

Imperial College London
Department of Chemical Engineering

Development of a predictive group-contribution
platform for the phase behaviour and
the effect of pH on the solubility of
pharmaceuticals and excipients

Malak Wehbe

A thesis submitted in fulfilment of the requirements for
the degree of Doctor of Philosophy and the Diploma of
Imperial College London, September 2022

Statement of Originality

I, Malak Wehbe, declare that this thesis titled, "Development of a predictive group-contribution platform for the phase behaviour and the effect of pH on the solubility of pharmaceuticals and excipients", and the research to which it refers, are the product of my own work. I confirm that, where I have used the work of others, published or otherwise, it is fully acknowledged in accordance with the standard referencing practices of the discipline.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

Predictive thermodynamic modeling tools play an important role in the design of pharmaceutical processes by focusing experimental efforts on promising areas and ultimately enabling a more efficient approach to drug product development and manufacturing. In this thesis, we explore the applicability of the SAFT- γ Mie GC approach, a heteronuclear GC version of the statistical associating fluid theory (SAFT) equation of state (EoS), for predicting the phase behaviour and the pH-dependent solubility of pharmaceutical mixtures.

An area of particular interest to the industry is enhancing the solubility of ionisable active pharmaceutical ingredients (APIs). Salt formation is commonly used for this purpose and is monitored via a pH-solubility profile of the API and its salt. Due to the extremely low solubilities of many APIs in water, experiments are difficult to perform, and reliable predictive tools can be especially useful in this context. The SAFT- γ Mie GC EoS is used to predict the phase diagrams and the pH-dependent solubility of two acidic *monoprotic* APIs: ibuprofen and ketoprofen. We consider ibuprofen and ketoprofen in inorganic and organic basic and acidic buffering agents. A new group, NH_3^+ , is characterised for the case of amine buffers and salts. Equilibrium constants for the dissociation of the APIs and the formation of salts ($\text{p}K_{\text{a}}$ and K_{sp} values) are also incorporated in the model using experimental values from literature. Predictions of the complete phase diagrams of the APIs in water are presented, including the vapour–liquid, liquid–liquid (oiling out), and solid–liquid (solubility) equilibria. The SAFT- γ Mie approach is shown to provide accurate predictions of the phase behaviour of the APIs, as well as the pH-solubility profiles of the APIs at 298.15 K and 310.15 K for the range of buffers and salts considered.

We extend the applicability of the SAFT- γ Mie GC approach by modeling the pH-solubility profile of the weakly basic *diprotic* API procaine, in aqueous solutions buffered by NaOH and HCl. Three new groups: aCNH_2 , aCCOO and aCNH_3^+ , in addition to their unlike interactions with other groups, are characterised to study the neutral and ionic forms of procaine within the SAFT- γ Mie framework. The newly developed groups and group interactions are shown to deliver accurate predictions of the intrinsic solubility and the pH-solubility profiles of the API; demonstrating the transferability of the SAFT- γ Mie group interactions, as no data related to procaine were used to obtain the group parameters.

Subsequently, we demonstrate the relevance of the GC approach for the prediction of the phase behaviour of more-complex pharmaceutical mixtures, by studying mixtures containing formaldehyde, a common degradation agent produced by some pharmaceutical excipients during the storage of drug products. Formaldehyde is usually processed in the form of aqueous solutions, with methanol added for stability. In these solutions, formaldehyde reacts with the solvents, to produce a variety of oligomers; and these chemical reactions have a significant influence on the fluid-phase equilibrium of formaldehyde mixtures. The SAFT- γ Mie GC approach is used to obtain the phase behaviour of binary and ternary mixtures of formaldehyde with water and methanol. The oligomerization reactions taking place in formaldehyde solutions are modeled implicitly using a physical approach, which is possible within the SAFT- γ Mie framework by adding association (reactive) sites to mediate the formation of the reaction products. A new group, CH_2O , characterising formaldehyde within the SAFT- γ Mie GC approach, is developed. Experimental data for the vapour-liquid equilibria (VLE) in binary mixtures of formaldehyde + water and formaldehyde + methanol are used to obtain the optimal unlike interaction parameters between the corresponding SAFT- γ Mie groups. The newly developed parameters are then used to predict the VLE of ternary formaldehyde + water + methanol mixtures, providing an excellent agreement with experimental data. Additionally, the nature of chemical speciation in formaldehyde + water, formaldehyde + methanol and formaldehyde + water + methanol mixtures is studied using this approach, obtaining accurate predictions of the distribution of reaction products (oligomers) in the binary and ternary mixtures.

The findings in this thesis demonstrate the applicability of the SAFT- γ Mie GC approach for thermodynamic property prediction of pharmaceutical mixtures. The SAFT- γ Mie EoS is proven to be a powerful predictive tool that can contribute to increasing the efficiency and lowering the cost of drug product development and manufacturing in the pharmaceutical industry.

Acknowledgements

The completion of my PhD would not have been possible without the support of many people. First and foremost, I would like to express my sincere gratitude to my supervisors, Prof. Amparo Galindo and Prof. George Jackson for their continuous support and guidance throughout this journey. I want to thank Amparo and George for their invaluable insights throughout my PhD research and for their kindness, patience, empathy, positivity and motivation throughout. I thank Dr Andrew Haslam for sharing his knowledge on science, British culture and the English language. I thank my industrial sponsor, Eli Lilly and Company, for funding my PhD, and Dr Sal García-Muñoz for his support and sharing his experiences from industry. I would also like to extend my gratitude to Prof. Claire Adjiman for her insight, guidance and leadership throughout my PhD collaboration with Lilly. You are all truly inspirational and I feel grateful and privileged to have worked with you.

I thank my friends from the MSE office, in no particular order: Suela, Sara, Dave, Tom, Spiros, Maz, Lauren, Oliver, Stefanos, Karl, Mohammed, Matthias, Griffin, Redz, Fabian, Saher, Tanuj, Harry, Yizu, Felipe, Thomas, Ahmed, Ben, Isaac, Ed, Jesus, Michele and Gustavo, for all the fun times, memories and discussions. Special thanks to Sara and Lauren for helping me ease into my PhD in the early days, thank you for the advice, guidance, kindness and humour. I will always cherish the memories we've made. I want to also thank my dear friends from the department, Anna-Maria, Anouk, Nerine and Yasmine, for their endless support during the ups and downs of this journey. You made it much more bearable and enjoyable. Yara and Reem, London would not have been the same without you, thank for being my second family away from home.

Finally, I thank my family for their endless love, support and encouragement. I thank Nidal for the life advice and, Omayma and Faten for their emotional support, encouragement and compassion. Most importantly, I am forever grateful to my amazing parents who believed in me and supported me at every moment during my PhD. I could not have done this without you. I hope I made you proud.

Dedication

To mom and dad, without whom none of this would be possible.

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Contents

Abstract	v
Acknowledgements	vii
1 Introduction	1
1.1 Motivation & objectives	1
1.2 Thesis overview	5
2 The SAFT equation of state for pharmaceutical mixtures	7
2.1 Introduction	7
2.1.1 Homonuclear GC approaches	8
2.1.2 Heteronuclear approaches	9
2.2 The SAFT- γ Mie group-contribution approach: molecular systems and electrolytes	12
2.2.1 Models and theory	12
2.2.1.1 Ideal term	15
2.2.1.2 Monomer term	16
2.2.1.3 Chain term	18

2.2.1.4	Association term	19
2.2.1.5	Born term	20
2.2.1.6	Ion term	21
2.2.1.7	Combining rules	22
2.2.1.8	Parameter estimation	23
2.2.2	Property prediction of relevance to pharmaceuticals	26
2.2.2.1	Fluid-phase equilibrium	27
2.2.2.2	Solid–liquid equilibria	28
2.3	Concluding remarks	33
3	Phase behaviour and pH-solubility profile prediction of aqueous buffered solutions of ibuprofen and ketoprofen	34
3.1	Introduction	34
3.2	Thermodynamic modelling of pH-dependent solubility	39
3.2.1	Intrinsic solubility of HA	40
3.2.2	Solubility at $\text{pH} < \text{pH}^{\text{max}}$	41
3.2.3	Salt formation: solubility at $\text{pH} > \text{pH}^{\text{max}}$	44
3.2.4	pH-solubility profile	45
3.3	SAFT- γ Mie theory and model development	47
3.3.1	Equation of state	47
3.3.2	Phase-equilibrium calculations	48
3.3.3	Model development	48

3.3.4	$\text{NH}_3^+ - \text{NH}_3^+$ and $\text{NH}_3^+ - \text{H}_2\text{O}$ interactions	56
3.3.5	Relative permittivity model	60
3.4	Results and discussion	61
3.4.1	Phase diagrams	61
3.4.2	pH-Solubility profiles	66
3.4.3	Temperature effect on the pH-solubility profile	69
3.4.4	Inorganic counter-ion effect on the pH-solubility profile	70
3.4.5	Organic counter-ion effect on the pH-solubility profile	71
3.5	Conclusions	74
4	Predicting the pH-dependent solubility of basic diprotic APIs	77
4.1	Introduction	77
4.2	Thermodynamic modelling of pH-solubility profiles	79
4.2.1	Intrinsic solubility of procaine	82
4.2.2	Solubility of procaine	83
4.2.3	Salt solubility	84
4.2.4	pH-Solubility profiles of procaine	85
4.3	Model development	86
4.3.1	Parameter estimation	93
4.3.2	The aCNH_2 group and the $\text{aCNH}_2 - \text{aCH}$ unlike interaction parameters .	95
4.3.3	The $\text{aCNH}_2 - \text{H}_2\text{O}$ unlike interaction parameters	99

4.3.4	The aCCOO group and its unlike interactions aCCOO–aCH, aCCOO–CH ₂ and aCCOO–CH ₃	101
4.3.5	The aCCOO–H ₂ O unlike interaction parameters	104
4.3.6	The aCNH ₂ –aCCOO, aCNH ₂ –CH ₂ and aCNH ₂ –CH ₃ unlike interaction parameters	107
4.3.7	The aCNH ₃ ⁺ group and the aCNH ₃ ⁺ –H ₂ O unlike interaction parameters	111
4.4	Solubility prediction	113
4.5	Conclusions	118
5	Thermodynamic modeling of the nature of speciation and phase behavior of binary and ternary mixtures of formaldehyde, water and methanol	120
5.1	Introduction	120
5.2	Theory & thermodynamic modelling	127
5.2.1	SAFT- γ Mie theory	127
5.2.2	Property calculations	129
5.2.3	Molecular models	130
5.2.4	Parameter estimation	134
5.3	Phase equilibria	135
5.3.1	Pure compounds	135
5.3.2	Binary formaldehyde + water and formaldehyde + methanol mixtures	136
5.3.3	Ternary mixtures of formaldehyde	142
5.4	Distribution of reaction species	148

5.4.1	Binary mixtures	150
5.4.1.1	Methylene glycol (MG_1)	150
5.4.1.2	Dioxymethylene glycol (MG_2)	151
5.4.1.3	Hemiformal (HMF_1)	152
5.4.1.4	Dioxymethylene hemiformal (HMF_2)	154
5.4.2	Ternary mixtures	155
5.4.2.1	MG_2	156
5.4.2.2	HMF_2	157
5.4.2.3	MG_1 and HMF_1	158
5.4.3	Speciation predictions	159
5.4.3.1	Binary mixtures	159
5.4.3.2	Ternary mixtures	163
5.5	Conclusions	165
6	Conclusions	168
6.1	Summary of key achievements	171
6.2	Directions for future work	173
6.3	Publications and conferences	174
6.3.1	Publications	174
6.3.2	Conference presentations and posters	175
	Bibliography	175

Chapter 1

Introduction

1.1 Motivation & objectives

Over the past two decades, the pharmaceutical industry has been facing unprecedented challenges in turning promising new molecular entities into affordable medicines, with a significant decline in research and development (R&D) productivity. R&D productivity is the relationship between the commercial and medical value created by a new medicine, and the costs incurred to develop that medicine. The main drivers for the drop in R&D productivity are the more-demanding regulatory requirements, more-complex diseases being targeted and expiring patents [1, 2]. In particular, the expiration of patents in recent years has led to plunging drug product revenues, as lower-priced generic substitutions have entered the market. To put this into perspective, the expiration of key patents between 2010 and 2014 resulted in a loss of \$113 billion for the industry [3]. Therefore, efforts in the pharmaceutical industry have lately been focused on the adoption of modern digital technologies, so that new medicines can be developed more efficiently, in reduced time frames and with lower costs.

There are two major trends currently in pharmaceutical manufacturing: the development of continuous manufacturing processes, to replace the conventional batch processes; and the introduction of greener manufacturing practices, which are more environmentally friendly approaches to drug development. It has been proposed that the adoption of process systems

engineering (PSE) concepts in the design, development, control, and optimisation of continuous pharmaceutical manufacturing processes can lead to greener, more-efficient and innovative manufacturing of small molecule drugs [4, 5]. A key aspect of PSE methodologies is the development of predictive models and design spaces for common unit operations in the industry such as solid-liquid separation, anti-solvent based crystallization, multi-phase reactors and distillation. In particular, pharmaceutical process design heavily relies on predictive thermodynamic modelling tools to gain an understanding of the physicochemical properties of the system of interest and to identify required experiments, ultimately reducing the time-consuming laboratory work.

At the same time, thermodynamic modelling tools are being continuously developed in order to meet the need for accurate property prediction in the pharmaceutical sector. Experiments are central to pharmaceutical R&D, however, they are also expensive, laborious and not always possible. By using modelling tools, experimental efforts can be focused on the most promising areas; thus enabling a more efficient approach to drug product development and manufacturing. Therefore, predictive approaches have an important role to play in the design of pharmaceutical processes. The requirements in property prediction of the pharmaceutical industry highlight that there is a pressing need for reliable predictive tools which are capable of covering a wide range of thermodynamic conditions and predicting properties such as phase equilibria and solubility of complex multi-phase, multi-component mixtures, including water, organic solvents, ionic species, dissolved solids and gases, proteins, and others [4, 6, 7].

Among the thermodynamic methods used for property prediction of pharmaceuticals are models based on the conductor-like screening model (COSMO) [8], i.e., COSMO for real solvents (COSMO-RS) [8–10] and COSMO-segment activity coefficient (COSMO-SAC) [11]. In COSMO, a quantum mechanical (QM) calculation is performed to obtain the energy, geometry and the screening charge density on the surface of a solute assuming it is found in an ideal conductor. For a given compound, this computationally expensive calculation is carried out only once. The screening charge profile of the structure can then be stored in a database and directly used from there. Various applications of COSMO-RS and COSMO-SAC to systems of relevance to the pharmaceutical industry have been presented in the literature [12–14].

Another predictive method, which is commonly used for solubility calculations of pharmaceuticals, is the non-random two-liquid segment activity coefficient model (NRTL-SAC) [15]. In this method, molecules are divided into segments of three different types: hydrophobic, hydrophilic and polar segments. The parameters describing the solvents are determined by fitting to experimental vapor-liquid equilibria (VLE) and liquid-liquid equilibria (LLE) data of a binary mixture of the solvent and a reference molecule of each category (hexane for hydrophobic segments, acetonitrile for hydrophilic segments and water for polar segments). After obtaining the solvent parameters, the parameters of the solute are determined using experimental solubility data. The NRTL-SAC parameters for the solute are transferable and can be applied to other solvents and solvent mixtures. Several case studies and applications of NRTL-SAC to pharmaceutical systems can be found in the literature [15–17].

An important category of thermodynamic methodologies with a high predictive power are group-contribution (GC) approaches. In GC methods, the properties of a system of interest (a pure compound or a mixture) are described in terms of the functional groups that make up the molecules. Each group is characterised by a set of group parameters which are typically developed from experimental data of simple compounds. The optimal group parameters obtained are transferable, and can be used to predict thermodynamic properties of more complex molecules and mixtures. The inherent assumption in GC methods is that combinations of these groups can lead to property prediction of the compounds or mixtures of interest.

One of the most commonly used GC thermodynamic methods for property prediction in the pharmaceutical industry, are activity coefficient based methods, in particular, the universal quasi-chemical functional group activity coefficient (UNIFAC) method which was first presented by Fredenslund et al. [18] in 1975. It is probably the most popular GC method for property prediction used by many academics and industrial groups due to the availability of an extensive group parameters table [19]. In UNIFAC, the molecules or mixtures of interest are first decomposed into distinct functional groups, in accordance with the general approach for representing molecules in GC methods, the activity coefficients of the mixtures can then be calculated. The UNIFAC group parameters can be obtained from published tables, and if a required group is not available, new group parameters may need to be developed by fitting to experimental data.

The primary application of UNIFAC is the prediction of phase behaviour, especially the VLE, of non-ideal mixtures. In pharmaceutical applications, UNIFAC has been predominantly used to predict the solubility of drugs in pure and mixed solvents [20–25], obtaining mostly satisfactory results. Some studies extended the UNIFAC solubility prediction of drugs to solvent selection applications [26–28]. Additionally, UNIFAC was applied for the prediction of the solubility of amino acids and peptides in water as a function of pH and temperature [29, 30].

Despite its popularity, UNIFAC is subject to a number of limitations which perhaps make it unsuitable for many applications in the pharmaceutical industry. UNIFAC can only be used to study the phase behaviour of a given system and cannot provide accurate calculations for other thermodynamic properties (e.g., densities, enthalpies, etc.). Phase equilibria calculations using UNIFAC cannot be used for systems containing ions, salts, polymers or non-condensable gases [19], which are frequently found in pharmaceutical solutions. Additionally, as UNIFAC is strongly dependent on the availability of experimental data to provide accurate predictions, poor predictions are often obtained for properties such as activity coefficients at infinite dilution, due to the lack of reliable experimental data. Furthermore, the group parameters in UNIFAC are obtained by regression to experimental VLE data and fail to predict LLE. Different parameters have to be used to predict the LLE of a mixture, which makes the method non-practical for systems that exhibit multiple phase equilibria. Moreover, in UNIFAC, as with other activity coefficient approaches, the vapour phase is assumed to be ideal, an assumption that is only valid at low or moderate pressures, i.e. below 10 atm.

With the aim of overcoming these limitations, GC methods have been coupled with equations of state such as the statistical associating fluid theory (SAFT) [31–34]. The advantages of using equations of state like SAFT, are that the vapour and liquid phases are treated equivalently, pressure effects are accounted for and a wide range of thermodynamic properties can be studied.

In this thesis we aim to demonstrate the applicability of the SAFT equation of state (EoS) for the prediction of thermodynamic properties of pharmaceutical systems. We particularly highlight the predictive capability of the SAFT- γ Mie GC approach [35–38], a recently devel-

oped heteronuclear GC version of the SAFT approach, for the prediction of the phase behaviour and the effect of pH on the solubility of pharmaceuticals and excipients.

1.2 Thesis overview

In Chapter 2, in addition to discussing the history of SAFT and its different versions, we provide a detailed discussion of the underlying theory of the SAFT- γ Mie EoS [35–38], a heteronuclear GC version of the SAFT approach. SAFT- γ Mie has gained increasing attention in recent years due its delivery of a high-level of accuracy in the modelling and prediction of various thermodynamic properties of many systems including those of relevance to the pharmaceutical industry [39–44].

In Chapter 3, we use the SAFT- γ Mie GC approach to provide predictions of the complete phase behaviour including VLE, LLE, and solid-liquid equilibria (SLE), of aqueous solutions of two non-steroidal anti-inflammatory active pharmaceutical ingredients (APIs), ibuprofen and ketoprofen. We extend the thermodynamic modelling to predict the pH-dependent solubility profile of the two monoprotic acidic APIs, ibuprofen and ketoprofen, in a range of aqueous buffered solutions. Additionally, we demonstrate the applicability of the predicted pH-solubility profiles for salt selection, to enhance the solubility of APIs, and for determining salt stability in terms of salt disproportionation. We note that Chapter 3 is an edited version of a peer-reviewed publication [44].

In Chapter 4, we use the SAFT- γ Mie predictive thermodynamic tool, to develop the pH-solubility profiles of diprotic APIs in aqueous buffered solutions. The framework is developed based on the weakly basic diprotic API procaine, a local anesthetic. Additionally, in this chapter, the parameter estimation workflow within the SAFT- γ Mie framework is demonstrated through the development of three new SAFT- γ Mie groups, including their like and unlike interaction parameters, to be used in the modelling of procaine mixtures.

In Chapter 5, we study the phase equilibria of binary and ternary mixtures of formaldehyde, methanol and water. Formaldehyde is known to lead to the degradation of pharmaceutical drug

products during storage when it is present as an impurity in the air or when it is generated by excipients. Therefore, modelling and understanding the VLE of aqueous and methanolic solutions of formaldehyde is of great importance to the industry. However, an added complexity in these solutions is that formaldehyde reacts with water and/or methanol to form oligomers in the liquid and vapour phases. We use the SAFT- γ Mie EoS to predict the VLE of the mixtures and to account for the oligomerization reactions that take place in the two phases, via a physical approach, in which the reaction products are treated implicitly and are considered to be aggregates of the reactants. Additionally, we study the nature of speciation in formaldehyde + water, formaldehyde + methanol and formaldehyde + water + methanol mixtures, where we predictively quantify the distribution of oligomers (reaction products) under various conditions of temperature and pressure.

We conclude the thesis in Chapter 6, where we provide a summary of the key achievements from the current work. We also discuss some directions for future work to build on the findings from this thesis.

Chapter 2

The SAFT equation of state for pharmaceutical mixtures

2.1 Introduction

The statistical associating fluid theory (SAFT) is an equation of state (EoS) which stems from the thermodynamic perturbation theory (TPT) of Wertheim [45–48] for associating and chain fluids. Its development in the late 1980s by Gubbins and co-workers [31–34], was motivated by the need for a rigorous algebraic framework to model properties of associating and polymeric fluids. Due to its successful application to complex fluid systems, such as surfactants, polymers, hydrogen-bonding molecules (e.g., water), micelles, polyfunctional molecules (e.g., amino acids and peptides) and electrolytes, the SAFT EoS received increasing attention in the academic and industrial community and, almost three decades after the publication of the original papers, a vast amount of research has been devoted to the development of numerous versions and applications of the method. For comprehensive reviews see refs [49–55].

More recently, work dedicated to developing SAFT-based group-contribution (GC) approaches involved the application of the group contribution concept to *homonuclear* and *heteronuclear* versions of SAFT. We devote the following sections to a more detailed discussion of these approaches.

2.1.1 Homonuclear GC approaches

In homonuclear GC approaches the molecular model employed is the traditional SAFT representation of molecular chains formed from identical bonded segments, with all segments described by the same set of parameters, so that the underlying theory remained unchanged. One of the first attempts to combine a group contribution method to the original SAFT EoS [33, 34] was by Lora et al. [56]. Several other SAFT-based GC approaches were consequently developed to study *n*-alkanes, alkylbenzenes, olefins, and alcohols [57, 58]. This was later extended to account for polarity in light and heavy esters [59], polycyclic hydrocarbons [60], ethers, aldehydes, ketones [61], amines [62] and long-chain multi-functional molecules such as alkanediols and alkanolamines [63].

A more sophisticated group contribution approach (GC-SAFT) was implemented later by Tamouza et al. [64], within the original SAFT [33, 34] and the SAFT-VR [65, 66] approaches. Even though the concept of groups building up to molecules is introduced explicitly in this approach, the underlying molecular model is still a homonuclear one in which group parameters need to be averaged to predict pure compounds and mixture properties. This method was used to study systems of alkanes, α -olefins, 1-alkanols, alkylbenzenes, alkyl cyclohexanes [64, 67] alkyl esters, formates and their mixtures [68, 69].

Group-contribution based approaches have also been studied within the perturbed chain SAFT (PC-SAFT) framework. The proposition of the PC-SAFT [70] approach is slightly different to other versions of SAFT in that the Helmholtz energy of a homonuclear hard chain fluid (as opposed to that of the monomer fluid) is first considered and the dispersion interactions are then described by a perturbation expansion with the hard-chain fluid as reference.

The PC-SAFT approach has been applied to several systems such as hydrofluoroethers [71]. The same method was used to study the ester series by including dipole-dipole interactions using the perturbation expansion of Twu and Gubbins [72]. In a later study, a quadrupolar-quadrupolar contribution [73], has also been used for the prediction of equilibrium properties of polycyclic aromatic hydrocarbons and their mixtures. The vapour pressures of 19 chemical

families including aliphatic, cyclic and aromatic hydrocarbons, alcohols, amines, nitriles, esters, ketones, ethers, and others have also been considered [74]. Tihic et al. [75] used PC-SAFT in combination with the group contribution of Constantinou et al. [76] to incorporate first- and second-order groups allowing the consideration of isomers and proximity effects. The PC-SAFT approach was later extended by Cameretti et al. [77] to account for charged molecules in the electrolyte PC-SAFT (ePC-SAFT) approach.

PC-SAFT (in all its versions) was more recently used to study the properties of systems relevant to the pharmaceutical industry, including active pharmaceutical ingredients (APIs), amino acids, and peptides. The methods were used to calculate the solubility, phase behaviour and relative humidity, among other properties, of API mixtures for various pharmaceutical applications as shown in Table 2.1. The application of PC-SAFT has also been extended to study mixtures of amino acids and peptides as summarized in Table 2.2.

Despite the success of homonuclear SAFT-based approaches in describing the thermodynamic properties of a wide range of systems, the predictive capability of these methods is limited to properties of a homologous series. No information on the nature of the unlike interactions between the groups or molecules is provided, as the underlying homonuclear model does not enable the differentiation of groups at the molecular level.

2.1.2 Heteronuclear approaches

Following the original first-order thermodynamic perturbation theory (TPT-1) of Wertheim [31, 32, 45–48], heteronuclear models can be proposed where the segments in a given molecule are arbitrarily different. Within the context of SAFT treatment, several versions have been developed based on a tangential heteronuclear model, in which the square-well (SW) [134–137] and Lennard-Jones (LJ) [138–140] potentials were implemented to describe the group-group interactions.

A heteronuclear model, which allows the free energy of the fluid to be written explicitly in terms of group parameters that are different is ideally suited to developing a molecular-

Table 2.1: Overview of the versions of the SAFT equation of state used to study the thermodynamic properties of various active pharmaceutical ingredients (APIs), including: amorphous solid dispersions (ASD), deep-eutectic solvents (DES), room temperature ionic liquids (RTIL), aqueous two phase system (ATPS), non-steroidal anti-inflammatory drugs (NSAIDs), and lipid-based drug delivery systems (LBDDS).

Property	SAFT version	Reference	Pharmaceutical application and/or system studied
Dissolution	PC-SAFT [70, 77]	Paus et al. [78] Ji et al. [79]	API + water ASD
Partition coefficients	SAFT- γ Mie [35–38] PC-SAFT	Hutacharoen et al.[39] Laube et al. [80]	Acidic drugs API + multiple solvents, at various T & pH
Osmotic pressure	SAFT- γ Mie	Di Lecce et al. [42, 43]	DES/RTIL
Vapour-liquid equilibria (VLE)	SAFT- γ Mie	Wehbe et al. [44] Di Lecce et al. [42, 43]	Ionizable acidic drugs DES/RTIL
Liquid-liquid equilibria (LLE)	SAFT- γ Mie	Wehbe et al. [44] Di Lecce et al. [42, 43]	Ionizable acidic drugs DES/RTIL
Solubility (solid-liquid equilibria (SLE))	PC-SAFT SAFT- γ Mie	Laube et al. [81] Wehbe et al. [44]	Extraction Ionizable acidic drugs
		Hutacharoen et al.[39] Febra et al.[41] Prudic et al. [82, 83] Paus et al. [84] Lehmkemper et al. [85] Dohrn et al. [86] Rueether et al. [87] Lange et al. [88] Brinkmann et al. [89] Brinkmann et al. [90] Mathers et al. [91] Paus et al. [92] Paus et al. [93] Ge et al. [94] Wehbe et al. [44]	Acidic drugs NSAIDs ASD ASD, solubility advantage ASD ASD, spray drying API in pure solvents & solvent mixtures Co-crystals, solvent & anti-solvent mixtures API + triglycerides, LBDDS API + edible oils, lipid-based formulations ASD, storage API + water, dissolution API + excipients + water API + water Ionizable acidic drugs
Multiple phase equilibria (SLE, LLE, VLE)	SAFT- γ Mie PC-SAFT	Prudic et al. [95] Veith et al. [96] Veith et al. [97] Borrmann et al. [98, 99] Prudic et al. [100] Luebbert et al. [101–103] Luebbert et al. [104] Luebbert et al. [105] Dohrn et al. [106] Dohrn et al. [107, 108] Wolbert et al. [109] Dohrn et al. [110] Iemtsev et al. [111] Chen et al. [112] Wehbe et al. [44] Cassens et al. [113] Paus et al. [114]	ASD ASD, stability, relative humidity, hydrates Co-crystals, stability, relative humidity ASD, relative humidity ASD, relative humidity ASD ASD, solvent selection ASD, stability ASD, solvent selection ASD ASD, stability, crystallization API + solvent mixtures, cosolvents ASD, stability API + excipients + water Ionizable acidic drugs Ionizable basic drugs Ionizable drugs, dissolution
pH-Solubility	SAFT- γ Mie PC-SAFT		

level group-contribution model. The advantage of a tangential model that follows directly from Wertheim’s theory for chain formation is that the detailed connectivity of the segments can be preserved at some level and hence some isomers can be differentiated. However, a

Table 2.2: Overview of the versions of the SAFT equation of state used to study the thermodynamic properties of various amino acids (AA) and peptides.

Property	SAFT version	Reference	Pharmaceutical application and/or system studied
Liquid-liquid equilibria (LLE)	SAFT of Huang & Radosz [115]	Nazem et al. [116]	AA + water + ionic liquids
Partition coefficients	PC-SAFT [70, 77]	Wysoczanska et al. [117]	AA + PEG + biodegradable salts
pH-Solubility	SAFT-VR [65, 66]	Anvari et al. [118]	AA + water, AA + AA, AA + salts
	original SAFT [31–34]	Seyfi et al. [119]	AA + water, AA + AA + water
	PC-SAFT	Daldrup et al. [120]	Multisolute AA + water
		Fuchs et al. [121]	AA + water, AA + alcohols, AA + water + alcohols
		Lee et al. [122]	AA + water
Solubility (solid-liquid equilibria (SLE))	PC-SAFT	Do et al. [123]	AA + water
		Do et al. [124]	AA + water + 2-propanol, crystallization, anti-solvents
		Cameretti et al. [125]	AA + water, polypeptides + water
		Daldrup et al. [126]	AA + water, AA + AA + water
		Do et al. [127]	Dipeptides + water
		Do et al. [128]	Tripeptides + water
		Held et al. [129]	AA + water
		Bang et al. [130]	AA + salt + water
		Held et al. [131, 132]	AA + salt + water
	SAFT of Huang & Radosz	Ji et al. [133]	AA + alcohol + water

tangential model is not capable of accurately describing multifunctional molecules [141]. A *fused* heteronuclear model, as proposed by the SAFT- γ SW GC approach [141, 142], is more suited to accurately describe real fluids.

In the SAFT- γ SW GC approach, heteronuclear molecules are represented as fused SW segments of different type with a shape factor to characterize the degree of overlap between the different groups. The method was shown to successfully describe the fluid phase behaviour of a wide range of pure components including *n*-alkanes, branched alkanes, alkylbenzenes, olefins, carboxylic acids, alcohols, ketones and amines [141, 142]. The method was also used to study aqueous solutions of alkanes and alcohols [143], in addition to systems of ionic liquids [144]. A closely related fused heteronuclear generalization of SAFT with attractive potentials of variable range (SAFT-VR) using SW potentials, the GC-SAFT-VR approach, was presented by Peng et al. [145]. The theories underlying the SAFT- γ GC and GC-SAFT-VR are identical, providing an equivalent description of the thermodynamic properties of heteronuclear diatomic and triatomic molecules, with the only difference being the treatment of the contribution due

to the formation of chain-like molecules from distinct functional groups [55]. A key advantage of both approaches is that mixtures can be treated fully predictively without the need to use combining rules and adjustable parameters; once the contributions from a set of functional groups are known they can be combined to predict the properties of any compound or mixture that comprises these groups.

A particularly useful and more recent heteronuclear GC SAFT approach is the SAFT- γ Mie EoS [35–38] which is a more general GC version of the earlier homonuclear SAFT-VR Mie EoS [146]. In the SAFT- γ Mie EoS, a Mie (generalized Lennard-Jones) potential [147] of variable repulsive and attractive ranges is used to describe the segment-segment group interactions. The increasing interest in the SAFT- γ Mie GC approach stems from the provision of a high-level of accuracy in the modelling and prediction of a variety of thermodynamic properties including phase equilibria, excess properties of mixing and second-derivative properties for a wide range of systems [35, 36, 39, 148].

In the following section, we provide a more detailed discussion of the underlying theory of the SAFT- γ Mie EoS, including its extension to treat electrolyte solutions. Following this, we discuss the applicability of the SAFT- γ Mie GC approach to pharmaceutical systems and its usefulness in predicting thermodynamic properties of relevance to the pharmaceutical industry.

2.2 The SAFT- γ Mie group-contribution approach: molecular systems and electrolytes

2.2.1 Models and theory

In the SAFT- γ Mie [35–38] GC EoS, molecules are modelled as heteronuclear chains of fused spherical segments which represent the functional groups that make up the molecule. The segments interact via Mie potentials of variable range. Short-range association sites may be added to the groups/segments, if necessary, to account for short-range directional forces. As

an example, in Figure 2.1, a non-steroidal anti-inflammatory drug (NSAID), ibuprofen, is divided into the corresponding groups needed to represent the molecule within the SAFT- γ Mie framework. The groups are: CH₃, aCH, aCCH, aCCH₂, CH and COOH.

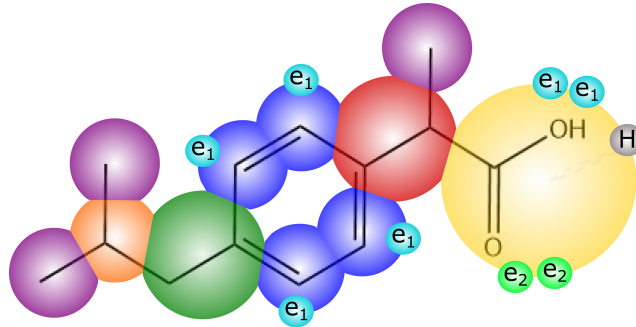


Figure 2.1: The representation of an ibuprofen molecule within the SAFT- γ Mie framework. Each of the coloured spheres corresponds to a different group with the small cyan, light green and grey spheres representing association sites on a given group.

In general, any compound i is represented by its different constituent groups k , where a group k is characterised by a number ν_k^* of spherical segments and a shape factor S_k which represents the contribution of each segment to the overall thermodynamic properties of the molecule considered. The interaction between two groups k and l is described using the Mie [147] potential

$$\Phi_{kl}^{\text{Mie}}(r_{kl}) = \mathcal{C}_{kl} \varepsilon_{kl} \left[\left(\frac{\sigma_{kl}}{r_{kl}} \right)^{\lambda_{kl}^r} - \left(\frac{\sigma_{kl}}{r_{kl}} \right)^{\lambda_{kl}^a} \right], \quad (2.1)$$

where r_{kl} is the distance between the centers of the segments, σ_{kl} the segment diameter, ε_{kl} the depth of the potential well (dispersion energy), and λ_{kl}^r and λ_{kl}^a are the repulsive and attractive exponents of the segment-segment interaction, respectively. The prefactor $\mathcal{C}_{kl}$ is a function of the Mie potential exponents

$$\mathcal{C}_{kl} = \frac{\lambda_{kl}^r}{\lambda_{kl}^r - \lambda_{kl}^a} \left(\frac{\lambda_{kl}^r}{\lambda_{kl}^a} \right)^{\frac{\lambda_{kl}^a}{\lambda_{kl}^r - \lambda_{kl}^a}}; \quad (2.2)$$

which ensures the minimum of the interaction potential is $-\varepsilon_{kl}$.

Hydrogen bonding and strong polar interactions are treated by incorporating short-range square-well association sites on any given segment when appropriate. A segment k can have a number of $N_{\text{ST},k}$ of different association site types, with $n_{k,a}$ sites of type $a = 1, \dots, N_{\text{ST},k}$. The

association interaction between two square-well sites of type a on segment k and b on segment l is given by

$$\Phi_{kl,ab}^{\text{HB}}(r_{kl,ab}) = \begin{cases} -\varepsilon_{kl,ab}^{\text{HB}} & \text{if } r_{kl,ab} \leq r_{kl,ab}^{\text{c}} \\ 0 & \text{if } r_{kl,ab} > r_{kl,ab}^{\text{c}} \end{cases}, \quad (2.3)$$

where $r_{kl,ab}$ is the distance between the centres of the two sites, $\varepsilon_{kl,ab}^{\text{HB}}$ is the association energy, and $r_{kl,ab}^{\text{c}}$ the cut-off range of the interaction between sites a and b , which can be equivalently described in terms of the bonding volume $K_{kl,ab}$ [31]. Each site is positioned at a distance $r_{kk,aa}^d$ or $r_{ll,bb}^d$ from the center of the segment on which it is placed.

If charged groups are present, the formalism is extended following an unrestricted primitive model to account for Coulombic interactions related to the charged species [149, 150]. Each ionic group is assigned a central point charge q and an additional parameter $\sigma_{kk}^{\text{Born}}$ is used to represent the Born cavity diameter, delivering a more accurate solvation free energy of the ions. The relative static permittivity ϵ_r for the dielectric medium presented by Schreckenberget al. [151] is used to replace the explicit representation of the polarity of the solvent.

In the SAFT- γ Mie EoS, the total Helmholtz free energy A of a mixture of associating heteronuclear chains of fused spherical segments is expressed as the sum of six contributions (cf. Figure 2.2), based on a perturbation approach:

$$A = A^{\text{ideal}} + A^{\text{mono.}} + A^{\text{chain}} + A^{\text{assoc.}} + A^{\text{Born}} + A^{\text{ion}}, \quad (2.4)$$

where A^{Ideal} corresponds to the ideal gas free energy, A^{Mono} is the contribution accounting for the interaction of monomeric segments through Mie potentials, A^{Chain} is the contribution to the free energy of the formation of chains of molecules from the fused Mie segments and A^{Assoc} accounts for molecular association through short-range directional interactions. The detailed expressions for each of the first four terms are given in the original publications [35–38]. The presence of charged species is accounted for by the last two terms [149, 150]: the A^{Ion} term represents the Coulombic ion-ion interactions, obtained using the primitive unrestricted Mean

Spherical Approximation (MSA) [152, 153] model, and the A^{Born} term accounts for the ion-solvent electrostatic interactions following the Born [154] model. The underlying expressions for each of the six terms in Equation (2.4) are summarized in the following sections.

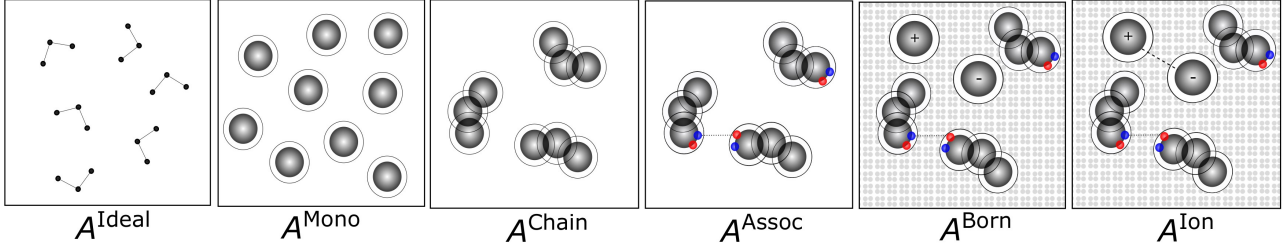


Figure 2.2: Schematic representation of each of the terms contributing to the Helmholtz free energy of the system within the SAFT- γ Mie GC approach: A^{Ideal} corresponds to the ideal-gas free energy, A^{Mono} accounts for the dispersion interaction of monomeric segments, A^{Chain} is the contribution to the free energy due to the formation of chains of molecules, A^{Assoc} accounts for molecular association through hydrogen bonding-like interactions, A^{Born} accounts for the ion-solvent electrostatic interactions following the Born [154] model and A^{Ion} represents the Coulombic ion-ion interactions.

2.2.1.1 Ideal term

The free energy corresponding to an ideal mixture of molecules is given by [155]

$$\frac{A^{\text{ideal}}}{Nk_{\text{B}}T} = \left(\sum_{i=1}^{N_{\text{C}}} x_i \ln (\rho_i \Lambda_i^3) \right) - 1, \quad (2.5)$$

where x_i is the mole fraction of component i in the mixture, $\rho_i = N_i/V$ is the number density with N_i the number of molecules of component i , V the total volume, $N = \sum_i^{N_{\text{C}}} N_i$ the total number of molecules, k_{B} the Boltzmann constant, T the absolute temperature and N_{C} is the total number of components in the mixture. Λ_i^3 is the thermal de Broglie volume which implicitly accounts for the effects of rotational, vibrational and translational contributions to the kinetic energy.

2.2.1.2 Monomer term

The monomer contribution to the free energy, which represents the segment-segment dispersion interactions, is obtained from a Barker-Henderson [156, 157] expansion up to third order [146], expressed as

$$\frac{A^{\text{mono.}}}{Nk_{\text{B}}T} = \frac{A^{\text{HS}}}{Nk_{\text{B}}T} + \frac{A_1}{Nk_{\text{B}}T} + \frac{A_2}{Nk_{\text{B}}T} + \frac{A_3}{Nk_{\text{B}}T} , \quad (2.6)$$

where A^{HS} is the free energy of a hard-sphere reference system of diameter d_{kk} [156]. This is given by [141]

$$\frac{A^{\text{HS}}}{Nk_{\text{B}}T} = \left(\sum_{i=1}^{N_{\text{C}}} x_i \sum_{k=1}^{N_{\text{G}}} \nu_{k,i} \nu_k^* S_k \right) a^{\text{HS}} , \quad (2.7)$$

where x_i is the mole fraction of component i , N_{G} is the number of types of groups present and $\nu_{k,i}$ is the number of groups of type k on component i . a^{HS} is the dimensionless contribution to the hard-sphere free energy per segment obtained following Boublík [158] and Mansoori et al. [159] as,

$$a^{\text{HS}} = \frac{6}{\pi \rho_{\text{s}}} \left[\left(\frac{\zeta_2^3}{\zeta_3^2} - \zeta_0 \right) \ln(1 - \zeta_3) + 3 \left(\frac{\zeta_1 \zeta_2}{1 - \zeta_3} \right) + \frac{\zeta_2^3}{\zeta_3(1 - \zeta_3)^2} \right] ,$$

where

$$\rho_{\text{s}} = \rho \left(\sum_{i=1}^{N_{\text{C}}} x_i \sum_{k=1}^{N_{\text{G}}} \nu_{k,i} \nu_k^* S_k \right) \quad (2.8)$$

and the moment densities ζ_m are given by

$$\zeta_m = \frac{\pi \rho_{\text{s}}}{6} \sum_{k=1}^{N_{\text{G}}} x_{\text{s},k} d_{kk}^m , \quad m = 0, 1, 2, 3 ; \quad (2.9)$$

$x_{\text{s},k}$ is the fraction of segments of a group of type k in the mixture,

$$x_{\text{s},k} = \frac{\sum_{i=1}^{N_{\text{C}}} x_i \nu_{k,i} \nu_k^* S_k}{\sum_{j=1}^{N_{\text{C}}} x_j \sum_{l=1}^{N_{\text{G}}} \nu_{l,j} \nu_l^* S_l} . \quad (2.10)$$

The remaining three terms in the expansion (Equation (2.6)) are evaluated using similar summations over the free-energy contributions per segment. The first-order, second-order and third-order terms are given by

$$\frac{A_q}{Nk_{\text{B}}T} = \left(\frac{1}{k_{\text{B}}T} \right)^q \left(\sum_{i=1}^{N_{\text{C}}} x_i \sum_{k=1}^{N_{\text{G}}} \nu_{k,i} \nu_k^* S_k \right) a_q, \quad q = 1, 2, 3. \quad (2.11)$$

The a_q are obtained by summing the pairwise $a_{q,kl}$ between groups k and l over all pairs of functional groups present in the system,

$$a_q = \sum_{k=1}^{N_{\text{G}}} \sum_{l=1}^{N_{\text{G}}} x_{\text{s},k} x_{\text{s},l} a_{q,kl}, \quad q = 1, 2, 3. \quad (2.12)$$

To calculate $a_{1,kl}$, appearing in the first perturbation term a_1 , knowledge of the radial distribution function (RDF) of the effective hard-sphere fluid for a range of separations is required. To derive an analytical expression for this contribution in a similar manner to the development of the original SAFT-VR [65], the mean-value theorem is used and a mapping of the real density of the fluid to an effective density is carried out for ranges of exponents of the Mie potentials $5 < \lambda_{kl} \leq 100$ [146].

The calculation of $a_{2,kl}$, appearing in the fluctuation term a_2 , depends on 2-, 3- and 4-body correlation functions. For this term, an analytical expression is obtained based on the macroscopic compressibility approximation (MCA) [160] combined with a correction similar to that proposed by Paricaud [161] for soft potentials [146]. A number of coefficients are obtained from the analysis of Monte Carlo molecular-simulation data of selected Mie fluids [146].

For $a_{3,kl}$, which appears in the third perturbation term a_3 , an empirical expression is used which includes coefficients obtained from phase-equilibrium and critical data of selected Mie fluids, such that higher-order terms of the expansion are implicitly considered. This results in an excellent description of experimental thermophysical-property data of real fluids, even in near critical regions [146]. The explicit expressions for the $a_{q,kl}$ are given in refs [35] and [146].

2.2.1.3 Chain term

The chain term in SAFT- γ Mie is the contribution to the Helmholtz free energy of the formation of chains of molecules from fused Mie segments. It is obtained using the formal TPT1 expression of Wertheim [32, 162]. To use this formalism, average molecular parameters ($\bar{\sigma}_{ii}$, $\bar{d}_{ii}$, $\bar{\varepsilon}_{ii}$ and $\bar{\lambda}_{ii}$) are first obtained for each molecular species.

The average molecular segment diameter $\bar{\sigma}_{ii}$ and the effective hard-sphere diameter $\bar{d}_{ii}$ are defined as

$$\bar{\sigma}_{ii}^3 = \sum_{k=1}^{N_G} \sum_{l=1}^{N_G} z_{k,i} z_{l,i} \sigma_{kl}^3, \quad (2.13)$$

and

$$\bar{d}_{ii}^3 = \sum_{k=1}^{N_G} \sum_{l=1}^{N_G} z_{k,i} z_{l,i} d_{kl}^3, \quad (2.14)$$

respectively. $z_{k,i}$ represents the molecular fraction of a given group k in a molecule i , which is given by

$$z_{k,i} = \frac{\nu_{k,i} \nu_k^* S_k}{\sum_{l=1}^{N_G} \nu_{l,i} \nu_l^* S_l}. \quad (2.15)$$

The average interaction energy $\bar{\varepsilon}_{ii}$ and exponents which characterize the range of the potential $\bar{\lambda}_{ii}$ are given as

$$\bar{\varepsilon}_{ii} = \sum_{k=1}^{N_G} \sum_{l=1}^{N_G} z_{k,i} z_{l,i} \varepsilon_{kl}, \quad (2.16)$$

and

$$\bar{\lambda}_{ii} = \sum_{k=1}^{N_G} \sum_{l=1}^{N_G} z_{k,i} z_{l,i} \lambda_{kl}, \quad (2.17)$$

respectively. Equation (2.17) is used for both the repulsive $\bar{\lambda}_{ii}^r$, and the attractive $\bar{\lambda}_{ii}^a$ exponents.

Consequently, the TPT1 expression for the chain term is

$$\frac{A^{\text{chain}}}{N k_B T} = - \sum_{i=1}^{N_C} x_i \left(\sum_{k=1}^{N_G} \nu_{k,i} \nu_k^* S_k - 1 \right) \ln g_{ii}^{\text{Mie}}(\bar{\sigma}_{ii}; \zeta_x), \quad (2.18)$$

where $g_{ii}^{\text{Mie}}(\bar{\sigma}_{ii}; \zeta_x)$ is the RDF evaluated at $\bar{\sigma}_{ii}$ in a hypothetical fluid of packing fraction $\zeta_x =$

$\frac{\pi}{6}\rho_s \sum_{k=1}^{N_G} \sum_{l=1}^{N_G} x_{s,k} x_{s,l} d_{kl}^3$, which is calculated using a second-order expansion [146] :

$$g_{ii}^{\text{Mie}}(\bar{\sigma}_{ii}; \zeta_x) = g_d^{\text{HS}}(\bar{\sigma}_{ii}; \zeta_x) \exp[\beta \bar{\epsilon}_{ii} g_1(\bar{\sigma}_{ii}) / g_d^{\text{HS}}(\bar{\sigma}_{ii}; \zeta_x) + (\beta \bar{\epsilon}_{ii})^2 g_2(\bar{\sigma}_{ii}) / g_d^{\text{HS}}(\bar{\sigma}_{ii}; \zeta_x)] . \quad (2.19)$$

$g_d^{\text{HS}}(\bar{\sigma}_{ii}; \zeta_x)$ represents the RDF of a system of hard spheres of diameter $\bar{d}_{ii}$ at distance $\bar{\sigma}_{ii}$ and packing fraction ζ_x ; which is calculated using the expression of Boublík [163]. $g_1(\bar{\sigma}_{ii}; \zeta_x)$ and $g_2(\bar{\sigma}_{ii}; \zeta_x)$, which appear in the first- and second-order terms respectively, are each approximated by their value at the contact distance $\bar{d}_{ii}$, such that, $g_q(\bar{\sigma}_{ii}; \zeta_x) \approx g_q(\bar{d}_{ii}; \zeta_x)$, for $q = 1, 2$. $g_1(\bar{d}_{ii}; \zeta_x)$ is obtained by the calculation of pressure from the virial and the free-energy routes [65, 66, 146] and an expression based on the MCA with a correction term is used for $g_2(\bar{d}_{ii}; \zeta_x)$. All the parameters and free energy terms required are evaluated using Equations (2.13)–(2.17). A more detailed description of the development of all the expressions can be found in the original SAFT- γ Mie paper [35] .

2.2.1.4 Association term

The association term represents the contribution to the free energy due to the association of molecules via short-range bonding sites. This follows from the original TPT1 expressions of Wertheim [31, 32, 45–48] as

$$\frac{A^{\text{assoc.}}}{Nk_B T} = \sum_{i=1}^{N_C} x_i \sum_{k=1}^{N_G} \nu_{k,i} \sum_{a=1}^{N_{\text{ST},k}} n_{k,a} \left(\ln X_{i,k,a} + \frac{1 - X_{i,k,a}}{2} \right) , \quad (2.20)$$

where N_G is the number of groups, $N_{\text{ST},k}$ is the number of site types on each group k , $n_{k,a}$ is the number of sites of type a on group k ($a = 1, \dots, N_{\text{ST},k}$) and $X_{i,k,a}$ is the fraction of molecules of component i that are not bonded at a site of type a on group k . $X_{i,k,a}$ is obtained from the solution of the mass action equations as [31, 48, 141]

$$X_{i,k,a} = \left[1 + \rho \sum_{j=1}^{N_C} x_j \sum_{l=1}^{N_G} \nu_{l,j} \sum_{b=1}^{N_{\text{ST},l}} n_{l,b} X_{j,l,b} \Delta_{ij,kl,ab} \right]^{-1} . \quad (2.21)$$

Here ρ is the molecular density and $\Delta_{ij,kl,ab}$ characterizes the overall strength of the association between a site of type a on a group of type k of component i and a site of type b on a group of type l of component j . This is approximated as

$$\Delta_{ij,kl,ab} = F_{kl,ab} K_{kl,ab} I_{ij} , \quad (2.22)$$

where $F_{kl,ab} = \exp(\varepsilon_{kl,ab}^{\text{HB}}/k_{\text{B}}T) - 1$, $K_{kl,ab}$ is the bonding-volume, and I_{ij} is a dimensionless polynomial temperature–density correlation of the association integral, expressed as

$$I_{ij} = \sum_{p=0}^{10} \sum_{q=0}^{10-p} c_{pq} (\rho_s \sigma_x^3)^p \left(\frac{k_{\text{B}}T}{\bar{\varepsilon}_{ij}} \right)^q . \quad (2.23)$$

The coefficients c_{pq} are given in refs [37, 38]. σ_x^3 and $\bar{\varepsilon}_{ij}$ are given by

$$\sigma_x^3 = \sum_{k=1}^{N_{\text{G}}} \sum_{l=1}^{N_{\text{G}}} x_{s,k} x_{s,l} \sigma_{kl}^3 , \quad (2.24)$$

and

$$\bar{\varepsilon}_{ij} = \frac{\sqrt{\bar{\sigma}_{ii}^3 \bar{\sigma}_{jj}^3}}{\bar{\sigma}_{ij}^3} \sqrt{\bar{\varepsilon}_{ii} \bar{\varepsilon}_{jj}} , \quad (2.25)$$

respectively. $\bar{\sigma}_{ii}$ and $\bar{\sigma}_{jj}$ are given by Equation (2.13), $\bar{\varepsilon}_{ii}$ and $\bar{\varepsilon}_{jj}$ by Equation (2.16), and

$$\bar{\sigma}_{ij} = \frac{\bar{\sigma}_{ii} + \bar{\sigma}_{jj}}{2} . \quad (2.26)$$

2.2.1.5 Born term

Within the Born model [154], a spherical cavity of diameter $\sigma_{ii}^{\text{Born}}$ is created for each ion in the dielectric medium, independent of any others, resulting in a change in the Helmholtz free energy given by

$$A^{\text{Born}} = -\frac{e^2}{4\pi\epsilon_0} \left(1 - \frac{1}{\epsilon_r} \right) \sum_{i=1}^{n_{\text{ion}}} \frac{N_i Z_i^2}{\sigma_{ii}^{\text{Born}}} . \quad (2.27)$$

Here index i runs over all n_{ion} charged species, $e = 1.602 \times 10^{-19}$ C is the elementary charge, $\epsilon_0 = 8.854 \times 10^{-12}$ C²J⁻¹m⁻¹ the static permittivity in vacuum, ϵ_r the relative static permittivity

and Z_i the valency of ion i .

2.2.1.6 Ion term

The ion term represents the contribution to the free energy due to the electrostatic interactions between charged species, given by [151]

$$A^{\text{ion}} = U^{\text{MSA}} + \frac{\Gamma^3 k_B T V}{3\pi}. \quad (2.28)$$

Here U^{MSA} is the MSA contribution [152, 153] to the internal energy U , V is the volume of the system, and Γ is the screening length of the electrostatic forces, and

$$U^{\text{MSA}} = -\frac{e^2 V}{(4\pi\epsilon_0)\epsilon_r} \left[\frac{\Gamma}{V} \sum_{i=1}^{n_{\text{ion}}} \left(\frac{N_i Z_i^2}{1 + \Gamma \sigma_{ii}} \right) + \frac{\pi}{2\Delta} \Omega P_n^2 \right], \quad (2.29)$$

where σ_{ii} is the diameter of each ion, Δ represents the packing fraction of the ions as

$$\Delta = 1 - \frac{\pi}{6V} \sum_{i=1}^{n_{\text{ion}}} N_i \sigma_{ii}^3. \quad (2.30)$$

P_n and Ω are coupling parameters, where P_n couples to the charge of the ions, and Ω relates to the packing fraction of the ions. Both coupling parameters are functions of the ionic parameters, as well as the screening length of the ions:

$$P_n = \frac{1}{\Omega V} \sum_{i=1}^{n_{\text{ion}}} \frac{N_i \sigma_{ii} Z_i}{1 + \Gamma \sigma_{ii}}, \quad (2.31)$$

$$\Omega = 1 + \frac{\pi}{2\Delta V} \sum_{i=1}^{n_{\text{ion}}} \frac{N_i \sigma_{ii}^3}{1 + \Gamma \sigma_{ii}}. \quad (2.32)$$

The screening length of the ions is a function of the relative static permittivity ϵ_r and the effective charge $Q_i(\Gamma)$ of the ions, leading to an implicit formulation through Q_i :

$$\Gamma^2 = \frac{\pi e^2}{(4\pi\epsilon_0)\epsilon_r k_B T V} \sum_{i=1}^{n_{\text{ion}}} N_i Q_i^2, \quad (2.33)$$

where the effective charge is related to the electric charge of the individual species and the P_n coupling parameter,

$$Q_i = \frac{Z_i - \sigma_{ii}^2 P_n (\pi / (2\Delta))}{1 + \Gamma \sigma_{ii}}. \quad (2.34)$$

2.2.1.7 Combining rules

Unlike group interaction parameters are required when using the SAFT- γ Mie GC approach due to the heteronuclear nature of the model. For example, the unlike interaction between the CH_3 and CH_2 groups is required to represent pure n -alkanes. Combining rules are generally used to determine the unlike group interactions, given that the like group parameters are known. If necessary, these unlike group interaction parameters can be further refined by estimation using experimental data

The unlike segment diameter σ_{kl} is obtained from the Lorentz rule [164]. The same combining rule is also used in the calculation of the unlike effective hard-sphere diameter, d_{kl} , i.e.,

$$\sigma_{kl} = \frac{\sigma_{kk} + \sigma_{ll}}{2} ; \quad d_{kl} = \frac{d_{kk} + d_{ll}}{2} . \quad (2.35)$$

In SAFT- γ Mie, the unlike dispersion energy can be calculated using an augmented geometric mean rule which accounts for the asymmetry in size of the segments [35]

$$\varepsilon_{kl} = \frac{\sqrt{\sigma_{kk}^3 \sigma_{ll}^3}}{\sigma_{kl}^3} \sqrt{\varepsilon_{kk} \varepsilon_{ll}} . \quad (2.36)$$

For the unlike attractive and repulsive exponents of the Mie potential, combining rules based on applying the geometric-mean criterion on the van der Waals energy are used [35]

$$\lambda_{kl}^y = 3 + \sqrt{(\lambda_{kk}^y - 3)(\lambda_{ll}^y - 3)}, \quad y = \text{a, r} . \quad (2.37)$$

In the case of interactions involving uncharged associating groups the unlike association

energy $\varepsilon_{kl,ab}^{\text{HB}}$ can be calculated using a geometric mean as [36]

$$\varepsilon_{kl,ab}^{\text{HB}} = \sqrt{\varepsilon_{kk,aa}^{\text{HB}} \varepsilon_{ll,bb}^{\text{HB}}} , \quad (2.38)$$

while the unlike bonding volume $K_{kl,ab}$ is calculated as [36]

$$K_{kl,ab} = \left(\frac{\sqrt[3]{K_{kk,aa}} + \sqrt[3]{K_{ll,bb}}}{2} \right)^3 . \quad (2.39)$$

Additionally, for the case of interactions involving charged groups, the dispersion energy between like or unlike ions, ε_{kl} , is calculated using combining rules based on the work of Hudson and McCoubrey [165]:

$$\varepsilon_{kl} = \frac{(\lambda_{kl}^{\text{r}} - 3)(\lambda_{kl}^{\text{a}} - 3)}{2C(\lambda_{kl}^{\text{r}} - \lambda_{kl}^{\text{a}})} \frac{\alpha_{0,k} \alpha_{0,l}}{(4\pi\epsilon_0)^2 \sigma_{kl}^6} \frac{I_k I_l}{I_k + I_l} , \quad (2.40)$$

where α_0 is the electronic polarisability and I is the ionisation potential of the ion.

When the use of combining rules does not lead to an accurate representation of the thermodynamic properties (or phase behaviour) of the systems of interest, the group parameters are further refined using experimental data as part of a parameter-estimation strategy, as discussed in the following section. The use of a heteronuclear model in the SAFT- γ Mie EoS, offers the advantage that pure-component data alone can, in many cases, be used to obtain an accurate estimate of the values of unlike group-interaction parameters.

2.2.1.8 Parameter estimation

To estimate the SAFT- γ Mie parameters for a given group, the following least-squares objective function is minimised:

$$\min_{\Theta} f_{\text{obj}} = \sum_{X=1}^{N^{\text{X}}} \sum_{i=1}^{N_X^{\text{D}}} w_{X,i} \left(\frac{R_{X,i}^{\text{exp}} - R_{X,i}^{\text{calc}}(\Theta)}{R_{X,i}^{\text{exp}}} \right)^2 , \quad (2.41)$$

where Θ is the vector of model parameters, N^X is the number of property types, N_X^D is the number of experimental data points for property X , $R_{X,i}^{\text{exp}}$ is the i th measured value of property X , $R_{X,i}^{\text{calc}}$ is the corresponding value calculated by the SAFT- γ Mie EoS and $w_{X,i}$ refers to the weight of data point i for property X .

In order to determine the accuracy of the theoretical description provided by the SAFT- γ Mie GC approach using the newly developed group interactions, compared to the experimental data, for a property X , the percentage absolute average deviation (%AAD [X]),

$$\% \text{AAD}[X] = \frac{100}{N_X^D} \sum_{i=1}^{N_X^D} \left| \frac{R_{X,i}^{\text{exp}} - R_{X,i}^{\text{calc}}}{R_{X,i}^{\text{exp}}} \right|, \quad (2.42)$$

and the absolute average deviation (AAD [X]),

$$\text{AAD}[X] = \frac{1}{N_X^D} \sum_{i=1}^{N_X^D} |R_{X,i}^{\text{exp}} - R_{X,i}^{\text{calc}}|, \quad (2.43)$$

are often used.

The increasing interest over the years in the SAFT- γ Mie GC approach for thermodynamic property prediction of various systems, led to the expansion of the corresponding groups database. Today, the SAFT- γ Mie group database features more than 60 groups, many of which are found in pharmaceutical compounds. The reader is referred to the GitHub repository (<https://github.ic.ac.uk/ce-mse-saft-database/database>) for the latest version of the SAFT- γ Mie groups database. A recent review by Haslam et al. [148] presented the SAFT- γ Mie group parameters table and showcased the performance of the GC approach for a wide range of applications. In Table 2.3 the SAFT- γ Mie group matrix relevant to the systems of interest in the current work is presented. Like and unlike group interaction parameters are obtained from previous work, where applicable [35–40, 42, 43, 150, 166, 167]. Additionally, we expand the matrix of groups by estimating like and unlike interactions for a number of new groups, as discussed in the following chapters. The complete set of SAFT- γ Mie group parameters used in the current work, including the newly developed ones, can be found in Tables A.1, A.2 and A.3 in the Appendix.

potential λ_{kk}^a is set to a value of 6 for most of the SAFT- γ Mie groups representative of the London dispersion interactions that occur [146]. Another factor considered when estimating group parameters is ensuring that for associating groups, the association energy $\varepsilon_{kk}^{\text{HB}}$ is larger than the group's dispersion energy ε_{kk} due to the nature of stronger intermolecular interactions taking place between hydrogen bonding (associating) sites as compared to other intermolecular forces.

In the following section we discuss the calculation of various thermodynamic properties using the SAFT- γ Mie EoS, highlighting the applicability of the GC approach for the prediction of properties of relevance to the pharmaceutical industry.

2.2.2 Property prediction of relevance to pharmaceuticals

In recent years, the SAFT- γ Mie GC approach has been successfully used to predict thermodynamic properties of pharmaceutical systems (cf. Tables 2.1 and 2.2). In SAFT- γ Mie, the Helmholtz free energy of a fluid mixture (Equation (2.4)) is a function of temperature T , volume V and the composition vector $\mathbf{N}$ (of i components). Knowledge of the pressure P , chemical potential μ_i and fugacity coefficient $\hat{\varphi}_i$ of a given component i , allows the calculation of fluid-phase behaviour and solution properties. These can be obtained from the Helmholtz free energy following standard thermodynamic relations [168]. The pressure is obtained as,

$$P = - \left. \frac{\partial A(T, V, \mathbf{x})}{\partial V} \right|_{T, \mathbf{N}}, \quad (2.44)$$

the chemical potential as,

$$\mu_i^{\text{Res}}(T, P, \mathbf{x}) = \left. \frac{\partial A^{\text{Res}}(T, V, \mathbf{x})}{\partial N_i} \right|_{T, V, N_{j \neq i}} - RT \ln Z(T, P, \mathbf{x}), \quad (2.45)$$

and the fugacity coefficient as

$$\ln \hat{\varphi}_i(T, P, \mathbf{x}) = \frac{\mu_i^{\text{Res}}(T, P, \mathbf{x})}{RT}, \quad (2.46)$$

where, μ_i^{Res} is the residual chemical potential of component i , $A^{\text{Res}} = A - A^{\text{Ideal}}$ is the residual free energy, $Z = Pv_P/(RT)$ is the compressibility factor, with $v_P = V_P/N$ as the molar volume at the specified pressure, N is the total number of moles and $\mathbf{x} = \mathbf{N}/N$.

2.2.2.1 Fluid-phase equilibrium

In general, once two or more phases are brought into contact with each other, there will be an exchange of their constituents until the composition of each phase reaches a constant value, i.e., phase equilibrium is achieved. The thermodynamic equilibrium of an isolated multi-phase mixture of i components depends on the temperature T , pressure P and total number of moles $\mathbf{x}$ of the mixture, and is based on the equality of the chemical potential of each component i , μ_i , in each phase (α or β),

$$\mu_i^\alpha(T, P, \mathbf{x}^\alpha) = \mu_i^\beta(T, P, \mathbf{x}^\beta) \quad , \quad \alpha \neq \beta \quad \text{and} \quad \forall i, \quad (2.47)$$

which can be obtained from the Helmholtz free energy of the system. Consequently, the composition $\mathbf{x}$ of each phase at equilibrium, for a given T and P , can be determined.

The SAFT- γ Mie GC approach has been shown to deliver a high-level of accuracy in the modelling and prediction of various fluid-phase equilibria, including liquid–liquid equilibria (LLE), vapor–liquid equilibria (VLE) and vapor–liquid–liquid equilibria (VLLE) of a wide range of mixtures [35, 36, 39, 148]. Of particular relevance to this work is the previous work by Hutacharoen et al. [39], in which the octanol-water partition coefficients $K_{i,\text{OW}}$ of multiple APIs were accurately predicted. $K_{i,\text{OW}}$ is based on the ratio of the molar concentrations of an infinitely dilute solute i distributed in the two equilibrium LLE phases formed by mixtures of 1-octanol and water. Knowledge of $K_{i,\text{OW}}$ gives an indication of the lipophilicity of pharmaceutical compounds and is useful in predicting the pharmacokinetic properties and toxicity of organic drug molecules [169–171].

In the following section, we present the calculation of the solid–liquid equilibria (SLE) of multi-component mixtures using the SAFT- γ Mie approach and demonstrate its proven

applicability for pharmaceutical systems.

2.2.2.2 Solid–liquid equilibria

The solid–liquid solubility of a compound i (the solute) in a solvent is defined as the maximum amount of the solid solute that can be dissolved in a given amount of the liquid solvent to form a homogeneous solution (at a given temperature T , and pressure P). This can be obtained by solving the solid–liquid equilibrium (SLE) of the system. Prausnitz et al. [25] presented a thermodynamic framework to solve the SLE of a mixture containing a solute, which is followed here, starting at the condition of thermodynamic phase equilibrium, i.e., the equality of the chemical potential of i in the solid (S) and the liquid (L) phases,

$$\mu_i^S(T, P, \mathbf{x}^S) = \mu_i^L(T, P, \mathbf{x}^L) , \quad (2.48)$$

which can equivalently be written in terms of the fugacity of component i f_i as

$$f_i^S(T, P, \mathbf{x}^S) = f_i^L(T, P, \mathbf{x}^L) . \quad (2.49)$$

Here, the solid phase is assumed to consist of pure solute, while the liquid phase is considered to be a mixture of the solute i and solvent(s). Therefore, Equation (2.49) becomes

$$f_i^{S,*}(T, P) = f_i^L(T, P, \mathbf{x}^L) , \quad (2.50)$$

where the asterisk denotes a pure component. We note that the equality of fugacities applies only to the solute i and not to all the components, as no other components are present in the solid phase.

The fugacity of i in the liquid phase is given by

$$f_i^L(T, P, \mathbf{x}^L) = f_i^{L,*}(T, P) x_i^L \gamma_i^L(T, P, \mathbf{x}^L) , \quad (2.51)$$

and therefore,

$$f_i^{S,*}(T, P) = f_i^{L,*}(T, P) x_i^L \gamma_i^L(T, P, \mathbf{x}^L), \quad (2.52)$$

where x_i^L is the composition of the solute in the liquid phase, γ_i^L is the corresponding activity coefficient of i and $f_i^{L,*}$ is the fugacity of the pure solute in the liquid phase at temperatures T below the melting point of the solute T_i^{fus} .

Rearranging Equation (2.52) gives:

$$\frac{f_i^{L,*}(T, P)}{f_i^{S,*}(T, P)} = \frac{1}{x_i^L \gamma_i^L(T, P, \mathbf{x}^L)}, \quad (2.53)$$

such that the left-hand side of Equation (2.53) is solely dependent on the properties of the solute, irrespective of the solvent(s) used, and can be calculated using the thermodynamic cycle shown in Figure 2.3, assuming a constant pressure P (a valid assumption for solubility calculations which are usually reported at ambient pressure). In this cycle, the solute in the solid phase at temperature T is heated to T_i^{fus} ($a \rightarrow b$); then the solute melts at T_i^{fus} ($b \rightarrow c$); and finally the liquid solute at T_i^{fus} is cooled down to T ($c \rightarrow d$).

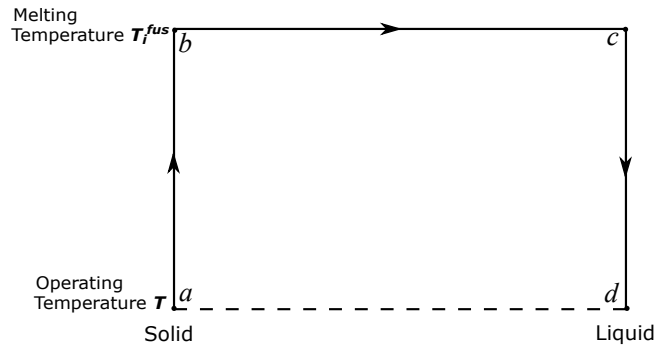


Figure 2.3: Thermodynamic cycle for the calculation of the fugacity of a pure subcooled liquid as outlined in Prausnitz et al. [25]

The molar Gibbs free energy of the solute going along the cycle from state a to state d , $\Delta g_{a \rightarrow d}$, can be related to the ratio of the fugacities of the solute in the liquid and solid phases (left-hand side of Equation (2.53)) by

$$\Delta g_{a \rightarrow d} = RT \ln \frac{f_i^{L,*}(T, P)}{f_i^{S,*}(T, P)}. \quad (2.54)$$

The molar Gibbs free energy is also related to the corresponding changes in molar enthalpy Δh and molar entropy Δs , from state a to state d , by

$$\Delta g_{a \rightarrow d} = \Delta h_{a \rightarrow d} - T \Delta s_{a \rightarrow d}. \quad (2.55)$$

The thermodynamic cycle in Figure 2.3 can be used to evaluate Δh and Δs along the path $a \rightarrow d$ by substituting the alternative path $a \rightarrow b \rightarrow c \rightarrow d$; which is permissible because enthalpy and entropy are state functions and are independent of the path. Hence, the change in molar enthalpy can be written as

$$\begin{aligned} \Delta h_{a \rightarrow d} &= \Delta h_{a \rightarrow b} + \Delta h_{b \rightarrow c} + \Delta h_{c \rightarrow d} \\ &= \int_T^{T_i^{\text{fus}}} c_{p,i}^{\text{S}} dT' + \Delta h_i^{\text{fus}}(T_i^{\text{fus}}, P) + \int_{T_i^{\text{fus}}}^T c_{p,i}^{\text{L}} dT' \\ &= \Delta h_i^{\text{fus}}(T_i^{\text{fus}}, P) + \int_{T_i^{\text{fus}}}^T \Delta c_{p,i}(T', P) dT', \end{aligned} \quad (2.56)$$

where Δh_i^{fus} is the molar enthalpy of fusion of the solute at the melting temperature T_i^{fus} , R is the universal gas constant, and $\Delta c_{p,i} = c_{p,i}^{\text{L}} - c_{p,i}^{\text{S}}$ is the difference between the isobaric heat capacity of i in the liquid and solid phases, respectively. Similarly, the change in molar entropy can be expressed as

$$\begin{aligned} \Delta s_{a \rightarrow d} &= \Delta s_{a \rightarrow b} + \Delta s_{b \rightarrow c} + \Delta s_{c \rightarrow d} \\ &= \int_T^{T_i^{\text{fus}}} \frac{c_{p,i}^{\text{S}}}{T} dT' + \Delta s_i^{\text{fus}}(T_i^{\text{fus}}, P) + \int_{T_i^{\text{fus}}}^T \frac{c_{p,i}^{\text{L}}}{T} dT' \\ &= \Delta s_i^{\text{fus}}(T_i^{\text{fus}}, P) + \int_{T_i^{\text{fus}}}^T \frac{\Delta c_{p,i}(T', P)}{T} dT', \end{aligned} \quad (2.57)$$

where Δs_i^{fus} is the molar entropy of fusion of the solute at the melting temperature T_i^{fus} , which can be related to Δh_i^{fus} by $\Delta s_i^{\text{fus}} = \frac{\Delta h_i^{\text{fus}}}{T_i^{\text{fus}}}$.

Combining Equations (2.54)–(2.57) with Equation (2.53) leads to the well-known solubility

equation [25],

$$\begin{aligned} \ln x_i^L(T, P) = & \frac{\Delta h_i^{\text{fus}}(T_i^{\text{fus}}, P)}{R} \left(\frac{1}{T_i^{\text{fus}}} - \frac{1}{T} \right) + \frac{1}{RT} \int_T^{T_i^{\text{fus}}} \Delta c_{p,i}(T', P) dT' \\ & - \frac{1}{R} \int_T^{T_i^{\text{fus}}} \frac{\Delta c_{p,i}(T', P)}{T'} dT' - \ln \gamma_i(T, P, \mathbf{x}^L). \end{aligned} \quad (2.58)$$

The SAFT- γ Mie GC approach can be used to calculate the activity coefficient γ_i in Equation (2.58) at the required T and P for the calculation of the solubility of an API. The remaining variables in the equation are the melting properties of the drug (Δh_i^{fus} , T_i^{fus}) and $\Delta c_{p,i}$ which are usually obtained from experimental data. It is noteworthy that the solubility equation is typically simplified by neglecting the terms which include $\Delta c_{p,i}$, due to the absence of experimental data for this property. It is assumed that the second and third terms on the right hand side of Equation (2.58) have approximately equal contributions and will cancel out due to their opposite signs. However, it has been suggested that the contribution of $\Delta c_{p,i}$ on the solubility calculation may not be negligible for some compounds and can introduce large errors to the calculation if neglected [172–174]. Therefore, in the absence of experimental data for $\Delta c_{p,i}$, the use of approximations for this property in Equation (2.58) is recommended [172].

Hutacharoen et al. [39, 40], Febra et al. [41] and Wehbe et al. [44] used the SAFT- γ Mie EoS to predict the solubilities of APIs in various solvents (cf. Table 2.1). The APIs studied include ibuprofen, ketoprofen, azelaic acid, phenylacetic acid, lovastatin, simvastatin, and mefenamic acid. These APIs are complex multifunctional compounds, some of which contain multiple aromatic rings as well as alcohol, ester, amine and acid functional groups. Accurate predictions of the solubility of these APIs in multiple solvents were obtained using the SAFT- γ Mie GC approach by modelling the compounds using the appropriate functionality within this GC framework. In Figure 2.4, an overview of the accuracy of the SAFT- γ Mie solubility calculations [39–41] compared to experimental data is shown. It is worth mentioning that the majority of the SAFT- γ Mie solubility calculations in Figure 2.4 correspond to predictions (shown as unfilled symbols), as the group and group-group interactions that were used to model the APIs within the SAFT- γ Mie framework were developed using experimental data for other compounds. For example, a large number of the groups and group-group interactions for mefenamic acid were

developed using experimental properties of benzoic acid and diphenylamine [41]. It can be seen in the figure that the SAFT- γ Mie solubility calculations fall within an order of magnitude of the experimental measurements, with the majority of the calculations falling within less than half an order of magnitude away from the experimental data. Two exceptions of this behaviour can be seen in Figure 2.4 corresponding to the solubility of mefenamic acid in 2-methyl-1-propanol and azelaic acid in water. However, within the context of predictive methods for solubility, the results shown in the figure can be considered an excellent level of agreement, particularly noting the ability of the SAFT- γ Mie calculations to capture the very low solubility ($x_i \approx 10^{-5}$) of some APIs (bottom left corner in the figure) as well as the significantly higher API solubilities ($x_i > 10^{-1}$) (top right corner in the figure) of other cases.

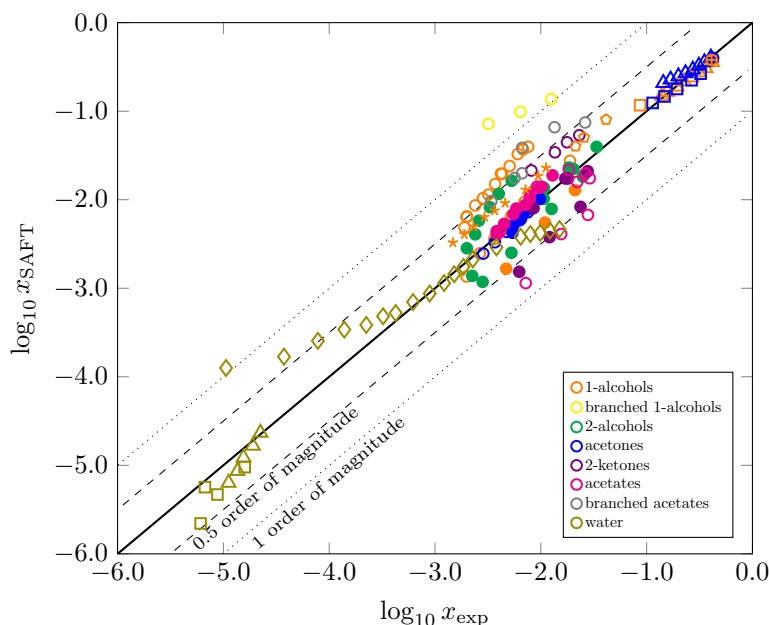


Figure 2.4: Parity plot of the solubility of mefenamic acid (circles), ibuprofen (triangles), ketoprofen (squares), azelaic acid (diamonds), simvastatin (pentagons), lovastatin (stars) and phenylacetic acid (oplus) in various solvents calculated with the SAFT- γ Mie group-contribution approach compared to experimental measurements. The experimental data used in the parameter estimation of the SAFT- γ Mie groups are represented by filled symbols and those not used by unfilled symbols. Thermodynamic conditions are detailed in refs. [39–41, 44]

2.3 Concluding remarks

The SAFT EoS has received increasing attention in the academic and industrial community in the last three decades, after the publication of the original papers, due to its successful application to a wide range of complex fluid systems. This led to the development of various versions of the SAFT approach over the years. In this chapter we have provided an overview of homonuclear and heteronuclear SAFT-based GC approaches which have been developed in recent years, including their application to pharmaceutical systems containing APIs (small molecules), amino acids and peptides (medium and large molecules).

The detailed theory of the SAFT- γ Mie GC approach which can be used to predict thermodynamic properties of molecular systems and electrolytes is also presented in this chapter. We demonstrate the methodology for calculating fluid phase equilibria and solid-liquid solubility using this approach. Additionally, we highlight previous work which shows the applicability of this GC approach for successfully predicting properties of relevance to pharmaceutical mixtures.

The SAFT- γ Mie GC approach is used in the following chapters to predict the phase behaviour and solubility of various pharmaceutical mixtures. A framework for predicting the phase behaviour, solubility and pH-solubility profiles of ionisable monoprotic APIs using the SAFT- γ Mie GC approach is developed in Chapters 3. The applicability of the GC approach is extended in Chapter 4 to predict the effect of pH on the solubility of diprotic APIs. In Chapter 5, we demonstrate the use of the SAFT- γ Mie EoS to model the phase behaviour of formaldehyde (a common degradation agent produced by some pharmaceutical excipients) in aqueous and methanolic solutions. A physical approach is used to model the oligomerization reactions that take place in formaldehyde mixtures, such that, the reaction products are treated implicitly and are considered to be aggregates of the reactants.

Chapter 3

Phase behaviour and pH-solubility profile prediction of aqueous buffered solutions of ibuprofen and ketoprofen

3.1 Introduction

Low water solubility is a common characteristic of many active pharmaceutical ingredients (APIs) that leads to challenges during drug discovery and process design. A review [175] of the solubility of the top 200 oral drugs in use in four regions in the world (US, GB, Spain, and Japan) indicated that around 40% of these are practically insoluble, with aqueous solubilities lower than 0.1 mg ml^{-1} , and approximately 33.3% of newly synthesized medicinal compounds have a solubility lower than 0.01 mg ml^{-1} in water, according to another study [176]. Furthermore, the solubility of a drug in water is directly relevant to its bioavailability, as the drug must be in aqueous solution at the site of absorption in the body [177, 178]. In aiming to increase the solubility of APIs, a range of techniques are commonly used, including particle-size reduction, the use of cosolvents, polymorphism and cocrystal formation, the addition of surfactants, and salt formation [178].

For ionisable APIs, which are in essence weak electrolytes that only partially dissociate in water, salt formation is the preferred method for solubility enhancement [179–181]. The degree of ionisation of the molecule in water depends on its dissociation constant (pK_a), such that for weak acids and weak bases, dissociation/ionisation increases at pH values higher and lower than the pK_a , respectively, leading to increased solubility, facilitated by favorable ion–dipole interactions, and eventual salt formation in most cases.

A number of experimental methods are available for measuring the pH-solubility profile of APIs [182]. The simplest is perhaps the saturation shake-flask (SSF) method, based on measuring the solubility of the drug at a single pH at equilibrium (for a given temperature and pressure) and repeatedly screening different pH solutions. The SSF method was used by Avdeef et al. [183] to obtain the pH-solubility profiles of twelve API molecules including ibuprofen (IBP) and ketoprofen (KTP). They reported aqueous solubilities at 298 K for IBP at pH 2.0, 5.1 and 8.0 of 2.84×10^{-4} , 1.39×10^{-3} and 0.24 mol L^{-1} , respectively, and for KTP at pH 2.5, 4.5 and 7.0 they reported solubilities of 4.22×10^{-4} , 1.90×10^{-3} and 0.34 mol L^{-1} , respectively. Microtitre plate methods are a modification of the SSF method with increased throughput and reduced API sample size using a robotic liquid dispenser, which have also been used for pH-solubility measurements of ibuprofen and ketoprofen [184], and have replaced the SSF method in several pharmaceutical companies [185–189].

More-automated approaches include potentiometric techniques, also known as dissolution template titration (DTT) methods [190], which use the pK_a and the octanol–water partition coefficient ($\log K_{OW}$) of the API as inputs to estimate the intrinsic solubility (the solubility of the neutral compound), using the pH electrode as the only detector. The intrinsic solubilities of IBP and KTP determined by DTT at the same temperature of 298 K were reported by Avdeef [184] as 2.38×10^{-4} and $4.64 \times 10^{-4} \text{ mol L}^{-1}$, respectively, in good agreement with solubility measurements using the SSF method. In the DTT method, the entire titration curve is calculated before the assay begins, such that the optimal amount of API to be used is known at each pH, which speeds the time required for the measurements and reduces API wastage.

In order to overcome very long equilibration times during solubility measurements of poorly

soluble APIs, the facilitated dissolution method (FDM) [191] can also be useful. Here, a small amount of immiscible organic solvent is added to the aqueous solution to facilitate the dissolution of the API. The method is especially useful for very low solubility measurements of about 1 ng ml^{-1} [192]. To use this method for ionizable compounds, such as ibuprofen and ketoprofen, a suitable buffered solution must be employed, as a poorly buffered (or buffer-free) solution can affect the pH-dependent reaction equilibria and alter the aqueous solubility of the API [184].

In summary, these experimental approaches require high precision due to the very low solubility of most APIs in water, poor wettability, the formation of aggregates, complexes and micelles, polymorphism, and several other factors [182]. Because of this, experiments can become time-consuming and costly, with numerous repeated evaluations often required to ensure reliability. In this context, computer-aided approaches provide an attractive alternative. In particular, predictive thermodynamic modelling can provide a useful mechanistic understanding of the solubility process, increasing productivity and reducing risks [5].

There is, however, a lack of previous work targeted at investigating predictive techniques capable of delivering accurate predictions of the pH-solubility profile of a given API. Empirical equations expressed in terms of the experimental pK_a and the intrinsic solubility of the drug, known as the Henderson-Hasselbalch (HH) equations [193], are commonly used to determine the solubility of a drug at a given pH, before salt formation [176, 184]. Bergström et al. [194] have used the HH equations for the prediction of pH-dependent solubility of multiple drugs and concluded that these relations provide only rough estimates of the profiles; (the predicted solubility of the drugs using the HH equations at pH 6.5 was, on average, 10 times greater than the experimental measurements). The deviations of the HH calculations from experimental data were expected, as the solubility of the salt cannot be calculated using these equations. Therefore, the HH equations are more reliably used to provide accurate correlations of known solubility data, but they are not appropriate to use predictively.

From a molecular-modelling perspective, a key requirement in the development of a model for the prediction or calculation of the solubility and pH-solubility profiles of ionisable APIs is to

account for the presence of charged species in solution. Activity-coefficient models are widely used thermodynamic models for the calculation of thermodynamic properties of electrolyte solutions, and have been employed in solubility modelling for pharmaceutical applications [195]. However, to our knowledge, the use of activity-coefficient approaches for the calculation of solubility as a function of pH has been limited to amino acids. These studies include the works of Figueiredo et al. [196], and of Pinho et al. [29]; in both of these studies, the universal functional activity-coefficient (UNIFAC) model [18] was combined with the Debye–Hückel [197] equation to determine the aqueous solubility of a number of amino acids and peptides. In addition, the non-random two liquid (NRTL) activity-coefficient model [198], was used by Tseng et al. [199] and by Lee et al. [200] to determine the pH-dependent solubility of several divalent and trivalent amino acids.

A more relevant study to our current work was presented by Cassens et al. [113], who used the electrolyte perturbed-chain statistical associating fluid theory (ePC-SAFT) [77] equation of state (EoS) to describe the pH-solubility profiles of a number of pharmaceutical bases and the formation of their salts. In their work, the APIs were modeled as weak electrolytes, by including the experimental pK_a value for each base, and the pH of the solution was varied by explicitly considering HCl (as a strong electrolyte) and H_3PO_4 (as a weak electrolyte) as the pH-changing agents. A good description of the calculated solubility as a function of pH was presented, although experimental data were used to adjust the model parameters for the uncharged species, the binary interactions for each solution, and the solubility product of the salts formed.

Here, we use the SAFT- γ Mie group-contribution (GC) EoS [35–38], in which molecules are modelled as heteronuclear chains formed from fused spherical segments, each representing the distinct chemical moieties (functional groups) characterising a molecule, to predict the solubility of ibuprofen and ketoprofen as a function of pH. The SAFT- γ Mie GC approach has been shown to deliver a high-level of accuracy in the modelling and prediction of a variety of thermodynamic properties, including the phase equilibria, excess properties of mixing, and second-derivative properties for a variety of mixtures [35, 36, 39, 148], and recently in the modelling of the solubility of mefenamic acid [41] (a non-steroidal anti-inflammatory drug). Of

special relevance here is previous work [39] in which accurate predictions of the liquid–liquid equilibria (LLE), vapor–liquid equilibria (VLE) and vapor–liquid–liquid equilibria (VLLE) of aqueous mixtures of *n*-alkanes and 1-alcohols were presented; accurate predictions were also made of the solid–liquid equilibria (SLE) of several APIs in different solvents, including aqueous solutions of ibuprofen and ketoprofen. We proceed in our current work to provide a prediction of the complete phase behaviour (VLE, LLE, and SLE) of these solutions and extend the thermodynamic modelling to predict the pH-dependent solubility of a range of ibuprofen and ketoprofen salts in various buffer solutions.

The SAFT EoS for variable range interactions with a Mie form (SAFT-VR Mie) [146] was extended to treat electrolyte solutions by including Coulombic and Born [154] contributions in previous work [150]. It was shown that only one unlike ion–solvent adjustable parameter is needed per ion–solvent pair in order to obtain an accurate description of a wide range of thermophysical properties of aqueous solutions of strong electrolytes [150], including the density, vapour pressure, osmotic coefficient, mean ionic activity coefficient, and the Gibbs free energy of solvation of the ions. A similar framework can be applied with the SAFT- γ Mie GC approach, accounting for the appropriate ionic free-energy contributions in a group-contribution framework [42, 43, 149]. The approach has been used in a study of complex ionic mixtures of choline and geranate (CAGE), a deep-eutectic-solvent mixture of therapeutic interest, in which predictions of osmotic coefficient data in very good agreement with experiment were presented [42, 43]. Recently, the same approach has been used to model a range of weak acids [201], combining chemical and phase equilibria in the model to deliver accurate predictions of the pH of saturated solutions of CO₂ as well as of the speciation of the aqueous solutions.

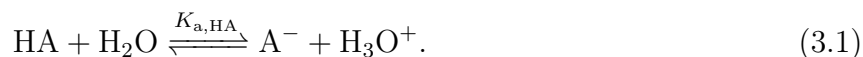
In this chapter, we extend the application of the SAFT- γ Mie GC approach to model the pH-solubility profile of aqueous solutions of ibuprofen and ketoprofen for a range of buffers. In particular, we use the approach to capture the complex speciation phenomena that take place under changing pH conditions, including the variation of the acidic buffering agents and counter ions (basic buffering agents). In the following section we present the thermodynamic framework. The development of models within the SAFT- γ Mie GC approach relevant to the aqueous solutions of APIs is provided in Section 3.3. We demonstrate the capability of the

approach in delivering accurate predictions of the phase diagrams and pH-solubility profiles of the APIs in Section 3.4, including a discussion on the effect of varying the temperature and counter ions in these solutions. Final remarks are provided in Section 3.5.

3.2 Thermodynamic modelling of pH-dependent solubility

The solid–liquid solubility of a compound (the solute) in a solvent is defined as the maximum amount of the solid solute that can be dissolved in a given amount of the liquid solvent to form a homogeneous solution (at a given temperature T , and pressure P). In the case of ionisable solutes, the solute may dissociate into the corresponding ionic species in the liquid phase. Moreover, the term "solubility" refers to the sum of the concentrations of the charged and uncharged forms of the API in a saturated solution, while "intrinsic solubility" refers to the concentration of only the neutral solute in a saturated solution [176].

For the purpose of the current chapter, the discussion is focused on a weakly acidic monoprotic API (labelled HA), which dissociates in water as



The equilibrium between the molecular (neutral) API (HA), in the solid and liquid phases is shown in Figure 3.1 (a), together with the dissociation that takes place in the liquid phase to give rise to the anionic form A^- and formation of hydronium H_3O^+ cations as determined by the value of the relevant thermodynamic equilibrium constant, $K_{\text{a,HA}}$. We also account for the dissociation of water into hydronium H_3O^+ and hydroxide OH^- ions, as given by the ionic product of water K_{W} . In addition, a strong acid HX and strong base $\text{B}_{\nu_{\text{B}}}\text{OH}_{\nu_{\text{OH}}}$ (where ν_i , $i=\{\text{B,OH}\}$ is the stoichiometric coefficient) are used to buffer the solution, and are modelled as fully dissociated into their respective ions; thus species X^- and $\text{B}^{z_{\text{B}}^+}$ (where z_{B}^+ is the charge number of B) are also present in the solution.

A schematic pH-solubility profile of a weak monoprotic acid is shown in Figure 3.1 (b). At

low pH, the solubility of the API corresponds to its intrinsic solubility (as no dissociation of HA occurs). As the pH increases with the addition of the counter ion (B^{z_B+}), the solubility increases due to the rise in the concentration of A^- in solution. At high pH, a salt $A_{\nu_A}B_{\nu_B}$ forms and eventually precipitates as the solubility limit of the salt, $K_{sp,A_{\nu_A}B_{\nu_B}}$, is reached. The pH value at the intersection between the solubility curves of the API and the salt determines the $pH^{\max}$, beyond which no further increase in solubility is possible. The value of $pH^{\max}$ also plays an important role in identifying the tendency for salt disproportionation [202] (the conversion of the salt to its free ions), a particularly undesirable process in pharmaceutical applications [203, 204] as it signals a decrease in solubility.

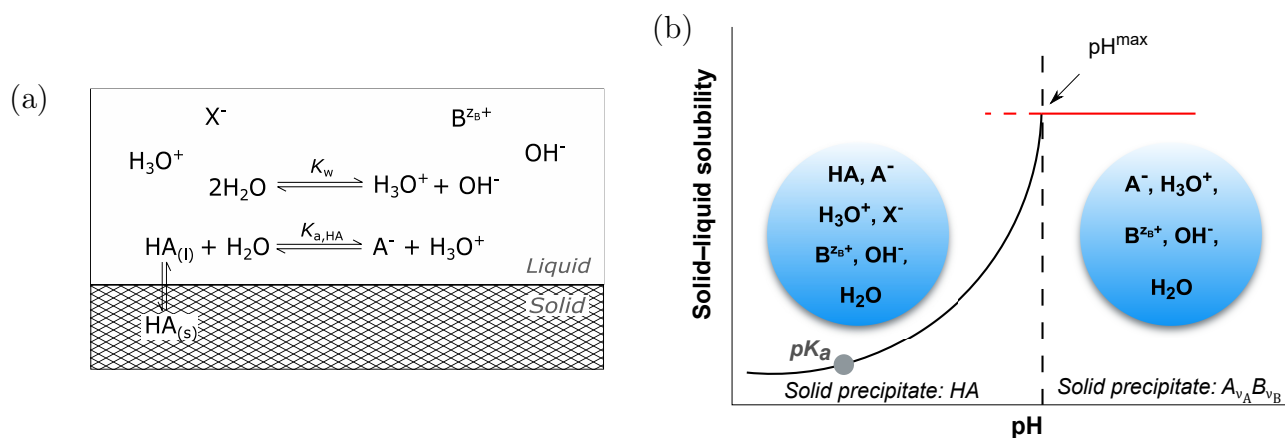


Figure 3.1: Schematic representation of the equilibria and speciation considered to model the pH-solubility profile of a weakly monoprotic acidic API (HA here) and its salt $A_{\nu_A}B_{\nu_B}$. (a) The three equilibria taking place before salt formation: the solid-liquid equilibrium of the neutral API, HA; the dissociation of HA in the liquid phase into its ionised form, A^- , regulated by the thermodynamic equilibrium constant $K_{a,HA}$, and the dissociation of water given by the ionic product of water K_w . HX and $B_{\nu_B}OH_{\nu_{OH}}$ are the strong acid and base buffers modelled as fully dissociated into X^- and H_3O^+ , and B^{z_B+} and OH^- ions, respectively. The solid phase consists of pure neutral HA. (b) The pH-solubility profile including the solubility of the free form of the API (black) and its salt $A_{\nu_A}B_{\nu_B}$ (red) intersecting at $pH^{\max}$. The blue regions are shown to indicate the components present in the liquid phase in the two regions of the plot.

3.2.1 Intrinsic solubility of HA

At a specified P and T the solid-liquid intrinsic solubility of the neutral API (here HA) can be calculated using the condition for phase equilibrium (the equality of the chemical potential of the API in the liquid and in the solid phase), assuming the solid phase to be pure HA, which

leads to the well-known solubility equation [25]:

$$\begin{aligned} \ln x_{\text{HA}}^{\text{L}}(T, P) = & \frac{\Delta h_{\text{HA}}^{\text{fus}}(T_{\text{HA}}^{\text{fus}}, P)}{R} \left(\frac{1}{T_{\text{HA}}^{\text{fus}}} - \frac{1}{T} \right) + \frac{1}{RT} \int_T^{T_{\text{HA}}^{\text{fus}}} \Delta c_{p,\text{HA}}(T', P) dT' \\ & - \frac{1}{R} \int_T^{T_{\text{HA}}^{\text{fus}}} \frac{\Delta c_{p,\text{HA}}(T', P)}{T'} dT' - \ln \gamma_{\text{HA}}(T, P, \mathbf{x}^{\text{L}}), \end{aligned} \quad (3.2)$$

where $\mathbf{x}^{\text{L}}$ is the composition vector of the liquid phase containing the solvent, buffer ions, and API (in molecular and dissociated form). $\Delta h_{\text{HA}}^{\text{fus}}$ is the molar enthalpy of fusion of the API at the melting temperature $T_{\text{HA}}^{\text{fus}}$, R is the universal gas constant, $\Delta c_{p,\text{HA}} = c_{p,\text{HA}}^{\text{L}} - c_{p,\text{HA}}^{\text{S}}$ is the difference between the isobaric heat capacity of the API in the liquid and solid phases, respectively, and γ_{HA} is the symmetric activity coefficient of HA in the liquid phase at the given T and P , calculated as

$$\gamma_{\text{HA}}(T, P, \mathbf{x}^{\text{L}}) = \frac{\hat{\varphi}_{\text{HA}}(T, P, \mathbf{x}^{\text{L}})}{\varphi_{\text{HA}}^*(T, P)}, \quad (3.3)$$

where $\hat{\varphi}_{\text{HA}}$ is the fugacity coefficient of HA in the mixture, and φ_{HA}^* the fugacity coefficient of pure HA at the same T and P .

3.2.2 Solubility at $\text{pH} < \text{pH}^{\text{max}}$

In solution, at pH below pH^{max} the API dissociates following Reaction (3.1). The Gibbs free energy change $\Delta_{\text{R}}G$ for this reaction is given by

$$\Delta_{\text{R}}G(T, P, \mathbf{x}) = \sum_{i=1}^N \nu_i \mu_i = 0, \quad (3.4)$$

where the summation is over the relevant species in the equilibrium, HA, H_2O , A^- , and H_3O^+ here, and μ_i and ν_i are the chemical potential and stoichiometric coefficient of species i , respectively, such that,

$$\Delta_{\text{R}}G(T, P, \mathbf{x}) = \mu_{\text{A}^-} + \mu_{\text{H}_3\text{O}^+} - \mu_{\text{HA}} - \mu_{\text{H}_2\text{O}} = 0. \quad (3.5)$$

The chemical potential of i can be written as the sum of the chemical potential at a reference

state μ_i° and a term capturing the properties of the real solution; i.e.,

$$\mu_i(T, P, \mathbf{x}) = \mu_i^\circ + RT \ln a_i(T, P, \mathbf{x}) , \quad (3.6)$$

where a_i is the activity of i . Inserting the normalization of Equation (3.6) in (3.5) leads to

$$\exp \left(-\frac{\mu_{A^-}^\circ + \mu_{H_3O^+}^\circ - \mu_{HA}^\circ - \mu_{H_2O}^\circ}{RT} \right) = \frac{a_{A^-} a_{H_3O^+}}{a_{HA} a_{H_2O}} . \quad (3.7)$$

The left-hand side of this equation corresponds to the thermodynamic equilibrium constant K_a [168]. Here,

$$K_{a,HA} = \exp \left(-\frac{\mu_{A^-}^\circ + \mu_{H_3O^+}^\circ - \mu_{HA}^\circ - \mu_{H_2O}^\circ}{RT} \right) , \quad (3.8)$$

and therefore,

$$K_{a,HA} \equiv \frac{a_{A^-} a_{H_3O^+}}{a_{HA} a_{H_2O}} . \quad (3.9)$$

In our model, we adopt the reference state for the chemical potential of the ion to be an infinitely dilute solution in water so that,

$$a_i = x_i \tilde{\gamma}_i = x_i \frac{\gamma_i}{\gamma_i^\infty} , \quad (3.10)$$

where $\tilde{\gamma}_i$ is the asymmetric activity coefficient of i and γ_i^∞ the symmetric activity coefficient (of i) at infinite dilution.

The dissociation constant K_a is often reported as the negative logarithm in base 10, pK_a [205]. Furthermore, in most experimental studies, the standard state considered for the determination of the pK_a is a concentrated solution of a strong inert electrolyte, usually a 0.15 M KCl or NaCl solution, referred to as a constant ionic medium [182]. The acid of interest is considered to be very dilute in the reference solution, such that its activity coefficient is usually assumed to be close to one, and the dissociation constant is reported as a ratio of concentrations instead of activities [117].

Here, we use the expression of the (true) thermodynamic equilibrium constant in molal units $K_{a,HA}^m = K_{a,HA} a_{H_2O}$, noting that the molal- and molarity-based constants can be treated as identical for very dilute aqueous solutions (where the density of the solution can be assumed

to be equal to that of pure water), so that:

$$K_{a,HA}^m = \left(\frac{m_{A^-} m_{H_3O^+}}{m_{HA}} \right) \left(\frac{\tilde{\gamma}_{m,A^-} \tilde{\gamma}_{m,H_3O^+}}{\tilde{\gamma}_{m,HA}} \right) , \quad (3.11)$$

where m_i is the molality of i and $\tilde{\gamma}_{m,i}$ is the molal asymmetric activity coefficient calculated as

$$\tilde{\gamma}_{m,i} = x_{H_2O} \tilde{\gamma}_i . \quad (3.12)$$

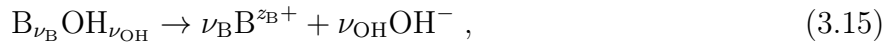
Within this framework pK_a is calculated as

$$pK_a = -\log_{10}(K_{a,HA}^m) . \quad (3.13)$$

In order to increase the pH of a solution containing an acidic compound to promote its dissociation, a base is added as a buffer and a source of counter ions (cf. Figure 3.1 (b)). If the source of counter ions is a weak base B, the ionisation of the base in aqueous solution takes place according to



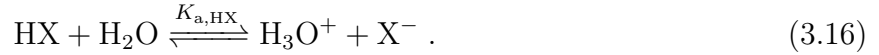
where BH^+ is the ionised form of the base and the counter ion in the solution, and the corresponding molal-based thermodynamic equilibrium constant of this dissociation K_{a,BH^+}^m is also given by Equation (3.11), but with the relevant species of Equation (3.14). In the case that a strong base $B_{\nu_B}OH_{\nu_{OH}}$ is used as the source of counter ions, complete dissociation is assumed, i.e.,



where B^{z_B+} is the counter ion in the model. The choice of counter ion depends on which base would shift the pH of the solution more effectively to promote salt formation with the associated increase in the solubility of the API [181].

Conversely, in order to decrease the pH of the solution below the pK_a , an acidic buffer (HX)

must be used:



The corresponding $K_{\text{a,HX}}^m$ is calculated from Equation (3.11) with the appropriate species. As with the case of the base, if a strong acid buffer is used, it is modelled as a fully dissociated strong electrolyte.

In addition, the dissociation of water,



is also treated in our model, with the corresponding molal thermodynamic ionic product K_{w} calculated as

$$K_{\text{W}} = (m_{\text{H}_3\text{O}^+} m_{\text{OH}^-}) (\tilde{\gamma}_{m,\text{H}_3\text{O}^+} \tilde{\gamma}_{m,\text{OH}^-}) , \quad (3.18)$$

which is more commonly reported in terms of $\text{p}K_{\text{W}}$:

$$\text{p}K_{\text{W}} = -\log_{10}(K_{\text{W}}) . \quad (3.19)$$

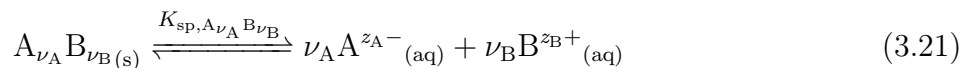
The solubility of the API, S_{API} , is given as the sum of the intrinsic solubility and the concentration of ionised API, often reported in molarity units (moles/volume), so that,

$$S_{\text{API}} = \rho M_{\text{w}} (m_{\text{HA}} + m_{\text{A}^-}) , \quad (3.20)$$

where ρ is the molar density of the solution and M_{w} is the molecular mass of water.

3.2.3 Salt formation: solubility at $\text{pH} > \text{pH}^{\text{max}}$

The API salt ($\text{A}_{\nu_{\text{A}}} \text{B}_{\nu_{\text{B}}}$) begins to precipitate when its solubility product ($K_{\text{sp},\text{A}_{\nu_{\text{A}}} \text{B}_{\nu_{\text{B}}}}$) is reached. Assuming that the salt dissociates in solution into its constituent solvated ions:



the solubility product is given by [25, 181]

$$K_{\text{sp}, A_{\nu_A} B_{\nu_B}} = m_{A_{\nu_A} B_{\nu_B}}^{(\nu_A + \nu_B)} \gamma_{m, \pm}^{(\nu_A + \nu_B)} \frac{\nu_A^{\nu_A} \nu_B^{\nu_B}}{m^{\circ(\nu_A + \nu_B)}} , \quad (3.22)$$

where, $m_{A_{\nu_A} B_{\nu_B}}$ is the solubility limit of the salt, m° is the standard-state molality of 1 mol kg^{-1} , and

$$\tilde{\gamma}_{m, \pm} = (\tilde{\gamma}_{m, A}^{\nu_A} \tilde{\gamma}_{m, B}^{\nu_B})^{1/(\nu_A + \nu_B)} \quad (3.23)$$

is the mean ionic activity coefficient (MIAC) in terms of the molal asymmetric activity coefficient of the ions.

The corresponding pH at which the salt starts to precipitate (the pH corresponding to the $K_{\text{sp}, A_{\nu_A} B_{\nu_B}}$) is often labelled as pH^{max} (see Figure 3.1 (b)). At this point the ionised API (A^-) and counter-ion (B^{z+}) concentrations are balanced. Any further addition of counter ions to increase the pH, leads to an excess of these ions, and to a drop in the solubility of the salt, (cf. Equations (3.22) and (3.24)); this is known as the common-ion effect. A more detailed discussion on the common-ion effect can be found in ref. [176]. The value of pH^{max} value is therefore expected to be a point of maximum solubility in the pH-solubility profile of a monoprotic API.

In this regime, the reported solubility, S_{API} , accounts only for the ionised form of the API and is given by

$$S_{\text{API}} = \rho M_W m_{A^-} . \quad (3.24)$$

3.2.4 pH-solubility profile

In order to compute the solubility-pH profile of weakly acidic APIs, we use experimental values for the relevant equilibrium constants, $K_{a, \text{HA}}^m$, K_w , and $K_{\text{sp}, A_{\nu_A} B_{\nu_B}}$, obtained from the literature [183, 184], and the SAFT- γ Mie GC EoS for the calculation of the activity coefficients and the density of the solution.

For a given temperature and pressure, the mole fraction of uncharged API in the liquid

phase x_{HA}^{L} is obtained using Equation (3.2). Note that this quantity is directly related to m_{HA} . To change the pH of the solution a buffering agent is added. If a weak acid HX is used to lower the pH, Equation (3.2), with $K_{\text{a,HA}}^m$, $K_{\text{a,HX}}^m$ and K_w , in addition to the condition of electroneutrality given by

$$\sum_{i=1}^N q_i m_i = 0 , \quad (3.25)$$

where q_i is the charge of species i and N is the total number of species in the system, are solved to obtain the concentrations of each of the species in the solution (in this case H_2O , H_3O^+ , OH^- , HX, X^- , HA and A^-), at $\text{pH} < \text{pH}^{\text{max}}$. If a weak base B is used to increase the pH of the solution and as a source of counter ions, Equation (3.2), with $K_{\text{a,HA}}^m$, $K_{\text{a,BH}^+}^m$, K_w , and the condition of electroneutrality, are solved to obtain the concentrations of the corresponding species in the solution (in this case H_2O , H_3O^+ , OH^- , BH^+ , B, HA and A^-), at $\text{pH} < \text{pH}^{\text{max}}$. However, if a strong acid or base is used to change the pH of the solution, each of which is modelled as a strong electrolyte, only $K_{\text{a,HA}}^m$ and K_w are solved, together with Equations (3.2) and (3.25) to obtain the concentrations of each of the species in the solution (in this case H_2O , H_3O^+ , OH^- , HA, A^- , and BH^+ or X^-), at $\text{pH} < \text{pH}^{\text{max}}$. The pH of the solution can be calculated from the corresponding hydronium-ion activity as [205]

$$\text{pH} = -\log_{10}(a_{\text{H}_3\text{O}^+}) . \quad (3.26)$$

As the pH changes, the solubility product of $\text{A}_{\nu_A}\text{B}_{\nu_B}$ is calculated (using Equation 3.22) to identify when the salt is formed and precipitated. During this process, the solubility of the API or the salt is calculated using either Equation (3.20) or (3.24) depending on which solid is precipitating, i.e., Equation (3.20) is used for a pH below the point at which $K_{\text{sp,A}_{\nu_A}\text{B}_{\nu_B}}^{\text{calc}} = K_{\text{sp,A}_{\nu_A}\text{B}_{\nu_B}}^{\text{exp}}$ and Equation (3.24) is used for a pH above this point.

3.3 SAFT- γ Mie theory and model development

3.3.1 Equation of state

In the SAFT- γ Mie [35–38] GC EoS, molecules are modelled as heteronuclear chains of fused spherical segments, which represent the functional groups that make up the molecule. The segments interact via Mie potentials of variable range, with association sites embedded in the groups/segments, if necessary, to account for short-range directional forces. A brief overview of the theory is provided here; the reader is referred to Chapter 2 (Section 2.2) and the original publications [35–38, 149, 150] for further details.

The Helmholtz free energy A of a mixture is written as a sum of terms, based on a perturbation approach:

$$A = A^{\text{Ideal}} + A^{\text{Mono}} + A^{\text{Chain}} + A^{\text{Assoc}} + A^{\text{Ion}} + A^{\text{Born}} , \quad (3.27)$$

where A^{Ideal} corresponds to the ideal-gas free energy, A^{Mono} accounts for the interaction of monomeric segments through Mie potentials, A^{Chain} is the contribution to the free energy due to the formation of chains of molecules from the fused Mie segments, and A^{Assoc} accounts for molecular association through short-range directional interactions. These first four terms describe the non-electrolyte contributions [35–38]. The presence of charged species is accounted for in the last two terms [149, 150], which are developed within a primitive-model formalism, with the electrostatic nature of the solvent described implicitly with a uniform dielectric medium. The A^{Ion} term represents the Coulombic ion–ion interactions, obtained using the primitive unrestricted Mean Spherical Approximation (MSA) [152, 153] model, and the A^{Born} term accounts for the ion–solvent electrostatic interactions following the Born [154] model.

3.3.2 Phase-equilibrium calculations

The Helmholtz free energy of a fluid mixture is given as a function of temperature, volume V , and the vector of moles $\mathbf{N}$ of i components. Knowledge of the pressure, chemical potential and fugacity coefficient of a given component i , allows for the calculation of fluid-phase behaviour and solution properties. These can be obtained from the Helmholtz free energy following standard thermodynamic relations [168], as given by Equations (2.44), (2.45), and (2.46), respectively.

3.3.3 Model development

In order to model the phase behaviour and pH-solubility profile of ibuprofen and ketoprofen using the SAFT- γ Mie GC approach, it is necessary to characterize the individual group parameters of all the compounds involved. The groups required to model these APIs are shown in Figure 3.2. These are: ($3 \times \text{CH}_3$, $4 \times \text{aCH}$, $1 \times \text{aCCH}$, $1 \times \text{aCCH}_2$, $1 \times \text{CH}$, $1 \times \text{COOH}$) for molecular ibuprofen; replacing the COOH group in the latter by COO^- for ionised ibuprofen; ($1 \times \text{CH}_3$, $9 \times \text{aCH}$, $1 \times \text{aCCH}$, $1 \times \text{aCCOaC}$, $1 \times \text{COOH}$) for molecular ketoprofen, with the additional COO^- group for ionised ketoprofen. For both APIs, the solvent considered is water ($1 \times \text{H}_2\text{O}$) and the dissociation of the acid produces a hydronium ion ($1 \times \text{H}_3\text{O}^+$) and a hydroxyl ion ($1 \times \text{OH}^-$).

The pH of the solutions is varied by adding acidic and basic buffers. In particular, acetate buffer is used to control the pH in the range of 4.0 – 4.6 [206], and is accounted for by adding acetic acid (which is a weak acid, $\text{p}K_{\text{a}} = 4.74$) in the molecular ($1 \times \text{CH}_3$, $1 \times \text{COOH}$) and ionised ($1 \times \text{CH}_3$, $1 \times \text{COO}^-$) forms, in addition to fully dissociated sodium acetate ($1 \times \text{Na}^+$, $1 \times \text{CH}_3$, $1 \times \text{COO}^-$), similar to the experimental buffer preparation [207]. A different acidic buffering agent, HCl (a strong acid), is used in the range of pH 2.0 – 5.0, and is modelled as fully dissociated into its respective ions: H_3O^+ and Cl^- .

Various strong basic buffering agents are used as the source of monovalent and divalent

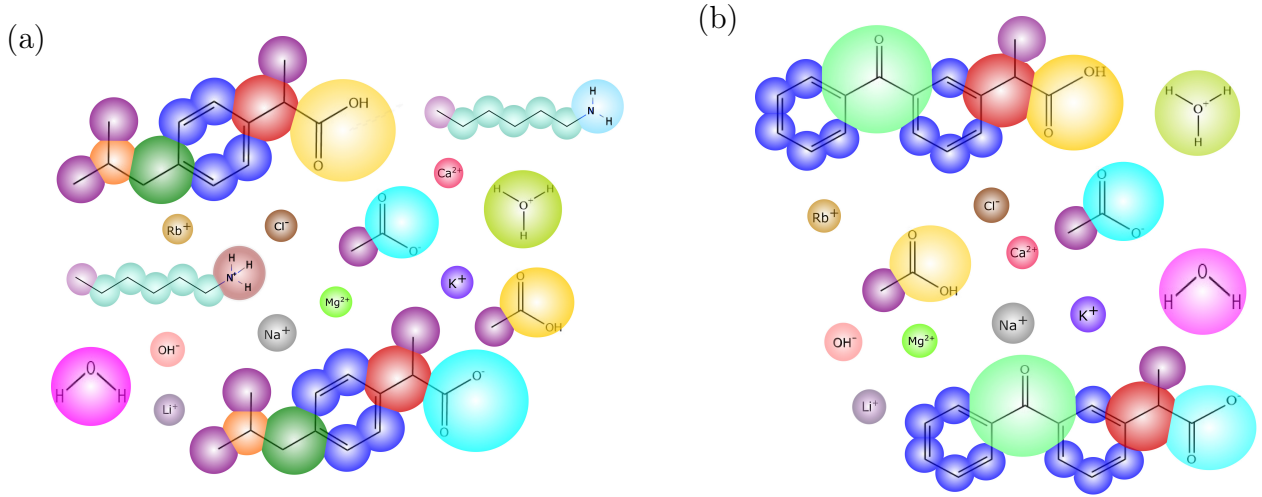


Figure 3.2: Chemical structures of the species considered in the solutions, including buffers, together with their corresponding SAFT- γ Mie representation. Each of the coloured spheres corresponds to a different group, and the same group is represented in the same colour, for aqueous buffered solutions of (a) ibuprofen and (b) ketoprofen. The color code for the groups is: purple for CH₃, turquoise for CH₂, navy blue for aCH, orange for CH, dark green for aCCH₂, red for aCCH, yellow for COOH, maroon for NH₃⁺, light blue for NH₂, magenta for H₂O, cyan for COO⁻, lime for H₃O⁺, mint green for aCCOaC, light brown for Rb⁺, brown for Cl⁻, light pink for OH⁻, grey for Na⁺, light green for Mg²⁺, pink for Ca²⁺, violet for K⁺ and lilac for Li⁺.

counter ions in the solutions (pH 5.0 – 12.0), specifically: NaOH, KOH, LiOH, RbOH, Ca(OH)₂, and Mg(OH)₂, which are also modelled as fully dissociated into OH⁻ and Na⁺, K⁺, Li⁺, Rb⁺, Ca²⁺, and Mg²⁺ ions, respectively. Weak bases including *n*-hexylamine ($pK_a = 10.64$ [208]; 1 $\times$ CH₃, 5 $\times$ CH₂, 1 $\times$ NH₂), *n*-octylamine ($pK_a = 10.65$ [209]; 1 $\times$ CH₃, 7 $\times$ CH₂, 1 $\times$ NH₂), benzylamine ($pK_a = 9.33$ [210]; 5 $\times$ aCH, 1 $\times$ aCCH₂, 1 $\times$ NH₂), and *tert*-butylamine ($pK_a = 10.68$ [209]; 3 $\times$ CH₃, 1 $\times$ C, 1 $\times$ NH₂), are also used as basic buffering agents and a source of counter ions (ionised amine) in the solution. The corresponding ionised amines are modeled using the same groups as the molecular amines replacing NH₂ with NH₃⁺.

In the SAFT- γ Mie approach, as discussed in Section 2.2, a group k is characterised by a number ν_k^* of identical spherical segments, a shape factor S_k , the segment energy of interaction ε_{kk} , the segment diameter σ_{kk} , the exponents of the Mie potential λ_{kk}^r and λ_{kk}^a and, where relevant, the parameters describing any site-site association interactions, $\varepsilon_{kk,ab}^{\text{HB}}$, $K_{kk,ab}$, $n_{k,a}$, $n_{k,b}$, $\dots$, $n_{k,N_{\text{ST}},k}$.

Charged groups are subject to additional (Coulombic) interactions. Each charged group is

assigned a central point charge and an additional parameter $\sigma_{kk}^{\text{Born}}$ is used in the Born term (A^{Born}) in order to allow for a more-accurate solvation free energy of the ions. The relative static permittivity model for the dielectric medium presented by Schreckenberget al. [151] is used to replace the explicit representation of the polarity of the solvent. Careful implementation of these ionic contributions is necessary, as slight modifications led to changes in the MIAC and osmotic coefficient calculations in our previous work [42, 43].

In Table 3.1 the SAFT- γ Mie group matrix relevant to the systems of interest here including the APIs, solvent, and buffers is presented. Like and unlike group interactions are obtained from previous work where applicable [35–37, 40, 42, 43, 150, 166, 167]. In this chapter we expand the matrix of groups by estimating unlike interaction parameters between the ionised amine group NH_3^+ and water, as discussed in the following section.

Combining rules are sometimes used to estimate unlike group interactions, assuming that the like group parameters are known, as outlined in Section 2.2.1.7 (cf. Equations (2.35)–(2.39) for the combining rules used in this chapter). Additionally, for the case of interactions involving charged groups, the dispersion energy ε_{kl} between like or unlike ions is calculated using combining rules based on the work of Hudson and McCoubrey [165] (cf. Equation (2.40)). The polarisabilities and ionisation potentials for the ions used in the current model can be found in Table 3.2. The molecular charged groups COO^- and NH_3^+ are found in weak acids and bases we consider. Following the methodology of Di Lecce et al. [42, 43], unlike dispersion energy parameters for these groups with uncharged, non-aromatic groups are given the same values as the analogue unlike interactions with the uncharged groups taking COOH in place of COO^- , and NH_2 in place of NH_3^+ .

The COO^- , H_3O^+ , OH^- , and NH_3^+ ionic groups also require association interaction parameters (see Tables 3.1 and 3.3). For these groups, we follow a similar methodology to that adopted by Eriksen et al. [150], in which the association parameters of the ions are related to their closest uncharged group: H_2O for OH^- and H_3O^+ , COOH for COO^- , and NH_2 for NH_3^+ . We demonstrate this by considering the example of association between the e_1 site on the aCH group and sites of type H on H_3O^+ . The bonding volume, $K_{e_1-\text{H}, \text{aCH}-\text{H}_3\text{O}^+}$ is scaled by the

[illegible]
$$\frac{K_{\text{e1-H, aCH-H}_3\text{O}^+}}{(\sigma_{\text{aCH-H}_3\text{O}^+})^3} = \frac{K_{\text{e1-H, aCH-H}_2\text{O}}}{(\sigma_{\text{aCH-H}_2\text{O}})^3}, \quad (3.28)$$
$$\frac{\varepsilon_{\text{e1-H, aCH-H}_3\text{O}^+}^{\text{HB}}}{K_{\text{e1-H, aCH-H}_3\text{O}^+}} = \frac{\varepsilon_{\text{e1-H, aCH-H}_2\text{O}}^{\text{HB}}}{K_{\text{e1-H, aCH-H}_2\text{O}}} . \quad (3.29)$$

Table 3.2: Values for the polarisabilities $\alpha_{0,k}$, and ionisation potentials I_k of the charged species k , taken from the references shown in the last row.

k	H_3O^+	Na^+	OH^-	Cl^-	COO^-	NH_3^+
$\alpha_{0,k}/(10^{-24} \text{ cm}^3)$	0.93300	0.17900	2.5200	3.6600	4.2042	1.3890
I_k/eV	14.230	47.286	1.8300	3.6127	3.4700	10.160
ref	[211, 212]	[209]	[211, 212]	[209]	[213, 214]	[215–218]

Two exceptions to this methodology are however made. Firstly, in the case of the interaction between e_1 sites of COO^- and sites of type H of the H_3O^+ group, the unlike association energy and bonding volume are scaled using the $e_1 - \text{H}$ interaction of the COO^- and H_2O groups, respectively, from the work of Di Lecce et al. [42, 43]. In the second exception, the association energy and bonding volume between sites of type H on NH_3^+ and e_1 sites on OH^- , are scaled using the association parameters between the H sites on NH_3^+ and the e_1 sites on H_2O obtained in the current chapter.

When the use of these combining and scaling rules do not lead to an accurate representation of the thermodynamic properties (or phase behaviour) of the systems of interest, the group parameters are refined further using experimental data as part of a parameter-estimation strategy, as discussed in the following section. The complete set of SAFT- γ Mie group parameters used in this chapter can be found in Tables 3.3, 3.4 and 3.5.

Table 3.3: SAFT- γ Mie like group parameters: ν_k^* is the number of segments in group k ; S_k and σ_{kk} are the shape factor and diameter of the segments in the group, respectively; $\sigma_{kk}^{\text{Born}}$ the Born diameter used for charged groups; λ_{kk}^r and λ_{kk}^a are the repulsive and attractive exponents of the Mie potential, respectively; ε_{kk} is the dispersion energy between two k groups; $N_{\text{ST},k}$ is the number of association site types on group k , with $n_{\text{H},k}$, $n_{\text{e}_1,k}$ and $n_{\text{e}_2,k}$ representing the number of association sites of type H, e₁ and e₂, respectively. † indicates parameters obtained in the current work.

Group k	ν_k^*	S_k	$\sigma_{kk}/\text{\AA}$	$\sigma_{kk}^{\text{Born}}/\text{\AA}$	λ_{kk}^r	λ_{kk}^a	$(\varepsilon_{kk}/k_{\text{B}})/\text{K}$	$N_{\text{ST},k}$	$n_{\text{H},k}$	$n_{\text{e}_1,k}$	$n_{\text{e}_2,k}$	ref
H ₂ O	1	1.0000	3.0063	-	17.020	6.0000	266.68	2	2	2	-	[37, 38]
CH ₃	1	0.57260	4.0772	-	15.050	6.0000	256.77	-	-	-	-	[35]
CH ₂	1	0.22930	4.8801	-	19.871	6.0000	473.39	-	-	-	-	[35]
aCH	1	0.32180	4.0578	-	14.756	6.0000	371.53	1	-	1	-	[36]
aCCH	1	0.20650	4.3128	-	8.0000	6.0000	61.325	1	-	1	-	[36]
aCCH ₂	1	0.20860	5.2648	-	8.5433	6.0000	591.56	1	-	1	-	[36]
CH	1	0.072100	5.2950	-	8.0000	6.0000	95.621	-	-	-	-	[36]
COOH	1	0.55590	4.3331	-	8.0000	6.0000	405.78	3	1	2	2	[36]
aCCOaC	3	0.18090	4.0670	-	9.8317	6.0000	656.71	-	-	-	-	[39]
C	1	0.040700	5.6571	-	8.0000	6.0000	50.020	-	-	-	-	[36]
NH ₂	1	0.79680	3.2477	-	10.254	6.0000	284.78	2	2	1	-	[167]
COO ⁻	1	0.55590	4.3331	4.6364	8.0000	6.0000	21.264	1	-	4	-	[42, 43]
NH ₃ ⁺	1	0.79680	3.2477	3.4750	10.254	6.0000	48.295	1	3	-	-	†
H ₃ O ⁺	1	1.0000	3.0063	3.0063	17.020	6.0000	68.190	1	3	-	-	[150]
OH ⁻	1	1.0000	2.4600	3.0063	17.020	6.0000	170.24	1	-	3	-	[150]
Na ⁺	1	1.0000	2.3200	3.3600	12.000	6.0000	31.711	-	-	-	-	[150]
K ⁺	1	1.0000	3.0400	4.3440	12.000	6.0000	90.097	-	-	-	-	[150]
Li ⁺	1	1.0000	1.8000	2.6320	12.000	6.0000	6.1039	-	-	-	-	[150]
Rb ⁺	1	1.0000	3.3200	4.6220	12.000	6.0000	130.35	-	-	-	-	[150]
Ca ²⁺	1	1.0000	2.2800	3.7240	12.000	6.0000	261.28	-	-	-	-	[150]
Mg ²⁺	1	1.0000	1.7200	2.9100	12.000	6.0000	89.264	-	-	-	-	[150]
Cl ⁻	1	1.0000	3.3400	3.8740	12.000	6.0000	113.77	-	-	-	-	[150]
Br ⁻	1	1.0000	3.6400	4.1740	12.000	6.0000	107.38	-	-	-	-	[150]

Table 3.4: Group dispersion interaction energies ε_{kl} and repulsive exponent λ_{kl}^r for use with the SAFT- γ Mie Equation of state. The unlike segment diameters σ_{kl} are calculated using the combining rule in Equation (2.35). CR indicates that combining rules are used to obtain ε_{kl} and/or λ_{kl}^r using Equations (2.36) and/or (2.37), respectively. † indicates parameters obtained in the current work.

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref	Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref
H ₂ O	H ₂ O	266.68	17.020	[37, 38]	CH	CH	95.621	8.0000	[36]
H ₂ O	CH ₃	358.18	100.00	[39]	CH	COOH	504.99	CR	[36]
H ₂ O	CH ₂	423.63	100.00	[39]	CH	C	2.0000	CR	[40]
H ₂ O	aCH	357.78	38.640	[36]	CH	COO ⁻	504.99	CR	†
H ₂ O	aCCH	314.03	CR	[40]	CH	NH ₃ ⁺	151.01	CR	†
H ₂ O	aCCH ₂	329.03	CR	[40]	COOH	COOH	405.78	8.0000	[36]
H ₂ O	CH	275.75	CR	[40]	COOH	NH ₂	CR	CR	†
H ₂ O	COOH	289.76	CR	[39]	COOH	COO ⁻	405.78	8.0000	[42, 43]
H ₂ O	C	420.82	CR	[40]	COOH	OH ⁻	CR	CR	†
H ₂ O	NH ₂	358.55	CR	[167]	COOH	H ₃ O ⁺	CR	CR	†
H ₂ O	COO ⁻	171.61	CR	[42, 43]	aCCOaC	aCCOaC	656.71	9.8317	[39]
H ₂ O	NH ₃ ⁺	450.21	CR	†	C	C	50.020	8.0000	[36]
H ₂ O	H ₃ O ⁺	391.04	CR	[150]	C	NH ₃ ⁺	106.49	CR	†
H ₂ O	OH ⁻	134.41	CR	[150]	NH ₂	NH ₂	284.78	10.254	[167]
H ₂ O	Na ⁺	539.68	CR	[150]	NH ₂	NH ₃ ⁺	284.78	CR	†
H ₂ O	K ⁺	376.25	CR	[150]	NH ₂	H ₃ O ⁺	CR	CR	†
H ₂ O	Li ⁺	1023.1	CR	[150]	NH ₂	OH ⁻	CR	CR	†
H ₂ O	Rb ⁺	354.23	CR	[150]	COO ⁻	COO ⁻	21.264	8.0000	[42, 43]
H ₂ O	Ca ²⁺	1460.8	CR	[150]	COO ⁻	NH ₃ ⁺	26.327	CR	†
H ₂ O	Mg ²⁺	2235.1	CR	[150]	COO ⁻	H ₃ O ⁺	27.740	CR	†
H ₂ O	Cl ⁻	95.406	CR	[150]	COO ⁻	OH ⁻	44.520	CR	†
H ₂ O	Br ⁻	112.01	CR	[150]	COO ⁻	Na ⁺	9.9125	CR	[42, 43]
CH ₃	CH ₃	256.77	15.050	[35]	COO ⁻	K ⁺	23.999	CR	[42, 43]
CH ₃	CH ₂	350.77	CR	[35]	COO ⁻	Li ⁺	2.6855	CR	†
CH ₃	aCH	305.81	CR	[36]	COO ⁻	Rb ⁺	31.864	CR	†

Table 3.4: Continued from previous page

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref	Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref
CH ₃	aCCH	455.85	CR	[36]	COO ⁻	Ca ²⁺	27.115	CR	†
CH ₃	aCCH ₂	396.91	CR	[36]	COO ⁻	Mg ²⁺	9.4414	CR	†
CH ₃	CH	387.48	CR	[36]	COO ⁻	Cl ⁻	21.265	CR	[42, 43]
CH ₃	COOH	255.99	CR	[36]	NH ₃ ⁺	NH ₃ ⁺	48.295	10.254	†
CH ₃	C	339.91	CR	[36]	NH ₃ ⁺	H ₃ O ⁺	56.951	CR	†
CH ₃	NH ₂	244.15	CR	[167]	NH ₃ ⁺	OH ⁻	62.231	CR	†
CH ₃	COO ⁻	255.99	CR	[42, 43]	NH ₃ ⁺	Cl ⁻	65.249	CR	†
CH ₃	NH ₃ ⁺	244.15	CR	†	NH ₃ ⁺	Br ⁻	61.726	CR	†
CH ₂	CH ₂	473.39	19.871	[35]	H ₃ O ⁺	H ₃ O ⁺	68.190	17.020	[150]
CH ₂	aCH	415.64	CR	[36]	H ₃ O ⁺	OH ⁻	66.439	CR	[150]
CH ₂	aCCH	345.80	CR	[36]	H ₃ O ⁺	Na ⁺	37.479	CR	†
CH ₂	aCCH ₂	454.16	CR	[36]	H ₃ O ⁺	K ⁺	72.869	CR	†
CH ₂	CH	506.21	CR	[36]	H ₃ O ⁺	Li ⁺	12.314	CR	†
CH ₂	COOH	413.74	CR	[36]	H ₃ O ⁺	Rb ⁺	89.271	CR	†
CH ₂	C	300.07	CR	[36]	H ₃ O ⁺	Ca ²⁺	104.69	CR	†
CH ₂	NH ₂	348.39	CR	[167]	H ₃ O ⁺	Mg ²⁺	44.541	CR	†
CH ₂	COO ⁻	413.74	CR	[42, 43]	H ₃ O ⁺	Cl ⁻	70.552	CR	[150]
CH ₂	NH ₃ ⁺	348.39	CR	†	OH ⁻	OH ⁻	170.24	17.020	[150]
aCH	aCH	371.53	14.756	[36]	OH ⁻	Na ⁺	27.898	CR	[150]
aCH	aCCH	429.16	CR	[36]	OH ⁻	K ⁺	54.733	CR	[150]
aCH	aCCH ₂	416.69	CR	[36]	OH ⁻	Li ⁺	9.1482	CR	[150]
aCH	CH	441.43	CR	[36]	OH ⁻	Rb ⁺	67.944	CR	[150]
aCH	COOH	331.61	9.0687	[36]	OH ⁻	Ca ²⁺	77.245	CR	[150]
aCH	H ₃ O ⁺	CR	CR	†	OH ⁻	Mg ²⁺	33.269	CR	[150]
aCCH	aCCH	61.325	8.0000	[36]	OH ⁻	Cl ⁻	123.21	CR	†
aCCH	aCCH ₂	462.04	CR	[36]	Na ⁺	Na ⁺	31.711	12.000	[150]
aCCH	CH	67.510	CR	[36]	K ⁺	K ⁺	90.097	12.000	[150]
aCCH	COOH	599.28	CR	[36]	Li ⁺	Li ⁺	6.1039	12.000	[150]
aCCH	H ₃ O ⁺	CR	CR	†	Rb ⁺	Rb ⁺	130.35	12.000	[150]

Table 3.4: Continued from previous page

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref	Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref
aCCH ₂	aCCH ₂	591.56	8.5433	[36]	Ca ²⁺	Ca ²⁺	261.28	12.000	[150]
aCCH ₂	CH	65.410	CR	[36]	Mg ²⁺	Mg ²⁺	89.264	12.000	[150]
aCCH ₂	COOH	473.66	CR	[36]	Cl ⁻	Cl ⁻	113.77	12.000	[150]
aCCH ₂	H ₃ O ⁺	CR	CR	†	Br ⁻	Br ⁻	107.38	12.000	[150]

3.3.4 NH₃⁺–NH₃⁺ and NH₃⁺–H₂O interactions

A SAFT- γ Mie model for the NH₃⁺ group is developed here based on the uncharged amine group (NH₂) characterised in previous work [167]. Assuming that the overall shape of the NH₂ group does not change considerably on gaining a proton, the number of spherical segments ν_k^* , shape factor S_k , segment diameter σ_{kk} , attractive and repulsive exponents of the Mie potential λ_{kk}^a and λ_{kk}^r , respectively, of the NH₂ group are used to model NH₃⁺, with the addition of a positive point-charge to the group ($z_k = +1$). In order to calculate the Born diameter $\sigma_{\text{NH}_3^+ - \text{NH}_3^+}^{\text{Born}}$, we follow the proposition of Rashin and Honig [219] of a 7% increase in the bare diameter to correct for the non-sphericity of the ion: $\sigma_{\text{NH}_3^+ - \text{NH}_3^+}^{\text{Born}} = 1.07\sigma_{\text{NH}_3^+ - \text{NH}_3^+}$. The like dispersion energy between NH₃⁺ groups $\varepsilon_{\text{NH}_3^+ - \text{NH}_3^+}$ is calculated using Equation (2.37) with the polarisability α_{0,NH_3^+} [215] and ionisation potential $I_{\text{NH}_3^+}$ [216–218] reported in Table 3.2. Based on quantum-mechanical calculations by Aue et al. [220] we assign three H proton-donor sites to the NH₃⁺ group corresponding to each of the hydrogen atoms, but no e-type acceptor site is included. Thus, self association between NH₃⁺ ions is not considered.

In order to characterise the unlike dispersion and association interaction between NH₃⁺ and H₂O groups, a parameter-estimation procedure is implemented, using experimental data of systems containing the groups of interest. More details on the parameter-estimation strategy used in the SAFT- γ Mie GC approach can be found elsewhere [36, 39, 150]. Here, the unlike parameters $\varepsilon_{\text{NH}_3^+ - \text{H}_2\text{O}}$, $\varepsilon_{\text{NH}_3^+ - \text{H}_2\text{O}, \text{H-e}_1}^{\text{HB}}$ and $K_{\text{NH}_3^+ - \text{H}_2\text{O}, \text{H-e}_1}^{\text{HB}}$ are estimated using experimental osmotic-coefficient data of dilute solutions of alkylammonium bromide salts at 298.15 K for $m_{\text{R-NH}_3^+ \text{Br}^-} \leq 0.65 \text{ mol kg}^{-1}$ where $\text{R} = \text{CH}_3 - (\text{CH}_2)_n$ with $3 \leq n \leq 6$, and of n -octylammonium

Table 3.5: Group association energies $\varepsilon_{kl,ab}^{\text{HB}}$ and bonding volume $K_{kl,ab}$ for use with the SAFT- γ Mie EoS. † indicates parameters obtained in the current work.

Group k	Site a on group k	Group l	Site b on group l	$(\varepsilon_{kl}^{\text{HB}}/k_{\text{B}})/\text{K}$	$K_{kl,ab}/\text{\AA}^3$	ref
H ₂ O	H	H ₂ O	e ₁	1985.4	101.69	[37, 38]
H ₂ O	H	aCH	e ₁	563.56	339.61	[36]
H ₂ O	H	aCCH	e ₁	563.56	339.61	[40]
H ₂ O	H	aCCH ₂	e ₁	563.56	339.61	[40]
H ₂ O	H	COOH	e ₁	1451.8	280.89	[39]
H ₂ O	H	COOH	e ₂	1252.6	150.98	[39]
H ₂ O	e ₁	COOH	H	2567.7	270.09	[39]
H ₂ O	e ₁	NH ₂	H	1988.3	55.824	[167]
H ₂ O	H	NH ₂	e ₁	1460.0	179.60	[167]
H ₂ O	H	COO ⁻	e ₁	802.21	52.555	[42, 43]
H ₂ O	e ₁	NH ₃ ⁺	H	2016.6	49.397	†
H ₂ O	e ₁	H ₃ O ⁺	H	1985.4	101.69	[150]
H ₂ O	H	OH ⁻	e ₁	1492.0	76.411	[150]
aCH	e ₁	H ₃ O ⁺	H	563.56	339.61	†
aCCH	e ₁	H ₃ O ⁺	H	563.56	339.61	†
aCCH ₂	e ₁	H ₃ O ⁺	H	563.56	339.61	†
COOH	H	COOH	H	6427.9	0.80620	[36]
COOH	H	NH ₂	e ₁	2623.5	20.765	†
COOH	e ₁	H ₃ O ⁺	H	1451.8	280.89	†
COOH	e ₂	H ₃ O ⁺	H	1252.6	150.98	†
COOH	H	OH ⁻	e ₁	2036.0	214.16	†
NH ₂	e ₁	NH ₂	H	1070.8	95.225	[167]
NH ₂	e ₁	H ₃ O ⁺	H	1460.0	179.60	†
NH ₂	H	OH ⁻	e ₁	1511.4	42.436	†
COO ⁻	e ₁	H ₃ O ⁺	H	802.21	52.555	†
NH ₃ ⁺	H	OH ⁻	e ₁	1533.0	37.550	†
H ₃ O ⁺	H	OH ⁻	e ₁	1492.0	76.411	[150]

bromide salts at 278.15 K, 288.15 K, and 298.15 K for $m_{\text{C}_8\text{H}_{17}\text{NH}_3^+\text{Br}^-} \leq 0.23 \text{ mol kg}^{-1}$. These salts are modelled as strong electrolytes fully dissociated in water.

The SAFT- γ Mie calculations [150] using the newly developed NH_3^+ group and its optimised interaction parameters with H_2O are compared to the experimental osmotic coefficient data considered in the estimation procedure in Figure 3.3. The model developed leads to an excellent description of the osmotic coefficient of the *n*-alkylammonium salts (Figure 3.3 (a)), but it is interesting to note that the experimental data [221] feature little change in the osmotic coefficient with increasing chain lengths of the alkylamines, while noticeable variation is seen in the description obtained using the SAFT- γ Mie approach. In our model each amine involves a different number of CH_2 groups (three for the *n*-butylamine salt, four for the *n*-pentylamine one, and so forth), such that the differences obtained in the calculation are not unexpected, although they seem to be more pronounced than suggested by the data. In the case of the solution of *n*-octylammonium salt (Figure 3.3 (b)) good overall agreement is observed, for the range of concentrations available, especially at 298.15 K.

In order to assess the robustness of the model it is also useful to consider thermodynamic properties not included in the parameter estimation. The liquid-phase density and MIAC of the alkylammonium bromide salts are compared to experimental data [221–223] in Figure 3.4. As can be seen in the figure, a good prediction of both properties is obtained with the model. The densities of the solutions slightly decrease with increasing chain length, a trend seen both in the experimental data, and in the predictions. We note that the small difference between the calculated densities and the data stems from our model of pure water ($m = 0$). In terms of the MIAC, it is reassuring to note the good agreement obtained with the predictions for the three temperatures assessed.

We consider also the accuracy of the theoretical description provided by the SAFT- γ Mie EoS by comparing the calculations (calc) to the experimental data (exp) in terms of the percentage average absolute deviation (%AAD) of a property X (cf. Equation (2.42)). A summary of the %AAD values for the properties used in the estimation and validation is provided in Table 3.6. We note the particularly low %AAD values for the MIAC and liquid-density calcu-

lations, even though these properties were not included in the parameter estimation.

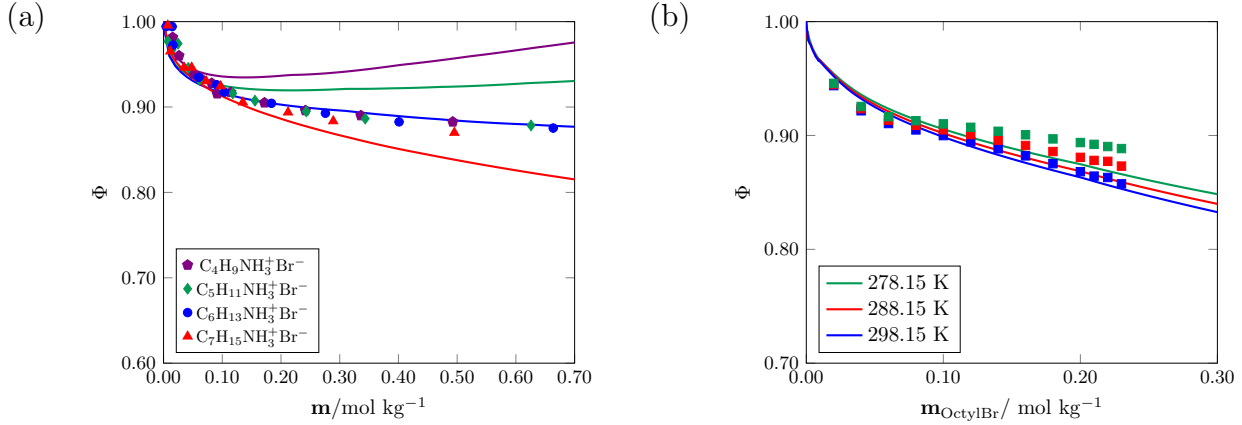


Figure 3.3: Concentration dependence of the osmotic coefficient Φ of: (a) aqueous alkylammonium bromide salts of different chain lengths at $T = 298.15$ K and $P = 0.101$ MPa; and (b) aqueous n -octylammonium bromide salts at $T = 278.15$ K, 288.15 K, and 298.15 K, and $P = 0.101$ MPa. The continuous curves correspond to SAFT- γ Mie calculations (carried out as outlined in ref. [150], using $\nu_+ = 1$ and $\nu_- = 1$), and the symbols to the experimental data [221] used for parameter estimation.

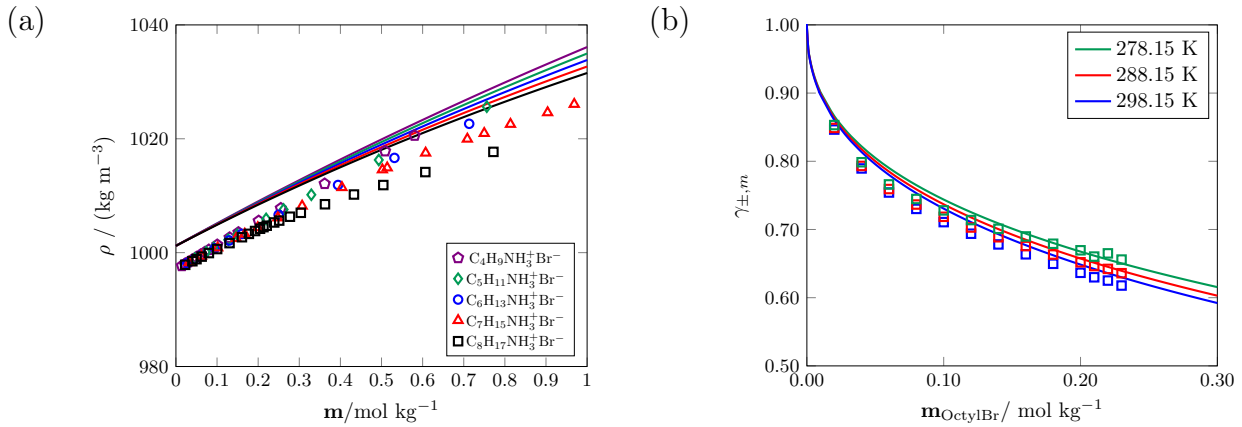


Figure 3.4: Concentration dependence of the: (a) liquid density ρ of aqueous alkylammonium bromide salts of different chain lengths at $T = 298.15$ K and $P = 0.101$ MPa; and (b) mean ionic activity coefficient $\gamma_{\pm,m}$ of aqueous n -octylammonium bromide salts at $T = 278.15$ K, 288.15 K, and 298.15 K and $P = 0.101$ MPa. The continuous curves correspond to SAFT- γ Mie predictions and the symbols to the experimental data [221–223] used for model validation (not considered in the parameter estimation).

Table 3.6: The percentage average absolute deviation % AAD between calculated and experimental values of the osmotic coefficient Φ , liquid density ρ and mean ionic activity coefficient $\gamma_{\pm,m}$ of aqueous alkylammonium bromide salts. Calculations are performed using the SAFT- γ Mie group-contribution approach; experimental data considered in the parameter estimation and model validation are taken from the references indicated.

Parameter estimation				
Salt	T / K	N_R	%AAD Φ	ref
<i>n</i> -Butylammonium bromide	298.15	12	3.00	[221]
<i>n</i> -Pentylammonium bromide	298.15	10	5.05	[221]
<i>n</i> -Hexylammonium bromide	298.15	10	0.92	[221]
<i>n</i> -Heptylammonium bromide	298.15	10	1.30	[221]
<i>n</i> -Octylammonium bromide	278.15–298.15	39	1.04	[221]
Model validation				
Salt	T / K	N_R	%AAD ρ	ref
<i>n</i> -Butylammonium bromide	298.15	17	0.36	[222]
<i>n</i> -Pentylammonium bromide	298.15	10	0.35	[223]
<i>n</i> -Hexylammonium bromide	298.15	10	0.38	[222]
<i>n</i> -Heptylammonium bromide	298.15	20	0.47	[223]
<i>n</i> -Octylammonium bromide	298.15	23	0.50	[222]
Salt	T / K	N_R	%AAD $\gamma_{\pm,m}$	ref
<i>n</i> -Octylammonium bromide	278.15–298.15	39	1.82	[221]

3.3.5 Relative permittivity model

In the implementation of the SAFT- γ Mie theory to charged systems, the relative permittivity, D , of the solution is implemented using the model of Schreckenberget al. [151], in which

$$D = 1 + \rho_{\text{solv}} d, \quad (3.30)$$

where $\rho_{\text{solv}} = N_{\text{solv}}/V$ is the number density of the solvent and d is obtained from:

$$d = d_V \left(\frac{d_T}{T} - 1 \right), \quad (3.31)$$

where d_V and d_T are component-specific parameters in units of $\text{dm}^3 \text{mol}^{-1}$ and K, respectively. The model allows one to implicitly account for changes in the dielectric constant brought about by changes in concentration, although only with qualitative accuracy. As an example, in

comparison to experimental data of aqueous solutions of 0.1 M NaCl, this model would underestimate the reduction in the D by 1.4%. Here, in order to incorporate the dielectric constant of the aqueous solution, the parameters used are $d_V = 0.3777 \text{ dm}^3 \text{ mol}^{-1}$ and $d_T = 1403.0 \text{ K}$ [151], which are appropriate for water.

3.4 Results and discussion

3.4.1 Phase diagrams

We start by carrying out calculations of the solid–liquid solubility of the two APIs of interest. We validate the SAFT- γ -Mie predictions for the SLE of ibuprofen and ketoprofen comparing the calculations with experimental data. Both the SLE of the neutral APIs (the intrinsic solubility) and that of the ionised forms are presented in Figure 3.5 and, as can be seen in the figure, the predicted solubilities are found to be in good overall agreement with the experimental data. In the case of the solubility of the neutral APIs, we follow the procedure of Hutacharoen et al., using the same groups as presented previously, confirming the previous results [39]. The values of Δh^{fus} , T^{fus} and Δc_p for ibuprofen and ketoprofen are taken from experimental data [224–226]; they are provided in Table 3.7. In the case of the calculations involving the ionised forms of the APIs the equilibrium constants needed ($\text{p}K_a$) are taken from the literature; these are also reported in Table 3.7. The importance of accurate values of these properties has been highlighted in the work of Chua et al. [227].

For ibuprofen, a discrepancy in the experimental data sets available in the literature is apparent (Figure 3.5 (a)). From our predictions, one can infer that three of the data sets presented (refs [229–231]) correspond to the solubility of the molecular form of ibuprofen, while the other two (refs [232, 233]) are likely to correspond to solubilities of ionised ibuprofen. Garzon et al. [232], for example, used UV/Vis spectroscopy (a technique capable of detecting charged species) to determine the solubility. In the case of ketoprofen, good agreement between our SAFT- γ Mie predictions and the two experimental data sets [225, 234] can be seen in Figure

Table 3.7: Properties of ibuprofen, ketoprofen, and sodium ibuprofen at 298.15 K and 310.15 K and 1 atm used for modelling the phase behaviour and pH-solubility profiles, including: molar melting enthalpies Δh^{fus} ; melting temperatures T^{fus} ; heat capacities Δc_p ; dissociation constants $\text{p}K_{\text{a}}$; and the solubility product $K_{\text{sp}, \text{A}\nu_{\text{A}} \text{B}\nu_{\text{B}}}$. * indicates properties that are calculated in the current chapter. Values of the dissociation constant $\text{p}K_{\text{w}}$ of water are added for completeness.

Property	Ibuprofen	Ketoprofen	Sodium Ibuprofen	Ref.
$\Delta h^{\text{fus}} / (\text{J mol}^{-1})$	25500	28226	-	[224, 225]
T_{m} / K	347.15	367.65	-	[224, 225]
$\Delta c_p(T) / (\text{J K}^{-1})$	$57.19 - 0.136T$	0.115	-	[224, 226]
$\text{p}K_{\text{a}}$ (298.15 K)	5.17	4.30	-	[184]
$\text{p}K_{\text{a}}$ (310.15 K)	5.16	4.34	-	*
$K_{\text{sp}, \text{A}\nu_{\text{A}} \text{B}\nu_{\text{B}}} (298.15 \text{ K}) / (\text{mol}^2 \text{L}^{-2})$	-	-	0.904	[183]
$K_{\text{sp}, \text{A}\nu_{\text{A}} \text{B}\nu_{\text{B}}} (310.15 \text{ K}) / (\text{mol}^2 \text{L}^{-2})$	-	-	0.932	*
$\text{p}K_{\text{w}}$ (298.15 K)	-	-	-	14.00 [228]
$\text{p}K_{\text{w}}$ (310.15 K)	-	-	-	13.60 [228]

3.5 (b). The good agreement confirms the validity of our SAFT- γ Mie model for the prediction of the solid–liquid equilibria of the APIs.

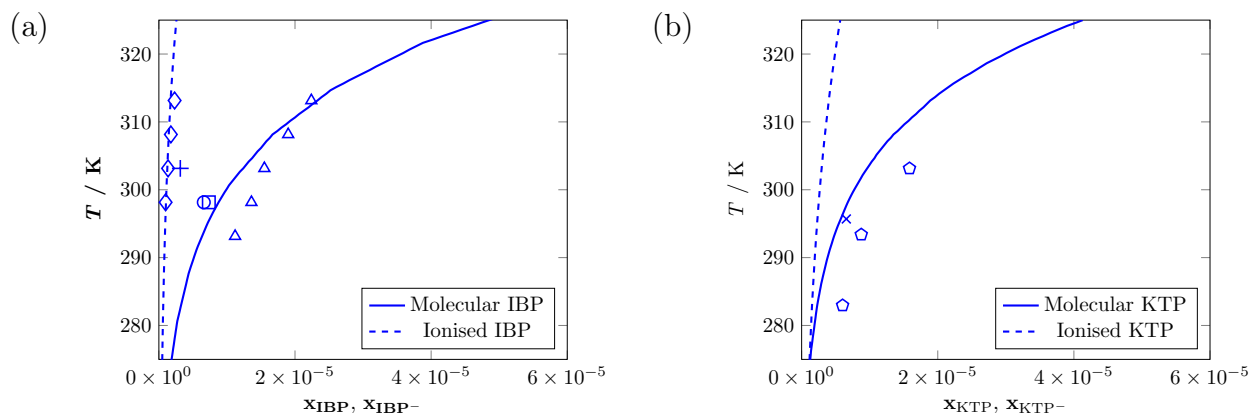


Figure 3.5: Isobaric solid–liquid equilibria (at $P = 0.101 \text{ MPa}$) of the molecular and ionised forms of (a) ibuprofen and (b) ketoprofen in water as a function of temperature. The symbols represent experimental data (triangles [229], circle [230], square [231], diamonds [232], plus [233], pentagons [225], and cross [234]), and the continuous and dashed curves represent predictions using the SAFT- γ Mie approach for the molecular and ionised forms, respectively.

We now turn our attention to the calculation of the complete isobaric temperature–composition phase diagrams for each of the API + water mixtures, including LLE and VLE, in addition to SLE. According to surveys in 2010 [6] and 2020 [7] (involving 28 and 37 companies, 14% and

8% of which were from the pharmaceutical industry, respectively), there is a critical need for predictive models that can simultaneously describe different types of phase equilibria with the same model and same parameter set, with a greater interest in accessing accurate LLE and SLE calculations. The prediction of phase equilibria over broad ranges of conditions can prove challenging for some GC approaches. In this context, the SAFT- γ Mie EoS has been shown to be well-suited to deliver accurate predictions for a wide range of types of phase equilibria, thermodynamic properties, and systems [36, 39, 148, 235].

In Figure 3.6 we present predictive calculations of the constant-pressure phase behaviour for (neutral) ibuprofen and ketoprofen in water. As can be seen in the figure, regions of VLE, LLE, and SLE are predicted for the mixtures. In more detail, in Figures 3.6 (a) and (b), at the lowest temperatures, a region of solid–solid phase separation ($S_{\text{IBP}} + S_{\text{w}}$) is indicated. It is important to note that the phase boundaries of this region are not calculated, as in our model (by construction), the solid phases are treated as pure in the corresponding component. A region of solid–liquid equilibrium has a lower bound in temperature corresponding to a triple-point temperature at 273.15 K (only slightly below the triple point of water), at which a eutectic point is calculated for a mole fraction of ibuprofen $x_{\text{IBP}} = 1.50 \times 10^{-6}$ (this extremely small mole fraction is not visible in the figures). As the temperature is increased above this eutectic point, two SLE regions are present: a large region to the right of the eutectic composition, clearly visible in Figures 3.6 (a) and (b) (labelled $S_{\text{IBP}} + L_1$), in which (pure) solid ibuprofen is predicted to be in equilibrium with a water-rich solution saturated in ibuprofen (i.e., with concentration corresponding to the SLE intrinsic solubility of ibuprofen); and a second SLE region to the left of the eutectic composition corresponding to the equilibrium of pure water crystals with liquid solutions containing very small amounts of ibuprofen ($x_{\text{IBP}} < 1.50 \times 10^{-6}$), not visible in the figure.

A second three-phase point is predicted at 335 K, corresponding to the phase equilibrium of a pure ibuprofen solid phase, in coexistence with a liquid solution rich in ibuprofen (L_2 in the diagram) and a liquid solution rich in water. It is especially interesting to note the presence of an LLE region ($L_1 + L_2$ in the figures), which corresponds to the so-called oiling-out of pharmaceuticals. The predicted temperature for this equilibrium (above 335 K) is in good

agreement with the LLE temperature range proposed by Codan et al. [236]. At temperatures above 373.63 K, a VLE region is calculated; we note, however, that ibuprofen is reported to begin decomposing at a temperature of 453 K [237].

The predicted phase behaviour of the water-rich side of the phase diagram is better seen in Figure 3.6 (b), where the various types of phase equilibria predicted to take place at very low concentrations of ibuprofen ($x_{\text{IBP}} < 2 \times 10^{-4}$) can be seen. These regions are difficult to confirm experimentally, emphasizing the importance of predictive tools such as the SAFT- γ Mie GC approach used here.

The phase diagram for the ketoprofen + water mixture (Figures 3.6 (c) and (d)) is qualitatively similar to that of ibuprofen, as can be expected given the similar chemical structure of the two APIs, and their similar melting points (347.15 K and 367.59 K for ibuprofen and ketoprofen, respectively). We predict the LLE temperature range for ketoprofen to be between 348.60 K and 373.63 K. The decomposition temperature of this API is reported as 508.15 K [238].

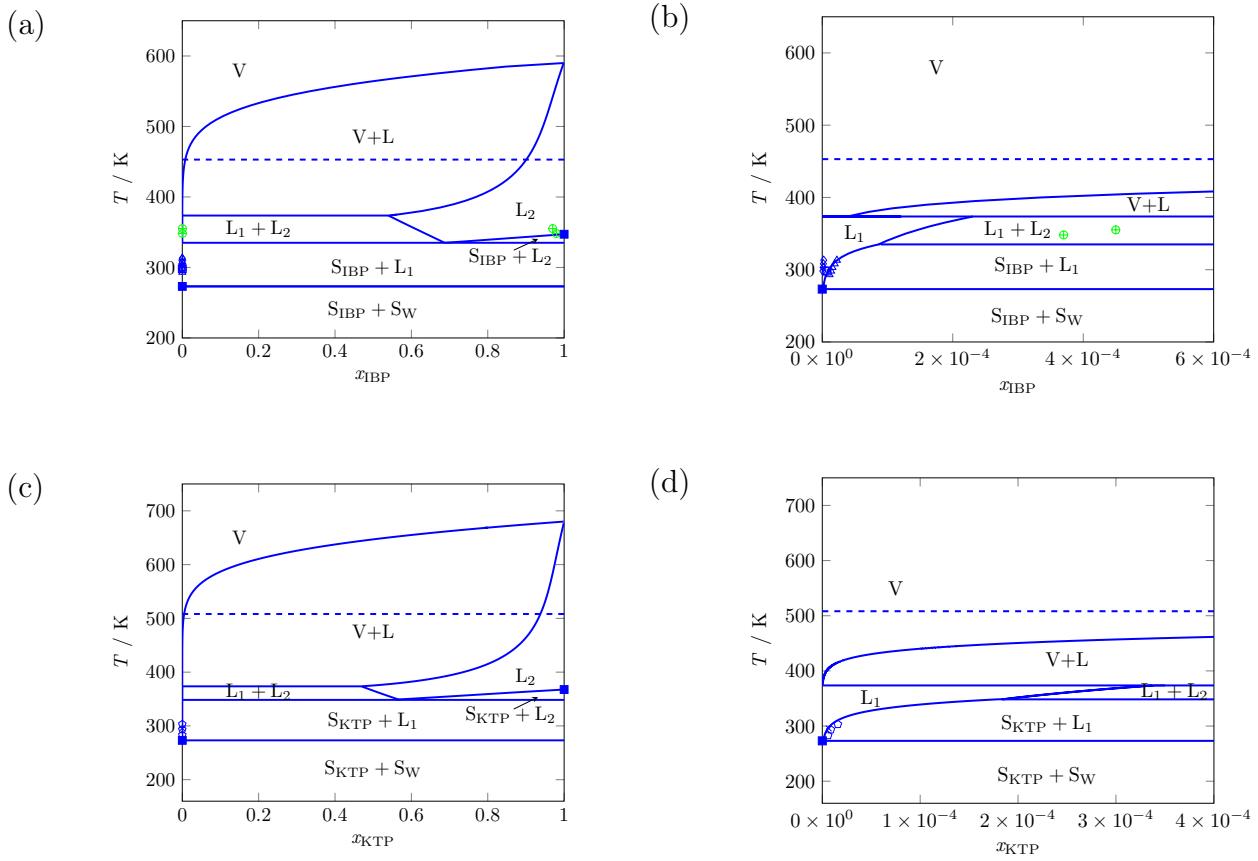


Figure 3.6: Isobaric ($P = 0.101$ MPa) temperature–composition (T – x) phase diagrams of the vapor–liquid equilibrium (VLE), liquid–liquid equilibrium (LLE), and solid–liquid equilibrium (SLE) regions of ibuprofen (IBP) + water ((a) and (b)), ketoprofen (KTP) + water ((c) and (d)) mixtures. The water-rich regions of the phase diagrams are highlighted in (b) and (d). The symbols represent the experimental data and the continuous curves the predictions obtained with the SAFT- γ Mie approach. The blue symbols correspond to SLE data for the APIs (triangles [229], circles [230], diamonds [232], plus [233], empty squares [231], pentagons [225], cross [234] and filled squares [224, 225, 239]), and the green symbols to the LLE data for ibuprofen + water [236]. The dashed lines indicate the temperatures at which ibuprofen and ketoprofen are reported to decompose [237, 238].

3.4.2 pH-Solubility profiles

Predicting the pH-dependent solubility of ibuprofen and ketoprofen using a GC approach poses a difficult challenge as, in addition to predicting the fluid-phase behaviour of the neutral APIs in solution, the speciation (ionisation) of the molecules and the presence of relevant ionic species in solution need to be taken into account. We assess now the prediction of the pH-solubility profile of the APIs at 298.15 K using the SAFT- γ Mie EoS parameters provided in Tables 3.3, 3.4. and 3.5. In order to model the complete profiles, including the salt precipitation at high values of pH where applicable, knowledge of the pK_a of the acid, $K_{sp,A\nu_A B\nu_B}$ of the salt, and pK_w of water, to be used in Equations (3.11), (3.22), and (3.19), respectively, is also required. The values of these constants are listed in Table 3.7.

In Figure 3.7 (a) the predicted pH-solubility profile of ibuprofen at 298.15 is compared to available experimental data [183]. HCl is used to lower the pH of the solution below the pK_a of ibuprofen to a pH of 1.0, while NaOH is used to buffer the solution to a pH of 11.6 and is also the source of Na^+ counter ions in the solution, leading to the formation of the sodium salt of ibuprofen at high pH [183, 184]. As is apparent from the figure, our SAFT- γ Mie predictions are in very good agreement with experimental data. We are able to capture the intrinsic solubility plateau at acidic pHs, the increase in solubility associated with adding Na^+ counter ions, and the salt solubility plateau at high pH. The pH at which solid ibuprofen and its sodium salt can coexist, pH^{max} , obtained from our SAFT- γ Mie prediction is 7.75, which compares well with the measured value of 8.55 [184]. We note that only one stable polymorph (form γ [240]) has been identified for sodium ibuprofen and that although the dihydrate form is reported by some authors [241, 242], the experimental data used in our work do not correspond to the dihydrate salt, therefore, only the anhydrous form of the salt is considered in our models.

In order to quantitatively assess the accuracy of the theoretical description provided by the SAFT- γ Mie GC approach compared to experimental data for the pH-dependant solubility of ibuprofen (Figure 3.7 (a)), the absolute average deviation (AAD) (cf. Equation (2.43)) is calculated to be 0.05 mol L^{-1} , which further validates the excellent predictive capability of our

model.

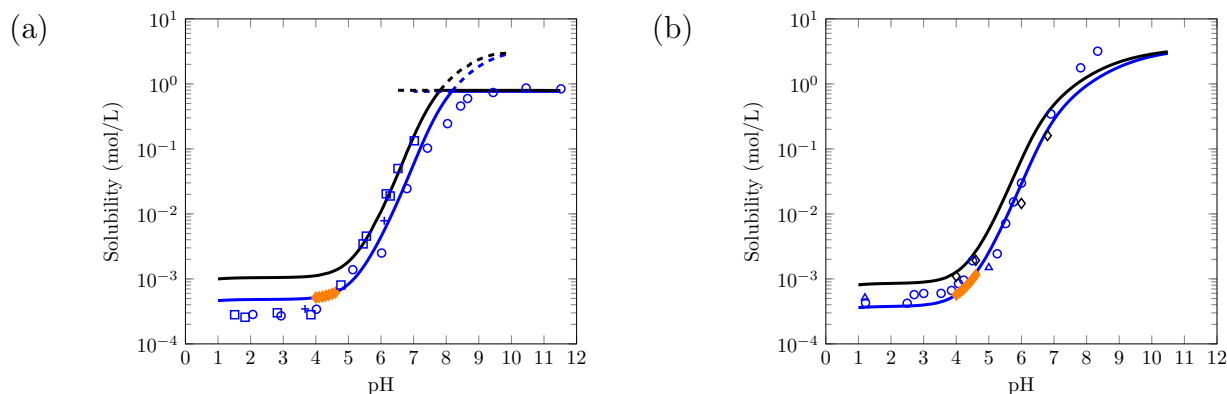


Figure 3.7: SAFT- γ Mie predictions of the pH-solubility profiles compared to experimental data of aqueous solutions of (a) ibuprofen and (b) ketoprofen at ambient pressure ($P = 0.101$ MPa) and given temperature. The open blue and black symbols represent experimental data and the continuous curves represent the predictions (blue curves correspond to 298.15 K and black curves to 310.15 K). In (a) the data for HCl and NaOH buffered solutions are shown as circles [184], squares [243], and crosses [244]. In (b) the data for HCl and NaOH buffered solutions are shown as circles [183], triangles [245], and empty diamonds [246]. The empty black diamonds correspond to experimental measurements at 310.15 K, all other experimental data were determined at 298.15 K. Predictions obtained using acetate as buffer are shown as filled orange diamonds. For the SAFT- γ Mie predictions, the pH is calculated using Equation (3.26), while the pH of the experimental data is taken as reported in the corresponding publications.

The SAFT- γ Mie prediction of the pH-solubility profile of ketoprofen at the same temperature of 298.15 K is compared to experimental data in Figure 3.7 (b). HCl and NaOH are again used to buffer the solution in the acidic and basic pH range, respectively. Na^+ counter ions are used, as sodium has been shown to lead to increases in the ionisation and solubility of the API [184]. The sodium salt of ketoprofen is reported to form only via a process of evaporation and is not expected to result from an aqueous solution of API and NaOH [184, 247]; as a consequence, this salt is not included in our model. As can be seen in the figure, our predictions are in very good agreement with the experimental data [183, 245], corresponding to an AAD of 0.17 mol L^{-1} .

A commonly used buffer to regulate the pH of solutions between 4.0 and 4.6 is prepared through the addition of acetic acid and sodium acetate [206, 207]. We use this buffer to predict the solubility of ibuprofen and ketoprofen in an aqueous solution over this pH range (and $T = 298.15$ K) using the SAFT- γ Mie GC approach. These predictions are also presented in Figures

3.7 (a) and (b). The solubilities of the APIs in the acetate buffer solution are found to be the same as in the buffer solutions of HCl, suggesting that the choice of buffer in this pH range has no effect on the solubility. This interesting result requires experimental confirmation.

As an added point here, we note how in modelling pH-solubility profiles it is crucial to use thermodynamic-based equilibrium constants pK_a , pK_w , and $K_{sp,A\nu_A B\nu_B}$ rather than concentration-based analogues, as the two definitions are not equivalent for concentrated solutions. In Figure 3.8 we depict the change in the thermodynamic-based equilibrium constant (calculated using the ratio of activities as given in Equation (3.9)) and in the concentration-based analogue (calculated using the ratio of concentrations) as a function of pH for the ibuprofen and ketoprofen solutions using HCl and NaOH as the buffering agents. As can be seen in the figures, similar values of pK_a are calculated at pH values close to the pK_a of each of the APIs, respectively. The difference between the thermodynamic- and concentration-based pK_a values becomes apparent as the buffering species are added to the solutions, resulting in more concentrated solutions. The ratio of activity coefficients in Equation (3.11) can no longer be assumed to be 1.0 for these solutions.

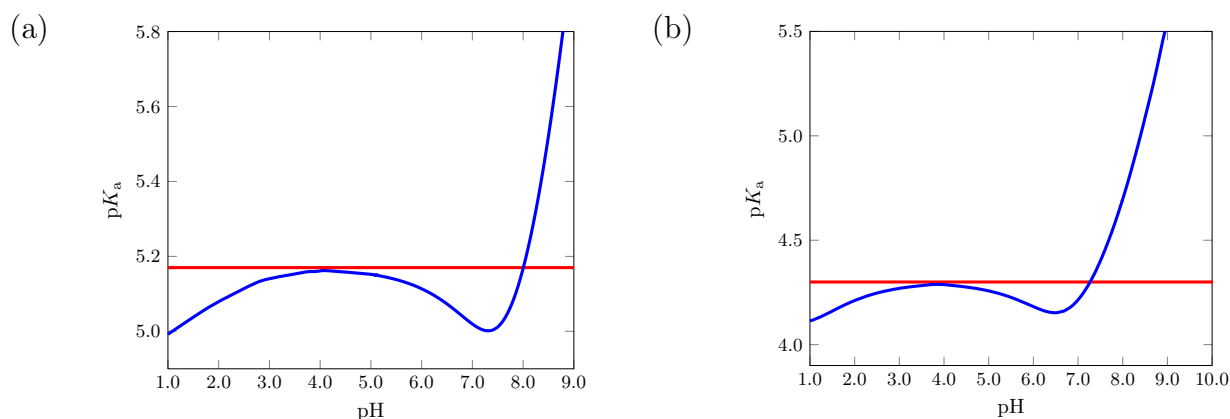


Figure 3.8: The calculated thermodynamic- (red) and concentration-based (blue) pK_a of (a) ibuprofen and (b) ketoprofen in aqueous solutions at 298.15 K, as a function of pH. HCl and NaOH are used as pH-changing agents. The pH is calculated using Equation (3.26).

3.4.3 Temperature effect on the pH-solubility profile

The pH solubility profiles of ibuprofen and ketoprofen at a higher temperature of 310.15 K (37°C) are also considered. The relevant dissociation constants at this temperature are reported in Table 3.7. The temperature dependence of pK_a is calculated using the well-known van't Hoff equation [182],

$$\left(\frac{\partial pK_a}{\partial T}\right)_P = \frac{-\Delta h^0(T)}{2.303 RT^2}, \quad (3.32)$$

in which the conversion factor of 2.303 is used, to convert the natural logarithm ($\ln K_a$) to a base-10 logarithm ($\log_{10} K_a$); the molar standard state enthalpy at given T can be calculated through integration of the isobaric heat capacity c_p . In the case of ibuprofen, Xu et al. [248] used a polynomial to represent c_p measurements obtained using an adiabatic calorimeter, as a function of scaled temperature, $\tilde{T}$, to obtain a correlation given as

$$\begin{aligned} c_p / (\text{JK}^{-1}\text{mol}^{-1}) = & 144.27 + 77.046 (\tilde{T}/\text{K}) + 3.5171 (\tilde{T}/\text{K})^2 + 10.925 (\tilde{T}/\text{K})^3 \\ & + 11.224 (\tilde{T}/\text{K})^4, \end{aligned} \quad (3.33)$$

where $\tilde{T}/\text{K} = (T/\text{K} - 206.201)/127.096$. This correlation, which is valid over the temperature range $79.105 \leq T/\text{K} \leq 333.397$, is used in our current work.

In the case of ketoprofen, a linear expression for pK_a as a function of T based on titration measurements has been reported by Avdeef [182] and by Avdeef and Tsinam [249] as

$$(\partial pK_a / \partial T)_P = 0.003/\text{K}. \quad (3.34)$$

This expression is applicable over the temperature range $298.15 \leq T/\text{K} \leq 310.15$.

The van't Hoff equation is also used in the calculation of the temperature dependence of the solubility product $K_{\text{sp}, A\nu_A B\nu_B}(T)$ of sodium ibuprofen. Manin and Perlovich [250] have used differential scanning calorimetry to report a correlation for c_p of sodium ibuprofen given as

$$c_p / \text{JK}^{-1}\text{mol}^{-1} = 315.12 - 0.522 (T/\text{K}) + 1.863 \times 10^{-3} (T/\text{K})^2. \quad (3.35)$$

Their correlation is valid in the temperature range $295.11 \leq T/\text{K} \leq 431.11$ and gives

$$(\partial \text{p}K_{\text{sp}, \text{A}_{\nu\text{A}} \text{B}_{\nu\text{B}}} / \partial T)_P = 0.0307/\text{K} \quad (3.36)$$

corresponding to $K_{\text{sp}, \text{A}_{\nu\text{A}} \text{B}_{\nu\text{B}}} = 0.932 \text{ mol}^2 \text{ L}^{-2}$ at 310.15 K.

In Figure 3.7 (a), the solubility of ibuprofen is found to be higher at 310.15 K than at 298.15 K, as expected. At higher temperature, the lower $\text{p}K_{\text{a}}$ leads to more ionisation of ibuprofen and to a higher concentration of ibuprofen anions in solution, increasing its solubility. Interestingly, the increase in solubility is most significant at the lowest pH. It is also evident from the pH-solubility profile that the solubility of the salt form ($\text{IBP}^- \text{Na}^+$) is almost unaffected by the increase in temperature. A small change in the value of pH^{max} is, however, predicted (from 7.75 at 298.15 K to 7.37 at 310.15 K). To the best of our knowledge, no experimental data are available for the pH-dependent solubility of ibuprofen at 310.15 K.

For ketoprofen, very good agreement between the SAFT- γ Mie predictions and the experimental data of Sheng et al. [246] at 310.15 K can be seen in Figure 3.7 (b) (with a corresponding $\text{AAD} = 0.06 \text{ mol L}^{-1}$). The solubility of ketoprofen is higher at 310.15 K compared to 298.15 K for the entire pH range. As in the case of ibuprofen, the solubility of ketoprofen increases most at the lowest pH. No predictions are made for the salt solubility of this API as the $\text{KTP}^- \text{Na}^+$ salt does not form in aqueous solution (as mentioned previously).

3.4.4 Inorganic counter-ion effect on the pH-solubility profile

The choice of counter ion plays a key role in determining the increase in solubility of the free form of the API and in the formation of its salt [181]. Analysis of the trends of counter ions published by the US Drug and Food Administration (FDA) in 2007 [251] indicates that inorganic counter ions are most-commonly used for acidic APIs, with the highest recurrence for sodium counter ions (75.3%). In this context, it is also of interest to predict the variations in solubility when other counter ions are considered as buffers. We take advantage of the predictive nature of our SAFT- γ Mie GC approach to predict the aqueous pH-solubility profile

of ibuprofen and ketoprofen at 298.15 K using various inorganic counter ions, and HCl, to buffer the solutions. Our predictions are shown in Figures 3.9 (a) and (b) for ibuprofen and ketoprofen, respectively. The strong bases: KOH, LiOH, RbOH, $\text{Ca}(\text{OH})_2$, and $\text{Mg}(\text{OH})_2$ are used to buffer the solution in the basic pH-range and as a source of K^+ , Li^+ , Rb^+ , Ca^{2+} , and Mg^{2+} counter ions, respectively (NaOH is also presented for comparison).

As can be seen in the figures, differences in solubility for each API are seen on addition of different counter ions only at higher pHs (6.8 and 6.2 for ibuprofen and ketoprofen, respectively). All counter ions lead to an increase in the solubility of the APIs, as expected, with monovalent counter ions (Na^+ , K^+ , Li^+ , Rb^+) leading to larger enhancement of the solubility compared to divalent cations. The order of decreasing solubility of ibuprofen and ketoprofen based on the counter ion used is found to be $\text{Rb}^+ > \text{K}^+ > \text{Na}^+ > \text{Li}^+ > \text{Ca}^{2+} > \text{Mg}^{2+}$, which interestingly follows the decreasing order of the base strength of the sources of these counter ions: $\text{p}K_{\text{a,RbOH}} > \text{p}K_{\text{a,KOH}} > \text{p}K_{\text{a,NaOH}} > \text{p}K_{\text{a,LiOH}} > \text{p}K_{\text{a,Ca}(\text{OH})_2} > \text{p}K_{\text{a,Mg}(\text{OH})_2}$. It should be noted here that no information on the strength of these bases is included in the model, as they are assumed to be fully dissociated (cf. Reaction (3.15)), such that it is especially reassuring that a physically meaningful trend is seen in our predictions.

No predictions are made as to the solubility of the salts of ibuprofen and ketoprofen formed with the different counter ions considered, as no data are available in the literature for the solubility or solubility product of these salts in aqueous solution. However, according to Stahl and Wermuth [181], it is expected that potassium salts will have a higher solubility than the corresponding sodium salts in aqueous solutions while calcium salts are expected to exhibit lower aqueous solubility compared to other salts.

3.4.5 Organic counter-ion effect on the pH-solubility profile

There is increased interest in the pharmaceutical sector in the use of more-diverse counter ions as more-challenging molecules are selected for development [251], provided that there is at least 3 units difference in the $\text{p}K_{\text{a}}$ of the API and its counter ion to form a stable salt

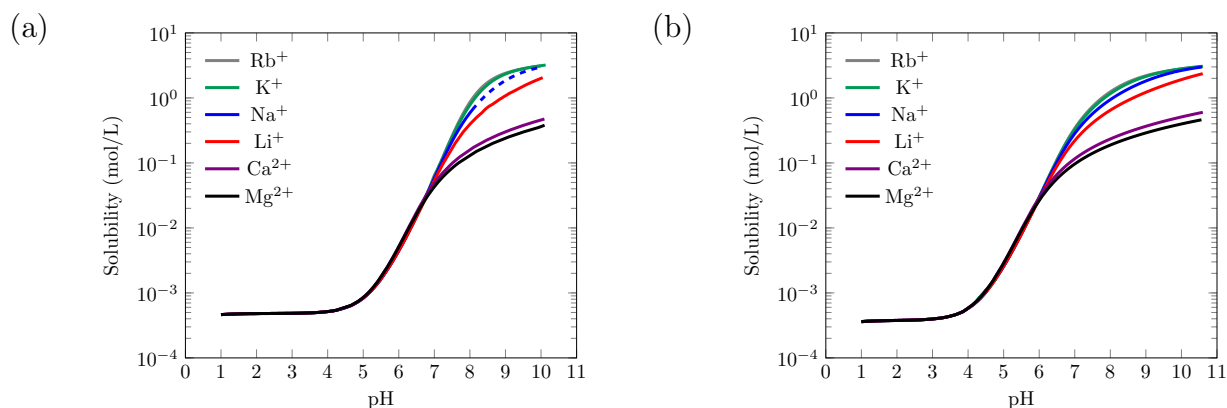


Figure 3.9: SAFT- γ Mie group-contribution predictions of the pH-solubility profiles of aqueous solutions of (a) ibuprofen and (b) ketoprofen at 298.15 K and 0.101 MPa using HCl and the strong bases NaOH, KOH, LiOH, RbOH, $\text{Ca}(\text{OH})_2$, and $\text{Mg}(\text{OH})_2$ to buffer the API solutions and as sources of the: Na^+ , K^+ , Li^+ , Rb^+ , Ca^{2+} , and Mg^{2+} cations, respectively. The colouring of the curves corresponds to different counter ions in solution as specified in the legends. The dashed blue curve in (a) represents the solubility of ibuprofen after the sodium salt of ibuprofen begins to precipitate (cf. Figure 3.7 (a)). The pH is calculated using Equation (3.26).

[252]. Here, we study multiple amines as the source of organic counter ions to develop the pH-solubility profile of ibuprofen, including the corresponding ibuprofen-amine salts. The SAFT- γ Mie predictions of the corresponding pH-solubility profiles using the weak bases *n*-hexylamine, *n*-octylamine, benzylamine, and *tert*-butylamine as buffering agents are shown in Figure 3.10. At $\text{pH} < \text{p}K_{\text{a,IBP}}$, HCl is used to buffer the solution, such that the solubility of ibuprofen in this pH-range is the same as that reported in Figure 3.7 (a), as no counter ions are added at this stage. In Figure 3.10, the solubility of ibuprofen at a higher pH is found to be the same regardless of the amine (counter ion) considered (before salt formation) which reflects the fact that these weak bases have similar values of $\text{p}K_{\text{a}}$ (≈ 10).

Once the solubility product $K_{\text{sp},\text{A}\nu_{\text{A}}\text{B}\nu_{\text{B}}}$ of the salts is reached the ibuprofen amine salts begin to precipitate. The thermodynamic solubility product $K_{\text{sp},\text{A}\nu_{\text{A}}\text{B}\nu_{\text{B}}}$ of each of the salts shown in Figure 3.10 is calculated based on the salt solubility measurements reported by David et al. [244], using Equation (3.22) and the density of the solution calculated by the SAFT- γ Mie EoS; the calculated values are given in Table 3.8. David et al. [244] have suggested the existence of two polymorphic forms for the benzylamine salt of ibuprofen, however, only the solubility of one of these is reported in their paper; this is the form considered in our calculations. Polymorphic

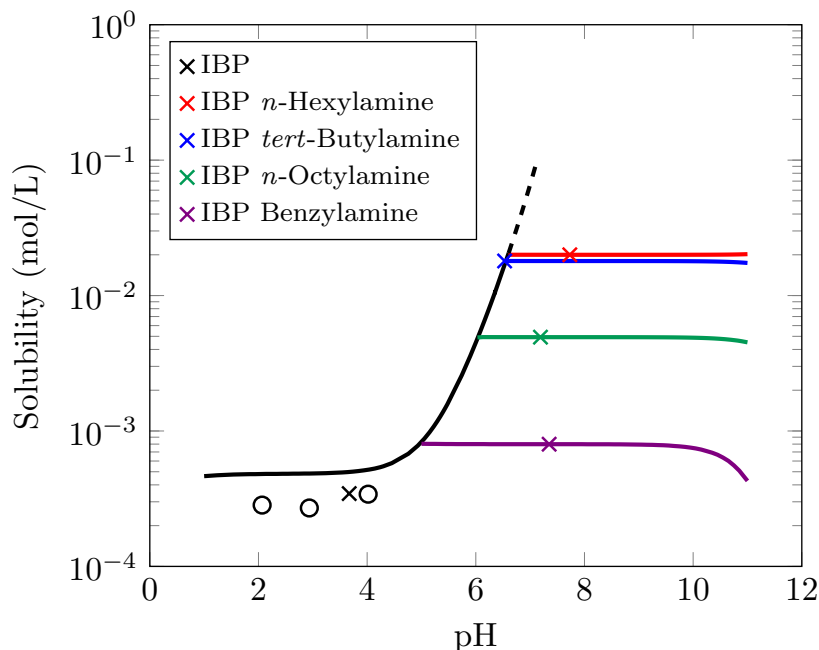


Figure 3.10: SAFT- γ Mie group-contribution predictions of the pH-solubility profile of aqueous solutions of ibuprofen and its amine salts compared to experimental data at 298.15 K and 0.101 MPa. HCl is used as the acidic buffering agent and *n*-hexylamine (red), *n*-octylamine (green), *tert*-butylamine (blue), and benzylamine (violet) as basic buffering agents as well as the source of organic counter ions. The symbols represent the experimental data [183, 244] and the continuous curves the predictions with the SAFT- γ Mie approach. The predicted solubility of the free form of ibuprofen (black) is the same for all of the counter ions used. For the SAFT- γ Mie predictions, the pH is calculated using Equation (3.26), while the pH of the experimental data is taken as reported in the corresponding publications.

forms for the remaining amine salts were not reported and the authors did not consider the presence of hydrated forms of any of the amine salts. From the SAFT- γ Mie calculations of the solubility of ibuprofen salts in Figure 3.10, a decrease in the solubility of the benzylamine salt is seen at $\text{pH} > 9.4$, due to the presence of excess counter ions; this phenomenon is known as the common-ion effect, and is more pronounced for salts with lower values of $K_{\text{sp}, \text{A}\nu_{\text{A}} \text{B}\nu_{\text{B}}}$ values [176]. This effect is not observed for the other salts, with higher $K_{\text{sp}, \text{A}\nu_{\text{A}} \text{B}\nu_{\text{B}}}$.

In the context of drug development, salts are typically used to enhance the solubility of the API [252]; therefore, counter ions that lead to higher salt solubilities are desirable. However, salts with higher solubilities also have a higher tendency to undergo salt disproportionation, which occurs when the salt converts to its free form, leading to a negative impact on the bioavailability of the API [203, 204]. The value of pH^{max} obtained from pH-solubility profiles

Table 3.8: Solubility products $K_{\text{sp},\text{A}_{\nu\text{A}}\text{B}_{\nu\text{B}}}$ and pH^{max} for ibuprofen amine salts obtained from SAFT- γ Mie predictions of the corresponding pH-solubility profiles (cf. Figure 3.10).

Ibuprofen salt	$K_{\text{sp},\text{A}_{\nu\text{A}}\text{B}_{\nu\text{B}}} / (\text{mol}^2 \text{L}^{-2})$	pH^{max}
<i>n</i> -Hexylamine	2.8892×10^{-4}	6.60
<i>n</i> -Octylamine	2.0453×10^{-5}	6.03
<i>tert</i> -Butylamine	2.4784×10^{-4}	6.53
Benzylamine	5.9826×10^{-7}	5.05

can give an indication of the likelihood of a salt to disproportionate [202], such that a salt of an acidic API with a higher pH^{max} is more likely to revert back to its free form than one with a lower pH^{max} . The pH^{max} values for the ibuprofen salts in Figure 3.10, obtained in our work, can be found in Table 3.8. The *n*-hexylamine salt of ibuprofen has the highest pH^{max} value of 6.60 and is most likely to undergo salt disproportionation compared to the other salts considered here.

3.5 Conclusions

Predictive models based on group-contribution (GC) approaches are an attractive prospect in the study of solutions containing active pharmaceutical ingredients (APIs). In particular, the solubility of APIs is of interest, as it can be especially challenging to measure accurately due to the low aqueous solubility of many drugs. In this chapter, we demonstrate the ability of the SAFT- γ Mie GC equation of state to predict the phase behaviour, including solid–liquid solubility and pH-solubility profiles of two acidic APIs: ibuprofen and ketoprofen in aqueous solution. Group interaction parameters presented in previous studies are used here, with the explicit intention to examine the predictive power of the approach.

We first present the intrinsic (neutral) solid–liquid equilibria (SLE) of the compounds, as well as the liquid–liquid (LLE), the so-called oiling-out, and the vapour–liquid equilibrium (VLE) regions of the phase diagrams of the binary mixtures. The various types of phase equilibria that take place across a wide temperature range, including VLE, LLE, and SLE at very low mole fractions of the APIs ($\approx 10^{-6}$) are predicted. Our calculations highlight the difficulty in

carrying out experimental measurements at such low concentrations and perhaps explain the lack of available literature data for many APIs. We obtain very good agreement with available SLE solubility data of ibuprofen and ketoprofen, and for the temperature range over which oiling-out (LLE) is observed in ibuprofen solutions.

By accounting for the speciation of the APIs, the acidic buffers, and the inorganic and organic counter ions in solution, we develop a model to predict the pH-solubility profile of the aqueous solutions of the APIs. These profiles are determined by combining experimental values for the equilibrium constant (pK_a of the acids, K_w of water, and $K_{sp,A\nu_A B\nu_B}$ of the salts) with molecular models using the SAFT- γ Mie GC EoS. Excellent agreement with experimental data for the pH-solubility profile predictions of aqueous solutions of ibuprofen and ketoprofen buffered by HCl and NaOH at 298.15 K, is shown. The SAFT- γ Mie GC approach accurately captures the intrinsic solubility plateau at acidic pH values, the increase in API solubility associated with adding NaOH, and, in the case of ibuprofen, the salt solubility plateau at high pH. We demonstrate that SAFT- γ Mie is a powerful predictive tool, considering that no data related to this model are used to obtain the group interaction parameters. The solubility and pH profiles at 310.15 K, using HCl and NaOH to buffer the solutions, are also predicted. The equilibrium constants, pK_a , $K_{sp,A\nu_A B\nu_B}$, and pK_w , are evaluated at this temperature using the van't Hoff equation. An increase in API solubility is observed at 310.15 K, with the enhancement being more apparent in the acidic pH range. Furthermore, the ability of the SAFT- γ Mie GC approach in representing buffer variation is showcased by using acetate as an acidic buffer in the pH range 4.0–4.6 at 298.15 K. Negligible differences are seen in the predicted solubilities with changes in pH when acetate is used as the buffer instead of HCl.

The choice of counter ion also plays an important role in determining the increase in the solubility of the API as the pH is changed and the solubility of its salt, if it forms. We carry out predictions of the pH-solubility profiles for ibuprofen and ketoprofen to include various counter ions, including several strong bases as a source of the inorganic counter ions (K^+ , Li^+ , Rb^+ , Ca^{2+} , and Mg^{2+}). Solubility predictions for both APIs indicate a trend of decreasing API solubility at $pH > 6.0$ following $Rb^+ > K^+ > Na^+ > Li^+ > Ca^{2+} > Mg^{2+}$, which interestingly follows the decreasing order of the strength of the base corresponding to the source of these

counter ions. These findings can be extended to identify the corresponding solubility of the API salt for each of the counter ions, once experimental salt solubility data become available.

Additionally, considering that the pharmaceutical industry is moving towards the use of more-diverse counter ions, we extend our predictions to treat several amines as the source of organic counter ions in aqueous ibuprofen solutions. In order to do this, we develop the relevant group interactions for a new NH_3^+ group using osmotic-coefficient data for the parameter estimation, and mean-ionic-activity-coefficient and saturated-liquid density data to validate the adequacy of the model. We demonstrate the capability of the SAFT- γ Mie GC approach in handling complex mixtures of molecular and ionic groups, where we use the predicted pH-solubility profiles to determine the values of the thermodynamic $K_{\text{sp}, A_{\nu_A} B_{\nu_B}}$ and pH^{max} for the ibuprofen amine salts.

The findings from our current work demonstrate the applicability of the SAFT- γ Mie GC approach for pharmaceutical systems, particularly for understanding the phase behaviour of APIs and the pH-solubility profiles, including the relevance of this tool for salt selection and disproportionation.

Chapter 4

Predicting the pH-dependent solubility of basic diprotic APIs

4.1 Introduction

In the work presented thus far, the performance of the SAFT- γ Mie GC approach in the prediction of the phase behaviour and pH-solubility profiles of ionisable active pharmaceutical ingredients (APIs) has been discussed. In Chapter 3, we outlined the thermodynamic modelling framework required for developing the pH-solubility profiles of weakly acidic monoprotic APIs, combining experimental values for the equilibrium constant for the dissociation of the APIs and the formation of salts, with molecular models using the SAFT- γ Mie GC EoS. The thermodynamic framework developed in Chapter 3, can also be applied for modelling the pH-solubility profiles of weakly *basic* monoprotic APIs, using the corresponding equilibrium constants for the APIs and their salts. In this chapter, after having successfully modelled the pH-solubility profiles of monoprotic APIs, we extend the application of the SAFT- γ Mie EoS to model the pH-solubility profiles of diprotic APIs.

In general, for ionisable APIs, which are weak electrolytes, salt formation is the preferred method for increasing the solubility of the APIs and improving their bioavailability [179–181]. Diprotic APIs particularly benefit from salt formation as their solubility is expected to increase

by a hundred units for a single unit change in pH, compared to a tenfold increase only in the solubility of monoprotic APIs, for every pH unit change [178, 192, 253]. Diprotic compounds contain two ionisable moieties in their molecules and can, in theory, form disalts, thus showing a significant increase in solubility with changing pH. The complex speciation exhibited by diprotic APIs in solution and the possible precipitation of multiple salt forms, make the determination of the pH-dependent solubility of diprotic APIs more challenging compared to that of monoprotic APIs.

A small number of experimental investigations studying the pH-solubility profiles of diprotic APIs can be found in the literature [192, 253–257]. Lakkaraju et al. [257] investigated the pH-solubility profile of avitriptan, a basic diprotic API, and found that only strongly acidic counter ions (e.g., hydrochloride) led to the formation of the disalt of the API, weaker acidic buffering agents (e.g., acetate) did not protonate the second basic moiety in the API thus forming the monosalts only. The pH-dependent solubility of another dibasic API was studied by Garren and Pyter [253], using the saturation shake-flask (SSF) method. They observed that despite using hydrochloride counter ions (strong acid) to lead to salt formation, the dihydrochloride salt of the diprotic base did not precipitate due to the high solubility product of the disalt. To the best of our knowledge only the work by Venkatesh et al. [192], in which the authors studied the pH-solubility profile of cosalane using the SSF method, explored the pH-solubility profiles of diprotic acids. The solubility of cosalane increased by more than 10^7 times, compared to the intrinsic solubility, following a pH adjustment of less than four pH units; although no salt was found to precipitate in the pH range considered.

In terms of computational approaches used to study the pH-solubility profiles of diprotic APIs, the literature [192, 253, 254, 256] on modelling these profiles is limited to the use of empirical equations, such as the Henderson-Hasselbalch (HH) equations [193], which have been found to only give rough estimates of the profiles, as discussed in Chapter 3. Molecular modelling techniques based on thermodynamic concepts have not been employed to calculate or predict the pH-solubility profiles of diprotic APIs.

In the current chapter, we aim to fill this gap in literature by developing a thermodynamic

framework for modelling the pH-solubility profiles of diprotic APIs using the SAFT- γ Mie GC approach. We demonstrate the applicability of the approach by modelling the pH-solubility profile of the weakly basic diprotic API procaine, a local anesthetic, in aqueous buffered solutions. In the following section we present the thermodynamic framework used. Model development within the SAFT- γ Mie GC approach relevant to aqueous solutions of procaine is provided in Section 4.3, including the estimated parameters for the new SAFT- γ Mie group interactions required for these mixtures. Predictions of the intrinsic solubility of procaine as well as its pH-solubility profile using the developed framework are demonstrated in Section 4.4. Final remarks are given in Section 4.5.

4.2 Thermodynamic modelling of pH-solubility profiles

Pharmaceutical drugs are classified as bases when they contain a nitrogen atom with lone pairs of electrons available to accept a proton [258]. Procaine contains two nitrogen atoms with free lone pairs of electrons: the first nitrogen atom (N1) is present in a tertiary amine in the diethylamino moiety and the second nitrogen atom (N2) is in the amine group attached to the benzene ring (cf. Figure 4.1 (a)). The presence of these two nitrogen atoms allows molecular (neutral) procaine (PC) to behave as a weak diprotic base in solution, i.e., it can accept two protons (cf. Figure 4.1 (c)).

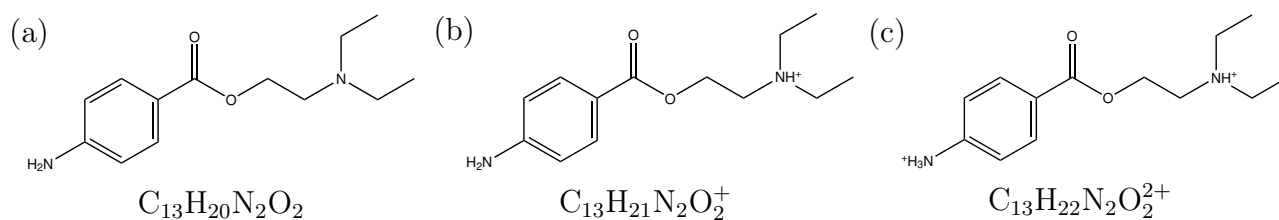
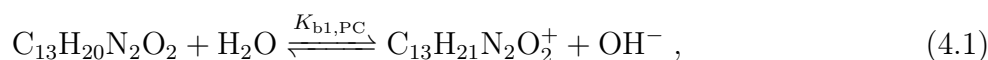
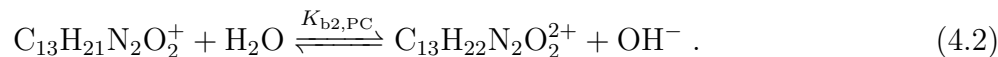


Figure 4.1: The chemical structure of (a) neutral procaine, (b) procaine monocation and (c) procaine dication.

In aqueous solution, procaine dissociates according to the following reactions:



and

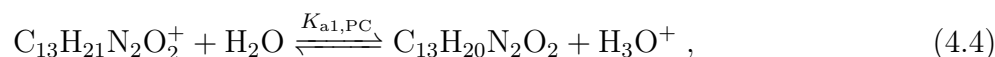


where $K_{\text{b}x,\text{PC}} = 10^{-\text{p}K_{\text{b}x,\text{PC}}}$ for $x = 1, 2$. $\text{p}K_{\text{b}x,\text{PC}}$ can be calculated as

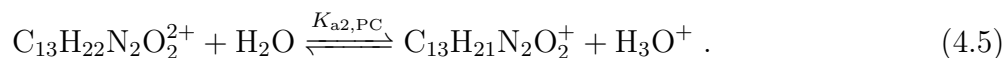
$$\text{p}K_{\text{b}x,\text{PC}} = \text{p}K_{\text{w}} - \text{p}K_{\text{a}x,\text{PC}} , \quad (4.3)$$

where $\text{p}K_{\text{a}x,\text{PC}}$ are the acid dissociation constants of procaine and $\text{p}K_{\text{w}}$ is the ionic product of water (cf. Equations (3.13) and (3.19) respectively). The first ionisation of procaine in Reaction (4.1) occurs when N1 on PC accepts a proton to produce the procaine monocation PC^+ (Figure 4.1 (b)). The lone pairs of electrons on N1 are concentrated on the nitrogen atom and are available to bond with a proton. This is in contrast to N2, where the lone pairs of electrons are delocalized into the benzene ring and are less available to accept a proton. The second ionisation of procaine occurs when N2 on PC^+ bonds to a hydrogen atom to form the dication of procaine PC^{2+} (Figure 4.1 (c)) as given by Reaction (4.2).

In general, it is more common in the literature to refer to the strength of basic APIs in terms of the $\text{p}K_{\text{a}}$ of the conjugate acid of the base. Therefore, we express Reactions (4.1) and (4.2) in terms of the conjugate acids as,



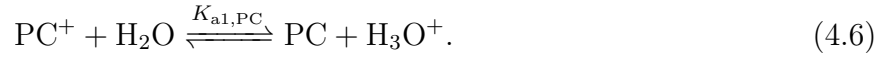
and



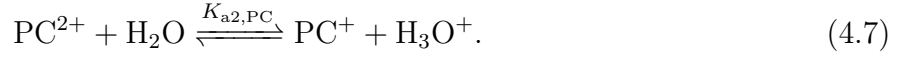
respectively, where $\text{p}K_{\text{a}1,\text{PC}}(298.15 \text{ K}) = 9.0$ and $\text{p}K_{\text{a}2,\text{PC}}(298.15 \text{ K}) = 2.5$ [258].

To develop the thermodynamic framework for modelling the pH-solubility profile of procaine in water, a preliminary step is to identify the phase equilibrium between the solid and liquid phases as well as the speciation occurring in the liquid phase. For simplicity, Reactions (4.4)

and (4.5) will be referred to as,



and



respectively, in this chapter. The equilibrium between molecular procaine in the liquid and solid phases is shown in Figure 4.2 (a), together with the dissociation of the API in aqueous solution to produce PC^+ (Reaction(4.6)), and the consequent dissociation of PC^+ in water to give rise to PC^{2+} (Reaction(4.7)). We also account for the dissociation of water into H_3O^+ and OH^- as given by the ionic product of water K_w . In addition, the strong base NaOH and the strong acid HCl are used to buffer the solution, and are modelled as fully dissociated into their respective ions; thus species Na^+ and Cl^- are also present in the liquid phase.

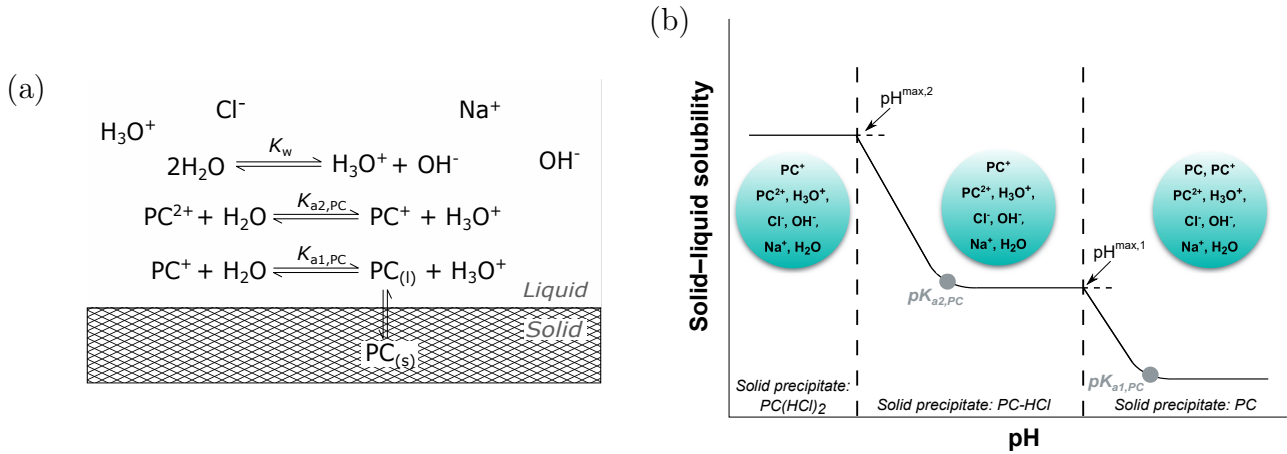


Figure 4.2: Schematic representation of the equilibria and speciation considered for modelling the pH-solubility profile of procaine (PC) and its salt forms: the hydrochloride (PC-HCl) and the dihydrochloride (PC(HCl)₂) salts. (a) The four equilibria taking place before salt formation: the solid-liquid equilibrium of neutral procaine, PC; the dissociation of PC in the liquid phase into its ionised form, PC^+ , regulated by the thermodynamic equilibrium constant $K_{a1,\text{PC}}$; the dissociation of PC^+ in water into the dication PC^{2+} accounted for by $K_{a2,\text{PC}}$; and the dissociation of water given by the ionic product of water K_w . HCl and NaOH are the strong acid and base buffers modelled as fully dissociated into Cl^- and H_3O^+ , and Na^+ and OH^- ions, respectively. The solid phase consists of pure neutral PC. (b) The pH-solubility profile including the solubility of the neutral form of the dibasic API and its salts PC-HCl and PC(HCl)₂. The green spheres show the components present in the liquid phase in the three regions of the plot.

A schematic of the pH-solubility profile of procaine is shown in Figure 4.2 (b). At high

pH, the solubility of procaine corresponds to its intrinsic solubility. As pH decreases below $pK_{a1,PC}$, with the addition of Cl^- counter ions, the solubility increases due to a rise in the concentration of the monoprotonated form of the API (PC^+) in solution. At lower pH values ($pK_{a2,PC} < pH < pK_{a1,PC}$), another plateau in solubility is expected due to the precipitation of the hydrochloride salt of the monoprotonated base ($PC-HCl$) after the first $pH^{\max}$ ($pH^{\max,1}$). The pH of the solution will continue to decrease with the addition of Cl^- counter ions. At pH below $pK_{a2,PC}$, a second sharp rise in solubility occurs due the protonation of the second basic moiety (N2) on procaine resulting in an increase in the concentration of the diprotonated form of the API (PC^{2+}) in solution. The solubility will continue to increase until reaching the second $pH^{\max}$ value ($pH^{\max,2}$), below which, the dihydrochloride salt of procaine ($PC(HCl)_2$) may form.

4.2.1 Intrinsic solubility of procaine

At a given temperature T and pressure P the solid-liquid intrinsic solubility of neutral procaine can be calculated as discussed in Section 3.2.1. A simplified form of Equation (3.2), which is commonly used in literature [25], is employed here:

$$\ln x_{PC}^L(T, P) = \frac{\Delta h_{PC}^{\text{fus}}(T_{PC}^{\text{fus}}, P)}{R} \left(\frac{1}{T_{PC}^{\text{fus}}} - \frac{1}{T} \right) - \ln \gamma_{PC}(T, P, \mathbf{x}^L), \quad (4.8)$$

where the second and third terms in Equation (3.2) are neglected. Due to the absence of experimental data for the difference in the isobaric heat capacity in the liquid and solid phases (Δc_p), it is assumed that the terms involving Δc_p in Equation (3.2) have approximately equal contributions and will cancel out due to their opposite signs.

In Equation (4.8), the melting properties of procaine $\Delta h_{PC}^{\text{fus}}$ and T_{PC}^{fus} are obtained from experimental data [259] as 26.1 kJ mol^{-1} and 334.15 K , respectively. The symmetric activity coefficient of procaine in the liquid phase γ_{PC} is calculated using Equation (3.3).

4.2.2 Solubility of procaine

In solution, procaine dissociates according to Reactions (4.6) and (4.7). The Gibbs free energy change $\Delta_R G$ for the Reaction (4.6) is given by Equation (3.4), by summing over the relevant species in the equilibrium, PC^+ , H_2O , PC , and H_3O^+ , which leads to,

$$\Delta_R G(T, P, \mathbf{x}) = \mu_{\text{PC}} + \mu_{\text{H}_3\text{O}^+} - \mu_{\text{PC}^+} - \mu_{\text{H}_2\text{O}} = 0 . \quad (4.9)$$

Using the expressions for the chemical potential of a component i in Equations (3.6) and (3.7), the thermodynamic equilibrium constant for the first dissociation step of procaine $K_{\text{a1,PC}}$ can be expressed as,

$$K_{\text{a1,PC}} = \exp \left(-\frac{\mu_{\text{PC}}^\circ + \mu_{\text{H}_3\text{O}^+}^\circ - \mu_{\text{PC}^+}^\circ - \mu_{\text{H}_2\text{O}}^\circ}{RT} \right) , \quad (4.10)$$

and hence,

$$K_{\text{a1,PC}} \equiv \frac{a_{\text{PC}} a_{\text{H}_3\text{O}^+}}{a_{\text{PC}^+} a_{\text{H}_2\text{O}}} . \quad (4.11)$$

Here, as for monoprotic acids in Chapter 3, we use the molal thermodynamic equilibrium constant, $K_{\text{a1,PC}}^m$ which is given by,

$$K_{\text{a1,PC}}^m = \left(\frac{m_{\text{PC}} m_{\text{H}_3\text{O}^+}}{m_{\text{PC}^+}} \right) \left(\frac{\tilde{\gamma}_{m,\text{PC}} \tilde{\gamma}_{m,\text{H}_3\text{O}^+}}{\tilde{\gamma}_{m,\text{PC}^+}} \right) ; \quad (4.12)$$

consequently,

$$\text{p}K_{\text{a1,PC}} = -\log_{10}(K_{\text{a1,PC}}^m) . \quad (4.13)$$

For the second dissociation step of procaine given by Reaction (4.7), a similar derivation of the thermodynamic equilibrium constant is employed, resulting in

$$K_{\text{a2,PC}}^m = \left(\frac{m_{\text{PC}^+} m_{\text{H}_3\text{O}^+}}{m_{\text{PC}^{2+}}} \right) \left(\frac{\tilde{\gamma}_{m,\text{PC}^+} \tilde{\gamma}_{m,\text{H}_3\text{O}^+}}{\tilde{\gamma}_{m,\text{PC}^{2+}}} \right) ; \quad (4.14)$$

and,

$$\text{p}K_{\text{a}2,\text{PC}} = -\log_{10}(K_{\text{a}2,\text{PC}}^m) . \quad (4.15)$$

In order to promote the dissociation of procaine, the pH of the solution is decreased by adding an acid as a buffering agent and as a source of counter ions. In this chapter, we use HCl, a strong acid, to buffer the solution. In the model, HCl is assumed to be completely dissociated in solution into its respective ions, H_3O^+ and Cl^- .

Analogously, the pH of the solution can be increased to determine the solubility of procaine in the alkaline pH range by adding a strong or weak base as a buffering agent. For solutions of procaine we use NaOH, a strong base, to increase the pH of the solution, modelled as fully dissociated into Na^+ and OH^- .

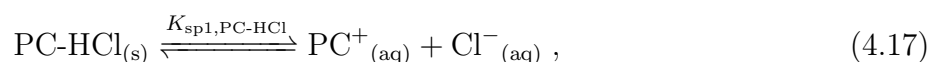
The solubility of procaine (in molarity (moles/volume) units), prior to salt formation, S_{PC} , is given as the sum of the intrinsic solubility and the concentrations of the ionised forms of procaine,

$$S_{\text{PC}} = \rho M_{\text{w}} (m_{\text{PC}} + m_{\text{PC}^+} + m_{\text{PC}^{2+}}) , \quad (4.16)$$

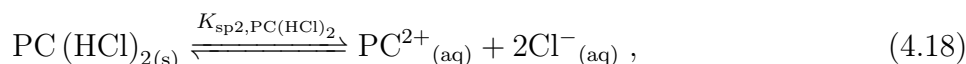
where ρ is the molar density of the solution and M_{w} is the molecular mass of water. We note that, the dissociation of water (Reaction (3.17)) is also accounted for in the model by incorporating the molal ionic product of water K_{w} (cf. Equation (3.18)).

4.2.3 Salt solubility

As a diprotic base, procaine can, in theory, precipitate as the hydrochloride (HCl) salt of the monoprotonated base, which dissociates in solution as,



or as the dihydrochloride salt of the diprotonated base, dissociating in solution as,



when the solubility products $K_{\text{sp1,PC-HCl}}$ or $K_{\text{sp2,PC(HCl)}_2}$, respectively, are reached. The solubility products of procaine salts can be calculated using Equation (3.22).

The corresponding pH at which the hydrochloride and dihydrochloride salts start to precipitate are labelled as $\text{pH}^{\text{max},1}$ and $\text{pH}^{\text{max},2}$, respectively (see Figure 4.2 (b)). Any further addition of counter ions to decrease the pH in the salt-precipitating regions of the pH-solubility profile, may lead to an excess of these ions, and to a drop in the solubility of the salt; characteristic of the common-ion effect [176].

The solubility reported after the onset of salt precipitation, corresponds to the ionised forms of procaine only, as given by,

$$S_{\text{PC}} = \rho M_{\text{w}} (m_{\text{PC}^+} + m_{\text{PC}^{2+}}) . \quad (4.19)$$

4.2.4 pH-Solubility profiles of procaine

In order to complete the pH-solubility profile of procaine, experimental values for $\text{p}K_{\text{a1,PC}}$, and $\text{p}K_{\text{a2,PC}}$ are obtained from ref. [259], with the SAFT- γ Mie EoS used for the calculation of the density of the solution and activity coefficients of the species. Additionally, $K_{\text{sp1,PC-HCl}}$ and $K_{\text{sp2,PC-(HCl)}_2}$ are required to determine the solubility of procaine salts; however, no information on these solubility products are found in literature. Alternatively, we obtain $K_{\text{sp1,PC-HCl}}$ by fitting Equation (4.19) to procaine hydrochloride solubility data reported by Guerrari et al. [260] at 298.15 K. We note that no data on the solubility of the dihydrochloride salt of procaine is found in literature; therefore, it is assumed that this salt does not precipitate. This is a valid assumption, in particular at $\text{pH} > \text{p}K_{\text{a2,PC}}$ where the concentration of PC^{2+} is expected to be very low.

In summary, prior to salt formation (at $\text{pH} > \text{pH}^{\text{max},1}$), the mole fraction of neutral procaine x_{PC} in solution, at a given temperature and pressure, is determined using Equation (4.8). As procaine is an ionisable drug, it will dissociate in aqueous solution according to Reactions (4.6) and (4.7) which are accounted for in the model by the thermodynamic equilibrium constants

in Equations (4.12) and (4.14), respectively. The pH of the solution is changed by adding the buffering agents HCl or NaOH, which are modelled as strong electrolytes. Hence, the species present in solution at $\text{pH} > \text{pH}^{\text{max},1}$ are: H_2O , H_3O^+ , OH^- , PC, PC^+ , PC^{2+} and Na^+ or Cl^- . To determine the concentration of each of the species, the equations for $\text{p}K_{\text{a}1,\text{PC}}$ (Equation (4.13)), $\text{p}K_{\text{a}2,\text{PC}}$ (Equation (4.15)) and K_w (Equation (3.18)) are solved together with Equation (4.8) and the condition of electroneutrality (Equation (3.25)).

As pH changes, $K_{\text{sp}1,\text{PC-HCl}}$ is calculated as (cf. Equation (3.22))

$$K_{\text{sp}1,\text{PC-HCl}} = (m_{\text{PC-HCl}})^2 (\tilde{\gamma}_{m,\text{PC}^+} \tilde{\gamma}_{m,\text{Cl}^-}) , \quad (4.20)$$

to identify when the salt is precipitated. During this process, the solubility of procaine is calculated using Equation (4.16) for pH above the point at which $K_{\text{sp}1,\text{PC-HCl}}^{\text{calc}} = K_{\text{sp}1,\text{PC-HCl}}^{\text{exp}}$ and the solubility of procaine hydrochloride is calculated using Equation (4.19) at pH below this point. To determine the concentration of the species present in solution (H_2O , H_3O^+ , OH^- , PC^+ , PC^{2+} and Na^+ or Cl^-) at $\text{pH} < \text{pH}^{\text{max},1}$, the equation for the solubility product $K_{\text{sp}1,\text{PC-HCl}}$ (Equation (4.20)) replaces Equation (4.8), to be used together with Equations (4.13), (4.15), (3.18) and (3.25).

4.3 Model development

In order to model the intrinsic solubility and pH-solubility profile of procaine in buffered aqueous solutions using the SAFT- γ Mie GC approach, it is necessary to characterize the individual group parameters of all the compounds involved. The groups required for modelling the neutral and ionised forms of procaine are shown in Figure 4.3. These are: ($2 \times \text{CH}_3$, $4 \times \text{CH}_2$, $1 \times \text{N}$, $4 \times \text{aCH}$, $1 \times \text{aCCOO}$ and $1 \times \text{aCNH}_2$) for molecular procaine, replacing the N group in the latter by N^+ for the procaine monocation PC^+ . The dication of procaine PC^{2+} is represented by an additional ionic group aCNH_3^+ (replacing the aCNH_2 group in PC^+), such that the groups used to model the dication are ($2 \times \text{CH}_3$, $4 \times \text{CH}_2$, $1 \times \text{N}^+$, $4 \times \text{aCH}$, $1 \times \text{aCCOO}$ and $1 \times \text{aCNH}_3^+$). The solvent considered for this system is water ($1 \times \text{H}_2\text{O}$) which dissociates to

produce a hydronium ion ($1 \times \text{H}_3\text{O}^+$) and a hydroxide ion ($1 \times \text{OH}^-$).

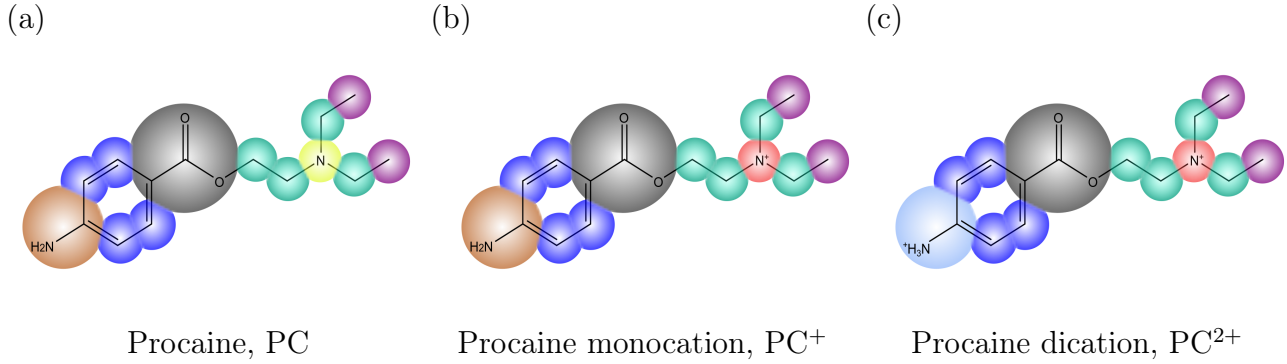


Figure 4.3: Chemical structures of the procaine species studied, together with their corresponding SAFT- γ Mie representation. Each of the coloured spheres corresponds to a different group, and the same group is represented in the same colour in figures of (a) neutral procaine PC, (b) procaine monocation PC⁺ and (c) procaine dication PC²⁺.

The pH of the solutions is varied by adding acidic and basic buffering agents. In particular, HCl (a strong acid), is used to decrease the pH of the solution and as a source of counter ion. It is modelled as fully dissociated into its respective ions: H_3O^+ and Cl^- . Similarly, NaOH (a strong base) is used as a buffering agent to increase the pH of the solution and is also modelled as fully dissociated into OH^- and Na^+ .

In the SAFT- γ Mie approach, as outlined in Chapter 2 (cf. Section 2.2), a group k is characterised by a number ν_k^* of identical spherical segments, a shape factor S_k , the segment energy of interaction ε_{kk} , the segment diameter σ_{kk} , the exponents of the Mie potential λ_{kk}^r and λ_{kk}^a and, where relevant, the parameters describing any site-site association interactions, $\varepsilon_{kk,ab}^{\text{HB}}$, $K_{kk,ab}$, $n_{k,a}$, $n_{k,b}$, $\dots$, $n_{k,N_{\text{ST}},k}$. For charged groups, two additional parameters are used, the charge q_k and the Born cavity diameter $\sigma_{kk}^{\text{Born}}$. The interactions between two groups k and l of different types are represented analogously by the corresponding unlike parameters ε_{kl} , σ_{kl} , λ_{kl}^r , λ_{kl}^a , $\varepsilon_{kl,ab}^{\text{HB}}$ and $K_{kl,ab}$. Additionally, due to the presence of charged groups, the relative permittivity, D , is required to complete the description of the system within the SAFT- γ Mie framework.

In Table 4.1 the SAFT- γ Mie group matrix relevant to the system of interest in this chapter including the API, solvent and buffering agents is presented. Like and unlike group interactions are obtained from previous work where applicable [35–37, 39, 42, 43, 150, 167]. In the current

Table 4.1: Group interactions required to model procaine in aqueous solutions containing HCl and NaOH as buffering agents within the SAFT- γ Mie GC approach. Purple shading indicates parameters obtained in the current chapter, blue shading interactions from previous work [35–37, 39, 42, 43, 150, 167], grey shading interactions calculated using the combining rules of Equations (2.36) and (2.40) for uncharged and charged groups, respectively, and yellow shading interactions developed in Chapter 3.

[illegible]

Table 4.2: SAFT- γ Mie like parameters for the species present in our model. ν_k^* is the number of segments making up group k , S_k is the shape factor, λ_{kk}^r and λ_{kk}^a are the repulsive and attractive exponents of the Mie potential, respectively, σ_{kk} is the segment diameter, $\sigma_{kk}^{\text{Born}}$ is the Born diameter, ε_{kk} is the dispersion energy between two k groups, $N_{\text{ST},k}$ is the number of association site types on group k , with $n_{\text{H},k}$, $n_{\text{e}_1,k}$ and $n_{\text{e}_2,k}$ representing the association sites of type H, e₁ and e₂, respectively. †† indicates parameters obtained in this chapter.

Group k	ν_k^*	S_k	$\sigma_{kk}/\text{\AA}$	$\sigma_{kk}^{\text{Born}}/\text{\AA}$	λ_{kk}^r	λ_{kk}^a	$(\varepsilon_{kk}/k_{\text{B}})/\text{K}$	$N_{\text{ST},k}$	$n_{\text{H},k}$	$n_{\text{e}_1,k}$	$n_{\text{e}_2,k}$	ref
H ₂ O	1	1.0000	3.0063	-	17.020	6.0000	266.68	2	2	2	-	[37, 38]
CH ₃	1	0.5726	4.0772	-	15.050	6.0000	256.77	-	-	-	-	[35]
CH ₂	1	0.2293	4.8801	-	19.871	6.0000	473.39	-	-	-	-	[35]
N	1	0.1507	3.0755	-	8.8971	6.0000	62.971	1	-	1	-	[167]
aCH	1	0.3218	4.0578	-	14.756	6.0000	371.53	1	-	1	-	[36]
aCCOO	2	0.6556	3.3851	-	36.616	6.0000	200.00	2	-	2	2	††
aCNH ₂	1	0.7152	3.7408	-	17.065	6.0000	670.91	2	2	1	-	††
N ⁺	1	0.1507	3.0755	3.2908	8.8971	6.0000	62.971	1	1	-	-	[42, 43]
aCNH ₃ ⁺	1	0.7152	3.7408	4.0027	17.065	6.0000	397.63	1	3	-	-	††
H ₃ O ⁺	1	1.0000	3.0063	3.0063	17.020	6.0000	68.181	1	3	-	-	[150]
Cl ⁻	1	1.0000	3.3400	3.8740	12.000	6.0000	113.76	-	-	-	-	[150]
Na ⁺	1	1.0000	2.3200	3.3600	12.000	6.0000	31.709	-	-	-	-	[150]
OH ⁻	1	1.0000	2.4600	3.0063	17.020	6.0000	170.16	1	-	3	-	[150]

Table 4.3: Group dispersion interaction energies ε_{kl} and repulsive exponent λ_{kl}^r for use with the SAFT- γ Mie equation of state, taken from the references indicated in the last column. The unlike segment diameter σ_{kl} are taken from the combining rule in Equation (2.35). CR indicates that combining rules are used to obtain ε_{kl} and/or λ_{kl}^r using equations (2.36) and/or (2.37), respectively. $\dagger\dagger$ indicates parameters obtained in this chapter.

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref	Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref
H ₂ O	H ₂ O	266.68	17.020	[37, 38]	aCH	aCH	371.53	14.756	[36]
H ₂ O	CH ₃	358.18	100.00	[39]	aCH	aCCOO	468.32	CR	$\dagger\dagger$
H ₂ O	CH ₂	423.63	100.00	[39]	aCH	aCNH ₂	564.39	23.276	$\dagger\dagger$
H ₂ O	N	1481.3	21.217	[167]	aCH	N ⁺	CR	CR	-
H ₂ O	aCH	357.78	38.640	[36]	aCH	aCNH ₃ ⁺	CR	CR	-
H ₂ O	aCCOO	339.28	CR	$\dagger\dagger$	aCH	H ₃ O ⁺	CR	CR	[44]
H ₂ O	aCNH ₂	438.284	CR	$\dagger\dagger$	aCH	Cl ⁻	CR	CR	-
H ₂ O	N ⁺	1481.3	21.217	[42, 43]	aCH	Na ⁺	CR	CR	-
H ₂ O	aCNH ₃ ⁺	1382.5	CR	$\dagger\dagger$	aCH	OH ⁻	CR	CR	-
H ₂ O	H ₃ O ⁺	391.04	CR	[150]	aCCOO	aCCOO	200.00	36.616	$\dagger\dagger$
H ₂ O	Cl ⁻	95.406	CR	[150]	aCCOO	aCNH ₂	494.91	CR	$\dagger\dagger$
H ₂ O	Na ⁺	539.68	CR	[150]	aCCOO	N ⁺	CR	CR	-
H ₂ O	OH ⁻	134.41	CR	[150]	aCCOO	aCNH ₃ ⁺	CR	CR	$\dagger\dagger$
CH ₃	CH ₃	256.77	15.050	[35]	aCCOO	H ₃ O ⁺	CR	CR	$\dagger\dagger$
CH ₃	CH ₂	350.77	CR	[35]	aCCOO	Cl ⁻	CR	CR	-
CH ₃	N	462.18	CR	[167]	aCCOO	Na ⁺	CR	CR	-
CH ₃	aCH	305.81	CR	[36]	aCCOO	OH ⁻	CR	CR	-
CH ₃	aCCOO	646.59	18.865	$\dagger\dagger$	aCNH ₂	aCNH ₂	670.91	17.065	$\dagger\dagger$
CH ₃	aCNH ₂	464.73	CR	$\dagger\dagger$	aCNH ₂	N ⁺	CR	CR	-
CH ₃	N ⁺	462.18	CR	[42, 43]	aCNH ₂	aCNH ₃ ⁺	CR	CR	$\dagger\dagger$
CH ₃	aCNH ₃ ⁺	CR	CR	-	aCNH ₂	H ₃ O ⁺	CR	CR	$\dagger\dagger$
CH ₃	H ₃ O ⁺	CR	CR	-	aCNH ₂	Cl ⁻	CR	CR	-
CH ₃	Cl ⁻	CR	CR	-	aCNH ₂	Na ⁺	CR	CR	-
CH ₃	Na ⁺	CR	CR	-	aCNH ₂	OH ⁻	CR	CR	$\dagger\dagger$
CH ₃	OH ⁻	CR	CR	-	N ⁺	N ⁺	62.971	8.8971	[42, 43]
CH ₂	CH ₂	473.39	19.871	[35]	N ⁺	aCNH ₃ ⁺	140.33	CR	$\dagger\dagger$

Table 4.3: Continued from previous page

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref	Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref
CH ₂	N	348.30	CR	[167]	N ⁺	H ₃ O ⁺	61.250	CR	††
CH ₂	aCH	415.64	CR	[36]	N ⁺	Cl ⁻	61.982	CR	[42, 43]
CH ₂	aCCOO	446.39	CR	††	N ⁺	Na ⁺	32.915	CR	††
CH ₂	aCNH ₂	497.03	CR	††	N ⁺	OH ⁻	58.782	CR	††
CH ₂	N ⁺	348.30	CR	[42, 43]	aCNH ₃ ⁺	aCNH ₃ ⁺	397.63	17.065	††
CH ₂	aCNH ₃ ⁺	CR	CR	-	aCNH ₃ ⁺	H ₃ O ⁺	155.38	CR	††
CH ₂	H ₃ O ⁺	CR	CR	-	aCNH ₃ ⁺	Cl ⁻	189.98	CR	††
CH ₂	Cl ⁻	CR	CR	-	aCNH ₃ ⁺	Na ⁺	CR	CR	-
CH ₂	Na ⁺	CR	CR	-	aCNH ₃ ⁺	OH ⁻	CR	CR	-
CH ₂	OH ⁻	CR	CR	-	H ₃ O ⁺	H ₃ O ⁺	68.181	17.020	[150]
N	N	62.971	8.8971	[167]	H ₃ O ⁺	Cl ⁻	70.545	CR	[150]
N	aCH	CR	CR	-	H ₃ O ⁺	Na ⁺	37.479	CR	[44]
N	aCCOO	CR	CR	-	H ₃ O ⁺	OH ⁻	66.420	CR	[150]
N	aCNH ₂	CR	CR	-	Cl ⁻	Cl ⁻	113.76	12.000	[150]
N	N ⁺	CR	CR	-	Cl ⁻	Na ⁺	27.937	CR	[150]
N	aCNH ₃ ⁺	CR	CR	-	Cl ⁻	OH ⁻	123.21	CR	[44]
N	H ₃ O ⁺	CR	CR	††	Na ⁺	Na ⁺	31.709	12.000	[150]
N	Cl ⁻	CR	CR	-	Na ⁺	OH ⁻	27.891	CR	[150]
N	Na ⁺	CR	CR	-	OH ⁻	OH ⁻	170.16	17.020	[150]
N	OH ⁻	CR	CR	-					

Combining rules, given by Equations (2.35)–(2.39), are sometimes used to estimate unlike group interactions, given that the like group parameters are known. Additionally, for the unlike interactions between charged groups, the Hudson and McCoubrey [165] combining rule (Equation (2.40)) can be used, with the corresponding polarisabilities and ionisation potentials of the ions found in Tables 3.2 and 4.5.

Additionally, for the mixtures studied in this chapter, the aCNH₃⁺ and H₃O⁺ ionic groups require association interaction parameters (see Tables 4.1 and 4.4). For these groups, we follow

Table 4.4: Group association energies $\varepsilon_{kl,ab}^{\text{HB}}$ and bonding volume $K_{kl,ab}$ for use with the SAFT- γ Mie EoS. †† indicates parameters obtained in this chapter.

Group k	Site a on group k	Group l	Site b on group l	$(\varepsilon_{kl}^{\text{HB}}/k_{\text{B}})/\text{K}$	$K_{kl,ab}/\text{\AA}^3$	ref
H ₂ O	H	H ₂ O	e ₁	1985.4	101.69	[37, 38]
H ₂ O	H	N	e ₁	2783.7	15.536	[167]
H ₂ O	H	aCH	e ₁	563.56	339.61	[36]
H ₂ O	H	aCCOO	e ₁	1586.2	92.159	††
H ₂ O	H	aCCOO	e ₂	1586.2	92.159	††
H ₂ O	e ₁	aCNH ₂	H	511.27	250.00	††
H ₂ O	H	aCNH ₂	e ₁	810.00	2496.8	††
H ₂ O	e ₁	N ⁺	H	2783.7	15.536	[42, 43]
H ₂ O	e ₁	aCNH ₃ ⁺	H	2550.8	501.61	††
H ₂ O	e ₁	H ₃ O ⁺	H	1985.4	101.69	[150]
H ₂ O	H	OH ⁻	e ₁	1492.0	76.411	[150]
N	e ₁	H ₃ O ⁺	H	2783.7	15.536	††
aCH	e ₁	H ₃ O ⁺	H	563.56	339.61	[44]
aCCOO	e ₁	aCNH ₂	H	1095.6	372.29	††
aCCOO	e ₂	aCNH ₂	H	1095.6	372.29	††
aCCOO	e ₁	aCNH ₃ ⁺	H	1095.6	372.29	††
aCCOO	e ₂	aCNH ₃ ⁺	H	1095.6	372.29	††
aCCOO	e ₁	H ₃ O ⁺	H	1586.2	92.159	††
aCCOO	e ₂	H ₃ O ⁺	H	1586.2	92.159	††
aCNH ₂	H	aCNH ₂	e ₁	700.00	241.64	††
aCNH ₂	e ₁	aCNH ₃ ⁺	H	700.00	241.64	††
aCNH ₂	e ₁	H ₃ O ⁺	H	810.00	2496.8	††
aCNH ₂	H	OH ⁻	e ₁	396.86	194.06	††
N ⁺	H	OH ⁻	e ₁	2098.9	11.714	††
H ₃ O ⁺	H	OH ⁻	e ₁	1492.0	76.411	[150]

Table 4.5: Values for the polarisabilities $\alpha_{0,k}$, and ionisation potentials I_k of the charged species k , taken from the references shown in the last row.

k	aCNH ₃ ⁺	N ⁺
$\alpha_{0,k}/(10^{-24} \text{ cm}^3)$	5.3600	1.1000
I_k/eV	9.3200	14.5340
ref	[215]	[209, 214]

a similar methodology to that adopted by Eriksen et al. [150] given in Equations (3.28) and (3.29) in Chapter 3, in which the association parameters of the ions are related to their closest uncharged group, i.e., H₂O for H₃O⁺, and aCNH₂ for aCNH₃⁺.

4.3.1 Parameter estimation

In general, to develop a new SAFT- γ Mie group, the group parameters are obtained using experimental data as part of a parameter-estimation strategy. The determination of group-interaction parameters is done in a successive manner, where a new group is characterised based on previously developed group parameters. Then, unlike interaction parameters for the new group can be obtained. In the SAFT- γ Mie GC approach the contribution of each functional group is considered to be independent of its position within the molecule, and the group interactions are transferable across systems. Hence, it is possible to develop groups separately before combining them to study the system of interest.

In the current work, to estimate the like and unlike interaction parameters for the new groups aCNH₂, aCCOO and aCNH₃⁺, we take advantage of the transferability of groups within the SAFT- γ Mie approach, by using pure and mixture experimental data for the simplest molecules containing the new groups, i.e., aniline, alkyl benzoates, alkyl 4-aminobenzoates and anilinium chloride (cf. Figure 4.4). Pure compound vapour pressure, enthalpy of vapourisation, and saturated liquid density data, as well as mixture excess enthalpy, isothermal vapor-liquid equilibrium (VLE) (calculated in terms of the bubble and dew pressures), isobaric VLE (calculated in terms of the bubble and dew temperatures), and liquid-liquid equilibrium (LLE) data of the simple molecules are used in characterising the aCNH₂, aCCOO and aCNH₃⁺ groups.

To estimate SAFT- γ Mie parameters in our work, a least-squares objective function (cf. Equation (2.41)) is minimised. The objective function is repeated here for completeness and is given by

$$\min_{\Theta} f_{\text{obj}} = \sum_{X=1}^{N^{\text{P}}} \sum_{X=1}^{N_X^{\text{D}}} w_{X,i} \left(\frac{R_{X,i}^{\text{exp}} - R_{X,i}^{\text{calc}}(\Theta)}{R_{X,i}^{\text{exp}}} \right)^2, \quad (4.21)$$

where Θ is the vector of model parameters, N^{X} is the number of property types, N_X^{D} is the number of experimental data points for property X , $R_{X,i}^{\text{exp}}$ is the i th measured value of property p , $R_{X,i}^{\text{calc}}$ is the corresponding value calculated by the SAFT- γ Mie EoS and $w_{p,i}$ refers to the weight of data point i for property X . Here, we use equal weights for all the data points of all the properties, i.e., $w_{X,i} = 1$, $\forall i, X$.

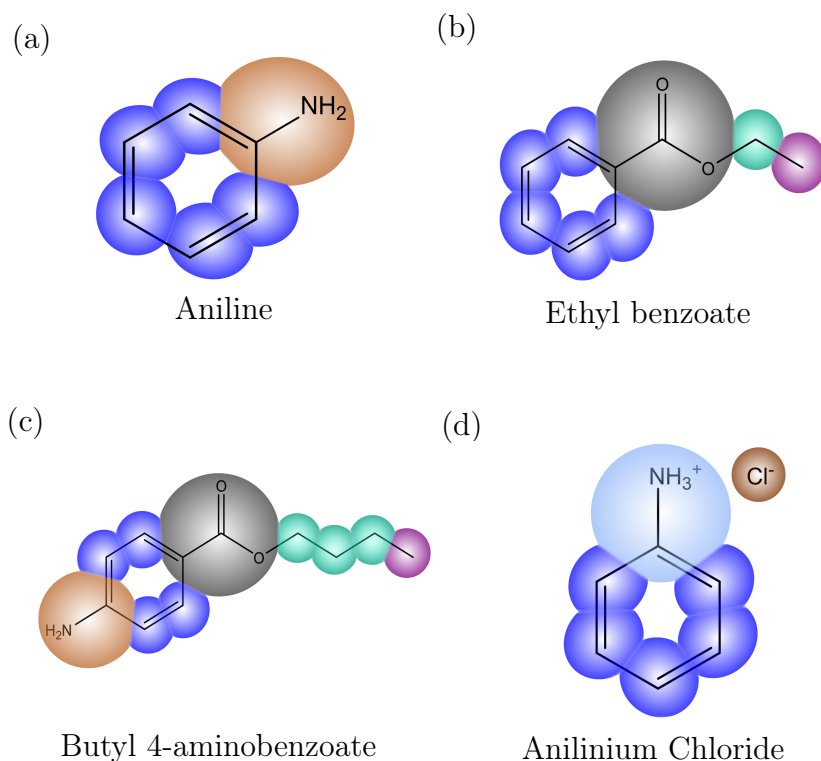


Figure 4.4: Chemical structures of the simplest molecules used in the parameter estimation of the new SAFT- γ Mie group interactions, together with their corresponding SAFT- γ Mie representation. Each of the coloured spheres corresponds to a different group, and the same group is represented in the same colour in figures of (a) aniline, (b) ethyl benzoate, (c) butyl 4-aminobenzoate and (d) anilinium chloride. Alkyl benzoates and alkyl 4-aminobenzoates with a different alkyl chain length can be represented as in (b) and (c), respectively, using the corresponding number of CH_2 groups.

Given that parameter estimation plays a key role in the application of the SAFT- γ Mie GC approach, it is important to determine the accuracy of the theoretical description provided by the EoS using the newly developed groups. Here, we compare the model calculations to experimental data for a property X , using the percentage absolute average deviation (%AAD) and the absolute average deviation (AAD) as given by Equations (2.42) and (2.43), respectively.

In the following sections we outline the successive approach used for the parameter estimation of the new like and unlike interaction parameters of relevance to procaine mixtures studied in this chapter. We proceed first to characterise the like parameters for the aCNH_2 group together with the aCNH_2 – aCH unlike interaction parameters using experimental data for pure aniline and mixtures of aniline + benzene. Similarly, the like interaction parameters for the new aCCOO group are obtained separately using experimental data for pure alkyl benzoates. The aCCOO – aCH , aCCOO – CH_3 and aCCOO – CH_2 unlike interactions are developed simulta-

neously, due to the presence of the aCH, CH₃ and CH₂ groups in alkyl benzoates. Following that, using the newly developed aCNH₂ and aCCOO groups, the aCNH₂-H₂O and aCCOO-H₂O unlike interactions are obtained using aniline + water and alkyl benzoates + water mixture data, respectively. Additionally, the unlike interactions between the two new groups aCNH₂ and aCCOO are characterised using experimental data for pure alkyl-4-aminobenzoates (cf. Figure 4.4 (c)), which are the simplest molecules containing both groups. The CH₃ and CH₂ groups are also used to model alkyl-4-aminobenzoates, therefore, the aCNH₂-CH₃ and aCNH₂-CH₂ unlike parameters are developed concurrently. Finally, we consider the development of the new charged group aCNH₃⁺, which is based on the uncharged aCNH₂ group. The aCNH₃⁺-H₂O unlike interaction parameters are then obtained using experimental mixture data for anilinium chloride + water.

4.3.2 The aCNH₂ group and the aCNH₂-aCH unlike interaction parameters

We start by developing the aromatic amine aCNH₂ group which is modelled with one segment ($\nu_{\text{aCNH}_2}^* = 1$) and two association site types ($N_{\text{ST,aCNH}_2} = 2$): two sites of type H ($n_{\text{aCNH}_2,\text{H}} = 2$) to represent the two hydrogen atoms and one site of type e₁ ($n_{\text{aCNH}_2,\text{e}_1} = 1$) to represent the lone pairs of electrons on the nitrogen atom, where only sites of different types are allowed to interact (i.e., $\varepsilon_{\text{aCNH}_2-\text{aCNH}_2,\text{e}_1-\text{e}_1}^{\text{HB}} = \varepsilon_{\text{aCNH}_2-\text{aCNH}_2,\text{H}-\text{H}}^{\text{HB}} = 0$). Of the remaining like parameters, $\lambda_{\text{aCNH}_2-\text{aCNH}_2}^{\text{a}} = 6$ is fixed, and S_{aCNH_2} , $\sigma_{\text{aCNH}_2-\text{aCNH}_2}$, $\varepsilon_{\text{aCNH}_2-\text{aCNH}_2}$, $\lambda_{\text{aCNH}_2-\text{aCNH}_2}^{\text{r}}$, $\varepsilon_{\text{aCNH}_2-\text{aCNH}_2,\text{H}-\text{e}_1}^{\text{HB}}$ and $K_{\text{aCNH}_2-\text{aCNH}_2,\text{H}-\text{e}_1}$ need to be determined.

To estimate the unknown like parameters of the aCNH₂ group, we use experimental data for aniline (Figure 4.4 (a)) which is represented within the SAFT- γ Mie GC approach by ($5 \times \text{aCH}$ and $1 \times \text{aCNH}_2$). Hence, the unlike dispersion energy between aCH and aCNH₂, $\varepsilon_{\text{aCNH}_2-\text{aCH}}$, is also determined. The aCH group was previously developed by Dufal et al. [36] and is modelled with an associating site of type e₁ to account for the interaction between the delocalised electrons in the aromatic ring (e.g., benzene) and a protic solvent (e.g., water). In our model, we do not account for association between the H site on aCNH₂ and the e₁ on

aCH, as it was suggested in literature that hydrogen bond formation through this interaction is negligible [261, 262].

The seven unknown parameters are estimated by the objective function given in Equation (4.21) using experimental data of pure aniline vapour pressure, saturated liquid density and enthalpy of vapourization, as well as excess enthalpy data of mixtures of aniline and benzene. We note that benzene is represented by $(6 \times \text{aCH})$, a group which is present in aniline so that no additional parameters need to be fitted. The parameters estimated in this section are presented in Tables 4.6, 4.7 and 4.8, and are also included in Tables 4.2, 4.3 and 4.4 together with the remaining like and unlike interaction parameters used in this chapter.

Table 4.6: SAFT- γ Mie like parameters developed in Section 4.3.2. ν_k^* is the number of segments making up group k , S_k is the shape factor, λ_{kk}^r and λ_{kk}^a are the repulsive and attractive exponents of the Mie potential, respectively, σ_{kk} is the segment diameter, ε_{kk} is the dispersion energy between two k groups, $N_{\text{ST},k}$ is the number of association site types on group k , with $n_{\text{H},k}$ and $n_{\text{e}_1,k}$ representing the association sites of type H and e_1 , respectively.

Group k	ν_k^*	S_k	$\sigma_{kk}/\text{\AA}$	λ_{kk}^r	λ_{kk}^a	$(\varepsilon_{kk}/k_{\text{B}})/\text{K}$	$N_{\text{ST},k}$	$n_{\text{H},k}$	$n_{\text{e}_1,k}$
aCNH ₂	1	0.71518	3.7408	17.065	6.0000	670.91	2	2	1

Table 4.7: Group dispersion interaction energies ε_{kl} and repulsive exponent λ_{kl}^r developed in Section 4.3.2 for use with the SAFT- γ Mie equation of state. The unlike segment diameter σ_{kl} are taken from the combining rule in Equation (2.35).

Group k	Group l	$(\varepsilon_{kl}/k_{\text{B}})/\text{K}$	λ_{kl}^r
aCNH ₂	aCH	564.39	23.276

Table 4.8: Group association energies $\varepsilon_{kl,ab}^{\text{HB}}$ and bonding volume $K_{kl,ab}$ developed in Section 4.3.2 for use with the SAFT- γ Mie EoS.

Group k	Site a on group k	Group l	Site b on group l	$(\varepsilon_{kl}^{\text{HB}}/k_{\text{B}})/\text{K}$	$K_{kl,ab}/\text{\AA}^3$
aCNH ₂	H	aCNH ₂	e_1	700.00	241.64

The SAFT- γ Mie calculations using the newly developed aCNH₂ group and its optimised unlike interactions with aCH are compared to experimental data [263–280] considered in the parameter-estimation procedure in Figure 4.5. The model developed results in an excellent agreement with experimental data, and it is particularly remarkable that the model can capture the critical temperature of aniline ($T_c = 702.6 \pm 3.6$ K [281]) (cf. Figure 4.5 (c)) which was not considered in the parameter estimation.

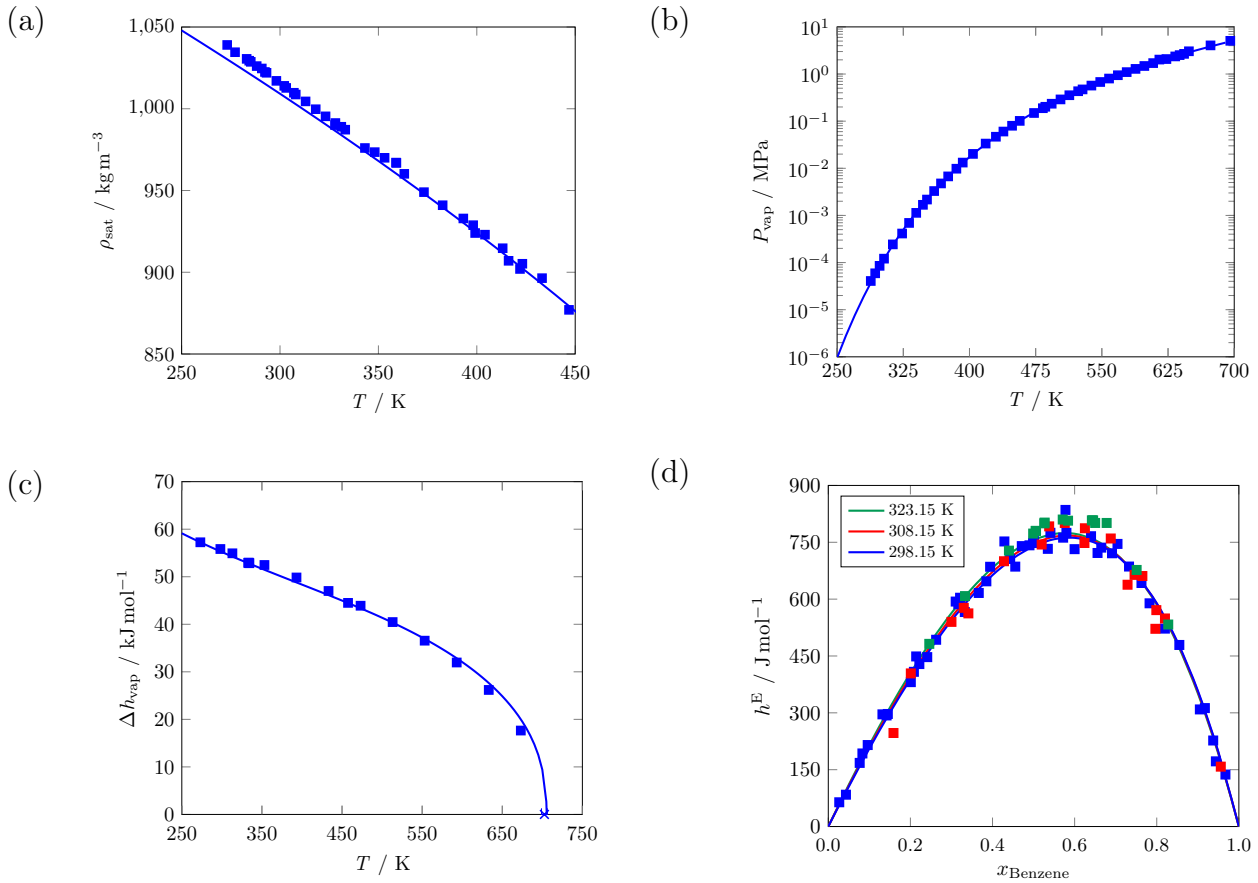


Figure 4.5: (a) Saturated density ρ_{sat} , (b) vapour pressure P_{vap} and (c) vaporization enthalpy Δh_{vap} of pure aniline and (d) excess enthalpy h^E of mixtures of aniline + benzene at $P = 0.101$ MPa. The curves correspond to the SAFT- γ Mie calculations and the symbols represent the experimental data [263–280] used for parameter estimation. The blue cross in (c) corresponds to the critical temperature of aniline [281] which was not used for the estimation of parameters.

In order to assess the validity of the model, we consider the prediction of thermodynamic properties not included in the parameter estimation. The SAFT- γ Mie predictions of the liquid-phase density and isobaric VLE of mixtures of aniline + benzene are compared to experimental data [276, 282, 283] in Figure 4.6. As shown in the figure, very good agreement between the model predictions and experimental data is obtained. It can be noted in Figure 4.6 (a) that the deviation of the model increases for higher concentrations of benzene, perhaps indicating that the aCH group parameters need to be refined.

The high accuracy of the developed model is also demonstrated in the low %AAD values (%AAD < 5 for all properties) given in Table 4.9. We note the especially low %AAD values (%AAD < 1) for the aniline + benzene mixture density, bubble temperature and dew temper-

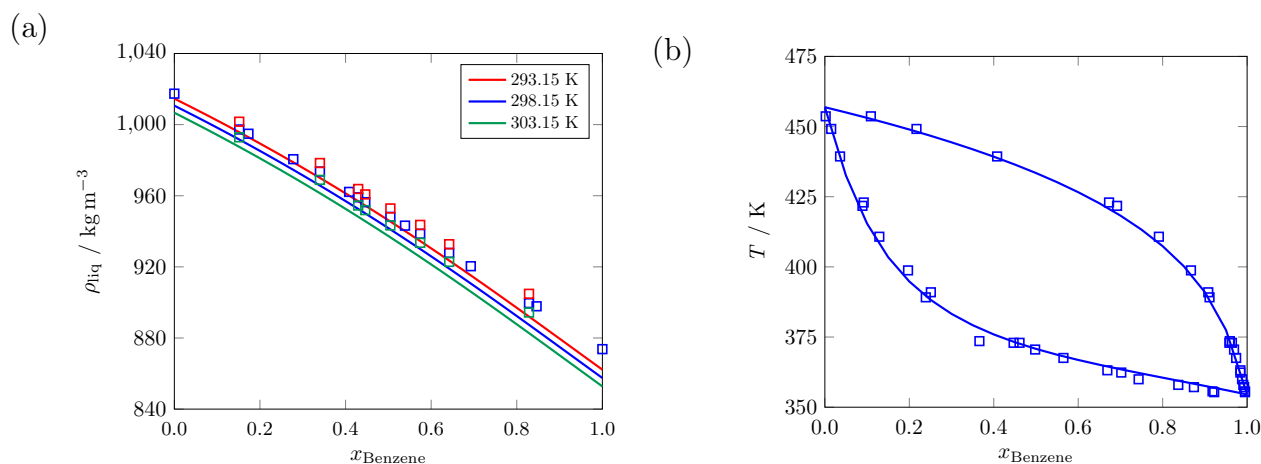


Figure 4.6: (a) Liquid-phase density ρ_{liq} and (b) $(T - x)$ isobars of binary mixtures of aniline + benzene at $P = 0.101$ MPa. The curves correspond to the SAFT- γ Mie calculations and the symbols represent the experimental data [276, 282, 283] used for model validation, as these were not considered in the parameter estimation.

ature, which were not included in the estimation of parameters.

Table 4.9: Overview of the accuracy of the SAFT- γ Mie GC approach using the newly developed aCNH₂ group and the aCNH₂ – aCH unlike interactions in the calculation of the saturated liquid density ρ_{sat} , vapor pressure P_{vap} and vaporization enthalpy Δh_{vap} for pure aniline, and the excess enthalpy h^E at $P = 0.101$ MPa, liquid density ρ_{liq} at $P = 0.101$ MPa, bubble temperature T_{bub} and dew temperature T_{dew} for binary mixtures of aniline and benzene. The experimental data considered in the parameter estimation and model validation are taken from the references indicated.

Parameter estimation						
System	Property X	T / K	N_X^D	%AAD [X]	AAD [X]	ref
Aniline	ρ_{sat} / kg m ⁻³	273–453	40	0.49	4.86	[263–268]
Aniline	P_{vap} / MPa	288–695	45	2.80	0.025	[269–272]
Aniline	Δh_{vap} / kJ mol ⁻¹	273–673	17	1.95	0.61	[273–275]
Aniline + benzene	h^E / J mol ⁻¹	298–323	92	4.49	19.98	[276–280]
Model validation						
System	Property X	T / K	N_X^D	%AAD [X]	AAD [X]	ref
Aniline + benzene	ρ_{liq} / kg m ⁻³	293–303	32	0.87	8.16	[276, 282]
Aniline + benzene	T_{bub} / K	293–303	21	0.46	1.79	[283]
Aniline + benzene	T_{dew} / K	293–303	21	0.30	1.16	[283]

4.3.3 The aCNH₂–H₂O unlike interaction parameters

In order to study the solubility of procaine in water using the SAFT- γ Mie GC approach, the unlike interactions between H₂O and the groups used to represent procaine need to be determined. In the current work, the aCNH₂ – H₂O unlike interaction parameters are determined using the previously developed aCNH₂ (Section 4.3.2) and H₂O [37, 38] groups. A total of six unlike interaction parameters need to be determined between the two groups, namely: $\varepsilon_{\text{aCNH}_2-\text{H}_2\text{O}}$, $\lambda_{\text{aCNH}_2-\text{H}_2\text{O}}^r$, $\varepsilon_{\text{aCNH}_2-\text{H}_2\text{O},\text{H}-\text{e}_1}^{\text{HB}}$, $K_{\text{aCNH}_2-\text{H}_2\text{O},\text{H}-\text{e}_1}$, $\varepsilon_{\text{aCNH}_2-\text{H}_2\text{O},\text{e}_1-\text{H}}^{\text{HB}}$ and $K_{\text{aCNH}_2-\text{H}_2\text{O},\text{e}_1-\text{H}}$. In this model, the repulsive exponent $\lambda_{\text{aCNH}_2-\text{H}_2\text{O}}^r$ is calculated by combining rules (cf. Equation (2.37)). Experimental liquid-phase density and, isobaric (at 0.101 MPa) and isothermal (at 373 K) VLE data for aniline + water mixtures are used in Equation (4.21) to determine the remaining parameters. Consequently, to assess the accuracy of the developed model, we compare the predictions of the phase diagrams (VLE and LLE) of aniline + water mixtures at temperature and pressure ranges not used for the estimation of parameters. The optimized unlike parameters can be found in Tables 4.10 and 4.11, and are also included in Tables 4.3 and 4.4 together with the remaining unlike interaction parameters used in this chapter.

Table 4.10: Group dispersion interaction energies ε_{kl} and repulsive exponent λ_{kl}^r developed in Section 4.3.3 for use with the SAFT- γ Mie equation of state. The unlike segment diameter σ_{kl} are taken from the combining rule in Equation (2.35). CR indicates that combining rules are used to obtain λ_{kl}^r using Equation (2.37).

Group k	Group l	$(\varepsilon_{kl}/k_B)/\text{K}$	λ_{kl}^r
aCNH ₂	H ₂ O	438.28	CR

Table 4.11: Group association energies $\varepsilon_{kl,ab}^{\text{HB}}$ and bonding volume $K_{kl,ab}$ developed in Section 4.3.3 for use with the SAFT- γ Mie equation of state.

Group k	Site a on group k	Group l	Site b on group l	$(\varepsilon_{kl}^{\text{HB}}/k_B)/\text{K}$	$K_{kl,ab}/\text{\AA}^3$
aCNH ₂	H	H ₂ O	e ₁	511.27	250.00
aCNH ₂	e ₁	H ₂ O	H	810.00	2496.8

The SAFT- γ Mie calculations using the newly developed aCNH₂ – H₂O unlike interactions show very good agreement to experimental data in terms of the density [284] and the phase

behaviour [285–290] of aniline + water mixtures in Figure 4.7. As can be seen in the insets of Figures 4.7 (b)–(e), the SAFT- γ Mie model predicts the formation of a heteroazeotrope in aniline + water mixtures in the water-rich regions of the isobaric and isothermal phase diagrams. This finding is validated by experimental data [288] reported at 0.09893 MPa (inset in Figure 4.7 (c)), which is outstanding given that this data was not considered in the parameter estimation procedure. Additionally, the model predicts the occurrence of a liquid-liquid phase separation (LLE) for these mixtures in Figures 4.7 (b)–(e), which is validated by the experimental data [289, 290] reported at 0.101330 MPa for $T < 372$ K as shown in Figure 4.7 (b). The SAFT- γ Mie approach provides a qualitative prediction of the LLE in the figure, showing an excellent agreement with experimental data for the composition of aniline in the water-rich region of the phase diagram; although, the model overpredicts the mole fraction of aniline in the organic-rich region of the diagram. In the case of the isothermal phase diagram at 373 K (Figure 4.7 (e)), the LLE region is predicted for the binary mixture of aniline and water with the corresponding three-phase vapor-liquid-liquid equilibrium (VLLE) occurring at 0.10133 MPa. Similarly, an LLE region is predicted in the isobaric phase diagrams of the binary mixture at 0.09893 MPa and 0.09933 MPa, with a triple phase coexistence at 372.0 K for the two isobars as shown in Figures 4.7 (c) and (d), respectively. We note that the accuracy of the model predictions of the LLE regions in Figures 4.7 (c), (d) and (e) could not be assessed due to the absence of experimental data for this phase equilibrium at these conditions.

In terms of %AAD and AAD, it is reassuring to note in Table 4.12 the small deviations of the model from experimental data for all the thermodynamic properties considered. It should be noted that, the high %AAD values evident in the table for the LLE compositions of aniline in the water-rich x_A^{aq} and organic-rich x_A^{org} regions of the phase diagram (Figure 4.7 (b)), are due to the division by a small number (< 1) in Equation (2.42); therefore, the AAD is more reliable in assessing the accuracy of the model for these properties.

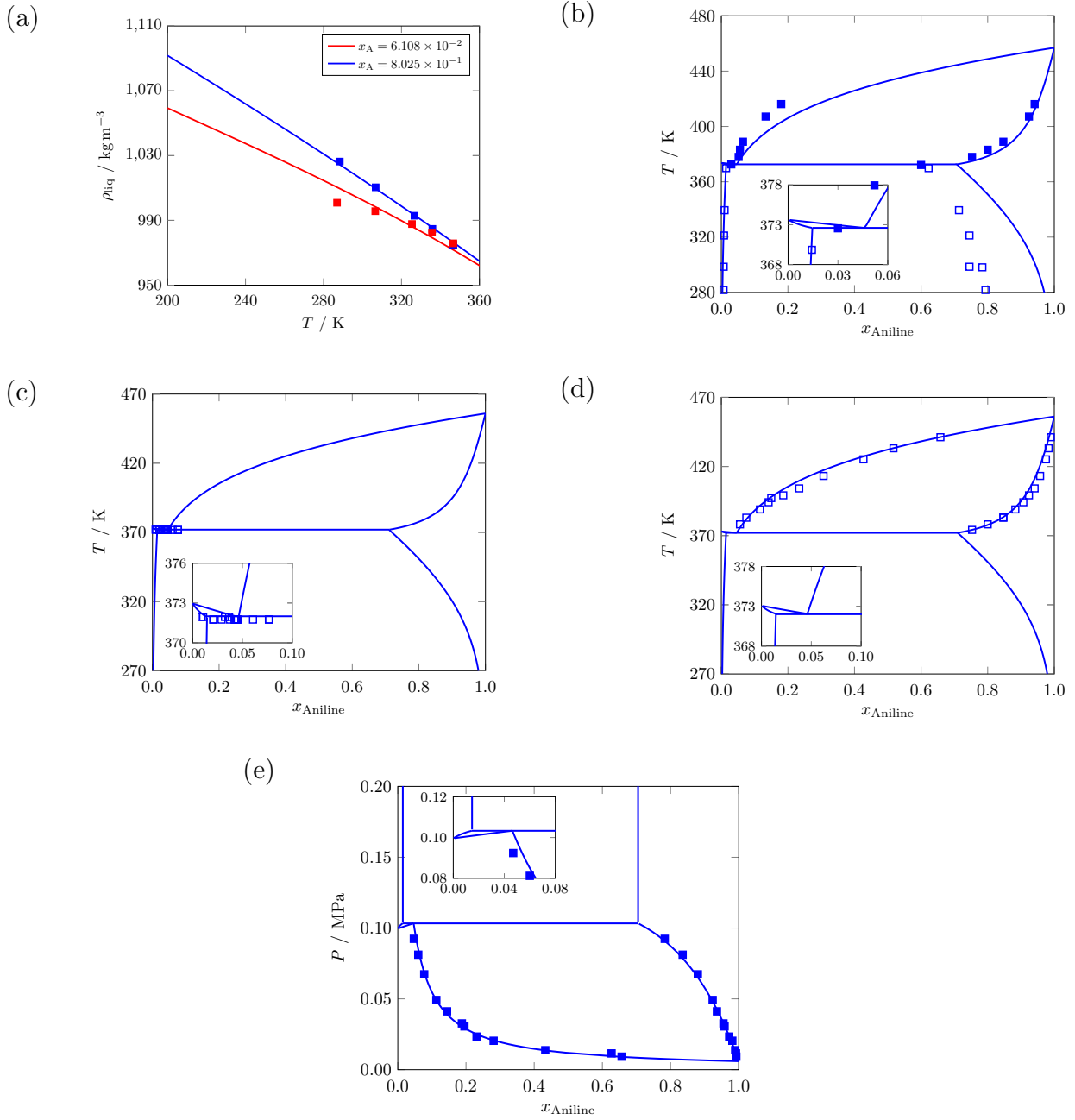


Figure 4.7: (a) Liquid-phase density ρ_{liq} at $P = 0.10133 \text{ MPa}$; $(T - x)$ isobars at (b) $P = 0.10133 \text{ MPa}$, (c) $P = 0.09893 \text{ MPa}$, (d) $P = 0.09933 \text{ MPa}$; and (e) $(P - x)$ isotherms at $T = 373.15 \text{ K}$ of binary mixtures of aniline + water. The curves correspond to the SAFT- γ Mie calculations, the symbols represent the experimental data used for parameter estimation (filled) [284–287] and those not used (unfilled) [287–290].

4.3.4 The aCCOO group and its unlike interactions aCCOO–aCH, aCCOO–CH₂ and aCCOO–CH₃

The aromatic ester aCCOO group is modelled with two segments ($\nu_{\text{aCCOO}}^* = 2$) and two association site types ($N_{\text{ST,aCCOO}} = 2$): two sites of type e_1 ($n_{\text{aCCOO},e_1} = 2$) corresponding to the

Table 4.12: The percentage average absolute deviation %AAD and the absolute average deviation AAD between the SAFT- γ Mie calculations and experimental data of the liquid-phase density ρ_{liq} , bubble and dew temperatures (T_{bub} , T_{dew}), bubble and dew pressures (P_{bub} , P_{dew}), and composition of aniline at the corresponding aniline-water LLE in the water-rich x_{A}^{aq} and organic-rich $x_{\text{A}}^{\text{org}}$ regions of the phase diagram (Figure 4.7 (b)) for mixtures of aniline and water. The SAFT- γ Mie calculations are performed using the newly developed aCNH₂ – H₂O unlike interaction parameters, and the experimental data considered in the parameter estimation and model validation are taken from the references indicated in the last column.

Parameter estimation						
Property X	P / MPa	T / K	N_X^{D}	%AAD [X]	AAD [X]	ref
ρ_{liq} / kg m ⁻³	0.10133	286–303	18	0.28	2.84	[284]
T_{bub} / K	0.10133	370–422	10	1.19	4.72	[285, 286]
T_{dew} / K	0.10133	370–422	10	2.35	5.18	[285, 286]
P_{bub} / MPa	0.0092–0.0924	373.15	12	4.28	1.20×10^{-3}	[287]
P_{dew} / MPa	0.0092–0.0924	373.15	12	5.22	2.09×10^{-3}	[287]
Model validation						
Property X	P / MPa	T / K	N_X^{D}	%AAD [X]	AAD	ref
T_{bub} / K	0.09893–0.09933	374–416	18	0.91	3.71	[287, 288]
T_{dew} / K	0.09893–0.09933	441–416	16	1.36	5.41	[287, 288]
x_{A}^{aq}	0.10133	298–370	6	26.19	1.95×10^{-3}	[289, 290]
$x_{\text{A}}^{\text{org}}$	0.10133	298–370	6	20.62	1.53×10^{-1}	[289, 290]

lone pairs of electrons on one oxygen atom and two sites of type e₂ ($n_{\text{aCCOO}, \text{e}_2} = 2$) to represent the lone pairs of electrons on the second oxygen atom. Thus, self association between aCCOO groups is not considered. The attractive exponent of aCCOO $\lambda_{\text{aCCOO}-\text{aCCOO}}^{\text{a}} = 6$ is fixed, while the remaining like parameters, S_{aCCOO} , $\sigma_{\text{aCCOO}-\text{aCCOO}}$, $\varepsilon_{\text{aCCOO}-\text{aCCOO}}$ and $\lambda_{\text{aCCOO}-\text{aCCOO}}^{\text{r}}$, need to be determined.

Here, we use pure and mixture experimental data for alkyl benzoates (C₆H₅CO₂C_{*n*}H_{2*n*+1} for $n = 1, 2, 3$ and 4), namely: methyl benzoate ($n = 1$), ethyl benzoate ($n = 2$), propyl benzoate ($n = 3$) and butyl benzoate ($n = 4$), to estimate the like interaction parameters of the aCCOO group using the objective function in Equation (4.21). These alkyl benzoates are represented in the SAFT- γ Mie framework by ($5 \times \text{aCH}$, $1 \times \text{aCCOO}$, $1 \times \text{CH}_3$ and $(n - 1) \times \text{CH}_2$) (cf. Figure 4.4 (b)). Therefore, we additionally determine the unlike interaction parameters of aCCOO–aCH, aCCOO–CH₃ and aCCOO–CH₂, which are: $\varepsilon_{\text{aCCOO}-\text{aCH}}$, $\lambda_{\text{aCCOO}-\text{aCH}}^{\text{r}}$, $\varepsilon_{\text{aCCOO}-\text{CH}_3}$, $\lambda_{\text{aCCOO}-\text{CH}_3}^{\text{r}}$, $\varepsilon_{\text{aCCOO}-\text{CH}_2}$ and $\lambda_{\text{aCCOO}-\text{CH}_2}^{\text{r}}$. To estimate the like and unlike inter-

action parameters of aCCOO, we use experimental data of the vapour pressure and saturated liquid density of pure methyl-, ethyl-, propyl- and butyl benzoate, the enthalpy of vapourization of pure methyl- and ethyl benzoate, together with the excess enthalpy data of methyl benzoate + benzene and ethylbenzoate + benzene mixtures. The parameters estimated in this section are presented in Tables 4.13 and 4.14, and are also included in Tables 4.2 and 4.3 together with the remaining like and unlike interaction parameters used in this chapter.

Table 4.13: SAFT- γ Mie like parameters developed in Section 4.3.4. ν_k^* is the number of segments making up group k , S_k is the shape factor, λ_{kk}^r and λ_{kk}^a are the repulsive and attractive exponents of the Mie potential, respectively, σ_{kk} is the segment diameter, ε_{kk} is the dispersion energy between two k groups, $N_{ST,k}$ is the number of association site types on group k , with $n_{e_1,k}$ and $n_{e_2,k}$ representing the association sites of type e_1 and e_2 , respectively.

Group k	ν_k^*	S_k	$\sigma_{kk}/\text{\AA}$	λ_{kk}^r	λ_{kk}^a	$(\varepsilon_{kk}/k_B)/K$	$N_{ST,k}$	$n_{e_1,k}$	$n_{e_2,k}$
aCCOO	2	0.65563	3.3851	36.616	6.0000	200.00	2	2	2

Table 4.14: Group dispersion interaction energies ε_{kl} and repulsive exponent λ_{kl}^r for use with the SAFT- γ Mie equation of state. The unlike segment diameter σ_{kl} are taken from the combining rule in Equation (2.35). CR indicates that combining rules are used to obtain λ_{kl}^r using Equation (2.37).

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r
aCCOO	aCH	468.32	CR
aCCOO	CH ₂	446.39	CR
aCCOO	CH ₃	646.59	18.865

The SAFT- γ Mie calculations using the newly characterised aCCOO group and its unlike interaction parameters with aCH, CH₂ and CH₃ are compared to experimental data [264, 270, 291–302] of alkyl benzoates in Figure 4.8, in which a good agreement is observed for all properties. The SAFT- γ Mie calculations of the thermodynamic properties of pure alkyl benzoates follow the same trend as experiments, i.e., decreasing liquid density (Figure 4.8 (a)), decreasing vapour pressure (Figure 4.8 (b)) and increasing enthalpy of vaporisation (Figure 4.8 (c)) with increasing chain length. Additionally, we note the accuracy of the model for the excess enthalpy of ethyl benzoate + benzene mixtures in Figure 4.8 (e), where the model captures the inflection in experimental data and the negative enthalpy values for $x_{\text{Benzene}} < 0.2$. Furthermore, to asses the robustness of the model developed we predict thermodynamic properties not considered in the parameter estimation procedure. In particular, we predict

the liquid-phase density of methyl benzoate + benzene (Figure 4.8 (f)) and ethyl benzoate + benzene (Figure 4.8 (g)) mixtures, which are shown to be in qualitative agreement with experimental data.

The performance of the model is also assessed in terms of the %AAD and AAD in Table 4.15. The agreement between experimental data and the SAFT- γ Mie calculations, in terms of %AAD, is generally lower than 7 % for the pure compounds' saturated density and enthalpy of vaporization, and for the liquid-phase density of the mixtures. Larger deviations are evident, however, for the vapour pressure of the pure compounds (e.g., 14.94% for ethylbenzoate) and the excess enthalpy of the mixtures (e.g. 19.23% for methyl benzoate + benzene). The larger deviations for vapour pressure are due to the wide range of pressures considered, ranging from an order of magnitude of 10^{-5} to 10^{-1} MPa. At very low pressure values, a high uncertainty in the experimental measurements and an increased sensitivity to the group parameters, is expected, resulting in a larger relative deviation (%AAD). Analogously, the large %AADs obtained for the excess enthalpy of the mixtures are due to the very low enthalpy values of these mixtures, for which higher uncertainties in the experimental measurements are probable. For these properties, it is useful to consider the absolute deviations (AAD) for an overall assessment of the accuracy of the model. The AAD for the vapour pressure of pure alkylbenzoates and the excess enthalpy of mixtures with benzene are shown to be relatively low. Therefore, based on the overall assessment of the model in Figure 4.8 and Table 4.15, the newly developed aCCOO group and its unlike interactions aCCOO-aCH, aCCOO-CH₃ and aCCOO-CH₂ are considered satisfactory.

4.3.5 The aCCOO-H₂O unlike interaction parameters

The unlike interaction parameters between the aCCOO group, developed in the previous section, and the H₂O group developed by Dufal et al. [37, 38], are required in order to study the solubility of procaine in water using the SAFT- γ Mie GC approach. The aCCOO - H₂O unlike interaction parameters to be determined are: $\varepsilon_{\text{aCCOO-H}_2\text{O}}$, $\lambda_{\text{aCCOO-H}_2\text{O}}^r$, $\varepsilon_{\text{aCCOO-H}_2\text{O},\text{e}_1-\text{H}}^{\text{HB}}$, $K_{\text{aCCOO-H}_2\text{O},\text{e}_1-\text{H}}$, $\varepsilon_{\text{aCCOO-H}_2\text{O},\text{e}_2-\text{H}}^{\text{HB}}$ and $K_{\text{aCCOO-H}_2\text{O},\text{e}_2-\text{H}}$. In this model, the repulsive expo-

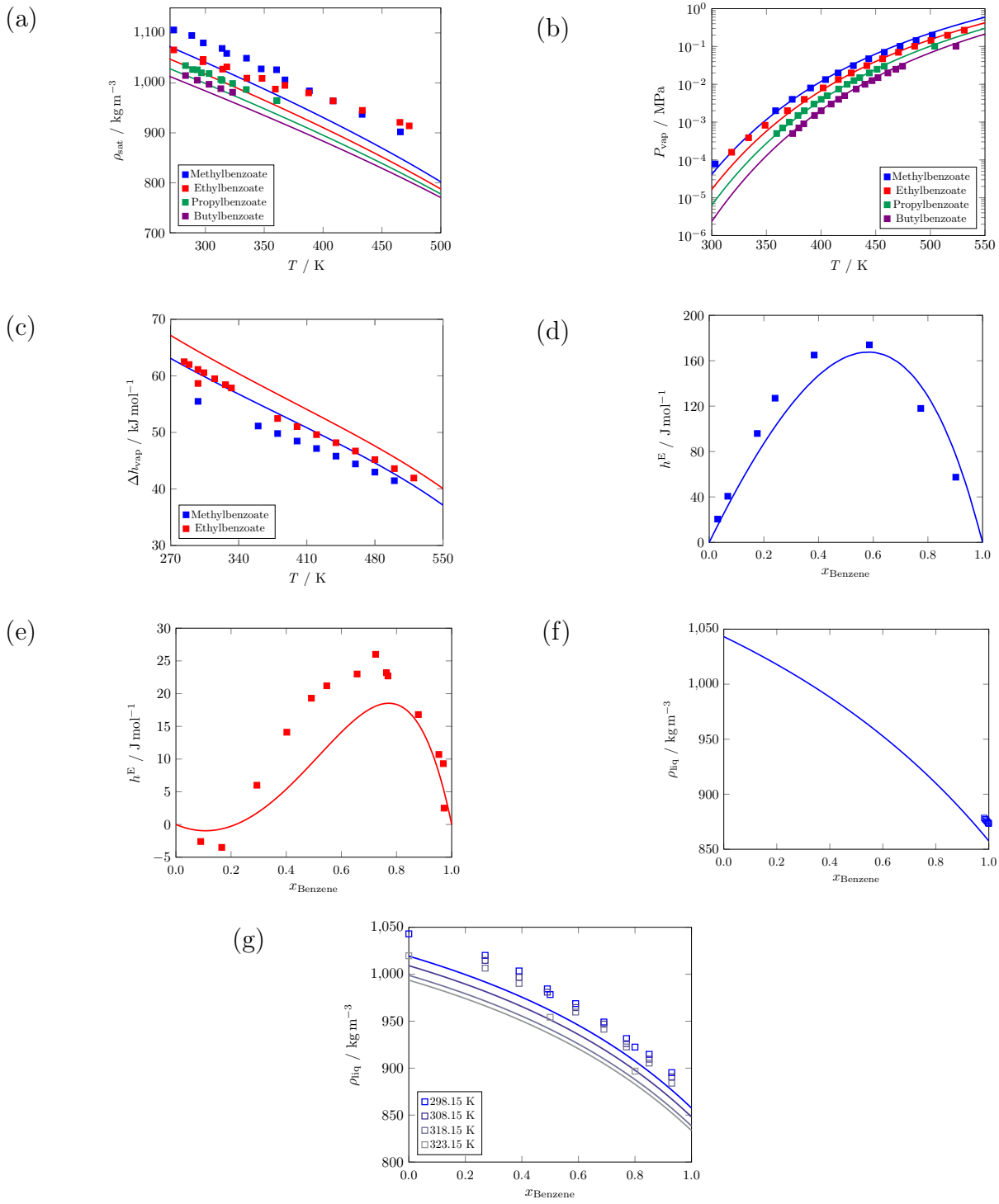


Figure 4.8: (a) Saturated density ρ_{sat} , (b) vapour pressure P_{vap} and (c) vaporization enthalpy Δh_{vap} of pure alkyl benzoates; (d) excess enthalpy h^{E} of binary mixtures of methyl benzoate + benzene at $T = 303.15$ K and $P = 0.101$ MPa; (e) excess enthalpy h^{E} of binary mixtures of ethyl benzoate + benzene at $T = 298.15$ K and $P = 0.101$ MPa; (f) liquid-phase density ρ_{liq} of binary mixtures of methyl benzoate + benzene at $T = 298.15$ K and $P = 0.101$ MPa; and (g) liquid-phase density ρ_{liq} of binary mixtures of ethyl benzoate + benzene at $P = 0.101$ MPa. The curves correspond to the SAFT- γ Mie calculations, the symbols represent the experimental data used for parameter estimation (filled) [264, 270, 291–299] and those not used (unfilled) [300–302].

Table 4.15: Overview of the accuracy of the SAFT- γ Mie GC approach using the newly developed aCCOO group and its unlike interaction parameters aCCOO – aCH, aCCOO – CH₂ and aCCOO – CH₃ in the calculation of the saturated liquid density ρ_{sat} , vapor pressure P_{vap} and vaporization enthalpy Δh_{vap} for pure methyl-, ethyl-, propyl- and butyl benzoate, and the excess enthalpy h^E and liquid density ρ_{liq} at $P = 0.101$ MPa for binary mixtures of methyl benzoate + benzene and ethyl benzoate + benzene. The experimental data considered in the parameter estimation and model validation are taken from the references indicated.

Parameter estimation					
System	T / K	N_X^D	%AAD ρ_{sat}	AAD ρ_{sat} / kg m ⁻³	ref
Methyl benzoate	273–360	13	4.15	42.21	[264, 291]
Ethyl benzoate	297–473	15	4.58	44.66	[264, 291]
Propyl benzoate	289–323	11	2.18	21.87	[291, 292]
Butyl benzoate	283–323	5	1.66	16.56	[292]
System	T / K	N_X^D	%AAD P_{vap}	AAD $\times 10^3 P_{\text{vap}}$ / MPa	ref
Methyl benzoate	358–501	14	15.23	6.58	[270, 293, 294]
Ethyl benzoate	318–531	15	14.94	4.47	[294, 295]
Propyl benzoate	359–504	16	4.71	1.00	[291, 293]
Butyl benzoate	374–523	16	8.72	2.08	[293, 296]
System	T / K	N_X^D	%AAD Δh_{vap}	AAD Δh_{vap} / kJ mol ⁻¹	ref
Methyl benzoate	298–500	9	5.89	2.87	[294]
Ethyl benzoate	298–520	16	6.02	3.25	[294, 297]
System	T / K	N_X^D	%AAD h^E	AAD h^E / J mol ⁻¹	ref
Benzene + methyl benzoate	303.15	8	19.23	15.53	[298]
Benzene + ethyl benzoate	298.15	12	41.93	5.53	[299]
Model validation					
System	T / K	N_X^D	%AAD ρ_{liq}	AAD ρ_{liq} / kg m ⁻³	ref
Benzene + methyl benzoate	298.15	10	1.81	2.87	[300]
Benzene + ethyl benzoate	298.15–323.15	29	2.57	24.66	[301, 302]

ment $\lambda_{\text{aCCOO-H}_2\text{O}}^r$ is calculated by combining rules (Equation (2.37)), and a symmetric association scheme is used, such that, $\varepsilon_{\text{aCCOO-H}_2\text{O},\text{e}_1-\text{H}}^{\text{HB}} = \varepsilon_{\text{aCCOO-H}_2\text{O},\text{e}_2-\text{H}}^{\text{HB}}$ and $K_{\text{aCCOO-H}_2\text{O},\text{e}_1-\text{H}} = K_{\text{aCCOO-H}_2\text{O},\text{e}_2-\text{H}}$.

Experimental LLE data of alkyl benzoates in water [303, 304] are used in the parameter estimation (cf. Equation (4.21)) of the aCCOO-H₂O unlike interaction parameters. The SAFT- γ Mie calculations of the LLE of methylbenzoate + water and ethyl benzoate + water mixtures at 0.10133 MPa are compared to experimental data in Figure 4.9. As shown in the figure, excellent agreement to experimental data is obtained in the water- and organic-rich regions of the phase diagrams. The accuracy of the model is also demonstrated by the low %AAD and AAD for the

liquid-phase compositions of the two alkylbenzoates in Table 4.16. Additionally, a VLE region is predicted by the SAFT- γ Mie GC approach for aqueous mixtures of methyl benzoate and ethyl benzoate at $T > 372.70$ K and $T > 373.15$ K, respectively. Moreover, the insets in Figures 4.9 (a) and (b) show the model prediction of a heteroazeotrope at $x_{\text{methylbenzoate}} = 0.0322$ and $x_{\text{ethylbenzoate}} = 0.0164$ for the methyl benzoate + water and ethyl benzoate + water mixtures, respectively.

The optimized aCCOO-H₂O unlike parameters can be found in Tables 4.17 and 4.18, and are also included in Tables 4.3 and 4.4 together with the remaining unlike interaction parameters used in this chapter.

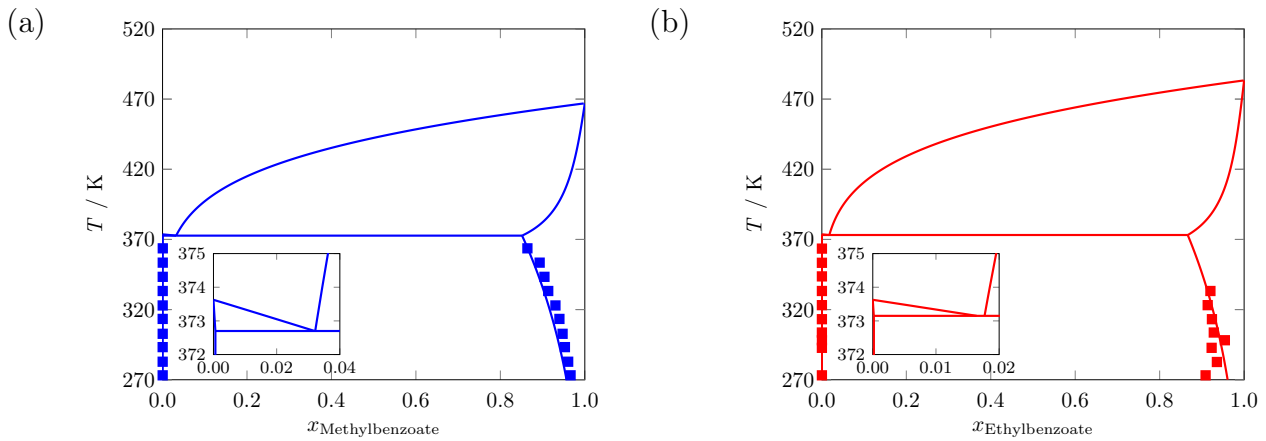


Figure 4.9: $(T-x)$ isobars at atmospheric pressure of binary mixtures of (a) methyl benzoate + water and (b) ethyl benzoate + water. The curves correspond to the SAFT- γ Mie calculations and the symbols represent the experimental data [303, 304] used for parameter estimation.

4.3.6 The aCNH₂-aCCOO, aCNH₂-CH₂ and aCNH₂-CH₃ unlike interaction parameters

To complete the set of unlike interactions required to describe neutral procaine using the SAFT- γ Mie GC approach, the unlike interaction parameters between the newly developed aCNH₂ (Section 4.3.2) and aCCOO (Section 4.3.4) groups need to be determined. Specifically, $\varepsilon_{\text{aCNH}_2-\text{aCCOO}}$, $\lambda_{\text{aCNH}_2-\text{aCCOO}}^r$, $\varepsilon_{\text{aCNH}_2-\text{aCCOO,H-e}_1}^{\text{HB}}$, $K_{\text{aCNH}_2-\text{aCCOO,H-e}_1}$, $\varepsilon_{\text{aCNH}_2-\text{aCCOO,H-e}_2}^{\text{HB}}$ and $K_{\text{aCNH}_2-\text{aCCOO,H-e}_2}$ are required. Here, the repulsive exponent $\lambda_{\text{aCNH}_2-\text{aCCOO}}^r$ is calculated by

Table 4.16: The percentage average absolute deviation %AAD and the absolute average deviation AAD between the SAFT- γ Mie calculations and experimental data of the composition of alkylbenzoates at the corresponding alkylbenzoate-water LLE in the water-rich x_1^{aq} and organic-rich x_1^{org} regions of the phase diagram (Figure 4.9) for mixtures of methylbenzoate + water and ethylbenzoate + water. The SAFT- γ Mie calculations are performed using the newly developed aCCOO – H₂O unlike interaction parameters, and the experimental data considered in the parameter estimation are taken from the references indicated in the last column.

Parameter estimation						
System s (1 + 2)	T / K	P / MPa	N_X^D	%AAD x_1^{aq}	AAD $\times 10^5$ x_1^{aq}	ref
Methylbenzoate + water	273–363	0.101	10	11.89	4.07	[303]
Ethylbenzoate + water	273–363	0.101	10	18.01	2.61	[303, 304]
System s (1 + 2)	T / K	P / MPa	N_X^D	%AAD x_1^{org}	AAD x_1^{org}	ref
Methylbenzoate + water	273–363	0.101	10	1.32	0.012	[303]
Ethylbenzoate + water	273–333	0.101	8	4.58	0.017	[303, 304]

Table 4.17: Group dispersion interaction energies ε_{kl} and repulsive exponent λ_{kl}^r developed in Section 4.3.5 for use with the SAFT- γ Mie equation of state. The unlike segment diameter σ_{kl} is taken from the combining rule in Equation (2.35). CR indicates that combining rules are used to obtain λ_{kl}^r using Equation (2.37).

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r
aCCOO	H ₂ O	339.28	CR

Table 4.18: Group association energies $\varepsilon_{kl,ab}^{\text{HB}}$ and bonding volume $K_{kl,ab}$ developed in Section 4.3.5 for use with the SAFT- γ Mie equation of state.

Group k	Site a on group k	Group l	Site b on group l	$(\varepsilon_{kl}^{\text{HB}}/k_B)/K$	$K_{kl,ab}/\text{\AA}^3$
aCCOO	e ₁	H ₂ O	H	1586.2	92.158
aCCOO	e ₂	H ₂ O	H	1586.2	92.158

combining rules. The remaining parameters are estimated by minimizing the objective function in Equation (4.21) and using saturated liquid density correlated experimental data from NIST [281] for pure alkyl 4-aminobenzoates ($\text{C}_6\text{H}_4\text{NH}_2\text{CO}_2\text{C}_n\text{H}_{2n+1}$ for $n = 3, 4, 5, 6, 7$ and 8), which are modelled as ($4 \times \text{aCH}$, $1 \times \text{aCCOO}$, $1 \times \text{aCNH}_2$, $1 \times \text{CH}_3$ and $(n - 1) \times \text{CH}_2$) in the SAFT- γ Mie framework. According to the groups used to model the alkyl 4-aminobenzoates, the unlike interactions between the aCNH₂ and, CH₂ and CH₃ groups can be refined together with the aCNH₂ – aCCOO interactions; therefore, two additional parameters are estimated, namely: $\varepsilon_{\text{aCNH}_2-\text{CH}_2}$ and $\varepsilon_{\text{aCNH}_2-\text{CH}_3}$. The repulsive exponents $\lambda_{\text{aCNH}_2-\text{CH}_2}^r$ and $\lambda_{\text{aCNH}_2-\text{CH}_3}^r$ are obtained using combining rules; such that a total of seven parameters remain to be obtained.

To reduce the number of parameters that need to be estimated, we consider a symmetric association scheme for the association interactions between the aCNH₂ and aCCOO groups, i.e., $\varepsilon_{\text{aCNH}_2-\text{aCCOO},\text{H}-\text{e}_1}^{\text{HB}} = \varepsilon_{\text{aCNH}_2-\text{aCCOO},\text{H}-\text{e}_2}^{\text{HB}}$ and $K_{\text{aCNH}_2-\text{aCCOO},\text{H}-\text{e}_1} = K_{\text{aCNH}_2-\text{aCCOO},\text{H}-\text{e}_2}$. The parameters determined in this section are presented in Tables 4.19 and 4.20, and are also included in Tables 4.3 and 4.4 together with the remaining unlike interaction parameters used in this chapter.

Table 4.19: Group dispersion interaction energies ε_{kl} and repulsive exponent λ_{kl}^r developed in Section 4.3.6 for use with the SAFT- γ Mie equation of state. The unlike segment diameters σ_{kl} calculated with the combining rule in Equation (2.35). CR indicates that combining rules are used to obtain λ_{kl}^r using Equation (2.37).

Group k	Group l	$(\varepsilon_{kl}/k_B)/\text{K}$	λ_{kl}^r
aCNH ₂	aCCOO	494.91	CR
aCNH ₂	CH ₂	497.03	CR
aCNH ₂	CH ₃	464.73	CR

Table 4.20: Group association energies $\varepsilon_{kl,ab}^{\text{HB}}$ and bonding volume $K_{kl,ab}$ developed in Section 4.3.6 for use with the SAFT- γ Mie equation of state.

Group k	Site a on group k	Group l	Site b on group l	$(\varepsilon_{kl}^{\text{HB}}/k_B)/\text{K}$	$K_{kl,ab}/\text{\AA}^3$
aCNH ₂	H	aCCOO	e ₁	1095.6	372.29
aCNH ₂	H	aCCOO	e ₂	1095.6	372.29

In Figure 4.10, we show the comparison between experimental data and SAFT- γ Mie calculations for the saturated liquid density of propyl-, butyl-, pentyl-, hexyl-, heptyl- and octyl 4-aminobenzoates. The model provides a qualitatively correct description of the saturated liquid density of the alkyl 4-aminobenzoates, although the theory is found to underestimate the saturated density of the pure compounds. However, the SAFT- γ Mie calculations remain within the uncertainty limits of the data at low temperatures (≤ 360 K). Furthermore, the %AAD and AAD for the saturated liquid density of the alkyl 4-aminobenzoates (Table 4.21) indicate an increase in the deviation of the model from experimental data with increasing alkyl chain length. The relatively high deviations observed are due to the large uncertainties in experimental data, which are higher for longer alkyl 4-aminobenzoates. Hence, a proper analysis of the performance of the model in terms of %AAD and AAD is not possible.

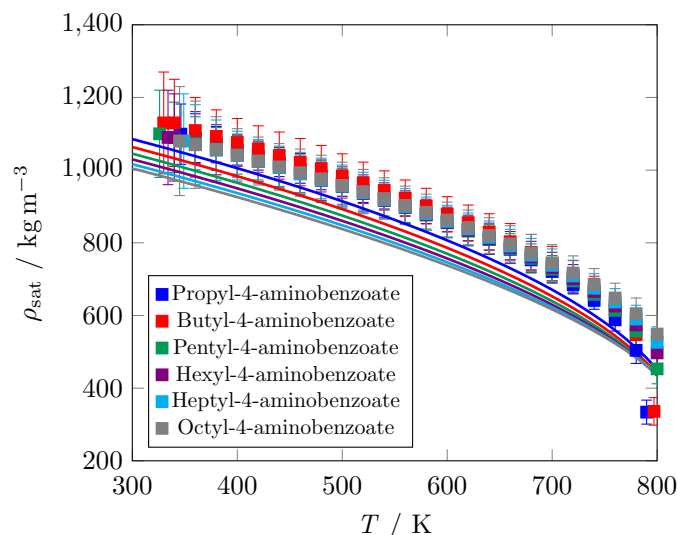


Figure 4.10: Saturated density ρ_{sat} of alkyl-4-aminobenzoates. The curves correspond to the SAFT- γ Mie calculations, the symbols represent NIST correlations [281] used for parameter estimation and the error bars represent the uncertainty in the experimental data.

Table 4.21: Overview of the accuracy of the SAFT- γ Mie GC approach using the newly developed unlike interaction parameters $\text{aCNH}_2 - \text{aCCOO}$, $\text{aCNH}_2 - \text{CH}_2$ and $\text{aCNH}_2 - \text{CH}_3$ in the calculation of the saturated liquid density ρ_{sat} for pure alkyl 4-aminobenzoates. The experimental data considered in the parameter estimation are taken from the references indicated in the last column.

System	T / K	N_X^D	%AAD ρ_{sat}	AAD $\rho_{\text{sat}} / \text{kg m}^{-3}$	ref
Propyl 4-aminobenzoate	346–790	24	7.00	51.07	[281]
Butyl 4-aminobenzoate	329–797	15	10.96	90.09	[281]
Pentyl 4-aminobenzoate	326–805	16	10.10	79.24	[281]
Hexyl 4-aminobenzoate	334–812	16	12.12	94.17	[281]
Heptyl 4-aminobenzoate	349–819	15	13.41	106.97	[281]
Octyl 4-aminobenzoate	345–826	16	14.63	116.71	[281]

Given the lack of experimental data with lower uncertainties, we consider that the the newly developed $\text{aCNH}_2 - \text{aCCOO}$, $\text{aCNH}_2 - \text{CH}_2$ and $\text{aCNH}_2 - \text{CH}_3$ unlike interaction parameters are satisfactory. Further refinement of these models will be necessary when new experimental data become available.

4.3.7 The aCNH_3^+ group and the $\text{aCNH}_3^+ - \text{H}_2\text{O}$ unlike interaction parameters

In order to model the second ionisation of procaine, the aCNH_3^+ group and the $\text{aCNH}_3^+ - \text{H}_2\text{O}$ unlike interactions are needed. A SAFT- γ Mie model for the aCNH_3^+ group is developed here based on the uncharged aromatic amine group (aCNH_2) characterised in Section 4.3.2. Assuming that the overall shape of the aCNH_2 group does not change considerably on gaining a proton, the number of spherical segments $\nu_{\text{aCNH}_3^+}^*$, shape factor $S_{\text{aCNH}_3^+}$, segment diameter $\sigma_{\text{aCNH}_3^+ - \text{aCNH}_3^+}$, attractive and repulsive exponents of the Mie potential $\lambda_{\text{aCNH}_3^+ - \text{aCNH}_3^+}^{\text{a}}$ and $\lambda_{\text{aCNH}_3^+ - \text{aCNH}_3^+}^{\text{r}}$, respectively, of the aCNH_2 group are used to model aCNH_3^+ , with the addition of a positive charge to the group ($z_k = +1$). The Born diameter, $\sigma_{\text{aCNH}_3^+ - \text{aCNH}_3^+}^{\text{Born}}$, is calculated as a 7% increase in the diameter of the ion $\sigma_{\text{aCNH}_3^+ - \text{aCNH}_3^+}$ according to the proposition of Rashin and Honig [219], i.e., $\sigma_{\text{aCNH}_3^+ - \text{aCNH}_3^+}^{\text{Born}} = 1.07\sigma_{\text{aCNH}_3^+ - \text{aCNH}_3^+}$. The like dispersion energy between aCNH_3^+ groups $\varepsilon_{\text{aCNH}_3^+ - \text{aCNH}_3^+}$ is calculated using Equation (2.40) with the polarisability $\alpha_{0,\text{aCNH}_3^+}$ and ionisation potential $I_{\text{aCNH}_3^+}$ as reported in Table 4.5. The aCNH_3^+ group is modelled with three H proton-donor sites corresponding to each of the hydrogen atoms [305, 306]. No self-association between aCNH_3^+ groups is considered as no e-type sites are assigned for the ion. The aCNH_3^+ group parameters determined in this section are reported in Table 4.22.

Table 4.22: SAFT- γ Mie like parameters developed in Section 4.3.7. ν_k^* is the number of segments making up group k , S_k is the shape factor, λ_{kk}^{r} and λ_{kk}^{a} are the repulsive and attractive exponents of the Mie potential, respectively, σ_{kk} is the segment diameter, $\sigma_{kk}^{\text{Born}}$ is the Born diameter, ε_{kk} is the dispersion energy between two k groups, $N_{\text{ST},k}$ is the number of association site types on group k , with $n_{\text{H},k}$ representing the association sites of type H.

Group k	ν_k^*	S_k	$\sigma_{kk}/\text{\AA}$	$\sigma_{kk}^{\text{Born}}/\text{\AA}$	λ_{kk}^{r}	λ_{kk}^{a}	$(\varepsilon_{kk}/k_{\text{B}})/\text{K}$	$N_{\text{ST},k}$	$n_{\text{H},k}$
aCNH_3^+	1	0.71518	3.7408	4.0027	17.065	6.0000	397.63	1	3

To characterise the unlike interaction parameters between aCNH_3^+ and H_2O , a parameter-estimation procedure is implemented. A total of four unlike interaction parameters need to be determined, namely: $\varepsilon_{\text{aCNH}_3^+ - \text{H}_2\text{O}}$, $\lambda_{\text{aCNH}_3^+ - \text{H}_2\text{O}}^{\text{r}}$, $\varepsilon_{\text{aCNH}_3^+ - \text{H}_2\text{O}, \text{H-e}_1}^{\text{HB}}$ and $K_{\text{aCNH}_3^+ - \text{H}_2\text{O}, \text{H-e}_1}$. In our model, the repulsive exponent $\lambda_{\text{aCNH}_3^+ - \text{H}_2\text{O}}^{\text{r}}$ is calculated by combining rules (cf. Equation (2.37)) and the remaining parameters are estimated using experimental liquid-phase density

data for anilinium chloride + water mixtures. In the SAFT- γ Mie framework, anilinium chloride is modelled as a strong electrolyte which is fully dissociated into anilinium ($5 \times \text{aCH}$ and $1 \times \text{aCNH}_3^+$) and chloride ($1 \times \text{Cl}^-$) ions (cf. Figure 4.4 (d)). We note, that the unlike dispersion energy between the aCNH_3^+ group and the aCH and Cl^- groups are calculated by the combining rules given in Equations (2.36) and (2.40), respectively. The unlike parameters developed in this section can be found in Tables 4.23 and 4.24, and are also included in Tables 4.3 and 4.4 together with the remaining unlike interaction parameters used in this chapter.

Table 4.23: Group dispersion interaction energies ε_{kl} and repulsive exponent λ_{kl}^r developed in Section 4.3.7 for use with the SAFT- γ Mie equation of state. The unlike segment diameter σ_{kl} is calculated from the combining rule in Equation (2.35). CR indicates that combining rules are used to obtain λ_{kl}^r using Equation (2.37).

Group k	Group l	$(\varepsilon_{kl}/k_B)/\text{K}$	λ_{kl}^r
aCNH_3^+	H_2O	1382.5	CR

Table 4.24: Group association energies $\varepsilon_{kl,ab}^{\text{HB}}$ and bonding volume $K_{kl,ab}$ developed in Section 4.3.7 for use with the SAFT- γ Mie equation of state.

Group k	Site a on group k	Group l	Site b on group l	$(\varepsilon_{kl}^{\text{HB}}/k_B)/\text{K}$	$K_{kl,ab}/\text{\AA}^3$
aCNH_3^+	H	H_2O	e_1	2550.8	501.61

The SAFT- γ Mie calculations using the newly developed aCNH_3^+ group and its optimised interaction parameters with H_2O are compared to the experimental liquid-phase density data of anilinium chloride and water mixtures [307] in Figure 4.11. The model developed leads to an excellent agreement to experimental data at all temperatures, including those not considered in the parameter estimation procedure (indicated by open symbols in the figure). The high-level of accuracy obtained by the model is also demonstrated by the low deviations reported in Table 4.25. The AADs follow the same trend as the %AADs for the liquid-phase density of the mixture, with the highest AAD and %AAD found to be at 333.2 K corresponding to 3.95 kg m^{-3} and 0.4%, respectively.

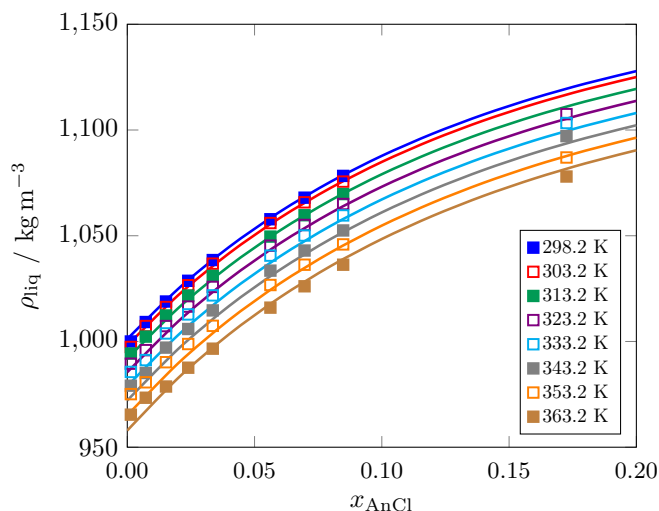


Figure 4.11: Liquid-phase density ρ_{liq} of anilinium chloride (AnCl) + water mixtures at $P = 0.101$ MPa and the given temperatures. The curves correspond to the SAFT- γ Mie calculations, the symbols represent the experimental data [307] used for parameter estimation (filled) and those not used (unfilled).

Table 4.25: Overview of the accuracy of the SAFT- γ Mie GC approach using the newly developed aCNH_3^+ group and its unlike interaction parameters $\text{aCNH}_3^+ - \text{H}_2\text{O}$ in the calculation of the liquid-phase density ρ_{liq} for binary mixtures of anilinium chloride and water. The experimental data considered in the parameter estimation and model validation are taken from the references indicated.

Parameter estimation						
T / K	P / MPa	N_X^D	%AAD ρ_{liq}	AAD ρ_{liq} / kg m ⁻³	ref	
298.2	0.101	8	0.085	0.87	[307]	
312.2	0.101	8	0.087	0.89	[307]	
343.2	0.101	9	0.36	3.61	[307]	
363.2	0.101	9	0.27	2.71	[307]	
Model Validation						
T / K	P / MPa	N_X^D	%AAD ρ_{liq}	AAD ρ_{liq} / kg m ⁻³	ref	
303.2	0.101	8	0.066	0.67	[307]	
323.2	0.101	9	0.20	2.01	[307]	
333.2	0.101	9	0.38	3.95	[307]	
353.2	0.101	9	0.37	3.70	[307]	

4.4 Solubility prediction

The solubility of procaine in aqueous solutions can be predicted using the SAFT- γ Mie GC approach by incorporating the newly developed groups and their unlike interactions (Tables

4.2, 4.3 and 4.4) combined with group interactions obtained in previous work [35–37, 39, 42, 43, 150, 167]. First, the solubility of neutral procaine (intrinsic solubility) and that of the ionised forms of the API are calculated using the model and compared to experimental data in Figure 4.12. The intrinsic solubility of the API is determined using Equation (4.8) together with the melting properties reported in Table 4.26. To calculate the solubility of PC^+ and PC^{2+} , $\text{p}K_{\text{a1,PC}}$ (Equation (4.13)) and $\text{p}K_{\text{a2,PC}}$ (Equation (4.15)), respectively, are needed. The equilibrium constants are reported in literature only at 298.15 K (cf. Table 4.26); however, the temperature dependence of the equilibrium constants is required to calculate the solubility of the ions at other temperatures. Hence, the van't Hoff equation (cf. Equation (3.32)) is used together with the isobaric heat capacity, c_p , polynomial of procaine in the liquid phase [281],

$$c_p/\text{JK}^{-1} \text{ mol}^{-1} = -37.027 + 1.38073(T/\text{K}) - 8.9608 \times 10^4 (T/\text{K})^2 + 2.1243 \times 10^{-7} (T/\text{K})^3 \\ - 2.47988 \times 10^{-20} (T/\text{K})^4, \quad (4.22)$$

which is valid in the range of temperatures $200 \leq T/\text{K} \leq 1000$, to evaluate $\text{p}K_{\text{a1,PC}}(T)$ and $\text{p}K_{\text{a2,PC}}(T)$.

Table 4.26: Properties of procaine at 298.15 K used for modelling the solid–liquid equilibrium (SLE) and pH-solubility profile, including: molar melting enthalpies Δh^{fus} ; melting temperature T^{fus} ; dissociation constants $\text{p}K_{\text{a}}$; and the solubility product $K_{\text{sp,A}\nu_{\text{A}}\text{B}\nu_{\text{B}}}$. * indicates properties calculated in the current work. The value of the water dissociation constant $\text{p}K_{\text{w}}$ is added for completeness.

Property	Procaine	Procaine hydrochloride	Ref.
$\Delta h^{\text{fus}}/\text{kJ mol}^{-1}$	26.1	-	[259]
T_{m}/K	334.15	-	[259]
$\text{p}K_{\text{a1,PC}}$	9.0	-	[258]
$\text{p}K_{\text{a2,PC}}$	2.5	-	[258]
$K_{\text{sp,PC-HCl}}/\text{M}^2$	-	10.87	*
$\text{p}K_{\text{w}}$	-	-	14.00 [228]

The predicted solubilities of neutral and ionic forms of procaine are shown in Figure 4.12. For the intrinsic solubility of neutral procaine, the SAFT- γ Mie predictions are compared to three sets of experimental data [308–310]. It is apparent that there are inconsistencies between the experimental data, perhaps due to the different measurement techniques used and to standard

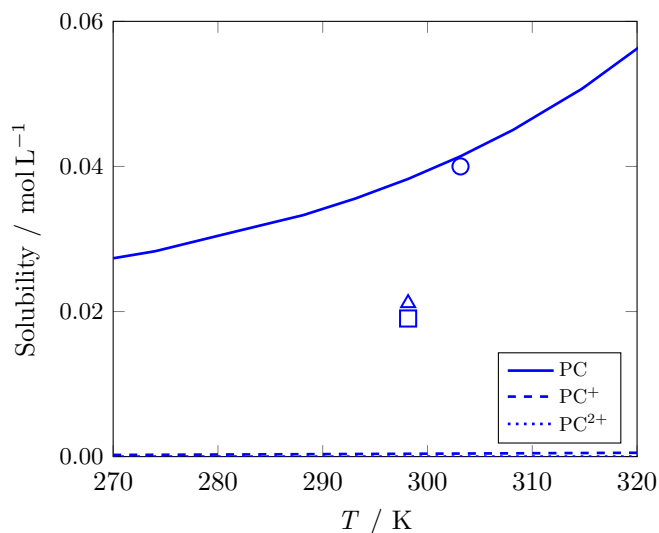


Figure 4.12: Constant-pressure solid–liquid equilibria (at $P = 0.101$ MPa) of the molecular and ionised forms of procaine as a function of temperature. The symbols represent experimental data (triangle [308], circle [309] and square [310]) and the continuous, dotted and dashed curves represent predictions using the SAFT- γ Mie approach for neutral procaine (PC), procaine monocation (PC^+) and procaine dication (PC^{2+}), respectively.

experimental error. Our predictions of the intrinsic solubility are of the same order of magnitude as experiments and agree particularly well with the data by Yalkowsky et al. [309]. In the case of the ionised forms of procaine, the solubility is clearly much lower than the neutral API. The solubility of PC^+ is in the order of magnitude of 10^{-4} mol/L, increasing with temperature (not visible in the figure), while the solubility of the dication PC^{2+} is several orders of magnitude smaller and is assumed to be negligible.

We now turn our attention to the prediction of the pH-solubility profile of procaine in water at 298.15 K. In order to model the complete profile, including the salt precipitation at low pH values, $\text{p}K_{\text{a}1,\text{PC}}$ and $\text{p}K_{\text{a}2,\text{PC}}$ of the base, $K_{\text{sp}1,\text{PC-HCl}}$ of the salt and $\text{p}K_{\text{w}}$ of water, are required, to be used in Equations (4.13), (4.15), (4.20) and (3.18), respectively. The values of these constants are listed in Table 4.26.

The predicted pH-solubility profile of procaine in water at 298.15 K is compared to experimental data [260, 308, 310–312] in Figure 4.13. NaOH is used to increase the pH of the solution above $\text{p}K_{\text{a}1,\text{PC}}$ to a pH of 14.0, while HCl is used as a buffering agent to lower the pH of the

solution to 1.0 and is also the source of Cl^- counter ions in the solution, leading to the formation of the hydrochloride salt of procaine at low pH [260]. It is evident in the figure that our SAFT- γ Mie predictions are in good agreement with experimental data. The intrinsic solubility plateau at high pH, as well as the sharp increase in solubility at pH values lower than $\text{p}K_{\text{a}1,\text{PC}}$, caused by the addition of Cl^- counter ions, are captured by the model. We note the scarcity of experimental pH-dependent solubility data for procaine in the literature, perhaps due to the complex ionisation of the API leading to difficulties in solubility measurements; emphasizing the importance of predictive tools such as the SAFT- γ Mie GC approach used here.

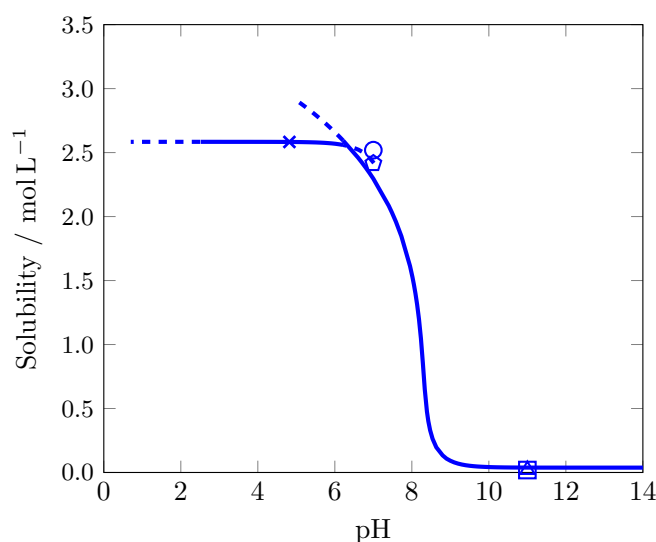


Figure 4.13: SAFT- γ Mie group contribution approach predictions of the pH-solubility profile compared to experimental data of aqueous solutions of procaine at ambient pressure ($P = 0.101$ MPa) and $T = 298.15$ K using HCl and NaOH as buffering agents. The symbols represent experimental data (triangle [308], square [310], circle [311], pentagon [312] and cross [260]) and the continuous curves represent predictions using the SAFT- γ Mie approach.

The hydrochloride salt of monoprotonated procaine begins to precipitate when $K_{\text{sp}1,\text{PC-HCl}}$ is reached. The thermodynamic solubility product $K_{\text{sp}1,\text{PC-HCl}}$ of PC-HCl is calculated based on the salt solubility measurements reported by Guerrieri et al. [260], using Equation (4.20); the calculated value is reported in Table 4.26. The pH value at which solid procaine and its hydrochloride salt can coexist, $\text{pH}^{\text{max},1}$, obtained from our SAFT- γ Mie prediction is 6.21 (cf. Figure 4.13). We note that, $K_{\text{sp}2,\text{PC}(\text{HCl})_2}$ is not used in our model due to the absence of experimental data for the solubility of the dihydrochloride salt of procaine. It is assumed that

this salt does not precipitate at $\text{pH} > \text{p}K_{\text{a}2,\text{PC}}$, which corresponds to the range of pH where the concentration of PC^{2+} is very low, thus $\text{pH}^{\text{max},2}$ is not determined. Refinement of the model to include the solubility product of the dihydrochloride salt of procaine $K_{\text{sp}2,\text{PC}(\text{HCl})_2}$, if it forms, should be considered, when experimental data become available.

According to the schematic pH-solubility profile for a diprotic base in Figure 4.2 (b) and to the profiles of other diprotic weak bases in literature [253–257], the pH-solubility profile of procaine in Figure 4.13 is expected to feature another sharp increase in solubility at $\text{pH} < \text{p}K_{\text{a}2,\text{PC}}$, but this is not predicted by our model. In order to model the increase in solubility at $\text{pH} < \text{p}K_{\text{a}2,\text{PC}}$, the properties of the salt in the solid state should be accounted for, as it is the precipitating solid in this pH range, i.e., using Equation (4.8) to calculate the intrinsic solubility of the salt. To use the solubility equation (Equation (4.8)), the melting properties of the salt are required in addition to the SAFT- γ Mie groups representing PC-HCl, which are needed to calculate the activity coefficient of the salt $\gamma_{\text{PC-HCl}}$ using the GC approach. The SAFT- γ Mie groups required for PC-HCl are: ($2 \times \text{CH}_3$, $4 \times \text{CH}_2$, $1 \times \text{N}$, $4 \times \text{aCH}$, $1 \times \text{aCCOO}$, $1 \times \text{aCNH}_2$ and $1 \times \text{HCl}$) (cf. Figure 4.14). However, the HCl group and its unlike interaction parameters with other groups are not currently available in the SAFT- γ Mie groups database and need to be developed. We note that the SAFT- γ Mie groups database contains the parameters for the H_3O^+ and Cl^- ionic groups, which have been used previously to model the fully dissociated HCl in an aqueous solution [44, 150]; these groups cannot be used to model the precipitating PC-HCl solid, as the HCl group does not dissociate in the solid phase, therefore, a neutral group for HCl is required for this model. Because of this limitation in the model, the pH-solubility profile of procaine in Figure 4.13 can only be predicted reliably at $\text{pH} > \text{p}K_{\text{a}2,\text{PC}}$. The development of the complete pH-solubility profile of procaine in water, including at $\text{pH} < \text{p}K_{\text{a}2,\text{PC}}$, will be the subject of future work.

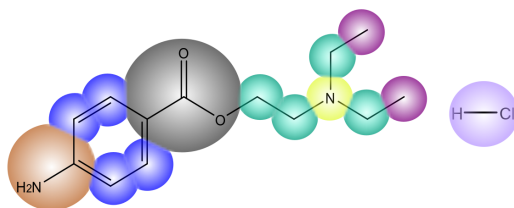


Figure 4.14: Chemical structure of procaine hydrochloride (PC-HCl) with its corresponding SAFT- γ Mie representation. Each of the coloured spheres corresponds to a different group.

4.5 Conclusions

In this chapter, we extend the application of the SAFT- γ Mie approach to model the solubility and pH-solubility profiles of diprotic APIs. These APIs are known to exhibit significant solubility enhancement due to changes in the pH of the solution which eventually lead to salt formation.

We have demonstrated the applicability of the SAFT- γ Mie GC equation of state by predicting the pH-solubility profile of procaine, a weakly basic diprotic API. In order to represent the molecular and ionised forms of procaine within the SAFT- γ Mie framework, we have expanded the matrix of groups by introducing three new groups, namely: the aCCOO, aCNH₂ and aCNH₃⁺ groups, in addition to their unlike interactions aCNH₂-aCCOO, aCNH₂-aCH, aCNH₂-CH₂, aCNH₂-CH₃, aCCOO-CH₂ and aCCOO-CH₃. We use previously reported experimental data for the simplest molecules containing the new groups (i.e., aniline, alkyl benzoates, alkyl 4-aminobenzoates and anilinium chloride) to estimate the like and unlike parameters of the groups and to assess the accuracy of the predictions of the models developed. Furthermore, to model the aqueous solubility of procaine as a function of temperature and pH, we develop the unlike interaction parameters between the new groups and the H₂O (water) group, using experimental data for binary mixtures of each of aniline, alkyl benzoates and anilinium chloride with water.

Using the newly developed groups together with group interaction parameters obtained in

previous work, we first present the SAFT- γ Mie intrinsic solid-liquid equilibria (SLE) prediction of procaine in water. The solubility predictions are shown to be in good agreement with experimental data, which demonstrates the transferability of the SAFT- γ Mie group interactions, considering that no data related to procaine were used to obtain the group interaction parameters.

A new thermodynamic framework to predict the pH-solubility profile of procaine has also been developed by extending the model presented for monoprotic APIs in Chapter 3. The pH-solubility profile of procaine in aqueous solution at 298.15 K using HCl and NaOH as buffering agents is obtained by combining experimental values for the equilibrium constants ($pK_{a1,PC}$ and $pK_{a2,PC}$ of the API, K_w of water, and $K_{sp1,PC-HCl}$ of the hydrochloride salt) with molecular models using the SAFT- γ Mie GC EoS. Very good agreement with experimental data is obtained for the pH-dependent solubility of procaine at $pH > 2.5$ (i.e., $pH > pK_{a2,PC}$), confirming that the SAFT- γ Mie approach is a useful predictive tool for the pH-solubility profiles of APIs.

We note that, as a diprotic API, procaine can, in theory, form disalts. However, when using chloride counter ions, only the procaine hydrochloride salt is reported to form. In our work, we assume that the dihydrochloride salt does not form, and $K_{sp2,PC-HCl}$ is not calculated. Moreover, the pH-solubility profile of procaine at $pH < 2.5$ was not predicted in our work, due to unavailability of SAFT- γ Mie group interactions that are required to model the intrinsic solubility of procaine hydrochloride, which is the precipitating solid in this pH range. This should be the subject of future work.

The findings from our current work demonstrate the transferability of the SAFT- γ Mie groups and the applicability of the approach for modelling mixtures of ionisable APIs and predicting their thermodynamic properties, in particular, the aqueous intrinsic solubility and pH-solubility profiles of diprotic APIs. Several refinements should be considered in future studies for modelling the pH-solubility profile of procaine, when experimental data become available, including: investigating the formation of the dihydrochloride salt of procaine as well as developing the group interactions required to model the intrinsic solubility of procaine hydrochloride.

Chapter 5

Thermodynamic modeling of the nature of speciation and phase behavior of binary and ternary mixtures of formaldehyde, water and methanol

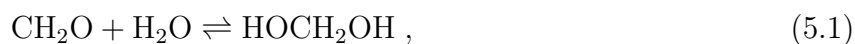
After having successfully applied the SAFT- γ Mie group-contribution (GC) approach for the prediction of the phase behaviour and pH-solubility profile of monoprotic and diprotic small molecule APIs in Chapters 3 and 4, in this chapter, we demonstrate the applicability of the GC approach for the phase behaviour prediction of more-complex mixtures of relevance to the pharmaceutical industry. In particular, we study the fluid-phase equilibrium of mixtures containing formaldehyde, a common degradation agent produced by some pharmaceutical excipients during the storage of drug products.

5.1 Introduction

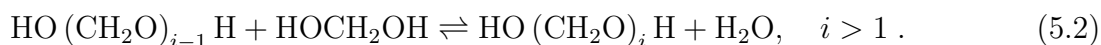
Formaldehyde is an important industrial chemical. There are numerous applications of formaldehyde due to its versatility and ability to react with organic and inorganic compounds [313]. In

the pharmaceutical industry, formaldehyde is used as a raw material for the synthesis of various solvents such as polyethylene glycol (PEG) and glycerol [314–316]. Due to its high reactivity however, formaldehyde can lead to the rapid degradation of pharmaceutical drug products during storage, significantly reducing the shelf-life of the drug [317]. In addition to being present as an impurity in air, it has been shown that formaldehyde can be generated by some excipients (e.g., magnesium stearate and lactose) during storage [318], negatively impacting the stability of drugs.

Pure monomeric formaldehyde is a colourless gas with a pungent smell, and is highly reactive. Due to its high reactivity, it is usually produced, stored, sold, and processed in the form of aqueous solutions. Methanol is sometimes added to enhance the stability of formaldehyde, to reduce the amount of water in solution, or to prevent the precipitation of polymers which may occur at low temperatures [313, 314]. In aqueous and methanolic solutions, formaldehyde reacts with the solvents to form a variety of reaction products, which only exist in solution and cannot be isolated. The reaction of formaldehyde with water leads to the formation of methylene glycol (MG₁),



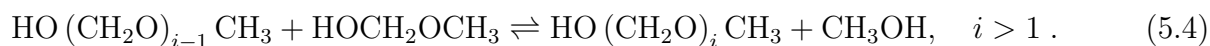
and poly(oxymethylene) glycols (MG_{*i*}),



Similarly, the reaction between formaldehyde and methanol leads to the formation of hemiformal (HMF₁) ,



and poly(oxymethylene) hemiformals (HMF_{*i*}),



In formaldehyde + water + methanol mixtures, the four reactions take place simultaneously. These reactions can occur in the liquid and vapour phase, depending on the operating conditions

used.

In aqueous and methanolic solutions, formaldehyde is predominantly chemically bound to the solvents at all conditions. Monomeric formaldehyde (CH_2O) makes up less than 0.04 wt%, even in highly concentrated solutions of formaldehyde [314]. Therefore, the chemical reactions (5.1)–(5.4) have a significant influence on the properties of the solutions and need to be considered in the thermodynamic modelling of these solutions. Although, other chemical reactions may take place in formaldehyde mixtures with water and methanol, they have a negligible impact on the properties of the solutions and are not considered in the current work. The reader is referred to the work of Walker [314] and Maurer [319] for more details on these reactions.

A large number of experimental measurements can be found in the literature for the vapour-liquid equilibria (VLE) of formaldehyde + water [320–326], formaldehyde + methanol [320, 326, 327] and of ternary formaldehyde + water + methanol mixtures [319, 320, 325, 326, 328–330]. A common experimental method used by several authors [319, 320, 328, 331], is the thin-film evaporator, in which a rotating coil is used to spread the liquid feed on the inner surface of a tube, surrounded by a heating jacket, causing the liquid to evaporate. After the separation of the liquid and vapour phases, the vapour phase is condensed. The two liquids are then collected in vials and analyzed. To ensure that phase equilibrium is reached, a small evaporation ratio (the ratio of the volume of the condensed vapour to the volume of the unevaporated liquid) and long residence time is maintained in the experiments. To analyze the composition of the coexisting phases, gas chromatography is used to determine the water and/or methanol content, while the formaldehyde content is determined using the sodium sulfite method [314]. At low temperatures (< 320 K), the VLE of formaldehyde mixtures is significantly influenced by reaction kinetics and special care must be taken to ensure that chemical phase equilibrium is attained. Hasse and Maurer [320] used a carrier-gas saturation technique to measure the VLE of formaldehyde + water, formaldehyde + methanol and formaldehyde + water + methanol mixtures. In this experimental method, an inert carrier-gas at a low flow rate is passed through, and is saturated with the formaldehyde liquid solution. During saturation, only a small fraction of the liquid evaporates, and the chemical equilibrium of the solution remains unaffected. The evaporated vapour is separated from the carrier gas in a cooling trap and is then condensed.

Similar to the thin-film evaporator technique, the coexisting phases are analyzed using the gas chromatography and sodium sulfite methods to determine their compositions.

For most industrial applications of formaldehyde, modelling the vapour-liquid equilibrium of aqueous and methanolic solutions of formaldehyde, is of great importance. Extensive work on the thermodynamic modelling of the VLE of formaldehyde aqueous and methanolic solutions can be found in the literature. Perhaps one of the earliest modelling attempts of the VLE of formaldehyde mixtures is the physiochemical model presented by Maurer [319] in 1986 to describe the VLE of binary mixtures of formaldehyde + water and formaldehyde + methanol using the universal functional activity coefficient (UNIFAC) [18] model. In their model, experimental data were used to determine the necessary UNIFAC parameters, the thermodynamic equilibrium constants and the vapour pressure of methylene glycol (or hemiformal) for each mixture. The performance of Maurer's model has been improved and its range of application extended, based on new experimental VLE data of formaldehyde mixtures, by several authors over the years [320, 331, 332]. Specifically, an extension of the Maurer model to represent the ternary formaldehyde + water + methanol mixture was presented by Albert et al. [328]. Using a different approach, Brandani and coworkers [333] modelled the VLE of ternary mixtures of formaldehyde + water + methanol using the Wilson [334] equation to calculate the activity coefficients of the three species. Later on, Brandani et al. [335] modelled the isothermal VLE of formaldehyde + water and formaldehyde + methanol mixtures using the universal quasi-chemical (UNIQUAC) [336, 337] equation to calculate the activity coefficients of the components in the mixtures. In addition to the unknown UNIQUAC parameters, the thermodynamic equilibrium constants for the reactions in the liquid phase were fitted to experimental VLE data. The Henry's constant of formaldehyde in each solvent was also required for their model. This model was later extended to represent the ternary mixtures of formaldehyde + water + methanol by Brandani et al. [338]. The models discussed so far, considered the vapour phase to be an ideal mixture of formaldehyde, water, methanol, methylene glycol and hemiformal only. The concentration of higher oligomers produced by Reactions (5.2) and (5.4) was assumed to be negligible in the vapour phase. The study by Brandani et al. [339] was the first to account for dioxymethylene hemiformal (HMF_2) in the vapour phase for a binary mixture of formaldehyde

+ methanol, and they showed improved predictions, compared to previous studies, of the VLE of binary and ternary mixtures of formaldehyde, water and methanol [339]. This suggests that, perhaps, accounting for the oligomerization reactions in the vapour phase, as well as in the liquid phase, can result in better models of the VLE of formaldehyde mixtures.

The equilibrium distribution of formaldehyde and its products with water (Reactions (5.1) and (5.2)) and methanol (Reactions (5.3) and (5.4)) influences the thermodynamic phase equilibrium, chemical reaction kinetics and transport properties. Therefore, quantitative information on the distribution of the reactions' products in aqueous and methanolic solutions of formaldehyde is crucial for industrial purposes and has been the focus of several research studies over the years [331, 332, 340–345]. Spectroscopic techniques can be used to determine the distribution and concentration of the species. UV/Vis spectroscopy can be used to study the formation of MG_1 and HMF_1 from monomeric formaldehyde (Reactions (5.1) and (5.3), respectively) [340, 341]. More commonly, however, high-frequency nuclear-magnetic-resonance (NMR) spectroscopy is used to determine the formation of MG_1 and HMF_1 as well as the formation of the oligomers formed in Reactions (5.2) and (5.4). In NMR spectroscopy, separate peaks are assigned from CH_2 groups in different MG_i and HMF_i in the mixture [342]. The peak areas from the NMR analysis are converted to concentrations assuming they are proportional to the mole numbers of the CH_2 groups in solution. Multiple studies on the speciation analysis for the binary mixture of formaldehyde + water in the liquid phase, using NMR spectroscopy, can be found in the literature [331, 332, 343, 344]. Only the work by Hahnenstein et al. [343] studied the species distribution for the binary mixture of formaldehyde + methanol in the liquid phase using the NMR technique. Maiwald et al. [345] later extended the work to investigate the distribution of species in the ternary mixture of formaldehyde + water + methanol in the liquid phase using also NMR spectroscopy. Note that, no peak could be assigned for monomeric formaldehyde in these NMR studies, due to its extremely low concentration in the liquid phase; hence, the true concentration of monomeric formaldehyde was considered to be negligible.

In an attempt to model the distribution of the species in a binary mixture of formaldehyde and water, Albert et al. [332] used NMR spectroscopy data for formaldehyde + water mixtures [343], in addition to experimental VLE data, to fit: the reaction equilibrium constants

(for the formation of the oligomers) and the unknown UNIFAC parameters in their work. Their model [332] provided a reliable description of the formation of different MG_i in liquid mixtures of formaldehyde and water for a temperature range of 290–420 K. The extended model by Albert et al. [328] showed a good agreement with NMR speciation data of Hahnenstein et al. [343] for the liquid mixture of formaldehyde + methanol. Maiwald et al. [345] used the model by Albert et al. [328] to predict the distribution of the reaction products in a ternary liquid mixture of formaldehyde, water and methanol. For ternary mixtures, this model [328] provided only a qualitative agreement to NMR experimental data [345]. The most significant deviations were observed for the concentrations of MG_i .

The literature discussed so far on the modelling attempts for the VLE and the nature of speciation in formaldehyde mixtures, follows a chemical approach to deal with the occurrence of chemical reactions and the formation of new species in each phase. In such methods, the reaction products are defined *a priori* and are treated explicitly within the thermodynamic model. An alternative to this methodology is the physical approach of dealing with reactions, in which the reaction products are treated implicitly and are considered to be aggregates of the reactants. In physical approaches, the reaction products do not need to be defined explicitly or known in advance, and their formation is driven by the presence of strong intermolecular interactions. It has been shown that physical approaches based on perturbation theories such as the statistical associating fluid theory (SAFT) [33, 34] provide equivalent results to the chemical or quasi-chemical theories for the formation of hydrogen bonding between species in a mixture [346]. In SAFT approaches, aggregates (reaction products) may form due to association interactions between the species, similar to the association approach used to model hydrogen bonding. The theory of chemical association within SAFT is due to a series of papers by Wertheim [45–48] in which the contribution to the Helmholtz free energy, due to association of a fluid composed of associating monomers, is evaluated. The short-range directional forces are accounted for by specifying off-center association sites on the monomers, which interact via a potential function. The parameters obtained in the physical theory to describe the association interaction between the reacting species can be related to the chemical reaction equilibrium constant in chemical theories [346]. An advantage of physical approaches to modelling reac-

tions over chemical approaches is that knowledge of the reaction equilibrium constants or the concentrations of the reaction products is not necessary for model development, which reduces the reliance on experimental data.

The physical approach to modelling reactions has been successfully implemented for carbon-capture processes within the SAFT framework, in previous work [347–350]. Mac Dowell et al. [347] and Rodríguez et al. [348] used the statistical associating fluid theory for potentials of variable range (SAFT-VR) [65, 66], with square-well (SW) potentials, to describe the phase behaviour of reactive aqueous mixtures of carbon dioxide (CO_2) with alkanolamines. SAFT-VR SW is a homonuclear version of the SAFT theory in which intermolecular parameters are used to represent whole molecules rather than specific functional groups. An accurate representation of the complex reactions that occur in these mixtures was provided [347, 348] by adding two association sites on the CO_2 molecule which can only interact with alkanolamines to form tightly bound aggregates, implicitly representing the formation of the main reaction products (carbamate and bicarbonate). Additionally, Rodríguez et al. [348] studied the degree of speciation in ternary mixtures of CO_2 , monoethanolamine (MEA) and water, providing an excellent agreement with experimental data of the true mole fraction of carbamate and bicarbonate for various CO_2 loadings in the mixture, without using any experimentally measured reaction equilibrium constants. To increase the predictive capability of these models, the SAFT- γ SW [141, 142] and the SAFT- γ Mie [35–38, 148] group-contribution approaches have been used to model the phase behaviour of mixtures containing CO_2 with different amines and solvents, using a physical approach to represent the reactions taking place [349, 350]. These physical models delivered an accurate representation of the fluid-phase behaviour and chemical speciation in CO_2 -containing mixtures of relevance to CO_2 capture.

In the current work, the SAFT- γ Mie GC approach is used to predict the VLE in binary and ternary formaldehyde mixtures with water and methanol. In particular, we follow a physical approach to account for the oligomerization reactions (Reactions (5.1)–(5.4)) that occur in the liquid and vapour phases, as part of the thermodynamic modelling of the fluid-phase behaviour of these mixtures. This is done by adding a reactive association site on the group representing formaldehyde within the SAFT- γ Mie framework. In our model, the vapour phase is considered

to be a mixture of real gases in which there is no limit on the length of oligomers that may form. Additionally, we investigate the nature of speciation in aqueous and methanolic mixtures of formaldehyde and predictively quantify the distribution of oligomers (reaction products) in various mixtures.

In the following section we present the thermodynamic modelling using the SAFT- γ Mie GC approach relevant to mixtures containing formaldehyde with water and methanol, including: a brief background on the SAFT- γ Mie theory, property calculations using the equation of state (EoS), the molecular models developed, and the parameter estimation strategy used within the SAFT- γ Mie GC approach. In Section 5.3, the phase equilibria calculations for pure formaldehyde, and its binary and ternary mixtures with water and methanol, are discussed. Following this, in Section 5.4, we present the methodology used to determine the distribution of reaction products in mixtures of formaldehyde, and the corresponding SAFT- γ Mie predictions of the speciation taking place in formaldehyde + water, formaldehyde + methanol and formaldehyde + water + methanol mixtures. Final remarks are provided in Section 5.5.

5.2 Theory & thermodynamic modelling

5.2.1 SAFT- γ Mie theory

The theoretical background of the equation of state (EoS) is presented in detail in Chapter 2 and in refs [35–38, 148]. Here we provide only a brief summary of the models.

The SAFT- γ Mie [35–38] EoS is a group-contribution (GC) approach in which molecules are modelled as heteronuclear chains of fused-spherical segments with association sites, to account for short-range directional forces, if necessary. As discussed in Section 2.2, within the SAFT- γ Mie framework, a group k is characterised by a number ν_k^* of identical spherical segments, a shape factor S_k , the segment energy of interaction ε_{kk} , the segment diameter σ_{kk} , the exponents of the Mie potential λ_{kk}^r and λ_{kk}^a and, where relevant, the parameters describing any site-site association interactions, $\varepsilon_{kk,ab}^{\text{HB}}$, $K_{kk,ab}$, $n_{k,a}$, $n_{k,b}$, $\dots$, $n_{k,N_{\text{ST}},k}$. The interactions between two

groups k and l of different types are represented equivalently by the corresponding unlike parameters ε_{kl} , σ_{kl} , λ_{kl}^r , λ_{kl}^a , $\varepsilon_{kl,ab}^{\text{HB}}$ and $K_{kl,ab}$.

In the SAFT- γ Mie EoS the Helmholtz free energy A of a mixture is written as a sum of terms, based on a perturbation approach, in which

$$A = A^{\text{Ideal}} + A^{\text{Mono}} + A^{\text{Chain}} + A^{\text{Assoc}} + A^{\text{Ion}} + A^{\text{Born}} , \quad (5.5)$$

where A^{Ideal} corresponds to the ideal gas free energy, A^{Mono} is the effect of the interaction of monomeric segments through Mie potentials, A^{Chain} is the contribution to the free energy of the formation of chains of molecules from the fused Mie segments and A^{Assoc} accounts for molecular association through short-range directional interactions. The first four terms describe the classical, non-electrolyte SAFT contributions [35–38], and the presence of charged species is accounted for by the last two terms [149, 150].

Of particular importance to the current chapter is the association contribution term, which follows from the TPT1 expressions of Wertheim [31, 32, 45–48] and is given by

$$\frac{A^{\text{assoc}}}{Nk_{\text{B}}T} = \sum_{i=1}^{N_{\text{C}}} x_i \sum_{k=1}^{N_{\text{G}}} \nu_{k,i} \sum_{a=1}^{N_{\text{ST},k}} n_{k,a} \left(\ln X_{i,k,a} + \frac{1 - X_{i,k,a}}{2} \right) , \quad (5.6)$$

where N_{C} is the number of components in the mixture, N_{G} is the number of groups, x_i is the mole fraction of species i and $n_{k,a}$ is the number of sites of type a on group k . $X_{i,k,a}$ is the fraction of molecules of component i that are not bonded at a site of type a on group k , obtained from the solution of the mass-action equations as [31, 48, 141]

$$X_{i,k,a} = \left[1 + \rho \sum_{j=1}^{N_{\text{C}}} x_j \sum_{l=1}^{N_{\text{G}}} \nu_{l,j} \sum_{b=1}^{N_{\text{ST},l}} n_{l,b} X_{j,l,b} \Delta_{ij,kl,ab} \right]^{-1} . \quad (5.7)$$

Here ρ is the molecular density and $\Delta_{ij,kl,ab}$ characterizes the overall strength of the association between a site of type a on a group k of component i and a site of type b on a group l of

component j . This is approximated as [31]

$$\Delta_{ij,kl,ab} = F_{kl,ab} K_{kl,ab} I_{ij} , \quad (5.8)$$

where $F_{kl,ab} = \exp(\varepsilon_{kl,ab}^{\text{HB}}/k_{\text{B}}T) - 1$ and I_{ij} is a polynomial temperature-density correlation for the association integral for a Lenard-Jones monomer [37, 38].

5.2.2 Property calculations

The SAFT- γ Mie EoS, expressed in terms of the total Helmholtz free energy in Equation (5.5), is a function of temperature T , volume V , and a vector $\mathbf{N}$ of the mole numbers of all the components in the mixture. Other properties can be calculated from standard thermodynamic relations and phase equilibrium conditions [25, 168].

The thermodynamic phase equilibrium of an isolated multi-phase mixture of i components, can be obtained by imposing the equality of the temperature T , pressure P and chemical potential of each component μ_i in each phase, i.e., [168]

$$\begin{aligned} T^\alpha &= T^\beta , \quad \alpha \neq \beta \\ P^\alpha &= P^\beta , \quad \alpha \neq \beta \\ \mu_i^\alpha &= \mu_i^\beta , \quad \alpha \neq \beta \quad \text{and} \quad \forall i , \end{aligned} \quad (5.9)$$

for all the α and β phases considered, to obtain the compositions of each phase. The pressure and chemical potential of component i can be calculated from the Helmholtz free energy, respectively, as

$$P = - \left(\frac{\partial A}{\partial V} \right)_{\mathbf{N}, T} , \quad (5.10)$$

and

$$\mu_i = \left(\frac{\partial A}{\partial N_i} \right)_{\mathbf{N}_{j \neq i}, T, V} . \quad (5.11)$$

The relevant equilibria and chemical reactions for a mixture of formaldehyde, water and methanol are shown in Figure 5.1. The liquid and vapour phases consist of methanol (CH_3OH),

water (H_2O) and monomeric formaldehyde (CH_2O) as well as the reaction products of formaldehyde with these solvents: hemiformal (HMF_1), poly(oxymethylene) hemiformals (HMF_i), methylene glycol (MG_1) and poly(oxymethylene) glycols (MG_i). In our model, we account for the four chemical reactions given by Reactions (5.1)–(5.4) in both the liquid and vapour phases.

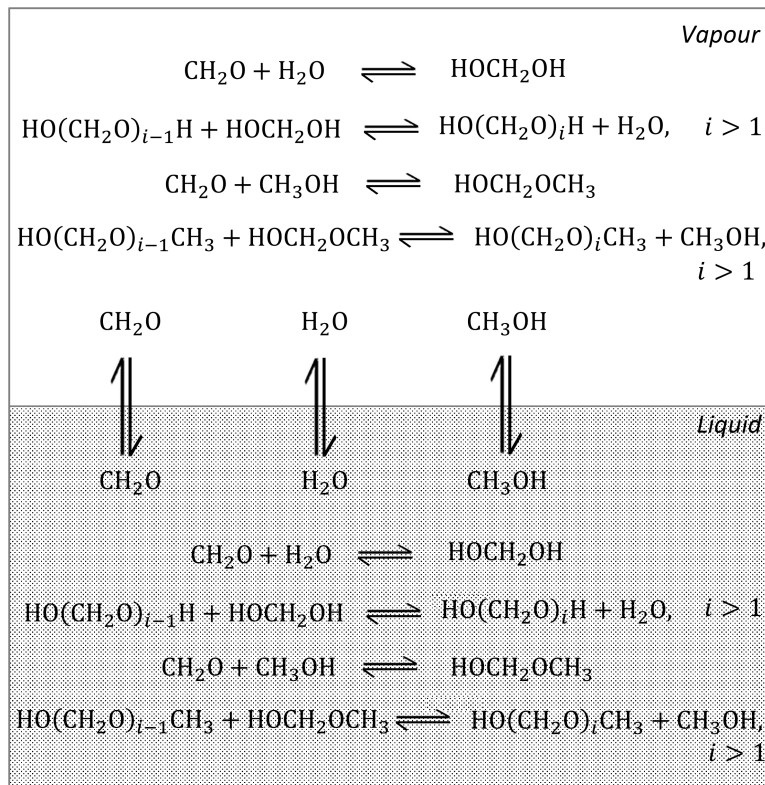


Figure 5.1: Schematic representation of the vapour-liquid equilibria in a ternary mixture of formaldehyde + water + methanol. The liquid and vapour phases consist of methanol (CH_3OH), water (H_2O) and monomeric formaldehyde (CH_2O). The reactions of formaldehyde with these solvents (Reactions (5.1)–(5.4)) are considered in the two phases, forming: hemiformal (HMF_1), poly(oxymethylene) hemiformals (HMF_i), methylene glycol (MG_1) and poly(oxymethylene) glycols (MG_i).

5.2.3 Molecular models

Using the SAFT- γ Mie GC approach, we need to characterize the groups that make up each of the compounds of interest. As formaldehyde, water and methanol are relatively small molecules; they are each modelled as a single molecular group. The groups required for modelling these compounds are presented in Figure 5.2. These are: $1 \times \text{CH}_2\text{O}$ for formaldehyde; $1 \times \text{CH}_3\text{OH}$ for methanol and $1 \times \text{H}_2\text{O}$ for water. Each group has two sites of type e which represent the

two electron lone-pairs of the oxygen atom, and sites of type H corresponding to each of the hydrogen atoms (that may participate in hydrogen bonding), such that, CH_3OH has one H site, H_2O has two H sites and CH_2O has two H sites. In our model, only sites of different types are allowed to interact, i.e., $\varepsilon_{kl,\alpha\alpha}^{\text{HB}} = 0$ for $k, l = \{\text{CH}_2\text{O}, \text{H}_2\text{O}, \text{CH}_3\text{OH}\}$ and $\alpha = \{e, H\}$.

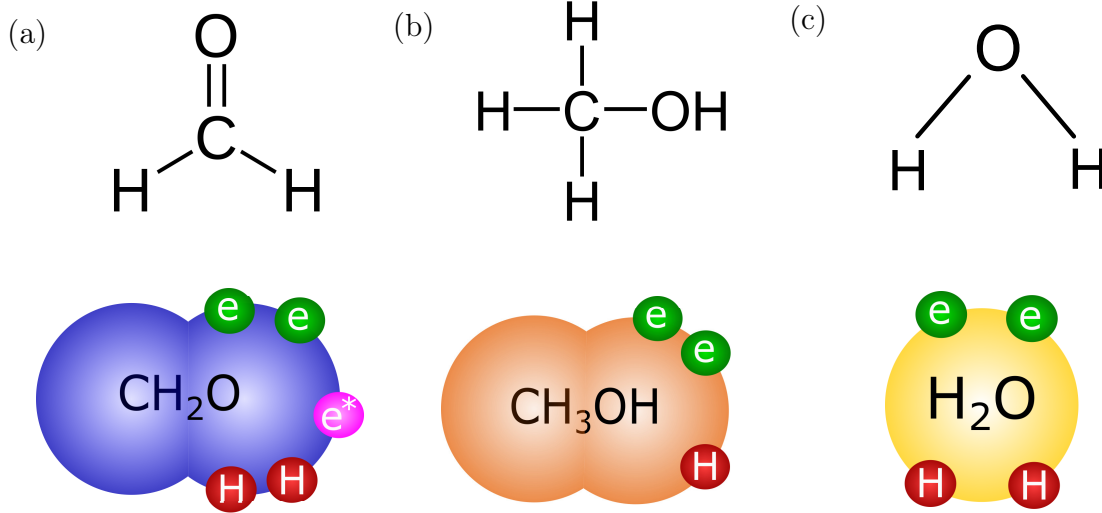


Figure 5.2: Chemical structures of (a) formaldehyde, (b) methanol and (c) water considered in the current work, together with their corresponding SAFT- γ Mie representation. Each of the large coloured spheres corresponds to a different group and the small spheres represent the association sites on the groups: H (red), e (green) and e^* (purple).

As discussed in Section 5.1, we use a physical approach in which the reaction products are considered to be aggregates of the reactants to model Reactions (5.1)–(5.4). This is possible within the SAFT- γ Mie framework by adding association (reactive) sites to mediate the formation of the reaction products. For a physical treatment of Reactions (5.1)–(5.4), we add a reactive e^* site to the CH_2O group which can only interact with the e sites on CH_3OH and H_2O (i.e., $\varepsilon_{\text{CH}_2\text{O}-\text{CH}_2\text{O},e^*-e^*}^{\text{HB}} = \varepsilon_{\text{CH}_2\text{O}-\text{CH}_2\text{O},e-e^*}^{\text{HB}} = \varepsilon_{\text{CH}_2\text{O}-\text{CH}_2\text{O},H-e^*}^{\text{HB}} = \varepsilon_{\text{CH}_2\text{O}-\text{H}_2\text{O},e^*-H}^{\text{HB}} = \varepsilon_{\text{CH}_2\text{O}-\text{CH}_3\text{OH},e^*-H}^{\text{HB}} = 0$). This reactive site is only active in a mixture but not in pure formaldehyde. Hence, the formation of methylene glycol (MG_1) (Reaction (5.1)) is mainly possible through the association between the e^* site on CH_2O and the e sites on H_2O (configuration I in Figure 5.3 (a)), and the formation of hemiformal (HMF_1) (Reaction (5.3)) mainly occurs through the association interaction between the e^* site on CH_2O and the e sites on CH_3OH (configuration I in Figure 5.3 (b)). It is important to note that, because the e sites on CH_3OH and H_2O also associate with other site types, and not exclusively to the e^* site on CH_2O , HMF_1 and MG_1 may form due to other association interactions in our model. MG_1 may also form by an association interaction

between the H site on CH_2O and the e site on H_2O (configuration II in Figure 5.3 (a)) or by the association interaction between the e site on CH_2O and the H site on H_2O (configuration III in Figure 5.3 (a)). Similarly, HMF_1 can form by the association between the H site on CH_2O and the e site on CH_3OH (configuration II in Figure 5.3 (b)) or between the e site on CH_2O and the H site on CH_3OH (configuration III in Figure 5.3 (b)).

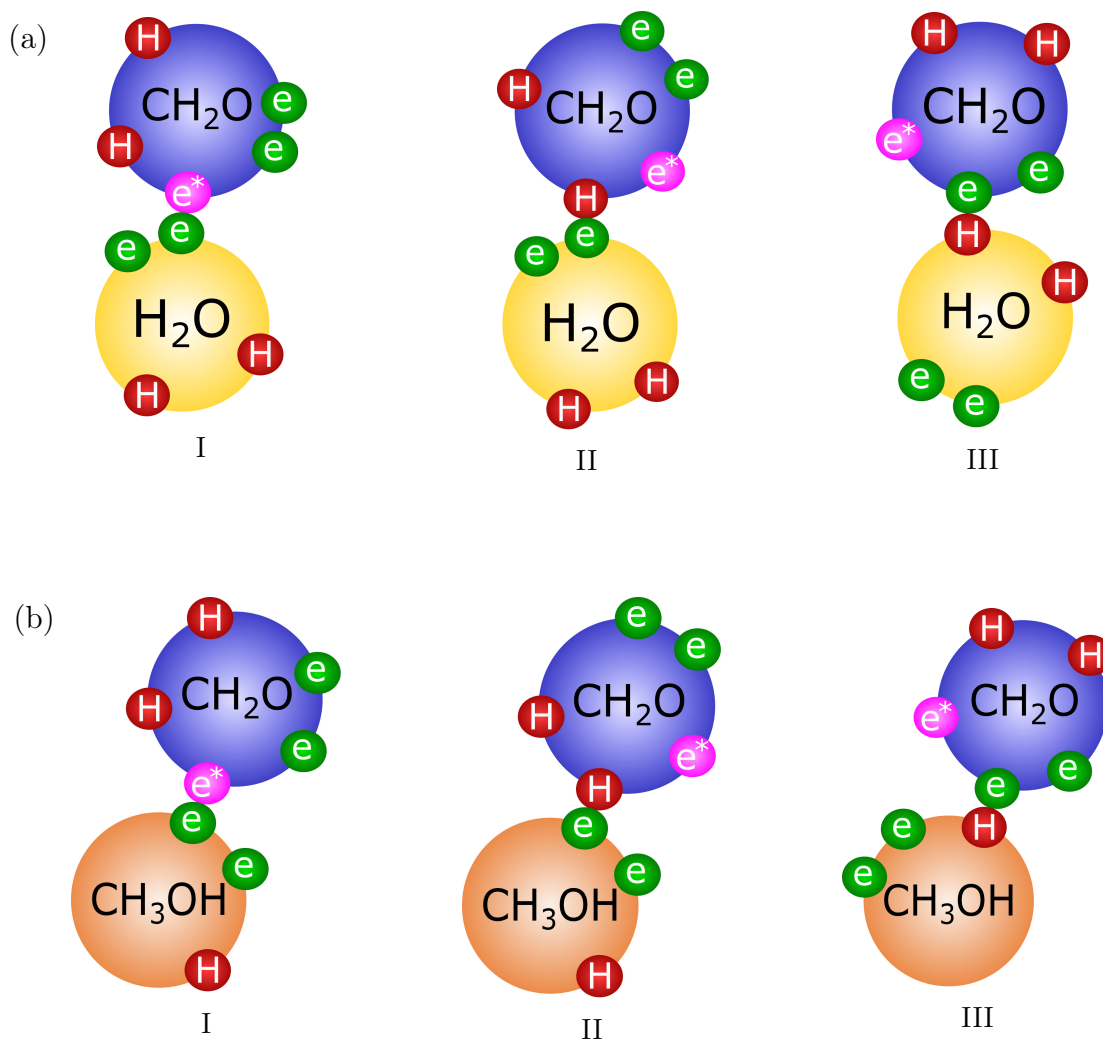
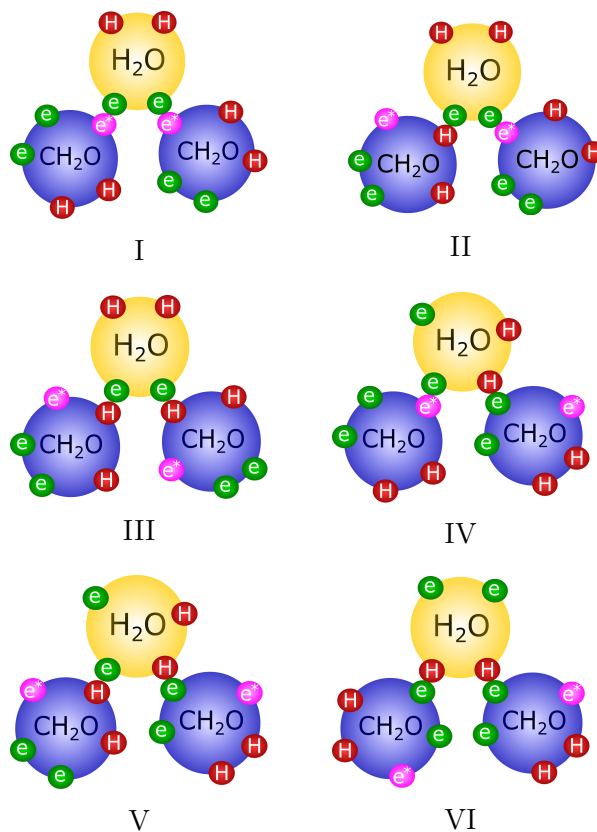


Figure 5.3: Configurations for the formation of (a) methylene glycol (MG_1) and (b) hemiformal (HMF_1) mediated by e -H, and e^* -H site-site bonds. e - e , e^* - e^* , H-H or H- e^* bonding is not allowed. We note that, for simplicity, the CH_2O and CH_3OH groups are depicted by a single sphere; refer to Figure 5.2 for the corresponding representation of the groups within the SAFT- γ Mie GC approach.

Our physical treatment, can also account for the formation of higher-order oligomers (Reactions (5.2) and (5.4)). For instance, there are six possible combinations of association interactions for the formation of dioxymethylene glycol (MG_2) in our model, these are shown in Figure 5.4 (a). Likewise, dioxymethylene hemiformal (HMF_2) can form by one of the five association

interaction combinations shown in Figure 5.4 (b).

(a)



(b)

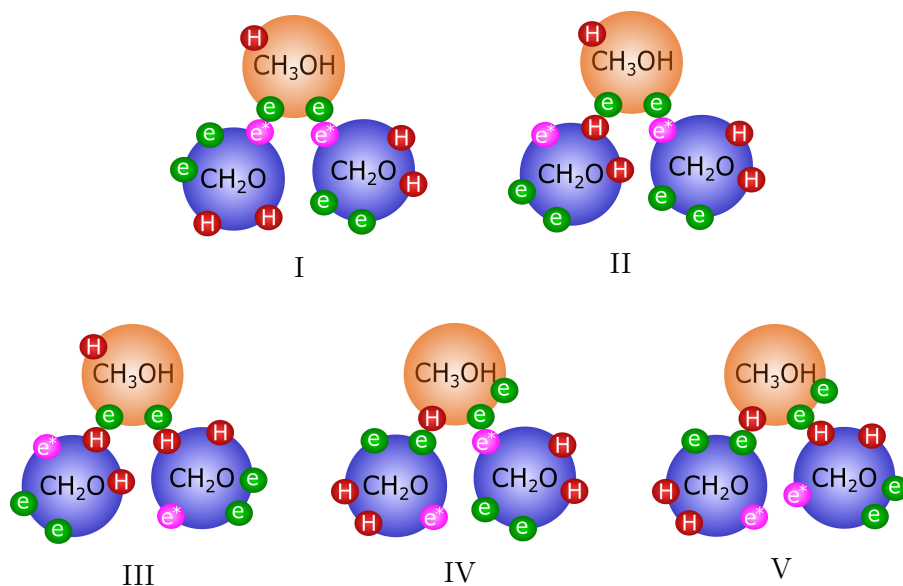


Figure 5.4: Configurations for the formation of (a) the dioxymethylene glycol (MG_2) and (b) the dioxymethylene hemiformal (HMF_2) mediated by e -H, and e^* -H site-site bonds. e - e , e^* - e^* , H-H or H- e^* bonding is not allowed. We note that, for simplicity, the CH_2O and CH_3OH groups are depicted by a single sphere; refer to Figure 5.2 for the corresponding representation of the groups within the SAFT- γ Mie GC approach.

The like and unlike SAFT- γ Mie group interactions for H_2O and CH_3OH are obtained from

previous work [36–38]. In this chapter we expand the group-interactions matrix by estimating the SAFT- γ Mie parameters for the CH₂O group and its unlike interactions with H₂O and CH₃OH. The determination of the group-interaction parameters is done in a successive manner. We first determine the like parameters for the CH₂O group, these parameters are then used to obtain the unlike interaction parameters with the other groups. Details of the parameter estimation strategy are outlined in the following section.

5.2.4 Parameter estimation

To estimate the SAFT- γ Mie parameters, the least-squares objective function given by Equation (2.41) is minimised. In order to determine the accuracy of the theoretical description compared to the experimental data for a property X , the percentage absolute average deviation (%AAD)

$$\%AAD[X] = \frac{100}{N_X^D} \sum_{i=1}^{N_X^D} \left| \frac{R_{X,i}^{\text{exp}} - R_{X,i}^{\text{calc}}}{R_{X,i}^{\text{exp}}} \right|, \quad (5.12)$$

and the absolute average deviation of (AAD)

$$AAD[X] = \frac{1}{N_X^D} \sum_{i=1}^{N_X^D} |R_{X,i}^{\text{exp}} - R_{X,i}^{\text{calc}}|, \quad (5.13)$$

are used.

The optimal SAFT- γ Mie parameters obtained for the CH₂O group are presented in Tables 5.1, 5.3 and 5.4 and are discussed in the following section. The unlike interaction parameters of CH₂O with H₂O and CH₃OH are provided in the tables, with a comparison of the SAFT- γ Mie predictions to experimental VLE data of binary mixtures of formaldehyde with water and methanol shown in Section 5.3.2. We also compare our SAFT- γ Mie predictions of the VLE for ternary formaldehyde + water + methanol mixtures with experimental data which were not used for parameter estimation in Section 5.3.3.

5.3 Phase equilibria

5.3.1 Pure compounds

To study the phase behaviour and speciation of formaldehyde solutions, the CH₂O group needs to be characterised within the SAFT- γ Mie framework. In the current work, the CH₂O group (cf. Figure 5.2) is modelled with two identical segments ($\nu_{\text{CH}_2\text{O}}^* = 2$) and with three association site types ($N_{\text{ST,CH}_2\text{O}} = 3$): two sites of type e which represent the electron lone-pairs on the oxygen atom, two sites of type H corresponding to each of the hydrogen atoms and one site of type e* as a *reactive* site. The attractive exponent of the Mie potential $\lambda_{\text{CH}_2\text{O}}^a$ is fixed to a value of 6.0. The remaining SAFT- γ Mie parameters for the CH₂O group are estimated using experimental data [273, 351] of pure formaldehyde. Specifically, vapour pressure, saturated liquid density and enthalpy of vapourisation data are used. The set of optimal parameters can be found in Tables 5.1 and 5.4, which also contain the previously published SAFT- γ Mie parameters for CH₃OH and H₂O [36–38]. The SAFT- γ Mie calculations are compared to experimental data [273, 351] in Figure 5.5, and the corresponding %AADs are presented in Table 5.2. As can be gleaned from the figure and %AADs, the model leads to a very good description of the thermodynamic properties of formaldehyde. Additionally, our model can closely predict the critical temperature of formaldehyde ($T_c^{\text{calc}} = 402$ K and $T_c^{\text{exp}} = 412 \pm 2.0$ K [352]) (cf. Figure 5.5 (c)), although data in the critical region, above $0.9 T_c$, were not included in the parameter estimation.

Table 5.1: SAFT- γ Mie like parameters for water, methanol and formaldehyde. ν_k^* is the number of segments making up group k , S_k is the shape factor, λ_{kk}^r and λ_{kk}^a are the repulsive and attractive exponents of the Mie potential, respectively, σ_{kk} is the segment diameter, $\sigma_{kk}^{\text{Born}}$ is the Born diameter, ε_{kk} is the dispersion energy between two k groups, $N_{\text{ST},k}$ is the number of association site types on group k , with $n_{\text{H},k}$, $n_{\text{e},k}$ and $n_{\text{e}^*,k}$ representing the association sites of type H, e and e*, respectively. † † † indicates parameters obtained in the current work.

Group k	ν_k^*	S_k	$\sigma_{kk}/\text{\AA}$	λ_{kk}^r	λ_{kk}^a	$(\varepsilon_{kk}/k_{\text{B}})/\text{K}$	$N_{\text{ST},k}$	$n_{\text{H},k}$	$n_{\text{e},k}$	$n_{\text{e}^*,k}$	ref
H ₂ O	1	1.0000	3.0063	17.020	6.0000	266.68	2	2	2	-	[37, 38]
CH ₃ OH	2	0.83517	3.2462	19.235	6.0000	307.69	2	1	2	-	[36]
CH ₂ O	2	0.78473	3.3131	36.423	6.0000	280.15	3	2	2	1	† † †

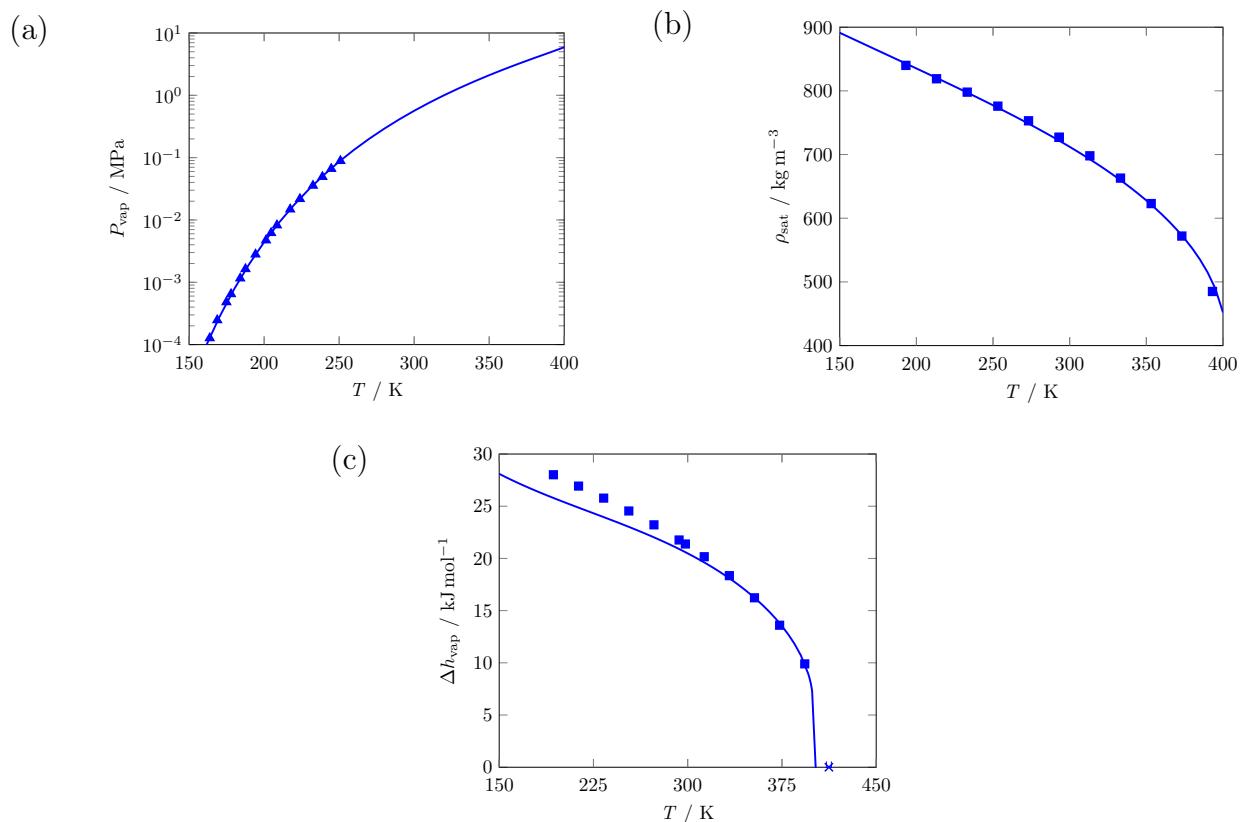


Figure 5.5: (a) Vapour pressure, (b) saturated liquid density and (c) vaporization enthalpy for pure formaldehyde. The curves correspond to the SAFT- γ Mie calculations and the symbols are experimental data [273, 351]. The cross in (c) corresponds to the critical temperature of formaldehyde [352].

Table 5.2: The percentage average absolute deviation %AAD between the SAFT- γ Mie calculations and experimental data of the saturated density ρ_{sat} , vapor pressure P_{vap} and vaporization enthalpy Δh_{vap} of formaldehyde.

Property X	T range/ K	N_X^D	%AAD [X]	ref
ρ_{sat}	193 – 393	12	0.64	[273]
P_{vap}	163 – 250	16	1.96	[351]
Δh_{vap}	193 – 393	12	3.95	[273]

5.3.2 Binary formaldehyde + water and formaldehyde + methanol mixtures

To model the mixture of formaldehyde + water we use the model described in Section 5.2.3, following a physical approach to account for Reactions (5.1) and (5.2) in the liquid and vapour phases. A total of eight unlike interactions need to be determined between the CH_2O and

H₂O groups, namely: $\varepsilon_{\text{CH}_2\text{O}-\text{H}_2\text{O}}$, $\lambda_{\text{CH}_2\text{O}-\text{H}_2\text{O}}^r$, $\varepsilon_{\text{CH}_2\text{O}-\text{H}_2\text{O},\text{H}-e}^{\text{HB}}$, $K_{\text{CH}_2\text{O}-\text{H}_2\text{O},\text{H}-e}$, $\varepsilon_{\text{CH}_2\text{O}-\text{H}_2\text{O},e-\text{H}}^{\text{HB}}$, $K_{\text{CH}_2\text{O}-\text{H}_2\text{O},e-\text{H}}$, $\varepsilon_{\text{CH}_2\text{O}-\text{H}_2\text{O},e^*-\text{e}}^{\text{HB}}$ and $K_{\text{CH}_2\text{O}-\text{H}_2\text{O},e^*-\text{e}}$. The repulsive exponent $\lambda_{\text{CH}_2\text{O}-\text{H}_2\text{O}}^r$ is calculated by combining rules. Experimental isobaric VLE data for formaldehyde + water mixtures [322] in a pressure range of 0.013–0.053 MPa are used for the estimation of the remaining parameters using the objective function in Equation (2.41). To reduce the number of parameters that need to be estimated, we consider a symmetric association scheme for the association interactions not involving the e* site, i.e., $\varepsilon_{\text{CH}_2\text{O}-\text{H}_2\text{O},\text{H}-e}^{\text{HB}} = \varepsilon_{\text{CH}_2\text{O}-\text{H}_2\text{O},e-\text{H}}^{\text{HB}}$ and $K_{\text{CH}_2\text{O}-\text{H}_2\text{O},\text{H}-e} = K_{\text{CH}_2\text{O}-\text{H}_2\text{O},e-\text{H}}$. The optimized unlike parameters can be found in Tables 5.3 and 5.4.

Table 5.3: Group dispersion interaction energies ε_{kl} and repulsive exponent λ_{kl}^r for use with the SAFT- γ Mie equation of state, taken from the references indicated in the last column. CR indicates that λ_{kl}^r is obtained from a combining rule [35] and † † † indicates parameters obtained in the current work.

Group k	Group l	$(\varepsilon_{kl}/k_B)/\text{K}$	λ_{kl}^r	ref
H ₂ O	H ₂ O	266.68	CR	[37, 38]
H ₂ O	CH ₃ OH	278.45	CR	[36]
CH ₃ OH	CH ₃ OH	307.69	CR	[36]
CH ₃ OH	CH ₂ O	320.98	CR	† † †
CH ₂ O	CH ₂ O	280.15	CR	† † †
CH ₂ O	H ₂ O	248.26	CR	† † †

Table 5.4: Group association energies $\varepsilon_{kl,ab}^{\text{HB}}$ and bonding volume $K_{kl,ab}$ for use with the SAFT- γ Mie EoS. † † † indicates parameters obtained in the current work. The last column shows the value of $F_{kl,ab}K_{kl,ab}$ at $T = 298.15$ K for each association interaction, which is representative of the association strength ($\Delta_{ij,kl,ab}$) for a given association kernel I_{ij} (cf. Equation (5.8)).

Group k	Site a on group k	Group l	Site b on group l	$(\varepsilon_{kl}^{\text{HB}}/k_B)/\text{K}$	$K_{kl,ab}/\text{\AA}^3$	ref	$F_{kl,ab}K_{kl,ab}$ (298 K)
H ₂ O	H	H ₂ O	e	1985.4	101.69	[37, 38]	7.9198×10^4
H ₂ O	H	CH ₃ OH	e	1993.5	104.11	[36]	8.3318×10^4
H ₂ O	e	CH ₃ OH	H	1993.5	104.11	[36]	8.3318×10^4
CH ₃ OH	H	CH ₃ OH	e	2062.1	106.57	[36]	1.0738×10^5
CH ₃ OH	H	CH ₂ O	e	4594.9	0.00100	† † †	4.9326×10^3
CH ₃ OH	e	CH ₂ O	H	4594.9	0.00100	† † †	4.9326×10^3
CH ₃ OH	e	CH ₂ O	e^*	5562.3	0.91815	† † †	1.1618×10^8
CH ₂ O	e	CH ₂ O	H	500.00	1381.9	† † †	6.0106×10^3
CH ₂ O	e	H ₂ O	H	2572.7	7.3729	† † †	4.1213×10^4
CH ₂ O	H	H ₂ O	e	2572.7	7.3729	† † †	4.1213×10^4
CH ₂ O	e^*	H ₂ O	e	2445.0	3147.1	† † †	1.1462×10^7

The SAFT- γ Mie calculations for the isobaric VLE of formaldehyde + water mixtures are

in excellent agreement to experimental data as shown in Figures 5.6 (a) and (b), with the corresponding %AAD for the bubble (T_{bub}) and dew temperatures (T_{dew}) of 0.12% and 0.12%, respectively (Table 5.5). The phase boundaries are reported in the figures in terms of the "overall" mole fraction of formaldehyde, always indicated by a ($\sim$) in the current work, $\tilde{x}_{\text{F}} = \frac{\tilde{N}_{\text{F}}}{\tilde{N}_{\text{tot}}} \left(\tilde{N}_{\text{tot}} = \sum_{i=\text{F,M,W}} \tilde{N}_i \right)$; the "overall" ($\tilde{x}_{\text{F}}$) and "true" (x_{F}) mole fractions of formaldehyde can be related as discussed in Section 5.4.

We assess the validity of the model by predicting isothermal VLE data [321] at temperature and pressure conditions not considered in the parameter estimation, and present a comparison between the SAFT- γ Mie predictions and experimental data in Figures 5.6 (c) and (d). Very good agreement to experimental data is shown in the figures. The corresponding %AAD is 2.37% and 1.39% for the bubble (P_{bub}) and dew (P_{dew}) pressures, respectively (Table 5.5). It is remarkable to see in Figures 5.6 (b) and (d) that the SAFT- γ Mie VLE calculations can capture the azeotropes for all the isotherms and isobars considered for this binary mixture. Additionally, our predictions of the saturation points of formaldehyde are in close agreement with experimental measurements [281], within experimental uncertainty, for all the temperatures and pressures in Figure 5.6.

Table 5.5: The percentage average absolute deviation %AAD between the SAFT- γ Mie calculations and experimental data of the bubble and dew temperatures (T_{bub} , T_{dew}) and pressures (P_{bub} , P_{dew}) for binary mixtures of formaldehyde + water (F + W) and formaldehyde + methanol (F + M).

Mixture	P range/ MPa	N_X^{D}	%AAD T_{dew}	N_X^{D}	%AAD T_{bub}	ref
F + W	0.013–0.053	34	0.12	38	0.12	[322]
F + M	0.027–0.101	12	0.67	12	1.94	[326]
Mixture	T range/ K	N_X^{D}	%AAD P_{dew}	N_X^{D}	%AAD P_{bub}	ref
F + W	313–363	38	1.39	43	2.37	[321]
F + M	333–353	27	9.32	27	9.49	[327]

A similar approach is followed to model the phase behaviour of the binary mixture of formaldehyde and methanol, using a physical approach to account for Reactions (5.3) and (5.4) in the liquid and vapour phases, as described in Section 5.2.3. Eight unlike interaction parameters are required for the CH_2O and CH_3OH groups: $\varepsilon_{\text{CH}_2\text{O}-\text{CH}_3\text{OH}}$, $\lambda_{\text{CH}_2\text{O}-\text{CH}_3\text{OH}}^{\text{r}}$,

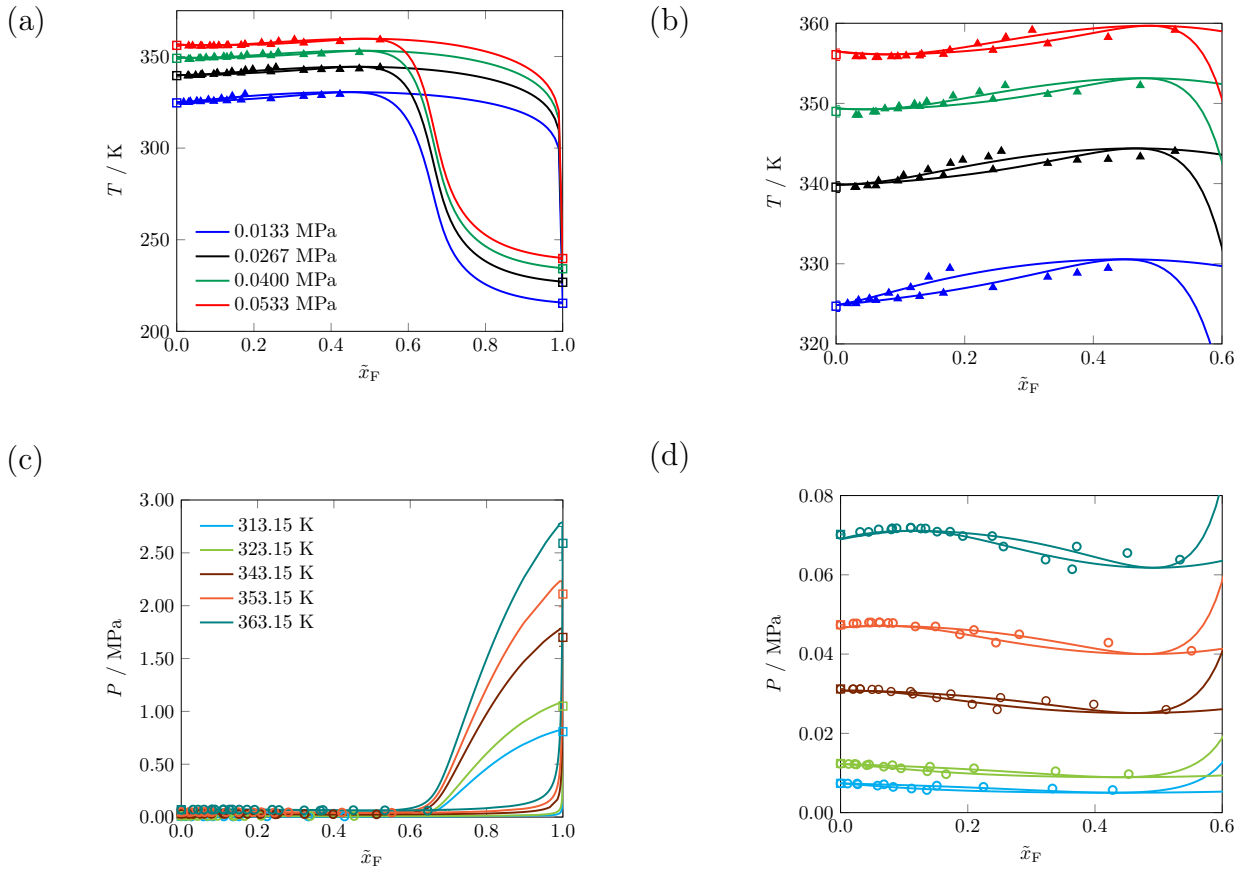


Figure 5.6: (a) Isobaric ($T-\tilde{x}$) and (c) isothermal ($P-\tilde{x}$) phase diagrams featuring the vapour-liquid equilibria of formaldehyde + water binary mixtures, with an enlargement of the water-rich regions of the (b) isobaric and (d) isothermal phase diagrams. The filled symbols represent experimental data [322] used in the parameter estimation and open symbols are experimental data (circles [321] and squares [281]) used for model validation only. The continuous curves represent calculations using SAFT- γ Mie. Error bars are shown for the experimental saturation points (squares [281]) indicating the uncertainty associated with the measurements.

$\varepsilon_{\text{CH}_2\text{O}-\text{CH}_3\text{OH},\text{H}-\text{e}}^{\text{HB}}$, $K_{\text{CH}_2\text{O}-\text{CH}_3\text{OH},\text{H}-\text{e}}$, $\varepsilon_{\text{CH}_2\text{O}-\text{CH}_3\text{OH},\text{e}-\text{H}}^{\text{HB}}$, $K_{\text{CH}_2\text{O}-\text{CH}_3\text{OH},\text{e}-\text{H}}$, $\varepsilon_{\text{CH}_2\text{O}-\text{CH}_3\text{OH},\text{e}^*-\text{e}}^{\text{HB}}$ and $K_{\text{CH}_2\text{O}-\text{CH}_3\text{OH},\text{e}^*-\text{e}}$. The repulsive exponent $\lambda_{\text{CH}_2\text{O}-\text{CH}_3\text{OH}}^{\text{r}}$ is calculated by combining rules and a symmetric association scheme is considered for the association interactions between CH_2O and CH_3OH not involving the reactive e^* site, such that, $\varepsilon_{\text{CH}_2\text{O}-\text{CH}_3\text{OH},\text{H}-\text{e}}^{\text{HB}} = \varepsilon_{\text{CH}_2\text{O}-\text{CH}_3\text{OH},\text{e}-\text{H}}^{\text{HB}}$ and $K_{\text{CH}_2\text{O}-\text{CH}_3\text{OH},\text{H}-\text{e}} = K_{\text{CH}_2\text{O}-\text{CH}_3\text{OH},\text{e}-\text{H}}$. The unlike parameters are obtained following the parameter estimation strategy outlined in Section 5.2.4, using experimental isobaric and isothermal VLE data for formaldehyde + methanol mixtures [326, 327]. The estimated parameters can be found in Tables 5.3 and 5.4.

A very good agreement is obtained between the SAFT- γ Mie calculations and experimen-

tal data for the VLE phase boundaries, as displayed in Figure 5.7. The model captures the azeotrope for all the isotherms and isobars with close agreement to experimental data (see Figures 5.7 (b) and (d)). The good agreement of our model to experimental data is also validated by the %AAD presented in Table 5.5. Very low %AAD values of 0.67% and 1.94% are obtained for the T_{dew} and T_{bub} , respectively, and slightly larger %AAD values are observed for the dew and bubble pressures with an overall %AAD of 9.62% and 7.54%, respectively. The largest deviations for P_{dew} and P_{bub} are for the data points at the highest $\tilde{x}_{\text{F}}$ of each isotherm, which correspond to the lowest pressure values, where larger uncertainties can be expected in the experimental data. For example, the %AAD for P_{dew} at $T = 353.15$ K and $\tilde{x}_{\text{F}} = 0.72$ is 27.83% while the %AAD for the same property at the same temperature and $\tilde{x}_{\text{F}} = 0.011$ is 2.99%. It can also be noted from Figure 5.7 (d) that, for the isotherms at 343.15 K and 353.15 K, the highest (at low $\tilde{x}_{\text{F}}$) and lowest (at high $\tilde{x}_{\text{F}}$) values for P_{dew} (or P_{bub}) differ by an order of magnitude, which explains the more significant %AAD values for small deviations at the lowest pressure values compared to the highest ones. Therefore, it is the deviations at the largest $\tilde{x}_{\text{F}}$ per isotherm that contribute to a higher overall %AAD for P_{dew} and P_{bub} . In the formaldehyde + water mixture, we did not observe large %AAD for P_{dew} and P_{bub} as the pressure values at the lowest and highest $\tilde{x}_{\text{F}}$ are of the same order of magnitude.

There are, to the best of our knowledge, no experimental data for the VLE of binary formaldehyde + methanol (F + M) or formaldehyde + water (F + W) mixtures at high concentrations of formaldehyde ($\tilde{x}_{\text{F}} > 0.53$ for F + W and $\tilde{x}_{\text{F}} > 0.72$ for F + M). Hence, the SAFT- γ Mie predictions in Figures 5.6 (a) and (c) and Figures 5.7 (a) and (c), provide insight of the VLE profile for these binary mixtures at high concentrations of formaldehyde for the given ranges of temperature and pressure.

In the last column of Table 5.4 the value of $F_{kl,ab}K_{kl,ab}$ for each association interaction, which is representative of the association strength ($\Delta_{ij,kl,ab}$) for a given association kernel I_{ij} (cf. Equation (5.8)), is presented at $T = 298.15$ K. It is evident that $F_{kl,ab}K_{kl,ab}$ of the interactions involving the reactive e^* site on the CH_2O group are several orders of magnitude larger than the other interactions shown. This suggests that the association interactions with the e^* site on CH_2O are favorable when the corresponding sites are present in a mix-

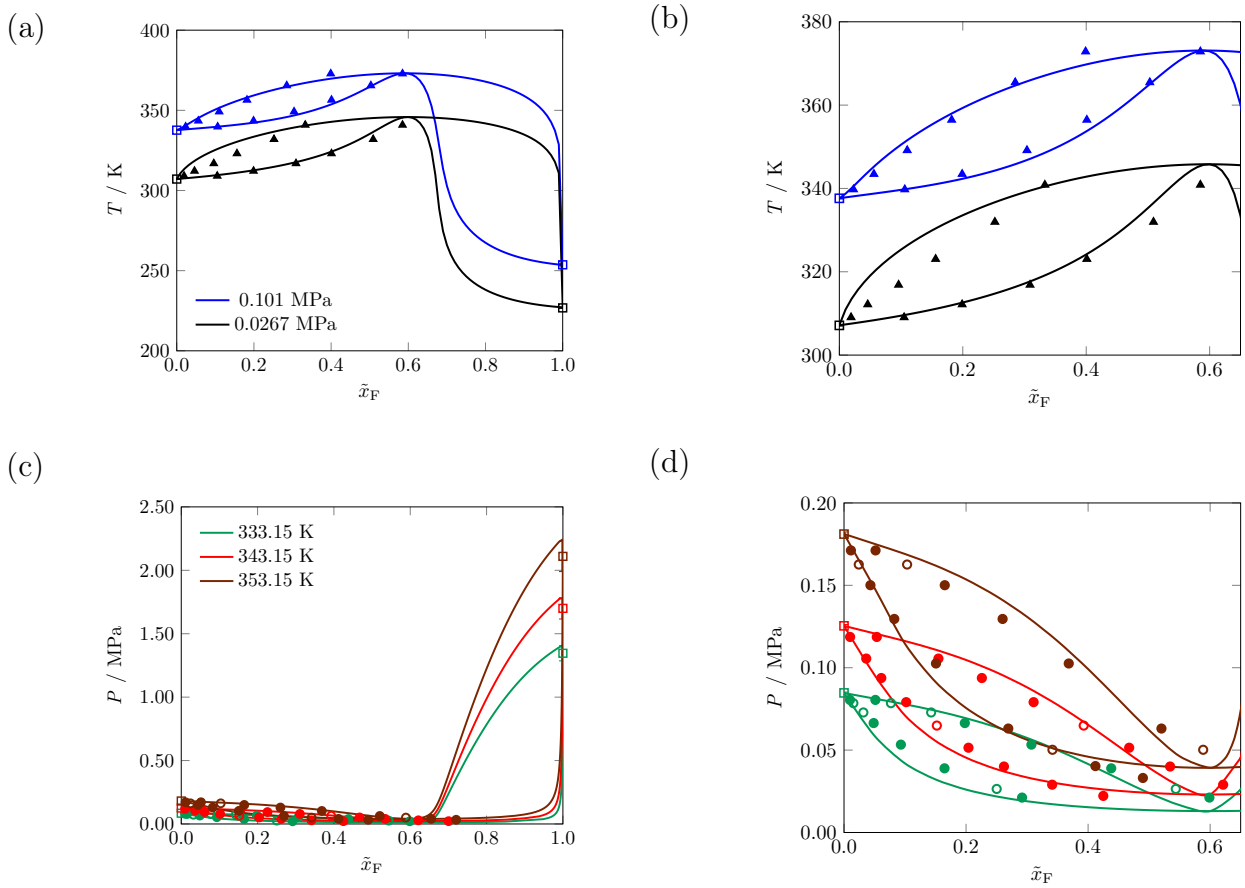


Figure 5.7: (a) Isobaric ($T - \tilde{x}$) and (b) isothermal ($P - \tilde{x}$) phase diagrams featuring the vapour-liquid equilibria of formaldehyde + methanol binary mixtures, with an enlargement of the methanol-rich regions of the (c) isobaric and (d) isothermal phase diagrams. The filled symbols represent experimental data (triangles [326] and circles [327]) used in the parameter estimation and open symbols are experimental data (circles [327] and squares [281]) used for model validation only. The continuous curves represent calculations using SAFT- γ Mie. Error bars are shown for the experimental saturation points (squares [281]) showing the uncertainty associated with the measurements.

ture, i.e., the e site on H_2O will preferentially associate with e^* on CH_2O compared to the other sites on CH_2O , H_2O and CH_3OH . Likewise, the e site on CH_3OH will favor the association with e^* on CH_2O compared to the other sites on CH_2O , H_2O and CH_3OH . We note that the value of $F_{\text{CH}_2\text{O}-\text{CH}_3\text{OH}, e^*-e} K_{\text{CH}_2\text{O}-\text{CH}_3\text{OH}, e^*-e}$ is an order of magnitude larger than $F_{\text{CH}_2\text{O}-\text{H}_2\text{O}, e^*-e} K_{\text{CH}_2\text{O}-\text{H}_2\text{O}, e^*-e}$, suggesting that, in a ternary formaldehyde + water + methanol mixture, the e^* site on CH_2O is more likely to associate with the e site on CH_3OH rather than the e site on H_2O . The preferential association of CH_2O with CH_3OH , as compared to H_2O observed in our model is in agreement with findings by other authors [314, 330] who suggest a higher affinity of formaldehyde for methanol as compared to water, in a ternary mixture.

5.3.3 Ternary mixtures of formaldehyde

The VLE for ternary mixtures of formaldehyde + methanol + water is predicted using the SAFT- γ Mie parameters given in Tables 5.1, 5.3 and 5.4. No additional parameters are needed to carry out the calculations for the ternary mixture.

A wide range of experimental data for the VLE of the ternary mixture is used in the validation of the predictive capability of the SAFT- γ Mie EoS (cf. Table 5.6). Generally, our SAFT- γ Mie model provides an accurate prediction of the VLE of the formaldehyde + methanol + water mixtures compared to all the experimental data sources reported in Table 5.6, with an overall %AAD of 3.78% for T_{bub} , T_{dew} , P_{bub} and P_{dew} . The deviation between calculations and experimental data, in terms of %AAD, for T_{bub} and T_{dew} is below 3%, which is excellent considering the wide range of pressures evaluated. The model developed also gives accurate predictions for the P_{bub} and P_{dew} , with %AAD typically below 8% and 11%, respectively. A larger deviation (31.5%) is observed between our SAFT- γ Mie calculations and experimental data of Hasse et al. [320] for P_{dew} of the ternary mixture. This larger deviation is in part due to the low values of pressure, such that small deviations from experimental data have a significant impact on the relative %AAD value. In this case, it is useful to consider the absolute deviations (AAD) for an overall assessment of the accuracy of the model. The low AAD for P_{dew} (2.94×10^{-3} MPa) indicates a good agreement between our calculations and the experimental data of Hasse et al. [320].

It is evident from Table 5.6 that there is some variation in the experimental data for the same property by different research groups. For example, the %AAD for P_{bub} between our calculations and the experimental data from Albert et al. [328] and Kogan et al. [329], who report data for a similar temperature range, is 1.92% and 7.60%, respectively. A similar discrepancy between these two data sources is observed for P_{dew} . This highlights the uncertainty in the experimental data measurements, which is often not reported, and the importance of comparing predictive models to multiple sources of experimental data for appropriate validation.

In Figure 5.8 we show a comparison between the SAFT- γ Mie VLE predictions of the

Table 5.6: The percentage average absolute deviation (%AAD) and the average absolute deviation (AAD) between the SAFT- γ Mie calculations and experimental data [319, 320, 325, 326, 328–330] for the bubble and dew temperatures (T_{bub} , T_{dew}) and pressures (P_{bub} , P_{dew}) of ternary mixtures of formaldehyde + water + methanol.

Ref.	P range/ MPa	N_X^D	%AAD T_{dew}	AAD T_{dew} / K
Blazhin et al. [326]	0.027–0.101	38	0.672	2.24
Ref.	P range/ MPa	N_X^D	%AAD T_{bub}	AAD T_{bub} / K
Blazhin et al. [326]	0.027–0.101	38	0.211	0.72
Olevski et al. [325]	0.008–0.101	450	2.52	8.30
Ref.	T range/ K	N_X^D	%AAD P_{dew}	AAD $\times 10^3$ P_{dew} / MPa
Albert et al. [328]	333–383	65	2.13	1.62
Hasse et al. [320]	293–313	8	31.5	2.94
Kogan et al. [329]	333–353	82	10.6	6.22
Maurer [319]	342–364	23	8.02	4.75
Ref.	T range/ K	N_X^D	%AAD P_{bub}	AAD $\times 10^3$ P_{bub} / MPa
Albert et al. [328]	333–383	65	1.92	1.70
Hasse et al. [320]	293–313	8	5.06	0.43
Kogan et al. [329]	333–353	82	7.60	3.47
Maurer [319]	342–364	23	5.46	2.75
Brandani et al. [330]	328–368	22	4.55	1.89
Total			3.78	3.10

ternary mixture and experimental data from Blazhin et al. [326] at $P = 0.026664$ MPa and $P = 0.101330$ MPa. For a given pressure and overall liquid phase composition, the predictions of the equilibrium vapour phase composition are found to be in good agreement with the experimental data. A closer agreement of the tie-lines connecting the liquid and vapour phases is evident for the isobaric data at 0.101330 MPa (Figure 5.8 (b)) than at 0.026664 MPa (Figure 5.8 (a)). This is also indicated by the lower AAD values in Table 5.7 for the equilibrium vapour composition at 0.101330 MPa compared to 0.026664 MPa. In general, however, the model performs well in predicting the equilibrium vapour composition for all the components at 0.026664 MPa (AAD $\tilde{y}_i < 0.05$ for $i = \text{F, M, W}$) and 0.101330 MPa (AAD $\tilde{y}_i < 0.035$ for $i = \text{F, M, W}$), which is

remarkable given that no ternary mixture data were included in the parameter estimation.

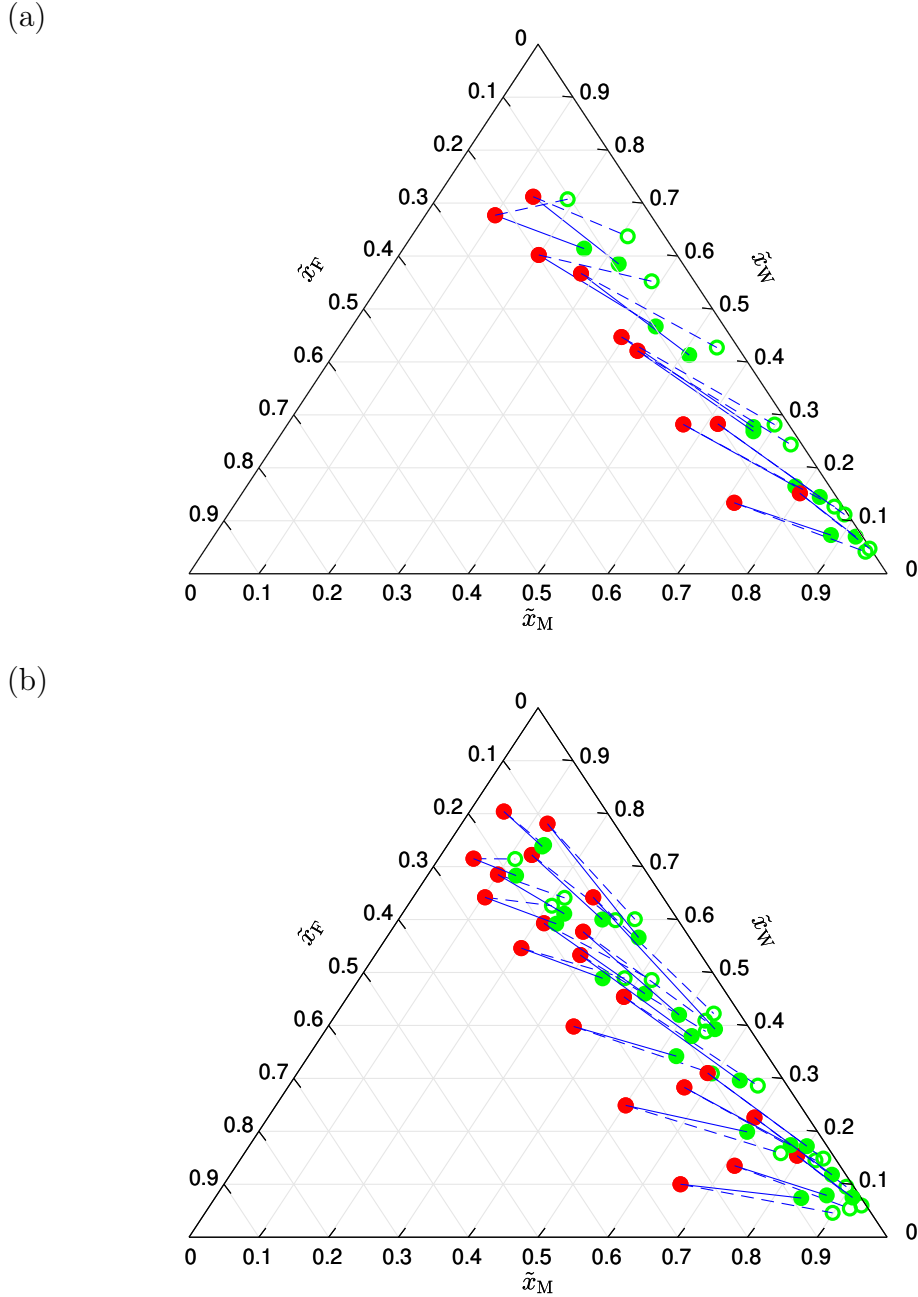


Figure 5.8: Isobaric phase diagrams featuring the vapour-liquid equilibria of formaldehyde + water + methanol ternary mixtures at (a) 0.026664 MPa (temperature range of 312–335 K) and (b) 0.101330 MPa (temperature range of 341–370 K). For a known overall composition of formaldehyde ($\tilde{x}_F$), water ($\tilde{x}_W$) and methanol ($\tilde{x}_M$) in the liquid phase (red), the corresponding overall composition of formaldehyde ($\tilde{y}_F$), water ($\tilde{y}_W$) and methanol ($\tilde{y}_M$) in the vapour phase is predicted using SAFT- γ Mie (open green circles) and is compared to experimental data [326] (filled green circles). Continuous and dashed blue tie-lines connect the overall liquid phase composition to the experimental and predicted overall vapour phase composition, respectively.

Similarly, in Figure 5.9 the ternary phase diagram SAFT- γ Mie predictions of the overall

Table 5.7: The average absolute deviation (AAD) between the SAFT- γ Mie calculations and experimental data from refs [326] and [329] for the equilibrium mole fraction of formaldehyde ($\tilde{y}_F$), water ($\tilde{y}_W$) and methanol ($\tilde{y}_M$) in the vapor phase of ternary mixtures of formaldehyde + water + methanol.

Blazhin et al. [326]					
P/MPa	T range/ K	N_X^D	AAD $\tilde{y}_M$	AAD $\tilde{y}_W$	AAD $\tilde{y}_F$
0.026664	312–334	10	0.048	0.040	0.032
0.101330	341–370	20	0.031	0.022	0.019
Kogan et al. [329]					
T/K	P range/ MPa	N_X^D	AAD $\tilde{y}_M$	AAD $\tilde{y}_W$	AAD $\tilde{y}_F$
333 K	0.017–0.078	27	0.048	0.049	0.041
343 K	0.028–0.114	29	0.038	0.043	0.036
353 K	0.045–0.161	26	0.029	0.037	0.034

vapour phase composition in equilibrium with a preset overall liquid phase composition, at 353 K, are shown compared to experimental data from Kogan et al. [329]. We note that the experimental data by Kogan et al. [329] was chosen for comparison in Figure 5.9 as this data set gives the highest deviations for P_{dew} and P_{bub} in Table 5.6 (with the exception of the %AAD of P_{dew} for the data reported by Hasse et al. [320] which was discussed previously). Additionally, the experimental data by Kogan et al. spans a wider formaldehyde composition range than that presented by the isobaric data of Blazhin et al. and can further validate the predictive capability of the SAFT- γ Mie GC approach for VLE predictions of the ternary mixture. Close agreement between our calculations and experimental data is evident for the isothermal data in Figure 5.9. It can be noted that the slope and length of the predicted tie-lines in the ternary phase diagrams is in close agreement to the experimental tie-lines, particularly at low $\tilde{x}_W$, low $\tilde{x}_F$ and high $\tilde{x}_M$ (bottom right vertex of the ternary plots). The performance of the SAFT- γ Mie model is also assessed in terms of the AAD for the equilibrium vapour phase compositions of the three components at 333 K, 343 K and 353 K, compared to the experimental data by Kogan et al. [329], in Table 5.7. The overall deviations are low (AAD $\tilde{y}_i < 0.05$ for $i = F, M, W$), following a trend of increasing AAD value with decreasing temperature for all the components.

The predictive capability of the SAFT- γ Mie GC approach is also validated in terms of

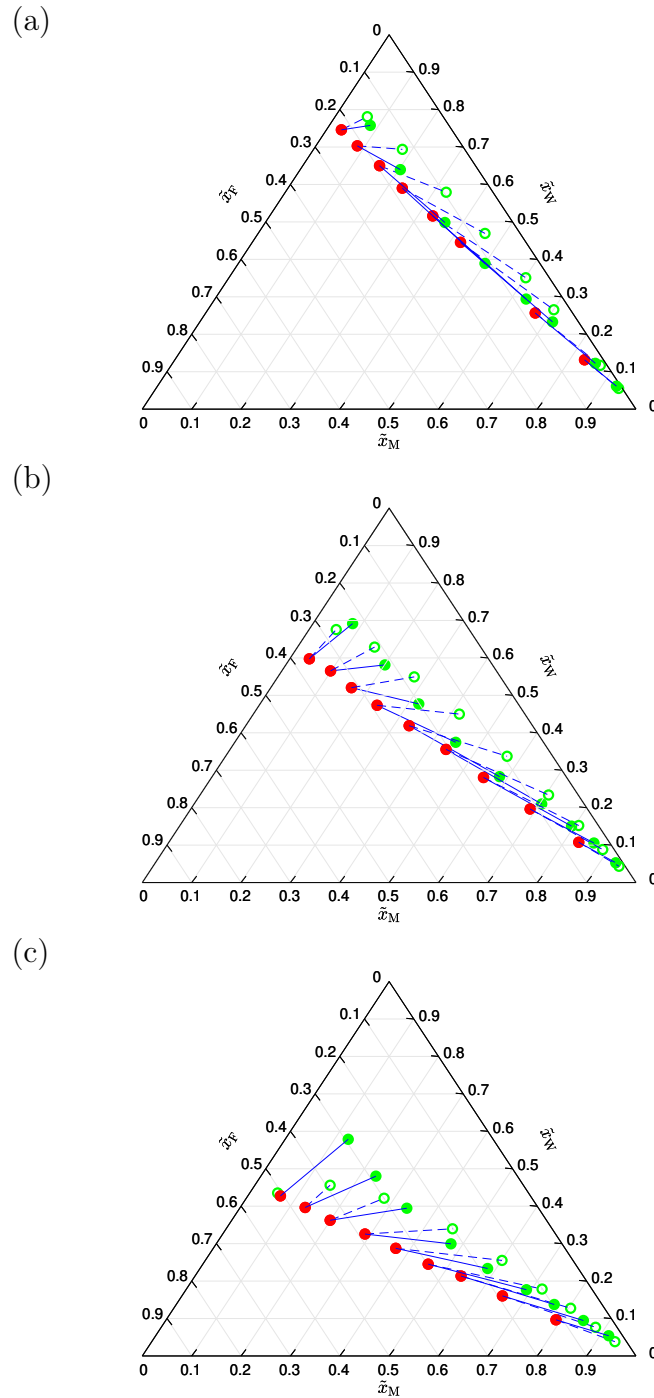


Figure 5.9: Isothermal phase diagrams featuring the vapour-liquid equilibria of formaldehyde + water + methanol ternary mixtures at 353 K (pressure range 0.0449–0.161 MPa). For a known overall composition of formaldehyde ($\tilde{x}_F$), water ($\tilde{x}_W$) and methanol ($\tilde{x}_M$) in the liquid phase (red circles), the corresponding overall composition of formaldehyde ($\tilde{y}_F$), water ($\tilde{y}_W$) and methanol ($\tilde{y}_M$) in the vapour phase is predicted using SAFT- γ Mie (open green circles) and is compared to experimental data [329] (filled green circles). Continuous and dashed blue tie-lines connect the overall liquid phase composition to the experimental and predicted overall vapour phase composition, respectively. In (a) $0.0383 \leq \tilde{x}_F \leq 0.2237$, $0.1317 \leq \tilde{x}_W \leq 0.7460$ and $0.0303 \leq \tilde{x}_M \leq 0.8300$; (b) $0.0621 \leq \tilde{x}_F \leq 0.3628$, $0.1076 \leq \tilde{x}_W \leq 0.5975$ and $0.0397 \leq \tilde{x}_M \leq 0.8303$; and (c) $0.1138 \leq \tilde{x}_F \leq 0.5065$, $0.0965 \leq \tilde{x}_W \leq 0.4275$ and $0.0660 \leq \tilde{x}_M \leq 0.7897$.

the bubble point predictions as shown in Figure 5.10, where we compare our calculations of the bubble pressure for a given temperature and equilibrium liquid phase composition, to the experimental data of Brandani et al. [330]. Excellent agreement between the calculated and measured P_{bub} is displayed for the compositions considered. Interestingly, the best agreement is obtained at the lowest temperature for each of the three liquid-phase compositions considered.

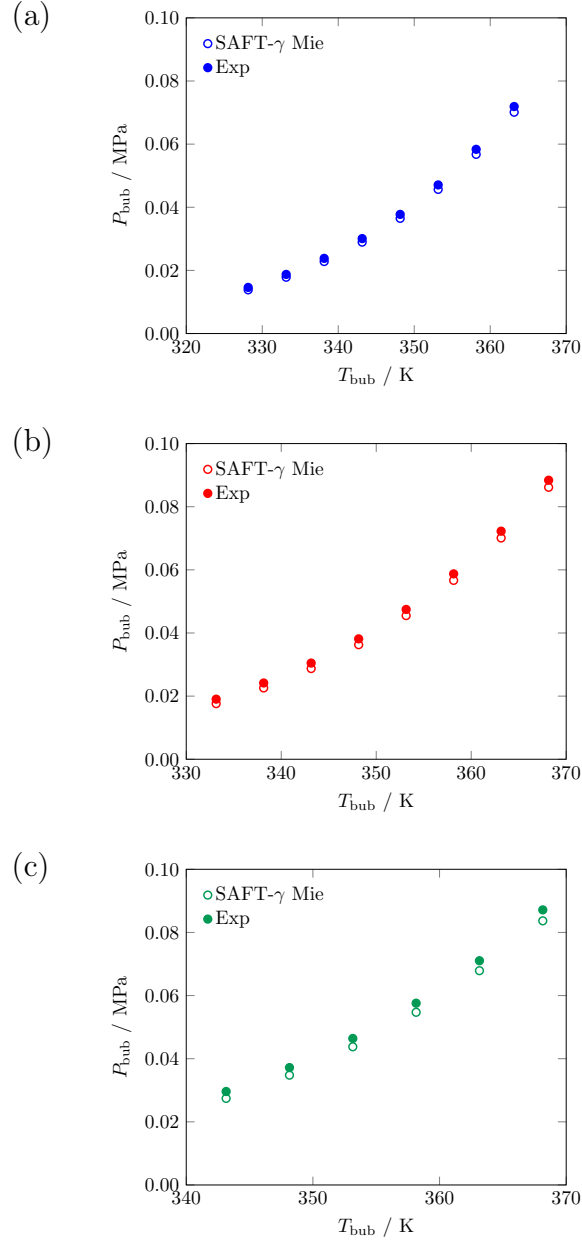


Figure 5.10: Constant-composition phase diagrams featuring the bubble point of formaldehyde + water + methanol ternary mixtures. The overall composition of formaldehyde ($\tilde{x}_F$), water ($\tilde{x}_W$) and methanol ($\tilde{x}_M$) in each solution is (a) $\tilde{x}_F = 0.234$, $\tilde{x}_W = 0.7642$, $\tilde{x}_M = 0.0018$; (b) $\tilde{x}_F = 0.2578$, $\tilde{x}_W = 0.7352$, $\tilde{x}_M = 0.007$; and (c) $\tilde{x}_F = 0.3216$, $\tilde{x}_W = 0.6735$, $\tilde{x}_M = 0.0049$. The filled symbols represent experimental data [330] and open symbols represent predictions using SAFT- γ Mie.

5.4 Distribution of reaction species

In a ternary mixture of formaldehyde, methanol and water, Reactions (5.1)–(5.4) take place simultaneously such that, formaldehyde (F), methanol (M), water (W), methylene glycol (MG_1), poly(oxy)methylene glycols (MG_i), hemiformal (HMF_1) and poly(oxy)methylene hemiformals (HMF_i) are present.

Suppose a ternary mixture is prepared by mixing $\tilde{N}_F$ moles of formaldehyde, $\tilde{N}_M$ moles of methanol and $\tilde{N}_W$ moles of water. These numbers of moles are used to calculate the "overall" mole fraction $\tilde{x}_i = \frac{\tilde{N}_i}{\tilde{N}_{\text{tot}}}$ ($\tilde{N}_{\text{tot}} = \sum_{i=F,M,W} \tilde{N}_i$) of the mixture. However, by virtue of the oligomerization reactions taking place (Reactions (5.1)–(5.4)), at a given T and P , the composition of the mixture is also specified by the "true" number of moles of F, M, W, MG_1 , MG_i , HMF_1 and HMF_i which we denote here as, N_F , N_M , N_W , N_{MG_1} , N_{MG_i} , N_{HMF_1} and N_{HMF_i} , respectively. The true and overall compositions of the components are related by mass balances,

$$\tilde{N}_F = N_F + \sum_{i=1}^{\infty} i \cdot N_{\text{MG}_i} + \sum_{i=1}^{\infty} i \cdot N_{\text{HMF}_i} , \quad (5.14)$$

$$\tilde{N}_W = N_W + \sum_{i=1}^{\infty} N_{\text{MG}_i} , \quad (5.15)$$

$$\tilde{N}_M = N_M + \sum_{i=1}^{\infty} N_{\text{HMF}_i} , \quad (5.16)$$

$$\tilde{N}_{\text{tot}} = \tilde{N}_F + \tilde{N}_W + \tilde{N}_M , \quad (5.17)$$

and

$$N_{\text{tot}} = N_F + N_W + N_M + \sum_{i=1}^{\infty} N_{\text{MG}_i} + \sum_{i=1}^{\infty} N_{\text{HMF}_i} . \quad (5.18)$$

Combining Equations (5.17) and (5.18) leads to

$$\tilde{N}_{\text{tot}} = N_{\text{tot}} + \sum_{i=1}^{\infty} i \cdot N_{\text{MG}_i} + \sum_{i=1}^{\infty} i \cdot N_{\text{HMF}_i} , \quad (5.19)$$

where $\tilde{N}_{\text{tot}}$ and N_{tot} are the overall and true total mole numbers, respectively. Equations

(5.14)–(5.19) can be used to determine the compositions of the binary mixtures too, whereby in mixtures of formaldehyde + water, $\tilde{N}_M = 0$ and, $N_M = N_{\text{HMF}_1} = N_{\text{HMF}_i} = 0$ and in formaldehyde + methanol mixtures, $\tilde{N}_W = 0$, and, $N_W = N_{\text{MG}_1} = N_{\text{MG}_i} = 0$.

The distribution of formaldehyde into the different oligomer species produced in Reactions (5.1)–(5.4) in the liquid and vapour phases can be determined using the SAFT- γ Mie GC approach through the fraction of molecules not bonded at a given site, $X_{i,k,a}$ [31, 32], which is calculated as part of the determination of the Helmholtz free energy in the SAFT EoS (Equation (5.7)).

Given that $X_{i,k,a}$ is the fraction of molecules of component i not bonded at a site of type a on group k , $(1 - X_{i,k,a})$ is the fraction of molecules of component i that are *bonded* at a site of type a on group k . For example, the proportion of formaldehyde molecules bonded at the e^* association site on the CH_2O group is, $(1 - X_{\text{F},\text{CH}_2\text{O},e^*})$ and the corresponding true mole fraction is obtained by multiplying the overall composition of formaldehyde: $\tilde{x}_F(1 - X_{\text{F},\text{CH}_2\text{O},e^*})$. Purely statistical arguments are used to estimate the true mole fraction of each species, where the bonded fractions are treated as probabilities.

It is worth mentioning that in order to determine the distribution of reaction species in the current work, the formation of oligomers up to $i = 2$ only in Reactions (5.2) and (5.4) is taken into account, for simplicity; although the SAFT- γ Mie molecular model presented for formaldehyde mixtures leads to the formation of longer oligomers. Hahnenstein et al. [343] and Albert et al. [331] reported that the average number of CH_2O segments in MG_i and HMF_i in aqueous and methanolic binary mixtures of formaldehyde is 2.1 and 1.1, respectively, at 293 K and an overall formaldehyde composition $\tilde{x}_F = 0.3$. In aqueous solutions, the average number of CH_2O segments in MG_i decreased with increasing temperature, while the effect of temperature on the average number of CH_2O segments in HMF_i was negligible in methanolic solutions [343]. Hence, in our analysis of the speciation occurring in formaldehyde mixtures, we assume that the true mole fraction of oligomers longer than MG_2 and HMF_2 , in the liquid (x_i) and vapour (y_i) phases, is zero, i.e., $x_{\text{MG}_i} = 0$, $y_{\text{MG}_i} = 0$, $x_{\text{HMF}_i} = 0$ and $y_{\text{HMF}_i} = 0$ for $i = 3, 4, \dots, \infty$.

In order to determine the composition of MG_1 , MG_2 , HMF_1 and HMF_2 , it is perhaps

more convenient to express the mass balances in Equations (5.14)–(5.19) in terms of the true, $x_j = \frac{N_j}{N_{\text{tot}}}$, and overall, $\tilde{x}_j = \frac{\tilde{N}_j}{\tilde{N}_{\text{tot}}}$ mole fractions of the components. The corresponding mass balances in terms of the mole fractions of the components for a maximum chain length of $i = 2$ for MG_i and HMF_i , are given as

$$\tilde{x}_{\text{F}} = \frac{1}{z} (x_{\text{F}} + x_{\text{MG}_1} + 2 x_{\text{MG}_2} + x_{\text{HMF}_1} + 2 x_{\text{HMF}_2}) , \quad (5.20)$$

$$\tilde{x}_{\text{W}} = \frac{1}{z} (x_{\text{W}} + x_{\text{MG}_1} + x_{\text{MG}_2}) , \quad (5.21)$$

and,

$$\tilde{x}_{\text{M}} = \frac{1}{z} (x_{\text{M}} + x_{\text{HMF}_1} + x_{\text{HMF}_2}) , \quad (5.22)$$

where

$$z = 1 + x_{\text{MG}_1} + 2 x_{\text{MG}_2} + x_{\text{HMF}_1} + 2 x_{\text{HMF}_2} . \quad (5.23)$$

5.4.1 Binary mixtures

In the analysis of formaldehyde + water and formaldehyde + methanol binary mixtures, we use statistical arguments including $X_{i,k,a}$, calculated by the SAFT- γ Mie EoS, to determine the true mole fraction of MG_1 , MG_2 , HMF_1 and HMF_2 using Equations (5.20)–(5.23).

5.4.1.1 Methylene glycol (MG_1)

Reaction (5.1) is a nucleophilic addition reaction in which water acts as a nucleophile, attracted by the partial positive charge of the carbon atom in the carbonyl group in formaldehyde, to form MG_1 ; this is represented by configuration I in Figure 5.3 (a). Here, however, we model all the possible association schemes that would lead to the formation of an MG_1 aggregate in the SAFT- γ Mie model (cf. Figure 5.3 (a)). In order to determine the true mole fraction of MG_1 , we sum the probability of each configuration forming. For each MG_1 configuration, we multiply the probability that the two corresponding association sites in Figure 5.3 (a) are bonded, and that no other sites on formaldehyde can bond to the bonded site-type on water.

Let us consider configuration I in Figure 5.3 (a). MG_1 forms when the e site on H_2O and the e^* site on CH_2O are bonded, which is given by $\tilde{x}_{\text{W}}(1 - X_{\text{W,H}_2\text{O,e}})\tilde{x}_{\text{F}}(1 - X_{\text{F,CH}_2\text{O,e}^*})$. Additionally, to ensure that no H sites on CH_2O bond to the free e site on H_2O , we account for the probability that the two H sites on CH_2O are not bonded, $(X_{\text{F,CH}_2\text{O,H}})^2$. Hence, the true mole fraction of MG_1 in configuration I, $x_{\text{MG}_{1,\text{I}}}$, in Figure 5.3 (a) is obtained as

$$x_{\text{MG}_{1,\text{I}}} = \tilde{x}_{\text{W}}(1 - X_{\text{W,H}_2\text{O,e}})\tilde{x}_{\text{F}}(1 - X_{\text{F,CH}_2\text{O,e}^*})(X_{\text{F,CH}_2\text{O,H}})^2. \quad (5.24)$$

Following the same approach for other configurations of MG_1 in Figure 5.3 (a), the true mole fraction of MG_1 x_{MG_1} , in a formaldehyde + water mixture is given by,

$$\begin{aligned} x_{\text{MG}_1} &= x_{\text{MG}_{1,\text{I}}} + x_{\text{MG}_{1,\text{II}}} + x_{\text{MG}_{1,\text{III}}} \\ &= \tilde{x}_{\text{W}}(1 - X_{\text{W,H}_2\text{O,e}})\tilde{x}_{\text{F}}(1 - X_{\text{F,CH}_2\text{O,e}^*})(X_{\text{F,CH}_2\text{O,H}})^2 \\ &\quad + \tilde{x}_{\text{W}}(1 - X_{\text{W,H}_2\text{O,e}})\tilde{x}_{\text{F}}(1 - X_{\text{F,CH}_2\text{O,H}})X_{\text{F,CH}_2\text{O,e}^*} \\ &\quad + \tilde{x}_{\text{W}}(1 - X_{\text{W,H}_2\text{O,H}})\tilde{x}_{\text{F}}(1 - X_{\text{F,CH}_2\text{O,e}}). \end{aligned} \quad (5.25)$$

5.4.1.2 Dioxymethylene glycol (MG_2)

MG_2 is produced via a condensation reaction between two MG_1 molecules according to Reaction (5.2) for $i = 2$. Here we model all the possible configurations for the formation of MG_2 (cf. Figure 5.4 (a)). In order to determine the true mole fraction of MG_2 , x_{MG_2} , in a formaldehyde + water binary mixture, we sum the probability that each configuration would form. For example, for MG_2 to form in configuration I or II, the two e sites on H_2O must be bonded and the two H sites on H_2O must not be bonded, to ensure that the bonding of e sites on H_2O is to an H site and/or an e^* site on CH_2O , we follow

$$x_{\text{MG}_{2,\text{I}}} + x_{\text{MG}_{2,\text{II}}} = \tilde{x}_{\text{W}}(1 - X_{\text{W,H}_2\text{O,e}})(1 - X_{\text{W,H}_2\text{O,e}})(X_{\text{W,H}_2\text{O,H}})^2, \quad (5.26)$$

where $x_{\text{MG}_{2,\text{I}}}$ and $x_{\text{MG}_{2,\text{II}}}$ are the true mole fractions of MG_2 in configurations I and II, respectively. By multiplying $X_{\text{W,H}_2\text{O,H}}$ in Equation (5.26) we ensure that the H_2O does not associate

with other H_2O or CH_2O groups via its H sites. For MG_2 to form in configuration III, IV, V or VI we express the probabilities in terms of the sites bonded and not bonded on CH_2O . Therefore,

$$\begin{aligned}
 x_{\text{MG}_2} &= x_{\text{MG}_{2,\text{I}}} + x_{\text{MG}_{2,\text{II}}} + x_{\text{MG}_{2,\text{III}}} + x_{\text{MG}_{2,\text{IV}}} + x_{\text{MG}_{2,\text{V}}} + x_{\text{MG}_{2,\text{VI}}} \\
 &= \tilde{x}_{\text{W}}(1 - X_{\text{W},\text{H}_2\text{O},\text{e}})^2 (X_{\text{W},\text{H}_2\text{O},\text{H}})^2 \\
 &\quad + \tilde{x}_{\text{F}}(1 - X_{\text{F},\text{CH}_2\text{O},\text{H}})^2 (X_{\text{F},\text{CH}_2\text{O},\text{e}})^2 X_{\text{F},\text{CH}_2\text{O},\text{e}^*} \\
 &\quad + \tilde{x}_{\text{F}}(1 - X_{\text{F},\text{CH}_2\text{O},\text{e}})(1 - X_{\text{F},\text{CH}_2\text{O},\text{e}^*}) (X_{\text{F},\text{CH}_2\text{O},\text{H}})^2 \\
 &\quad + \tilde{x}_{\text{F}}(1 - X_{\text{F},\text{CH}_2\text{O},\text{e}})(1 - X_{\text{F},\text{CH}_2\text{O},\text{H}}) X_{\text{F},\text{CH}_2\text{O},\text{e}^*} \\
 &\quad + \tilde{x}_{\text{F}}(1 - X_{\text{F},\text{CH}_2\text{O},\text{e}})^2 (X_{\text{F},\text{CH}_2\text{O},\text{H}})^2 X_{\text{F},\text{CH}_2\text{O},\text{e}^*}.
 \end{aligned} \tag{5.27}$$

We note that in each term in Equation (5.27) we trace whether the reactive e^* site on CH_2O is bonded or not bonded. For configurations III–VI this is done by explicitly specifying whether the reactive e^* site on CH_2O is bonded or not bonded. For configurations I and II, we assume that the two e sites on H_2O are bonded to the e^* site on CH_2O and not the H site on H_2O since $F_{\text{CH}_2\text{O}-\text{H}_2\text{O},\text{e}^*-\text{e}} K_{\text{CH}_2\text{O}-\text{H}_2\text{O},\text{e}^*-\text{e}} > F_{\text{H}_2\text{O}-\text{H}_2\text{O},\text{H}-\text{e}} K_{\text{H}_2\text{O}-\text{H}_2\text{O},\text{H}-\text{e}}$ as shown in Table 5.4.

In Table 5.8 a breakdown of the bonded (tick) and not bonded (cross) sites on H_2O and CH_2O that are considered in the calculations of x_{MG_1} and x_{MG_2} in Equations (5.25) and (5.27), respectively, are shown.

5.4.1.3 Hemiformal (HMF_1)

Reaction (5.3) is a nucleophilic addition reaction, similar to Reaction (5.1), in which methanol acts as a nucleophile, attracted by the partial positive charge of the carbon atom in the carbonyl group in formaldehyde, to form HMF_1 ; this is represented by configuration I in Figure 5.3 (b). As for MG_1 , in our current work, we model all the possible configurations for the formation of HMF_1 in our SAFT- γ Mie model (cf. Figure 5.3 (b)). The true mole fraction of HMF_1 , x_{HMF_1} , in a formaldehyde + methanol mixture, can be determined by summing the probability that each configuration would form. For each HMF_1 configuration, we multiply the probability that the two corresponding association sites in Figure 5.3 (b) are bonded. For example, in configuration I in Figure 5.3 (b), HMF_1 will form when the e site on CH_3OH and the e^* site

Table 5.8: The association sites on formaldehyde, F, and water, W, which are bonded (✓) and not-bonded (✗) for each configuration of MG₁ and MG₂ (cf. Figures 5.3 (a) and 5.4 (a)), in a formaldehyde + water binary mixture, as expressed in Equations (5.25) and (5.27), respectively.

Oligomer	Config.	Site on F			Site on W	
		H	e	e*	H	e
MG ₁	I	✗✗		✓		✓
	II	✓		✗		✓
	III		✓		✓	
MG ₂	I, II				✗✗	✓✓
	III	✓✓	✗✗	✗		
	IV	✗✗	✓	✓		
	V	✓	✓	✗		
	VI	✗✗	✓✓	✗		

on CH₂O are bonded, which is given by $\tilde{x}_M (1 - X_{M,CH_3OH,e}) \tilde{x}_F (1 - X_{F,CH_2O,e*})$. Additionally, to ensure that no H sites on CH₂O bond to the free e site on CH₃OH, we account for the probability that the two H sites on CH₂O are not bonded, $(X_{F,CH_2O,H})^2$. Hence, the true mole fraction of HMF₁ forming in configuration I, $x_{HMF_{1,I}}$, in Figure 5.3 (b), becomes

$$x_{HMF_{1,I}} = \tilde{x}_M (1 - X_{M,CH_3OH,e}) \tilde{x}_F (1 - X_{F,CH_2O,e*}) (X_{F,CH_2O,H})^2 . \quad (5.28)$$

Therefore, x_{HMF_1} in a formaldehyde + water mixture is given by,

$$\begin{aligned}
 x_{HMF_1} &= x_{HMF_{1,I}} + x_{HMF_{1,II}} + x_{HMF_{1,III}} \\
 &= \tilde{x}_M (1 - X_{M,CH_3OH,e}) \tilde{x}_F (1 - X_{F,CH_2O,e*}) (X_{F,CH_2O,H})^2 \\
 &\quad + \tilde{x}_M (1 - X_{M,CH_3OH,e}) \tilde{x}_F (1 - X_{F,CH_2O,H}) X_{F,CH_2O,e*} \\
 &\quad + \tilde{x}_M (1 - X_{M,CH_3OH,H}) \tilde{x}_F (1 - X_{F,CH_2O,e}) .
 \end{aligned} \quad (5.29)$$

5.4.1.4 Dioxymethylene hemiformal (HMF₂)

HMF₂ is produced via a condensation reaction between two HMF₁ molecules according to Reaction (5.4) for $i = 2$. Analogous to modelling MG₂ (Section 5.4.1.2), we model all the possible configurations for the formation of HMF₂ (cf. Figure 5.4 (b)). The true mole fraction of HMF₂, x_{HMF_2} , in a formaldehyde + methanol binary mixture can be determined by considering the probability that HMF₂ will form via one of the configurations in Figure 5.4 (b). For example, for HMF₂ to form in configuration I or II, the two e sites on CH₃OH must be bonded and the H site on CH₃OH must not be bonded, such that

$$x_{\text{HMF}_{2,\text{I}}} + x_{\text{HMF}_{2,\text{II}}} = \tilde{x}_{\text{M}}(1 - X_{\text{M,CH}_3\text{OH,e}})^2 X_{\text{M,CH}_3\text{OH,H}} , \quad (5.30)$$

where $x_{\text{HMF}_{2,\text{I}}}$ and $x_{\text{HMF}_{2,\text{II}}}$ are the true mole fractions of HMF₂ in configurations I and II, respectively. By specifying $X_{\text{M,CH}_3\text{OH,H}}$ we ensure that the CH₃OH group does not associate with other CH₃OH or CH₂O groups via the H site. For HMF₂ to form in configuration III, IV or V we express the probabilities in terms of the sites bonded and not bonded on CH₂O. Therefore,

$$\begin{aligned} x_{\text{HMF}_2} &= x_{\text{HMF}_{2,\text{I}}} + x_{\text{HMF}_{2,\text{II}}} + x_{\text{HMF}_{2,\text{III}}} + x_{\text{HMF}_{2,\text{IV}}} + x_{\text{HMF}_{2,\text{V}}} \\ &= \tilde{x}_{\text{M}}(1 - X_{\text{M,CH}_3\text{OH,e}})^2 X_{\text{M,CH}_3\text{OH,H}} \\ &\quad + \tilde{x}_{\text{F}}(1 - X_{\text{F,CH}_2\text{O,H}})^2 (X_{\text{F,CH}_2\text{O,e}})^2 X_{\text{F,CH}_2\text{O,e}^*} \\ &\quad + \tilde{x}_{\text{F}}(1 - X_{\text{F,CH}_2\text{O,e}})(1 - X_{\text{F,CH}_2\text{O,e}^*}) (X_{\text{F,CH}_2\text{O,H}})^2 \\ &\quad + \tilde{x}_{\text{F}}(1 - X_{\text{F,CH}_2\text{O,e}})(1 - X_{\text{F,CH}_2\text{O,H}}) X_{\text{F,CH}_2\text{O,e}^*} . \end{aligned} \quad (5.31)$$

We note that in each term in Equation (5.31), we trace the bonding status of the reactive e* site on CH₂O. For configurations III–V this is done by explicitly specifying whether the reactive e* site on CH₂O is bonded or not-bonded. For configurations I and II, we assume that the two e sites on CH₃OH are bonded to the e* site on CH₂O and not the H site on CH₃OH, since $F_{\text{CH}_2\text{O}-\text{CH}_3\text{OH,e}^*-\text{e}} K_{\text{CH}_2\text{O}-\text{CH}_3\text{OH,e}^*-\text{e}} > F_{\text{CH}_3\text{OH}-\text{CH}_3\text{OH,H}-\text{e}} K_{\text{CH}_3\text{OH}-\text{CH}_3\text{OH,H}-\text{e}}$ as shown in Table 5.4.

In Table 5.9 a breakdown of the bonded (tick) and not-bonded (cross) sites on CH₃OH

and CH_2O that are considered in the calculations of x_{HMF_1} and x_{HMF_2} in Equations (5.29) and (5.31), respectively, are shown.

Table 5.9: The association sites on formaldehyde, F, and methanol, M, which are bonded ($\checkmark$) and not-bonded ($\times$) for each configuration of HMF_1 and HMF_2 (cf. Figure 5.3 (b) and 5.4 (b)), in a formaldehyde + methanol mixture, as expressed in Equations (5.29) and (5.31), respectively.

Oligomer	Config.	Site on F			Site on M	
		H	e	e*	H	e
HMF_1	I	$\times\times$		$\checkmark$		$\checkmark$
	II	$\checkmark$		$\times$		$\checkmark$
	III		$\checkmark$		$\checkmark$	
HMF_2	I, II				$\times$	$\checkmark\checkmark$
	III	$\checkmark\checkmark$	$\times\times$	$\times$		
	IV	$\times\times$	$\checkmark$	$\checkmark$		
	V	$\checkmark$	$\checkmark$	$\times$		

5.4.2 Ternary mixtures

In order to predict the distribution of the reaction species in mixtures of formaldehyde + water + methanol, the component mass balances are used as well as information on the fraction of molecules not bonded, $X_{i,k,a}$, from the SAFT- γ Mie EoS. The complexity of estimating the distribution of species increases for a ternary mixture due to the presence of an additional component in the solution which competes for the association sites on CH_2O . Therefore, expressions involving $X_{i,k,a}$ of the components are more specific for ternary mixtures, compared to binary ones, where information on the sites bonded on the *two* components making up the oligomer are specified for each configuration; that is in contrast to specifying the sites bonded on only one of the components for binary mixtures (cf. Equations (5.27) and (5.31)).

In the current work, x_{HMF_2} and x_{MG_2} are predicted using expressions of $X_{i,k,a}$ then, x_{HMF_1} and x_{MG_1} are calculated from the material balances given in Equations (5.20)–(5.23).

5.4.2.1 MG₂

The true mole fraction of MG₂, x_{MG_2} , in a formaldehyde + water + methanol ternary mixture, is determined by the probability that MG₂ will form via one of the configurations in Figure 5.4 (a). For example, for MG₂ to form in configuration I, the two e sites on H₂O must be bonded and the two H sites on H₂O must not be bonded, in addition to the e* sites on the two CH₂O molecules being bonded i.e.,

$$x_{\text{MG}_{2,\text{I}}} = \tilde{x}_{\text{W}} (1 - X_{\text{W,H}_2\text{O,e}})^2 (X_{\text{W,H}_2\text{O,H}})^2 \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,e}^*})^2 . \quad (5.32)$$

Following a similar approach to calculate the probability of the formation of the remaining MG₂ configurations, the true mole fraction of MG₂ x_{MG_2} is given by

$$\begin{aligned} x_{\text{MG}_2} &= x_{\text{MG}_{2,\text{I}}} + x_{\text{MG}_{2,\text{II}}} + x_{\text{MG}_{2,\text{III}}} + x_{\text{MG}_{2,\text{IV}}} + x_{\text{MG}_{2,\text{V}}} + x_{\text{MG}_{2,\text{VI}}} \\ &= \tilde{x}_{\text{W}} (1 - X_{\text{W,H}_2\text{O,e}})^2 (X_{\text{W,H}_2\text{O,H}})^2 \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,e}^*})^2 \\ &\quad + \tilde{x}_{\text{W}} (1 - X_{\text{W,H}_2\text{O,e}})^2 (X_{\text{W,H}_2\text{O,H}})^2 \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,e}^*}) (1 - X_{\text{F,CH}_2\text{O,H}}) \\ &\quad + \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,H}})^2 (X_{\text{F,CH}_2\text{O,e}})^2 X_{\text{F,CH}_2\text{O,e}^*} \tilde{x}_{\text{W}} (1 - X_{\text{W,H}_2\text{O,e}})^2 \\ &\quad + \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,e}}) (1 - X_{\text{F,CH}_2\text{O,e}^*}) (X_{\text{F,CH}_2\text{O,H}})^2 \tilde{x}_{\text{W}} (1 - X_{\text{W,H}_2\text{O,e}}) (1 - X_{\text{W,H}_2\text{O,H}}) \\ &\quad + \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,e}}) (1 - X_{\text{F,CH}_2\text{O,H}}) X_{\text{F,CH}_2\text{O,e}^*} \tilde{x}_{\text{W}} (1 - X_{\text{W,H}_2\text{O,e}}) (1 - X_{\text{W,H}_2\text{O,H}}) \\ &\quad + \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,e}})^2 (X_{\text{F,CH}_2\text{O,H}})^2 X_{\text{F,CH}_2\text{O,e}^*} \tilde{x}_{\text{W}} (1 - X_{\text{W,H}_2\text{O,H}})^2 . \end{aligned} \quad (5.33)$$

By comparing the calculation of x_{MG_2} in binary and ternary mixtures, given by Equations (5.27) and (5.33), respectively, we note that for x_{MG_2} in ternary mixtures we use information of the sites bonded or not bonded for the two components (formaldehyde and water) in each term in Equation (5.33). This is in contrast to using the fraction of bonded sites for one of the components in each term in Equation (5.27) for the formaldehyde + water binary mixture. In the ternary formaldehyde + water + methanol mixture, the third component (methanol in this case) is competing with water to associate with formaldehyde, hence, we follow a more specific formulation in Equation (5.33) to ensure that the calculation is accounting for the formation of

MG₂ only. The differences in the formulations of x_{MG_2} in binary and ternary mixtures is also shown by comparing Tables 5.8 and 5.10 for MG₂.

Table 5.10: The association sites on formaldehyde, F, water, W, and methanol, M, which are bonded (✓) and not-bonded (✗) for each configuration of MG₂ and HMF₂ (cf. Figures 5.4 (a) and (b)), in a formaldehyde + water + methanol mixture, as expressed in Equations (5.33) and (5.35), respectively.

Oligomer	Config.	Site on F			Site on W		Site on M	
		H	e	e*	H	e	H	e
MG ₂	I			✓✓	✗✗	✓✓		
	II	✓		✓	✗✗	✓✓		
	III	✓✓	✗✗	✗		✓✓		
	IV	✗✗	✓	✓	✓	✓		
	V	✓	✓	✗	✓	✓		
	VI	✗✗	✓✓	✗	✓✓			
HMF ₂	I			✓✓			✗	✓✓
	II	✓		✓			✗	✓✓
	III	✓✓	✗✗	✗				✓✓
	IV	✗✗	✓	✓			✓	✓
	V	✓	✓	✗			✓	✓

5.4.2.2 HMF₂

The true mole fraction of HMF₂, x_{HMF_2} , in a formaldehyde + water + methanol ternary mixture, is determined by the probability that HMF₂ will form via one of the configurations in Figure 5.4 (b). Hence, we sum the probability that each configuration would form. For example, for HMF₂ to form in configuration I, the two e sites on CH₃OH must be bonded, the H site on CH₃OH must not be bonded and the e* sites on the two CH₂O molecules should be bonded, i.e.,

$$x_{\text{HMF}_2, \text{I}} = \tilde{x}_{\text{M}} (1 - X_{\text{M}, \text{CH}_3\text{OH}, \text{e}})^2 X_{\text{M}, \text{CH}_3\text{OH}, \text{H}} \tilde{x}_{\text{F}} (1 - X_{\text{F}, \text{CH}_2\text{O}, \text{e}^*})^2 . \quad (5.34)$$

Following a similar approach to calculate the probability of the formation of the remaining HMF_2 configurations, x_{HMF_2} is given by

$$\begin{aligned}
 x_{\text{HMF}_2} &= x_{\text{HMF}_{2,\text{I}}} + x_{\text{HMF}_{2,\text{II}}} + x_{\text{HMF}_{2,\text{III}}} + x_{\text{HMF}_{2,\text{IV}}} + x_{\text{HMF}_{2,\text{V}}} \\
 &= \tilde{x}_{\text{M}} (1 - X_{\text{M,CH}_3\text{OH,e}})^2 X_{\text{M,CH}_3\text{OH,H}} \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,e}^*})^2 \\
 &\quad + \tilde{x}_{\text{M}} (1 - X_{\text{M,CH}_3\text{OH,e}})^2 X_{\text{M,CH}_3\text{OH,H}} \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,e}^*}) (1 - X_{\text{F,CH}_2\text{O,H}}) \\
 &\quad + \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,H}})^2 (X_{\text{F,CH}_2\text{O,e}})^2 X_{\text{F,CH}_2\text{O,e}^*} \tilde{x}_{\text{M}} (1 - X_{\text{M,CH}_3\text{OH,e}})^2 \\
 &\quad + \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,e}}) (1 - X_{\text{F,CH}_2\text{O,e}^*}) (X_{\text{F,CH}_2\text{O,H}})^2 \tilde{x}_{\text{M}} (1 - X_{\text{M,CH}_3\text{OH,e}}) (1 - X_{\text{M,CH}_3\text{OH,H}}) \\
 &\quad + \tilde{x}_{\text{F}} (1 - X_{\text{F,CH}_2\text{O,e}}) (1 - X_{\text{F,CH}_2\text{O,H}}) X_{\text{F,CH}_2\text{O,e}^*} \tilde{x}_{\text{M}} (1 - X_{\text{M,CH}_3\text{OH,e}}) (1 - X_{\text{M,CH}_3\text{OH,H}}) .
 \end{aligned} \tag{5.35}$$

By comparing the calculation of x_{HMF_2} in binary and ternary mixtures, given by Equations (5.31) and (5.35), respectively, we note that for x_{HMF_2} in ternary mixtures we use information of the sites bonded or not bonded for the two components (formaldehyde and water) in each term in Equation (5.35). This is in contrast to using the fraction of bonded sites for one of the components in each term in Equation (5.31) for the formaldehyde + methanol binary mixture. In the ternary mixture, the third component (water in this case) is competing with methanol to associate with formaldehyde, hence, we follow a more specific formulation in Equation (5.35) to ensure that the calculation is accounting for the formation of HMF_2 only. The differences in the formulations of x_{HMF_2} in binary and ternary mixtures is also shown by comparing Tables 5.8 and 5.10 for HMF_2 .

5.4.2.3 MG_1 and HMF_1

As mentioned previously, x_{HMF_1} and x_{MG_1} are determined by solving the mass balances in Equations (5.20)–(5.23) simultaneously, and using x_{MG_2} and x_{HMF_2} obtained from Equations (5.33) and (5.35), respectively. However, to solve Equations (5.20)–(5.23) the true mole fraction of formaldehyde, x_{F} , and the true mole fraction of methanol, x_{M} , or water, x_{W} , must be known.

Pure formaldehyde predominantly forms when the e^* site on CH_2O is not bonded hence, x_{F}

is given by

$$x_F = \tilde{x}_F X_{F,CH_2O,e^*} . \quad (5.36)$$

Additionally, methanol forms when none of the sites on the CH_3OH group are bonded or, when all the sites (the two e sites and one H site) on a CH_3OH group associate with sites on other CH_3OH groups. Consequently, the true mole fraction of methanol, x_M , can be calculated as

$$x_M = \tilde{x}_M [X_{M,CH_3OH,e} X_{M,CH_3OH,H} + (1 - X_{M,CH_3OH,e})^2 (1 - X_{M,CH_3OH,H})] . \quad (5.37)$$

5.4.3 Speciation predictions

To determine the distribution of species in binary and ternary mixtures containing formaldehyde we follow the methodology outlined in Sections 5.4.1 and 5.4.2, and compare our predictions to experimental data to assess the validity of our calculations. The experimental data in the literature [328, 332, 343, 345, 353] are reported for the liquid phase only. Hence, the predictions by our model are only shown for the liquid phase. It is important to note however, that the same methodology is applicable for predicting the distribution of species at any P , T , x , including in the vapour phase.

5.4.3.1 Binary mixtures

In formaldehyde + water mixtures, methylene glycol (MG_1) and poly(oxy)methylene glycols (MG_i) are formed according to Reactions (5.1) and (5.2). The expressions given in Equations (5.25) and (5.27) are used to calculate x_{MG_1} and x_{MG_2} , respectively. Good agreement between our predictions and experimental data by Albert et al. [332] can be seen in Figures 5.11 (a) and (b) at 338.15 K and 368.15 K, respectively. Very small deviations, in terms of AAD, for x_{MG_1} and x_{MG_2} are obtained at both temperatures (Table 5.11). This level of agreement is especially remarkable considering the SAFT- γ Mie results are fully predictive.

Furthermore, the mole fraction of each of the configurations of MG_1 (cf. Figures 5.3 (a)) are presented in Figures 5.11 (c) and (d). The total true mole fraction of MG_1 , x_{MG_1} , is the

sum of the mole fractions of each of the configurations. As can be seen, configuration III is found in highest concentration, followed by configuration I; the mole fraction of configuration II, in which the H site on CH₂O and the e site on H₂O are bonded, is found to be negligible. Similarly, in Figure 5.11 (e) and (f) the true mole fraction of each configuration of MG₂ (cf. Figures 5.4 (a)) at 338.15 K and 368.15 K, respectively, are shown. It is evident from Figures 5.11 (e) and (f) that configurations I, II and IV, in which the e* site on at least one of the CH₂O groups is bonded to an e site on H₂O, are present in the highest mole fractions. The mole fractions of configurations III, V and VI, which do not involve the association of any e* on CH₂O, are found to be negligible at the two temperatures.

We note that the model predictions for x_{MG_2} at 338.15 K (Figures 5.11 (a) and (e)) show that $x_{\text{MG}_2} = 8.44 \times 10^{-3}$ (instead of zero) for $\tilde{x}_F = 0$, this is due to the numerical values of $X_{\text{W,H}_2\text{O,e}}$ and $X_{\text{W,H}_2\text{O,H}}$ which are found to be slightly different from 1 and 0, respectively (cf. Equation (5.27)). This calculated value of x_{MG_2} is accounted for in the determination of the deviation of the model from experimental data (where $x_{\text{MG}_2}^{\text{exp}} = 0$ for $\tilde{x}_F = 0$) at 338.15 K (Table 5.11). A similar behaviour can be observed at 368.15 K (Figures 5.11 (b) and (f)). The corresponding calculated value of x_{MG_2} for $\tilde{x}_F = 0$ is also considered in the calculation of AAD x_{MG_2} at 368.15 K (Table 5.11).

Table 5.11: The average absolute deviation (AAD) between the SAFT- γ Mie calculations and experimental data [328, 332] for the true mole fraction of MG₁ (x_{MG_1}) and MG₂ (x_{MG_2}) and, of HMF₁, (x_{HMF_1}) and HMF₂ (x_{HMF_2}) in binary mixtures of formaldehyde + water and formaldehyde + methanol, respectively.

T / K	P / MPa	N_X^{D}	AAD x_{MG_1}	AAD x_{MG_2}	ref
338.15	0.101	7	0.0555	0.00732	[332]
368.15	0.101	11	0.0436	0.01350	[332]
T / K	P / MPa	N_X^{D}	AAD x_{HMF_1}	AAD x_{HMF_2}	ref
276	0.101	12	0.3169	0.0257	[328]

Considering the case of formaldehyde + methanol mixtures, hemiformal (HMF₁) and poly(oxy)methylene hemiformals (HMF_{*i*}) are formed according to Reactions (5.3) and (5.4), respectively. In our approach, the true mole fraction of HMF₁ and HMF₂ can be obtained using Equations (5.29) and (5.31), respectively. Very good agreement between our predictions

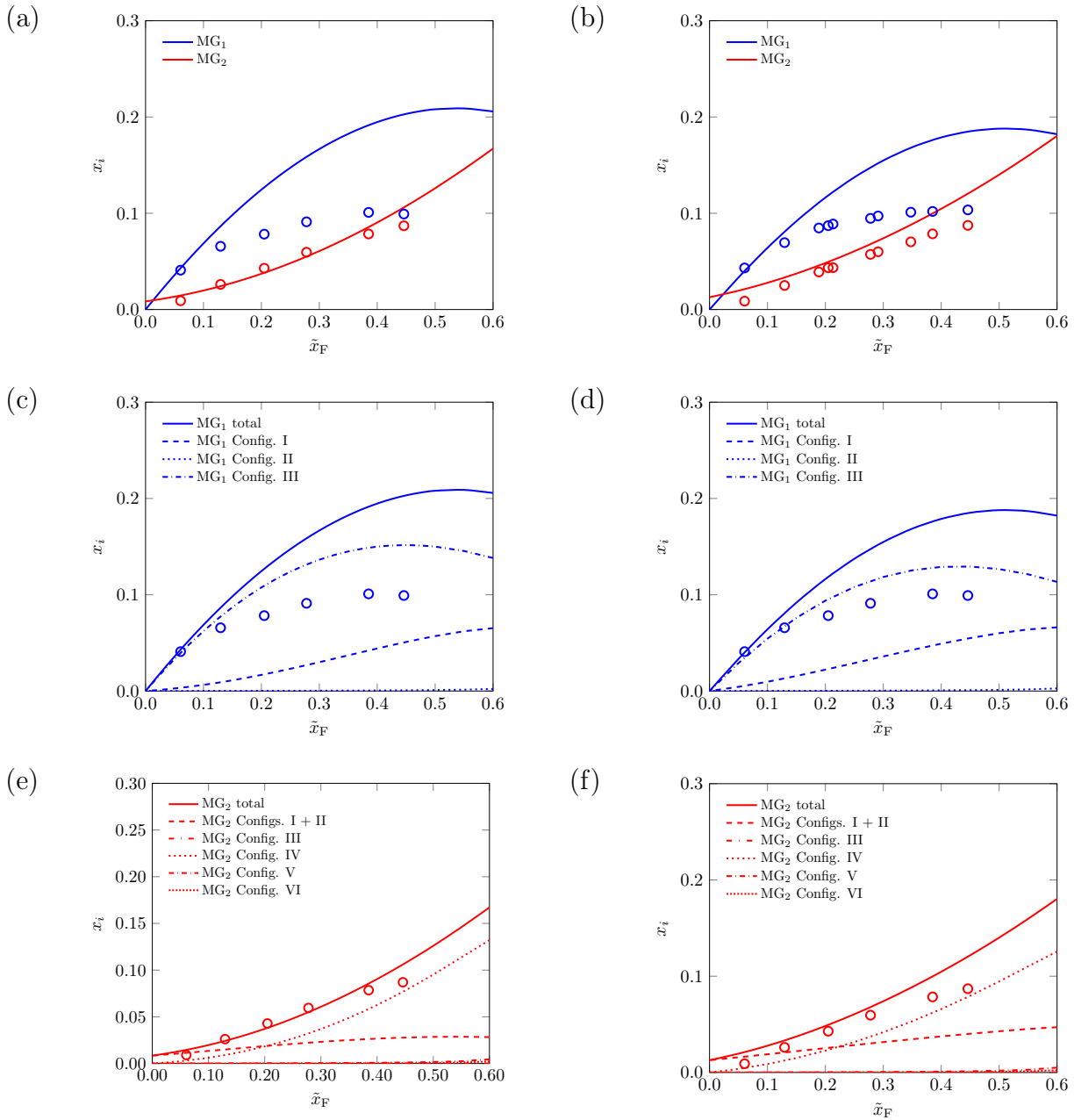


Figure 5.11: The true liquid mole fraction of MG₁ and MG₂ for a given overall formaldehyde liquid-phase mole fraction in a formaldehyde + water binary mixture at 0.101 MPa and (a) 338.15 K, and (b) 368.15 K. Symbols represent experimental data [332] and continuous curves are SAFT- γ Mie predictions. In (c) and (d), at 338.15 K and 368.15 K, respectively, dashed curves are SAFT- γ Mie predictions of the true mole fraction of the configurations of MG₁ (cf. Figure 5.3 (a)). In (e) and (f), at 338.15 K and 368.15 K, respectively, dashed curves are SAFT- γ Mie predictions of the true mole fraction of the configurations of MG₂ (cf. Figure 5.4 (a)). The continuous curves in (c)–(f) are SAFT- γ Mie predictions of the total true mole fraction of the species (the summation of the dashed curves).

and experimental data of Albert et al. [328] is shown in Figure 5.12 for x_{HMF_2} , which is also evident by the low AAD reported in Table 5.11. However, higher deviations between the SAFT- γ

Mie calculations and experimental data for x_{HMF_1} are observed in Table 5.11.

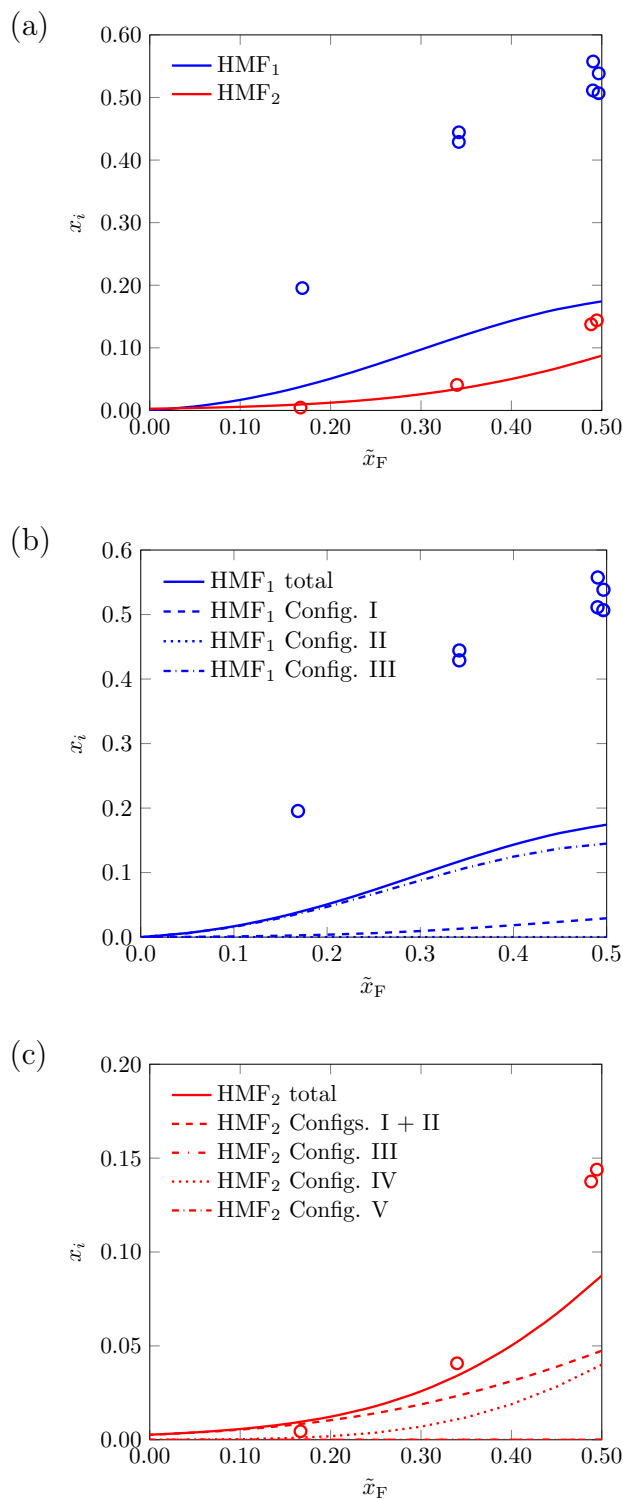


Figure 5.12: The true liquid mole fraction of HMF₁ and HMF₂ for a given overall formaldehyde liquid-phase composition in a formaldehyde + methanol binary mixture at 276 K and 0.101 MPa. The symbols correspond to experimental data [328, 343], and the curves to SAFT- γ Mie predictions of the different aggregates as indicated in each figure.

The concentration of each of the configurations of HMF_1 and HMF_2 shown in Figures 5.3 (b) and 5.4 (b) are presented in Figure 5.12 together with the corresponding total true mole fractions at each $\tilde{x}_F$. For HMF_1 , configuration III is seen in the highest mole fraction, followed by configuration I; the mole fraction of configuration II is found to be negligible. It is shown in Figure 5.12 (c), that for HMF_2 , configurations I, II and IV have the highest mole fractions, with the remaining configurations having negligible mole fractions in this mixture.

We note, that the model predictions for x_{HMF_2} at 276 K (Figures 5.12 (a) and (c)) show that $x_{\text{HMF}_2} = 2.69 \times 10^{-3}$ (instead of zero) for $\tilde{x}_F = 0$, this is due to the numerical values of $X_{\text{M,CH}_3\text{OH,e}}$ and $X_{\text{M,CH}_3\text{OH,H}}$ which are found to be slightly different from 1 and 0, respectively (cf. Equation (5.31)). This calculated value of x_{HMF_2} is accounted for in the determination of the deviation of the model from experimental data (where $x_{\text{HMF}_2}^{\text{exp}} = 0$ for $\tilde{x}_F = 0$) in terms of the AAD of x_{HMF_2} at 276 K (Table 5.11).

5.4.3.2 Ternary mixtures

In the ternary (formaldehyde + water + methanol) system, methylene glycol MG_1 , hemiformal HMF_1 , poly(oxy)methylene glycols MG_i and poly(oxy)methylene hemiformals HMF_i are formed. The SAFT- γ Mie EoS can be used to predict the true mole fractions of MG_1 , MG_2 , HMF_1 and HMF_2 , x_{MG_1} , x_{MG_2} , x_{HMF_1} and x_{HMF_2} , respectively. The methodology discussed in Section 5.4.2 and Equations (5.20)–(5.23) are used to predict the distribution of reaction species in the current work.

In Figure 5.13 a comparison of the SAFT- γ Mie predictions for the distribution of species is compared with experimental data. Two experimental data sets [345] (A and B) are considered. The experimental data are obtained using ^{13}C NMR spectroscopy, and the difference between the two experimental data sets is the NMR techniques and instruments used [345]. Very good agreement between our calculations and experimental data can be seen in the figures. The corresponding AADs for all the mixtures are shown in Table 5.12. Particularly accurate predictions are obtained for x_{MG_2} , x_{HMF_1} and x_{HMF_2} with overall AADs $x_i < 0.02$. However, an over-estimation of x_{MG_1} can be seen in all the ternary mixtures of Figure 5.13 which leads to a

relatively higher AAD for x_{MG_1} (Table 5.12). This is, in part, due to the assumption used in our calculation of the concentration of species', that the maximum oligomer length of methylene glycols that form, is MG_2 , such that $x_{\text{MG}_i} = 0$ for $i = 3, 4, \dots, \infty$. The over estimation of x_{MG_1} indicates that perhaps longer MG_i oligomers have non-negligible mole fractions. To overcome this limitation and confirm the formation of longer oligomers, the lattice-based Flory-Huggins theory [354, 355], which is commonly used to study the thermodynamic behaviour of polymer solutions, can be used in future work to determine the concentration of all the oligomers that form (i.e. x_{MG_i} and x_{HMF_i} for $i = 1, 2, \dots, \infty$).

Table 5.12: The average absolute deviation (AAD) between the SAFT- γ Mie predictions and experimental data [345] for the true mole fraction of MG_1 (x_{MG_1}), MG_2 (x_{MG_2}), HMF_1 (x_{HMF_1}) and HMF_2 (x_{HMF_2}) in the liquid phase of formaldehyde + water + methanol ternary mixtures at $P = 0.101$ MPa, corresponding to Figure 5.13. Here, one data point is used to calculate the AADs in each case ($N_X^{\text{D}} = 1$).

T / K	$\tilde{x}_{\text{F}}$	$\tilde{x}_{\text{M}}$	AAD x_{MG_1}	AAD x_{MG_2}	AAD x_{HMF_1}	AAD x_{HMF_2}
348	0.3047	0.2087	9.48×10^{-2}	9.98×10^{-4}	5.02×10^{-3}	2.10×10^{-3}
348	0.3323	0.2933	8.08×10^{-2}	3.19×10^{-3}	2.55×10^{-3}	3.60×10^{-3}
298	0.3047	0.2087	1.60×10^{-1}	3.08×10^{-5}	4.71×10^{-2}	1.37×10^{-2}
All			1.12×10^{-1}	1.41×10^{-3}	1.82×10^{-2}	6.46×10^{-3}

The true mole fraction of each configuration considered for MG_2 and HMF_2 in the ternary mixtures calculated from the fraction of molecules not bonded at given sites from the SAFT- γ Mie EoS using Equations (5.33) and (5.35), respectively, are shown in Figure 5.13. As can be seen, both in MG_2 and HMF_2 , configurations I, II and IV, which involve the association of the e^* site of CH_2O , are present in the highest mole fractions. For HMF_2 and MG_2 this result is consistent with the findings for the binary mixtures, and is expected given the relatively higher association strength of the e sites of CH_3OH and H_2O with the e^* site on CH_2O (cf. Table 5.4).

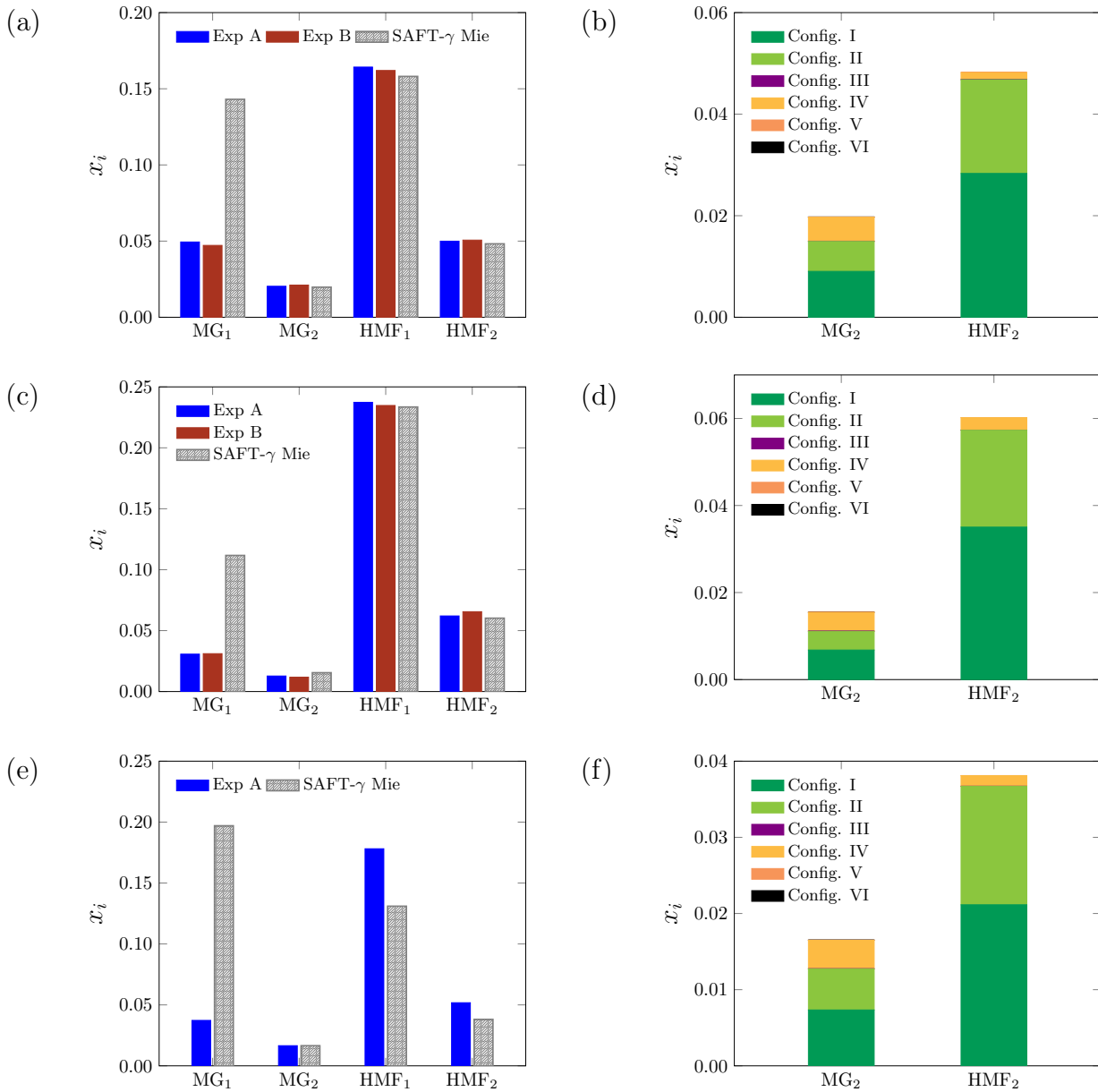


Figure 5.13: The true liquid-phase mole fraction ((a), (c) and (e)) of MG_1 , MG_2 , HMF_1 and HMF_2 and ((b), (d) and (f)) of each of the configurations of MG_2 (cf. Figure 5.4 (a)) and HMF_2 (cf. Figure 5.4 (b)) in ternary mixtures of formaldehyde + water + methanol. In (a) and (b) the ternary mixture is at $T = 348$ K, $P = 0.101$ MPa and overall composition $\tilde{x}_F = 0.3047$, $\tilde{x}_W = 0.4866$ and $\tilde{x}_M = 0.2087$. In (c) and (d) the ternary mixture is at $T = 348$ K, $P = 0.101$ MPa and overall composition $\tilde{x}_F = 0.3323$, $\tilde{x}_W = 0.3684$ and $\tilde{x}_M = 0.2933$. In (e) and (f) the ternary mixture is at $T = 298$ K, $P = 0.101$ MPa and overall composition of $\tilde{x}_F = 0.3047$, $\tilde{x}_W = 0.4866$ and $\tilde{x}_M = 0.2087$. The blue and red bars correspond to experimental data [345] and gray bars to SAFT- γ Mie predictions. The colored bars in (b), (d) and (f) are SAFT- γ Mie predictions.

5.5 Conclusions

We have presented the development of SAFT- γ Mie group-contribution (GC) models to predict the thermodynamic fluid-phase equilibria in binary and ternary mixtures containing formalde-

hyde with water and methanol. In the modelling of the phase equilibria of these mixtures, we account for the oligomerization reactions that occur in the liquid and vapour phases. We follow a physical approach to modelling the reactions, in which the reaction products are treated implicitly and are considered to be aggregates of the reactants.

We develop the SAFT- γ Mie parameters for the new CH_2O group representing formaldehyde by using experimental data of the saturated liquid density, vapour pressure and enthalpy of vaporization in pure formaldehyde for the parameter estimation. Previously published parameters for CH_3OH and H_2O are used in the current work. The unlike interaction parameters between CH_2O and H_2O , and between CH_2O and CH_3OH , are obtained by fitting to experimental vapour-liquid equilibria (VLE) data of binary mixtures of formaldehyde + water and formaldehyde + methanol, respectively. The newly obtained parameters are validated by comparing the SAFT- γ Mie predictions of the VLE in binary mixtures containing formaldehyde to experimental data which are not included in the parameter estimation. Excellent agreement with experimental data is obtained, and the model is able to capture the azeotropes in the VLE profiles of these binary mixtures. The VLE in ternary mixtures of formaldehyde + water + methanol is predicted using the SAFT- γ Mie GC approach without fitting any additional parameters; obtaining a close agreement to experimental data with an overall %AAD of 3.78%.

In our model, the oligomerization reactions of formaldehyde with water and methanol are considered to take place in the liquid and vapour phases. The vapour phase is treated as a non-ideal gas, with the thermodynamic properties calculated using the SAFT- γ Mie equation of state (EoS). The reactions taking place in the liquid and vapour phases are modelled using a reactive, e^* , site on the formaldehyde group (CH_2O), to mediate the formation of reaction products. The reactive e^* site is only allowed to associate with the e sites on the methanol (CH_3OH) and water (H_2O) groups, and it remains inactive in solutions of pure formaldehyde. Other association interactions, which represent the hydrogen bonding that can take place between the groups in a mixture, are considered as usual within the SAFT- γ Mie GC approach.

The nature of speciation and distribution of reaction products in binary and ternary mixtures of formaldehyde with water and methanol is investigated using information of the fraction

of molecules i not bonded at a given site a on group k , $X_{i,k,a}$, obtained from the SAFT- γ Mie EoS. Statistical arguments are used to determine the true mole fraction of methylene glycol (MG_1), dioxymethylene glycol (MG_2), hemiformal (HMF_1) and dioxymethylene hemiformal (HMF_2) in formaldehyde + water, formaldehyde + methanol and formaldehyde + water + methanol mixtures, where the bonded fractions are treated as probabilities. Very good agreement to experimental data is obtained. Using our methodology, it is possible to determine the configuration which is most likely to form for each oligomer. Interestingly, we find that the MG_2 and HMF_2 configurations involving the association of the reactive e^* site of CH_2O exhibit the highest mole fractions in the ternary mixtures.

The findings from our current work demonstrate the applicability of the SAFT- γ Mie GC approach for studying the complex mixtures containing formaldehyde, in particular, for understanding the phase behaviour and the nature of speciation in binary and ternary mixtures of formaldehyde with water and methanol; which can influence the chemical reaction kinetics and transport properties of these mixtures. This can be done without requiring any experimental information for these mixtures. The predictive SAFT- γ Mie model developed in the current work is relevant to various industrial applications involving formaldehyde, and is particularly useful for enhancing the stability and extending the shelf-life of drugs in the pharmaceutical industry.

Chapter 6

Conclusions

In recent years, efforts in the pharmaceutical industry have been focused on the adoption of more-digital technologies in every stage of drug product development, from the early stages of drug discovery to the manufacturing and supply of the drug product. Lately, it has become apparent in the industry that using process systems engineering (PSE) concepts can speed up the development of continuous and greener manufacturing practices.

Pharmaceutical process design heavily relies on predictive thermodynamic modelling tools to gain an understanding of the physicochemical properties of the system of interest and to identify required laboratory work. Experiments are fundamental for pharmaceutical R&D, however, through using modelling tools experimental efforts can be focused on the most promising areas, enabling more efficient drug product development and manufacturing. Some of the commonly used thermodynamic modelling tools for this purpose in the industry include, COSMO-RS, COSMO-SAC and NRTL-SAC. Additionally, group contribution (GC) approaches have increased in popularity in the industry due to their high predictive power, with the use of UNIFAC, an activity coefficient based method, being the most common. Another category of GC approaches that have been gaining increasing interest for pharmaceutical applications are equations of state, in particular, the statistical associating fluid theory (SAFT).

An overview of the history of SAFT is provided in Chapter 2, in addition to a description of the various homonuclear and heteronuclear SAFT approaches that have been developed over

the years. We present a review of the application of these SAFT approaches to pharmaceutical systems containing APIs (small molecules), amino acids and peptides (medium and large molecules). Following this, we introduce the theory of the SAFT- γ Mie GC approach which can be used to predict thermodynamic properties of molecular systems and electrolytes. The methodology for calculating fluid phase equilibria and solid-liquid solubility using this GC approach is then discussed.

In Chapter 3, we demonstrate the applicability of the SAFT- γ Mie GC equation of state (EoS) for predicting the phase behaviour and pH-solubility profiles of two acidic ionisable APIs, ibuprofen and ketoprofen, in aqueous solution. Group interaction parameters presented in previous work are used, with the explicit intention to examine the predictive power of the approach. The various types of phase equilibria that take place at atmospheric pressure and across a wide temperature range, including vapour-liquid equilibria (VLE), liquid-liquid equilibria (LLE), and solid-liquid equilibria (SLE) at very low mole fractions of the APIs ($\approx 10^{-6}$) are predicted. Our calculations highlight the difficulty in carrying out experimental measurements at such low concentrations and perhaps explain the lack of available literature data for many APIs. We obtain very good agreement with available SLE solubility data of ibuprofen and ketoprofen, and for the temperature range over which LLE is observed in ibuprofen solutions. By considering the speciation of the APIs, the acidic buffering agents, and the inorganic and organic counter ions in solution, we develop a model to predict the pH-solubility profile of the APIs in aqueous solutions. Excellent agreement with experimental data for the pH-solubility profile predictions of aqueous solutions of ibuprofen and ketoprofen buffered by HCl and NaOH at 298.15 K and 310.15 K, is obtained. In view of the pharmaceutical industry moving towards the use of more-diverse counter ions, we extend our predictions to treat several amines as the source of organic counter ions in aqueous ibuprofen solutions. We demonstrate the capability of the SAFT- γ Mie GC approach in handling complex mixtures of molecular and ionic groups. The findings from this chapter demonstrate the applicability of the SAFT- γ Mie GC approach for pharmaceutical systems, particularly for understanding the phase behaviour of APIs and the pH-solubility profiles, including the relevance of this tool for salt selection and disproportionation.

Subsequently, the model presented for predicting the pH-solubility profiles of monoprotic

APIs in Chapter 3 is extended in Chapter 4 to study the solubility of ionisable diprotic APIs. The SAFT- γ Mie GC approach is used to model the intrinsic and pH-dependent solubility of procaine, a weakly basic diprotic API. First, in order to model the molecular and ionised forms of procaine within the SAFT- γ Mie GC framework, we expand the SAFT- γ Mie matrix of groups by introducing new groups and their unlike interactions with other groups. Hence, the parameter estimation workflow within the SAFT- γ Mie approach is demonstrated in this chapter through the development of three new groups: the aCCOO, aCNH₂ and aCNH₃⁺ groups, in addition to their unlike interactions with other groups in the mixtures. Consequently, the intrinsic solubility of procaine in aqueous solutions is predicted using the SAFT- γ Mie GC approach and is shown to be in good agreement with experimental data. By extending the model presented in Chapter 3, a thermodynamic framework to predict the pH-solubility profile of diprotic APIs is developed, using procaine as an example. Overall good agreement to experimental data is obtained using HCl and NaOH as buffering agents to develop the pH-solubility profile of the API, including the precipitation of the procaine hydrochloride salt. The findings from this chapter demonstrate the transferability of the SAFT- γ Mie groups and the applicability of the approach for predicting thermodynamic properties of diprotic APIs. Furthermore, we present an outlook on the limitations of the developed framework and discuss the additional experimental data required for accounting for the precipitation of the dihydrochloride salt of procaine.

Following that, in Chapter 5 we study the phase behaviour of more-complex pharmaceutical mixtures containing formaldehyde, an undesirable degradation agent affecting the shelf-life of pharmaceutical drugs. Among other sources, formaldehyde can be produced by some excipients during storage, negatively impacting the stability of drugs. Therefore, in this chapter we study the thermodynamic fluid-phase equilibria of mixtures containing formaldehyde. In particular, we study the phase behaviour of binary and ternary mixtures of formaldehyde with water and methanol. These mixtures are of interest as formaldehyde, due to its high reactivity, is usually stored, sold and processed in the form of aqueous solutions, with methanol added for enhanced stability. The SAFT- γ Mie GC approach is used to model the VLE of these mixtures, accounting for the oligomerization reactions that occur in the liquid and vapour phases. We follow a

physical approach to modelling the reactions, in which the reaction products are treated implicitly and are considered to be aggregates of the reactants. In terms of VLE calculations, excellent agreement with experimental data is obtained for binary mixtures of formaldehyde + water and formaldehyde + methanol. Following that, the VLE in ternary mixtures of formaldehyde + water + methanol is accurately predicted using the GC approach, without fitting any additional parameters. Moreover, the distribution of reaction products in binary and ternary mixtures of formaldehyde with water and methanol is investigated using information of the fraction of molecules not bonded at a given association site on a specific group, which can be obtained from the SAFT- γ Mie EoS. Consequently, it is possible to determine the concentration of the oligomers at various conditions of temperature and pressure, showcasing the applicability of the SAFT- γ Mie GC approach for studying the speciation occurring in complex mixtures.

6.1 Summary of key achievements

The key contributions presented in this thesis can be summarized as follows:

- The development of a predictive framework to determine the pH-dependent solubility of ionisable active pharmaceutical ingredients (APIs). The SAFT- γ Mie group-contribution (GC) approach is used predictively to obtain the pH-solubility profiles of two monoprotic APIs, ibuprofen and ketoprofen, in aqueous solution containing various buffering agents. Equilibrium constants for the dissociation of the APIs and the formation of salts (pK_a and K_{sp} values) are also incorporated in the model using experimental values from literature. Excellent agreement with experimental data is obtained for the pH-solubility profile predictions of the APIs in aqueous solutions buffered by HCl and NaOH at 298.15 K and 310.15 K.
- The capability of the SAFT- γ Mie GC approach in handling complex mixtures of molecular and ionic groups is demonstrated by using various inorganic bases as buffering agents and as sources of counter ions to develop the pH-solubility profiles of ibuprofen and ketoprofen in water. Additionally, by using several amines as the source of organic counter

ions in aqueous ibuprofen solutions, we could use the predicted pH-solubility profiles to determine the values of the thermodynamic K_{sp} and pH^{max} for the ibuprofen amine salts; highlighting the relevance of the SAFT- γ Mie GC approach for salt selection and disproportionation.

- The prediction of the complete phase diagrams of ibuprofen and ketoprofen in aqueous solutions using the SAFT- γ Mie GC approach is demonstrated, including the solid–liquid equilibria (SLE), the liquid–liquid equilibria (LLE) (also known as oiling-out or liquid–liquid phase separation) and the vapour–liquid equilibrium (VLE) at very low mole fractions of the APIs ($\approx 10^{-6}$). Our calculations emphasize the difficulty in performing experimental measurements at such low concentrations, possibly explaining the lack of available data for many APIs in the literature.
- The SAFT- γ Mie approach is used to study thermodynamic properties of diprotic compounds for the first time. In particular, the GC approach is used to calculate the intrinsic solubility and the pH-solubility profile of the diprotic basic API procaine. Predictions of the temperature- and pH-dependent solubility of the API in water are in very good agreement to experimental data; highlighting the applicability of the approach for studying diprotic, in addition to monoprotic, APIs.
- The SAFT- γ Mie matrix of groups has been extended in this thesis to study various mixtures of ionisable APIs in aqueous solutions. Four new groups are introduced, namely, the NH_3^+ group, to study mixtures of ibuprofen with amine buffers, and the $aCCOO$, $aCNH_2$ and $aCNH_3^+$ groups to model procaine mixtures; in addition to their unlike interactions with other groups present in the systems of interest. The group parameters are estimated using experimental data of simpler compounds containing the groups after which they are used to predict properties of the APIs within the SAFT- γ Mie framework. The good agreement between experimental data and the SAFT- γ Mie predictions for API mixtures, presented in this thesis, showcase the transferability of the SAFT- γ Mie group interactions, considering that no data related to the APIs are used to obtain the group interaction parameters.

- The thermodynamic fluid-phase equilibria in binary and ternary mixtures containing formaldehyde (an API degradation agent produced by some excipients) with water and methanol is predicted using the SAFT- γ Mie GC approach. The oligomerization reactions taking place in the liquid and vapour phases are considered in the modelling of the phase behaviour of these mixtures. A physical approach to modelling the reactions is followed, in which the reaction products are treated implicitly and are considered to be aggregates of the reactants. Excellent agreement with experimental data is obtained for the VLE phase boundaries of binary and ternary mixtures of formaldehyde with the solvents. Moreover, the nature of speciation and the distribution of reaction products in formaldehyde + water, formaldehyde + methanol and formaldehyde + water + methanol mixtures is successfully predicted using the SAFT- γ Mie equation of state.

6.2 Directions for future work

In this thesis, the SAFT- γ Mie GC approach has been shown to deliver accurate property predictions for pharmaceutical systems, which can be useful for focusing experimental efforts on the necessary experiments, ultimately saving time and costs for the industry. To continue applying this modelling tool for pharmaceutical mixtures however, an important avenue for future work, is the expansion of the SAFT- γ Mie groups database. Pharmaceutical compounds are multifunctional molecules which exhibit complex speciation in solutions. Therefore, the continuous development of new SAFT- γ Mie group parameters is necessary to include the molecular and ionic functional groups often found in pharmaceutical mixtures.

In Chapters 3 and 4 a predictive thermodynamic framework for developing the pH-solubility profile of small molecule APIs (monoprotic and diprotic) was presented. In future work, this framework can be used to study the pH-dependent solubility of medium to large molecules such as amino acids and peptides. In recent years, large molecules (e.g. peptides), which are composed of a defined sequence of amino acids, have gained increased attention from the pharmaceutical industry in order to develop new therapies for unmet medical needs [356].

Amino acids and peptides are zwitterionic in nature and are predominantly found in the ionised form in solution. Hence, accurate solubility predictions of amino acids and peptides are only possible when pH effects, and the consequent ionisation of the species, are considered.

Another interesting area for future work is modelling the oligomerization reactions of formaldehyde with water and methanol via the chemical approach, using the SAFT- γ Mie EoS. In the chemical approach, all the reaction products have to be accounted for explicitly and the reactions are represented in terms of the reaction equilibrium constants. Following that, a comparison can be done with the physical approach used for modelling these reactions in Chapter 5, to show, and further confirm, the equivalence of the two approaches for modelling the oligomerization reactions when studying the phase behaviour and the nature of speciation in formaldehyde mixtures.

6.3 Publications and conferences

6.3.1 Publications

- **Wehbe, M.**, Haslam, A. J., García-Muñoz, S., Jackson, G. & Galindo, A. Thermodynamic modelling of the nature of speciation and phase behavior of binary and ternary mixtures of formaldehyde, water and methanol (submitted to Thermodynamics 2022 Special Issue in *Mol. Phys.*).
- **Wehbe, M.**, Haslam, A. J., Jackson, G. & Galindo, A. Phase behaviour and pH-solubility profile prediction of aqueous buffered solutions of ibuprofen and ketoprofen. *Fluid Phase Equilib.* **560**, 113504 (2022).
- Haslam, A. J., González-Pérez, A., Di Lecce, S., Khalit, S. H., Perdomo, F. A., Kournopoulos, S., Kohns, M., Lindeboom, T., **Wehbe, M.**, Febra, S., Jackson, G., Adjiman, C. S. & Galindo, A. Expanding the Applications of the SAFT- γ Mie Group-Contribution Equation of State: Prediction of Thermodynamic Properties and Phase Behavior of Mixtures. *J. Chem. Eng. Data* **65**, 5862–90 (2020).

6.3.2 Conference presentations and posters

- Thermodynamic modelling of the nature of speciation and phase behavior of binary and ternary mixtures of formaldehyde, water and methanol. AIChE Annual Meeting 2022 in Phoenix, AZ, USA.
- Thermodynamic modelling of the nature of speciation and phase behavior of binary and ternary mixtures of formaldehyde, water and methanol. Thermodynamics Conference 2022 in Bath, England, UK.
- Thermodynamic modelling of the nature of speciation and phase behavior of binary and ternary mixtures of formaldehyde, water and methanol. 32nd European Symposium of Applied Thermodynamics (ESAT) 2022 in Graz, Austria.
- Predicting the phase behaviour and pH-solubility profiles of active pharmaceutical ingredients using the SAFT- γ Mie group-contribution approach. 21st Symposium on Thermophysical Properties 2021 in Colorado, USA (virtual).
- Prediction of phase diagrams and pH-solubility profiles of active pharmaceutical ingredients Using the SAFT- γ Mie group-contribution approach. 31st European Symposium on Applied Thermodynamics (ESAT) 2021 in Paris, France (virtual).
- Predicting optimal salt forms for active pharmaceutical ingredients using the SAFT- γ Mie equation of state. AIChE Annual Meeting 2020 in San Francisco, USA (virtual).
- Prediction of the solubility of active pharmaceutical ingredients under varying conditions of pH and temperature using the SAFT- γ Mie equation of state. AIChE Annual Meeting 2019 in Orlando, FL, USA.

Bibliography

- [1] A. Schuhmacher, O. Gassmann, M. Hinder, Changing R&D models in research-based pharmaceutical companies, *Journal of Translational Medicine* 14 (1) (2016) 1–11. doi:10.1186/s12967-016-0838-4.
- [2] J.-P. Garnier, Rebuilding the R&D Engine in Big Pharma., Tech. Rep. 5 (2008).
- [3] S. M. Paul, D. S. Mytelka, C. T. Dunwiddie, C. C. Persinger, B. H. Munos, S. R. Lindborg, A. L. Schacht, How to improve RD productivity: The pharmaceutical industry’s grand challenge, *Nature Reviews Drug Discovery* 9 (3) (2010) 203–214. doi:10.1038/nrd3078.
- [4] K. V. Gernaey, A. E. Cervera-Padrell, J. M. Woodley, A perspective on PSE in pharmaceutical process development and innovation, *Computers and Chemical Engineering* 42 (2012) 15–29. doi:10.1016/j.compchemeng.2012.02.022.
- [5] P. Mckenzie, S. Kiang, J. Tom, A. E. Rubin, M. Futran, Can Pharmaceutical Process Development Become High Tech?, *AIChE Journal* 52 (12) (2006) 3990–3994. doi:10.1002/aic.11022.
- [6] E. Hendriks, G. M. Kontogeorgis, R. Dohrn, J.-C. de Hemptinne, I. G. Economou, L. F. Žilnik, V. Vesovic, Industrial Requirements for Thermodynamics and Transport Properties, *Industrial & Engineering Chemistry Research* 49 (22) (2010) 11131–11141. doi:10.1021/ie101231b.
- [7] G. M. Kontogeorgis, R. Dohrn, I. G. Economou, J.-C. de Hemptinne, A. t. Kate, S. Kuitunen, M. Mooijer, L. F. Žilnik, V. Vesovic, Industrial requirements for thermo-

- dynamic and transport properties: 2020, *Industrial & Engineering Chemistry Research* 60 (13) (2021) 4987–5013. doi:10.1021/acs.iecr.0c05356.
- [8] A. Klamt, Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena, *The Journal of Physical Chemistry* 99 (7) (1995) 2224–2235.
- [9] A. Klamt, F. Eckert, COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids, *Fluid Phase Equilibria* 172 (1) (2000) 43–72. doi:10.1016/S0378-3812(00)00357-5.
- [10] A. Klamt, V. Jonas, T. Bürger, J. C. Lohrenz, Refinement and parametrization of COSMO-RS, *Journal of Physical Chemistry A* 102 (26) (1998) 5074–5085. doi:10.1021/JP980017S.
- [11] S.-T. Lin, S. I. Sandler, A Priori Phase Equilibrium Prediction from a Segment Contribution Solvation Model, *Industrial & Engineering Chemistry Research* 41 (5) (2002) 899–913. doi:10.1021/IE001047W.
- [12] A. Klamt, F. Eckert, M. Hornig, M. E. Beck, T. Bürger, Prediction of aqueous solubility of drugs and pesticides with COSMO-RS, *Journal of Computational Chemistry* 23 (2) (2002) 275–281. doi:10.1002/JCC.1168.
- [13] A. Klamt, F. Eckert, M. Hornig, COSMO-RS: A novel view to physiological solvation and partition questions, *Journal of Computer-Aided Molecular Design* 15 (4) (2001) 355–365. doi:10.1023/A:1011111506388.
- [14] H. H. Tung, J. Tabora, N. Variankaval, D. Bakken, C. C. Chen, Prediction of Pharmaceutical Solubility Via NRTL-SAC and COSMO-SAC, *Journal of Pharmaceutical Sciences* 97 (5) (2008) 1813–1820. doi:10.1002/JPS.21032.
- [15] C.-C. Chen, Y. Song, Solubility Modeling with a Nonrandom Two-Liquid Segment Activity Coefficient Model, *Industrial & Engineering Chemistry Research* 43 (26) (2004) 8354–8362. doi:10.1021/IE049463U.

- [16] P. Crafts, The role of solubility modeling and crystallization in the design of active pharmaceutical ingredients, in: N. Ka M., G. Rafiqul, D.-J. Kim (Eds.), *Computer Aided Chemical Engineering*, Vol. 23, Elsevier, Oxford, 2007, Ch. 2, pp. 23–85. doi:10.1016/S1570-7946(07)80005-8.
- [17] C.-C. Chen, P. A. Crafts, Correlation and prediction of drug molecule solubility in mixed solvent systems with the Nonrandom Two-Liquid Segment Activity coefficient (NRTL-SAC) model, *Industrial & Engineering Chemistry Research* 45 (13) (2006) 4816–4824. doi:10.1021/ie051326p.
- [18] A. Fredenslund, R. L. Jones, J. M. Prausnitz, Group-contribution estimation of activity coefficients in nonideal liquid mixtures, *AIChE Journal* 21 (6) (1975) 1086–1099. doi:10.1002/aic.690210607.
- [19] V. Papaioannou, C. S. Adjiman, G. Jackson, A. Galindo, Group Contribution Methodologies for the Prediction of Thermodynamic Properties and Phase Behavior in Mixtures, in: E. N. Pistikopoulos, M. Georgiadis, V. Dua, C. Adjiman, A. Galindo (Eds.), *Process Systems Engineering*, Vol. 6, John Wiley & Sons, London, 2011, Ch. 4, pp. 135–172. doi:10.1002/9783527631315.ch4.
- [20] T. C. Frank, J. R. Downey, S. K. Gupta, Quickly screen solvents for organic solids, *Chemical Engineering Progress* 95 (12) (1999) 41–61.
- [21] J. Abildskov, J. P. O’Connell, Predicting the solubilities of complex chemicals I. Solutes in different solvents, *Industrial & Engineering Chemistry Research* 42 (22) (2003) 5622–5634.
- [22] J. Abildskov, J. P. O’Connell, Prediction of solubilities of complex medium-sized chemicals. II. Solutes in mixed solvents, *Molecular Simulation* 30 (6) (2004) 367–378.
- [23] J. Abildskov, J. P. O’Connell, Thermodynamic method for obtaining the solubilities of complex medium-sized chemicals in pure and mixed solvents, *Fluid Phase Equilibria* 228–229 (2005) 395–400.

- [24] S. Gracin, T. Brink, A. C. Rasmuson, Prediction of solubility of solid organic compounds in solvents by UNIFAC, *Industrial & Engineering Chemistry Research* 41 (20) (2002) 5114–5124.
- [25] J. M. Prausnitz, R. N. Lichtenthaler, E. Gomes de Azevedo, *Molecular thermodynamics of fluid-phase equilibria*, 3rd Edition, Prentice-Hall Inc., New Jersey, 1998.
- [26] P. Kolář, J.-W. Shen, A. Tsuboi, T. Ishikawa, Solvent selection for pharmaceuticals, *Fluid Phase Equilibria* 194-197 (2002) 771–782.
- [27] R. Gani, J. Abildskov, G. Kontogeorgis, Application of property models in chemical product design, in: *Computer Aided Property Estimation for Process and Product Design*, Elsevier, 2004.
- [28] H. Modarresi, E. Conte, J. Abildskov, R. Gani, P. Crafts, Model-based calculation of solid solubility for solvent selection- A review, *Industrial & Engineering Chemistry Research* 47 (15) (2008) 5234–5242.
- [29] S. P. Pinho, C. M. Silva, E. A. Macedo, Solubility of Amino Acids: A Group-Contribution Model Involving Phase and Chemical Equilibria, *Industrial & Engineering Chemistry Research* 33 (1994) 1341–1347.
- [30] R. B. Gupta, R. A. Heidemann, Solubility models for amino acids and antibiotics, *AIChE Journal* 36 (3) (1990) 333–341. doi:10.1002/AIC.690360303.
- [31] G. Jackson, W. G. Chapman, K. E. Gubbins, Phase equilibria of associating fluids spherical molecules with multiple bonding sites, *Molecular Physics* 65 (1) (1988) 1–31. doi:10.1080/00268978800100821.
- [32] W. G. Chapman, G. Jackson, K. E. Gubbins, Phase equilibria of associating fluids chain molecules with multiple bonding sites, *Molecular Physics* 65 (5) (1988) 1057–1079. doi:10.1080/00268978800101601.
- [33] W. G. Chapman, K. E. Gubbins, G. Jackson, M. Radosz, SAFT: Equation-of-state

- solution model for associating fluids, *Fluid Phase Equilibria* 52 (1989) 31–38. doi:10.1016/0378-3812(89)80308-5.
- [34] W. G. Chapman, K. E. Gubbins, G. Jackson, M. Radosz, New Reference Equation of State for Associating Liquids, *Industrial & Engineering Chemistry Research* 29 (8) (1990) 1709–1721.
- [35] V. Papaioannou, T. Lafitte, C. Avendaño, C. S. Adjiman, G. Jackson, E. A. Müller, A. Galindo, Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments, *Journal of Chemical Physics* 140 (5) (2014) 054107. doi:10.1063/1.4851455.
- [36] S. Dufal, V. Papaioannou, M. Sadeqzadeh, T. Pogiatis, A. Chremos, C. S. Adjiman, G. Jackson, A. Galindo, Prediction of thermodynamic properties and phase behavior of fluids and mixtures with the SAFT- γ mie group-contribution equation of state, *Journal of Chemical & Engineering Data* 59 (10) (2014) 3272–3288. doi:10.1021/je500248h.
- [37] S. Dufal, T. Lafitte, A. J. Haslam, A. Galindo, G. N. I. Clark, C. Vega, G. Jackson, The A in SAFT: Developing the contribution of association to the Helmholtz free energy within a Wertheim TPT1 treatment of generic Mie fluids, *Molecular Physics* 113 (9-10) (2015) 948–984. doi:10.1080/00268976.2015.1029027.
- [38] S. Dufal, T. Lafitte, A. J. Haslam, A. Galindo, G. N. I. Clark, C. Vega, G. Jackson, Corrigendum, *Molecular Physics* 116 (2) (2018) 283–285. doi:10.1080/00268976.2017.1402604.
- [39] P. Hutacharoen, S. Dufal, V. Papaioannou, R. M. Shanker, C. S. Adjiman, G. Jackson, A. Galindo, Predicting the Solvation of Organic Compounds in Aqueous Environments: From Alkanes and Alcohols to Pharmaceuticals, *Industrial & Engineering Chemistry Research* 56 (38) (2017) 10856–10876. doi:10.1021/acs.iecr.7b00899.
- [40] P. Hutacharoen, Prediction of Partition Coefficients and Solubilities of Active Pharmaceutical Ingredients with the SAFT- γ Mie Group-contribution Approach, Ph.D. thesis, Imperial College London (2017).

- [41] S. A. Febra, T. Bernet, C. Mack, J. McGinty, I. I. Onyemelukwe, S. J. Urwin, J. Seifick, J. H. ter Horst, C. S. Adjiman, G. Jackson, A. Galindo, Extending the SAFT- γ Mie approach to model benzoic acid, diphenylamine, and mefenamic acid: Solubility prediction and experimental measurement, *Fluid Phase Equilibria* 540 (2021) 113002. doi:10.1016/j.fluid.2021.113002.
- [42] S. Di Lecce, G. Lazarou, S. H. Khalit, C. S. Adjiman, G. Jackson, A. Galindo, L. McQueen, Modelling and prediction of the thermophysical properties of aqueous mixtures of choline geranate and geranic acid (CAGE) using SAFT- γ Mie, *RSC Advances* 9 (65) (2019) 38017–38031. doi:10.1039/C9RA07057E.
- [43] S. Di Lecce, G. Lazarou, S. H. Khalit, D. Pugh, C. S. Adjiman, G. Jackson, A. Galindo, L. McQueen, Correction: Modelling and prediction of the thermophysical properties of aqueous mixtures of choline geranate and geranic acid (CAGE) using SAFT- γ Mie, *RSC Advances* 10 (2020) 19463–19465. doi:10.1039/d0ra90058c.
- [44] M. Wehbe, A. J. Haslam, G. Jackson, A. Galindo, Phase behaviour and pH-solubility profile prediction of aqueous buffered solutions of ibuprofen and ketoprofen, *Fluid Phase Equilibria* 560 (2022) 113504. doi:10.1016/J.FLUID.2022.113504.
- [45] M. S. Wertheim, Fluids with highly directional attractive forces. I. Statistical thermodynamics, *Journal of Statistical Physics* 35 (1) (1984) 19–34. doi:10.1007/BF01017362.
- [46] M. S. Wertheim, Fluids with highly directional attractive forces. II. Thermodynamic perturbation theory and integral equations, *Journal of Statistical Physics* 35 (1) (1984) 35–47. doi:10.1007/BF01017363.
- [47] M. S. Wertheim, Fluids with highly directional attractive forces. III. Multiple attraction sites, *Journal of Statistical Physics* 42 (3) (1986) 459–476. doi:10.1007/BF01127721.
- [48] M. S. Wertheim, Fluids with highly directional attractive forces. IV. Equilibrium polymerization, *Journal of Statistical Physics* 42 (3) (1986) 477–492. doi:10.1007/BF01127722.
- [49] I. G. Economou, Statistical associating fluid theory: A successful model for the calculation

- of thermodynamic and phase equilibrium properties of complex fluid mixtures, *Industrial & Engineering Chemistry Research* 41 (5) (2002) 953–962. doi:10.1021/ie0102201.
- [50] E. A. Müller, K. E. Gubbins, Associating fluids and fluid mixtures, in: J. V. Sengers, R. F. Kayser, C. J. Peters, H. J. White (Eds.), *Equations of State for Fluids and Fluid Mixtures*, Elsevier, Amsterdam, 2000, Ch. 12, pp. 435–478.
- [51] E. A. Müller, K. E. Gubbins, Molecular-based equations of state for associating fluids: A review of SAFT and related approaches, *Industrial & Engineering Chemistry Research* 40 (10) (2001) 2193–2211. doi:10.1021/ie000773w.
- [52] P. Paricaud, A. Galindo, G. Jackson, Recent advances in the use of the SAFT approach in describing electrolytes, interfaces, liquid crystals and polymers, *Fluid Phase Equilibria* 194–197 (2002) 87–96. doi:10.1016/S0378-3812(01)00659-8.
- [53] S. P. Tan, H. Adidharma, M. Radosz, Recent advances and applications of statistical associating fluid theory, *Industrial & Engineering Chemistry Research* 47 (21) (2008) 8063–8082. doi:10.1021/ie8008764.
- [54] G. M. Kontogeorgis, G. K. Folas, *Thermodynamic Models for Industrial Applications: From Classical and Advanced Mixing Rules to Association Theories*, John Wiley & Sons, Chichester, 2009. doi:10.1002/9780470747537.
- [55] C. McCabe, A. Galindo, SAFT Associating Fluids and Fluid Mixtures, in: A. R. H. Goodwin, J. V. Sengers, C. J. Peters (Eds.), *Applied Thermodynamics of Fluids*, Royal Society of Chemistry, Cambridge, 2010, Ch. 8, pp. 215–279. doi:10.1039/9781849730983-00215.
- [56] M. Lora, F. Rindfleisch, M. A. McHugh, Influence of the alkyl tail on the solubility of poly (alkyl acrylates) in ethylene and CO₂ at high pressures: Experiments and modeling, *Journal of Applied Polymer Science* 73 (10) (1999) 1979–1991.
- [57] S. Benzaghrou, J. P. Passarello, P. Tobaly, Predictive use of a SAFT EOS for phase equilibria of some hydrocarbons and their binary mixtures, *Fluid Phase Equilibria* 180 (1–2) (2001) 1–26. doi:10.1016/S0378-3812(00)00481-7.

- [58] J. P. Passarello, P. Tobaly, Modeling some alcohol/alkane binary systems using the SAFT equation of state with a semipredictive approach, *Industrial and Engineering Chemistry Research* 42 (21) (2003) 5383–5391. doi:10.1021/ie030306p.
- [59] T. X. N. Thi, S. Tamouza, P. Tobaly, J.-P. Passarello, J.-C. de Hemptinne, Application of group contribution SAFT equation of state (GC-SAFT) to model phase behaviour of light and heavy esters, *Fluid Phase Equilibria* 238 (2) (2005) 254–261.
- [60] D. N. Huynh, M. Benamira, J.-P. Passarello, P. Tobaly, J.-C. de Hemptinne, Application of GC-SAFT EOS to polycyclic aromatic hydrocarbons, *Fluid Phase Equilibria* 254 (1-2) (2007) 60–66.
- [61] D. Nguyen-Huynh, J.-C. de Hemptinne, R. Lugo, J.-P. Passarello, P. Tobaly, Modeling liquid–liquid and liquid–vapor equilibria of binary systems containing water with an alkane, an aromatic hydrocarbon, an alcohol or a gas (methane, ethane, CO₂ or H₂S), using group contribution polar perturbed-chain statistical associating fluid, *Industrial & engineering chemistry research* 50 (12) (2011) 7467–7483.
- [62] J. Rozmus, J.-C. de Hemptinne, P. Mougin, Application of GC-PPC-SAFT EoS to amine mixtures with a predictive approach, *Fluid Phase Equilibria* 303 (1) (2011) 15–30.
- [63] J. Rozmus, J.-C. de Hemptinne, N. Ferrando, P. Mougin, Long chain multifunctional molecules with GC-PPC-SAFT: Limits of data and model, *Fluid Phase Equilibria* 329 (2012) 78–85.
- [64] S. Tamouza, J.-P. Passarello, P. Tobaly, J.-C. de Hemptinne, Group contribution method with SAFT EOS applied to vapor liquid equilibria of various hydrocarbon series, *Fluid Phase Equilibria* 222-223 (2004) 67–76.
- [65] A. Gil-Villegas, A. Galindo, P. J. Whitehead, S. J. Mills, G. Jackson, A. N. Burgess, Statistical associating fluid theory for chain molecules with attractive potentials of variable range, *Journal of Chemical Physics* 106 (10) (1997) 4168–4186. doi:10.1063/1.473101.
- [66] A. Galindo, L. A. Davies, A. Gil-Villegas, G. Jackson, The thermodynamics of mixtures

- and the corresponding mixing rules in the SAFT-VR approach for potentials of variable range, *Molecular Physics* 93 (2) (1998) 241–252. doi:10.1080/002689798169249.
- [67] S. Tamouza, J.-P. Passarello, P. Tobaly, J.-C. de Hemptinne, Application to binary mixtures of a group contribution SAFT EOS (GC-SAFT), *Fluid Phase Equilibria* 228 (2005) 409–419.
- [68] D. Nguyen-Huynh, A. Falaix, J.-P. Passarello, P. Tobaly, J.-C. de Hemptinne, Predicting VLE of heavy esters and their mixtures using GC-SAFT, *Fluid Phase Equilibria* 264 (1-2) (2008) 184–200.
- [69] D. Nguyen-Huynh, J.-P. Passarello, P. Tobaly, J.-C. de Hemptinne, Application of GC-SAFT EOS to polar systems using a segment approach, *Fluid Phase Equilibria* 264 (1-2) (2008) 62–75.
- [70] J. Gross, G. Sadowski, Perturbed-Chain SAFT: An equation of state based on a perturbation theory for chain molecules, *Industrial and Engineering Chemistry Research* 40 (4) (2001) 1244–1260. doi:10.1021/ie0003887.
- [71] J. Vijande, M. M. Pineiro, D. Bessieres, H. Saint-Guirons, J. L. Legido, Description of PVT behaviour of hydrofluoroethers using the PC-SAFT EOS, *Physical Chemistry Chemical Physics* 6 (4) (2004) 766–770.
- [72] K. E. Gubbins, C. H. Twu, Thermodynamics of polyatomic fluid mixtures—I theory, *Chemical Engineering Science* 33 (7) (1978) 863–878.
- [73] C. H. Twu, K. E. Gubbins, Thermodynamics of polyatomic fluid mixtures—II: Polar, quadrupolar and octopolar molecules, *Chemical Engineering Science* 33 (7) (1978) 879–887.
- [74] F. S. Emami, A. Vahid, J. R. Elliott Jr, F. Feyzi, Group Contribution Prediction of Vapor Pressure with Statistical Associating Fluid Theory, Perturbed-Chain Statistical Associating Fluid Theory, and Elliott-Suresh-Donohue Equations of State, *Industrial & Engineering Chemistry Research* 47 (21) (2008) 8401–8411.

- [75] A. Tihic, G. M. Kontogeorgis, N. Von Solms, M. L. Michelsen, L. Constantinou, A predictive group-contribution simplified PC-SAFT equation of state: application to polymer systems, *Industrial & Engineering Chemistry Research* 47 (15) (2008) 5092–5101.
- [76] L. Constantinou, R. Gani, New group contribution method for estimating properties of pure compounds, *AIChE Journal* 40 (10) (1994) 1697–1710.
- [77] L. F. Cameretti, G. Sadowski, J. M. Mollerup, Modeling of aqueous electrolyte solutions with perturbed-chain statistical associated fluid theory, *Industrial & Engineering Chemistry Research* 44 (9) (2005) 3355–3362. doi:10.1021/ie0488142.
- [78] R. Paus, Y. Ji, F. Braak, G. Sadowski, Dissolution of crystalline pharmaceuticals: Experimental investigation and thermodynamic modeling, *Industrial & Engineering Chemistry Research* 54 (2) (2015) 731–742.
- [79] Y. Ji, R. Paus, A. Prudic, C. Lübbert, G. Sadowski, A novel approach for analyzing the dissolution mechanism of solid dispersions, *Pharmaceutical Research* 32 (8) (2015) 2559–2578.
- [80] F. Laube, T. Klein, G. Sadowski, Partition coefficients of pharmaceuticals as functions of temperature and pH, *Industrial & Engineering Chemistry Research* 54 (15) (2015) 3968–3975.
- [81] F. S. Laube, G. Sadowski, Predicting the Extraction Behavior of Pharmaceuticals, *Industrial & Engineering Chemistry Research* 53 (2) (2014) 865–870. doi:10.1021/IE403284Y.
- [82] A. Prudic, T. Kleetz, M. Korf, Y. Ji, G. Sadowski, Influence of Copolymer Composition on the Phase Behavior of Solid Dispersions, *Molecular Pharmaceutics* 11 (11) (2014) 4189–4198. doi:10.1021/MP500412D.
- [83] A. Prudic, A. K. Lesniak, Y. Ji, G. Sadowski, Thermodynamic phase behaviour of indomethacin/PLGA formulations, *European Journal of Pharmaceutics and Biopharmaceutics* 93 (2015) 88–94. doi:10.1016/J.EJPB.2015.01.029.

- [84] R. Paus, Y. Ji, L. Vahle, G. Sadowski, Predicting the Solubility Advantage of Amorphous Pharmaceuticals: A Novel Thermodynamic Approach, *Molecular Pharmaceutics* 12 (8) (2015) 2823–2833.
- [85] K. Lehmkemper, S. O. Kyeremateng, M. Degenhardt, G. Sadowski, Influence of Low-Molecular-Weight Excipients on the Phase Behavior of PVPVA64 Amorphous Solid Dispersions, *Pharmaceutical Research* 35 (25) (2018) 1–14.
- [86] S. Dohrn, P. Rawal, C. Luebbert, K. Lehmkemper, S. O. Kyeremateng, M. Degenhardt, G. Sadowski, Predicting process design spaces for spray drying amorphous solid dispersions, *International Journal of Pharmaceutics: X* 3 (2021) 100072. doi:10.1016/J.IJPX.2021.100072.
- [87] F. Ruether, G. Sadowski, Modeling the Solubility of Pharmaceuticals in Pure Solvents and Solvent Mixtures for Drug Process Design, *Journal of Pharmaceutical Sciences* 98 (11) (2009) 4205–4215. doi:10.1002/JPS.21725.
- [88] L. Lange, S. Heisel, G. Sadowski, Predicting the Solubility of Pharmaceutical Cocrystals in Solvent/Anti-Solvent Mixtures, *Molecules* 21 (5) (2016) 593. doi:10.3390/MOLECULES21050593.
- [89] J. Brinkmann, F. Huxoll, C. Luebbert, G. Sadowski, Solubility of pharmaceutical ingredients in triglycerides, *European Journal of Pharmaceutics and Biopharmaceutics* 145 (2019) 113–120. doi:10.1016/J.EJPB.2019.10.012.
- [90] J. Brinkmann, F. Rest, C. Luebbert, G. Sadowski, Solubility of Pharmaceutical Ingredients in Natural Edible Oils, *Molecular Pharmaceutics* 17 (7) (2020) 2499–2507.
- [91] A. Mathers, F. Hassouna, M. Klajmon, M. Fulem, Comparative Study of DSC-Based Protocols for API-Polymer Solubility Determination, *Molecular Pharmaceutics* 18 (4) (2021) 1742–1757. doi:10.1021/acs.molpharmaceut.0c01232.
- [92] R. Paus, Y. Ji, Modeling and predicting the influence of variable factors on dissolution of crystalline pharmaceuticals, *Chemical Engineering Science* 145 (2016) 10–20. doi:10.1016/J.CES.2016.01.040.

- [93] R. Paus, A. Prudic, Y. Ji, Influence of excipients on solubility and dissolution of pharmaceuticals, *International Journal of Pharmaceutics* 485 (1-2) (2015) 277–287. doi:10.1016/j.ijpharm.2015.03.004.
- [94] K. Ge, R. Paus, V. Penner, G. Sadowski, Y. Ji, Theoretical modeling and prediction of biorelevant solubility of poorly soluble pharmaceuticals, *Chemical Engineering Journal* 444 (2022) 136678. doi:10.1016/j.cej.2022.136678.
- [95] A. Prudic, Y. Ji, G. Sadowski, Thermodynamic phase behavior of API/polymer solid dispersions, *Molecular Pharmaceutics* 11 (7) (2014) 2294–2304.
- [96] H. Veith, E. Turan, C. Luebbert, G. Sadowski, Hydrate formation in polymer-based pharmaceutical formulations, *Fluid Phase Equilibria* 521 (2020) 112677. doi:10.1016/J.FLUID.2020.112677.
- [97] H. Veith, M. Zaeh, C. Luebbert, N. Rodríguez-Hornedo, G. Sadowski, Stability of Pharmaceutical Co-Crystals at Humid Conditions Can Be Predicted, *Pharmaceutics* 13 (3) (2021) 433. doi:10.3390/PHARMACEUTICS13030433.
- [98] D. Borrmann, A. Danzer, G. Sadowski, Measuring and Modeling Water Sorption in Amorphous Indomethacin and Ritonavir, *Molecular Pharmaceutics* 19 (3) (2022) 998–1007. doi:10.1021/acs.molpharmaceut.1c00984.
- [99] D. Borrmann, A. Danzer, G. Sadowski, Predicting the Water Sorption in ASDs, *Pharmaceutics* 14 (6) (2022) 1181. doi:10.3390/pharmaceutics14061181.
- [100] A. Prudic, Y. Ji, C. Luebbert, G. Sadowski, Influence of humidity on the phase behavior of API/polymer formulations, *European Journal of Pharmaceutics and Biopharmaceutics* 94 (2015) 352–362. doi:10.1016/J.EJPB.2015.06.009.
- [101] C. Luebbert, G. Sadowski, Moisture-induced phase separation and recrystallization in amorphous solid dispersions, *International Journal of Pharmaceutics* 532 (1) (2017) 635–646. doi:10.1016/J.IJPHARM.2017.08.121.

- [102] C. Luebbert, F. Huxoll, G. Sadowski, G. Van Den Mooter, H. Grohgan, Amorphous-Amorphous Phase Separation in API/Polymer Formulations, *Molecules* 22 (2) (2017) 296. doi:10.3390/MOLECULES22020296.
- [103] C. Luebbert, M. Wessner, G. Sadowski, Mutual Impact of Phase Separation/Crystallization and Water Sorption in Amorphous Solid Dispersions, *Molecular Pharmaceutics* 15 (2) (2018) 669–678. doi:10.1021/ACS.MOLPHARMACEUT.7B01076/SUPPL{_}FILE/MP7B01076{_}SI{_}001.PDF.
- [104] C. Luebbert, D. Real, G. Sadowski, Choosing Appropriate Solvents for ASD Preparation, *Molecular Pharmaceutics* 15 (11) (2018) 5397–5409.
- [105] C. Luebbert, E. Stoyanov, G. Sadowski, Phase behavior of ASDs based on hydroxypropyl cellulose, *International Journal of Pharmaceutics: X* 3 (2021) 100070. doi:10.1016/J.IJPX.2020.100070.
- [106] S. Dohrn, C. Luebbert, K. Lehmkemper, S. O. Kyeremateng, M. Degenhardt, G. Sadowski, Solvent influence on the phase behavior and glass transition of Amorphous Solid Dispersions, *European Journal of Pharmaceutics and Biopharmaceutics* 158 (2021) 132–142. doi:10.1016/J.EJPB.2020.11.002.
- [107] S. Dohrn, C. Luebbert, K. Lehmkemper, S. O. Kyeremateng, M. Degenhardt, G. Sadowski, Phase behavior of pharmaceutically relevant polymer/solvent mixtures, *International Journal of Pharmaceutics* 577 (2020) 119065. doi:10.1016/J.IJPHARM.2020.119065.
- [108] S. Dohrn, P. Reimer, C. Luebbert, K. Lehmkemper, S. O. Kyeremateng, M. Degenhardt, G. Sadowski, Thermodynamic Modeling of Solvent-Impact on Phase Separation in Amorphous Solid Dispersions during Drying, *Molecular Pharmaceutics* 17 (7) (2020) 2721–2733.
- [109] F. Wolbert, I.-K. Fahrig, T. Gottschalk, C. Luebbert, M. Thommes, G. Sadowski, Factors Influencing the Crystallization-Onset Time of Metastable ASDs, *Pharmaceutics* 14 (2) (2022) 269. doi:10.3390/PHARMACEUTICS14020269.

- [110] S. Dohrn, C. Luebbert, K. Lehmkemper, S. O. Kyeremateng, M. Degenhardt, G. Sadowski, Solvent mixtures in pharmaceutical development: Maximizing the API solubility and avoiding phase separation, *Fluid Phase Equilibria* 548 (2021) 113200. doi:10.1016/J.FLUID.2021.113200.
- [111] A. Iemtsev, F. Hassouna, A. Mathers, M. Klajmon, M. Dendisová, L. Malinová, T. Školáková, M. Fulem, Physical stability of hydroxypropyl methylcellulose-based amorphous solid dispersions: Experimental and computational study, *International Journal of Pharmaceutics* 589 (2020) 119845. doi:10.1016/j.ijpharm.2020.119845.
- [112] Q. Chen, Y. Ji, K. Ge, Influence of excipients on thermodynamic phase behavior of pharmaceutical/solvent systems: Molecular thermodynamic model prediction, *Chemical Engineering Science* 244 (2021) 116798. doi:10.1016/j.ces.2021.116798.
- [113] J. Cassens, A. Prudic, F. Ruether, G. Sadowski, Solubility of pharmaceuticals and their salts as a function of pH, *Industrial & Engineering Chemistry Research* 52 (7) (2013) 2721–2731. doi:10.1021/ie302064h.
- [114] R. Paus, E. Hart, Y. Ji, A novel approach for predicting the dissolution profiles of pharmaceutical tablets, *European Journal of Pharmaceutics and Biopharmaceutics* 96 (2015) 53–64. doi:10.1016/j.ejpb.2015.06.029.
- [115] S. H. Huang, M. Radosz, Equation of state for small, large, polydisperse, and associating molecules, *Industrial & Engineering Chemistry Research* 29 (11) (1990) 2284–2294. doi:10.1021/ie00107a014.
- [116] H. Nazem, C. Ghotbi, M. H. Zare, S. Shirazian, Experimental investigation and thermodynamic modeling of amino acids partitioning in a water/ionic liquid system, *Journal of Molecular Liquids* 260 (2018) 386–390. doi:10.1016/J.MOLLIQ.2018.03.108.
- [117] K. Wysoczanska, B. Nierhauve, G. Sadowski, E. A. Macedo, C. Held, Solubility of DNP-amino acids and their partitioning in biodegradable ATPS: Experimental and ePC-SAFT modeling, *Fluid Phase Equilibria* 527 (2021) 112830. doi:10.1016/j.fluid.2020.112830.

- [118] M. H. Anvari, G. Pazuki, S. S. Kakhki, B. Bonakdarpour, Application of the SAFT-VR equation of state in estimation of physiochemical properties of amino acid solutions, *Journal of Molecular Liquids* 184 (2013) 24–32. doi:10.1016/j.molliq.2013.04.024.
- [119] S. Seyfi, G. Pazuki, S. F. Aghamiri, M. Beheshti, On the prediction of equilibrium phase behavior of amino acids in aqueous and aqueous-electrolyte solutions using SAFT equation of state, *Fluid Phase Equilibria* 287 (1) (2009) 15–22. doi:10.1016/J.FLUID.2009.08.020.
- [120] J.-B. G. Daldrup, C. Held, G. Sadowski, G. Schembecker, Modeling pH and Solubilities in Aqueous Multisolute Amino Acid Solutions, *Industrial & Engineering Chemistry Research* 50 (6) (2011) 3503–3509. doi:10.1021/IE1010367.
- [121] D. Fuchs, J. Fischer, F. Tumakaka, G. Sadowski, Solubility of Amino Acids: Influence of the pH value and the Addition of Alcoholic Cosolvents on Aqueous Solubility, *Industrial & Engineering Chemistry Research* 45 (19) (2006) 6578–6584. doi:10.1021/IE0602097.
- [122] B. S. Lee, K. C. Kim, Study on the activity coefficients and solubilities of amino acids in aqueous solutions with perturbed-chain statistical associating fluid theory, *Korean Journal of Chemical Engineering* 2009 27:1 27 (1) (2010) 267–277. doi:10.1007/S11814-009-0351-Z.
- [123] H. T. Do, Y. Z. Chua, A. Kumar, D. Pabsch, M. Hallermann, D. Zaitsau, C. Schick, C. Held, Melting properties of amino acids and their solubility in water, *RSC Advances* 10 (72) (2020) 44205–44215. doi:10.1039/D0RA08947H.
- [124] H. T. Do, P. Franke, S. Volpert, M. Klinksiek, M. Thome, C. Held, Measurement and modelling solubility of amino acids and peptides in aqueous 2-propanol solutions, *Physical Chemistry Chemical Physics* 23 (18) (2021) 10852–10863. doi:10.1039/d1cp00005e.
- [125] L. F. Cameretti, G. Sadowski, Modeling of aqueous amino acid and polypeptide solutions with PC-SAFT, *Chemical Engineering and Processing: Process Intensification* 47 (6) (2008) 1018–1025. doi:10.1016/J.CEP.2007.02.034.

- [126] J.-B. G. Daldrup, C. Held, F. Ruether, G. Schembecker, G. Sadowski, Measurement and modeling solubility of aqueous multisolute amino-acid solutions, in: *Industrial and Engineering Chemistry Research*, Vol. 49, American Chemical Society, 2010, pp. 1395–1401. doi:10.1021/ie900913c.
- [127] H. T. Do, Y. Z. Chua, J. Habicht, M. Klinksiek, M. Hallermann, D. Zaitsau, C. Schick, C. Held, Melting properties of peptides and their solubility in water. Part 1: Dipeptides based on glycine or alanine, *RSC Advances* 9 (56) (2019) 32722–32734. doi:10.1039/c9ra05730g.
- [128] H. T. Do, Y. Z. Chua, J. Habicht, M. Klinksiek, S. Volpert, M. Hallermann, M. Thome, D. Pabsch, D. Zaitsau, C. Schick, C. Held, Melting Properties of Peptides and Their Solubility in Water. Part 2: Di- And Tripeptides Based on Glycine, Alanine, Leucine, Proline, and Serine, *Industrial & Engineering Chemistry Research* 60 (12) (2021) 4693–4704. doi:10.1021/acs.iecr.0c05652.
- [129] C. Held, L. F. Cameretti, G. Sadowski, Measuring and Modeling Activity Coefficients in Aqueous Amino-Acid Solutions, *Industrial & Engineering Chemistry Research* 50 (1) (2011) 131–141. doi:10.1021/IE100088C.
- [130] C. H. Bang, H. K. Choi, B. S. Lee, Modeling of activity coefficients of amino acid and electrolyte in aqueous solutions, *Journal of Molecular Liquids* 223 (2016) 1–9. doi:10.1016/J.MOLLIQ.2016.08.016.
- [131] C. Held, T. Reschke, S. Mohammad, A. Luza, G. Sadowski, ePC-SAFT revised, *Chemical Engineering Research and Design* 92 (12) (2014) 2884–2897. doi:10.1016/j.cherd.2014.05.017.
- [132] C. Held, T. Reschke, R. Müller, W. Kunz, G. Sadowski, Measuring and modeling aqueous electrolyte/amino-acid solutions with ePC-SAFT, *Journal of Chemical Thermodynamics* 68 (2014) 1–12. doi:10.1016/j.jct.2013.08.018.
- [133] P. Ji, J. Zou, W. Feng, Effect of alcohol on the solubility of amino acid in water, *Journal*

- of Molecular Catalysis B: Enzymatic 56 (2-3) (2009) 185–188. doi:10.1016/J.MOLCATB.2008.06.008.
- [134] H. Adidharma, M. Radosz, Prototype of an engineering equation of state for heterosegmented polymers, *Industrial & Engineering Chemistry Research* 37 (11) (1998) 4453–4462.
- [135] H. Adidharma, M. Radosz, Square-well SAFT equation of state for homopolymeric and heteropolymeric fluids, *Fluid Phase Equilibria* 158 (1999) 165–174.
- [136] C. McCabe, A. Gil-Villegas, G. Jackson, F. Del Rio, The thermodynamics of heteronuclear molecules formed from bonded square-well (BSW) segments using the SAFT-VR approach, *Molecular Physics* 97 (4) (1999) 551–558.
- [137] Y. Peng, H. Zhao, C. McCabe, On the thermodynamics of diblock chain fluids from simulation and heteronuclear statistical associating fluid theory for potentials of variable range, *Molecular Physics* 104 (4) (2006) 571–586.
- [138] M. Banaszak, C. K. Chen, M. Radosz, Copolymer SAFT equation of state. Thermodynamic perturbation theory extended to heterobonded chains, *Macromolecules* 29 (20) (1996) 6481–6486.
- [139] M. Banaszak, M. Radosz, Molecular dynamics study on homonuclear and heteronuclear chains of Lennard–Jones segments, *Fluid Phase Equilibria* 193 (1-2) (2002) 179–189.
- [140] F. J. Blas, L. F. Vega, Thermodynamic behaviour of homonuclear and heteronuclear Lennard-Jones chains with association sites from simulation and theory, *Molecular Physics* 92 (1) (1997) 135–150. doi:10.1080/002689797170707.
- [141] A. Lymperiadis, C. S. Adjiman, A. Galindo, G. Jackson, A group contribution method for associating chain molecules based on the statistical associating fluid theory (SAFT- γ), *Journal of Chemical Physics* 127 (23) (2007) 4168. doi:10.1063/1.2813894.

- [142] A. Lympieriadis, C. S. Adjiman, G. Jackson, A. Galindo, A generalisation of the SAFT- γ group contribution method for groups comprising multiple spherical segments, *Fluid Phase Equilibria* 274 (1-2) (2008) 85–104. doi:10.1016/J.FLUID.2008.08.005.
- [143] V. Papaioannou, C. S. Adjiman, G. Jackson, A. Galindo, Simultaneous prediction of vapour–liquid and liquid–liquid equilibria (VLE and LLE) of aqueous mixtures with the SAFT- γ group contribution approach, *Fluid phase equilibria* 306 (1) (2011) 82–96.
- [144] S.-S. Ashrafmansouri, S. Raeissi, Modeling gas solubility in ionic liquids with the SAFT- γ group contribution method, *The Journal of Supercritical Fluids* 63 (2012) 81–91.
- [145] Y. Peng, K. D. Goff, M. C. dos Ramos, C. McCabe, Developing a predictive group-contribution-based SAFT-VR equation of state, *Fluid Phase Equilibria* 277 (2) (2009) 131–144.
- [146] T. Lafitte, A. Apostolakou, C. Avendaño, A. Galindo, C. S. Adjiman, E. A. Müller, G. Jackson, Accurate statistical associating fluid theory for chain molecules formed from Mie segments, *Journal of Chemical Physics* 139 (15) (2013) 1–37. doi:10.1063/1.4819786.
- [147] G. Mie, Zur kinetischen Theorie der einatomigen Körper, *Annalen der Physik* 316 (8) (1903) 657–697. doi:10.1002/andp.19033160802.
- [148] A. J. Haslam, A. González-Pérez, S. Di Lecce, S. H. Khalit, F. A. Perdomo, S. Kournopoulos, M. Kohns, T. Lindeboom, M. Wehbe, S. Febra, G. Jackson, C. S. Adjiman, A. Galindo, Expanding the Applications of the SAFT- γ Mie Group-Contribution Equation of State: Prediction of Thermodynamic Properties and Phase Behavior of Mixtures, *Journal of Chemical and Engineering Data* 65 (12) (2020) 5862–90. doi:10.1021/acs.jced.0c00746.
- [149] G. Lazarou, Development of the SAFT- γ Mie equation of state for predicting the thermodynamic behaviour of strong and weak electrolyte solutions, Ph.D. thesis, Imperial College London (2017).

- [150] D. K. Eriksen, G. Lazarou, A. Galindo, G. Jackson, C. S. Adjiman, A. J. Haslam, Development of intermolecular potential models for electrolyte solutions using an electrolyte SAFT-VR Mie equation of state, *Molecular Physics* 114 (18) (2016) 2724–2749. doi:10.1080/00268976.2016.1236221.
- [151] J. M. Schreckenber, S. Dufal, A. J. Haslam, C. S. Adjiman, G. Jackson, A. Galindo, Modelling of the thermodynamic and solvation properties of electrolyte solutions with the statistical associating fluid theory for potentials of variable range, *Molecular Physics* 112 (17) (2014) 2339–2364. doi:10.1080/00268976.2014.910316.
- [152] L. Blum, Solution of a model for the solvent-electrolyte interactions in the mean spherical approximation, *The Journal of Chemical Physics* 61 (5) (1974) 2129–2133. doi:10.1063/1.1682224.
- [153] L. Blum, Mean spherical model for asymmetric electrolytes I. Method of solution, *Molecular Physics* 30 (5) (1975) 1529–1535. doi:10.1080/00268977500103051.
- [154] M. Born, Volumen und Hydratationswärme der Ionen, *Zeitschrift für Physik* 1 (1) (1920) 45–48. doi:10.1007/BF01881023.
- [155] L. L. Lee, *Molecular Thermodynamics of Nonideal Fluids*, Butterworths, Boston, 1988.
- [156] J. A. Barker, D. Henderson, Perturbation Theory and Equation of State for Fluids. II. A Successful Theory of Liquids, *The Journal of Chemical Physics* 47 (11) (1967) 4714. doi:10.1063/1.1701689.
- [157] J. A. Barker, D. Henderson, What is "liquid"? Understanding the states of matter, *Reviews of Modern Physics* 48 (4) (1976) 587. doi:10.1103/RevModPhys.48.587.
- [158] T. Boublik, Hard-Sphere Equation of State, *The Journal of Chemical Physics* 53 (1) (1970) 471. doi:10.1063/1.1673824.
- [159] G. A. Mansoori, N. F. Carnahan, K. E. Starling, T. W. Leland, Equilibrium Thermodynamic Properties of the Mixture of Hard Spheres, *The Journal of Chemical Physics* 54 (4) (1971) 1523. doi:10.1063/1.1675048.

- [160] B. J. Zhang, Calculating thermodynamic properties from perturbation theory: I. An analytic representation of square-well potential hard-sphere perturbation theory, *Fluid Phase Equilibria* 154 (1) (1999) 1–10. doi:10.1016/S0378-3812(98)00431-2.
- [161] P. Paricaud, A general perturbation approach for equation of state development: Applications to simple fluids, ab initio potentials, and fullerenes, *The Journal of Chemical Physics* 124 (15) (2006) 154505. doi:10.1063/1.2181979.
- [162] M. S. Wertheim, Thermodynamic perturbation theory of polymerization, *The Journal of Chemical Physics* 87 (12) (1987) 7323–7331. doi:10.1063/1.453326.
- [163] T. Boublík, Background correlation functions in the hard sphere systems, *Molecular Physics* 59 (4) (1986) 775–793. doi:10.1080/00268978600102391.
- [164] J. S. Rowlinson, F. L. Swinton, *Liquids and liquid mixtures*, 3rd Edition, Butterworths, Norwich, 1982.
- [165] G. H. Hudson, J. C. McCoubrey, Intermolecular forces between unlike molecules. A more complete form of the combining rules., *Transactions of the Faraday Society* 56 (1960) 761–766.
- [166] M. Sadeqzadeh, V. Papaioannou, S. Dufal, C. S. Adjiman, G. Jackson, A. Galindo, The development of unlike induced association-site models to study the phase behaviour of aqueous mixtures comprising acetone, alkanes and alkyl carboxylic acids with the SAFT- γ Mie group contribution methodology, *Fluid Phase Equilibria* 407 (2016) 39–57. doi:10.1016/j.fluid.2015.07.047.
- [167] F. A. Perdomo, S. H. Khalit, C. S. Adjiman, A. Galindo, G. Jackson, Description of the thermodynamic properties and fluid-phase behavior of aqueous solutions of linear, branched, and cyclic amines, *AIChE Journal* 67 (3) (2021) e17194. doi:10.1002/AIC.17194.
- [168] M. Michelsen, M. JM, J. Mollerup, *Fundamentals & computational aspects*, Tie-Line Publications, Holte, 2007.

- [169] A. Rogers, M. Ierapetritou, Challenges and opportunities in modeling pharmaceutical manufacturing processes, *Computers & Chemical Engineering* 81 (2015) 32–39.
- [170] R. Dohrn, O. Pfohl, Thermophysical properties - Industrial directions, in: *Fluid Phase Equilibria*, Vol. 194-197, Elsevier, 2002, pp. 15–29. doi:10.1016/S0378-3812(01)00791-9.
- [171] D.-B. Ren, Z.-H. Yang, Y.-Z. Liang, Q. Ding, C. Chen, M.-L. Ouyang, Correlation and prediction of partition coefficient using nonrandom two-liquid segment activity coefficient model for solvent system selection in counter-current chromatography separation, *Journal of Chromatography A* 1301 (2013) 10–18.
- [172] G. D. Pappa, E. C. Voutsas, K. Magoulas, D. P. Tassios, Estimation of the Differential Molar Heat Capacities of Organic Compounds at Their Melting Point, *Industrial & Engineering Chemistry Research* 44 (10) (2005) 3799–3806. doi:10.1021/IE048916S.
- [173] I. Tsivintzelis, I. G. Economou, G. M. Kontogeorgis, Modeling the solid-liquid equilibrium in pharmaceutical-solvent mixtures: Systems with complex hydrogen bonding behavior, *AIChE Journal* 55 (3) (2009) 756–770. doi:10.1002/aic.11716.
- [174] S. A. Febra, Ring Formation in a Statistical Associating Fluid Theory Framework, Ph.D. thesis, Imperial College London (2018).
- [175] T. Takagi, C. Ramachandran, M. Bermejo, S. Yamashita, L. X. Yu, G. L. Amidon, A provisional biopharmaceutical classification of the top 200 oral drug products in the United States, Great Britain, Spain, and Japan, *Molecular Pharmaceutics* 3 (6) (2006) 631–643. doi:10.1021/mp0600182.
- [176] A. T. Serajuddin, Salt formation to improve drug solubility, *Advanced Drug Delivery Reviews* 59 (7) (2007) 603–616. doi:10.1016/j.addr.2007.05.010.
- [177] K. T. Savjani, A. K. Gajjar, J. K. Savjani, Drug Solubility: Importance and Enhancement Techniques, *ISRN Pharmaceutics* 2012 (2012) 1–10. doi:10.5402/2012/195727.

- [178] H. D. Williams, N. L. Trevaskis, S. A. Charman, R. M. Shanker, W. N. Charman, C. W. Pouton, C. J. H. Porter, Strategies to Address Low Drug Solubility in Discovery and Development, *Pharmacological Reviews* 65 (1) (2013) 315–499. doi:10.1124/pr.112.005660.
- [179] S. Sweetana, M. J. Akers, Solubility principles and practices for parenteral drug dosage form development., *PDA journal of pharmaceutical science and technology* 50 (5) (1996) 330–342.
- [180] S. M. Berge, L. D. Bighley, D. C. Monkhouse, Pharmaceutical Salts, *Journal of Pharmaceutical Sciences* 66 (1) (1977) 1–19. doi:10.1002/JPS.2600660104.
- [181] P. H. Stahl, C. G. Wermuth, *Handbook of Pharmaceutical Salts: Properties, Selection, and Use*, John Wiley & Sons, Zurich, 2002. doi:10.1515/ci.2002.24.3.20a.
- [182] A. Avdeef, *Absorption and Drug Development: Solubility, Permeability, and Charge State*, 2nd Edition, Wiley-Interscience, Hoboken, 2012. doi:10.1002/9781118286067.
- [183] A. Avdeef, C. M. Berger, C. Brownell, pH-metric solubility 2, *Pharmaceutical research* 17 (1) (2000) 85–89.
- [184] A. Avdeef, Solubility of sparingly-soluble ionizable drugs, *Advanced Drug Delivery Reviews* 59 (7) (2007) 568–590. doi:10.1016/j.addr.2007.05.008.
- [185] A. Avdeef, S. Bendels, O. Tsinman, K. Tsinman, M. Kansy, Solubility-Excipient Classification Gradient Maps, *Pharmaceutical Research* 24 (3) (2007) 530–545. doi:10.1007/s11095-006-9169-0.
- [186] J. Alsenz, E. Meister, E. Haenel, Development of a Partially Automated Solubility Screening (PASS) assay for early drug development, *Journal of Pharmaceutical Sciences* 96 (7) (2007) 1748–1762. doi:10.1002/jps.20814.
- [187] H. Chen, Z. Zhang, C. McNulty, C. Olbert, H. J. Yoon, J. W. Lee, S. C. Kim, M. H. Seo, H. S. Oh, A. V. Lemmo, S. J. Ellis, K. Heimlich, A High-Throughput Combinatorial

- Approach for the Discovery of a Cremophor EL-Free Paclitaxel Formulation, *Pharmaceutical Research* 20 (8) (2003) 1302–1308.
- [188] T.-M. Chen, H. Shen, C. Zhu, Evaluation of a Method for High Throughput Solubility Determination using a Multi-wavelength UV Plate Reader, *Combinatorial Chemistry & High Throughput Screening* 5 (7) (2002) 575–581. doi:10.2174/1386207023330075.
- [189] L. Pan, Q. Ho, K. Tsutsui, L. Takahashi, Comparison of chromatographic and spectroscopic methods used to rank compounds for aqueous solubility, *Journal of Pharmaceutical Sciences* 90 (4) (2001) 521–529. doi:10.1002/1520-6017(200104)90:4<521::AID-JPS1009>3.0.CO;2-B.
- [190] A. Avdeef, C. M. Berger, pH-metric solubility.: 3. Dissolution titration template method for solubility determination, *European Journal of Pharmaceutical Sciences* 14 (4) (2001) 281–291. doi:10.1016/S0928-0987(01)00190-7.
- [191] T. Higuchi, F. L. Shih, T. Kimura, J. H. Rytting, Solubility determination of barely aqueous-soluble organic solids, *Journal of Pharmaceutical Sciences* 68 (10) (1979) 1267–1272. doi:10.1002/jps.2600681019.
- [192] S. Venkatesh, J. Li, Y. Xu, R. Vishnuvajjala, B. D. Anderson, Intrinsic solubility estimation and pH-solubility behavior of cosalane (NSC 658586), an extremely hydrophobic diprotic acid, *Pharmaceutical Research* 13 (10) (1996) 1453–1459. doi:10.1023/A:1016059008464.
- [193] K. A. Hasselbalch, Die Berechnung der Wasserstoffzahl des Blutes aus der freien und gebundenen Kohlensäure desselben, und die Sauerstoffbindung des Blutes als Funktion der Wasserstoffzahl., *Biochemische Zeitschrift* 78 (1916) 112–144.
- [194] C. A. Bergström, K. Luthman, P. Artursson, Accuracy of calculated pH-dependent aqueous drug solubility, *European Journal of Pharmaceutical Sciences* 22 (5) (2004) 387–398. doi:10.1016/j.ejps.2004.04.006.
- [195] C. C. Chen, Toward development of activity coefficient models for process and product

- design of complex chemical systems, *Fluid Phase Equilibria* 241 (1-2) (2006) 103–112. doi:10.1016/j.fluid.2006.01.006.
- [196] B. R. Figueiredo, F. A. Da Silva, C. M. Silva, Non-ideality and Solubility Modeling of Amino Acids and Peptides in Aqueous Solutions: New Physical and Chemical Approach, *Industrial & Engineering Chemistry Research* 52 (2013) 16044–16056. doi:10.1021/ie402575c.
- [197] P. Debye, E. Huckel, De la theorie des electrolytes. I. abaissement du point de congelation et phenomenes associes, *Physikalische Zeitschrift* 24 (9) (1923) 185–206.
- [198] H. Renon, J. M. Prausnitz, Local compositions in thermodynamic excess functions for liquid mixtures, *AIChE Journal* 14 (1) (1968) 135–144. doi:10.1002/aic.690140124.
- [199] H. C. Tseng, C. Y. Lee, W. L. Weng, I. M. Shiah, Solubilities of amino acids in water at various pH values under 298.15 K, *Fluid Phase Equilibria* 285 (1-2) (2009) 90–95. doi:10.1016/j.fluid.2009.07.017.
- [200] C.-Y. Lee, J.-T. Chen, W.-T. Chang, I.-M. Shiah, Effect of pH on the solubilities of divalent and trivalent amino acids in water at 298.15 K, *Fluid Phase Equilibria* 343 (2013) 30–35. doi:10.1016/j.fluid.2013.01.010.
- [201] M. Kohns, G. Lazarou, S. Kournopoulos, E. Forte, F. A. Perdomo, G. Jackson, C. S. Adjiman, A. Galindo, S. Kournopoulos, G. Jackson, C. S. Adjiman, A. Galindo, Predictive models for the phase behaviour and solution properties of weak electrolytes: nitric, sulphuric, and carbonic acids, *Physical Chemistry Chemical Physics* 22 (27) (2020) 15248–15269. doi:10.1039/c9cp06795g.
- [202] Y. L. Hsieh, J. M. Merritt, W. Yu, L. S. Taylor, Salt Stability - The Effect of pHmax on Salt to Free Base Conversion, *Pharmaceutical Research* 32 (9) (2015) 3110–3118. doi:10.1007/s11095-015-1691-5.
- [203] G. A. Stephenson, A. Aburub, T. A. Woods, Physical stability of salts of weak bases in the solid-state, *Journal of Pharmaceutical Sciences* 100 (5) (2011) 1607–1617. doi:10.1002/jps.22405.

- [204] J. M. Merritt, S. K. Viswanath, G. A. Stephenson, Implementing quality by design in pharmaceutical salt selection: a modeling approach to understanding disproportionation., *Pharmaceutical research* 30 (1) (2013) 203–217. doi:10.1007/s11095-012-0863-9.
- [205] P. W. Atkins, J. De Paula, *Elements of Physical Chemistry*, 5th Edition, no. 5, Oxford University, Oxford, 2009.
- [206] National Center for Biotechnology Information. PubChem Database, Acetic acid, CID=176.
- [207] D. Perrin, B. Dempsey, *Buffers for pH and Metal Ion Control*, Springer Science & Business Media, London, 1974. doi:10.1007/978-94-009-5874-6.
- [208] D. D. Perrin, *Dissociation constants of organic bases in aqueous solution*, Butterworths, London, 1965.
- [209] D. R. Lide, *CRC Handbook of Chemistry and Physics*, 86th Edition, Taylor and Francis, Boca Raton, 2005.
- [210] R. C. Weast (Ed.), *CRC Handbook of Chemistry and Physics*, 66th Edition, CRC press, Boca Raton, 1985.
- [211] S. Klein, E. Kochanski, A. Strich, Electric properties of the oxonium ion in its ground and two lowest excited states, *Chemical Physics Letters* 260 (1-2) (1996) 34–42. doi:10.1016/0009-2614(96)00818-4.
- [212] M. I. Al-Joboury, D. W. Turner, Molecular photoelectron spectroscopy. Part VI. Water, methanol, methane, and ethane, *Journal of the Chemical Society B: Physical Organic* (0) (1967) 373–376. doi:10.1039/J29670000373.
- [213] Z. Lu, R. E. Continetti, Dynamics of the acetyloxyl radical studied by dissociative photodetachment of the acetate anion, *Journal of Physical Chemistry A* 108 (45) (2004) 9962–9969. doi:10.1021/jp040355v.
- [214] N. C. Pyper, C. G. Pike, P. P. Edwards, The polarizabilities of species present in ionic solutions, *Molecular Physics* 76 (2) (1992) 353–372. doi:10.1080/00268979200101381.

- [215] J. L. Franklin, J. G. Dillard, H. M. Rosenstock, J. T. Herron, K. Draxl, F. H. Field, Ionization potentials, appearance potentials and heats of formation of gaseous positive ions, 26th Edition, National Bureau of Standards, Gaithersburg, 1969.
- [216] T. M. Reed, The ionization potential and the polarizability of molecules, *Journal of Physical Chemistry* 59 (5) (1955) 428–432. doi:10.1021/j150527a012.
- [217] R. Chang, J. W. J. Thoman, Ion-Induced Dipole and Dipole-Induced Dipole Interactions, in: *Physical Chemistry for the Chemical Sciences*, Royal Society of Chemistry, 2014, Ch. 17.3.3.
- [218] J. R. Tessman, A. H. Kahn, W. Shockley, Electronic Polarizabilities of Ions in Crystals, *Physical Review* 92 (4) (1953) 890–895.
- [219] A. A. Rashin, B. Honig, Reevaluation of the Born Model of Ion Hydration, *Journal of Physical Chemistry* 89 (26) (1985) 5588–5593.
- [220] D. H. Aue, H. M. Webb, M. T. Bowers, Quantitative Proton Affinities, Ionization Potentials, and Hydrogen Affinities of Alkylamines, *Journal of the American Chemical Society* 98 (2) (1976) 311–317. doi:10.1021/ja00418a001.
- [221] R. De Lisi, G. Perron, J. Paquette, J. E. Desnoyers, Thermodynamics of micellar systems: activity and entropy of sodium decanoate and n -alkylamine hydrobromides in water, *Canadian Journal of Chemistry* 59 (13) (1981) 1865–1871. doi:10.1139/v81-278.
- [222] P. A. Leduc, J. L. Fortier, J. E. Desnoyers, Apparent molal volumes, heat capacities, and excess enthalpies of n-alkylamine hydrobromides in water as a function of temperature, *Journal of Physical Chemistry* 78 (12) (1974) 1217–1225. doi:10.1021/j100605a017.
- [223] J. E. Desnoyers, M. Arel, Apparent molal volumes of n -alkylamine hydrobromides in water at 25°C: hydrophobic hydration and volume changes, *Canadian Journal of Chemistry* 45 (4) (1967) 359–366. doi:10.1139/v67-064.
- [224] S. Gracin, A. C. Rasmuson, Solubility of phenylacetic acid, p-hydroxyphenylacetic acid,

- p-aminophenylacetic acid, p-hydroxybenzoic acid, and ibuprofen in pure solvents, *Journal of Chemical and Engineering Data* 47 (6) (2002) 1379–1383. doi:10.1021/je0255170.
- [225] F. Espitalier, B. Biscans, C. Laguérie, Physicochemical Data on Ketoprofen in Solutions, *Journal of Chemical and Engineering Data* 40 (6) (1995) 1222–1224. doi:10.1021/je00022a016.
- [226] T. R. Kommuru, M. A. Khan, I. K. Reddy, Racemate and enantiomers of ketoprofen: Phase diagram, thermodynamic studies, skin permeability, and use of chiral permeation enhancers, *Journal of Pharmaceutical Sciences* 87 (7) (1998) 833–840. doi:10.1021/js9704644.
- [227] Y. Z. Chua, H. T. Do, C. Schick, D. Zaitsau, C. Held, New experimental melting properties as access for predicting amino-acid solubility, *RSC Advances* 8 (12) (2018) 6365–6372. doi:10.1039/C8RA00334C.
- [228] J. A. Dean, *The Landolt Börnstein Database* (2008).
- [229] J. Martínez, F. Manrique, Solubility of Ibuprofen in Some Ethanol + Water Cosolvent Mixtures at Several Temperatures, *Latin American Journal of Pharmacy* 26 (3) (2007) 344–354.
- [230] A. Fini, G. Fazio, G. Feroci, Solubility and solubilization properties of non-steroidal anti-inflammatory drugs, *International Journal of Pharmaceutics* 126 (1-2) (1995) 95–102. doi:10.1016/0378-5173(95)04102-8.
- [231] S. Soltanpour, A. Jouyban, Solubility of acetaminophen and ibuprofen in polyethylene glycol 600, propylene glycol and water mixtures at 25 °C, *Journal of Molecular Liquids* 155 (2-3) (2010) 80–84. doi:10.1016/J.MOLLIQ.2010.05.011.
- [232] L. C. Garzón, F. Martínez, Temperature Dependence of Solubility for Ibuprofen in Some Organic and Aqueous Solvents, *Journal of Solution Chemistry* 33 (11) (2004) 1379–1395. doi:10.1007/s10953-004-1051-2.

- [233] S. H. Yalkowsky, S. C. Valvani, T. J. Roseman, Solubility and partitioning VI: Octanol solubility and octanol–water partition coefficients, *Journal of Pharmaceutical Sciences* 72 (8) (1983) 866–870. doi:10.1002/jps.2600720808.
- [234] C. A. Bergström, U. Norinder, K. Luthman, P. Artursson, Experimental and computational screening models for prediction of aqueous drug solubility, *Pharmaceutical Research* 19 (2) (2002) 182–188. doi:10.1023/A:1014224900524.
- [235] V. Papaioannou, F. Calado, T. Lafitte, S. Dufal, M. Sadeqzadeh, G. Jackson, C. S. Adjiman, A. Galindo, Application of the SAFT- γ Mie group contribution equation of state to fluids of relevance to the oil and gas industry, *Fluid Phase Equilibria* 416 (2016) 104–119. doi:10.1016/j.fluid.2015.12.041.
- [236] L. Codan, S. Casillo, M. U. Bäbler, M. Mazzotti, Phase diagram of a chiral substance exhibiting oiling out. 2. Racemic compound forming ibuprofen in water, *Crystal Growth and Design* 12 (11) (2012) 5298–5310. doi:10.1021/cg300890u.
- [237] B. Tita, A. Fulas, M. Stefanescu, E. Marian, D. Tita, Kinetic study of sodium diclofenac under isothermal conditions, *Revista de Chimie* 62 (1) (2011) 31–36.
- [238] D. Tița, A. Fulaș, B. Tița, Thermal stability of ketoprofen: Part 2. Kinetic study of the active substance under isothermal conditions, *Journal of Thermal Analysis and Calorimetry* 111 (3) (2013) 1979–1985. doi:10.1007/s10973-011-2147-8.
- [239] M. W. Chase, NIST-JANAF Thermochemical Tables, Fourth Edition, *J. Phys. Chem. Ref. Data*, Monograph 9 (1998) 1–1951.
- [240] G. G. Zhang, S. Y. Paspal, R. Suryanarayanan, D. J. Grant, Racemic species of sodium ibuprofen: Characterization and polymorphic relationships, *Journal of Pharmaceutical Sciences* 92 (7) (2003) 1356–1366. doi:10.1002/jps.10393.
- [241] R. Censi, V. Martena, E. Hoti, L. Malaj, P. Di Martino, Sodium ibuprofen dihydrate and anhydrous: Study of the dehydration and hydration mechanisms, *Journal of Thermal Analysis and Calorimetry* 111 (3) (2013) 2009–2018. doi:10.1007/s10973-012-2194-9.

- [242] T. Lee, Y. Wen Wang, Initial salt screening procedures for manufacturing ibuprofen, *Drug Development and Industrial Pharmacy* 35 (5) (2009) 555–567. doi:10.1080/03639040802459452.
- [243] L. R. Shaw, W. J. Irwin, T. J. Grattan, B. R. Conway, The Effect of Selected Water-Soluble Excipients on the Dissolution of Paracetamol and Ibuprofen, *Drug Development and Industrial Pharmacy* 31 (6) (2005) 515–525. doi:10.1080/03639040500215784.
- [244] S. E. David, P. Timmins, B. R. Conway, Impact of the counterion on the solubility and physicochemical properties of salts of carboxylic acid drugs, *Drug Development and Industrial Pharmacy* 38 (1) (2012) 93–103. doi:10.3109/03639045.2011.592530.
- [245] M. Yazdanian, K. Briggs, C. Jankovsky, A. Hawi, The "High Solubility" Definition of the Current FDA Guidance on Biopharmaceutical Classification System May Be Too Strict for Acidic Drugs, *Pharmaceutical Research* 21 (2) (2004) 293–299. doi:10.1023/B:PHAM.0000016242.48642.71.
- [246] J. J. Sheng, N. A. Kasim, R. Chandrasekharan, G. L. Amidon, Solubilization and dissolution of insoluble weak acid, ketoprofen: Effects of pH combined with surfactant, *European Journal of Pharmaceutical Sciences* 29 (3-4 SPEC. ISS.) (2006) 306–314. doi:10.1016/j.ejps.2006.06.006.
- [247] G. E. Hildebrand, C. C. Müller-Goymann, Ketoprofen sodium: Preparation and its formation of mixed crystals with ketoprofen, *Journal of Pharmaceutical Sciences* 86 (7) (1997) 854–857. doi:10.1021/js960278i.
- [248] F. Xu, L.-X. Sun, Z.-C. Tan, J.-G. Liang, R.-L. Li, Thermodynamic study of ibuprofen by adiabatic calorimetry and thermal analysis, *Thermochimica Acta* 412 (1-2) (2004) 33–57. doi:10.1016/j.tca.2003.08.021.
- [249] A. Avdeef, O. Tsinman, Miniaturized Rotating Disk Intrinsic Dissolution Rate Measurement: Effects of Buffer Capacity in Comparisons to Traditional Wood's Apparatus, *Pharmaceutical Research* 25 (11) (2008) 2613–2627. doi:10.1007/s11095-008-9679-z.

- [250] N. G. Manin, G. L. Perlovich, Thermodynamic properties of aqueous solutions of sodium ibuprofen at 293.15–318.15 K, *Russian Journal of Physical Chemistry A* 89 (4) (2015) 644–653. doi:10.1134/S0036024415040160.
- [251] G. Steffen Paulekuhn, J. B. Dressman, C. Saal, Trends in active pharmaceutical ingredient salt selection based on analysis of the orange book database, *Journal of Medicinal Chemistry* 50 (26) (2007) 6665–6672. doi:10.1021/jm701032y.
- [252] R. J. Bastin, M. J. Bowker, B. J. Slater, Salt selection and optimisation procedures for pharmaceutical new chemical entities, *Organic Process Research and Development* 4 (5) (2000) 427–435. doi:10.1021/op000018u.
- [253] K. W. Garren, R. A. Pyter, Aqueous solubility properties of a dibasic peptide-like compound, *International Journal of Pharmaceutics* 63 (2) (1990) 167–172. doi:https://doi.org/10.1016/0378-5173(90)90167-3.
- [254] E. Shoghi, E. Fuguet, E. Bosch, C. Ràfols, Solubility–pH profiles of some acidic, basic and amphoteric drugs, *European Journal of Pharmaceutical Sciences* 48 (1-2) (2013) 291–300. doi:10.1016/J.EJPS.2012.10.028.
- [255] Z. Wang, L. S. Burrell, W. J. Lambert, Solubility of E2050 at various pH: A case in which apparent solubility is affected by the amount of excess solid, *Journal of Pharmaceutical Sciences* 91 (6) (2002) 1445–1455. doi:10.1002/jps.10107.
- [256] A. T. M. Serajuddin, M. Pudipeddi, Salt-Selection Strategies, in: P. H. Stahl, C. G. Wermuth (Eds.), *Handbook of Pharmaceutical Salts: Properties, Selection, and Use*, John Wiley & Sons, Zurich, 2002, Ch. 6, pp. 135–167.
- [257] A. Lakkaraju, H. Joshi, S. Varia, A. T. M. Serajuddin, pH-solubility relationship of avitriptan, a dibasic compound, as a function of counterion, *Pharm. Res* 14 (1997) S–228.
- [258] D. Cairns, Physicochemical properties of drugs, in: *Essentials of Pharmaceutical Chemistry*, 4th Edition, Pharmaceutical Press, Aberdeen, 2012, Ch. 3, p. 70.

- [259] N. Wyttenbach, W. Kirchmeyer, J. Alsenz, M. Kuentz, Theoretical Considerations of the Prigogine-Defay Ratio with Regard to the Glass-Forming Ability of Drugs from Undercooled Melts, *Molecular Pharmaceutics* 13 (2016) 241–250. doi:10.1021/acs.molpharmaceut.5b00688.
- [260] P. Guerrieri, K. Jarring, L. S. Taylor, Impact of Counterion on the Chemical Stability of Crystalline Salts of Procaine, *Journal of pharmaceutical sciences* 99 (2010) 3719–3720. doi:10.1002/jps.22009.
- [261] K. B. Whetsel, J. H. Lady, Infrared Studies of Amine Complexes. III. Association of Aniline and N-Methylaniline with Benzene,-Dimethylaniline, Pyridine, and ,-Dimethylcyclohexylamine1, *The Journal of Physical Chemistry* 69 (5) (1965) 1596–1602.
- [262] H. Wolff, D. Mathias, Hydrogen bonding and Fermi resonance of aniline, *Journal of Physical Chemistry* 77 (17) (1973) 2081–2084. doi:10.1021/j100636a010.
- [263] J. Timmermans, Hennaut-Roland, Work of the International Bureau of Physicochemical Standards. VII. Physical Constants of 20 Organic Compounds., *J. Chim. Phys. Rev. Gen. Coll.* 32 (1935) 617. doi:10.1051/JCP/1935320589.
- [264] F. M. Jaeger, Über die Temperaturabhängigkeit der molekularen freien Oberflächenenergie von Flüssigkeiten im Temperaturbereich von - 80 bis + 1650C, *Zeitschrift für anorganische und allgemeine Chemie* 101 (1) (1917) 1–214. doi:https://doi.org/10.1002/zaac.19171010102.
- [265] F. B. Thole, Viscosity and association. Part IV. The viscosity of the aromatic amines, *Journal of the Chemical Society, Transactions* 103 (1913) 317–323. doi:10.1039/CT9130300317.
- [266] A. Bramley, The study of binary mixtures. Part I. The densities and viscosities of mixtures containing phenol, *Journal of the Chemical Society, Transactions* 109 (1916) 10–45. doi:10.1039/CT9160900010.
- [267] E. W. Hough, D. M. Mason, B. H. Sage, Heat Capacities of Several Organic Liquids,

- Journal of the American Chemical Society 72 (12) (1950) 5775–5777. doi:10.1021/JA01168A531.
- [268] C. A. Buehler, J. H. Wood, D. C. Hull, E. C. Erwin, A study of molecular organic compounds. IV. The molecular organic compounds of phenol, their parachors and refractivities, Journal of the American Chemical Society 54 (6) (1932) 2398–2405. doi:10.1021/ja01345a030.
- [269] G. W. A. Kahlbaum, C. G. von Wirkner, Studies on vapor stress measurements, Monograph (1897) 1–222.
- [270] A. A. Ravdel', V. V. Danilov, Determination of the vapor pressure over some binary solutions, Izv. Vyssh. Uchebn. Zaved. Khim. Khim. Tekhnol. 11 (6) (1968) 642–645.
- [271] A. M. Lastovtsev, The investigation of P-T dependence for saturated vapor of organic heat carrier, Khim. Mashinostr. Moscow 6 (3) (1937) 19–22.
- [272] N. B. Vargaftik, Handbook of thermophysical properties of gases and liquids, Moscow, 1972.
- [273] G. Liessmann, W. Schmidt, S. Reiffarth, Recommended thermophysical data, Data compilation of the Saechsische Olefinwerke Boehlen Germany 1 (1995).
- [274] W. E. Hatton, D. L. Hildenbrand, G. C. Sinke, D. R. Stull, Chemical Thermodynamic Properties of Aniline., Journal of Chemical & Engineering Data 7 (2) (1962) 229–231. doi:10.1021/JE60013A021.
- [275] D. Hopfe, Thermophysical data of pure substances, Data Compilation of FIZ CHEMIE Germany (20) (1990).
- [276] D. D. Deshpande, M. V. Pandya, Thermodynamics of binary solutions. Part 1.—Heats of mixing of aniline in benzene, carbontetrachloride and chlorobenzene, Transactions of the Faraday Society 61 (1965) 1858–1868. doi:10.1039/TF9656101858.
- [277] K. Merkel, Formation of Over-Molecules in Liquids, Nova Acta Leopoldina 9 (1940) 243–303.

- [278] R. K. Nigam, P. P. Singh, K. C. Singh, Thermodynamics of molecular interactions in aniline + benzene mixtures, *Canadian Journal of Chemistry* 57 (17) (1979) 2211–2216. doi:10.1139/v79-354.
- [279] G. Pannetier, L. Abello, Contribution a l'étude de la liaison hydrogene N-N-N, *Bull. Soc. Chim. Fr.* (32) (1965) 2048–2058.
- [280] I. Nagata, K. Tamura, Excess enthalpies of (aniline + acetonitrile or benzene) and of (aniline + acetonitrile + benzene) at the temperature 298.15 K, *The Journal of Chemical Thermodynamics* 24 (6) (1992) 613–617. doi:[https://doi.org/10.1016/S0021-9614\(05\)80032-1](https://doi.org/10.1016/S0021-9614(05)80032-1).
- [281] K. Kroenlein, C. D. Muzny, A. F. Kazakov, V. Diky, R. D. Chirico, J. W. Magee, I. Abdulagatov, M. Frenkel, NIST/TRC Web Thermo Tables (WTT) NIST Standard Reference Subscription Database 2-Lite Edition Version 2, 2011.
- [282] T. Takagi, Ultrasonic speeds and thermodynamics of (benzene + aniline) and (chlorobenzene + aniline) under high pressures, *The Journal of Chemical Thermodynamics* 12 (3) (1980) 277–286. doi:[https://doi.org/10.1016/0021-9614\(80\)90047-6](https://doi.org/10.1016/0021-9614(80)90047-6).
- [283] G. Kortüm, D. Moegling, F. Woerner, Die Leistungsfähigkeit von Umlaufapparaturen zur Bestimmung von Siede- und Dampfdruckdiagrammen binärer Flüssigkeitsgemische, *Chemie Ingenieur Technik* 22 (21) (1950) 453–457. doi:10.1002/CITE.330222102.
- [284] F. Schwers, La densité et l'indice de réfraction des solutions. Nouvelles contributions à l'étude des solutions, *J. Chim. Phys.* 9 (1911) 15–100.
- [285] L. K. Zharikov, K. S. Krylova, G. M. Kopylevich, N. K. Tikhonova, G. K. Oparina, L. A. Serafimov, Phase Equilibria in the Systems Water-Aniline, Water-Cyclohexanol, Ethanol-Aniline and Ethanol-Cyclohexanol, *Journal of Applied Chemistry of the USSR* 48 (6) (1975) 1306–1308.
- [286] G. Y. Kolyuchkina, V. S. Timofeev, L. A. Serafimov, Study on the Vapor-Liquid Equilibrium in the System Heptane - Water - Aniline, *Uch. Zap. Mosk. Inst. Tonkoi Khim. Tekhnol.* 1 (1971) 65–72.

- [287] J. Griswold, D. Andres, E. F. Arnett, F. M. Garland, Liquid–Vapor Equilibrium of Aniline–Water, *Industrial & Engineering Chemistry* 32 (6) (1940) 878–880. doi:10.1021/ie50366a029.
- [288] I. Horyna, Vapor-Liquid Equilibrium of Aqueous Aniline Solutions at Higher Pressures, *Collect. Czech. Chem. Commun.* 24 (1959) 3253–3259.
- [289] A. N. Campbell, The System Aniline-Phenol-Water, *Journal of the American Chemical Society* 67 (6) (1945) 981–987. doi:10.1021/ja01222a029.
- [290] D. J. Donahue, F. E. Bartell, The Boundary Tension at Water-Organic Liquid Interfaces, *The Journal of Physical Chemistry* 56 (4) (1952) 480–484. doi:10.1021/j150496a016.
- [291] A. I. Vogel, Physical properties and chemical constitution. Part XV. The phenyl group, *Journal of the Chemical Society (Resumed)* (1948) 654–658 doi:10.1039/JR9480000654.
- [292] B. Garcia, J. C. Ortega, Excess viscosity, excess volume, and excess free energy of activation at 283, 293, 303, 313, and 323 K for mixtures acetonitrile and alkyl benzoates, *Journal of Chemical & Engineering Data* 33 (2) (1988) 200–204. doi:10.1021/je00052a041.
- [293] H. Katayama, Vapor pressures of methyl, ethyl, n-propyl, isobutyl, and n-butyl benzoates at reduced pressures, *Journal of Chemical & Engineering Data* 33 (2) (1988) 75–77. doi:10.1021/je00052a002.
- [294] W. V. Steele, R. D. Chirico, A. B. Cowell, S. E. Knipmeyer, A. Nguyen, Thermodynamic Properties and Ideal-Gas Enthalpies of Formation for Methyl Benzoate, Ethyl Benzoate, (R)-(+)-Limonene, tert-Amyl Methyl Ether, trans-Crotonaldehyde, and Diethylene Glycol, *Journal of Chemical & Engineering Data* 47 (4) (2002) 667–688. doi:10.1021/je0100847.
- [295] W. Hieber, E. Reindl, Bildungswärmen und Konstitution von Molekülverbindungen des Zinntetrachlorids, *Zeitschrift für Elektrochemie und angewandte physikalische Chemie* 46 (10) (1940) 559–570.

- [296] J. Timmermans, La température de congélation des substances organiques, IV: Nouvelles déterminations expérimentales, *Bulletin de la Société chimique de Belgique* 30 (1921) 62–72.
- [297] T. V. Vasiltsova, S. P. Verevkin, E. Bich, A. Heintz, R. Bogel-Lukasik, U. Domańska, Thermodynamic Properties of Mixtures Containing Ionic Liquids. 7. Activity Coefficients of Aliphatic and Aromatic Esters and Benzylamine in 1-Methyl-3-ethylimidazolium Bis(trifluoromethylsulfonyl) Imide Using the Transpiration Method, *Journal of Chemical & Engineering Data* 51 (1) (2006) 213–218. doi:10.1021/je050334+.
- [298] N. N. Kedrina, L. V. Semenov, A. A. Gaile, Enthalpy of mixing of hydrocarbons with polar solvents, *Deposited Doc Oniitekhim* (1981) 1–10.
- [299] J.-P. E. Grolier, K. Sosnkowska-Kehiaian, H. V. Kehiaian, Enthalpies de melange des chlorures organiques avec des hydrocarbures, *Journal de Chimie Physique* 70 (1973) 367–373. doi:10.1051/jcp/1973700367.
- [300] W. J. Svirbely, J. E. Ablard, J. C. Warner, Molar Polarizations in Extremely Dilute Solutions. The Dipole Moments of d- Limonene, d-Pinene, Methyl Benzoate and Ethyl Benzoate¹, *Journal of the American Chemical Society* 57 (4) (1935) 652–655. doi:10.1021/ja01307a015.
- [301] A. E. Lutsky, E. Obukhova, Molecular constants and macro-physical properties of mixtures. 2. the density of liquid binary systems, *Zhurnal Fizicheskoi Khimii* 31 (9) (1957) 1964–1975.
- [302] T. Thenappan, U. Sankar, Study of correlation factors and dipolar excess free energies of esters in benzene, *Journal of Molecular Liquids* 126 (1) (2006) 23–28. doi:10.1016/j.molliq.2004.09.004.
- [303] R. Stephenson, J. Stuart, Mutual binary solubilities: water-alcohols and water-esters, *Journal of Chemical & Engineering Data* 31 (1) (1986) 56–70. doi:10.1021/je00043a019.

- [304] I. L. Krupatkin, M. F. Glagoleva, Liquid-Phase Equilibrium in Systems of Furfurol-Water High-Boiling Esters, *Zhurnal Prikladnoi Khimii* 45 (8) (1972) 1872–1875.
- [305] A. Khaoulani Idrissi, M. Rafiq, P. Gougeon, R. Guerin, Anilinium Hydrogen Monofluorophosphate, $C_6H_8N^+.HPO_3F^-$, *Acta Crystallographica Section C* 51 (7) (1995) 1359–1361. doi:10.1107/s010827019401214x.
- [306] Z. Boutobba, A. Direm, N. Benali-Cherif, Anilinium hydrogen sulfate, *Acta Crystallographica Section E* 66 (3) (2010) 595–596. doi:10.1107/S1600536810004782.
- [307] H. S. Hu, Y. L. Wu, M. D. Yang, Determination of saturated vapour pressure of aniline hydrochloride and vapour-liquid equilibrium of the water- aniline hydrochloride system, *Fluid Phase Equilibria* 414 (2016) 133–142. doi:10.1016/j.fluid.2015.12.053.
- [308] P. Guerrieri, A. C. Rumondor, T. Li, L. S. Taylor, Analysis of relationships between solid-state properties, counterion, and developability of pharmaceutical salts, *AAPS PharmSciTech* 11 (3) (2010) 1212–1222. doi:10.1208/s12249-010-9499-4.
- [309] S. Yalkowsky, Y. He, P. Jain, *Handbook of aqueous solubility data*, 2nd Edition, CRC press, Boca Raton, 2016.
- [310] K. Box, J. Comer, Using Measured pKa, LogP and Solubility to Investigate Supersaturation and Predict BCS Class, *Current Drug Metabolism* 9 (9) (2008) 869–878. doi:10.2174/138920008786485155.
- [311] D. R. Delgado, E. F. Vargas, F. Martínez, Thermodynamic Study of the Solubility of Procaine HCl in Some Ethanol + Water Cosolvent Mixtures, *Journal of Chemical & Engineering Data* 55 (8) (2010) 2900–2904. doi:10.1021/je900958z.
- [312] I. Labastidas, F. Martinez, Aspectos termodinámicos de la solubilidad acuosa de algunas sales orgánicas de interés farmacéutico, *Acta Farmacéutica Bonaerense* 25 (1) (2006) 55–63.
- [313] R. N. Hader, R. D. Wallace, R. W. McKinney, Formaldehyde from Methanol, *Industrial & Engineering Chemistry Research* 44 (7) (1952) 1508–1518.

- [314] J. F. Walker, Formaldehyde, American Chemical Society Monograph Series, 3rd Edition, Reinhold Publishing Corporation, London, 1964.
- [315] X. J. Liu, S. Ahlgren, H. Korthout, L. F. Salome-Abarca, L. M. Bayona, R. Verpoorte, Y. H. Choi, Broad range range chemical profiling of natural deep eutectic solvent extracts using a high performance thin layer chromatography-based method, *Journal of Chromatography A* 1532 (2018) 198–207. doi:10.1016/j.chroma.2017.12.009.
- [316] H. Matsuda, K. Mori, M. Tomioka, N. Kariyasu, T. Fukami, K. Kurihara, K. Tochigi, K. Tomono, Determination and prediction of solubilities of active pharmaceutical ingredients in selected organic solvents, *Fluid Phase Equilibria* 406 (2015) 116–123. doi:10.1016/j.fluid.2015.07.032.
- [317] Y. Wu, J. Levons, A. S. Narang, K. Raghavan, V. M. Rao, Reactive impurities in excipients: Profiling, identification and mitigation of drug-excipient incompatibility, *AAPS PharmSciTech* 12 (4) (2011) 1248–1263.
- [318] M. Fujita, T. Ueda, T. Handa, Generation of formaldehyde by pharmaceutical excipients and its absorption by meglumine, *Chemical and Pharmaceutical Bulletin* 57 (10) (2009) 1096–1099. doi:10.1248/CPB.57.1096.
- [319] G. Maurer, Vapor-liquid equilibrium of formaldehyde-and water-containing multicomponent mixtures, *AIChE Journal* 32 (6) (1986) 932–948. doi:10.1002/aic.690320604.
- [320] H. Hasse, G. Maurer, Vapor-liquid equilibrium of formaldehyde-containing mixtures at temperatures below 320 K, *Fluid Phase Equilibria* 64 (1991) 185–199. doi:10.1016/0378-3812(91)90013-W.
- [321] L. V. Kogan, Y. M. Blazhin, S. K. Ogorodnikov, V. V. Kafarov, Liquid-Vapor Equilibrium in the System Formaldehyde - Water, *Zhur Prikl Khim* 50 (12) (1977) 2552–2556.
- [322] Y. L. Bondarenko, I. I. Sabylin, M. A. Kharisov, K. I. Strel'nikov, O. D. Soboleva, Y. M. Blazhin, S. K. Ogorodnikov, Mathematical Modeling of Phase Equilibrium and Distillation for the System Formaldehyde - Water, *Promst. Sint. Kauch.* (8) (1981) 2–4.

- [323] V. Brandani, G. Di Giacomo, P. U. Foscolo, Isothermal vapor-liquid equilibria for the water-formaldehyde system. A predictive thermodynamic model., *Industrial & Engineering Chemistry Process Design and Development* 19 (1) (1980) 179–185.
- [324] S. J. Green, R. E. Vener, Vapor-liquid equilibria of formaldehyde-methanol-water, *Industrial & Engineering Chemistry* 47 (1) (1955) 103–109.
- [325] V. M. Olevskii, I. F. Golub'ev, Vapor-Liquid Equilibrium in the Methanol-Formaldehyde-Water at Atmospheric Pressure and at Vacuum Conditions, *Tr. Gos. Nauchno Issled. Proektn. Inst. Azotn. Promst. Prod. Org. Sin.* 4 (1954) 36–59.
- [326] Y. M. Blazhin, L. V. Kogan, L. K. Vagina, V. E. Pastor, A. I. Morozova, S. K. Ogorodnikov, Liquid-Vapor Equilibrium in the System Formaldehyde-Methanol-Water Under Atmospheric and Reduced Pressure, *Journal of Applied Chemistry of the USSR* 49 (174) (1976) 167–170.
- [327] L. V. Kogan, S. K. Ogorodnikov, Liquid-Vapor Equilibrium in the System Formaldehyde-Methanol, *Journal of Applied Chemistry of the USSR* 53 (1) (1980) 98–102.
- [328] M. Albert, B. Coto García, C. Kuhnert, R. Peschla, G. Maurer, Vapor-liquid equilibrium of aqueous solutions of formaldehyde and methanol, *AIChE Journal* 46 (8) (2000) 1676–1687. doi:10.1002/aic.690460818.
- [329] L. V. Kogan, S. K. Ogorodnikov, Liquid-Vapor Equilibrium in the System Formaldehyde-Methanol-Water, *Journal of Applied Chemistry of the USSR* 53 (1) (1980) 102.
- [330] V. Brandani, G. Di Giacomo, Effect of small amounts of methanol on the vapour-liquid equilibrium for the water-formaldehyde system, *Fluid Phase Equilibria* 24 (3) (1985) 307–333. doi:10.1016/0378-3812(85)85011-1.
- [331] M. Albert, I. Hahnenstein, H. Hasse, G. Maurer, Vapor-Liquid Equilibrium of Formaldehyde Mixtures: New Data and Model Revision, *AIChE Journal* 42 (6) (1996) 1741–1752. doi:10.1002/aic.690420625.

- [332] M. Albert, B. G. Garcia, C. Kreiter, G. Maurer, Vapor-liquid and chemical equilibria of formaldehyde-water mixtures, *AIChE Journal* 45 (9) (1999) 2024–2033. doi:10.1002/AIC.690450919.
- [333] V. Brandani, G. Di Giacomo, V. Mucciante, Vapor-liquid equilibrium of formaldehyde-water-methanol mixtures, *Quaderni dell'ingegnere chimico italiano* 23 (7-8) (1987) 2–4.
- [334] G. M. Wilson, Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing, *Journal of the American Chemical Society* 86 (2) (1964) 127–130. doi:10.1021/JA01056A002.
- [335] S. Brandani, V. Brandani, G. Di Giacomo, A physical theory superimposed onto the chemical theory for describing vapor-liquid equilibria of binary systems of formaldehyde in active solvents, *Industrial & Engineering Chemistry Research* 30 (2) (1991) 414–420.
- [336] D. S. Abrams, J. M. Prausnitz, Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems, *AIChE Journal* 21 (1) (1975) 116–128. doi:10.1002/AIC.690210115.
- [337] G. Maurer, J. M. Prausnitz, On the derivation and extension of the uniquac equation, *Fluid Phase Equilibria* 2 (2) (1978) 91–99. doi:10.1016/0378-3812(78)85002-X.
- [338] S. Brandani, V. Brandani, G. D. Giacomo, The System Formaldehyde-Water-Methanol: Thermodynamics of Solvated and Associated Solutions, *Industrial & Engineering Chemistry Research* 31 (7) (1992) 1792–1798.
- [339] S. Brandani, V. Brandani, I. Tarquini, Vapor-liquid equilibrium of formaldehyde mixtures containing methanol, *Industrial & Engineering Chemistry Research* 37 (8) (1998) 3485–3489. doi:10.1021/ie980020y.
- [340] A. V. Rudnev, E. P. Kalyazin, K. S. Kalugin, G. V. Kovalev, Formaldehyde. 2. Desolvation Kinetics Of Formaldehyde in Aqueous and Methanol Solutions, *Journal of Physical Chemistry* 51 (10) (1977) 2603–2606.

- [341] H.-G. Schecker, G. Schulz, Untersuchungen zur Hydratationskinetik von Formaldehyd in wäßriger Lösung, *Journal of Physical Chemistry* 65 (1969) 221–224.
- [342] D. J. Le Botlan, B. G. Mechin, G. J. Martin, Proton and Carbon-13 Nuclear Magnetic Resonance Spectrometry of Formaldehyde in Water, *Analytical Chemistry* 55 (3) (1983) 587–591. doi:10.1021/ac00254a041.
- [343] I. Hahnenstein, H. Hasse, C. G. Kreiter, G. Maurer, ¹H and ¹³C-NMR Spectroscopic Study of Chemical Equilibria in Solutions of Formaldehyde in Water, Deuterium Oxide, and Methanol, *Industrial & Engineering Chemistry* 33 (4) (1994) 1022–1029.
- [344] A. L. Balashov, S. M. Danov, A. Golovkin, V. L. Krasnov, A. N. Ponomarev, Equilibrium mixture of polyoxymethylene glycols in concentrated aqueous formaldehyde solutions, *Russian Journal of Applied Chemistry* 69 (2) (1996) 190–192.
- [345] M. Maiwald, H. H. Fischer, M. Ott, R. Peschla, C. Kuhnert, C. G. Kreiter, G. Maurer, H. Hasse, Quantitative NMR spectroscopy of complex liquid mixtures: methods and results for chemical equilibria in formaldehyde-water-methanol at temperatures up to 383 K, *Industrial & Engineering Chemistry Research* 42 (2) (2003) 259–266. doi:10.1021/ie0203072.
- [346] I. G. Economou, M. D. Donohue, Chemical, quasi-chemical and perturbation theories for associating fluids, *AIChE Journal* 37 (12) (1991) 1875–1894. doi:10.1002/aic.690371212.
- [347] N. Mac Dowell, F. Llovel, C. S. Adjiman, G. Jackson, A. Galindo, Modeling the fluid phase behavior of carbon dioxide in aqueous solutions of monoethanolamine using transferable parameters with the SAFT-VR approach, *Industrial & Engineering Chemistry Research* 49 (4) (2010) 1883–1899. doi:10.1021/ie901014t.
- [348] J. Rodriguez, N. Mac Dowell, F. Llovel, C. S. Adjiman, G. Jackson, A. Galindo, Modelling the fluid phase behaviour of aqueous mixtures of multifunctional alkanolamines and carbon dioxide using transferable parameters with the SAFT-VR approach, *Molecular Physics* 110 (11-12) (2012) 1325–1348. doi:10.1080/00268976.2012.665504.

- [349] A. Chremos, E. Forte, V. Papaioannou, A. Galindo, G. Jackson, C. S. Adjiman, Modelling the phase and chemical equilibria of aqueous solutions of alkanolamines and carbon dioxide using the SAFT- γ SW group contribution approach, *Fluid Phase Equilibria* 407 (2016) 280–297. doi:10.1016/J.FLUID.2015.07.052.
- [350] F. Perdomo, S. Khalit, E. Graham, F. Tzirakis, A. Papadopoulos, I. Tsvintzelis, A. Papadopoulos, P. Seferlis, A. Galindo, G. Jakson, C. Adjiman, Thermodynamic modelling of CO₂ absorption in cyclic amines, alkyl polyamines, alkanolamines and phase-change amines using the SAFT- γ Mie group contribution approach (2022) In preparation.
- [351] R. Spence, W. Wild, 114. The vapour-pressure curve of formaldehyde, and some related data, *Journal of the Chemical Society (Resumed)* (1935) 506–509doi:10.1039/JR9350000506.
- [352] H. R. Gerberich, G. C. Seaman, Formaldehyde, in: *Kirk-Othmer Encyclopedia of Chemical Technology*, American Cancer Society, Hoboken, NJ, USA, 2013. doi:10.1002/0471238961.0615181307051802.A01.PUB3.
- [353] M. Ott, H. H. Fischer, M. Maiwald, K. Albert, H. Hasse, Kinetics of oligomerization reactions in formaldehyde solutions: NMR experiments up to 373 K and thermodynamically consistent model, *Chemical Engineering and Processing: Process Intensification* 44 (6) (2005) 653–660. doi:10.1016/j.cep.2003.07.004.
- [354] P. J. Flory, Thermodynamics of High Polymer Solutions, *The Journal of Chemical Physics* 10 (1) (1942) 51–61. doi:10.1063/1.1723621.
- [355] M. L. Huggins, Thermodynamic properties of solutions of long-chain compounds, *Annals of the New York Academy of Sciences* 43 (1) (1942) 1–32. doi:10.1111/J.1749-6632.1942.TB47940.X.
- [356] A. Gautam, X. Pan, The changing model of big pharma: Impact of key trends (3 2016). doi:10.1016/j.drudis.2015.10.002.

Appendix

Table A.1: SAFT- γ Mie like group parameters used in the current work: ν_k^* is the number of segments in group k ; S_k and σ_{kk} are the shape factor and diameter of the segments in the group, respectively; $\sigma_{kk}^{\text{Born}}$ the Born diameter used for charged groups; λ_{kk}^r and λ_{kk}^a are the repulsive and attractive exponents of the Mie potential, respectively; ε_{kk} is the dispersion energy between two k groups; $N_{\text{ST},k}$ is the number of association site types on group k , with $n_{\text{H},k}$, $n_{\text{e}_1,k}$ and $n_{\text{e}_2,k}$ representing the number of association sites of type H, e₁ and e₂, respectively. †† indicates parameters obtained in Chapter 4, † † † interactions developed in Chapter 5 and † correspond to interactions developed in Chapter 3 which are also published in ref [44].

Group k	ν_k^*	S_k	$\sigma_{kk}/\text{\AA}$	$\sigma_{kk}^{\text{Born}}/\text{\AA}$	λ_{kk}^r	λ_{kk}^a	$(\varepsilon_{kk}/k_{\text{B}})/\text{K}$	$N_{\text{ST},k}$	$n_{\text{H},k}$	$n_{\text{e}_1,k}$	$n_{\text{e}_2,k}$	ref
H ₂ O	1	1.0000	3.0063	-	17.020	6.0000	266.68	2	2	2	-	[37, 38]
CH ₃	1	0.57255	4.0772	-	15.050	6.0000	256.77	-	-	-	-	[35]
aCH	1	0.32184	4.0578	-	14.756	6.0000	371.53	1	-	1	-	[36]
aCCH	1	0.20650	4.3128	-	8.0000	6.0000	61.325	1	-	1	-	[36]
aCCH ₂	1	0.20859	5.2648	-	8.5433	6.0000	591.56	1	-	1	-	[36]
CH	1	0.07210	5.2950	-	8.0000	6.0000	95.621	-	-	-	-	[36]
COOH	1	0.55593	4.3331	-	8.0000	6.0000	405.78	3	1	2	2	[36]
aCCOaC	3	0.18086	4.0670	-	9.8317	6.0000	656.71	-	-	-	-	[39]
CH ₂	1	0.22932	4.8801	-	19.871	6.0000	473.39	-	-	-	-	[35]
NH ₂	1	0.79675	3.2477	-	10.254	6.0000	284.78	2	2	1	-	[167]
C	1	0.04072	5.6571	-	8.0000	6.0000	50.020	-	-	-	-	[36]
N	1	0.15069	3.0755	-	8.8971	6.0000	62.971	1	-	1	-	[167]
aCCOO	2	0.65563	3.3851	-	36.616	6.0000	200.00	2	-	2	2	††
aCNH ₂	1	0.71518	3.7408	-	17.065	6.0000	670.91	2	2	1	-	††
CH ₂ O	2	0.78473	3.3131	-	36.423	6.0000	280.15	3	2	2	1	† † †

Table A.1: Continued from previous page

Group k	ν_k^*	S_k	$\sigma_{kk}/\text{\AA}$	$\sigma_{kk}^{\text{Born}}/\text{\AA}$	λ_{kk}^r	λ_{kk}^a	$(\varepsilon_{kk}/k_B)/\text{K}$	$N_{\text{ST},k}$	$n_{\text{H},k}$	$n_{\text{e1},k}$	$n_{\text{e2},k}$	ref
CH ₃ OH	2	0.83517	3.2462	-	19.235	6.0000	307.69	2	1	2	-	[36]
COO ⁻	1	0.55593	4.3331	4.6364	8.0000	6.0000	21.264	1	-	4	-	[42, 43]
NH ₃ ⁺	1	0.79675	3.2477	3.4750	10.254	6.0000	48.295	1	3	-	-	[44] †
N ⁺	1	0.15069	3.0755	3.2908	8.8971	6.0000	62.971	1	1	-	-	[42, 43]
aCNH ₃ ⁺	1	0.71518	3.7408	4.0027	17.065	6.0000	397.63	1	3	-	-	††
H ₃ O ⁺	1	1.0000	3.0063	3.0063	17.020	6.0000	68.190	1	3	-	-	[150]
OH ⁻	1	1.0000	2.4600	3.0063	17.020	6.0000	170.24	1	-	3	-	[150]
Cl ⁻	1	1.0000	3.3400	3.8740	12.000	6.0000	113.77	-	-	-	-	[150]
Br ⁻	1	1.0000	3.6400	4.1740	12.000	6.0000	107.38	-	-	-	-	[150]
Na ⁺	1	1.0000	2.3200	3.3600	12.000	6.0000	31.711	-	-	-	-	[150]
K ⁺	1	1.0000	3.0400	4.3440	12.000	6.0000	90.097	-	-	-	-	[150]
Li ⁺	1	1.0000	1.8000	2.6320	12.000	6.0000	6.1039	-	-	-	-	[150]
Rb ⁺	1	1.0000	3.3200	4.6220	12.000	6.0000	130.35	-	-	-	-	[150]
Ca ²⁺	1	1.0000	2.2800	3.7240	12.000	6.0000	261.28	-	-	-	-	[150]
Mg ²⁺	1	1.0000	1.7200	2.9100	12.000	6.0000	89.264	-	-	-	-	[150]

Table A.2: Group dispersion interaction energies ε_{kl} and repulsive exponent λ_{kl}^r for use with the SAFT- γ Mie equation of state in the current work. The unlike segment diameters σ_{kl} are calculated using the combining rule in equation (2.35). CR indicates that combining rules are used to obtain ε_{kl} and/or λ_{kl}^r using equations (2.36) and/or (2.37), respectively. $\dagger\dagger$ indicates parameters obtained in Chapter 4, $\dagger\dagger\dagger$ interactions developed in Chapter 5 and $\dagger$ correspond to interactions developed in Chapter 3 which are also published in ref [44].

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref	Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref
H ₂ O	H ₂ O	266.68	17.020	[37, 38]	COOH	NH ₃ ⁺	CR	CR	-
H ₂ O	CH ₃	358.18	100.00	[39]	COOH	H ₃ O ⁺	CR	CR	[44] $\dagger$
H ₂ O	aCH	357.78	38.640	[36]	COOH	OH ⁻	CR	CR	[44] $\dagger$
H ₂ O	aCCH	314.03	CR	[40]	COOH	Cl ⁻	CR	CR	-
H ₂ O	aCCH ₂	329.03	CR	[40]	COOH	Na ⁺	CR	CR	-
H ₂ O	CH	275.75	CR	[40]	COOH	K ⁺	CR	CR	-
H ₂ O	COOH	289.76	CR	[39]	COOH	Li ⁺	CR	CR	-
H ₂ O	aCCOaC	CR	CR	-	COOH	Rb ⁺	CR	CR	-
H ₂ O	CH ₂	423.63	100.00	[39]	COOH	Ca ²⁺	CR	CR	-
H ₂ O	NH ₂	358.55	CR	[167]	COOH	Mg ²⁺	CR	CR	-
H ₂ O	C	420.82	CR	[40]	aCCOaC	aCCOaC	656.71	9.8317	[39]
H ₂ O	N	1481.3	21.217	[167]	aCCOaC	COO ⁻	CR	CR	-
H ₂ O	aCCOO	339.28	CR	$\dagger\dagger$	aCCOaC	H ₃ O ⁺	CR	CR	-
H ₂ O	aCNH ₂	438.2839	CR	$\dagger\dagger$	aCCOaC	OH ⁻	CR	CR	-
H ₂ O	CH ₂ O	248.26	CR	$\dagger\dagger\dagger$	aCCOaC	Cl ⁻	CR	CR	-
H ₂ O	CH ₃ OH	278.45	CR	[36]	aCCOaC	Na ⁺	CR	CR	-
H ₂ O	COO ⁻	171.61	CR	[42, 43]	aCCOaC	K ⁺	CR	CR	-
H ₂ O	NH ₃ ⁺	450.21	CR	[44] $\dagger$	aCCOaC	Li ⁺	CR	CR	-
H ₂ O	N ⁺	1481.3	21.217	[42, 43]	aCCOaC	Rb ⁺	CR	CR	-
H ₂ O	aCNH ₃ ⁺	1382.5	CR	$\dagger\dagger$	aCCOaC	Ca ²⁺	CR	CR	-
H ₂ O	H ₃ O ⁺	391.04	CR	[150]	aCCOaC	Mg ²⁺	CR	CR	-
H ₂ O	OH ⁻	134.41	CR	[150]	CH ₂	CH ₂	473.39	19.871	[35]
H ₂ O	Cl ⁻	95.406	CR	[150]	CH ₂	NH ₂	348.39	CR	[167]
H ₂ O	Br ⁻	112.01	CR	[150]	CH ₂	C	300.07	CR	[36]
H ₂ O	Na ⁺	539.68	CR	[150]	CH ₂	N	348.30	CR	[167]

Table A.2: Continued from previous page

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref	Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref
H ₂ O	K ⁺	376.25	CR	[150]	CH ₂	aCCOO	446.39	CR	††
H ₂ O	Li ⁺	1023.1	CR	[150]	CH ₂	aCNH ₂	497.03	CR	††
H ₂ O	Rb ⁺	354.23	CR	[150]	CH ₂	CH ₃ OH	341.41	17.050	[36]
H ₂ O	Ca ²⁺	1460.8	CR	[150]	CH ₂	COO ⁻	413.74	CR	[42, 43]
H ₂ O	Mg ²⁺	2235.1	CR	[150]	CH ₂	NH ₃ ⁺	348.39	CR	[44] †
CH ₃	CH ₃	256.77	15.050	[35]	CH ₂	N ⁺	348.30	CR	[42, 43]
CH ₃	aCH	305.81	CR	[36]	CH ₂	aCNH ₃ ⁺	CR	CR	-
CH ₃	aCCH	455.85	CR	[36]	CH ₂	H ₃ O ⁺	CR	CR	-
CH ₃	aCCH ₂	396.91	CR	[36]	CH ₂	OH ⁻	CR	CR	-
CH ₃	CH	387.48	CR	[36]	CH ₂	Cl ⁻	CR	CR	-
CH ₃	COOH	255.99	CR	[36]	CH ₂	Br ⁻	CR	CR	-
CH ₃	aCCOaC	CR	CR	-	CH ₂	Na ⁺	CR	CR	-
CH ₃	CH ₂	350.77	CR	[35]	NH ₂	NH ₂	284.78	10.254	[167]
CH ₃	NH ₂	244.15	CR	[167]	NH ₂	C	CR	CR	-
CH ₃	C	339.91	CR	[36]	NH ₂	COO ⁻	CR	CR	-
CH ₃	N	462.18	CR	[167]	NH ₂	NH ₃ ⁺	284.78	CR	[44] †
CH ₃	aCCOO	646.59	18.865	††	NH ₂	H ₃ O ⁺	CR	CR	-
CH ₃	aCNH ₂	464.73	CR	††	NH ₂	OH ⁻	CR	CR	-
CH ₃	CH ₃ OH	275.76	15.537	[36]	NH ₂	Br ⁻	CR	CR	-
CH ₃	COO ⁻	255.99	CR	[42, 43]	C	C	50.020	8.0000	[36]
CH ₃	NH ₃ ⁺	244.15	CR	[44] †	C	COO ⁻	CR	CR	-
CH ₃	N ⁺	462.18	CR	[42, 43]	C	NH ₃ ⁺	106.49	CR	[44] †
CH ₃	aCNH ₃ ⁺	CR	CR	-	C	H ₃ O ⁺	CR	CR	-
CH ₃	H ₃ O ⁺	CR	CR	-	C	OH ⁻	CR	CR	-
CH ₃	OH ⁻	CR	CR	-	N	N	62.971	8.8971	[167]
CH ₃	Cl ⁻	CR	CR	-	N	aCCOO	CR	CR	-
CH ₃	Br ⁻	CR	CR	-	N	aCNH ₂	CR	CR	-
CH ₃	Na ⁺	CR	CR	-	N	N ⁺	CR	CR	-
CH ₃	K ⁺	CR	CR	-	N	aCNH ₃ ⁺	CR	CR	-

Table A.2: Continued from previous page

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref	Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref
CH ₃	Li ⁺	CR	CR	-	N	H ₃ O ⁺	CR	CR	††
CH ₃	Rb ⁺	CR	CR	-	N	OH ⁻	CR	CR	-
CH ₃	Ca ²⁺	CR	CR	-	N	Cl ⁻	CR	CR	-
CH ₃	Mg ²⁺	CR	CR	-	N	Na ⁺	CR	CR	-
aCH	aCH	371.53	14.756	[36]	aCCOO	aCCOO	200.00	36.616	††
aCH	aCCH	429.16	CR	[36]	aCCOO	aCNH ₂	494.91	CR	††
aCH	aCCH ₂	416.69	CR	[36]	aCCOO	N ⁺	CR	CR	-
aCH	CH	441.43	CR	[36]	aCCOO	aCNH ₃ ⁺	CR	CR	††
aCH	COOH	331.61	9.0687	[36]	aCCOO	H ₃ O ⁺	CR	CR	††
aCH	aCCOaC	CR	CR	-	aCCOO	OH ⁻	CR	CR	-
aCH	CH ₂	415.64	CR	[36]	aCCOO	Cl ⁻	CR	CR	-
aCH	NH ₂	CR	CR	-	aCCOO	Na ⁺	CR	CR	-
aCH	C	CR	CR	-	aCNH ₂	aCNH ₂	670.91	17.065	††
aCH	N	CR	CR	-	aCNH ₂	N ⁺	CR	CR	-
aCH	aCCOO	468.32	CR	††	aCNH ₂	aCNH ₃ ⁺	CR	CR	††
aCH	aCNH ₂	564.39	23.276	††	aCNH ₂	H ₃ O ⁺	CR	CR	††
aCH	COO ⁻	CR	CR	-	aCNH ₂	OH ⁻	CR	CR	††
aCH	NH ₃ ⁺	CR	CR	-	aCNH ₂	Cl ⁻	CR	CR	-
aCH	N ⁺	CR	CR	-	aCNH ₂	Na ⁺	CR	CR	-
aCH	aCNH ₃ ⁺	CR	CR	-	CH ₂ O	CH ₂ O	280.15	36.423	†††
aCH	H ₃ O ⁺	CR	CR	[44] †	CH ₂ O	CH ₃ OH	320.98	CR	†††
aCH	OH ⁻	CR	CR	-	CH ₃ OH	CH ₃ OH	307.69	19.235	[36]
aCH	Cl ⁻	CR	CR	-	COO ⁻	COO ⁻	21.264	8.0000	[42, 43]
aCH	Na ⁺	CR	CR	-	COO ⁻	NH ₃ ⁺	26.327	CR	[44] †
aCH	K ⁺	CR	CR	-	COO ⁻	N ⁺	24.277	CR	[42, 43]
aCH	Li ⁺	CR	CR	-	COO ⁻	H ₃ O ⁺	28.111	CR	[44] †
aCH	Rb ⁺	CR	CR	-	COO ⁻	OH ⁻	44.520	CR	[44] †
aCH	Ca ²⁺	CR	CR	-	COO ⁻	Cl ⁻	21.265	CR	[42, 43]
aCH	Mg ²⁺	CR	CR	-	COO ⁻	Na ⁺	9.9125	CR	[42, 43]

Table A.2: Continued from previous page

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref	Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref
aCCH	aCCH	61.325	8.0000	[36]	COO ⁻	K ⁺	23.999	CR	[42, 43]
aCCH	aCCH ₂	462.04	CR	[36]	COO ⁻	Li ⁺	2.6855	CR	[44] †
aCCH	CH	67.510	CR	[36]	COO ⁻	Rb ⁺	31.864	CR	[44] †
aCCH	COOH	599.28	CR	[36]	COO ⁻	Ca ²⁺	27.115	CR	[44] †
aCCH	aCCOaC	CR	CR	-	COO ⁻	Mg ²⁺	9.4414	CR	[44] †
aCCH	CH ₂	345.80	CR	[36]	NH ₃ ⁺	NH ₃ ⁺	48.295	10.254	[44] †
aCCH	NH ₂	CR	CR	-	NH ₃ ⁺	H ₃ O ⁺	56.951	CR	[44] †
aCCH	C	CR	CR	-	NH ₃ ⁺	OH ⁻	62.231	CR	[44] †
aCCH	COO ⁻	CR	CR	-	NH ₃ ⁺	Cl ⁻	65.249	CR	[44] †
aCCH	NH ₃ ⁺	CR	CR	-	NH ₃ ⁺	Br ⁻	61.726	CR	[44] †
aCCH	H ₃ O ⁺	CR	CR	[44] †	N ⁺	N ⁺	62.971	8.8971	[42, 43]
aCCH	OH ⁻	CR	CR	-	N ⁺	aCNH ₃ ⁺	140.33	CR	††
aCCH	Cl ⁻	CR	CR	-	N ⁺	H ₃ O ⁺	61.25	CR	††
aCCH	Na ⁺	CR	CR	-	N ⁺	OH ⁻	58.782	CR	††
aCCH	K ⁺	CR	CR	-	N ⁺	Cl ⁻	61.982	CR	[42, 43]
aCCH	Li ⁺	CR	CR	-	N ⁺	Na ⁺	32.915	CR	††
aCCH	Rb ⁺	CR	CR	-	aCNH ₃ ⁺	aCNH ₃ ⁺	397.63	17.065	††
aCCH	Ca ²⁺	CR	CR	-	aCNH ₃ ⁺	H ₃ O ⁺	155.38	CR	††
aCCH	Mg ²⁺	CR	CR	-	aCNH ₃ ⁺	Cl ⁻	189.98	CR	[44] †
aCCH ₂	aCCH ₂	591.56	8.5433	[36]	aCNH ₃ ⁺	Na ⁺	CR	CR	-
aCCH ₂	CH	65.410	CR	[36]	H ₃ O ⁺	H ₃ O ⁺	68.190	17.020	[150]
aCCH ₂	COOH	473.66	CR	[36]	H ₃ O ⁺	OH ⁻	66.439	CR	[150]
aCCH ₂	CH ₂	454.16	CR	[36]	H ₃ O ⁺	Cl ⁻	70.552	CR	[150]
aCCH ₂	NH ₂	CR	CR	-	H ₃ O ⁺	Br ⁻	65.800	CR	[150]
aCCH ₂	C	CR	CR	-	H ₃ O ⁺	Na ⁺	37.479	CR	[44] †
aCCH ₂	COO ⁻	CR	CR	-	H ₃ O ⁺	K ⁺	72.869	CR	[44] †
aCCH ₂	NH ₃ ⁺	CR	CR	-	H ₃ O ⁺	Li ⁺	12.314	CR	[44] †
aCCH ₂	H ₃ O ⁺	CR	CR	[44] †	H ₃ O ⁺	Rb ⁺	89.271	CR	[44] †
aCCH ₂	OH ⁻	CR	CR	-	H ₃ O ⁺	Ca ²⁺	104.69	CR	[44] †

Table A.2: Continued from previous page

Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref	Group k	Group l	$(\varepsilon_{kl}/k_B)/K$	λ_{kl}^r	ref
aCCH ₂	Cl ⁻	CR	CR	-	H ₃ O ⁺	Mg ²⁺	44.541	CR	[44] †
aCCH ₂	Na ⁺	CR	CR	-	OH ⁻	OH ⁻	170.24	17.020	[150]
aCCH ₂	K ⁺	CR	CR	-	OH ⁻	Cl ⁻	123.21	CR	[44] †
aCCH ₂	Li ⁺	CR	CR	-	OH ⁻	Na ⁺	27.898	CR	[150]
aCCH ₂	Rb ⁺	CR	CR	-	OH ⁻	K ⁺	54.733	CR	[150]
aCCH ₂	Ca ²⁺	CR	CR	-	OH ⁻	Li ⁺	9.1482	CR	[150]
aCCH ₂	Mg ²⁺	CR	CR	-	OH ⁻	Rb ⁺	67.944	CR	[150]
CH	CH	95.621	8.0000	[36]	OH ⁻	Ca ²⁺	77.245	CR	[150]
CH	COOH	504.99	CR	[36]	OH ⁻	Mg ²⁺	33.269	CR	[150]
CH	CH ₂	506.21	CR	[36]	Cl ⁻	Cl ⁻	113.77	12.000	[150]
CH	NH ₂	CR	CR	-	Cl ⁻	Na ⁺	27.937	CR	[150]
CH	C	2	CR	[40]	Cl ⁻	K ⁺	61.007	CR	[150]
CH	COO ⁻	504.99	CR	[44] †	Cl ⁻	Li ⁺	8.2899	CR	[150]
CH	NH ₃ ⁺	151.01	CR	[44] †	Cl ⁻	Rb ⁺	78.250	CR	[150]
CH	H ₃ O ⁺	CR	CR	-	Cl ⁻	Ca ²⁺	76.931	CR	[150]
CH	OH ⁻	CR	CR	-	Cl ⁻	Mg ²⁺	29.599	CR	[150]
CH	Cl ⁻	CR	CR	-	Br ⁻	Br ⁻	107.38	12.000	[150]
CH	Na ⁺	CR	CR	-	Br ⁻	Na ⁺	24.988	CR	[150]
CH	K ⁺	CR	CR	-	Br ⁻	K ⁺	56.589	CR	[150]
CH	Li ⁺	CR	CR	-	Br ⁻	Li ⁺	7.1798	CR	[150]
CH	Rb ⁺	CR	CR	-	Br ⁻	Rb ⁺	73.485	CR	[150]
CH	Ca ²⁺	CR	CR	-	Br ⁻	Ca ²⁺	68.642	CR	[150]
CH	Mg ²⁺	CR	CR	-	Br ⁻	Mg ²⁺	25.497	CR	[150]
COOH	COOH	405.78	8.0000	[36]	Na ⁺	Na ⁺	31.711	12.000	[150]
COOH	aCCOaC	CR	CR	-	K ⁺	K ⁺	90.097	12.000	[150]
COOH	CH ₂	413.74	CR	[36]	Li ⁺	Li ⁺	6.1039	12.000	[150]
COOH	NH ₂	CR	CR	[44] †	Rb ⁺	Rb ⁺	130.35	12.000	[150]
COOH	C	CR	CR	-	Ca ²⁺	Ca ²⁺	261.28	12.000	[150]
COOH	COO ⁻	405.78	8.0000	[42, 43]	Mg ²⁺	Mg ²⁺	89.264	12.000	[150]

Table A.3: Group association energies $\varepsilon_{kl,ab}^{\text{HB}}$ and bonding volume $K_{kl,ab}$ for use with the SAFT- γ Mie EoS in the current work. $\dagger\dagger$ indicates parameters obtained in Chapter 4, $\dagger\dagger\dagger$ interaction developed in Chapter 5 and $\dagger$ correspond to interactions developed in Chapter 3 which are also published in ref [44].

Group k	Site a on group k	Group l	Site b on group l	$(\varepsilon_{kl}^{\text{HB}}/k_{\text{B}})/\text{K}$	$K_{kl,ab}/\text{\AA}^3$	ref
H ₂ O	H	H ₂ O	e ₁	1985.4	101.69	[37, 38]
H ₂ O	H	aCH	e ₁	563.56	339.61	[36]
H ₂ O	H	aCCH	e ₁	563.56	339.61	[40]
H ₂ O	H	aCCH ₂	e ₁	563.56	339.61	[40]
H ₂ O	H	COOH	e ₁	1451.8	280.89	[39]
H ₂ O	H	COOH	e ₂	1252.6	150.98	[39]
H ₂ O	e ₁	COOH	H	2567.7	270.09	[39]
H ₂ O	e ₁	NH ₂	H	1988.3	55.824	[167]
H ₂ O	H	NH ₂	e ₁	1460.0	179.60	[167]
H ₂ O	H	N	e ₁	2783.7	15.536	[167]
H ₂ O	H	aCCOO	e ₁	1586.2	92.159	$\dagger\dagger$
H ₂ O	H	aCCOO	e ₂	1586.2	92.159	$\dagger\dagger$
H ₂ O	e ₁	aCNH ₂	H	511.27	250.00	$\dagger\dagger$
H ₂ O	H	aCNH ₂	e ₁	810.00	2496.8	$\dagger\dagger$
H ₂ O	e ₁	CH ₂ O	H	2572.7	7.3729	$\dagger\dagger\dagger$
H ₂ O	H	CH ₂ O	e ₁	2572.7	7.3729	$\dagger\dagger\dagger$
H ₂ O	e ₁	CH ₂ O	e ₂	2445.0	3147.1	$\dagger\dagger\dagger$
H ₂ O	e ₁	CH ₃ OH	H	1993.5	104.11	[36]
H ₂ O	H	CH ₃ OH	e ₁	1993.5	104.11	[36]
H ₂ O	H	COO ⁻	e ₁	802.21	52.555	[42, 43]
H ₂ O	e ₁	NH ₃ ⁺	H	2016.6	49.397	[44] $\dagger$
H ₂ O	e ₁	N ⁺	H	2783.7	15.536	[42, 43]
H ₂ O	e ₁	aCNH ₃ ⁺	H	2550.8	501.61	$\dagger\dagger$
H ₂ O	e ₁	H ₃ O ⁺	H	1985.4	101.69	[150]
H ₂ O	H	OH ⁻	e ₁	1492.0	76.411	[150]
aCH	e ₁	H ₃ O ⁺	H	563.56	339.61	[44] $\dagger$

Table A.3: Continued from previous page

Group k	Site a on group k	Group l	Site b on group l	$(\varepsilon_{kl}^{\text{HB}}/k_{\text{B}})/\text{K}$	$K_{kl,ab}/\text{\AA}^3$	ref
aCCH	e ₁	H ₃ O ⁺	H	563.56	339.61	[44] †
aCCH ₂	e ₁	H ₃ O ⁺	H	563.56	339.61	[44] †
COOH	H	COOH	H	6427.9	0.80620	[36]
COOH	H	NH ₂	e ₁	2623.5	20.765	[44] †
COOH	e ₁	H ₃ O ⁺	H	1451.8	280.89	[44] †
COOH	e ₂	H ₃ O ⁺	H	1252.6	150.98	[44] †
COOH	H	OH ⁻	e ₁	2036.0	214.16	[44] †
NH ₂	e ₁	NH ₂	H	1070.8	95.225	[167]
NH ₂	e ₁	H ₃ O ⁺	H	1460.0	179.60	[44] †
NH ₂	H	OH ⁻	e ₁	1511.4	42.436	[44] †
N	e ₁	H ₃ O ⁺	H	2783.7	15.536	††
aCCOO	e ₁	aCNH ₂	H	1095.6	372.29	††
aCCOO	e ₂	aCNH ₂	H	1095.6	372.29	††
aCCOO	e ₁	aCNH ₃ ⁺	H	1095.6	372.29	††
aCCOO	e ₂	aCNH ₃ ⁺	H	1095.6	372.29	††
aCCOO	e ₁	H ₃ O ⁺	H	1586.2	92.159	††
aCCOO	e ₂	H ₃ O ⁺	H	1586.2	92.159	††
aCNH ₂	H	aCNH ₂	e ₁	700.00	241.64	††
aCNH ₂	e ₁	aCNH ₃ ⁺	H	700.00	241.64	††
aCNH ₂	e ₁	H ₃ O ⁺	H	810.00	2496.8	††
aCNH ₂	H	OH ⁻	e ₁	396.86	194.06	††
CH ₂ O	H	CH ₂ O	e ₁	500.00	1381.9	† † †
CH ₂ O	H	CH ₃ OH	e ₁	4594.9	0.0010000	† † †
CH ₂ O	e ₁	CH ₃ OH	H	4594.9	0.0010000	† † †
CH ₂ O	e ₂	CH ₃ OH	e ₁	5562.3	0.91815	† † †
CH ₃ OH	H	CH ₃ OH	e ₁	2062.1	106.57	[36]
COO ⁻	e ₁	H ₃ O ⁺	H	802.21	52.555	[44] †
NH ₃ ⁺	H	OH ⁻	e ₁	1988.3	55.824	[44] †
N ⁺	H	OH ⁻	e ₁	2098.9	11.714	††

Table A.3: Continued from previous page

Group k	Site a on group k	Group l	Site b on group l	$(\varepsilon_{kl}^{\text{HB}}/k_{\text{B}})/\text{K}$	$K_{kl,ab}/\text{\AA}^3$	ref
H_3O^+	H	OH^-	e_1	1492.0	76.411	[150]