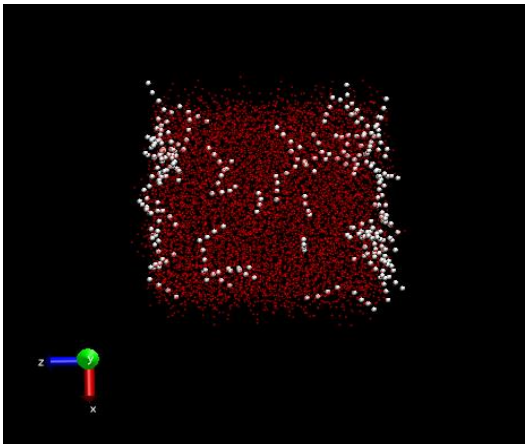
TÉCNICO LISBOA



Self-assembly of Perfluoroalkylalkanes

A Molecular Dynamics Simulation Study

Luís Maria Fabião de Campos Belchior

Thesis to obtain the Master of Science Degree in

Chemical Engineering

Supervisors:

Prof. Doctor George Jackson

Prof. Doctor Eduardo Jorge Morilla Filipe

Examination Committee

Chairperson: Prof. Doctor Francisco Manuel da Silva Lemos

Supervisor: Prof. Doctor Eduardo Jorge Morilla Filipe

Member of the Committee: Prof. Doctor José Nuno Aguiar Canongia Lopes

October of 2017

ABSTRACT

In this work, Molecular Dynamics (MD) simulations were used to study the interfacial behaviour and the self-assembly of mixtures of Perfluoroalkylalkanes (PFAAs) in hydro- and fluorocarbon solvents. PFAAs are linear di-block copolymers formed from an alkane and a perfluoroalkane chains bonded together, with the general formula $F(CF_2)_n(CH_2)_mH$ (shortened for F_nH_m). The distinct amphiphilic character imparted by these two mutually phobic chains (importantly named *primitive surfactants*) is known to contribute to promote supramolecular organization.

A recently developed heteronuclear *coarse-grained* (CG) force field of the SAFT-γ Mie family was refined. A new set of fluorinated beads was defined and the intramolecular terms of the force field determined via Direct Boltzmann Inversion from United-Atom simulations. Additionally, the SAFT-γ Mie Equation of State was used to obtain the intermolecular Mie potentials between the novel CG groups from thermodynamic data.

The theory accurately described bulk properties, as well as critical points, of perfluoroalkanes and PFAAs, even for compounds not included in the modelling procedure. Furthermore, supramolecular organization in the referred mixtures was detected, with the model – despite its simplicity – capturing the subtleties arising from the interactions between PFAAs and the solvent. In so doing, this work can be used to further refine the industrial practice of fluorinated compounds and provide the means to advance the understanding of aggregation of this important surfactant family.

Keywords:

Perfluoroalkylalkanes – Coarse-Grained – Molecular Dynamics – SAFT-γ Mie

RESUMO

Neste trabalho, simulações por dinâmica molecular foram conduzidas para estudar propriedades interfaciais e a organização de misturas de perfluoroalquilalcanos (PFAAs) em diversos solventes (alcanos e perfluoroalcanos). PFAAs são copolímeros lineares constituídos por duas cadeias (alcano e perfluoroalcano) ligadas, de fórmula genérica $F(CF_2)_n(CH_2)_mH$ (abreviada para F_nH_m). O acentuado carácter anfipático que a presença de duas cadeias que pouco interagem impõe tem contribuído para uma série de fenómenos de organização supramolecular.

Aperfeiçoou-se um *force field* heteronuclear *coarse-grained* (CG) da família SAFT-γ Mie publicado recentemente. Propôs-se um novo conjunto de grupos fluorados, com os respectivos parâmetros intramoleculares a serem determinados por inversão de Boltzmann a partir de simulações *United-Atom*. Já a equação de estado SAFT-γ Mie foi utilizada para obter os potenciais intermoleculares Mie entre os novos grupos CG a partir de propriedades termodinâmicas.

A teoria descreveu rigorosamente as diversas propriedades termodinâmicas de perfluoroalcanos e PFAAs consideradas, assim como pontos críticos, mesmo para compostos não considerados durante a modelação. Em seguida, detectou-se organização supramolecular nas misturas em causa, com o modelo, embora pouco detalhado, a conseguir capturar as subtis interacções entre PFAAs e o solvente. Nesse sentido, o modelo CG proposto pode não só afirmar-se, pelo seu potencial preditivo, como uma mais-valia na indústria de fluorados, como também contribuir para aprofundar o conhecimento sobre organização nesta família de *detergentes primitivos*.

Palavras-chave:

Perfluoroalquilalcanos – *Coarse-Grained* – Dinâmica Molecular – SAFT-γ Mie

“Every day you may make progress. Every step may be fruitful. Yet there will stretch out before you an ever-lengthening, ever-ascending, ever-improving path. You know you will never get to the end of the journey. But this, so far from discouraging, only adds to the joy and glory of the climb.”

Sir Winston Churchill

Contents

Table of Figures

Acknowledgements

This work would have not been possible without the help of many people who I hereby want to acknowledge.

Foremost, I want to thank my supervisors: Professor Eduardo Filipe, for this wonderful opportunity and, broadly, all his friendship and trust throughout the years; and Professor George Jackson, for accepting me and giving me the chance to work with his most talented group and for his tireless support and care. I am truly indebted to both.

To all my friends and colleagues at Imperial College, I want to express my gratitude for the caring and supportive environment, the way I was welcomed and treated, and the stimulating, challenging atmosphere that defines this group. It was a privilege to work side by side with such bright people. I also want to thank Andrew Haslam for our statistical mechanics discussions and for the incredible work he does in this group.

Also, I ought to thank to Eng. Gonçalo Silva and Doctor Pedro Morgado, for the preparation and fruitful discussions throughout this thesis, without which my work would have certainly been much harder.

Because this thesis was, overall, much more than the specific work I have done, I owe a special word to my flatmates and dear friends Susana, Joana and Ricardo. Their genuine sympathy, care and sense of humour put these six months among the finest of my life.

I also want to acknowledge the generous scholarship I received from the European Union.

Finally, I wish to thank to my parents Helena and Paulo and my sister Maria do Mar, for having created the conditions for me to succeed in every possible dimension in London and, more broadly, in every moment of my life.

I - Introduction

1. Properties of Perfluoroalkanes and Perfluoroalkylalkanes

Perfluoroalkanes (or perfluorocarbons, shortened from C_nF_{2n+2} to "F_n") are fully saturated carbon chains solely comprising carbon and fluorine – they resemble a normal alkane in which all the hydrogen atoms were swapped to fluorine atoms (see Figure 1, left). The first synthesis of tetrafluoromethane, the shortest perfluoroalkane (PFA) dates back to the late 19th Century, but only after the Second World War was the production of perfluorocompounds truly accelerated [1].

Figure 1 - Sketched representation of perfluoroalkanes (left) and perfluoroalkylalkanes (right).

Perfluoroalkanes are chemically inert (as in non-reactive) compounds [2]. This stability supports the wide array of applications in which these compounds have been used: lubricants, refrigerants, solvents, surfactants, anticorrosive and antifriction components, flame retardants, among many others [3, 4]. In addition to the referred chemical inertia, perfluoroalkanes are biologically harmless and can dissolve large amounts of gases like oxygen and carbon dioxide: they can thus be medically used as artificial blood, in eye surgery and in the treatment of burns [5].

From the molecular point of view, the forces between PFA are solely of dispersion. The surface tension of fluorocarbons is about half of that of hydrocarbons, indicating weaker *intermolecular* interactions – in contrast with the strong *intramolecular* interactions that the strong bond C–F induces. In addition, the boiling point of fluorocarbons containing more than five carbons is lower than that of the equivalent alkane [6].

Despite the apparent similarities between hydrocarbons and fluorocarbons (both exhibit only dispersion forces, both share a carbon chain, etc.), it is well-known that binary mixtures of both are highly non-ideal, with tremendous positive excess properties (as enthalpy and volume) and large regions of liquid-liquid immiscibility, clearly demonstrating that the unlike interactions are much weaker than the like ones [7, 8]. In addition, not only the fluorinated molecules are much thicker than their hydrogenated counterparts, thanks to the larger size of the fluorine atom, but also their backbone is far stiffer, as will be demonstrated in the following chapters (which grants PFAs a so-called "pencil-like" shape).

The strange behaviour of mixtures of alkanes and perfluoroalkanes motivates the study of another, more intriguing, family of compounds: perfluoroalkylalkanes (PFAAs) or, more loosely, semifluorinated[1] alkanes (SFAs). As shown in Figure 1 (right), PFAAs consist of two chemically bonded chains – one is a fluorocarbon and the other is a hydrocarbon. PFAAs (of general formula $F(CF_2)_n(CH_2)_mH$, shortened for F_nH_m) have been interpreted as "chemical mixtures of two mutually phobic segments" [7] that, in the absence of the chemical bond attaching them, would segregate in two

[1] The prefix "semi-" does not necessarily mean that the molecule is divided into two carbon chains of equal length.

phases. For that reason, they have been termed *primitive surfactants* [9], inasmuch as they not only contain the mutually phobic tails but also lack a permanent molecular dipole – a system of surfactants was built, after all, without the usual hydrophobic/hydrophilic opposition [10]. Nevertheless, the fluorocarbon chain is even more hydrophobic than the hydrocarbon one, and tends to stay at the surface when a PFAA is added to a water/air interface [11]. Figure 2 purveys a more graphic representation of PFAAs.

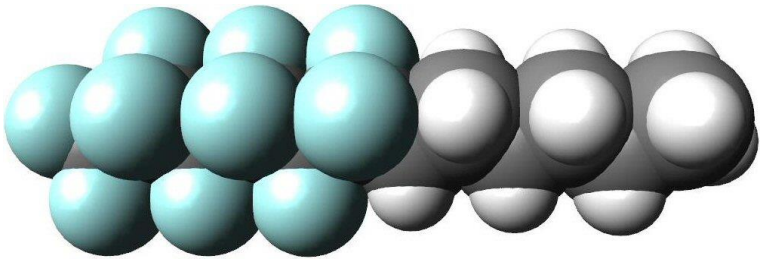


Figure 2 - Representation of a perfluoroalkylalkane (PFAA). On the left, the fluorinated tail stands out as much thicker than its hydrogenated counterpart (on the right). From [12].

This extraordinary feature inspires interesting questions: what happens when a PFAA molecule is introduced in a perfluoroalkane or in an alkane solvent? How do the properties of PFAAs relate to the properties of both alkanes and perfluoroalkanes? How do the properties of PFAAs depend on parameters such as: the overall chain length; the length of both hydrogenated and fluorinated chains; the relative proportion of H/F atoms; the solvent in which they are solubilized (if any)? The intriguing behaviour of such molecules would make this an endless list.

Unlike in common surfactants, in which hydrogen or ionic bonds may stand out as main drivers of self-assembly phenomena, in PFAAs, the exclusive presence of dispersion forces, allied to the mutual phobicity between the two blocks, creates a much more subtle landscape. It is this delicate, sophisticated setting that makes the study of these molecules such a fascinating topic.

Interesting interfacial phenomena such as surface freezing (the creation of a quasi-solid monolayer above the melting point [13]), formation of Langmuir monolayers at interfaces such as water/air and water/hydrocarbon [11] and organization into liquid crystals of pure PFAAs [14] have been proven. It becomes evident that the behaviour of PFAAs near interfaces (even in the absence of any other compound) is not trivial at all, and to be able to predict it with simulation techniques would be beyond question an important achievement.

This work is dedicated to the interfacial behaviour of PFAAs and their organization in different solvents (hydro- and fluorocarbons). In particular, it focuses on how what promotes this self-assembly how it depends on several factors such as: temperature, composition and the fluorine content of PFAAs. But since the present study was accomplished using molecular dynamics simulations, coarse-graining procedures and SAFT as equation of state, a short overview of the main concepts is presented hereby.

2. Molecular Simulation

As an alternative or complimentary technique to the traditional laboratory experiments, molecular simulation is becoming rapidly a major tool for scientific research. The possibilities offered by

molecular simulation extend far beyond the reach of physical experiments, naturally constrained by very high (or very low) temperatures and pressures; furthermore, they provide a unique opportunity to study a wide range of phenomena that would otherwise be tremendously time and resources-consuming. Besides this ability to cope with extreme conditions, computer simulations may be used to observe details in the motion and structure of molecules (e.g.: catalysis, ion conduction, enzymatic reactions) known to be hard to probe experimentally [15].

This ability to study a system at the "*nano*" scale, providing impressive detail on what happens at the very interactions between atoms and molecules, has been pushing boundaries. With the continuous evolution of computational throughput [16], molecular simulation is becoming able to study larger and more complex systems every day. A crucial application of this technique is biochemistry, for its obvious complexity and for the size of the (macro)molecules it investigates, such as proteins, lipids or even DNA. Molecular simulation can also be used to pre-screen compounds for a given purpose (e.g. pharmaceuticals) or to choose the best chemical to be used as a solvent in a reaction [17].

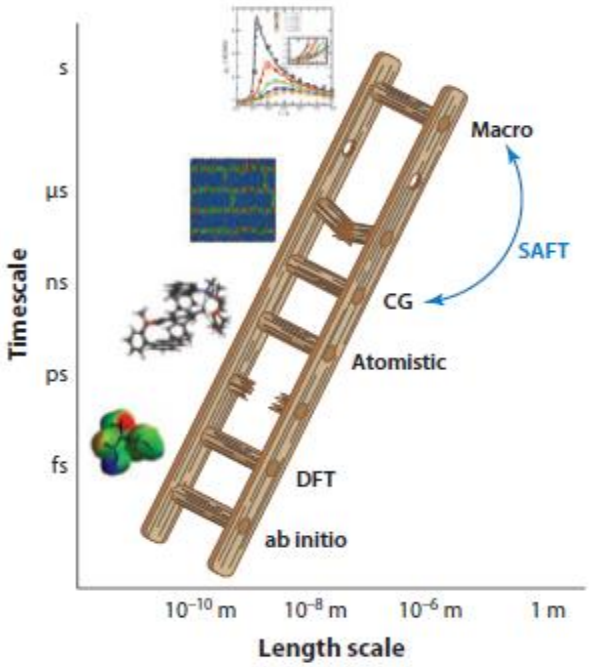


Figure 3 – The computer simulation ladder, in which the different levels of simulation are sketched as a function of the time and length scales they manage to cope with. From [18].

The following discussion is based on the enlightening review of Jackson and Müller [18], to which the interested reader is referred for more information. The simulation ladder (Figure 3) notably sketches the different options presently available for chemical systems simulation. One should always bear in mind the aim of the current research and choose accordingly the best-suited method – this research, in particular, is focused in the phase equilibria of pure compounds and their mixtures and, on an upper scale, in the supramolecular organization of perfluoroalkylalkanes in hydro- or fluorocarbon solvents. Computational calculations based on quantum mechanics (*first principles*) lie on the bottom of the ladder, followed by Density Functional Theory (DFT). Though highly accurate, given the few simplifications made from the quantum mechanical description of matter, both are useless to study

molecular systems: *ab initio*[2] calculations can study just a few atoms and DFT is limited by hundreds of them. This level of detail is virtually impossible to work with if one wants to study systems with a few hundreds or thousands of molecules.

Atomistic and coarse-grained models stay above in the ladder. The broken rung between them and the previous methods is supposed to represent the fact that one does not get to atomistic simulations by somehow simplifying furthermore DFT methods. Those are completely different approaches, whose focus lies on the atomistic scale. In atomistic simulations, two types can be discerned: all-atom (AA) and united-atom (UA). Whilst the first treat explicitly every single atom, the second looks only to "heavy" atoms (all but hydrogen). Even this subtle simplification of moving from AA to UA models can grant substantial improvements in terms of simulation time: for hexane, for instance, an AA simulation would perform calculations for 20 atoms, while a UA would deal only with 6 "atoms" (four CH_2 plus two CH_3). Coarse-graining pushes the boundaries of atomistic simulations by grouping more atoms into one "superatom" or site. The referred hexane molecule could be, as an example, treated as two connected propyl groups. Naturally, this "zoom out" greatly reduces the computation time at expenses of less detail and a less sophisticated simulation. Clearly, the trade-off that needs to be recognized is the one between detail, on one hand, and resources consumed, on the other. The chemical systems and processes of interest determine the best approach.

Nevertheless, increasing the size of the coarse-grained groups indefinitely is no solution at all – and understanding chemical processes that surpass the timescale of the nanosecond is still a challenge, particularly relevant for biochemists, polymer scientists and other researchers. Another challenge arises, thus, from the lack of connection between the coarse-graining world and the macroscopic properties experimentally measured.

The Statistical Fluid Associating Theory (SAFT), presented later in this Introduction, is one key to overcome this problem and provide a reliable pathway between macroscopic properties and molecular parameters. The aforementioned review [18] deepens the analysis of this challenges and the solutions that have been thus far presented to overcome them.

2.1 Theory, Simulation and Experiment

Before molecular simulation started, theories could only be tested with experimental data. This comparison, however, raised concerns because differences between the prediction and the experiment could be ascribed to both the approximations in the theory and the model (the intermolecular potential description) in which it was based, let alone the ubiquitous errors associated with any experimental measurement. Likewise, if theories required fitting parameters from experimental data, apparent accuracy could hide true flaws of the theory [19].

When computer simulations became available, the implementation of a model became possible – it is impossible to make experiments with hard spheres in a laboratory, but it is straightforward to implement them in a simulation. Thereby, the approximations in a theory could be directly evaluated: if

[2] (Latin) Literally "from first principles" – from known equations without any simplifications or approximations.

the prediction matched the simulation result, the theory was sound (that is, the eventual approximations made while developing the theory were valid). Furthermore, molecular simulations can be confronted with experimental data to assess the model. For instance, two similar simulations grounded on two different models would yield diverse results that could be compared to experimental data: the model getting closer to this data could be taken as more accurate [15] (see Figure 4).

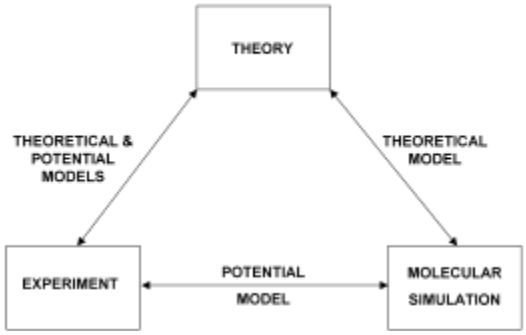


Figure 4 - The theory-simulation-experiment triangle, with the role of molecular simulation in testing both theories and models. From [19].

Finally, molecular simulations can provide an "exact" result (in other words, as exact as one wants it to be) which can be then compared with experimental data. Granted that a realistic model is used (not the case of hard spheres, naturally), simulation results can supplement experimental data in accuracy, not to mention in effort and resources [19].

2.2 Molecular Dynamics

In molecular simulation of chemical systems, there are fundamentally two different, widely spread techniques: Molecular Dynamics (MD), used in this work, and Monte Carlo (MC). The reader is referred to [16, 20], the last of which is followed in this section.

Quite similar to laboratory experiments, in MD simulations one prepares a box with N particles (usually hundreds or thousands of particles) and then applies Newton's Laws of Motion to compute the individual trajectory of each particle. Of paramount importance is the parameterised potential U_i, for its derivative is used to compute the resultant force (F_{x_i}) in each direction (Equation 1.1), which in turn is used in the numerical integration of the position $x_i(t)$ – (Equation 1.2). For the sake of simplicity, the equations are presented for one dimension, but one can easily extend them to a manifold.

$$\frac{dU_i}{dx} = F_{x_i} \quad (1.1) \qquad \frac{d^2x_i(t)}{dt^2} = a_i(t) = \frac{F_{x_i}(t)}{m_i} \quad (1.2)$$

Though the algorithms involved in MD effectively compute the trajectory of each single particle during the simulation time, what ends up being most valuable is the statistical averages collected after a sufficiently long time. In light of the ergodic hypothesis (explained above), the simulated sub-ensemble average equals, after a sufficiently long time, the ensemble average (the macroscopic property measured in a macroscopic experiment).

The integration of the equations of motion is solved using a finite difference method [16]. The Verlet integration[3] method [21], given its wide application, is described as an example. For a time step δt, the position of the particle at time $t + \delta t$ is given by the set of equations (Equations 2, 3 and 4).

$$v_i\left(t + \frac{1}{2}\delta t\right) = v_i(t) + \frac{1}{2}\delta t a_i(t) \quad (2)$$

$$r_i(t + \delta t) = r_i(t) + \delta t v_i\left(t + \frac{1}{2}\delta t\right) \quad (3)$$

$$v_i(t + \delta t) = v_i\left(t + \frac{1}{2}\delta t\right) + \frac{1}{2}\delta t\, a_i(t + \delta t) \quad (4)$$

Importantly, the Verlet algorithm is *exactly time reversible*, that is, if applied as it is, allows the calculation of a "backwards trajectory"; additionally, it is compatible with long time steps and quite easy to program [22].

The overarching capacity of MD simulations to study molecular systems in a reasonable time span is partly due to the *periodic boundary conditions* (PBC). The reader is asked to imagine a box inside which are inserted some hundreds or thousands of molecules. Since the box dimensions have the same length scale than the molecular phenomena to be studied (nanometres), the wall effects are much more relevant than in a normal laboratory experiment (where they are virtually inconsequential). Thus, if the wall effects were to be considered (and mitigated), one would need a much larger box in order to have a proper "bulk", insensible to the wall influence. Periodic boundary conditions assure that when a molecule crosses one of the boundaries, an exact copy ("image") of that same molecule appears on the opposite side entering the box [15]. Fundamentally, the bulk phase is extended to an infinite lattice of small boxes, allowing a modest simulation of a few hundreds or thousands of molecules to provide insight into thermodynamic (so called "bulk") properties such as density, heat of vaporization, vapour pressure, among others.

The normal procedure in a MD simulation comprises essentially four steps [16]:

1) A box with the desired number of molecules is generated;

2) An energy minimization algorithm is run to "shake" the system: this is an essential step since the random insertion of the particles inevitably leads to the overlapping of some molecules, which drives the potential energy to very high values (due to the repulsive part of the potential at short distances).

3) An equilibration run follows, in which the molecular dynamics algorithm is used, to enable the system to reach equilibrium from the starting configuration.

4) With the system equilibrated, a final run with a sufficiently long simulation time is used to collect data for the properties of interest. In theory, the longer the simulation time, the less relevant are the statistical errors.

[3] Actually, there are several, equivalent versions of the Verlet algorithm. The one described is the "velocity Verlet".

A MD simulation is normally run in the *NPT* or the *NVT* ensembles, that is, either with fixed number of molecules, pressure and temperature, or with fixed number of molecules, volume and temperature. While the *NPT* is better suited for saturated density calculations (for example), the *NVT* allows the formation of an interface between a vapour and a liquid phase (clearly depending on the conditions) and, by means of that, the calculation of bulk properties such as vapour pressures and interfacial properties such as surface tension, among others. To assure that both temperature and pressure are constant when required, a thermostat and a barostat are used, respectively. The former permanently controls the temperature by correcting the kinetic energy of the particles; the latter regulates the pressure by varying the volume of the box [20].

3. Coarse-graining

Despite the witnessed achievements in computational power, the need to access longer time and length scale phenomena than those offered by the current capabilities of atomistic simulations led to the development of *coarse-grained* (CG) simulations [23]. *Coarse-graining* means grouping several atoms in *sites* (or *superatoms*) [24]. An effective coarse-graining truly simplifies the representation and study of soft matter and biomolecules [25], primarily because the number of particles is reduced. Moreover, since the energy landscape is smoothed (as atomistic details are integrated out), an acceleration of molecular dynamics is achieved [23]. In fact, in MD simulations, the evaluation of the force applied in every single particle is the most demanding part of a molecular dynamics simulation, so there is a clear rationale for reducing the number of particles.

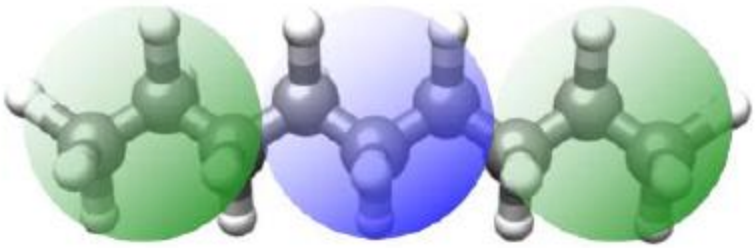


Figure 5 – Example of a coarse-grained model of n*-nonane in three beads (each one comprising three carbon atoms). From [26] (p. 32). Notice how a 29-atom molecule can be described with a "3-superatom" CG molecule.*

In fact, coarse-graining is effectively *integrating out* some details that, if the CG model proves accurate, are not critical for the overall description of the system. The correspondence between atomistic sites and CG sites is called *mapping*; importantly, different mappings may yield similar results – as such, the mapping is not the only decisive factor imparting the accuracy of the novel CG model, and the assumption that a more detailed mapping earns more exact outcomes is not trivial.

The process of mapping is inevitably a trade-off between, on the one hand, computational efficiency derived from the reduction of interacting sites and, on the other hand, the necessary chemical representability to accurately reproduce the phenomena of interest [27]. Marrink *et al.* suggested [28] a four-to-one ratio[4], and although that is usually a "rule of thumb", the choice of the sites depends intrinsically on the phenomena intended to be studied: for instance, a *coarser* coarse-graining may be

[4] For heavy atoms, that is, excluding hydrogen.

well-suited if the target is some sort of qualitative observation (such as aggregation), but may prove unreliable if the aim is the quantitative prediction of some thermodynamic properties [23]. Ideally, the mapping would incorporate one or more functional groups but, as Lobanova judiciously explains [29], this very mapping is guided in practice by chemical intuition, and is definitely not trial-and-error-free.

It is straightforward by now that two important questions arise in coarse-graining: (i) what is the methodology, if any, to choose the new *superatoms*; and (ii) how are defined the interactions between them – and there are no linear answers for both.

A CG model should not be judged *a priori*, since its quality must be inferred by three different attributes: **representability** – the ability to predict properties at different state points than those used in the modelling procedure; **robustness** – the force field should provide reliable predictions for various (structural, thermodynamic or transport) properties at conditions not used in the parameter estimation; and **transferability** – the ability to use a modelled group - for instance, a CH_2 – in different molecules, such as short and long alkanes). In particular, the success of any group contribution (GC) method relies heavily on good transferability [18]. In consequence, a sound CG model is accurate across a wide array of states (representable), properties (robust) and molecules (transferable).

G. A. Voth [23] formalizes (through Equation 5) the connection between atomistic and CG models through statistical mechanics ("(…) *a venerable and remarkable theoretical framework (…)*"):

$$\exp\left(-\frac{A}{k_B T}\right) = (const.)\int dx\ \exp\left[-\frac{V(x)}{k_B T}\right] \approx (const.')\int dx_{CG}\ \exp\left[-\frac{V_{CG}(x_{CG})}{k_B T}\right] \quad (5)$$

where A is the Helmholtz free energy, x and V are, respectively, the coordinates and potential of all atoms, k_B is the Boltzmann constant, T is the temperature; and x_{CG} and V_{CG} are, respectively, the coordinates and potential of the CG sites (this last one often called *potential of mean force*). Naturally, there are fewer CG than atomistic coordinates. Equation 5 is rarely solved directly; instead, the integral is evaluated using MD or MC techniques, which allow the computation of the Helmholtz free energy and, in consequence, of the thermodynamic variables that can be thereby derived.

The potential of mean force (PMF) is not a usual potential energy function, but must rather be interpreted as a "configuration-dependent free energy function", reflecting both energetic and entropic contributions [24]. This quantity is related to the mapping through Equations 6 and 7:

$$V_{CG}(x_{CG}) = -k_B T\ \ln Z(x_{CG}) \quad (6)$$

where $Z(x_{CG})$ is given by Equation 7:

$$Z(x_{CG}) = \int dx\ \exp\left[-\frac{V(x)}{k_B T}\right]\delta(M(x) - x_{CG}) \quad (7)$$

The last equation deserves a more comprehensive explanation. $M(x)$ is the *mapping function* whose inputs are the atomistic space coordinates and whose outputs are the CG space coordinates. The Dirac delta function argument is null for the positions of the atomistic configuration space that are mapped in one specific CG site; therefore, the integrand function is not zero only on those specific

coordinates. Thus, the PMF weights each CG configuration (x_{CG}) according to the atomistic Boltzmann factor and can then be used to sample a canonical distribution of CG configurations that is equivalent to that implied in the initial atomistic model [24].

Not surprisingly, the coarse-graining procedure inevitably leads to a loss of entropy (as the number of particles is reduced and some atomistic details are integrated out). To compensate for this, CG potentials have weaker attractions and repulsions, which is translated in an inherent inability to capture the true physics of the system. As a consequence, both bonded and non-bonded potentials are, strictly speaking, *effective* potentials [30], taking into account both entropic and enthalpic effects [24].

3.1 Coarse-graining methodologies

Numerous paths can be followed to obtain a CG potential; briefly, and attending to the applications shown in this thesis, this section will focus on two main procedures: *bottom-up* and *top-down*. In the bottom-up approaches, the interactions between CG sites are developed from more detailed (e.g. atomistic) models; top-down approaches, in turn, are developed from macroscopic properties.

Both methodologies share an important feature: facing the inherent complexity and intractability of the PMF, they represent it as a summation of interactions, each one dependent on a single mechanical degree of freedom: bond stretching, angle bending, torsion are described, respectively, as functions of bond lengths, bond angles, dihedral angles; van der Waals interactions are translated via an intermolecular potential, primarily a function of the distance between two atoms; and Coulomb interactions are modelled (when applied) as a function of the electrical charges in the different atoms as well as their distance. This may be represented as in Equation 8:

$$U(\vec{R}) = \sum_{\omega} \sum_{\delta} U_{\omega}(\psi_{\omega}(\vec{R}_{\delta})) \quad (8)$$

where ω characterizes a certain interaction modelled by a potential U_{ω}, which in turn is a composed function of the positions of a specific set of positions $\vec{R}_{\delta}$ through the scalar function ψ_{ω}. The chapter on force fields will develop on the different contributions to the bonded potential.

3.1.1 Bottom-up

Bottom-up strategies start from quantum-mechanical or atomistic-level models and integrate out some of the degrees of freedom of the system [18] (ideally, a bottom-up model could even be developed from fundamental first principles) [24]. All bottom-up approaches, however, rely equally upon information from a more detailed model which, in the end, brings one back to quantum-mechanical calculations as the “most fundamental” starting point.

The treatment of Brini *et al.* [25] is followed. By coarse-graining, one is indirectly distinguishing between “slow” and “fast” degrees of freedom: while the CG model still retains the “slow” degrees of freedom (designated by $\vec{R}$), the fast ones ($\vec{r}$) are integrated out. The partition functions of the *fine-grained* (Q_{FG}) and *coarse-grained* models (Q_{CG}) are formalized in Equations 9 and 10:

$$Q_{FG} = c \int d\vec{R} \int d\vec{r}\ \exp(-\beta U^{AA}(\vec{r}, \vec{R})) \quad (9)$$

where c accounts for the kinetic contribution of the Hamiltonian and U^{AA} is the atomistic potential;

$$Q_{CG} = \int d\vec{R}\ \exp(-\beta U^{CG}(\vec{R})) \quad (10)$$

where U^{CG} is the effective coarse-grained potential (already called *potential of mean force*). Again, this effective potential is not a purely energetic measure, but a free energy that also accounts for the entropic loss that coarse-graining encompasses.

To assure the procedure is thermodynamically consistent, both partition functions must be equivalent, thus rendering the following equality (Equation 11):

$$U^{CG}(\vec{R}) = -\, k_B T \ln\left[c \int d\vec{r}\ \exp(-\beta U^{AA}(\vec{r}, \vec{R}))\right] \quad (11)$$

Equation 11 demonstrates how this PMF depends not only on the temperature but also on the volume (through the integration over the configuration space). Noteworthy is the fact that the efficiency and success of the coarse-graining procedure deeply relies on the decoupling of the interactions related with the "fast" and the "slow" degrees of freedom [31].

Unfortunately, computing this integral is prohibitively demanding as it is highly multidimensional; *per si*, this expression wouldn't be much helpful in evaluating the interactions between coarse-grained sites (or *beads*). The aforementioned *pairwise additivity* is essential to keep the simulation computationally inexpensive.

As already discussed, the potential can be safely (usually!) assumed to be the summation of two independent contributions: the bonded and non-bonded potentials. Addressing how bottom-up approaches can provide a pathway to both potentials is what the next section is all about.

Several methods have been proposed so far to address the practical problems of evaluating the potential of mean force. This work will focus on those most related with the developed research (the interested reader is referred to [25, 32, 33] for more detail). Briefly, bottom-up approaches can be divided into structure-based or force-based. The latter are designed to match the forces distributions on the beads in the CG models with the Boltzmann average of the forces in the atomistic sites averaged over the CG sites – simply put, matching the fine-grained and coarse-grained force distributions. The former were conceived to reproduce the fluid structure (in the form of the radial distribution function, RDF). A precious contribution to structure-based methods is the Henderson Uniqueness Theorem, which states that "*there is only one pair potential able to exactly reproduce a given RDF*" [34].

The simplest structure-based method is the *Direct Boltzmann Inversion* (DBI). Deemed the most straightforward approach, it inverts, in one single step, the atomistic radial distribution function obtained through atomistic simulations and the CG mapping (stated below as $g_{AA}(r)$) and computes a coarse-grained potential [35], as stated in Equation 12.

$$U^{CG}(r) = -k_B T \ln[g_{AA}(r)] \quad (12)$$

This equation, however, demands a cautious approach. While the atomistic RDF takes into account the positions of all atoms, the subsequent *pair potential* is just a function of the distance between two (CG) particles. This simplification is partially compensated by using the DBI to compute each of the potentials associated with the different degrees of freedom introduced in Equation 8 – the bond stretching, the angle bending and the torsions. By so doing, each term of the CG potential is given by Equation 13 (following the notation presented in Equation 8):

$$U_\omega(\vec{R}_\delta) = -k_B T \ln\left[p_\omega(\vec{R}_\delta)/J_\omega(\vec{R}_\delta)\right] \quad (13)$$

where $J_\omega(\vec{R}_\delta)$ is the Jacobian of the coordinates transformation and $p_\omega(\vec{R}_\delta)$ the probability of occurrence of a certain configuration (be it a bond length, an angle or a dihedral). Though arguably simplistic, DBI has provided accurate results for biomolecules coarse-graining and is particularly useful for bonded potentials for stiff CG bonds [25]. As already said (and this issue will be stressed later on), this is the case of the fluorocarbon chains.

Along with DBI, widely spread is the *Iterative Boltzmann Inversion* (IBI), which presents an algorithm to overcome the impossibility to obtain a (multi-body) PMF through a pair potential as explained before. An iterative procedure (translated in Equations 14.1 and 14.2) is followed to obtain the intermolecular CG potential:

$$\begin{cases} u_{CG}^0 = -k_B T \ln[g_{AA}(r)] & (14.1) \\ u_{CG}^{i+1} = u_{CG}^i + k_B T \ln\left[\dfrac{g_i(r)}{g_{AA}(r)}\right] & (14.2) \end{cases}$$

where the superscript *i* denotes the iteration number [15]. The atomistic RDF is used to compute the initial guess and, given Henderson Uniqueness Theorem, universal convergence is assured independently of u_{CG}^0. In addition, because there are in practice many pair potentials able to yield an RDF similar to the target, thermodynamic properties can be used to refine the iterative scheme [25]. The IBI method has managed to develop accurate and successful models for complex liquids and polymers [24].

Overall, bottom-up models that rely purely on atomistic simulations (as a more detailed model) can only aspire to have, at best, the same accuracy of the atomistic models used in the first place (though most likely they perform worse). Furthermore, since the RDF is the only input, the model will hardly be robust; and because the RDF changes with temperature, it may also lack representability. These concerns motivate the mixed approach of bottom-up methods with some macroscopic properties input, as a way of building a structurally coherent model that accurately predicts experimental data.

3.1.2 Top-Down

Top-down approaches are based on macroscopic properties and the models thereby derived lay their focus on reproducing thermodynamic properties [36]. Ironically, a top-down model could be considered a “fine-grained” [24] as it is built on even *coarser* observables. Unlike a bottom-up model, a top-down one *cannot* be said to purely predict properties, for it intrinsically requires experimental data

to be created in the first place. Rather important is the fact that the interactions modelled in a top-down approach are *truly effective*, since they are fitted to represent a given property, and not built upon a detailed, more rigorous knowledge of molecular structure.

Traditionally, top-down methodologies were associated with a heavy computational burden because the intermolecular parameters had to be manually, "empirically" optimized [37, 38]. Bulk thermodynamic properties are usual targets, namely saturation liquid density, heat of vaporisation and vapour pressures. Notwithstanding, some CG models were already developed with surface tensions [30], excess volumes or Upper Critical Solution Temperatures (UCST) [8]. Again, *a priori*, there is no right or wrong set of properties to use in the modelling; rather, it is deeply related with the phenomena to investigate and the ultimate goals of the coarse-graining itself.

With SAFT as an equation of state, a direct link between intermolecular parameters and macroscopic properties was established – and that is why it is so widely used and so important for the modelling part of this work. For instance, in 2001, Müller and Gubbins [39] managed to relate the Helmholtz free energy with the Lennard-Jones (LJ) intermolecular potential parameters using the SAFT equation fitted to saturated fluid densities and pressures. In so doing, they started from Wertheim's perturbation theory and developed a theoretical treatment that allows for the computation of the Helmholtz free energy from statistical mechanical grounds, thus providing a theoretical robustness unseen in the family of cubic equations of state (the first of which was van der Waals'). This achievement endorsed the SAFT EoS as an impressive coarse-graining instrument, not only because of the aforementioned link but also because the set of equations it comprises allows for a much faster (and less "empirical") fitting of the parameters. The SAFT methodology has already been brought to bear on the modelling of substances such as CO_2 [40], water [41], benzene [37], alkanes [42], among many other pure compounds and their mixtures. More recently, and with an obvious importance for this work, it was successfully applied to pure perfluoroalkylalkanes [43]. The theoretical background behind SAFT will be explored in detail in the next chapter.

A widely spread force field developed in a top-down approach is the MARTINI [44], designed to study biomolecules. Even though biochemistry is far from the topic of this work, it should nevertheless be stressed the promising results obtained from a mixed procedure: the authors combined a top-down approach for the non-bonded interactions (based on thermodynamic data [27]) with a bottom-up methodology aiming at the bonded interactions (based on atomistic simulations). This particular tactic of modelling with both the atomistic structure and the macroscopic behaviour of the systems via bottom-up and top-down approaches has revealed quite successful, and was used indeed throughout this work.

4. Statistical Mechanics and Equations of State

The link between the microscopic and macroscopic description of a chemical system has its theoretical foundations in Statistical Mechanics. Though this introduction does not pretend to have the detail of a textbook on the subject, ([45-47] are recommended for more detail) some basic concepts ought to be presented to give an overall idea of the connection between both simulation and equations of state (addressed later).

First, it is introduced the $6N$-*Phase Space*, or *Gibbs Phase Space* (denoted by the greek letter Γ) which is the multidimensional space of both positions and momenta coordinates. It is, therefore, a $6N$-dimensional space (for each of the N particles, there are three positions (x, y, z) and three momenta (p_x, p_y, p_z) coordinates). One moves then to one of the main postulates of Statistical Mechanics – the *ergodic hypothesis*. Behind this concept is the principle[5] that, if a system with a defined energy is given sufficiently long time, then each and every phase point compatible with that energy will ultimately be crossed [48]. Hence, the ensemble average of a physical quantity (the one experimentally measured) equals the *long-time* average of any given "sub-ensemble". This is remarkably helpful for molecular simulation, as it empowers the simulation of a few thousands of molecules to capture the behaviour of a whole macroscopic system (with ~ 10^{23} molecules).

Another way of explaining the *ergodic hypothesis* is to admit that "*every accessible point in configuration space can be reached in a finite number of (…) steps from any other point*" [20]. Provided that a molecular simulation is sufficiently long, any configuration can be achieved and, therefore, the time average of a specific property for the sub-ensemble equals the ensemble average. Equation 8 formalizes this hypothesis. In every time step, the position and momenta of the molecules change, and with that also changes the property of interest A. If a sufficiently long simulation is carried out (that is, if the observation time t_{obs} is sufficiently large), one allows for all microscopic states to be crossed, and thus the average of this property over the whole simulation time in the sub-ensemble simulated will equal the observable value of that very property in the macroscopic scale (A_{obs}). For the sake of clarity, it should be stressed that the ergodic hypothesis is essentially responsible for the first equality of Equation 15; and whilst the first member is referred to the *ensemble* (or macroscopic system), the following members concern the *sub-ensemble* simulated that is intended to replicate the macroscopic system.

$$A_{obs} = \langle A \rangle_{time} = \langle A(\Gamma(t)) \rangle_{time} = \lim_{t_{obs} \to \infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} A\big(\Gamma(t)\big)\, dt \quad (15)$$

An ensemble can be defined by three state fixed variables, with the most common ensemble being the canonical one (with fixed *N*, *V* and *T*, and accordingly called *NVT*). Other ensembles include the *NVE* (microcanonical), the *μVT* (grand canonical ensemble) and the *NPT* (isothermal-isobaric ensemble) [49].

It follows then that if a link between the macroscopic and microscopic worlds ought to be established, one has to correlate the thermodynamic properties of a system with the corresponding probability distributions functions of finding the system in every specific configuration. The Boltzmann distribution provides the probability of finding a subsystem in a particular set of coordinates in the phase space – more precisely, of finding a subsystem with certain positions r^N and momenta p^N (Equation 16).

$$P\Big(H(r_i^N, p_i^N)\Big) = \frac{\exp(-\beta H(r_i^N, p_i^N))}{\sum_k \exp(-\beta H(r_k^N, p_k^N))} \quad (16)$$

[5] Also known as *Principle of Equal* a priori *probabilities*.

where $H(r_i^N, p_i^N)$ is the classical Hamiltonian operator and $\beta = 1/(k_B T)$. The denominator, which is a summation extended to all states of the system, is named *(canonical) partition function* (hereafter Q_{NVT}). In general (Haslam suggests "*above 50 K*" [45]), energy levels are so closely spaced that the energy can be regarded as a continuous variable, thus allowing the substitution of the discrete summation by the (continuous) integral over the *whole phase space*; bearing in mind that the particles are identical, the partition function is now given by Equation 17.

$$Q_{NVT} = \frac{1}{N!\, h^{3N}} \int \exp\left(-\beta H(r_k^N, p_k^N)\right) dr^N\, dp^N \quad (17)$$

While the $N!$ factor takes into account the indistinguishability of the particles, the h^{3N} factor results from the Heisenberg uncertainty principle and grants that the partition function is consistent in form with quantum statistical mechanics [47]. The Hamiltonian, nonetheless, can be expanded into its two contributions: the kinetic and the potential energies. Fortunately, the kinetic energy depends exclusively on the momenta coordinates, whilst the potential energy is just a function of the position coordinates. This fact allows for the separation of both contributions and, which is equally important, to *independent integrations* [45]:

$$Q_{NVT} = \frac{1}{N!\, h^{3N}} \int \exp\left(-\beta V(r_k^N)\right) dr^N \int \exp\left(-\beta K(p_k^N)\right) dp^N = Q^{id} Q^{res} \quad (18)$$

where Q^{id} and Q^{res} are, respectively, the ideal and residual contributions to the partition function. If one considers an ideal gas, for which the potential is null, the partition function can be resumed in Equation 19.

$$Q_{NVT} = Q^{id} = \frac{V^N}{N!\, \Lambda^{3N}}, \qquad \Lambda = \sqrt{\frac{h^2}{2\pi m k_B T}} \quad (19)$$

where Λ is the *de Broglie thermal wavelength*, h is the Planck constant, m is the mass of the particle, k_B is the Boltzmann constant and T is the absolute temperature in Kelvin. The de Broglie wavelength conveys whether or not quantum mechanics play an important role at describing macroscopic behaviour: at low temperatures, the wavelength is higher, comparable with the average distance between particles and therefore relevant; at high temperatures, the wavelength is much smaller than the length-scale of particles motion [48].

Hence, it is clear that the potential term, which accounts for the interactions between the particles, actually imparts the *correction* to the ideal state. This correspondence, on the one hand, between the kinetic and ideal terms and, on the other hand, between the potential and residual terms, is truly enlightening. For any fluid other than the ideal gas, the partition function is given by Equation 20:

$$Q_{NVT} = \frac{\int \exp\left(-\beta V(r^N)\right) dr^N}{N!\, \Lambda^{3N}} = \frac{Z}{N!\, \Lambda^{3N}} \quad (20)$$

where Z is the *configurational integral*, as the integral over all position coordinates of the potential contribution is to be named hereafter. Describing the interactions between the particles is the

key, as shown, to the partition function which, as stated through the *Massieu Bridge* (Equation 21), is related with the Helmholtz free energy, A.

$$A = -k_B T \ln Q_{NVT} \quad (21)$$

Since $dA = -SdT - pdV + \mu dN$, it is possible to draw immediate relationships between the molecular potentials (through the configurational integral and the partition function) and the macroscopic thermodynamic properties – Equations 22.1, 22.2 and 22.3.

$$S = -\left(\frac{\partial A}{\partial T}\right)_{V,N} \quad (22.1) \qquad p = -\left(\frac{\partial A}{\partial V}\right)_{T,N} \quad (22.2) \qquad \mu = \left(\frac{\partial A}{\partial N}\right)_{V,T} \quad (22.3)$$

In chemical engineering problems, perhaps as useful as the *NVT* ensemble is the *NPT* ensemble, since many industrial blocks and units are operated at constant pressure, and not at constant volume. In such circumstance [15], the probability density is proportional to:

$$\exp(-\beta(H(r^N, p^N) + PV)) \quad (23)$$

with the pertaining partition function being given by Equation 24:

$$Q_{NPT} = \frac{\beta P}{N!\, h^{3N}} \int dV \int \exp(-\beta(H(r^N, p^N) + PV))\ dr^N dp^N = \beta P \int \exp(-\beta PV)\, Q_{NVT}\, dV \quad (24)$$

βP can be interpreted as an inverse volume chosen to render a dimensionless partition function [47]. The Gibbs free energy is, therefore, written as in Equation 25:

$$G = -k_B T \ln Q_{NPT} \quad (25)$$

The last equation clearly states that the evaluation of the *NPT*-ensemble average involves calculating the canonical-ensemble average in the first place and then averaging over V with a weight factor $\exp(-\beta PV)$ [47]. Again, it is easy to relate macroscopic properties with derivatives of the Gibbs free energy, leading to Equations 26.1, 26.2 and 26.3:

$$S = -\left(\frac{\partial G}{\partial T}\right)_{P,N} \quad (26.1) \qquad V = -\left(\frac{\partial G}{\partial p}\right)_{T,N} \quad (26.2) \qquad \mu = \left(\frac{\partial G}{\partial N}\right)_{P,T} \quad (26.3)$$

4.1 Statistical Associating Fluid Theory (SAFT)

Since van der Waals first proposed the equation for real gases that today bears his name, many Equations of State (EoS) have been developed to provide accurate predictions of single-phase properties, activity coefficients and phase equilibria [50]. Many followed the van der Waals structure and became generally known as *cubic equations of state* – including the renowned Redlich-Kwong and Peng-Robinson [51]. However, these proposals did not solve the intrinsic inconsistencies of the van der Waals EoS, and not only their accuracy is restricted to a limited range of conditions but they also exhibit important limitations when applied to associating and non-spherical molecules, as well as when used to predict the so-called “second derivative properties” – heat capacity, speed of sound, etc.

The SAFT Equation of State is one example of a *perturbation theory*. These theories represent a given property of the system (such as distribution functions, pressure, free energy, among others) as

a perturbation from a reference system [52]. The calculation of the "real property", so to speak, comprises evaluating that very property for the reference fluid and then estimating the "perturbations". The reference fluid, must be said, is normally derived from a simpler model whose properties are known.

Advances in the field of applied statistical mechanics throughout the 20th Century fostered the SAFT EoS, first proposed by Jackson, Chapman and Gubbins [53]. Anchored in Wertheim's theory of association (TPT) [54-57], the original SAFT EoS expresses the residual Helmholtz free energy as the sum of several contributions (see Equation 27): with an hard sphere reference fluid, the perturbations entail terms referring to the formation of chain and to association (through Hydrogen bonding) [58].

$$\frac{A^{res}}{Nk_BT} = \frac{A}{Nk_BT} - \frac{A^{ideal}}{Nk_BT} = \frac{A^{ref}}{Nk_BT} + \frac{A^{pert}}{Nk_BT} = \frac{A^{mono}}{Nk_BT} + \frac{A^{chain}}{Nk_BT} + \frac{A^{assoc}}{Nk_BT} \quad (27)$$

where the numerous superscripts in the Helmholtz free energy (A) stand, from left to right, for: residual; ideal; reference fluid; perturbations; monomer; chain formation; association. The second term, without superscript, refers to the "real" Helmholtz free energy of the system [50].

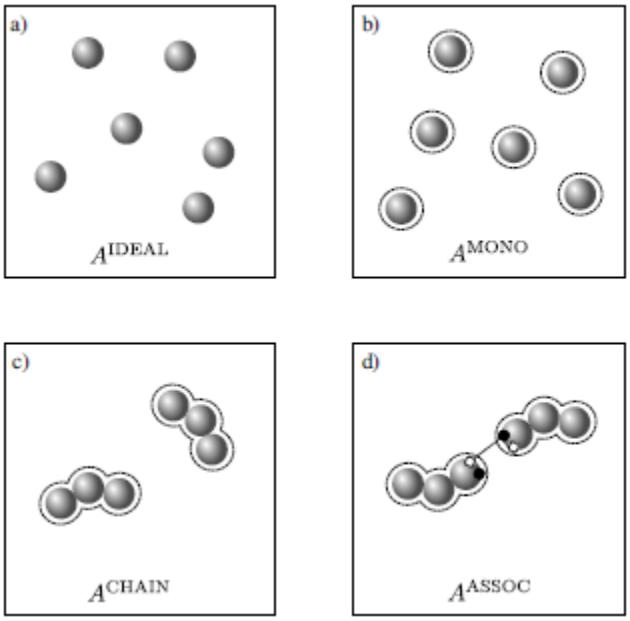


Figure 6 – The Helmholtz free energy can be evaluated by adding, to the ideal term, the three different perturbations (monomer, or hard spheres; chain formation; association). From [26] (p. 17).

Figure 6 shows the gradual evolution of the perturbation scheme: (a) the ideal fluid is constituted by particles with no volume and no interactions; (b) then, the volume of atoms is considered, as well as the repulsive and attractive interactions that are established between atoms; (c) further, atoms are gathered in chain molecules and their interactions – at both intramolecular and intermolecular levels – are considered; (d) finally, association sites, if required, are added to account for phenomena like hydrogen bonding. Obviously, this particular treatment can only be used for chain molecules (which did not limit at all the present work, entirely focused on *n*-alkanes, *n*-perfluoroalkanes and *n*-perfluoroalkylalkanes).

A wide plethora of versions of SAFT has been published since the original version first appeared. However, all of SAFT 'flavours' share the feature of being a molecular Equation of State that relates parameters for the intermolecular potential with macroscopic properties. Since this thesis was developed using the SAFT-γ Mie Equation of State, the description of the several terms of the SAFT perturbation equation will summarize the approach pursued in the original SAFT-gamma Mie paper [59]. Notwithstanding, a brief overview of the other incarnations of SAFT is also provided in the end.

IDEAL TERM

The Helmholtz free energy of an ideal gas can be derived from the *Massieu Bridge* (Equation 21) and the ideal gas equation [45], yielding Equation 28.

$$A^{ideal} = -k_B T \ln Q_{ideal} = -N k_B T \left(\ln\left(\frac{V}{N}\right) + 1 - \ln(\Lambda^3) \right) \quad (28)$$

From here, it follows that, for a mixture with Nc components, each with a mole fraction of x_i and a numerical density of $\rho_i = N_i/V$, the ideal term is given by Equation 29:

$$\frac{A^{ideal}}{N k_B T} = \left(\sum_{i=1}^{N_C} x_i \ln(\rho_i \Lambda_i^3) \right) - 1 \quad (29)$$

MONOMER TERM

The monomer term accounts for the interactions between each segment, or monomer (whether it is an atom, a molecule, or a CG bead), and for that reason requires that an intermolecular potential is defined [58]. Initially, Barker and Henderson proposed a perturbation expansion in powers of β to second order [60] in which the reference term is the repulsive hard sphere and the perturbation terms add attractive interactions; in the SAFT-γ Mie framework, however, a third order expansion is used, as translated in Equation 30:

$$\frac{A^{mono}}{N k_B T} = \frac{A^{HS}}{N k_B T} + \frac{A_1}{N k_B T} + \frac{A_2}{N k_B T} + \frac{A_3}{N k_B T} \quad (30)$$

In the SAFT-γ group contribution treatment, the contribution of each segment to the free energy is given by Equation 31:

$$\frac{A^{HS}}{N k_B T} = \left(\sum_{i=1}^{N_C} x_i \sum_{k=1}^{N_G} \nu_{k,i} \nu_k^* S_k \right) a^{HS} \quad (31)$$

where N_C is the number of components, N_G is the number of group types, $\nu_{k,i}$ is the number of occurrences of group type k in molecule i, S_k is the shape factor (a free energy weighting factor) of group k, ν_k^* stands for the number of identical segments that form group k and a^{HS} is derived from the expression of Boublik [61] and Mansoori [62].

For the attractive perturbation terms, A_1 is associated with the mean attractive energy per molecule (and is evaluated through the integral mean value theorem) whereas A_2 accounts for the local

fluctuations in the attractive interactions (and is estimated using the local compressibility approximation). The more recent use of a third order expansion has allowed for a greatly improved approximation to the reference radial distribution function (RDF) [45].

CHAIN TERM

By including the chain term, the SAFT EoS accounts for the free energy contributions due to the formation of chains made of monomer segments. For a mixture, the chain term is given by Equation 32:

$$\frac{A^{chain}}{Nk_BT} = -\sum_{i=1}^{N_C} x_i \left(\sum_{k=1}^{N_G} \nu_{k,i} \nu_k^* S_k - 1 \right) \ln g_{ii}^{Mie}(\bar{\sigma}_{ii}, \xi_x) \quad (32)$$

where the different parameters have the same meaning presented in Equation 31 and $g_{ii}^{Mie}(\bar{\sigma}_{ii}, \xi_x)$ represents the RDF evaluated at a distance $\bar{\sigma}_{ii}$ in a hypothetical fluid (whose interactions are given by the Mie potential) of packing fraction ξ_x. For a mixture, the chain term is built upon the approximation that the heteronuclear chain can be modelled as a homonuclear chain with effective/average parameters. The RDF for the Mie fluid is in turn obtained as a correction of the hard sphere RDF.

ASSOCIATION TERM

The association contribution has its roots in Wertheim's first-order thermodynamic perturbation theory (shortened for TPT1) and it is motivated precisely because some molecules tend to "stick together" [45] for various reasons – water in the liquid phase, carboxylic acids in the gas phase, NO, NO_2, among others [39]. Not surprisingly, hydrogen bonding plays a fundamental role in the association of molecules, and has been a molecular phenomenon that other well-established equations of state have failed to accurately describe. The reason behind this poor predictive capability comes from the complex set of forces that governs the interactions in such compounds that are not considered in an explicit way (Coulombic forces, complexing forces, induction forces, even forces related to the chain flexibility, to name just a few).The interested reader is referred to [39] for an in-depth discussion of the association contribution in SAFT.

The contribution due to the association via short-range bonding sites follows can be translated into a triple summation (condensed in Equation 33): over the number of components (N_C), the number of groups in each component (N_G) and the number of sites on each group ($N_{ST,k}$):

$$\frac{A^{assoc}}{Nk_BT} = \sum_{i=1}^{N_C} x_i \sum_{k=1}^{N_G} \nu_{k,i} \sum_{a=1}^{N_{ST,k}} n_{k,a} \left(\ln X_{i,k,a} + \frac{1 - X_{i,k,a}}{2} \right) \quad (33)$$

where $n_{k,a}$ stands for the number of sites of type a on group 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. This fraction is obtained through a set of equations called *mass-action equations*, which essentially result from a mass balance

to association sites. Interestingly, the chain term can be interpreted as a limiting case of an infinitely strong association between molecules – which turns out to form an "unbreakable" bond [45].

Despite the fact that SAFT-γ Mie has been used recently to model the behaviour of more complex mixtures, one should bear in mind that this is indeed a late version of SAFT, distinguished for two important reasons: because it includes the Mie potential, a more sophisticated expression for the intermolecular potential ("SAFT-γ **Mie**"); and because it embodies the idea of a group contribution of fused heteronuclear segments (SAFT-**γ** Mie). Other versions of SAFT include: PC-SAFT (perturbed chain), first proposed by Gross and Sadowski in the early 2000's [63], which used a hard-chain reference system instead of a hard-sphere; soft-SAFT, initially developed by Johnson [64] and extended by Blas and Vega [65] , which treats chain molecules as Lennard-Jones segments; among many, many others. The reader is referred to a thorough review by McCabe and Galindo [66] for more information.

By now, it should have become clear that the SAFT EoS (really, any of the equations under this label) cannot be used in one step to compute properties. The fact that the parameters comprising each equation are, one after the other, obtained through other expressions makes the implementation of SAFT a far from trivial task. The analytical procedure of relating thermodynamic properties to the microscopic parameters is virtually impossible to be done "manually". The software gSAFT®, developed by the company Process Systems Enterprise[6], has overcome this difficulty and allows its users to effectively take advantage of the powerful predictive capacity of this breakthrough theory.

5. Force Fields

The background information in this introduction could not exclude the definition of force fields, even though, in previous sections, methodologies to develop them were presented. It is well-known that every single macroscopic property of a pure component or a mixture is governed by the interactions between particles [67]. Understanding the set of interactions established in a chemical system is therefore the key to capture the physics of the system and the first step to attain sound property predictions.

The forces ruling the motion of particles are given by the negative derivative of the potential (Equation 34), thus establishing a direct relationship between the molecular potentials and the evolution of the molecules' trajectory (via Newton's Second Law of Motion) in the chemical system.

$$F_i(r^N) = -\nabla_{r_i} u_i(r^N) \quad (34)$$

where $F_i(r^N)$ represents the force acting on particle i, $u_i(r^N)$ is the potential of particle i, and r^N stands for the position of all the N particles in the system. Not surprisingly, the potential depends on the space coordinates of all particles, which is rather challenging to evaluate in a computer simulation. Because of this, the potential $V(r^N)$ of a system comprising N spherical particles is normally split as a summation of the pairwise interaction between molecules $u(r_i, r_j)$, the three-body $u(r_i, r_j, r_k)$, and many-body contributions [15]. The functional form of the potential between particles is called *force field*.

[6] https://www.psenterprise.com/

$$V(r^N) = \sum_i u(r_i) + \sum_i \sum_{j>1} u(r_i, r_j) + \sum_i \sum_{j>1} \sum_{k>j>i} u(r_i, r_j, r_k) + \cdots \quad (35)$$

The first term represents the eventual effect of an external field (for instance, electrical fields or container walls). The pairwise interaction is, by far, the most important term; the three-body contribution accounts for about 10% and the many-body contributions can be neglected [15]. Besides much less relevant, the many-body terms are computationally very demanding, which has led to the widespread utilization of an *effective pair-wise potential*.

$$V(r^N) \approx \sum_i \sum_{j>1} u^{eff}(r_i, r_j) \approx \sum_i \sum_{j>1} u^{eff}(r_{ij}) \quad (36)$$

Above, the assumption of an isotropic system (as a liquid phase usually is[7]) allows for the effective potential to be rewritten as a function of the distance between two particles instead of the position of both particles [45]. Moreover, because of the summation, the potential is also assumed *pair-wise additive.* For molecules that only exhibit van der Waals forces (like the alkanes and perfluoroalkanes on which this thesis focuses), this assumption is safe to use; however, it has been studied that, for associating fluids such as water, cooperative effects (that is, the non-additivity of the pair-wise interactions) are quite significant [68].

It should be noted that Equations 35 and 36 are suited for *systems of isolated, spherical particles* – they do not account for the intramolecular interactions. Because even a quite simple system involves solving a (unsolvable) many-body problem, a procedure named *parameterisation of the force field* needs to be undertaken.

Parameterisation refers to the process of choosing a functional form for the force field and defining how the different interactions are expressed. Molecular force fields necessarily include different contributions to account for the diverse array of phenomena happening at the molecular scale (see Figure 7). One can detach the interactions between different molecules and interactions within the same molecule – these will be called, respectively, *non-bonded* (belonging to the *intermolecular contributions*) and *bonded* interactions (belonging to the *intramolecular contributions*). Building a force field requires, thereby, a joint description of both intramolecular and intermolecular potentials. A brief analysis of those potentials follows.

[7] Liquid crystals may be regarded as an exception due to the arrangements of the molecules.

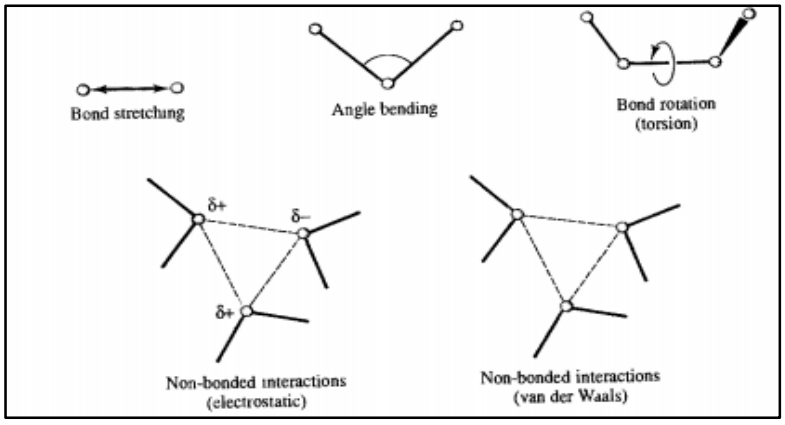


Figure 7 – Represented are the bonded *(top) and the* non-bonded *(bottom) interactions. From [16] (p.166).*

5.1 Intramolecular forces

Bonded interactions include bond stretching, angle bending and dihedral rotation forces. As a first approximation, the bond stretching and angle bending forces are translated by the Hooke's Law[8] (also called *harmonic law*). The dihedral or torsional potential can be written in several different ways; in particular, the Ryckaert-Bellemans (RB) model [69] describes the potential as a power series of cosines. The bonded contribution can be outlined in Equation 37, with the summations extended to all the bonds, angles and dihedral configurations in the system.

$$U^{bonded}(r,\theta,\varphi) = \sum_{bonds} \frac{1}{2} k_{bond}(r - r_0)^2 + \sum_{angles} \frac{1}{2} k_{angle}(\theta - \theta_0)^2 + \sum_{dihedrals} \sum_{n=0}^{5} C_n \cos^n(\varphi) \quad (37)$$

where k_{bond} and k_{angle} are the stretching and bending constants, respectively; r_0 and θ_0 are the equilibrium bond length and equilibrium angle, respectively; and C_n is the set of Ryckaert-Bellemans constants for the torsional term. It may prove useful to know that the RB potential can be exactly converted in the form $\sum_{n=0}^{5} b_n\,(1 + (-1)^n \cos(n\varphi))$; hence, attention should be given to the source of parameters to assure coherence. Figure 8 illustrates the different contributions for the intramolecular forces.

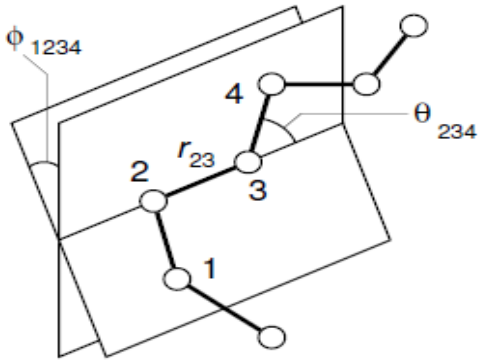


Figure 8 – Geometry of a linear molecule, highlighting bond stretching, angle bending and dihedral rotation. From [22] (p. 4).

[8] Named after the British physicist Robert Hooke (1635 – 1703).

Notice that the functional form in Equation 37 implicitly states that the bond stretching depends only on the bond length; the angle bending exclusively on the angle; and so on. Of course, this is not the real case, but it is the practical impossibility to use “rigorous” potentials, allied to the usually good agreement that these parameterisation achieves, that drives this simplification [16].

Finally, it must be stressed that bonded interactions focus on interactions between atoms 1-2, 1-3 and 1-4. For a molecule with more than three bonds, the interactions between atoms more than three bonds away is given by the intermolecular part of the force field (thus treating them as atoms of “different molecules”).

As abovementioned, in this thesis, the intramolecular contribution to the force field was determined through Direct Boltzmann Inversion from atomistic simulations. This procedure quantified not only the spring constants and equilibrium values for bond stretching and angle bending, but also determined the energetic barriers pertaining the torsional part of the potential.

5.2 Intermolecular forces

The history of intermolecular potentials dates back to Newton’s *Principia*, and was investigated by both Clausius and Maxwell, just to name a few [70]. In the dawn of the 20th Century, Gustav Mie proposed the renowned *Mie potential* [71] and, in 1924, John Lennard-Jones [72] wrote the potential that got ubiquitous under his name. Extensive literature about the different potentials is available and can be found in [67]. In the following chapter, attention will be drawn to the most useful potentials for simulation purposes.

It is well-known that the most rudimentary potentials – hard-spheres, square-well, Sutherland – are more difficult to compute due to the discontinuity on their derivative (which is necessary to calculate the force and, subsequently, the trajectory), let alone their inaccuracy [16]. Consequently, expressions such as those for the Lennard-Jones [72] or the Mie potential [71] have been successfully implemented for molecular simulations. The functional forms of the potentials are as follows (Equations 38, for the Lennard-Jones, and 39 and 40, for the Mie).

$$u_{LJ}(r) = 4\varepsilon\left[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6}\right] \quad (38)$$

$$u_{MIE}(r) = C\varepsilon\left[(\frac{\sigma}{r})^{\lambda_r} - (\frac{\sigma}{r})^{\lambda_a}\right] \quad (39)$$

$$C(\lambda_a, \lambda_r) = \left(\frac{\lambda_r}{\lambda_r - \lambda_a}\right)\left(\frac{\lambda_r}{\lambda_a}\right)^{\frac{\lambda_a}{\lambda_r - \lambda_a}} \quad (40)$$

For both potentials, the pre-factor grants that the minimum of the potential has an energy of $-\varepsilon$, with that factor being calculated through the expression of Equation 40. Easily, one can confirm the “4” in the LJ potential by substituting both the attractive and repulsive exponents by 6 and 12, respectively.

The Mie potential, where both exponents are adjustable parameters, offers far more versatility than the Lennard-Jones [40], an attribute that has proven to significantly improve the prediction of thermodynamic properties such as vapour pressure, speed of sound, heat capacity and compressibility.

The intermolecular parameters were, in this thesis, determined with the theoretical treatment of SAFT-γ Mie – which explains why it can safely be called a *force field of the SAFT-γ Mie family*. It is this successful feature of SAFT that made it so popular to model intermolecular potentials (from hard-spheres and square-well to Lennard-Jones and Mie).

In addition to the van der Waals interactions, Coulomb forces often play an important role in the definition of the force field, namely in the presence of important partial charges [15]. Nonetheless, the increased computational cost of the simulation thus implied supports the usual decision of omitting electrostatic interactions in coarse-grained models, for the sake of simplicity [42]. This approach was followed in this work.

As a final note, it should be emphasized that the different contributions above listed are not equally important; and that it is actually possible to rank them by "relative strength" (Equation 41) [73].

$$U_{stretching} \rightarrow U_{bending} \rightarrow U_{non-bonded} \rightarrow U_{torsional} \quad (41)$$

In fact, the torsional contribution is often marginal, namely in coarse-grained models. Subsequent calculations will support the decision of neglecting that term.

6. Simulations of Fluorocompounds: a State-of-the-Art

The interesting properties of fluorocarbons have motivated experimental research since the 1950s but, more recently, they have also driven computational studies through molecular simulations. This last section of the Introduction aims at presenting the most significant work done in terms of computer simulations of fluorocarbons. Built on the concepts, previously covered, of force fields, SAFT, molecular dynamics and coarse-graining, it prepares the ground for future results.

Previous force fields [74] often addressed PFA modelling as a way to investigate polytetrafluoroethylene (PTFE, the basic compound of Teflon® [75]); others dedicated to some of the shortest fluoroalkanes, such as perfluoromethane [76] and perfluoropropane [77], motivated by concerns with refrigerants such as CFC and HFC. This brief overview of proposed force fields does not pretend to be exhaustive, and the reader is referred to Jorgensen *et al.* [74] for a broader review.

However, essentially none of the previous force fields had been developed to specifically address long fluorocarbon chains, which obviously hampered the computational study of those compounds. The first force field effectively aimed at these very chains dates back to 1998, when Siepmann *et al.* [78] proposed a UA force field (based on QM data) in which the intermolecular interactions were described by a Lennard-Jones (LJ) potential. This model managed to reproduce, quite accurately, saturated liquid and vapour densities, as well as predict, within less than 4%, critical temperatures. In 2001, also from *ab initio* studies, Jorgensen *et al.* [74] proposed a set of parameters comprehended in the OPLS-AA project Jorgensen himself directed. With their AA model, the authors correctly predicted, through MC simulation, liquid densities and heats of vaporization for linear, branched and cyclic perfluoroalkanes (with no more than six carbon atoms), using a LJ potential for the non-bonded interactions too. Nevertheless, as Potoff *et al.* later recognised [79], the LJ potential made it impossible to simultaneously and accurately reproduce both saturated liquid densities and vapour

pressures. This limitation motivated their subsequent United-Atom modelling of perfluoroalkanes using a Mie (or, more precisely, an *n-6* Lennard-Jones) Potential. The accuracy of this description largely surpassed the previous proposals, with the authors reporting not only excellent agreement for pure properties of both alkanes and perfluoroalkanes (saturated liquid and vapour densities, vapour pressures and heats of vaporization) but also predicting with remarkable precision the vapour-liquid equilibrium (VLE) of ethane – perfluoroethane, a highly non-ideal mixture. It is impressive indeed how the models of Jorgensen (under the label OPLS-AA) and Siepmann (later included on the TraPPE-UA initiative) fail resoundingly in the VLE prediction.

More recently, in a paper that may be considered the starting point of this thesis, Morgado *et al.* [43] performed coarse-grained simulations of semifluorinated alkanes. The model mimicked the procedure followed for the alkanes [42], that is, a hybrid bottom-up / top-down CG method. For the intermolecular interactions of the fluorinated beads, in particular, an *n-6* Mie potential was used. This breakthrough research showed that not only the properties of pure PFAA can be accurately predicted using a CG methodology but also that such a model can be used, in MD simulations, to predict interfacial properties in what is truly a pure prediction (since no interfacial property whatsoever was used in the modelling).

On hindsight, this introductory chapter offered the reader a wide (yet non-exhaustive) panorama of the chemistry of both perfluoroalkanes and perfluoroalkylalkanes; major concepts in molecular simulation, applied statistical mechanics and coarse-graining; and what can arguably be called the "state-of-the-art" for the simulation of fluorocompounds. With this in mind, the pathway for this work is described as follows.

At first, a global reassessment of the force field proposed in [43] is undertaken, in an effort to identify its flaws; then, with those very imperfections listed, a CG force field (based on the previous one, but adopting important changes) is developed, with both intra- and intermolecular contributions. An analysis of the predictive capabilities of the force field follows, to ensure it is reliable in predicting thermodynamic and interfacial data for pure components as well as mixtures. Finally, MD simulations of mixtures of PFAAs in hydro- and fluorocarbon solvents are performed, ultimately aiming at supramolecular organization.

II - Improvements on a Recent Force Field

Despite the reassuring results obtained in the model by Morgado *et al.* [43], there are flaws that ought to be improved. To begin with, the intramolecular interactions for perfluoroalkanes are admitted to be governed by the same set of parameters of the alkanes, with the sole exception of the equilibrium angle (159.9 degrees for the alkanes, 180.0 degrees for the PFA). This value, though, is not grounded in any experimental or computational data but only in the qualitative judgement that the fluorinated chains, due to the large size of the fluorine atoms, are highly stiff and therefore have a "pencil-like" shape. Therefore, the first step to improve the previous set of beads was to develop an intramolecular part for the force field. This step was accomplished through DBI of MD simulations using the UA force field by Potoff *et al.* [79].

In addition, the set of CG beads proposed has a questionable asymmetry. As a matter of fact, while the fluorinated beads have only one or two carbon atoms (CF_3 and CF_2-CF_2), the hydrogenated beads, originally developed for the modelling of long chain alkanes [42], have three or four carbons (CH_2-CH_2-CH_2, CH_2-CH_2-CH_3 and CH_2-CH_2-CH_2-CH_3). In so doing, the authors envisioned a more balanced mapping in terms of molecular weight. This imbalance, though, led to a physical perversion of the SFA structure: even though SAFT purveys a chain picture of a molecule (as opposed to the spherical one implicated in the cubic equations of state), each bead is still treated as a tangent sphere. Because the hydrogenated beads encompass more atoms, their radius (as evaluated through the σ involved in the Mie potential) is larger than that of the fluorinated beads. With the inclusion of many more atoms in the hydrogenated groups than in the fluorinated ones, inevitably the cross-area of the alkyl chain becomes larger than the perfluoroalkyl's – thus subverting the whole visual comprehension of the SFA molecules that is accepted today (as portrayed by Figure 9).

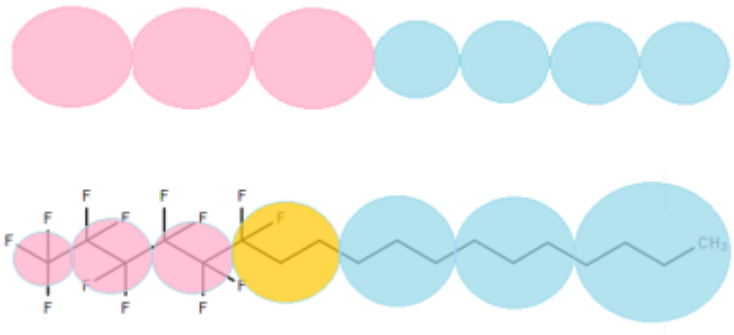


Figure 9 - "Natural" representation of a PFAA (above) and the twisted representation that the asymmetric mapping purveys (below). Circles colour scheme: pink stands for fluorinated beads; blue for hydrogenated beads; orange for the linker that connects both tails.

From here, it seems that this twisted depiction arises from the very uneven size of the fluorinated and hydrogenated beads. Truly: *to change, or not to change, the set of beads: that is the question.*

Two important considerations should be borne in mind on this topic. It is well-known that the merits of a CG force field should be appraised based on the accuracy of the theoretical and computational predictions that it upholds; still and all, as pointed out in the background information, good

predictions are not sufficient to support a CG model – it has to be physically meaningful and, as Noid [24] famously summarized, it "*should not only provide the correct answer, but also provide this answer for the correct reason*".

Furthermore, if the long-term goal of such research is to grasp the behaviour of mixtures with fluorinated and hydrogenated chains, a deep understanding of the real structure of the molecules is undoubtedly critical. The way in which those chains interact with each other – and the way those very interactions nurture organization and diffusion towards the surface – relies heavily on the accuracy of the global set of forces felt by the particles (specifically, the CG beads). A useful, easily measured and directly linked to the potential energy via statistical mechanics is the radial distribution function (RDF), which relates the density of molecules, or atoms, or CG beads, in a given shell to the overall density of the liquid phase (or the density of a shell with infinite radius). If the RDF derived through the CG model and the atomistic model are similar, then the short-range structure of the liquid phase is preserved, thus mitigating the inherent losses of detail the coarse-graining methodology encompasses. In turn, the radius of gyration offers an insight, yet perhaps incipient, on the overall length of the molecule. Again, matching this measurement is an *a priori* victory of the CG model – even if it does not guarantee a more powerful predictive capability, at least a sounder description of the structure may be accomplished.

But before discussing in detail the preliminary calculations of the RDF, let one consider a parallel aspect that is no less relevant. As explained, one of the core assumptions of this model is to use the σ optimized through the SAFT-γ Mie framework (or subsequent combining rules) as equilibrium bond length (r_0) between beads; in other words, one is using an intermolecular parameter in the intramolecular potentials to ensure that the beads comprising the chain are effectively *tangent spherical beads*. This is absolutely indispensable because even though the SAFT-γ Mie theoretical framework allows for non-integer beads[9] (it is just another parameter in the set of the SAFT equations), computer simulations are limited to an integer count of beads, from which arises the handicap of coarse-grained models to describe a wide range of compounds.

Experience has shown [42] that the "non-bonded" σ is consistently higher than the "bonded" r_0. For this reason, the pertaining group is kept further apart from the bonded beads than in reality it is, which tends to produce a "stretched" molecule. By taking a CF_3 as a CG group, a tremendous imbalance evolves: Direct Boltzmann Inversion of atomistic simulations of perfluoroalkanes has shown that the average length of a $CF_3 – (CF_2)_2$ bond (note that the length is measured between the centres of mass) is roughly 0.19 nm – in stark contrast with the σ = 0.39 nm reported! As a comparison, for the alkanes, the average length of a $C_3H_6 – C_3H_6$ bond is around 0.35 nm, much closer to the 0.43 nm optimized with SAFT.

Bearing this in mind, one focuses now on the RDF and the radius of gyration of perfluorohexane (formed by the group sequence $CF_3 – (CF2)_2 – (CF2)_2 – CF_3$), as obtained for an equilibrated system in

[9] This is accomplished via the *Shape factor*, which allows for "fused" segments – another way of looking at non-integer beads.

the liquid phase and compared to similar experiments for atomistic simulations (UA, with the molecule being formed by six groups). The results are shown on Figures 10 and 11.

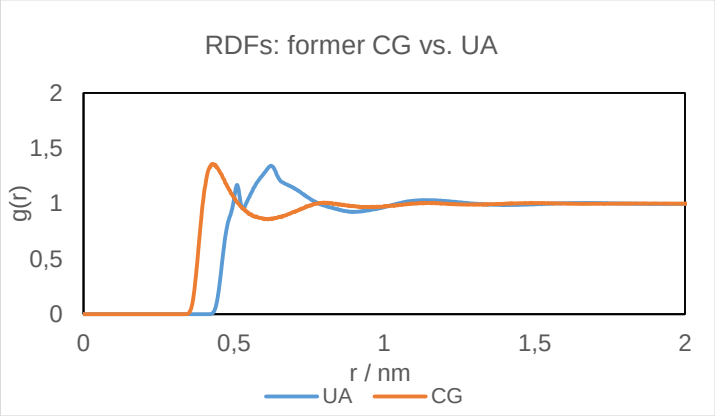


Figure 10 - Comparison of the RDF (evaluated between the centres of mass) of F_6 at 373.15K and 5 bar, for both UA and CG models.

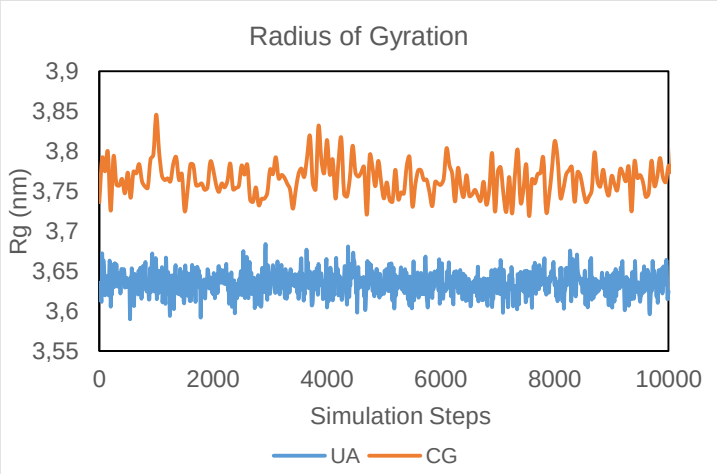


Figure 11 - Comparison of the RDF (evaluated between the centres of mass, left) and the radius of gyration (right) of F_6 at 373.15K and 5 bar, for both UA and CG models.

Looking at the RDF's exhibited, it is noticeable that not only the peaks do not match but, which is far more serious, the CG $g(r)$ increases from null values significantly earlier than its atomistic counterpart. This fact, let alone the slight deviation in the radius of gyration, hints for a poor comprehension of the short-range structure by the set of fluorinated groups proposed in [43]

The especially high difference between the "intramolecular" r_0 and the "intermolecular" σ witnessed for the fluorinated beads may help to explain the disparity between these conflicting RDF's. In a smaller extent, they may also be related with the differences confirmed in the measurement of the radius of gyration.

As a first attempt to mitigate these asymmetries, and because the work on the fluorinated chains was not as advanced as the one already done for the alkanes, the fluorinated groups were slightly enlarged: the terminal bead (FE) was converted from CF_3 to CF_3-CF_2; and the middle bead (FM) was changed from $(CF_2)_2$ to $(CF_2)_3$. While the substitution of the terminal bead was the most urgent (given

the fact that it had only one carbon atom, with the negative implications already highlighted), maintaining a two-carbon middle group would prevent the simulation of SFAs with an even number of fluorinated carbons, which constitute the large majority of molecules reported in the literature; for that reason, the middle group was increased as well. The respective RDF and radius of gyration of perfluoroheptane are exhibited in Figures 12 and 13.

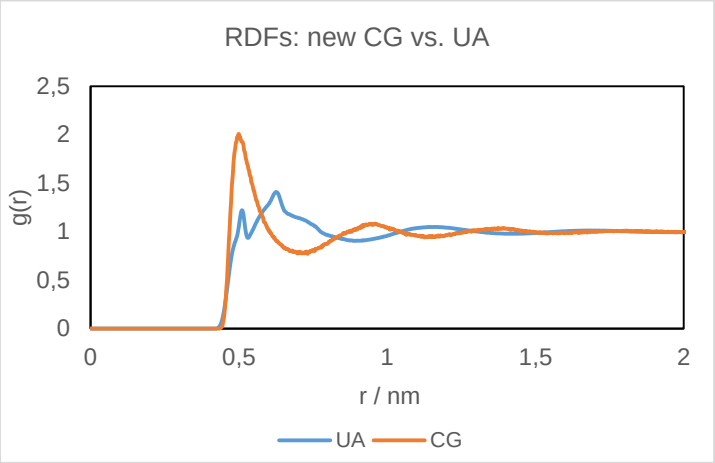


Figure 12 - Comparison of the RDF (evaluated between the centres of mass) of F_7 at 373.15K and 5 bar, for both UA and CG models.

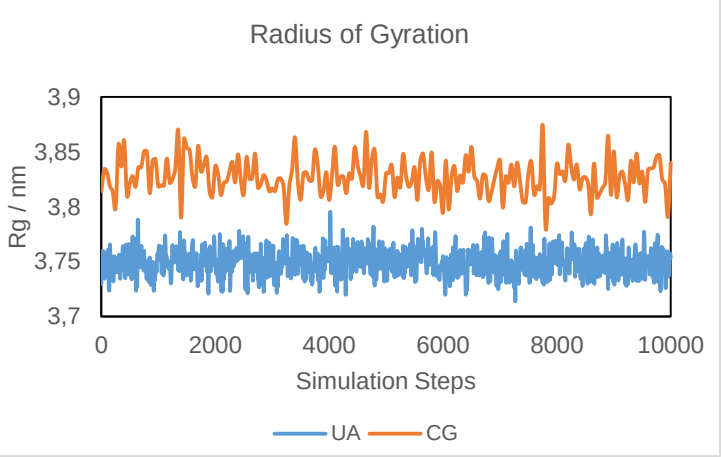


Figure 13 - Comparison of the Radius of Gyration (R_g) of F_7 at 373.15K and 5 bar, for both UA and CG models.

Two improvements must be recognised, even if different molecules are being considered (the different sets of beads do not allow for the "construction" of the same molecules). Whilst in the previous model there was a clear failure to predict the radius of the first coordination sphere (that is to say, the first peak) and the CG RDF actually "started" before its atomistic counterpart, the new one allows for a much better resemblance despite the overshooting of the first peak. Concerning the radius of gyration, the changes are not so relevant, but a relative reduction on the deviation between the atomistic and CG calculations scores one more point for the new set of beads.

Even though the next results will be presented in detail later in this thesis, they are briefly signalled hereby to further embolden this group reshuffle. For instance, adequate simulations indicated an average C_2F_5 – C_3F_6 bond length of 0.320 nm, compared to a cross σ of 0.469 nm. The gains in

similarity are equally witnessed for the C_3F_6 – C_3F_6: whilst the bond length was estimated at 0.370 nm, the associated σ is just 0.474 nm. For now, however, these results are deemed sufficiently convincing to proceed with further studies with the newly developed fluorinated coarse-grained groups.

In short, this very action contributed to curtail the discrepancy between the equilibrium bond length and the optimized intermolecular σ. In this very moment, though, one stands at a crossroads. The Iterative Boltzmann Inversion is a trustable procedure to guarantee that the CG potential leads to an identical radial distribution function. If one is too concerned with the conservation of the short-range order during the coarse-graining procedure, why not using this recognised, widely used tool? The answer lies on the evidence that strategies adopted to reproduce structure are often at odds with others focused on reproducing thermodynamic properties. A complete focus on the reproduction of the RDF could lead to a disastrous prediction of macroscopic properties, thus rendering useless the proposed model. Not surprisingly, a trade-off must be made, and that compromise was finally achieved combining the DBI procedure with the SAFT-γ Mie theoretical framework.

This preliminary result is by no means a confirmation that the model to-be-developed is better than the previous one – namely because the concept of “better” is far too ambiguous in this field: better in predicting properties? And which properties - and in which thermodynamic states? In reproducing spectroscopic data? It all depends on the desired applications and the resources intended to be spent studying them. Only future results may provide a hint on whether this replacement is advisable.

To sum up, there are several ways of continuing the work started on the cited paper, among which are: to change the set of beads to provide a more realistic depiction of the SFAs; to improve the intramolecular forces description; and to apply the model, given its early achievements (namely on interfacial properties), to more challenging assignments such as mixtures. These three vectors light the way followed in this thesis.

III - Force Field Development

1. Intramolecular Parameters

The intramolecular contribution of the force field is obtained through what has been previously named a bottom-up approach, in particular via Direct Boltzmann Inversion (DBI). This procedure accepts three inputs: the atomistic simulation results; the definition of the CG groups (that is, the mapping); and the functional form of the intramolecular potential, which we assume to be the one expressed in Equation 4.

The legitimacy of the DBI technique is confirmed by Noid [24] for "*bonded potentials if the CG bonds are 'stiff'* ", a rather fair assumption for the fluorinated chains given the steric hindrance created by the fluorine atoms (this assumption is essential in [43]).

As detailed in the previous chapter, Potoff's UA force field [79] for alkanes and perfluoroalkanes stands out, in the literature, as the most accurate and complete atomistic force field for perfluorocarbons. It should be stressed that this section pretends solely to provide a rigorous intramolecular potential for the fluorinated chains; the parameters for the alkanes are taken from [42] and are kept unchanged.

In terms of mapping, and as followed from the previous reflection on the set of beads, the fluorinated groups were enlarged from CF_3 to C_2F_5 (labelled FE, after "F-End") and from C_2F_4 to C_3F_6 (FM, after "F-Middle"). This segments constrained the molecules to simulate – it ought not to be forgotten that even if the theory can be safely used for non-integer beads, molecular simulations can only deal, at least under the SAFT-γ Mie framework, with integer CG groups. Therefore, since parameters for bond stretching (two atoms involved), angle bending (three atoms) and torsions (four atoms) were needed, and in an effort to reduce computational effort, two different molecules were studied: perfluoroheptane (C_7F_{16}, or F_7) and perfluorodecane ($C_{10}F_{22}$, or F_{10}). Notice that these compounds are "built" with the sequences FE-FM-FE and FE-FM-FM-FE, respectively.

One of the main handicaps of such *bottom-up* procedures is the state-dependency of the results. In fact, there is no guarantee that the results obtained for the simulated conditions can be safely used for different temperature, pressure, physical state and (if that was the case) composition. In order to circumvent this limitation, simulations of both pure compounds were performed at both 373.15K and 473.15K, in both the liquid and the gas phase, following the procedure done for the alkanes by Sadia Rahman [26]. The temperatures are necessarily below the critical so not to render useless the distinction between the gas and the liquid phases.

For the sake of reproducibility of results, MD simulation details are presented hereby. Systems comprising of 1000 molecules were simulated with the package GROMACS 4.5.5; the time-step was set to 1 fs; the cut-off radius was set to 1.4 nm, roughly nine times the bond length defined by Potoff *et al.* (1.54 nm). The procedure for the liquid phase comprised an energy minimization period, followed by 5 ns of equilibration, followed by 10 ns of production, both in the NPT ensemble. For the gas phase, and starting from the final configuration on the liquid phase, the steps ensued were: edit the configuration of the box from a cubic *L x L x L* to an (approximate) *L x L x 3L*; undergo an energy minimization period,

after which an NVT simulation (5 ns) was carried on; then, based on the gas phase density in the box (on both sides of the liquid slab), a new box was generated that would allow for roughly the same density, keeping constant the number of molecules; finally, after the usual energy minimization stage and the equilibration period (again of 5 ns), a final simulation of 10 ns (intended for production) was completed.

To assure that the initial simulations were effectively performed in the liquid phase, for both F_7 and F_{10}, pressures of 5.0 bar (for T = 373.15K) and of 20.0 bar (for T = 473.15K) were used. It should be stressed that the first steps in the liquid phase are shared by both the liquid (obviously) and the gas phase simulations (as the initial part before moving on to the NVT ensemble).

It is not irrelevant the fact that the authors of the force field proposed fixed bond lengths, rather than providing a spring constant for the bond stretching. This constraint makes the simulation more likely to crash, and can rarely be used immediately after an energy minimization stage. Therefore, the equilibration periods were usually conducted without this constraint ("*all-bonds*", as it is named in the GROMACS language) and only then would this constraint be activated. This extra step was never needed in the liquid phase, and when done in the gas phase, it was always assured that all the different systems undertook the same uninterrupted equilibration period before moving on to production.

Figure 14 shows the coarse-graining of F_7 and F_{10}. Black lines represent the "pseudo-bonds" between the FE and FM beads, whilst the blue line symbolises the "pseudo-bond" between two FM beads. All bonds are meant to start and end in the centres of mass of the respective CG beads.

The idea behind DBI is to compute, based on the trajectory of every molecule during the simulation period, the centre of mass of each bead (that is why the mapping is treated as an input of the DBI). With the positions of two adjacent centres of mass, the bond length distribution can be computed; with the positions of three consecutive centres of mass, through the calculation of the bond vectors, an angle distribution can be obtained; and finally, with the positions of four centres of mass (as only happens for the F_{10}), a dihedral distribution can be worked out (this angle is measured by the intersection of the planes defined by bonds 1-2 and 2-3 and bonds 2-3 and 3-4).

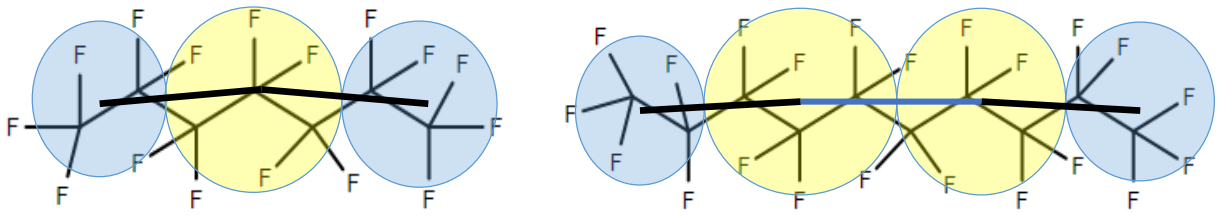


Figure 14 - Coarse-grained representations of F_7 (3 beads, left) and F_{10} (4 beads, right). Yellow circles represent the middle groups (FM: $-C_3F_6-$) and the blue ones stand for the terminal groups (FE: $-C_2F_5$).

Although the DBI provides a histogram of the frequency of each bond length (or bond angle, or torsional angle), these results still need to be converted into spring constants[10] and equilibrium values. Leach [16] offers an insightful discussion on the difference between the "reference" and the "equilibrium" values. While the reference (bond length, for instance) is the value adopted by an isolated molecule,

[10] To be more rigorous, spring constants are only the constants related with the bond stretching and the angle bending potentials. For the torsional potential, there is what can be called a *set of constants* (often 5), which are not spring constants for they do not promote a Hookean Law.

when no other forces are applied, the equilibrium value reflects the minimum energy configuration in which all the other interactions (in that case, the angle bending, the torsions, and even electrostatics and van der Waals') are taken into account. It may well happen that the equilibrium value is slightly different from the reference value if it promotes, overall, a reduction in the total energy of the system.

Apart from the torsions, which are treated differently, this chapter focuses now on the bond stretching and angle bending phenomena. According to the functional form of the bonded potential, both interactions are described via a harmonic law; mathematically, it can be shown that a potential that follows such a law is equivalent to a distribution that follows a Gaussian Bell (what is usually called a Normal distribution). Therefore, one shall expect both Equations 42 and 43:

$$P(X = x) \sim \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (42)$$

$$U(x) \sim \frac{k_x}{2} (x-\mu)^2 \quad (43)$$

where *X* stands for any measured variable (be it bond length or bond angle), μ for the average value, σ for the standard deviation and k_x for the spring constant. The distributions and the associated potentials are interconnected by Equations 44 and 45 [80]:

$$U_{bond}(r) \propto -k_B T \ln[P(r)/r^2] \quad (44)$$

$$U_{angle}(\theta) \propto -k_B T \ln[P(\theta)/\sin\theta] \quad (45)$$

where $P(r)$ and $P(\theta)$ are the bond length and angle distributions, respectively; and $U_{bond}(r)$ and $U_{angle}(\theta)$ are the bond stretching and angle bending potentials, respectively. The reader may have noticed the fact that both distributions are being divided by either r^2 or $\sin\theta$; this is due to the change of coordinates from Cartesian to spherical, a change that requires the inclusion of the Jacobian (confront with Equation 40). Finally, it can be proven that the spring constants and the standard deviation of the distribution are related via Equation 46[11].

$$k_x = \frac{RT}{\sigma^2} \quad (46)$$

For the sake of conciseness, and because the charts are very similar between each other, only the angle and bond length potentials are presented for perfluoroheptane, at 373.15K and in the liquid phase (see below Figure 15).

[11] The reader should be very careful with this equation. Writing the potential using *half of the spring constant* has physical roots on the equipartition of energy; however, many authors prefer to use only what may be named a $k' = k_x/2$. Regardless of preferences, a careful analysis of the supporting equations is absolutely critical before comparing different results. Throughout this thesis, the spring constants are always shown in the first format.

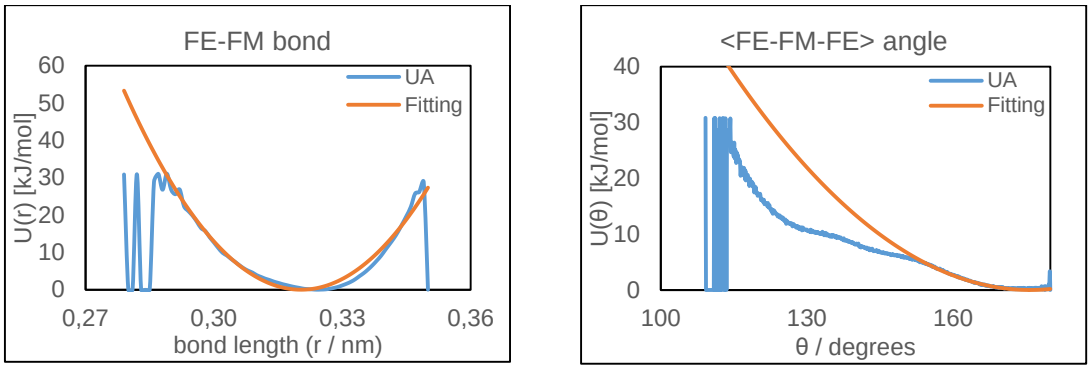


Figure 15 - Bond stretching (for bond FE-FM) and angle bending (for angle FE-FM-FE) potentials for F_7, at 373.15K, in the liquid phase.

It should be noted that the fitting of the <FE-FM-FM> angle bending potential is just accurate from around 150 to 180 degrees. One might ask whether or not the poor depiction of the UA potential for smaller angles may compromise the overall description of the entailed phenomena. Nonetheless, the major principle behind the accuracy of the harmonic potential is that the "spring", whatever it is, spends most of its time in the close surroundings of the equilibrium value. It is not critical, then, if the fitted potential overestimates the energy barrier to access angles between 110 and 150 degrees for the molecules will most often exhibit angles larger than 150 degrees, and most likely much closer to the equilibrium value (around 180 degrees, as will be seen later). Any attempt to accurately cover the whole curve would result in a worse representation of the essential region around the equilibrium value, where the potential is closer to zero.

Now that the procedure of the DBI has been detailed and some of its idiosyncrasies have been highlighted, the results for both the F_7 and the F_{10} are shown below. For the sake of simplicity, the units of all the parameters are only detailed in Table 1.

Table 1 - Intramolecular parameters for F_7 from atomistic simulations.

F_7	T/K	r(FE-FM) / nm	k_{bond} / kJ.mol^{-1}.A^{-2}	<FE-FM-FE> / deg	k_{angle} / kJ.mol^{-1}.rad^{-2}
Liquid	373.15	0.320	622.55	175.85	68.61
	473.15	0.320	652.12	179.18	60.65
Vapour	373.15	0.321	573.69	171.90	109.30
	473.15	0.319	612.88	178.88	65.66*
Averages		**0.320**	**615.31**	**176.45**	**76.06**

Table 2 - Intramolecular parameters for F_{10} from atomistic simulations.

F_{10}	T/K	r(FE-FM)	k_{bond}	r(FM-FM)	k_{bond}	<FE-FM-FM>	k_{angle}
Liquid	373.15	0.320	636.03	0.370	508.16	179.50	48.03
	473.15	0.319	627.02	0.370	534.07	181.44**	49.48
Vapour	373.15	0.320	639.30	0.370	508.24	178.27	55.82
	473.15	0.321	621.40	0.370	508.93	179.50	55.69
Averages		**0.320**	**630.94**	**0.370**	**514.85**	**178.96**	**52.26**

The consistency of the results does not come as a surprise, and is in total agreement with the qualitative understanding of the fluorinated chains. One shall notice:

1) How close the spring constants for the FE-FM bond stretching calculated for the two compounds are; and also how comparatively smaller is the FM-FM bond stretching spring constant, which was expected given the tendency of gradually larger CG groups being increasingly softer; in addition, these values are not only one order of magnitude larger than those obtained previously for the alkanes (thus contradicting the guess, in the referred paper, to use the same spring constants for alkanes and perfuoroalkanes) but are also so high that the bonds could be almost taken as fixed ones (though they were not);
2) How the angles are remarkably close to the original guess of 180 degrees suggested by Morgado *et al.*; besides, it was expected that the angle <FE-FM-FM> was larger than the <FE-FM-FE>, since it is harder for larger, more voluminous groups to be arranged in a non-straight line.

The reason why UA bond angles of 114 degrees are compatible with CG bond angles of almost 180 is not trivial. Figure 14 shows how a carbon chain can be actually treated as a straight line of CG groups, provided that these groups are large enough to integrate out the bond orientation in every carbon atom. Actually, in the DBI procedure, all the calculations are based on the position of the centres of mass – and one can easily see that the centres of mass look quite aligned. As an extra information, atomistic simulations of fluorocarbons using the former set of beads (that is, using the CG groups CF_3 and C_2F_4) shown that the angle formed by the sequence $CF_3 – C_2F_4 – CF_3$ was around 160 degrees – supporting the idea of a "volume threshold" of the fluorinated beads for the pencil-like shape to be actually verified – to put it another way, a minimum bead volume above which the bond angle is very close to 180 degrees.

One entry in each of the tables was marked with a star. In the first table, it is noticeable that the k_{angle} for the vapour phase at 473.15K is much closer to the values for the liquid phase rather than to its vapour phase counterpart at 373.15K. Such apparent deviation is motivated by the proximity to the critical temperature (which, for F_7, is reported to be 475.3K). Even though the theoretical predictions (using the SAFT-γ Mie) slightly overestimate this value, the vapour phase that was simulated resembled much more a supercritical fluid (with all its similarities with the liquid phase behaviour) than with a "common" vapour phase, so to speak. It is arguably a matter of debate whether or not should any of the values for the vapour phase be included in the computed average. To neglect the 109.30, because it diverges too much from the others, would lead to a loss of information about an actual vapour phase (not a "vapour-almost-supercritical" one), something especially important for NVT simulations, in which vapour-liquid equilibria is formed; but to neglect the 65.66, alleging that it does not purvey an accurate value for the vapour phase, would disregard the proximity between the working temperatures reported in the literature (30ºC – 90ºC) to the critical temperature, thus biasing an analysis and making it more difficult to follow the same procedure for perfluorodecane. Therefore, even though this is not a straightforward call, the average was evaluated using all the four values.

The starred entry on Table 2 draws attention to the equilibrium value obtained in the fitting is slightly higher than 180 degrees. Though it may seem impossible, it cannot be ignored that for such a chain molecule there is what may safely be termed a symmetry axis around 180 degrees (i.e. a bond angle of + 240º is equivalent to an angle of + 120º). Hence, the value is not only vested of physical meaning but must also be included in the computation of the average as long as, instead of 181.44º, 360 – 181.44 = 178.56º is used (to keep the reasoning coherent).

At this point, the intramolecular component of the force field is able to rigorously describe both hydrocarbon and fluorocarbon chains. However, since this thesis equally aims to semifluorinated alkanes, parameters for the "linker" – the CG group ($CF_2 – C_2H_4$) that bridges both hydrogenated and fluorinated chains – are also necessary.

In light of that, two approaches were considered. The first one would resemble deeply what has been done for the purely fluorinated molecules: different compounds would be selected and simulated at different temperatures and physical states, and an overall average would be computed. This procedure would lead to the development of an individual set of parameters just for the linker. It ought not to be forgotten, though, that since the range of SFAs reported in the literature varies widely from 6 to 30 atoms [43, 81], a systematic analysis of short and long SFAs would need to be undertaken.

Another option would attempt a description the intramolecular motion of the linker exclusively based in the parameters obtained for either the alkanes or the perfluoroalkanes. Undeniably, there is physical support for such a description: semifluorinated alkanes have been interpreted as *chemically bonded mixtures* of alkanes and perfluoroalkanes (see, for instance, [7, 82]). Creating, thus, a potential that was purely based on the parameters of both hydro- and fluorocarbons would add physical meaning to the force field at the expenses of a more rigorous approach that dealt uniquely with semifluorocompounds. Moreover, it would allow for a less exhaustive recollection of SFAs as only a qualitative agreement (with one or another family of compounds) would be required.

Taking everything into account, and not ignoring the time limitations entailed in a master thesis, the second methodology was pursued. For that purpose, simulations in the liquid phase of F_6H_6 (built as (C_2F_5)-(C_3F_6)-$(CF_2C_2H_4)$-(C_4H_9)), at 373.15K and 473.15K, were performed. The vapour phase was not studied because systems involving mixtures of SFAs with other solvents (the final targets of this research) are *mostly*, as far as the author is aware of, in the liquid phase (mainly because the longer SFAs, more used as surfactants, have a melting point close to room temperature). The results are exhibited in Tables 3 and 4. As done for tables 1 and 2, units were omitted from Table 4.

Table 3 - Intramolecular parameters for F_6H_6 – bond stretching.

T/K	r(FH-FM) / nm	k_{bond} / kJ.mol^{-1}.A^{-2}	r(FH-C4) / nm	k_{bond} / kJ.mol^{-1}.rad^{-2}
373.15	0.308	356.50	0.450	36.40
473.15	0.307	410.82	0.445	46.01
AVERAGES	**0.308**	**383.66**	**0.447**	**41.21**
Alkanes	-	61.30	-	**61.30**
Perfluoroalkanes	-	**514.85**	-	514.85

Table 4 - Intramolecular parameters for F_6H_6 – angle bending and torsional barriers.

T/K	<FE-FM-FH>	k_{angle}	<FM-FH-C4>	k_{angle}	$\Delta U_{dihedral}$ / k_BT
373.15	180.30	62.22	157.29	27.07	~ 0.8
473.15	179.98	65.09	157.12	31.37	~ 0.8
AVERAGES	**179.84**	**63.66**	**157.21**	**29.19**	**~ 0.8**
Alkanes	159.90	17.66	**159.90**	**17.66**	-
Perfluoroalkanes	**178.96**	**52.26**	178.96	52.26	-

For the subsequent comparison, motivated by simplicity concerns, the notation H-bead and F-bead shall be adopted for purely hydrogenated and fluorinated CG groups, respectively.

By confronting with the last rows, which recall the intramolecular parameters for both alkanes (taken from [43]) and PFA, it can be extrapolated that:

1) Bonds (linker – F-bead) behave as F-bead – F-bead bonds in terms of stiffness (assessed via k_{bond});
2) Bonds (linker – H-bead) behave as H-bead – H-bead bonds in terms of stiffness;
3) Angles (F-bead – F-bead – linker) behave as angles in a perfluorocarbon chain, both in arrangement (angle) and "flexibility"[12] (k_{angle});
4) Angles (F-bead – linker – H-bead) behave as angles in a hydrocarbon chain, both in arrangement (angle) and "flexibility" (k_{angle});
5) *It will be assumed that* angles (linker – H-bead – H-bead) behave as angles in a hydrocarbon chain as well;
6) Torsional barriers may be neglected, since the largest barrier of the whole torsional profile ($\Delta U_{dihedral}$) is smaller than $0.8\ k_BT$.

The intramolecular part of the force field has been finally concluded. All in all, a *bottom-up* methodology was followed to quantify the bonded interactions between CG groups from UA simulations. Whilst the parameters for hydrocarbons had already been proposed [42], further computational study of fluorocarbons was stalled because this concrete information was lacking for perfluoroalkanes. This research has also proposed, via a different approach, intramolecular parameters for the linker, yet recognising the fact that it is limited to the very group previously designated (CF_2 – C_2H_4).

2. Intermolecular Parameters

Bottom-up approaches offer a legitimate solution to concerns about the structure. Their ability to reproduce, to a certain extent (depending on the particular method applied), the structure purveyed by more detailed models alleviates the fears of a completely misunderstood molecule. Nonetheless, no matter how accurate the atomistic force field, there is little guarantee that thermodynamic properties will be correctly predicted – and the phase envelope of a substance be captured – if no macroscopic data

[12] Perhaps "flexibility" is a rather loose term; it means, in this context, the ability of two consecutive bonds to bend and approach, reducing the angle formed between them. Intuitively, the higher the spring constant, the more difficult it is for the bonds to move apart from the equilibrium value.

(from experimental or computational sources) is used in the modelling of the force field. The progressive loss of detail entailed in the upwards move in the simulation ladder (Figure 3) prevents the predictive capabilities to be sustained as the time and length scales increase.

From the hybrid technique developed already in previous SAFT-γ papers [42, 83], the intermolecular interactions are described via the framework embedded in this novel group contribution theory. Experimental data is used as input, obviously along with the mapping (definition of CG groups), and the combining rules, if that is the case. Mie potential parameters are the final goal of this process, executed with the software gSAFT®.

On the one hand, the CG groups are the following:

Table 5 - CG group definition (the mapping*). Adapted from the table presented in [43].*

Name	All-atom
FE	C_2F_5 –
FM	– C_3F_6 –
FH	– $CF_2C_2H_4$ –
CE	C_3H_7 –
CM	– C_3H_6 –
C4	C_4H_9 –

On the other hand, the targeted parameters are:

Table 6 - Targeted parameters in the SAFT-γ Mie modelling.

Like Interactions	Unlike Interactions
• σ	• cross σ
• λ_{rep}	• cross λ_{rep}
• ε	• cross ε

The need to evaluate cross parameters motivates the use of combining rules, of which the Lorentz-Berthelot (LB) comprise some of the most widely used (Equations 47.1 and 47.2). As will be explained, these were used carefully and selectively; again, it should be underscored that the attractive exponent is not a variable – instead, it is always fixed at 6.

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \quad (47.1)$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \quad (47.2)$$

The validity of the combining rules has motivated several studies (for instance, [84, 85], just to name a few), with some actually focused on the application of such expressions to mixtures of alkanes and perfluoroalkanes [86]. The reader is referred to the cited papers for an extensive analysis; here, it should suffice to clarify in which conditions such rules can be safely used.

The SAFT-γ Mie framework also includes a set of combining rules for the unlike parameters in mixtures of chains formed from Mie segments [87] (which is the case of the present work). Besides the arithmetic average for the size parameter (which is usually taken without modifications), Lafitte *et al.* proposed Equation 48.1 (for the cross repulsive exponent) and Equation 48.2, with k_{ij} is set to zero, for the cross $\boldsymbol{\varepsilon}$.

$$\lambda_{ij} - 3 = \sqrt{(\lambda_{ii} - 3)(\lambda_{jj} - 3)} \quad (48.1)$$

$$\varepsilon_{ij} = \left(1 - k_{ij}\right) \frac{\sqrt{\sigma_{ii}^3 \sigma_{jj}^3}}{\sigma_{ij}^3} \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \quad (48.2)$$

It is relevant indeed to understand the presence of the parameter k_{ij} in the combining rule. If that very expression was used to obtain the cross depth potential, a procedure to calculate k_{ij} would be required *a priori*. Whilst some authors have optimized a specific k_{ij} to apply over several similar interactions [79], the SAFT-γ Mie framework links directly the determination of the cross $\boldsymbol{\varepsilon}$ with the thermodynamic data provided as input, in which case that expression is not actually used – rather, the cross $\boldsymbol{\varepsilon}$ itself is optimized at once.

Because of the known non-ideality of alkane – perfluoroalkane mixtures, it has been shown that applying Berthelot-like combining rules (grounded on the geometric average of the depth) to these mixtures yields severe failure to predict mixed properties such as mixed second virial coefficients, excess energetics, solubilities of gaseous PFA in liquid alkanes, vapour-liquid equilibria diagrams, among others [79, 86]). For such interactions, there are clear benefits from estimating directly the cross depth from experimental data where both hydro- and fluorocarbons coexist. That being said, the LB combining rules – even that for the energetics, deemed more “problematic” – are accurate for ideal mixtures, where the like interactions between the two components are similar (and thus similar to the cross interactions). They just fail for these mixtures because, informally, they assume that the hydrocarbon – fluorocarbon intermolecular forces should be “strong” *because* those between two hydrocarbons and between two fluorocarbons are as “strong”. It is this premise, formalized in Equation 47.2 (or even in Equation 48.2 with k_{ij} set to zero), that justly requires a novel approach to estimate the cross interactions.

In a nutshell, combining rules must be used responsibly, in a way that acknowledges their limitations in the description of the specific mixtures this thesis deals with. The performed optimization used, when legitimate, Equations 47.1, 48.1 and 48.2 (with k_{ij} set to zero) as combining rules. Bearing everything said in mind, Table 7 resumes how the different parameters are obtained.

Table 7 - Optimization procedure – the role of combining rules and experimental data in parameter estimation.

Parameters	CR used to calculate	Exp. data used to estimate
All like interactions: • every σ • every λ_{rep} • every ε All unlike interactions • every cross σ • every cross λ_{rep} • every cross ε	• every cross σ • every cross λ_{rep} • cross ε between two hydrocarbon groups or two fluorocarbon groups	• All like interactions parameters • cross ε between an hydrocarbon and a fluorocarbon groups • every cross ε involving the linker ($CF_2 – C_2H_4$)

To conclude this short comment, it may be enlightening to quote Potoff's cited paper: "*For unlike molecules, it may be advantageous to use different combining rules to improve the agreement of simulation with experimental data. However, there is no way to know a priori which combining rules will give the best results*". Again, it is from the ability of the theory to correctly predict several thermodynamic properties, for both pure components and mixtures, that one may infer its usefulness and accuracy. The several steps undertaken during the optimization of the parameters just underscore how difficult it is to strictly compare two theories, even if generally based on the same theoretical background.

Since the fluorinated groups were changed, new like parameters are needed, as well as all the cross interactions in which they play a role. Moreover, because the linker is the last group to be modelled, its parameters also need a revaluation – a joint effort of no less than 20 (!) parameters to determine.

The risks of such a multivariable modelling are well-known: the purely mathematical fitting may accomplish its goal at the expenses of reasonable values for the physical parameters, not to speak about the redundancy that becomes more likely as the number of variables increase. The procedure of inputting all the experimental data available and all the parameters proved unfeasible – not because of the optimization tools, which returned the desired parameters, but because of the lack of physical meaning that some of them presented. To bypass this difficulty, a sequential modelling was undertaken, as described in Table 8.

Table 8 - Successive steps undertaken during the sequential modelling, with the experimental data used in each step.

Steps	Targeted Parameters	Experimental Data Used	Reference
1.	• FE: σ, λ_{rep}, ε • FM: σ, λ_{rep}, ε	Saturated Liquid Densities and Vapour Pressures of F_4 and F_7 in the range $0.5T_c$ – $0.9T_c$	[88]
2.	• FE – CM: ε • FE – CE: ε • FE – C4: ε • FM – CM: ε • FM – CE: ε • FM – C4: ε	Vapour – Liquid Equilibria (composition of both phases) at constant temperature and pressure of: $F_5 + H_6$, $F_6 + H_5$, $F_6 + H_6$, $F_6 + H_7$, $F_6 + H_8$, $F_7 + H_6$, $F_8 + H_6$.	[89]
3.	• FH: σ, λ_{rep}, ε • FH – FE: ε • FH – FM: ε • FH – CE: ε • FH – CM: ε • FH – C4: ε	Saturated Liquid Densities and Vapour Pressures of F_4H_5, F_4H_6, F_4H_8, F_6H_6; Saturated Liquid Density of F_6H_8.	[7, 82]

First, the parameters for the perfluorinated groups (FE and FM) are estimated; then, once these are known (and since those for the alkanes had been previously determined), mixture data is exploited to model the cross interactions between hydrogenated and fluorinated groups (notice that the interactions between two fluorinated groups and two hydrogenated groups were already evaluated through the combining rules); finally, thermodynamic bulk data for SFAs is used to discover not only the like parameters for the linker but also every cross ε between it and the remaining groups.

As a matter of fact, the linker could not be considered neither a purely fluorinated nor a purely hydrogenated CG bead, hence the estimation of every cross ε. Besides, treating the linker as a specific bead allows to place the permanent dipole exhibited by PFAs in one bead, in a somehow *effective* way. Although bulk data for PFAs was selected in the range $0.5T_c$ – $0.9T_c$ (which is a rather common choice) for the modelling of intermolecular interactions, due to scarce experimental data, the semifluorinated alkanes modelling was restricted to the shorter-range data available.

The last step, surely due to the number of parameters involved, was impossible to perform at once, that is, providing all the input data and obtaining all the desired outputs. In fact, when it was attempted to optimize in a single step all the eight parameters, some of the cross ε lied on physically meaningless regions (such as around 50K or below, when common values range between 300K and 450K). Therefore, and since all the SFAs were not built with every single CG group, a so-called “partial optimization” was undertaken, in which three or four parameters were optimized while the others were fixed; then, the former were fixed in the newly found values and the latter were optimized; and so on,

until the objective function barely changed. The author is aware that this is not a perfect procedure, but trusts it as the best possible choice given the singularities of the system.

The objective function (OF) changed depending on the optimization step. In any case, though, the implemented OF does not value more one property over other; instead, it treats equally any experimental point – implicitly weighting more the properties for which more data is available.

For the first and third steps, where bulk data of pure components was used, the optimization was given by the minimization of the function $F(\alpha)$, as follows in Equation 49:

$$\min_{\alpha} F(\alpha) = \min_{\alpha} \sum_{i=1}^{N_\rho} \left(\frac{\rho_L(T_i, \alpha) - \rho_L^{exp}}{\rho_L^{exp}} \right)^2 + \sum_{k=1}^{N_P} \left(\frac{P_{vap}(T_k, \alpha) - P_{vap}^{exp}}{P_{vap}^{exp}} \right)^2 \quad (49)$$

where α is the vector containing all the SAFT parameters (every pertaining σ, ε and λ_{rep}); T is the temperature; ρ_L is the saturated liquid density; P_{vap} is the vapour pressure; and N_ρ and N_P are the number of experimental points for the saturated liquid density and for the vapour pressure, respectively.

For the second step, which involved agreement with the vapour phase composition at fixed temperature and pressure, the objective function $G(\alpha)$ was implemented. Equation 50 formalizes that second step.

$$\min_{\alpha} G(\alpha) = \min_{\alpha} \sum_{i=1}^{N_\rho} \left(\frac{y(T_i, P_i, x_i) - y^{exp}}{y^{exp}} \right)^2 \quad (50)$$

The goal of transferability is central in the context of SAFT-γ Mie. As a group contribution theory, it attempts to describe families of compounds using the functional groups upon which they are built. The more compounds are used in the modelling of such functional groups, the more likely it is for the theory to cover accurately the whole targeted family of compounds instead of just those inputted. Thereby, namely for the vapour-liquid equilibria, many molecules that could not be built with integer beads were nevertheless used, in an effort to widen the basis of compounds inputted and thus the potential transferability of the model.

The final results for the parameters are shown on Table 9.

Table 9 - Parameters for the like (on the left) and unlike interactions (on the right).

Like Interactions				
MW (g/mol)	**Group**	**σ / Å**	**ε / K**	**λ_{rep}**
119.01	FE	4.637	322.14	25.29
150.02	FM	4.739	380.80	22.26
78.06	FH	4.533	330.09	17.11
42.08	CM	4.184	377.14	16.43
43.09	CE	4.501	358.37	15.95
57.11	C4	5.001	473.62	24.00

Unlike interactions				
		σ / Å	**ε / K**	**λ_{rep}**
FE	FM	4.688	350.18	23.72
FE	FH	4.585	393.13	20.73
FE	CM	4.410	349.26	20.30
FE	CE	4.569	322.03	19.99
FE	C4	4.819	348.52	24.64
FM	FH	4.636	302.86	19.49
FM	CM	4.461	355.34	19.08
FM	CE	4.620	341.12	18.79
FM	C4	4.870	420.74	23.11
FH	CM	4.358	283.30	16.77
FH	CE	4.517	331.81	16.52
FH	C4	4.767	397.72	20.21
CM	CE	4.343	366.90	16.19
CM	C4	4.593	417.63	19.79
CE	C4	4.751	410.27	19.49

IV - Application of the New CG Force Field

With the force field defined in the previous chapter (which took into account the redefinition of the fluorinated beads), its performance can now be assessed. The following sections will cover: theoretical calculations of thermodynamic properties (of pure compounds and mixtures); prediction of interfacial properties through MD simulations; and an analysis of interfacial behaviour and self-assembly of mixtures of PFAAs in hydro- and fluorocarbon solvents.

1. SAFT-Calculation of Thermodynamic Properties

Following the structure of a paper of Papaioannou *et al.* [58], a comparison between theoretical predictions and experimental data is presented, covering compounds both included and not included in the modelling. Starting by bulk properties of fluoroalkanes and semifluorinated alkanes (the work for alkanes has been already done in [58], and is not the focus of this thesis) such as vapour pressure and saturated liquid densities, then moving on to critical properties and vapour-liquid and liquid-liquid equilibria (LLE), this section aims to underscore the performance of the newly proposed CG model. The software gSAFT® is used to make the theoretical calculations. These computations do not require the intramolecular potential – in fact, the latter is only applied when molecular simulations are performed. The information is presented in tables, and not in charts, for the sake of compactness.

The agreement with experimental data is evaluated through the *Average Absolut Deviation* – $AAD(\%)$ –, defined as follows in Equation 51:

$$AAD(\%) = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{X_i^{pred}(T) - X_i^{exp}}{X_i^{exp}}\right| \quad (51)$$

where N is the total number of data points or compounds analysed, X is the property of interest, X_i^{exp} is the experimental value and X_i^{pred} is the predicted value (with SAFT).

Table 10 summarizes the agreement with experimental data achieved for the vapour pressure and saturated liquid density of both perfluoro- and semifluorinated alkanes.

Table 10 - Theoretical predictions for the vapour pressure and saturated liquid density of fluoroalkanes and SFAs.

	Vapour Pressure					Saturated Liquid Density				
	T range (K)	N	$AAD(\%)$	Ref.	Comments	T range (K)	N	$AAD(\%)$	Ref.	Comments
F_4	189 - 349	33	0.36%	[88]	Input in modelling	189 - 349	33	1.45%	[88]	Input in modelling
F_5	*Not available*				-	273.15 - 293.15	5	0.45%	[7]	Prediction
F_6	256.43 - 447.08	20	1.00%	[90]	Prediction	278.15 - 323.15	10	0.53%	[7]	Prediction
F_7	280 - 425	30	0.57%	[88]	Input in modelling	280 - 425	30	0.82%	[91]	Input in modelling
F_8	310.47 - 378.91	17	0.58%	[88]	Prediction	278.15 - 353.15	16	1.37%	[7]	Prediction
F_9	288.18 - 333.15	10	1.93%	[92]	Prediction	283.15 - 353.15	15	1.63%	[7]	Prediction
F_4H_5	278.02 - 327.78	21	0.69%	[93]	Input in modelling	278.15 - 353.15	16	0.81%	[7]	Input in modelling
F_4H_6	278.12 - 327.90	21	1.85%	[93]	Input in modelling	278.15 - 353.15	16	0.57%	[7]	Input in modelling
F_4H_8	297.98 - 327.97	13	1.40%	[93]	Input in modelling	278.15 - 353.15	16	0.23%	[7]	Input in modelling
F_6H_6	288.06 - 327.93	17	3.94%	[93]	Input in modelling	273.15 - 353.15	17	0.36%	[82]	Input in modelling
F_6H_8	*Not available*				-	273.15 - 353.15	17	0.48%	[82]	Input in modelling

Theoretical predictions for compounds not used in the modelling are, overall, in good agreement with experimental data.

For the perfluorocarbons, an assessment of the critical properties was also carried out. It is noteworthy that critical data is not used in the modelling, since experimental measurements of both F_4 and F_7 (which modelled both groups FE and FM) were in the range $0.5T_c$ – $0.9T_c$. Critical properties include critical temperature, critical density and critical pressure – the results are shown below on Table 11. Experimental values are taken from [94].

Table 11 – SAFT-calculation of critical properties of perfluoroalkanes and comparison with experimental data.

	T_c / K			d_c / kg.m-3			P_c / MPa		
Compound	Exp.	Prediction	%Dev.	Exp.	Prediction	%Dev.	Exp.	Prediction	%Dev.
F_4	386.4	390.7	1.1	630	663.0	5.2	2.314	2.583	11.6
F_5	421.4	424.3	0.7	622	648.0	4.2	2.037	2.233	9.6
F_6	450.6	453.3	0.6	621	624.3	0.5	1.877	1.947	3.7
F_7	475.3	478.8	0.7	619	596.2	3.7	1.62	1.717	6.0
F_8	498.2	501.7	0.7	611	567.0	7.2	1.548	1.531	1.1
F_9	524.0	522.5	0.3	-	539.7	-	1.560	1.380	11.5
F_{10}	542.4	541.5	0.2	-	515.3	-	1.450	1.255	13.4
Averages			**0.6**			**4.2**			**8.1**

As Table 11 displays, agreement with critical temperature is achieved with high accuracy, even though SAFT-type equations of state are generally deemed to be unreliable in the prediction of pure components and mixtures critical data [95]. The predictions for critical densities and pressures are not as good, though the exponential dependence of the latter with the temperature has inevitably a multiplying effect on the deviations (shortened for "*%Dev.*").

It should be stressed that, even though Morgado *et al.* [43] proposed the first CG model recently, no thorough study of its predictive capabilities had been completed, thus hindering any quantitative comparison between these new force field and the previous (noting that the only change brought by this work concerned the fluorinated groups, which were enlarged).

The theory could be also used to predict mixture properties, especially difficult to forecast for the non-ideal alkane-perfluoroalkane mixtures. For vapour-liquid equilibria, the results are shown in Table 12. The source of experimental measurements is [89].

Table 12 - Theoretical predictions for the vapour-liquid equilibria of alkanes - perfluoroalkanes mixtures. All the VLE data presented was used as an input in the modelling.

	T/K	x(mol) range	*n*	average |dy| (%)
F_5+H_6	293.15	0.035 - 0.955	15	2.73
F_6+H_6	298.15	0.026 - 0.972	16	4.24
F_6+H_7(C4)	317.65	0.038 - 0.953	11	0.99
F_6+H_7(CM)	317.65	0.038 - 0.953	11	3.39
F_6+H_8(C4)	313.15	0.024 - 0.979	14	0.89
F_6+H_8(CM)	313.15	0.024 - 0.979	14	1.66
F_7+H_6	303.15	0.021 - 0.985	18	4.43
F_8+H_6	313.15	0.049 - 0.955	10	3.01
F_8+H_7(C4)	298.15	0.020 - 0.969	10	2.98
F_8+H_7(CM)	298.15	0.020 - 0.969	10	3.36
Average				2.77

At this point, perhaps it is useful to recall the interactions that the vapour-liquid equilibria sought to describe. Having two fluorinated (FE and FM) and three hydrogenated groups (CE, CM and C4), 6 cross depths were needed. However, since both CE and C4 are terminal beads, they are often not used simultaneously to build the same molecule. For that reason, some hydrocarbons were defined with two different structures that were, each at a time, used in the prediction of the vapour-liquid equilibria. For instance, even though the most natural definition of the *n*-octane is C4 – C4, because the theory allows for non-integer beads, it could also be coarse-grained via CE – ⅔ CM – CE. The same thing can be done with the n-heptane (either a more natural CE – C4 or a CE – ⅓ CM – CE). As such, where in Table 12 is written "H_7 (C4)" it is meant that the n-heptane was defined in the first way listed.

Finally, the agreement assessment was not made with the usual $AAD(\%)$ but with what was called "average |dy| (%)". In fact, the $AAD(\%)$ is a relative measurement which penalizes small quantities. If, for the vapour phase content, the experimental value was y=0.10 and SAFT predicted y=0.09, the $AAD(\%)$ would equal 10%; if, instead, that very composition was y=0.90 and SAFT advanced y=0.89, the $AAD(\%)$ would be slightly higher than 1% - yet the theory would have been equally accurate in both cases. For that reason, an absolute calculation was preferred, given by Equation 52.

$$average\ |dy|(\%) = \sum_{i=1}^{N} \left| y_i^{pred}(T,P,x) - y_i^{exp} \right| \times 100\% \quad (52)$$

The expression above tells, essentially, what is the average deviation between the predicted vapour composition and the experimental value. Since the computed value was 2.77%, then SAFT predictions are, on average, for a vapour phase composition that is just 2.77% far from the experimental values. For such non-ideal mixtures, and given the known difficulty [96] that several force fields, even

more detailed ones, have faced to predict the multi-phase behaviour of these mixtures, this accuracy is considered acceptable.

As an example, some of the VLE charts (showing the dew point lines) are exhibited below, in Figures 16 and 17.

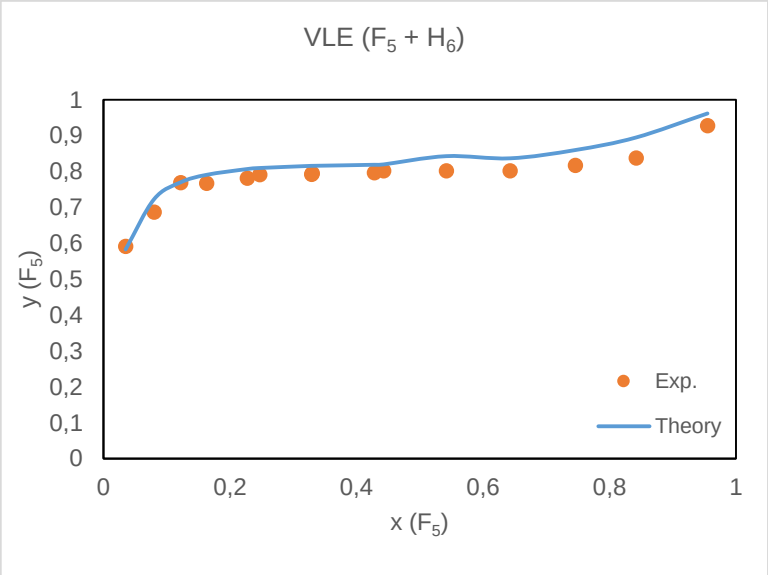


Figure 16 - Vapour-liquid equilibria for the F_5 + H_6 mixture, at 293.15K.

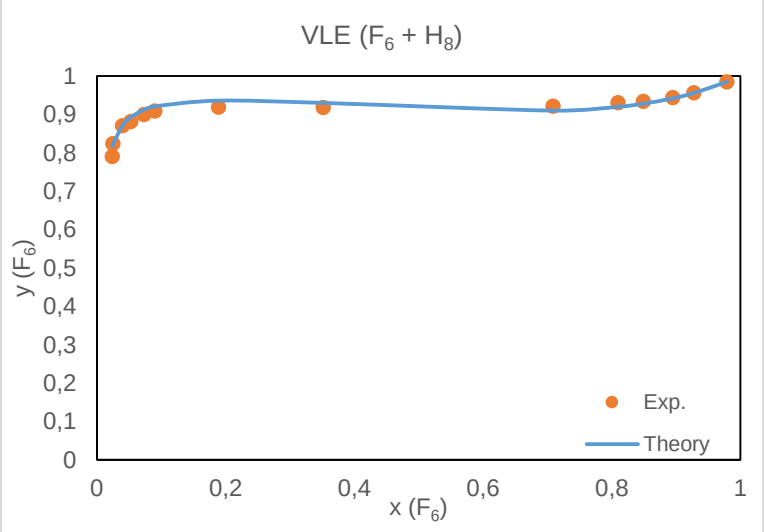


Figure 17 - Vapour-liquid equilibria for the F_6 + H_8 mixture, at 313.15K, with H_8 built as two C4 beads.

Finally, liquid-liquid equilibria was calculated. It is not as trivial to quantify deviations in LLE, so a more qualitative approach is undertaken. Matsuda *et al.* [96] report that, although the trends are shared by different pieces of research, an agreement of experimental data is not always possible because of the "*very difficult properties*" showed by these mixtures. In this section, the focus lies more on the ability of the theory to capture the UCST and the levelling verified around the critical concentration. Below are shown the results for four mixtures: F_6+H_6, F_6+H_7, F_6+H_8 and F_8+H_8.

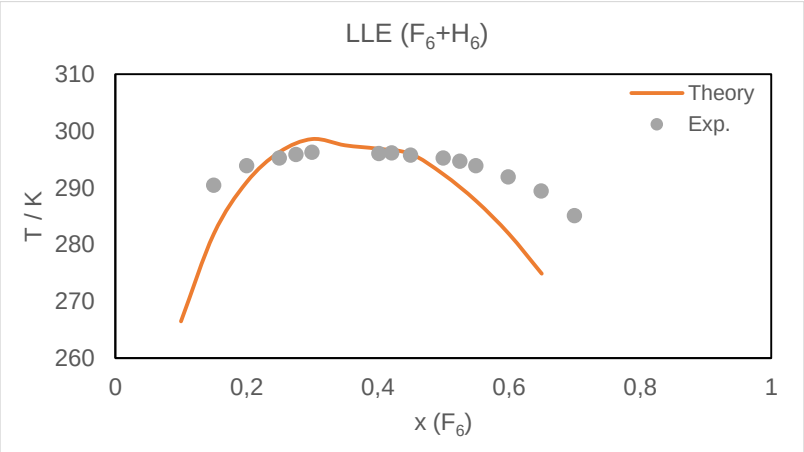


Figure 18 - Liquid-liquid equilibria for the F_6 + H_6 mixture. Experimental data taken from [96].

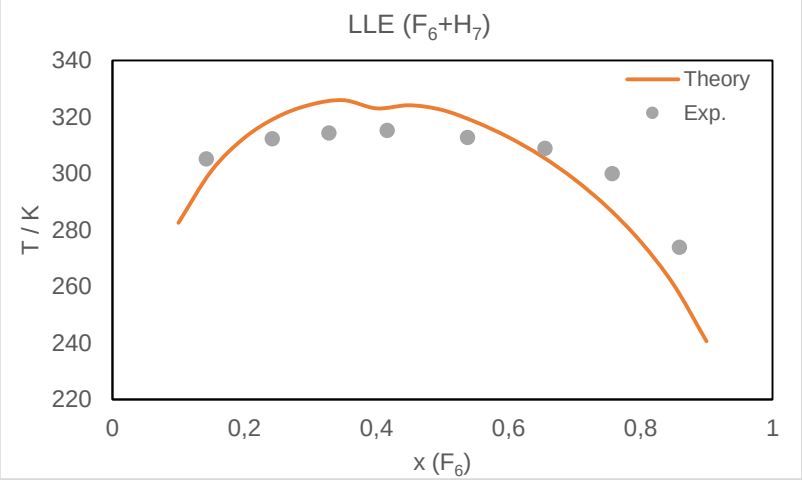


Figure 19 - Liquid-liquid equilibria for the F_6 + H_7 mixture. Experimental data taken from [89].

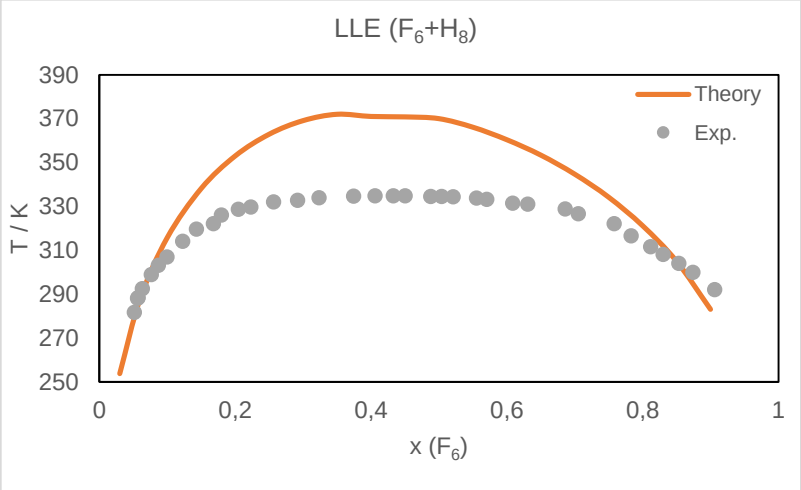


Figure 20 - Liquid-liquid equilibria for the F_6 + H_8 mixture. Experimental data taken from [96].

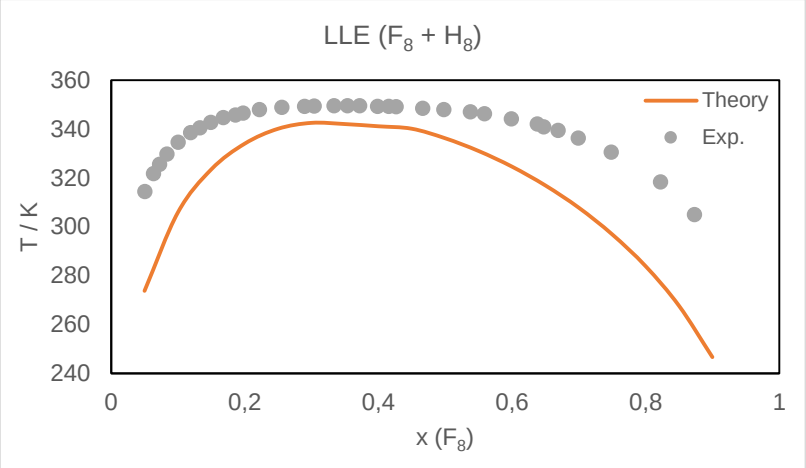


Figure 21 - Liquid-liquid equilibria for the F_8 + H_8 mixture. Experimental data taken from [96].

The Upper Critical Solution Temperatures (UCST) and the critical compositions are presented in Table 13, for further comparison.

Table 13 - Predictions and experimental data for the critical composition and critical temperature of alkane-PFA mixtures.

	Experimental		Theory		
Mixture	x_{UCST}	UCST	x_{UCST}	UCST	Reference
F_6+H_6	0.3621	296.41	0.30	298.6	[96]
F_6+H_7	0.416	315.2	0.35	325.9	[89]
F_6+H_8	0.4203	334.84	0.35	372.0*	[96]
F_8+H_8	0.3506	349.52	0.30	342.5	[96]
* Done at P = 5 bar to separate the LLE and VLE envelopes in the phase diagram, and bearing in mind the virtual independence of the LLE on the pressure.					

It should be stressed that the calculations via SAFT-γ Mie were done for compositions of multiples of 5% molar – the critical concentration declared is, hence, the composition for which the critical temperature was the highest, though it cannot safely be taken as a "true prediction" of the UCST. The same reasoning extends to the "theoretical" UCST – the highest temperature found for the tested compositions. For the temperature, however, the uncertainty caused by this procedure should not be a matter of concern since, around the UCST, the T – x curve tends to level.

Looking at the results obtained, the accuracy is far from consistent. Whilst the theory presents accurate estimates for mixtures of alkanes and perfluoroalkanes with the same number of carbon atoms, it fails for the F_6+H_7 and F_6+H_8 mixtures (with the deviations being particularly pronounced in the latter).

Theoretically predict LLE encompasses, in this case, two difficulties: in the first place, LLE data was not used in the modelling; secondly, liquid mixtures of alkanes and perfluoroalkanes are extremely non-ideal (the very high excess volume, of around + 5 cm^3 / mol, is paradigmatic), which adds to the complexity of the task.

Several research has been done in the modelling and prediction of the LLE of perfluoroalkanes and alkanes. Back in 1998, McCabe *et al.* [97] modelled these compounds (in CG beads) using the SAFT-VR approach (predecessor of the SAFT-γ Mie). Optimizing the unlike interaction parameter solely for the perfluorobutane + butane mixture, the authors managed to predict quite accurately the critical lines of liquid-liquid immiscibility perfluoromethane and *n*-alkane (*n* = 1, 2, 3). In 2005, Morgado *et al*. [8] modelled the phase behaviour of such mixtures using excess properties and UCST's instead (with data for hexane + perfluorohexane); this methodology allowed the authors to subsequently predict accurately the excess volume of other mixtures not used in the modelling. That set of parameters was later applied to VLE calculations, with equally reassuring results.

Back in 1964, Munson showed how the *T-x* charts for different fluorocarbon/hydrocarbon mixtures collapse in one single curve when the reduced temperature is plotted against the volume fraction of one component [98]. This discovery asserted some type of universal phase separation process (though this "universal behaviour" is still a matter of debate [99]), shared by all mixtures independently of the chain length of its components.

However, addressing the subtleties arising from mixtures of aliphatic compounds with different chain lengths, Matsuda *et al.* [96] commented on the much higher solubility of hexane in perfluorohexane than in octane – showing how equally long compounds tend to establish much stronger interactions than compounds with a different number of carbon atoms. What's more, as Lo Nostro pointed out [99], because the London forces (the sole type of interactions in these mixtures) are directly proportional to the polarizabilities, the critical temperatures of fluorocarbon/hydrocarbon mixtures increase almost linearly with the polarizability of the alkanes.

The theoretical predictions appear to overestimate the interactions between molecules with different chain lengths, thus overshooting the UCST. It could be the fact that both F_6 and F_8 are described with only two beads (and F_7 with 2⅓ beads) is a *too coarse* coarse-graining and, in this particular analysis, depicts too large molecules - consequently, more polarizable than it is in reality. Higher polarizability is related, as Lo Nostro *et al.* showed, with higher critical temperatures. However, the quest for the prediction of liquid-liquid immiscibility of alkanes and perfluoroalkanes has been solved with an even less sophisticated version of SAFT (SAFT-VR Square Well) [97] than the one supporting this research, with a coarse-graining *similarly coarser*, so to speak: hexane (2.67 spheres vs. 2 spheres in the present work); octane (3.33 vs 2); perfluorohexane (2.85 vs. 2.67) and perfluorooctane (3.59 vs. 3.33). The burden for the deviations does not seem to lie, therefore, on the coarse-graining.

Nonetheless, what McCabe *et al.* did in the referred paper was to fit the cross interaction energy parameter to the critical liquid-liquid and gas-liquid lines of the perfluoromethane + *n*-butane mixture. It is perhaps more likely that a fitting that takes critical lines into account helped to predict the critical lines of the mixtures. Fitting VLE data, though more abundant, may not support the model to predict the critical behaviour of mixtures as accurately as effectively fitting to the critical lines.

However, at this point, these are just guesses and how they contribute to the large deviations witnessed, if they actually do, is open to discussion, with more work needed to discern the causes of

this overestimation. After being used in the modelling procedure, VLE was accurately predicted; the fact that the VLE can be correctly tested and, at the same time, the LLE cannot shows that calling the "unlike interactions" one tends to unify for all mixture data may actually differ from a LLE to a VLE situation.

The question now arising of whether or not the modelling of the cross interactions should had been done using other experimental data – namely, LLE – plays, in general, a crucial role in any coarse-graining procedure. There is no guarantee, *a priori*, that modelling, say, property A will allow for a correct (pure) prediction of property B, and vice-versa. McCabe *et al.* were able to predict (though not as precisely as the LLE) the VLE of similar mixtures, above the UCST. Nonetheless, their study – unlike this thesis – focused on the high-pressure, near-critical region, and not at room pressures such as this thesis. It remains an open discussion on which modelling path may entail, globally, the best prediction capabilities – and if, in this particular case, other options (inclusively taking into account both LLE and VLE at the same time) could have brought a sounder set of parameters.

2. Simulations - Interfacial Properties Prediction

The applications of a force field are not confined to the prediction of bulk properties or even mixture behaviour. Through molecular simulation, dynamic properties and interfacial properties can be computed – in fact, these properties cannot be assessed from theory directly.

Surface tension has its roots in the energetic deficit faced by molecules at the surface of a liquid. While molecules in the bulk are completely surrounded by their counterparts, thus benefiting from a certain potential energy, molecules at the surface find themselves only "half-surrounded" by similar molecules (and "half-surrounded" by air/vacuum). As a consequence, an energetic imbalance between the bulk and the surface is established, thus creating a driving force for the molecules to reduce the surface area. Indeed, surface tension translates this differential.

It follows that the referred energetic imbalance becomes more relevant if the intramolecular interactions between molecules increase in value. For water, for instance, because of hydrogen bonds, the surface tension is very high (around 72 mN/m) [100]; for alkanes, which are only gathered by dispersion forces, surface tension stays between 20 and 30 mN/m [101]; and for perfluoroalkanes, that value is even lower – around 10 mN/m [4]!

In MD simulations, the surface tension can be evaluated via a mechanical route involving the pressure tensor [43]. For a box with two interfaces normal to the *Oz* axis, Equation 53 is applied.

$$\gamma = \frac{1}{2}\int_0^{L_z}\left(P_{zz} - \frac{1}{2}\left(P_{xx} + P_{yy}\right)\right)dz \approx \frac{L_z}{2}\left(\overline{P_{zz}} - \frac{1}{2}\left(\overline{P_{xx}} + \overline{P_{yy}}\right)\right) \quad (53)$$

where $P_{\alpha\alpha}$ are the Cartesian $\alpha\alpha$ components of the pressure tensor, $\overline{P_{\alpha\alpha}}$ is the average value of that very component during the simulation time and L_z is the *z*-dimension of the simulation box (or whatever dimension is set normal to the interfaces). In practice, it was confirmed that when simulations, ran in the NVT ensemble, are well equilibrated (and can thus be used to provide precise estimations of

surface tensions), the *xx* and *yy* components of the pressure tensor are very similar, and roughly one order of magnitude higher than its *zz* counterpart.

Unlike theoretical calculations, which can be done for a non-integer number of beads, molecular simulations are constrained by an integer number of segments. Because the fluorinated groups are C_2F_5 and C_3F_6, it is perfluoroheptane (C_2F_5 – C_3F_6 – C_2F_5, named F_7) the first compound that combines being liquid at room temperatures (normal boiling point of around 80ºC) and having experimental data available. Hence, molecular simulations of F_7 were performed, at different temperatures, with the results being shown on Figure 22. In addition, the surface tension of alkanes was tested with *n*-dodecane at 50ºC. The results of both alkane and perfluoroalkanes are gathered on Table 14.

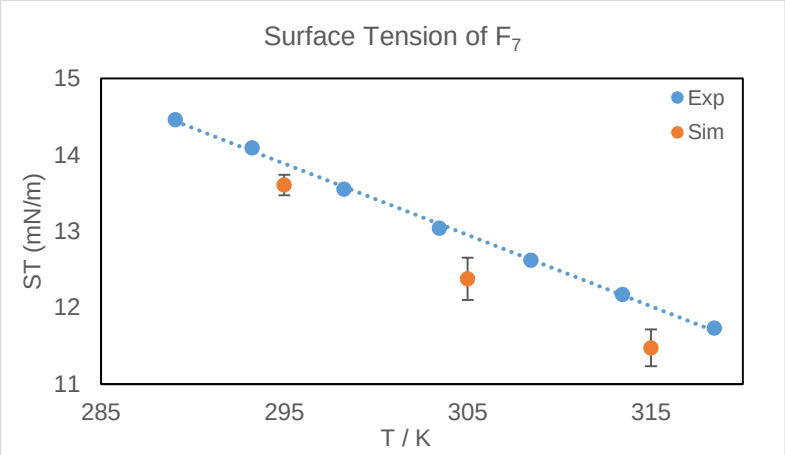


Figure 22 - Surface tension (ST) of F_7. Orange filled circles represent simulation ("Sim") results and the blue line encompasses experimental data ("Exp") from [4].

Table 14 - Surface tension results from simulations and comparison with experimental data (from [4] for PFA and from [101] for alkanes).

Compound	T/K	Sim. ST (mN/m)	Stat. Err.	Experimental	Deviation (mN/m)	Deviation (%)
F_7	295	13.60	0.14	13.87*	- 0.27	- 1.9%
	305	12.38	0.28	12.94*	- 0.56	- 4.3%
	315	11.47	0.24	12.01*	- 0.54	- 4.5%
H_{12}	323.15	26.40	0.30	22.70	+ 3.70	+ 16.3%
	*** Calculated with the linear trend line applied to experimental data**					

Foremost, it should be recognised that no interfacial data was used in the modelling, making these results *pure predictions*. Whilst the surface tension of PFAs was predicted with reasonable accuracy (average deviations below 0.5 mN/m), that of *n*-dodecane was significantly overestimated [101]. Assuming that these deviations extend to other alkanes and other temperatures, their force field [42] must be used carefully in the computation of this interfacial property.

As a matter of fact, coarse-graining entails a choice on the precise properties that one wants to be studied; and it is not a failure of a force field modelled with bulk properties to fail to predict interfacial ones. For instance, the force field for alkanes proved to be accurate in predicting not only bulk properties

but also viscosities [26]; as such, for future work, it could be interesting to assess whether the force field for fluoroalkanes is well-suited to predict dynamic properties as well.

The previous set of beads had not been applied to evaluate the surface tension of perfluorocompounds (only SFAs, as detailed below). To assess the performance of both force fields, simulations of perfluorohexane (F_6) and perfluorooctane (F_8) were carried on (two temperatures for each compound). Unfortunately, the different set of beads did not allow for the "construction" of the same molecules. The results are shown on Table 15 and Figure 23, with experimental data taken, again, from [4].

Table 15 - Surface tensions of F_6 and F_8, evaluated with the previous force field.

	T/K	Sim. ST (mN/m)	Stat. Err.	Experimental (mN/m)	Deviation (mN/m)	Deviation (%)
F_6	295	11.74	0.37	12.71	- 0.97	- 7.6%
	305	10.53	0.11	11.55	- 1.02	- 8.8%
F_8	295	14.27	0.16	14.77	- 0.50	- 3.3%
	315	12.26	0.50	13.00	- 0.74	- 5.7%

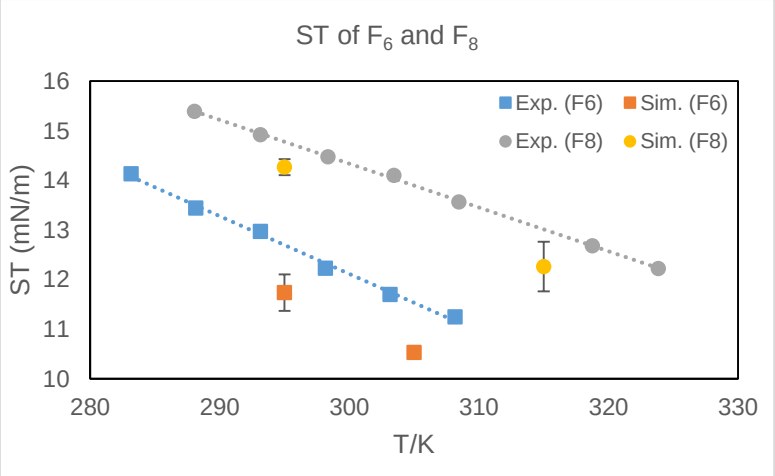


Figure 23 - Surface tension of F_6 and F_8. Circles concern F_8 and squares represent data for F_6.

From the relative deviations of both models, it can be concluded that the novel definition of CG groups leads to predictions of similar accuracy.

Subsequently, molecular simulations for PFAAs were performed. Because the new set of beads did not allow for the simulation of all compounds, only F_6H_6 and F_6H_8 were investigated.

Table 16 - Simulation data for the surface tension of F_6H_6 and F_6H_8 (with the new CG force field), with comparison with experimental data.

	T/K	Sim. ST (mN/m)	Stat. Error	Experimental	Deviation (mN/m)	Deviation (%)
F_6H_6	299.2	18.95	0.29	17.81	+ 1.14	+ 6.4%
	320.3	18.04	0.90	16.05	+ 1.99	+ 12.4%
F_6H_8	296.2	21.03	0.50	19.20	+ 1.83	+ 9.5%
	314.3	19.73	0.28	17.74	+ 1.99	+ 11.2%

Though not as accurate as for perfluoroalkanes, the simulated surface tension for two PFAAs is nevertheless reasonably close (average deviation of 1.74 mN/m and comparable relative deviations) to experimental measurements. Because Morgado *et al.* [43] had simulated these compounds, too, with the previous set of beads, a comparative analysis could be most useful to understand if one is better-suited for this purpose – see Figure 24, where the filled circles are the results of Table 16, the open circles represent the MD simulations obtained in the referred paper and the geometric figures translate experimental results (united by the dashed trend lines).

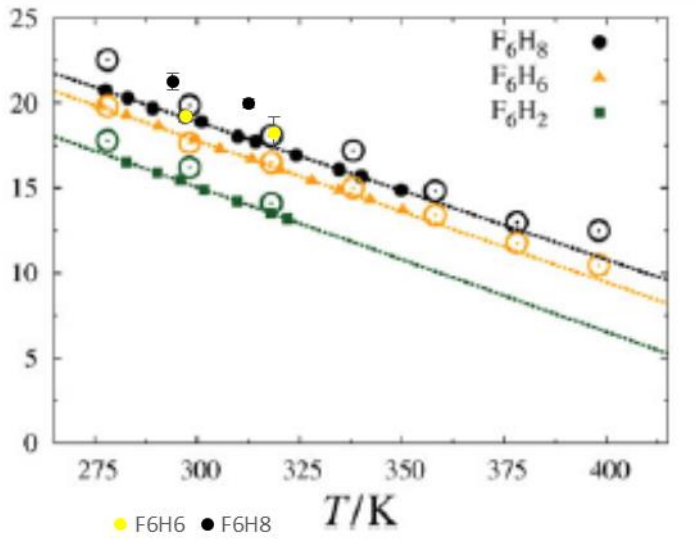


Figure 24 - Surface tension of PFAAs (adapted from [43]). The yellow and black circles represent the simulation results for F_6H_6 and F_6H_8, respectively, with the new CG beads.

Unfortunately, the cited paper does not provide quantitative values for the simulated surface tensions, so an analytical comparison in terms of deviation to the experimental values is not possible. Still and all, the values obtained with the previous force field do look somewhat closer to the dashed lines, but the difference is presumably around 1 mN/m. Given the intrinsic experimental and simulation uncertainties, both models in terms of surface tension for PFAAs exhibit alike accuracy.

Before concluding, and again to safeguard the reproducibility of the results, MD simulation details are described. A system comprising of 1000 molecules was simulated, with a time-step of 5 fs and with a cut-off radius of 3.0 nm. The procedure was: energy minimization; running on the NPT ensemble for 5 ns to assure equilibration; editing configuration to a box with dimensions L x L x 3L

(where L was the size of the cubic box of the final configuration after the NPT run) and performing 5 ns of NVT to equilibrate; and finally 10 ns of NVT of production, to obtain results. This methodology was followed for all MD simulations aiming at surface tension.

Up to now, the CG model proved to be accurate in the prediction of thermodynamic properties (liquid density and vapour pressure), vapour-liquid and liquid-liquid equilibria (with some exceptions for the latter). Surface tensions were accurately predicted for both perfluoro and semifluorinated alkanes, but showed a larger deviation for *n*-dodecane. Compared to the previous proposal, the model appears to be slightly better for the surface tension of perfluoroalkanes and slightly worse for that of semifluorinated alkanes, though both models have, frankly, a very akin accuracy. These tests, however, were not the aim but the means to tackle the important supramolecular phenomena that have been observed (and described in this thesis, above) when SFAs are mixed with alkanes or perfluoroalkanes. The next section will address the behaviour of such mixtures.

3. Organization of PFAAs in Hydrocarbons and Fluorocarbons

Over the last thirty years, there has been an important volume of research (by experimental and computational means) in several aspects of the behaviour of SFAs both as pure compounds and as solutes in the referred solvents. Interesting features reported include: the adsorption of SFAs at the surface [81, 102]; the arrangement in micellar aggregates of SFAs when added to certain solvents [10, 103]; surface freezing of films of SFAs at the surface of alkanes [13, 104, 105]; formation of smectic liquid crystals of pure SFAs [14]; and the impressive change of the LLE of alkane + perfluoroalkane mixtures when a SFA is added [103]. All the phenomena listed before show just how diverse can be the self-assembly of SFAs.

This chapter aims at these mixtures, focusing primarily on the surface tension [106] and self-assembly evidence. In particular, the influence of three different parameters will be investigated: temperature, concentration of SFA and fluorine content of the SFA (equivalent to the proportion of fluorinated carbon atoms in the SFA).

In order to have a sound starting point, some of the mixtures prepared in [10] were simulated (with the needed changes if the set of beads proved insufficient). $F_{12}H_{14}$ in H_{12} (at 1% and 4% molar concentration, at 50ºC) and F_9H_{15} in F_7 (at 5% and 10% molar concentration, at 45ºC)[13] were simulated. From here, the effect of SFA concentration could be analysed.

For these mixtures, in particular, the temperature was increased by 20ºC – hence allowing for a study of the temperature dependence of the ST. Furthermore, a parallel set of simulations (at the same temperature, concentrations and solvent) was performed with a range of different solutes, always keeping the total chain length: $F_{12}H_{14}$ was replaced by F_3H_{23}, F_9H_{17}, and $F_{21}H_5$; F_9H_{15} was substituted by F_3H_{21}, $F_{12}H_{12}$ and $F_{18}H_6$. The impact of the fluorine content could, then, be examined. All MD simulations were done with 2000 molecules.

[13] The authors used F_8H_{16}, which cannot be reproduced with the present CG groups.

3.1 PFAAs in *n*-dodecane

It is well known that perfluoroalkanes have a lower surface tension than *n*-alkanes. The known mutual phobicity between both families of compounds, allied to the differences in surface tensions, can then drive the adsorption of PFAAs at the surface, with the fluorinated tails oriented outwards to stabilize the surface. This ability to adsorb will be the first feature to be investigated.

Often, surfactant systems promote aggregation above a critical concentration (if above the Krafft temperature). For such systems, solute molecules initially adsorb at the surface and, at higher concentrations, they start aggregating. However, in systems like these, where van der Waals' forces are the only intermolecular interactions, it has not been reported a clear break at the CMC – instead, the physical properties of these mixtures change gradually [102]. Besides, the weak interactions between the solute molecules do not drive strongly aggregation, which explains why the same papers report very low aggregation numbers. Thus, aggregation will be another analysed feature.

Bearing in mind this discussion on the surfactant character of PFAAs, one starts by looking at the influence of temperature and concentration of a $F_{12}H_{14}$ in H_{12} system (see Figure 25). At constant temperature, thanks to the surfactant effect of SFAs, the addition of solute progressively reduces the surface tension; a rise in temperature, in turn, inspires an overall lower ST, as expected for both pure solvent and mixtures. In fact, the surface tension of pure solvents is negatively correlated with temperature; but when a surfactant is added, another layer must be considered: with higher thermal energy, it is increasingly less likely that the surfactant molecules stay concentrated at the surface – the surfactant effect is thereby reduced at higher temperatures, because the entropic contribution to the free energy increases with temperature. The two trends are, all in all, captured by the model. It is hard to infer, though, whether the ST (at each temperatures) has indeed stabilized, as the surface tension at both concentrations is identical if the error bars are considered. A stabilization would hint that the critical concentration had been surpassed.

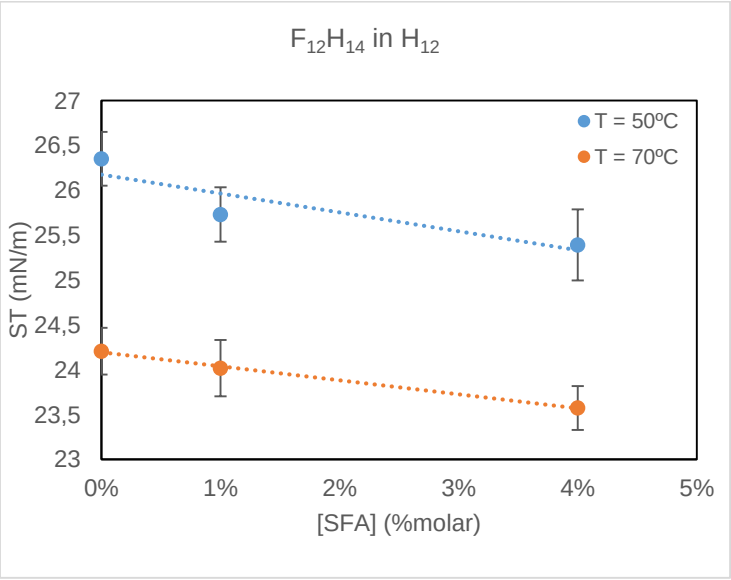


Figure 25 - Surface tension of mixtures of $F_{12}H_{14}$ in H_{12} at 50ºC (blue circles) and 70ºC (orange circles).

Before proceeding to other SFAs, and since the simulation conditions match exactly those in the paper by Binks *et al.* [10], an immediate comparison was drawn to evaluate the agreement with experimental data (see Figure 26).

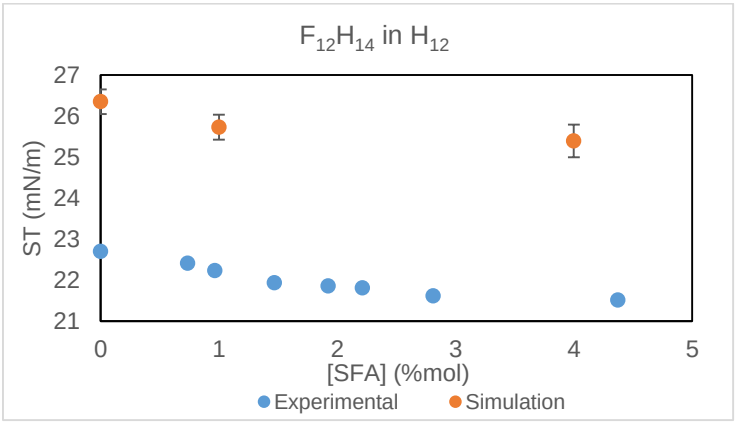


Figure 26 – Comparison of simulation results (orange dots) with experimental measurements (blue dots). Adapted from [10].

The contrast between simulation results for this mixture and experimental data is expected to be rooted in the force field for alkanes. As shown in Table 14, MD simulations pointed out to a ST roughly 4 mN/m above the experimental value for pure *n*-dodecane. From Figure 26, it can be concluded that whilst these simulations predict a higher ST, they can still capture the overall trend that was observed experimentally, thus showing qualitative agreement.

Lastly, one proceeds then to the influence of concentration of SFA and its fluorine content. Four different solutes were analysed: F_3H_{23}, F_9H_{17}, $F_{12}H_{14}$ and $F_{21}H_5$, always at 1%mol and 4%mol, as exhibited in Figure 27.

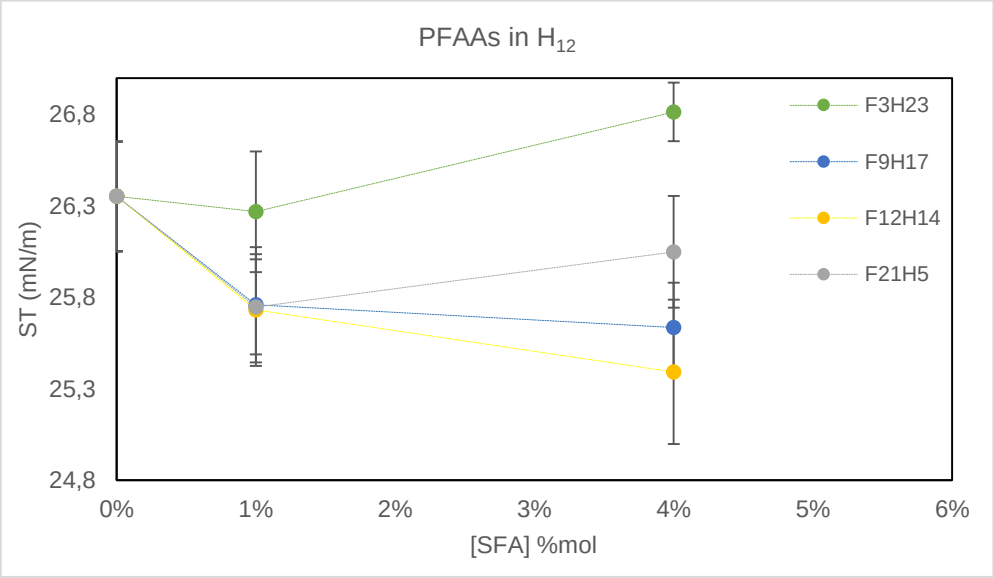


Figure 27 - Surface tensions of mixtures of different PFAAs in H_{12} at 50ºC.

For the more diluted mixtures, the surface tension of the F_3H_{23} mixture stands out as it promotes the lowest decrease in surface tension of the four mixtures while the ST of the other mixtures is virtually

equal (with their differences far smaller than the statistical error from simulations). In turn, for the more concentrated mixtures, there are two different sets of points: those who registered a further decrease in surface tension (F_9H_{17} and $F_{12}H_{14}$) and those who did not (F_3H_{23} and $F_{21}H_5$).

This analysis focuses, in the first place, in the mixtures at 1%mol. To help understand the trends in organization, snapshots taken at the end of the simulations are presented in Figure 28. Notice that the simulation box is not cubic – instead, it has dimensions *L x L x 3L* (though most of the molecules are located in the liquid slab in the middle of the simulation box). Whilst an adsorption of the fluorinated ends is shared by all mixtures, as hypothesized, the extension of this adsorption varies throughout the four images. This may happen because the driving force – the mutual phobicity between fluorinated and hydrogenated tails – is reduced when the fluorinated tail is shortened. For F_3H_{23} in particular, the very short alkane chain (comprising only one CG bead) causes a very slight reduction in surface tension; for the others, though, it is unlikely that they are so similar: one would expect that the surface tension of $F_{21}H_5$, being the most fluorinated SFA, would rank as the lowest – only future work can confirm or deny this guess.

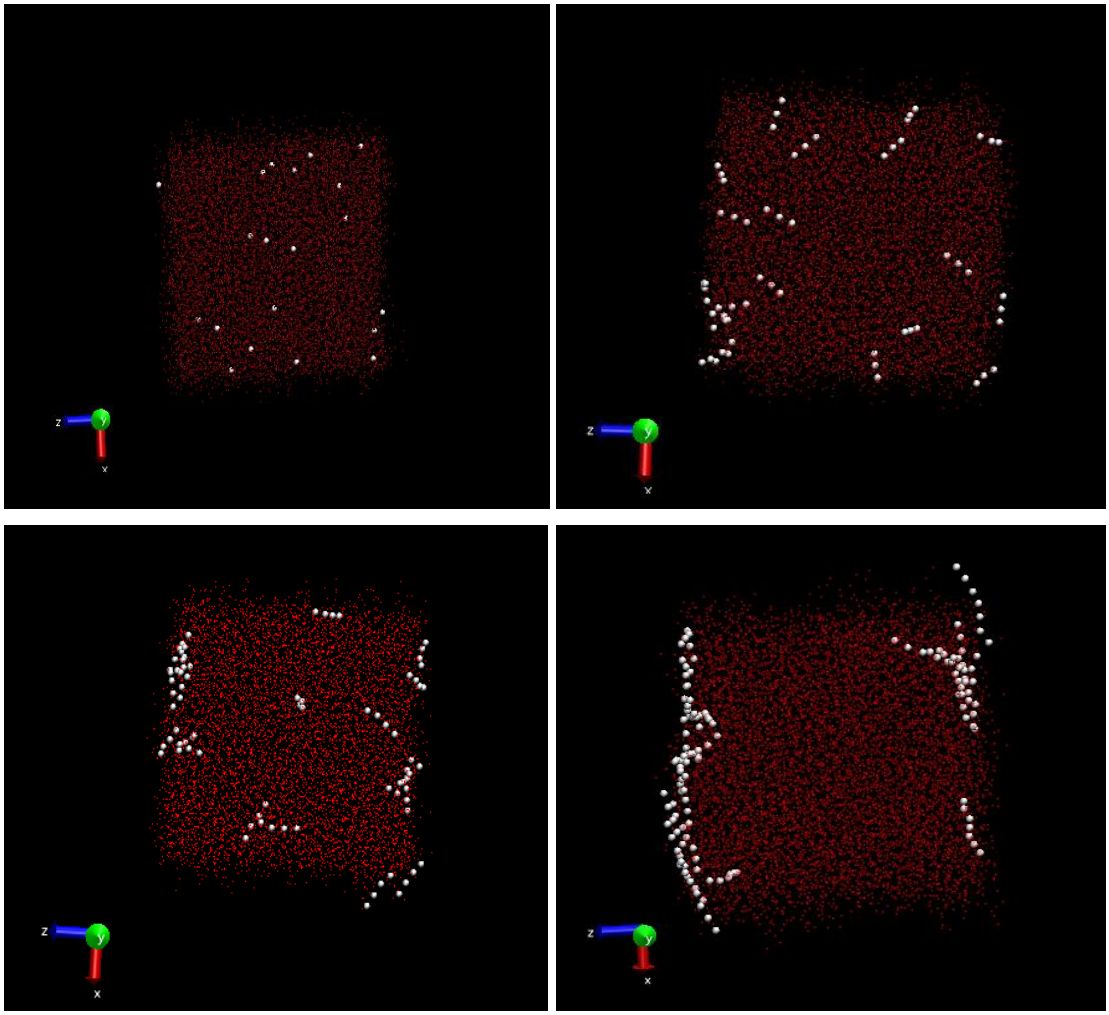


Figure 28 - Snapshots for the mixtures of F_3H_{23} (top, left), F_9H_{17} (top, right), $F_{12}H_{14}$ (bottom, left) and $F_{21}H_5$ (bottom, right) in H_{12} (with concentrations 1%mol). White spheres represent the fluorinated ends; red dots symbolise the solvent molecules.

On a different note, notice how the white spheres, on the snapshots of the three last compounds, tend to lay at the surface and not perpendicular to it – that is, the more fluorinated the solute, the more parallel to the surface is the adsorbed SFA molecule. What's more, the snapshots clearly demonstrate the relationship between the length of the fluorinated tail and the driving force to adsorb: whilst for F_3H_{23} there are several white spheres in the bulk, when the fluorine content rises the "bulk" (as opposed to "surface") concentration gets progressively lower; ultimately, the snapshot for $F_{21}H_5$ shows that there is virtually no solute molecule on the bulk. It is concluded that adsorption is gradually enhanced with the fluorine content. To sum up, the organization of SFAs in the analysed mixtures appears to be based on *positive adsorption*.

A valuable tool to quantify the distribution of molecules (or CG groups) along the *Oz* axis is the density profile. To take a closer look at adsorption, density profiles were computed for the above mixtures (only those for $F_{12}H_{14}$ and $F_{21}H_5$ are presented, as the remaining are quite similar to $F_{12}H_{14}$), as shown in Figure 29.

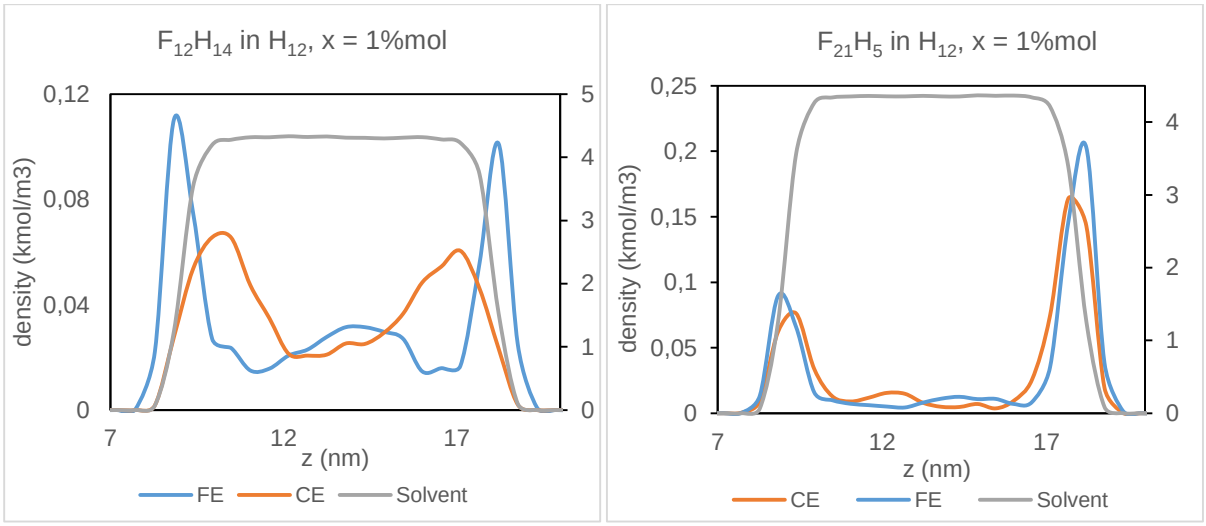


Figure 29 – Density profiles for $F_{12}H_{14}$ (left) and $F_{21}H_5$ (right) in H_{12}, at 1%mol. Solvent densities are always referred to the secondary axis. FE and CE are, respectively, the terminal fluorinated and hydrogenated groups of the SFA.

Regarding $F_{12}H_{14}$ in H_{12}, the gap between groups FE and CE proves this positive adsorption – the fluorinated group is definitely oriented outwards, in a segregation effort that leads to the stabilization of the surface. About $F_{21}H_5$, what must be highlighted is the absence of the aforementioned gap. In fact, because the molecule lays parallel to the surface, both terminal groups of the solute stand almost at the same *z* coordinate, with the peaks of both curves (FE and CE) matching for both surfaces. Moreover, the average molar density (equivalent to numerical density, and comparable since the concentration is the same) is higher at the surfaces for $F_{21}H_5$ than for $F_{12}H_{14}$ – thus supporting the greater extension of adsorption predicted for more fluorinated solutes.

One would expect that charts like those above would be symmetrical, for no interface should be preferred to other. At least two reasons are likely to contribute, yet to an uncertain extent, to the opposite. To begin with, the simulations could be not well equilibrated; however, an already significant effort was done to achieve a fairly equilibrated simulation: after 5 ns in the *NPT* ensemble and other 5 ns in the

NVT ensemble (each equivalent to 1.000.000 steps), the system ran another 10 ns (2.000.000 steps) for full production. What's more, the averages of all properties (not only surface tension but the potential energy of the system, density, etc.) were completely stabilized at the end of the production run. Furthermore, the size of the system may be too small – something like a few molecules starting closer to an interface may be enough to create an imbalance between the interfaces, even though there is virtually no physical reason that supports a preferential adsorption. This root cause is far more likely – and one should not exclude the wall effects that may be present, since in a box of roughly 10 nm, one can roughly fit 3 molecules, 3 nm long each. With only 20 solute molecules, the presence of a couple more in the "wrong" surface would suffice to uneven the peaks. More troubling than the inequality in the height of the peaks, though, is the absence of a region that could frankly be named "bulk" – in which the density profiles stabilized. Unfortunately, for the time being, this thesis could not afford to incur in larger simulations (the sequence of simulations to generate these charts took an average of 2-3 days). These results pretend ultimately to guide future work, which could include, among other things, the same simulations but with larger systems and longer simulation times.

Now that the results for more diluted mixtures are explained, this chapter proceeds with the same mixtures at 4%mol, discussing, each at a time, the two sets of mixtures specified before. For the first group (F_9H_{17} and $F_{12}H_{14}$), the results follow what was expected: a further reduction on the surface tension, with the more fluorinated SFA ($F_{12}H_{14}$) having a stronger surfactant effect. The fact that the surface tension did in fact reduce hints that the CMC is above 1%mol for both solutes, though more points would be needed to make a more precise estimate of the critical concentration. The respective snapshots (in Figure 30) seem to point towards a greater adsorption achieved at higher concentrations, when compared to those from Figure 28, though, from what the eye can tell, it is hard to identify aggregates, as underlined in [102] . The enrichment of the bulk appears, again, to be correlated with the fluorine content, as the more fluorinated solutes show a partition between bulk and surface more shifted to the surface side. From here, one could point out the same conclusions predicted on the beginning of this chapter: higher fluorine content produce a higher surfactant effect, which in turn contributes to a greater decrease in surface tension.

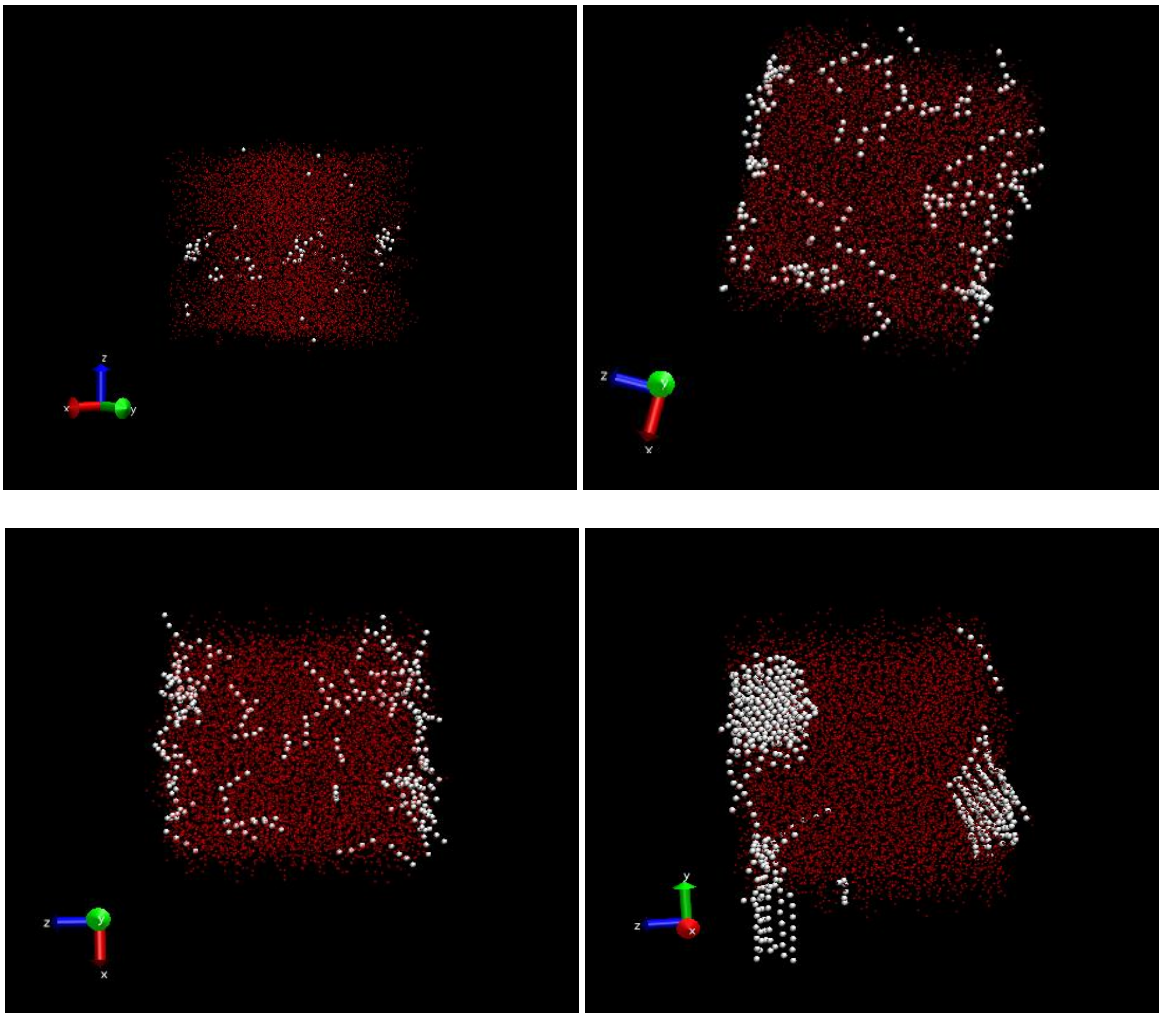


Figure 30 - Snapshots for the mixtures of F_3H_{23} (top, left), F_9H_{17} (top, right), $F_{12}H_{14}$ (bottom, left) and $F_{21}H_5$ (bottom, right) in H_{12} (with concentrations 4%mol). Colour scheme is the same of Figure 28. The first snapshot has its axis system purposely revolved to highlight the aggregates.

For the remaining mixtures, yet, this reasoning does not hold at all, as the surface tension happened to rise at 4%mol – and the snapshots, unlike for the previous pair, are far from similar. The $F_{21}H_5$ mixture snapshot hints what could be some aggregates (or precipitates), likely in the solid phase (given its packing), suggesting that the solubility limit may have been crossed between 1%mol and 4%mol. The surface tension value is hence meaningless – it would be like calculating the surface tension of a "water + ice" mixture; the presence of those aggregates at the surface thus prevents any conclusion from the computation, at least if executed via mechanical route (recall Equation 53).

The transition from aggregates to precipitates is very subtle, and it would be hard to tell, from the simulation, the exact moment in which this transition occurs. Even so, the snapshots for the $F_{21}H_5$ mixtures justify a suspicion that some solid-liquid transition has been undertaken. Crystallization is known to be rooted in a more compacted packing of the solute. Decisive to this packing is the conformation of the chains about to freeze. Whilst alkanes have an *all-trans* flat conformation, perfluoroalkanes chains grow according to a 15/7 or 13/6 helix[14] [11]. If one SFA is balanced (that is, if

[14] Broniatowski reports that the 15/7 helix is observed above room temperature while the 13/6 one happens below room temperature.

the hydrogenated and fluorinated chains are similar in length), it is harder for it to pack given the two different conformations coexisting – with more free spaces, the crystallization is hindered and the solute remains solubilized. This may help explaining why at least this mixture – of one very unbalanced solute – most likely underwent solidification.

Concerning the F_3H_{23} mixture, the respective snapshot indicates, in the bulk, what could be some aggregates or, at least, some "proto-aggregates", so to speak. The adsorbed quantity remains similar to the one witnessed (not strictly speaking), which hints that the CMC is actually between 1%mol and 4%mol. Nonetheless, if the CMC had been reached already, the surface tension should have achieved the explained plateau – a concentration above which it changes no more; instead, it augmented to a value higher than that of pure H_{12} (even considering the error bars).

Density profiles provide a better hint of what is happening indeed in both simulations with F_3H_{23} (see Figure 31). Whilst at 1%mol the concentration at the surface is still comparable to (but lower than) that in the bulk, at 4%mol what is seen is a huge segregation of the solute towards the bulk, with almost no SFA at the surface.

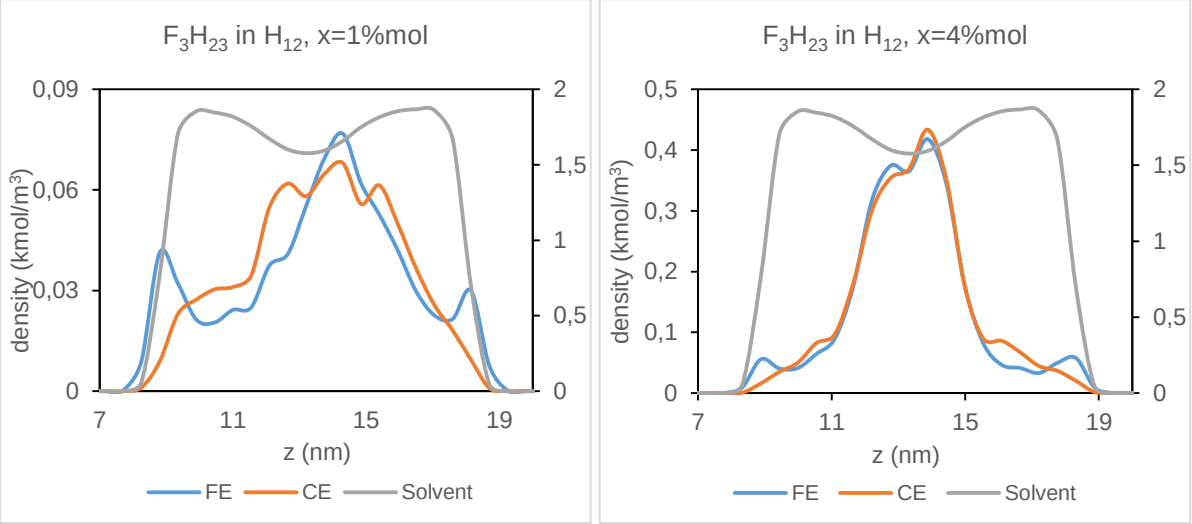


Figure 31 – Density charts for F_3H_{23} in H_{12} at 1%mol (left) and 4%mol (right).

What's more, if one defines the surface as roughly the distance from the "lift off" to the first minimum (on both sides, it encompasses on average a distance of 2.2 nm), it is clear that the increase in the quantity of SFA at the surface is outpaced by that in the bulk, as Table 17 shows. The enrichment numbers were actually insensitive to the precise position of the frontier.

Table 17 – Quantitative analysis (via number of molecules) of the distribution of F_3H_{23}.

	1%mol	4%mol	Enrichment (out of added 60 molecules)
Bulk	14.85	71.49	56.64
Surface	5.15	8.51	3.36

Not only, even at 1%mol, is most of solute placed at the bulk, but also the large majority (around 57) out of 60 molecules added from 1%mol to 4%mol goes to the bulk. This suggests that the surface-

bulk equilibrium for F_3H_{23} is particularly shifted to the bulk side, and that the ability of this SFA to adsorb is hampered, right from the beginning.

Dealing with F_3H_{23} – a "quasi-alkane" – brings new challenges that more "balanced" SFAs had not raised. The addition of a heavier alkane to a lighter alkane was proven to increase the surface tension [107] – an effect that could be achieved by hypothetically adding H_{26} to H_{12}. It is proposed that the weak surfactant effect of F_3H_{23} is only felt at small concentrations (at least until 1%mol); for more concentrated mixtures, the ability of this solute to decrease the surface tension is outstripped by the inevitable effect that adding a heavier compound to a lighter one would have if the former had no marked surfactant character.

This discussion raises the question of whether or not the ability of SFAs to stabilize the surface is concentration dependent, as well as if there is a minimum fluorine content above which a plateau in surface tension is achievable.

Unfortunately, due to lack of time, a proper tool for quantitatively studying these aggregates was not developed. However, at least to provide a hint on whether or not the reasoning described is legitimate, a script prepared by an Imperial College PhD student, Maziar Fayaz-Torshizi (to whom I am deeply grateful for his valuable collaboration), was used to compute a rough estimate of aggregation numbers. This programme analyses several frames of the trajectory file; in each frame, it focuses individually on each given group; and of each group, it counts the number of similar groups closer than a certain cut-off. In so doing, it computes the number of groups in each (or, equivalently, the average number of solute molecules) in each aggregate. Even though there was no time to properly test the sensitivity of this cut-off, or to make a thorough study of the shape of these aggregates, they nevertheless provided a quantitative measurement (though perhaps inexact) of the extension of the aggregation.

The programme takes, as an input, the groups prone to aggregate. Because both FE and FM are groups capable of aggregating (both nurture a mutual phobicity with the solvent), on a first attempt, those groups were considered equal for aggregation purposes – that is, it was assumed that for an FE group was equally advantageous (in terms of free energy) to aggregate to another FE or to an FM group (and vice-versa). The results of such analysis (which, again, is only presented to support or deny the anticipated hypothesis) are shown on Table 18.

Table 18 - Aggregation extension for PFAAs in H_{12}. In this analysis, both FE and FM groups were taken into account. "Avg." means "average".

Groups / Molecule	PFAA	Avg. Groups / Aggregate		Avg. Molecules / Aggregate	
		x=1%mol	x=4%mol	x=1%mol	x=4%mol
1	F_3H_{23}	1.01	1.27	1.01	1.27
3	F_9H_{17}	3.05	3.22	1.02	1.07
4	$F_{12}H_{14}$	4.22	4.79	1.06	1.20
7	$F_{21}H_5$	7.20	27.69	1.03	3.96

It must be stressed that “Avg. Groups / Aggregate” is not necessarily equivalent to “Avg. Molecules / Aggregate”, as aggregate may not comprise exclusively “full tails” instead of a few beads of a longer tail.

The predictions made before appear to be consistent with these aggregation calculations. $F_{21}H_5$ reports the most pronounced aggregation (which was expected, given the fact that it has the longest fluorinated chain). The script also captured an inflexion in the aggregation of F_3H_{23} (it increases slightly when compared to F_9H_{17}). Snapshots in Figure 30, however, had already portrayed how that solute barely adsorbs – instead, it tends to move to the bulk and adsorb there. This inflexion may reflect this different arrangement, but the differences in the numbers prevent a cabal conclusion.

A more thorough coverage of the whole range (potentially including both F_{26} and H_{26}) would contribute decisively to understand the true scope of the highlighted mechanisms. Among the very few scenarios investigated, there were: mixtures in which the surfactant effect led to a progressive decrease in surface tension, ruled by the fluorine content; aggregation in the bulk and negligible surfactant character; and crystallization at the surface. Questions such as: what is the Krafft temperature for each of the SFAs, and how does it relate to variables such as fluorine content?; how does aggregation and adsorption interact with each other depending on the fluorine content and concentration?; and what is the influence of the solvent’s chain length on the whole equilibrium?; among many others are left to future work.

3.2 PFAAs in perfluoroheptane

Let one consider the new case of a PFAA solute in a fluorocarbon solvent. Unlike in the previous scenario, the migration towards the surface of PFAAs is now hindered by the very same reasons it was promoted before: as the alkanes have a higher surface tension than perfluoroalkanes, the adsorption of PFAAs (with the alkane tails oriented outwards) would destabilize the surface. In fact, Binks *et al.* [10, 102] report that SFAs *do not adsorb* at fluorocarbon interfaces. But if these hydrogenated chains cannot adsorb and still cope with the mutual phobicity between them and the solvent, they might promote an arrangement where the solute molecules aggregate. Drawing parallels to the previous scenario, it seems reasonable to argue that this eventual aggregation would benefit from long hydrogenated tails, since maximizing the antipathy towards the solvent would increase the driving force for aggregation.

As done for hydrocarbon solvents, the influence of temperature, composition and %H (equivalent to %F) is investigated.

To begin with, Figure 32 shows the impact of increasing the temperature for the F_9H_{15} in F_7 mixture which, as before, reduces the surface tension of both pure solvent and mixtures. More importantly, the addition of PFAAs contributes to an increase of the surface tension. The first conclusion here is that PFAAs do not have surfactant effect (as in the ability to stabilize the surface) when mixed with fluorocarbon solvents.

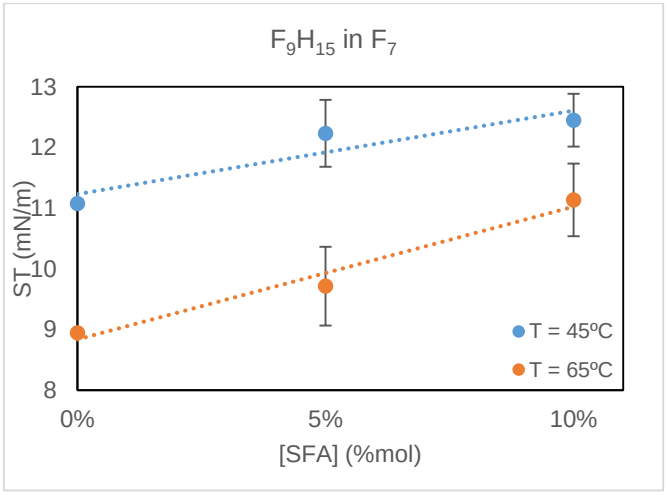


Figure 32 - Surface tension of mixtures of F_9H_{15} in F_7 at 45ºC and 65ºC (simulation data).

One proceeds then to the influence of concentration of SFA and its fluorine content. Four different solutes were analysed: F_3H_{21}, F_9H_{15}, $F_{12}H_{12}$ and $F_{18}H_6$, always at 5%mol and 10%mol, as exhibited in Figure 33.

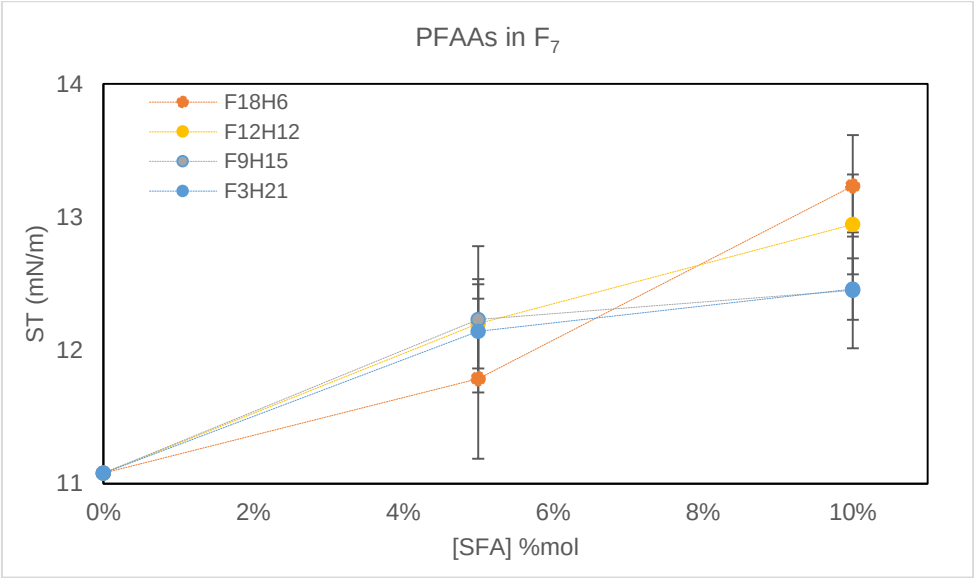


Figure 33 - Surface tensions of different mixtures of PFAAs in F_7, at 45ºC.

For the more diluted mixtures, the value for $F_{18}H_6$ is detached from the remaining three (which can be deemed statistically identical), that is, the more fluorinated solute imparted the smallest increase in surface tension. Regarding the more concentrated mixtures, the whole sequence of points appears to invert; the more fluorinated PFAA leads to the highest surface tension, whilst the less fluorinated SFAs impact the least this property. This inversion of sequence suggests that a different mechanism is ruling what happens at 5%mol from what happens at 10%mol.

Let one focus first on the mixtures at 5%mol. The results confirm the influence of the fluorine content on this property: because alkanes have a higher ST than perfluoroalkanes, a more hydrogenated SFA is related to a higher rise the ST. Again, the overlapping of the remaining three points

is considered unlikely, as one can expect that experimental results would rank them according to decreasing fluorine content. Complementing this discussion, the four snapshots of those mixtures at 5%mol are presented in Figure 34.

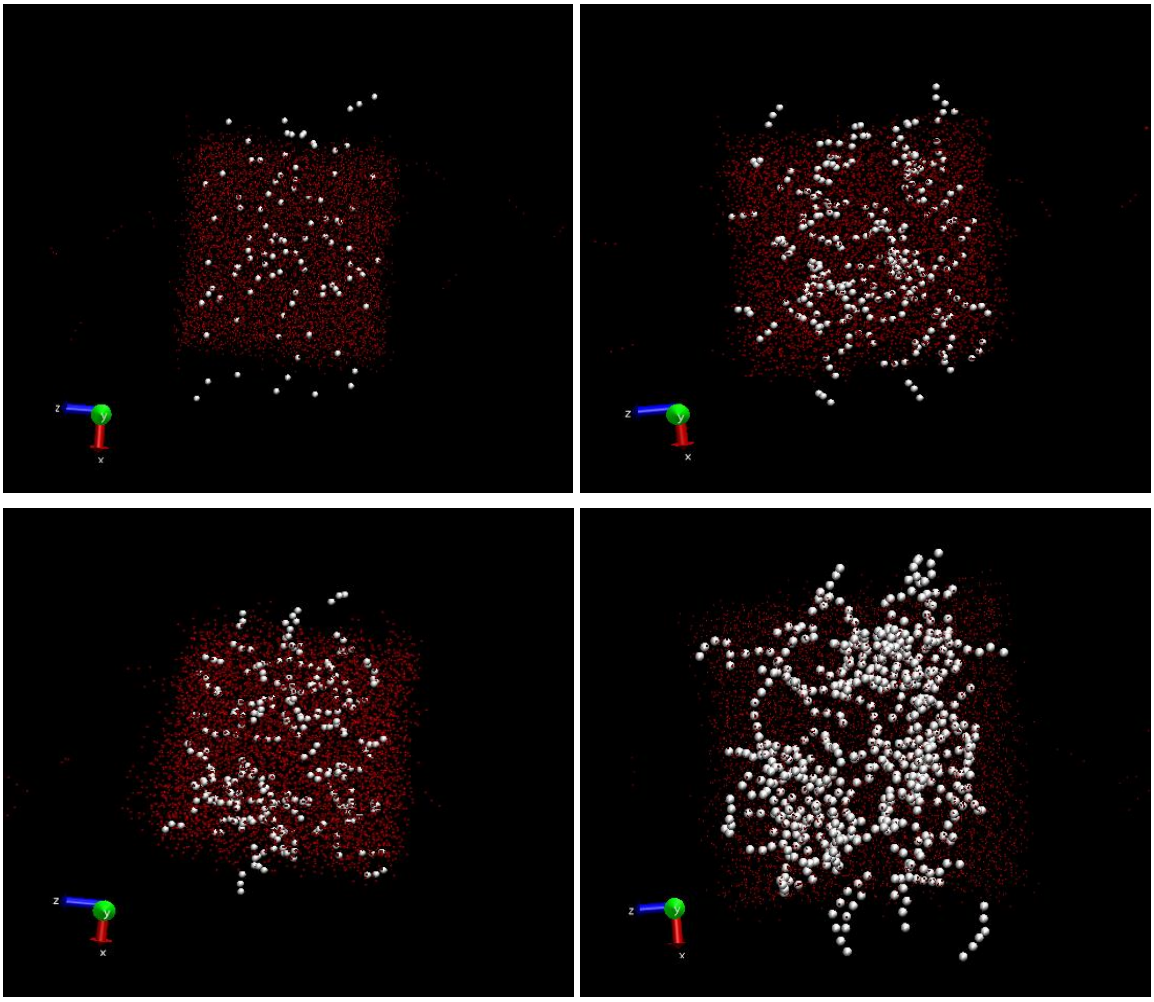


Figure 34 – Snapshots for the mixtures of $F_{18}H_6$ (top, left), $F_{12}H_{12}$ (top, right), F_9H_{15} (bottom, left) and F_3H_{21} (bottom, right) in F_7 (with concentrations 5%mol). White spheres represent the hydrogenated ends; red dots symbolise the solvent molecules.

As can be seen, there are no white spheres (forming the alkane tail) at the surface. This *negative adsorption* is shared by the four images, confirming the experimental observation reported in [10, 102] that SFAs do not adsorb in fluorocarbon solvents. Furthermore, whilst some regions look more populated than others (mainly in the two snapshots at the bottom), it is not obvious whether there is aggregation or not. A more comprehensive analysis of aggregation will follow in this chapter.

To better grasp this negative adsorption, density profiles were evaluated once again (see Figure 35). Notice how the "lift off" of the solvent density precedes that of the terminal groups of the solute, thus showing that, in this case, the solute moves further from the surface to eventually prevent an increase on the surface free energy. Whilst the two charts below are referred to F_9H_{15} and $F_{18}H_6$, the charts for the other mixtures were very similar, hence their omission for the sake of conciseness.

These density profiles validate, then, the *negative adsorption* expected for PFAAs in fluorocarbon solvents. Even $F_{18}H_6$, the solute more chemically similar to the solvent, presents a very similar distribution when compared to F_9H_{15}, with both ends of the solute chain concentrating in the bulk at the expenses of the surface.

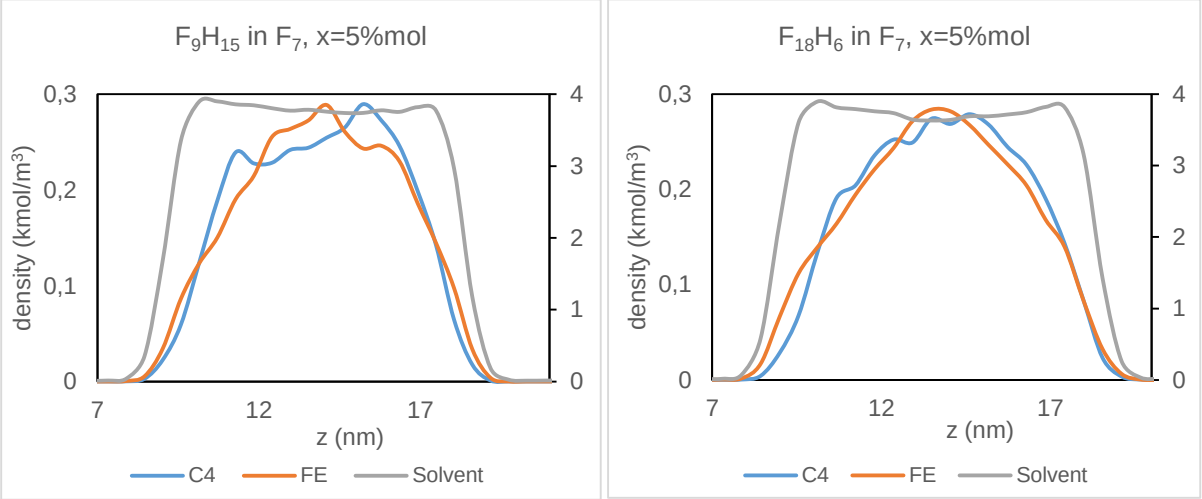


Figure 35 – Group density charts for mixtures of F_9H_{15} and $F_{18}H_6$ in F_7 at 5%mol.

Regarding the more concentrated mixtures, the whole sequence of points appears to invert, as already said: the more fluorinated compound now leads to the highest surface tension, whilst the less fluorinated SFAs impact the least this property (see again Figure 33).

The snapshots at 10%mol were inconclusive, inasmuch as the large number of hydrogenated beads does not allow do conclude whether or not they are gathered in aggregates (as opposed to concentrated single molecules); and the density profiles purveyed exactly the same shape of the charts above. However, because concentration is known to promote aggregation, analysing it may provide a better understanding of what is happening to the solute molecules in the bulk. Accordingly, the aggregation script referred above was applied, taking now, as groups prone to aggregate, CM, CE and C4. The results are shown on Table 19.

Table 19 - Aggregation extension for PFAAs in F_7. In this analysis, groups CM, CE and C4 were taken into account.

Groups / Molecule	PFAA	Avg. Groups / Aggregate		Avg. Molecules / Aggregate	
		x=5%mol	x=10%mol	x=5%mol	x=10%mol
6	F_3H_{21}	10.88	20.90	1.81	3.48
4	F_9H_{15}	5.03	6.45	1.26	1.61
3	$F_{12}H_{12}$	3.17	3.38	1.06	1.13
1	$F_{18}H_6$	*Not detected*			

Table 19 shows that aggregation gets progressively more relevant as the hydrogenated tail length increases. For $F_{18}H_6$, aggregation was not consistently detected, meaning that having only one hydrogenated bead (in this case, containing 4 carbon atoms) did not suffice to promote the formation of aggregates.

Looking at these results, it becomes clearer that aggregation plays a role in the self-assembly of PFAAs in fluorocarbon solvents. To better understand on what grounds the sequence inversion occurred, it is useful to focus on the $F_{18}H_6$ in F_7 mixture.

The $F_{18}H_6$ molecule comprises a single hydrogenated bead (C4, $–C_4H_9$). Eventually, this very short alkane tail hinders the formation of aggregates as the driving force is lower than for other PFAAs. Longer alkane tails increase the free energy gains from the aggregation of several molecules; and one could thus expect that F_3H_{21} would exhibit the most pronounced aggregation (which Table 19 confirms).

By forming aggregates, though, the system is preventing the individual solute molecules to move to (or to spend so much time at) the surface – actually dampening the direct effect of adding a heavier solute to the fluorinated solvent. Effectively hiding the alkane tails from the surface, the system is integrating the added solute without destabilizing (as much) the surface.

It appears, then, that the evolution of the surface tension of $F_{18}H_6$ in F_7 mixtures is one not perturbed by aggregation – instead, it is due to the addition of a heavier compound (with a higher surface tension than the solvent) to perfluoroheptane. Even in a context of negative adsorption (as the density profiles and the snapshots showed), single $F_{18}H_6$ molecules must affect the surface tension, at least because statistically they spend some time near the surface.

The other solutes, however, promote a different response. With increasingly longer alkane tails, the mixtures containing $F_{12}H_{12}$, F_9H_{15} and F_3H_{21} are able to form aggregates in the bulk with each other in order to minimize the exposure to the solute. In so doing, they tend to form supramolecular structures that are less likely to affect the overall “free energetic” balance between the bulk and the surface – the ultimate cause of surface tension itself.

To sum up, without aggregation (a scenario that can be assumed for low concentrations, in which the extension of such phenomenon falls short), a more hydrogenated SFA would contribute to a higher surface tension. With aggregation considered, though, a more hydrogenated SFA may be able to aggregate and henceforth hampering its impact on the ST, unlike a more fluorinated one. In other words, the fluorinated content could reduce the impact on the surface tension *as long as the molecules stayed separated from each other*. But since they also aggregate (following: the longer the hydrogenated chain, the larger the extent of the aggregation, and the higher the aggregation numbers), the F_9H_{15} mixture will contribute, in the end, to a smaller increase on the surface tension than its $F_{12}H_{12}$ counterpart.

This hypothesis would also explain why, at 5%mol, the $F_{18}H_6$ mixture had the smallest surface tension of all. Not that there was no aggregation – but the negative effects of longer, more hydrogenated chains were already felt in the surface tension, and the concentration was not high enough to assure a high degree of aggregation of the other SFAs. Notice that one could affirm that the negative adsorption (and its increasing effect on surface tension, the so-called *statistical effect*) is always there; upon it, there may be (or not) fulfilled the requirements to achieve a significant aggregation (intimately related with the length of the alkane tail). With a highly fluorinated solute, the driving force to aggregate is less significant, and the addition of the solute contributes “entirely” to the verified increase in the surface

tension. With a more balanced – and then a highly hydrogenated – solute, the ability that SFAs have to gather and self-assembly is more pronounced and the influence on the surface energy is *dampened*.

This hypothesis lacks formal evidence and even the statistical errors prevented the author from taking cabal conclusions. Unfortunately, the limited time resources available for a master thesis hindered further tests; still and all, some analysis left to future work include: a more rigorous aggregation number evaluation for all the mixtures; a more detailed study of this range (with more solutes, to understand the consequences imparted from the change of a single carbon atom from fluorinated to hydrogenated); and a thorough evaluation of the possible supramolecular structures that can be formed (namely micelles, as reported in [9]).

V - Final Remarks and Future Work

This thesis started with the tuning of an existent CG force field for fluorocompounds which had proved to be accurate in the prediction of bulk and interfacial properties of semifluorinated alkanes. The changes applied were meant to add a layer of physical coherence to the structure purveyed by the set of beads. The predictive capabilities of the new model are highly accurate for bulk and critical properties (namely temperatures) of pure perfluorocompounds (with deviations around 1%); in addition, a correct prediction of the VLE of alkane-perfluoroalkane mixtures (differences for experimental data around 2-4%) was accomplished, yet the same cannot be stated for the LLE of such mixtures (namely when the chain length of the two compounds differed).

This force field was subsequently applied, in MD simulations, to evaluate interfacial properties of perfluoro- and semifluorinated alkanes as well as *n*-dodecane. For the first family, the predictions are very accurate (average deviations below 0.5 mN/m); for SFAs, they are also reliable, though not as accurate (average deviations of around 1.70 mN/m); concerning *n*-dodecane, simulations overestimated the surface tension. The previous set of beads showed akin accuracies concerning both PFAs and PFAAs.

Furthermore, MD simulations were carried out aiming at supramolecular organization in mixtures of SFA in hydro- or fluorocarbon solvents. Observed phenomena included adsorption (positive and negative) as well as the formation of micelle-like aggregates. Even with a very elementary script, the very low aggregation numbers found (between 1 and 4) are consistent with results reported in [10], which range approximately between 2 and 6 (for similar mixtures and temperatures). $F_{12}H_{14}$ in H_{12}, being the only mixture that can be exactly compared, does not show, in simulation, any significant aggregation, suggesting an agreement with the experimental observation that the Krafft temperature is above 50ºC.

Even though the surface tension, due to overestimations ultimately rooted on the alkane parameterisation, could not be precisely computed, the results nevertheless allowed to take conclusions from a more qualitative approach. Thanks to the simulation of four PFAAs for each solvent that pretty much covered the whole range of fluorine content, it was possible to hypothesize what is the general influence of the addition of these compounds to hydrocarbon or fluorocarbon solvents.

In hydrocarbon solvents, conclusions differed between “unbalanced” and “balanced” SFAs. At 1%mol, all four solutes exhibited surfactant effect (decreasing surface tension, and doing so in an extent positively correlated with their fluorine content). At 4%mol, however, the “unbalanced” SFAs promoted a rise in surface tension, on which the influence of solubility matters was not fully clarified. Nonetheless, it hinted that “unbalanced” PFAAs tend to aggregate and/or crystallize at relatively low concentrations, perhaps due to an easier packing that the predominance of one conformation in the structure of the solute promotes, thus having an inconsistent surfactant effect. For more balanced SFAs, in which the packing is disfavoured by the different conformations coexisting in the two tails of the copolymer, the more fluorinated SFAs tend to have a more pronounced surfactant effect; this was frankly expectable, since perfluoroalkanes have a lower surface tension than alkanes – thus the presence of fluoroalkyl tails would contribute to stabilize the surface more effectively than the presence of their alkyl counterparts.

For these SFAs in particular, the surfactant effect was progressive and coherent with the fluorine content; but it remains to be studied if their trajectory would invert, for higher concentrations, as did the one for F_3H_{23}.

In fluorocarbon solvents, the surprising results showed what can be defined as a context regulated by aggregation and (negative) adsorption. First of all, the addition of longer compounds to a shorter fluorocarbon, in the absence of surfactant effect, would inevitably lead to a destabilization of the surface. For low concentrations, whilst the extension of aggregation is still trifling (solute molecules are sparse throughout the solvent), the (negative) adsorption is felt – and its extension is positively correlated with the hydrogenated content of the SFAs. In fact, the more hydrogenated the solute, the greater the ability to increase the surface tension, as they have indeed a larger surface tension themselves. When the concentration increases, aggregation becomes progressively more important – and the length of the hydrogenated chains plays a decisive role. The longer this chain, the larger the driving force for aggregation. With the SFAs closer to each other in the bulk (in which they tend to concentrate due to the negative adsorption), the hydrogenated chains start coalescing and forming *aggregates* – and they tend to produce higher aggregation numbers for longer hydrogenated chains. This organization ultimately *dampens* the effect that a more hydrogenated solute would have on the surface tension, with the effective consequence of a more hydrogenated solute (able to aggregate more) contributing less to the surface tension than a more fluorinated one (unable to aggregate). This counterintuitive idea fully explains the obtained the results. Nevertheless, this reasoning poses very intriguing questions such as: if the increasing effect in the surface tension brought by an addition of SFA may be potentially dampened by the aggregation of these very SFA molecules, is there an asymptote for the surface tension, that is, a threshold concentration above which the surface tension does not increase anymore? And until which concentration is “longer hydrogenated chain” correlated with “higher mixture surface tension”? The never-ending list of questions related to the surprising behaviour of SFAs in hydro- and fluorocarbon solvents demonstrates the complexity derived from the extremely subtle interactions established in such systems.

Above all, this work aimed to present a valuable tool for further studies of these systems – to support novel investigations with a reliable force field. Even though the surface tension of alkanes is out of the predictive capabilities of the model (something this thesis cannot be held accountable for), it still provides the means to investigate the variation of interfacial properties, aggregation, adsorption, in one word, *supramolecular organization* motivated by the marked amphiphilic character that so boldly characterizes SFAs.

For future work, at least two research avenues can be followed: the first would be an experimental study that could prove (or disprove) the presented forecasts. This study would focus not only on the measurement of surface tension of mixtures but also on the roots of those very values, with a clear emphasis on aggregation numbers and the supramolecular structures. This study could globally spread over more variables than the ones considered throughout this work: different solvents (longer or shorter chains); the behaviour of different SFAs in which the proportion of fluorinated to hydrogenated carbons is kept constant; among others.

Another research pathway could place effort on extending the results obtained via MD simulations. There are, obviously, more simulations to do, and a priority would undoubtedly be the coverage of the whole range of SFAs with a given carbon number (so to better understand the continuous evolution of the trends hypothesized above). Larger systems could be built – with the immediate advantage of having a "bulk" clearly detached from the surfaces (which was not completely achieved with the number of molecules used); in addition to this, a larger system would allow the simulation of an extraordinary result: the upwards shift of liquid-liquid curves of perfluoroalkane – alkane systems when a SFA is added [103]. Unfortunately, possible solid-liquid transitions ought not to be forgotten, since they place an important challenge in simulation procedures.

There is no such "last step" thing in science, but it seems fair to say that this thesis was a "second first step". After the successful model launched by Morgado *et al.* [43], it provided not only the structural coherence but also the tests (from bulk to interfacial properties, from VLE to LLE and critical points) that now support its acceptance. It finally moved on to mixtures, providing more insight over what has been the frankly challenging behaviour of SFAs in different solvents, as well as the mechanisms ruling their organization.

VI - References

1. Banks, R.E., B.E. Smart, and J.C. Tatlow, *Organofluorine chemistry : principles and commercial applications*. Topics in applied chemistry. 1994, New York: Plenum. xxv, 644 p.
2. Sandford, G., *Perfluoroalkanes.* Tetrahedron, 2003. **59**(4): p. 437-454.
3. Dias, A., et al., *Thermodynamic characterization of pure perfluoroalkanes, including interfacial and second order derivative properties, using the crossover soft-SAFT EoS.* Fluid Phase Equilibria, 2009. **286**(2): p. 134-143.
4. Freire, M.G., et al., *Surface tension of liquid fluorocompounds.* Journal of Chemical & Engineering Data, 2006. **51**(5): p. 1820-1824.
5. McCabe, C., et al., *Transport properties of perfluoroalkanes using molecular dynamics simulation: Comparison of united-and explicit-atom models.* Industrial & engineering chemistry research, 2003. **42**(26): p. 6956-6961.
6. Haszeldine, R. and F. Smith, *129. Organic fluorides. Part VI. The chemical and physical properties of certain fluorocarbons.* Journal of the Chemical Society (Resumed), 1951: p. 603-608.
7. Morgado, P., et al., *Systems involving hydrogenated and fluorinated chains: volumetric properties of perfluoroalkanes and perfluoroalkylalkane surfactants.* The Journal of Physical Chemistry B, 2011. **115**(50): p. 15013-15023.
8. Morgado, P., C. McCabe, and E.J. Filipe, *Modelling the phase behaviour and excess properties of alkane+ perfluoroalkane binary mixtures with the SAFT–VR approach.* Fluid phase equilibria, 2005. **228**: p. 389-393.
9. Turberg, M.P. and J.E. Brady, *Semifluorinated hydrocarbons: primitive surfactant molecules.* Journal of the American Chemical Society, 1988. **110**(23): p. 7797-7801.
10. Binks, B., et al., *Semifluorinated alkanes as primitive surfactants in apolar hydrocarbon and fluorocarbon solvents.* Journal of molecular liquids, 1997. **72**(1-3): p. 177-190.
11. Broniatowski, M. and P. Dynarowicz-Łątka, *Semifluorinated alkanes—primitive surfactants of fascinating properties.* Advances in colloid and interface science, 2008. **138**(2): p. 63-83.
12. Filipe, E.J.M., *Modeling the Interfacial Behaviour of Perfluoroalkylalkanes Surfactants using the SAFT-γ Coarse-Grained Force Field*. 2017, Heidelberg, Germany: SAFT 2017 Conference.
13. Sloutskin, E., et al., *Thermal expansion of surface-frozen monolayers of semifluorinated alkanes.* Langmuir, 2002. **18**(6): p. 1963-1967.
14. Mahler, W., D. Guillon, and A. Skoulios, *Smectic liquid crystal from (perfluorodecyl) decane.* Molecular Crystals and Liquid Crystals Letters, 1985. **2**(3-4): p. 111-119.
15. Allen, M.P. and D.J. Tildesley, *Computer simulation of liquids*. 1987, Oxford, United Kingdom: Clarendon Press, Oxford University Press.
16. Leach, A.R., *Molecular modelling : principles and applications*. 2nd ed. 2001, Harlow, England ; New York: Prentice Hall.
17. Frenkel, D., *Simulations: The dark side.* The European Physical Journal Plus, 2013. **128**(1): p. 10.
18. Müller, E.A. and G. Jackson, *Force-Field Parameters from the SAFT-γ Equation of State for Use in Coarse-Grained Molecular Simulations.* Annual Review of Chemical and Biomolecular Engineering, 2014. **5**(1): p. 405-427.
19. Gubbins, K.E. and J.D. Moore, *Molecular modeling of matter: Impact and prospects in engineering.* Industrial & Engineering Chemistry Research, 2010. **49**(7): p. 3026-3046.
20. Frenkel, D. and B. Smit, *Understanding molecular simulation : from algorithms to applications*. 2nd ed. Computational science series. 2002, San Diego: Academic Press. xxii, 638 p.
21. Verlet, L., *Computer" experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules.* Physical review, 1967. **159**(1): p. 98.
22. Allen, M.P., *Introduction to molecular dynamics simulation.* Computational soft matter: from synthetic polymers to proteins, 2004. **23**: p. 1-28.
23. Voth, G.A., *Coarse-graining of condensed phase and biomolecular systems*. 2009, Boca Raton: CRC Press.
24. Noid, W., *Perspective: Coarse-grained models for biomolecular systems.* The Journal of chemical physics, 2013. **139**(9): p. 09B201_1.
25. Brini, E., et al., *Systematic coarse-graining methods for soft matter simulations–a review.* Soft Matter, 2013. **9**(7): p. 2108-2119.
26. Rahman, S., *PhD Thesis*, in *Department of Chemical Engineering*. 2016, Imperial College London, United Kingdom.

27. Marrink, S.J. and D.P. Tieleman, *Perspective on the Martini model.* Chemical Society Reviews, 2013. **42**(16): p. 6801-6822.
28. Marrink, S.J., A.H. De Vries, and A.E. Mark, *Coarse grained model for semiquantitative lipid simulations.* The Journal of Physical Chemistry B, 2004. **108**(2): p. 750-760.
29. Lobanova, O., *PhD Thesis*, in *Department of Chemical Engineering*. 2014, Imperial College London: United Kingdom.
30. Nielsen, S.O., et al., *A coarse grain model for n-alkanes parameterized from surface tension data.* The Journal of chemical physics, 2003. **119**(14): p. 7043-7049.
31. Riniker, S., J.R. Allison, and W.F. van Gunsteren, *On developing coarse-grained models for biomolecular simulation: a review.* Physical Chemistry Chemical Physics, 2012. **14**(36): p. 12423-12430.
32. Monticelli, L. and E. Salonen, *Biomolecular simulations : methods and protocols*. Methods in molecular biology. 2013, New York: Humana Press ; Springer. xiii, 702 p.
33. Müller-Plathe, F., *Coarse-graining in polymer simulation: from the atomistic to the mesoscopic scale and back.* ChemPhysChem, 2002. **3**(9): p. 754-769.
34. Henderson, R., *A uniqueness theorem for fluid pair correlation functions.* Physics Letters A, 1974. **49**(3): p. 197-198.
35. Ingólfsson, H.I., et al., *The power of coarse graining in biomolecular simulations.* Wiley Interdisciplinary Reviews: Computational Molecular Science, 2014. **4**(3): p. 225-248.
36. Peter, C. and K. Kremer, *Multiscale simulation of soft matter systems–from the atomistic to the coarse-grained level and back.* Soft Matter, 2009. **5**(22): p. 4357-4366.
37. Lafitte, T., et al., *SAFT-γ force field for the simulation of molecular fluids: 3. Coarse-grained models of benzene and hetero-group models of n-decylbenzene.* Molecular Physics, 2012. **110**(11-12): p. 1189-1203.
38. Cornell, W.D., et al., *A second generation force field for the simulation of proteins, nucleic acids, and organic molecules.* Journal of the American Chemical Society, 1995. **117**(19): p. 5179-5197.
39. Müller, E.A. and K.E. Gubbins, *Molecular-based equations of state for associating fluids: A review of SAFT and related approaches.* Industrial & engineering chemistry research, 2001. **40**(10): p. 2193-2211.
40. Avendano, C., et al., *SAFT-γ force field for the simulation of molecular fluids. 1. A single-site coarse grained model of carbon dioxide.* The Journal of Physical Chemistry B, 2011. **115**(38): p. 11154-11169.
41. Lobanova, O., et al., *SAFT-γ force field for the simulation of molecular fluids: 4. A single-site coarse-grained model of water applicable over a wide temperature range.* Molecular Physics, 2015. **113**(9-10): p. 1228-1249.
42. Rahman, S., et al., *SAFT-γ force field for the simulation of molecular fluids. 5. Hetero-group coarse-grained models of linear alkanes and the importance of intra-molecular interactions.* (in press), 2014.
43. Morgado, P., et al., *SAFT-γ force field for the simulation of molecular fluids: 8. Hetero-segmented coarse-grained models of perfluoroalkylalkanes assessed with new vapour–liquid interfacial tension data.* Molecular Physics, 2016. **114**(18): p. 2597-2614.
44. Marrink, S.J., et al., *The MARTINI force field: coarse grained model for biomolecular simulations.* The journal of physical chemistry B, 2007. **111**(27): p. 7812-7824.
45. Haslam, A.J., *An Introduction to Statistical Thermodynamics and its Application in Equation-of-State Modelling of Fluids*, Imperial College London, 2017.
46. Hill, T.L., *Statistical mechanics : principles and selected applications*. 1987, New York: Dover Publications.
47. Hansen, J.-P. and I.R. McDonald, *Theory of simple liquids*. 3rd ed. 2007, Amsterdam ; Boston: Elsevier / Academic Press.
48. Pathria, R.K. and P.D. Beale, *Statistical mechanics*. 3rd ed. 2011, Amsterdam ; Boston: Elsevier/Academic Press.
49. Atkins, P.W. and J. De Paula, *Atkins' Physical chemistry*. 8th ed. 2006, Oxford ; New York: Oxford University Press.
50. Economou, I.G., *Statistical associating fluid theory: a successful model for the calculation of thermodynamic and phase equilibrium properties of complex fluid mixtures.* Industrial & engineering chemistry research, 2002. **41**(5): p. 953-962.
51. Rowlinson, J.S., *Cohesion : a scientific history of intermolecular forces*. 1st Ed. 2005, Cambridge ; New York: Cambridge University Press.

52. Jackson, G., *Liquid State Theory*. 2016: National Training School in Theoretical Chemistry - Oxford University.
53. Chapman, W.G., G. Jackson, and K.E. Gubbins, *Phase equilibria of associating fluids: chain molecules with multiple bonding sites.* Molecular Physics, 1988. **65**(5): p. 1057-1079.
54. Wertheim, M., *Fluids with highly directional attractive forces. I. Statistical thermodynamics.* Journal of statistical physics, 1984. **35**(1): p. 19-34.
55. Wertheim, M., *Fluids with highly directional attractive forces. II. Thermodynamic perturbation theory and integral equations.* Journal of statistical physics, 1984. **35**(1-2): p. 35-47.
56. Wertheim, M., *Fluids with highly directional attractive forces. III. Multiple attraction sites.* Journal of statistical physics, 1986. **42**(3): p. 459-476.
57. Wertheim, M., *Fluids with highly directional attractive forces. IV. Equilibrium polymerization.* Journal of statistical physics, 1986. **42**(3): p. 477-492.
58. Papaioannou, V., et al., *Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments.* The Journal of chemical physics, 2014. **140**(5): p. 054107.
59. Dufal, S., et al., *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, 2014. **59**(10): p. 3272-3288.
60. Barker, J.A. and D. Henderson, *Perturbation theory and equation of state for fluids. II. A successful theory of liquids.* The Journal of chemical physics, 1967. **47**(11): p. 4714-4721.
61. Boublík, T., *Hard-Sphere Equation of State.* The Journal of chemical physics, 1970. **53**(1): p. 471-472.
62. Mansoori, G., et al., *Equilibrium thermodynamic properties of the mixture of hard spheres.* The Journal of Chemical Physics, 1971. **54**(4): p. 1523-1525.
63. Gross, J. and G. Sadowski, *Perturbed-chain SAFT: An equation of state based on a perturbation theory for chain molecules.* Industrial & engineering chemistry research, 2001. **40**(4): p. 1244-1260.
64. Johnson, J.K., E.A. Mueller, and K.E. Gubbins, *Equation of state for Lennard-Jones chains.* The Journal of Physical Chemistry, 1994. **98**(25): p. 6413-6419.
65. Felipe J, B. and V. Lourdes F, *Thermodynamic behaviour of homonuclear and heteronuclear Lennard-Jones chains with association sites from simulation and theory.* Molecular Physics, 1997. **92**(1): p. 135-150.
66. McCabe, C. and A. Galindo, *SAFT associating fluids and fluid mixtures*. 2010: Royal Society of Chemistry: London.
67. Prausnitz, J.M., R.N. Lichtenthaler, and E.G.d. Azevedo, *Molecular thermodynamics of fluid-phase equilibria*. 3rd ed. Prentice-Hall international series in the physical and chemical engineering sciences. 1999, Upper Saddle River, N.J.: Prentice Hall PTR.
68. Marshall, B.D. and W.G. Chapman, *Molecular theory for self assembling mixtures of patchy colloids and colloids with spherically symmetric attractions: The single patch case.* The Journal of Chemical Physics, 2013. **139**(10): p. 104904.
69. Ryckaert, J.-P. and A. Bellemans, *Molecular dynamics of liquid alkanes.* Faraday Discussions of the Chemical Society, 1978. **66**(0): p. 95-106.
70. Jackson, G., *Molecular Moddeling of Fluids - Molecular Simulation Approaches*. 2017: Imperial College London, United Kingdom.
71. Mie, G., *Zur kinetischen Theorie der einatomigen Körper.* Annalen der Physik, 1903. **316**(8): p. 657-697.
72. Jones, J.E. *On the determination of molecular fields. II. From the equation of state of a gas*. in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. 1924. The Royal Society.
73. Reith, D., H. Meyer, and F. Müller-Plathe, *Mapping atomistic to coarse-grained polymer models using automatic simplex optimization to fit structural properties.* Macromolecules, 2001. **34**(7): p. 2335-2345.
74. Watkins, E.K. and W.L. Jorgensen, *Perfluoroalkanes: Conformational analysis and liquid-state properties from ab initio and Monte Carlo calculations.* The Journal of Physical Chemistry A, 2001. **105**(16): p. 4118-4125.
75. Okada, O., et al., *Molecular Dynamics Studies of Amorphous Poly (tetrafluoroethylene).* Molecular simulation, 1999. **21**(5-6): p. 325-342.
76. Gough, C.A., S.E. Debolt, and P.A. Kollman, *Derivation of fluorine and hydrogen atom parameters using liquid simulations.* Journal of computational chemistry, 1992. **13**(8): p. 963-970.

77. Yamamoto, R., O. Kitao, and K. Nakanishi, *Monte Carlo simulation of fluoro propane.* Fluid phase equilibria, 1995. **104**: p. 349-361.
78. Cui, S.T., et al., *Intermolecular potentials and vapor–liquid phase equilibria of perfluorinated alkanes.* Fluid Phase Equilibria, 1998. **146**(1): p. 51-61.
79. Potoff, J.J. and D.A. Bernard-Brunel, *Mie potentials for phase equilibria calculations: Application to alkanes and perfluoroalkanes.* The Journal of Physical Chemistry B, 2009. **113**(44): p. 14725-14731.
80. Shinoda, W., R. DeVane, and M.L. Klein, *Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants.* Molecular Simulation, 2007. **33**(1-2): p. 27-36.
81. Gaines Jr, G.L., *Surface activity of semifluorinated alkanes: F (CF2) m (CH2) nH.* Langmuir, 1991. **7**(12): p. 3054-3056.
82. Morgado, P., et al., *Liquid phase behavior of perfluoroalkylalkane surfactants.* The Journal of Physical Chemistry B, 2007. **111**(11): p. 2856-2863.
83. Lobanova, O., et al., *SAFT-γ force field for the simulation of molecular fluids 6: binary and ternary mixtures comprising water, carbon dioxide, and n-alkanes.* The Journal of Chemical Thermodynamics, 2016. **93**: p. 320-336.
84. Delhommelle, J. and P. Millié, *Inadequacy of the Lorentz-Berthelot combining rules for accurate predictions of equilibrium properties by molecular simulation.* Molecular Physics, 2001. **99**(8): p. 619-625.
85. Duh, D.-M., D. Henderson, and R.L. Rowley, *Some effects of deviations from the Lorentz-Berthelot combining rules for mixtures of Lennard-Jones fluids.* Molecular Physics, 1997. **91**(6): p. 1143-1147.
86. Song, W., P.J. Rossky, and M. Maroncelli, *Modeling alkane+ perfluoroalkane interactions using all-atom potentials: Failure of the usual combining rules.* The Journal of chemical physics, 2003. **119**(17): p. 9145-9162.
87. Lafitte, T., et al., *Accurate statistical associating fluid theory for chain molecules formed from Mie segments.* The Journal of chemical physics, 2013. **139**(15): p. 154504.
88. E. W. Lemmon, M.O.M., D. G. Friend. *Thermophysical Properties of Fluid Systems in NIST Chemistry WebBook, NIST Standard Reference Database Number 69 Eds. Linstrom, P.J.; Mallard, W.G.* . National Institute of Standards and Technology, Gaithersburg MD, 20899, retrieved 2013; Available from: http://webbook.nist.gov.
89. Duce, C., et al., *VLE and LLE of perfluoroalkane+ alkane mixtures.* Fluid Phase Equilibria, 2002. **199**(1): p. 197-212.
90. Crowder, G.A., et al., *Vapor pressures and triple point temperatures for several pure fluorocarbons.* Journal of Chemical and Engineering Data, 1967. **12**(4): p. 481-485.
91. Steele, W., et al., *Vapor pressure, heat capacity, and density along the saturation line, measurements for cyclohexanol, 2-cyclohexen-1-one, 1, 2-dichloropropane, 1, 4-di-tert-butylbenzene,(±)-2-ethylhexanoic acid, 2-(methylamino) ethanol, perfluoro-n-heptane, and sulfolane.* Journal of Chemical & Engineering Data, 1997. **42**(6): p. 1021-1036.
92. Dias, A.M., et al., *Densities and vapor pressures of highly fluorinated compounds.* Journal of Chemical & Engineering Data, 2005. **50**(4): p. 1328-1333.
93. Morgado, P., et al., *Vapor pressure of perfluoroalkylalkanes: the role of the dipole.* The Journal of Physical Chemistry B, 2015. **119**(4): p. 1623-1632.
94. Vandana, V., D.J. Rosenthal, and A.S. Teja, *The critical properties of perfluoro n-alkanes.* Fluid phase equilibria, 1994. **99**: p. 209-218.
95. Privat, R., et al., *Are safe results obtained when SAFT equations are applied to ordinary chemicals? Part 2: Study of solid–liquid equilibria in binary systems.* Fluid Phase Equilibria, 2012. **318**: p. 61-76.
96. Matsuda, H., et al., *Liquid–liquid equilibrium data for binary perfluoroalkane (C 6 and C 8)+ n-alkane systems.* Fluid Phase Equilibria, 2010. **297**(2): p. 187-191.
97. McCabe, C., et al., *Predicting the high-pressure phase equilibria of binary mixtures of perfluoro-n-alkanes+ n-alkanes using the SAFT-VR approach.* The Journal of Physical Chemistry B, 1998. **102**(41): p. 8060-8069.
98. Munson, M., *Solutions of fluorochemicals and hydrocarbons.* The Journal of Physical Chemistry, 1964. **68**(4): p. 796-801.
99. Lo Nostro, P., L. Scalise, and P. Baglioni, *Phase separation in binary mixtures containing linear perfluoroalkanes.* Journal of Chemical & Engineering Data, 2005. **50**(4): p. 1148-1152.
100. Vargaftik, N., B. Volkov, and L. Voljak, *International tables of the surface tension of water.* Journal of Physical and Chemical Reference Data, 1983. **12**(3): p. 817-820.

101. Jasper, J.J. and E.V. Kring, *The Isobaric Surface Tensions and Thermodynamic Properties of the Surfaces of a Series of n-Alkanes, C5 to C18, 1-Alkenes, C6 to C16, and of n-Decylcyclopentane, n-Decylcyclohexane and n-Dcylbenzene.* The Journal of Physical Chemistry, 1955. **59**(10): p. 1019-1021.
102. Binks, B., et al., *Adsorption and aggregation of semifluorinated alkanes in binary and ternary mixtures with hydrocarbon and fluorocarbon solvents.* Langmuir, 1997. **13**(25): p. 6669-6682.
103. Nostro, P.L., et al., *Effect of a semifluorinated copolymer on the phase separation of a fluorocarbon/hydrocarbon mixture.* The Journal of Physical Chemistry, 1995. **99**(27): p. 10858-10864.
104. Marczuk, P., et al., *Gibbs films of semi-fluorinated alkanes at the surface of alkane solutions.* Langmuir, 2002. **18**(18): p. 6830-6838.
105. Gang, O., et al., *Surface phases of semi-fluorinated alkane melts.* EPL (Europhysics Letters), 2000. **49**(6): p. 761.
106. Pierce, F., et al., *Interfacial properties of semifluorinated alkane diblock copolymers.* The Journal of chemical physics, 2008. **128**(21): p. 214903.
107. Rolo, L.I., et al., *Surface tension of heptane, decane, hexadecane, eicosane, and some of their binary mixtures.* Journal of Chemical & Engineering Data, 2002. **47**(6): p. 1442-1445.