Molecular modeling of organic phases after plutonium extraction

Abdelmounaim Failali

University of Lille
Material Sciences, Radiation and Environment
Doctoral School 104

Thesis prepared at the French Alternative Energies and Atomic Energy Commission (CEA)

&

Laboratory of Physics of Lasers, Atoms and Molecules (PhLAM)

for obtaining the degree of

Doctor of Philosophy

Speciality: Physics and Theoretical Chemistry

Molecular modeling of organic phases after plutonium extraction

presented and defended publicly by

Abdelmounaim Failali

on 26 January 2021 in front of the jury composed of

Carine Clavaguéra	CNRS, Université Paris-Saclay, France	Rapporteur
Mathieu Salanne	Université de Sorbonne, France	Rapporteur
Francesca Ingrosso	Université de Lorraine, France	Examinateur
Frédéric Affouard	Université de Lille, France	Examinateur
Florent Réal	Université de Lille, France	Co-Directeur de thèse
Valérie Vallet	CNRS, Université de Lille, France	Co-Directrice de thèse
Dominique Guillaumont	CEA - Marcoule, France	Co-Directeur de thèse
Eléonor Acher	CEA - Marcoule, France	Co-Directeur de thèse

Université de Lille
Sciences de la Matière, du Rayonnement et de l'Environnement
École Doctorale 104

Thèse préparée au Commissariat à l'énergie atomique et aux énergies alternatives (CEA)
&
Laboratoire de Physique des Lasers, Atomes et Molécules (PhLAM)

pour obtenir le grade de

Docteur de l'Université de Lille

Spécialité : Physique et Chimie théorique

Modélisation moléculaire des phases organiques après extraction de plutonium

Présentée et soutenue publiquement par

Abdelmounaim Failali

le 26 Janvier 2021 devant un jury composé de :

Carine Clavaguéra	CNRS, Université Paris-Saclay, France	Rapporteur
Mathieu Salanne	Université de Sorbonne, France	Rapporteur
Francesca Ingrosso	Université de Lorraine, France	Examinateur
Frédéric Affouard	Université de Lille, France	Examinateur
Florent Réal	Université de Lille, France	Co-Directeur de thèse
Valérie Vallet	CNRS, Université de Lille, France	Co-Directrice de thèse
Dominique Guillaumont	CEA - Marcoule, France	Co-Directeur de thèse
Eléonor Acher	CEA - Marcoule, France	Co-Directeur de thèse

Abstract

The nuclear fuel after its dwell time in reactor still bears a substantial amount of recoverable U and Pu. The recovery and purification of these actinides is achieved using a hydro-metallurgical process known as PUREX (Plutonium Uranium Recovering by EXtraction). Based on Liquid-Liquid extraction techniques, this process requires the use of a specific molecule to extract Pu and U, the tri-n-butylphosphate TBP. N,N-dialkylamides (monoamides) are regarded as an alternative family of extractants to TBP, as they are well-known for their strong extraction ability of Pu(IV) and U(VI) elements. In addition to this, they show some interesting features, such as, the strong dependence of the extraction properties (distribution coefficient and selectivity) on the ligands structure as well as chemical conditions. In order to propose the best extracting molecule design for future fuel reprocessing plants, it is crucial to understand the relationship between the structure and the extraction ability. However, the radioactivity of these elements combined with their chemical complexity make the study of these phases experimentally a real challenge. Hence, molecular modeling appears to be the golden solution for getting new insights on this issue.

In the first part of this thesis, a relativistic density functional theory study was performed to investigate the influence of the monoamides alkyl chain nature on the relative stability of Pu(IV) complexes. It was possible to reach a better understanding of the strong influence of amide structure on plutonium extraction. For both investigated amide-plutonium-nitrate complexes (inner and outer-sphere complexes), it was found that the introduction of a bulky alkyl group on the carbonyl side has a major impact on the complexation energy. The impact of the polarity of the solution was also investigated and found to be significant.

In the second part, within the aim of studying more realistic systems, i.e systems containing long alkyl chains monoamides, heavy elements and other counter ions, and to go beyond the static picture of QM/DFT optimized geometries with molecular dynamics simulations, we have developed a consistent polarizable FF model for the solvent molecules (alkanes, monoamides) based solely on quantum chemical calculations. The chosen ab initio parameterization approach as well as the final force field are presented. Then, the results of molecular dynamics simulations were compared to available experimental macroscopic thermodynamics and structural properties, and show an excellent agreement, making the predictions of properties of pure monoamides reliable. Finally, preliminary MD simulations results for monoamides-dodecane mixtures (DEHiBA/dodecane and DEHBA/dodecane) are presented.

Mots clés: Plutonium, ab initio, Extraction, Separation, N,N-dialkylamide, Dynamique moléculaire, Champ de force, Paramétrage, DFT, MP2.

Résumé

Après son séjour au sein d'un réacteur nucléaire, le combustible contient encore une quantité importante de matières valorisables qu'il est intéressant de récupérer, à savoir le plutonium et l'uranium. La récupération et la purification de ces actinides sont réalisées à l'aide d'un procédé hydrométallurgique appelé PUREX (Plutonium Uranium Recovering by Extraction), basé sur les techniques d'extraction liquide-liquide. Ce procédé nécessite l'utilisation d'une molécule spécifique pour extraire Pu et U, le phosphate de tri-n-butyle TBP. Les N, N-dialkylamides (monoamides) sont considérés comme une famille alternative d'agents d'extraction au TBP en raison de leur forte capacité d'extraction des éléments Pu(IV) et U(VI). De plus, ces molécules présentent des caractéristiques intéressantes, telles que la forte dépendance des propriétés d'extraction (coefficient de distribution et sélectivité) à la structure des ligands ainsi qu'aux conditions chimiques. Afin de proposer le meilleur design de molécule d'extraction pour les futures usines de retraitement de combustible, il est crucial de comprendre la relation entre la structure et la capacité d'extraction. Cependant, le caractère radioactif de ces éléments combinés à leur complexité chimique rendent les études expérimentales de ces phases complexes. Par conséquent, la modélisation moléculaire semble être la solution idéale pour obtenir de nouvelles informations à l'échelle moléculaire.

Dans la première partie de cette thèse, une étude quantique relativiste scalaire utilisant la théorie fonctionnelle de la densité a été réalisée pour déterminer l'influence de la nature de la chaîne alkyle monoamides sur la stabilité des complexes Pu(IV). Il a été possible de mieux comprendre la forte influence de la structure amide sur l'extraction du plutonium. Pour les deux complexes d'amide-plutonium-nitrate étudiés (complexes

de sphères interne et externe), il a été constaté que l'introduction d'un groupe alkyle volumineux du côté carbonyle a un impact majeur sur l'énergie de complexation. L'impact de la polarité de la solution a été également étudié et jugé significatif.

Dans le but d'étudier des systèmes plus réalistes, contenant des monoamides avec des longues chaînes alkyles, des actinides et des contre-ions, et d'aller au-delà de l'image statique de géométries optimisées au niveau QM/DFT avec des simulations de dynamique moléculaire classique, nous avons développé des champs de force polarisables pour les molécules de solvant (alcanes et monoamides) ajustés uniquement sur des calculs de chimie quantique.

L'approche *ab initio* retenue pour le paramétrage ainsi que le champ de force résultant et détaillés nous permettent d'obtenir des propriétés macroscopiques comparables aux données expérimentales (thermodynamiques et structurales). L'excellent accord nous permet d'avoir confiance quant à la précision des prédictions réalisées sur les systèmes pures de monoamides. Enfin, les résultats préliminaires de simulations des mélanges monoamides-dodécane (DEHiBA / dodécane et DEHBA / dodécane) sont présentés.

Mots clés: Plutonium, ab initio, Extraction, Separation, N,N-dialkylamide, Dynamique moléculaire, Champ de force, Paramétrage, DFT, MP2.

Acknowledgments

Undertaking this PhD has been a truly life-changing experience for me and it would not have been possible to do without the support and guidance that I received from many people. Here, I would like to thank all those who have contributed to this thesis, directly or indirectly, and who have supported me during these years.

First of all, I would like to thank the members of the jury for my thesis defense, for reviewing my manuscript and for participating to the defense of this thesis. Your professional comments and fruitful discussions have improved the quality of the manuscript.

A very big thank you to all my supervisors Dr. Florent Réal, Pr. Valérie Vallet, Dr. Dominique Guillaumont and Dr. Eléonor Acher for giving me the opportunity to be a part of this interesting project. A huge thank for all the consistent support and encouragement during the three years of my PhD. Without their guidance and constant feedback this PhD would not have been achievable. Their expertise was invaluable in formulating the research questions and methodology. Your insightful feedback pushed me to sharpen my thinking and brought my work to a higher level. Their vast experience and valuable advices have been fundamental in my growth as a scientist. A big thank you for the many discussions we had, allowing me to progress scientifically which largely contributed to the outcome of this work. Moreover, in addition to being one of the experts in the field, I wish to thank them for being such a friendly and kind characters at all times. I am truly thankful to have you as my supervisors.

I extend my sincere thanks to all members of the PhLAM Laboratory-University of Lille, Céline Toubin, Denis Duflot, Andre S. Pereira, Vaibhav Jaiswal, Hanna Oher, Antoine Roose, Sophie Kervazo ... and LILA Laboratory-CEA-Marcoule, Philippe Guilbaud, Philippe Moisy, Marie-Christine Charbonnel, Thomas Dumas, Nathalie Boubals, Claude Berthon, Laurence Berthon, Amoudji Félicité, Elizabith Makombe, Thomas Dirks ... for their welcome and their kindness, and for making these years of thesis, a period as pleasant as it was instructive.

I also thank my friends who have frequently asked me the one-million-dollar question " When are you going to defend your thesis?", Bayaza, Abdelatif, Fadil, Zoheir, Mohammed, Ali, Zakaria ... Despite distressing in times of frequent doubts, it allowed me not to deviate from my ultimate goal. My thanks go to all those who came to support me on January 26th, 2021, and do not go to *Mark Zuckerberg*, who made me waste a lot of time.

I would like to thank all my family for encouraging me and for having always believed in me during my university studies, a very special thank to my mother for her wise counsel and sympathetic ear. You are always there for me. Finaly, I would like to thank my late father who was always supportive, he was the first who supported me and encouraged my passion for science, I owe him everything .... We miss you ...

Contents

List of Figures

List of Tables

Dedicated to my late father ... We miss you ...

Chapter 1

Introduction

> *Nuclear power is one hell of a way to boil water.*
>
> Albert Einstein,

1.1. Nuclear fuel reprocessing

France is considered as one of the leading countries in nuclear power generation, as up to 75% of French electricity production is of nuclear origin, via the exploitation of 56 reactors distributed throughout the territory [1]. This is a result of the 1974 decision (after the 1973 oil crisis). France now has a cognizable level of energy independence and almost the lowest cost for electricity in Europe [2]. France and many "nuclearized" countries have opted for the reprocessing and recycling of spent fuel at the output of nuclear power plants. The principal reason for reprocessing used fuel is to reduce the volume and the level of radioactivity of high-level wastes before their disposal. Another reason is to recover the unused

uranium, along with plutonium, in the used fuel elements and thereby achieve a closed nuclear fuel cycle and plutonium multirecycling. A significant amount of plutonium is recovered from used fuel, and it is then mixed with depleted uranium oxide in a MOX fabrication plant to make fresh fuel.

Nuclear fuel reprocessing dates from the beginning of World War II. During that period, the reactors were designed for the production of plutonium for military uses, in particular to make atomic weapons. Therefore, the only reprocessing required was the extraction of the plutonium (free of fission-product contamination) from the spent natural uranium fuel. The very first process used for this goal, called the bismuth phosphate process, was developed at the US Oak Ridge National Laboratory (ORNL) in the early of 1943. Because of the overwhelming success of next-generation nuclear reactors upon their capability of using uranium from used nuclear fuel, originally designed to use only natural uranium, a great intention was given to the reprocessing of nuclear fuel during the '50s. In 1949, the first successful solvent extraction process for the recovery of pure uranium and plutonium was developed: the PUREX process ?Plutonium and Uranium Recovery by Extraction [3]. Till now, it is the most developed and widely used process in the nuclear industry, currently implemented in the reprocessing factories, namely that of La Hague in France and almost every nuclear fuel reprocessing facility in the world. However, this technique presents certain disadvantages, most of them are coming from the use of TBP (Tri-Butyl-Phosphate) as the extracting molecule:

- The solvent TBP does not respect the CHON principle[4][1]: the presence of a phosphorus atom impose special treatment of the elements produced during the incineration of TBP, secondary phosphorus waste [5].
- Hydrolysis and radiolysis: TBP undergoes significant hydrolytic and radiolytic degradation that can form other complexes with uranium and plutonium soluble in aqueous phase, which are not extractable by TBP [6].
- The U(VI)/Pu(IV) separation: to separate plutonium and uranium from each other, one needs to reduce plutonium, either using as reducing agents tetravalent uranium or hydrazinium nitrate [7]. However,

[1] CHON is an acronym for the four most common elements in living organisms: Carbon, Hydrogen, Oxygen, and Nitrogen. These elements can be incinerated easily, with no ash or acidic gases being produced, in contrast instead of phosphorous or sulfur-containing ones that release toxic gases and need special handling precautions.

PUREX fails to handle solution with high Pu(IV) concentrations, making it necessary to run multiple U(VI)/Pu(IV) separation cycles.

These inconveniences impose the need for multiple cycles to achieve maximum purification of plutonium and uranium. Hence, over the years, many processes have been developed by the modification of different PUREX steps [8, 9]; e.g., they were attempts to eliminate the Pu(IV) reduction step and to reduce the number of solvent extraction cycles[10, 11]. However, most of these processes are quite similar to PUREX and present more or less same limitations. Therefore, there is a strong ongoing research effort, in particular at the CEA, to design new extraction and separation processes, still based on liquid-liquid solvent extraction methods.

1.2. Liquid-Liquid extraction technique

Liquid-liquid extraction, also known as solvent extraction, is a separation process consisting of the transfer of a solute from one solvent to another, the two solvents being immiscible or partially miscible with each other (polar and non-polar). It involves a net transfer of one or more species from one liquid into another liquid phase, generally from aqueous to organic solution. It is a technique widely used in the nuclear field, from the front end to the back end of the nuclear fuel cycle. It is used for uranium recovery from ores and other sources and also the reprocessing of spent nuclear fuel. In the latter industry, the transfer of metal cations, mainly actinides, from an acidic aqueous phase to an organic phase is ensured by a special insoluble extracting molecule. However, this technique is also used in other fields, such as chemical and mining industries and in the downstream recovery of fermentation products (antibiotics, amino acids, steroids) [12, 13].

To characterize the efficiency of a solvent extraction process, a fundamental quantity has been defined, the distribution coefficient, which represents the extracting power of the metallic species M to be extracted by an extracting ligand L:

$$D = \frac{[\mathrm{M}]_{org}}{[\mathrm{M}]_{aq}}, \tag{1.1}$$

where $[\mathrm{M}]_{org}$ and $[\mathrm{M}]_{aq}$ are the equilibrium concentrations of the solute in the organic and aqueous phase,

respectively. The equilibrium is reached when the chemical potential of the extractable solute is the same in both phases.

Another property often used to compare the selectivity of the element of interest M_1 with respect to another element M_2, is equal to the ratio of the distribution coefficients of each element, also known as the selectivity factor (SF), defined as follows:

$$SF_{M_1/M_2} = \frac{D_{M_1}}{D_{M_2}} \tag{1.2}$$

The efficiency of a liquid-liquid extraction process can be characterized and improved by modifying the distribution coefficient; hence, the extracting molecule used.

1.3. Extraction properties of monoamides

To improve and/or develop new separation processes, over the years, different molecules have been used as extracting molecules depending on the field of applications, such as food and pharmaceutical industries, recovery of pollutants in factory effluents and lastly our field of interest the nuclear domain. In the latter, long-chain aliphatic amines [14], dinonylnaphthylsulfonic acid [15], tertiary phosphine sulphides [16], Tri-n-Butyl-Phosphate (TBP) [3], amide derivatives [17, 18], and many other molecules have been used to recover different metals present in the spent fuel for decades. As already mentioned, till now, the most used molecule is the TBP. It is used in PUREX process to extract uranium and plutonium, independent of each other, from the fission products.

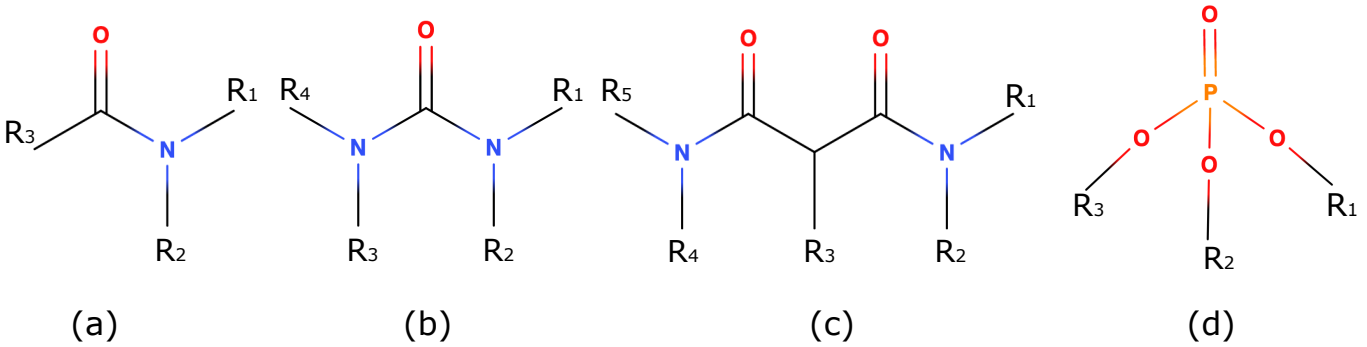


FIGURE 1.1: Structural formula of some extractant ligands used for uranium and plutonium reprocessing, **(a)** Monoamides (N,N-dialkylamides), **(b)** Carbamides (also known as Urea family) **(c)** Maloamides and **(d)** Tri-Butyl-Phosphate (TBP and where $R_{1,2,3,4,5}$ represent organic alkyl groups or hydrogen atoms.)

Considering the very aggressive chemical conditions encountered in the recovery of U and Pu – high concentration of nitric acid are used to dissolve the spent fuel and the strong radiolysis and hydrolysis effects due to the presence of radioactive elements – amide derivatives are regarded as the most promising alternative extractant family to TBP, namely monoamides, maloamides and carbamides [17–21]. They represent an important class of molecules to selectively recover metal ions such as platinum-group metals, gold and the 5f elements [22–26]. Moreover, the monoamides family exhibits some impressive features that overcome the known limitations of TBP; they respect the CHON principle with less radiolysis and hydrolysis problems; their main degradation products are carboxylic acids and secondary amines which hardly affect the separation of U(VI) and Pu(IV) from fission products (FP). Most interestingly, they offer the possibility to (i) co-extract uranium and plutonium from nitric acid solutions and (ii) achieve their separation by decreasing the nitric acid concentration without reducing plutonium(IV). In fact, there is a strong dependence of the extraction properties (distribution coefficient and selectivity) on the chemical conditions, namely the nitric acid and extractant concentrations. In addition, their extraction properties are strongly influenced by their structural features, especially by the size and bulkiness of their alkyl groups. Hence, they offer the possibility to achieve their separation simply by adjusting the nitric acid concentration or tuning the alkyl organic group [27–30].

Because of the keen industrial interest, a large number of N,N-dialkylamides have been synthesized and an extensive database of U(VI) and Pu(IV) distribution ratio between organic and aqueous phases and U(VI)/Pu(IV) selectivity factors is available in literature [28, 31]. Recently, it was shown that carbamides are also capable of such U(VI) and Pu(IV) extraction and separation with comparable features to those of monoamides [32]. This has consequently triggered the interest in amides ligands as good candidates for new and more efficient nuclear reprocessing processes [17–21, 23–26].

1.4. Why molecular modeling?

The understanding and improvement of separation processes require a fundamental knowledge of the chemical mechanisms that come into play at the molecular scale, thus a better knowledge of actinide speciation and molecular organization in the organic and aqueous phases. In fact, it has been shown that most of the physico-chemical properties are linked to the molecular organization of solutions [33]. Experimentalists

have studied these phases with the different existing methods to obtain information on the complexes formed, such as infrared spectroscopy, UV-visible spectrophotometry, time-resolved laser fluorescence spectroscopy (TRLFS), X-ray absorption spectroscopy (EXAFS) and electrospray ionization mass spectrometry. Yet, the radioactive character of the solutions (containing actinides) combined with the complexity of their chemistry (many accessible oxidation states, high coordination numbers, etc.) makes such experimental studies very challenging. Moreover, the number of experimental data available remains limited, and the interpretation of these is often remarkably arduous, due to the possible presence of several species, as well as to the complexity and flexibility of the cations coordination sphere. Hence, molecular modeling has been shown to be very useful for this kind of situation. Thanks to the tremendous advances of algorithms and computer power, molecular simulations have undergone considerable development over the past years, making them complementary to experimental methods. Moreover, we may also carry out simulations on the computer that are difficult or impossible in the laboratory (for example, working at extreme conditions of temperature or pressure). Two types of molecular modeling methods that are widely used to model the liquid state are based either on static quantum mechanical (QM) calculations, or on numerical sampling simulations (molecular dynamics or Monte Carlo simulations). In this work, we are interested particularly on *ab initio* quantum calculations and classical molecular dynamics simulations.

Static *ab initio* methods have been widely used to calculate electronic structures properties of systems containing actinides, including structures, thermochemistry, spectroscopic quantities of various types, and responses to external perturbations. Such information is highly relevant to the understanding of the separation of the radionuclides and their behavior in the gaseous and liquid phase. Experimental methods are also able to measure such properties but with less fundamental understanding of the physical processes taking place at the atomic scale [34, 35]. In the context of the solvent extraction processes under development, such methods have been largely used as to perform complementary studies to calculate and determine the structures of the different complexes present in solution, and characterize the interaction of molecular species with electromagnetic radiation, based on different spectroscopic approaches (IR, Raman, UV-Vis, Fluorescence, EXAFS, etc.) [29, 30, 36, 37], as well as thermodynamics studies to study the stability of the complexes present in solutions [38, 39].

However, the use of such quantum methods is limited to relatively small systems containing up to hundred

atoms in the first coordination sphere. In our systems of interest, the extractant molecules (monoamides) contains long alkyl chains, resulting in a large number of configurations that quantum chemical methods may not be handled straightforwardly, as the computational cost may become prohibitive. Consequently, this restricts QM methods to the study of model systems (ligands with short alkyl chains) only. Moreover, the "solvent" effects are taken into account in an implicit way with the so-called Polarizable Continuum Model (PCM) [40] that describes the solvent as a polarizable dielectric medium. Nevertheless, quantum methods are mostly used to calculate static properties; otherwise the computational cost becomes unaffordable. Further, in our case, when the goal is modeling the migration of actinide ions from one phase to another, and this dynamic process that requires long simulation times. As with classical molecular dynamics simulations, in which the system is treated in a classical way without taking into account the electronic motion, we have the possibility to study much larger systems containing a significant number of atoms and hence the solvent effects are handled explicitly, by including the real solvent molecules. Moreover, MD simulations are considered to be the best alternative to calculate the structural and dynamical properties of actinides in solution [41, 42]. Another interesting feature of MD simulations is the possibility to study the molecular organization beyond the first coordination sphere (known as supra-molecular organization), predict the solvation properties of actinide cations present in solutions, capture the behaviour of the species at a liquid-liquid interface. However, this kind of simulation requires a full description of the interactions between all the elements present in solution, also known as "Force-Field".

1.5. Thesis organization

This PhD work involves the use and development of theoretical approaches to provide insights into the organization at the molecular level of the organic phase containing plutonium in the presence of monoamide extractant molecules. In a first study, we have studied the influence of monoamides alkyl chains nature and length on the stability of plutonium nitrates complexes using static QM simulations based on density functional theory. Experimentally, two complex structures with amide ligands have been characterized in the solid state by crystallography X-ray diffraction (XRD) and in solution by EXAFS, inner- and outer-sphere complexes ($[Pu(NO_3)_4\mathbf{L}_2]$ and $[Pu(NO_3)_6](H\mathbf{L}_2)$[29, 30]. In this study, we have considered both structures,

inner- and outer-sphere complexes. To reduce the computational cost, a series of model ligands has been used, and various alkyl groups are considered in order to compare the influence the alkyl group on the inner- and outer-sphere interactions. The results of this study are presented in Chapter 3.

In order to move toward classical MD simulations for complexes of plutonium (IV) in aqueous and organic phase, parameters describing different interactions between molecules present in solution, shown in Figure 1.2, need to be derived. This is the so-called force-field models (FF). In literature, over the years, many FF describing lanthanides(III), actinides(III), uranyl (UO_2^{2+}) and plutonyl (PuO_2^{2+}) ions, mainly in inorganic phase and few in organic one, have been developed [43–47]. Recently, Acher et al. have developed a consistent FF for the tetravalent actinide series in aqueous phase (An^{4+} – H_2O)[48], but unfortunately, no FFs for Pu(IV) with nitrate and extracting ligands are available. Hence, the ultimate project consists in developing force field (FF) models for Pu(IV) in aqueous and organic phases. This project is a long-term work because of the complexity of the FF parametrization process, especially for the system of interest, containing plutonium and polar molecules (monoamides + nitrates), which requires incorporating special kind of interactions, such as polarization and charge transfer. Acher et al. initiated this effort, with the development of FF for the tetravalent actinide series in aqueous phase, using accurate quantum calculations [48], the latter was motivated by the previous work of Réal et al. [47].

The second part of my PhD work has thus focused on building-up the description of the remaining interactions. More specifically I will present the development of FF models for alkanes and monoamides based on QM calculations as reference data. The evaluation of these potentials was achieved by performing MD simulations of alkanes, monoamides and alkanes + monoamides in the liquid state and comparing the results to experimental physical properties (density, heat of vaporization, diffusion coefficient).

In chapter 2, we describe the principles of quantum chemistry methods (also called *ab initio*) necessary for the calculation of electronic structures. The theoretical study on plutonium complexes is then presented in chapter 3. The theoretical background of molecular dynamics (MD) simulations and FF models as well as the parameterization process is detailed in chapter 4. The development and evaluation of FFs for alkanes and monoamides is discussed in chapter 5. General conclusions, future work directions and perspective are then presented at the end of this manuscript.

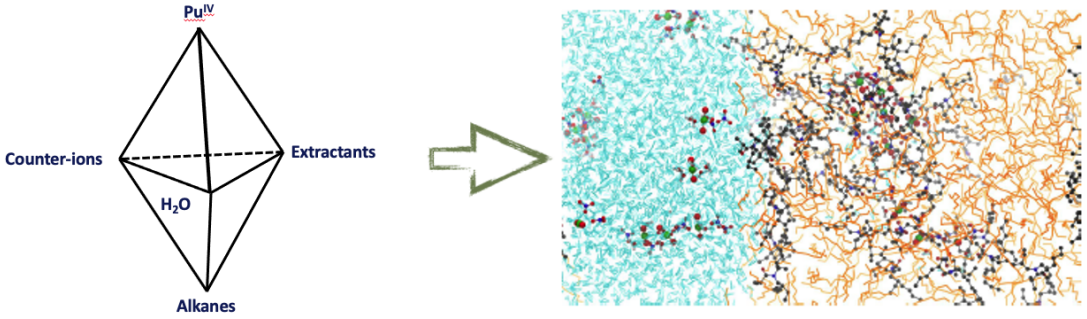


FIGURE 1.2: A figure to illustrate the project of modeling the Pu(IV) extraction mechanisms; on the left, shows the set of pair interactions potentials between the constituents needed for a full-phase simulation; on the right, a figure showing the ultimate goal, modeling the migration of Pu(IV) ion from one phase to another; (in blue: the aqueous phase; in orange: the organic phase, oxygen in red, nitrogen in blue, the actinide ion in green.

Chapter 2

Quantum chemistry methods

> *Quantum mechanics has explained all of chemistry and most of physics.*
>
> Paul Dirac,

In this chapter, we describe the principles of quantum chemical methods (also called *ab initio*) necessary for the calculation of electronic structures. *Ab initio* methods are based on the solution of the stationary Schrödinger equation. These methods allow us to access to the majority of the information of the studied system such as equilibrium geometries, energy, electronic, magnetic, vibrational and thermodynamic properties, etc.

In this work, quantum chemistry calculations allowed us to study the influence of amides alkyl chains nature and length on the stability of Pu(IV) complexes, and also to generate reference data to parametrize the force fields that will be used in classical molecular dynamics simulations. These calculations were carried out with the Turbomole [49] and GAUSSIAN16 [50] quantum chemistry packages.

After having recalled the various electronic structure methods, such as Hartree-Fock HF (Sec. 2.2), the theory of perturbation Møller-Plesset MP (Sec. 2.2.3), and the density functional theory DFT (Sec. 2.3), we will present the atomic orbital basis functions (Sec. 2.4), and relativistic pseudopotentials (Sec. 2.5) used for the calculations. For further details about quantum chemistry methods one could refer, for example, to the textbook ***Szabo and Ostlund*** *"Modern quantum chemistry: introduction to advanced electronic structure theory"* [51].

2.1. The Schrödinger equation

Quantum chemistry calculations allow us to calculate the energy levels and the associated wave function of any molecular system, by solving the time-independent Schrödinger equation:

$$H\Psi(\vec{R},\vec{r}) = E\Psi(\vec{R},\vec{r}), \tag{2.1}$$

with H being the non-relativistic Hamiltonian, E the total energy of the system, Ψ the wave function of the system. $\vec{R}$ and $\vec{r}$ denote, respectively, the nuclear and electronic coordinates in the molecular coordinate system considered.

For a molecular system, the total Hamiltonian sums up different operators:

$$H = T_e + T_N + V_{ee} + V_{NN} + V_{eN}, \tag{2.2}$$

T_e and T_N denoting the kinetic energies of electrons and nuclei, respectively, V_{eN} corresponding to the electrostatic attraction of the electrons by the field of nuclei, and V_{ee}, V_{NN} being the electrostatic repulsion between electrons and nuclei respectively.

The analytical resolution of this equation (2.1) is impossible beyond two particles. It is thus necessary to make a certain number of approximations to be able to solve it in an approximate way.

In the early of the 1920s, approximations were introduced into quantum theory. Among them is the Born-Oppenheimer (BO) approximation. The aim of this approximation is to simplify the Hamiltonian of the

molecular system by separating the movements of the electrons from that of the nuclei, based on the fact that the electrons (mass m_e) are way lighter than the protons (mass m_p) $\left(\frac{m_p}{m_e} = 1836\right)$. This implies that nuclei have a much slower motion than electrons, and leads to consider that the electrons move in a field created by the immobile nuclei. Consequently, the Hamiltonian of the system can be decomposed into two parts, an electronic H_e and a nuclear one H_N:

$$H = H_e + H_N. \tag{2.3}$$

The total wave function of the system can be rewritten as the product of electronic and nuclear wave functions:

$$\Psi(\vec{R}, \vec{r}) = \Psi_e(\vec{r}, \vec{R}).\Psi_N(\vec{R}). \tag{2.4}$$

The electronic wave function is a function of the coordinates of all the electrons in the system but also the positions of nuclei. This means that, for different arrangements of the nuclei, we have a different wave function. From this decomposition we come across two equations, an electronic one and a nuclear one, to be solved separately:

$$He\Psi_e(\vec{R}, \vec{r}) = E_e\Psi_e(\vec{R}, \vec{r}), \tag{2.5a}$$

$$H_N\Psi_N(\vec{R}) = E_N\Psi_N(\vec{R}). \tag{2.5b}$$

In this approximation, the electronic Hamiltonian contains three terms:

$$H_e = T_e + V_{eN} + V_{ee}. \tag{2.6}$$

Since the electronic wave function depends on the position of the nuclei, the associated energies also depend on the position of the nuclei, $E_e = E_e(\vec{R})$. The total energy can be expressed as:

$$E_{tot}(\vec{R}) = E_e(\vec{R}) + V_{NN}. \tag{2.7}$$

If the electronic part is known (it will be discussed later), we can then solve the nuclear part in the same way as the electronic part. As the electrons move much faster than nuclei, it is then possible as a first

approximation, to replace the electronic coordinates by average coordinates on the electronic wave function. The Hamiltonian for the movements of the nuclei in the mean electronic field is then obtained by:

$$H_N = E_e(\vec{R}) + V_{NN} + T_N(\vec{R}), \tag{2.8a}$$

$$= T_N(\vec{R}) + E_{tot}(\vec{R}). \tag{2.8b}$$

The total energy $E_{tot}(\vec{R})$ allows us to obtain the potential describing the movements of the nuclei and allows to build the potential energy surface, a concept commonly used in molecular simulations. The Schrödinger equation describing the motion of nuclei is then written as:

$$H_N \Psi_N(\vec{R}) = \epsilon \Psi_N(\vec{R}), \tag{2.9}$$

where ϵ is the total energy in the BO approximation, including the electronic, vibrational, rotational and translational energy contributions.

However, to solve the electronic Schrödinger equation (Eq. 2.5a), we can distinguish two categories of electronic structure methods :

- Wave Function Theory (WFT),
- Density Functional Theory (DFT).

2.2. Wave Function Theory methods

2.2.1 Hartree-Fock Method

In 1928, Hartree proposed to construct the $n-$electron wave function $\Psi_e(\vec{R}, \vec{r})$ as a product of n mono-electronic orbitals ϕ_i, which leads us to n mono-electronic equations called "Hartree's equations". This approximation is also known as "the orbital approximation".

This approximation has a constraint; it is valid only for a model of independent particles, i.e., the third term of equation 2.6 is neglected. Then, the $n-$electron Schrödinger equation can be divided into n mono-electronic

equations:

$$h_i\phi_i = \varepsilon_i\phi_i, \tag{2.10}$$

where $h_i = T_e(i) + V_{Ne}(i)$ and $\phi_i = \psi_i(\vec{r_i}).\sigma_i(\omega_i)$ is a so-called spin orbital, with $\psi_i(\vec{r_i})$ the spatial function and $\sigma_i(\omega_i)$ the spin function. Thus, the electronic energy of the system for the Hamiltonian $H = \sum_i h_i$ is then $E_e = \sum_i \varepsilon_i$.

Unfortunately, within this approximation, the total wave function does not satisfy the Pauli exclusion principle [1] and it is not antisymmetric by permutation of two particles, here two electrons. To solve this problem, in 1930 Fock and Slater proposed to write the wave function as a determinant constructed from n spin orbital $\phi_n(n)$, known as the "Slater determinant". The peculiarity of this determinant is that for two identical columns which represent two electrons in the same spin orbital, which in reality is impossible (Pauli exclusion principle), would correspond to a zero wave function.The Slater determinant is written as :

$$\Psi_{elec}(r_1, r_2, r_3, ..., r_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(r_1) & \phi_2(r_1) & \phi_3(r_1) & \dots & \phi_n(r_1) \\ \phi_1(r_2) & \phi_2(r_2) & \phi_3(r_2) & \dots & \phi_n(r_2) \\ & & \dots\dots\dots\dots\dots & & \\ \phi_1(r_n) & \phi_2(r_n) & \phi_3(r_n) & \dots & \phi_n(r_n) \end{vmatrix}, \tag{2.11}$$

$\frac{1}{\sqrt{n!}}$ is the normalization factor and where the spin orbitals ϕ_i are functions of the coordinates r_i and the spin coordinates of a single electron.

However, the Hartree-Fock wave function thus defined is only valid for a system of non-interacting particles, as the term of inter-electronic repulsions are neglected. This is far from being the case in a real molecular system. To take this into account, Fock considered that each electron moves in a potential created by the immobile nuclei and by the $(n-1)$ remaining electrons (mean electron field) and therefore the mono-electronic Hamiltonian in equation 2.10 takes the form of the so-called Fock operator F which, for an electron j, is written as:

$$F_i = h_i + V_i^{HF} \tag{2.12}$$

[1] The Pauli exclusion principle is the quantum mechanical principle which states that two or more identical fermions (particles with half-integer spin) cannot simultaneously occupy the same quantum state within a quantum system.

where h_i is the Hartree mono-electronic Hamiltonian and V_i^{HF} is the Hartree-Fock potential describing the interaction of the electron i with the mean field created by the remaining electrons. Since the Fock operator is nonlocal, the Hartree equations become Hartree-Fock equations (HF):

$$F_i\phi_i = \varepsilon_i\phi_i, \tag{2.13}$$

with F_i and ε_i being respectively, the Fock Hamiltonian and the HF energy for the electron i in the spin orbital ϕ_i.

These equations can be interpreted as being the Schrödinger equations for electrons evolving in the field of the nuclei and the other electrons of the system, and whose eigenvalues are the mono-electronic energies ε_i associated with their eigenfunctions, the spin orbitals. If one writes the total wave function as a Slater's determinant, a new term appears, K called the exchange energy term. Using the variational principle, we find that the repulsion energy between the electrons is expressed in the form:

$$E_{rep} = J_{ij} - K_{ij}, \tag{2.14}$$

where J_{ij} is the Coulomb integral which represents the average Coulomb interaction between two electrons located in the orbitals $|\phi_i\rangle$ and $|\phi_j\rangle$ regardless of their spin. K_{ij} is the exchange integral which reduces the Coulomb interaction between the two electrons located in the $|\phi_i\rangle$ and $|\phi_j\rangle$ orbitals having parallel spins (same).

Although the HF approximation represents a great simplification of the Schrödinger equation, these equations remain complicated to solve for most chemical systems. Therefore, each spin orbital is represented by a Linear Combination of Atomic Orbitals (LCAO):

$$|\phi_i\rangle = \sum_{k=1}^{M} C_{ik}\,|\chi_k\rangle \tag{2.15}$$

with M being the number of atomic orbitals (AO) used to describe $|\phi_i\rangle$ (the size of the AO basis) and i takes the values from 1 to n.

It is possible to solve the Hartree-Fock equations by constructing the Fock operator and the Fock matrix[2] which, after diagonalization, lead to the one-electron orbitals. It is thus clear that one would need to know the solutions of the Hartree-Fock equations before one can define the operators that are needed to build these equations. The solutions are therefore obtained by an iterative process. This procedure is repeated until the constructed orbital $|\phi_k^n\rangle$ does not change in the next iteration $\left|\phi_k^{n+1}\right\rangle$; this is the so-called Self-Consistent-Field method (SCF).

2.2.2 Electronic correlation

Even if the Hartree Fock method recovers up to 99% of the exact energy of the system, there is still 1%, which corresponds to the so-called correlation energy, that corresponds to the difference between the exact (non-relativistic) energy of the molecule and the HF energy:

$$E_{corr} = E_{exact} - E_{HF} \tag{2.16}$$

This energy is not taken into account in the Hartree-Fock method since the latter has expressed the N-electron wave function with only one Slater determinant, which is a rather poor representation of a many-electron system's state: in certain cases an electronic state is only well described by a linear combination of more than one Slater determinants. Two categories of electronic correlation effects exist, the dynamic and static correlation. The dynamic one (also known as short-range correlation) refers to the correlation between the movement of electrons coming from Coulomb's inter-electronic repulsion. In fact, the SCF method, in which the inter-electronic repulsions are effectively averaged, ignores this effect. Non-dynamic (static) correlation (long-range correlation) refers to other deficiencies of the HF wave function, such as the inability to describe molecular dissociation, or systems in which the electronic state has to be described by a linear combination of several nearly degenerate Slater determinants.

The evaluation of such correlation energy is one of the major challenges in quantum chemistry. There are two types of methods which take into account electronic correlation, namely post-Hartree-Fock methods in the framework of WFT, and Density Functional Theory methods.

[2]This matrix can be written as : $F_{pq} = \langle \xi_p | F | \xi_q \rangle$

2.2.3 Post-Hartree-Fock methods

There are three main post-Hartree-Fock methods, Møller-Plesset perturbative methods (MPn), Coupled Cluster approaches and configuration interaction (CI) methods. We present here, briefly, the second-order Møller-Plesset perturbation method (MP2) [52], which we have preferably used in this thesis. This method estimates the correlation of electrons as a perturbative correction to the HF problem. The MP theory is based on the Rayleigh-Schrödinger perturbation theory (RS-PT), considering an undisturbed Hamiltonian (H_0) to which a small (weak) perturbation (V) is added:

$$H = H_0 + \lambda V \tag{2.17}$$

Here λ is an arbitrary real parameter that controls the magnitude of the perturbation. In MP theory, the zeroth-order wave function is the exact eigenfunction of the Fock operator. The perturbation therefore shall recover the correlation energy. If one writes the total wave function $|\Psi\rangle$ and the total energy E as Taylor series of the λ perturbative factor:

$$|\Psi\rangle = \left|\Psi^{(0)}\right\rangle + \lambda \left|\Psi^{(1)}\right\rangle + \lambda^{(2)} \left|\Psi^{(2)}\right\rangle + ... \tag{2.18a}$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^{(2)} E^{(2)} + ..., \tag{2.18b}$$

the electronic Schrödinger equation (eq. (2.5a)) becomes:

$$\begin{aligned} &(H_0 + \lambda V)\left(\left|\Psi^{(0)}\right\rangle + \lambda \left|\Psi^{(1)}\right\rangle + \lambda^2 \left|\Psi^{(2)}\right\rangle + ...\right) \\ =&(E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + ...)\left(\left|\Psi^{(0)}\right\rangle + \lambda \left|\Psi^{(1)}\right\rangle + \lambda^2 \left|\Psi^{(2)}\right\rangle + ...\right) \end{aligned} \tag{2.19}$$

Looking at this equation order by order in λ, we obtain

$$H_0 \left|\Psi^{(0)}\right\rangle = E^{(0)} \left|\Psi^{(0)}\right\rangle \tag{2.20a}$$

$$H_0 \left|\Psi^{(1)}\right\rangle + V \left|\Psi^{(0)}\right\rangle = E^{(0)} \left|\Psi^{(1)}\right\rangle + E^{(1)} \left|\Psi^{(0)}\right\rangle \tag{2.20b}$$

$$H_0 \left|\Psi^{(2)}\right\rangle + V \left|\Psi^{(1)}\right\rangle = E^{(0)} \left|\Psi^{(2)}\right\rangle + E^{(1)} \left|\Psi^{(1)}\right\rangle + E^{(2)} \left|\Psi^{(0)}\right\rangle \tag{2.20c}$$

The MP perturbation theory consists in applying the perturbation computation in the case where H_0 is defined as the sum of monoatomic Fock operators:

$$H_0 = \sum_i F_i \tag{2.21}$$

By projecting the equations 2.20b and 2.20c on the bra state $\left\langle \Psi^0 \right|$, we obtain

$$E^{(1)} = \left\langle \Psi^{(0)} \middle| V \middle| \Psi^{(0)} \right\rangle \tag{2.22a}$$

$$E^{(2)} = \left\langle \Psi^{(0)} \middle| V \middle| \Psi^{(1)} \right\rangle \tag{2.22b}$$

Projecting now the equation 2.20b on the bra state $\left\langle \Psi^{(n)} \right|$, for $n \neq 0$ and using $\left\langle \Psi^{(n)} \right| H = E^{(n)} \left\langle \Psi^{(n)} \right|$ and $\left\langle \Psi^{(n)} \middle| \Psi^{(0)} \right\rangle = 0$ we arrive at:

$$\left| \Psi^{(1)} \right\rangle = - \sum_{n \neq 0} \frac{\left\langle \Psi^{(n)} \middle| V \middle| \Psi^{(0)} \right\rangle}{E^{(n)} - E^{(0)}} \left| \Psi^{(0)} \right\rangle \tag{2.23}$$

The second-order energy correction $E^{(2)}$ will therefore be finally expressed as:

$$E^{(2)} = E^{(2)}_{MP} = \sum_{n \neq 0} \frac{| \left\langle \Psi^{(n)} \middle| V \middle| \Psi^{(0)} \right\rangle |^2}{E^{(0)} - E^{(n)}} \tag{2.24}$$

Finally, the second-order Møller-Plesset (MP2) total energy is expressed as the sum of the HF energy and the MP2 correction:

$$E_{MP2} = E_{HF} + E^{(2)}_{MP} \tag{2.25}$$

Higher orders of perturbation such as MP3 or MP4 (with the third and the fourth order energy) allow to include more dynamical correlation but they are rarely used due to their cost. However, it has been shown that high-order MP theory may not always converge, converge slowly or even oscillate. This is due to the fact that the Hamiltonian may not represent a small perturbation, which would lead to a very erratic or nonexistent convergence if HF is not a good starting point, and from the non-variational nature of the method, which leads to oscillations in the calculated properties that make it difficult to extrapolate [53]. In general, they

have a similar cost to other approaches such as coupled cluster or configuration interaction (not discussed).

2.3. Density functional theory DFT

Density Functional Theory offers a different approach to deal with electronic correlation. It uses the electron density $\rho(r)$ instead of the $n-$electron wave function as the system variable. Electron density is, in principle, an observable physical property of molecules, unlike the wave function which is a mathematical tool without any physical meaning. The electronic energy, $E[\rho]$, is a functional (a function of a function) of electron density, with ρ the electron density, a function of three variables (x, y, z).

The Density Functional Theory is based on the two theorems of Hohenberg and Kohn [54]. In the first theorem, they showed that there is an exact correspondence between the electron density and the external potential of a given physical system; the external potential (V_{ext}) is determined with a single density to within a constant.

$$E[\rho] = \int V_{ext}(r).\rho(r).dr + F(\rho(r)), \tag{2.26}$$

where $\rho(r)$ is the electron density, $F(\rho(r))$ is a universal density function which contains the kinetic and Coulomb contribution to the energy.

In the second theorem (variational principle), they proved the existence of a universal functional $E[\rho(r)]$ expressing the total energy as a function of the electron density, valid for any external potential $V_{ext}(r)$. The energy of the ground state of the system is the value which minimizes this functional and the corresponding density is the exact density of the ground state. Therefore the total energy of the system which is a functional of the electron density takes the following form :

$$E[\rho] = T[\rho] + \int V_{ext}(r).\rho(r).dr + \frac{1}{2}\int \frac{\rho(r)\rho(r')}{|r-r'|}drdr' + E_{XC}[\rho], \tag{2.27}$$

where $T[\rho]$ and $E_{XC}[\rho]$ are kinetic energy and exchange-correlation energy, respectively.

The Hohenberg-Kohn theorems basically state that an exact density functional exists, however, they don't say anything about how to get it. The formulas of kinetic energy and exchange-correlation remain unknown.

The Kohn-Sham formalism is used for this purpose, using mono-electronic equations analogous to the Hartree-Fock equations [55]. From these, it is in principle possible to obtain the electron density of a system and therefore its total energy. The energy of a system can be written as follows:

$$E[\rho] = -\frac{1}{2}\sum_{i=1}^{n_{occ}} \int \theta_i^*(r)\nabla^2\theta_i(r)dr - \sum_{I=1}^{N}\int \frac{Z_I}{|r-R_I|}\rho(r)dr + \frac{1}{2}\int \frac{\rho(r)\rho(r')}{|r-r'|}drdr' + E_{XC}[\rho], \tag{2.28}$$

where θ_i are the Kohn-Sham orbitals (KS), solutions of the equations given below (Eq. 2.30). The exact charge density of the ground state is given by:

$$\rho(r) = \sum_{i=1}^{n_{occ}} |\theta_i(r)|^2. \tag{2.29}$$

The first term of the equation 2.28 is the kinetic energy of the electrons. The second term is electron-nucleus attraction, the third term represents the Coulomb interaction between the total charge distribution at positions r and r'. The last term is the exchange-correlation term which is also a functional of the density, and takes into account electron-electron interactions.

The KS orbitals are obtained via solving the Kohn-Sham equations by applying a variational principle to the energy $E[\rho]$. In general they are solved by the self-consistent field method (SCF). These equations take the following form:

$$\left[-\frac{1}{2}\nabla^2 - \sum_{I=1}^{N}\frac{Z_I}{|r-R_I|} + \frac{1}{2}\int \frac{\rho(r')}{|r-r'|}dr' + V_{XC}[r]\right]\theta_i(r) = \varepsilon_i\theta_i(r), \tag{2.30}$$

Where the ε_i are the energies of the KS orbitals, and V_{XC} is the exchange-correlation potential which is the functional derivative of the exchange-correlation energy $E_{XC}[\rho]$. The analytical form of this functional is still unknown. It can be approximated using different physical arguments, leading to three main categories of exchange-correlation functionals, the Local Density Approximation (LDA), the Generalized Gradient Approximation (GGA) and the Hybrid Approximation (HA).

The LDA functional family is, historically, the first approximation that has been proposed. LDA method

relies on the fact that the local density, within a small volume, can be assumed as homogeneous, and that the contribution in the total exchange-correlation energy could be calculated as the product of the small volume and the exchange-correlation energy density from the homogeneous gas theory that is calculated inside the small volume(uniform electron gas model). It is a pretty popular and practical method, employed extensively by solid-state physicists in *ab initio* DFT studies to interpret electronic and magnetic interactions in materials and semiconductor materials. Still, LDA has been shown in the literature to yield poor agreements with experimental data for liquid state studies because of the bad treatment of weak interactions [56]. The exchange-correlation energy ($E_{XC}[\rho]$) for a particle in a uniform electron gas is expressed as:

$$E_{XC}[\rho] = E_X[\rho] + E_C[\rho], \tag{2.31}$$

where $E_X[\rho]$ exchange energy per particle of the uniform electron gas can be analytically expressed as:

$$E_X[\rho] = C_X \rho(r)^{1/3} \tag{2.32}$$

with $C_X = -\frac{9}{4\pi}$. As for the correlation energy per particle $E_C[\rho]$ of the uniform electron gas, it cannot be calculated analytically. This quantity has been obtained numerically for a number of densities ρ using accurate quantum Monte Carlo calculations, and fitted to a parametrized function of ρ satisfying the known high- and low-density expansions [57]. Depending on the parametrization protocol, we can find different correlation potentials, such as Vosko, Wilk, and Nusair (VWN) [58] and the one of Perdew and Wang (PW92)[59].

The second XC category is the GGA, which was introduced to correct the defaults of LDA, notably taking into account the non-uniformity of the electron density. It is also widely used in the solid-state field, and it gives good results for molecular geometries and ground-state energies [56, 60]. GGA XC potentials depend not only on the density, but also its first derivative (gradient):

$$E_{XC}^{GGA}[\rho] = E_X[\rho] + E_C[\rho] = E_{XC}^{LDA}[\rho] + f_{XC}(\rho, \nabla\rho), \tag{2.33}$$

where $f_{XC}(\rho, \nabla\rho)$ is the energy function taking into account the density and its gradient. There are many

different prescriptions for choosing the function f, each of one leading to distinct GGA's. We can quote the Becke functional (B88), the Lee, Yang and Parr functional (LYP) and the Perdew, Burke and Ernzerhof functional (PBE). Potentially more accurate than the GGA functionals are the meta-GGA functionals that includes the second derivative of the electron density (the Laplacian).

The last XC category, HA, stands for hybrid functional. It was first introduced by Becke [61]. This kind of functional incorporates a portion of exact exchange from Hartree-Fock theory with the rest of the exchange-correlation energy from other sources (LDA and GGA). They are one of the most used methods in the literature, giving a very good agreement with experimental properties, especially in the condensed phase. The best known within these hybrid functionals are the B3LYP [62] and PBE0 [63] functionals:

$$E_{XC}^{B3LYP} = (1-a)E_X^{LSDA} + aE_X^{HF} + b\Delta E_X^{B88} + (1-c)E_C^{LSDA} + cE_C^{LYP}, \tag{2.34}$$

with the parameters $a = 0.2, b = 0.7$ and $c = 0.8$, corresponding to 20% of HF exchange. The PBE0 functional consists on adding 25% of the HF exchange energy to the PBE functional:

$$E_{XC}^{PBE0} = \frac{1}{4}E_X^{HF} + \frac{3}{4}E_X^{PBE} + E_C^{PBE}, \tag{2.35}$$

During the mid-2000s, it was recognized that the local (semi-local) density functionals do not properly capture dispersion interactions [64]. Three classes of approach currently are in use to correct the DFT energy, the nonlocal van der Waals Density Functional (vdW-DFs) [65], dispersion-correcting atom-centered one-electron potentials [66] and last the DFT-D method proposed by Grimme et al. [67] to account for dispersion interactions with a computationally cheap and simple solution. The latter correction, which is used in our work, consists in adding a standard semi-empirical correction term to any of the approximate functionals already discussed, leading to the so-called dispersion-corrected DFT (DFT-D) [67]. The dispersion correction term reads

$$E_{disp} = -s\sum_{\alpha<\beta} f(R_{\alpha\beta})\frac{C_6^{\alpha\beta}}{R_{\alpha\beta}^6}, \tag{2.36}$$

where $C_6^{\alpha\beta}$ are the pairwise atomic dispersion coefficients, $R_{\alpha\beta}$ is the distance between each pair of atoms, $f(R_{\alpha\beta})$ is a parametrized damping function which converges to 1 at large $R_{\alpha\beta}$ distance values and zero for

small $R_{\alpha\beta}$ values. Lastly, s is a possible scaling parameter that pairs with each approximate XC functional, e.g., $s = 0.75$ for PBE, $s = 1.2$ for BLYP, $s = 1.05$ for BP86, and $s = 1.05$ for B3LYP [68]. Different DFT-D methods have been developed since the 2000s, the difference between these being either in the use of different damping functions or by adding dispersion corrections of higher orders than R^6, such as Chai and Head-Gordon method [69]. Grimme et al. have developed two methods GD2 [67] and GD3 [70]; the first method includes only $C_6^{\alpha\beta}$ two-body terms while GD3 also includes the three-body terms $C_9^{\alpha\beta\gamma}$. Different versions of these methods are available depending on the parameters and functions employed.

2.4. Atomic basis sets

The numerical resolution of the equations describing the electronic structure requires the expansion of atomic orbitals χ_k on a basis of simple analytical functions f_v:

$$\chi_k(r, R) = \sum_v C_{kv} f_v(r, R), \tag{2.37}$$

where C_{kv} are coefficients and f_v are the atomic basis functions. The most commonly used basis sets for quantum chemistry calculations consist of a radial part and an angular part, for which there are two types: the Slater type basis STO (Slater Type Orbitals) and the Gaussian type basis GTO (Gaussian Type Orbitals):

$$f^{STO}(r, R) = g(\theta, \varphi) \exp(-\xi r) \tag{2.38a}$$

$$f^{GTO}(r, R) = g(\theta, \varphi) \exp\left(-\alpha r^2\right), \tag{2.38b}$$

that differ by their radial part. In both STOs and GTOS, the angular part $g(\theta, \varphi)$ is represented by a spherical harmonic $Y_{ml}(\theta, \varphi)$ and where (θ, φ) are the angular coordinates and $r = |r - R_P|$ where R_P corresponds, in general, to the position of the atom $(R_P \equiv R)$. ξ and α are constant and kept fixed during the calculation of the electronic structure. Only the orbital coefficients C_{kv} need to be optimized. The advantage of GTOs over STOs is that integrals are calculated faster with GTOs than with STOs. However, the behavior of these GTO basis functions is qualitatively not correct at the nucleus level as well as at long distance, due to their analytical forms (see Fig.2.1). To approximate the analytical form of STOs, several GTOs basis

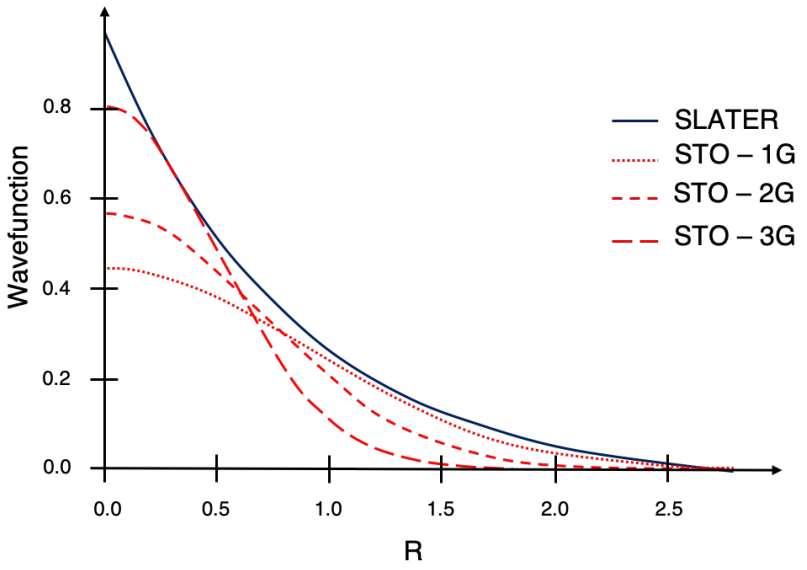


FIGURE 2.1: Illustration of STO and GTO and the approximation of a Slater function by a two and three Gaussians.

are, in general, grouped together thus forming contracted Gaussian functions. Each of these contracted Gaussian functions is a linear combination of primitive Gaussian functions centered on the same atom. Dozens basis sets are available on the literature, by increasing the size with additional functions to describe polarization and diffusion of the electron density of the atom in molecules, giving a controlled way to obtain more accurate solutions but at a higher cost. Such as Pople basis sets [71], Karlsruhe basis sets [72] and Correlation-consistent basis sets [73]. However, the choice of which basis set to use for the calculations is usually a compromise between precision and cost. In our case, we have preferably used the correlation-consistent polarized basis sets of Dunning and coworkers, aug-cc-pVTZ [74] for the generation of quantum chemistry reference data and Karlsruhe basis sets, def2-TZVP, for the theoretical study of plutonium nitrates complexes.

2.5. Relativistic effects and Effective core potentials

The regions near the nuclei, the "core regions", are composed primarily of tightly bound core electrons which respond very little to the presence of neighboring atoms, while the remaining volume contains the valence electron density which is involved in the interatomic bonding. Hence, in order to reduce the computational cost of all-electron calculations for elements containing a lot of electrons, namely heavy elements, the

core electrons and the potential due to the charge of the nuclei can be replaced by an effective potential. This approach consists in assuming that there is no significant overlap between the core and valence wave functions. The side advantage is that it is possible to get rid of the atomic basis functions describing the core orbitals, thus reducing the computational cost. In the effective core potential approach, only the chemically active valence electrons are dealt with explicitly, because they are the only ones involved in the establishment of chemical bonds, while the core electrons are 'frozen'. There are several pseudopotential formalisms which differ in whether or not conserving the charge in the core region, for further details see [75–78]

When dealing with heavy atoms, relativistic effects become quite relevant [79]. There are namely two types of relativistic effects: the scalar relativistic effects and the spin-orbit coupling. In a simplified view, the scalar relativistic effects result from the variation of mass with the velocity of the electrons, which is roughly proportional to the nuclear charge Z. Thus in heavy elements, the scalar relativistic effects lead to a contraction of the core s and p orbitals and in the expansion of the d and f orbitals. Accordingly, a mass-velocity correction is applied to the kinetic energy of the electrons.[3]. The spin-orbit coupling originates from the interaction of the spin of the electron with this magnetic field due to the relative motion of the charges.

For a heavy atom (as heavy we refer to all atoms beyond the last two rows of the periodic table), the pseudopotential that replaces the core electrons must account for these relativistic effects. These are the so-called Relativistic Effective Core Potentials [77, 78], the parameters of which are adjusted to reproduce all-electron relativistic calculations. The RECPs differ by the choice of the all-electron reference method and the number of core electrons modeled. Namely in the case of Pu, one can choose a "small-core" RECP, ECP60MWB, that was adjusted to multiconfigurational Wood-Boring relativistic calculations and replaces 60 core electrons, leaving the valence 5s, 5p, 5d, 5f, 6s, 6p, 7s electrons active. Alternatively, the "large 5f-in-core" pseudopotential have been optimized for the tetravalent oxidation state of Pu, that incorporates the four unpaired $5f^4$ electrons into the core, thus turning Pu(IV) into a fictive closed-shell atom. The latter ECP82MWB was used in Chapter 3, knowing it yields very good accuracy for structural parameters and energetics [80].

[3]The relativistic mass increases as a function of the particle velocity: $m = \frac{m_0}{\sqrt{1-(v/c)^2}}$

2.6. Solvent effects

The majority of chemical reactions take place in solution, and the effects due to the solvent can be very significant on the electronic structure. Hence, it is important to take them into account, either explicitly by placing solvent molecules around the molecule or implicitly by using what's known as Polarizable Continuum Model (PCM). In such models the solute molecule is placed in a cavity surrounded by a dielectric polarizable continuum, whose response modifies the energy and the properties of the solute (Fig 2.2).

The molecular free energy of solvation is therefore computed as the sum of three terms:

$$G_{solv} = G_{elec} + G_{vdw} + G_{cav} \tag{2.39}$$

where G_{elec} accounts for the interaction energy between the solute and solvent, G_{vdw} for the Van der Waals interactions and lastly the G_{cav} for the free energy required to create a cavity in the solvent.

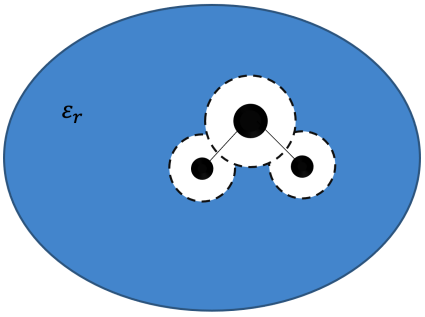


FIGURE 2.2: The Polarizable continuum model. The solute molecule (water) is embedded in a cavity which is formed in the continuum with a dielectric constant of ϵ_r .

A variety of solvent continuum models exist in literature, depending on the form of cavity used and the manner with which the different terms of the equation 2.39 are evaluated [81]. In this work, we have used the PCM method implemented in Gaussian16 package [82].

Chapter 3

Quantum chemical study of plutonium nitrates complexes with amides derivatives

Amide derivatives represent an important class of molecules to recover selectively metal ions such as platinum-group metals, gold or 5f elements by solvent extraction [23–26]. Tertiary N,N-dialkyl amides are particularly well known for their strong extraction ability of uranium(VI) and plutonium(IV) and are regarded as a promising alternative extractant family to Tri-n-Butyl-Phosphate (TBP) [17–20, 83]. A key property of amide derivatives is that it is possible to co-extract uranium and plutonium from nitric acid solutions and to further achieve U/Pu separation by decreasing the nitric acid concentration [27]. More recently, it was shown that carbamides are also capable of achieving such U(VI) and Pu(IV) extraction and separation [32]. Because of their keen industrial interest, a large number of N,N-dialkylamides have been synthesized and an extensive database of U(VI) and Pu(IV) distribution ratio between organic and aqueous phases and U(VI)/Pu(IV) selectivity factors is available. The strong dependence of the extraction properties on the chemical conditions, namely nitric acid and extractant concentrations, has been established.

Moreover, one of their most interesting feature is that it is possible to adjust their physico-chemical and extraction properties by altering the alkyl chains on nitrogen or carbonyl sides. Extraction properties are particularly influenced by the size and bulkiness of their alkyl groups adjacent to the carbonyl group. Variations of extraction properties are particularly strong for plutonium. This was first discovered in the 1960s by Siddall et al. [18]. This opens up the possibility to tune their extracting strength and selectivity toward metal ions by altering the length and branching of amide alkyl groups. Siddall et al. [18] have demonstrated that a simple ramification of the acyl group can suppress Pu(IV) extraction [29, 84, 85]. In the series R = CH_3, C_2H_5, $CH(CH_3)_2$, $C(CH_3)_3$, Pu(IV) distribution ratio D_{Pu} continuously drops from 21 (R = CH_3) to 0.001 (R = $C(CH_3)_3$), while U(VI) distribution coefficients decrease much less significantly from 9.9 to 0.6 (upon extraction from a 3 M nitric acid aqueous solution to a toluene organic phase with 0.5 M of amide extractant) [18]. Accordingly, U/Pu separation factor increases from 0.5 to ~600. Since this early work, monoamides have been investigated extensively for the separation of hexa- and tetravalent actinide ions. Monoamides with long alkyl chains that prevent third phase formation have been developed. DEHiBA (N,N-di(2-ethylhexyl)isobutyramide) and DEHBA (N,N-di(2-ethylhexyl)butyramide) are probably the most investigated derivatives so far [83, 86–90]. As reported for the earlier monoamides with shorter alkyl chains, the branching on the acyl group for DEHiBA suppress plutonium(IV) extraction compared to DEHBA.

Notwithstanding their industrial interest, the strong influence of N,N-dialkylamide structure on extracting properties is yet to be rationalised. To the best of our knowledge, there is no clear explanation for what makes the monoamide structure have such a great impact on the extraction ability. For decades, it was assumed that amide derivatives operate through a solvating mechanism as in the PUREX process with TBP, two amide ligands bind Pu(IV) and U(VI) in the inner coordination sphere, which are predominantly extracted as $Pu(NO_3)_4L_2$ and $UO_2(NO_3)_2L_2$ [20]. The strong variation of Pu distribution coefficient as a function of the bulkiness of the acyl group was attributed to steric hindrance in the Pu coordination sphere. However, recent studies have revealed that plutonium extraction with amide derivatives is more complex than with TBP [29, 91]. By coupling UV–vis spectroscopy, EXAFS and quantum chemistry calculations, it has been shown that the plutonium hexanitrato complex $Pu(NO_3)_6^{2-}$ can also be extracted in the organic phase with no amide ligand present in the first coordination sphere. It was deduced that two protonated amide ligands bind in the outer coordination sphere to ensure the charge neutrality and the stability of the complex in the

organic phase. However the detailed structure of the outer-sphere species is not known and could not be characterized from experimental data. It was shown that the Pu(IV) coordination structure can switch from inner to outer complexation either by varying the experimental conditions (such as nitric acid concentration) or by slightly altering the amide alkyl group [29, 30, 92, 93]. Furthermore, the recent study of Pu(IV) with linear alkyl chains carbamides demonstrated that carbamide are strong extractants at high nitric acidity while promoting outer-sphere coordination. As a result, the strong variation of extraction properties cannot be solely rationalized from steric hindrance in the inner coordination sphere. As found for the recovery of gold with amides, outer-sphere interactions appear to be of primary importance [25]. Understanding the relationship between the structure of the ligands and their extraction ability is crucial to design new extraction processes.

The aim of this work is to explore and rationalize Pu(IV) inner and outer-sphere interactions with amide or carbamide derivatives from Density Functional Theory (DFT) calculations, that have proven to be reliable at predicting the geometries of Pu(IV) complexes [30, 94–96]. While the structures of the inner-sphere complexes with amide and carbamide ligands have been well characterized in the solid state from XRD and in solution from EXAFS, the structures of outer-sphere complexes are largely unknown, though carbamides are believed to promote the formation of outer-sphere complexes [32]. DFT geometry optimizations shall help shedding light on their 3D structures.

Herein, we consider that the ligands extract Pu(IV) from nitric acid according to the following reactions:

$$\mathrm{Pu^{4+}(aq) + 4\,NO_3^-(aq) + 2\,L\,(org) \rightleftharpoons [Pu(NO_3)_4L_2](org)} \tag{3.1}$$

$$\mathrm{Pu^{4+}(aq) + 6\,NO_3^-(aq) + 2\,H^+\,(aq) + 2\,L\,(org) \rightleftharpoons [Pu(NO_3)_6](HL)_2(org)} \tag{3.2}$$

in which the complexing ligand is noted L; (aq) and (org) stand for the aqueous phase and the organic phase, respectively. Reaction 3.1 corresponds to the formation of an inner-sphere complex $[Pu(NO_3)_4L_2](org)$ with four bidentate nitrate anions and two ligands. Reaction 3.2 corresponds to the formation a Pu(VI)-hexanitrato core $Pu(NO_3)_6^{2-}$ which is charge-balanced by protonated HL^+ attached by hydrogen bonds (outer-sphere complexation). Note that we disregard the more hypothetical possibility of forming a complex with a hybrid coordination $Pu(NO_3)_5L(HL)$, corresponding to $Pu(NO_3)_5L^-$ with one HL^+ protonated ligand

hydrogen-bonded to one nitrate ion.

To rationalize the influence of the alkyl groups on Pu(IV) complexation, the nature of the alkyl groups in the monoamide (R_1–R_3) and the carbamide (R_1–R_4) ligands drawn in Figure 3.1 have been systematically varied following the list and labelling given in Table 3.1. PEE and IEE ligands can be considered as model ligands for the two isomeric DEHBA and DEHiBA monoamides.

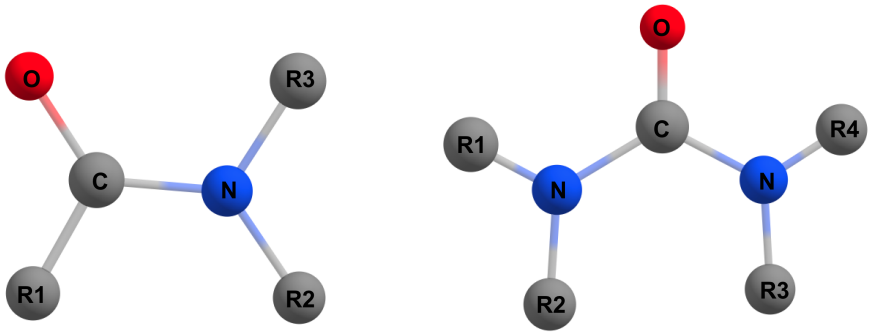


FIGURE 3.1: Labelling of the R_1, R_2, R_3 and R_4 alkyl groups in the monoamide (left) and carbamide (right) ligands, as listed in Table 3.1.

TABLE 3.1: List of aliases for the substituted monoamide and carbamide ligands.

Ligands	R_1	R_2	R_3	R_4
MMM	CH_3	CH_3	CH_3	-
EMM	C_2H_5	CH_3	CH_3	-
MME	CH_3	CH_3	C_2H_5	-
MEE	CH_3	C_2H_5	C_2H_5	-
EEE	C_2H_5	C_2H_5	C_2H_5	-
PEE	C_3H_7	C_2H_5	C_2H_5	-
IEE	$CH(CH_3)_2$	C_2H_5	C_2H_5	-
C4M	CH_3	CH_3	CH_3	CH_3
C4E	C_2H_5	C_2H_5	C_2H_5	C_2H_5

3.1. Computational Methods

All molecular geometries have been fully optimized using Gaussian quantum chemistry package at the Density Functional Theory (DFT) level [50], employing the PBE0 functional [63], that has been shown to yield accurate structural parameters of the $[Pu(NO_3)_6]^{2-}$ complex [96, 97]. For each ligand, 4 to 10 initial

conformations for $[Pu(NO_3)_4\mathbf{L}_2]$ and $[Pu(NO_3)_6](H\mathbf{L}_2)$ complexes were constructed and optimized in the gas phase in order to determine the lowest energy structure. Def2-TZVP basis set has been used for light atoms (H, N, O, and C) [98]. For plutonium, the 82 inner-shell core electrons (78 inner-shell electrons as well as the $5f^4$ unpaired electrons) were replaced by the relativistic 5f-in-core large-core pseudopotential (LPP) noted ECP82MWB [99], together with the corresponding valence basis set ECP82MWB-AVTZ [99]. The use of an LPP for actinides eliminates the difficulties arising from the open 5f-shells ($5f^4$ valence configuration for Pu(IV)) that may yield a multi-reference character for the ground state with spin contamination issues. It is also trustworthy as it offers satisfactory results for actinide-ligand bond distances and binding energies, with deviations to small-core results not larger than 0.025 Å (1.2%), and 0.92 eV (0.6%), for actinocene complexes [100]. To confirm the applicability of LPP to our systems, test calculations were performed using a plutonium small-core pseudopotential noted ECP60MWB and accompanying basis set.35Energy variations for reaction (1) and (2) are compared in Table S1 in annex A section. The values differ by less than 5 $kJ \cdot mol^{-1}$ and confirm the judicious choice of a 5f-in-core pseudopotential for the systems of interest.

Core polarization effects described through a core-polarization potential (CPP) were shown to be small [99] and can be safely disregarded. We have also neglected spin-orbit as Sulka et al. [96] found a small effect of 0.014 Å on the $Pu-O_{NO3}$ bond length in the Pu(IV)-hexanitrato complex. Furthermore, as we are interested in relative complexation energies, we can forecast that the absolute errors induced by neglecting core-polarization effects and spin-orbit coupling will compensate when taking energy differences.

When calculating relative complexation free energies for various monoamide or carbamide ligands, it is relevant to account for long-range dispersion contributions, that might be sizable in the outer-sphere complexes ($[Pu(NO_3)_6](HL)_2(org)$). We have thus performed single-point energy calculations with two approaches (i) by adding to the PBE0 energies the Grimme's DFT-D3 dispersion correction [101] (ii) by using the MP2 wave-function based method as implemented in the Turbomole quantum chemistry package [102]. The later calculations used aug-cc-pVTZ [74] basis sets for the light atoms and the ECP82MWB-AVTZ [99] for Pu, together with the resolution of identity [103]with the corresponding auxiliary basis sets. [98].

To mimic experimental conditions, in which the Pu metal ions are extracted into an organic phase that contains the extractant molecules, nitric acid and aliphatic solvent, the impact of solvation effects on the complex

geometries and relative complexation energies was explored using a Polarizable Continuum solvation Model (PCM) as implemented in Gaussian package (See Ref. [104] for a review on solvation models), for two solvents *n*-dodecane ($\epsilon_r = 2.02$) and Dimethylacetamide DMA ($\epsilon_r = 37.78$).

Vibrational frequencies calculation for the ligands and complexes were performed to verify that geometries were minima and to compute zero point energies and thermal corrections using the rigid rotor-harmonic oscillator approximation, as implemented in Gaussian, either in the gas phase or in solution.

Finally, the affinity of the model extractants is compared by calculating the energy for the following ligand exchange reactions :

$$[\mathrm{Pu(NO_3)_4(\mathbf{L1})_2}] + 2\,\mathbf{L2} \rightleftharpoons [\mathrm{Pu(NO_3)_4(\mathbf{L2})_2}] + 2\,\mathbf{L1} \tag{3.3}$$

$$[\mathrm{Pu(NO_3)_6}](\mathrm{H}\mathbf{L1})_2 + 2\,\mathbf{L2} \rightleftharpoons [\mathrm{Pu(NO_3)_6}](\mathrm{H}\mathbf{L2})_2 + 2\,\mathbf{L1} \tag{3.4}$$

For instance, the relative complexation energy $\Delta\Delta E$ reads

$$\Delta\Delta E = E(\mathrm{Pu}_{\mathrm{NO_3}}\mathbf{L2}) - E(\mathrm{Pu}_{\mathrm{NO_3}}\mathbf{L1}) - 2(E(\mathbf{L2}) - E(\mathbf{L1})) \tag{3.5}$$

3.2. Results and Discussion

All Pu(IV) complexes with amide ligands reported in Table 3.1 were considered. The calculations were performed in the gas phase for all the ligands and in solution for some selected ligands. In the industrial process, the organic phase contains extractant molecules in an aliphatic solvent. After contact with the acidic aqueous phase, some various amounts of extracted nitric acid is also present in the organic phase. *n*-dodecane is the best model for the aliphatic solvent but the presence of amides and nitric acid is expected to increase the polarity of the solution. Therefore, the DMA solvent model was also used to investigate the effect of increasing the polarity of the solution.

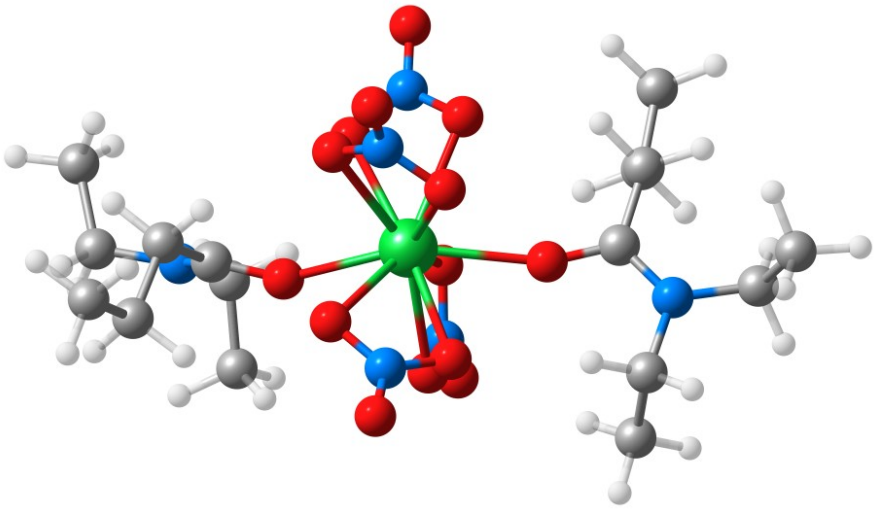


FIGURE 3.2: Optimized geometry of the inner-sphere $Pu(NO_3)_4(L)_2$ complex in the gas phase with PEE ligand

3.2.1 Inner-sphere complexation with monoamides

The optimized structure of $[Pu(NO_3)_4\mathbf{L}_2]$ is represented in Figure 3.2 for PEE ligand and structural parameters are listed in Table 3.2 for all the ligands. Additional structural parameters are given in annex A section. In the most stable geometry of these systems, the coordination number of plutonium is 10, with the nitrate groups acting as bidentate ligands. Amide ligands act as monodentate ligands, which interact with plutonium through the carbonyl oxygen. Changing the ligand alkyl chains does not significantly alter the structure. In the gas phase, the $Pu - O_L$ distance varies from 2.34 to 2.35 Å while the mean value of $Pu - O_{nit}$ varies from 2.47 to 2.49 Å. Ligand interatomic distances, such as C-O or C-N distances, also remain similar for all the ligands. In the polar DMA solvent $Pu - O_L$ distances are shortened by 0.03 - 0.04 Å compared to the gas phase while $Pu - O_{nit}$ distances are lengthened by 0.02 - 0.03 Å. These calculated distances are all longer than the solid phase distances reported from XRD of 2.26 Å and 2.44 Å respectively for $Pu - O_L$ and $Pu - O_{NO_3}$ distances in the Pu(IV)-tetranitrate complex with N,N-dibutyl-butanamide (DBBA) [29].

The variation of the complexation energies between the ligands and the tetra-nitrate plutonium complex are reported in Table 3.3. As detailed in computational methods section, energies are given as ligand exchange complexation energies. Energy differences were also calculated by adding empirical dispersion with Grimme's D3 method to account for dispersion interactions. On the carbonyl side, substituting the methyl by an ethyl group in R_1 position (from MMM to EMM or from MEE to EEE) has a small destabilizing

TABLE 3.2: Selected interatomic distances (average values in Å) calculated in $[Pu(NO_3)_4\mathbf{L}_2]$ complexes in the gas phase and with a solvent model[a]

Ligand	d_{Pu-O_L}	$d_{Pu-O_{nit}}$	$d_{Pu-N_{nit}}$	d_{O-C}	d_{C-N}
gas					
MMM	2.352	2.474	2.932	1.250	1.327
EMM	2.347	2.477	2.931	1.250	1.328
MME	2.356	2.483	2.930	1.249	1.327
MEE	2.354	2.480	2.931	1.250	1.327
EEE	2.348	2.485	2.933	1.251	1.326
PEE	2.347	2.483	2.934	1.252	1.327
IEE	2.343	2.486	2.937	1.252	1.327
solution *n*-dodecane					
MMM	2.337	2.489	2.934	1.252	1.325
EEE	2.332	2.495	2.940	1.256	1.323
PEE	2.332	2.495	2.940	1.257	1.323
IEE	2.327	2.498	2.943	1.256	1.325
solution DMA					
MMM	2.313	2.508	2.941	1.259	1.319
EEE	2.315	2.506	2.942	1.261	1.320
PEE	2.315	2.506	2.942	1.261	1.320
IEE	2.300	2.513	2.962	1.263	1.320

effect (+7 kJ $\cdot$ mol^{-1}) which nearly disappears with dispersion corrections at the D3 level. Lengthening the alkyl chain further from ethyl to *n*-propyl stabilizes the complex at both levels of calculation (EEE $\rightarrow$ PEE). In the reverse, increasing the branching from *n*-propyl to *iso*-propyl destabilizes strongly the complexation and has the most significant effect. On the nitrogen side, in R_2 and R_3 positions, changing only one methyl into one ethyl has a small destabilizing effect while further changing the second methyl group into ethyl has the reverse stabilizing effect. The inclusion of D3 dispersion correction systematically increases the stability of the complexation with longer alkyl chains by a few kJ $\cdot$ mol^{-1} as given by the comparison between PBE0 and PBE0-D3 results. This is indicative of the dispersion interactions induced by alkyl groups.

If we focus on PEE and IEE, which can be considered as model ligands for DEHBA (N,N-di(2-ethylhexyl)butyramide) and DEHiBA (N,N-di(2-ethylhexyl)isobutyramide) that have been largely studied for U(VI) and Pu(IV) extraction [20, 84], we can observe a +18 kJ $\cdot$ mol^{-1} free energy increase induced by the ramification at the PBE0 level and +14 kJ $\cdot$ mol^{-1} at the PBE0-D3 level. For these two isomeric ligands, calculations at the MP2 level and with *n*-dodecane and DMA solvent models were also performed. As they are of comparable shape and size, the errors associated with the solvent model are expected to largely cancel.

TABLE 3.3: Free energy variations corresponding to inner-sphere ligand exchange reaction (3.3), values in kJ · mol^{-1} calculated in the gas phase at the PBE0 and PBE0-D3 level

	L1 → L2	PBE0	PBE0-D3
R_1	MMM → EMM	7	2
	MEE → EEE	7	-2
	EEE → PEE	-5	-9
	PEE → IEE	22	14
R_2, R_3	MMM → MME	-3	-4
	MME → MEE	-4	-5

The relative complexation free energies between PEE and IEE are given in table 3.4. The 18 kJ · mol^{-1} free energy difference calculated for the gas phase is conserved in *n*-dodecane and slightly decreased, by only 3 kJ · mol^{-1}, in the more polar DMA. MP2 calculations are consistent with PBE0-D3 results and give a slightly lower energy difference than PBE0. According to the calculations, the free energy variation systematically indicates a less favorable complexation for the branched alkyl group than for the linear alkyl one in the case of [Pu(NO$_3$)$_4$**L**$_2$] formation. These results are consistent with the sharp diminution of the distribution ratio moving from DEHBA to DEHiBA observed experimentally for Pu(IV) extraction at low nitric acidity where the inner-sphere complex is the predominant species. They confirm the destabilizing steric hindrance effect, induced by the presence of voluminous branched alkyl group at the amide R_1 moiety, on plutonium tetranitrate complexation.

TABLE 3.4: Variation of complexation free energies for the formation of [Pu(NO$_3$)$_4$**L**$_2$] when going from PEE to IEE ligand, values in $kJ.mol^{-1}$ calculated in the gas phase and in solution

	$\Delta\Delta G(PEE \rightarrow IEE)$
Gas PBE0	22
Gas PBE0-D3	14
Gas MP2	14
Dodecane PBE0	18
DMA PBE0	15

3.2.2 Outer-sphere complexation with monoamides

In the outer-sphere Pu(NO$_3$)$_6$(H**L**)$_2$ structures, the interaction of the amide ligands with plutonium nitrates complexes occurs via the oxygen of the carbonyl group by hydrogen bonding. The coordination number of

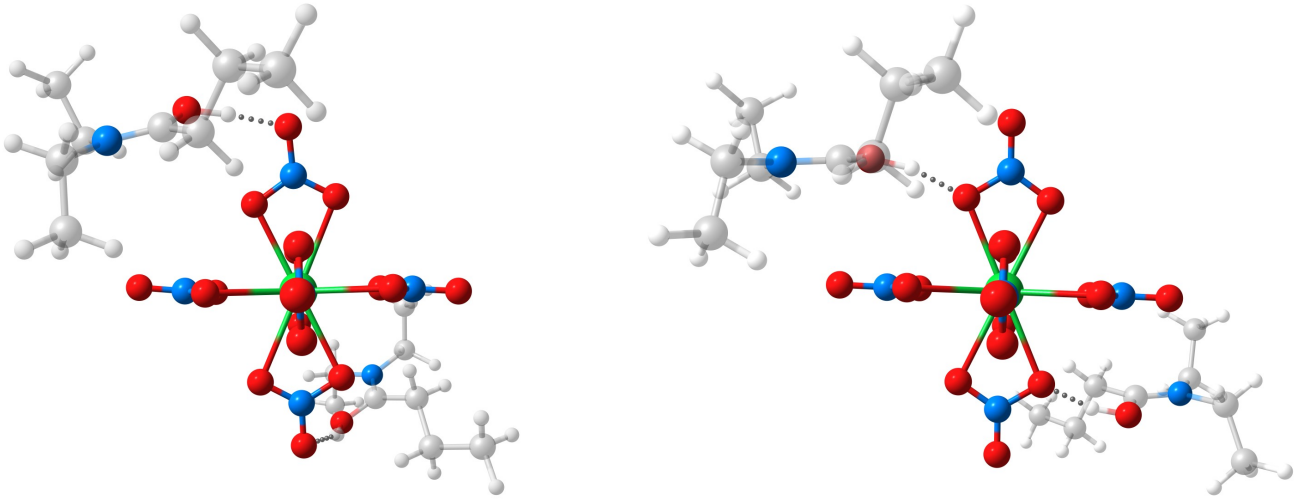


FIGURE 3.3: Optimized geometries of outer-sphere $[Pu(NO_3)_6](H\mathbf{PEE})_2$ in geometries I (left) and II (right) in the gas phase.

plutonium is twelve with six bidentate nitrate ions in the first coordination sphere and the two protonated amides form hydrogen bonds with nitrate ions in the outer coordination sphere. Two possible hydrogen bonds interactions were considered resulting in two different geometries described in Figure 3.3 for PEE ligand. In the first geometry (I), hydrogen bonds are formed with oxygen nitrate of the outer sphere. This structure was described in our previous work [29]. In geometry II, hydrogen atoms are bound to nitrate oxygen atoms that are also coordinated to plutonium. This coordination mode can be deduced from the electrostatic potential which was determined for $Pu(NO_3)_6^{2-}$ anion [105]. It was shown that first-shell oxygen atoms may be considered as possible acceptor sites for non-covalent interactions even though such interactions were not observed in the crystal structure. In addition, such coordination mode has been shown to exist for uranyl sulfate complexes with tertiary amines. Hydrogen bonds between the protonated amines and coordinated oxygen sulfate atoms were found in the solid state in crystal structure [106] and in *n*-dodecane from molecular dynamics simulations [107]. According to DFT calculations, Geometry II is significantly more stable than geometry I in the gas phase and in solution (Table 3.5) for all the ligands except MMM. For this ligand, geometry I and II are found at similar energy in the gas phase ($\Delta G = 1$ kJ $\cdot$ mol^{-1}) but geometry II is stabilized in solution (by 14 kJ $\cdot$ mol^{-1} in *n*-dodecane).

In both geometries, the proton is located on the carbonyl group rather than on a nitrate ion. In the gas phase, the protonated ligands wrap the hexanitrato plutonium complex. Calculated interatomic distances are collected in Table 3.6 for the most stable geometry II. Interatomic distances for geometry I are given in annex

TABLE 3.5: Free energy differences between the two outer-sphere geometries I and II $[Pu(NO_3)_6](HL_2)$, values in kJ · mol^{-1} calculated in the gas phase, in *n*-dodecane (Dod.) and in DMA

Ligand	MMM	EMM	MME	MEE	EEE	PEE	IEE
Gas PBE0	1	-17	-5	-11	-20	-22	-17
Gas PBE0-D3	1		-4		-20	-23	-18
Dod. PBE0	-14				-24	-23	-26
DMA PBE0	-22				-13	-23	-26

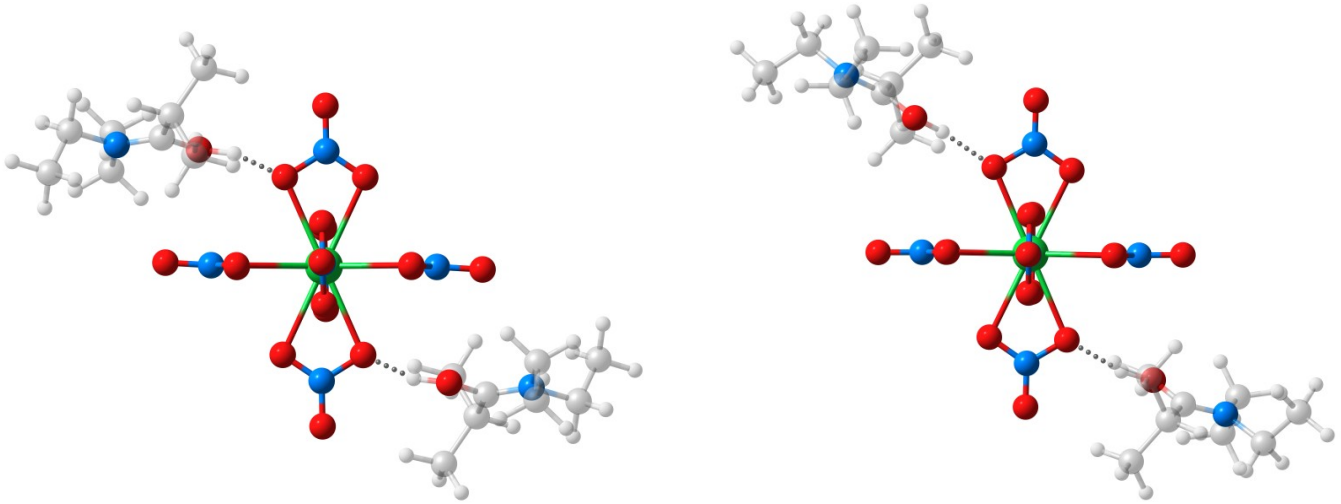


FIGURE 3.4: Optimized geometries of outer-sphere $[Pu(NO_3)_6](HIEE)_2$ in geometries II in the gas phase (left) and in DMA (right).

A section. The inner $Pu(NO_3)_6^{2-}$ coordination shell is very slightly altered by the nature of the outer-sphere ligand. The average Pu-O distance with nitrate ions which are not involved in hydrogen bonding ($Pu - O_{nit1}$) is 2.52 Å for all the complexes. This distance is close to the average value reported in $Pu(NO_3)_6^{2-}$ crystal structures (2.49 Å) [108]. The average distance between Pu and the oxygen nitrates involved in the hydrogen bonds ($Pu - O_{nit2}$) is 0.04-0.05 Å longer than the distance with other nitrates ($Pu - O_{nit1}$). The longest $Pu - O_{nit2}$ distance is found with IEE ligand. The lengths of the hydrogen bonds between the nitrate anions and the protonated ligand increase with the lengthening of the alkyl group in any R_1, R_2 or R_3 positions, as indicated from $O_{nit} - O_L$ and $O_L - H$ distances. On the reverse, the hydrogen bond distance decreases from *n*-propyl to *iso*-propyl. This could indicate a stronger H-bond but the OHO angle also diminishes from 164 to 160°. This may indicate that the hydrogen bond is more constrained with IEE ligands. Another geometric difference arises from the position of the two amide groups in the complexes. In the optimized structures in the gas phase, each planar amide group is nearly parallel to one of the six trigonal planar nitrate anions

(Figure 3.3) except with the bulky isopropyl group. As depicted on Figure 3.4, with IEE the planar orientation of the amide group deviates from the parallel orientation with the planar nitrate ion. As a consequence the distance between Pu and the nitrogen atom of the amide function ($Pu - N_L$) increases by about 0.4 Å for IEE compared to other ligands in the gas phase.

As shown in Table S3 in annex A section, a series of short interatomic close contacts are found between C-H oxygen and NO_3^- oxygen atoms (< 2.80 Å), involving C-H groups from alkyl chains of both nitrogen and carbonyl sides.

The inclusion of a solvent model does not alter plutonium-oxygen distances in the inner sphere but modifies the outer-sphere structure: the two outer-shell ligands remain hydrogen bonded by the protonated amide group but draw apart $Pu(NO_3)_6^{2-}$. This effect increases with the solvent dielectric constant. As a consequence, the $Pu - N_L$ distances increase by $\simeq$ 0.2 Å in *n*-dodecane and by $\simeq$ 0.9 Å in DMA compared to the gas phase except for IEE. The solvent effect is weaker for IEE ligands because the deviation of the amide group from the parallel planar orientation is already strong in the gas phase. The hydrogen bond distance increases slightly in dodecane (by 0.04 to 0.07 Å) and strongly in DMA (by 0.13 to 0.22 Å). The largest increase corresponds to IEE ligands. It should be noted that only the short close contacts between C-H from alkyl groups attached to the carbonyl side remain in polar solution.

The complexation energies were calculated for ligand exchange reaction (3.4) corresponding to the formation of $[Pu(NO_3)_6](H\mathbf{L2})_2$ relative to $[Pu(NO_3)_6](H\mathbf{L1})_2$. Calculated values are reported in Table 3.7 for the most stable geometry II. On the carbonyl side, in R_1 position, substituting the methyl by an ethyl group or an ethyl by a *n*-propyl has very similar effects than those determined for inner-sphere complexation. Changing from MMM to EMM and MEE to EEE has a small destabilizing effect that is reversed when D3 dispersion corrections are included. From EEE to PEE, a stabilizing effect is calculated at both levels of calculations. Increasing the branching from *n*-propyl to *iso*-propyl has a strong destabilizing effect ($\Delta\Delta G_{Gas}$ = 39 kJ $\cdot$ mol^{-1} at the PBE0-D3 level). Such a destabilizing effect was also found for inner complexation but is unexpectedly much larger for this coordination mode while steric hindrance is expected to be smaller in the outer coordination sphere.

On the nitrogen side, the effects are different than those found for inner complexation. Changing one methyl

TABLE 3.6: Selected interatomic distances (average values d in Å)[a] and Hydrogen bond angle (α_{O-H-O}) calculated in [Pu(NO_3)$_6$](H***L***$_2$) geometry II in the gas phase and in solution.[a] O_{nit1} denotes oxygen from nitrate ions which are not involved in hydrogen bonds. O_{nit2} denotes oxygen from nitrate ions which are involved in hydrogen bonds.

	$d_{Pu-O_{nit1}}$	$d_{Pu-O_{nit2}}$	d_{Pu-N_L}	$d_{O_{nit2}-O_L}$	$d_{O_{nit}-H}$	α_{OHO}
gas						
MMM	2.524	2.565	5.973	2.504	1.506	163°
EMM	2.523	2.566	6.031	2.511	1.513	164°
MME	2.524	2.564	6.031	2.511	1.515	164°
MEE	2.524	2.564	6.006	2.517	1.521	164°
EEE	2.523	2.566	6.047	2.522	1.528	164°
PEE	2.523	2.565	6.085	2.524	1.530	164°
IEE	2.520	2.573	6.363	2.498	1.514	160°
solution *n*-dodecane						
MMM	2.524	2.562	6.178	2.546	1.555	165°
EEE	2.524	2.562	6.210	2.558	1.570	165°
PEE	2.524	2.562	6.232	2.559	1.571	165°
IEE	2.523	2.567	6.551	2.546	1.585	158°
solution DMA						
MMM	2.524	2.560	6.856	2.644	1.651	174°
EEE	2.524	2.560	6.811	2.650	1.660	173°
PEE	2.524	2.561	6.840	2.654	1.664	172°
IEE	2.525	2.559	6.993	2.666	1.732	156°

into in an ethyl group has a strong stabilizing effect up to 19 kJ · mol^{-1} from MMM to MME at the PBE0-D3 level. Such effect was not found for inner-sphere complexation. The further addition of one ethyl group has no significant effect. As for inner-sphere complexes, the complexation energy in the solvent phase was calculated for isomeric PEE and IEE derivatives. Relative energies in the gas phase and in solution are given in Table 3.8. The free energy difference between the two ligands is similar in the gas phase and in *n*-dodecane but decreases from 34 kJ · mol^{-1} in the gas phase to 14 kJ · mol^{-1} in the more polar DMA solution. These energy variations follow the structural changes from the gas phase to polar solution with the diminution of close-contact interactions between C-H groups and nitrate ions. The energy difference between PEE and IEE is significant and is consistent with the sharp diminution that has been measured for Pu(IV) distribution ratio from DEHBA to DEHiBA in high nitric acid concentration. The solvent effect is more significant for outer than inner-sphere complexation. This is consistent with solvent effects which are stronger on the outer-sphere structure. In the gas phase, the preferred conformation of the outer-sphere complex with nearly parallel nitrate and amide groups is disfavored with the bulky IEE. In polar solution, the ligands draw apart

TABLE 3.7: Free energies variations corresponding to outer-sphere Ligand exchange reaction (3.4), values in kJ · mol^{-1} calculated in the gas phase

	L1 → L2	PBE0	PBE0-D3
R_1	MMM → EMM	3	-1
	MEE → EEE	3	-3
	EEE → PEE	-6	-11
	PEE → IEE	34	39
R_2, R_3	MMM → MME	-6	-10
	MME → MEE	-3	0

and steric constraint becomes smaller.

TABLE 3.8: Variation of complexation free energies for the formation of [Pu(NO$_3$)$_6$(H**L**)$_2$] when going from PEE to IEE ligands, values in kJ · mol^{-1} calculated in solution

	$\Delta\Delta G(PEE \rightarrow IEE)$	
	PBE0	PBE0-D3
Gas	34	39
Dodecane	29	35
DMA	14	19

3.2.3 Complexation with carbamides

On one hand, a carbamide is considered as being more basic than an amide function and is expected to bind a metal cation more strongly. Calculated proton affinities confirm the stronger electron donor capability of the carbonyl oxygen in carbamide compared to amides (in annex A section). On the other hand, tetra-alkyl carbamide is bulkier than tri-alkyl monoamides for similarly sized-alkyl groups and are expected to induce larger steric hindrance effect in plutonium coordination sphere. In the present work, inner and outer-sphere complexations were investigated for two carbamide derivatives with methyl and ethyl groups (C4M and C4E).

In the inner-sphere complex, plutonium-oxygen bond distances are close to those calculated for monoamides within 0.02 Å (Table 3.9). The $Pu - O_L$ distance is slightly shorter with C4M than with any other monoamides while $Pu - O_L$ distance with C4E is the longest one. This is an indication of the competition between electronic donor and steric hindrance effects which are expected to be stronger with carbamide than with monoamides.

For outer-sphere complexation, the energy differences between geometry I and II are not as large than for monoamides (see Table 3.11). At the PBE0 level, the preferred coordination mode is geometry I in the gas phase and II in *n*-dodecane.The inclusion of dispersion corrections has more significant effects for the tetra-ethyl carbamide than for monoamides derivatives. Dispersion attractive interactions stabilize geometry I over geometry II in the gas phase and in *n*-dodecane. This is attributed to the weak interactions that take place between the alkyl chains and nitrate ions. Such interactions are becoming more important as the number of alkyl groups increases and also more important in the gas phase and in *n*-dodecane than in the more polar DMA solution which counterbalances their strength. Finally, considering the small energy difference in solution it is not possible discriminate between both structures which may coexist in solution. The structural parameters are reported in Table 3.10 for geometry II and in SI section for geometry I. The hydrogen bond distances are slightly longer with the carbamide than with the monoamides (in the gas phase, the lengthening of $O_{nit} - H$ distance is 0.03 Å from C4E to EEE and 0.05 Å for the $O_{nit} - O_L$ distance). In geometry I, this is the reverse, the hydrogen bond distance is shorter with carbamide, O-H distance decreases by 0.08 Å from C4E to EEE.

TABLE 3.9: Selected interatomic distances (average values in Å) calculated in $[Pu(NO_3)_4\mathbf{L}_2]$ complexes with carbamides in the gas phase

Ligand	d_{Pu-O_L}	$d_{Pu-O_{nit}}$	$d_{Pu-N_{nit}}$	d_{O-C}	d_{C-N}
C4M	2.346	2.490	2.936	1.256	1.354
C4E	2.358	2.493	2.939	1.252	1.349
solution *n*-dodecane					
C4M	2.326	2.497	2.941	1.260	1.346
C4E	2.345	2.499	2.943	1.267	1.352
solution DMA					
C4M	2.298	2.510	2.952	1.268	1.342
C4E	2.323	2.508	2.951	1.273	1.348

Complexation energies are given relative in Table 3.12. For outer-sphere complexes, the lowest energy geometry was taken at each level of calculation. Energies were calculated in the gas phase and in solution. Results are compared with the tri-ethyl EEE monoamide. For inner-sphere complexation, results are strongly dependent on dispersion corrections. At the PBE0 level, when the length of the carbamide alkyl groups increases (from C4M to C4E) or from monoamide to carbamide (PEE to C4E), inner-sphere complexation is destabilized by 9 to 17 kJ $\cdot$ mol^{-1}, in the gas phase as well as in solution. In the presence of dispersion

TABLE 3.10: Selected interatomic distances (average values d in Å)[a] and Hydrogen bond angle (α_{O-H-O}) calculated in [Pu(NO$_3$)$_6$](H***L***$_2$) geometry II in the gas phase and in solution.[a] O_{nit1} denotes oxygen from nitrate ions which are not involved in hydrogen bonds. O_{nit2} denotes oxygen from nitrate ions which are involved in hydrogen bonds.

	$d_{Pu-O_{nit1}}$	$d_{Pu-O_{nit2}}$	d_{Pu-N_L}	d_{Pu-N_L}	$d_{O_{nit2}-O_L}$	$d_{O_{nit}-H}$	α_{OHO}
gas							
C4M	2.523	2.563	5.178	6.190	2.557	1.552	169°
CAE	2.516	2.580	5.608	7.016	2.573	1.561	174°
solution *n*-dodecane							
C4M	2.524	2.560	5.234	6.296	2.600	1.606	169°
C4E	2.529	2.571	5.695	7.071	2.613	1.613	174°
solution DMA							
C4M	2.525	2.561	5.513	6.667	2.693	1.712	169°
C4E	2.525	2.559	5.863	7.178	2.692	1.711	170°

TABLE 3.11: Free energy differences between the two outer-sphere geometries I and II [Pu(NO$_3$)$_6$](H***L***$_2$) , values in kJ $\cdot$ mol^{-1} calculated in the gas phase, in *n*-dodecane and in DMA

	C4M	C4E
Gas PBE0	4	7
Dodecane PBE0	-4	-2
DMA PBE0	1	-3
Gas PBE0-D3	6	17
Dodecane PBE0-D3	-3	7
DMA PBE0-D3	-1	-3

correction, this is the reverse, inner sphere is stabilized, attractive dispersion forces overbalance steric hindrance effects. Dispersion interactions are likely to be overestimated from such calculations since the competitive dispersion interactions with alkyl chains from the solvent are not properly taken into account. However, these results emphasize the key role of weak interactions in these systems. For outer-sphere complexation, the lengthening of the alkyl group and destabilizes the complexation in the gas phase. Solvent effects counterbalance steric hindrance effects in the outer sphere and favor C4E complexation compared to C4M. There is no strong influence of dispersion corrections. The replacement of PEE by C4E slightly favor outer sphere complexation in the gas phase and in *n*-dodecane. Including dispersion corrections or increasing the solvent polarity favor outer-sphere complexation.

TABLE 3.12: Complexation free energies when going from C4M to C4E and from PEE to C4E ligands for reaction (3.3) and (3.4), values in kJ · mol^{-1} calculated in the gas phase and in solution at the PBE0 ($\Delta\Delta G_{Gas}$) and PBE0-D3 level $\Delta\Delta G^{D3}_{gas}$.

	inner-sphere		outer sphere	
	C4M → C4E	PEE → C4E	C4M → C4E	PEE → C4E
Gas PBE0	9	17	11	5
Dod. PBE0	11	10	4	6
DMA PBE0	13		-2	-8
Gas PBE0-D3	-13	-3	17	-1
Dod. PBE0-D3	-11	-9	6	-3
DMA PBE0-D3	-9		-6	-10

3.3. Conclusion

From this work it is possible to get a better understanding of the strong influence of amide structure on plutonium extraction. Experimentally, a simple ramification of the alkyl chain attached to the carbonyl group can suppress plutonium extraction at low and high acidity even though amides operate through different coordination modes. Complexation energies were calculated for inner and outer-sphere complexation, which take place in weakly acidic and strong acidic conditions, respectively. For both inner and outer-sphere complexes, it is found that the introduction of a bulky alkyl group on the carbonyl side diminishes strongly the complexation energy. This is fully consistent with monamide extraction properties. The influence of the bulkiness of the alkyl group is as or even more important for outer than for inner-sphere interactions. This was unexpected when considering that there is less flexibility and stronger steric constraints in the inner sphere. This is attributed to specific electrostatic interactions between the two outer-sphere amide ligands and two nitrate ions of $Pu(NO_3)_6^{2-}$. By increasing the polarity of the solution, such interactions diminish and the outer-sphere ligands expand out of $Pu(NO_3)_6^{2-}$. As a consequence, solvent effects are very significant for outer-sphere complexation while there are small for inner-sphere complexation. It gives the key possibility to tune the substituent effect by changing the polarity of the solution. When changing the ligand, from monoamide to carbamide, the results show that weak interactions such has dispersion forces may have a remarkable impact on ligand binding affinities and on plutonium(IV) extraction. In order to predict ligand affinity of different families, it will be essential to go beyond the DFT approximation and to perform explicit solvent simulations.

Chapter 4

Molecular dynamics simulations

> *Everything in life can be understood in terms of jigglings and wigglings of atoms.*
>
> Richard Feynman,

The aim of this chapter is to introduce the basic concepts of molecular dynamics (MD) simulations and how one can use these simulations to understand and predict the physical properties of molecular systems in the condensed phase.

4.1. Molecular Dynamics simulations

Molecular Dynamics (MD) is one of the most used computational methods for evaluating the equilibrium and transport properties of assemblies of molecules, along with their thermodynamics, structural and dynamic properties, by solving the classical equations of motion of a molecular system (liquid, gas or solid) with the aim of obtaining the time evolution of the system. Each particle (atom of the molecule) is considered as a

point mass whose movement is determined by the set of forces exerted on it by the other atoms as a function of time (time scale: 1 fs = 1×10^{-15} s). The molecule is therefore described as a dynamic entity whose atoms change their spatial positions over time. For instance, the atomic movements of the molecular system correspond to the vibrations around the minimum or the passage from a minimum to another minimum of energy. In the Newton's notation, each particle obeys the following differential equation:

$$\overrightarrow{F_i} = m_i \overrightarrow{\gamma_i}, \tag{4.1}$$

where the $\overrightarrow{F_i}$ corresponds to the sum of different forces acting on the particle i, assimilated to a point mass, and $\overrightarrow{\gamma_i}$ its acceleration in a Galilean reference.

The forces $\overrightarrow{F_i}$ exerted on an atom i with coordinates $\overrightarrow{r_i}$ at time t are determined by the derivative of the potential energy function $U(r_i)$.

$$\overrightarrow{F_i} = -\overrightarrow{\nabla} U(\overrightarrow{r_i}) = -\frac{\partial U(\overrightarrow{r_i})}{\partial r_i}, \tag{4.2}$$

Hence, the velocity $\overrightarrow{v}(\overrightarrow{r_i})$ and positions $\overrightarrow{r_i}$ of the particle i are linked to the potential $U(\vec{r_i})$ with :

$$\frac{d\overrightarrow{v}(\overrightarrow{r_i})}{dt} = \frac{1}{m_i}\overrightarrow{F_i}, \tag{4.3a}$$

$$\frac{d^2\overrightarrow{r_i}}{dt^2} = \frac{d\overrightarrow{v}(\overrightarrow{r_i})}{dt}. \tag{4.3b}$$

The goal is to determine the coordinates $\overrightarrow{r_i}(t + \delta t)$ of the atom at the instant $t + \delta t$ knowing its acceleration at the previous instant $\overrightarrow{\gamma_i}(t)$ that we calculated knowing the interaction forces $\overrightarrow{F_i}(t)$.

Since we are dealing with several thousands of atoms, these coupled equations are solved numerically. Many algorithms that predict positions (and velocities) of atoms at a later point in time have been developed providing an approximate solution to equation 4.2, such as the Velocity-Verlet and the Leap-Frog algorithms [41]. For example, the Velocity-Verlet algorithm integrates the equations of motion (4.2, 4.3) and the positions;

the velocities of the particles are then given by:

$$r_i(t+\delta t) = r_i(t) + v_i(t)\delta t + \frac{F_i(t)}{2m_i}(\delta t)^2, \tag{4.4a}$$

$$v_i(t+\delta t) = v_i(t) + \frac{F_i(t) + F_i(t+\delta t)}{2m_i}\delta t. \tag{4.4b}$$

Depending on the manner the forces acting on atoms are computed, one can distinguish between classical and quantum " *ab initio*" molecular dynamics.

In *ab initio* MD, which takes into account the electronic structure of the particles and the quantum nature of the interactions involved, the forces are obtained via Hellmann–Feynman theorem [109]. According to this theorem, once the spatial distribution of the electrons has been determined by solving the Schrödinger equation, all the forces in the system can be calculated using :

$$\vec{F_i} = -\frac{\partial E}{\partial \vec{r_i}} = -\left\langle \Psi \middle| \frac{\partial H}{\partial \vec{r_i}} \middle| \Psi \right\rangle, \tag{4.5}$$

Ψ being the electronic wave function of the system. However, the need to calculate the forces at each time step of the simulation makes *ab initio* MD extremely time-consuming and its application limited to a small number of atoms and short trajectories, since one needs to solve the Schrödinger equation, for an electronic system at each time iteration. Moreover, most *ab initio* MD simulations are carried out using DFT with LDA or GGA functionals to keep the computational cost affordable. However, such functionals are known to be inaccurate for describing weak intermolecular interactions, especially for atoms containing a large number of electrons; further details can be found in Ref. [110] and discussed in Section 2.3.

In classical MD, the forces are computed as derivatives of parameterized interaction potentials, also known as force fields (FF) (see Section 4.2). This MD approach is quite fast and applicable to much larger systems, since the forces are calculated at each iteration of the simulation as analytical or numerical derivatives of analytical FFs. Despite the many advantages of classical MD, it has its drawbacks that deserve to be mentioned. First, most traditional FF employ preset bonding arrangements (unlike the *ab initio* dynamics) and are thus unable to model the process of any chemical bond breaking and reactions explicitly (excluding the Reactive Force Field methods). Second, electrostatic forces are described using partial charges, and

they rely on point charges to reproduce the electrostatic potential around molecules has proven to be a good approximation and made huge success in several fields, yet most molecular systems have anisotropic charge distribution. To capture these features of the electronic structure (such as electronic polarization) additional terms are meant to be included (e.g., polarizable force fields). By the same token, Van der Waals interactions are also described using fixed parameters, while they are strongly environment-dependent in principle. Since these forces originate from interactions of induced and "instantaneous" dipoles. Lastly, in every force field, the analytical formula that describes a specific type of interaction contains a certain set of parameters that might be specific to a given molecular system, in literature one may find many special FFs derived for individual molecules. Hence, before using any FF parameters set, one should validate them by careful and suitable comparisons with some other experimental results (not used in the parameterization process). However, those FF parameters are, generally, derived from either experimental data or *ab initio* "quantum mechanical" calculations and sometimes both.

4.2. Molecular mechanics: the Force-Field Models

In molecular mechanics, the atoms are considered as point masses and bonds as springs, and the interactions are modelled by classical potential energy functions. Describing a molecular system using classical mechanics may seem surprising and perhaps unexpected, after all, the atomic scale is where quantum mechanical effects are supposedly important. Still, this is not an issue in the case of molecular systems. Indeed, by calculating the thermal de Broglie wavelength of the constituent atoms at the temperature of interest T, $\lambda_{th} = \sqrt{\frac{\beta\hbar^2}{2\pi m}}$ with m the atomic mass and $\beta = 1/k_B T$, k_B and $\hbar$ are the Boltzmann and the Planck constants, respectively. The thermal wavelength is usually very small for the atoms, meaning that no significant quantum effects are expected and classical mechanics can be considered [111].

In molecular mechanics, the mathematical form of the potential energy includes on the one side bonded terms (also called intramolecular) for the interactions of atoms that are linked by covalent bonds, and on the other side non-bonded (also called intermolecular) terms that describe the long-range electrostatic and other forces, such as Van der Waals and polarization interactions. The general form of the total potential energy is

thus written as:

$$U = U_{bonded} + U_{nonbonded}, \tag{4.6}$$

with:

$$U_{bonded} = U_{bond} + U_{angle} + U_{dihedral}, \tag{4.7a}$$

$$U_{nonbonded} = U_{qq\cdot} + U_{rep} + U_{disp} + U_{specific-terms}. \tag{4.7b}$$

In Eq. 4.7a, the $U_{bond}, U_{angle}, U_{dihedral}$ represent bond stretching, angles bending and dihedral torsion energies, respectively. In Eq. 4.7b, $U_{qq\cdot}$ accounts for the Coulomb electrostatic interactions, U_{rep} and U_{disp} represent the repulsion and dispersion interactions, respectively. The specific decomposition of the terms included in the analytical forms depends on the areas of application and the desired level of accuracy. Certain specific terms can be added for more precision, such as cross many-body terms, polarization, hydrogen bonds, and charge transfer, which will be discussed later in Section 4.2.2.4.

4.2.1 Bonded interactions

Bonded (intramolecular) interactions are not exclusively pair interactions but also include 3- and 4-body interactions as well. There are bond stretching (2-body), bond angle (3-body), and dihedral angle (4-body) interactions. In general, the harmonic approximation is used to describe the internal motions of molecules, stretching, and bending angles. In the field of biochemistry or the studies of vibrational spectroscopy, where the intramolecular forces play a crucial role [112], specific terms are meant to be added for more sophisticated FF, namely, "cross-terms" that describe the coupling of different internal degrees of freedom, such as angles and bond lengths and improper torsions.

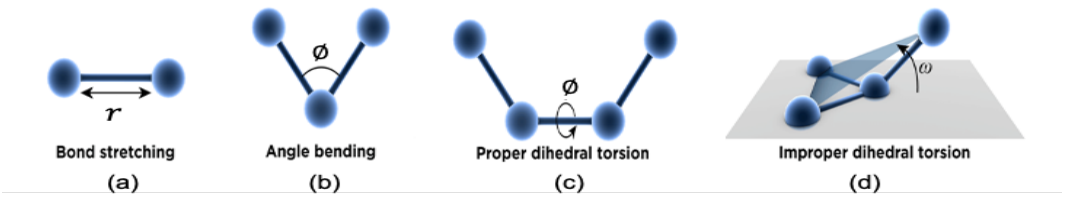


FIGURE 4.1: Schematic illustrations of the main bonded terms.

4.2.1.1 Bond Stretching

Bonds within molecules are generally described by simple functions that depend only on the interatomic distance. The most used simple form is a harmonic function:

$$U_{bond} = \frac{1}{2} K_b (r - r_0)^2, \tag{4.8}$$

that solely depends on two parameters, K_b the force constant, and r_0 the equilibrium bond length.

To account for the potential anharmonic nature of the chemical bond, a better model would include higher-order expansions of the interatomic potential or a Morse potential. Yet, under most circumstances the harmonic approximation is reasonably good, most of the existing potentials use the simpler harmonic function.

4.2.1.2 Angle bending

Bending energy potentials are usually treated very similarly to stretching potentials; the energy is assumed to increase quadratically with the displacement of the bond angle θ from the equilibrium value θ_0:

$$U_{angle} = \frac{1}{2} K_a (\theta - \theta_0)^2. \tag{4.9}$$

This expression has only two parameters, K_a and θ_0. Same comment as for the bond stretching can be said about including the anharmonicity to the potential.

4.2.1.3 Proper dihedral torsion

In any molecule containing more than four atoms in a row, we need to include also a dihedral or torsional term. Torsion angle twisting describes the twisting between the planes formed by the first three and last three atoms of a consecutively bonded (i, j, k, l)-quadruple of atoms. Most existing force fields use a 3-term Fourier series to take into account the $1-4$ interactions:

$$U_{dihedral} = \frac{1}{2} \sum_n K_n (1 + cos(n\phi + \phi_0)), \tag{4.10}$$

where ϕ is the dihedral angle between the (i, j, k, l) atoms, and ϕ_0 is the phase shift angle and K_n is the multiplicative constant with n being a non-negative integer constant that indicates the periodicity (often called the multiplicity; it usually takes values between 1 and 3. We note that here, $2n$ parameters define the torsion potential (K_n, ϕ_0).

4.2.1.4 Improper dihedral torsion

For planar groups, an additional improper torsion term may be needed to ensure planarity, such as for sp^2 hybridized carbons in carbonyl groups or in aromatic rings. It describes the positive contribution to the energy of these out-of-plane motions:

$$U_{imp} = \frac{1}{2} K_{imp}(\omega - \omega_0)^2, \tag{4.11}$$

where ω is the improper angle corresponding to the deviation from planarity and K_{imp} is the improper force constant and ω_0 is the equilibrium improper dihedral angle.

4.2.1.5 Cross terms

It is possible to add cross terms to account for the coupling between the stretching, bending, and torsion motions. They bring corrections to the intramolecular energy and allow to better reproduce the intramolecular vibrational spectra. Examples of such cross-terms are

$$U_{bond-bond} = \frac{1}{2} K_{bb}(r - r_0)(r' - r'_0), \tag{4.12a}$$

$$U_{bond-bend} = \frac{1}{2} K_{ba}[(r - r_0)(r' - r'_0)](\theta - \theta_0). \tag{4.12b}$$

These cross-terms have been discarded in our work as we do not aim at modeling vibrational spectroscopy.

4.2.2 Non-bonded interactions

The non-bonded potential terms involve interactions between all (i, j)-pairs of atoms, excluding, all pairs of atoms already involved in bonded terms. It is known that the cost of computing the non-bonded interactions

(e.g., electrostatic, Van der Waals forces, polarization) dominates the computational effort required at each time step of an MD simulation since we are computing interaction between all (i, j)-pairs of atoms (excluding, a certain number of atoms after a cutoff distance).

Polarizable FF usually account for three types of non-bonded interactions (See Eq. 4.6): electrostatic, Van der Waals and polarization interactions. It has to be mentioned that the more terms we add to describe the potential energy the more the computational cost.

4.2.2.1 Electrostatic interactions

Electrostatic interactions are generally a major contribution to the non-bonded potential energy in any system. They arise from the unequal distribution of charge in a molecule or the net charge of ions. This uneven distribution of charge is usually modeled by placing point charges at each of the atomic sites. The most common way to obtain reliable partial charges consists in performing an *ab initio* electronic structure calculation and then deriving them from the molecular electrostatic potential. Many different methods have been developed to determine them and they differ in the predicted value of the distribution of partial charges. Unfortunately, they cannot be derived unambiguously because atomic charges are not experimental observables. In a simpler way, we aim at reducing the electrostatic description with charges distributed continuously in space into localized point charges. This affects the computed potential energy of the simulated system, which determines the state probabilities. Ultimately, the equilibrium properties of the system depend strongly on the partial charge calculation scheme. Anyhow, the interaction between these point charges is generally modeled by a Coulomb potential:

$$U_{qq'} = \sum_{i \neq j} \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}, \tag{4.13}$$

where ϵ_0 is the permittivity of free space (vacuum), q_i and q_j are atomic charges of atoms i and j, and r_{ij} is the distance between i and j.

4.2.2.2 Van der Waals interactions

Non-bonded interactions which are not electrostatic are labeled Van der Waals interactions. In atomic systems, they arise from the balance between repulsive and attractive forces (also called dispersion forces). The repulsion is due to the overlap of the electron clouds of both atoms, while dispersion arises from the interactions between instantaneously induced dipoles. The Lennard-Jones (L-J) potential is the simplest mathematical model that accounts for such interactions. Namely the repulsive term varies as r^{-12}, while the attractive term varies as r^{-6}:

$$U_{vdw} = 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right), \tag{4.14}$$

Another form of Lennard-Jones potential, in which the repulsive interactions acting at short distances are represented by a r^{-9} term is also often used:

$$U_{vdw} = \epsilon_{ij}\left(2\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{9} - 3\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right). \tag{4.15}$$

For both functional forms, $r_{ij} = \|\overrightarrow{r_j} - \overrightarrow{r_i}\|$ gives the distance between the pair of atoms. ϵ_{ij} is the Van der Waals well depth for atoms i and j, and σ_{ij} is the distance at which $E_{vdw} = 0$. As for ϵ_{ij} and σ_{ij}, they are generally found from the ϵ_{ii} and σ_{ii} using the mixing rules, for example, the Berthelot rule [113, 114]:

$$\begin{aligned} \epsilon_{ij} &= \sqrt{\epsilon_{ii}\epsilon_{jj}}, \\ \sigma_{ij} &= \frac{\sigma_{ii} + \sigma_{jj}}{2}. \end{aligned} \tag{4.16}$$

where σ_{ii} and ϵ_{ii} are the only parameters adjusted when developing a new FF with a LJ potential, they describe the interaction between atoms of the same type, e.g., C - C or H- H. From the analytical form of the L-J potential, we can see that it approaches 0 rapidly as r_{ij} increases. Hence, in practice, there are many techniques to truncate the potential to save computing time and satisfy the minimum image convention when using periodic boundary conditions. It is usually truncated (smoothly shifted) to 0 after a cutoff radius or shifting up the whole potential to the value of the L-J at the cutoff to avoid the discontinuity.

A more rigorous choice for the repulsive part of the potential would be an exponential term, as the electron

density falls off roughly exponentially with distance from the nuclei [60, 115]. This leads to the Buckingham mathematical model:

$$U_{vdw}^{Buck} = A_{ij} \exp\{(-B_{ij} r_{ij})\} - \frac{C_{ij}}{r_{ij}^6}, \tag{4.17}$$

where A_{ij} and B_{ij} define the intensity and the range of the repulsion, respectively. E_{vdw}^{Buck} retains the long range r^{-6} dependence of the Lennard-Jones potential. As $r \to 0$ the Buckingham potential goes through a maximum then rapidly falls to $-\infty$. Although Buckingham potentials have been used extensively in molecular dynamics simulations, and considered more accurate than L-J potential but this comes with a price in terms of the number of parameters to be fitted. Owing to the use of the combining rules for the parameters for Lennard-Jones potentials, the latter are more practical and convenient than the Buckingham ones [116, 117].

4.2.2.3 Polarization

In the case of systems composed of multi-charged ions and/or polar molecules, the properties of these systems are strongly influenced by their environment. Describing interactions only via pair potentials then becomes insufficient. In fact, effective partial fixed charges are unable to describe the variation in electrostatics due to many-body polarization effects, which have been shown to be a significant component of intermolecular forces in such systems (multi-charged ions, polar molecules) [47, 117, 118]. For example, the partial charges used to describe a molecule in the liquid state will not be adequate to describe the same molecule in the gas phase, because the surrounding particles will induce a charge redistribution that needs to be accounted for. In a simpler view, the polarization can be defined as the redistribution of a molecule's electron density due to an electric field exerted by other molecules. Thus, explicit many-body polarization effects are required to accurately describe the system.

Many classical electrostatic models that take into account polarization have been developed, such as

- Fluctuating charges [119]: where charges are allowed to fluctuate according to the environment, so that the charge flows between atoms until their instantaneous electro-negativities are equal.
- Shell model (Drude particles) [120]: the electronic polarization is incorporated by representing the atom as the sum of a charged core and a charged shell connected by a harmonic spring whose force

constant is inversely proportional to the atomic polarizability. The relative displacement of both charges depends on the electrostatic field created by the environment.

- The induced point dipoles model: This model being used in our work will be discussed in more details. In this approach, each atom i has an isotropic polarizability α_i assigned to it. Thereby, each atom in the system will, at any given moment, have an induced dipole moment. The atomic-induced dipole moments $\overrightarrow{\mu_i}$ obey

$$\overrightarrow{\mu_i} = \alpha_i \left(\overrightarrow{E_i^q} + \sum_{j=1}^{N} T_{ij} \overrightarrow{\mu_j} \right) = \alpha_i \overrightarrow{E_i}, \tag{4.18}$$

where $\overrightarrow{E_i}$ is the total electric field acting on atom i. So, the induced dipoles depend both on the static charges, via the permanent electric field $\overrightarrow{E_i^q}$ due to partial charges, and the induced dipoles of the rest of the atoms in the system ($\sum_{j \neq i}^{N} T_{ij} \overrightarrow{\mu_j}$). Therefore, polarization is non-additive and often solved with a self-consistent field method, by solving the following equations iteratively until the values for all the induced dipoles converge. Thus, the values of the dipole moments are those that minimize the polarization energy (U_{pol}):

$$\overrightarrow{E_i} = \overrightarrow{E_i^q} + \sum_{j=1}^{N} T_{ij} \overrightarrow{\mu_j}. \tag{4.19}$$

with T_{ij}, the dipolar tensor defined as:

$$T_{ij} = \frac{1}{4\pi\epsilon_0} \left(\frac{f_5(r_{ij})}{r_{ij}^5} \begin{bmatrix} x^2 & xy & xz \\ xy & y^2 & yz \\ xz & yz & z^2 \end{bmatrix} - \frac{f_3(r_{ij})}{r_{ij}^3} I_3 \right), \tag{4.20}$$

Here, I_3 is the identity matrix. The f_5 and f_3 functions are introduced to account for short-range damping effects to prevent the so-called "polarization catastrophe", proposed by Thole [121] as:

$$f_3(r_{ij}) = 1 - exp(-a_{ij} r_{ij}^3), \tag{4.21}$$

$$f_5(r_{ij}) = 1 - (1 + a_{ij} r_{ij}^3) exp(-a_{ij} r_{ij}^3), \tag{4.22}$$

where a_{ij} is an adjustable damping parameter which depends on the nature of atoms i and j.

The polarization energy U_{pol} can be thus expressed as:

$$U_{pol} = -\frac{1}{2}\sum_{i=1}^{N}\overrightarrow{\mu_i}\cdot\overrightarrow{E_i}. \tag{4.23}$$

The induced point dipole method is the most used approach for molecular polarization. It is incorporated into several popular force fields, such as OPLS/PFF, AMOEBA.

To conclude, incorporating polarization effects significantly improves the description of non-bonded interactions, enhances the agreement with experimental or high-level *ab initio* calculations. Moreover, it has already been demonstrated that for systems with charged ions, ligands, polarized solvent molecules, it is compulsory to include polarization effects [47, 118, 122, 123].

4.2.2.4 Specific terms

Beside the cross terms already discussed in Section 4.2.1.5, other specific terms can be added to account for some special and specific interactions, such as charge transfer and hydrogen bonding. A part of the latter interactions is already taken into account by electrostatic, Van der Waals and polarization terms but it may not be sufficient for an accurate and more satisfying description of some systems. It is well known that force fields with fixed point charges cannot take into account all the non-negligible correlation between atomic charges and structure changes, even when they include polarization effects [47, 118]. Thus, for a more reliable description of the interaction potential function, two more contributions might need to be explicitly taken into account: the charge-transfer and hydrogen-bond contributions.

- Charge transfer (CT) Charge transfer refers to the partial electron transfer between a metallic center and its ligands. This contribution is introduced principally to describe better the charge fluctuation effects, alongside the polarization contribution [124, 125]. In the literature, charge transfer can be either considered as a non-additive or an additive contribution (For further details see Ref. [123]).

- Hydrogen bond (HB)

Hydrogen bond is often described as a partially electrostatic attraction between a hydrogen (H) atom which is bound to a more electronegative atom such as nitrogen (N), oxygen (O), or fluorine (F), etc. A ubiquitous

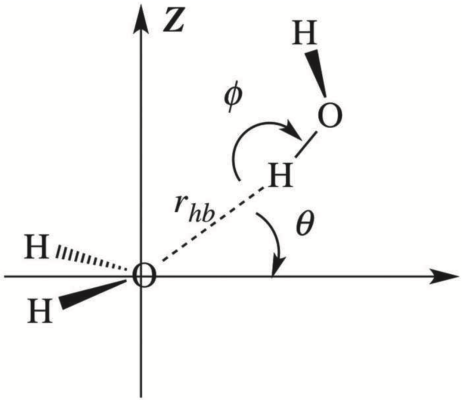


FIGURE 4.2: Definition of the geometrical parameters r_{hb}, θ and ϕ of the energy term U_{HB}. The axis X is the bisector of the angle -H-O-H of the water molecule whose oxygen accepts the hydrogen atom in a hydrogen bond. The axis Z is orthogonal to the plane HOH of the latter water molecule [118].

example of hydrogen bonding is found between water molecules. In fact, the very high boiling point, melting point, and viscosity of water are mainly due to the high number of hydrogen bonds each molecule can form. It also plays an important role in determining the three-dimensional structures and the properties of many synthetic and natural proteins. The simplest method to account for hydrogen bonds into implicitly incorporate the HB by parameterizing the L-J and the electrostatic potential to give accurate representation of the electrostatics of the specific system environment. Over time, several authors have introduced more refined hydrogen-bond potential functions reported in Refs. [126, 127]). In the TCPEp water model developed by our collaborator M. Masella [118], a short-range anisotropic potential is used to account for the directionality of the HB as displayed in Figure 4.2. The HB term is written as:

$$U_{HB} = \sum F(r_{hb}).G(\theta, \phi). \tag{4.24}$$

It sums up all the water HB pairs. r_{hb} corresponds to the HB bond length, θ may be interpreted as the angle between one water lone pair and the $O-H$ bond of the second water molecule participating to the HB, ϕ is $H....O-H$ angle. Lastly, F and G are Gaussian functions of the latter geometrical parameters (more details are available elsewhere in Ref. [118]).

4.2.2.5 1 - 4 interactions

The interactions between atoms separated by three bonds are known as 1 - 4 interactions. They are often treated in a very special way, as if they were bonded and non-bonded interactions simultaneously. Indeed, it is the combination of the dihedral potential with the Van der Waals and electrostatic interactions that determines the dihedral rotational barriers. Therefore, the intermolecular part is partially excluded by introducing a scaling factor. Generally, three ways are used to treat these interactions:

- By considering that these interactions are appropriately treated by intramolecular terms. 1-4 interactions are completely excluded (scaling = 0).
- Using the same parameters used to describe the Van der Waals interactions with a scaling factor (e.g., in amber FF, the Lennard-Jones parameters are used with a scaling factor of 0.5).
- Last, these interactions can be included by using a scaling factor on both electrostatic charges and Van der Waals parameters.

Regardless of the adopted method, it is crucial to keep in mind that the scaling can be quite different from one force field to another [41, 128–130]. This has to be taken into account if parameters coming from different force fields are mixed, so in general it is not advised to modify them or to combine parameters coming from different sources.

4.2.3 Parameterization of a Force Field

Before addressing the parameterization part, an extended definition of what is called “Force Field” (FF) is required. A FF can be defined as the union of three critical concepts: 1) The analytical form of the potential energy, 2) the parameter set for each interaction, and 3) the atom types listing. The potential energy function defines the functional form to compute each energetic contribution. In fact, each specific term contains one or more "parameters" which must be defined for each atom type or each pair of atoms within a molecular system. For example, for every bond between two particular atom types in the molecular system, there must be a corresponding force constant and bond equilibrium distance. The collection of these values defines a FF parameter set and the process by which they are derived is called the FF "parameterization". The

construction of a force field and its parameterization is known to be a crucial task, as it determines the accuracy, and transferability of the FF model. The transferability implies that the force field parameters for a given interaction site should be transferable between different molecules (e.g., identical parameters should be used for the alkyl group in n-hexane, 1-hexene, or 1-hexanol) and that can be used to simulate different properties (e.g., thermodynamic, structural, or transport) in the different state points (e.g., pressure, temperature).

4.2.3.1 Overview of the parameterization approach

Generally, four basic steps have to be followed for constructing/developing a new reliable FF:

- *Step 1: The choice of the functional forms to describe the system.* It should be both computationally efficient and flexible enough to capture the relevant physical interactions in the thermodynamic accessible regions of phase space. This choice will determine the accuracy, the speed of execution, the number of parameters to be adjusted. At this point one needs to decide whether to work with a polarizable or non-polarizable force field. It is worth mentioning that the more interaction terms added in the potential energy functional the more complex the FF becomes and hence its parameterization. However, this choice is deeply linked to the field of application of the FF model and the molecular system to be modeled.

- *Step 2: The choice of the reference data to which the parameters will be adjusted.* Two kinds of reference data can be used, separately or combined, in the optimization and development of FF models, namely experimental data, and QM calculations. In the early days of molecular dynamics, FF were mostly empirical and adjusted to experimental data, such as solvation free energies, interatomic distances, coordination number and density measured for solid and/or liquid phases. However, there are many systems for which experimental data are rare or difficult to obtain (expensive and time consuming). Additionally, they may be biased as they inherit the experimental uncertainties and most measurements generally correspond to an average behavior in the phase space area explored by the system. Sometimes we may have access to data on interaction distances and vibration frequencies which can be used for parameterization. Still, these data do not provide direct information on the interactions between particles and lack of explicit physical basis. On the other side, quantum chemistry

calculations offer the possibility of generating a large number of reference data essential to the parameterization of force fields (atomic/molecular data, dissociation curves, interaction energies, etc.), along with an understanding of the interactions, since the energy surfaces are calculated/computed rather than 'probed' using molecular properties. Nowadays, most of the newly developed FFs are adjusted to QM data and free from any experimental inputs.

- *Step 3: The adjustment of the FF parameters.* This step is very sensitive and important and needs to be done very carefully, since most of the parameters are coupled and can compensate each other; the change in one value may affect some other parameters. This is due to the ample parameter space, non-linear interdependencies of parameters, and limitations in the amount and quality of experimental and/or *ab initio* reference data. So they are very often refined in various stages. Deriving accurate FF parameters is exceptionally dependent on the quantity/quality of the reference data used. For instance, in polarizable FFs, there is a substantial risk to balance the dispersion parameters with the polarization ones if the reference data set is not large enough to explore the largest number of points in phase space (geometries where the two interactions will differentiate their behavior).
- *Step 4: Validation of the final set of parameters* by computing physical - chemistry properties of systems not employed in the parameterization procedure.

Generally speaking, the parameterization protocols followed for the development of molecular potentials depend strongly on the reference data set used and the areas of application, as well as on the accuracy needed. One should always keep in mind that while parameterizing a FF one is not looking for the "exact" solution, but only for a physically-chemically sound, and reliable set of parameters, with which one can pursue MD simulations with a sufficient level of confidence (i.e., reproduction of thermodynamics and structural data). After all, it is a classical model describing purely quantum system.

4.2.3.2 The followed parameterization approach

The applied parameterization strategy, in the case of using QM reference data, is often labeled "Bottom-Up", meaning that the parameters are derived to match the atomic/molecular-scale information. Later these

parameters will be used to perform MD simulations and the results will be confronted to the higher-scale experimental data. This strategy is described in Figure 4.3.

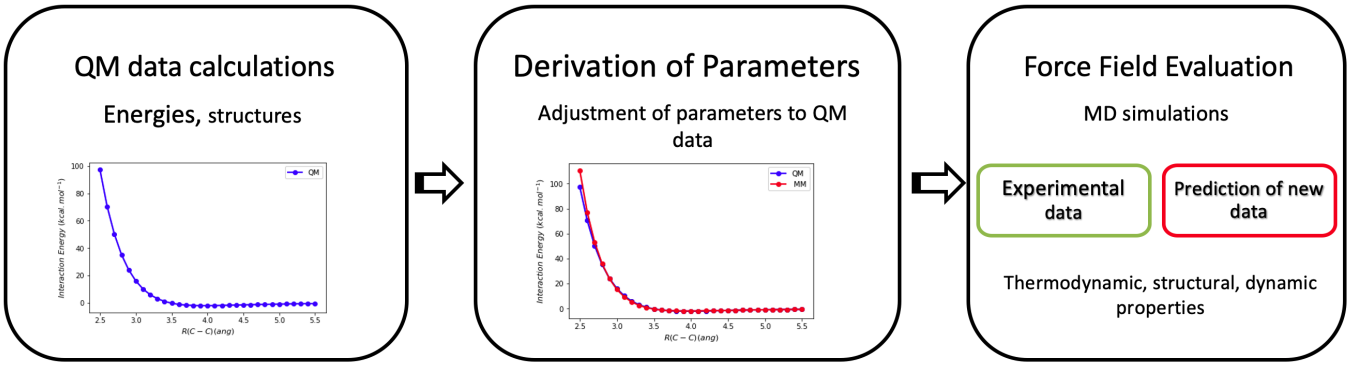


FIGURE 4.3: The followed parameterization procedure.

This protocol involves the production of reference data using QM methods (intramolecular, binding interactions, dissociation energies, etc.). These data serve as reference to adjust the MM potentials. In our case we use the PEST optimization code [131] to refine the parameters. In the latter we have the possibility to either fix the parameters or optimize them within a user-defined range of values. Starting from initial parameters, we calculate the same quantities (i.e., interaction energies, dissociation curves, etc.) for the target QM structures using Polaris-MD simulation package [132]. The PEST code iteratively changes the parameters to match the reference energy data using non-linear least squares method, till the best possible parameters set is found. While using PEST, one should always keep in mind that PEST is aware neither of the physical sense of the parameters values or nor of the physical phenomena occurring, meaning that the program will search for the parameters which will best reproduce the reference data and not the best ones describing the real physics behind. For this reason, the choice of the initial values is crucial and one needs to consider as many QM reference data as possible. At last, MD simulations are performed to evaluate the new parameters set by predicting condensed phase properties namely the density and heat of vaporization, diffusion coefficient, etc. By comparing these simulation results to experimental ones, one validates or not the accuracy of the FF.

4.3. Molecular Dynamics in Various Statistical ensembles

As already mentioned, MD simulations provide the means to solve the equation of motion of the particles by generating information at the microscopic level of a molecular system, such as atomic positions and velocities. The conversion of these microscopic quantities to macroscopic observables such as pressure, energy, density, heat of vaporization, heat capacities, etc., requires the use of statistical mechanics. The latter is the only possible way to connect microscopic simulations to macroscopic properties as it provides the rigorous mathematical expressions that relate macroscopic properties to the distribution and motion of the atoms and molecules of a N-body molecular system.

In general, measuring or calculating the properties of any particular system must be carried out by taking into account and controlling the thermodynamic variables, such as T temperature, P pressure, V volume. The definition and control of these parameters lead to what is known as statistical (thermodynamics) ensemble. In MD simulations it is possible to work within different types of thermodynamic ensembles using special algorithms and methods to control these thermodynamic quantities.

Before discussing in more details how temperature and pressure are controlled in MD simulations, it is extremely important to keep in mind that, in MD simulations, the properties are calculated as time averages, but experimental observables are assumed to be ensemble averages. This leads to a fundamental hypothesis of statistical mechanics, the ergodic hypothesis, which states that the time average equals the ensemble average. Hence, the system should be allowed to evolve in the state space for an appropriately long time, so that the system will eventually explore all possible states.

In a classical MD simulation, the natural evolution of a system of N particles in a volume V imposes that the total energy E is a constant of motion, making this simulation equivalent to the N,V,E micro-canonical ensemble. This kind of ensemble can only be used for isolated systems. However, it is often advantageous and demanded to perform simulations in other ensembles such as N,V,T and N,P,T. This is possible only by controlling the temperature and pressure during the simulations. At first sight this seems impossible but fortunately, over the past years these problems have been overcome. In the following part, we will discuss briefly the methods to control temperature and pressure control in MD simulations.

4.3.1 Constant temperature: canonical ensemble (NVT)

Modeling a system at constant temperature T, so that the system trajectory would visit all regions in the phase space according to the canonical distribution (NVT) was a challenge over years. Many approaches have been proposed to achieve this, such as rescaling the atomic velocities to force the system to be at the required temperature, or supplementing the equations of motion of atoms by an artificial "equation of motion" for the total kinetic energy which drives it to the correct value corresponding to the preset temperature [133]. However, these two approaches have failed to successfully generate the canonical distribution since the kinetic energy of the particles actually fluctuates during the simulation which makes the desired temperature difficult to reach, even so, those two approaches are still being used in special cases and studies [42].

At present, the most common and widely used manner to modulate the temperature of a system is the coupling to a heat bath, so-called "thermostat". A thermostat can be seen as a modification of the Newtonian MD scheme with the purpose of generating a statistical ensemble at a constant temperature. A variety of thermostat methods are available to add and remove energy from an MD system in a realistic way, approximating the canonical ensemble, such as Anderson thermostat [134], Berendsen thermostat [133], Nosé-Hoover thermostat [135], and Langevin (stochastic) thermostat [41]. Since, in the used MD code Polaris-MD [132], the temperature is modulated by a Nosé-Hoover thermostat, the latter will be the only one discussed in more details.

In 1984, Nosé and Hoover have proposed a technique that can generate a canonical ensemble by adding to the energy function a fictitious variable s that "stores" the kinetic energy and removes or adds kinetic energy/temperature within the degrees of freedom. So that the system evolves according to the micro-canonical distribution in the extended phase space; at the same time, in the actual phase space the distribution is exactly canonical, making the Nosé-Hoover thermostat a deterministic scheme to generate the canonical ensemble. In this case, the Hamiltonian H of the N-particle system can be written as:

$$H(\overrightarrow{r}^N, \overrightarrow{p}^N) = \sum_i^N \frac{\overrightarrow{p_i}^2}{2m_i} + V(\overrightarrow{r_i}) + \frac{Q\zeta^2}{2} + g\frac{ln(s)}{\beta}, \tag{4.25}$$

where $\zeta = \frac{ds}{dt}$, s is the additional fictive variable acting as external system, $g = 3N$ is the total number of

degrees of freedom and Q plays the role of fictitious mass associated to s and $\beta = \frac{1}{k_B T}$. The corresponding equations of motion are

$$\frac{d\overrightarrow{r_i}}{dt} = \frac{\overrightarrow{p_i}}{m_i}, \tag{4.26a}$$

$$\frac{d\overrightarrow{p_i}}{dt} = -\frac{\partial V(\overrightarrow{r_i})}{\partial \overrightarrow{r_i}} - \zeta \overrightarrow{p_i}, \tag{4.26b}$$

$$\frac{d\zeta}{dt} = \frac{1}{Q}\left[\sum_i^N \frac{\overrightarrow{p_i^2}}{m_i} - \frac{g}{\beta}\right]. \tag{4.26c}$$

To perform a NVT simulation, these equations of motion of the Nosé-Hoover thermostat can be implemented by a small modification of the velocity Verlet algorithm or any other MD propagation algorithm (see Refs. [41, 42]). It has to be mentioned that the use of any special thermostat depends solely on the application field. Furthermore, every existing thermostat, till now, has its own advantages and disadvantages as reviewed in more details in Ref. [42].

4.3.2 Constant pressure and temperature: NPT ensemble

Generally, most real macroscopic systems maintain a constant pressure instead of constant volume, the volume being a dynamic variable that changes over time. Thus, MD studies at experimentally relevant conditions require the simulations not only to be performed at an imposed temperature, but often also at constant pressure. Various algorithms have been developed to impose a constant pressure along an MD trajectory. The ideas were pretty analogous to those used for the constant temperature, by using the so-called "barostats". We can name the Berendsen barostat, Gauss barostat, Langevin barostat. However, the most famous and widely used barostat in MD simulation codes, is the weak coupling method of Andersen et al. [134]. The reason behind its celebrity is the easiness of combining it with other thermostats to build the NPT ensemble [41, 42, 136].

In the Andersen barostat the system is coupled to an external variable V, the volume of the simulation box. The coupling of the system to the volume is made by a piston that applies an isotropic (uniform) expansion or compression to the system. This piston of “mass” M_p is associated with a kinetic energy $E_V = \frac{1}{2} M_p \dot{V}^2$ and the potential energy of the volume V is $U_V = PV$ where P is the desired pressure. The potential energy

U and the kinetic energy E of the system's particles are re-expressed as a function of scaled coordinates $\overrightarrow{s}$:

$$\overrightarrow{s_i} = \frac{1}{V^{\frac{1}{3}}}\overrightarrow{r_i}. \tag{4.27}$$

The rescaling of the cell tensor (volume) is then carried out at every integration step to maintain the difference between the instantaneous internal pressure and the external stress applied on the cell box. The Hamiltonian of the system can then be written as:

$$H(\overrightarrow{s}^N, \overrightarrow{\dot{s}}^N) = \frac{V^{2/3}}{2}\sum_i^N m_i \overrightarrow{\dot{s}_i}^2 + U_{pot}(V^{\frac{1}{3}}\overrightarrow{s}^N) + \frac{1}{2}M_p\dot{V}^2 + PV, \tag{4.28}$$

where $\overrightarrow{\dot{s}}^N = \frac{d\overrightarrow{s}^N}{dt}$. From this Hamiltonian, and after simple math, the equations of motion can be rewritten as:

$$\frac{d\overrightarrow{r_i}}{dt} = \frac{\overrightarrow{p_i}}{m_i} + \frac{1}{3}\frac{\dot{V}}{V}\overrightarrow{r_i}, \tag{4.29a}$$

$$\overrightarrow{p_i} = m_i V^{1/3}\frac{d\overrightarrow{s_i}}{dt} \tag{4.29b}$$

$$\frac{d\overrightarrow{p_i}}{dt} = \overrightarrow{F_i} - \frac{1}{3}\frac{\dot{V}}{V}\overrightarrow{p_i}, \tag{4.29c}$$

$$\ddot{V} = \frac{1}{M_p}[P(t) - P], \tag{4.29d}$$

where V is the volume and $P(t)$ is the instantaneous pressure P is the imposed (desired) pressure, and $\vec{r}$, $\vec{p}$, m, and $\vec{F}$ are the position, momentum conjugate to s, mass, and force, respectively, for each atom.

The solution to these equations produces trajectories in the isobaric – isoenthalpic (NPH) ensemble where the particle number, pressure, and enthalpy of the system are constant. Nosé-Hoover thermostats are combined with the Andersen barostat (see Section 4.3.1) to enable simulations in the NPT-ensemble. Thus, the real variables of the system are derived from the virtual parameters by scaling simultaneously the coordinates by $V^{1/3}$ and the time step δt by the parameter s, which represents the thermal heat bath (thermostat) [137].

4.4. Periodic Boundary Conditions (PBC)

Using MD simulations to calculate the properties of bulk gasses, liquids, crystals, or mixtures is limited by the restricted number N of particles included in the model system. The latter depends strongly on the computer memory and the speed of execution. This often leads to select only a small finite sample to represent an extended system. Hence, by simulating a finite-size system in a box, the particles hitting the surfaces bounce back inside which leads to significant edge effects (boundary condition artifacts). These contributions of surfaces affect any physical property of the simulated fluid. In order to overcome this problem, one can impose what is known as periodic boundary conditions (PBC). The latter ones mimic an infinite bulk surrounding our N-particle model system as illustrated by Figure 4.4. It consists of an endless replication of the box in space according to the three space directions. If, during the simulation, a particle leaves the box, it is replaced by its image on the opposite side of the box and makes the system an infinite one. However, only one – the nearest – image of each particle is considered for short-range non-bonded interaction terms, namely Van der Waals interactions. For long-range electrostatic interactions, this would not be accurate enough. Therefore, most MD codes incorporate lattice summation methods such as Ewald Sum, and Particle Mesh Ewald (PME) [138], where the long-range interactions are efficiently evaluated by calculations in the Fourier space, assuming that the periodicity of the target molecular system is infinitely repeated.

Lastly, it is crucial to keep in mind that the size of the simulation box should be large enough to prevent periodic artifacts from occurring due to the non-physical topology of the simulation. If a box is too small, a molecule can interact with its own image in a nearby cell; it is equivalent to the interaction of the "head" of a molecule with its own "tail" (further details are available elsewhere [42, 139]).

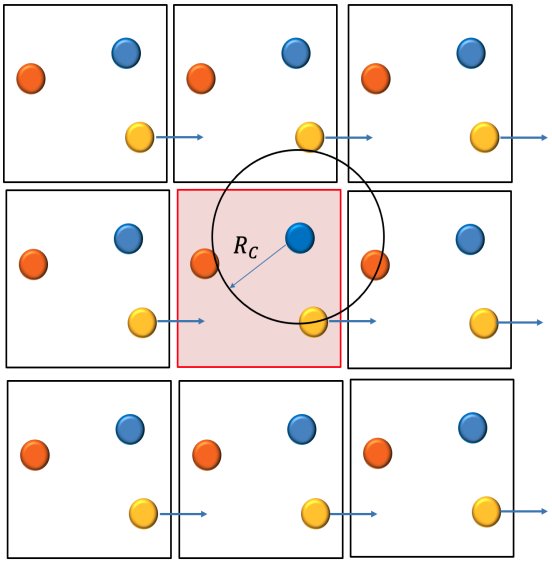


FIGURE 4.4: Diagram of the periodic conditions during a simulation by molecular dynamics. The central square (in red) represents the primary simulation box surrounded by its periodic images. If a particle leaves the central box, it is replaced by its image by the opposite side. The circle surrounding the "blue" particle represents the "cut-off" distance up to which short-range non-bonded interactions are calculated.

4.5. Physical properties derived from MD simulations

When we perform a real experiment, we proceed as follows. We prepare a sample of the material that we wish to study. We fix the experimental conditions, we connect this sample to a measuring instrument and we measure the property of interest during a certain time interval. In an MD simulation, we follow exactly the same approach. First, we choose an initial configuration of the system, a starting point at $t = 0$; the choice of the initial configuration must be done carefully as this can influence the quality of the simulation, and we set the "experimental" conditions at which we wish to perform the simulation (temperature, pressure). Secondly, the equilibration step is performed to heat and calibrate the system so that it reaches the desired simulation conditions. Once the desired conditions are reached, the simulation of the system continues until several properties converge on average with respect to time, such as temperature, pressure, energy, and structure. The final step of the simulation is the "production" phase for the desired time length. This can last from several hundred of picoseconds to tens of nanoseconds or more depending on the desired property. It is during this phase that all thermodynamics or physicochemical properties can be calculated. In order to compute or simulate a certain property, we must first of all be able to express this observable as a function of the positions and momenta of the particles in the system (selecting the statistical ensemble), the conditions

of the simulations should be the same (which is true with accurate FFs) as those normally encountered in experiments. In the following part, a brief description of some physical properties that can be obtained from MD simulations.

4.5.1 The bulk density

Bulk density, also known as specific mass, is defined as the particles mass divided by the total volume they occupy. NPT ensemble is the best choice to calculate this property, since, as the molecular system evolves during the simulation in such a way that the internal pressure and temperature match the corresponding external values, hence reaching the equilibrium state.

The average bulk density can be computed, by performing NPT simulations, using the formula:

$$\langle \rho \rangle = \frac{m}{\langle V \rangle} = \frac{N_{mol} M}{\langle V \rangle \, N_A}, \tag{4.30}$$

where $\langle V \rangle$ is the average volume of the simulation box, N_{mol} is the number of molecules in the simulation box, M is the molar mass of the molecule and N_A is the Avogadro constant. MD codes are most likely based on molar mass and atomic count to estimate density. Generally, this property is printed out in the output files of most MD simulations codes. It is also one of the most used reference data for developing the first generation empirical FFs.

4.5.2 Heat of Vaporization

The enthalpy (or heat) of vaporization, which represents the amount of enthalpy that must be added to a liquid substance to transform it into a gas. It can be calculated with equation 4.31, where $H(P,T)_{gas}$ and $H(P,T)_{liquid}$ are the enthalpies in the gas and liquid phases, respectively. $E(T)_{gas}$ and $E(T)_{liquid}$ are the total energies in the gas and liquid phases, respectively:

$$\Delta H_{vap}(T) = H(P,T)_{gas} - H(P,T)_{liquid}, \tag{4.31a}$$

$$\Delta H_{vap}(T) = E(T)_{gas} - E(T)_{liquid}. + P(V_{gas} - V_{liquid}) \tag{4.31b}$$

In the framework of ideal gas theory, the kinetic energies of a molecule in gas and liquid phases are identical, and the liquid volume is negligible compared to the gas one. Hence, Eq. 4.31 can be simplified:

$$\Delta H_{vap}(T) = U_G^{pot} - U_L^{pot} + k_B T, \tag{4.32}$$

where, $U_L^{pot} = \frac{U_T^{pot}}{N_{mol}}$ is the potential energy of a molecule in the liquid phase and U_T^{pot} is total potential energy of the system, and U_G^{pot} the potential energy of one molecule in the gas phase (in a vacuum). A correction term C can be added to accounts for the difference in vibration energy-calculated quantum mechanically and classically, as well as the polarization and non-ideal gas effects. In most cases this term is very small and neglected. The term U_G^{pot} can also be obtained by performing an optimization of the molecule in the gas phase:

$$U_G^{pot} = U_{gas}^{min} + \frac{1}{2} k_B T (3N_{atm} - 6 - N_{cons}), \tag{4.33}$$

where U_{gas}^{min} is the minimized energy, N_{atom} and N_{cons} are the number of atoms in the molecules and the number of the constrained degrees of freedom, respectively. The second term represents the vibrational energy of the molecule.

One can make a further approximation and assume that the intramolecular (bonded) energy in the liquid phase is the same as that in the gas phase. As a result equation 4.32 can be further simplified to :

$$\Delta H_{vap}(T) = -U_{inter} + k_B T, \tag{4.34}$$

where U_{inter} is the mean intermolecular energy per molecule along a MD trajectory. This method is sufficient and recommended only for small molecules. For large ones, this approximation is no longer valid since the intramolecular energy in the gas phase is quite different than the liquid phase (especially the torsion part).

As far as we can tell, the best way to compute such a property using MD methods is to perform a NVT simulation and to use the standard equation 4.32. For further details on these methods and applications, we refer to Ref. [140].

In our work on organic solvent discussed in Chapter 5, we denoted the first protocol (P1), when Eq. 4.32 is applied to calculate ΔH_{vap}, P2 when Eq. 4.33 is used to calculate the potential energy of the gas phase, and lastly, P3 for Eq. 4.34.

4.5.3 Radial distribution function (RDF)

Radial distribution function (RDF), also known as pair correlation function, can be seen as a measure of the probability of finding a particle at a distance of r away from a given reference particle. It is an effective way of describing the average structure of disordered molecular systems, the general algorithm involves on determining how many particles are within a distance of r and $r + dr$ away from a reference particle. This general idea is illustrated in Figure 4.5, where the blue particle is our reference particle, and the red particles are those whose centers are within the circular shell, dotted in orange.

The distance corresponding to the maximum of a peak indicates a higher probability for the two atoms considered to be at this distance from each other than the mean probability.

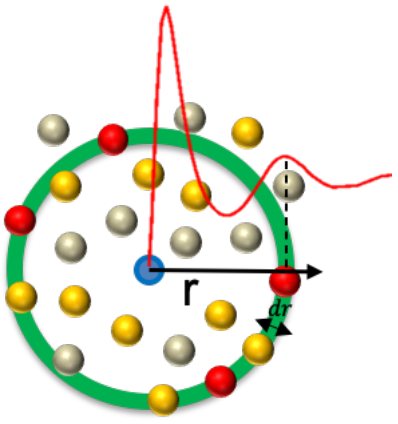


FIGURE 4.5: Illustration of the radial distribution function calculation.

The RDF can be expressed in the following way :

$$g(r) = \frac{1}{N\rho} \left\langle \sum_{i}^{N} \sum_{j \neq i} \delta(r - r_{ij}) \right\rangle . \tag{4.35}$$

where ρ stands for the average density of particles ($\rho = N/V$, number of particles in a volume V).

4.5.4 Diffusion coefficient

Molecular dynamics simulations have the advantage of explicitly including the notion of time. We can therefore use it to characterize the dynamics of the molecular system, by calculating what is known as the diffusion coefficient. The latter reflects the mobility of the molecule within the fluid. It is due to the Brownian (random) movement of molecules. There are two equivalent methods for determining this coefficient [41]. The first is based on Einstein's relation linking the mean square of the displacement and the observation time:

$$D = \frac{1}{6} \frac{\lim_{\tau \to \infty} \left\langle [\overrightarrow{r_i}(t+\tau) - \overrightarrow{r_i}(t)]^2 \right\rangle}{\tau}, \tag{4.36}$$

where r_i stand for the position of the i^{th} molecule and τ is the observation time. During the simulation, the displacement of each particle of its original position is a function of time. By plotting the average of the MSDs of all particles as a function of the simulation time, the diffusion coefficient D is calculated from the slope of the line obtained at long times. Hence, much longer simulation is required to get a reliable diffusion coefficient, up to several ns.

The second method to determine the diffusion coefficient from the integral of the velocity autocorrelation function, known as the Green-Kubo relation:

$$D = \frac{1}{3} \int_0^\infty \langle \overrightarrow{v}_i(t+\tau) . \overrightarrow{v}_i(t) \rangle \, d\tau = \frac{1}{3} V ACF(t), \tag{4.37}$$

where $\overrightarrow{v_i}$ stand for the velocity of the i^{th} molecule. This method is considered, in literature, better than the previous one, as it yields better agreement with experimental data and does not require long simulation [141].

The velocity autocorrelation function (VACF) is an interesting property since it indicates the fundamental nature of the dynamic processes of a molecular system, the time scale for changes in atomic motion. A code was written in python to do the transformation of the trajectories (velocities of the atoms at different instances of time) to the velocity auto correlation function (VACF) following Eq. 4.38:

$$VACF(t) = \int_0^\infty \langle \overrightarrow{v}_i(t+\tau).\overrightarrow{v}_i(t)\rangle \, d\tau. \tag{4.38}$$

Taking into consideration that MD simulations provide atomic information, it can be computed for any atom type of the molecular system. However, the best and recommended way to calculate such property is to be considered the center of mass's motion of the molecule.

Chapter 5

Force field for alkanes and amides derivatives

As already mentioned, in the nuclear field, amide derivatives can be used as extracting molecules for actinide ions for nuclear fuel solutions as alternatives to the TBP molecules, with advantages already listed in the introduction. Alkanes such as dodecane, TPH (Tetra Propylene Hydrogenated), kerosene,and so forth [84, 85, 142] are hydrocarbons suitable as solvents to dilute TBP and/or amide derivatives and to achieve the extraction of U and Pu from the irradiated nuclear fuel. However, the use of these molecules is not limited to the nuclear field, but impacts other research areas, such as biology, medicine and petrology. In these latter, the structural, thermodynamics and transport properties are of interest for the design and study of artificial or biological membranes and also play a major role in the recovery and refining of crude oil [143–150]. One way to obtain such information at the molecular level involves the use of atomistic molecular dynamics simulations. In this regard, several force fields have been developed over the years, such as CHARMM, GROMOS, MM4, OPLS and AMBER [128–130, 151, 152]. Generally, the hydrocarbons (alkanes) parameters are used to describe alkyls chains regardless of whether they are amides, acids, amines chains or lipids, peptides and some proteins tails in biology applications. Most of these FFs have been derived focusing on individual aspects and based on the reproduction of experimental data, such

as enthalpies of vaporization, vapor pressures and densities. However, the FF parameters are not always transferable to all the molecular group series, and they are reoptimized before use in most cases. As a result, different versions of AMBER, CHARMM, OPLS FFs for hydrocarbons have been developed over the years for a better description of the macroscopic properties [153–155]. For example, the first version of OPLS FF for hydrocarbons (labeled OPLS-AA) is quite successful for short hydrocarbons but not for long [152]. This comes from the fact that they used experimental data measured for small molecules to develop such FF, and hence inducing a lack of physical meaning of the interactions for longer chains. Therefore, Siu et al. have reoptimized the OPLS-AA parameter set for long hydrocarbons, termed L-OPLS [154]. Generally, FF parameters obtained for small molecules do not always represent the real interactions between atoms, since the goal was to reproduce the reference experimental data in an average way. In summary, in the literature exists a zoo of FFs, the latter derived for individual molecules and specific cases. Again, this makes the prediction of new features of newly designed molecules challenging and questionable.

For amides derivatives, most available FFs have been derived for small molecules and then combined with the hydrocarbons parameters to describe the alkyl chains of the long amides. As in the case of hydrocarbons, there are several FFs available in the literature but they are not always transferable and accurate enough for direct use. Hence the re-optimization of the FF parameters is often required [156–158]. For amides derivatives, it was shown that most of the transferability problems arise *i)* from the fact that amides are polar molecules and thus FFs describing this kind of molecules should include these effects explicitly [158]; *ii)* and, of course, from the parameters used to describe the alkyls chains.

Nowadays, the use of *ab initio* data for the parameterization of FFs has become increasingly common, since a solid physical/mathematical foundation provides a better understanding of the physics and chemistry of the systems to be investigated. Moreover, most transferability problems encountered for force fields with empirical parameters are related to the reference data used. If the systems used in the parameterization process significantly differ from the ones being investigated and/or if the amount of data set used for parameterization is small, limited to certain kinds of data at restricted temperature and pressure conditions, the parameters may not be as trustworthy. The available experimental data are often limited to certain molecules at specific experimental conditions, while, quantum chemistry methods in combination with the availability of significant computational resources offer the possibility to generate a large number of reference

data essential to the development of force fields (atomic/molecular data, dissociation curves, interaction energies, etc.). However, we should also mention that even quantum chemistry methods have limitations, they are limited to the inclusion of a small number of atoms and therefore does not take into account the influence of the environment in the calculation.

Herein, we have developed a new set of FF for alkanes and amides derivatives based solely on quantum chemistry calculations. Since alkanes are known to be a non-polar molecules, polarization effects were neglected and only Coulomb, repulsion and dispersion interactions were considered. However, it should be mentioned that the alkanes do not induce polarization for alkanes-alkanes interactions but that they are polarizable in the model for alkanes-amides interactions. As for the amides, polar molecules for which polarization forces play an important role, polarization was incorporated alongside Coulomb, repulsion and dispersion interactions to ensure the transferability of the parameters for longer amides, as well as the correct description of the intermolecular interactions.

First, parameters for alkanes and amides model molecules have been derived following the same approach and based on the same kind of data, namely *ab initio* data. Then, we have combined the two sets of parameters to describe the large amides, for which the alkyl chains have been described with the derived alkane parameters. The newly proposed parameter sets were validated on physical properties of interest, namely density, heat of vaporization as well as on the distributions of trans and gauche conformation for alkanes.

5.1. Parametrization methodology

The functional form of the force field embeds six components. Covalent interactions between atoms are modeled using harmonic bond stretching and angle bending parameters, while rotations around a bond are described by anharmonic 4-body torsional terms (See Section 4.7a). Non-bonded interactions are described by a sum of Coulombic interactions between atom-centered point charges (See Section 4.2.2.1), a physically motivated Buckingham interaction, which combines a short-range repulsive $\left(Ae^{-Br}\right)$ potential with a longer-range attractive $\left(\frac{C}{r^6}\right)$ interaction (See Section 4.2.2.2), and lastly polarization was incorporated with the induced point dipoles model (see Section 4.2.2.3).

The bonded and non-bonded parameters of this work were obtained following the parameterization methodology discussed in section 4.2.3.2. The strategy is as follows. First, different charge calculation methods have been tested to select and compute atomic partial charges. Second, atomic polarizabilities were derived by partitioning the molecular polarizability, the latter being computed at the MP2 level of theory. Third, Van der Waals parameters, namely A, B and C in the Buckingham potential, were subsequently adjusted to reproduce interactions energies for different dimers orientations and for different molecules. Simultaneously, torsion scans using *ab initio* gas phase calculations were performed to adjust the torsion parameters. Lastly, we proceed to the validation of the resulting parameter sets by performing MD simulations and comparing the simulated macroscopic data to experiment.

The choice of the quantum chemical level to calculate the reference data used for the development of any FF is fundamental, as to ensure that the different interactions are accurately treated. For instance, dihedral potentials are crucial in order to reproduce the (temperature-dependent) fraction of trans and gauche isomers, as the energy difference between the two conformations is small (about 0.6 kcal $\cdot$ mol^{-1}) and must be well described. In this work, all the structural and energetic data, needed for the parametrization of our force field, were calculated using the Molpro quantum chemistry package at the MP2 level (the Møller-Plesset Perturbation Theory) [52] with the correlation consistent aug-cc-pVTZ basis set of Dunning [74]. This level of theory was chosen based on two arguments (1) the geometries of simple alkanes are known to be less sensitive to the size of the basis set than the energies itself [159] (2) by performing a benchmarking of MP2 interaction energies with respect to Coupled Cluster "CCSD(T)" [160] (the gold standard) (see Figure B.3 in Annex B). In conclusion, the MP2 approach proved to provide a good compromise between the computational cost and the accuracy of the computed interaction energies.

For different dimer orientations, the MP2 interaction energies were computed using the super-molecule approach taking into account the basis set superposition error (BSSE) :

$$IE = E_{AB}(AB) - E_{AB}(A) - E_{AB}(B), \tag{5.1}$$

where, $E_{AB}(AB)$ is the energy of their interacting assembly (dimer). $E_{AB}(A)$ and $E_{AB}(B)$ denote the total energies of monomers A and B, respectively, computed with the dimer AB basis sets, i.e. in the calculation

of monomer A including the basis set of the B monomer (but neglecting the nuclear charges and electrons of B).

5.2. Results of parametrization: bonded and non bonded parameters

Before presenting the parameters of different FF contributions, we show in Table 5.1 the atom types used in this work. We define HA as the hydrogens in the aliphatic compounds and HN for hydrogen atom bonded to the special carbon atom CT2-N. For carbons, we introduce four types: methyl carbon (CT3), methylene carbon (CT2), alkane branched carbon (CT1) and (CT2-N) for methylene carbon bonded to amide nitrogen atom. The oxygen of the carbonyl group is denoted O, the carbon atom double-bonded to an oxygen atom (C), the nitrogen atom of the amide function (N), respectively.

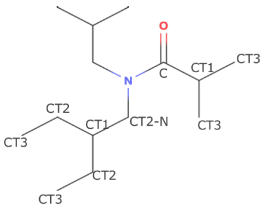
No.	Atom Type	description	Illustration of atom type introduced
1	CT3	Alkane methyl carbon $-C-\mathbf{C}H_3$	
2	CT2	Alkane methylene carbon $-C-\mathbf{C}H_2-C$	
3	CT1	Alkane branched carbon $-C-\mathbf{C}H-C$	
4	CT2-N	Methylene carbon bonded to amide nitrogen atom	
5	HN	Hydrogen atom bonded to CT2-N	
6	HA	Aliphatic Hydrogen (for alkane and amides)	
7	N	Amide nitogen atom	
8	O	Amide carbonyl oxygen	
9	C	Amide carbonyl carbon	

TABLE 5.1: Atom types and their definitions in Polaris-MD.

Bond and angle parameters

In most available FFs in the literature, bond stretching and angle bending are described using harmonic potentials that contain the bond and angle force constants and equilibrium bond lengths and angles (See Eqs 4.9,4.8). Historically, reasonable values for the equilibrium bond lengths r_0 can be obtained from X-ray diffraction experiments, while the spring constants may be estimated from infrared or Raman spectra. An alternative route to obtain these parameters is to use quantum chemistry methods, in which case the equilibrium bond lengths and angles are taken from the optimized geometry of the molecule, and the motion force constants can be derived either directly from the QM computed Hessian matrix, or indirectly by

performing scans over the bond/angle of interest, considering how the energy of a bond/angle changes with its length/angle. In this work we have opted for the second approach to derive the bond and angle force constants for alkanes, and we have compared our parameters with that proposed by the CHARMM and Amber FFs. It turned out that for alkanes, the CHARMM FF parameters matched very well the QM scans, and could be taken without further refinements, see Annex B. Thereby, we have used the angle and bond parameters described in CHARMM FF for both alkane and amide derivatives.

Charges and atomic polarizabilities:

Partial charges are known to be the first intermolecular parameters to be defined in the development of a new FF from scratch. However, there are several methods to calculate such atomic charges (NBO, ESP, RESP, Mulliken, CM5, Hirshfeld, etc.) [161–166]. They have been always developed with the aim of providing the most realistic description of the system, yet atomic charges are not quantum observables. Electron density can be easily calculated and studied, but, there are no operators to unambiguously determine the charge associated to each atom. In fact, assigning accurate partial atomic charges has long been a significant challenge in the general use of modeling methods. The choice of the method depends mainly on the reproduction of the electrostatic interaction between molecules and the class of the force field developed (static or polarizable). In general, Mulliken charges are avoided, as they exhibit a strong dependence on the method and basis set used. For NBO (Natural Bond Orbital Analysis) and RESP (Restrained Electrostatic Potential) they are often used in CHARMM, AMBER and OPLS FFs. Recently, CM5 (Charge Model 5, latest update of the CMx series), a method that evaluates the partial charges, was developed by Cramer, Truhlar, and co-workers, seems to give good results. It uses the charges obtained from a Hirshfeld population analysis (of a wave function obtained with density functional calculation) as a starting point. The charges are then varied based on some specific parameters, derived originally by fitting to gas-phase dipole moments of several molecular structures. Jorgensen et al., developers of the OPLS FF series [167], stated that CM5 charges yielded the best agreement with experiments in pure liquid simulations, with the extra advantage of being essentially basis set independent.

To assess our methodology, we have calculated the atomic charges for a series of alkanes (ethane, butane, heptane) by different methods (NBO, RESP, Mulliken, CM5) and they are reported them in Table 5.2. We

Molecule	Mulliken			RESP			NBO			CM5		
Atom types	CT3	HA	CT2	CT3	HA	CT2	CT3	HA	CT2	CT3	HA	CT2
Ethane	-0.85	+0.28	—	-0.20	+0.07	–	-0.23	+0.07	–	-0.22	+0.07	–
Butane	-0.95	+0.26	-0.39	-0.17	+0.04	-0.20	-0.22	+0.07	-0.12	-0.21	+0.07	-0.14
Heptane	-0.97	+0.26	-0.20	-0.09	+0.03	-0.18	-0.22	+0.07	-0.13	-0.20	+0.07	-0.14

TABLE 5.2: Partial charges calculated for a series of alkanes using different methods (in a.u.).

can see clearly the difference between the values obtained from the different methods. Mulliken charges seem to be overestimated as compared to other methods, while the three others give very similar values, especially for NBO and CM5. It is clear from these values that, for alkanes, electrostatic charges are small and polarization effects can be neglected. Electrostatics potential along with Van der Waals interactions are sufficient to model long-range interactions for alkanes. Anyway, for all the reasons discussed above, CM5 charges are used in the FFs of alkanes and amides.

To include polarization effects within amide derivatives in our force field, relying on an induced dipole moment approach including Thole's damping effects, we have considered only the non-hydrogen atoms as polarizable centers, assigning to each center an atomic polarizability. There are several procedures to decompose the molecular polarizability into atomic polarizabilities [168]. They differ on whether molecular polarizability was obtained from experimental refractive indices or from quantum mechanics. In this work, we have opted for the method proposed by Marenich et al. [169] for partitioning the molecular polarizability into atomic contributions by the use of Hirshfeld population analysis [170], involving the numerical differentiation of the dipole moments computed for different values of the applied external electric field.

To sum up, Table 5.3 gathers the partial charges and atomic polarizabilities assigned for each atom of the molecules of interest.

Buckingham parameters

To derive the Van der Waals parameters, interactions energies were computed using the super molecule approach (Eq. 5.1) for several different dimer relative orientations and distances, starting from ethane, propane, butane, heptane and $n-$dodecane for alkanes. The same work was done to for the amides derivatives VdW parameters, we have selected the N,N-DiEthyl-PropanAmide (DEPA) and N,N-diethyl-2-methyl-PropAmide

TABLE 5.3: Partial charges (in a.u.) and atomic polarizabilities (α in Å^3) used for alkanes and amides derivatives

Atom Type	Partialchargesq(CM5)	atomicpolarizabilityα (Å^3)
CT3	–0.21	2.0
CT2	–0.14	2.0
CT1	–0.07	2.0
CT2-N	–0.02	2.0
HN	0.10	0.0
HA	0.07	0.0
N	–0.34	1.3
O	–0.42	1.3
C	0.32	1.0

DEMPA (see Figure 5.1 and Annex B for more details). In this work, in order to later speed up the MD simulations and ease the parametrization process, coulomb, polarization, repulsion and dispersion contributions were only considered for interactions between carbon, nitrogen, oxygen atoms. As for all interactions that involve hydrogen atoms, only the coulomb and repulsion terms were kept. Consequently, a total number of of 77 parameters were adjusted, with partial charges and atomic polarizabilities been fixed (18 parameters).

The size and the flexibility of these systems generate a large number of degrees of freedom to explore the potential energy surfaces, and making a random generation of all possible configurations appears to be meaningless. For that purpose, some structures have been chosen wisely so that we can selectively enhance some pair interactions and decrease some others. This allows to explore/examine specific interactions and fit the desired pair interactions. The parameters were optimized iteratively in stages. At the final stage, after adjusting the interactions for each pair, we combine all the energies and using the PEST optimization program we get the final set of parameters that fits well all the QM reference data. In Figure B.5 we have reported some dimer conformations of ethane, dodecane and DEPA on the side of the corresponding interaction energy curves as a function of the distance between the two molecules. The other curves are available in annex B, Figures B.4,B.5. The final sets of intermolecular parameters are also listed in annex B as Table B.2. The total procedure including charge and polarizability estimations as well as the determination of the other parameters allows to reproduce the QM energies, with a maximum deviation of $\simeq 0.6\,\text{kcal}\cdot\text{mol}^{-1}$.

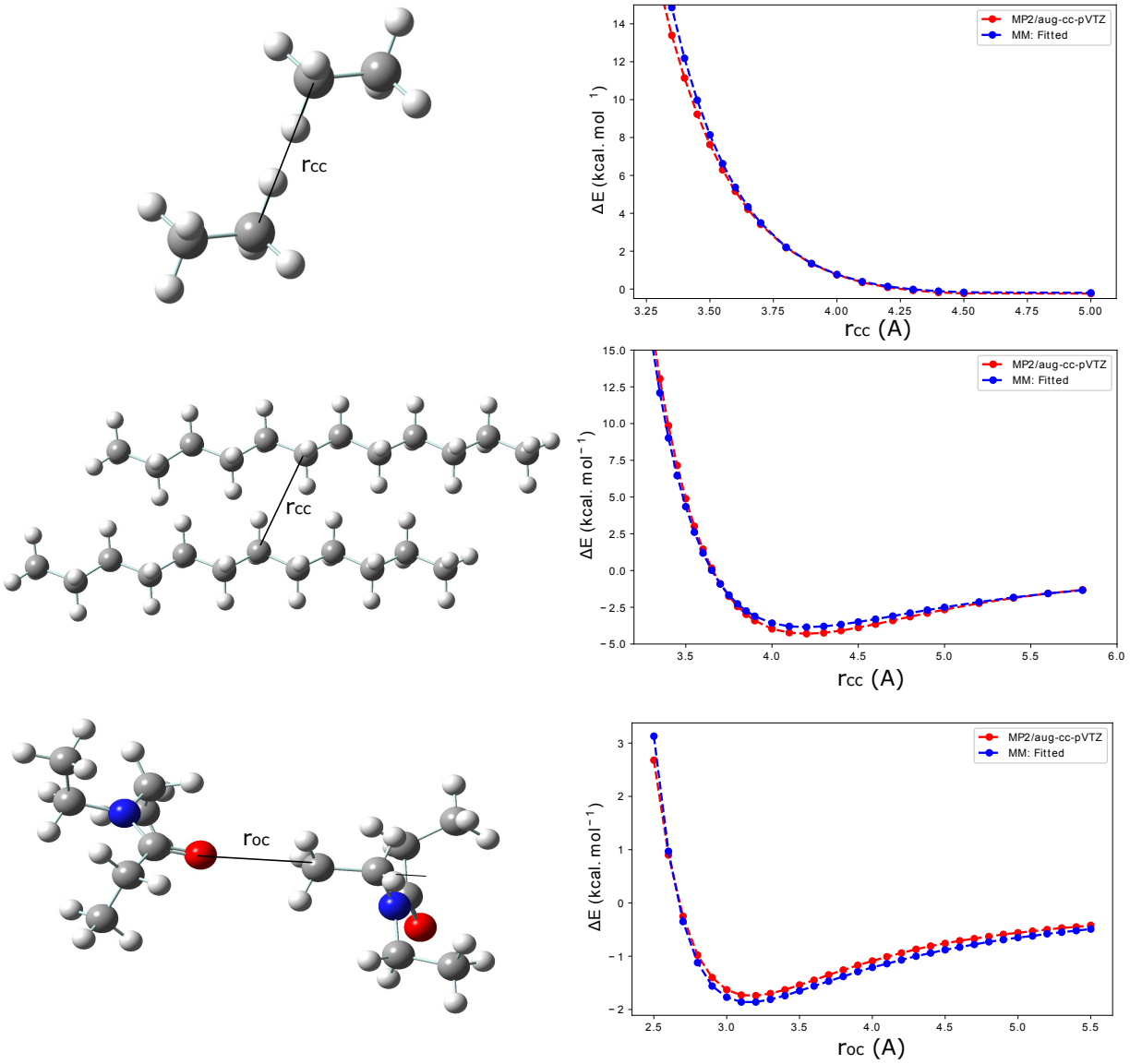


FIGURE 5.1: Dimer structures on the left with the corresponding interaction energy curves as a function of the inter-molecular distance for ethane (top), dodecane (middle) and DEPA (bottom). The MP2 *ab initio* energies are drawn in red and the FF values in blue (oxygen atoms in red, nitrogen in blue, carbon atoms in grey and hydrogens in white).

Torsional parameters

The torsion angle motions embed contributions from both the nonbonded (Van der Waals, electrostatic and 1-4 interactions) terms, as well as the angle bending. The torsional parameters are therefore intimately coupled to the nonbonded and bonded parameters. For this reason, the dihedral parameters are chronologically the last ones to be derived during the development of a FFs. In this work, the parameters for dihedral angles are introduced with the standard type of torsion potential, the cosine expansion (See Eq. 4.10).

Usually, torsional parameters are derived to reproduce the energy differences of conformations and rotational profiles, based on high-level *ab initio* data. In this work, complete QM scans of the dihedral torsions were performed as a sequence of constrained optimization in which the torsion angle is varied in steps of 10°. For alkanes, all possible torsions CTx–CT2–CT2–CTx with $x = 2, 3$ were scanned and resulted in quite similar torsion *ab initio* energy profiles (difference between maxima $\leq 0.2\,\text{kcal} \cdot \text{mol}^{-1}$ as shown in Figure 5.2). For most available FFs in literature, the same torsional parameters are used for both torsion involving terminal methyl and middle torsion. In this work, since we differentiate the non-bonded parameters of CT3 from those of CT2, and since the dihedral parameters do dependent on the non-bonded ones, we have derived different parameters for each dihedral torsion (CT3–CT2–CT2–CT2 and CT2–CT2–CT2–CT2).

The coefficients were obtained by minimizing the difference of the relative total potential energies calculated from the force field and from *ab initio* calculations. The comparisons of torsional profiles obtained from *ab initio* and force field calculations are shown in Figure 5.2. The almost prefect superposition of the classical and QM curves makes us confident to explore further the different trans-gauche populations of the various species.

The same approach was followed for amides, the optimizations of the torsional parameters was carried out for the different dihedral angles.

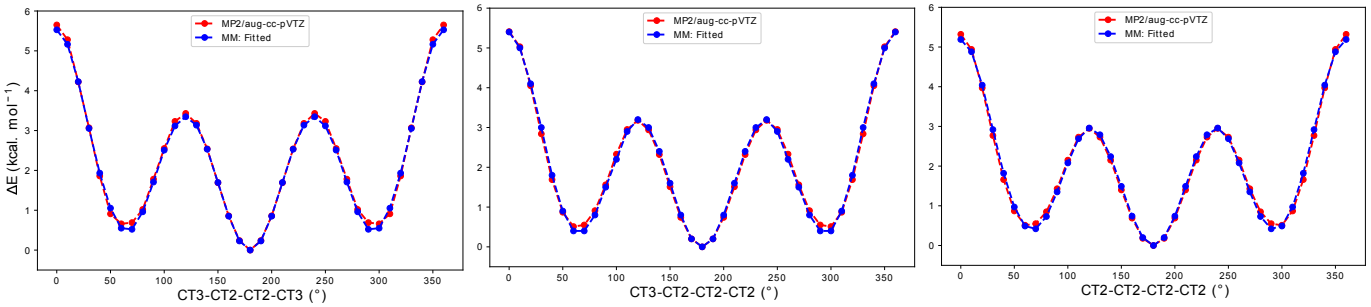


FIGURE 5.2: Relative energy profiles of the dihedral angle scan of butane (left) and heptane (middle and right).

To demonstrate the quality of our *ab initio* FFs, we will now turn to the computation of macroscopic properties from MD simulations.

5.3. Force-field evaluation: physical properties

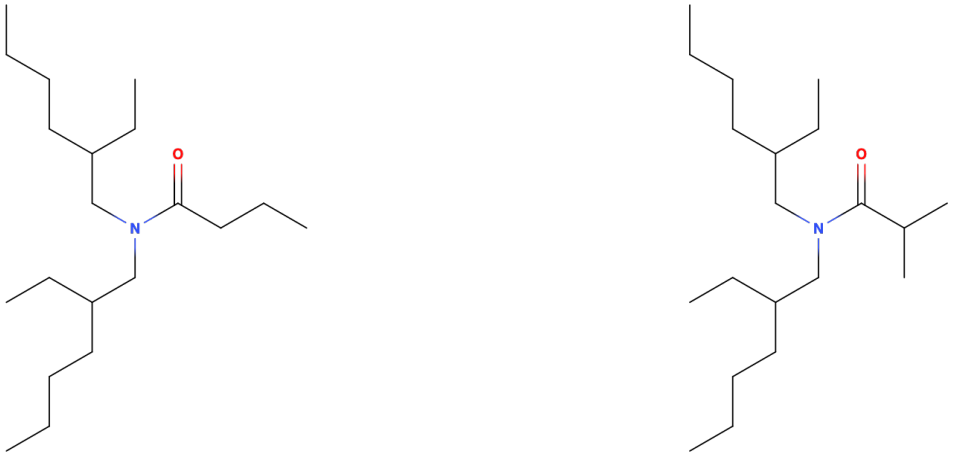
In this section we present the condensed phase properties obtained by performing MD simulations with the developed FFs. First, the densities and thermodynamic properties of alkanes are presented and compared to available experimental data as well as to the two version of OPLS FFs. Secondly, we will discuss their structural properties such as gauche-trans states ratios in alkanes and RDFs for both class of molecules, and lastly the self diffusion coefficients. The choice of the temperature at which we carried out our simulations is based on, i) the availability of experimental data at such conditions ii) to get as close as possible to the experimental conditions of liquid-liquid extraction processes (room temperature, 1 atm).

To compute these liquid properties, MD simulations were carried out considering periodic boundary conditions; first, in the isothermal-isobaric ensemble (NPT) to equilibrate the systems, and then the production part in the NVT ensemble. The temperature was maintained by a Nose-Hoover thermostat and the system pressure isotropically by an Andersen barostat. Verlet leapfrog scheme was employed with the integration time step of the dynamic equations of motion set to 1 fs. The equilibration period of our systems varies between 3 ns for pure phases (alkanes, monoamides) and 5 ns for alkanes/amides mixtures, with a temperature scaling interval each 10 steps. The production run is about 10 ns, which was enough considering the small size of the systems that we simulated. The C-H structural parameters were constrained to their bulk equilibrium values thanks to the RATTLE algorithm (the convergence criteria is set to 1×10^{-6} Å). The fast multipole method was used for computing Coulomb electrostatic forces and polarization interactions [171]. The Polaris-MD code [132] was used to carry out all MD simulations in this work.

The details about the method of calculation of the physical properties, such as density, heat of vaporization and self-diffusion coefficient have been given in a previous chapter in Section 4.5.

5.3.1 Density and heat of vaporization

We have computed the density and the enthalpy of vaporization of the series of alkanes (ethane, propane, butane, heptane, dodecane) and two monoamides DEHiBA and DEHBA (Figure 5.3) and the primary amide N,N-Dimethylacetamide (DMA) . The results are reported in Table 5.4. The bulk density values agree well with the experimental ones, as the largest deviation is 5% for dodecane and 6% for DEHiBA. This confirms



(a) (b)

FIGURE 5.3: Structure of (a) di-2-ethylhexyl-butyramide (DEHBA) and (b) di-2-ethylhexylisobutyramide (DEHiBA).

the correct description of the phase equilibria, the position of the minima of potential wells. For the heat of vaporization, we can solely consider alkanes in the comparison to experimental data. The largest deviation is less than 1 kcal $\cdot$ mol^{-1} for dodecane. For the simulated monoamides no experimental heat of vaporization is available in the literature, but since the same parametrization approach was used for both systems, we can safely predict them. To our knowledge, we are the first group to report the value of the heat of vaporization of DEHiBA and DEHBA via computational methods. The heats of vaporization of DEHiBA and DEHBA seem to be quite close for both molecules ($\simeq$ 23.5 kcal $\cdot$ mol^{-1} for DEHiBA and $\simeq$ 25 kcal $\cdot$ mol^{-1} for DEHBA), this was expected since both molecules are quite similar as they only differ by the branching of the alkyl chain attached to the carbonyl group (butyl vs. iso-butyl). Overall, the computed properties yield a great agreement with experiment, and the errors are within the error bars of alkanes traditional FF developed based on density and heat of vaporization, about 5% for the densities and up to 1 kcal $\cdot$ mol^{-1} for the heats of vaporization [128, 140, 152, 154].

Some FFs may give very accurate densities and vaporization heats for some molecules, compared to our work. This is directly related to the fact that these molecules were used in the parameterization process or being

TABLE 5.4: Densities (in $kg \cdot m^{-3}$) and heats of vaporization (in $kcal \cdot mol^{-1}$) of alkanes and monoamides (DEHiBA and DEHBA) using the parameters derived in this work. Δ values correspond to either the relative errors in % or absolute errors in $kcal \cdot mol^{-1}$ between the simulated and experimental values.[a] Experimental values are from Haynes et al. [172]. Error estimates were obtained by block averaging (see annex B for further details).

Name	T(K)	P(atm)	Density			Heat of Vaporization		
			This work	expt[a]	$\Delta(\%)$	expt[a]	This work (P1)	$\Delta(kcal \cdot mol^{-1})$
Ethane	185	1	566	544	4%	3.60	3.65 ±0.06	0.05
Propane	225	1	604	587	3%	4.40	3.57 ±0.07	0.83
Butane	273	1	596	601	1%	5.36	4.66 ±0.10	0.70
Heptane	298	1	700	677	3%	8.60	7.69 ± 0.10	0.91
	371	1	637	616	3%	7.60	7.84±0.05	0.24
Decane	298	1	761	730	4%	12.30	11.90 ± 0.10	0.40
Dodecane	298	1	784	745	5%	14.70	15.43 ± 0.19	0.72
	490	1	604	590	3%	10.54	11.51 ± 0.20	0.97
Isopentane	301	1	648	617	5%	5.88	6.32 ±0.04	0.44
DMA	298	1	937	900	4%	10.90	10.21 ±0.11	0.69
DEHiBA	298	1	898	865	4%	-	23.34 ±0.15	-
	308	1	892	858	4%	-	23.13 ±0.15	-
	318	1	883	851	4%	-	21.46 ±0.15	-
DEHBA	298	1	916	861	6%	-	24.93 ± 0.21	-
	308	1	905	854	6%	-	24.44 ± 0.21	-
	318	1	895	847	6%	-	22.56 ± 0.21	-

very close to them. However, as already mentioned, these FFs are not always transferable for all molecular groups. For instance, the dodecane heat of vaporization obtained with OPLS FF is about $\simeq 22.3\,kcal \cdot mol^{-1}$ with a deviation of $\simeq 7.6\,kcal \cdot mol^{-1}$ from the experiment, and a density of $839\,kg \cdot m^{-3}$ overestimating the experimental value of $745\,kg \cdot m^{-3}$. These deviations pair with the fact that the OPLS force field was optimized to reproduce liquid densities and enthalpies of vaporization of short alkanes (ethane, propane, and butane). As a result, Siu et al. have re-optimized these parameters for longer alkanes (labeled L-OPLS [154]), following the same parameterization approach of the original OPLS (based on experimental densities and vaporization heats). Still, this may lead to the same disadvantages for much longer molecules than the ones used in the parameterization. It should also be mentioned that, even though cis-9 octadecene molecule was used in the parameterization process of L-OPLS, a deviation of $2\,kcal \cdot mol^{-1}$ from reference data was accepted which is more than the maximum deviation in our work ($1\,kcal \cdot mol^{-1}$). All of this confirms the strength and reliability of our parametrization approach and the developed FFs in simulating/predicting

thermodynamic properties of alkanes and monoamides.

5.3.2 Structural properties

Chain conformation: the gauche and trans states in alkanes:

The chain conformation is also evaluated by analyzing the population of the gauche and trans states in series of alkanes (heptane, decane and dodecane), which is characterized by the ratio **D** defined as $\frac{\text{population of trans}}{\text{population (gauche+trans)}}$ for each specific torsion angle (see Figure 5.5). This kind of analysis serves to determine the preferred conformations in the liquid phases, and among the fundamental structural units in biology and chemistry in general [173–175].

To illustrate these gauche (G) and trans (T) states in alkanes, in Figure 5.4 we show the trans and gauche conformations for the butane molecule. These states (T and G) exist and are easily accessible in the liquid phase, as the energy difference between trans and gauche conformations in alkanes is small (about $0.6\,\text{kcal}\cdot\text{mol}^{-1}$) and the energy barrier between the two states is about $3.5\,\text{kcal}\cdot\text{mol}^{-1}$, hence, it can be explored readily during the dynamics of the system.

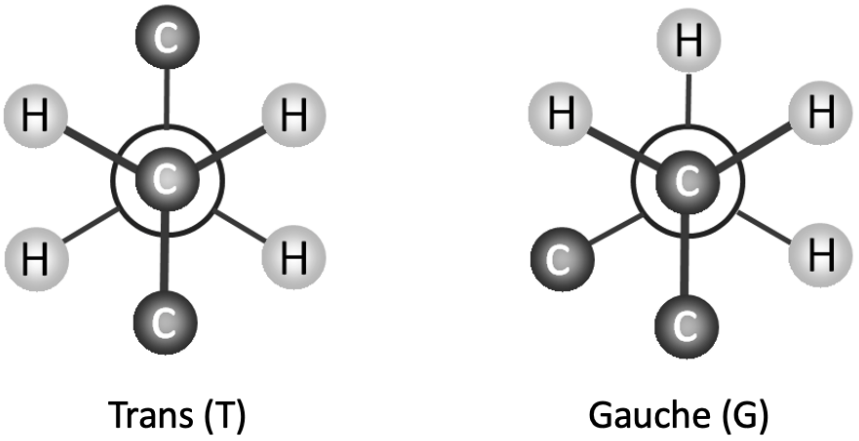


FIGURE 5.4: Newman projections of butane: Gauche and Trans conformations

In Table 5.5, we reported the obtained results from our FF, the available experimental data and some other FFs (OPLS-AA and L-OPLS). We can clearly observe the good agreement of our work with the experimental

data, as a difference of 6 *pp* (percentage point[1]) for the terminal torsion populations (D1 and D9) and 1 *pp* for the torsion in the middle of the chain (Dx with x$\in [3..7]$) population was noted for dodecane. The largest deviation was observed for heptane, about 7 *pp* for the terminal torsion population and 1 *pp* for the middle torsion ones.

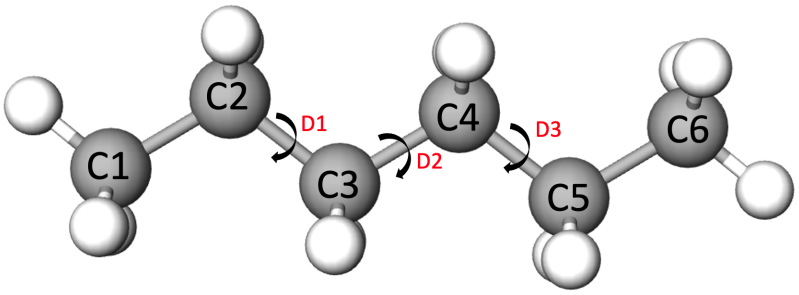


FIGURE 5.5: Representation of dihedral population notations for hexane molecule.

One may notice that we are comparing our results to the experimental values for tridecane ($C_{13}H_{28}$), this is because $i)$ it is the only available experimental values in the literature, $ii)$ the simulated alkanes are quite close to tridecane, especially, decane ($C_{10}H_{22}$) and n-dodecane ($C_{12}H_{26}$).

It is also worth noticing that for the literature results (OPLS FFs and experimental), the trans fraction is clearly dependent on the position within the carbon chain; the interior of the chain is almost constant while the chain terminals have a larger flexibility, mainly because of the influence of intermolecular interactions within the molecule. In our work, this dependence on the position within the carbon chain is quite small (about 3 *pp*). The results obtained with the OPLS-AA FF for dodecane shows a divergence of about 22 *pp* with experimental ones for the chain terminal populations and 13 *pp* for the trans populations of the interior of the chain. This again confirms the inability of OPLS-AA to correctly describe the trans-gauche populations for alkanes. The more recent L-OPLS FF results are in quite good agreement with the experimental data reported by Casal et al., with a deviation of about 3 pp, and may be considered as good as with our FF (maximum deviation 6 pp) for the simulated alkanes.

Lastly, the ability of our FF in investigating/predicting trans-gauche states have been demonstrated over the series of investigated alkanes.

[1]percent point is the unit for the arithmetic difference of two percentages. For example, moving up from 40% to 44% is a 4 percentage point increase.

TABLE 5.5: Trans populations of heptane, decane, dodecane and tridecane as a function of dihedral angle noted Dn along the carbon chain (in %). [a] Experimental values for the first and the sixth carbon torsion of tridecane from Casal et al. [176], [b] values from MD simulations with L-OPLS [154] and [c] values from Monte Carlo (MC) simulations [177] with the original OPLS-AA FF [152].

Molecule	Dihedral angle	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
heptane (298 K)	This work	65	68	68	65						
	OPLS-AA [c]	73	82	78	76						
decane (300 K)	This work	65	69	66	66	66	69	65			
	OPLS-AA [c]	78	81	84	82	81	81	79			
dodecane (298 K)	This work	64	68	65	65	65	65	65	68	64	
	OPLS-AA [c]	80	83	79	80	78	80	81	83	81	
dodecane (490 K)	This work	56	61	57	57	57	57	57	61	56	
tridecane (298 K)	L-OPLS[b]	55	66	64	64	64	64	64	64	66	55
	exp[a]	58	-	-	-	-	67	-	-	-	58

Radial distribution function analysis: The structural organization of the pure alkanes and monoamides (DEHiBA and DEHBA) solvents was investigated by calculating the radial distribution functions (RDFs, $g(r)$), as they reveal the distribution of neighboring molecules and the long-range solvent organization. The $g(r)$ for carbon atoms in alkanes are plotted in Figure 5.6.

	C_1-C_2		C_1-C_3		gC_1-C_4		tC_1-C_4	
	This work	expt	This work	expt	This work	expt	This work	expt
Butane 273 K	1.54	1.55	2.55	2.56	3.10	3.07	3.88	3.93
Heptane 298 K	1.53	1.55	2.54	2.56	3.10	3.14	3.87	3.92
Decane 298 K	1.53	1.54	2.55	2.57	3.10	3.15	3.88	3.95
Dodecane 298 K	1.53	-	2.55	-	3.10	-	3.88	-

TABLE 5.6: The molecular parameters (Maxima of the RDFs) in Å obtained from RDFs opposite to experimental work of Habenschuss et al. [178].

The RDF formula takes into account the system volume and the total number of particles (atoms), i.e., the density, as well as the number of available atoms of the type for which the RDF is calculated. Hence, considering the difference in the system densities, the RDF gives insights on the distances at which the atomic pairs are found. The absolute RDF values for the different systems are not comparable/interpretable due to the different available volume.

We have drawn the C–C RDFs starting at 2 Å, hence, the first C-C direct interaction at 1.53 Å is not visible.

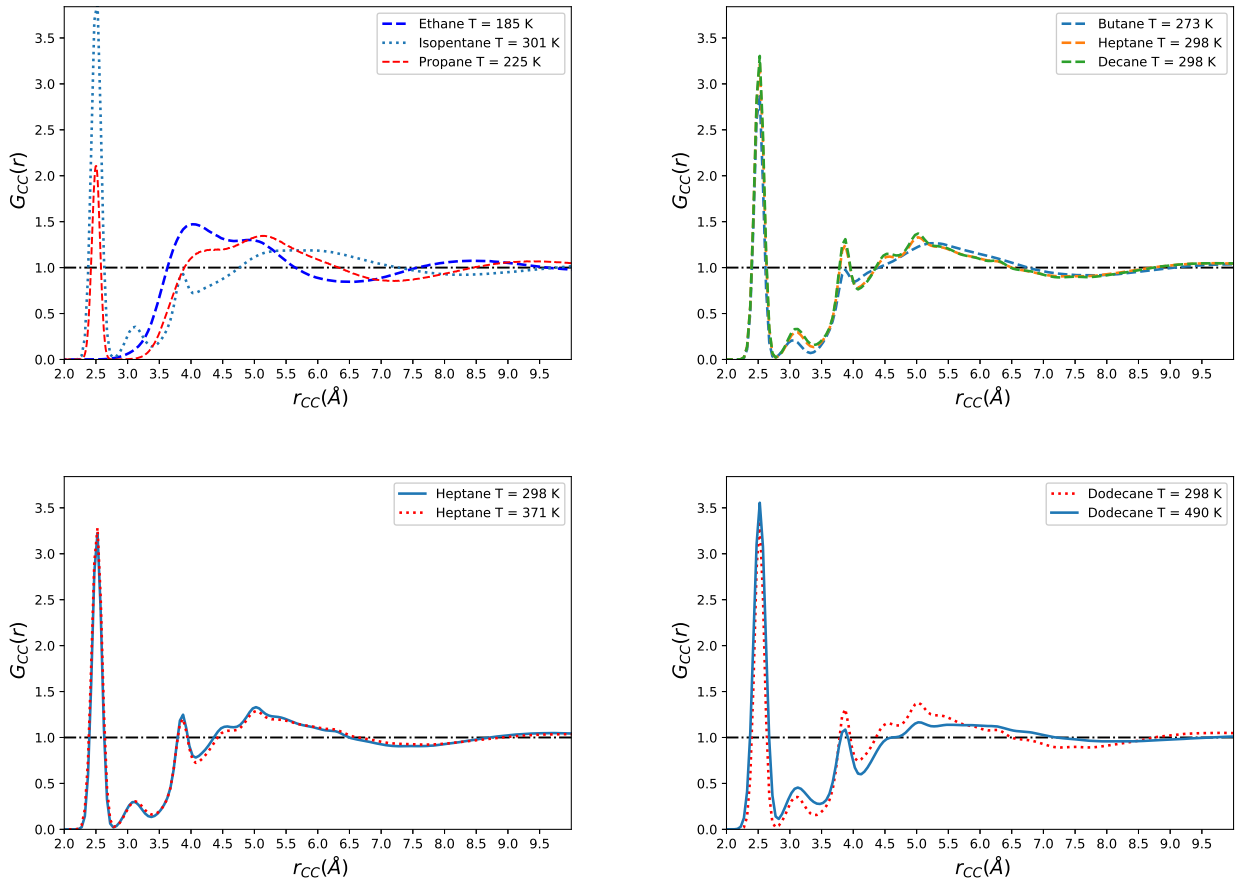


FIGURE 5.6: Radial distribution functions of carbon atoms for selected alkanes: ethane, propane, butane, isopentane, and heptane, decane and dodecane.

The peak at 2.55 Å corresponding to the intramolecular $C_1 - C_3$ second-neighbor interaction is present for all alkanes, except, of course for the ethane molecule (C_2H_6). The third peak at $\simeq$3.1 Å corresponds to the gauche $C_1 - C_4$ molecular segments. At the distance of $\simeq$ 3.88 Å appears a peak that can be attributed to the trans $C_1 - C_4$ molecular segments. These results are quite consistent with the experimental X-ray diffraction results of Habenschuss et al. [178].

We can also notice the influence of temperature upon the RDFs, for dodecane at two different temperature (298 K and 490 K). It is readily apparent in Figure 5.6 (d), that the intensity of the 3rd peak increases while the 4th peak simultaneously decreases, implying that the gauche population increases with temperature and the reverse for the trans population. This was confirmed by the calculation of the populations as presented in Table 5.5. This is simply because by increasing the temperature we are more likely to populate the state with higher energy (gauche).

For both investigated monoamides (DEHBA,DEHiBA), no structural experimental data is available to compare with, but since we have used the same parametrization approach as for alkanes, we can safely predict them. The RDFs at various temperatures (298 K,308 K and 318 K) are reported in Figures 5.7 and B.7. The analysis reveals that RDF(O−O), RDF(O−N) and RDF(N−N) are very similar for both monoamides, meaning that the structuring is quite similar for both ligands. From Figure 5.7, we can also notice the impact of the temperature on the RDFs of DEHBA and DEHiBA, the intensities for all RDFs in DEHiBA and DEHBA remain constant as the temperature increases, the influence of temperature is minor for this interval [298 K ; 318 K].

5.4. Molecular properties of dodecane/monoamides mixtures

After the validation of our FF in the previous section, herein we present the results obtained for DEHBA and DEHiBA mixtures with dodecane. This study was performed to gain insight and to visualize molecular-level behavior of these extraction systems (alkane-monoamides mixtures), for instance, to understand the monoamide structure impact on the molecular organization of alkane-monoamides mixtures, more precisely the branching of the alkyl chain bonded to the carbonyl function. At the end, this may give insights about the sharp difference of the Pu(IV) extraction between DEHiBA and DEHBA, since these mixtures are the ones used for the extraction of actinides cation from nuclear fuel waste.

In order to approach as much as possible experimental conditions, the simulations boxes where constructed based on available experimental data for DEHiBA (densities) at room temperature and atmospheric pressure (298 K,1 atm). Since the experimental data were only available for DEHiBA, we chose the same simulation conditions for the DEHBA/dodecane mixtures to be able to compare the solvent mixtures, as a function of the nature of the monoamide.

The densities of different DEHiBA mixtures were calculated using the NPT ensemble and compared favorably, as showed by the small error in Table 5.7 (a maximum deviation of 5% ($28\,\mathrm{kg}\cdot\mathrm{m}^{-3}$)). The densities of DEHBA/dodecane mixture appear to be quite similar to the DEHiBA/dodecane ones; this was again expected since the structures of the two ligands are quite close (See Figure 5.3).

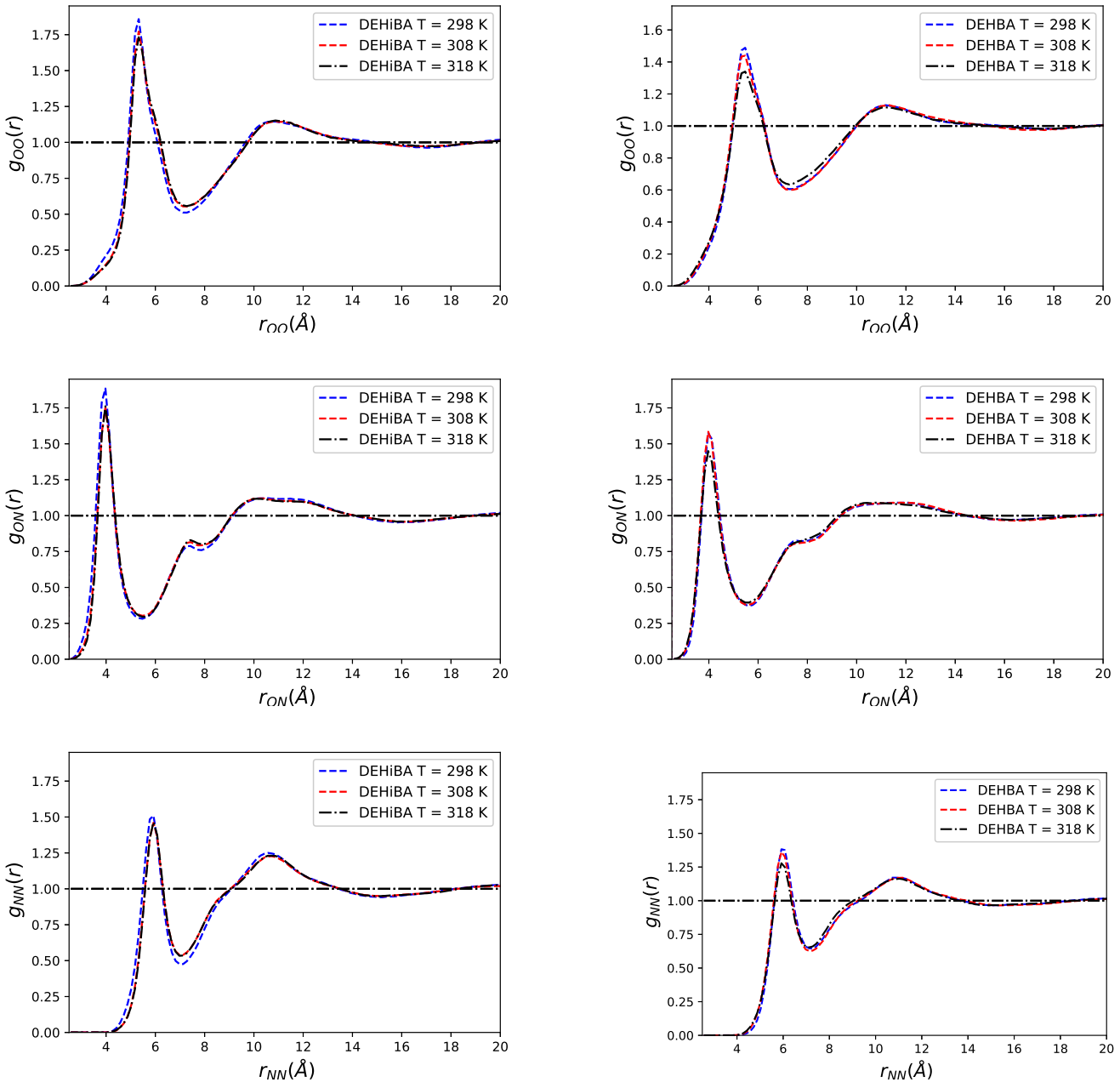
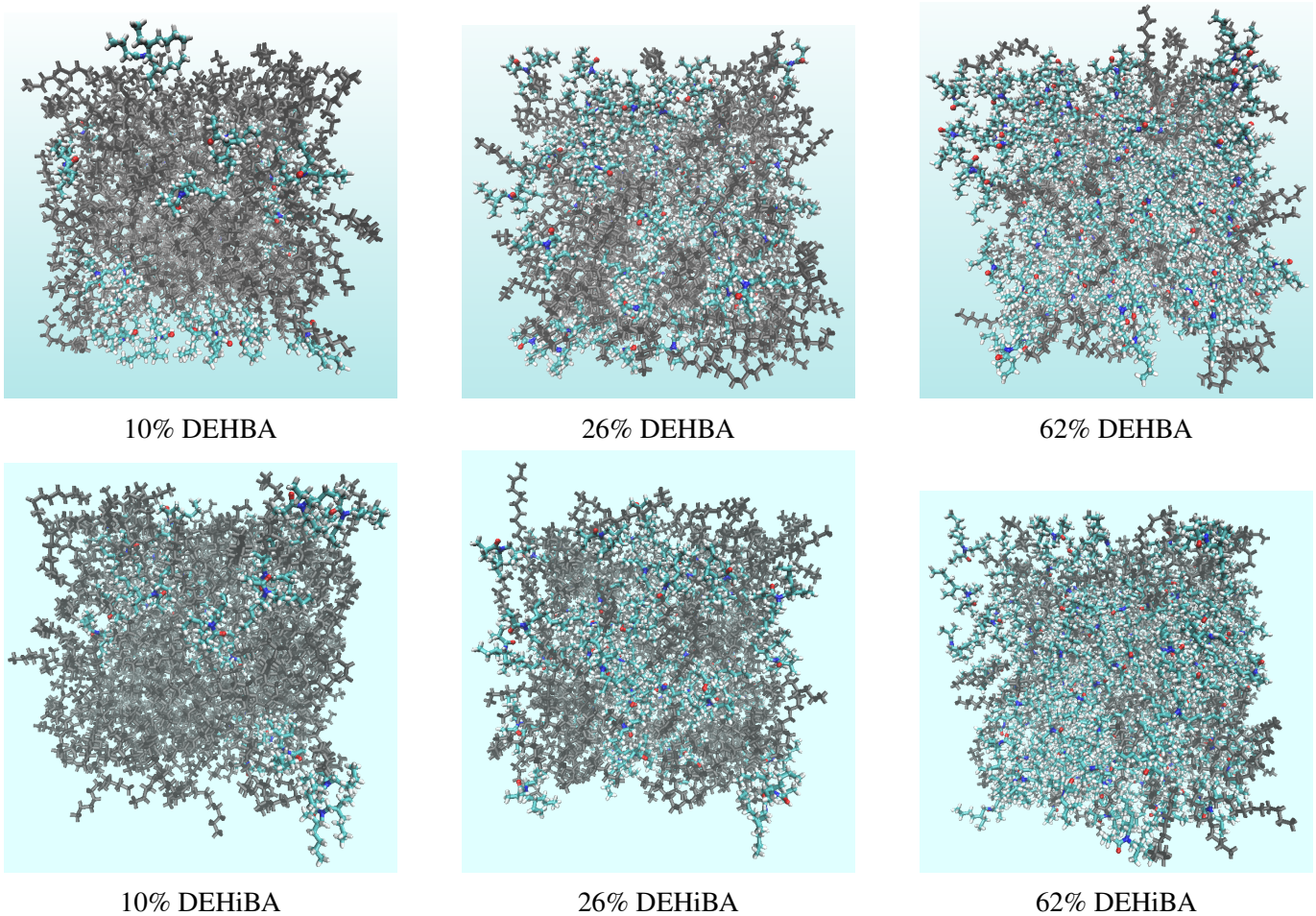


FIGURE 5.7: Radial distribution functions of oxygen, nitrogen atoms in DEHiBA (left figures) and DEHBA (right figures) at 298, 308 and 318 K.

From Figure 5.8, it can be seen that the molecular organization for both systems (DEHiBA/dodecane and DEHBA/dodecane) seem to be quite similar, the DEHiBA and DEHBA molecules form some kind of aggregates, as the concentration of DEHiBA/DEHBA increases we can also notice some sort of aggregation for dodecane. Somehow, this may be explained by taking in consideration that monoamides molecules

TABLE 5.7: Composition of DEHiBA/dodecane and DEHBA/dodecane mixtures, associated simulation boxes and densities, L denote the extractant ligands.

Monoamide Concentration (mol/L)				Number of molecules			density			
expt		MD					DEHiBA		DEHBA	
DEHiBA	DEHBA	DEHiBA	DEHBA	L	dodecane	% L	expt	MD	expt	MD
0.40	-	0.42	0.43	34	316	10%	765	792	-	795
1.00	-	1.04	1.07	92	258	26%	790	815	-	830
2.00	-	2.07	2.03	217	133	62%	832	866	-	859



10% DEHBA 26% DEHBA 62% DEHBA

10% DEHiBA 26% DEHiBA 62% DEHiBA

FIGURE 5.8: Final snapshots of the DEHBA/dodecane (top) and DEHiBA/dodecane (bottom) systems at the end of the MD trajectories, at 298 K for different concentrations 10% (34 Ligands and 316 dodecane), 26% (92 Ligands and 258 dodecane) and 62% (217 Ligands and 133 dodecane). In black dodecane molecules and monoamides ligands are colored, oxygen atoms in red, nitrogen in blue, carbons in aqua and hydrogen atoms in white.

are polar, and prefer interacting with each other rather than alkanes. In general, alkanes are known to be "insoluble" in polar solvent (such as water). They are soluble only in non-polar and slightly polar solvents.

In our case, DEHiBA and DEHBA are polar molecules with a dipole moment of $\simeq$3.64 debye[2] (water : $\simeq$ 1.85 debye), yet, the fact that they are long molecules with lipophilic tails makes them soluble in alkanes.

To gain insights into the structures of DEHiBA/DEHBA in the liquid phase as well as the organization of DEHiBA/DEHBA within the dodecane solvent, the RDFs between atoms belonging to different molecules were computed using a bin width of 0.10 Å and a cutoff distance of 25 Å (about half of the simulation box length), averaged over the last 8 ns of the MD trajectories. The RDFs of DEHiBA/DEHBA at different concentrations are reported in Figures B.6 and B.8. We can clearly see that for both systems, the RDF(O−N), RDF(O−C) and RDF(N−N) are very similar for both monoamides, with a small difference for the RDF(O−O), as for pure phase of DEHiBA and DEHBA. Even when the concentration of the amide ligand increases, the positions of the peaks do not vary.

5.5. Conclusions

In this work, a FF model for short- and long- alkane chains and amides derivatives was constructed by solely considering quantum chemical calculations (at the MP2 level of theory) and taking explicitly into account polarization effects, which are essential for a good representation of the dispersion forces and polarization interactions. The different intra-molecular and inter-molecular parameters, such as atomic charges, atomic polarizabilities, Buckingham parameters, force constants and equilibrium values, were derived and validated by performing MD simulations to calculate liquid-phase thermodynamics and structural properties. The reported simulation results are in great agreement with experimental data. These high-quality results for both molecular families suggest that this purely *ab initio* parametrization methodology is promising and can be applied to any molecules of interest.

The developed FFs were also used to simulate monoamides-dodecane mixtures (DEHiBA/dodecane and DEHBA/dodecane), revealing that for both mixtures, amides ligands tends to self-assemble in the organic solution. The RDFs calculations showed that the molecular organization for both systems is quite similar. Further investigations for a better characterization of the molecular organization and aggregation of these phases are in progress,namely the orientation of the carbonyl group (C=O) is under investigation. Finally,

[2]Obtained with quantum calculations at the MP2 level and aug-cc-pVTZ basis set.

the calculation of Small-angle X-ray scattering (SAXS) spectra will give us the opportunity for a direct comparison with experimental data.

Conclusions and perspectives

The work presented in this thesis is directly linked to the research work carried out for the development of new technologies for spent nuclear fuel reprocessing, namely the liquid-liquid extraction techniques. The studied extracting molecules, monoamides, are being considered for use in future nuclear fuel recycling processes. The study of solutions representative of the organic phases used in these processes enables a better understanding of the organization of these solutions and the effect of the structuration/aggregation on their physico-chemical properties. The main objective of this thesis is the development and use of theoretical approaches to provide information on the organization at the molecular level of the organic phase containing plutonium in the presence of the different other molecules (alkanes, water, extractants, and nitric acid), namely quantum chemistry and molecular dynamics simulations methods.

To reach this goal, in the first part, we have studied the stability of plutonium nitrates complexes in the organic phase using quantum chemistry methods (DFT, MP2). This study allowed us to get a better understanding of the strong influence of the amide structure on plutonium extraction at the atomistic level. The calculated complexation energies for inner and outer-sphere complexations showed that the introduction of a bulky alkyl group on the carbonyl side (branching) leads to an increase in the complexation energy and therefore the complexation is weakened, which is fully consistent with the measured monoamide extraction properties (distribution coefficient). Moreover, changing the polarity of the solution (dodecane to DMA) also showed

that solvent effects are very significant for outer-sphere complexation and rather small for inner-sphere complexation.

In our systems of interest, the extractants (monoamides) contain long alkyl chains, which generate numerous degrees of freedom and the possibility to interact further with solvent molecules. Thus, being aware of it and wishing to go beyond the static picture proposed by DFT quantum chemical calculations, we proposed to use classical molecular dynamics to perform explicit solvent simulations. However, as there are no force fields to describe plutonium(IV), the parametrization of new FF models is required. Nevertheless, being able to propose a complete picture requires a step-by-step development of each sophisticated possible interaction potential.

Thus, we have first concentrated our efforts on the proper description of the physico-chemical properties of the solvent molecules, i.e., the alkanes and the monoamides, and their mixtures. In the second part of this thesis, a polarizable FF model for the solvent molecules (alkanes + monoamides) was constructed. Within the aim of developing a new generation of accurate and robust FF models based on ab initio data, the derivation of the different FF parameters is based solely on quantum chemical calculations. Classical simulations were carried out with the developed potentials to calculate macroscopic thermodynamics and structural properties, such as density, heat of vaporization, and trans-gauche populations. The MD results obtained were in good agreement with experimental ones which allowed us to validate the parametrization methodology. The transferability of the developed potentials was demonstrated by calculating physical properties for different molecules of different sizes, ethane to dodecane, DMA to DEHiBA. We should mention that the calculation of other properties such as the diffusion coefficient and the viscosity are also in progress.

After the validation of the FF models, MD simulations of monoamides-dodecane mixtures (DEHiBA/dodecane and DEHBA/dodecane) were carried out, first to check the ability of our potentials to model correctly these mixtures, second to investigate the influence of the branching of the monoamide alkyl chain bonded to the carbonyl group on the structuring of the phase mixtures. The experimental bulk density of the mixtures fits well with the experimental one. The radial distribution functions calculated for both systems showed that for both investigated extractants, DEHiBA and DEHBA, the structuration seems to be quite similar. Furthermore, the analysis of these phases shows that aggregation phenomena may be suspected and additional

investigations for a better characterization of these phases are underway. Finally, the force-field potentials developed in this work are not limited to our study, they can also be used to study other properties of long alkanes or lipids in other fields, such as biology and petrology.

As already mentioned, the main objective is to be able to simulate the Pu(IV) in the organic/aqueous phase. However, the route to modeling the Pu(IV) extraction mechanisms; the migration of such an ion from one phase to another (from aqueous to organic phase); using classical MD is still long, mainly because the parameterization of FF models, to describe all the interactions presented in Figure 1.2, is a laborious and time-consuming task. Till now, FFs models for water (Réal et al. [118]), water with Pu(IV) (Acher et al. [48]), alkanes, monoamides and alkanes with monoamides have been developed based solely on QM calculations. This has shown that the development of potentials based such kind of data (QM calculations) is quite efficient and promising. Afterwards, it would be interesting to continue the description of the other interactions listed in Figure 1.2.

At the short term, the next logical step is the derivation of parameters describing the interactions between alkanes and water. Since FFs models for water and alkanes are now developed, electrostatic and polarization parameters are already available and fixed (charges, atomic polarizabilities), the only parameters to be adjusted are the Van der Waals ones (Buckingham parameters). The QM reference data (interaction energies) needed for the parameterization are being obtained and the derivation of the Buckingham parameters is in progress. The experimental data for the validation part are available, such as solvation free energies [179]. For monoamides with water, where partial charges and atomic polarizabilities are also fixed, the derivation of the potential would be a bit strenuous compared to that of alkanes-water. This is essentially because of the necessity of taking into account the hydrogen bond interactions that are highly important for describing correctly the interactions between the two molecules. Accordingly, describing the interaction between alkanes, monoamides and water will give us the possibility in modelling the extraction of water by monoamides ligands. These studies will be worthwhile in obtaining insights about the molecular organization of such phases, as well as illustrating the impact of structure and concentration of monoamides on the extraction.

The development of FFs to describe the interactions between the preceding molecules and nitric acid would be the next step, these interactions, of course, will contain specific terms describing hydrogen bond interaction;

which will make the parametrization of such interactions a bit tricky especially for water with nitric acid. The difficulty is related to the quantization of each contribution, in particular attaining the right balance between polarization, dispersion and hydrogen bond interactions. However, deriving these potentials to model the extraction of water and nitric acid occurring in solvent extraction techniques will be based on the neglect of the nitric acid protonation (deprotonation) that current models cannot take it into account.

The most arduous and expensive part will be is the development of force fields to describe the interactions of Pu(IV) with the rest of the molecules. These potentials will take into account special terms, such as charge transfer, hydrogen bonds, and polarization; and the parameterization of these terms is expensive and strenuous, in particular the treatment of N-body effects, whether they are taken into account explicitly or not. Most of these laboriousness problems come largely from the high charge of Pu(4+) which induces polarization and charge transfer effects pushing our simple models to the limits of their relevance. All of the above make the ultimate goal of this work a long-term one and a lot of development work remains to be done.

To overcome the difficult and long-term process of parameterizing the Pu(IV) - nitrates/monoamide interactions. Meanwhile, it would be quite interesting to develop a FF models for Pu nitrate complexes (plutonium hexanitrate $Pu(NO_3)_6{}^{2-}$ and plutonium tetranitrate $Pu(NO_3)_4$) interacting with other molecules. This may be quite helpful in studying the molecular organization beyond the first coordination sphere and the study of aggregation phenomena as well as third phase formation during the extraction of Pu(IV). For the outer-sphere complex, we should take in consideration that monoamides interact with the plutonium hexanitrates via a hydrogen bond, where the proton is located between the oxygen of the nitrates and the oxygen of the ligand carbonyl function. From quantum chemistry calculations, we have shown that this proton tends to prefer monoamides (see chapter 3), yet in liquid state this proton may be more flexible and swing between nitrates and the ligand. Therefore, deriving a potential for the whole complex (hexanitrates + two protonated ligands $Pu(NO_3)_6(HL)_2$ with other molecules would be the best way to correctly model the outer-sphere complex, giving the green light to study the physico-chemical properties of the phase.

Last, the FF parameters developed here may be used to study other aspects and other phenomena in other fields. The FF models can be easily implemented in other MD simulations codes that supports the Buckingham potential and the induced point dipoles model for treating the polarization, such as Tinker [180].

References

[1] *Nuclear Power in France*. https://www.world-nuclear.org/information-library/country-profiles/countries-a-f/france.aspx. Accessed : 20 September 2020 (cit. on p. 1).

[2] *Electricity tariffs in France*. https://en.selectra.info/energy-france/guides/electricity/tariffs. Accessed : 20 September 2020 (cit. on p. 1).

[3] H. Mckay. "The PUREX process". In: *Science and technology of tribuytl phosphate*. 1990 (cit. on pp. 2, 4).

[4] Y. Yang et al. "REE Recovery from End-of-Life NdFeB Permanent Magnet Scrap: A Critical Review". In: *J. Sustain. Met.* 3.1 (Mar. 2017), pp. 122–149. DOI: 10.1007/s40831-016-0090-4 (cit. on p. 2).

[5] E. Abonneau, P. Baron, C. Berthon, L. Berthon, A. Beziat, I. Bisel, L. Bonin, E. Bosse, B. Boullis, J. Broudic, et al. *Treatment and recycling of spent nuclear fuel. Actinide partitioning-Application to waste management*. France: CEA and Editions du Moniteur, 2008 (cit. on p. 2).

[6] L. Berthon and M.-C. Charbonnel. "Radiolysis of solvents used in nuclear fuel reprocessing". In: *Ion exchange and solvent extraction*. CRC Press, 2009, pp. 443–527 (cit. on p. 2).

[7] J. Rydberg. "The reduction of plutonium by tetravalent uranium". In: *J. Inorg. Nucl. Chem.* 5.1 (1957), pp. 79–86. DOI: 10.1016/0022-1902(57)80084-0 (cit. on p. 2).

[8] J. E. Birkett, M. J. Carrott, O. D. Fox, C. J. Jones, C. J. Maher, C. V. Roube, R. J. Taylor, and D. A. Woodhead. "Recent developments in the Purex process for nuclear fuel reprocessing: Complexant based stripping for uranium/plutonium separation". In: *CHIMIA International Journal for Chemistry* 59.12 (2005), pp. 898–904 (cit. on p. 3).

[9] S. Tachimori, Y. Sasaki, and S.-i. Suzuki. "Modification of TODGA-n-dodecane solvent with a monoamide for high loading of lanthanides (III) and actinides (III)". In: *Solvent Extr. Ion Exch.* 20.6 (2002), pp. 687–699. DOI: 10.1081/SEI-120016073 (cit. on p. 3).

[10] R. Taylor et al. "The EURO-GANEX Process: Current Status of Flowsheet Development and Process Safety Studies". In: *Procedia Chemistry* 21 (Dec. 2016), pp. 524–529. DOI: 10.1016/j.proche.2016.10.073 (cit. on p. 3).

[11] S. Tachimori, Y. Sasaki, Y. Morita, and S.-i. Suzuki. "Recent progress of partitioning process in JAERI: development of amide-based ARTIST process". In: *Proceedings 7th OECD/NEA Information Exchange Meeting on Actinide and Fission Product Partitioning and Transmutation*. 2002 (cit. on p. 3).

[12] Z. Berk. *Food Process Engineering and Technology*. Elsevier, 2018 (cit. on p. 3).

[13] B. Boullis et al. "List of contributors". In: *Reprocessing and Recycling of Spent Nuclear Fuel*. Ed. by R. Taylor. Woodhead Publishing Series in Energy. Oxford: Woodhead Publishing, 2015, pp. xiii–xv. DOI: 10.1016/B978-1-78242-212-9.09992-5 (cit. on p. 3).

[14] R. Bush. "Recovery of platinum group metals from high level radioactive waste". In: *Platin Met Rev.* 35.4 (1991), pp. 202–208 (cit. on p. 4).

[15] "Solvent extraction studies with dinonyl naphthalene sulphonic acid: Separation of protactinium(V) from the other mixed metal ions". In: *J. Inorg. Nucl. Chem.* 37.9 (1975), pp. 1983–1987. DOI: 10.1016/0022-1902(75)80929-8 (cit. on p. 4).

[16] Q. Jia, Q. Shang, and W. Zhou. "Synergistic extraction of lanthanum (III) from a nitrate medium by mixtures of 1-phenyl-3-methyl-4-benzoyl-pyrazalone-5 and triisobutylphosphine sulfide". In: *Ind. Eng. Chem.* 43.21 (2004), pp. 6703–6707. DOI: 10.1021/ie049831u (cit. on p. 4).

[17] T. Siddall III. "USAEC Report DP - 54". In: *E. I. Du Pont de Nemours and Co. Alken* SC.3 (1961), pp. 342–351 (cit. on pp. 4, 5, 29).

[18] T. Siddall III. “Effects of structure of N,N-disubstituted amides on their extraction of actinide and zirconium nitrates and of nitric acid”. In: *J. Phys. Chem.* 14.64 (1960), pp. 1863–1866. DOI: 10.1021/j100841a014 (cit. on pp. 4, 5, 29, 30).

[19] C. Pohlandt and J. S. Fritz. “Extraction of metal ions from chloride solution with N, N-dioctylacetamide”. In: *Talanta* 26.5 (1979), pp. 395–399. DOI: 10.1016/0039-9140(79)80201-5 (cit. on pp. 5, 29).

[20] N. Condamines and C. Musikas. “The extraction by N.N.-dialkylamides. II. Extraction of actinide cations”. In: *Solvent Extr. Ion Exch.* 10.1 (1992), pp. 69–100. DOI: 10.1080/07366299208918093 (cit. on pp. 5, 29, 30, 36).

[21] E. Dukes and T. Siddall III. “Tetrabutyl urea as an extractant for nitric acid and some actinide nitrates”. In: *J. Inorg. Nucl. Chem.* 28.10 (1966), pp. 2307–2312. DOI: 10.1016/0022-1902(66)80121-5 (cit. on p. 5).

[22] B. Mahanty, A. S. Kanekar, S. A. Ansari, A. Bhattacharyya, and P. K. Mohapatra. “Separation of neptunium from actinides by monoamides: a solvent extraction study”. In: *Radiochimica Acta* 107.5 (2019), pp. 369–376. DOI: 10.1515/ract-2018-3074 (cit. on p. 5).

[23] C. Pohllandt and J. Fritz. “Separation of metal ions on a new amide resin”. In: *J. Chromatogr. A* 176.2 (1979), pp. 189–197. DOI: 10.1016/S0021-9673(00)85650-2 (cit. on pp. 5, 29).

[24] H. Narita, M. Tanaka, K. Morisaku, and T. Abe. “Extraction of gold(III) in hydrochloric acid solution using monoamide compounds”. In: *Hydrometallurgy* 81.3 (2006), pp. 153–158. DOI: 10.1016/j.hydromet.2005.10.008 (cit. on pp. 5, 29).

[25] E. D. Doidge et al. “Evaluation of Simple Amides in the Selective Recovery of Gold from Secondary Sources by Solvent Extraction”. In: *ACS Sustain. Chem. Eng.* 7.17 (Sept. 2019), pp. 15019–15029. DOI: 10.1021/acssuschemeng.9b03436 (cit. on pp. 5, 29, 31).

[26] Mahanty Bholanath, Kanekar Avinash S., Ansari Seraj A., Bhattacharyya Arunasis, and Mohapatra Prasanta K. “Separation of neptunium from actinides by monoamides: a solvent extraction study”. In: *Radiochimica Acta* 107.5 (2019), p. 369. DOI: 10.1515/ract-2018-3074 (cit. on pp. 5, 29).

[27] G. Milanole, E. Russello, C. Marie, M. Miguirditchian, and C. Sorel. "Dissymmetric N, N-dialkylamides used particularly for separating uranium (VI) from plutonium (IV), synthesis thereof and uses of same". In: *WO2018138441 (A1)* (2018) (cit. on pp. 5, 29).

[28] K. McCann, J. A. Drader, and J. C. Braley. "Comparing Branched versus Straight-chained Monoamide Extractants for Actinide Recovery". In: *Sep. Purif. Rev.* 47.1 (2018), pp. 49–65. DOI: 10.1080/15422119.2017.1321018 (cit. on p. 5).

[29] E. Acher, Y. Hacene Cherkaski, T. Dumas, C. Tamain, D. Guillaumont, N. Boubals, G. Javierre, C. Hennig, P. L. Solari, and M.-C. Charbonnel. "Structures of Plutonium(IV) and Uranium(VI) with N,N-Dialkyl Amides from Crystallography, X-ray Absorption Spectra, and Theoretical Calculations". In: *Inorg. Chem.* 55.11 (June 2016), pp. 5558–5569. DOI: 10.1021/acs.inorgchem.6b00592 (cit. on pp. 5–7, 30, 31, 35, 38).

[30] E. Acher, T. Dumas, C. Tamain, N. Boubals, P. L. Solari, and D. Guillaumont. "Inner to outer-sphere coordination of plutonium(iv) with N,N-dialkyl amide: influence of nitric acid". In: *Dalton Trans.* 46 (12 2017), pp. 3812–3815. DOI: 10.1039/C7DT00031F (cit. on pp. 5–7, 31).

[31] J. R. Kumar, J.-S. Kim, J.-Y. Lee, and H.-S. Yoon. "A Brief Review on Solvent Extraction of Uranium from Acidic Solutions". In: *Sep. Purif. Rev.* 40.2 (2011), pp. 77–125. DOI: 10.1080/15422119.2010.549760 (cit. on p. 5).

[32] C. Berger, C. Marie, D. Guillaumont, E. Zekri, and L. Berthon. "Extraction of Uranium(VI) and Plutonium(IV) with Tetra-Alkylcarbamides". In: *Solvent Extr. Ion Exch.* 37.2 (2019), pp. 111–125. DOI: 10.1080/07366299.2019.1630095 (cit. on pp. 5, 29, 31).

[33] Y. Meridiano. "Organisation des molecules extractantes de type diamide : lien avec les propriétés extractantes ?" Thèse de doctorat dirigée par Zemb, Thomas Physique. Physicochimie Paris 11 20092009PA112017. PhD thesis. 2009 (cit. on p. 5).

[34] V. Vallet, P. Macak, U. Wahlgren, and I. Grenthe. "Actinide chemistry in solution, quantum chemical methods and models". In: *Theor. Chem. Acc.* 115.2-3 (2006), pp. 145–160. DOI: 10.1007/s00214-005-0051-7 (cit. on p. 6).

[35] P. Söderlind, G. Kotliar, K. Haule, P. M. Oppeneer, and D. Guillaumont. "Computational modeling of actinide materials and complexes". In: *MRS bulletin* 35.11 (2010), pp. 883–888. DOI: 10.1557/mrs2010.715 (cit. on p. 6).

[36] A. Kovács, R. J. Konings, J. K. Gibson, I. Infante, and L. Gagliardi. "Quantum chemical calculations and experimental investigations of molecular actinide oxides". In: *Chem. Rev.* 115.4 (2015), pp. 1725–1759. DOI: 10.1021/cr500426s (cit. on p. 6).

[37] M. Pepper and B. E. Bursten. "The electronic structure of actinide-containing molecules: a challenge to applied quantum chemistry". In: *Chem. Rev.* 91.5 (1991), pp. 719–741. DOI: 10.1021/cr00005a005 (cit. on p. 6).

[38] S. O. Odoh and G. Schreckenbach. "Theoretical study of the structural properties of plutonium (IV) and (VI) complexes". In: *J. Phys. Chem. A* 115.48 (2011), pp. 14110–14119. DOI: 10.1021/jp207556b (cit. on p. 6).

[39] A. Prestianni, L. Joubert, A. Chagnes, G. Cote, and C. Adamo. "A density functional theory study of uranium (VI) nitrate monoamide complexes". In: *Phys. Chem. Chem. Phys.* 13.43 (2011), pp. 19371–19377. DOI: 10.1039/C1CP22320H (cit. on p. 6).

[40] V. Barone and M. Cossi. "Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model". en. In: *J. Phys. Chem. A* 102.11 (Mar. 1998), pp. 1995–2001. DOI: 10.1021/jp9716997 (cit. on p. 7).

[41] M. P. Allen and D. J. Tildesley. *Computer simulation of liquids*. Oxford university press, 2017 (cit. on pp. 7, 48, 60, 65, 66, 73).

[42] D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications*. Vol. 1. Elsevier, 2001 (cit. on pp. 7, 65, 66, 68).

[43] R. Spezia, V. Migliorati, and P. D'Angelo. "On the development of polarizable and Lennard-Jones force fields to study hydration structure and dynamics of actinide (III) ions based on effective ionic radii". In: *J. Chem. Phys.* 147.16 (2017), p. 161707. DOI: 10.1063/1.4989969 (cit. on p. 8).

[44] V. A. Cocalia, K. E. Gutowski, and R. D. Rogers. "The coordination chemistry of actinides in ionic liquids: A review of experiment and simulation". In: *Coord. Chem. Rev.* 250.7-8 (2006), pp. 755–764. DOI: 10.1016/j.ccr.2005.09.019 (cit. on p. 8).

[45] N. Rai, S. P. Tiwari, and E. J. Maginn. "Force field development for actinyl ions via quantum mechanical calculations: an approach to account for many body solvation effects". In: *J. Phys. Chem. B* 116.35 (2012), pp. 10885–10897. DOI: 10.1021/jp3028275 (cit. on p. 8).

[46] K. A. Maerzke, G. S. Goff, W. H. Runde, W. F. Schneider, and E. J. Maginn. "Structure and dynamics of uranyl (VI) and plutonyl (VI) cations in ionic liquid/water mixtures via molecular dynamics simulations". In: *J. Phys. Chem. B* 117.37 (2013), pp. 10852–10868. DOI: 10.1021/jp405473b (cit. on p. 8).

[47] F. Réal, M. Trumm, V. Vallet, B. Schimmelpfennig, M. Masella, and J.-P. Flament. "Quantum Chemical and Molecular Dynamics Study of the Coordination of Th(IV) in Aqueous Solvent". In: *J. Phys. Chem. B* 114.48 (2010). PMID: 21070066, pp. 15913–15924. DOI: 10.1021/jp108061s (cit. on pp. 8, 56, 58).

[48] E. Acher, M. Masella, V. Vallet, and F. Réal. "Properties of the tetravalent actinide series in aqueous phase from a microscopic simulation self-consistent engine". In: *Phys. Chem. Chem. Phys.* 22.4 (2020), pp. 2343–2350. DOI: 10.1039/C9CP04912F (cit. on pp. 8, 99).

[49] R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel. "Electronic structure calculations on workstation computers: The program system turbomole". In: *Chem. Phys. Lett* 162.3 (1989), pp. 165–169. DOI: 10.1016/0009-2614(89)85118-8 (cit. on p. 11).

[50] M. J. Frisch et al. *Gaussian 16 Revision C.01*. Gaussian Inc. Wallingford CT. 2016 (cit. on pp. 11, 32).

[51] A. Szabo and N. S. Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012 (cit. on p. 12).

[52] M. J. Frisch, M. Head-Gordon, and J. A. Pople. "A direct MP2 gradient method". In: *Chem. Phys. Lett* 166.3 (1990), pp. 275–280. DOI: 10.1016/0009-2614(90)80029-D (cit. on pp. 18, 78).

[53] M. L. Leininger, W. D. Allen, H. F. Schaefer III, and C. D. Sherrill. "Is Mo/ller–Plesset perturbation theory a convergent ab initio method?" In: *J. Chem. Phys* 112.21 (2000), pp. 9213–9222. DOI: 10.1063/1.481764 (cit. on p. 19).

[54] P. Hohenberg and W. Kohn. "Inhomogeneous electron gas". In: *Phys. Rev.* 136.3B (1964), B864 (cit. on p. 20).

[55] W. Kohn and L. J. Sham. “Self-consistent equations including exchange and correlation effects”. In: *Phys. Rev.* 140.4A (1965), A1133. DOI: 10.1103/PhysRev.140.A1133 (cit. on p. 21).

[56] C. Fiolhais, F. Nogueira, and M. A. Marques. *A primer in density functional theory*. Vol. 620. Springer Science & Business Media, 2003 (cit. on p. 22).

[57] D. M. Ceperley and B. J. Alder. “Ground state of the electron gas by a stochastic method”. In: *Phys. Rev. Letters* 45.7 (1980), p. 566. DOI: 10.1103/PhysRevLett.45.566 (cit. on p. 22).

[58] S. H. Vosko, L. Wilk, and M. Nusair. “Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis”. In: *Can. J. Phys.* 58.8 (1980), pp. 1200–1211. DOI: 10.1139/p80-159 (cit. on p. 22).

[59] J. P. Perdew and Y. Wang. “Pair-distribution function and its coupling-constant average for the spin-polarized electron gas”. In: *Phys. Rev. B* 46.20 (1992), p. 12947. DOI: 10.1103/PhysRevB.46.12947 (cit. on p. 22).

[60] F. Jensen. *Introduction to computational chemistry*. John wiley & sons, 2017 (cit. on pp. 22, 56).

[61] A. D. Becke. “A new mixing of Hartree–Fock and local density-functional theories”. In: *J. Chem. Phys* 98.2 (1993), pp. 1372–1377. DOI: 10.1063/1.464304 (cit. on p. 23).

[62] C. Lee, W. Yang, and R. G. Parr. “Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density”. In: *Phys. Rev. B* 37.2 (1988), p. 785. DOI: 10.1103/PhysRevB.37.785 (cit. on p. 23).

[63] C. Adamo and V. Barone. “Toward reliable density functional methods without adjustable parameters: The PBE0 model”. In: *J. Chem. Phys* 110.13 (1999), pp. 6158–6170. DOI: 10.1063/1.478522 (cit. on pp. 23, 32).

[64] A. O. de la Roza and G. A. DiLabio. *Non-covalent Interactions in Quantum Chemistry and Physics: Theory and Applications*. Elsevier, 2017 (cit. on p. 23).

[65] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist. “Van der Waals density functional for general geometries”. In: *Physical review letters* 92.24 (2004), p. 246401. DOI: 10.1103/PhysRevLett.92.246401 (cit. on p. 23).

[66] O. A. Von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani. “Optimization of effective atom centered potentials for London dispersion forces in density functional theory”. In: *Physical*

review letters 93.15 (2004), p. 153004. DOI: 10.1103/PhysRevLett.93.153004 (cit. on p. 23).

[67] S. Grimme. "Accurate description of van der Waals complexes by density functional theory including empirical corrections". In: *J. Comput. Chem* 25.12 (2004), pp. 1463–1473. DOI: 10.1002/jcc.20078 (cit. on pp. 23, 24).

[68] S. Grimme. "Semiempirical GGA-type density functional constructed with a long-range dispersion correction". In: *J. Comput. Chem* 27.15 (2006), pp. 1787–1799. DOI: 10.1002/jcc.20495 (cit. on p. 24).

[69] J.-D. Chai and M. Head-Gordon. "Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections". In: *Phys. Chem. Chem. Phys.* 10.44 (2008), pp. 6615–6620. DOI: 10.1039/B810189B (cit. on p. 24).

[70] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu". In: *J. Chem. Phys* 132.15 (2010), p. 154104. DOI: 10.1063/1.3382344 (cit. on p. 24).

[71] R. Ditchfield, W. J. Hehre, and J. A. Pople. "Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules". In: *J. Chem. Phys* 54.2 (1971), pp. 724–728. DOI: 10.1063/1.1674902 (cit. on p. 25).

[72] J. Zheng, X. Xu, and D. G. Truhlar. "Minimally augmented Karlsruhe basis sets". In: *Theor. Chem. Acc.* 128.3 (2011), pp. 295–305. DOI: 10.1007/s00214-010-0846-z (cit. on p. 25).

[73] T. H. Dunning Jr. "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen". In: *J. Chem. Phys* 90.2 (1989), pp. 1007–1023. DOI: 10.1063/1.456153 (cit. on p. 25).

[74] R. A. Kendall, T. H. Dunning, and R. J. Harrison. "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions". In: *J. Chem. Phys* 96.9 (1992), pp. 6796–6806. DOI: 10.1063/1.462569 (cit. on pp. 25, 33, 78).

[75] V. Heine. "The pseudopotential concept". In: *Solid state physics*. Vol. 24. Elsevier, 1970, pp. 1–36 (cit. on p. 26).

[76] W. E. Pickett. “Pseudopotential methods in condensed matter applications”. In: *Comput. Phys. Rep.* 9.3 (1989), pp. 115–197 (cit. on p. 26).

[77] M. Dolg and X. Cao. “Relativistic Pseudopotentials: Their Development and Scope of Applications”. In: *Chem. Rev.* 112 (2012), pp. 403–480. DOI: 10.1021/cr2001383 (cit. on p. 26).

[78] M. Dolg. “Relativistic Effective Core Potentials”. In: *Handbook of Relativistic Quantum Chemistry*. Ed. by W. Liu. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 449–478. DOI: 10.1007/978-3-642-40766-6{_}5 (cit. on p. 26).

[79] P. Pyykkö. “Relativistic effects in chemistry: more common than you thought”. In: *Annu. Rev. Phys. Chem.* 63 (2012). DOI: 10.1146/annurev-physchem-032511-143755 (cit. on p. 26).

[80] A. Moritz, X. Cao, and M. Dolg. “Quasirelativistic energy-consistent 5f-in-core pseudopotentials for divalent and tetravalent actinide elements”. In: *Theor. Chem. Acc.* 118 (2007), pp. 845–854. DOI: 10.1007/s00214-007-0330-6 (cit. on p. 26).

[81] J. Tomasi, B. Mennucci, and R. Cammi. “Quantum mechanical continuum solvation models”. In: *Chem. Rev.* 105 (2005), pp. 2999–3094. DOI: 10.1021/cr9904009 (cit. on p. 27).

[82] G. Scalmani and M. J. Frisch. “Continuous surface charge polarizable continuum models of solvation. I. General formalism”. In: *J. Chem. Phys* 132.11 (2010), p. 114110. DOI: 10.1063/1.3359469 (cit. on p. 27).

[83] K. McCann, J. A. Drader, and J. C. Braley. “Comparing Branched versus Straight-chained Monoamide Extractants for Actinide Recovery”. In: *Sep. Purif. Rev.* 47.1 (2018), pp. 49–65. DOI: 10.1080/15422119.2017.1321018 (cit. on pp. 29, 30).

[84] Y. Ban, S. Hotoku, N. Tsutsui, A. Suzuki, Y. Tsubata, and T. Matsumura. “Uranium and Plutonium Extraction by N,N-dialkylamides Using Multistage Mixer-settler Extractors”. In: *Procedia Chemistry* 21 (2016). ATALANTE 2016 International Conference on Nuclear Chemistry for Sustainable Fuel Cycles, pp. 156–161. DOI: 10.1016/j.proche.2016.10.022 (cit. on pp. 30, 36, 75).

[85] Y. Ban, S. Hotoku, N. Tsutsui, A. Suzuki, Y. Tsubata, and T. Matsumura. “Uranium and Plutonium Extraction by N,N-dialkylamides Using Multistage Mixer-settler Extractors”. In: *Procedia Chemistry* 21 (2016). ATALANTE 2016 International Conference on Nuclear Chemistry for Sustainable Fuel Cycles, pp. 156–161. DOI: 10.1016/j.proche.2016.10.022 (cit. on pp. 30, 75).

[86] N. Condamines and C. Musikas. “THE EXTRACTION BY N.N-DIALKYLAMIDES. II. EXTRACTION OF ACTINIDE CATIONS”. In: *Solvent Extr. Ion Exch.* 10.1 (1992), pp. 69–100. DOI: 10.1080/07366299208918093 (cit. on p. 30).

[87] D. Prabhu, G. Mahajan, and G. Nair. “Di (2-ethyl hexyl) butyramide and di (2-ethyl hexyl) isobutyramide as extractants for uranium (VI) and plutonium (IV)”. In: *J. Radioanal. Nucl. Chem.* 224.1-2 (1997), pp. 113–117. DOI: 10.1007/bf02034622 (cit. on p. 30).

[88] V. Manchanda and P. Pathak. “Amides and diamides as promising extractants in the back end of the nuclear fuel cycle: an overview”. In: *Sep. Purif. Technol.* 35.2 (2004), pp. 85–103. DOI: 10.1016/j.seppur.2003.09.005 (cit. on p. 30).

[89] B. Mahanty, A. Bhattacharyya, A. S. Kanekar, and P. Mohapatra. “Selective Separation of Neptunium from an Acidic Feed Containing a Mixture of Actinides Using Dialkylamides”. In: *Solvent Extr. Ion Exc.* 38.3 (2020), pp. 290–303. DOI: 10.1080/07366299.2020.1720953 (cit. on p. 30).

[90] F. Rodrigues et al. “New insights into the extraction of uranium(VI) by an N,N-dialkylamide”. In: *Mol. Phys.* 112.9-10 (2014), 1362–1374. DOI: {10.1080/00268976.2014.902139} (cit. on p. 30).

[91] K. McCann, J. A. Drader, and J. C. Braley. “Comparing branched versus straight-chained monoamide extractants for actinide recovery”. In: *Sep. Purif. Rev.* 47.1 (2018), pp. 49–65. DOI: 10.1080/15422119.2017.1321018 (cit. on p. 30).

[92] P. G. Allen, D. K. Veirs, S. D. Conradson, C. A. Smith, and S. F. Marsh. “Characterization of Aqueous Plutonium(IV) Nitrate Complexes by Extended X-ray Absorption Fine Structure Spectroscopy”. In: *Inorg. Chem.* 35.10 (Jan. 1996), pp. 2841–2845. DOI: 10.1021/ic9511231 (cit. on p. 31).

[93] K. McCann, S. I. Sinkov, G. J. Lumetta, and J. C. Shafer. “Inner versus outer sphere metal-monoamide complexation: ramifications for tetravalent and hexavalent actinide selectivity”. In: *New J. Chem.* 42 (7 2018), pp. 5415–5424. DOI: 10.1039/C7NJ04851C (cit. on p. 31).

[94] S. Pahan, A. Boda, and S. M. Ali. “Density functional theoretical analysis of structure, bonding, interaction and thermodynamic selectivity of hexavalent uranium (UO22+) and tetravalent plutonium (Pu4+) ion complexes of tetramethyl diglycolamide (TMDGA)”. In: *Theor. Chem. Acc.* 134.4 (Mar. 2015), p. 41. DOI: 10.1007/s00214-015-1641-7 (cit. on p. 31).

[95] S. O. Odoh and G. Schreckenbach. "Theoretical Study of the Structural Properties of Plutonium(IV) and (VI) Complexes". In: *J. Phys. Chem. A* 115.48 (2011). PMID: 22040181, pp. 14110–14119. DOI: 10.1021/jp207556b (cit. on p. 31).

[96] M. Sulka, L. Cantrel, and V. Vallet. "Theoretical Study of Plutonium(IV) Complexes Formed within the PUREX Process: A Proposal of a Plutonium Surrogate in Fire Conditions". In: *J. Phys. Chem. A* 118.43 (2014). PMID: 25290588, pp. 10073–10080. DOI: 10.1021/jp507684f (cit. on pp. 31–33).

[97] P. G. Allen, D. K. Veirs, S. D. Conradson, C. A. Smith, and S. F. Marsh. "Characterization of Aqueous Plutonium(IV) Nitrate Complexes by Extended X-ray Absorption Fine Structure Spectroscopy". In: *Inorg. Chem.* 35.10 (1996), pp. 2841–2845. DOI: 10.1021/ic9511231 (cit. on p. 32).

[98] F. Weigend. "Accurate Coulomb-fitting basis sets for H to Rn". In: *Phys. Chem. Chem. Phys.* 8 (9 2006), pp. 1057–1065. DOI: 10.1039/B515623H (cit. on p. 33).

[99] A. Moritz, X. Cao, and M. Dolg. "Quasirelativistic energy-consistent 5f-in-core pseudopotentials for divalent and tetravalent actinide elements". In: *Theor. Chem. Acc.* 118.5 (Dec. 2007), pp. 845–854. DOI: 10.1007/s00214-007-0330-6 (cit. on p. 33).

[100] A. Moritz, X. Cao, and M. Dolg. "Quasirelativistic energy-consistent 5f-in-core pseudopotentials for divalent and tetravalent actinide elements". In: *Theor. Chem. Acc.* 118.5-6 (2007), pp. 845–854. DOI: 10.1007/s00214-007-0330-6 (cit. on p. 33).

[101] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu". In: *J. Chem. Phys* 132.15 (2010), p. 154104. DOI: 10.1063/1.3382344 (cit. on p. 33).

[102] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, and F. Weigend. "Turbomole". In: *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 4.2 (2014), pp. 91–100. DOI: 10.1002/wcms.1162 (cit. on p. 33).

[103] M. Sierka, A. Hogekamp, and R. Ahlrichs. "Fast evaluation of the Coulomb potential for electron densities using multipole accelerated resolution of identity approximation". In: *J. Chem. Phys.* 118.20 (2003), pp. 9136–9148. DOI: 10.1063/1.1567253 (cit. on p. 33).

[104] J. Tomasi, B. Mennucci, and R. Cammi. “Quantum mechanical continuum solvation models”. In: *Chem. Rev.* 105.8 (2005), pp. 2999–3094. DOI: 10.1021/cr9904009 (cit. on p. 34).

[105] R. G. Surbella, L. C. Ducati, J. Autschbach, K. L. Pellegrini, B. K. McNamara, J. M. Schwantes, and C. L. Cahill. “Plutonium chlorido nitrato complexes: ligand competition and computational metrics for assembly and bonding”. In: *Chem. Commun.* 54 (85 2018), pp. 12014–12017. DOI: 10.1039/C8CC05578E (cit. on p. 38).

[106] M. B. Doran, A. J. Norquist, and D. O’Hare. “Exploration of Composition Space in Templated Uranium Sulfates”. In: *Inorg. Chem.* 42.22 (2003). PMID: 14577764, pp. 6989–6995. DOI: 10.1021/ic034540j (cit. on p. 38).

[107] T. Sukhbaatar, M. Duvail, T. Dumas, S. Dourdain, G. Arrachart, P. L. Solari, P. Guilbaud, and S. Pellet-Rostaing. “Probing the existence of uranyl trisulfate structures in the AMEX solvent extraction process”. In: *Chem. Commun.* 55 (53 2019), pp. 7583–7586. DOI: 10.1039/C9CC02651G (cit. on p. 38).

[108] C. E. Ruggiero, J. H. Matonic, S. D. Reilly, and M. P. Neu. “Dissolution of plutonium (IV) hydroxide by desferrioxamine siderophores and simple organic chelators”. In: *Inorg. Chem.* 41.14 (2002), pp. 3593–3595. DOI: 10.1021/ic015591o (cit. on p. 39).

[109] R. Stanton. “Hellmann-Feynman Theorem and Correlation Energies”. In: *J. Chem. Phys* 36.5 (1962), pp. 1298–1300. DOI: 10.1063/1.1732731 (cit. on p. 49).

[110] D. Marx and J. Hutter. *Ab initio molecular dynamics: basic theory and advanced methods*. Cambridge University Press, 2009 (cit. on p. 49).

[111] C. Kittel and H. Kroemer. *Thermal physics*. American Association of Physics Teachers, 1998 (cit. on p. 50).

[112] J. M. Berg. “Biochemistry”. In: *Chem. Eng. News* 79.13 (2001), pp. 130–130 (cit. on p. 51).

[113] R. J. Good and C. J. Hope. “New combining rule for intermolecular distances in intermolecular potential functions”. In: *J. Chem. Phys* 53.2 (1970), pp. 540–543. DOI: 10.1063/1.1674022 (cit. on p. 55).

[114] D. Berthelot. “Sur le melange des gaz”. In: *Compt. Rendus* 126 (1898), pp. 1703–1706 (cit. on p. 55).

[115] R. Buckingham. “a.(1938) Proc. R. Soc”. In: *A Math. Phys. Eng. Sci* 168.933 (), pp. 264–283 (cit. on p. 56).

[116] G. Sutmann. “Molecular Dynamics-Extending the Scale from Microscopic to Mesoscopic”. In: *Multiscale Simulations Methods in Molecular Sciences* 42 (2009), pp. 1–50 (cit. on p. 56).

[117] J. N. Israelachvili. *Intermolecular and surface forces*. Academic press, 2011 (cit. on p. 56).

[118] F. Réal, V. Vallet, J.-P. Flament, and M. Masella. “Revisiting a many-body model for water based on a single polarizable site: From gas phase clusters to liquid and air/liquid water systems”. In: *J. Chem. Phys* 139.11 (2013), p. 114502. DOI: 10.1063/1.4821166 (cit. on pp. 56, 58, 59, 99).

[119] P. Cieplak, F. Y. Dupradeau, Y. Duan, and J. Wang. “Polarization effects in molecular mechanical force fields”. In: *J. Phys. Condens. Matter* 21.33 (2009), p. 333102. DOI: 10.1088/0953-8984/21/33/333102 (cit. on p. 56).

[120] G. Lamoureux and B. Roux. “Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm”. In: *J. Chem. Phys* 119.6 (2003), pp. 3025–3039. DOI: 10.1063/1.1589749 (cit. on p. 56).

[121] B. Thole. “Molecular polarizabilities calculated with a modified dipole interaction”. In: *Chem. Phys.* 59.3 (1981), pp. 341–350. DOI: 10.1016/0301-0104(81)85176-2 (cit. on p. 57).

[122] C. Clavaguera Sarrio, V. Brenner, S. Hoyau, C. J. Marsden, P. Millie, and J.-P. Dognon. “Modeling of Uranyl Cation Water Clusters”. In: *J. Phys. Chem. B* 107.13 (2003), pp. 3051–3060. DOI: 10.1021/jp0273833 (cit. on p. 58).

[123] F. Réal, M. Trumm, B. Schimmelpfennig, M. Masella, and V. Vallet. “Further insights in the ability of classical nonadditive potentials to model actinide ion?water interactions”. In: *J. Comput. Chem* 34.9 (), pp. 707–719. DOI: 10.1002/jcc.23184 (cit. on p. 58).

[124] L. Hemmingsen, P. Amara, E. Ansoborlo, and M. J. Field. “Importance of Charge Transfer and Polarization Effects for the Modeling of Uranyl- Cation Complexes”. In: *J. Phys. Chem. A* 104.17 (2000), pp. 4095–4101. DOI: 10.1021/jp994395o (cit. on p. 58).

[125] E. Acher. “Etude du plutonium(IV) en solution en couplant approches theoriques et experimentales”. PhD thesis. University of sciences and Technologies Lille, 2017 (cit. on p. 58).

[126] J. R. Damewood, R. A. Kumpf, W. C. Muhlbauer, J. J. Urban, and J. E. Eksterowicz. “Parametrization of molecular mechanics calculations for the accurate description of hydrogen-bonding interactions”. In: *J. Phys. Chem.* 94.17 (1990), pp. 6619–6626. DOI: 10.1021/j100380a019 (cit. on p. 59).

[127] M. Korth. “Empirical Hydrogen-Bond Potential Functions?An Old Hat Reconditioned”. In: *ChemPhysChem* 12.17 (2011), pp. 3131–3142. DOI: 10.1002/cphc.201100540 (cit. on p. 59).

[128] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods. “The Amber biomolecular simulation programs”. In: *J. Comput. Chem* 26.16 (), pp. 1668–1688. DOI: 10.1002/jcc.20290 (cit. on pp. 60, 75, 86).

[129] B. R. Brooks et al. “CHARMM: The biomolecular simulation program”. In: *J. Comput. Chem* 30.10 (), pp. 1545–1614. DOI: 10.1002/jcc.21287 (cit. on pp. 60, 75).

[130] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, and W. F. van Gunsteren. “The GROMOS Biomolecular Simulation Program Package”. In: *J. Phys. Chem. A* 103.19 (1999), pp. 3596–3607. DOI: 10.1021/jp984217f (cit. on pp. 60, 75).

[131] J. Doherty. “PEST: Model Independent Parameter Estimation”. In: (Jan. 2010) (cit. on p. 63).

[132] *The Polaris-MD code*. http://biodev.cea.fr/polaris/. Accessed: 2020-10-30 (cit. on pp. 63, 65, 85).

[133] H. J. Berendsen, J. v. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. “Molecular dynamics with coupling to an external bath”. In: *J. Chem. Phys* 81.8 (1984), pp. 3684–3690. DOI: 10.1063/1.448118 (cit. on p. 65).

[134] H. C. Andersen. “Molecular dynamics simulations at constant pressure and/or temperature”. In: *J. Chem. Phys* 72.4 (1980), pp. 2384–2393. DOI: 10.1063/1.439486 (cit. on pp. 65, 66).

[135] S. Nosé. “A unified formulation of the constant temperature molecular dynamics methods”. In: *J. Chem. Phys* 81.1 (1984), pp. 511–519. DOI: 10.1063/1.447334 (cit. on p. 65).

[136] G. Raabe. “Molecular Simulation Studies on Thermophysical Properties”. In: *With Application to Working Fluids (Molecular Modeling and Simulation)(Singapore: Springer)* (2017) (cit. on p. 66).

[137] S. Melchionna, G. Ciccotti, and B. Lee Holian. “Hoover NPT dynamics for systems varying in shape and size”. In: *Molec. Phys* 78.3 (1993), pp. 533–544. DOI: 10.1080/00268979300100371 (cit. on p. 67).

[138] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. “A smooth particle mesh Ewald method”. In: *J. Chem. Phys* 103.19 (1995), pp. 8577–8593. DOI: 10.1063/1.470117 (cit. on p. 68).

[139] M. S. Reed and K. M. Flurchick. “Investigation of artifacts due to periodic boundary conditions”. In: *Comput. Phys. Commun* 95.1 (1996), pp. 39–46. DOI: 10.1016/0010-4655(95)00140-9 (cit. on p. 68).

[140] J. Wang and T. Hou. “Application of molecular dynamics simulations in molecular property prediction. 1. Density and heat of vaporization”. In: *J. Chem. Theory Comput* 7.7 (2011), pp. 2151–2165. DOI: 10.1021/ct200142z (cit. on pp. 72, 86).

[141] J. Wang and T. Hou. “Application of molecular dynamics simulations in molecular property prediction II: diffusion coefficient”. In: *J. Comput. Chem* 32.16 (2011), pp. 3505–3519. DOI: 10.1002/jcc.21939 (cit. on p. 73).

[142] W. W. Schulz and J. D. Navratil. “Science and technology of tributyl phosphate”. In: (1987) (cit. on p. 75).

[143] K. D. Papavasileiou, L. D. Peristeras, A. Bick, and I. G. Economou. “Molecular Dynamics Simulation of Pure n-Alkanes and Their Mixtures at Elevated Temperatures Using Atomistic and Coarse-Grained Force Fields”. In: *J. Phys. Chem. B* 123.29 (2019), pp. 6229–6243. DOI: 10.1021/acs.jpcb.9b02840 (cit. on p. 75).

[144] F. Goodsaid-Zalduondo and D. Engelman. “Conformation of liquid N-alkanes”. In: *Biophysical journal* 35.3 (1981), pp. 587–594 (cit. on p. 75).

[145] P. L. Yeagle. *The structure of biological membranes*. CRC press, 2004 (cit. on p. 75).

[146] J. Murgich. “Molecular simulation and the aggregation of the heavy fractions in crude oils”. In: *Molecular Simulation* 29.6-7 (2003), pp. 451–461. DOI: 10.1080/0892702031000148762 (cit. on p. 75).

[147] C. B. G. J. X. Chunming. “The Applications of Molecular Simulation Technology in the Fields of Petroleum [J]”. In: *Progress In Chemistry* 2 (2004) (cit. on p. 75).

[148] P. Ungerer, V. Lachet, and B. Tavitian. “Applications of molecular simulation in oil and gas production and processing”. In: *Oil. Gas. Sci. Technol.* 61.3 (2006), pp. 387–403. DOI: 10.2516/ogst:2006040a (cit. on p. 75).

[149] X. Chen, L. Hou, X. Wei, and D. Bedrov. “Transport Properties of Waxy Crude Oil: A Molecular Dynamics Simulation Study”. In: *ACS omega* 5.30 (2020), pp. 18557–18564. DOI: 10.1021/acsomega.0c00070 (cit. on p. 75).

[150] A. A. Petrov. *Petroleum hydrocarbons*. Springer Science & Business Media, 2012 (cit. on p. 75).

[151] N. L. Allinger, Y. H. Yuh, and J. H. Lii. “Molecular mechanics. The MM3 force field for hydrocarbons. 1”. In: *J. Am. Chem. Soc* 111.23 (1989), pp. 8551–8566. DOI: 10.1021/ja00205a001 (cit. on p. 75).

[152] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. “Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids”. In: *J. Am. Chem. Soc* 118.45 (1996), pp. 11225–11236. DOI: 10.1021/ja9621760 (cit. on pp. 75, 76, 86, 90).

[153] I. V. Vorobyov, V. M. Anisimov, and A. D. MacKerell. “Polarizable empirical force field for alkanes based on the classical drude oscillator model”. In: *J. Phys. Chem. B* 109.40 (2005), pp. 18988–18999. DOI: 10.1021/jp053182y (cit. on p. 76).

[154] S. W. Siu, K. Pluhackova, and R. A. bockmon. “Optimization of the OPLS-AA force field for long hydrocarbons”. In: *J. Chem. Theory Comput* 8.4 (2012), pp. 1459–1470. DOI: /10.1021/ct200908r (cit. on pp. 76, 86, 87, 90).

[155] A. M. Nikitin, Y. V. Milchevskiy, and A. P. Lyubartsev. “A new AMBER-compatible force field parameter set for alkanes”. In: *J. Mol. Model* 20.3 (2014), p. 2143. DOI: 10.1007/s00894-014-2143-6 (cit. on pp. 76, 127).

[156] A. Jagielska, Y. A. Arnautova, and H. A. Scheraga. “Derivation of a new force field for crystal-structure prediction using global optimization: nonbonded potential parameters for amines, imidazoles, amides, and carboxylic acids”. In: *J. Phys. Chem. B* 108.32 (2004), pp. 12181–12196. DOI: 10.1021/jp040115f (cit. on p. 76).

[157] J.-H. Lii and N. L. Allinger. “The MM3 force field for amides, polypeptides and proteins”. In: *J. Comput. Chem* 12.2 (1991), pp. 186–199. DOI: 10.1002/jcc.540120208 (cit. on p. 76).

[158] A. Hagler, E. Huler, and S. Lifson. “Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals”. In: *J. Am. Chem. Soc* 96.17 (1974), pp. 5319–5327. DOI: 10.1021/ja00824a004 (cit. on p. 76).

[159] J. B. Klauda, B. R. Brooks, A. D. MacKerell, R. M. Venable, and R. W. Pastor. “An ab initio study on the torsional surface of alkanes and its effect on molecular simulations of alkanes and a DPPC bilayer”. In: *J. Phys. Chem. B* 109.11 (2005), pp. 5300–5311. DOI: 10.1021/jp0468096 (cit. on p. 78).

[160] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon. “A fifth-order perturbation comparison of electron correlation theories”. In: *Chem. Phys. Lett* 157.6 (1989), pp. 479–483. DOI: 10.1016/S0009-2614(89)87395-6 (cit. on p. 78).

[161] F. Weinhold. “Natural bond orbital analysis: A critical overview of relationships to alternative bonding perspectives”. In: *J. Comput. Chem* 33.30 (), pp. 2363–2379. DOI: 10.1002/jcc.23060 (cit. on p. 80).

[162] C. Campañá, B. Mussard, and T. K. Woo. “Electrostatic Potential Derived Atomic Charges for Periodic Systems Using a Modified Error Functional”. In: *J. Chem. Theory Comput* 5.10 (2009). PMID: 26631798, pp. 2866–2878. DOI: 10.1021/ct9003405 (cit. on p. 80).

[163] W. D. Cornell, P. Cieplak, C. I. Bayly, and P. A. Kollman. “Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation”. In: *J. Am. Chem. Soc* 115.21 (1993), pp. 9620–9631. DOI: 10.1021/ja00074a030 (cit. on p. 80).

[164] S. Huzinaga and S. Narita. “Mulliken Population Analysis and Point Charge Model of Molecules”. In: *Israel Journal of Chemistry* 19.1?4 (), pp. 242–254. DOI: 10.1002/ijch.198000027 (cit. on p. 80).

[165] A. V. Marenich, S. V. Jerome, C. J. Cramer, and D. G. Truhlar. “Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases”. In: *J. Chem. Theory Comput* 8.2 (2012). PMID: 26596602, pp. 527–541. DOI: 10.1021/ct200866d (cit. on p. 80).

[166] J. F. Harrison. "A Hirshfeld-I interpretation of the charge distribution, dipole and quadrupole moments of the halogenated acetylenes FCCH, ClCCH, BrCCH, and ICCH". In: *J. Chem. Phys* 133.21 (2010), p. 214103. DOI: 10.1063/1.3511784 (cit. on p. 80).

[167] L. S. Dodda, J. Z. Vilseck, K. J. Cutrona, and W. L. Jorgensen. "Evaluation of CM5 Charges for Nonaqueous Condensed-Phase Modeling". In: *J. Chem. Theory Comput* 11.9 (2015). PMID: 26575922, pp. 4273–4282. DOI: 10.1021/acs.jctc.5b00414 (cit. on p. 80).

[168] D. Bedrov, J.-P. Piquemal, O. Borodin, A. D. MacKerell Jr, B. Roux, and C. Schröder. "Molecular dynamics simulations of ionic liquids and electrolytes using polarizable force fields". In: *Chem. rev.* 119.13 (2019), pp. 7940–7995. DOI: 10.1021/acs.chemrev.8b00763 (cit. on p. 81).

[169] A. V. Marenich, C. J. Cramer, and D. G. Truhlar. "Reduced and quenched polarizabilities of interior atoms in molecules". In: *Chem. Science* 4.6 (2013), pp. 2349–2356. DOI: 10.1039/C3SC50242B (cit. on p. 81).

[170] F. L. Hirshfeld. "Bonded-atom fragments for describing molecular charge densities". In: *Theoretica chimica acta* 44.2 (1977), pp. 129–138. DOI: 10.1007/BF00549096 (cit. on p. 81).

[171] J. P. Coles and M. Masella. "The fast multipole method and point dipole moment polarizable force fields". In: *J. Chem. Phys.* 142.2 (2015), p. 024109. DOI: 10.1063/1.4904922 (cit. on p. 85).

[172] W. M. Haynes. *CRC handbook of chemistry and physics*. CRC press, 2014 (cit. on p. 87).

[173] M. Dettenmaier. "Conformation of n-alkane molecules in the melt and in cyclohexane solution studied by small-angle neutron scattering". In: *J. Chem. Phys.* 68.5 (1978), pp. 2319–2322. DOI: 10.1016/S0006-3495(81)84814-X (cit. on p. 88).

[174] L. S. Bartell and D. Kohl. "Structure and rotational isomerization of free hydrocarbon chains". In: *J. Chem. Phys.* 39.11 (1963), pp. 3097–3105. DOI: 10.1063/1.1734149 (cit. on p. 88).

[175] F. Menger and L. D'Angelo. "Does the conformation of hydrocarbon chains depend on solvation?" In: *J. Am. Chem.* 110.24 (1988), pp. 8241–8242. DOI: 10.1021/ja00232a049 (cit. on p. 88).

[176] H. Casal and H. Mantsch. "Positional dependence of solvent effects on the conformation of liquid n-alkanes. An infrared spectroscopic study". In: *J. Mol. Struct* 192.1-2 (1989), pp. 41–45. DOI: 10.1016/0022-2860(89)87004-8 (cit. on p. 90).

[177] L. L. Thomas, T. J. Christakis, and W. L. Jorgensen. “Conformation of alkanes in the gas phase and pure liquids”. In: *J. Phys. Chem. B* 110.42 (2006), pp. 21198–21204. DOI: /10.1021/jp064811m (cit. on p. 90).

[178] A. Habenschuss and A. Narten. “X-ray diffraction study of some liquid alkanes”. In: *J. Chem. Phys.* 92.9 (1990), pp. 5692–5699. DOI: 10.1063/1.458500 (cit. on pp. 90, 91).

[179] A. Ben-Naim and Y. Marcus. “Solvation thermodynamics of nonionic solutes”. In: *J. Chem. Phys.* 81.4 (1984), pp. 2016–2027. DOI: 10.1063/1.447824 (cit. on p. 99).

[180] J. A. Rackers, Z. Wang, C. Lu, M. L. Laury, L. Lagardère, M. J. Schnieders, J.-P. Piquemal, P. Ren, and J. W. Ponder. “Tinker 8: software tools for molecular design”. In: *J. Chem. Theory Comput.* 14.10 (2018), pp. 5273–5289. DOI: r/10.1021/acs.jctc.8b00529 (cit. on p. 100).

[181] J. P. Foster and F. Weinhold. “Natural hybrid orbitals”. In: *J. Am. Chem. Soc* 102.24 (Nov. 1980), pp. 7211–7218. DOI: 10.1021/ja00544a007 (cit. on p. 121).

[182] M. M.-N. (Mautner). “The proton affinity scale, and effects of ion structure and solvation”. In: *Int. J. Mass Spectrom.* 227.3 (2003). Thermochemistry and Solvation of Gas Phase Ions, pp. 525–554. DOI: 10.1016/S1387-3806(03)00100-3 (cit. on p. 122).

[183] A. D. MacKerell Jr, D. Bashford, M. Bellott, R. L. Dunbrack Jr, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, et al. “All-atom empirical potential for molecular modeling and dynamics studies of proteins”. In: *J. Phys. Chem. B* 102.18 (1998), pp. 3586–3616. DOI: 10.1021/jp973084f (cit. on p. 127).

[184] G. ANDEREGG. *MANDEL J-STATISTICAL ANALYSIS OF EXPERIMENTAL DATA*. 1966 (cit. on p. 128).

Appendix A

Supporting Information

The Cartesian coordinates of all structures are available on the free platform Zenodo at the following link with DOI 10.5281/zenodo.4068222

A.1. Intrinsic properties of the ligands

Intrinsic properties of the ligands, such as basicity and electronic properties are important parameters for understanding their ability to bind a metal ion. We hereafter report atomic charges, molecular polarizabilities and proton affinities of the ligands, computed as the enthalpy change associated with the following gas-phase reaction:

$$\mathrm{LH^+(gas)} \rightleftharpoons \mathrm{L(gas)} + \mathrm{H^+(gas)} \tag{A.1}$$

The computed atomic charges and molecular polarizabilities for the different ligands are reported in Table A.1. Atomic charges were determined from a Natural Population Analysis (NPA) [181]. The charges of the

complexing oxygen atom are slightly affected by the lengthening of the alkyl chains, they vary from -0.61 to -0.63 for monoamides and from -0.63 to -0.65 for carbamides. Carbon and nitrogen electronic charges vary between monoamides and carbamides but do not vary significantly when changing the alkyl groups. The molecular polarizability shows a notable increase, from 8.9 $Å^3$ for MMM to 15.9 $Å^3$ for PEE. This rise was expected, since the increase in polarizability is equal to expected addition of polarizability for $-CH_2-$ groups ($\simeq$ 1.8 $Å^3$ per group). As polarizability increases, the dispersion forces are expected to become stronger.

TABLE A.1: Electronic atomic charges (NPA charges q_O, q_N and q_C), molecular polarizabilities (α in $Å^3$) and proton affinity (PA) for the different amide ligands in kJ $\cdot$ mol^{-1}, relative to that of the MMM ligand.

Charges	q_O	q_N	q_C	α	PA
MMM	–0.613	–0.409	0.653	8.9	0
MME	–0.614	–0.410	0.658	10.6	8
EMM	–0.617	–0.411	0.652	10.6	10
EEM	–0.618	–0.413	0.657	12.3	16
MEE	–0.621	–0.415	0.657	12.3	14
EEE	–0.625	–0.414	0.659	14.1	22
PEE	–0.623	–0.413	0.662	15.9	27
IEE	–0.628	–0.415	0.665	15.8	28
C4M	–0.631	–0.441	0.761	11.8	23
C4E	–0.652	–0.445	0.778	18.9	40

The proton affinities (PA) of the ligands relative to that of the MMM ligand are reported in Table A.1. For monoamides, the protonation energy increases when substituting a methyl by an ethyl group, by 10 kJ $\cdot$ mol^{-1} in R_1 position and by 8 kJ $\cdot$ mol^{-1} in R_2 or R_3 positions. As expected, the proton affinities tend to increase with the oxygen electronic charge and with the size of the alkyl chain. A similar observation can be made for carbamides, the protonation affinity increased from 23 up to 40 kJ $\cdot$ mol^{-1} by substituting methyl groups into ethyl. As reported in literature, for organic molecules the proton affinities tend to increase with the size and polarizability of the alkyl substituent [182].

A.2. The structural parameters for geometry I

TABLE A.2: Selected interatomic distances (average values d in Å)[a] and hydrogen bond angle ($\alpha_{\mathrm{O-H-O}}$) in ° calculated in the [Pu(NO_3)$_6$](H***L***$_2$) complex in geometry I optimized the gas phase. [a] O_{nit1} denotes oxygen from nitrate ions which are not involved in hydrogen bonds. O_{nit} denotes oxygen from nitrate ions which are involved in hydrogen bonds.

Ligand	$d_{\mathrm{O_{nit}-O_L}}$	$d_{\mathrm{Pu-O_{nit1}}}$	$d_{\mathrm{Pu-N_{nit2}}}$	$d_{\mathrm{Pu-N_L}}$	$d_{\mathrm{Pu-C_{R1}}}$	α_{HOH}
MMM	2.568	2.544	2.976	5.605	4.910	169.8
MME	2.598	2.541	2.972	5.536	4.905	169.4
EMM	2.603	2.526	2.992	5.411	5.176	165.3
EEM	2.595	2.514	2.959	5.386	5.129	163.0
MEE	2.566	2.544	2.978	5.720	4.897	170.1
EEE	2.589	2.537	2.968	5.703	5.167	163.4
PEE	2.594	2.537	2.974	5.839	5.122	163.4
IEE	2.551	2.537	2.969	5.682	5.996	158.4
C4M	2.592	2.529	2.957	5.427	5.793	171.8
C4E	2.573	2.540	2.972	5.584	5.947	168.2

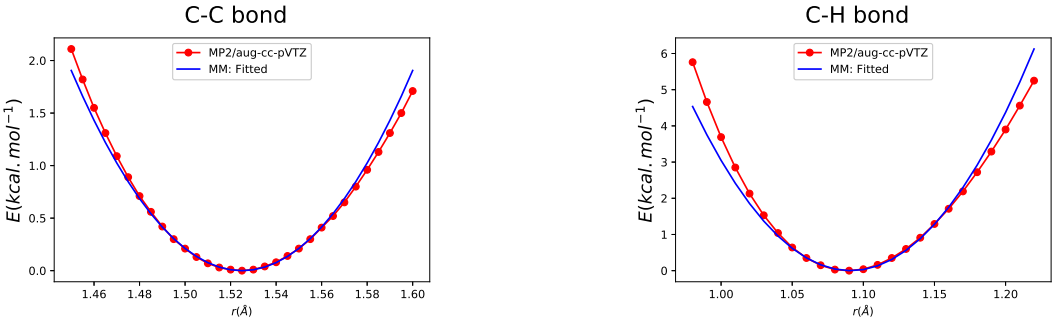
Appendix B

Supporting Information

B.1. Bond and Angle parameters

Bond stretching: (K_b, r_0)

For alkanes, we have two types of bonds, C-H and C-C. After performing the scans, and fitting the binding energy analytic formula to the QM data using a python script, we obtain the values of the harmonic force constants. The fittings of the C−C and C−H bond stretching illustrated in Figure B.1.



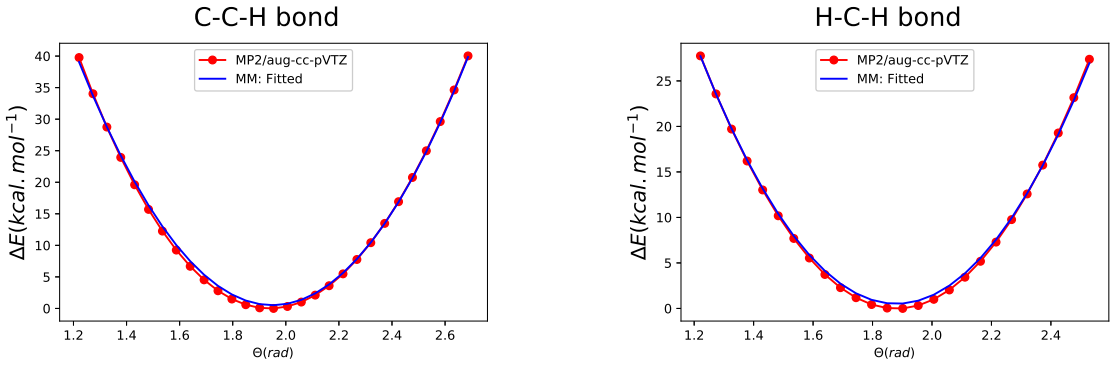
(a) (b)

FIGURE B.1: Potential energy curves for the C-C (a) and C-H (b) bond stretching. QM data are in red filled circle and the MM fit in blue.

It is too important to restrict the stretching scan in an interval close to the bond equilibrium value, within the harmonic range, in order not to break the bond otherwise the fit will not be possible.

Angle bending:(k_a, θ_0)

In alkanes, we have three angles (C−C−C, C−C−H and H−C−H). In order to calculate the bending constants, $i)$ scans over the angles were performed. $ii)$ fitting the obtained QM data to the classical analytical functions (harmonic potential 4.9), describing the contribution of angles. The fitting results are presented in Figure B.2.



(a) (b)

FIGURE B.2: Potential energy curves for the C-C-H (a), H-C-H (b) bending angles. QM data are in red filled circle and the MM fit in blue.

Synopsis

To sum up this section, our optimized FF parameters are given in Table B.1, facing that obtained from other FFs, namely CHARMM and AMBER.

If we compare all the parameters of Table B.1, we can notice that there is a small difference between the parameters of CHARMM and those of Amber, this is due to the fact that the ones in Amber are also derived using QM methods unlike the CHARMM ones which are based on experimental data. However, the bond and angle intramolecular parameters of amber or CHARMM FF can be used with no worries, the bond and angles parameters are known to be less critical; since at room temperature, bond and angle vibrations typically does not become high enough to have a qualitatively important influence on the dynamics and

angle / Bond	Source	r_0(Å) or $\phi_0(^\circ)$	$k(cm^{-1})$
C-C	CHARMM FF	1.530	222
	This work	1.525	242
	Amber FF	1.526	310
C-H	CHARM FF	1.09	322
	Amber FF	1.09	340
	This work	1.09	353
C-C-H	CHARM FF	110.1	37
	Amber FF	109.5	50
	This work	112.0	44
H-C-H	CHARM FF	108.4	35
	Amber FF	109.5	35
	This work	108.1	32
H-C-C-H	CHARMM FF	n=3, Φ_0 =0	0.25 (kcal $\cdot$ mol^{-1})
	Amber FF	n=3, Φ_0 =0	1.40 (kcal $\cdot$ mol^{-1})
	This work	n=3, Φ_0 =0	1.41 (kcal $\cdot$ mol^{-1})

TABLE B.1: Intramolecular FF parameters for alkanes, taken from the CHARMM [183], Amber [155] FFs and this work.

energetics. Meanwhile, they are critical when performing vibrational spectroscopy studies, for which the intramolecular motions play a vital role.

However, for the dihedral torsion motions, we can clearly see marked differences between the Amber and CHARMM FFs. This is due to the fact that dihedral torsion parameters are quite dependent of intermolecular parameters (Van der Waals, electrostatic and 1-4 interactions). In CHARMM, identical Van der Waals parameters are used for the three types of carbon atoms ($-CH_3, -CH-$ or $-CH_2-$). Hence, they have the same parameters for the (CTx$-$CT2$-$CT2$-$CTx with $x = 2, 3$) torsions, while we have considered them to be different for a better transferability of the parameters over the series of alkanes. We have thus adjusted all the dihedral parameters for alkanes and amides. The final set of dihedral parameters used in this work is not reported, there are a lot of parameters, but available on demand.

B.2. Error estimations

The approach used to estimate the errors in this work is based entirely on the statistical nature of the results. Assumimg that an MD simulation is performed for a total period of t_{tot} (production time). We select the last 20% of the production (noted t_{st}, we can divide t_{st} into N segments with the end point of each segment being $t_i = i\Delta t$ (with $i = 1, 2, \ldots, N$) where $\Delta t = t_{st}/N$. Any time-averaged property can be calculated for each of the time intervals $\Delta t_i = t_i - t_{i-1} = \Delta t$, and as a result, each MD simulation will produce N values of the property P. If we denote each estimate of P to be P_i (with $i = 1, 2, \ldots, N$), the best estimate of the property can be calculated as

$$\overline{P} = \frac{\sum_{i=1}^{N} P_i}{N}. \tag{B.1}$$

The uncertainty of the samples P_i can be quantified by the sample standard deviation defined as :

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (\overline{P} - P_i)^2}{N}}. \tag{B.2}$$

The best estimate $\overline{P}$ also associated with an uncertainty $\overline{\sigma}$, this latter is reduced from σ through the well-known relationship [184]:

$$\overline{\sigma} = \frac{\sigma}{\sqrt{N}}. \tag{B.3}$$

B.3. Dimer structures used to derive Buckingham parameters

Herein, we reported just some of the dimer structures with its' corresponding interaction energy curves as a function of distance between the two molecules. In total, we fitted over 40 other structures and overall the MM energy curves fits well all the QM reference data.

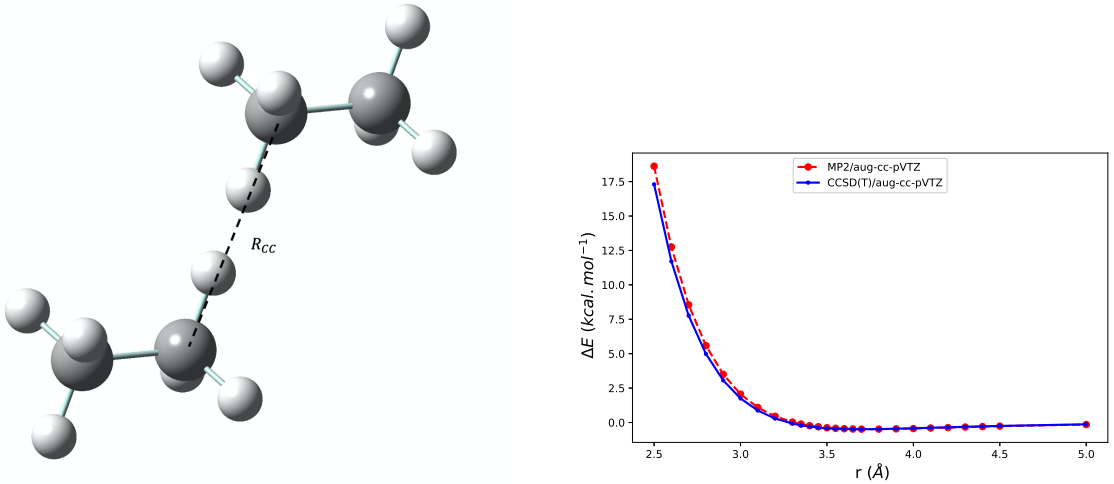


FIGURE B.3: Ethane dimer structure on the left with the corresponding interaction energy curve a function of the intermolecular distance computed at the MP2 and CCSD(T) levels of theory.

B.4. Radial distribution functions

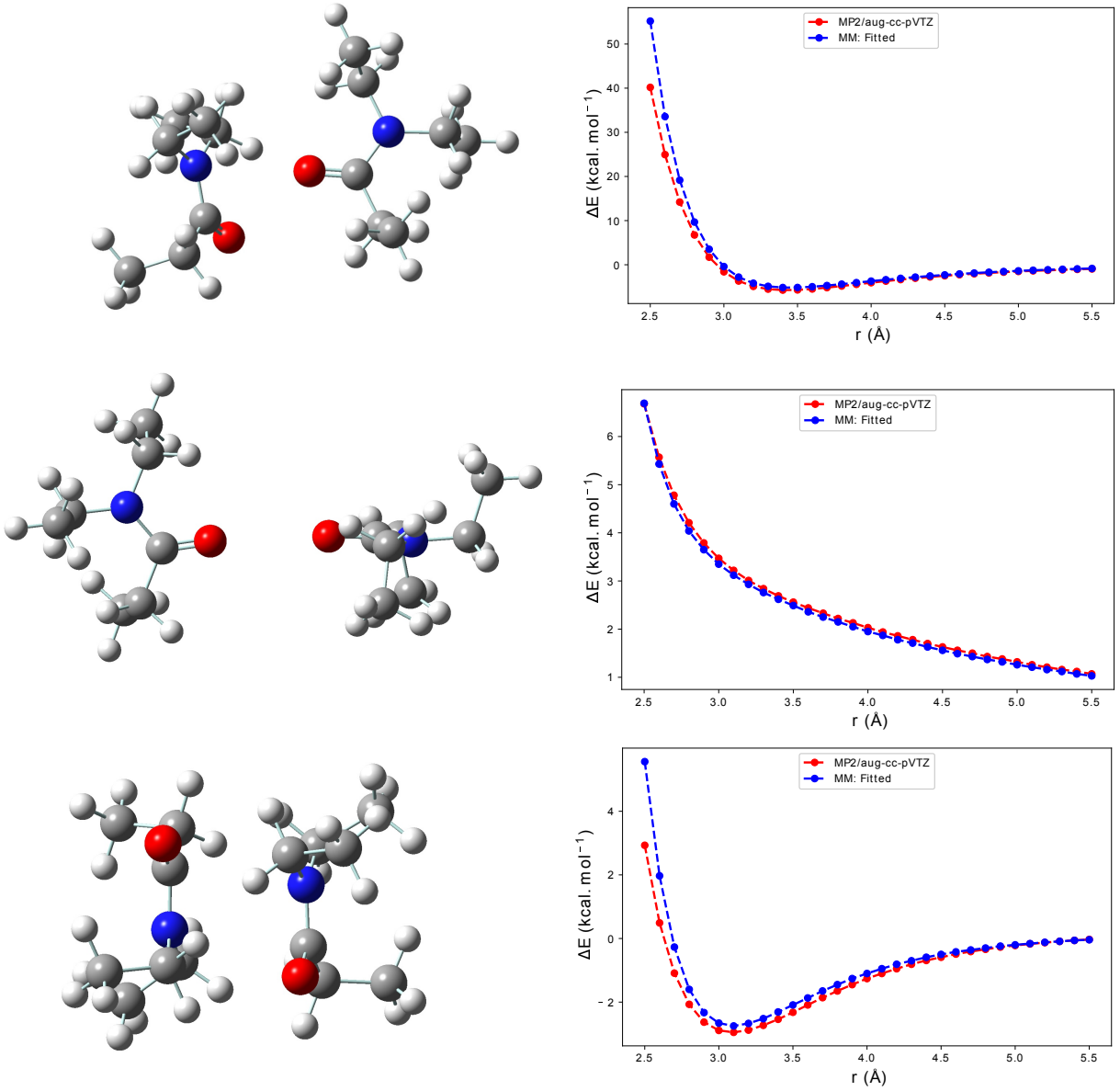


FIGURE B.4: Selected DEPA dimer structures on the left with the corresponding interaction energy curves as a function of distance between the two molecules. MP2 values are in red and the fitted FF in blue.

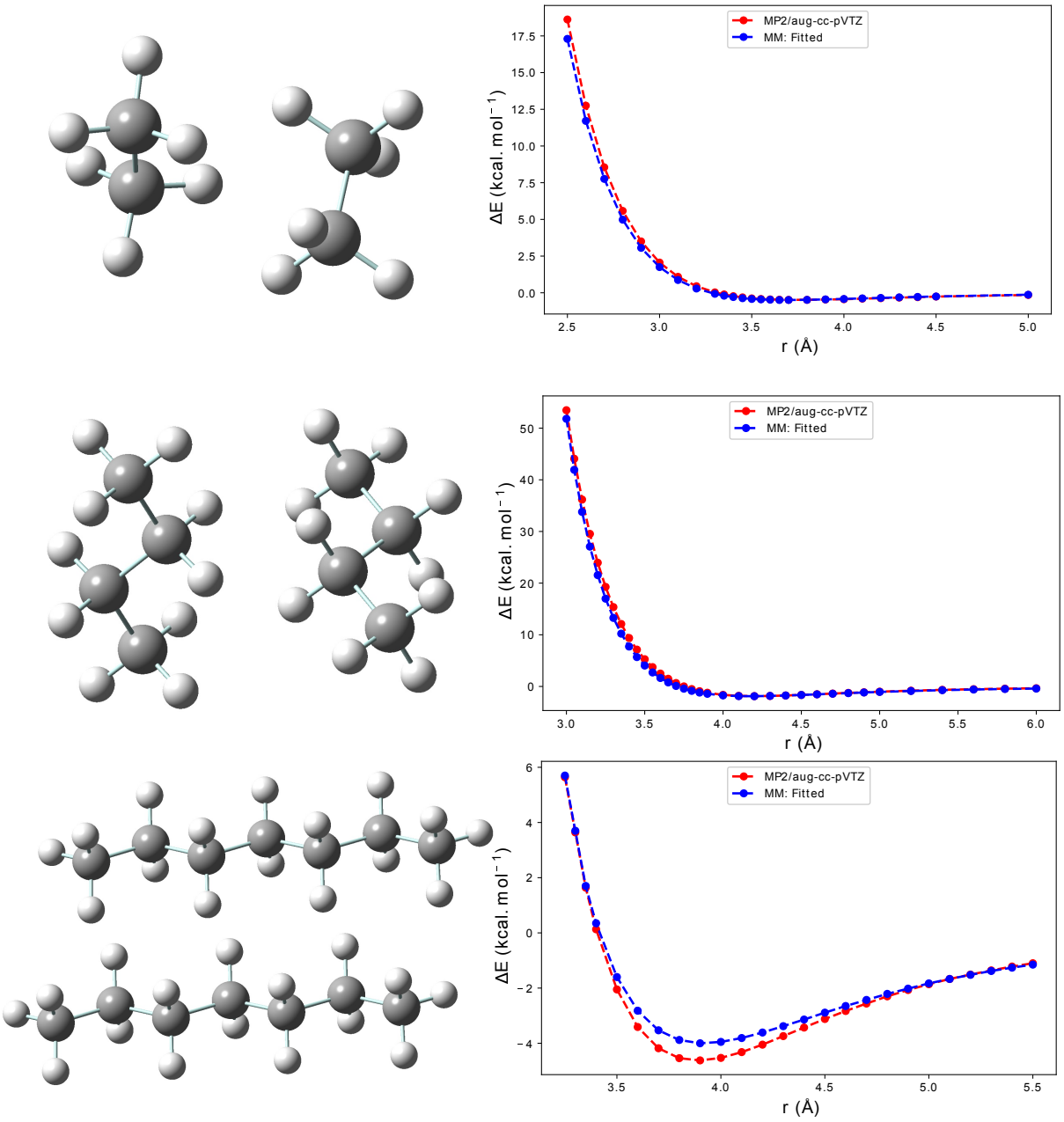


FIGURE B.5: Selected dimer structures (for ethane, butane, and heptane) opposite to their interaction energy curve as a function of distance between the two molecules. MP2 values are in red and the fitted FF in blue.

i	j	A_{ij}	B_{ij}	damping	C_{ij}	correc tor (1-4)	
CT3	CT3	3.86	320643	0.3	1850	4.82	1000000
CT3	CT2	3.80	320000	0.3	1500	4.82	1000000
CT3	NCT3	4.45	320000	0.3	1500	5.30	1000000
CT3	C	3.33	71347	0.3	1550	4.70	1000000
CT3	O	4.25	300000	0.3	0	4.16	1000000
CT3	HA	3.84	5440	0.5	0	7.50	75000
CT3	N	4.65	600000	0.3	700	4.90	1000000
CT2	CT2	4.55	320000	0.3	1400	4.82	1000000
CT2	NCT3	4.50	1004868	0.3	500	4.82	1000000
CT2	C	3.65	72000	0.3	1500	4.82	1000000
CT2	O	4.40	1000000	0.3	500	4.16	1000000
CT2	HA	3.70	5700	0.5	0	7.50	75000
CT2	N	4.15	304000	0.3	700	4.82	1000000
NCT3	C	3.43	101347	0.3	700	4.82	1000000
NCT3	O	4.25	500000	0.3	700	4.16	1000000
NCT3	HA	3.23	1476	0.5	0	7.50	75000
NCT3	N	4.75	500000	0.3	700	5.30	1000000
C	C	4.50	304000	0.3	500	4.82	1000000
C	O	4.20	500000	0.4	1000	5.30	1000000
C	HA	3.23	1476	0.5	0	7.50	75000
C	N	4.30	1000000	0.3	700	5.30	800000
O	O	5.10	900000	0.3	0	4.80	1000000
O	HA	7.00	75000	0.5	0	8.00	50000
O	N	5.20	500000	0.5	1200	5.10	1000000
HA	HA	3.65	1000	0.5	0	8.00	50000
HA	N	5.50	350000	0.5	0	8.00	50000
N	N	4.70	9000000	0.3	700	5.60	500000

TABLE B.2: The Buckingham parameters derived and used in this work,

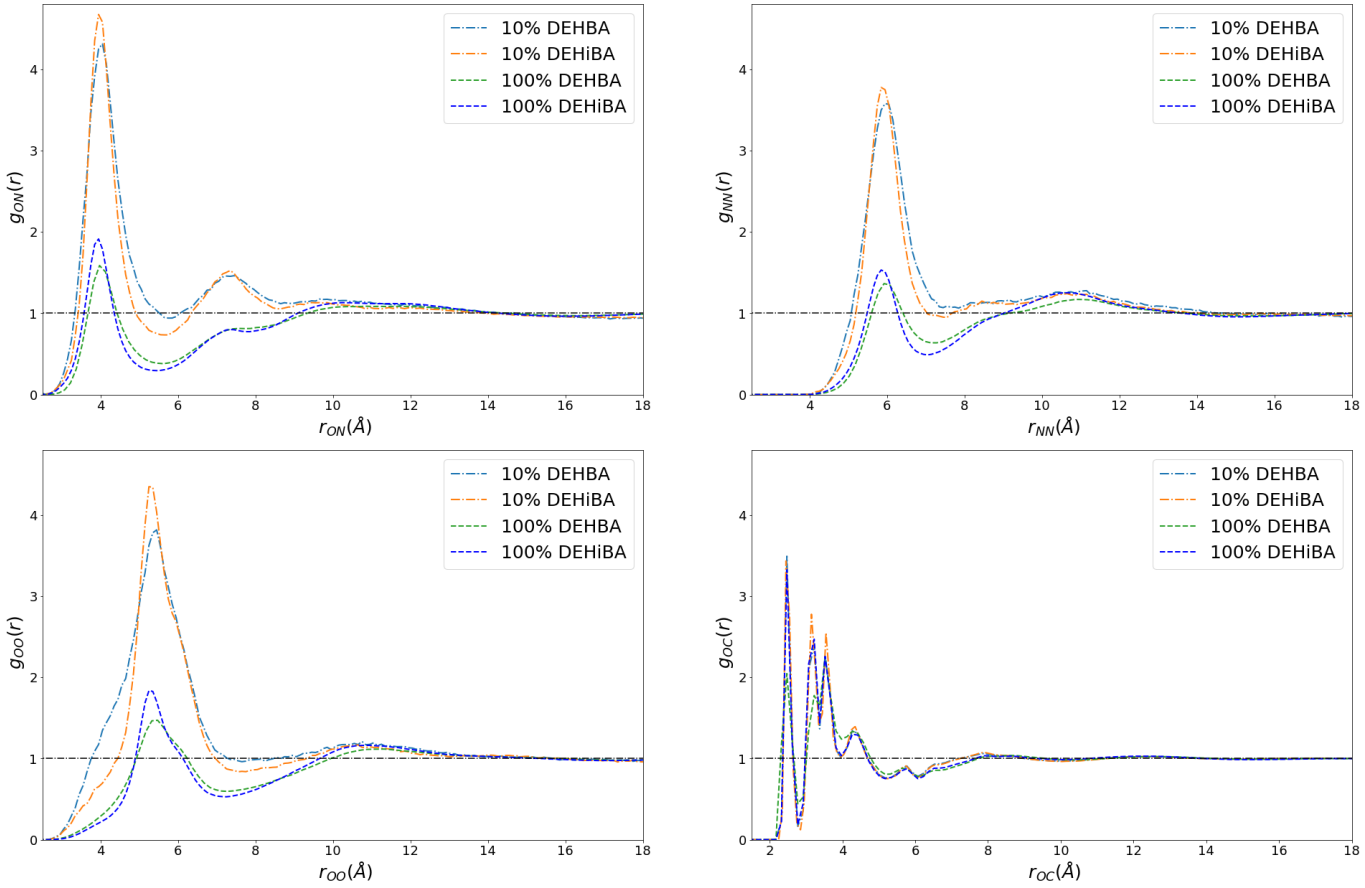


FIGURE B.6: Radial distribution functions of oxygen, carbon and nitrogen atoms in DEHBA/dodecane and DEHiBA/dodecane at 298 K and pure monoamides (DEHBA,DEHiBA)

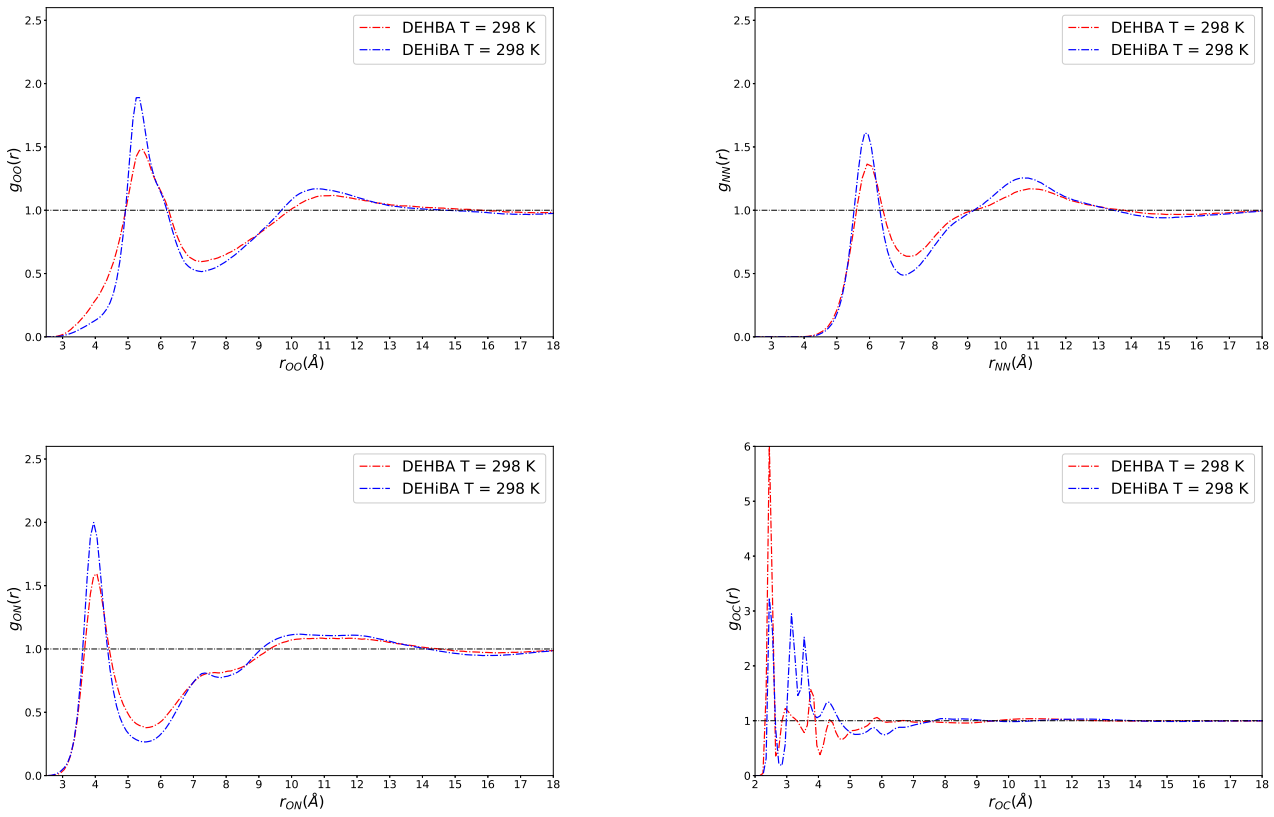
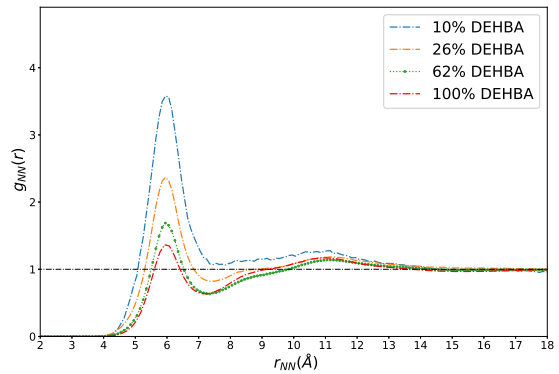
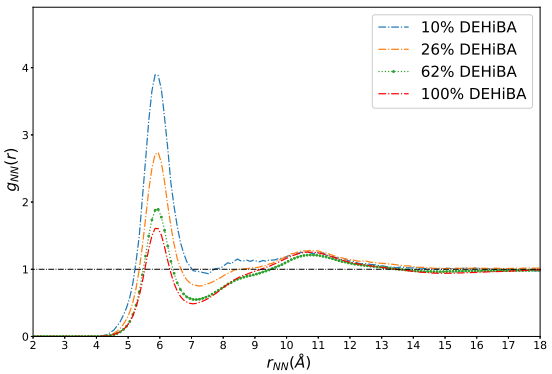
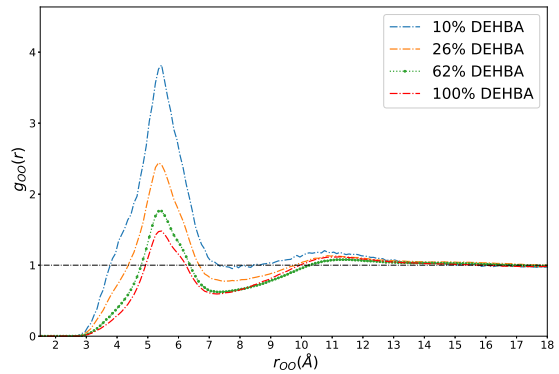
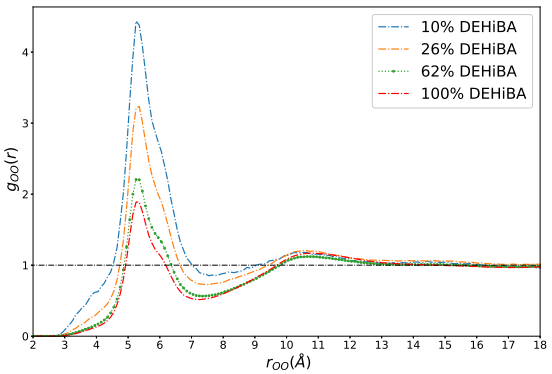


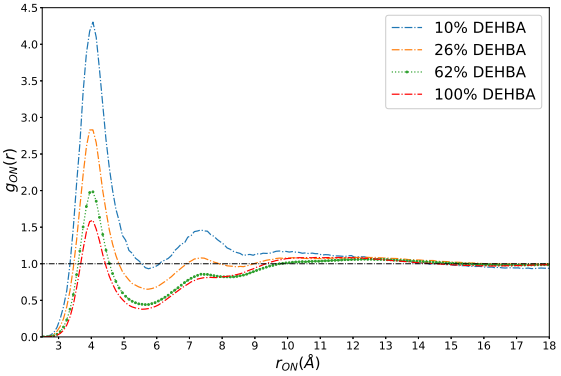
FIGURE B.7: Radial distribution functions of oxygen, carbon and nitrogen atoms in DEHiBA and DEHBA at 298 K.

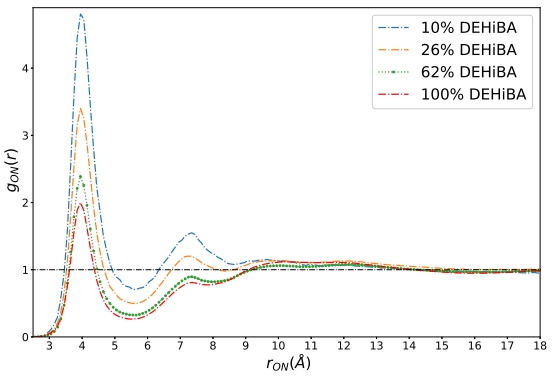






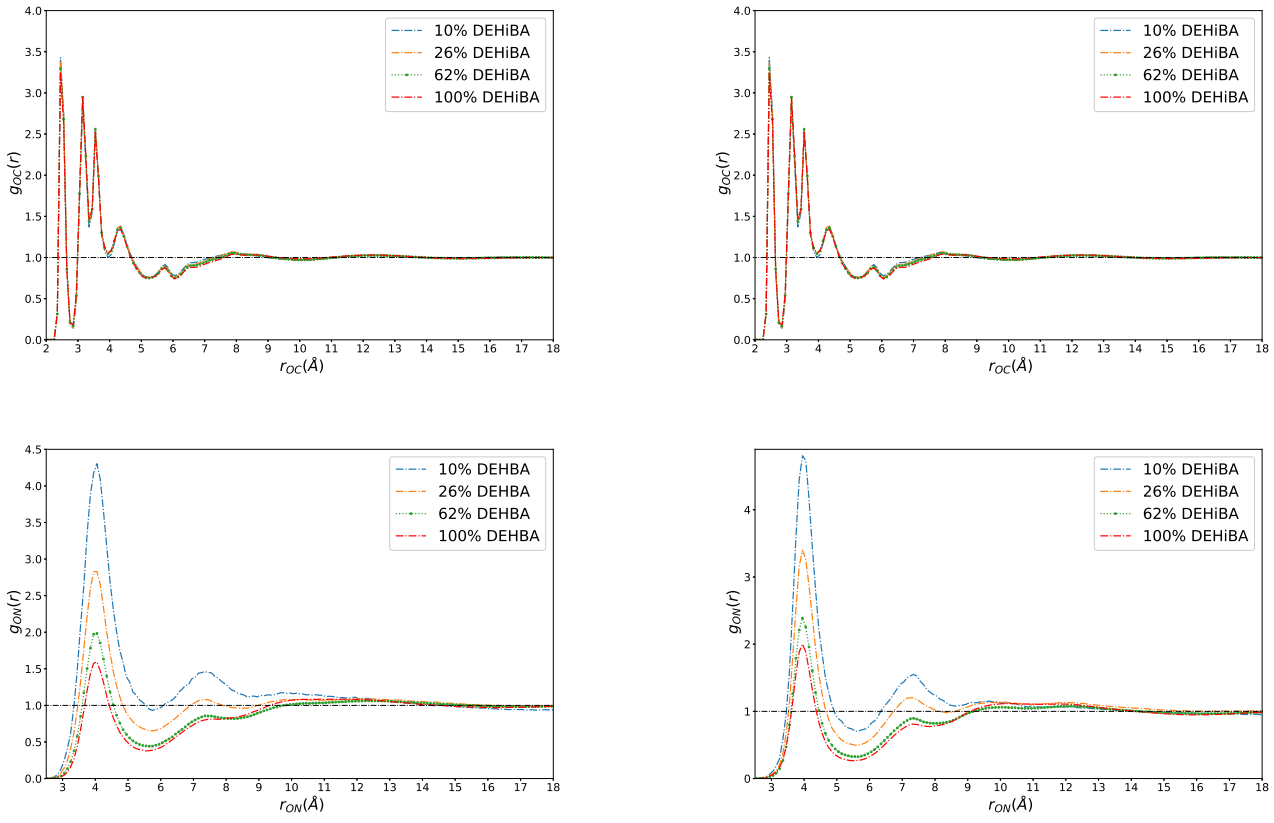






DEHBA/dodecane DEHiBA/dodecane

FIGURE B.8: Radial distribution functions of oxygen, carbon and nitrogen atoms in DEHBA/dodecane (left figures) and DEHiBA/dodecane (right figures) at 298 K.



DEHBA/dodecane DEHiBA/dodecane

FIGURE B.9: Radial distribution functions of oxygen, carbon and nitrogen atoms in DEHBA/dodecane (left figures) and DEHiBA/dodecane (right figures) at 298 K.