SAC parameters for the studied solutes, those containing ethyl acetate and 2-butanone are the more difficult to be correlated. Furthermore, predicted solubilities values for binary systems containing 1-propanol usually have lower deviations than binary systems containing dimethylformamide.

As shown in Table G.2 from Appendix G, almost all the systems containing 1-propanol and DMF presented estimated activity coefficients lower than 1 (the only exception was for the binary system formed by α-resorcylic acid and 1-propanol). Due to the high experimental solubility presented by those systems, the interactions between the studied phenolic acids and the mentioned solvents are strong, causing the activity coefficients to be lower than 1.

Chapter 5 Conclusions and Future Work

In this work, the solubility of gallic acid, protocatechuic acid, gentisic acid and α-resorcylic acid was experimentally measured in water and different organic solvents (methanol, ethanol, 1-propanol, isopropanol, 2-butanone, ethyl acetate, acetonitrile and dimethylformamide), at 298.2 and 313.2 K. The shake-flask methodology was applied using the gravimetric method as analysis method with good results. In general, an increase in the temperature indicates an increment in the solubility of the binary systems.

The coefficients of variation of all the experimental assays were lower than 2%, which are very acceptable. The solubility data generally exhibited the same pattern for alcohols, presenting the highest solubilities for those having the shortest carbon chain. An exception occurred for the system composed by α-resorcylic acid and 1-propanol, for which the solubility values were much higher than those obtained for binary systems containing ethanol and isopropanol. Further experiments should be performed in the future to corroborate the values obtained in this work.

Melting points and enthalpies of fusion were also determined by Differential Scanning Calorimetry (DSC) for the phenolic compounds addressed in this work. The results for the melting temperature were consistent with literature values. Regarding the melting enthalpy, a high uncertainty is associated with the measured values as solid phase transitions or sublimation phenomena may interfere with the values obtained. .

The second part of this work consisted in the thermodynamic modelling of the experimental data measured in this work by applying the Non-Random Two Liquid Segment Activity Coefficient (NRTL-SAC) model. NRTL-SAC was selected because this framework has shown great robustness and accuracy to predict the solid-liquid equilibria of a wide range of systems. In this work, the model presented acceptable correlation results with average relative deviation (ARD) varying between 25 and 34%. After, the model was used to predict the solubility in 1-propanol and dimethylformamide and the ARD% were 70 and 78%, respectively.

Those values are satisfactory for semi-predictive models, using a limited set of solvents. From the solvents employed in the simulations, the solubility in alcohols and acetonitrile was better correlated by the model, presenting the lowest ARD values.

Considering that no previous simulations were reported until this moment for the studied phenolic acids, the NRTL-SAC segment descriptors can contribute for future predictions in different systems. For future work, further experimental solubility measurements in different binary and multicomponent systems are suggested, in order to provide more robustness to the optimized parameters. In addition, other thermodynamic frameworks, such as UNIFAC, UNIQUAC and their variations could also be implemented.

In order to better understand the solubility behavior of those phenolic acids in water and organic solvents, measurements of the potential of hydrogen (pH) can be performed in the future, as well as solid-phase characterizations. By knowing more detailed about the solid-solid transactions, a better understating of the solubility behavior of the studied compounds could be accomplished. In addition, studies on the solubility of other phenolic acids can also be performed to aid in the understanding of molecular structure's influence in the solubility of those compounds.

Bibliographic References

Abildskov, J. & O'Connell, J.P., 2003. Predicting the solubilities of complex chemicals I. Solutes in different solvents. *Industrial & Engineering Chemistry Research*, 42(22), pp.5622–5634. Available at: papers2://publication/doi/10.1021/ie030210v.

Advanced Chemistry Development, I. (ACD/Labs), 2017. Percepta Predictors.

Apley, M. et al., 2015. Determination of thermodynamic solubility of active pharmaceutical ingredients for veterinary species: A new USP general chapter. *Dissolution Technologies*, 41, pp.36–39.

Aspen Technology, 2010. Thermodynamic Property Models. *Aspen Physical Property System - Physical Property Models*, pp.12–184.

Baka, E., Comer, J.E.A. & Takács-Novák, K., 2008. Study of equilibrium solubility measurement by saturation shake-flask method using hydrochlorothiazide as model compound. *Journal of Pharmaceutical and Biomedical Analysis*, 46(2), pp.335–341.

Chen, C.-C. & Song, Y., 2004. Solubility modeling with NRTL segment activity coefficient model. *Ind. Eng. Chem. Res*, 43, pp.8354–8362.

Chen, C.C. et al., 2008. Correlation and prediction of phase behavior of organic compounds in ionic liquids using the nonrandom two-liquid segment activity coefficient model. *Industrial and Engineering Chemistry Research*, 47(18), pp.7081–7093.

Chen, C.C. & Crafts, P.A., 2006. Correlation and prediction of drug molecule solubility with the NRTL-SAC model. *Computer Aided Chemical Engineering*, 21(C), pp.859–864.

Dali, I. et al., 2016. Correlation and semi-empirical modeling of solubility of gallic acid in different pure solvents and in binary solvent mixtures of propan-1-ol + water, propan-2-ol + water and acetonitrile + water from (293.2 to 318.2) K. *Journal of Molecular Liquids*, 222, pp.503–519.

Daneshfar, A., Ghaziaskar, H.S. & Homayoun, N., 2008. Solubility of Gallic Acid in Methanol , Ethanol , Water , and Ethyl Acetate. *Journal of Chemical and Engineering Data*, 53, pp.776–778.

David R. Lide, 2003. *Handbook of Chemistry and Physics* 84th ed., CRC Press.

Fakhraian, H. et al., 2016. Ternary Phase Diagram Modeling of Chiral Medetomidine Salts. *Physical Chemistry Research*, 4(4), pp.663–671.

Farajnezhad, A. et al., 2016. Correlation of interaction parameters in Wilson, NRTL and UNIQUAC models using theoretical methods. *Fluid Phase Equilibria*, 417, pp.181–186.

Feng, Y. et al., 2014. (Solid + liquid) phase equilibria of tetraphenyl piperazine-1, 4-diyldiphosphonate in pure solvents. *Journal of Chemical Thermodynamics*, 78(34670), pp.143–151.

Ferguson, L.R., Zhu, S. & Harris, P.J., 2005. Antioxidant and antigenotoxic effects of plant cell wall hydroxycinnamic acids in cultured HT-29 cells. , 450, pp.585–593.

Ferreira, O. & Pinho, S.P., 2012. Solubility of Flavonoids in Pure Solvents. *Industrial & Engineering Chemistry Research*, 51(September 2012), pp.6586–6590.

Ferreira, O., Schröder, B. & Pinho, S.P., 2013. Solubility of hesperetin in mixed solvents. *Journal of Chemical and Engineering Data*, 58(9), pp.2616–2621.

Frank, T.C., Downey, J.R. & Gupta, S.K., 1999. Quickly screen solvents for organic solids. *Chemical engineering progress*, 95(12), pp.41–61.

Gao, C. & Yan, D., 2004. Hyperbranched polymers: From synthesis to applications. *Progress in Polymer Science (Oxford)*, 29(3), pp.183–275.

Hansen, C.M., 2013. *Hansen Solubility Parameters: A User's Handbook* Second Edi., CRC Press.

Hatzipanayioti, D. & Petropouleas, P., 2010. Theoretical and experimental investigation of the semiquinone forms of protocatechuic acid. The effect of manganese. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 75(3), pp.997–1007.

Hefter, G.T. & Tomkins, R.P.T., 2003. *The Experimental Determination of Solubilities* Volume 6., WILEY.

Herzog, K.A. & Swarbrick, J., 1971. Drug Permeation through Thin Model Membranes II: Permeation Characteristics of a Polymeric Model Biomembrane. *Journal of Pharmaceutical Sciences*, 60(3), pp.393–395.

Jr, W.A., Chickos, J.S. & Chickos, J.S., 2016. Phase Transition Enthalpy Measurements of Organic and Organometallic Compounds . Sublimation , Vaporization and Fusion Enthalpies From 1880 to 2015 . Part 1 . C1 − C10 Phase Transition Enthalpy Measurements of Organic and Organometallic Compounds . Sublim. , 33101. Available at: http://dx.doi.org/10.1063/1.4948363.

Kampa et al., 2004. Antiproliferative and apoptotic effects of selective phenolic acids on T47D

human breast cancer cells: potential mechanisms of action. *Breast cancer research : BCR*, 6(2), pp.R63-74.

Kim, S. et al., 2016. PubChem Substance and Compound databases. Available at: https://pubchem.ncbi.nlm.nih.gov [Accessed June 20, 2017].

Letcher, T.M., Pinho, S.P. & Macedo, E.A., 2007. *Developments and Applications in Solubility* First Edit., PSC Publishing.

Lin, M., Tesconi, M. & Tischler, M., 2009. Use of 1H NMR to facilitate solubility measurement for drug discovery compounds. *International Journal of Pharmaceutics*, 369(1–2), pp.47–52.

Liu, C. et al., 2012. 3,5-Dihydroxybenzoic acid, a specific agonist for hydroxycarboxylic acid 1, inhibits lipolysis in adipocytes. *The Journal of pharmacology and experimental therapeutics*, 341(3), pp.794–801.

de Lorgeril, M. et al., 1999. Mediterranean diet, traditional risk factors, and the rate of cardiovascular complications after myocardial infarction: final report of the Lyon Diet Heart Study. *Circulation*, 99(6), pp.779–785.

Lu, L.L. & Lu, X.Y., 2007. Solubilities of gallic acid and its esters in water. *Journal of Chemical and Engineering Data*, 52(1), pp.37–39.

Mansour, S.H. et al., 2005. Hyperbranched polyesters based on polycondensation of 1,3,5-tris(2-hydroxyethyl) cyanuric acid and 3,5-dihydroxybenzoic acid. *Journal of Polymer Science, Part A: Polymer Chemistry*, 43(15), pp.3278–3288.

Martins, C.R., Lopes, W.A. & Andrade, J.B. de, 2013. Solubilidade das Substâncias Orgânicas. *Quimica Nova*, 36(8), pp.1248–1255.

Monte, M.J.S., Gonçalves, M. V. & Ribeiro Da Silva, M.D.M.C., 2010. Vapor pressures and enthalpies of combustion of the dihydroxybenzoic acid isomers. *Journal of Chemical and Engineering Data*, 55(6), pp.2246–2251.

Mota, F.L. et al., 2012. Calculation of drug-like molecules solubility using predictive activity coefficient models. *Fluid Phase Equilibria*, 322–323, pp.48–55.

Mota, F.L. et al., 2011. Solubility of drug-like molecules in pure organic solvents with the CPA EoS. *Fluid Phase Equilibria*, 303(1), pp.62–70.

Mota, F.L. et al., 2009. Temperature and solvent effects in the solubility of some pharmaceutical compounds: Measurements and modeling. *European Journal of Pharmaceutical Sciences*,

37(3–4), pp.499–507.

Mota, F.L., Maria, E.R.M. & Pinho, S.P., 2010. *Solubility Studies with Pharmaceutical Applications*. Faculdade Engenharia Universidade do Porto.

Mota, L., Queimada, J. & Pinho, P., 2008. Aqueous Solubility of Some Natural Phenolic Compounds Fa. *Ind. Eng. Chem. Res. 2008*, 47(5182–5189), pp.5182–5189.

Nafees, S. et al., 2012. Modulatory effects of gentisic acid against genotoxicity and hepatotoxicity induced by cyclophosphamide in Swiss albino mice. *Journal of Pharmacy and Pharmacology*, 64(2), pp.259–267.

Nouar, A. et al., 2016. Solubility Prediction of Active Pharmaceutical Compounds with the UNIFAC Model. *International Journal of Thermophysics*, 37(3), pp.1–15.

Noubigh, A. et al., 2013. Measurement and correlation of the solubility of gallic acid in methanol plus water systems from (293.15 to 318.15) K. *Journal of Molecular Liquids*, 187, pp.226–229.

Noubigh, A., Aydi, A. & Abderrabba, M., 2015. Experimental Measurement and Correlation of Solubility Data and Thermodynamic Properties of Protocatechuic Acid in Four Organic Solvents. *Journal of Chemical and Engineering Data*, 60, pp.514–518.

Obied, H.K. et al., 2005. Investigation of Australian Olive Mill Waste for Recovery of Biophenols. *Journal of Agricultural and Food Chemistry*, 53(26), p.10.

De Oliveira, D.M. & Bastos, D.H.M., 2011. Biodisponibilidade de ácidos fenólicos. *Quimica Nova*, 34(6), pp.1051–1056.

Ow, Y.-Y. & Stupans, I., 2003. Gallic acid and gallic acid derivatives: effects on drug metabolizing enzymes. *Current drug metabolism*, 4, pp.241–248.

Paula, J.T. et al., 2016. Solubility of protocatechuic acid , sinapic acid and chrysin in supercritical carbon dioxide. *The Journal of Supercritical Fluids*, 112, pp.89–94.

Pistikopoulos, E.N., Galindo, A. & Dua, V., 2010. *Molecular Systems Engineering*,

Prausnitz, J.M., Lichtenthaler, R.N. & de Azevedo, E.G., 1999. 6.11 Wilson, NRTL, and UNIQUAC Equations. *Molecular Thermodynamics of Fluid-Phase Equilibria*, pp.260–261.

Price, D.M., Bashir, S. & Derrick, P.R., 1999. Sublimation properties of x,y-dihydroxybenzoic acid isomers as model matrix assisted laser desorption ionisation (MALDI) matrices. *Thermochimica Acta*, 327(1–2), pp.167–171. Available at: http://wrap.warwick.ac.uk/14754/%5Cnhttp://linkinghub.elsevier.com/retrieve/pii/S004060

3198006066.

Qin, Y. et al., 2015. Separation of phenolic acids and flavonoids from Trollius chinensis Bunge by high speed counter-current chromatography. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, 1001, pp.82–89.

Queimada, A.J. et al., 2009. Solubilities of biologically active phenolic compounds: Measurements and modeling. *Journal of Physical Chemistry B*, 113(11), pp.3469–3476.

Ren, D.B. et al., 2013. Correlation and prediction of partition coefficient using nonrandom two-liquid segment activity coefficient model for solvent system selection in counter-current chromatography separation. *Journal of Chromatography A*, 1301, pp.10–18.

Ren, D.B. et al., 2015. Systematic and practical solvent system selection strategy based on the nonrandom two-liquid segment activity coefficient model for real-life counter-current chromatography separation. *Journal of Chromatography A*, 1393, pp.47–56.

Renon, H. & Prausnitz, J.M., 1968. Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE Journal*, 14(1), pp.135–144.

Rowlinson, J.., 1970. *Molecular thermodynamics of fluid-phase equilibria*,

Sarma, B., Sanphui, P. & Nangia, A., 2010. Polymorphism in isomeric dihydroxybenzoic acids. *Crystal Growth and Design*, 10(5), pp.2388–2399.

Sevillano, D.M. et al., 2013. Model comparison for the prediction of the solubility of green tea catechins in ethanol/water mixtures. *Industrial and Engineering Chemistry Research*, 52(17), pp.6039–6048.

Shefter, E. & Higuchi, T., 1963. Dissolution behavior of crystalline solvated and nonsolvated forms of some pharmaceuticals. *Journal of Pharmaceutical Sciences*, 52(8), pp.781–791.

Sheikholeslamzadeh, E., Chen, C.C. & Rohani, S., 2012. Optimal solvent screening for the crystallization of pharmaceutical compounds from multisolvent systems. *Industrial and Engineering Chemistry Research*, 51(42), pp.13792–13802.

Sheikholeslamzadeh, E. & Rohani, S., 2012. Solubility prediction of pharmaceutical and chemical compounds in pure and mixed solvents using predictive models. *Industrial and Engineering Chemistry Research*, 51(1), pp.464–473.

Song, Y. & Chen, C.C., 2009. Symmetric nonrandom two-liquid segment activity coefficient model for electrolytes. *Industrial and Engineering Chemistry Research*, 48(11), pp.5522–5529.

Spieweck, F. & Bettin, H.&, 1992. *Review: Solid and liquid density determination*,

Srinivas, K. et al., 2009. Optimization of subcritical fluid extraction of bioactive compounds using hansen solubility parameters. *Journal of Food Science*, 74(6), pp.342–354.

Strupat, K.; Karas, M.; Hillenkamp, F., 1991. 2,5- Dihydroxybenzoic Acid: A New Matrix for Laser Desorption-Ionization Mass Spectrometry. *Int. J. Mass Spectrom.*, 111, pp.89–102.

Tang, W. et al., 2014. Experimental determination and computational prediction of androstenedione solubility in alcohol + water mixtures. *Industrial and Engineering Chemistry Research*, 53(28), pp.11538–11549.

Valavi, M., Svard, M. & Rasmuson, A.C., 2016. Prediction of the Solubility of Medium-Sized Pharmaceutical Compounds Using a Temperature-Dependent NRTL-SAC Model. *Industrial and Engineering Chemistry Research*, 55(42), pp.11150–11159.

Vecchio, S., 2013. Thermogravimetric method for a rapid estimation of vapor pressure and vaporization enthalpies of disubstituted benzoic acids: An attempt to correlate vapor pressures and vaporization enthalpies with structure. *Structural Chemistry*, 24(6), pp.1821–1827.

Vecchio, S. & Brunetti, B., 2011. Thermochemical study of 2,4-, 2,6- and 3,4-dihydroxybenzoic acids in the liquid phase using a TG apparatus. *Thermochimica Acta*, 515(1–2), pp.84–90. Available at: http://dx.doi.org/10.1016/j.tca.2011.01.001.

Vermerris, W. & Nicholson, R., 2009. *Phenolic Compound Biochemistry*, Springer.

Vrsalović, L. et al., 2010. Corrosion inhibition of Al-Mg alloy by gentisic acid. *Indian Journal of Chemical Technology*, 17(2), pp.89–94.

Widenski, D.J., Abbas, A. & Romagnoli, J.A., 2010. Comparison of different solubility equations for modeling in cooling crystallization. *Chemical Engineering and Processing: Process Intensification*, 49(12), pp.1284–1297.

Widenski, D.J., Abbas, A. & Romagnoli, J.A., 2011. Use of predictive solubility models for isothermal antisolvent crystallization modeling and optimization. *Industrial and Engineering Chemistry Research*, 50(13), pp.8304–8313.

Yalkowsky, S.H., He, Y. & Jain, P., 2010. *Handbook of Aqueous Solubility Data* 2nd ed., New York.

Appendix

Appendix A: Solubility data collected in literature

The following tables expose solubility data found in literature for gallic acid and protocatechuic acid in water and organic solvents.

Table A.5: Solubility in g/L of gallic acid in water.

Solvent	Temperature range (K)	Solubility (g/L)	Reference
Water	288	9.1 ± 0.7	(Mota et al. 2008)
	293	11.9	(Yalkowsky et al. 2010)
	298	14.7 ± 0.8	(Mota et al. 2008)
	303	18.6 ± 0.9	(Mota et al. 2008)
	313	22.5 ± 0.62	(Mota et al. 2008)
	323	38.9 ± 2.1	(Mota et al. 2008)

Table A.6: Solubility in weight fraction of gallic acid in water and organic solvents.

Solvent	Temperature range (K)	Solubility*100 (g/g)	Reference
Water	293.15	0.96 ± 0.01	(Lu & Lu 2007)
	298.15	0.10 ± 0.02	
	303.15	1.38 ± 0.01	
	308.15	1.79 ± 0.01	
	313.15	2.36 ± 0.02	
	318.15	3.07± 0.05	
	323.15	4.02 ± 0.06	
	328.15	5.15 ± 0.06	
	333.15	6.86 ± 0.06	
		1.516 ± 0.021	
	298.2		(Daneshfar et al. 2008)
	303.2	1.615 ± 0.021	
	308.2	2.367 ± 0.021	
	313.2	2.540 ± 0.023	
	318.2	3.429 ± 0.026	
	323.2	3.820 ± 0.026	
	328.2	4.787 ± 0.026	
	333.2	7.378 ± 0.027	
Methanol	298.2	27.93± 0.29	(Daneshfar et al. 2008)
	303.2	28.83 ± 0.29	
	308.2	29.30 ± 0.30	
	313.2	29.59 ± 0.30	
	318.2	30.13 ± 0.30	

	323.2	30.48 ± 0.32	
	328.2	31.07 ± 0.33	
	333.2	31.74 ± 0.33	
	298.2	18.9 ± 0.22	
	303.2	18.94 ± 0.22	
	308.2	19.00 ± 0.22	
	313.2	19.17 ± 0.23	
Ethanol	318.2	19.55 ± 0.24	(Daneshfar et al. 2008)
	323.2	20.02 ± 0.24	
	328.2	20.45 ± 0.24	
	333.2	20.93 ± 0.25	
	298.2	1.276 ± 0.020	
	303.2	1.29 ± 0.021	
	308.2	1.303 ± 0.020	
	313.2	1.335 ± 0.021	(Daneshfar et al. 2008)
	318.2	1.438 ± 0.023	
Ethyl acetate	323.2	1.544 ± 0.023	
	328.2	1.598 ± 0.024	
	333.2	1.6898 ± 0.024	

Table A.7: Solubility in mole fraction of gallic acid in organic solvents.

Solvent	**Temperature range (K)**	**Mole Fraction Solubility*1000**	**Reference**
	293.15	1.0099	
	298.15	1.3527	
	303.15	1.706	
	308.15	2.2661	(Noubigh et al. 2013)
	313.15	2.8901	
	318.15	3.7984	
Water			
	293.15	1.269	
	298.15	1.478	
	303.15	2.04	
	308.15	2.524	
	313.15	3.205	(Dali et al. 2016)
	318.15	3.797	
	293.15	63.872	
	298.15	67.163	
Methanol	303.15	70.006	
	308.15	72.072	(Noubigh et al. 2013)
	313.15	75.466	
	318.15	76.815	
	293.15	33.169	
	298.15	41.295	
1 – Propanol	303.15	42.919	(Dali et al. 2016)
	308.15	45.788	
	313.15	47.362	
	318.15	48.563	
	293.15	27.979	
2 – Propanol	298.15	34.834	(Dali et al. 2016)
	303.15	36.204	

	308.15	38.615	
	313.15	39.952	
	318.15	40.947	
	293.15	27.403	
	298.15	34.117	
	303.15	35.459	
Acetonitrile	308.15	37.820	(Dali et al. 2016)
	313.15	39.130	
	318.15	40.461	

Table A.8: Solubility in g/L of protocatechuic acid in water.

Solvent	Temperature range (K)	Solubility (g/L)	Reference
	288	7.6 ± 0.6	
	298	12.7 ± 0.2	
Water	303	17.4 ± 0.1	(Queimada et al. 2009)
	313	28.1 ± 0.9	
	323	49.3 ± 0.5	

Table A.9: Solubility in mole fraction of protocatechuic acid in organic solvents.

Solvent	Temperature range (K)	Mole Fraction Solubility*1000	Reference
	293.15	33.74	
	298.15	45.08	
	303.15	65.38	
Methanol	308.15	92.8	(Noubigh et al. 2015)
	313.15	126.11	
	318.15	167.21	
	293.15	30.24	
	298.15	40.99	
	303.15	49.55	
Ethanol	308.15	66.5	(Noubigh et al. 2015)
	313.15	84.65	
	318.15	108.12	
	293.15	5.63	
	298.15	8.43	
	303.15	11.5	
Methyl Acetate	308.15	16.65	(Noubigh et al. 2015)
	313.15	22.54	
	318.15	30.03	
	293.15	4.78	
	298.15	6.71	
Ethyl Acetate	303.15	9.06	(Noubigh et al. 2015)
	308.15	13	
	313.15	17.91	
	318.15	23.29	

Appendix B: NRTL-SAC conceptual parameters

Table B.2: NRTL-SAC local binary parameters and non-randomness factors (Chen & Song 2004).

Segment 1	X	X	Y-	Y+	X
Segment 2	Y-	Z	Z	Z	Y+
τ_{12}	1.643	6.547	-2.000	2.000	1.643
τ_{21}	1.834	10.949	1.787	1.787	1.834
$\alpha_{12} = \alpha_{21}$	0.2	0.2	0.3	0.3	0.2

Table B.2: NRTL-SAC Molecular Parameters for Common Solvents (Chen & Crafts 2006).

solvent name	*X*	*Y*−	*Y*+	*Z*
acetic acid	0.048	0.222	0.195	0.206
acetone	0.131	0.109	0.513	
acetonitrile	0.018	0.131	0.883	
anisole	0.536	0.010	0.653	
benzene	0.615		0.281	
1-butanol	0.425	0.004		0.490
2-butanol	0.343	0.069		0.393
N-butyl acetate	0.317	0.030	0.330	
methyl *tert*-butyl ether	0.483	0.105	0.142	
carbon tetrachloride	0.739	0.027	0.142	
chlorobenzene	0.727	0.024	0.484	
chloroform	0.393		0.167	
cumene	1.161			
cyclohexane	0.892			
1,2-dichloroethane	0.394		0.691	
1,1-dichloroethylene	0.529		0.208	
1,2-dichloroethylene	0.188		0.832	
dichloromethane	0.459		0.427	0.038
1,2-dimethoxyethane	0.277	0.030	0.077	0.057
N,*N*-dimethylacetamide	0.160	0.778	0.193	
N,*N*-dimethylformamide	0.180	0.752	0.254	
dimethyl sulfoxide		1.114		
1,4-dioxane	0.154	0.086	0.401	
ethanol	0.251	0.030		0.630
2-ethoxyethanol	0.179	0.121	0.106	0.323
ethyl acetate	0.339	0.058	0.441	
ethylene glycol		0.343		0.852
diethyl ether	0.387	0.028	0.177	
ethyl formate	0.256	0.305		
formamide		0.089	0.341	0.252
formic acid		0.090		0.420
N-heptane	1.152			
N-hexane	1.000			
isobutyl acetate	0.620	0.183	0.541	
isopropyl acetate	0.552	0.154	0.498	
methanol	0.090	0.139		0.594
2-methoxyethanol	0.082	0.095	0.180	0.361
methyl acetate	0.239		0.338	
3-methyl-1-butanol	0.419		0.538	0.314
methyl buty ketone	0.673	0.224	0.469	
methylcyclohexane	1.053		0.246	
methyl ethyl ketone	0.261	0.095	0.463	
methyl isobutyl ketone	0.673	0.224	0.469	

isobutanol	0.566		0.067	0.485
N-methyl-2-pyrrolidone	0.252	0.790	0.281	
nitromethane	0.122		1.032	0.051
N-pentane	0.898			
1-pentanol	0.458	0.024		0.491
1-propanol	0.374	0.013		0.530
isopropyl alcohol	0.332			0.636
N-propyl acetate	0.514	0.134	0.587	
pyridine	0.135		0.305	0.249
sulfolane	0.209	0.089		0.708
tetrahydrofuran	0.235	0.040	0.320	
1,2,3,4-tetrahydronaphthalene	0.924		0.865	
toluene	0.604		0.304	
1,1,1-trichloroethane	0.548		0.287	
trichloroethylene	0.552		0.262	
M-xylene	0.758	0.021	0.316	
water				1.000
triethylamine	0.403	0.030		
1-octanol	0.867			0.534
N-octane	1.253			

Appendix C: Shaking and settling times

Table C.1: Shaking and settling times for each solubility experiment

System	Shaking-Time	Settling Time	Temperature (K)
Solutes plus water	30	10	298.15
	40	8	313.15
Solutes plus methanol	27	8	298.15
	33	8	313.15
Solutes plus ethanol	38	8	298.15
	39	9	313.15
Solutes plus isopropanol	30	9	298.15
	39	8	313.15
Solutes plus 1-propanol	32	12	298.15
	46	8	313.15
Solutes plus 2-butanone	27	8	298.15
	33	8	313.15
Solutes plus ethyl acetate	30	8	298.15
	33	8	313.15
Solute plus acetonitrile	47	8	298.15
	33	8	313.15
Solutes plus DMF	36	8	298.15
	39	8	313.15

Appendix D: DSC Thermograms for the Addressed Substances

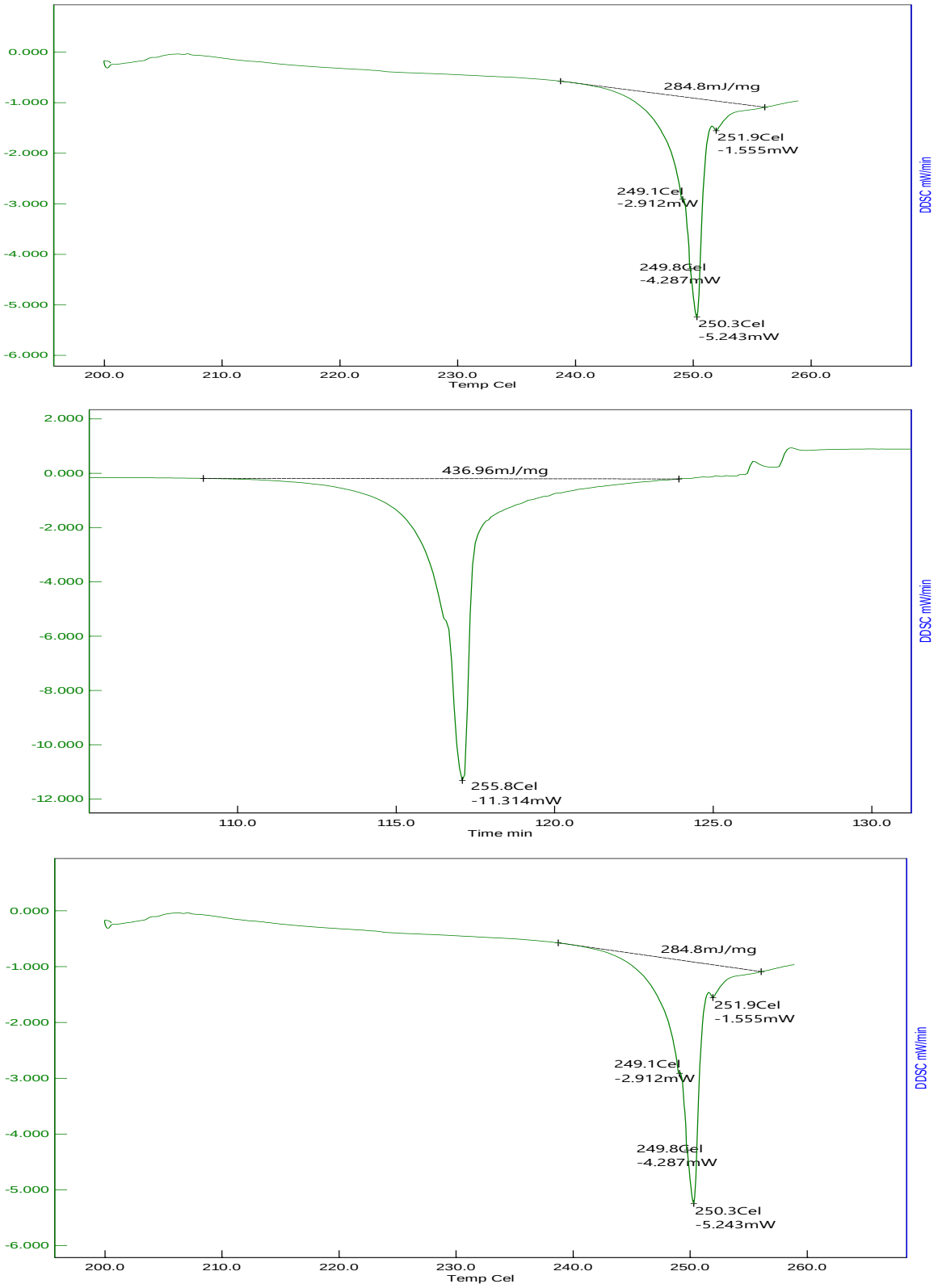


Figure D.1: Thermograms of three DSC analyses performed to gallic acid.

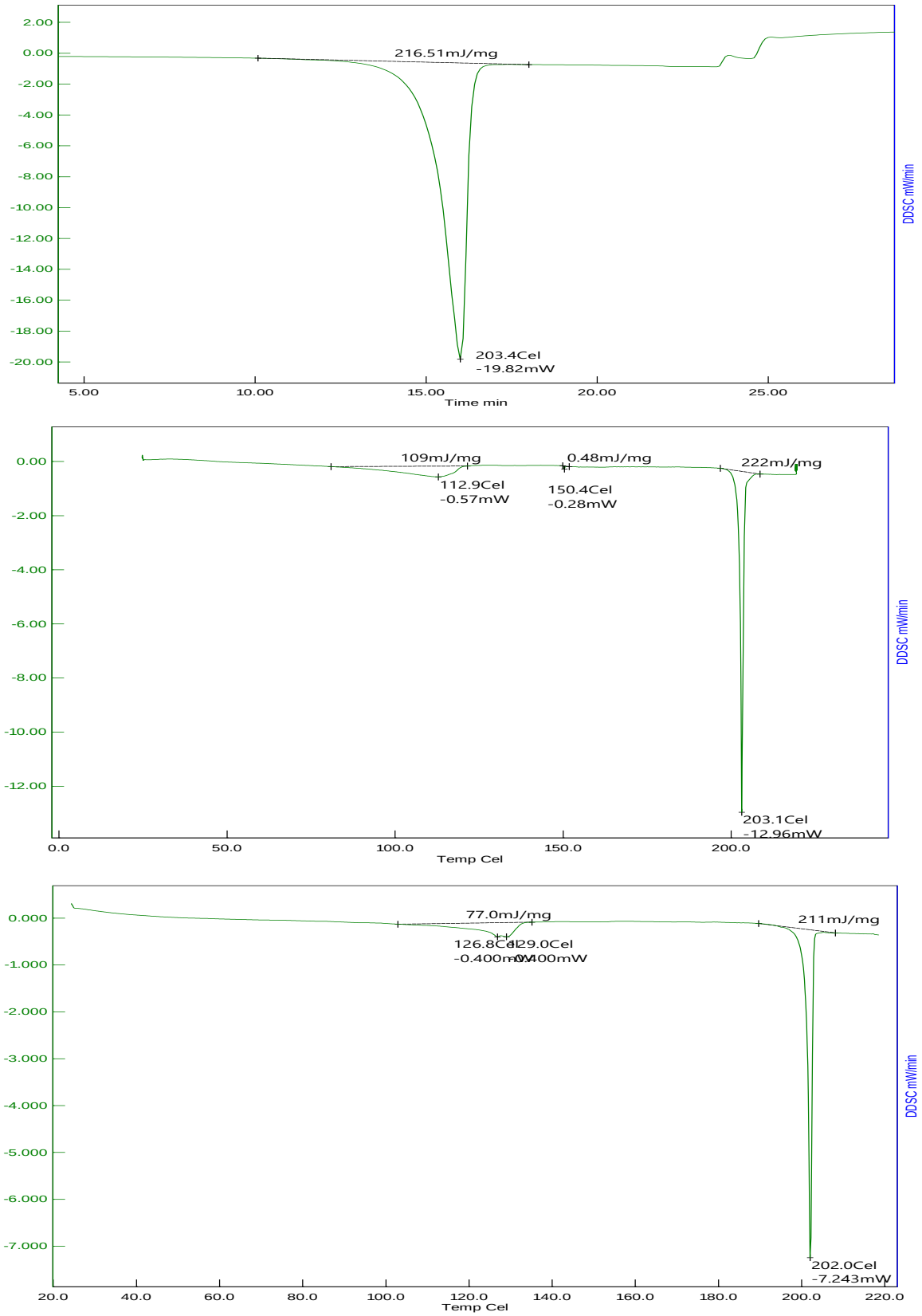


Figure D.2: Thermograms of three DSC analyses performed to protocatechuic acid.

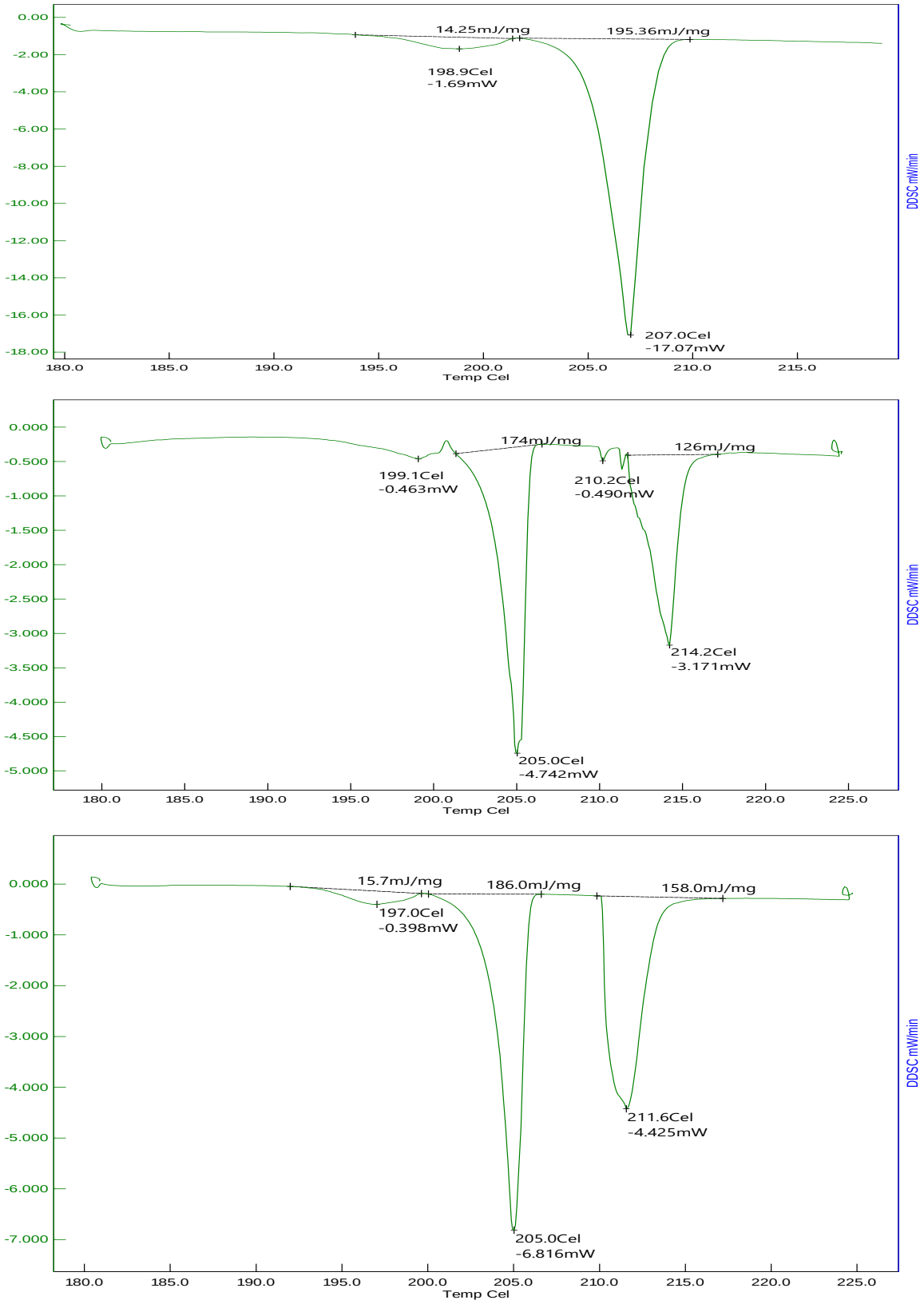


Figure D.3: Thermograms of three DSC analyses performed to gentisic acid.

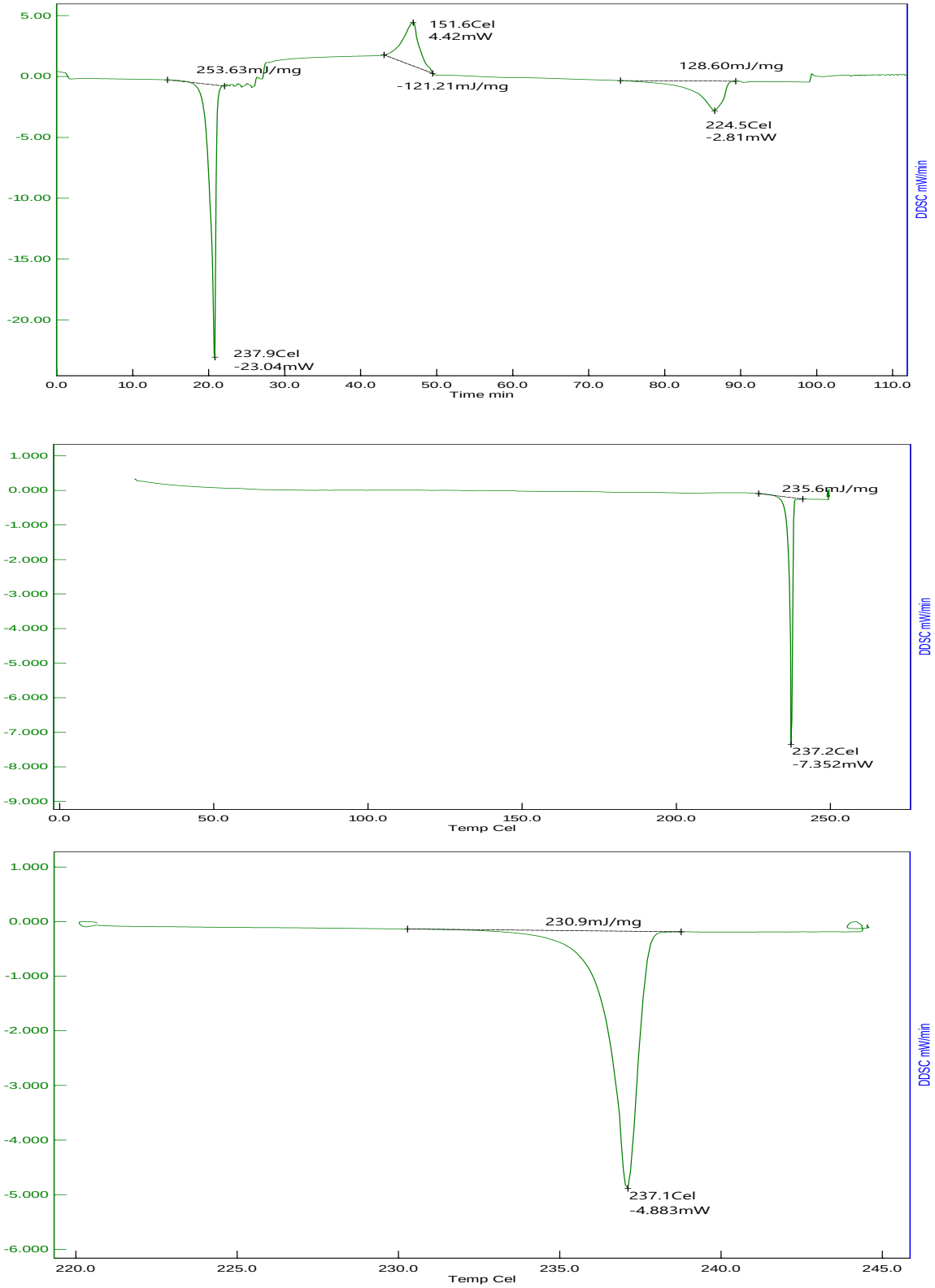


Figure D.4: Thermograms of three DSC analyses performed to α-resorcylic acid.

Appendix E: Experimental solubility data and statistical parameters of the analyzed assays

Table E1: Experimental solubilities of gallic acid at 298.15 K, number of samples analyzed, in the selected pure solvents and statistical parameters of the experimental assays.

Solvent	Solubility at 298.15 K (g of solute/100 g of solvent)	Standard Deviation (s)*10^2	Coefficient of Variation (%)	Number of Analyzed Samples
Water	1.072	0.1	0.350	3
Methanol	38.623	0.2	0.508	3
Ethanol	23.732	0.1	0.037	3
Isopropanol	13.007	0.1	0.051	3
1-Propanol	10.585	0.1	0.832	3
2-Butanone	6.196	0.1	0.116	3
Ethyl Acetate	0.996	0.1	0.137	3
Acetonitrile	0.492	0.1	0.829	3
DMF	44.513	0.8	1.987	4

Table E.2: Experimental solubilities of gallic acid at 313,15 K in the selected pure solvents and statistical parameters of the experimental assays.

Solvent	Solubility at 313.15 K (g of solute/100 g of solvent)	Standard Deviation (s)*10^{-2}	Coefficient of Variation (%)	Number of Analyzed Samples
Water	2.417	0.1	1.603	3
Methanol	41.471	0.1	0.202	3
Ethanol	24.522	0.1	0.218	3
Isopropanol	14.422	0.1	0.281	3
1-Propanol	11.697	0.1	0.183	3
2-Butanone	6.027	0.1	0.089	3
Ethyl Acetate	1.096	0.1	0.885	3
Acetonitrile	0.699	0.1	0.939	3
DMF	49.403	0.2	0.452	4

Table E.3: Experimental solubilities of protocatechuic acid at 298.15 K in selected the pure solvents and statistical parameters of the experimental assays.

Solvent	Solubility at 298.15 K (g of solute/100 g of solvent)	Standard Deviation (s)*10^{-2}	Coefficient of Variation (%)	Number of Analyzes Samples
Water	1.293	0.1	0.491	3

Methanol	79.193	0.7	0.887	3
Ethanol	55.577	0.2	0.281	3
Isopropanol	45.146	0.2	0.440	3
1-Propanol	40.904	0.1	0.0120	3
2-Butanone	49.272	0.1	0.072	4
Ethyl Acetate	7.894	0.1	0.089	3
Acetonitrile	5.910	0.1	0.166	3
DMF	60.728	0.2	0.403	4

Table E.4: Experimental solubilities of protocatechuic acid at 313,15 K in the selected pure solvents and statistical parameters of the experimental assays.

Solvent	Solubility at 313.15 K (g of solute/100 g of solvent)	Standard Deviation (s)*10^{-2}	Coefficient of Variation (%)	Number of Analyzed Samples
Water	3.046	0.1	0.207	3
Methanol	92.404	0.3	0.360	3
Ethanol	57.988	0.1	0.235	3
Isopropanol	50.261	0.1	0.123	3
1-Propanol	43.987	0.1	0.0120	3
2-Butanone	50.444	0.1	0.155	3
Ethyl Acetate	12.991	0.1	0.147	3
Acetonitrile	10.785	0.1	0.124	3
DMF	67.401	0.4	0.580	4

Table E.5: Experimental solubilities of gentisic acid at 298.15 K in selected the pure solvents and statistical parameters of the experimental assays.

Solvent	Solubility at 298.15 K (g of solute/100 g of solvent)	Standard Deviation (s)*10^{-2}	Coefficient of Variation (%)	Number of Analyzed Samples
Water	2.196	0.1	0.203	3
Methanol	67.565	0.1	0.051	3
Ethanol	45.503	0.1	0.171	3
Isopropanol	33.156	0.2	0.664	3
1-Propanol	35.277	0.1	0.007	3
2-Butanone	40.529	0.1	0.004	3
Ethyl Acetate	11.222	0.1	0.060	3
Acetonitrile	7.680	0.1	0.081	3
DMF	73.025	0.7	0.982	4

Table E.6: Experimental solubilities of gentisic acid at 313,15 K in the selected pure solvents and statistical parameters of the experimental assays.

Solvent	Solubility at 313.15 K (g of solute/100 g of solvent)	Standard Deviation (s)*10^{-2}	Coefficient of Variation (%)	Number of Analyzed Samples
Water	5.137	0.1	0.074	3
Methanol	78.613	0.4	0.465	3
Ethanol	51.607	0.3	0.529	3
Isopropanol	44.943	0.1	0.054	3
1-Propanol	40.499	0.1	0.328	3
2-Butanone	40.529	0.1	0.004	3
Ethyl Acetate	18.261	0.1	0.005	3
Acetonitrile	10.989	0.1	0.125	3
DMF	78.101	1.1	1.389	3

Table E.7: Experimental solubilities of α-resorcylic acid at 298.15 K in selected the pure solvents and statistical parameters of the experimental assays.

Solvent	Solubility at 298.15 K (g of solute/100 g of solvent)	Standard Deviation (s)*10^{-2}	Coefficient of Variation (%)	Number of Analyzed Samples
Water	10.176	0.1	0.251	3
Methanol	43.376	0.1	0.193	3
Ethanol	13.068	0.1	0.046	3
Isopropanol	12.823	0.1	0.368	3
1-Propanol	34.512	0.1	0.056	3
2-Butanone	14.469	0.1	0.009	3
Ethyl Acetate	3.317	0.1	0.108	3
Acetonitrile	3.271	0.1	0.226	3
DMF	47.774	0.4	0.772	4

Table E.8: Experimental solubilities of α-resorcylic acid at 313,15 K in the selected pure solvents and statistical parameters of the experimental assays.

Solvent	Solubility at 313.15 K (g of solute/100 g of solvent)	Standard Deviation (s)*10^{-2}	Coefficient of Variation (%)	Number of Analyzed Samples
Water	22.452	0.1	0.065	3
Methanol	52.207	0.1	0.243	3
Ethanol	16.485	0.1	0.131	3

Isopropanol	16.088	0.1	0.205	3
1-Propanol	37.816	0.1	0.110	3
2-Butanone	4.395	0.1	0.173	3
Ethyl Acetate	5.064	0.1	0.074	3
Acetonitrile	5.343	0.1	0.158	3
DMF	53.318	0.3	0.526	4

Appendix F: Polarity parameters of the compounds employed in this work

Table F.1: Predicted polar properties data of the analyzed solutes.

Compound	Topological Polar Surface Area (A^2)	Hydrogen Bond Donor Count	Hydrogen Bond Acceptor Count	LogP	Polarizability (cm^3)
Gallic Acid	98	4	5	0.91	$15.4 \pm 0.5\ 10^{-24}$
Protocatechuic Acid	77.8	3	4	1.16	$14.6 \pm 0.5\ 10^{-24}$
Gentisic Acid	77.8	3	4	1.56	$14.6 \pm 0.5\ 10^{-24}$
A-Resorcylic Acid	77.8	3	4	1.12	$14.6 \pm 0.5\ 10^{-24}$

Data obtained from ChemSpider and PubChem.

Table F.2: Polar properties data of the selected solvents.

Compound	Topological Polar Surface Area (A^2)[c]	Hydrogen Bond Donor Count[a]	Hydrogen Bond Acceptor Count[a]	LogP[a]	Polarizability* 10^{-24} (cm^3)	Dipole Moment (D)[b]	Dielectric Constant[b]
Water	0	2	1	-1.38	1.45 [b]	1.86	80.20
Methanol	20.2	1	1	-0.78	3.29 [b]	1.70	33.00*
Ethanol	20.2	1	1	-0.19	5.1 [b]	1.69	25.3*
1-Propanol	20.2	1	1	0.34	6.7 [b]	1.55	20.8*
Isopropanol	20.2	1	1	0.16	6.97 [b]	1.56	20.18*
Acetonitrile	23.8	0	1	-0.45	4.40 [b]	3.92	36.64*
Ethyl Acetate	26.3	0	2	0.71	9.7 [b]	1.78	6.08*
2-Butanone	17.1	0	1	0.37	8.13 [b]	2.78	18.56*
DMF	20.3	0	2	-1.01	7.81 [b]	3.82	38.25*

[a]Data obtained from (Advanced Chemistry Development 2017). [b]Data obtained from (David R. Lide 2003). [c]Data obtained from (Kim et al. 2016). *Data obtained at 292.15 K. **Data obtained at 298.15 K. ***Data obtained at 303.15 K.

Appendix G: Average Relative Deviation (ARD) of each binary system obtained through NRTL-SAC simulations

Table G.1: ARD (%) obtained from the simulations performed to determine the NRTL-SAC segment parameters for gallic, protocatechuic, gentisic and α-resorcylic acids.

Solute	Reference Solvent	Solvent	Temperature (K)	Experimental Mole Fraction Solubility	Predicted Mole Fraction Solubility	Predicted Activity Coefficient	ARD (%)
Gallic Acid	Acetonitrile	Water	298.2	0.001	0.001	14.576	20
			313.2	0.003	0.002	14.021	
		Methanol	298.2	0.068	0.059	0.329	9
			313.2	0.073	0.074	0.359	
		Ethanol	298.2	0.060	0.049	0.388	11
			313.2	0.062	0.064	0.417	
		Isopropanol	298.2	0.044	0.043	0.445	10
			313.2	0.049	0.057	0.474	
		Ethyl Acetate	298.2	0.005	0.000	85.199	95
			313.2	0.006	0.000	84.644	
		Acetonitrile	298.2	0.001	0.001	16.124	0
			313.2	0.002	0.002	15.950	
Protocatechuic Acid	Acetonitrile	Water	298.2	0.002	0.002	38.695	10
			313.2	0.003	0.003	32.891	
		Methanol	298.2	0.141	0.115	0.510	10
			313.2	0.161	0.163	0.579	
		Ethanol	298.2	0.143	0.155	0.378	22
			313.2	0.148	0.201	0.470	
		Isopropanol	298.2	0.150	0.161	0.363	17
			313.2	0.164	0.206	0.458	
		2-Butanone	298.2	0.187	0.036	1.641	71
			313.2	0.191	0.075	1.269	
		Ethyl Acetate	298.2	0.043	0.020	2.954	43
			313.2	0.069	0.047	2.003	
		Acetonitrile	298.2	0.016	0.016	3.783	0
			313.2	0.028	0.028	3.388	
Gentisic Acid	Isopropanol	Water	298.2	0.003	0.000	1426527.394	99.9
			313.2	0.006	0.000	1425926.160	

		Methanol	298.2	0.123	0.116	0.453	8
			313.2	0.141	0.154	0.500	
		Ethanol	298.2	0.120	0.118	0.445	8
			313.2	0.134	0.152	0.504	
		Isopropanol	298.2	0.115	0.115	0.458	0
			313.2	0.149	0.149	0.516	
		2-Butanone	298.2	0.145	0.096	0.544	25
			313.2	0.159	0.130	0.592	
		Ethyl Acetate	298.2	0.060	0.066	0.798	8
			313.2	0.095	0.101	0.761	
		Acetonitrile	298.2	0.020	0.020	2.684	4
			313.2	0.028	0.030	2.571	
α-Resorcylic Acid	Acetonitrile	Water	298.2	0.012	0.012	3.344	11
			313.2	0.026	0.021	3.112	
		Methanol	298.2	0.083	0.048	0.857	33
			313.2	0.098	0.074	0.867	
		Ethanol	298.2	0.038	0.040	1.008	22
			313.2	0.047	0.064	1.009	
		Isopropanol	298.2	0.048	0.038	1.085	11
			313.2	0.059	0.059	1.085	
		2-Butanone	298.2	0.017	0.007	5.563	48
			313.2	0.020	0.012	5.353	
		Ethyl Acetate	298.2	0.019	0.003	11.887	81
			313.2	0.028	0.006	11.537	
		Acetonitrile	298.2	0.009	0.009	4.708	0
			313.2	0.014	0.014	4.595	

Table G.2: ARD (%) obtained from solubility estimations for binary systems containing gallic acid, protocatechuic acid, gentisic acid and α-resorcylic acid as solutes and 1-propanol and DMF as solvents.

Solute	**Solvent**	**Temperature (K)**	**Experimental Mole Fraction Solubility**	**Predicted Mole Fraction Solubility**	**Predicted Activity Coefficient**	**ARD (%)**
Gallic Acid	1-Propanol	298.2	0.036	0.005	0.445	85
		313.2	0.040	0.007	0.450	
	DMF	298.2	0.161	0.016	0.140	89
		313.2	0.175	0.021	0.146	

Protocatechuic Acid	1-Propanol	298.2	0.138	0.080	0.229	35
		313.2	0.146	0.106	0.290	
	DMF	298.2	0.224	0.023	0.786	87
		313.2	0.242	0.039	0.796	
Gentisic Acid	1-Propanol	298.2	0.121	0.024	0.241	77
		313.2	0.136	0.034	0.260	
	DMF	298.2	0.257	0.091	0.064	61
		313.2	0.270	0.115	0.078	
α-Resorcylic Acid	1-Propanol	298.2	0.119	0.017	1.219	82
		313.2	0.129	0.027	1.216	
	DMF	298.2	0.185	0.040	0.518	74
		313.2	0.202	0.062	0.531	