Tracking the degradation of organic chlorine compounds through reactive molecular dynamics

Matthieu Wolf

UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR

École doctorale des sciences exactes et leurs applications

Tracking the degradation of organic chlorine compounds through reactive molecular dynamics

THÈSE

présentée et soutenue publiquement le 08/12/2022

pour l'obtention du

Doctorat de l'Université de Pau et des Pays de l'Adour

(Spécialité : Chimie-Physique)

par

Matthieu Wolf

Composition du jury

Directeur :	Germain Salvato Vallverdu	Maître de conférences, IPREM-UPPA
Rapporteures :	Aurélie Perrier-Pineau	Maître de conférences, Université Paris Cité
	Carine Michel	Chargée de recherche CNRS, ENS Sciences Lyon
Examinateurs :	Jose Luis Sanchez Cebrian	Professeur des Universités, Université De Saragosse
	Pascal Brault	Directeur de recherche, Université d'Orléans
	Patrice Bordat	Maître de conférences, UPPA
	Dider Bégué	Professeur des Universités, IPREM-UPPA

Remerciements

Cher(è)s ami(e)s,

Je ne peux exprimer en mots suffisants ma gratitude pour chacun(e) d'entre vous qui ont rendu ma vie si riche et m'ont accompagné dans ce parcours académique jusqu'ici.

Je commence par remercier Germain pour avoir été le meilleur directeur de thèse que l'on puisse avoir. Notamment pour l'ensemble des choses qu'il m'a enseigné, que ce soit scientifiquement, pédagogiquement, ou encore Pythonesquement. Mais également pour tous les bons moments que nous avons passés ensemble au laboratoire et en dehors, les discussions riches, que ce soit sur mes idées, sur le volley, le rugby, la vie en général, la liste est longue… Et enfin pour ses conseils qui m'ont été précieux tout au long de ma thèse et continuent de l'être.

Je tiens à exprimer ma gratitude envers le laboratoire IPREM ainsi que son directeur, Ryszard Lobinski, pour m'avoir accueilli durant les trois années que j'ai passées à Pau. Je voudrais également remercier les membres du jury, José Luis Sanchez Cebrian, Pascal Brault (pour nos échanges enrichissants), Patrice Bordat et Didier Bégué. Je tiens à adresser un remerciement tout particulier à Aurélie Perrier-Pineau et Carine Michel, qui ont accepté de rapporter ma thèse.

Je poursuis en adressant mes sincères remerciements à toute l'équipe de chimie théorique, avec laquelle j'ai pu partager de précieux échanges et des pauses café stimulantes : Didier, William, Giovanna, Marc, Julia, Jeremy, Alexia, et Bhuvan. Je remercie également mon binôme de bureau, Carlos, avec qui j'ai trouvé le temps de procrastiner, discuter et passer de bons moments.

Je remercie également l'équipe CAPT, Panos, Sandra, Vincent, avec qui j'ai eu le plaisir de participer, de planifier et d'assister à des séminaires. Je tiens également à remercier Jef, avec qui j'ai commencé ma thèse et passé de bonnes soirées, Mathieu, mon partenaire d'entraînement de tennis, avec qui j'ai également passé des super soirées Trampol, Yann pour tous ses conseils et les agréables pauses déjeuner, et Danielle pour toute la positivité qu'elle apporte au laboratoire.

Un remerciement spécial à mon ami Gab, avec qui j'ai partagé trois années de thèse formidables, à faire des pauses café agréables, des soirées mémorables et des leçons de portugais améliorables ;-). Je remercie également tous les Bobby Lovers, la Team Barbeuk, etc., qui, pendant ces trois ans, ont été des acteurs majeurs dans le rechargement de mes batteries mentales, avec Parnian et Hélo et les formidables soirées que nous avons pu faire chez eux, ou encore les super festivals. Laure et Benjamin avec qui j'ai eu la chance de partir en Italie, en Corse, et bien sûr aux fêtes de Bayonne. Marine avec qui j'ai commencé ma thèse, et grâce à qui j'ai failli lancer une start-up sur le horse-ball. Camille avec qui j'ai pu explorer une grotte, ou encore voir le feu d'artifice de Biarritz. Céline et Manon qui sont venues mettre une ambiance d'enfer pour mon anniversaire. Et enfin, Matheus, pour les balades et les barbecues brésiliens !

Je profite de l'occasion pour glisser un mot de remerciement pour Tonks, pour tous ses accueils chaleureux, ainsi que pour Bobby (R.I.P) pour sa résilience.

Si je remonte un peu dans le temps, j'aimerais remercier mes amis d'enfance Simon et Cédric, qui ont toujours été à mes côtés et avec qui j'ai passé des moments inoubliables. J'adresse également mes remerciements à tous mes amis du lycée, Constant, Hugo[2], Mélanie, Jean, etc., ainsi qu'à Mme Degorce, sans qui je n'aurais jamais suivi le parcours universitaire que j'ai eu.

Je tiens à remercier tous les amis de la fac/PCCP, Chloé, Jean, Karim, Eliott, David, Guillaume, Simon, Luc, Jessica, Valentin, Manu, Benoit, etc. pour tous les fous rires, les belles rencontres et les bons moments passés en sirotant notre café noisette. Je tiens également à adresser un remerciement particulier à mes deux frères de PCCP, Maxou et Florian, en qui j'ai une confiance absolue, avec lesquels j'ai partagé tant de projets et d'heures de travail à l'espace Lavoisier, et je suis convaincu que nous créerons une super start-up !

Je remercie chaleureusement l'ensemble des enseignants du Master PCCP pour leur soutien, leur encadrement et leur enseignement passionné qui ont su éveiller en moi la passion pour la chimie physique et la recherche scientifique.

Je tiens également à remercier mes deux meilleurs amis, Jocelyn et Salomé, qui ont toujours été présents pour m'écouter, me soutenir, m'apporter des conseils quand j'en avais besoin et qui ont toujours répondu à l'appel de la fête.

Je tiens à remercier toute ma famille, en particulier Manou chez qui j'ai passé ma première année de thèse, pour toutes nos riches discussions et les bons moments passés ensemble.

Je remercie mes deux frères, Théo et Samuel, pour tous les moments de complicité que nous avons partagés depuis notre enfance, à tous nos fous rires, nos sessions jeux de sociétés, notre éclectitude (PS : j'ai inventé le mot je crois) musicale, notre temps passé à hacker l'ordinateur des parents, et pour tous les moments où vous avez été là pour moi quand il le fallait.

Je me devais de finir avec cette phrase pour qu'elle ne s'oublie pas : je remercie mes parents de m'avoir élevé jusqu'ici. Je suis fier de tout ce qu'ils ont pu m'apprendre socialement, scientifiquement, culturellement, etc., de tout ce qu'ils ont fait pour moi et je leur suis extrêmement reconnaissant de tout le soutien qu'ils m'ont apporté pendant ces trois années et depuis toujours.

Pour terminer, je tiens à remercier Marina, celle qui a partagé mon quotidien de thésard stressé en fin de thèse et qui est restée à mes côtés. Les moments passés ensemble à croquer la vie à pleines dents, que ce soit en voyage, en festival, en soirée, en match de volley, pendant toutes ces années, ont été exceptionnels. Je te remercie pour ton aide, tes conseils, ton soutien moral, ta bonne humeur et ton humour. Je t'aime.

Enfin, je souhaite remercier toutes les personnes que j'ai pu rencontrer au cours de ma thèse, même celles que je n'ai peut-être pas mentionnées dans cette liste. Votre soutien, votre collaboration et vos encourage-

ments ont été essentiels pour moi tout au long de ces années de recherche.

Merci infiniment à vous tous !

Contents

General Introduction **ix**

1 Force Field Review and Presentation of Molecular Dynamics **1**

1.1 Introduction . . . 1
1.2 History of Computer Simulations associated with molecular dynamics . . . 1
1.3 Force Field review . . . 3
1.3.1 Concepts of Force Field . . . 3
1.3.2 Classical Force Field . . . 4
1.3.3 Polarisable Force Field . . . 7
1.3.4 Coarse-grained Force Field . . . 8
1.3.5 Machine Learning Force Field . . . 8
1.4 Reactive Force field review . . . 13
1.4.1 History of Reactive Force Field (ReaxFF) . . . 13
1.4.2 Chlorine atom description in ReaxFF . . . 15
1.4.3 Reactive Force Field potential . . . 15
1.5 Introduction to Molecular Dynamics . . . 29
1.5.1 Thermodynamic ensembles . . . 31
1.5.2 Thermostat . . . 32
1.5.3 Initial configuration . . . 33
1.5.4 Boundary conditions . . . 33
1.5.5 Integration of Newton's equation . . . 34
1.5.6 Timestep . . . 37

2 Overview of organochlorides in an environmental context **39**

2.1 Introduction . . . 39
2.2 Presentation and history of PCBs, DDT and TCDD . . . 40
2.2.1 DDT . . . 40
2.2.2 TCDD . . . 41
2.2.3 PCB . . . 41
2.2.4 Toxicity of PCBs, DDT and TCDD . . . 45
2.3 Main transport routes of organochlorides . . . 46
2.4 Physical and chemical properties of PCBs . . . 49
2.5 Computational studies related to the reactivity of PCBs . . . 50

2.5.1 Biphenyl molecular properties . . . 51
2.5.2 PCB molecular properties . . . 52
2.6 Removal of organochlorines . . . 56
2.7 Conclusion . . . 57

3 Development of a new reactive force field 59
3.1 Introduction . . . 59
3.2 Training and Validation sets . . . 60
3.2.1 The choice of molecules . . . 60
3.2.2 PES scan and QM methods . . . 62
3.2.3 Creation of the Training set with all the files associated . . . 68
3.3 Optimization of the Force field . . . 70
3.3.1 Choice of the optimizer . . . 70
3.3.2 CMA-ES Initialization . . . 71
3.3.3 CMA-ES optimization of Reactive force field . . . 77
3.3.4 Optimization procedure . . . 81
3.4 Summary . . . 87
3.5 New ReaxFF force field performance . . . 88
3.6 Article 1 : Development of a novel ReaxFF reactive potential for organochloride molecules . . . 90
3.6.1 Introduction . . . 91
3.6.2 Computational methods . . . 92
3.6.3 Results and discussions . . . 96
3.6.4 Conclusion . . . 111
3.7 Conclusion . . . 118

4 Pyrolysis investigation using ReaxFF 119
4.1 Introduction . . . 119
4.2 Article 2 . . . 120
4.2.1 Introduction . . . 121
4.2.2 Computational methods . . . 122
4.2.3 Results and discussions . . . 125
4.2.4 Conclusion . . . 138
4.3 Conclusion . . . 141

5 Development of the new software SnookRMol 143
5.1 Introduction . . . 143
5.2 Article 3 . . . 144
5.2.1 Introduction . . . 145
5.2.2 Software description . . . 146
5.2.3 Computational details . . . 149
5.2.4 Results and discussions . . . 150
5.2.5 Conclusion . . . 158

5.3 Conclusion 159

General conclusion **161**

Appendixs **163**

A ReaxFF force field published **163**

B Removal of organochlorines **167**

List of Figures **171**

Bibliography **181**

General Introduction

This thesis presents a study of the degradation processes of organochloride compounds using theoretical chemistry methods and, in particular reactive molecular dynamics simulations. We developed a new reactive potential, including the chloride atom, adapted to organochloride compounds, and implemented reactive molecular simulations of the degradation of plastics and organochloride compounds of higher interest for environmental concern.

In theoretical chemistry, a simple view of the available methodologies leads to two main families: quantum chemistry and molecular mechanics calculations[1]. The two families allow computing the energy of molecular systems using either approximations or empirical potential.

The quantum chemistry calculations are based on the resolution of the Schrödinger equation providing the electronic states of the systems characterized by their wave functions and energies[2]. This resolution provides a complete description of the system by explicitly taking into account the electron density, which opens various fields of investigation, including electronic or optical electron transitions, electron transfer, chemical bonding analysis and chemical reactivity. Unfortunately, the accuracy of these methods is related to their computational cost, so obtaining qualitative results can take a long time or is done at the cost of a drastic simplification of the real system. In terms of the molecular system size, due to this limitation of time consumption, using methodologies such as density functional theory on high-performance computing centres allows quantum chemistry calculations to be applied to systems of maximum few hundred atoms.

On the other hand, breaking this limitation of molecular system size, molecular mechanics calculations, using empirical laws often based on analytic expressions, can compute the energies and forces of large systems[1]. These empirical laws or potentials are commonly named force fields and are parameterized from quantum calculations or experimental data. In this scope, the electrons are usually implicitly described as partial charges or dipoles and are sometimes neglected. With these empirical approximations, molecular mechanics can be efficiently implemented in molecular dynamics simulations to compute the trajectory of large molecular systems over time. With the technological improvement of computers over the last decades, using molecular mechanics, it is possible to run simulations with hundreds of thousands of atoms. This allows for increasing the complexity of the considered systems in terms of chemical diversity and thermodynamic

conditions.

The major thread of this thesis is the investigation of the reactivity of molecular systems in the gas phase and in condensed matter. When it comes to investigating a molecule's reactivity from theoretical chemistry, quantum chemistry calculations are usually a good starting point[1,3,4]. Quantum chemistry is indeed able to catch the electronic aspects of the bonds, and such calculations bring valuable information on the reaction mechanisms. On the contrary, using classical molecular mechanics calculations, the connectivity between atoms is fixed because it is a part of the force field definition, usually called the topology. Thus, it is impossible to investigate the chemical reactivity of molecular systems using molecular dynamics which implement classical force fields[1]. New types of force fields such as ReaxFF, published in 2001 by van Duin et al.[5], were designed to enable the simulation of dynamics bonds using molecular mechanics calculations by including the possibility for bond forming and breaking. These reactive force fields were created to bridge the gap between quantum and molecular mechanics calculations, allowing the investigation of the chemical reactivity of complex molecular systems from reactive molecular dynamics simulations.

From the reading of chapter I of this thesis, the reader will learn the different types of existing force fields, as well as a presentation of the molecular dynamics simulation method. Special attention on the ReaxFF force field will be made, and it will be presented that ReaxFF, since its first publication in 2001[5], has considerably evolved with many different applications. Figure 1 gives an overview of the diversity of the investigations implementing ReaxFF simulations. Notably, ReaxFF is divided into two major branches, the combustion and the aqueous, where the differentiation came from a good description of liquid water in the aqueous force field.

Now that the reader knows ReaxFF, chapter II presents the topics on which the new force field we developed will be applied. Therefore, since the beginning of the 1900s, a new type of compound family, the organochlorides, started to be used. Among them, we can find PCBs (polychlorinated biphenyls), massively produced worldwide with more than 1.325.810 tonnes[6], which are persistent, toxic and used in many different industrial fields. The TCDD ((2,3,7,8-tetrachlorodibenzo-p-dioxin), the most toxic molecule of these families, was used in the Vietnam war as a by-product of pesticides. Finally, the last compound presented in chapter II is DDT (dichlorodiphenyltrichloroethane), which was used as a pesticide and is still used for fighting the disease of malaria. Due to the environmental concerns with organochlorides, it has unfortunately become an important research field. That is why, due to their size and their complexity, it was chosen to develop a new ReaxFF force field to investigate their degradation in different environments.

Chapter III presents a complete description of the methodology used to develop the new reactive potential with constructing the training and validation sets from quantum mechanics calculations reference data. Then, the new force field is optimized to implement the chlorine atoms extending the very last version of the ReaxFF force field, allowing it to describe water and biomolecules adequately. After the entire optimization procedure

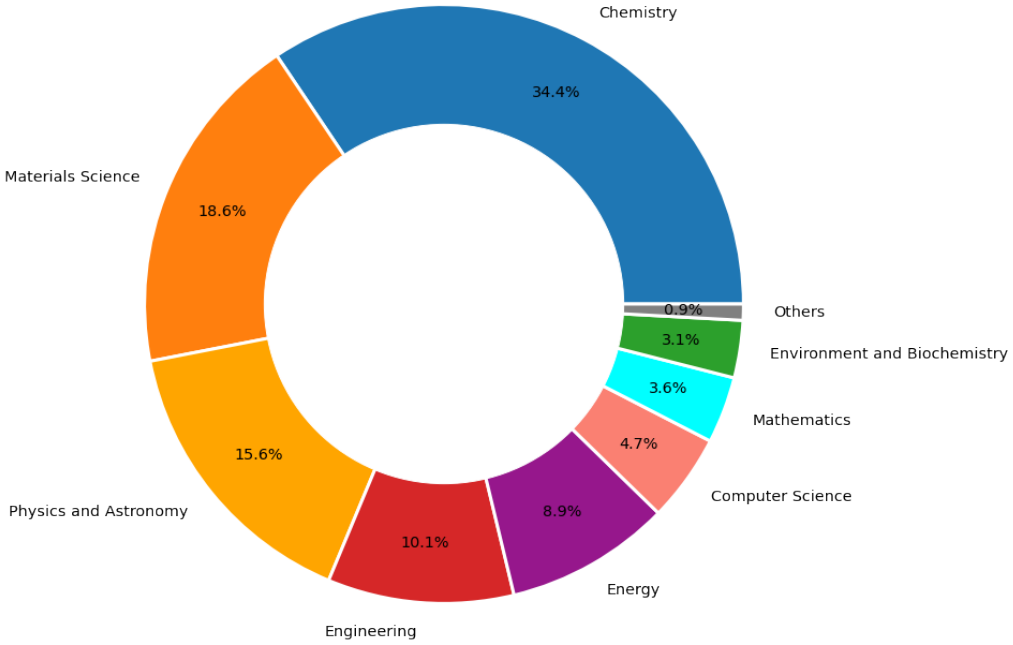


Figure 1: Representation of the subject area of ReaxFF publications, from 4594 articles.

depicted, it is also presented in this chapter the validation of the force field and its accuracy evaluation through the publication of the force field in the Journal of Chemical Physics[7].

Afterwards, Chapter IV focuses on plastic pyrolysis simulations and presents a performance comparison of the two major branches of ReaxFF potentials, namely, the aqueous and the combustion branches. This chapter uses the new aqueous ReaxFF force field we developed and compares its efficiency in describing the combustion process of plastics with additives against two other combustion ReaxFF potentials. The three polymers investigated are polyethylene (PE), polypropylene (PP) and polystyrene (PS), and the two additives used are chlorobenzene and toluene.

The last chapter presents SnookRMol, a novel Python-based framework that aims to probe the reactivity of molecules with simple reactants such as hydroxyl radicals. In the continuity of the thesis, the newly developed force field was used to investigate the reactivity of three organochlorides, DDT, TCDD and the PCB 169, as well as the worldwide used drugs diclofenac, with hydroxyl radical. The workflow of the software and the recommendation for usage are described in the article, and the results obtained using SnookRMol are compared with the literature.

Chapter 1

Force Field Review and Presentation of Molecular Dynamics

1.1 Introduction

This chapter starts with a historical presentation on the origin of computer simulation related to molecular dynamics. It leads to the presentation of the different molecular mechanics' force fields used in molecular dynamics to compute the total energy of a molecular system. First, the classical force field is introduced with three examples. Then, brief presentations of the polarisable and coarse-grained force fields are given. A more actual kind of force field is presented after, which is about the machine learning force fields with two examples. After the presentation of all these force fields, as most of the force fields presented before cannot track the reactivity, we present the reactive force field. These kinds of force fields are able to simulate bond breaking and forming. It starts with the history of ReaxFF (Reactive force field), followed by a complete presentation of the force field potential and the presence of the chlorine atom along ReaxFF. For tracking the reactivity of molecular reactions using ReaxFF, molecular dynamics is used. That is why the last section of this chapter is an introduction to molecular dynamics.

1.2 History of Computer Simulations associated with molecular dynamics

The story of computer simulation takes its root at the end of the second world war. Indeed, many new technologies were developed for war purposes. To cite a few examples, the Enigma machine was invented in 1919 for encrypting confidential information or the famous electromechanics machine built by Allan Turing to decrypt it (which started to work in 1943). The first electromechanical computer Zuse 3, was created in 1941 and was the first digital programmable computer used to compute complex matrices. It was followed by the Harvard Mark I in 1944. At this time, John Von Neumann, who was working on the Manhattan Project,

implemented a program in the Mark I to model the process of a nuclear detonation[8]. Nevertheless, the technology made a giant step forward in 1945, with the appearance of the fully electronic computers created for the nuclear arms race. The first one created was the ENIAC (Electronic Numerical Integrator And Computer) in the US, on which John Mauchly and Presper Eckert demonstrated in 1946 that thermonuclear weapon construction was viable. From the astonishing capability of the new electronic computer, the mathematician Stanislaw Ulam contacted John Von Newmann to discuss the use of statistical sampling together[9]. After several exchanges, John Von Newmann wrote in 1947 the first version of what we now call the Monte-Carlo algorithm to be used on the ENIAC[10]. Enrico Fermi originally invented this method in the early thirties, but it could not be applied then. One year later, Newmann and a physicist called Nicholas Metropolis, with their team, programmed the ENIAC to run the first Monte Carlo simulation of weapon core fission[11]. Thus, just like this, in April 1948, the story of computer simulation began, and the first computerized Monte Carlo simulations were made[11]. From that point, everything starts to accelerate in the computer simulation field. In our case, we will only follow the development of the computer simulation evolution related to molecular dynamics.

Some years later, in 1953, the Monte-Carlo algorithm was used by Nicholas Metropolis and its team to perform the first simulation of a liquid on the new electronic computer, the MANIAC I (Mathematical Analyzer Numerical Integrator and Automatic Computer Model I)[12].

During the same period, published in 1955[13], Fermi, Pasta and Ulam created the first force field to simulate the dynamics of an anharmonic one-dimensional crystal. One year later, the first force field for molecular dynamics, and so, the first molecular dynamics was published by Alder and Wainwright[14]. In this first MD simulation, Alder and Wainwright investigate the dynamics of an assembly of hard spheres. Almost fifty years later, in 2001, Berni Alder was rewarded with the Boltzmann medal for "inventing the technique of molecular dynamics simulation and showing that with such 'computer experiments' important discoveries in the field of statistical mechanics can be made, in particular, the melting/crystallization transition of hard spheres and the long-time decay of autocorrelation functions in fluids"[15]. Following the molecular dynamics of Alder et al.[15], the first real material (not using conceptual particle) MD was done in 1960 to investigate radiation damage[16] and the first MD of a real liquid was made in 1964[17]. Even if, in the beginning, scepticism existed about computer simulations and molecular Dynamics, Alder is not the only scientist to get rewarded for working in this field. Indeed, in 2013, Martin Karplus and his team were rewarded with the chemistry Nobel prize 2013 for the first molecular dynamics simulation on biomolecule, which came back in 1977[18].

1.3 Force Field review

As presented in the history of Computer simulation associated with molecular dynamics, the first force field dates back to 1955. Since the development of the molecular mechanics (MM) method, many different force fields have been developed, allowing the energy computation of molecular system. Originally designed to study small organic molecules and determine molecular structures and vibrational spectra[19]. The development of force fields was classified into four different groups. These four groups are the classical, the polarisable, the coarse-grained and the reactive force fields. We mainly focus on the classical force field and the reactive ones with some examples, and short reviews on the polarisable and the coarse-grained are presented.

1.3.1 Concepts of Force Field

A simple definition of a force field (FF) is a mathematical expression describing the potential energy of a system as a function of the system particles. The interatomic potential energy U(r_i) becomes an analytical equation with different parameters inside. The formalism of the potential comes from molecular mechanics, which signifies that inversely to quantum mechanics, the electrons are implicitly described, such as point charge or dipoles. The potential parameters are usually fitted from quantum mechanics computations or experimental data. Numerous force fields exist with different prospects and degrees of complexity. Nevertheless, the equation 1.1 gives the simplest view of a molecular FF[19].

$$\begin{aligned} U(r_i) = &\sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 + \sum_{torsions} \frac{V_n}{2}(1 + cos(n\phi - \delta)) \\ &+ \sum_{impropers} \frac{k_{imp}}{2}(1 + cos(2\omega - \pi)) + \sum_{nj} 4\epsilon_{nj} \left(\frac{\sigma_{nj}^{12}}{r_{nj}^{12}} - \frac{\sigma_{nj}^{6}}{r_{nj}^{6}} \right) + \sum_{elec} \frac{q_n q_j}{r_{nj}} \end{aligned} \tag{1.1}$$

The force field is divided into two kinds of energetical contributions, the intramolecular contributions represented by the first fourth terms of the equation 1.1 and the intermolecular with the last two.

- The first term of this equation corresponds to intramolecular contributions of the bonds stretching. In this simple case, the potential is harmonic controls the length of the covalent bond with reasonable values for the parameter r_0 the equilibrium length, r the bond length after stretching, and the spring constant k_b. However, a more realistic description of a covalent bond can be obtained with the Morse potential[20] defined by the potential of the equation 1.2.

$$V(r) = D_e(1 - e^{-a(r - r_0)})^2 \tag{1.2}$$

With D_e the dissociation energy of the corresponding bond.

- The second term, similarly to the bond stretching potential, represents a harmonic potential for the bending angle, which controls, for example, three atoms linked in a molecule, with θ_0 the angle of reference and k_a the spring constant.

- The dihedral bending potential is represented in the third potential, with ϕ the dihedral angle between four different atoms, δ the phase, n the number of minima and maxima between 0 and 2π and V_n the height of the potential barrier.

- The fourth is the improper bending potential which ensures the planarity of the system with ω the improper angle implying a deviation of planarity.

- The two last terms correspond to intermolecular interactions, firstly the Van der Waals interactions represented with the 12-6 Lennard Jones potential[21] in the equation 1.1, with ϵ the well depth and σ the particle diameter, and r the separation between the atoms. When $r \leq \sigma$, it induces a repulsion interaction preventing overlapping, and when $\sigma \leq r$, it creates an attractive interaction which tends rapidly to zero.

- Finally, the last term corresponds to the electrostatic interactions represented by a coulomb equation, with q the charge of the different atoms.

Obviously, the equation 1.1 presents an example of such potentials. After a quick review of the different kinds of force fields, some of the more used force fields are introduced with their potential.

1.3.2 Classical Force Field

This group of force fields is similar to the equation 1.1 in their conception. Indeed, they are composed of simple different potentials to describe different kinds of interactions inside the system. Some examples are given in the following section, with a presentation of the AMBER force field (Assisted Model Building with Energy Refinement)[22], UFF (Universal force field)[23], and CHARMM (Chemistry at HARvard Molecular Mechanics)[24].

But it also exists different other notorious force fields developed with the same idea but with different parameterization, like

- CVFF (Consistent valence force field)[25], its parameters were parametrized to represent amino acids, water and several other functional groups including small organic crystals and gas-phase geometries. It can represent well the potential of peptides and proteins (It is implemented in the Discover program).

- MMFF94 (Merck Molecular force field)[26] and its predecessor the MM3[27] were optimized through quantum mechanics computations. It was not developed to simulate one kind of molecule, even if it has great accuracy on "organic/protein" complexes energies computations, but to represent the process of different organic chemistry calculations.

- Compass (Condensed-phase optimized molecular potentials for atomistic simulation studies)[28] was parametrized

from a wide range of experimental observables. Moreover, It is the first force field using condensed-phase properties. So, it can predict properties, like structure, and conformation, in isolated and condensed phase conditions.

GROMOS (GROningen MAchine for Chemical Simulations)[29], similarly to CHARMM, shared a similar name with the software GROMACS in which it is implemented. It allows doing simulations of biomolecules with protein, nucleotide, and polymers. One of the specificities of the GROMOS FF is the united atoms (groups of atoms can be represented as united atoms).

- OPLS (Optimized Potentials for Liquid Simulations)[30] is extremely similar in the functional form of the potential as the one of AMBER. However, in this force field, two models are presented, one named united atom, where aliphatic hydrogens are not explicitly represented, which helps save simulation time. The other, entitled all-atom, where the hydrogens are explicitly described.

1.3.2.1 AMBER Force Field

One of the first FF developed to investigate macromolecules as DNA and protein is the AMBER (Assisted Model Building with Energy Refinement) [[22] force field developed by Peter Kollman. It was originally fitted from quantum mechanics computation and experimental data as neutron diffusions or RMN. It becomes a family of force fields. Even if the first published is one of the oldest classical FF, this family is still popular and has been updated numerous times. The force field functional form of the AMBER family is similar to the one we presented in the equation 1.1 with the potential presented in the equation 1.3.

$$\begin{aligned} U(r_i) = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 + \sum_{torsions} \frac{V_n}{2} (1 + cos(n\phi - \delta)) \\ + \sum_{j=1}^{i-1} \sum_{n=j+1}^{i} \left(\epsilon_{nj} \left[\frac{r_{0nj}^{12}}{r_{nj}^{12}} - 2(\frac{r_{0nj}^{6}}{r_{nj}^{6}}) \right] + \frac{q_n q_j}{4\pi\epsilon_0 r_{nj}} \right) \end{aligned} \tag{1.3}$$

With ϵ_0 the permittivity of free space, and r_{0nj} the distance at which the interaction energy between the two atoms is zero. It is an equivalent expression of the Lennard-Jones potential where $r_{0nj} = 2^{1/6}(\sigma)$.

1.3.2.2 UFF Force Field

The universal force field (UFF) was published in 1992[23]. The main idea of this new potential was to create a new force field to investigate almost all periodic table elements with an atomic number lower than 103. Indeed, using only the type of element, its hybridization and its connectivity, they provide parameters and formulas to have a reliable potential for the full periodic table.

$$U(r_i) = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} \frac{k_a}{n^2} [1 - cos(n\theta)] + \sum_{torsions} \frac{V_n}{2} [1 - cos(n\phi_0) cos(n\phi)] + \sum_{impropers} k_{imp} (C_n + C_{n+1} cos(\omega) + C_{n+2} cos(2\omega)) + \sum_{j=1}^{i-1} \sum_{n=j+1}^{i} \left(\epsilon_{nj} \left[\frac{r_{0nj}^{12}}{r_{nj}^{12}} - 2(\frac{r_{0nj}^{6}}{r_{nj}^{6}}) \right] + \frac{q_n q_j}{4\pi\epsilon_0 r_{nj}} \right) \quad (1.4)$$

Where C_n are coefficients of a cosine Fourier series chosen to satisfy physically justified boundary conditions, ϕ_0 is the equilibrium dihedral angle. The principle difference with the other force field is the expression of all the angle-related potentials through a small cosine Fourier series instead of the harmonic potential.

1.3.2.3 CHARMM Force Field

The CHARMM (Chemistry at HArvard Molecular Mechanics) force field shared its name with a macromolecular mechanics program[24]; it was optimized through ab initio energies and geometries of small organic models. It was initially designed to investigate biomolecules like peptides, protein, and nucleic acid. The force field is represented in the equation 1.5.

$$U(r_i) = \sum_{bonds} k_b (r - r_0)^2 + \sum_{angles} k_a (\theta - \theta_0)^2 + \sum_{torsions} k_\phi (1 + cos(n\phi - \delta)) + \sum_{impropers} k_\omega (\omega - \omega_0)^2 + \sum_{Urey-Bradley} k_u (u - u_0)^2 + \sum_{j=1}^{i-1} \sum_{n=j+1}^{i} \left(\epsilon_{nj} \left[\frac{r_{0nj}^{12}}{r_{nj}^{12}} - 2(\frac{r_{0nj}^{6}}{r_{nj}^{6}}) \right] + \frac{q_n q_j}{4\pi\epsilon_0 r_{nj}} \right) \quad (1.5)$$

The difference here is the use of a harmonic potential for the potential terms of improper bending and the Urey-Bradley potential with u, the distance between atoms 1,3. This potential is a cross-term that accounts for 1,3 non-bonded interactions not taken into account by the bond and bending angles terms.

Even if the classical force field has great advantages, such as simple analytical expressions, chemically intuitive, and efficient scaling, they cannot describe chemical reactivity as the connectivity is fixed. After a short review of the polarisable, coarse-grained and machine-learning force fields, it is presented the reactive force field created to investigate chemical reactivity.

1.3.3 Polarisable Force Field

Some properties that the classical force field does not take into account are the electrostatic polarisation and the inter and intramolecular charge transfers. Thus, another kind of force field appeared in 1976 with the work of Warshel et al.[31], named the polarisable force field. In this first work, they investigated enzymic reactions by considering an enzyme-substrate complex surrounded by water solvent. They used a dielectric model to investigate the solvation energy due to the solvent water molecules' polarization of the enzyme atoms. This model defines different regions, with a region for the water and another for the complex. In both regions, dipole-induced moments are computed in response to an electrical field and allow to increase in the accuracy of the enzyme-substrate interaction.

From the classical force fields presented above, AMBER was updated in 2001 into a polarisable force field using a simple isotropic polarization model. In this model, atomic point charges polarize other atoms and interact with their induced atomic dipole moments. Thus, the electrostatic field inducing the dipole moment on a particular atom is iteratively and self-consistently computed, and then the polarization energy is computed with the equation 1.6.

$$V_{polarization} = -\frac{1}{2}\sum^{i} \alpha_i E_i^0 . E_i \tag{1.6}$$

With α the polarisability, E_i the electric field on atom i due to charges and induced dipole and $E_i{}^0$ is the electrostatic field on atom i^{th} due to permanent atomic charges only.

Moreover, to represent the electrons in polarisable force fields, in addition with the atomic point charges, the atoms also have a slight dipole moment or are defined with variable charges.

CHARMM classical force field[32] was also updated with the fluctuating charge method to become polarisable, which is a model where the charge is redistributed among the atoms of a molecule to equalize the electronegativity. The electrostatic energy of an N-atom molecule in a vacuum is described by the equation 1.7.

$$V_{electrostatic} = \sum_{i=1}^{N} \chi_i Q_i + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \eta_{ij} Q_i Q_j + \sum_{i=1}^{N} \Phi_i Q_i \tag{1.7}$$

With Φ an external electrostatic potential at site i, Q_i the charges, η_{ij} the hardness and χ_i the electronegativity.

Finally, the last polarizable force field discussed here is the AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications) FF[33]. Initially designed to investigate water, it gets updated, and it can now simulate organic molecules and proteins with great interaction energy description[34 35]. The functional

form of the AMOEBA force field is represented in the equation 1.8.

$$U_{r_i} = U_{bonds} + U_{angles} + U_{torsions} + U_{improper} + U_{bond-angle} + U_{vdW} + U_{ele}^{perm} + U_{ele}^{ind} \tag{1.8}$$

It has some differences with the classical force field potentials with a bond-angle cross-term and a 14-7 Lennard Jones potential for the Van der Waals interaction. However, the major difference is that instead of the fixed charge atom model, they introduce a combination of fixed and polarizable multipoles. Thus, the electrostatic energy is computed with permanent and induced multipoles, which allows a good description of the molecular electrostatic potential.

1.3.4 Coarse-grained Force Field

Another type of FF are the coarse-grained force fields. The simulation of large systems, like biological systems, is extremely time-consuming due to the huge number of atoms. These force fields allow simplifying the system representing atoms groups by unique spheres or beads. Thus, the number of interactions is radically reduced. So, the choice of the extracted properties is essential to avoid the risk of an oversimplification of the system. One of the first CGFFs was published by Levitt et al.[36], in their work, they succeeded in simulating a folding process of a protein. To do so, they replaced the main and lateral chains of the protein structure with chains of pseudoatoms and used statistical sampling to arrange them. This simulation shows that the main forces at the origin of the folded structure were the packing and the pairwise interactions of the side chains. Nowadays, it exists numerous different CGFFs, and one of the most famous potentials is called the MARTINI[37], which was initially developed in 2004 to study lipids, but it evolved to simulate diverse biomolecules such as proteins and DNA. The conceptual idea of this FF is the use of four main concepts: charge, polar, nonpolar, and apolar, which are also divided into different types. Then different kinds of interactions are used between the different types of beads.

1.3.5 Machine Learning Force Field

More recently, a new type of FF has appeared, taking benefits from the development of machine learning, entitled machine learning force field (MLFF), which is greatly reviewed by Unke et al.[38]. In this concept, the main idea was to close the gap between ab initio methods and the efficiency of the classical force fields. In contrast to the other force fields, it does not use a functional form. Indeed, they are built from reference data through learning algorithms. Thus, one of the major advantages of these FFs is that there is no need to describe chemical properties because they are learned from the reference data. The associated drawback is the limited transferability of these force fields[39]. These force fields were successfully used to investigate for example atomization energies[40], predict unknown spin states[41], predict thermodynamics properties[42],

determine electronic ground and excited-state properties[43], study reactive molecular collisions[44] or follow protein folding[45]. It is important to notice that plenty of MLFFs are in development, continuously bringing new possibilities. One key point in this kind of FF is that the necessary chemical properties used to train the MLFF should be carefully made and selected[38]. Usually, to have accurate training data, high-level ab initio calculations are used to create them[38]. The main idea of these force fields is that they learn the statistical relation between chemical structure and potential energy from these data without the need for predefined connectivity or interactions. Depending on the force field objectives, they require at least 200 chemistry data in the training set to create the force field. Nevertheless, it can go up to 1000 data for complex systems[38]. However, a recent new model used to train ML-FF named Bravais-Inspired Gradient-Domain Machine Learning BIGMDL[46] was developed. It produces reliable MLFF with a relatively small training set (between 10 and 200 data).

These force fields are usually based on two different learning methods, Kernel-Based and Artificial Neural Networks, which are presented in the following sections.

1.3.5.1 Kernel-Based Method

This method uses a linear classifier to solve a non-linear problem, such as energy predictions from nuclear positions. To achieve this, it transforms the initial system of non-linear data, like the structure and the associated energy, into a higher dimension to obtain linear separable data. Then, it uses the so-called 'kernel trick' to avoid the heavy computation of all the new coordinates in this feature space, and it computes only in this space all the inner-product of the images of all pairs of data. In order to give a better understanding of this principle, a representation is presented in figure 1.1.

With N references values y_i and the associated inputs of the kernel x_i, this method aims to estimate y_{new} from unknown x_{new}. In the case of MLFF, if we take the example of a force field investigating potential energy surfaces (PES), y_i can be potential energy and x_i chemical and structural information, as positions charges[38].

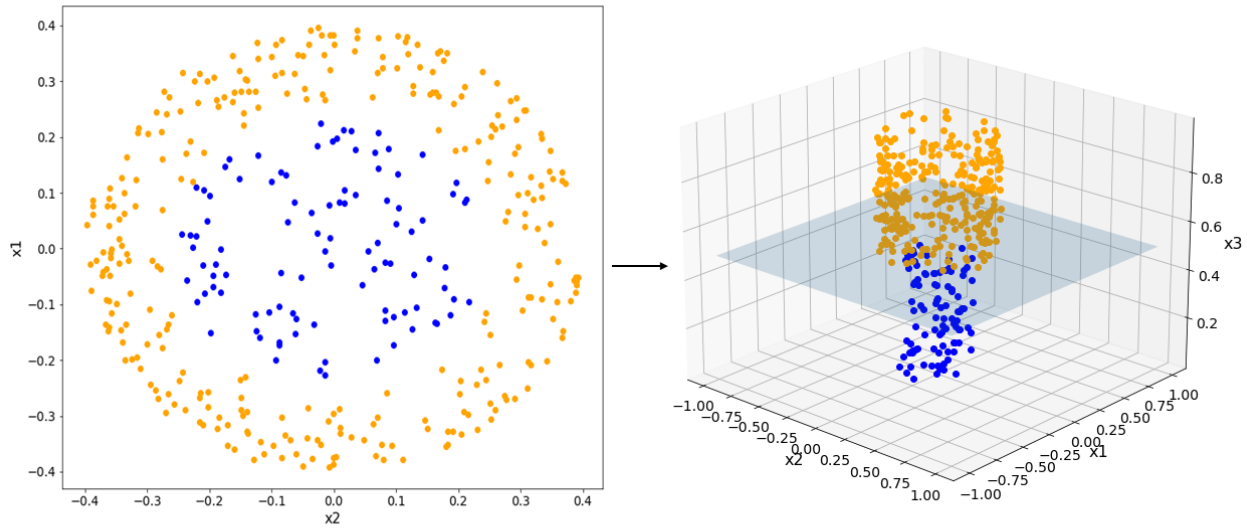
The formalism of this method, is represented in the equation 1.9 1.10 and1.11[38].

$$y = f(x) + \epsilon \tag{1.9}$$

With ϵ the noise measurement, and $f(x)$ which can be approximate as a linear combination presented in the equation 1.10.

$$f(x) \approx \sum_{i=1}^{N} \alpha_i K(x_i, x') \tag{1.10}$$

With α_i being a coefficient, and K the similarity kernel function, that allows computing the inner prod-



a) Linearly non-separable b) Linearly separable

Figure 1.1: In this schematization in figure a), the blue and yellow points are inseparable in this two-dimensional space (x_1, x_2). However, it can be transformed into a higher dimensional feature space thanks to a suitable kernel function represented by the black arrow. Thus it leads to a linearly separable space (x_1, x_2, x_3) presented in figure b) where the yellow points and blue points can be linearly separated by a plane.

uct between the new x' and x_i from the training set to measure the similarity of the two inputs, and so to predict the new x'. It exists different type of kernel function, like the Polynomial kernel[47], Fisher kernel[48], Convolution kernel[49] and Gaussian kernel . Which are different in terms of complexity and usage.

Thus, with the example of the Polynomial kernel, the function is define with the equation :

$$K(x_i, x') = (\langle x_i, x' \rangle + c)^d \tag{1.11}$$

With $\langle x_i, x' \rangle$ the inner product, d a hyperparameter defining the polynomial degree, and c a free-parameter trading of the influence of lower and higher degree polynomials terms.

1.3.5.2 Artificial Neural Networks

The second method presented here is the Artificial Neural Networks method. This concept used artificial neurons (or nodes) to mimic biological neurons network[50] to predict new output from a set of inputs.

Thus, we propose first presenting the most simple case where there is only one artificial neuron in the network, also called a logistic regression depicted in the figure 1.2. In this case, we have a set of observations $X \in \mathbb{R}^i$ represented by the vector $(x_1, \ldots, x_i)$ from which we want to predict $Y \in \mathbb{R}^j$ the output vector $(y_1, \ldots, y_j)$ using the equation 1.12.

$$Y = WX + b \tag{1.12}$$

With W the synaptic weight vector $\in \mathbb{R}^{i \times j}$, and the biases b $\in \mathbb{R}^{j}$.

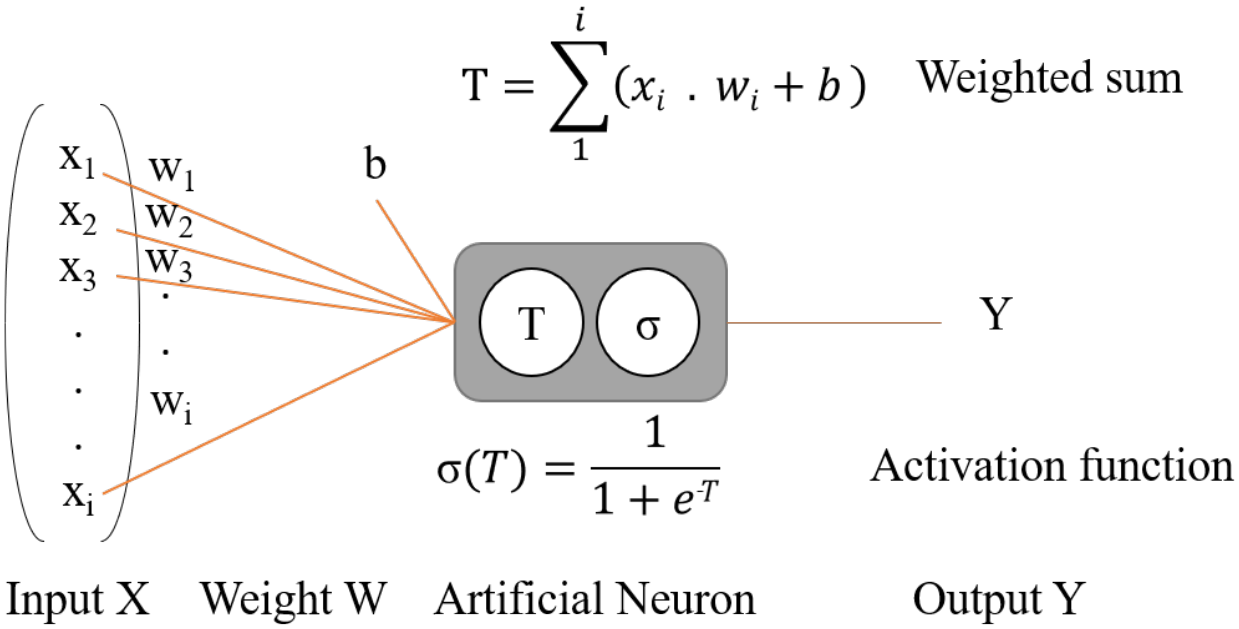


Figure 1.2: In this schematization of a simple neural network with one neuron, it is represented the different layer with the input X, the weight associated W, the biases term b, the artifical neuron composed of the weighted sum and the activation function (an example of activation function is represented here), and finaly the output predicted Y.

From the initial vector input X, a weighted sum T is computed using the synaptic weights w_i and the bias term b. Then, the weighted sum T goes through an activation function σ, which allows to accept or refuse the predicted output Y.

In the case of the artificial neuron network, it is similar to the one neuron case. However, it exists several layers including several artificial neurons. Thus we have the input layers with X, followed by Z hidden layers of variable dimension of artificial neurons leading to the final Y output layer. A representation is presented in the figure 1.3.

The equation 1.12, becomes now the iterative equation 1.13 :

$$\begin{aligned} h_1 &= \sigma(W_1 X + b_1) \\ h_2 &= \sigma(W_2 h_1 + b_2) \\ &\dots \\ h_Z &= \sigma(W_Z h_{Z-1} + b_Z) \\ Y &= W_{Z+1} h_Z + b_{Z+1} \end{aligned} \tag{1.13}$$

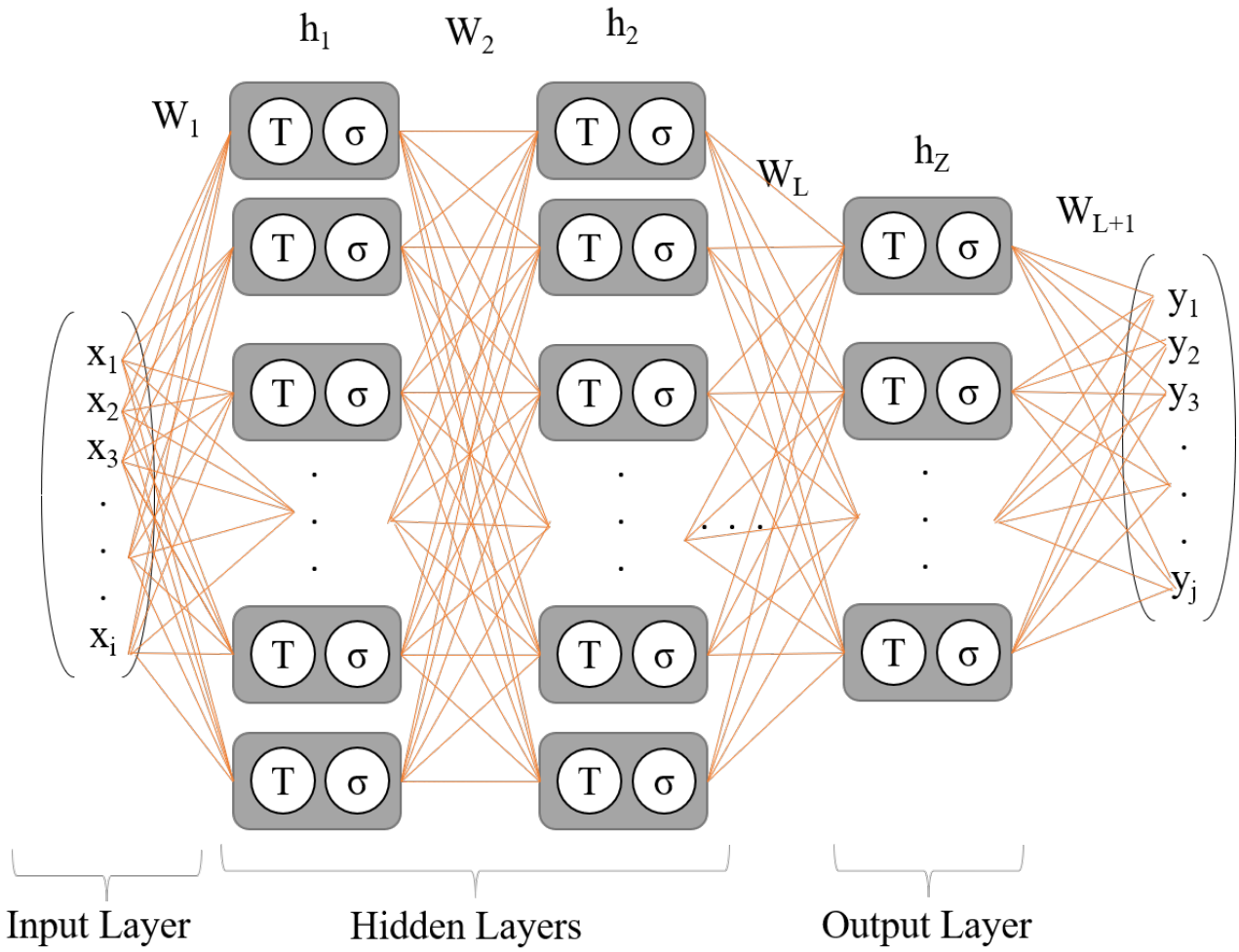


Figure 1.3: Representation of a neural network, with the initial input X being the initial layer, followed by several Hidden Layers, which the last one being the output layer predicting the output Y. The neurons are indicated in grey, similar to the figure 1.2 with the layer denoted as h_i, the synaptic connections in orange with the weight vector W associated.

Finally, we can compare the predicted output Y to the reference output using a loss function. Then, by applying a gradient minimization to this loss function, we can do a back-propagation procedure and update all the weights and biases terms. With these new terms, new outputs can be predicted, representing how the artificial neural networks are constructed to predict outputs.

An example of a force field developed with this method is, for example, the force field of Yoo et al.[51], who published an MLFF able to investigate the decomposition and reaction of the high-energy nitramine 1,3,5-trinitroperhydro-1,3,5-triazine (RDX). It is indeed able to investigate the reactivity of these specific molecules without requiring connectivity. It allows the introduction of the last kind of force field, the reactive force field, which is presented in the next section.

1.4 Reactive Force field review

The reactive force field is the last type of FF presented in this chapter. Indeed, with most of the force fields presented before, breaking or creating a bond is impossible. That is why these force fields appeared, with the possibility of simulating chemical reactions. In the following parts, it is presented a review of the ReaxFF force field. Before, ReaxFF, reactive force fields based on the cut-off method based on distance criteria[52] were used. For example, they were used to investigate polymerization processes. However, this method does not include dynamic bonds, which is essential for a good description of reactivity.

1.4.1 History of Reactive Force Field (ReaxFF)

In 2001 Adri van Duin, William A. Goddard, III, and co-workers published a new reactive force field named ReaxFF[53]. Initially designed to investigate hydrocarbons, they developed a new empirical potential for studying reactivity through molecular dynamics at a low computational cost. They introduced the concept of bond order in their intramolecular potentials, which permit bond breaking and creating. Moreover, the force field is also based on the unique atom type principle, which means that there is only one set of parameters to describe one element; for the carbon atom, there are no differences if the carbon is sp2 or sp3, the parameterisation and the functional form of the force field include all carbon environments. It did not take long before new versions of this potential appeared, with a new force field extending the force field of hydrocarbons in 2003 to investigate Si and SiO_2 system reactivity[54]. It was followed in 2005 by a new ReaxFF potential developed to investigate thermal-induced reaction in RDX[55]. Then, everything started to accelerate in 2008, with the famous C/H/O force field published by Chenoweth et al.[56], which was trained against more than 1600 chemical data. This publication presents the best current description of the ReaxFF potential, which we will use in the next part, along with the different extensions.

Moreover, Chenoweth et al.[56] optimized this force field using the full training set of carbon and hydrogen from the 2001 force field, completing it with the addition of oxygen. This force field is mainly famous for its reliability in computing high-temperature reactions such as pyrolysis. It is at the origin of many other force fields, which are classified in the combustion branches of ReaxFF.

Indeed, an important aspect of the ReaxFF development is the appearance of two main branches during its development. The combustion branch appears with the RDX force field of 2005. However, its denomination really appears with the C/H/O force field of 2008. It comes from the great accuracy of the C/H/O ReaxFF to describe water in the gas phase but lacks a description of the liquid phase. Hence, the combustion branch is composed of force fields with oxygen and hydrogen parameters that describe well-water molecules in the gas phase. In contrast, aqueous ReaxFF that appeared in 2008[57], describes water molecules accurately in the liquid phase. The evolution of these two branches is fully detailed in figure 1.6 which is inspired by the ReaxFF

review of Senftle et al.[58].

Nevertheless, inside each of the two branches, they are force field subgroups that do not have the same parameterisation of oxygen and hydrogen.

For example, at the beginning of the combustion branch, one subgroup was created following the RDX force field with the C/H/O of 2008, with its development presented in the figure 1.4, and another, made from the Si-O force field of 2003 leading to the 2010 Si/O/H. In this branch, we can find different prospects for these force fields, like studying hydrocarbon oxidation[56], following the combustion of ammonia borane dehydrogenation process[59], investigating the decomposition of Explosive α-HMX[60], modelisation of solid oxide fuel cell[61], evaluating the carbonation of polymers[62].

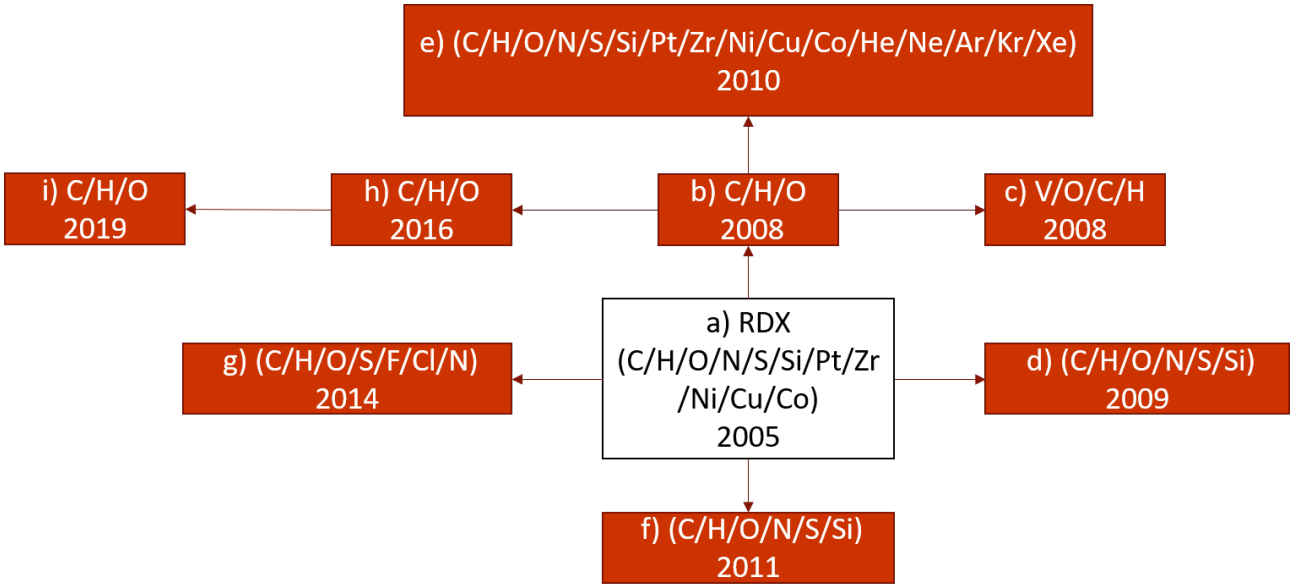


Figure 1.4: Here it is represented a schematization of a combustion subgroup ReaxFF development from the RDX ReaxFF of 2005. Each box represents the year of publication with the element of the force field published and the letters for references with a[55], b[56], c[63], d[64], e[65], f[66], g[60], h[67], i[62].

On the aqueous side, a great example of subgroups is, for example, the CHON-2017_weak force field[68] which has one of the best descriptions of the water density and the weak interaction associated with it, and on the other hand, the force field of Zhang 2017[69] which described the acid-base reactivity accurately in aqueous solution with different pH. As well as for the combustion, the subgroups of the aqueous branch have different prospects like, investigate biomolecules with the example of the subgroups of the figure figure 1.5, DNA-damage by hydroxyl radicals[70], proteins conformation[71], polymers conformation[72], or investigating different surfaces Zn oxides[57], Fe[73,74], Si[75], Al[76], Ni[77], and Ti[78] but also to investigate Zeolite reactions[79].

Finally, before going into the details on the functions of the ReaxFF potential, it is important to notice that it also exists plenty of force fields which are not part of none of the two main branches but are dedicated to specific cases. Indeed, they were not developed to investigate water in any case. In the different specific cases, for example, it exists ReaxFF potentials to investigate different interfaces with Al[81], Pt[82], Ni[83], or cluster

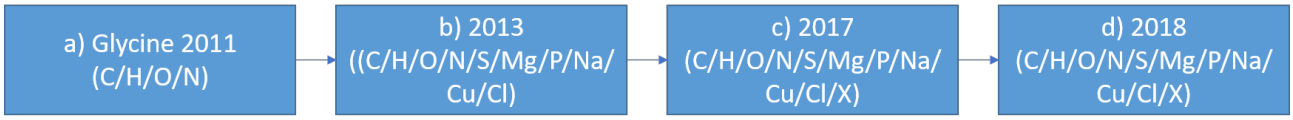


Figure 1.5: Here it is represented a schematization of an aqueous subgroup for biomolecules investigation, developed from the glycine ReaxFF of 2011[80]. Each box represents the year of publication with the element of the force field published and the letters for references with a)[80], b) named protein-2013[71], c) entitled CHON_2017_weak[68], and d) the CHON_2017_weak_bb[72].

fragmentation and desorption of molecular hydrogen with Mg[84], Na[85], Pd[86] or different nanoparticles as TiO_2[87], gold nanoparticles[88].

We can see from the figure 1.6t hat the field of applications of ReaxFF is extremely wild. Thus, it is necessary to have a tunable potential. This flexibility is achieved through many empirical parameters presented in the next part and usually denoted as p_i (for partial energy contribution). Concerning the transferability of these parameters between the different branches, Senftle et al.[58] indicate in their review that they can be transferred inside the aqueous branch and the combustion branch. However, depending on the origin of the two merged force fields, if both are in the same branch and have the general parameters in common, they may probably merge just fine. However, two force fields of the same branch but with different general parameters will likely result in bad efficiency. In any case, merging two force fields should be done carefully; usually, reoptimizing the force field may give better results than merging.

1.4.2 Chlorine atom description in ReaxFF

In this work, we are interested in investigating the reactivity of organochlorides in diverse environments and, thus, by force fields describing the chlorine atoms correctly. In the existing ReaxFFs, some of them have parameterized the chlorine atoms, such as all the force fields in the aqueous branch presented in figure 1.5. However, it was used as a counter ion and could not simulate chlorinated molecules. It was also parameterized in 2014 in the combustion force field[60] g) of figure 1.4, with parameters for the carbon chlorine bond but not for the bond with other elements. So, no real ReaxFF was available to investigate the degradation of organochlorides, that is why, as it is presented in chapter III of the thesis, we optimize a new aqueous force field, continuing the evolution of figure 1.5.

1.4.3 Reactive Force Field potential

This section presents the full ReaxFF potential from 2005 formalism implemented in different software such as LAMMPS[89] and ADF[90], as well as the different variations that have appeared since 2008. A complete description of the force field was given by Chenoweth et al.[56]in 2008, and this section takes its root from this

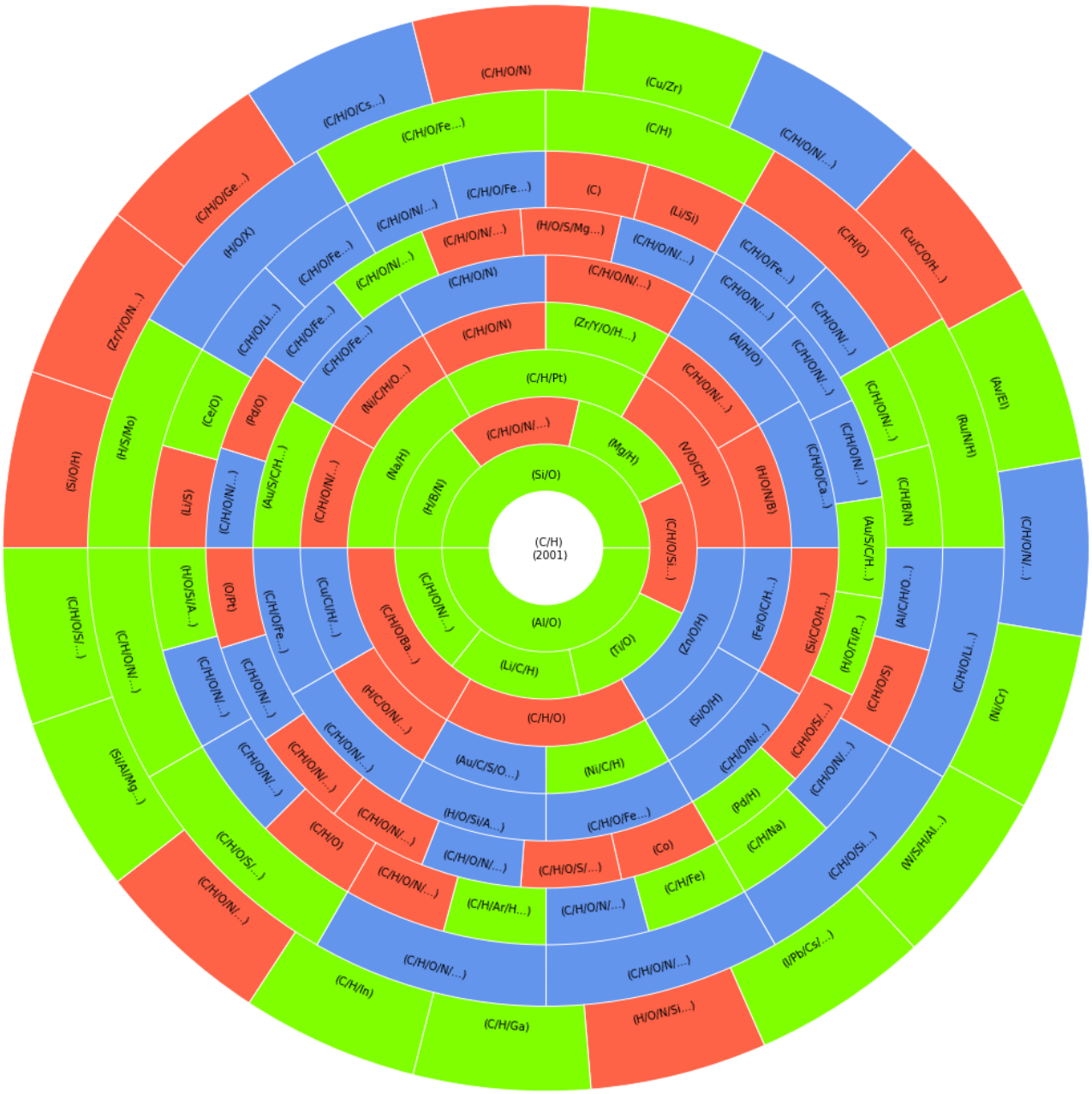


Figure 1.6: Here it is represented a schematization of the ReaxFF development from the 2001 initial ReaxFF. At the centre of the chart is represented the initial force field of 2001, then each extension of the circle corresponds to two more years, starting with 2003/2004, followed by 2005/2006..., except the last one which contains ReaxFF from 2019, 2020 and 2021. The ReaxFF in red parts belong to the combustion branch, those in blue to the aqueous branch, and in green, they are independent (usually, the O/H parameters are not included). The 116 ReaxFF references depicted here are available in the annexe.

article.

1.4.3.1 ReaxFF total energy formalism

The equation 1.14 represents the potential to compute the total energy of the system with all the different energetical terms.

$$E_{system} = \underline{E_{bond}} + \underline{E_{lp}} + \underline{E_{over}} + \underline{E_{under}} + \underline{E_{val}} + \underline{E_{pen}} + \underline{E_{coa}} + \underline{E_{C2}} + \underline{E_{triple}} + \underline{E_{tors}} + \underline{E_{conj}} + \underline{E_{Hbond}} + E_{vdW} + E_{Coulomb} \quad (1.14)$$

All the energetical terms are presented in the following. The underlined terms are those using the bond order to compute the energy. Indeed, as we mentioned above, ReaxFF introduces the bond order concept in its potential, which allows for creating and breaking of bonds. Its formalism is directly presented in the next part.

1.4.3.2 Bond order

The bond order BO'_{ij} between a pair of atom, is computed from the interatomic distance r_{ij} as it is shown in the equation 1.15.

$$BO'_{ij} = exp\left[p_{bo,1}.\left(\frac{r_{ij}}{r_0^{\sigma}}\right)^{p_{bo,2}}\right] + exp\left[p_{bo,3}.\left(\frac{r_{ij}}{r_0^{\pi}}\right)^{p_{bo,4}}\right] + exp\left[p_{bo,5}.\left(\frac{r_{ij}}{r_0^{\pi\pi}}\right)^{p_{bo,6}}\right] \quad (1.15)$$

With the first term being the contribution of the sigma bonds, the second pi-bonds and the last one the double pie-bonds. The p_i are complete empirical parameters, and r_0^x are parameters defined.

In order to get a better understanding of the bond order in ReaxFF, it is presented in the figure 1.7 the evolution of the bond order of the carbon-carbon bond in the C/H/O force field from 2008.

From this BO', a correction is applied to obtain the correct BO. Indeed, from the uncorrected BO', the over-coordinatation is computed with the equation 1.16.

$$\Delta'_i = -Val_i + \sum_{j=1}^{neighbours(i)} BO'_{ij} \quad (1.16)$$

With the second term being the total BO' of the atom around and Val_i being the bonding electrons of the atom (equal to 4 for carbon and 2 for oxygen). This correction can compute the corrected BO with the equations 1.17, 1.18, 1.19, 1.20, 1.21, 1.22 and 1.23.

$$BO_{ij} = BO_{ij}^{\sigma} + BO_{ij}^{\pi} + BO_{ij}^{\pi\pi} \quad (1.17)$$

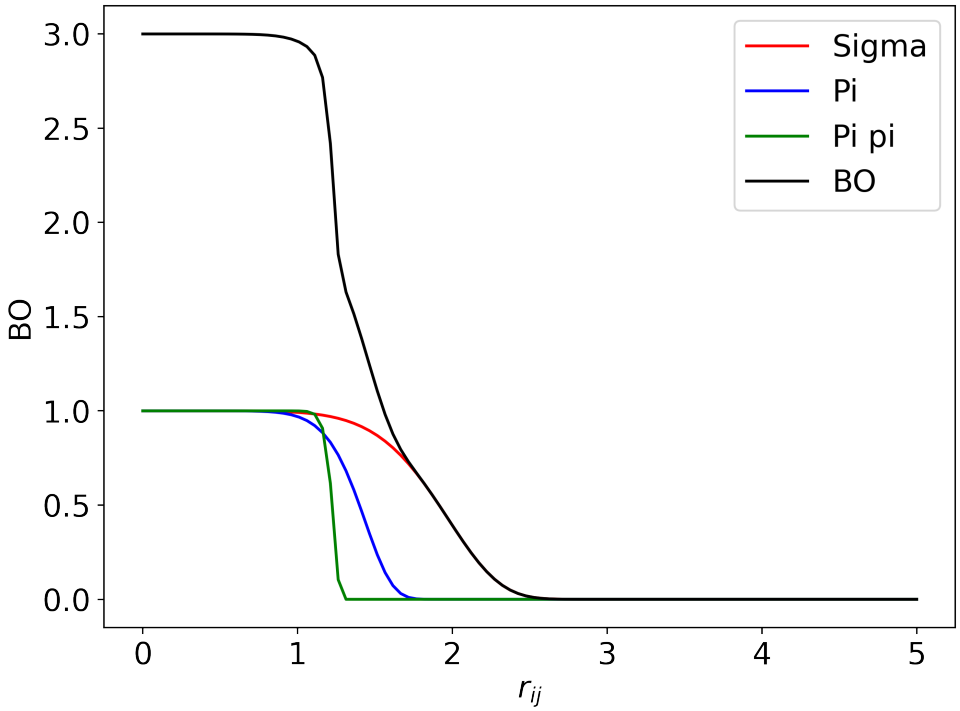


Figure 1.7: Here, the evolution of the BO' is presented as a function of the interatomic distance r_{ij}. In Black, it is the BO', in red the sigma contribution, in blue the pie contribution and in green the piepie contribution.

$$\begin{aligned} BO_{ij}^{\sigma} &= BO_{ij}^{'\sigma}.f_1(\Delta_i^{'},\Delta_j^{'}).f_4(\Delta_i^{'},BO_{ij}^{'}).f_5(\Delta_j^{'},BO_{ij}^{'}) \\ BO_{ij}^{\pi} &= BO_{ij}^{'\pi}.f_1(\Delta_i^{'},\Delta_j^{'}).f_1(\Delta_i^{'},\Delta_j^{'}).f_4(\Delta_i^{'},BO_{ij}^{'}).f_5(\Delta_j^{'},BO_{ij}^{'}) \\ BO_{ij}^{\pi\pi} &= BO_{ij}^{'\pi\pi}.f_1(\Delta_i^{'},\Delta_j^{'}).f_1(\Delta_i^{'},\Delta_j^{'}).f_4(\Delta_i^{'},BO_{ij}^{'}).f_5(\Delta_j^{'},BO_{ij}^{'}) \end{aligned} \tag{1.18}$$

$$f_1(\Delta_i^{'},\Delta_j^{'}) = \frac{1}{2}.\left(\frac{Val_i + f_2(\Delta_i^{'},\Delta_j^{'})}{Val_i + f_2(\Delta_i^{'},\Delta_j^{'}) + f_3(\Delta_i^{'},\Delta_j^{'})} + \frac{Val_j + f_2(\Delta_i^{'},\Delta_j^{'})}{Val_j + f_2(\Delta_i^{'},\Delta_j^{'}) + f_3(\Delta_i^{'},\Delta_j^{'})}\right) \tag{1.19}$$

$$f_2(\Delta_i^{'},\Delta_j^{'}) = exp(-p_{boc1}.\Delta_i^{'}) + exp(-p_{boc1}.\Delta_j^{'}) \tag{1.20}$$

$$f_3(\Delta_i^{'},\Delta_j^{'}) = -\frac{1}{p_{boc2}}.ln\left[\frac{1}{2}.\left(exp(-p_{boc2}.\Delta_i^{'}) + exp(-p_{boc2}.\Delta_j^{'})\right)\right] \tag{1.21}$$

$$f_4(\Delta_i^{'},BO_{ij}^{'}) = \frac{1}{1 + exp(-p_{boc3}.(p_{boc4}.BO_{ij}^{'}.BO_{ij}^{'} - \Delta_i^{'boc}) + p_{boc5}} \tag{1.22}$$

$$f_5(\Delta_i^{'}, BO_{ij}^{'}) = \frac{1}{1 + exp(-p_{boc3}.(p_{boc4}.BO_{ij}^{'}.BO_{ij}^{'} - \Delta_j^{'boc}) + p_{boc5}} \tag{1.23}$$

With $\Delta_j^{'boc}$, another overcoodination correction is created for atoms with lone electron pairs, like nitrogen and oxygen, which help break these electron pairs and create bonds. The formula for this overcoordination is described in the equation 1.24.

$$\Delta_i^{'boc} = -Val_i^{boc} + \sum_{j=1}^{neighbours(i)} BO_{ij}^{'} \tag{1.24}$$

With Val_i^{boc}, the number of electrons associated with this valency equals four for oxygen and four for carbon.

1.4.3.3 Bond Energy

This term is define by E_{bond} in the equation 1.14, it is computed using the different contribution σ, π and $\pi\pi$ of the BO, with the dissociation energies parameters associated with these BO types, D_e^{σ}, D_e^{π} and $\text{D}_e^{\pi\pi}$. The equation of the bond energy is presented in the equation 1.25.

$$E_{bond} = -D_e^{\sigma}.BO_{ij}^{\sigma}.exp\left[p_{be1}\left(1 - (BO_{ij}^{\sigma})^{p_{be2}}\right)\right] - D_e^{\pi}.BO_{ij}^{\pi} - D_e^{\pi\pi}.BO_{ij}^{\pi\pi} \tag{1.25}$$

1.4.3.4 Lone pair energy

In the equation 1.14 it referees to E_{lp}. This penalty energy is due to the deviation of the computed number of lone pairs from the optimal number. It uses the valence electrons Val_i^e and the bond order. Its equation is presented in the equation 1.26.

$$E_{lp} = \frac{p_{lp2}.\Delta_i^{lp}}{1 + exp(-75.\Delta_i^{lp})} \tag{1.26}$$

With Δ_i^{lp} of the equation 1.27, being the difference between the optimal number of lone pairs $\text{n}_{lp,opt}$ (equal to 2 for oxygen and 0 for hydrogen) and the computed number of lone pairs around an atom $\text{n}_{lp,i}$, the last term calculated with the equation 1.28.

$$\Delta_i^{lp} = n_{lp,opt} - n_{lp,i} \tag{1.27}$$

$$n_{lp,i} = int\left(\frac{\Delta_i^e}{2}\right) + exp\left[-p_{lp1}.\left(2 + \Delta_i^e - 2.int\left(\frac{\Delta_i^e}{2}\right)\right)^2\right] \tag{1.28}$$

With Δ_i^e being the difference between the valence electrons and the total BO of an atom, defined by the equation 1.29.

$$\Delta_i^e = -Val_i^e + \sum_{j=1}^{neighbours(i)} BO_{ij} \tag{1.29}$$

It uses the valence electron valence electrons Val_i^e, which equals 6 for Oxygen. From these equations, if an oxygen atom has normal coordination with a BO = 2, Δ_{ij}^e will be equal to 4, and $n_{lp,i}$ = 2 and the lone pair Energy vanished. However, if the BO starts to be higher than 2, it causes the break of a lone pair, and the penalty is applied.

1.4.3.5 Overcoordination Energy

This energy, defines as E_{over} in the equation 1.14, is also a penalty energy which applies for overcoordinate atoms. It uses the Val_i valency of an atom i and the BO order to compute an overcoordination term with the equation 1.30.

$$\Delta_i = -Val_i + \sum_{j=1}^{neighbours(i)} BO_{ij} \tag{1.30}$$

From this term, it computes a corrected overcoordination taking into account the deviation from lone pair electrons Δ_i^{lp}, using the equation 1.31.

$$\Delta_i^{lpcorr} = \Delta_i - \frac{\Delta_i^{lp}}{1 + p_{ovun3}.exp\left(p_{ovun4}.\left[\sum_{j=1}^{neighbours(i)}(\Delta_j - \Delta_j^{lp}).(BO_{ij}^{\pi} + BO_{ij}^{\pi\pi})\right]\right)} \tag{1.31}$$

Thus, with this corrected overcoordination term, the penalty energy of overcoordination is computed with the equation1.32.

$$E_{over} = \frac{\sum_{j=1}^{nbond} p_{ovun1}.D_e^{\sigma}.BO_{ij}}{\Delta_i^{lpcorr} + Val_i}.\Delta_i^{lpcorr}.\left[\frac{1}{1 + exp(p_{ovun2}.\Delta_i^{lpcorr}}\right] \tag{1.32}$$

1.4.3.6 Undercoordination Energy

Similarly, with the overcoordination, there is also an energetical term E_{under}, used for undercoordination when Δ_i is negative. Indeed, for under-coordinated atoms with π-electron, this term takes into account the resonance of this π-bond through the equation 1.33 .

$$E_{under} = -p_{ovun5}.\frac{1 - exp(p_{ovun6}.\Delta_i^{lpcorr})}{1 + exp(-p_{ovun2}.\Delta_i^{lpcorr})}.\frac{1}{1 + p_{ovun7}.exp\left[p_{ovun8}.\left(\sum_{j=1}^{n(i)}(\Delta_j - \Delta_j^{lp})(BO_{ij}^{\pi} + BO_{ij}^{\pi\pi})\right)\right]} \tag{1.33}$$

1.4.3.7 Valence Angle Energy

Denoted as E_{val} in the equation 1.14, it is an energy contribution from the deviations in valence angle from its equilibrium value (determined by the sum of π-bond orders SBO). Thus, it is calculated with the equation 1.34.

$$E_{val} = f_7(BO_{ij}).f_7(BO_{jk}).f_8(\Delta j).\left[p_{val1} - p_{val1}exp\left(-p_{val2}(\Theta_0(BO) - \Theta_{ijk})^2\right)\right] \tag{1.34}$$

With Θ_{ijk} the angle considered, and the function f_7 and f_8, and Θ_0 (BO) presented in the equations 1.35, 1.36 and 1.38.

$$f_7(BO_{ij}) = 1 - exp(-p_{val3}.BO_{ij}^{p_{val4}}) \tag{1.35}$$

$$f_8(\Delta j) = p_{val5} - (p_{val5} - 1).\frac{2 + exp(p_{val6}.\Delta_j^{angle})}{1 + exp(p_{val6}.\Delta_j^{angle}) + exp(-p_{val7}.\Delta_j^{angle})} \tag{1.36}$$

With the Δ_j^{angle} computed with the equation 1.37 to take into account over and undercoordination, and Val_i^{angle} is equal to the Val_i^{boc} of the equation 1.24.

$$\Delta_j^{angle} = -Val_j^{angle} + \sum_{n=1}^{neighbours(j)} BO_{ij} \tag{1.37}$$

$$\Theta_0(BO) = \pi - \Theta_{0,0}.[1 - exp(-p_{val10}.(2 - SBO2))] \tag{1.38}$$

With $\Theta_{0,0}$ the equilibrium angle.

And finally, the term SBO2 is obtained from the SBO term, the sum of π-bond orders with the equations 1.39 and 1.40.

$$SBO = \sum_{n=1}^{neighbours(j)} (BO_{jn}^{\pi} + BO_{jn}^{\pi\pi}) + \left[1 - \prod_{n=1}^{neighbours(j)} exp(-BO_{jn}^{8})\right].(-\Delta_j^{angle} - p_{val8}.n_{lp,j}) \tag{1.39}$$

$$\begin{aligned} SBO2 &= 0 \text{ if } SBO \leq 0 \\ SBO2 &= SBO^{p_{val9}} \text{ if } 0 < SBO < 1 \\ SBO2 &= 2 - (2 - SBO)^{p_{val9}} \text{ if } 1 < SBO < 2 \\ SBO &= 2 \text{ if } SBO > 0 \end{aligned} \tag{1.40}$$

With these equations, for a carbon atom j, the equilibrium angle is changing with the π-bond order, when its is equal to 0 it has an equilibrium angle of 109.47 and the carbon is hybridized sp^3, to 120 in sp^2 and the π-bond order is equal to 1, and 180 in sp with π-bond order equal to 2.

1.4.3.8 Penalty Energy

Another penalty energy E_{pen} in the equation 1.14 is added to the ReaxFF potential. It uses, for case when an atom share double bonds with different neighbours, like in allene molecule, described by the equation 1.41.

$$E_{pen} = p_{pen1}.f_9(\Delta_i).exp\Big[- p_{pen2}.(BO_{ij} - 2)^2\Big].exp\Big[- p_{pen2}.(BO_{jk} - 2)^2\Big] \tag{1.41}$$

With the function f_9 defines by the equation 1.42.

$$f_9(\Delta i) = \frac{2 + exp(-p_{pen3}.\Delta_i)}{1 + exp(-p_{pen3}.\Delta_i) + exp(p_{pen4}.\Delta_i)} \tag{1.42}$$

1.4.3.9 Three-body conjugation Energy

With the example of the -NO_2 group, it is necessary to have an additional term to describe three-body conjugation, allowing great stability for this kind of conjugation. This term, E_{coa}, is computed with the equation 1.43.

$$\begin{aligned} E_{coa} = p_{coa1}.\frac{1}{1 + exp(p_{coa2}.\Delta_i)}.exp\Big[- p_{coa3}.\Big(- BO_{ij} + \sum_{n=1}^{neighbours(i)} BO_{in}\Big)^2\Big] \\ exp\Big[- p_{coa3}.\Big(- BO_{jk} + \sum_{n=1}^{neighbours(i)} BO_{kn}\Big)^2\Big] \\ exp[-p_{coa4}.(BO_{ij} - 1.5)^2].exp[-p_{coa4}.(BO_{jk} - 1.5)^2] \end{aligned} \tag{1.43}$$

1.4.3.10 Torsion angle Energy

E_{tors} corresponds to the energy contribution from torsion angles, calculated by taking into account the dependence of the energy of torsion angles with the BO through the equation 1.44.

$$E_{tors} = f_{10}(BO_{ij}, BO_{jk}, BO_{kl}).sin(\Theta_{ijk}).sin(\Theta_{jkl}).$$
$$\left[\frac{1}{2}V_1.(1 + cos\omega_{ijkl}) + \frac{1}{2}V_2.exp\left(p_{tor1}.(BO_{jk}^{\pi} - 1 + f_{11}(\Delta_j, \Delta_k))^2\right)\right. \tag{1.44}$$
$$\left.(1 - cos2\omega_{ijkl}) + \frac{1}{2}V_3.(1 + cos3\omega_{ijkl})\right]$$

Where ω_{ijkl} is the torsion angle, V_1, V_2 and V_3 are torsion barriers, and the functions f_{10} and f_{11} are defined by the equations 1.45 and 1.46.

$$f_{10}(BO_{ij}, BO_{jk}, BO_{kl}) = [1 - exp(-p_{tor2}.BO_{ij}].[1 - exp(-p_{tor2}.BO_{jk}].[1 - exp(-p_{tor2}.BO_{kl}] \tag{1.45}$$

$$f_{11}(\Delta_j, \Delta_k)) = \frac{2 + exp[-p_{tor3}.(\Delta_j^{angle}, \Delta_k^{angle})]}{1 + exp[-p_{tor3}.(\Delta_j^{angle}, \Delta_k^{angle})] + exp[p_{tor4}.(\Delta_j^{angle}, \Delta_k^{angle})]} \tag{1.46}$$

A correction of these energetic terms is denoted as torsions 2013 in the ADF software. In function 10, to ensure correct asymptotic behaviour of the BO, each function's three-term is replaced by the functional form of the equation 1.47.

$$[1 - exp(-p_{tor2}.BO_{ij}] \rightarrow [1 - exp(-2p_{tor2}.BO^2] \tag{1.47}$$

1.4.3.11 Four body conjugation Energy

In order to describe effectively aromatic system, an energitecal contribution is added to the ReaxFF potential of the equation 1.14 named E_{conj}, which represents the contribution of conjugation effects and it is represented in the equation 1.48.

$$E_{conj} = f_{12}(BO_{ij}, BO_{jk}, BO_{kl}).p_{cot1}.\left[1 + (cos^2\omega_{ijkl} - 1).sin(\Theta_{ijk}).sin(\Theta_{jkl})\right] \tag{1.48}$$

With f_{12} being described in the equation 1.49.

$$f_{12}(BO_{ij}, BO_{jk}, BO_{kl}) = exp\left[-p_{cot2}.(BO_{jk} - 1\frac{1}{2})^2\right].exp\left[-p_{cot2}.(BO_{kl} - 1\frac{1}{2})^2\right]exp\left[-p_{cot2}.(BO_{ij} - 1\frac{1}{2})^2\right] \tag{1.49}$$

Thus, with these equations, the conjugation energy is maximal when the BO is equal to 1.5.

As well as for the torsion it exists a correction for the function 12 in the torsions 2013 presented in the

equation 1.50.

$$exp\Big[- p_{cot2}.(BO_{jk} - 1\frac{1}{2})^2\Big] \rightarrow sin(\frac{\pi}{3}BO)^4 \tag{1.50}$$

1.4.3.12 Hydogen bond Energy

In the ReaxFF potential, a term for the hydrogen bond interactions E_{Hbond} is also implemented. It is obtained from the interactions between hydrogen linked to an atom X-H, creating a system: X-H - - Z. It uses the hydrogen bond equilibrium distance r^0_{hb}, the distance r_{HZ} and the angle of the system Θ_{XHZ}, and the equation is shown on the equation 1.51.

$$E_{Hbond} = p_{hb1}.[1 - exp(p_{hb2}.BO_{XH})].exp\Big[p_{hb3}\Big(\frac{r^0_{hb}}{r_{HZ}} + \frac{r_{HZ}}{r^0_{hb}} - 2\Big)\Big].sin^8\Big(\frac{\Theta_{XHZ}}{2}\Big) \tag{1.51}$$

1.4.3.13 Correction Energy

E_{C2} is an energy penalty added in case of the triple bond between two carbons to destabilise the bond, which is overstabilised by ReaxFF. This term is described by the equation 1.52 and depends on the BO and the overcoordination term.

$$\begin{aligned} E_{C2} = k_{c2}.(BO_ij - \Delta_i - 0.04.\Delta_i^4 - 3)^2 \text{ if } BO_{ij} - \Delta_i - 0.04.\Delta_i^4 > 3 \\ E_{C2} = 0 \text{ if } BO_{ij} - \Delta_i - 0.04.\Delta_i^4 \leq 3 \end{aligned} \tag{1.52}$$

1.4.3.14 Triple Bond Energy correction

However, paradoxically, for carbon monoxide, it is necessary to have the potential to stabilize this special triple bond. This term E_{triple} in the 1.14 , only affects the C-O bonded pairs, and it is represented in the equation 1.53.

$$E_{trip} = p_{trip1}.exp[-p_{trip2}(BO_{ij} - 2.5)^2] \frac{exp\Big[- p_{trip4}.\Big(\sum_{k=1}^{neighbours(i)} BO_{ik} - BO_{ij}\Big)\Big] + exp\Big[- p_{trip4}.\Big(\sum_{k=1}^{neighbours(j)} BO_{jk} - BO_{ij}\Big)\Big]}{1 + 25.exp[p_{trip3}(\Delta_i + \Delta_j)]} \tag{1.53}$$

1.4.3.15 Taper correction

In order to describe the intermolecular interactions of Van der Waals and the coulombic forces, a taper correction from Vos Burchart (1995)[56] was introduced in ReaxFF to avoid energy discontinuities at the range of the non-bonded cutoff radius. Thus, this Taper correction is a 7^{th} order polynomial depending on the distance described by the equation 1.54.

$$Tap = Tap_7.r_{ij}^7 + Tap_6.r_{ij}^6 + Tap_5.r_{ij}^5 + Tap_4.r_{ij}^4 + Tap_3.r_{ij}^3 + Tap_2.r_{ij}^2 Tap_1.r_{ij} + Tap_0 \tag{1.54}$$

With Tap_i are terms calculated from the non-bonded cutoff radius R_{cut} with the equation 1.55.

$$\begin{aligned} Tap_7 &= 20/R_{cut}^7 \\ Tap_6 &= -70/R_{cut}^6 \\ Tap_5 &= 84/R_{cut}^5 \\ Tap_4 &= -35/R_{cut}^4 \\ Tap_3 &= 0 \\ Tap_2 &= 0 \\ Tap_1 &= 0 \\ Tap_0 &= 1 \end{aligned} \tag{1.55}$$

1.4.3.16 Van der Waals Energy

In the equation 1.14, the E_{vdW} energy is the contribution of the Van der Waals interactions described by a distance-corrected Morse-potential. It is expressed using the f_{13}, which is a shielded interaction which prevents high repulsion between atoms and shared valence angles. The equation 1.56 represents the Van der Waals Energy, and 1.57 the f_{13}.

$$E_{vdWaals} = Tap.D_{ij}.\left[exp\left(\alpha_{ij}.\left(1 - \frac{f_{13}(r_{ij})}{r_{vdW}}\right)\right) - 2.exp\left(\frac{1}{2}\alpha_{ij}.\left(1 - \frac{f_{13}(r_{ij})}{r_{vdW}}\right)\right)\right] \tag{1.56}$$

$$f_{13}(r_{ij}) = \left[r_{ij}^{p_v dw1} + \left(\frac{1}{\gamma_w}\right)^{p_{vdw1}}\right]^{\frac{1}{p_{vdw1}}} \tag{1.57}$$

With α_{ij} and γ_w being empirical van der Walls parameters, D_{ij} the van der Waals dissociation energy,

r_{vdW} the van der Waals radius.

1.4.3.17 Coulomb Energy

The last term of the ReaxFF potential of the equation 1.14 is $E_{Coulomb}$ which is the contribution of the Coulomb interactions described by a shielded Coulomb-potential presented in the equation 1.58. The potential depends on the atomic charges q_i and q_j, the distance r_{ij} and γ_{ij}, the shielding parameter.

$$E_{coulomb} = Tap.C.\frac{q_i.q_j}{\left[r_{ij}^3 + (1/\gamma_{ij})^3\right]^{1/3}} \tag{1.58}$$

1.4.3.18 Electronegativity Equalization Method

The Electronegativty Equalization Method (EEM)[91 92].

$$\chi_i = \chi_i^* + 2\eta_i^* q_i + \sum_{j \neq i} \frac{q_j}{r_{ij}} \tag{1.59}$$

With χ_i the electronegativty of the atom i, χ_i^* and η_i^* the EEM electronegativity and hardness.

1.4.3.19 Long-range-correction

In 2011, Liu et al.[66], added to the ReaxFF potential, a long-range-correction energy term describe in the equation 1.60. This term is created to represent the London dispersion of the van der Waals interactions.

$$E_{lgd} = -\sum_{ij,i<j}^{N} \frac{C_{lg,ij}}{r_{ij}^6 + dR_{eij}^6} \tag{1.60}$$

With R_{eij} is the equilibrium Van der Waals distance, and $C_{lg,ij}$ is the dispersion energy correction parameter.

1.4.3.20 Furman and Wales corrections

In 2019 and 2020, David Furman and David J. Wales published two articles fixing some discontinuities in the Standalone ReaxFF potential presented above[93 94]. This section presents the corrections that they had developed.

In their first publication to improve the numerical stability of ReaxFF[93], they introduce a taper correction on the BO to cancel the short-range interactions define in the equation 1.61.

$$\begin{aligned} S(x) = 0 \text{ if } x \leq x_{min} \\ S(x) = Tap^{7th}(x), \text{ if } x_{min} < x < x_{max} \\ S(x) = 1 \text{ if } x \geq x_{max} \end{aligned} \tag{1.61}$$

With S(x) the new taper correction, where x_{min} and x_{max} are cutoff values chosen in order to respect that S(x_{min}) corresponds to a non-bonded environment and S(x_{max}) to a bonded environment, Tap$^{7th}(x)$ is a 7^{th} polynomial similar to the equation 1.54 and describes in the equation 1.62 and 1.63.

$$Tap^{7th}(x) = S_7.x^7 + S_6.x^6 + S_5.x^5 + S_4.x + S_3.x^3 + S_2.x^2 S_1.x + S_0 \tag{1.62}$$

$$\begin{aligned} S_7 &= 20/x_r^7 \\ S_6 &= -70/x_r^6 \\ S_5 &= 84/x_r^5 \\ S_4 &= -35/x_r^4 \\ S_3 &= 140/x_r^3 \\ S_2 &= -210/x_r^2 \\ S_1 &= 140/x_r^1 \\ S_0 &= 1 \\ x_r &= x_{max} - x_{min} \end{aligned} \tag{1.63}$$

With this new taper correction, they applied this function on the BO, which transformed several intramolecular energetical terms presented in the equation 1.64.

$$
\begin{aligned}
BO'_{ij} = S(BO'^{\sigma}_{ij}).BO'^{\sigma}_{ij} + S(BO'^{\pi}_{ij}).BO'^{\pi}_{ij} + s(BO'^{\pi\pi}_{ij}).BO'^{\pi\pi}_{ij} \\
E_{bond} = E_{bond} \\
E_{val} = E_{val}.S(BO_{ij}).S(BO_{ik}) \\
E_{pen} = E_{pen}.S(BO_{ij}).S(BO_{ik}) \\
E_{coa} = E_{coa}.S(BO_{ij}).S(BO_{ik}) \\
E_{hbond} = E_{hbond}.S(BO_{ij}).S(r_{ik}) \\
E_{tors} = E_{tor}.S(BO_{ij}).S(BO_{ik}).S(BO_{jl}) \\
E_{conj} = E_{conj}.S(BO_{ij}).S(BO_{ik}).S(BO_{jl})
\end{aligned}
\tag{1.64}
$$

This correction seems to increase the energetical conservation of ReaxFF, which might be helpful for geometrical optimization using ReaxFF. However, it is essential to indicate that this correction is only available in the ADF, non-free software.

In their second publication in 2020, Furman and Wales[94], introduces also taper correction to the lone pair energetically terms, and especially the computation of number of lone pairs around an atom $n_{lp,i}$ presented in the equation 1.28, with the new formalism presented in the equation 1.65 that ensure a continuous function.

$$
\begin{aligned}
n_{lp,i} = f_L(\Delta^e_i) \text{ if } \Delta^e < l_1 \\
n_{lp,i} = f_L(\Delta^e_i).Tap^{7th}(\Delta^e_i), \text{ if } l_1 \leq \Delta^e_i \leq l_2 \\
n_{lp,i} = 0 \text{ if } l_2 \leq \Delta^e \leq r_1 \\
n_{lp,i} = f_R(\Delta^e_i).Tap^{7th}(\Delta^e_i), \text{ if } r_1 \leq \Delta^e_i \leq r_2 \\
n_{lp,i} = f_R(\Delta^e_i) \text{ if } \Delta^e_i > l_1
\end{aligned}
\tag{1.65}
$$

With $f_R(\Delta^e_i)$ and $f_R(\Delta^e_i)$ functions presented in the equations 1.66 and 1.67, and l_1, l_2, r_1 and r_2 parameters that verify $l_1 < l_2$ and $r_1 < r_2$.

$$
f_L(\Delta^e_i) = \Delta^e_i - \frac{1}{2} - \frac{1}{\pi} tan^{-1}\left[\frac{\lambda sin(2\pi\Delta^e_i)}{\lambda cos(2\pi\Delta^e_i) - 1}\right] \tag{1.66}
$$

$$
f_R(\Delta^e_i) = \Delta^e_i + \frac{1}{2} - \frac{1}{\pi} tan^{-1}\left[\frac{\lambda sin(-2\pi\Delta^e_i)}{\lambda cos(2\pi\Delta^e_i) - 1}\right] \tag{1.67}
$$

With λ a fitting parameter that verify $0 < \lambda < 1$. In their publication they proposed values for their

parameters with λ = 0.8, l_1=-2.3, l_2=-1.5, r_1=1.5 and r_2=2.3.

Moreover, they also applied a taper correction to the SBO2 computation of the equation 1.40.It is shown in the equation 1.68, 1.69 and 1.70.

$$\begin{aligned} SBO2 = 0 \text{ if } SBO \leq 0 \\ SBO2 = 2 * Tap^{5th}(SBO), \text{ if } 0 < SBO < 2 \\ SBO2 = 2 \text{ if } SBO \geq 2 \end{aligned} \tag{1.68}$$

$$Tap^{5th}(x) = \frac{s_5.x^5 + s_4.x + s_3.x^3 + s_2.x^2 s_1.x + s_0}{(x_r - x_l)^5} \tag{1.69}$$

$$\begin{aligned} s_5 &= -6 \\ s_4 &= 15(x_r + x_l) \\ s_3 &= -10(x_r^2 + 4x_r x_l + x_l^2) \\ s_2 &= 30(x_r^2 x_l + x_r x_l^2) \\ s_1 &= -30(x_r^2 x_l^2) \\ s_0 &= 10x_r^2 x_l^3 - 5x_r x_l^4 + x_l^5 \end{aligned} \tag{1.70}$$

With x_l and x_r parameters, which are used in the publication of Furman et al. with x_l = -0.2 and x_r = 2.2.

Finally, the last correction that Furman et al.[94] developed concerned the torsion energetical terms. Indeed they found that when a bending angle approaches 0° or 180°, the cos(ω_{ijkl} terms tend to 0 in the equation 1.44. To fix this, they transform the product of the sins by elevating it to the power 3, presented in the equation 1.71.

$$sin(\Theta_{ijk}).sin(\Theta_{jkl}) \rightarrow \left[sin(\Theta_{ijk}).sin(\Theta_{jkl})\right]^3 \tag{1.71}$$

Again, as for the correction of 2019, this correction is only available in the ADF software.

1.5 Introduction to Molecular Dynamics

In this part, the molecular dynamics simulation method is presented. This method allows investigating the dynamic evolution of a molecular system depending on the time by generating the most probable trajectory of the system from a given number of particles, which can be atoms[95,96]. A simple trajectory definition is a

collection of phase points containing all the system particles' positions and velocities.

The molecular dynamics simulation depends on statistical mechanics. Thus, the system investigated, at equilibrium, depends on probability functions, which allow the connection of microscopic properties to macroscopic thermodynamic quantities (such as temperature and free energy). Different thermodynamics ensembles can be used with molecular dynamics and are presented in the following sections. However, the main concept is that when the system respects the ergodic hypothesis, using the trajectory, it allows for determining the time average (constituted of all the phase points sampled) and, then, the ensemble average.

In order to obtain the trajectory, the molecular dynamics method uses an iterative algorithm depicted in figure 1.8 that will allow us to get into more details. One key component to investigate the motion of the system is the forces that will be applied between each particle in the simulation, and for that, it is necessary to determine the forces that apply on each particles.

Indeed the classical molecular dynamics method is defined by classical mechanics where each particle is represented by its mass and interacts with each other using potentials. Thus, the simulation of the particles is made thanks to Newton's second law[97] presented in the equation 1.72 :

$$\vec{F_i} = m_i \vec{a_i} = m_i \frac{\delta \vec{v_i}}{\delta t} = m_i \frac{\delta^2 \vec{r_i}}{\delta t^2}, \quad i = 1, ..., N \tag{1.72}$$

With N the total number of particles, m_i the particle mass, $\vec{a_i}$ its acceleration, $\vec{v_i}$ its velocity, $\vec{r_i}$ its position and F_i the total forces acting on it, and t defining the time. This forces is also defined by the opposite derivative of the potential energy through the equation :

$$\sum F_i = -\frac{\delta U(r_i)}{\delta r_i} \quad i = 1, ..., N \tag{1.73}$$

With U, the potential energy and r_i the ensemble of the position vector of the N particles. Thus, with these two equations, it is possible iteratively, knowing the initial positions r_i(t=0) and velocities v_i(t=0), and the forces acting on each particle (F_1 (t=0),..., F_N (t=0)) from the potential energy, to integrate and obtain our new phase point with the new positions and the new velocities. So, after the chosen number of steps for the simulation, the trajectory is obtained.

However, resolving this newton equations is not trivial and needs the use of an integration algorithm.

So, before running a molecular dynamics simulation, we must choose the model that we want for the simulation. Indeed the model would determine the nature and number of particles inside our system, the thermodynamics conditions, the boundary conditions implemented, the interaction potential used to compute the system's potential energy and to obtain the forces, and the algorithm used to solve Newton's equation. All of these different aspects of the model are presented in the following.

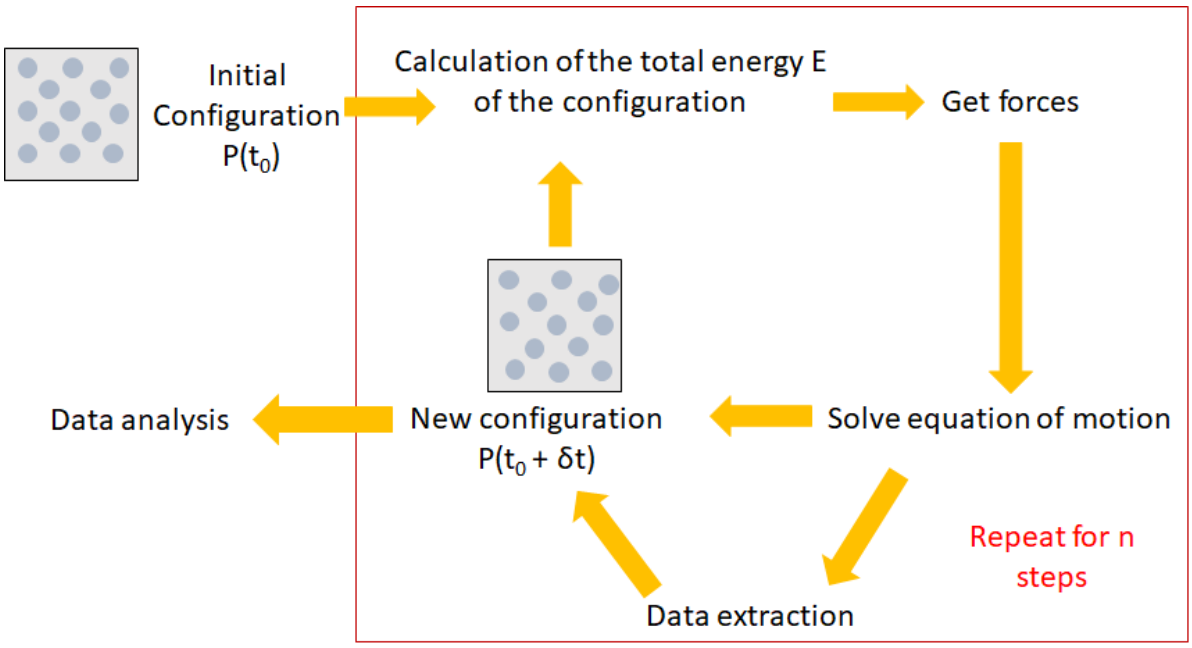


Figure 1.8: Schematization of the Molecular Dynamics algorithm workflow. It starts with an initial phase point denoted as $P(t_0)$, and then the system's total energy is computed, the force is obtained, and with the equation of motion resolved, it gives a new configuration. The iterative cycle is repeated as a function of the number of steps chosen n.

1.5.1 Thermodynamic ensembles

The model for the molecular simulation also contain thermodynamic properties such as the temperature, the total energy or the enthalpy, the pressure or the volume, and the number of particles or the chemical potential. It exists different thermodynamic ensembles as: the microcanonical ensemble (NVE), in which the number of particles (N), the volume (V) and the energy (E) are fixed to defined values. It is used in molecular dynamics to investigate adiabatic processes with no heat exchange. In this ensemble, the Hamiltonian corresponding to the total energy of the system is conserved through the equation :

$$\mathcal{H}_{NVE} = \sum_{i=1}^{N} \frac{1}{2} \frac{\vec{p_i}^2}{m_i} + U(\vec{r_i}) \tag{1.74}$$

With the first term being the system's kinetic energy and $\vec{p_i}$ the momentum equal at the velocity time the mass, and the second term corresponds to the potential energy.

The canonical ensemble (NVT), with N, V and the temperature (T) are constant. With this ensemble, it is possible to observe endothermic and exothermic processes where the system is connected to a thermostat with which it can exchange energy. With this ensemble, the temperature of the system is defined by the next equation :

$$T = \frac{1}{g k_B} \sum_{i=1}^{N} \frac{\vec{p_i}^2}{m_i} \tag{1.75}$$

with g the degree of freedom and k_b the Boltzmann constant.

In the Isothermal-Isobaric ensemble (NPT), N, the pressure (P) and T are kept constant. It is usually denoted as an ensemble corresponding to laboratory conditions with ambient temperature and pressure. These three first ensembles are the most commonly used in molecular dynamics simulation in chemistry.

Thus, to maintain the temperature constant, different methods acting as a thermostat are used and presented below.

1.5.2 Thermostat

It exists diverse thermostats to equilibrate the system at a specific temperature. thereafter only two will be presented which are the Berendsen[98] and the Nosé-Hoover thermostats[99 100].

1.5.2.1 Berendsen thermostat

The Berendsen thermostat is based on the usage of scaling factor for the velocities. Indeed, using this thermostat, the system is coupled to an external heat bath with a temperature T_0. At each step the gap between the temperature of the system T and T_0 is corrected by using a scaling factor on the velocities named λ represented in the equation :

$$\lambda^2 = 1 + \frac{\delta t}{\tau}\left(\frac{T_0}{T} - 1\right) \tag{1.76}$$

With δ t the timestep of the simulation, τ is a coupling parameter that defines how tight the bath and system temperature are. Indeed if τ is too high, the thermostat is inactive. If it is too small, it will create unrealistic low-temperature fluctuations. Usually, in molecular dynamics, the value is fixed at 0.1ps. This thermostat can efficiently relax a system at a temperature. Still, to extract properties when the equilibrium is reached, the Nosé-Hoover is generally better[101].

1.5.2.2 Nosé-Hoover thermostat

Another method is represented by the Nosé-Hoover thermostat. In this methodology, the idea is to implement a new degree of freedom (s) acting like a heat reservoir. With this reservoir, it controls the system temperature and it leads to a new Hamiltonian :

$$\mathcal{H}(P, R, p_s, s) = \mathcal{H}_\prime + \frac{p_s^2}{2Q} + g k_B T \ln(s) \tag{1.77}$$

Where P and R represent all coordinates $\vec{r_i}$ and $\vec{p_i}$, but they are both virtual, and related to the real coordinates as R = $\vec{r_i}$ and P = $\frac{\vec{p_i}}{s}$, g being the number of independent momentum degrees of freedom of the system,

and Q is an imaginary mass. Thus in this equation, the first term correspond to the kinetic and potential energy of the system. Then the third and fourth terms represent the kinetic and potential energy of the artificial coordinate s. It allows a constant temperature and ensures energy conservation.

1.5.3 Initial configuration

The first step of the Molecular Dynamics simulation is the initial configuration of the system, denoted here as $P(t_0)$, containing the position and the velocities of our particles. Depending on the thermodynamics ensemble chosen, an equilibrium procedure is usually done before the simulation to avoid drastic energy changes during the simulation and generate the most probable phase point.

1.5.4 Boundary conditions

Another essential parameter of the model that should be well-considered is the boundary conditions. Indeed, these conditions will drive the simulation by defining the environment around the simulation. In this section, two of them are presented.

1.5.4.1 Periodic boundary conditions

The most used boundary conditions are the periodic boundary conditions (PBC)[96]. It corresponds to a set of boundary conditions that simulate an "infinite" system with a finite number of particles using a unit cell. This unit cell is repeated infinitely in all the dimensions, usually in three dimensions. Only one will be the simulation box, the others being named "images". It means the simulation box that when a particle crosses the border of the box, it re-appears on the opposite side with the same velocity. These conditions ensure that the same number of particles in the simulation box will always allow condensed phase simulations. Thus, during the simulation, only the properties of the simulated box are recorded and propagated. Moreover, the PBC are usually combined with the minimum-image convention, which implies that each particle in the simulation interacts with the closest image of the real particles in the system. A representation in two-dimension of the PBC is presented in figure 1.9.

Moreover, due to the periodic boundary conditions, the simulation box should have a propagable topology, which means it cannot be a sphere, for example. It also important, to choose wisely the size of the simulation to prevent unwanted interactions. As an example, if the size of the simulation box is too small for the simulation of a macromolecule, it may interact with its own image.

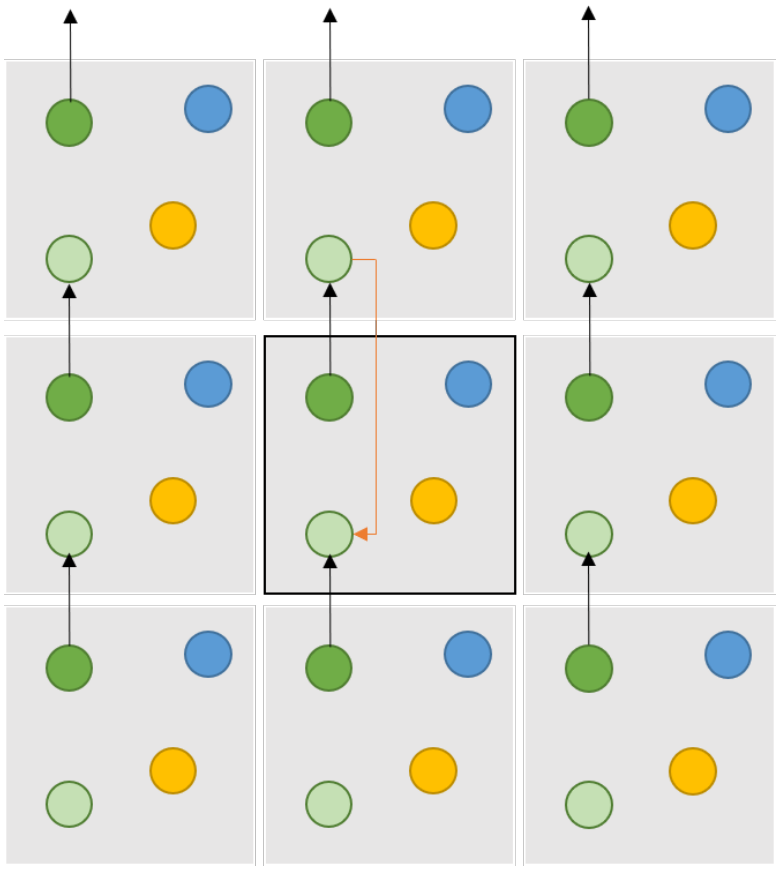


Figure 1.9: Schematization of PBC in two dimensions with, in black, the simulation box and around the different images. The green particles represent particles going out of the boundary of the simulation box, and the images, re-appearing on the other side, are represented by lighter green particles.

1.5.4.2 Lees-Edwards boundary conditions

The PBC is useful for molecular simulations of systems in thermodynamic equilibrium, but for non-equilibrium, it is not suitable [102]. Thus, in this particular case, the Lees-Edwards boundary conditions (LEBC) are used. It is based on a simple shear flow. Indeed, with these conditions, the upper and lower images are sliding in different directions with constant velocity. A representation is given in figure 1.10.

So with this conditions, it main difference with the PBC, is for particles going out of the box on the y axis. Indeed, the same process as for the PBC is applied for the y-coordinate, but the x coordinate and its velocity should be shifted. Thus, the velocity U_v is directly proportional to the y direction depicted as L in the figure 1.10.

1.5.5 Integration of Newton's equation

Integrating the forces using the second law of Newton's described in the equation 1.72 leads to a system of k times N differential equation, with k corresponding to the dimensional space. These kinds of differential equations are not solvable with a computer, so there is no analytical solution. However, integration algo-

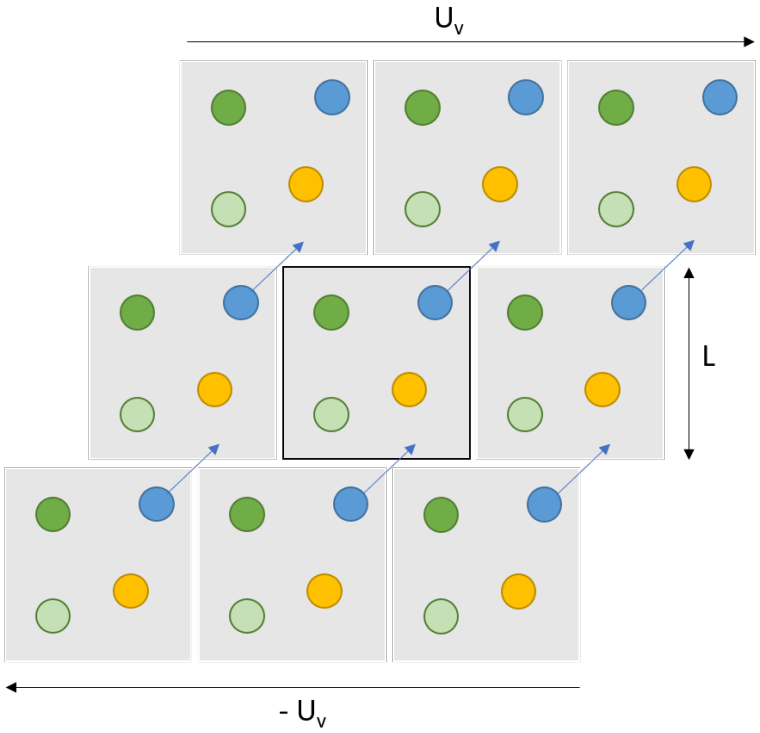


Figure 1.10: Schematization of LEBC in two dimensions with, in black, the simulation box and around the different images. In these conditions, the system is defined in three layers; In the upper is moving periodic images with constant velocities U_v, and the lower layer with velocities -U_v. The last layer is the medium layer, one with the simulation box. The images on this layer are named stationary periodic images. L indicates the length on the y-axis of the simulation box.

rithms are used to solve this problem, which transforms the differential equations of the equation 1.72 into an algebraic expression using the Taylor series expansion into the equation 1.78.

$$r_i(t+\delta t) = r_i(t) + v_i(t)\delta t + \frac{1}{2!}a_i(t)\delta t^2 ... \tag{1.78}$$

Where δt being the timestep of the simulation. With this equation, at the time (t+δt), the new position can be expressed as a sum of the actual position adding the different differential orders, with two orders represented here. Among all the different integration algorithms that exist as the Euler [103], the Runge-Kutta [104], the simple Verlet [105], we are presenting the Velocity Verlet (VV) [106] and the Leapfrog Verlet (LFV) [107].

1.5.5.1 Velocity Verlet algorithm

The Velocity Verlet algorithm is an algorithm of second-order integration. It is one of the most stable, fastest and more accurate available. This kind of numerical integration considers that the forces acting on particles are constant during one step. It stops the Taylor development in the equation 1.78 at the second order, and define the velocity v_i (t + $\frac{1}{2}$ δt) presented in the equation 1.79.

$$v_i(t + \frac{1}{2}\delta t) = v_i(t) + \frac{1}{2}a_i(t)\delta t \tag{1.79}$$

Thus, replacing $v_i(t)$ in the equation 1.78, we obtain :

$$r_i(t + \delta t) = r_i(t) + v_i(t + \frac{1}{2}\delta t)\delta t \tag{1.80}$$

Moreover, as it was presented in the equation 1.73, the sum of forces (t+δt) can be computed from the positions (t+δt), which leads us to summarize the following procedure :

$$\begin{aligned}
v_i(t + \frac{1}{2}\delta t) &= v_i(t) + \frac{\sum F_i(t)}{2m_i}\delta t \\
r_i(t + \delta t) &= r_i(t) + v_i(t + \frac{1}{2}\delta t)\delta t \\
\text{Computation of the forces } \sum F_i(t + \delta t) &\text{ from the new positions } r_i(t + \delta t) \\
v_i(t + \delta t) &= v_i(t + \frac{1}{2}\delta t) + \frac{\sum F_i(t + \delta t)}{2m_i}\delta t \\
\text{equivalent to : } v_i(t + \delta t) &= v_i(t) + \frac{\sum F_i(t + \delta t) + \sum F_i(t)}{2m_i}\delta t
\end{aligned} \tag{1.81}$$

All this procedure is depicted in the figure 1.11.

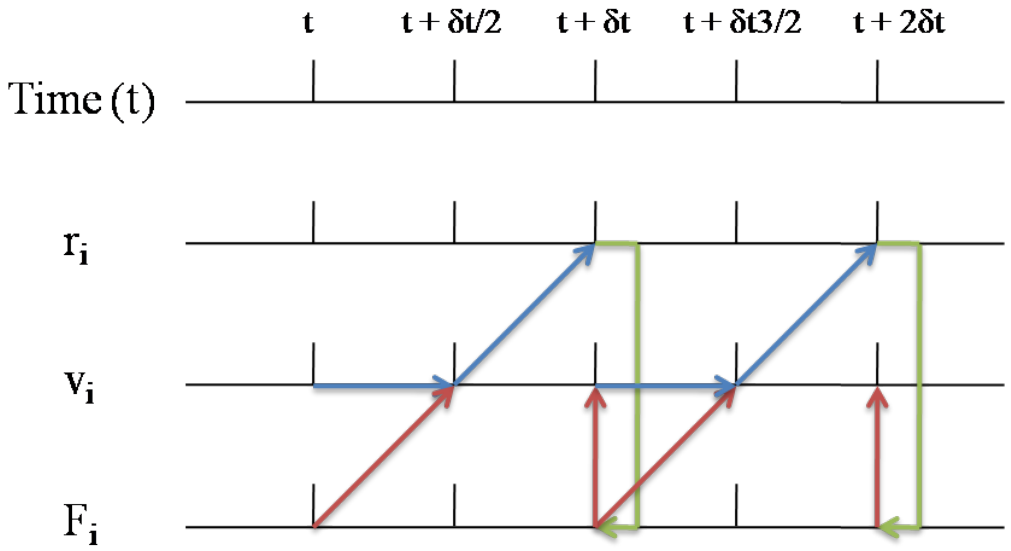


Figure 1.11: Schematization of the Velocity Verlet algorithms. In red correspond the use of the total forces, in blue the velocity and green the computation of the new forces from the new positions.

This algorithm allows to obtain at each step r_i (t+δt), v_i (t+δt) and $\sum F_i$ (t+δt). It respects the trajectory invariance by the transformation t=-t, which means that the dynamic of the system is independent of the time direction, and also it conserves the energy during the simulation. Another version of the Verlet Algorithm is the Leapfrog, where the difference is that the calculations of positions and forces happen at the time t or t+δt,

where the velocities are computed at the time t+$\frac{1}{2}\delta$t.

1.5.6 Timestep

As we have seen with the integration of Newton's equation, the last key parameter that needs to be defined is the timestep. It should be well-considered before running a simulation. For the two Verlet algorithms presented above, based on the second order of the Taylor expansion, the timestep δt should be small enough to neglect high-order differential terms. However, if the timestep is too small, the simulation time would be much higher. One way of choosing the timestep consists of defining the properties we want to extract for our system. Indeed, we might consider the different bond vibrations and the diverse degree of freedom of the system, which means that we should take a timestep lower than the smallest vibration of the system to observe it. For example, when using a reactive force field, the timestep chosen is usually between 0.1 fs and 0.25 fs.

Chapter 2

Overview of organochlorides in an environmental context

2.1 Introduction

This chapter presents a global overview of polychlorinated biphenyls (PCBs), a family of molecules being part of the organochlorides (OC), and a presentation of two major organochlorides, DDT (dichlorodiphenyltrichloroethane) and TCDD (2,3,7,8-tetrachlorodibenzo-p-dioxin). These three compounds are used in the following chapters to investigate their reactivity which is one major reason for their presentation. The OC are organic compounds with at least one covalent bond between a carbon atom and a chlorine atom. They are compounds that have been widely used in industry or pharmaceuticals for the past half-century. Due to their chemical stability, these OC can be measured in different environments, such as soil, water, atmosphere, living organisms and food. Unfortunately, they are compounds of interest due to their environmental impacts. Thus, DDT and TCDD are both organochlorine pesticides, which are highly toxic and persistent. Along with PCBs (polychlorinated biphenyls), they have received worldwide attention due to their environmental impact and high toxicity. Indeed, they have been identified as toxic, persistent compounds by the United Nations Environment Programme and have been banned since 2001. However, they are still used in some countries. This chapter introduces different OC and mainly focuses on PCBs because it is a major compound used to develop the new ReaxFF force field created for organochlorides in the next chapters. In the next section, a short presentation and a historical introduction are given on the production and use of these three OC. Afterwards, a brief presentation of their toxicity is given. It is followed by a presentation of their distribution in the environment. Then, it focused on the physical and chemical properties of the PCB, with then, which will be extremely useful in the following chapter, an overview of the theoretical investigation made on the PCBs. This chapter finishes with the different techniques developed to remove organochlorides related to the

following ReaxFF investigations.

2.2 Presentation and history of PCBs, DDT and TCDD

2.2.1 DDT

The DDT $C_{14}H_9Cl_5$ is synthesized from chloral and chlorobenzene molecules, which leads to an organochloride with five chlorine atoms and two aromatic cycles. The representation of its formation is presented in the figure 2.1. In total, DDT used in pesticides was composed of 14 compounds, with most of them being the p,p-DDT (1,1,1-trichloro-2,2-bis(p-chlorophenyl) ethane) presented in the figure 2.1, and the second being the o,p-DDT (1,1,1-tricholoro-2-(p-chlorophenyl)-2-(o-chlorophenyl) ethane), with one chlorine atom in ortho position on one phenyl ring. These two compounds are the most persistent compounds among the DDTs pesticides with a reported half-life between 2 and 15 years[108] depending of the environment.

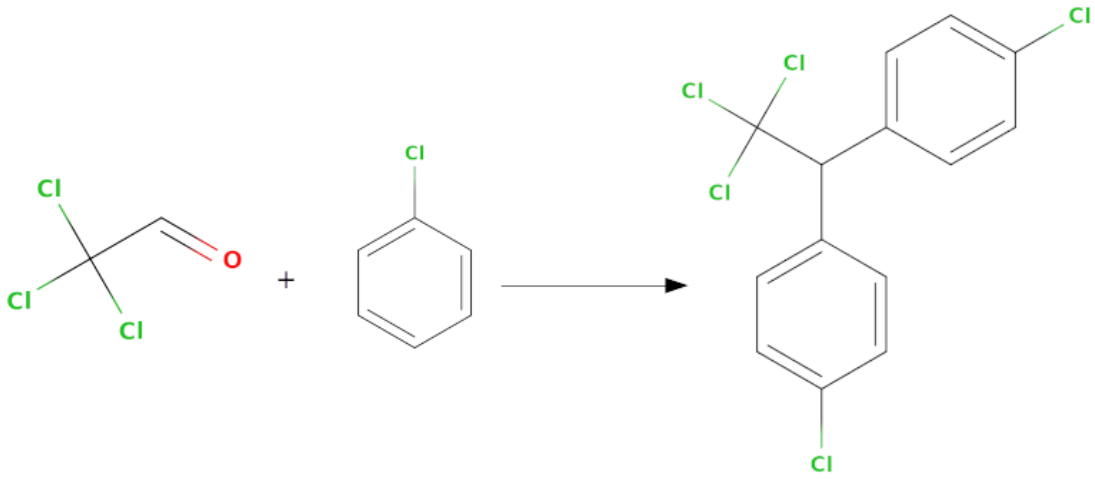


Figure 2.1: It schematizes the synthesis of the DDT ($C_{14}H_9Cl_5$) molecule represented on the right, from the chloral (C_2HCl_3O), the first molecule and the chlorobenzene (C_6H_5Cl), the second. The chlorine is represented in green and the oxygen in red.

Historically, in 1874, the first synthesis of DDT was made by Othmar Zeidler, an Austrian Pharmacologist[109]. However, the compound did not receive much attention until a swiss chemist named Paul Hermann Müller experimented with insecticides for Switzerland's major food shortages investigation. Thus, in September 1939, Müller experimented on DDT and identified its values as insecticides. From there, DDT was massively used around the world[108], and it helped eradicate malaria from many countries. This capability of the DDT to fight against disease allowed Müller to be rewarded with the Nobel Prize in 1948 for Physiology or Medicine. However, despite its good efficiency against diseases, some worries started to appear at the beginning of the sixties. For example, in the book 'silent spring' by Rachel Carson, initially published in 1962[110],

where she started to question the potential impact that DDT could have on the environment and health and presented results showing that it might be correlated to cancer and was dangerous for the wildlife. Thus, ten years after the book, DDT agricultural use was banned in the United States[108].

However, DDT was still used worldwide until 2001, UNEP (the United Nations Environment Programme) decided internationally to ban certain POPs ("Persistent Organic Pollutants"), including PCBs and DDT, by signing "The Stockholm Convention on Persistent Organic Pollutants" in May 2001[111].

Nowadays, some countries are still using DDT as herbicides[112], even if we are aware of its environmental impacts. Despite its banishment, it is still measurable in European countries, validating its persistence[113]. However, it is also used in some countries for malaria fighting disease, with the World Health Organization (WHO) agreement[114].

2.2.2 TCDD

The TCDD molecule was never industrially produced like the DDTs pesticides or the PCBs. It is formed as a secondary product when synthesising different herbicides such as chlorophenols or some of the phenoxy herbicide family. A representation of an example of chlorophenol and phenoxy acetic acid (based structure of most of the phenoxy herbicide) and the TCDD are presented in the figure 2.2. Similarly to DDT, the TCDD molecule's half-life depends on the environment. In the article of Oslon et al.[115], it is written that on soil surfaces exposed to the sun, the half-life is about 1 or 3 years; when buried in tropical subsoils, the half-life rise between 20 and 50 years and more than 100 years in water sediments.

Most of the reasons for the investigation made on the TCDD came back to the Vietnam War[115]. Indeed, the TCDD was a byproduct of the compounds used by the US army between 1961 and 1971 in the orange agent, used in their herbicidal warfare program, which is still now measurable in Vietnam[115]. Moreover, another unfortunate event with TCDD also happened in Italy in 1976, when an accident happened in a chemical manufacturing plant, which is the highest population contaminated by TCDD[116] and caused the death of more than 3000 animals. It leads to new standardization of chemical manufacturing with new safety regulations, notably the 'Seveso Directive' of the European Union.

2.2.3 PCB

The last family introduces here is the PCBs. They come from the biphenyls (two phenyl rings linked with a carbon bond), where between 1 and 10 hydrogens are substituted by chlorine atoms, which gives the chemical formula $C_{12}H_{(0-9)}Cl_{(1-10)}$. In total, they are 209 molecules which are named following the IUPAC (International Union of Pure and Applied Chemistry) nomenclature[117].

With this nomenclature, numbers are associated to the topological positions of carbons atoms (1, 1', 2, 2',

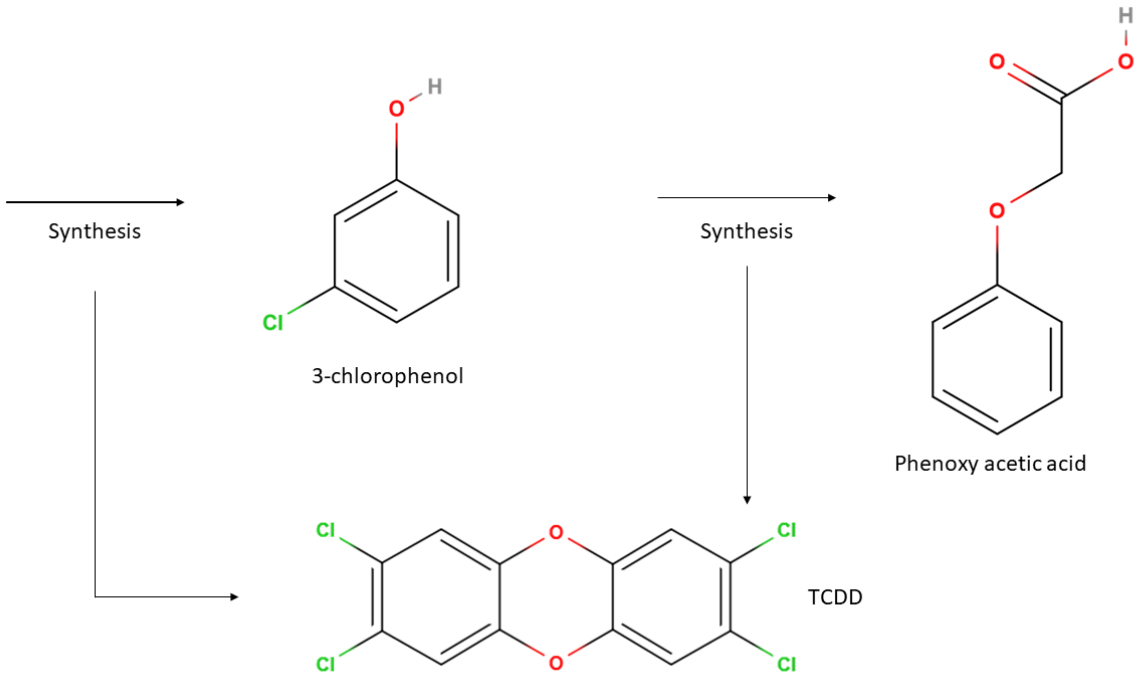


Figure 2.2: Schematization of the formation of the TCDD (C14H9Cl5) molecule from the synthesis of chlorophenols (here with the example of the 3-chlorophenol) or the phenoxy acetic acid. The chlorine is represented in green, and the oxygen in red.

3, 3', 4, 4', 5, 5', 6, 6'), which allows to localise the hydrogen and/or chlorine atoms. The atoms linked to the carbons 2, 2' and 6, 6' are in the ortho position, 3,3' and 5,5' in meta, and 4,4' in para. A representation is given in the figure 2.3. Moreover, in 1980, Ballschmiter and Zell invented a different nomenclature specifically for PCBs by giving them a 'BZ number'[118]. This number depends on the number of chlorine atoms substituted and their positions. An example is given with the PCB-1 in the figure 2.4, which in IUPAC nomenclature is the 2-chlorobiphenyl, which indicates the substitution of hydrogen linked to carbon 2 by a chlorine atom.

The PCBs are divided in 10 subgroups, depending of the number of chlorine atoms in the molecule. Thus, they are 3 mono-chlorine, 12 di-chlorine, 13 tri-chlorine, 42 tetra-chlorine, 46 penta-chlorine, 42 hexa-chlorine, 24 hepta-chlorine, 12 octa-chlorine, 3 ennea-chlorine, and 1 deca-chlorine.

Concerning their history, it started in 1865, when the PCBs were discovered during tar analysis[119]. However, it was firstly synthesis only sixteen years after, in 1881, by two Deutsch chemists, Hermann Schmidt and Gustav Schultz, who published the first article on the PCBs synthesis[120]. The intensive production started in 1929[121]. Indeed, they were used for their great thermal and electrical capacity allowing them to be used in a large field of applications. For example, PCBs were used as dielectric fluids, transformers, and capacitors, but they were also used in textile fabrication, and painting[120].

In the history of PCBs, with the different large fields of application, the different production of PCBs could have been a mixture of one conformer. However, for the majority, it was complex mixtures of different PCBs. The goal of this mixture was to obtain a chlorination rate (between 21% and 68%) to fulfil their usages[117].

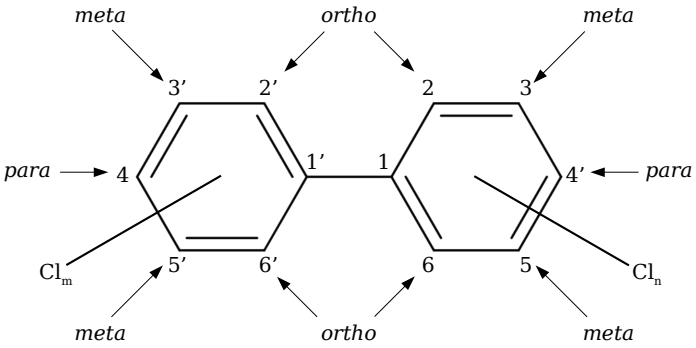


Figure 2.3: Representation of the PCBs structure using the IUPAC nomenclature. The positions ortho, meta and para are indicated. The letters m and n are associated with the chlorine atoms representing the number of substituted hydrogens.

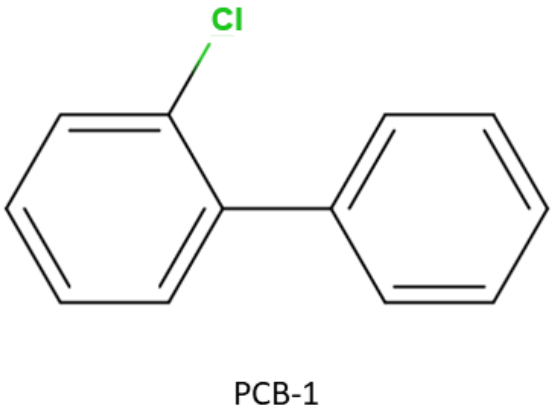


Figure 2.4: Here is represented an example of the PCB-1 using the nomenclature of Ballschmiter and Zell.

The fabrication process consists in adding anhydrous chlorine at a high temperature (150°C) with a catalyser (iron or chlorine iron) to chlorate melted biphenyls [121]. The product is then purified with 0.3% of alkaline and then vacuum distilled. The chlorination rate is more or less elevated depending on the contact time.

Thus, in the different PCBs mixtures the most used, we can find the Aroclor, Clophen, Phenochlor, Kanechlor, Pyralene, Fenclor, Sovol, and Delor [117]. They are usually composed of between 100 and 140 different kinds of PCBs, mono or non-ortho substituted [122].

According to Breivik et al. [6], in 2007, 1.325.810 tonnes of PCBs were produced since the 1930s. Among the largest producers (production of more than 2.3 % of the total production) is the company Monsanto (Aroclor) in the United States with 48.4% of the total production, then Bayer AG (Clophen) in Germany with 12% followed by Orgsteklo with 10. 7% and Orgsintez 2.4% in the U.S.S.R. (sovol), Prodelec with 10.2% in France (Phenochlor and Pyralene), again Monsanto in England with 5% and finally Kanegafuchi (Kanechlor) in Japan with 4.2%.

In 1936 [123], an article by Louis Schwartz entitled "Dermatitis from Synthetic Resins and Waxes" describes the symptoms of workers involved in the chlorination of biphenyl and, more specifically, in purification. Thus, it is written that some workers are subject to skin problems, digestive problems, impotence, eye burns and haematuria. These symptoms were the first signs of more in-depth investigations into PCBs. In the spring of 1936, three workers at the 'Halowax Corporation' in a PCB factory died of extreme jaundice. This tragedy led to a study on rats carried out by Drinker et al. [124], who showed that the toxicity of PCBs increased with the degree of chlorination and that the liver was more likely to be affected. It was then established that PCBs, even in low concentrations, can have severe repercussions. Their family of compounds is probably one of the most dangerous of the different families of chlorinated hydrocarbons studied. Despite several controversies surrounding this subject [119], PCB production continued worldwide.

It was in 1969 that a Swedish scientist published an alarming article [125]. Indeed, he showed the omnipresence of PCBs around us in various animal samples. He also demonstrated a higher concentration in animals at the top of the food chain (eagles in this case), thus showing the ability of PCBs to bio-accumulate throughout the food chain.

In October 1968, PCB poisoning from contaminated rice affected over 1200 people in Japan [126]. However, the link with PCBs was not immediate and became known as "oil disease" or "Yusho" as a kind of food poisoning. Two years later, in the autumn of 1970, two studies from Kyoto University and Ehime demonstrated the presence of PCBs in salt water and fresh fish in Japan. It proves Japan's contamination of PCBs. Rapidly after, between March and June 1971, PCB production in Japan was stopped. Unfortunately, after this incident, several studies on the effect of PCBs on health will be published [127] [128] [129]. They identified numerous symptoms with physical abnormalities. However, in contrast to the abnormalities detected following "Yusho", the

presence of liver abnormalities is not prevalent and seems more related to the presence of PCDFs (Polychlorinated Dibenzofurans) in the contaminated rice. The blood test showed a link between the plasma and certain abnormalities. Indeed, exposure to PCBs in some subjects appears to have reduced the mean half-life of antipyrin, suggesting an effect on the enzymatic metabolism of antipyrine in humans. In the continuation of the studies on the symptoms linked to PCBs, in 1981, an article by Maroni et al.[130] showed that the absorption of PCBs happens by skin contact and that protection must be used.

As concern about the omnipresence and dangerousness of PCBs grew, Sweden, after Japan, decided to ban their production in 1972[131]. This decision was followed, in 1976, by a measure of the American Congress to ban the production of PCBs, but it was not until 1978 that all factories stopped producing them[121]. In Europe, restrictions began in the 1970s. Moreover, countries began to ban PCBs, starting with Norway in 1980, followed by Finland in 1985 and Denmark in 1986[131]. Nevertheless, it was not until 1990, during the Third International Ministerial Conference on the Protection of the North Sea, that the European PCB-producing countries decided to eliminate and destroy the remaining PCBs in use by 1999.

As mentioned previously, in 2001, UNEP (the United Nations Environment Programme) banished PCBs with "The Stockholm Convention on Persistent Organic Pollutants"[111]. This choice was made because of their capacity to bioaccumulate, their toxicities, their persistence and their capacity for long-range atmospheric transport[6]. Nevertheless, in 2010, in old fabrication sites of PCBs, the concentration is still measurable with quantities between 1 and 100 pg/m^3 in the air and between 100 and 1000 pg/g on the ground. In 2011, PCBs were still considered the sixth leading national cause of water quality problems in the United States[132].

2.2.4 Toxicity of PCBs, DDT and TCDD

From the presentation of this three OC, the reader might have understand that they are all toxic and persistent compounds. An index, named TEF (toxicity equivalent factor) was created, originally in 1997[133] at the "The World Health Organization" (WHO) assembly in Stockholm, and then reevaluated in 2006[134] to classified the OC as a function of their toxicity. This index, is a comparison of the TCDD toxicity where the TEF of TCDD is equal to 1, which is the most oxic of all the OC, known to have dermatological effects, immunotoxicity, cancer predisposition, reproductive, growth, and endocrine effects[134]. Thus, the TEF index is determined based on the so-called REP (relative effect potency), which corresponds to a power level of a single compound from a single study *in vivo* or *in vitro*, depending on the available data during the evaluation date. Then, to fix a TEF to a compound, it should produce a biochemical response or toxic when linked to aryl hydrocarbon receptors[133]. Concerning PCBs, twelve of them got a TEF index, and are considered as the most toxic PCBs. They are named dioxin-like PCBs, or DL-PCBs. These DL-PCBs have four chlorine atoms and are all mono or non-ortho-substituted. The twelve DL-PCBs and their TEFs are presented in the table 2.1.

N° PCB	IUPAC	WHO TEF 2005
PCB 126	3,3',4,4',5-PCBP	0.1
PCB 169	3,3',4,4',5,5'-HCBP	0.03
PCB 81	3,4,4',5-TCBP	0.0003
PCB 105	2,3,3',4,4'-PCBP	0.0003
PCB 114	2,3,4,4',5-PCBP	0.0003
PCB 118	2,3',4,4',5-PCBP	0.0003
PCB 123	2',3,4,4',5-PCBP	0.0003
PCB 156	2,3,3',4,4',5-HCBP	0.0003
PCB 157	2,3;3',4,4',5'-HCBP	0.0003
PCB 167	2,3',4,4',5,5'-HCBP	0.0003
PCB 189	2,3,3',4,4',5,5'-HPCBP	0.0003
PCB 77	3,3',4,4'-TCBP	0.0001

Table 2.1: Table Summary of different TEF indexes of the twelve DL-PCBs, with the reference index being the TCDD TEF equal to 1. It indicates the BZ number and the IUPAC name with the TEF index.

Among the twelve DL-PCBs, PCB 126 is evaluated as the most toxic, and to confirm this, in 1995, it was shown that it might be responsible for tumours[135]. It also as been demonstrate that PCBs have impact on animals health[136 137], and plants by reducing their ability to assimilate soil elements[136] and their biosynthesis[138,139].

DDT does not have a TEF index, which signifies that it is less toxic than TCDD and DL-PCBs. However, different studies were made on animals like birds, they identified hyperplasia of the thyroid, changes in plasma steroids, and brain weight[140]. The effect of DDT on human health is a controversial topic[141]. Indeed, thanks to its capacity to fight malaria and the fact that each year 880,000 people die from it, the health effect that it might provide is maybe worth the use.

2.3 Main transport routes of organochlorides

This section presents the main transports routes of the organochlorides. In reason of their large field of applications, this section focuses mainly on the PCBs transport. A presentation of all the different transport routes of PCB is given in the figure 2.5.

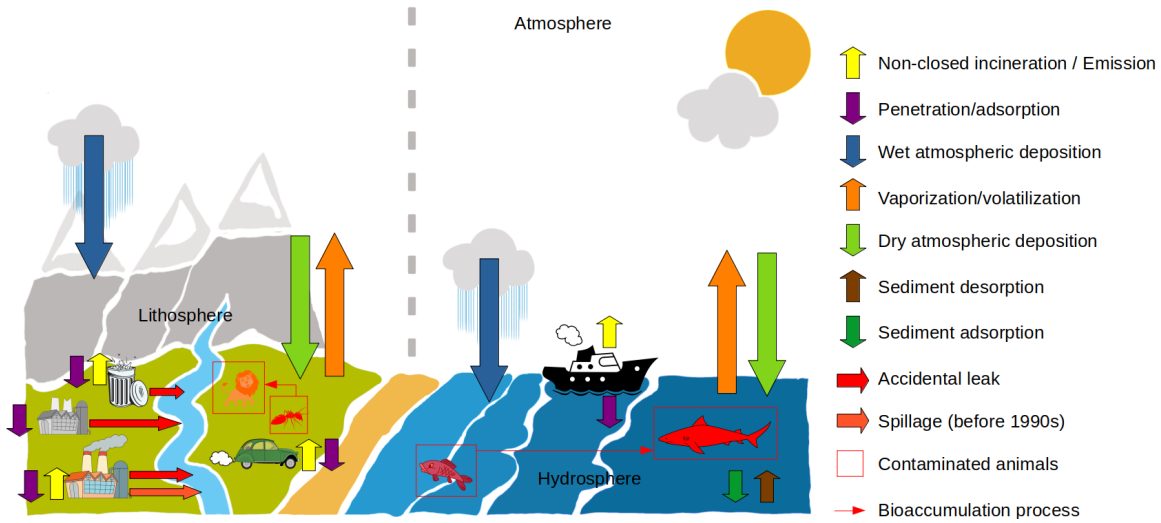


Figure 2.5: Schematization of the diverse transport routes of the PCBs in the environment. On the right, it is indicated a legend of all the different ways of transport. In order to simplify the visualization of the representation, the bin symbolizes a landfill, a closed factory is represented in black and white, and a working factory is also represented, the transport is indicated by a car on the lithosphere and by a boat on the hydrosphere. The bioaccumulation is indicated along the trophic chain with the ant and the lion on the lithosphere and the fish with the shark on the hydrosphere.

Firstly, DDT, TCDD and PCB are three organochlorides which, thanks to their chemical properties, are extremely persistent in both biotic and abiotic environments [136 142 108]. Before their production stopped, PCBs could have been released into the air, water and soil through accidental leaks, losses during transport, and spills into oceans, rivers and streams [143]. DDT was used principally as herbicides, so it has been released into the environment like the PCBs [144 145]. The TCDD is never industrially produced; it spreads in the environment mainly by the orange agent used in Vietnam [146] or by synthesising other herbicides. Thus, its persistency, like the other organochlorides, makes it spread in different environments [147].

Focusing on the PCBs, their main transport route is the atmospheric [148]. Indeed, even today, despite major restrictions on PCBs, they can be released into the atmosphere due to poor recycling of old materials/products. Thus, the destruction by non-closed incineration of this waste is at the origin of the dispersion in the atmosphere of PCBs and also of even more dangerous secondary products [143 136 148]. A study in 2003 showed that in 5 different municipal incinerators, PCB concentrations between 0.01 and 1.5 mg/kg could be measured. Leakages of old products such as capacitors or transformers [143 131], can also volatilise PCBs into the air from landfills. Unfortunately, these are not the only sources of PCB contamination in the atmosphere. During their manufacture and even afterwards, evaporation from paints, coatings, plastics, accidental leaks during transport, and contaminated water have also contributed to the dispersion of PCBs in the atmosphere [131 136 149].

Concerning the PCBs penetration on soil, it happens mainly on the PCBs elimination sites and after atmospherical deposition [136]. Thus, when the soil is rich in carbon, the adsorption of the PCBs is greater. However, inverted, the vaporisation from the soil to the atmosphere happens when there is a low carbon concentration

in the soil. Moreover, the vaporisation increases after rain thanks to a co-evaporation with water[150].

Looking at their concentration, PCBs are found in the atmosphere in higher concentrations close to their production site when they were still opened, or now from the old factory[148]. Moreover, the PCB concentration is higher in urban areas than in rural areas, with a factor of 20, according to Reddy et al.[136]. For example, an American city has a concentration of 0.5mg/m^3 against 0.02ng/m^3 in the land. Finally, the indoor concentration is generally ten times higher than outdoors[148]. Measurements have also shown that PCBs are also present in the Arctic and the Antarctic poles with concentrations around 0.2ng/m^3, whereas they have never been used there.

Due to PCBs' volatility and solubility differences, there is a redistribution on the water surface and in the sediments in the hydrosphere. According to Soedergren's study in 1990[151], the concentration of PCBs at the surface can be up to 500 times higher than in deeper areas. If the PCBs dissolved on the surface are in high concentration, combined with a strong evaporation effect during a sunny day, the PCBs can be vaporised into the atmosphere. In 1986, a study on Lake Michigan showed that vaporisation was the most important process for removing PCBs from the lake[152]. In 1990, Maclachlan et al.[153] also showed that this vaporisation phenomenon was amplified when the water was more aerated, for example, in waterfalls.

The most chlorinated PCBs in water can be adsorbed on the surface of sediments, but also on organic matter in suspension and on clay[136 131]. This adsorption significantly reduces their vaporisation and moves the PCBs from water to a solid state. In 1995, a study showed a two-to-one ratio between PCBs associated with suspended particles and PCBs dissolved in water[154]. This transition to the solid state by adsorption, followed by sedimentation, leads to the immobilisation of PCBs for very long periods in the sediments. Thus, aquatic sediments are natural reservoirs of PCBs, releasing them progressively by re-dissolution, which can re-evaporate into the air[131 136]. These compounds are much less volatile than surface PCBs, so this process occurs mainly in dry periods.

In case of heavy rainfall, PCBs can also be transported from the lithosphere to the hydrosphere by erosion[131].

About PCBs present in the atmosphere, they can be redistributed in the environment by two different types of deposition, dry and wet[155 156]. Wet deposition is the main deposition source, accounting for 58% of the total deposition flux, occurs through rain, snow and fog[131 136]. As a result of the meteorological conditions, wet deposition occurs mainly by recovery of PCBs fixed on particles (98%) rather than from the gas phase (2%)[157]. As for dry deposition, it is mainly done by gravitational settling after the PCBs have attached themselves to particles such as dust or soot. An example of the amount deposited in 1990 on the Chesapeake Bay in the United States, an annual concentration of 1.9μg/m^3 was recorded by dry and wet deposition[156].

Finally, the last pathway of contamination of PCBs in the environment, which is also the largest source of

human contamination, is bioaccumulation along the food chain[143,158]. Indeed, thanks to their high solubility in lipids, PCBs can accumulate in the various tissues of animals and thus extend along the food chain[136]. Of course, bioaccumulation will depend on the feeding area of the animals. For example, for marine animals, those feeding at the surface will be more likely to accumulate more PCBs than those feeding at depth. Similarly, since the concentration in sediment is much higher than in water, fish feeding in this area will bioaccumulate more. Studies of the bio-concentration factor have shown that in the PCB family, the factor increased for hexachlorobiphenyls, with some of them among the more toxic[159,160]. Moreover, it has been shown that in the lithosphere, insects play a non-negligible role in the PCBs bioaccumulation[161]. Furthermore, for humans, the contamination is mainly done through their food via meat, fish, and dairy products[136]. For example, in water containing 1 ng/L of PCBs, a fish living in this area will contain 10^5 ng/kg. Thus, a human consuming about 1000L per year will be exposed to 1000ng of PCBs, compared to 10kg of fish for a total exposure of 10^9ng of PCBs. Thus fish consumption will be 1000 times more contaminant for humans than water consumption[158].

In conclusion, PCBs can be found in many different environments, such as water, atmosphere, sediments, soil, at low and high temperature. Their degradation in these different environments can be investigated through the implementation of various computational strategies using several theoretical chemistry methodologies. Some investigations on the degradation in atmospherical conditions were done through quantum chemistry calculations, in gas phases or in vacuum. However, quantum chemistry calculations cannot be applied to other environmental degradations due to the size or complexity of the system if one wants to include an explicit solvent or the surface or interface with a solid material. For these systems, usually, molecular mechanics calculations are used, and to investigate the reactivity, reactive molecular dynamics simulations are usually implemented. As was presented in the first chapter, it is able to simulate the reactivity of complex systems in different environments. That is why developing a new ReaxFF force field for OC will allow studying their reactivity in all these different environments, and it might help design new removal methodologies in the future. The creation of this new ReaxFF potential is the main objective of the following chapter. Moreover, to develop this model, we need to understand the different PCBs properties presented in the following sections and, more importantly, the theoretical investigations already made.

2.4 Physical and chemical properties of PCBs

In this section, the physicochemical properties of PCBs will be presented. A product consisting of a single PCB will be colourless and odourless, but mixtures of PCBs are generally a little yellowish, and sometimes a little darker[148,162].

It is not easy to generalise properties for PCBs due to their diversity and the compositions of the mixtures.

However, the number of substituted chlorine atoms for single PCBs, or the degree of chlorination for mixtures, play significant roles in their properties.[162 131].

A fairly common characteristic of most PCBs is their ubiquity in the environment[125]. In fact, they are said to be non-reactive by being resistant to acids, bases and heat[136 162]. This is why they are difficult to biodegrade[122] and can be maintained in ecosystems for a very long time[163].

The degree of chlorination in the PCB mixtures impacts the viscosity. The higher it is, the more viscous the solution will be[148 162]. Similarly, mixtures with low chlorination levels will have higher vapour pressures and be more volatile. The densities of the mixtures are significant with values between 1.18g/cm^3 and 1.56g/cm^3 at 25°C[162]. In 2000 Johnson et al.[122] published an article in which the composition of the most popular PCB mixtures is indicated, composed of between 100 and 140 PCBs.

When looking at PCBs alone, it is possible to make similar conclusions. That is, the solubility in water decreases with the number of substituted chlorines with LogK_{ow} values ranging from 5 to 6.9[148]. However, the PCBs with the highest solubilities are those with chlorines in ortho positions[162].

As mentioned previously, PCBs are highly lipophilic, allowing them to pass through the food chain by bio-accumulating in fatty tissues and blood lipids[164].

The PCBs with the least chlorine atoms are also those with the highest vapour pressures[162 131]. It implies a dominance of these PCBs in the air. However, even those with the highest vapour pressures remain relatively low compared to benzene and naphthalene[148]. Thus, from their characteristics, even if the favoured route of transport is atmospheric, it can be assumed that a majority of PCBs are found in soils. In 2001, Erickson et al.[148] suggested that 99% of the global mass of PCB contamination in the environment is located in soils and sediments.

PCBs also have high molecular weights; taking ^{35}Cl as a reference, we obtain values ranging from 188 amu for mono-chlorinated PCBs to 493.7 amu for the ten-chlorinated PCBs[148].

Presenting the last physicochemical properties, they have high boiling temperatures ranging from 285°C to 456°C[148] and are fire resistant with flash points (lowest temperature at which a liquid can form a mixture with air capable of embracing) between 141°C and 196°C[162], both increasing with the number of substituted chlorine atoms. However, at very high temperatures, they can be combustible. They might produce harmful by-products, such as hydrogen chloride, PCDFs (Polychlorinated dibenzofurans)[162 131] and PCDDs (Polychlorinated dibenzodioxins) like the TCDD.

2.5 Computational studies related to the reactivity of PCBs

The objective of this section is to present the molecular properties of PCBs by different factors, thus allowing us to discuss their reactivities. That is why it is first presented some molecular properties of the

biphenyls, which will be helpful to detailed PCBs molecular properties.

2.5.1 Biphenyl molecular properties

The biphenyl is at the origin of PCBs. It is interesting to identify the different molecular properties usually investigated and do the same for PCBs, are compared them. In this idea, in 2001, Arulmozhiraja et al.[165] published a review of the molecular properties of biphenyl to validate the use of DFT (Density-functional theory) to study PCBs. In this review, it is compared various molecular properties obtained experimentally and theoretically. It includes intramolecular bond distances, angles and torsions. We are mainly interested in the central carbon-carbon bond between the two aromatic rings and the torsion angle between these two rings, which will be named ϕ. In terms of energy, the rotational barriers necessary to obtain the plane conformation of the biphenyl (ϕ = 0°) and the perpendicular one (ϕ = 90°) were investigated. Finally, the ionisation energy and the electronic affinity are studied. The computation of the rotational barrier is done thanks to the two equations 2.1 and 2.2. A representation of the perpendicular and planar structure of the biphenyl is given in figure 2.6.

$$\Delta E^0 = [E(\phi = 0) - E(\phi = \phi_{eq})] \tag{2.1}$$

With ΔE^0 the rotational barrier to reach the planar conformation, with $E(\phi = 0)$ the energy of the planar conformation and E($\phi = \phi_{eq}$) the energy of the equilibrium structure.

$$\Delta E^{90} = [E(\phi = 90) - E(\phi = \phi_{eq})] \tag{2.2}$$

With ΔE^{90} the rotational barrier to reach the perpendicular conformation, with $E(\phi = 90)$ the energy of the perpendicular conformation.

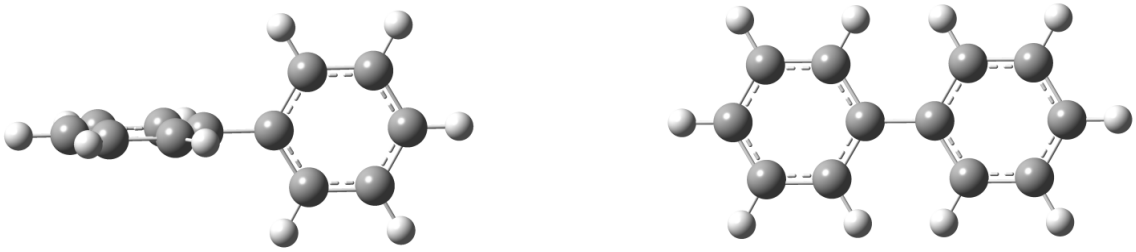


Figure 2.6: Here it schematizes on the left the perpendicular structure of the biphenyl and on the right the planar structure. The carbon atoms are represented in grey and hydrogen white.

In table 2.2, the different molecular properties have been summarised with indications of the method used and, if necessary for the theoretical studies, the basis set used. For simplifications, only 3 of the multiple theoretical calculations that were made by Arulmozhiraja et al.[165] are shown in the table 2.2.

Reference	Method	Basis set	C-C bond	ϕ	ΔE^0	ΔE^{90}	EA	IP
[166]	IR Exp.				4-8	4-8		
[167]	Raman Exp.				6.0	6.0		
[168]	ED Exp.		1.507	44.3±1.2	6.0±2.1	6.5±2.0		
[165]	HF	6-311G(2d,2p)		46.0	13.18	5.49		
[165]	MP2	6-311G(d,p)	1.479	45.3	15.78	6.57		8.21
[165]	DFT B3LYP	6-311+G(2d,2p)	1.484	40.1	8.07	8.02	0.028	7.86

Table 2.2: Table summarising the various studies carried out on biphenyl, with IR: Infrared, ED: Electron diffraction, C-C bond: distance (Å) of the two central carbons connecting the aromatic rings, ϕ: equilibrium torsion angle between the two aromatic rings (in degrees), ΔE^0: rotational barrier leading to the planar conformation (kJ/mol), ΔE^{90}: rotational barrier for the perpendicular conformation (kJ/mol), EA: electronic affinity (eV), IP: ionisation potential (eV).

All the results on the biphenyl torsions ϕ agree that they are all non-planar. The equilibrium angle around 45° results from a competition between the repulsion of the hydrogen in ortho positions favouring the perpendicular structure and the electronic delocalisation favouring the planar structure [169].

Moreover, as was concluded by Arulmozhiraja et al. [165], it seems that the DFT gives the most similar results to the experimental values. Calculation the ratio $\Delta E^{90}/\Delta E^0$, the experimental values and the DFT values are similar with ratio of 1.06 in DFT and 1.08 for experimental values [168]. The values obtained with HF and MP2 are in agreement with the ΔE^{90} experimental results but quite different from the ΔE^0. In their publication, they precise that increasing the basis set for MP2 leads to slightly better accuracy of the results.

Furthermore, the ionisation potential obtained with the DFT is in good agreement with the value of Rubio et al. [170] of 7.91eV obtained from the high level of computation CASPT2.

Thus, the good behaviour of the DFT on these calculations validates the used of this method for the PCB, and as the reader will see in the next chapter, this is one of the reasons we used the B3LYP method for the calculations of potential energy surfaces of torsion scans.

2.5.2 PCB molecular properties

From the results obtained on the biphenyl and the conclusion obtained, it is now presented a similar investigation on certain PCBs.

Before the main article studying biphenyl by Arulmozhiraja et al. in 2001 [165], some theoretical and experimental studies were already made on the PCBs. Indeed, an x-ray [171] experimental study in 1981 had shown that the most stable structure in the solid state of PCB with or without ortho substitution was non-planar, even if the planar structure has a better physiological activity. Thus they hypothesised that the toxicity might not be exclusively dependent on the structure planar structure of the PCBs. Moreover, they also supposed that one other determinant factor of toxicity is the presence of 4 chlorine atoms in lateral positions (non-ortho). This study is followed by a theoretical investigation in 1983 by McKinney et al. [172] with the Gaussian

N° PCB	IUPAC	C-C bond	ϕ	ΔE^0	ΔE^{90}	EA	IP	ϕ Cation	ϕ Anion
PCB 77	3,3',4,4'-TCBP	1.482	38.7	7.25	8.28	0.92	8.07	19.1	0.0
PCB 52	2,2',5,5'-TCBP	1.491	95.5	69.04	0.26	0.69	8.34	114.1	40.6
PCB 126	3,3',4,4',5-PCBP	1.483	39.2	7.39	8.33	1.04[a]	8.16	20.0	5.2
PCB 101	2,2',4,5,5'-PCBP	1.490	99.1	66.92	0.19	0.84	8.36	51.6	40.6
PCB 123	2,3',4,4',5- PCBP	1.487	58.8	27.66	2.31	0.95	8.20	40.3	30.8
PCB 169	3,3',4,4',5,5'-HCBP	1.483	38.9	7.17	7.75	1.27	8.26	20.5	0.0

Table 2.3: Table Summary of different studies on selected PCBs by Arulmozhiraja et al.[173,174], with the values computed at the level B3LYP/6-311+G(2d,2p) expect for the PCB-126 [a] where the basis set 6-311G(d,p) was sued, with the BZ number indicated and the IUPAC named, the C-C bond indicating the distances between the two central carbons bond (Å), ϕ: equilibrium torsion angle between the two aromatic rings (°), ΔE^0: rotational barrier leading to the planar conformation (kJ/mol), ΔE^{90}: rotational barrier for the perpendicular conformation (kJ/mol), EA: electronic affinity (eV), IP: ionisation potential (eV), and ϕ cation and anion the equilibrium torsion angle of the cationic and anionic structure (°).

80 software and the basis set STO-3G. They investigated the rotational barriers ΔE^0 and ΔE^{90} similarly to the biphenyl investigation of five PCBs. This early study concluded that the PCBs with no chlorine ortho-substituted have all similar equilibrium torsion angles ϕ around 42° close to the biphenyl value. They obtained similar rotational barriers with ΔE^0 = 3.6 kcal/mol and ΔE^{90} = 2.3 kcal/mol. However, most importantly, they conclude that the PCB with chlorine ortho-substituted have a shift of the equilibrium ϕ value towards the perpendicular structure increasing with the number of chlorine atoms ortho-substituted (80° for the PCB-4 di-ortho substituted, 57° for PCB-6 mono-ortho substituted), and also a huge increased of the rotational barrier through the planar structure with ΔE^0 = 48 kcal/mol. This rotational barrier increase was used as an explanation of the reduced biological activities of the PCB ortho-substituted with the impossibility of reaching the planar structure. In the same study series, in 1985, Bastiansen et al.[169] investigated through electronic diffraction the molecular properties of the PCB-15 non-ortho-substituted (4,4'-Dichlorobiphenyl). As for the previous study, they concluded that the torsion ϕ is almost not impacted by the chlorine substituted in meta and para positions with ϕ = 45.1°$\pm$1.5. They obtained similar rotational barrier ΔE^0 = 8.5Kj/mol and ΔE^{90} = 8.3Kj/mol.

Thus from this point, it can be concluded that the PCBs equilibrium geometry is not the planar structure, but as a function of chlorine-ortho-substituted, it can be the perpendicular structure. Moreover, the rotational barrier is an indicator of the flexibility of the PCBs and their capability to reach the planar structure, increasing their physiological activity associated with the number of chlorine atoms.

After the results of Arulmozhiraja in 2001[165], the choice of using DFT formalism for PCBs investigation was made based on the biphenyl investigation. In 2002, two articles from Arulmozhiraja et al. were published in which they presented the molecular properties of 6 PCBs obtained with B3LYP, and the basis set 6-311+G(2d,2p)[173,174] . The summary of the values from the two articles[173,174] are in the table 2.3.

These results concluded that the para and meta substitution have a low influence on the rotational barrier,

similarly to the early studies. The substitution of chlorine atoms in ortho positions leads to steric repulsions impacting the torsion angle and driving it towards the perpendicular structure, related to high ΔE^0 values, preventing the planar structure from being reached. The three PCBs 77, 126 and 169 with no-substituted chlorine atoms in ortho positions have almost planar anionic structures. In correlation with the positive electronic affinity, it suggests that the PCBs interact with biological cells as electron acceptors. It leads them to the anionic structure and, for the three PCBs cited above, leads to a planar structure, allowing them to be easily linked with the cells and, thus, more toxic. Another piece of information that can be extracted from table 2.3 is that the bond distance between the two central carbon atoms is not impacted much by the number of chlorine atoms substituted.

One year later, in 2003, Pathasarathi et al.[175] published a theoretical study on the PCBs 52 and 126 with B3LYP and the basis set 6-31G*. They show that the charge transfer between nucleic acid bases with PCBs occurs from the bases toward the PCBs. This demonstrates the electron acceptors' behaviour of the PCBs, as hypothesised previously.

In 2005, Padmanabhan et al.[176] published a similar article based on the work of Pathasarathi et al.[175] on the molecular properties of PCB 169 depicted figure 2.7 representing it. It gives complementary information on the origin of the titled ϕ angle resulting in the competition of the π orbitals of the two aromatic cycles and the repulsion of the atoms in ortho positions. It is concluded that in the positions ϕ = 0°, 90° and 180°, the reactivity of the PCB 169 increased due to a higher instability by computing the hardness, the electrophilicity, the polarisability and the electronegativity. Again, they demonstrate the acceptor electron behaviour of the PCB 169 and add that the charge transfer is more important in the three positions ϕ = 0°, 90° and 180°.

They went into more detail on the reactivity of PCB 169 by computing the local electrophilicity of the molecule to identify the most reactive sites.

Here, the electrophilicity of carbon atom 4 (and thus also 10 by symmetry) is higher than any other carbon. This indicates a greater reactivity in these two sites, and they correlate it with toxicity by concluding that the toxicity is higher in the two sites, chlorine atoms 21 and 18.

Some years later, in 2014, Gorbunova et al.[177] investigated the reactivity of all the 209 PCBs theoretically with B3LYP and the basis set 6-31G(d). They demonstrate that the nucleophilic substitution involving PCB occur through a reaction of hard acid hard base. On different PCBs, they investigated the charge distribution. They conclude that the nucleophilic attacks are favoured on the aromatic cycles which are the most substituted due to an increase of the carbon atoms' charges.

From all these investigations, non of them have at equilibrium a planar structure, but the PCBs, ortho-substituted, can have an equilibrium perpendicular structure. The equilibrium structure depends on the competition between maximizing the overlap between the two aromatic cycles' π orbitals and the ortho atoms'

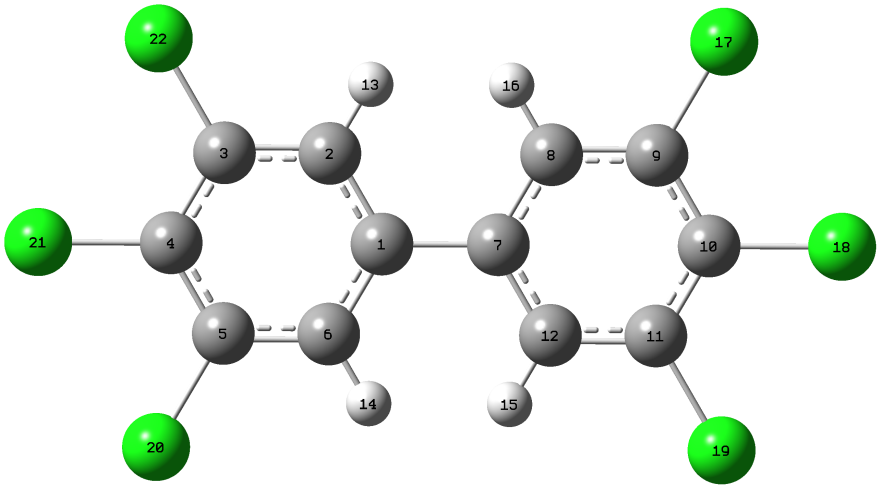


Figure 2.7: Representation of PCB-169 with the numbering of the atoms. The grey atoms are carbon, the white ones are hydrogen, and the green ones are chlorine.

steric repulsions. The number of chlorine atoms substituted does not impact the central carbon bond distances. The toxicity is governed by the capability of the PCB to be flexible, have low rotational barriers to reach the planar and perpendicular structures, and the number of chlorine atoms. Finally, the charge distribution can help to determine the favoured nucleophilic sites, and the electrophilicity showed that for PCB 169, the most reactive sites and, thus, the most toxics are the para-sites.

It is also important that some molecular dynamics investigations were done on the PCBs. For example, the work of Jensen et al.[178], who investigated the selectivity of PCB adsorption from fish oil onto activated carbon and concluded that the planar orientation of PCB molecule is energetically favoured with the graphite sheet thanks to the π-cloud overlap. Alternatively, the work of Chana et al.[179], who used MD for PCB conformations searches. These works are usually interested in the PCBs' structural arrangement and use classical molecular mechanics force field, which is the most adequate for this. Thus, to our knowledge, there is no investigation of PCBs reactivity through molecular dynamics with the complete degradation process tracked. It allows us to introduce the following chapter with the development and publication of the new ReaxFF force field developed to investigate the reactivity of OC in different environments. Moreover, in the conclusion of this reactivity investigation through the molecular properties of the PCBs, we determine key parameters as the torsion angle ϕ or the rotational barriers ΔE^0 and ΔE^{90}, that the new ReaxFF force field should be able to reproduce correctly. They will be mainly used in the validation and performance evaluation of the new force field.

2.6 Removal of organochlorines

This last section presents different methods used to remove organochlorides from an aqueous environment, atmosphere and under high temperatures. The methodologies presented below were chosen due to their link with the theoretical investigations that were made with the new ReaxFF force field. However, a detailed description of different removal methodologies of the organochlorides is presented in the annexe.

It exists many different removal methodologies to remove organochlorides. Usually, in the case of water treatment, these processes include the generation of strong oxidizing active species, such as hydroxyl radicals (OH), able to oxidize the pollutants until their transformation into non-toxic compounds. Different methodologies exist to generate OH radicals. For example, one can use the Fenton process[180], which uses ferrous iron to catalyze the decomposition of the hydrogen peroxide (H_2O_2) following the reaction 2.3.

$$H_2O_2 + Fe^{2+} \rightarrow OH^{\bullet} + OH^{-} + Fe^{3+} \tag{2.3}$$

The different methodologies used to generate are presented in the review of Oturan et al.[180]. Nevertheless, The use of OH radical generation, with an advance Fenton process name electro-Fenton, was tested by Dominguez et al.[181] on the OC lindane, which seems to be efficient in degrading it. The evolution of this method made them available at the industrial level, and a better understanding of the degradation mechanisms will be of great interest. In a similar context, in the literature, using molecular dynamics with a ReaxFF force field, the tracking of the degradation process of the paracetamol using OH radicals in aqueous solutions was made by Brault et al.[182]. However, there is no investigation of the OC reactivity in aqueous environments with hydroxyl radicals through theoretical investigations. It motivates the creation of the new ReaxFF force field presented in the next chapter III to track the degradation of OC in an aqueous environment. Hence, after the force field creation, it was tested at the end of chapter III by an investigation of the chlorobenzene degradation process through reactive molecular dynamics in the presence of hydroxyl radicals in an aqueous environment.

Considering now the removal of OC in the atmosphere, with the example of chlorobenzene, it is possibly removed from the atmosphere by its reactions with OH, NO_3, and O_3, photolysis and deposition[183]. It was shown by Atkinson et al.[184] that the radicals OH seem to be the major molecules responsible for the removal of chlorobenzene in the atmosphere. That is why, the atmospherical degradation of chlorobenzene was investigated in chapter III in the presence of hydroxyl radicals.

Finally, as was mentioned in the previous section on the transport routes of organochlorides, PCBs can still be measured in municipal incinerators[131]. Thus, the combustion of the PCBs is, nowadays, probably an important removal methodology for OC. Moreover, it was shown that pyrolysis was an efficient method to

remove OC[185]. However, the investigation of the products formed high temperature degradation of OC is of great interest to verify that they do not degrade in more toxic compounds. Thus, in chapter IV, to evaluate the performance of the new force field in that kind of environment, pyrolysis investigations were done using the new reactive force field, with a dedicated section on the chlorobenzene OC.

2.7 Conclusion

This chapter presents three organochlorides with a major focus on the PCB. This focus was made because of its utilisation in the next chapter to develop the new reactive force field for organochlorides. A presentation and a historical overview were presented in the first section of the three organochlorides finishing with their toxicities. It is followed by their transport routes, indicating their presence in all the different environments. Then a presentation of the physical and chemical properties was proposed for the PCBs. Afterwards, from the computational studies led on the PCBs, we identify the most interesting chemical properties that define the PCB, such as the central torsion angle between the two aromatics ring of the PCBs, the rotational barrier energies, and the usage of the electrophilicity and the charge distribution to predict reactivity. These fundamental elements are used in the following chapter to obtain an optimal ReaxFF force field for OC. This chapter ends with a review of several technologies used to remove organochlorides, which are directly related to the following investigations made with the new force field in the next chapters.

Chapter 3

Development of a new reactive force field

3.1 Introduction

This chapter presents the methods used to optimize a new set of parameters of a reactive force field (ReaxFF). As it was presented in the first chapter, in 2001, Adri Van Duin et al.[53] published the first ReaxFF force field. Indeed, they allowed to investigate the reactivity of molecules through molecular dynamics. They developed this new force field based on the so-called bond order, breaking and creating bonds. It was initially designed to study the combustion of organic compounds with carbon and hydrogen atoms. However, it rapidly increased its field of application in 2008 after the CHO reactive force field publication by Chenoweth et al.[63].

In this thesis, a new ReaxFF was developed to investigate the degradation of organochlorides presented in the second chapter. From the different force fields already published, the CHON-2017_weak force field[68], the latest aqueous force field coming from the glycine force field of 2011[80] with an accurate description of water, was looking as the best starting point. Indeed, this force field only needs to include the parametrization of the chlorine atom, which was only present in the force field as a counter ion.

This chapter is divided into different sections and presents the methodology applied to optimize the chlorine parameters compatible with carbon, hydrogen and oxygen of the ReaxFF CHON-2017_weak. The first section will be dedicated to the description of the training and validation sets. The second will present the different optimization implemented to optimize the parameters and how they were validated to create the new CHONCl-2022_weak force field. Much information is given in this chapter, and thus a final general workflow of the force field optimization procedure, which can be used in parallel with the chapter reading, is presented in figure 3.26. Finally, the performance of the force field and its final validation were published in 'The Journal of Chemical Physics, and the article is included at the end of this chapter.

3.2 Training and Validation sets

As it was made before in the creation of reactive force field,[53,63], the parameters are optimized against quantum mechanics calculations. Indeed, the parameters are fitted to match the quantum data. That is why it is necessary to have a well-made training set of reference data to forge the parameters to replicate them correctly. Also, we need a validation set to verify that the parameters are not over-fitted on the training set data and reproduce well all the reactivity of the chosen compounds. So, multiple choices should be made to fabricate the two sets. The first one, next presented, is the choice of molecules used. Then, the choice of the different QM methods that we used to produce reference data from these molecules and the different data types. And finally, we present the concatenation of all these data on the two sets.

3.2.1 The choice of molecules

The choice of molecules is the first step of the training and validation sets fabrication. Indeed, these sets should have all the elements as bonds, angles, and torsions necessary to reproduce the system we want to study. Here, by looking at the organochlorides, with the example shown in the figure 3.1 of a PCB, it is possible to set up a list of all the bonds, angles and dihedrals associated with the chlorine atom that we need to have in our training set. Moreover, as the final goal of this force field would be to investigate the reactivity of organochlorides, it is also essential to add to the sets molecules or chemical fragments that may appear during those reactions if the initial force field does not already describe them.

A list of all the elements that we should describe in our training set, is shown in the table 3.1.

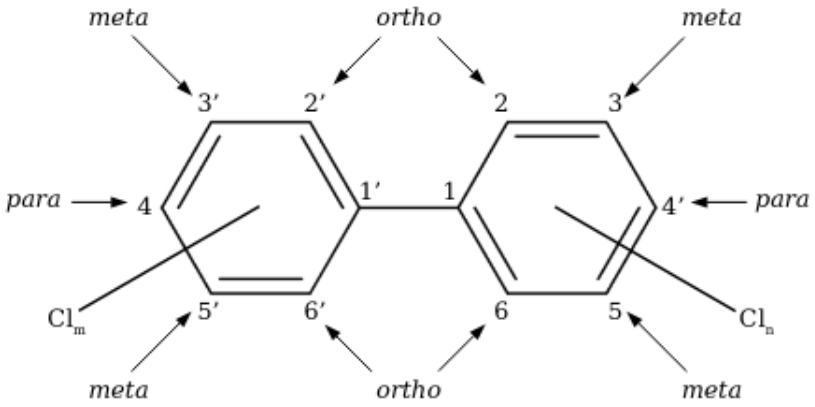


Figure 3.1: Schematisation of a PCB, the number represents the position of the hydrogen or chlorine atoms. Position 6 or 2 are named ortho, 3 and 5 meta and 4 the para.

From this list, it is possible to select molecules that contain all of these internal coordinates. The molecules are shown in the table 3.2 and represented in the figure 3.2 and were included in the training set or in the validation set.

Bond	C-Cl	H-Cl	Cl-Cl	O-Cl	
Angle	H-C-Cl	C-C-Cl	Cl-C-O	Cl-O-C	Cl-O-H
Dihedral	H-C-C-Cl	C-C-C-Cl	H-O-C-Cl	Cl-C-C-Cl	O-C-C-Cl

Table 3.1: List of the internal coordinates to include in the training set.

Names of the molecules constituting the training set or the validation set		
Chloroethane (C_2H_5Cl)	PCB5 ($C_{12}H_8Cl_2$)	Dichlorobenzene-para ($C_6H_4Cl_2$)
Chloroethanol ($HOCH_2CH_2Cl$)	Dichlore (Cl_2)	Dichlorobenzene-ortho ($C_6H_4Cl_2$)
Chloromethane (CH_3Cl)	Hydrogen chloride (HCl)	Dichlorobenzene-meta ($C_6H_4Cl_2$)
PCB12 ($C_{12}H_8Cl_2$)	Chloromethanol (CH_2ClOH)	Methyl hypochlorite (CH_3OCl)
Chlorobenzene (C_6H_5Cl)	PCB4 ($C_{12}H_8Cl_2$)	Hypochlorous acid (HOCl)
PCB101	PCB169	cis/trans-Dichloroethene (CH_2Cl_2)

Table 3.2: Table of all the different molecules used in the training set or validation set. For each, there is the name of the molecule followed by the molecular formula.

Moreover, some other molecules, such as some PCBs, were also used to validate the force field but were not involved in all steps using the validation set. They are presented in the next article for the validation of the force field.

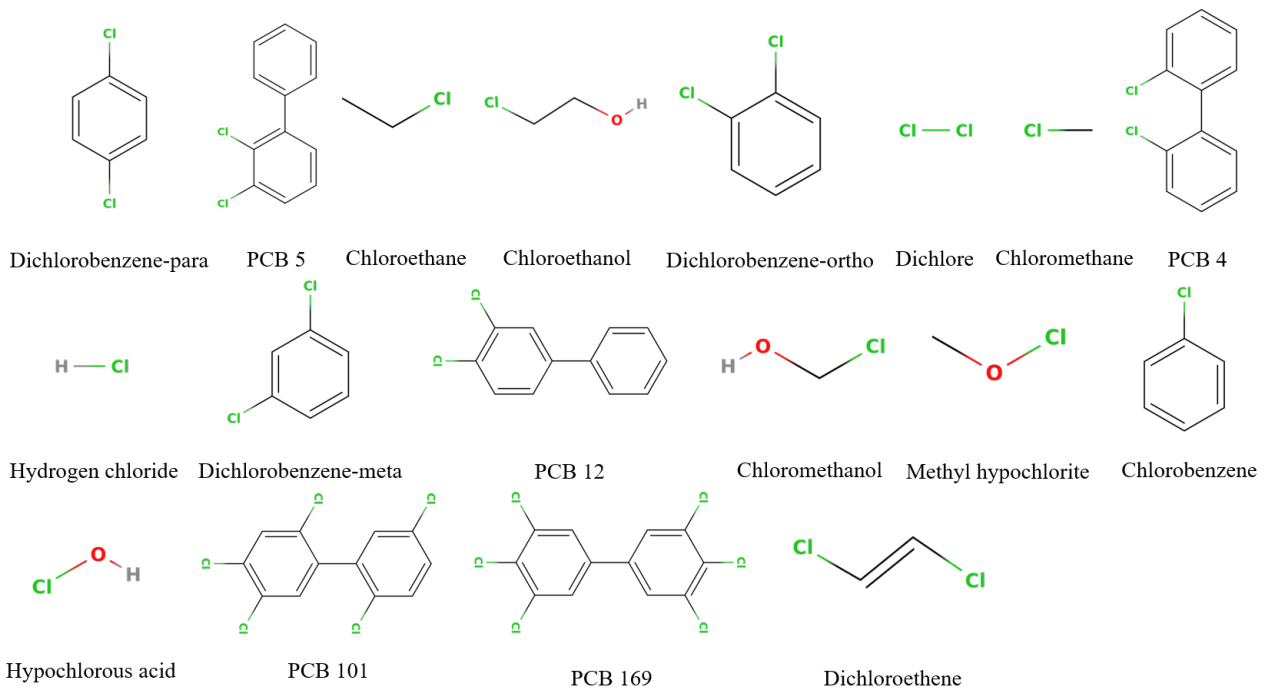


Figure 3.2: Schematization of all the different molecules directly used in the training set or the validation set. Chlorine atoms are represented in green, hydrogen grey, oxygen red, and carbon black.

Now that the molecules are chosen, it is necessary to decide what kind of data would be used to optimise the force field. The following section will detail the quantum calculations with the different methods linked to these data and the association with all these molecules.

3.2.2 PES scan and QM methods

As previously discussed, to forge the ReaxFF parameters, they are fitted against quantum mechanics data. This section reviews the different references used for the QM calculations. It is possible to obtain optimised molecule geometry, potential energy surfaces (PES), and many different properties through QM calculations. Here, quantum energies and geometries of the training set molecules will be computed to optimise the force field. Hence, first, the methodology used to compute the PES of bond dissociations is presented. It is followed by the methodology used for the potential energy surfaces computations of angle and torsion distortions. Afterwards, quantum geometries optimisations are presented, and finally, the method used to obtain PES of small reactions.

3.2.2.1 Bond Dissociation

One of the most essential data that should be well reproduced by the force field is the bond dissociations between the different couples of atoms. However, at the dissociations limit, the bonding and antibonding molecular orbitals become degenerates, which one electronic configuration cannot describe. Thus, singe-references methodlogies like density functional theory, or Hatree-fock calculations cannot describe accuratly. That is why, to fulfil this criterion, we decided to use the high level of quantum calculation CASSCF (complete active space self-consistent field)[186] method, which uses a multireference wavefunction to compute the accruate PES of the bond dissociation energies that we needed in our training set[187]. Moreover, the NEVPT2 (n-electron valence state second-order perturbation theory)[188] was used to take care of the perturbation theory.

Finally, we investigated seven bond dissociations using the CASSCF method. All the molecules that we used are listed in the table 3.3 with the bond associated.

Molecule Name	Bond investigated	Set
Dichlore	Cl-Cl	T
Chlorobenzene	C-Cl	T
Chloromethane	C-Cl	T
Chloromethanol	C-Cl	V
Chloroethan	C-Cl	V
Chloroethanol	C-Cl	V
Methyl hypochlorite	O-Cl	T
Hydrogen chloride	H-Cl	V

Table 3.3: Table of all the different bond dissociations investigated. T indicates that the data are in the training set and V in the validation set.

To do those calculations, we used the ORCA software version 4.2.1[189] , with the basis set cc-pVTZ[190], and followed those different steps :

1. Optimize with the MP2[191] method all the structures of each different molecule to get the optimized geometry.

2. By stretching the bond studied, we optimized with the MP2 method the fixed bond geometries necessary to make the PES.

3. Then we reused the finals orbitals obtained by the MP2 method to do non-relaxed CASSCF calculations.

4. The orbitals of the fixed CASSCF are then used to do relaxed CASSCF calculations.

5. Finally, from the results of these calculations, final CASSCF non-relaxed calculations with the NEVPT2 correction are done.

When doing a CASSCF calculation, the choice of the active space (number of electrons and the different orbitals) is essential and should be fixed for all the calculations of the PES. In the active space, it is important to have all different orbitals that may have an important role along with the bond dissociation. Usually, those orbitals are close to the HOMO and LUMO orbitals, but they can vary for extreme bond distances. The molecule's chemical investigation should help determine the orbitals that should be part of the active space.

However, one methodology that may help choose the orbitals is to do large CASSCF at extreme bond distances (short and long distances) for the molecule investigated and the optimized geometry and inspect the occupation of the different orbitals. The occupied orbitals with an occupancy superior to 1.98 should always be considered, as well as the orbitals with an occupancy superior to 0.02. Nevertheless, this methodology being time-consuming or sometimes inaccessible was only used for the smallest molecules.

In the table 3.4 it is presented all the different active space chosen for the CASSCF calculations.

Molecule	Bond	Active space
Dichlore	Cl-Cl	(12,10)
Chlorobenzene	C-Cl	(12,10)
Chloromethane	C-Cl	(10,9)
Chloromethanol	C-Cl	(14,14)
Chloroethan	C-Cl	(14,14)
Chloroethanol	C-Cl	(14,14)
Methyl hypochlorite	O-Cl	(14,14)
Hydrogen chloride	H-Cl	(12,10)

Table 3.4: This table presents the active spaces selected for the different CASSCF computations using the cc-pVTZ basis set. For the active space, the first number indicates the number of valence electrons distributed and the second the number of molecular orbitals.

3.2.2.2 Example of the CASSCF calculation for the chlorobenzene

The chlorobenzene is one of the molecules for which the large CASSCF did not work because of insufficient memory for the calculation. So, an investigation was made. In the literature, different chlorobenzene CAS calculations were made[192,193,194]. Some CAS(6,6) (6 electrons and six orbitals in the active space) were made

using only the π orbitals of the benzene ring, and a CAS(12,10) adding the two σ orbitals bonding and anti-bonding of the C-Cl bond and the π orbitals of the chlorine atoms. Furthermore, the publication of Liu et al.[194] presents the dissociation energy curve of the C-Cl bond in the chlorobenzene obtained by the CASSCF method with the CASPT2[195] correction, using the cc-pVDZ basis for the carbon and hydrogen atom and an effective core potential with seven valence electrons for the chlorine, which might explain the slight difference. To keep the consistency of the CASSCF calculations in this thesis, instead of CASPT2, we will use NEVPT2, and for all the atoms, the basis set cc-pVTZ is used. However, even with these slight differences, comparing the results obtained with Liu et al.[194] and the results computed with our method will allow us to validate our results.

So, it was decided to use the largest active space CAS(12,10). First, the orbitals of the chlorobenzene were analyzed with an example presented in figure 3.3 using the MP2 level of calculation and the cc-pVTZ basis set. For all the different distances of the PES, we identified the π orbitals of the benzene ring, the lone pair π orbitals of the chlorine atoms and the σ bonding and anti-bond orbitals of the C-Cl bond to include them in the active space. This study was made for all the different points of the PES. It usually implies rotating some orbitals to have those in the active space. For example, the LUMO+7 was switched with the LUMO+2 to be in the active space for the CAS(12,10).

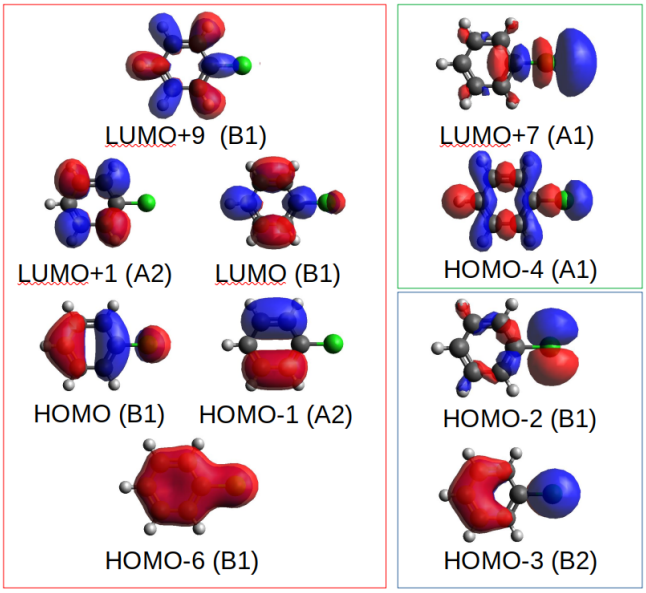


Figure 3.3: We decided to add all the orbitals in the active space for the CASSCF at the equilibrium point of the PES. In the red rectangle, there are the six π orbitals of the benzene ring; in the green rectangle, the σ bonding and anti-bonding orbitals of the C-Cl bond; and in the blue one, the lone pair π orbitals of the chlorine atom. Bellow each orbital is indicated its position and, in parenthesis, its symmetry.

Then, we follow the different step of calculations and we obtain the CASSCF/NEVPT2 PES of the C-Cl bond dissociation of the chlorobenzene.

3.2.2.3 Validation of the CASSCF calculation method

Finally, to validate our CASSCF calculations, it was possible to compare PES obtain from the C-Cl bond dissociation of the chlorobenzene with value obtain by Liu et al.[194].

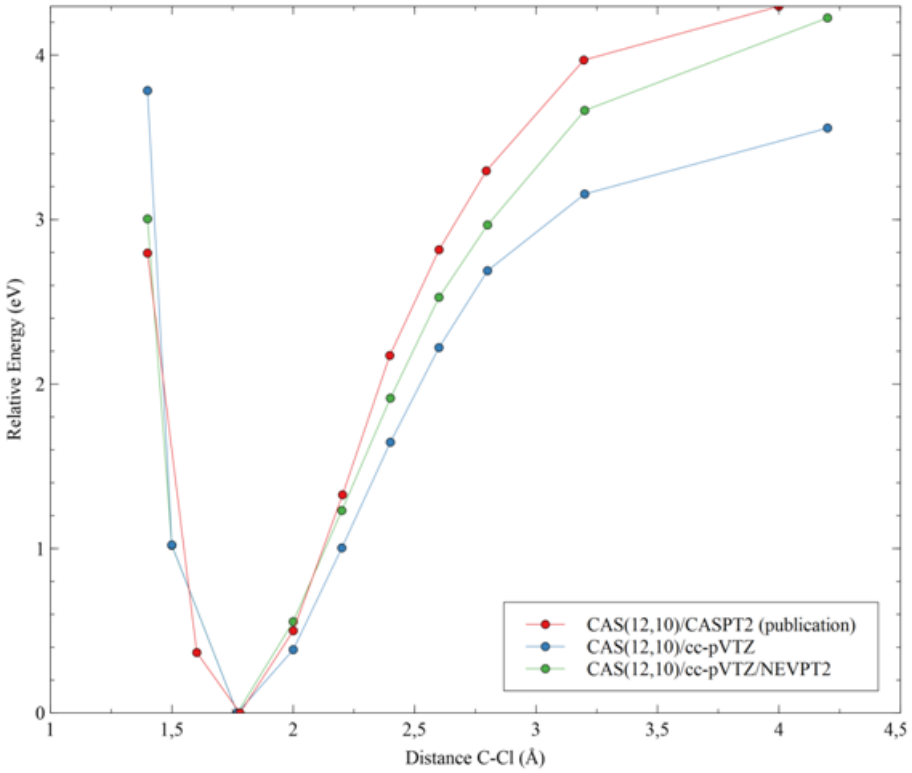


Figure 3.4: Comparison of the PES of the C-Cl bond dissociation of the chlorobenzene obtained from the CASSCF calculations in blue, with the value obtained from the CASSCF with the NEVPT2 correction in green and the value published by Liu et al.[194] with correction CASPT2 in red. The evolution of the Relative energy (eV) as a function of the distance C-Cl (in Angstrom).

Figure 3.4 presents the results obtained by Liu et al.[194], in green the results obtained from the CASSCF calculations with NEVPT2, and in blue without NEVPT2. We can see that the three curves have the same minimum for the bond distance C-Cl. Moreover, the computations with the perturbation theory have higher energies at the dissociation limits than the results without showing its contribution. More importantly, we can see that the two curves of the Liu et al.[194] results and the results obtained here with NEVPT2 are similar, with the slight difference coming from the slight different basis set and NEVPT2 instead of CASPT2, which can allow us to validate the methodology that we used to obtain our CASSCF results.

3.2.2.4 QM calculations benchmark

The force field must correctly reproduce the evolution of the different angles and torsion during the simulations. So, it must have quantum data in the training set that can train the torsion and angle parameters of the force field. We decided to produce these results using DFT (density functional theory). As it was mentioned in chapter II, most of the molecular properties computed for organochlorides were obtained with the B3LYP

functional[196,197,198,199], which is usually a good wave-function in terms of accuracy and time-consumption. Thus, the B3LYP functional was chosen for our DFT calculations.

However, a benchmark of basis set was made, and an example is presented in figure 3.5 with the three basis sets 6-311g(d,p), 6-311+g(d,p) and 6-31+g(d)[200,201,202], by investigating the bond dissociation of CH_3Cl. Thus, from this figure, the bond dissociations PES obtained with these three basis sets are similar. However, the basis set 6-31+G* reduced the time consumption and was chosen as the basis set.

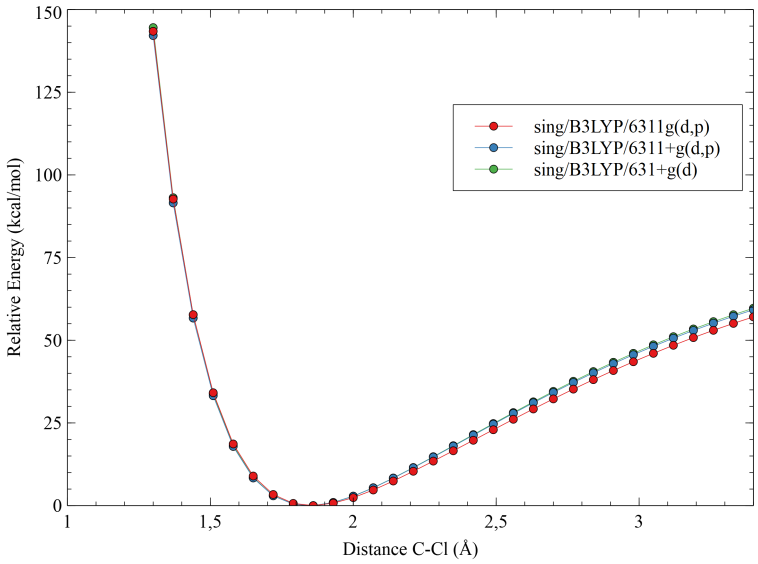


Figure 3.5: Comparison of basis sets for the C-Cl bond dissociation of CH_3Cl. Relative energy (in kcal/mol) as a function of the distance C-Cl (in Angstrom). In red is a singlet PES scan with the B3LYP method and the 6-311g(d,p) basis set, in blue the 6-311+g(d,p) basis set, and in green the 6-31+g(d).

Moreover, with the functional B3LYP and the basis set 6-31+g(d), to obtain better accuracy, the empirical dispersion correction GD3BJ[203] was used.

3.2.2.5 Torsion and angle scan

As for the bond, it is needed in the training set to have energetic descriptions of the different angles and torsions constituting the organochlorides. Hence, 16 additional scans of torsion and angle were made for the training set and the validation at the B3LYP/6-31+g(d) level of theory with GD3BJ correction. They are listed below in the table 3.5.

3.2.2.6 Geometry optimization

As mentioned before, it is necessary to add geometrical parameters and not only energetical terms in the training and validation sets to optimize the force field. The table 3.6 summarises all the different values added to the training set or the validation set.

The geometrical bond values were extracted from geometry optimization with the CASSCF level of theory

Angle		
Molecule Name	Angle or torsion investigated	Set
Dichlorobenzene-ortho	C-C-Cl	V
Chloroethane	Cl-C-C	V
Dichlorobenzene-para	Cl-C-C	T
Chloromethanol	O-C-Cl	T
Chlorobenzene	C-C-Cl	T
Chloromethane	Cl-C-H	T
Torsion		
Chloromethanol	H-O-C-Cl	T
Chlorobenzene	Cl-C-C-H	T
Dichlorobenzene-para	H-C-C-Cl	V
Dichlorobenzene-ortho	Cl-C-C-Cl	T
Chloroethanol	Cl-C-C-H	V
Dichlorobenzene-ortho	H-C-C-Cl	T
Dichlorobenzene-meta	Cl-C-C-H	T
Chloroethanol	O-C-C-Cl	T
PCB101	Cl-C-C-Cl	T
PCB169	Cl-C-C-Cl	V

Table 3.5: Table of all the different scans of angle and torsion made for the training or validation set. T indicates that the PES was added to the training set and V to the validation set.

(except Hydrochloride, obtained with B3LYP). For the angle and torsion geometrical values, it depends on the molecule. If the molecules appeared in Table 3.3 and were investigated at the CASSCF level of theory, they were extracted from these calculations. Otherwise, they were obtained from B3LYP calculations with the GD3BJ correction. The basis set used with the CASSCF method was cc-pVTZ and 6-31+G* with B3LYP.

3.2.2.7 PES of simple reaction

Finally, the last kind of quantum data added to the training set are PES of simple reactions involving bond formation that we to describe with the force field. The PES presented below will be used to validate the force field, except the PES concerning HOCl which was used in the training set.

In the table 3.7, it is presented the simple reaction that we added to the validation set or to the training set.

The energetic profiles of these reactions were obtained using the B3LYP/6-31+g(d) level of theory with GD3BJ empirical dispersion correction and using a using an in-house chain-of-States methodology[204]. The following figure 3.6 shows the only energetic profile of the reaction HOCl to HClO added to the training set.

Molecule Name	Parameter	Atoms involved	Value	Set
Hydrochloride	bond	H-Cl	1.29	V
Chlorobenzene	bond	C-Cl	1.77	T
Chloromethane	bond	C-Cl	1.81	T
Chloromethanol	bond	C-Cl	1.83	V
Chloroethane	bond	C-Cl	1.83	V
Chloroethanol	bond	C-Cl	1.82	V
Dichlore	bond	Cl-Cl	2.02	T
Hypochlorous acid	bond	O-Cl	1.71	T
Chlorobenzene	angle	C-C-Cl	119.3	V
Chloromethane	angle	H-C-Cl	108.6	V
Hypochlorous acid	angle	H-O-Cl	102.7	T
Chloromethanol	angle	O-C-Cl	112.7	T
PCB101	angle	C-C-Cl (ortho)	120.2	V
PCB101	angle	C-C-Cl (meta)	117.2	V
Dichlorobenzene-ortho	angle	C-C-Cl	118.8	V
Chloroethane	angle	C-C-Cl	111.4	V
Methyl hypochlorite	angle	C-O-Cl	110.4	V
Chlorobenzene	torsion	C-C-C-Cl	180.0	V
Chlorobenzene	torsion	H-C-C-Cl	0.0	T
Dichlorobenzene-ortho	torsion	Cl-C-C-Cl	360.0	V
PCB5	torsion	Cl-C-C-Cl	358.15	T
PCB4	torsion	C-C-C-Cl	181.35	T
PCB4	torsion	H-C-C-Cl	359.33	T
PCB12	torsion	Cl-C-C-Cl	0.56	T
Chloromethanol	torsion	H-O-C-Cl	67.4	V
p-dichlorobenzene	Cl-C-C-H	360.0	360.0	V
Chloroethanol	Cl-C-C-O	179.7	180.0	V
Methyl hypochlorite	H-C-O-Cl	180.0	180.0	V
Chloroethan	Cl-C-C-H	299.6	299.6	V
PCB101	Cl(para)-C-C-C	179.9	183.2	V
PCB169	Cl-C-C-Cl	359.5	359.9	V

Table 3.6: Table of all the geometrical elements added to the training or validation set indicated by T or V, indicating what kind of parameter it is, the atom concerned, and its value.

Reaction
$CH_3Cl + H2 \longrightarrow CH_4 + HCl$
$ClCH2OH \longrightarrow CH_2O + HCl$
$HOCl \longrightarrow HClO$
$Chlorobenzene + H_2 \longrightarrow Benzene + HCl$
$CH_2CHCl + H_2 \longrightarrow CH_2CH_2 + HCl$
$OCHCl + H_2O \longrightarrow HOC(O)H + HCl$
$CH_2CHCl + O_2 \longrightarrow OOCHClCH_2$

Table 3.7: Table of all the different reactions made to validate or add to the training set (only the reaction with HOCl is part of the training set).

3.2.3 Creation of the Training set with all the files associated

Finally, to create this training set and validation set, four different kinds of files were created. Two files named "geo" was constructed with all the different coordinates of all the molecules from the training or the

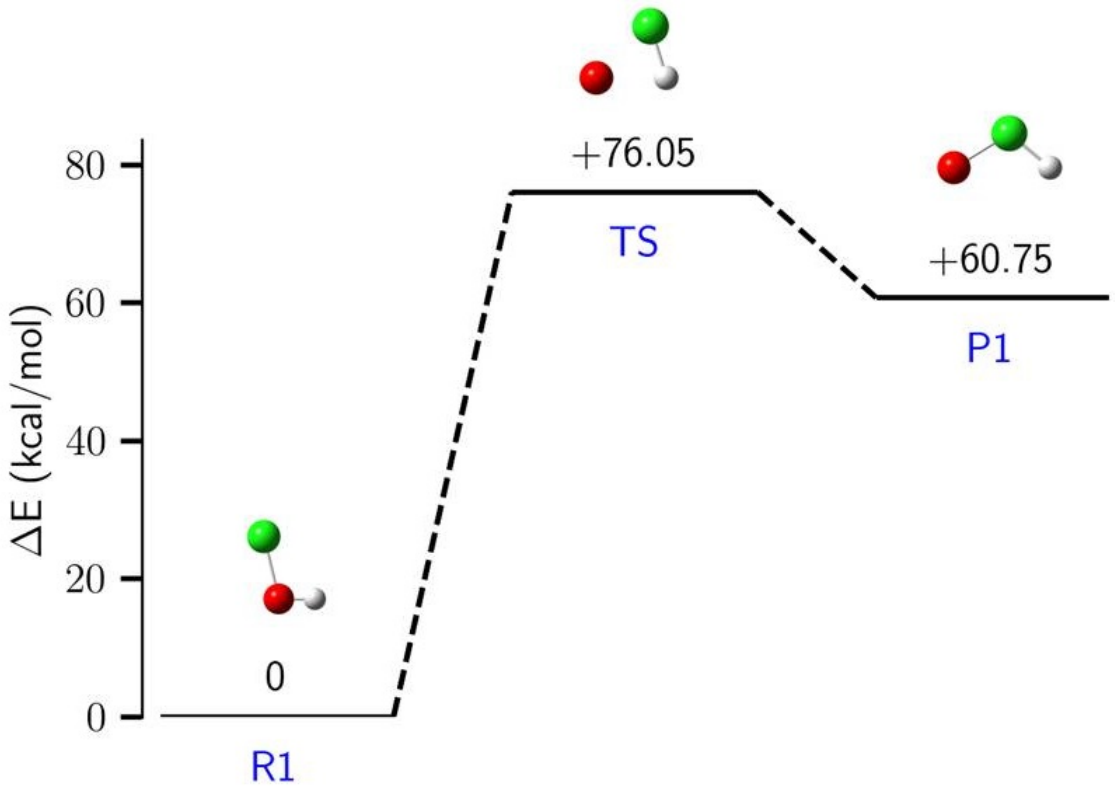


Figure 3.6: Here is the energetic profile following the reaction starting with HOCl (R1) to HClO (P1) with the transition state (TS). Above each level, the energy difference is indicated (in comparison with the R1 value) in kcal/mol and the image of the molecule (in green is the chlorine atom, in red is the oxygen, and white is the hydrogen).

validation set. The "trainset" (or validation) file contains all the different energetic and geometrical values which will train the force field parameters. To create the trainset and geo file, we develop a python script using the Pymatgen package[205], allowing us to extract the data from the different PES of bond dissociations, angle and torsions distortions, and the geometrical optimizations. In order to avoid overfitting the force field parameters, we automatically extracted a subset of points from the PES. Thus, if a PES has more than 15 points, the number of points was reduced by selecting only the extrema of the scans and adding a certain (until at least 15 points, which tends to describe the PES behaviour correctly) number of points between each extremum. An example of selection for an angle and torsion is given in figure 3.7.

To optimize the force field, we also need to define the initial parameters of the force field and, thus, the initial "force field" file. The last file needed is the "params" file which contains all the different parameters we want to optimize. We obtain all the necessary data to optimize the force field with all four files. It allows us to introduce the second main part of this chapter, force field optimization.

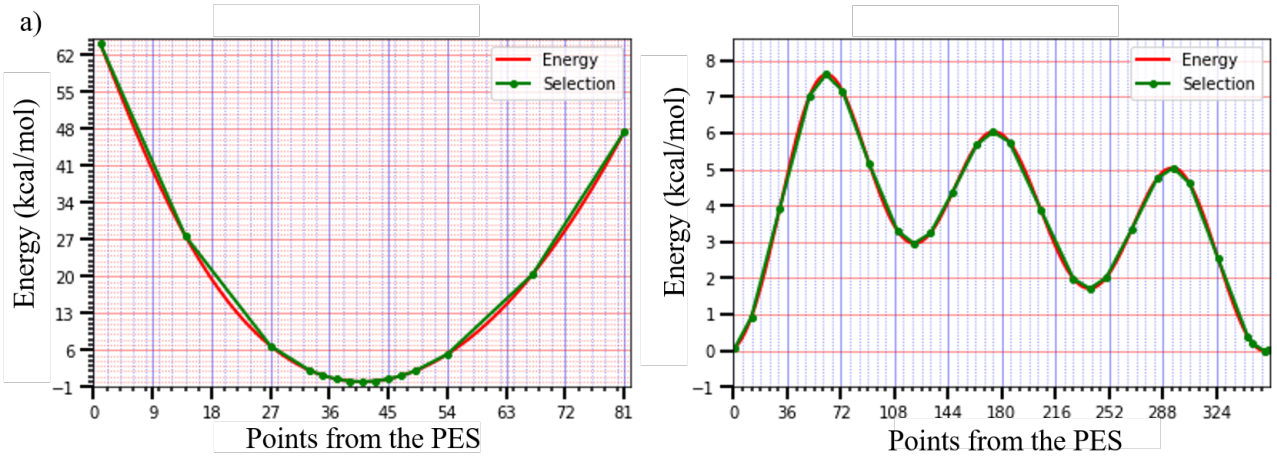


Figure 3.7: Representation of two examples of selected points from two different scans. On the left is the choice from the angle Cl-C-C of the dichlorobenzene meta and on the right is the torsion Cl-C-C-H of the chloroethanol. The red curve represents the full PES, and the green is the selected curve.

3.3 Optimization of the Force field

In this section, we present the different processes involved in optimising a reactive force field. We were faced with different problems, such as considering the correlation of specific parameters linked to the bond order or deciding the initial value of new parameters. We solved all the different problem encounters and presented our strategies in this section.

3.3.1 Choice of the optimizer

The 'easiest' way to optimize a reactive force field is to use a single-parameter search optimization, which was, for example, used by Chenoweth et al.[63]. Thus using a simple fitness (also named Error) presented in the equation 3.1, each parameter is optimized successively, reducing the difference between the reference values and the values computed with ReaxFF.

$$Fitness = \sum_{i=1}^{n} \left[\frac{x_{i,TS} - x_{i,ReaxFF}}{\sigma_i} \right]^2 \tag{3.1}$$

On the equation, 3.1 $x_{i,TS}$ is the reference value and $x_{i,ReaxFF}$ is the calculated value with ReaxFF, and σ is a parameter allowing to affect a weight to the different reference values (which usually depends on the accuracy of the reference value). Since 2008, different novel algorithms have been used to optimize reactive force fields, which are greatly reviewed by Shchygol et al.[206]. In this review, they compared the efficiency of three different optimizers, with one genetic algorithm optimizer (GA) developed by Dittner et al.[207], the second being a Monte Carlo force field parameter optimizer MCFF[208], and the last one the covariance matrix adaptation evolutionary strategy CMA-ES[209 210]. The first conclusion of this review is that none of these methods is reliable by doing only one optimization of parameters. Moreover, they all seem to have their

strength and weakness. The CMA-ES optimizer gives the lowest error for two of the three training sets they tried, but it can converge to a higher local minimum depending on the initial condition. GA seems to compute more reproducible errors from single runs, and both CMA-ES and GA give better results than MCFF. That is why we decided to use the CMA-ES optimizer, which is fully detailed in the next section. We consider the conclusion of the review of Shchygol et al.[206] and ensure that the optimizer does not converge to a higher local minimum by doing an iterative process with multiple CMA-ES optimizations. Moreover, the advantage of CMA-ES, by the correlation matrix, it can take into account the multiple correlations in the ReaxFF potential, with, for example, the correlation of all the intramolecular energetical terms with the bond order. This final note consolidates our choice of using the CMA-ES.

3.3.2 CMA-ES Initialization

Before beginning the CMA-ES optimization, several inputs need to be chosen. The choice of this starting point is crucial when using an optimization algorithm because it may condition the solution. Thus, it has to be chosen carefully. In the case of CMA-ES, it needs an initial set of weights for the different reference calculations of the training set to compute the fitness during the optimization. It also needs the initial parameter values and associated limits. Moreover, the CMA-ES optimization also needs a step size for each parameter that drives the exploration of the parameters space. All of these CMA-ES inputs are presented in this section.

3.3.2.1 Choice of weight

CMA-ES is a stochastic optimizer, and like all the evolutionary algorithms, CMA-ES is partly based on biological evolution and tends to obtain the optimal population, with the lowest fitness, iteratively with the equation 3.1.

In this equation, the differences between the values of reference calculations (TS) $x_{i,TS}$ and $x_{i,ReaxFF}$ the values calculated by the new force field are weighted over σ_i. In our case, the σ_i value is manually chosen as a function of the "importance" of the value associated. For example, the weight associated with the CASSCF calculations should be higher than those corresponding to DFT due to the higher level of theory. As was previously discussed, the number of points of each energetical scan should not be higher than 20 points. Otherwise, the number of points of the PES could bias the fitness by artificially increasing the weight of this PES and creating an over-fitting. So, it tends to obtain an optimal set of parameters that reproduces the reference values from the training set well. The table 3.8 presents the different values used for the different kinds of data.

Type of data	Weight
Energetical data	
CASSCF	1.0
PES	2.5
Structural data	
Angle and Torsion	5.0
Bond distances	0.1

Table 3.8: This table presents the weights used in the CMA-ES optimization as a function of the different data.

3.3.2.2 Choice of initial parameter

In a reactive force field, they are meaningful parameters like the bond dissociation energy or the bond radius of an atom, which can be determined by knowing their value in comparison with the other value of the other bond present in the force field. However, there are also "non-meaningful" parameters that are not directly related to a physical property and are usually correlated with other parameters, like all the bond order terms to obtain the bond order values. They are usually denoted with the letter 'p'. In this later case, if the parameter was never described before in the literature, or if the force field gives non-reliable results for the subject investigated, it will be necessary to choose an initial value. If possible, it could be a good idea to plot the energy term depending of these parameters. Indeed it will help to investigate their behaviour and might help to choose an initial value. As an example, for the bond order parameters, one can plot the bond order for the investigated bonds as a function of the distance and try to find an initial couple of parameters that give a reasonable result based on the knowledge of the correspondiong bond length. For example, in the figure 3.8, we reproduce the evolution of the bond order of the C-Cl bond described by the force field CHOSFClN[60].

Another option to select the starting values is to look for similar systems or elements in the literature. If the parameters describe a property similar to another chemical system close to it, the mimetism could be a good option for the initial guess.

In our case, we optimized 68 parameters presented in table 3.9. We used the methods mentioned above to determine the initial parameters starting with the aqueous force field CHON-2017_weak force field[68]. We optimize all the necessary parameters to create a reactive force field that can reproduce the different bonds, angles and dihedral of organochloride molecules and their reactivity. Some parameters were already defined in the CHON-2017_weak force field[68], but the chlorine atom was used as a counter ion, and these parameters were generally unreliable. We modified parameters from the CHOFSClN[60] for the bond parameter of C-Cl as initial parameters. For the rest of the parameters, we usually used mimetism from other elements in the CHON-2017_weak force field to guess the initial values. Moreover, we did some small simulations to find an initial guess that seemed to give reasonable geometries of the targetted molecules, and we used it at the initial force field.

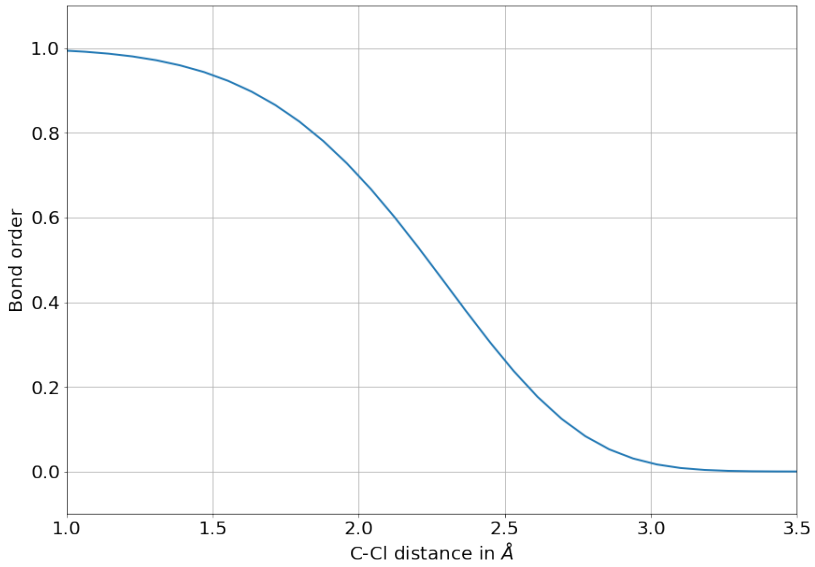


Figure 3.8: This graph represents the evolution of the bond order as a function of the C-Cl bond from the combustion reactive force field CHOSFClN[60]. The equation plotted is $\exp\left(-0.0827 \cdot \left(\frac{x}{1.5598}\right)^{5.9023}\right)$.

3.3.2.3 Choice of initial limit

After choosing the initial guess of the force field parameters, it is important to choose the associated boundaries. These boundaries define the intervals from which the parameter are drawn. For the initial parameter value, it was arbitrarily decided than between 6 and -6, we used the formulas 3.2 and 3.3.

$$min_{value} = x - \frac{x}{x * 3} \tag{3.2}$$

$$max_{value} = x + \frac{x}{x * 3} \tag{3.3}$$

With x the parameter value. For the initial parameter values superior at 6 or inferior at -6, the formulas 3.4 and 3.5 were used.

$$min_{value} = x - (x * 0,1) \tag{3.4}$$

$$max_{value} = x + (x * 0,1) \tag{3.5}$$

With these formulas, we determined the limits of the first steps to obtain the reactive force field. However, they should be sharpened to obtain an optimized force field, and it is presented after the CMA-ES optimization.

Index	Nom	Équation	Comment	intra/inter	OPT
General					
1	p_boc1	4c	Overcoordination parameter	intra	
2	p_boc2	4d	Overcoordination parameter	intra	
3	-p_coa2	15	Valency angle conjugation parameter	intra	
4	p_trip4	20	Triple bond stabilization parameter	intra	
5	p_trip3	20	Triple bond stabilization parameter	intra	
6	k_c2	19	C2-correction	intra	
7	p_ovun6	12	Undercoordination parameter	intra	
8	p_trip2	20	Triple bond stabilization parameter	intra	
9	p_ovun7	12	Undercoordination parameter	intra	
10	p_ovun8	12	Undercoordination parameter	intra	
11	p_trip1	20	Triple bond stabilization energy	intra	
12	n/a	21	Lower Taper-radius	inter	
13	R_cut	21	Upper Taper-radius	inter	
14	p_fe1	6a	Fe dimer correction		
15	p_val6	13c	Valency undercoordination	intra	
16	p_lp1	8	Valency angle/lone pair parameter	intra	
17	p_val9	13f	Valency angle parameter	intra	
18	p_val10	13g	Valency angle parameter	intra	
19	p_fe2	6a	Fe dimer correction		
20	p_pen2	14a	Double bond/angle parameter	intra	
21	p_pen3	14b	Double bond/angle parameter: overcoord	intra	
22	p_pen4	14b	Double bond/angle parameter: overcoord	intra	
23	p_fe3	6a	Fe dimer correction		
24	p_tor2	16b	Torsion/BO parameter	intra	
25	p_tor3	16c	Torsion overcoordination	intra	
26	p_tor4	16c	Torsion overcoordination	intra	
27	p_elho	26a	eReaxFF	inter	
28	p_cot2	17b	Conjugation	intra	
29	p_vdW1	23b	VdW shielding	inter	
30	cutoff * 100	3a,b	Cutoff for bond order (* 100)	intra	
31	p_coa4	15	Valency angle conjugation parameter	intra	
32	p_ovun4	11b	Overcoordination parameter	intra	
33	p_ovun3	11b	Overcoordination parameter	intra	
34	p_val8	13d	Valency/lone pair parameter	intra	
35	X_soft	25	ACKS2 softness parameter	inter	
36	unused	n/a	n/a		
37	p_val	27 via n_el	eReaxFF	inter	
38	n/a	13d	If =1 remove delta_j term for non C-C-C angles where none of the atoms is N		
39	p_coa3	15	Valency angle conjugation parameter		
40	n/a	20	Condition to turn triple bond option : vpar(40) == 1		
41	n/a	26 via Tap(R)	Ereax-specific taper radius for interactions with/between electrons and holes		

Index	Nom	Équation	Comment	Intra/inter	OPT
Atoms					
1	r_0^sigma	2	Sigma bond covalent radius	intra	y
2	Val_i	3a , 4b,5 b,9a	Valency	intra	
3	n/a	9a	Atomic mass		
4	r_vdW	23a	van der Waals radius	inter	y
5	D_ij	23a	van der Waals dissociation energy	inter	y
6	gamma_i	24	gammaEEM; EEM shielding	inter	y
7	r_0^pi	2	Pi bond covalent radius	intra	
8	Val_i^e	7,8,9	Number of valence electrons	intra	
9	alpha_ij	23b	van der Waals parameter	inter	y
10	1/gamma_w	23b	van der Waals shielding	inter	y
11	Val_j^angle	16c,13c	Valency for 1,3-BO correction	intra	
12	p_ovun5	12	Undercoordination energy	intra	y
13	p_i^xel2	26	eReaxFF, atom type parameter	inter	
14	chi_i	24,25	EEM electronegativity	inter	y
15	eta_i	24,25	EEM hardness	inter	y
16	n/a	n/a	Donor or acceptor switch in H-bonds	inter	
17	r_0^pi	2	Double pi bond covalent radius	intra	
18	p_lp2	10	Lone pair energy	intra	y
19	n/a	n/a	Atomic heat of formation	intra	y
20	p_boc4	4e,f	Bond order correction	intra	y
21	p_boc3	4e,f	Bond order correction	intra	y
22	p_boc5	4e,f	Bond order correction	intra	y
23	C_i	25	VdW Radius for Lg dispersion correction	inter	
24	alpha, alpha_i	26, 26a	eReaxFF, constant, dependent on atom type	inter	
25	p_ovun2	12	Valence angle parameter	intra	y
26	p_val3	13b =>13a	Valence angle parameter	intra	y
27	beta, beta_i	26a	eReaxFF, constant, dependent on atom type	inter	
28	Val_i^'boc	3b	Number of lone pairs	intra	
29	p_val5	13b	Valence angle parameter	intra	y
30	p_c1	23c	Inner wall vdW repulsion parameter	inter	
31	p_c2	23C	Inner wall vdW repulsion parameter	inter	
32	p_c3	23c	Inner wall vdW repulsion parameter	inter	y
33	C_i	23d	Lg dispersion parameter	inter	y
34	R_eij	23d	VdW Radius for Lg dispersion correction	inter	

Index	Nom	Équation	Comment	Intra/inter	OPT
bonds					
1	D_e^sigma	6, 11a	Sigma-bond dissociation energy	intra	y
2	D_e^pi	6	Pi-bond dissociation energy	intra	
3	D_e^pipi	6	Double pi-bond dissociation energy	intra	
4	p_be1	6	Bond energy parameter	intra	y
5	p_bo5	2	Double pi bond parameter	intra	y
6	Val'_i^boc	3b	1,3-Bond order correction	intra	
7	p_bo6	2	Double pi bond order	intra	y
8	p_ovun1	11a	Overcoordination penalty	intra	y
9	p_be2	6	Bond energy parameter	intra	y
10	p_bo3	2	Pi bond order parameter	intra	y
11	p_bo4	2	Pi bond order parameter	intra	y
13	p_bo1	2	Sigma bond order	intra	y
14	p_bo2	2	Sigma bond order	intra	y
15	delta'_i	3a	Uncorrected BO overcoordination	intra	
16	p_ij^xel1	27	eReaxFF param : for adjusting number of electrons vailable to host atom	inter	

Index	Nom	Équation	Comment	Intra/inter	OPT
Off diag					
1	D_ij	23a	VdW energy	inter	y
2	r_vdW	23a	VdW radius	inter	y
3	alpha_ij	23A	VdW parameter	inter	y
4	r_0^sigma	2	Sigma bond length	intra	y
5	r_0^pi	2	Pi bond length	intra	
6	r_0^pipi	2	PiPi bond length	intra	
7	C_i, C_lg,ij	23d	Lg dispersion parameter	inter	

Index	Nom	Équation	Comment	Intra/inter	OPT
Angles					
1	Theta_0,0	13g	180o-(equilibrium angle)	intra	y
2	p_val1	13a	Valence angle parameter	intra	y
3	p_val2	13a	Valence angle parameter	intra	y
4	p_coa1	15	Valence conjugation	intra	
5	p_val7	13c	Undercoordination	intra	y
6	p_pen1	14b ->14a	Penalty energy	intra	
7	p_val4	13b	Valence angle parameter	intra	y

Index	Nom	Équation	Comment	Intra/inter	OPT
Torsions					
1	V_1	16a	V1-torsion barrier	intra	y
2	V_2	16a	V2-torsion barrier	intra	y
3	V_3	16a	V3-torsion barrier	intra	y
4	p_tor1	16a	Torsion angle parameter	intra	y
5	p_cot1	17a	Conjugation energy	intra	y

Index	Nom	Équation	Comment	Intra/inter	OPT
H bond					
1	r_hb^0	18	Hydrogen bond equilibrium distance	inter	
2	p_hb1	18	Hydrogen bond energy	inter	
3	-p_hb2	18	Hydrogen bond/bond order	inter	
4	-p_hb3	18	Hydrogen bond parameter	inter	

Table 3.9: Table representing all the different parameters included in a reactive force field. The column 'Opt' indicates if the parameters were optimized, and the column intra/inter indicates if the parameter is involved in intra-molecular interaction on inter-molecular ('inter' is indicated in yellow). A brief comment is given to describe the parameters.

3.3.2.4 Choice of step-size

The limits of the parameters will define the total parameter space used during the CMA-ES optimization. However, as it is presented below, the CMA-ES is an iterative process, and to define the parameter space used at each step, the algorithm uses a step-size σ. In this work, the step-size was define for each parameter with the equation 3.6.

$$\sigma = x \times 10^{-4} \times 5 \tag{3.6}$$

With x the parameter value.

3.3.3 CMA-ES optimization of Reactive force field

Now that the starting point is defined, we have the four files: trainset.in with the reference values defined by x and the geo files with the coordinates of the different molecules, the force field file with the initial parameters values, and the params file with the limit chosen and the step-size σ. In this section, it is presented how the CMA-ES optimizes the parameters. All the different optimization were done with the train ADF module of the Amsterdam Modeling Suite (AMS) software [90], which implements the CMA-ES algorithm.

The optimization starts with the initial force field provided at the beginning of the optimization and the parameters chosen to be optimized. Then, CMA-ES samples some force fields respecting the initial step-size value of each parameter given in the 'params' file by drawing the parameters from a multivariate normal distribution shown in the equation 3.7.

$$x_i \sim \mathcal{N}(CMA_{value}, \sigma^2 I) \tag{3.7}$$

With x_i the force field sampled, CMA_{value} the initial parameters values of the force field, I the identity matrix, and σ the step-size.

The number of sampled force fields λ should be higher than :

$$\lambda = 4 + 3 \times ln(N) \tag{3.8}$$

With the number of parameters optimized N (in this work 68). In our case, we used 20 sampled force fields. These new force fields are then ranked as a function of the fitness values obtained from the equation 3.1, and from this ranking a mean distribution is calculated through the equation 3.9 :

$$m_{(k+1)} = \sum_{i=1}^{\mu} w_i * x_{i:\lambda} \tag{3.9}$$

With w_i the positive weights used to rank the force fields, μ the number of sampled force field kept, and x_i the force fields sampled.

In figure 3.9 it is plotted the first step of the CMA-ES optimization is to obtain the next generation of force field with an example of two parameters optimized.

For the understanding of figure 3.9, the parameter values from the force field to the optimizer are transformed through these formulas 3.10 3.11.

$$x = \left[\frac{\frac{value}{abs(CMA_{value}) + 1} + 1}{2} \right] \tag{3.10}$$

$$ReaxValue = (max_{value} - min_{value}) * x + minValue \tag{3.11}$$

With CMA$_{value}$ being the value of the optimiser, max and min value corresponding to the limit chosen in the params file for the optimised parameters and REAX$_{value}$ being the value of the force field.

So, for example, the value of the example presented in figure 3.9, are respectively for the best force field obtain CMA1$_{value}$ = 15.9403 and CMA2$_{value}$ = 5.02226, which correspond respectively to REAX1$_{value}$ = -0.0525 and REAX2$_{value}$= 1.9075.

Then, after calculating the mean distribution, the step-size and the covariance matrix are updated. As it was discussed previously in the choice of CMA-ES, the covariance matrix described the correlation between the parameters. After that, a new iteration begins with the mean distribution and its covariance matrix. New samples force fields are produced according to a multivariate normal distribution denoted as :

$$x_i \sim \mathcal{N}(m_k, \sigma^2 C_k) \tag{3.12}$$

With C_k being the covariance matrix associated with the mean distribution m_k.

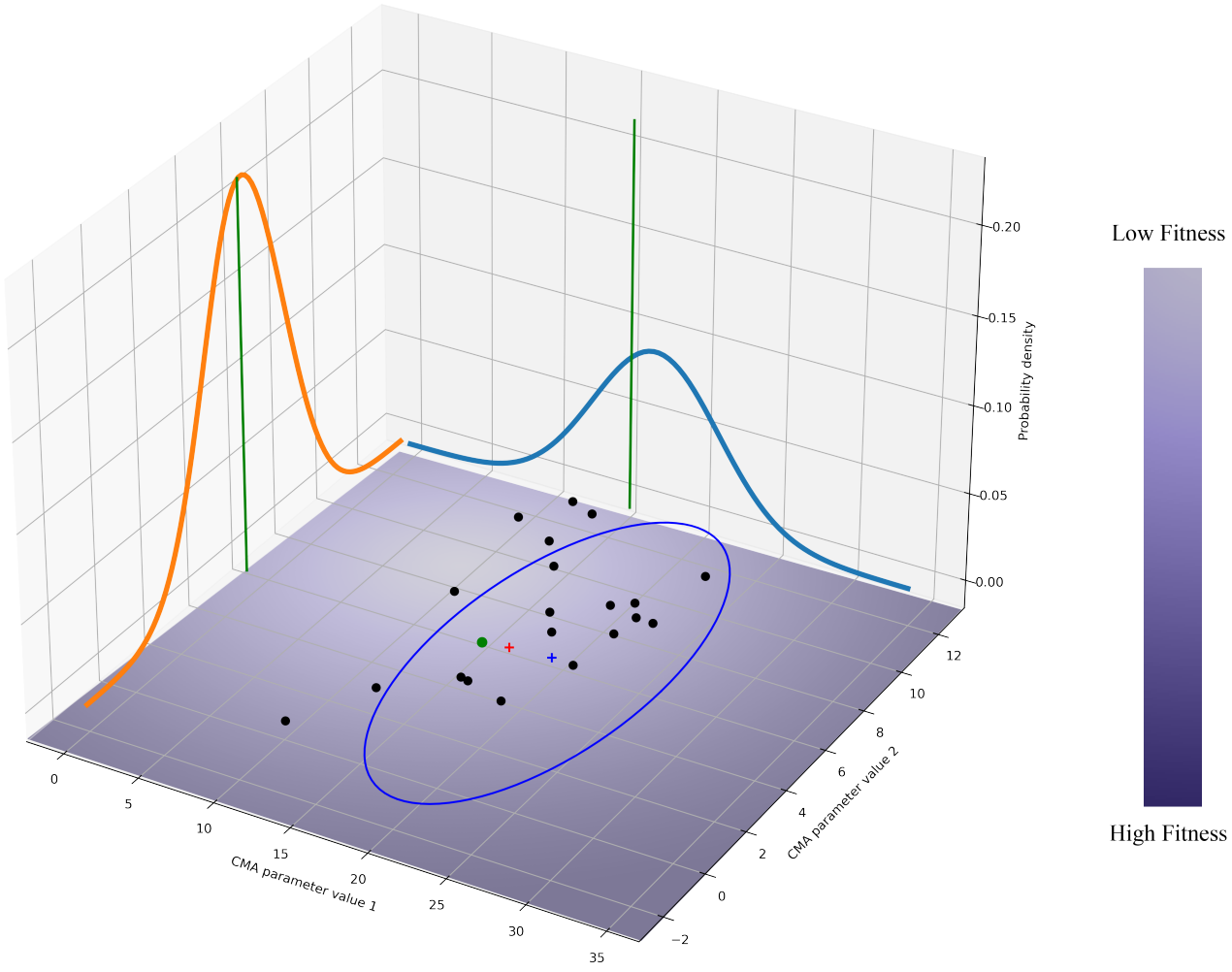


Figure 3.9: Here is the 3D representation of the first iteration of the CMA-ES optimizer to obtain a new generation. The surface at the bottom of the 3D representation is a heat map schematizing the fitness evolution in the parameters space with the legend on the rigth side of the graph. The blue cross represents the first initial values of the two parameters in the equation 3.7, and the blue dotted circle is the covariance associated with these parameters. The red cross represents the mean distribution obtained from the equation 3.9, which will be the next initial step for the next iteration. Black points represent the force field generated at this step, with the green point representing the best force field with the lowest fitness. The yellow and blue curves correspond to the two distributions of parameters 1 and 2. The two green axes of the distributions correspond to the two values of these parameters, giving the best fitness among all the different force fields generated. The green point represents this best force field in the parameters space.

After that, the optimizer operates the same first steps by recalculating the mean distribution, evaluating the step-size, and then the covariance matrix until it reaches the global optimum or the number of max iterations.

An example of these steps are presented in the next figures from 3.10 to 3.18.

The mean distribution and the covariance matrix are indicated by a red cross and a red dotted circle. The fitness evolution is based on the lightness, the lowest fitness is the most lightning spot, and here, it is localized

on the top right of the figures at the centre of the orange circles and correspond to the global optimum. (one axis correspond to one parameter optimized, in the subsequent figures, it is an example where only two parameters are optimized simultaneously with the limit associated).

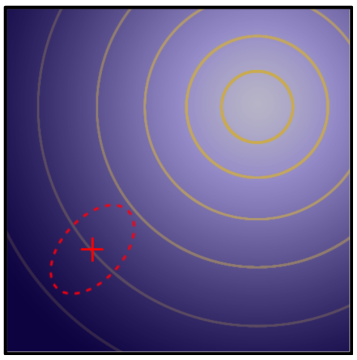


Figure 3.10: Here is the next step after the initial one presented in the figure 3.9, with the mean distribution of the parameters depicted by the red cross and the covariance matrix by the red dotted circle.

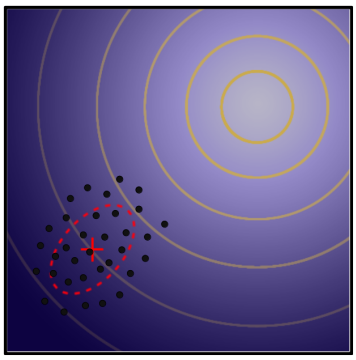


Figure 3.11: Here is the generation step of different force fields represented by black points.

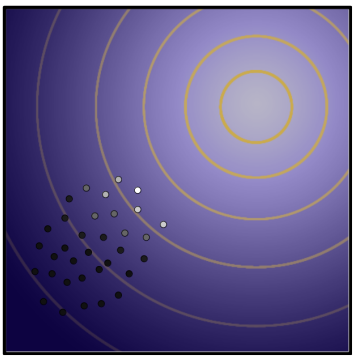


Figure 3.12: In this step, the different force fields are ranked as a function of the fitness obtained from the equation 3.1. The whiter the point is, the lowest the fitness is.

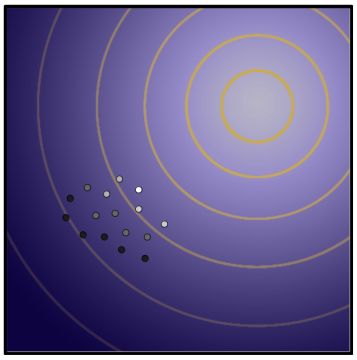


Figure 3.13: In this step, the lowest fitness force fields are kept, and the others discarded.

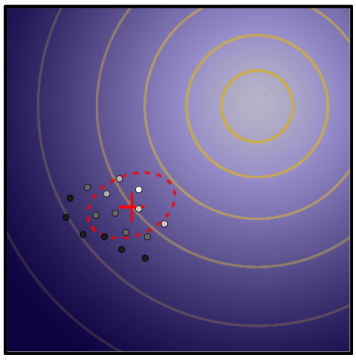


Figure 3.14: In this step, the mean distribution of the parameters is updated as the covariance matrix associated.

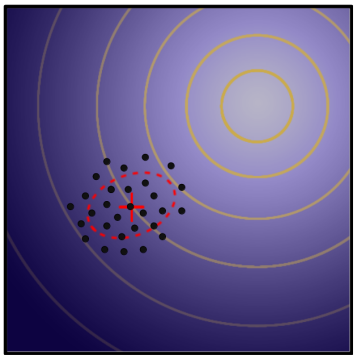


Figure 3.15: It is shown a new step of generation of force fields with the new mean distribution and covariance matrix.

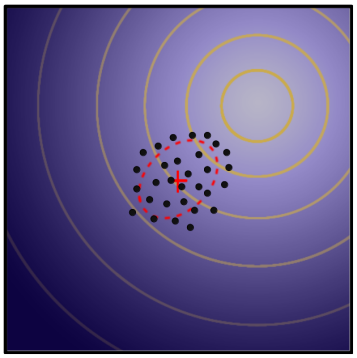


Figure 3.16: Following the same procedure as shown in the figure 3.12, 3.13 and 3.14 with the starting points of the figure 3.15, generation of new force fields with new mean distribution and new covariance matrix.

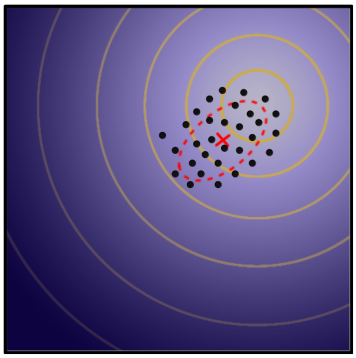


Figure 3.17: Here is a generation step of new force fields from the new mean distribution and covariance matrix obtained from the previous generation.

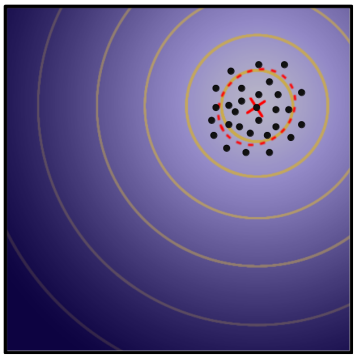


Figure 3.18: This step is the last generation step of force fields close to the global optimum.

To summarize the procedure of the CMA-ES optimization, it starts with an initial state defined by initial parameters, their limits and its covariance matrix shown in the figure 3.10. Then, force fields are generated through a multivariate normal distribution. The fitness is calculated for each force field, allowing us to rank them. The force fields with the highest fitness are discarded, and the others are weighted and then used to determine the new mean distribution and its covariance matrix. This procedure is done as shown in the figure 3.15 to 3.18 until the global optimum is reached. As we now understand how the optimizer operates, it is presented in the new section, its utilisation to develop the new ReaxFF force field.

3.3.4 Optimization procedure

Now that the reader understands the choice of CMA-ES inputs and the CMA-ES algorithms, it is presented in this section the optimization procedure developed to obtain a final optimized reactive force field, with the different investigations created.

The optimization procedure, is divided in multiple steps, in which, one hundred CMA-ES optimizations are done for statistical data, which leads to three different investigations. All the different investigations are presented in the following, and a final workflow summarizing all of them is presented in the figure 3.26. It might help the reader to follow step by step the optimization procedure.

3.3.4.1 Choice of the number of CMA-ES optimizations step

As we are doing a lot of CMA-ES optimizations, we needed to define the number of optimization steps needed to reach convergence. We obtained the figure 3.19 doing ten CMA-ES optimizations using the training

set and the first set of parameters and their limits.

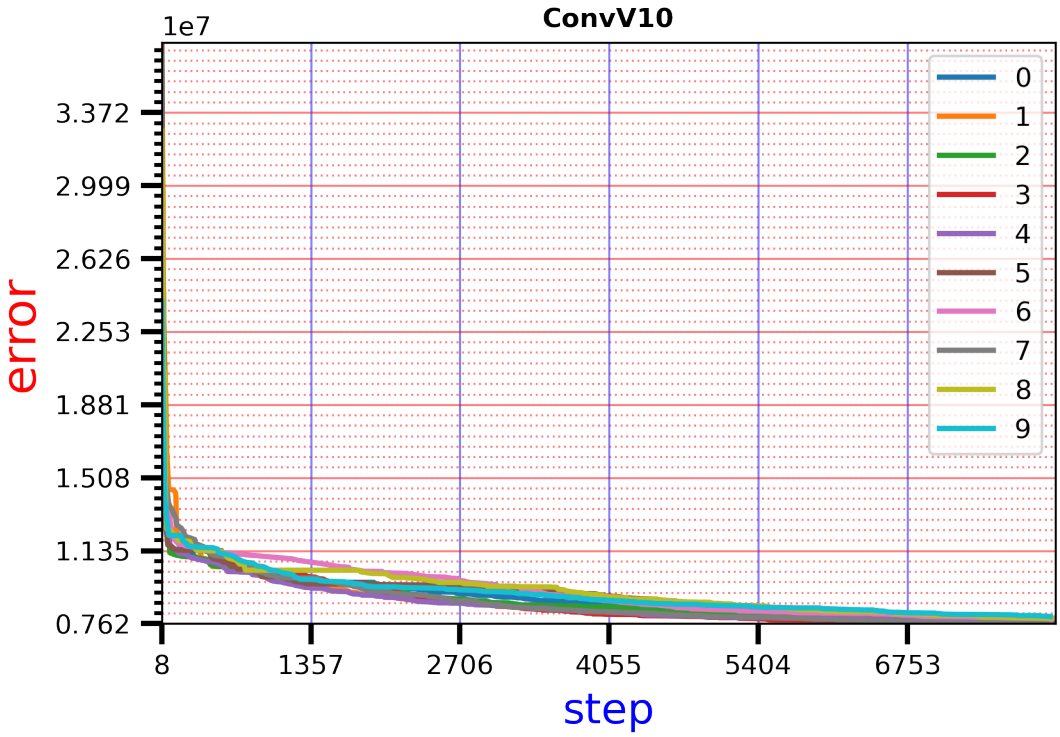


Figure 3.19: 10 evaluations of the convergence as a function of the step of the optimization. Each graph contains 10 runs of the optimization which are all denoted by a number.

We can see that the optimization gets, for each run, a quick convergence in the 1000th first steps. Then, it continues to converge slowly. That is why to have a reasonable time consumption and accuracy ratio, we did optimizations with 5000 steps.

3.3.4.2 Parameters distribution and correlation

In order to obtain our final reactive force field, we did numerous CMA-ES optimizations. We divided our optimization into two central parts; the first corresponds to optimizing all the parameters involving the chlorine atoms with hydrogen and carbon. The second part is the optimization of all the parameters linking chlorine atoms with oxygen atoms. Moreover, these two parts were divided into different optimization steps to reduce the fitness and obtain a well-suited force field. All these steps to obtain the optimal set of parameters include one hundred CMA-ES optimizations. The results obtained from this hundred of optimization are statistically investigated to define a new starting point for the parameters, along with their boundaries for the next optimization step. Thereafter, we describe this statistical investigation.

3.3.4.3 Distribution investigation

At the end of the one hundred CMA-ES runs, we have a set of one hundred force fields along with the associated parameters. From these data, it is possible to plot the distribution of each parameter. This information allows us to determine if the parameter is well defined or if the limits should be changed. The ideal case would be to obtain a normal centred distribution on the force field with the lowest fitness. In the figure 3.20 and 3.21 are presented two examples of these distributions.

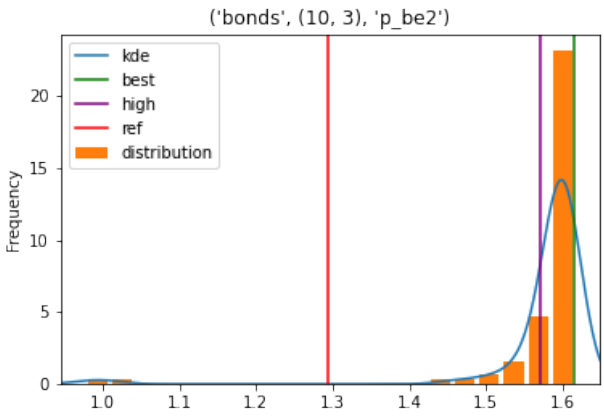


Figure 3.20: Distribution of the parameter of the block bonds involving the atoms 10 (Chlorine) and 3 (Oxygen) named "pbo2". In red, the initial value, in green the best force field and in purple, the force field value with the highest fitness.

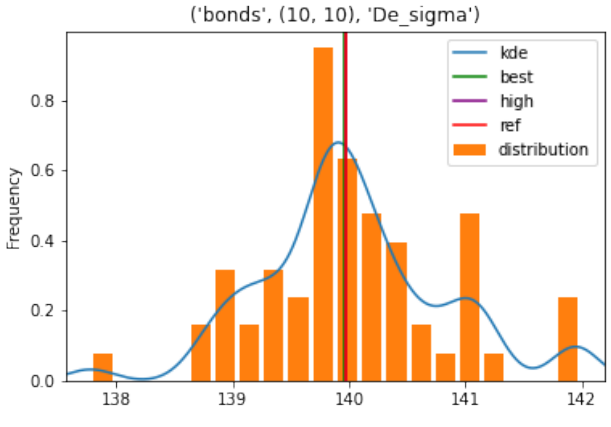


Figure 3.21: Distribution of the parameter of the block bonds involving the atoms 10 (Chlorine) and 10 (Chlorine) named "De sigma" (Sigma-bond dissociation energy). In red, the initial value, in green the best force field and in purple, the force field value with the highest fitness.

Figure 3.20 is an example of an inadequate distribution and the figure 3.21 of a good one. Indeed, on figure 3.20, the distribution is highly stacked on the maximum limits of the parameter, which indicates a wrong choice of the limits which should be changed to a higher value. On the other hand, in the figure 3.21, we can see that the best force field with the lowest fitness and the initial force field have almost the same value, and lie around the maximum of the distribution. From this observations, limits can be updated, or the force field validated if the majority of the distributions are well shaped. Obviously, due to the high number of parameters, all the distributions won't have all the best force field values at the centre. Still, if the distribution shape is similar to a Gaussian function, it indicates that the limits are all-right and can be sharpened. In parallel of the parameter distributions, it is also interesting to investigate the correlation between the parameters and try to avoid non-suitable correlation. The correlation study will be the next topic of the following section.

3.3.4.4 Correlation investigation

One of the main interests of using CMA-ES optimizer is its ability to consider the correlation between the parameters through the covariance matrix all along the optimisation procedure. The functional form of the

ReaxFF force field implies correlations between some parameters. For example, the energetic bond parameters should correlate with each other. So, using CMA-ES wisely allows taking into account these correlations. However, due to the fact that many parameters are optimized at the same time the formation of a non-suitable correlation might appear during these optimizations. An investigation of these correlations might help to validate the force field or upgrade the initial parameters values and their boundaries for the following optimizations. Figure 3.22 and 3.23 are two heat maps indicating the correlation between the parameters obtained after two distinct optimization steps from one hundred CMA-ES optimizations.

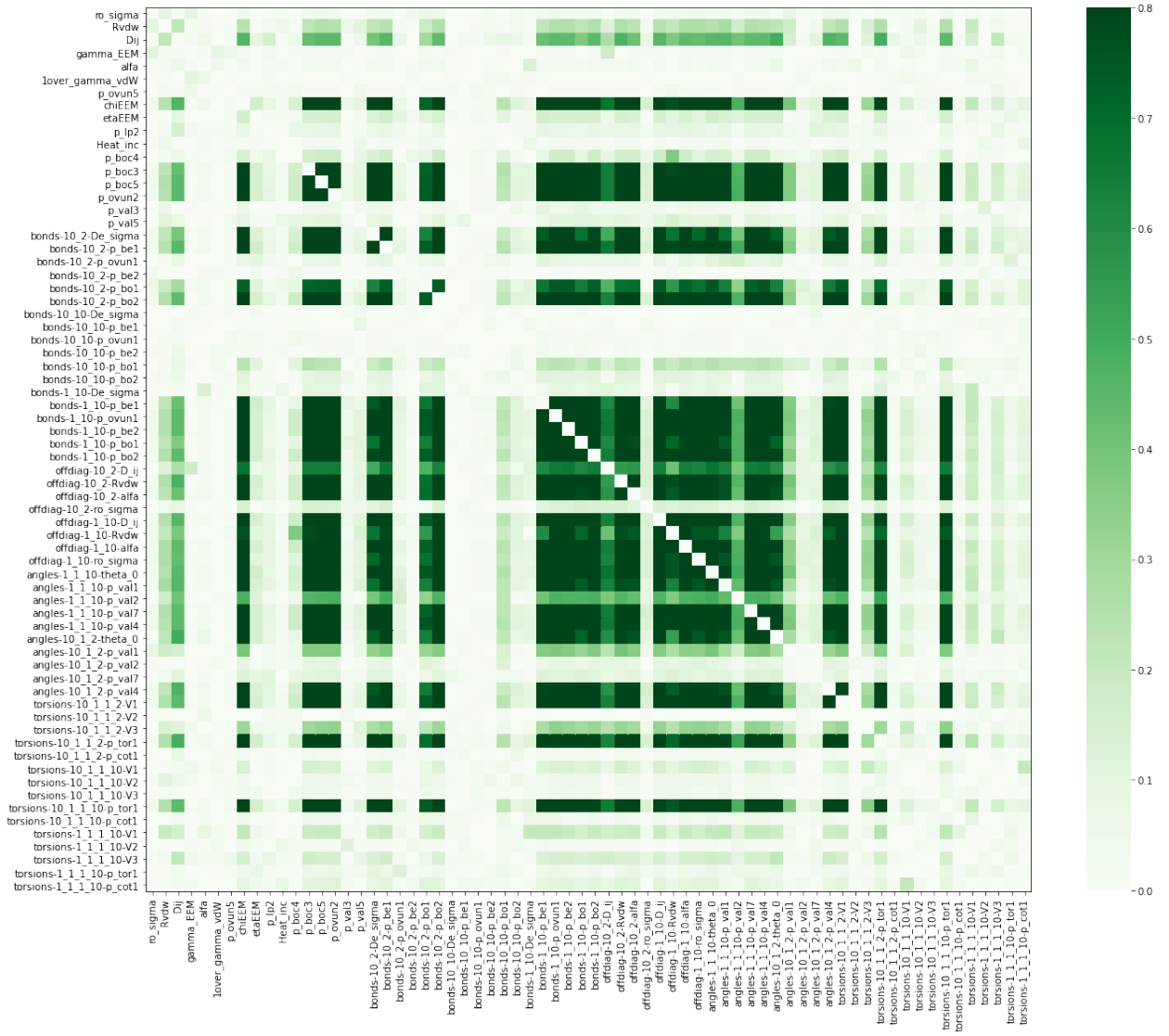


Figure 3.22: This heat map represents the correlation between all the parameters optimized from one hundred CMA-ES optimizations. The scale on the right indicates the correlation rate between the parameters, with 0.8 being the highest correlation.

First of all, by comparing these two heat maps, we can see that there are many more highly correlated parameters in figure 3.22 than in figure 3.23. The difference between these two heat maps is that the first one was one of the first optimization steps, whereas the second was one of the last. The first has non-suitable correlations, which means parameters are correlated without meaningful sense. In contrast, the second one

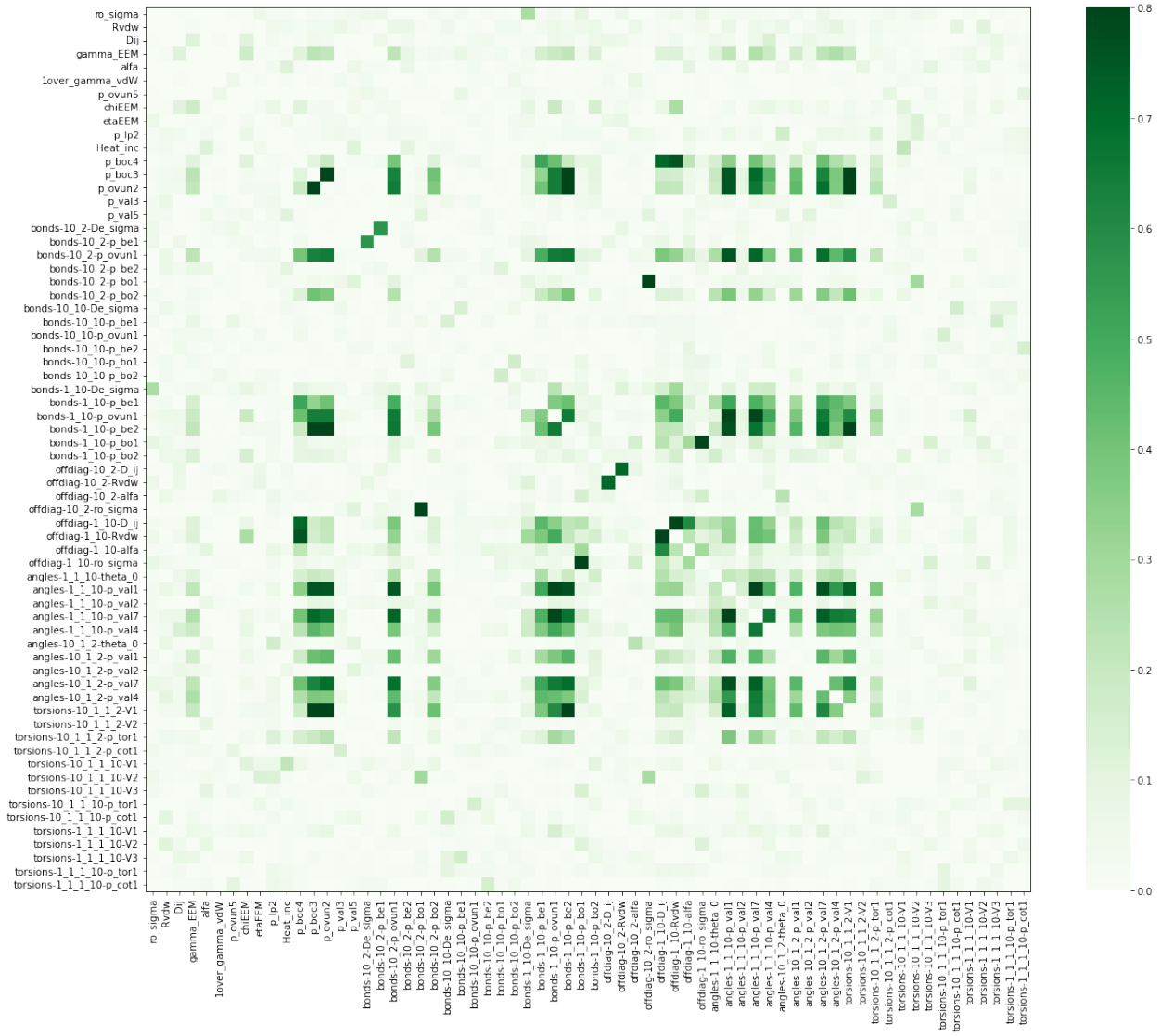


Figure 3.23: This heat map represents the correlation between all the parameters optimized from one hundred CMA-ES optimizations. The scale on the right indicates the correlation rate between the parameters, with 0.8 being the highest correlation.

has almost only meaningful correlations. An example of non-suitable correlations is the high correlations of the parameter "chiEEM" with many parameters implied in the intramolecular interactions. Indeed, being used in the Coulomb interactions equation, it should be only correlated to other parameters of the same equation, or at least in the same kind of interactions. On the contrary, by looking at figure 3.23, we can see that almost all of the high correlation values are meaningful, linking parameters of the same equations or playing a role in the same kind of interactions.

So, verifying that non-meaningful correlations are suppressed along the different optimization steps is crucial to better define the parameters and to obtain a well and efficient reactive force field at the end.

3.3.4.5 Identification of force fields families

The last investigation made to improve the parameters along the parameters optimizations procedure is the identification of force field families. Indeed, due to the existence of non-suitable correlations, the lowest fitness force field obtained from the hundred CMA-ES optimizations is sometimes not the best choice. Thus a good way to determine the next initial set of parameters is to find a family of force fields combining the lowest fitness and an acceptable correlation available.

Figure 3.24 presents the distribution of the parameters presented in the previous section and the distribution of the fitness over all the considered force fields.

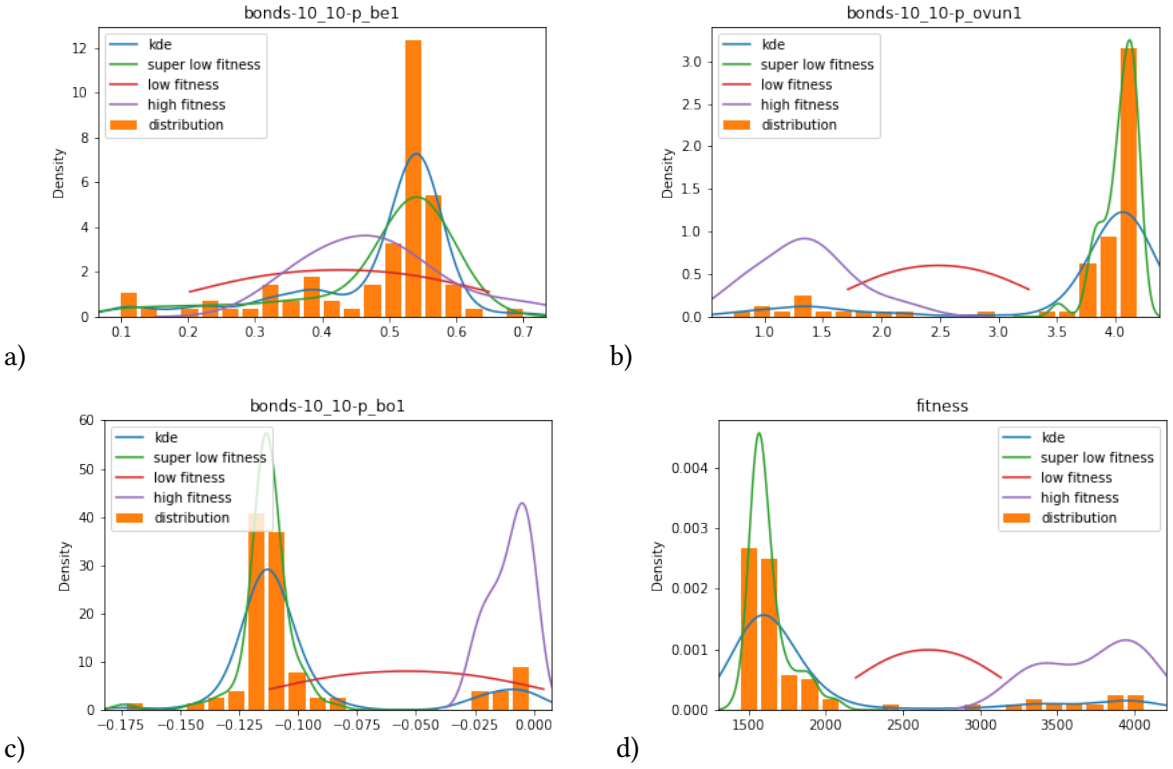


Figure 3.24: The distribution of three different parameter is presented in Figures a) b) and c). A blue curve indicates the KDE (Kernel density estimation) curve on each graph. The distributions of the parameters taking into account only force fields with high fitness (purple curve), low fitness (red curve) and super low fitness (green curve) are presented. The figure d), represents the distribution of the force field fitness, with super low fitness between 1300 and 2100, low fitness between 2100 and 3050 and high fitness more than 3050.

In the this figures, we divide the parameters depending on the fitness values of the force field and we define families of force field depending on the fitness values. Another way to represent this association is presented in the figure 3.25.

We presented an ideal case where the different force field parameters families are easily identifiable in the example presented in the two figures 3.24 and 3.25. However, at the beginning of the optimization procedure, it is much harder to determine these families. The family identification was principally used during the last optimizations, and it helps us to determine the final force fields to be tested against the validation set to

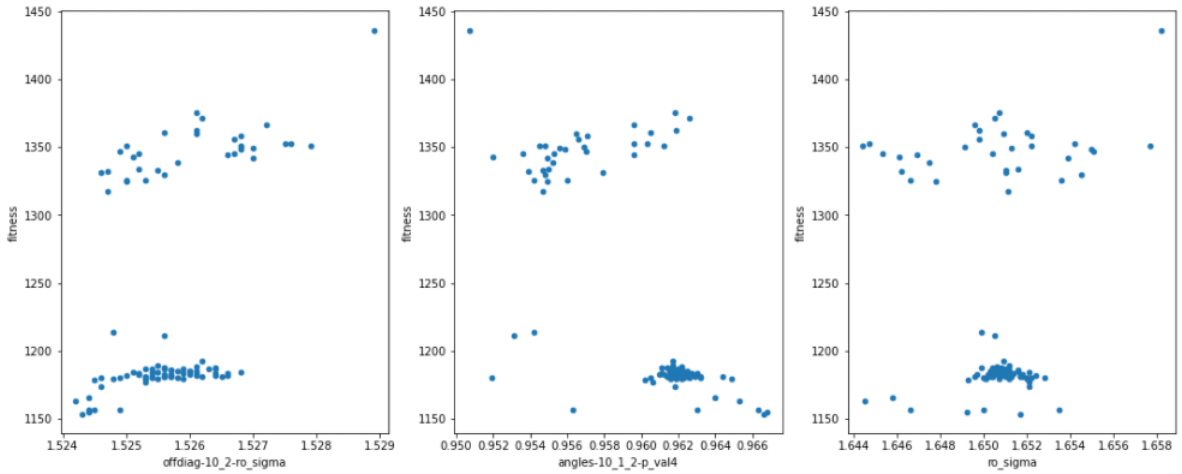


Figure 3.25: Here it is represented the values of three different parameters as a function of the fitness values.

identify the best of them.

3.3.4.6 Validation of the force field

After all the optimization procedures, we finally end up with a couple of force fields with a low fitness and well-correlated parameters. The final evaluation of these force fields would be made by comparing the fitness values computed with the validation set. The validation of our force field is described in the final section of this chapter associated to its publication. It is also important to notice that more than 6000 CMA-ES optimizations were done to obtain the final force field to optimize the 68 parameters.

3.4 Summary

To give a global overview of the procedure that was followed to optimize the new force field, we propose to comment on the figure 3.26. First, we built the training and the validation sets with energetical and structural data, like PES and bond distances. We determined the initial conditions as the optimizer used with the number of optimization steps and the initial force field. We chose the optimized parameters and their boundaries. From this point, we obtain the start of the optimization cycle. A representation of the fitness space with the different force fields is indicated, with an example initial force field in red, the boundaries associated in purple and the global minimum that we want to reach in green. Then, the training set is used in a series of CMA-ES optimizations leading to 100 sets of parameters. From there, we explore the parameters space by doing three different investigations. The first one is the parameter distributions, which allow us to update the boundaries correctly and identify if the parameters are well-defined. Then, we have the correlation investigation, which is an indicator of the good optimization of the force field and helps to remove non-suitable correlations again by updating the limits of the parameters. Finally, we can identify force fields' families using the CMA-ES

optimization's fitness values. From these investigations, we continue the optimization cycle if the fitness is still progressing and bad correlations still exist. The procedure stops when the fitness is stagnant and the parameters are well-defined. From this last optimization, one force field of each of the different families extracted is kept (with the lowest fitness). They are evaluated with the validation set by a simple fitness computation. The force field with the best fitness is then kept and validated, and several performance tests are made, which are presented in the next section corresponding to the publication of the force field.

3.5 New ReaxFF force field performance

In the continuity of the force field development, it is now presented in this section, the article published of the new CHONCl-2022_weak force field. This article presented the training and validation set construction as it was done at the beginning of this chapter. However, afterwards, different performance tests are performed on chlorinated molecules, with, for example, the capability of the force field to represent the central torsions of the PCBs or their charge distribution. Moreover, it finishes with a new PES investigation of the oxidation of the chlorobenzene molecule in the presence of dioxygen.

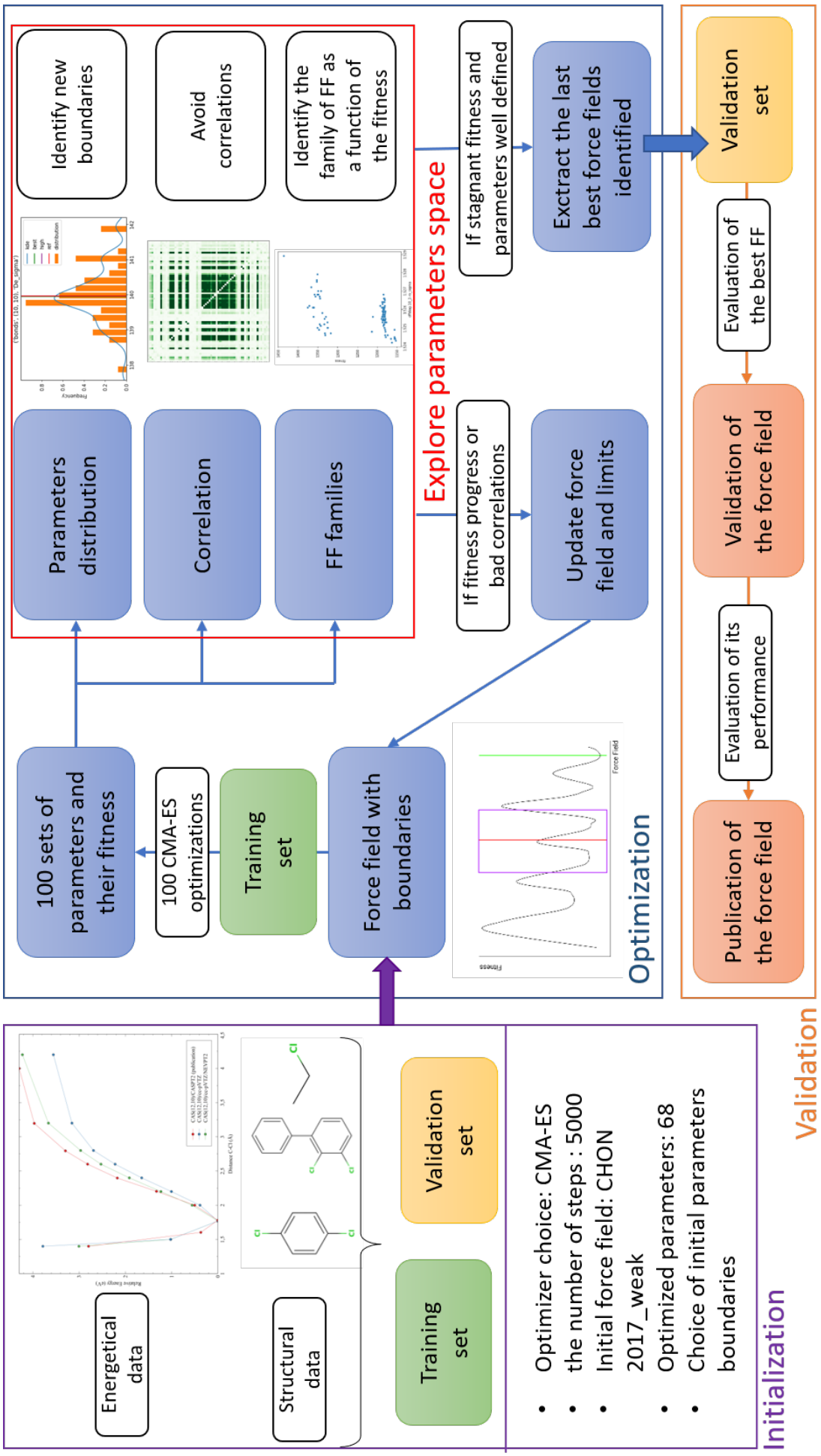


Figure 3.26: Summary of the full method developed and used to obtain our new reactive force field.

3.6 Article 1 : Development of a novel ReaxFF reactive potential for organochloride molecules

Development of a novel ReaxFF reactive potential for organochloride molecules

Matthieu Wolf, Didier Bégué, Germain Salvato Vallverdu

germain.vallverdu@univ-pau.fr

Universite de Pau et des Pays de l'Adour, E2S UPPA, CNRS, IPREM, UMR 5254, Pau France

https://doi.org/10.1063/5.0120831

Abstract

This article presents a new reactive potential in the ReaxFF formalism. It aims to include the chlorine element, and opens up the fields of use of ReaxFF to the whole class of organochloride compounds including conjugated or aromatic groups. Numerous compounds in this family raise global awareness due to their environmental impact, and such a reactive potential will help investigate their degradation pathways. The new force field, named CHONCl-2022_weak, belongs to the aqueous branch. The force field parameters were fitted against high-level quantum chemistry calculations, including CASSCF/NEVPT2 calculations and density functional theory calculations, and its accuracy was evaluated using a validation set. The root means square deviation against quantum mechanics energies is $0.38\,\mathrm{eV}$ ($8.91\,\mathrm{kcal\,mol^{-1}}$). From a structural point of view, the root means square deviation is about 0.06Å for the bond lengths, $11.86°$ for the angles and $4.12°$ for the dihedral angles. With CHONCl-2022_weak new force field, we successfully investigated the regioselectivity for nucleophilic or electrophilic attacks on polychlorinated biphenyls (PCB), which are toxic and permanent pollutants. The rotation barriers along the bond linking the two benzene rings, which is crucial in the toxicity of these compounds, are well reproduced by CHONCl-2022_weak. Then our new reactive potential is used to investigate the chlorobenzene reactivity in the presence of hydroxyl radicals in atmospheric condition or in aqueous solution. The reaction pathways computed with ReaxFF agree with the quantum mechanics results. We showed that, in the presence of dioxygen molecules, in atmospheric condition, the oxidation of chlorobenzene likely leads to the formation of highly oxygenated compounds after the abstraction of hydrogen radicals. In water the addition of an hydroxyl radical leads to the formation of chlorophenol or phenol molecule, as already predicted from plasma-induced degradation experiments.

3.6.1 Introduction

Computational chemistry provides powerful methodologies in order to investigate chemical systems, including classical molecular mechanics approaches, the fastest and quantum mechanics (QM) methodologies, the most accurate. Considering the investigation of the chemical reactivity of complex systems, a methodology accurate enough to describe bonds breaking and formation is needed at an atomic resolution. Bridging the gap between molecular mechanics and a more sophisticated approach such as quantum chemistry, ReaxFF is a reactive potential based on the so-called bond order. It allows chemical reactions to be considered in molecular dynamics simulations of large systems. ReaxFF, developed since 2001 by van Duin et al.[5], has proven its efficiency in describing reactive systems of various domains and different phases, including a wide variety of elements[58]. It was initially designed to study organic compounds with carbon and hydrogen atoms to focus on combustion reactions[5]. However, it rapidly increased its field of application since the CHO reactive force field publication by Chenoweth et al.[211] in 2008. Many variations appear from this new force field, and a wide variety of force fields were published divided into two main branches called the aqueous branch and the combustion branch[58].

One of these force fields called Protein-2013, coming from the aqueous branch, was a step forward when it was published[71]. Indeed it allows for the first time to investigate the reactivity of biomolecules through molecular dynamics simulations, including water as a solvent. All the elements needed to consider amino acids were parameterized, including, in particular, nitrogen, sulfur and counter ions. This force field has been updated twice in 2017[69] and 2018[68] with a better description of the weak forces and has become the well-none force field CHON-2017_weak force field used in numerous works[212,213,214].

This new version of ReaxFF can be implemented to investigate the degradation pathways of organic molecules. One point of particular interest in the nowadays environmental context would be the ability to consider the chlorine element. Indeed, organochlorides are compounds widely used in the last half-century in an industrial or pharmaceutical context[215,216,217]. Because of their chemical stability, these organochlorides are measurable in different environments[218], such as soil[219], water[220], the atmosphere[221], living-organism[216] and food[222]. Among the organochlorides, organochloride pesticides and PCBs (polychloride biphenyls) raise global awareness due to their environmental impact and their hight toxicity[223,224]. The DDT (dichlorodiphenyltrichloroethane) and TCDD (2,3,7,8-Tétrachlorodibenzo-p-dioxin) are both organochloride pesticides with high toxicity and persistence[225,226,227]. Even if they were identified as toxic, persistent compounds and prohibited since 2001 by the United Nations Environment Program[224], they are still used in some countries[215,218]. PCBs were highly produced with around 1,325,810 tons[6] since 1930 in different manufacturing products. Their worldwide presence was proven in all environments and capable of biomagnifying along the trophic chain[158,143]. It was identified that 12 of the 209 PCBs were particularly toxic with an important TEF (toxic-

ity equivalent factor)[133] due to a low torsional angle between the two cycles allowing them to arrange their structures easily[176].

The investigation of these compounds' reactivity in different environments is of great interest to understand their degradation, which will allow the development of new treatment processes and a better understanding of the toxic effect on the environment and health. In the last decade, the reactivity of organochlorides and in particular the different enantiomers of the DDT, TCDD and PCBs, have been investigated through computational approach[228,229,230,231,232,233,234]. The degradation processes were investigated from molecular dynamics simulation on the DDT in a water-sediment environment, and several studies have implemented atmospheric reactions initiated by radicals such as OH, CH or molecules such as O_3 and H_2O_2.

ReaxFF parameters for the chlorine element were first provided in the Protein-2013 force field where chlorine was included as a counter ion[71]. It was also parameterized in two different force fields to deal with oxidative degradation of organic matter in geological formations[235] and explosive compounds[60]. In both cases, the parameter set is not transferable enough to consider organochloride compounds, particularly for conjugated or aromatic molecules.

In this work, we describe the parametrization of a new reactive force field, named CHONCl-2022_weak, based on the CHON-2017_weak force field. It aims to include the chlorine element, and opens up the fields of use of ReaxFF to the whole class of organochloride compounds including conjugated or aromatic groups. We decided to start from the CHON-2017_weak as it was the last version of a CHON reactive force field in the aqueous branch. After a quick review of the ReaxFF methodology, we describe the training set and the validation set implemented in this work and the computational details linked to reference calculations. The parameters are obtained against the QM calculations. The comparison with the validation set is detailed, considering bonds, angles and dihedrals internal degrees of freedom along with charge distribution and reaction pathways. The validation set includes PCB molecules. Then the reaction pathways of the chlorobenzene oxidation are investigated from QM calculations and ReaxFF simulations, including the regioselectivity of the reaction with the hydroxyl radical.

3.6.2 Computational methods

3.6.2.1 A novel ReaxFF reactive potential

ReaxFF is a force field based on the bond order (BO), which makes that quantity the key point that tunes the various energy terms. The main consequence is that ReaxFF is able to describe bond dissociation and enable the investigation of large reactive systems[211]. The following equation represents the complete equation with

all energy terms which contribute to the total energy:

$$E_{system} = E_{bond} + E_{lp} + E_{over} + E_{under} + E_{val} + E_{pen} + E_{coa} + E_{C2} + E_{triple} + E_{tors} + E_{conj} + E_{Hbond} + E_{vdW} + E_{Coulomb} \quad (3.13)$$

All these energetic terms are divided into two main groups: intramolecular interactions (11 terms) and intermolecular interactions (3 terms). The common factor in intramolecular interactions are their dependence upon the BO. Following the terms in equation (3.13), the intramolecular terms describe the bond, lone pair, overcoordination, undercoordination, valence angle, correction of double bonds angle valency issues, correction for NO_2 group stability, correction to destabilize triple bond for carbon, a particular correction for carbon monoxide, torsions and conjugation energies. On the other hand, the intermolecular interactions include the Van der Waals, Coulombic and hydrogen-bond interactions. More details are available in the literature, and we invite the reader to look at the article of Chenoweth et al.[211] for further information.

The CHON-2017_weak version was used as the initial force field by including the chlorine atoms and the associated bond, angle and torsion parameters. A total of 68 parameters have been parameterized including parameters associated to valence and non-bonded terms of the equation (3.13). Even for a single element, this represents a huge number of parameters to optimize, leading to a complex optimization procedure, particularly due to correlations between parameters. In the following, we give some hints by describing the steps we followed during this optimization.

As done in previous works[211], the required parameters are optimized against quantum mechanics (QM) reference calculations. The computational details associated with the different methods used to implement these calculations are presented in the following subsections.

Primarily, a training set and a validation set were made. They include a set of molecules and small reaction pathways representing all the internal coordinates as bonds, angles, and torsions necessary to mimic the reactivity of organochloride. For example, the potential energy surface (PES) of the bond dissociation of the C-Cl bond in the chlorobenzene, as well as the geometrical parameter of its optimized structure, were added to the training set for the C-Cl bond parametrization. The training set and the validation set files, including energy values and geometries of the compounds along PESs are given in the supplementary information.

The stochastic optimizer CMA-ES (Covariance Matrix Adaptation Evolution Strategy)[210], implemented in the REAX module of Amsterdam Modeling Suite (AMS)[90], was used to optimize the parameters. This algorithm considers the correlation between the parameters thanks to the implementation of covariance behaviour. Like all the evolutionary algorithms, CMA-ES is partly based on biological evolution and tends to

obtain the optimal population iteratively by minimizing an error function, the fitness, defined as followed:

$$Fitness = \sum_{i=1}^{n} \left[\frac{x_{i,train} - x_{i,ReaxFF}}{\sigma_i} \right]^2 \tag{3.14}$$

where the differences between the reference values, $x_{i,ref}$ and the ReaxFF values, $x_{i,ReaxFF}$, are weighted by σ_i values. These σ_i weights are manually chosen depending on the accuracy of the reference or its utmost importance. For example, the weight of the accurate CASSCF calculations can be higher than that made with DFT. Moreover, it is necessary to consider the number of points in each PES in order to avoid overfitting this coordinate.

One crucial step before the CMA-ES optimization is to build an initial set of parameters and their individual boundaries. Thus, the parameter values were chosen by mimetism with the same parameters of a different element in the CHON-2017_weak force field or compared with the values of identical parameters in different force fields containing chlorine. For the first boundaries, broad limits were chosen. However, in order to manage the correlation between the parameters, the optimization was split with no more than 25 parameters optimized simultaneously, starting by optimizing C-Cl and Cl-H bond parameters with the angle and torsion parameters containing these bonds, then Cl-O and finally Cl-Cl.

It leads to a process of optimization of the parameters in three successive parts. Starting from the previous set of parameters, we proceed as follows in each part. Several cycles of hundreds of CMA-ES optimizations were produced with broad boundaries, leading to sets of one hundred optimized parameters. Then, from this set, we analyzed the distribution of each parameter in order to narrow the boundaries and avoid bimodal distributions. Moreover, we also investigated the distribution of all the different fitness obtained, which helps identify groups of force fields that are not well optimized. On the other hand, we analyzed the correlation between the parameters in the population of optimized force fields. We used it as an indicator of well-defined boundaries for the parameters by reducing the correlation between them as much as possible. Considering the parameter distributions, the fitness distribution and the correlation analyzes, the parameter boundaries were manually improved in each cycle of CMA-ES optimization runs until we stabilize a set of parameters.

Finally, at the end of the three parts of the last parameters optimization, the new CHONCl-2022_weak force field was chosen according to the one with the lowest fitness and by comparison with the validation set. The file, including the parameters, is provided in the supporting information in the format suitable for LAMMPS[89].

3.6.2.2 Computational details, quantum chemistry methods

This part presents the computational details associated with the quantum chemistry methods implemented to build the training set and the validation set. These sets consist of PES, relative energies and geo-

metrical data such as bond lengths, bending angles and dihedral angles.

One essential reference values associated with ReaxFF are the bond dissociation energies (BDE). In this work, we used complete active space self-consistent field (CASSCF) calculations[186] to compute the BDEs. Indeed, at the dissociation limit, the multireference character of the wavefunction needs to implement such high level methodologies to gather accurate results.[187]. Details about the CASSCF calculations are presented in the supporting information in the table S1 with the selected active spaces associated. Moreover, NEVPT2[188] calculations, including a perturbation theory correction at the second-order were considered. The calculations were done with ORCA software version 4.2.1[189] and the cc-pVTZ basis set[190].

Density functional theory (DFT) calculations were made apart from the BDE calculations to obtain the PES along bending angles and dihedral angles. The calculations were driven with Gaussian 16[236], using the B3LYP functional[196,197,198,199], the 6-31+g(d) basis set[200,201,202] and the GD3BJ empirical dispersion model[203].

The training and validation sets include reaction pathway and transition states geometries obtained using an in-house Chain-of-States methodology[204] coupled to Gaussian 16 program. In order to keep the consistency with the PES calculations, the B3LYP functional with the 6-31+g(d) basis set and GD3BJ empirical dispersion were used. After the Chain-of-States process, the transition state geometry was optimized. The frequencies were computed, and we verified the presence of one single imaginary frequency. Moreover, using the same methodology, we computed transition states energies using the M06-2X DFT functional and the basis set 6-311++g(2df,2p) for further comparison with available results in the literature.

Bond lengths were computed at the MP2[237,238,239,240,241] level with the cc-pVTZ basis set using the ORCA software version 4.2.1[189]. This calculation level is well known for its accuracy in reproducing this geometrical data.

3.6.2.3 Molecular mechanics and molecular dynamics

All the calculations at the ReaxFF level were done using either the Large-scale Atomic/Molecular Massively Parallel Simulator[89] code (LAMMPS) version October2021, or the REAX module of the AMS software[90]. For the comparison of BDE, all PES along internal coordinates and reaction pathways, the ReaxFF energies were computed on the geometries obtained at the QM level of calculations included in the training set or the validation set. Defaults LAMMPS cutoff parameters were used in all ReaxFF calculations, in particular nbrhood_cutoff equal to 5Å and hbond_cutoff equal to 7.5Å.

Considering molecular dynamics simulations in water or in atmospheric condition a timestep of 0.1 fs was used for all the simulations, in the NVT ensemble using the Nose-Hoover thermostat[99,100]. In the atmospheric investigation, constant temperature simulations were implemented, whereas, in the water degradation, we used a ramp of temperature between $550\,\mathrm{K}$ and $750\,\mathrm{K}$ over a simulation time of 2 ns.

3.6.3 Results and discussions

3.6.3.1 Force field validation against QM calculations

The new CHONCl-2022_weak force field was validated by comparing the ReaxFF energies with energies obtained from high-level quantum chemistry calculations at the CASSCF/NEVPT2 level or calculations at the B3LYP/6-31+g(d) level. The root mean square deviations (RMSD) of the energy differences are $8.36\,\mathrm{kcal\,mol^{-1}}$ ($0.36\,\mathrm{eV}$) and $8.91\,\mathrm{kcal\,mol^{-1}}$ ($0.39\,\mathrm{eV}$ in the training set and the validation set, respectively. For comparison, Hur et al.[235] obtained a RMSD of the energy of about $1.57\,\mathrm{eV}$ for their force field optimized to investigate the degradation of compounds with oxychlorine.

Considering references to molecular structures, the RMSD on bond lengths is about 0.06 Å and considering bending angles and dihedral torsions, the RMSD are $11.86°$ and $4.12°$ respectively. The RMSD obtained on angles is surprisingly much higher than the one on torsion. This is mainly due to the Cl-C-H angle in chloroethane and chloroethanol. Indeed, this angle is highly overestimated by the CHONCl-2022_weak force field and represents more than 30% of the total error between the QM values against the ReaxFF ones.

3.6.3.2 Bond dissociation energy and bond lengths

This subsection compares the PES and BDE energies computed at the CASSCF/NEVPT2 level to ReaxFF results. In the figure 3.27, there are the four PES along the different bonds parametrized in the new force field. The PES of the dichloride (Cl_2) and the acetyl chloride (CH_3OCl) molecules belong to the training set, while the chloromethanol (CHOHCl) and the hydrogen-chloride (HCl) molecules belong to the validation set. From a general point of view, the ReaxFF PES follow with acceptable accuracy the QM PES for the four molecules. Some differences appear at large distances. For chloroethane and chloroethanol, differences appear above 2.6Å, as the results obtained by Hur et al.[235]. This can be explained by the overestimation of the Cl-C-H angle. For the dichloride molecule, the ReaxFF PES seems to reach a plateau for distances larger than 3Å. This is due to the values of the BO that vanish at this distance, and thus only non-bonded interactions remain for such distance.

The BDE values are gathered in the table 3.10. CHONCl-2022_weak's BDE complies with the BDE computed at the CASSCF/NEVPT2 level, both in the training and the validations set, even if the BDE tend to be slightly underestimated. For the C-Cl bond in chlorobenzene, chloromethane and chloromethanol, BDE differences between QM and ReaxFF are less than $0.15\,\mathrm{eV}$. For the Cl-Cl bond, the difference is around $0.24\,\mathrm{eV}$, and for the Cl-O bond, the difference is equal to $0.26\,\mathrm{eV}$, which are acceptable values comparing other ReaxFF force fields in the literature. Concerning the C-Cl bond in chloroethane and chloroethanol, the differences in the BDE are 1.14 and $1.49\,\mathrm{eV}$, respectively, which is larger than the RMSD. In H-Cl, the BDE is $0.68\,\mathrm{eV}$ which is also larger than the RMSD. Nevertheless, these differences are still lower than the energy RMSD of $1.57\,\mathrm{eV}$

Molecule	Bond	CASSCF/NEVPT2	ReaxFF	set
Dichlore	Cl-Cl	2.81	2.57	T
Chlorobenzene	C-Cl	4.23	4.19	T
Chloromethane	C-Cl	3.48	3.33	T
Chloromethanol	C-Cl	3.56	3.44	V
Chloroethan	C-Cl	3.63	2.18	V
Chloroethanol	C-Cl	3.85	2.36	V
Methyl hypochlorite	O-Cl	1.92	1.66	T
Hydrogen chloride	H-Cl	4.63	3.95	V

Table 3.10: Bond dissociation energies (in eV) associated with the parameterized bonds in the considered molecule at the QM CASSCF/NEVPT2 level and the ReaxFF CHONCl-2022_weak level. The last column indicates if the corresponding bond is in the training set (T) or the validation set (V).

		QM				
Molecule	Bond	CASSCF	B3LYP	MP2	ReaxFF	set
Dichlore	Cl-Cl	2.02	2.05	2.00	2.20	T
Chlorobenzene	C-Cl	1.77	1.76	1.74	1.74	T
Chloromethane	C-Cl	1.81	1.81	1.78	1.82	T
Chloromethanol	C-Cl	1.83	1.84	1.80	1.87	V
Chloroethan	C-Cl	1.83	1.82	1.79	1.85	V
Chloroethanol	C-Cl	1.82	1.81	1.78	1.88	V
Methyl hypochlorite	O-Cl	1.71	1.73	1.69	1.81	T
Hydrogen chloride	H-Cl	1.27	1.29	1.27	1.29	V

Table 3.11: Comparison of the bond lengths obtained at several QM levels (CASSCF/NEVPT2, DFT, MP2) and the CHONCl-2022_weak force field. The last column indicates if the corresponding bond is in the training set (T) or the validation set (V).

of Hur et al.[235] and still acceptable.

Then we focus on the molecular geometry and compare the bond lengths obtained with ReaxFF and those obtained at several QM levels. The bond lengths are reported in the table 3.11. The bond lengths obtained with CHONCl-2022_weak well reproduced the QM values for bonds both in the training set and the validation set with a slight overestimation. The largest deviation is observed for the chloroethanol with 0.1Å.

CHONCl-2022_weak overestimates the Cl-Cl bond distance with a maximum error of $0.2\,\text{Å}$. Here, we reached the limitation of the optimization starting from a previous force field. Indeed, we could have improved the distances between the chlorine atom, but it has considerable repercussions. For example, shortening the Cl-Cl bond distance leads to high difficulties in reproducing the density of different chloride solvents. CHONCl-2022_weak reproduces well the densities of common chlorinated solvents, see table S2 in the supporting information. It also makes aromatic molecules with two chlorine atoms in ortho positions unstable, for example, in PCB structures. This small error on the Cl-Cl bond is thus a compromise and is still tolerable.

PES along bending angles

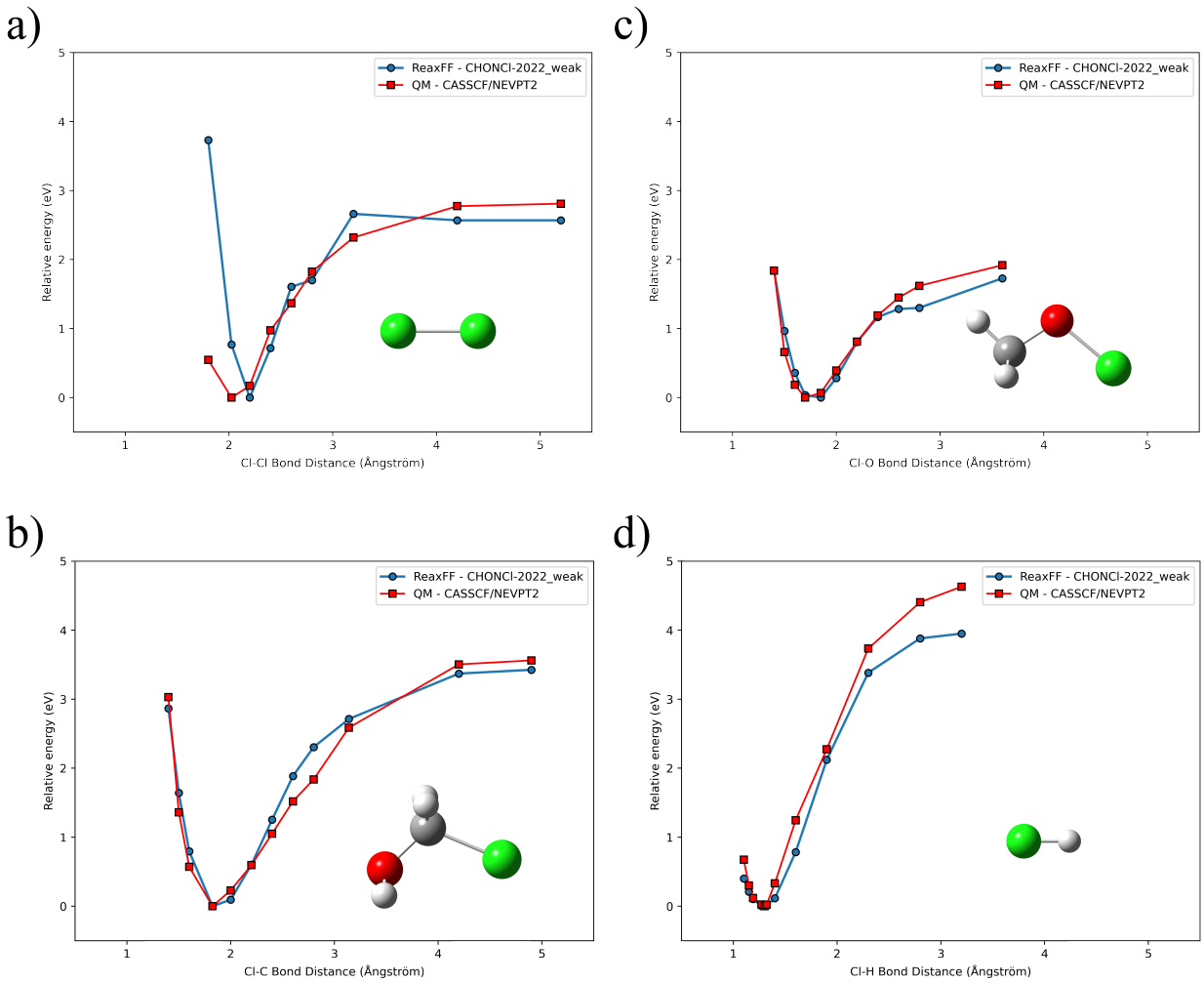


Figure 3.27: Comparison between PES (in eV) along bond dissociations obtained with ReaxFF (blue circle) and at the QM CASSCF/NEVPT2 level (red square) as a function of the bond distance (in Å). a) Cl-Cl bond in dichloride molecule; b) Cl-C bond in chloromethanol; c) Cl-O bond in acetyl chloride and d) Cl-H bond in hydrogen chloride. The molecules are depicted on the different graphics with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red.

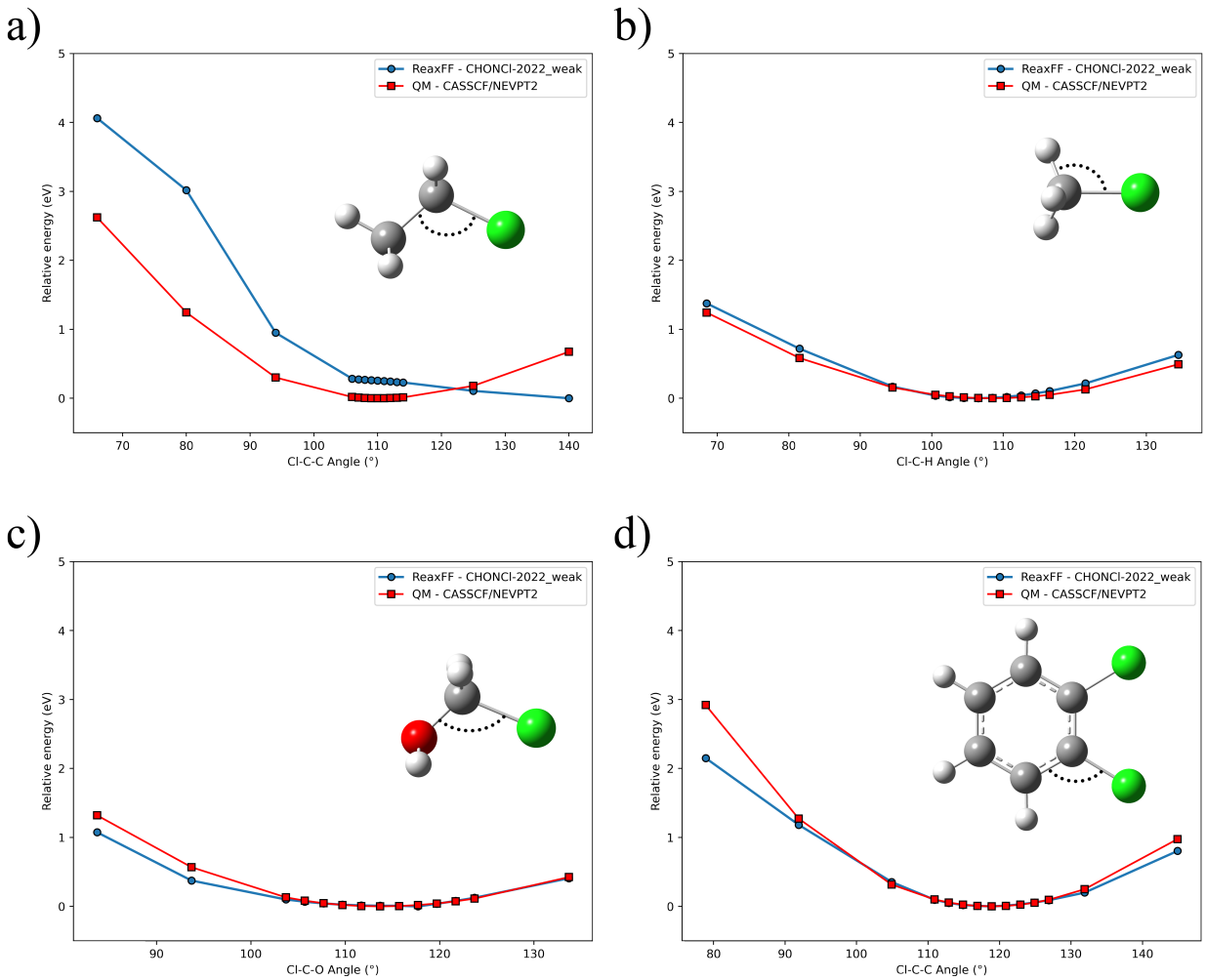


Figure 3.28: ReaxFF (blue circle) and QM-B3LYP/6-31+g(d)/D3BJ (red square) comparison of the relative energy (eV) as a function of the angle (in degrees) for the Cl-C-C angle of the chloroethane, the Cl-C-H of the chloromethane, the Cl-C-O of chloromethanol and the Cl-C-C of the ortho-dichlorobenzene. There is a representation of these molecules on the different graphics with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red. Black dotted points show the angle concerned by the PES.

Here PESs at the B3LYP/6-31+g(d) level of several bending angles are considered and compared to the one computed with CHONCl-2022_weak, see figure 3.28. From a general point of view ReaxFF PESs agree well with the DFT results. Concerning the Cl-C-C angle of the chloroethane, as in the case of BDE, it is responsible for the largest contribution of the RMSD for bending angle , which is 11.86°. In a non-conjugated system, the force field tends to overestimate the repulsion induced by the chlorine atom, leading to a shift of the PES of this bending angle. However, the difference coming from the table 3.12 is about 19°, which is still acceptable for the force field.

From the table 3.12, there is a good agreement between the angles obtained on the full optimized structure at the B3LYP/6-31+g(d) level and with CHONCl-2022_weak. Nevertheless, the C-C-Cl bending angle of the PCB101 with a chlorine atom in the ortho position is also overestimated at around 17°. This difference is due

Molecule	Angle	B3LYP	ReaxFF	set
PCB101	C-C-Cl (ortho)	120.2	137.8	V
PCB101	C-C-Cl (meta)	117.2	118.9	V
Chlorobenzene	C-C-Cl	119.3	121.5	V
Chloromethane	Cl-C-H	108.6	105.6	V
di-Chlorobenzene-ortho	C-C-Cl	118.8	118.7	V
Chloroethan	C-C-Cl	111.4	130.6	V
Methyl hypochlorite	C-O-Cl	110.4	119.3	V

Table 3.12: Comparison of the angle (in degrees) obtained at the B3LYP/6-31+g(d)/D3BJ level and the CHONCl-2022_weak force field. The last column indicates if the corresponding angle is in the training set (T) or the validation set (V).

to the dihedral angle around the central bond between the two cycles, which is small in the case of CHONCl-2022_weak. This leads to higher repulsion between the chlorine atom with the opposite hydrogen in the ortho position. However, when the two aromatic cycles of the PCB101 become more tilted, this bending angle rapidly decreases around $120°$, similar to the QM value.

PES along dihedral angles

This subsection compares the PESs of torsions along dihedral angles obtained at the B3LYP/6-31+g(d)/D3BJ level against the CHONCl-2022_weak values. In the case of PCBs, the torsion around the biphenyl central bond is of high interest as it plays a significant role in their toxicity[176]. The value of this central torsion for different PCBs, as well as some activation barriers are included in the validation. The PCBs considered in this work are depicted in figure S1 in the supporting information. PES along four dihedral angles are depicted in figure 3.29. The PES of the chloromethanol and PCB101 were in the training set, while the p-dichlorobenzene and the PCB169 were in the validation set.

For the dichlorobenzene and the chloromethanol, CHONCl-2022_weak reproduces almost identically the QM PES along the Cl-C-C-H and the Cl-C-O-H angles. At the same time, small differences are obtained between the two PCBs. For the PCB101, ReaxFF reproduces well the repulsion between the chlorine and hydrogen atoms when the C-C-C-C torsion is equal to $0°$ and $360°$. The position of the minimum is shifted between QM and ReaxFF PES, but it belongs to a flat part of the PES and is acceptable. Concerning the PCB169, the PES is slightly overestimated when it has a planar structure at $0°$ and $360°$ and a little underestimated when the two cycles are perpendicular, $90°$. Such deviation can be improved by using the torsions 2013 version available in the AMS Software[90] and the tapered bond orders of Furman and Wales[93], see table 3.14.

Table 3.13 presents the values of several kinds of dihedrals angles that were considered in the validation set. The values obtained with CHONCl-2022_weak are in good agreement with the QM ones and present a deviation lower than the global RMSD over dihedral angles of $4.12°$.

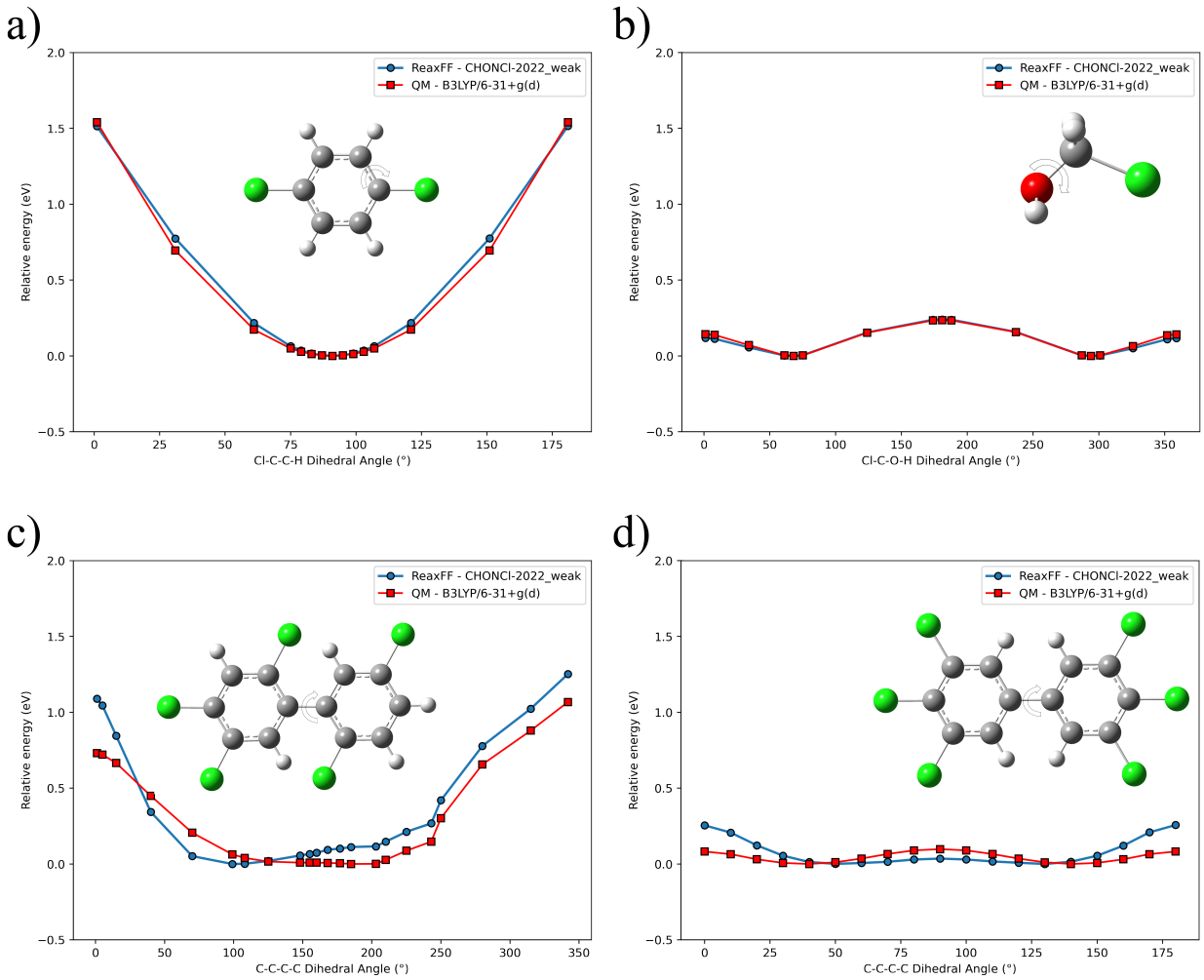


Figure 3.29: CHONCl-2022_weak ReaxFF (blue circle) and QM-B3LYP/6-31+g(d)/D3BJ (red square) PES (in eV) along several dihedral angles (in degrees). a) the Cl-C-C-H torsion of the p-dichlorobenzene, b) the Cl-C-O-H of the chloromethanol, c) the C-C-C-C torsion around the central bond of the PCB101 and the PCB169 (d). There is a representation of these molecules on the different graphics with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red. Black arrows indicate the considered dihedral angle.

Molecule	dihedral	QM	ReaxFF	Set
p-dichlorobenzene	Cl-C-C-H	360.0	360.0	V
Chlorobenzene	Cl-C-C-C	180.0	180.0	V
Chloromethanol	Cl-C-O-H	67.4	68.6	V
Chloroethanol	Cl-C-C-O	179.7	180.0	V
Methyl hypochlorite	H-C-O-Cl	180.0	180.0	V
Chloroethan	Cl-C-C-H	299.6	299.6	V
o-dichlorobenzene	Cl-C-C-Cl	360.0	360.0	V
PCB101	Cl(para)-C-C-C	179.9	183.2	V
PCB169	Cl-C-C-Cl	359.5	359.9	V

Table 3.13: Comparison of the dihedral angle (in degrees) obtained at the QM-B3LYP/6-31+g(d)/D3BJ level and the CHONCl-2022_weak force field. The last column indicates if the corresponding dihedral angle is in the training set (T) or the validation set (V).

	Dihedral angles (degrees)		
PCBs	QM	ReaxFF	Exp.
PCB15	39.3[b], 37.9[e]	33.0	45.1 ± 1.5[a]
PCB52	95.5[c], 87.3[d], 81.3[e]	76.8	
PCB77	38.7[c], 37.7[e]	32.7	
PCB101	99.1[c], 77.2[d], 81.0[e]	92.4	
PCB126	39.2[c], 37.9[e]	37.7	
PCB169	38.9[c], 37.9[e]	43.1	

	Energy barriers (in eV)			
	ΔE^0		ΔE^{90}	
PCBs	QM	ReaxFF	QM	ReaxFF
PCB15	0.12[b], 0.09[a], 0.09[e]	0.04	0.11[b], 0.09[a], 0.11[e]	0.09
PCB52	0.72[c], 1.24[e]	1.55	0.003[c], 0.001[e]	0.04
PCB77	0.08[c], 0.08[e]	0.05	0.09[c], 0.11[e]	0.11
PCB101	0.69[c], 1.23[e]	0.49	0.002[c], 0.001[e]	0.06
PCB126	0.08[c], 0.08[e]	0.04	0.09[c], 0.11[e]	0.08
PCB169	0.07[c], 0.08[e]	0.08	0.09[c], 0.11[e]	0.08

Table 3.14: Dihedral angles (in degrees) of the central C-C-C-C torsion along the biphenyl bond at equilibrium geometries from the literature or B3LYP/6-31+g(d)/D3BJ calculation level. The energy barrier of the rotation around the biphenyl bond (in eV) between the equilibrium geometry and the planar geometry, ΔE^0, or the perpendicular geometry ΔE^{90}. The values in the ReaxFF columns are obtained with the CHONCl-2022_weak force field.
[a]: Experimental value of Bastiansen and Samdal[169] ; [b]: this work B3LYP/6-31+g(d)/D3BJ ; [c]: Arulmozhiraja and Fujii[165] ; [d]: Gorbunova et al.[177]; [e]: Dorofeeva et al.[242], or with the new reactive force field. Rotational barrier ΔE^0 (planar structure) and ΔE^{90} (vertical structure) from QM are also compared with the ReaxFF values (in eV).

The rotational barrier of several PCBs was then considered to discuss the validity of the CHONCl-2022_weak force field. Hereafter the rotational energy barriers and the optimized values of the dihedral angle between the two phenyl rings are considered. Two quantities were computed: ΔE^0 is the energy barrier to reach a planar conformation from the optimized geometry, and ΔE^{90} is the energy barrier needed to reach a perpendicular conformation, with an angle of $90°$ between the two phenyl rings, from the optimized geometry. The values obtained with CHONCl-2022_weak are compared to those available in the literature or computed at several QM levels of calculation, see table 3.14. In 1981, McKinney and Singh[171] explained that the PCBs considered as toxic, have at least four laterals chlorine atoms. The PCBs with the lowest rotational energy barrier are more flexible, leading to better associations with bio-molecules and increasing their toxicity[176]. These rotational barriers are strongly impacted by the presence of chlorine atoms in ortho positions, which increases the needed energy to reach the planar conformation. In consequence, PCBs 126 and 169 presented in the SI figure S2 are among the more toxic and present only one or no chlorine atom in the ortho position.

Looking at table 3.14, the dihedral angle values of the torsion between the two phenyl rings are well reproduced as well as the relative values between PCBs, the largest error being a deviation of $5°$ for PCB77 with the closest reference value. These results enforce the validity of the Van der Waals parameters optimized for

the chlorine atom which mainly drive the equilibrium values along this torsional motion. Globally, CHONCl-2022_weak tends to underestimate the values slightly. The same conclusion can be drawn considering the energy barriers. CHONCl-2022_weak reproduces well the significant increase of the energy barrier ΔE^0 in the case of the presence of chlorine atoms in ortho positions. For example, one can consider the comparison of ΔE^0 and ΔE^{90} in the case of PCB101 and PCB52 both wearing chlorine atoms in ortho positions. On the contrary, the PCBs that do not have ortho chlorine atoms have a low rotational barrier leading to high flexibility. For example, for the most toxic PCBs 126 and 169, energy barriers are in agreement with reference values provided in table 3.6 and literature investigations of these PCBs assuming their flexibility.

Charge distribution

In figure 3.35 we focus on the carbon charge distributions of PCBs 52, 101 and 22. Literature values obtained with the Mulliken approximation from DFT calculations [177] are compared against ReaxFF ones using CHONCl-2022_weak to determine potential centres of nucleophilic and electrophilic additions. The variations of the charges are consistent over the carbon atoms when comparing the two levels of calculation. For the PCB22, presented in the figure 3.30 a), ReaxFF and DFT show a large positive charge on the carbon atoms 2, 3 and 4', which might be considered the most favourable nucleophilic sites. Moreover, the carbon 5,6 and 6' for ReaxFF and DFT show a negative charge leading to a most favourable zone for electrophilic additions. Continuing this comparison for the PCB 52 in the figure 3.30 b), the carbon atoms 2 and 2' show the highest positive values for both methods, leading to the most favourable nucleophilic sites. Finally, for the PCB 101 figure 3.30 c), carbon atoms 4 and 5 exhibit the two highest positive charges for the two methods.

Validation against reaction pathways

Finally, we considered reaction pathways of small chlorinated molecules to compare the ReaxFF energies obtained with CHONCl-2022_weak with QM energies. Figure 3.31 displays four reaction pathways obtained at the B3LYP/6-31+G(d)/D3BJ level using a chain of state approach: a) the reaction between chloromethane and dihydrogen $CH_3Cl + H_2 \longrightarrow CH_4 + HCl$; b) the reaction between hydrogen-chloride and formic acid $HCOOH + HCl \longrightarrow HCOCl + H_2O$; c) the recombination of the HOCl molecule in HClO and d) the dissociation of chloromethanol in formaldehyde and hydrogen-chloride $CH_2Cl-OH \longrightarrow H_2CO + HCl$. The energy barriers obtained using the CHONCl-2022_weak force field are in the same order of magnitude as those obtained at the QM level, the largest error being $0.54\,\mathrm{eV}$ for the dissociation of chloromethanol. Moreover, it shows similar energetic differences between reactants and products with the maximum deviation of $0.57\,\mathrm{eV}$.

In the case of the acid-base reaction presented in figure 3.31 b), ReaxFF underestimate the energy barrier

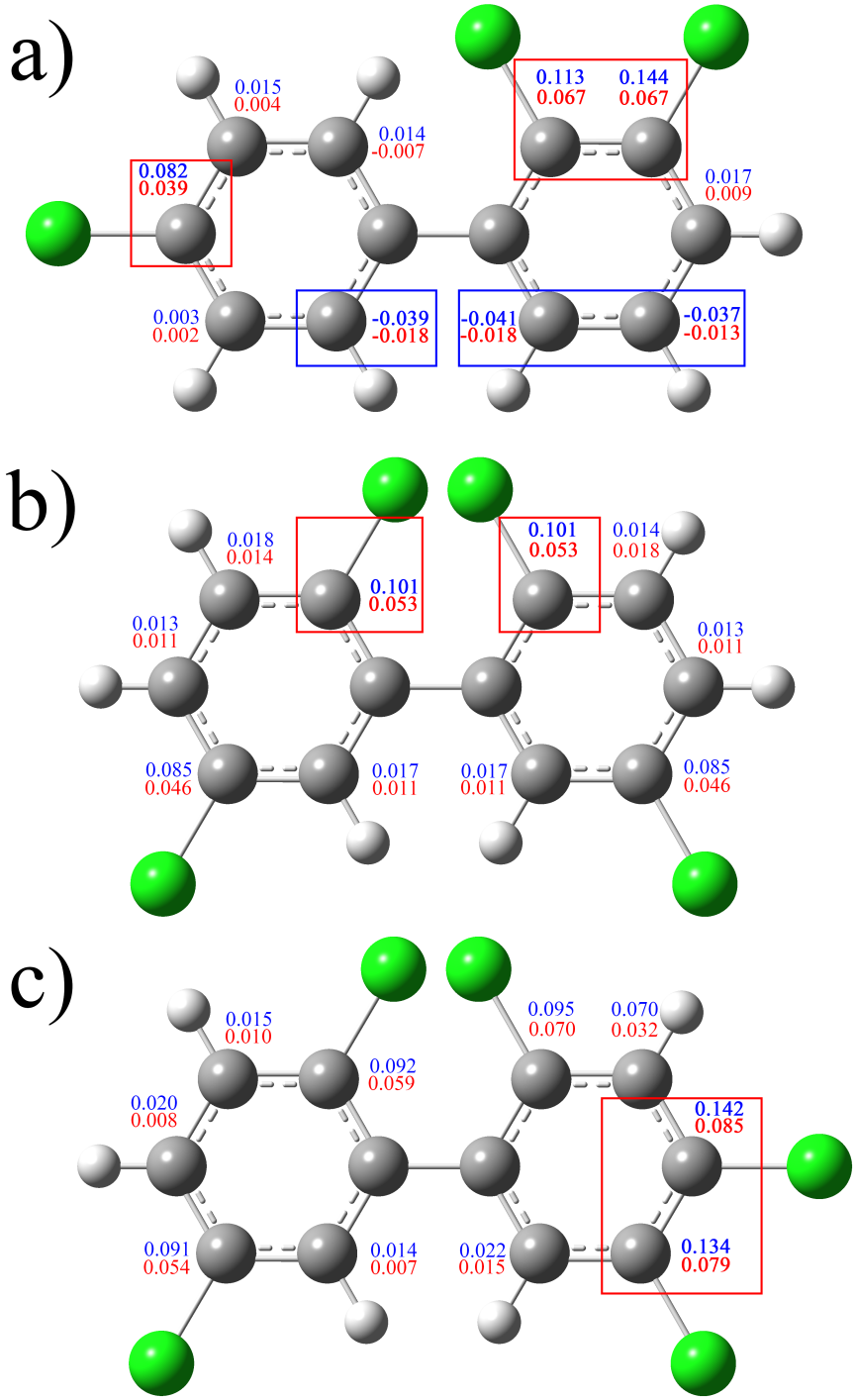


Figure 3.30: Carbon atomic charges of several PCBs, a) PCB 22, b) PCB 52, and c) PCB 101, computed with ReaxFF CHONCl-weak_2022 (in blue) and DFT from the reference 177 (in red). Blue rectangles indicate the favourable carbon atoms for electrophilic additions and in red for nucleophilic.

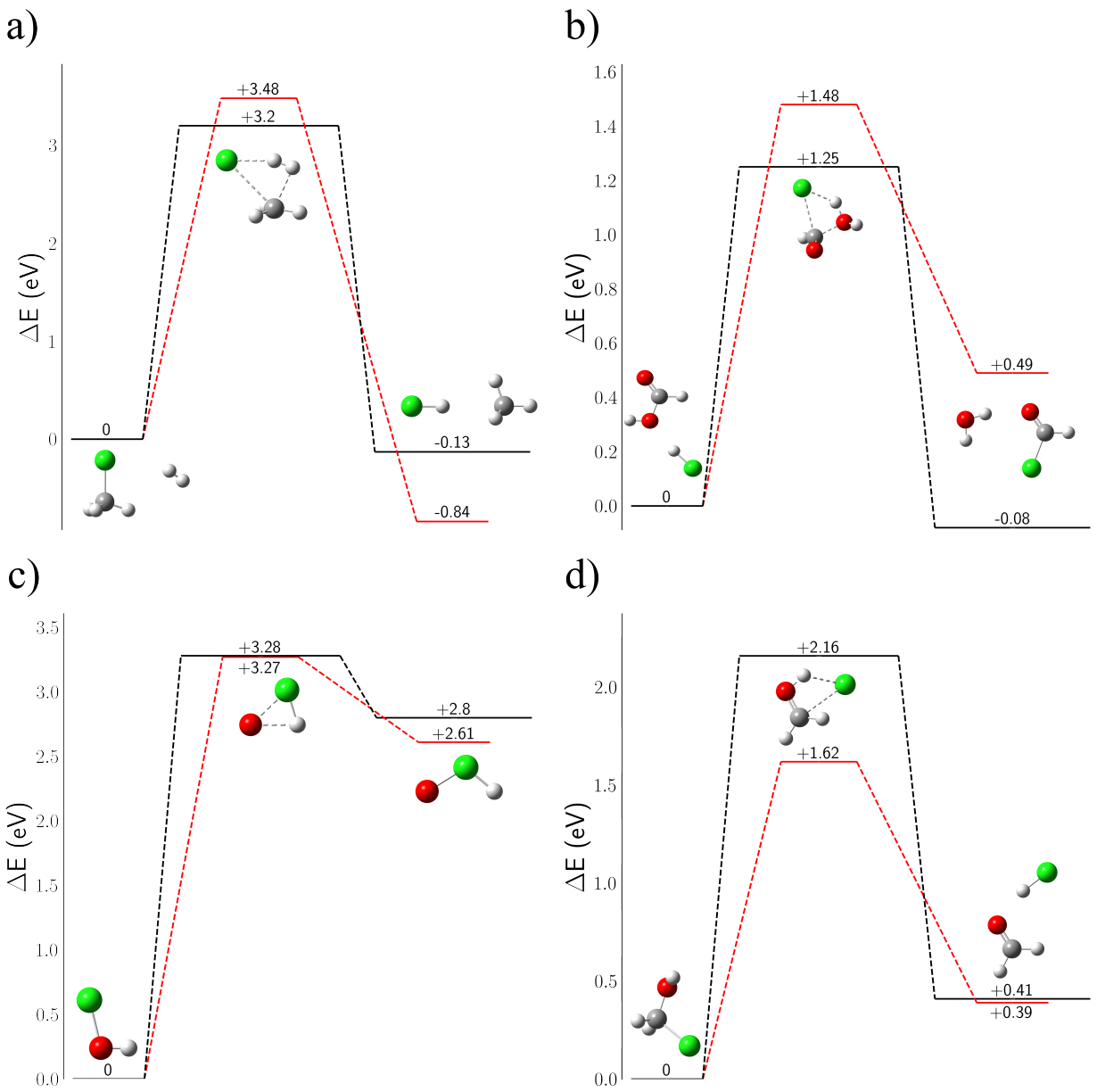


Figure 3.31: Reaction pathways obtained at the B3LYP/6-31+g(d)/D3BJ level (red) and with CHONCl-2022_weak on the QM geometries (black) for the reactions a) $CH_3Cl + H_2 \longrightarrow CH_4 + HCl$, b) $HCOOH + HCl \longrightarrow HCOCl + H_2O$, c) the recombination of HOCl in HClO and d) the dissociation $CH_2Cl-OH \longrightarrow H_2CO + HCl$. There is a representation of these molecules on the different diagrams with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red.

of the reaction and the energy of the products. The gap overestimation between the transition state and the product might be due to the BO calculations in the case of acyl chloride systems with a carbon atom bearing both a chlorine atom and a C-O double bond. Indeed, the BO of C-Cl being lower than the BO of the C-O bond, the BO of the C-Cl bond is weakened at the expense of the BO of the C-O double bond. However, the difference between the QM energies and ReaxFF is about $0.57\,\mathrm{eV}$ and is still tolerable.

3.6.3.3 Atmospheric oxidation of chlorobenzene

In order to evaluate the performance of the new CHONCl-2022_weak force field, the oxidation of chlorobenzene initiated by hydroxyl radical is investigated. This subsection is divided into two parts. Firstly, reactive pathways of chlorobenzene obtained at the M062X/6-311++G(2df,2p) in relation to the work of Wu et al. [243] are compared with ReaxFF calculations with the new CHONCl-2022_weak force field. Kinetic and thermodynamics products are identified. Secondly, the regioselectivity of the hydroxyl radical oxidation process on the chlorobenzene is investigated, and the results are compared with the kinetic results of Bryukov et al. [244].

Investigation of the reactive pathway of chlorobenzene oxidation by hydroxyl radical in the presence of dioxygen

In this part, we consider the reactivity of chlorobenzene at the M062X/6-311++G(2df,2p) calculation level, consistently with the work of Wu et al. [243]) and using ReaxFF calculations with the CHONCl-2022_weak force field. The atmospheric gas-phase oxidation of the chlorobenzene initiated by hydroxyl radicals is investigated and presented in the figure 3.32. In the figure 3.32 a), the diagram corresponds to the reaction between the hydroxyl radical and the chlorobenzene in the ipso position corresponding to the carbon linked to the chlorine atom. Panel b), c), and d) correspond to the reaction between the hydroxyl radical and the chlorobenzene in ortho, meta and para positions, respectively, followed by the addition of dioxygen. For each of the different reaction pathways, the initial point, in the middle of each scheme, is the chlorobenzene with a hydroxyl radical. ReaxFF energies were computed using the QM optimized geometries. Two pathways were investigated by considering either the abstraction of a hydrogen atom leading to the formation of H_2O (the right side of each diagram of figure 3.32); or the addition of the hydroxyl radical on the aromatic rings (the left side of each diagram of figure 3.32). The only exception is the ipso reaction, where the right side of the diagram corresponds to the substitution of the chlorine atom by the hydroxyl radical. For ortho, meta and para reaction pathways, after both the addition or the H-abstraction, an addition of an O_2 molecule is then implemented to obtain final products, usually named highly oxygenated organic molecules (HOM) or chlorophenols. The energy barriers are in electron-volt and relative to the initial reactants except for the last products. After adding a dioxygen molecule, the barriers are relative to the previous states. QM calculations associated with the hydroxyl radical

addition (left side), were reported from the work of Wu et al.[243]. Globally CHONCl-2022_weak reproduces well the relative values of the QM energies of the reactive pathways but tends to overestimate the energy barriers of the reactions with the radicals.

Concerning the reaction of the OH radical in the ipso position (figure 3.32 a), the addition exhibit the largest energy barriers of all the reactive pathways, with both ReaxFF ($1.12\,\mathrm{eV}$ and $1.02\,\mathrm{eV}$) and QM levels of calculations ($0.24\,\mathrm{eV}$ and $0.24\,\mathrm{eV}$). These results agree with the conclusion of Wu et al.[243] and Bryukov et al.[244] that the ipso addition is, at ambient temperature, the least favoured reactive pathway. Moreover, the energy barrier of the reaction of the chlorine atom substitution (right side) is also a non-favoured pathway.

For the reactions in positions ortho, meta and para, figure 3.32 b, c and d, respectively, ReaxFF calculations reproduce the trends obtained at the QM level of calculation. The first tendency that we can notice, without considering the presence of dioxygen, is that both ReaxFF and QM calculations show that the addition of OH on the chlorobenzene (left side) is thermodynamically favoured. Indeed, in the three cases, the energies of products are lower than the energies of the products obtained after the hydrogen abstractions. However, we can observe a difference in the kinetic products obtained from ReaxFF and QM calculations. Considering the energy barriers, differences might be explained by the π-stacking arrangement of the hydroxyl radical above the benzene ring in the transition states geometries obtained from QM calculations which are not well described with ReaxFF due to the CHO parametrization. However, one specificity of these pathways is that the energy barriers on the two sides are closed to each other with differences at most about $0.1\,\mathrm{eV}$ in the range of the accuracy of the methodology. Our new ReaxFF potential well reproduces this point; thus, from a kinetic point of view, it is difficult to conclude on the most favoured pathway.

Now, if we consider the presence of the dioxygen molecule, we observe a good agreement between ReaxFF and QM calculations on products' energies and energy barriers. On the left side, after OH addition, the final products are the chlorophenol molecule in ortho, meta or para positions and the HO_2 molecule as suggested by Wu et al.[243]. On the right side, both QM and ReaxFF calculation levels do not present energy barriers from the products of the hydrogen abstraction to the formation of HOM with a dioxygen molecule crafted on the chlorobenzene ring. In the presence of dioxygen, ReaxFF with the new CHONCl-2022_weak force field predict that the formation of these HOM compounds are thus likely to happen[245]. These results complete the previous work of Wu et al.[243] on the chlorobenzene oxydation.

Regioselectivity of the chlorobenzene oxidation

Bryukov et al.[244] investigated the kinetics of the OH additions and the hydrogen abstractions on the chlorobenzene molecule from a combination of experimental and computational approaches. They agree with the QM conclusion of Wu et al.[243] by assuming that the primary reaction occurring at low temperatures

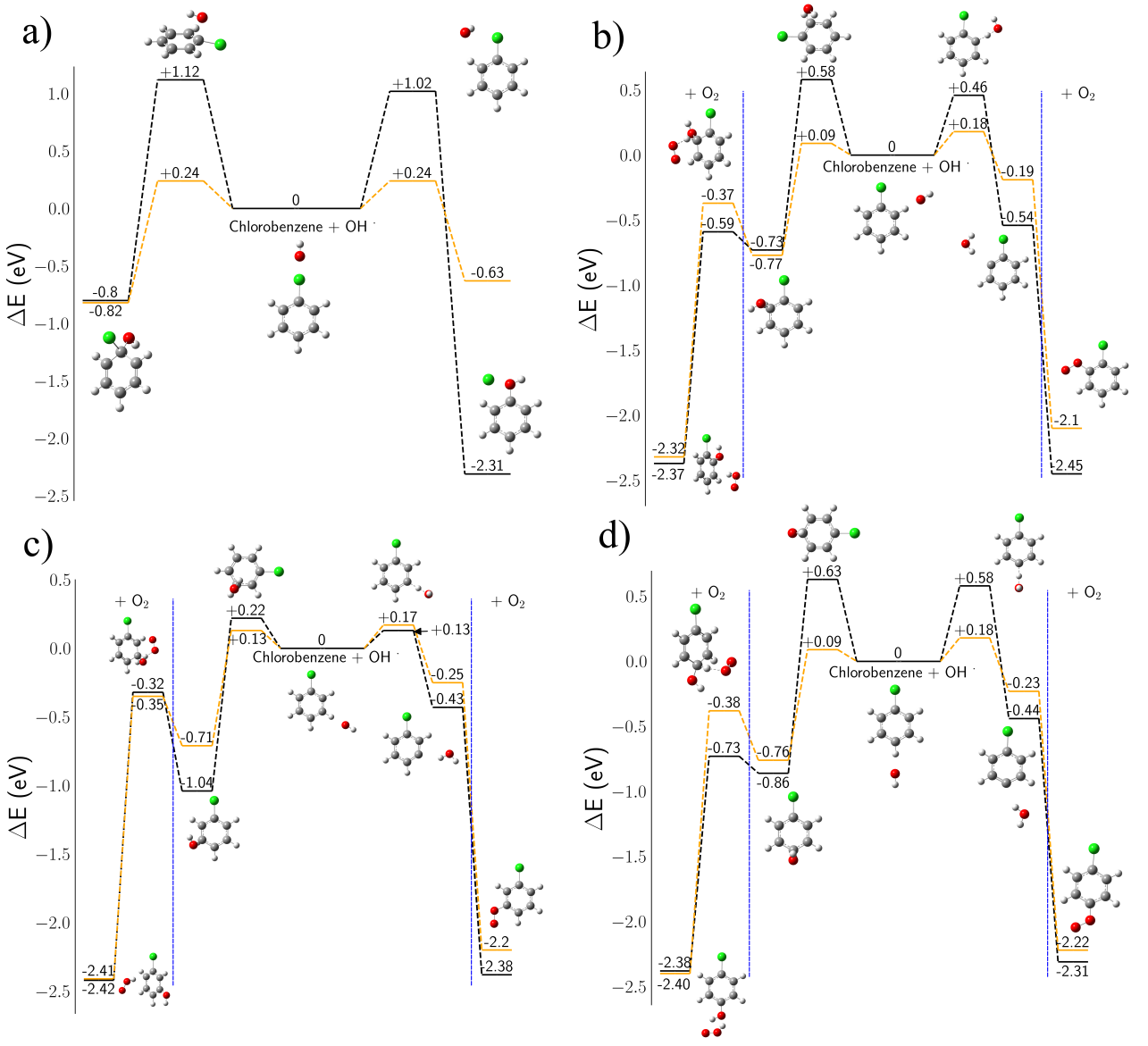


Figure 3.32: Reaction pathways obtained at the M06-2X/6-311++g(2df,2p) level (orange from reference 243 or this work) and using ReaxFF with CHONCl-2022_weak (black). For all schemes, the initial reactants are chlorobenzene and hydroxyl radical in the middle. The left side corresponds to the addition of the hydroxyl radical, while the right side corresponds to the abstraction of a hydrogen atom of the phenyl ring. The reaction is done on the substituent in position a) ipso, b) ortho, c) meta and d) para. The final products are obtained after the reaction with dioxygen. The energy barriers are in electron-volt and relative to the initial reactants except for the last products, after the addition of dioxygen materialized by the blue vertical dashed lines, where in these cases, the barrier is relative to the previous state. There is a representation of all the different molecules on the different diagrams with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red.

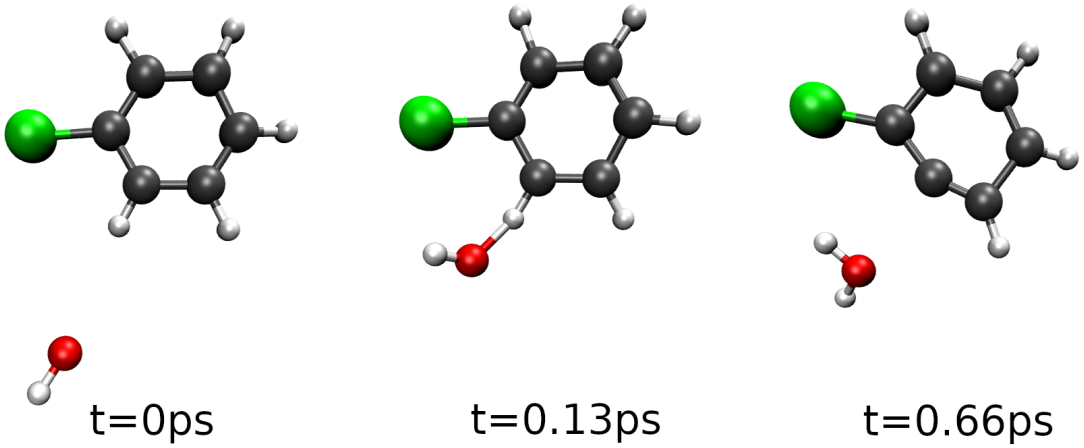


Figure 3.33: Representation of key snapshots along one reaction pathway leading to the abstraction of an hydrogen atom in ortho position. The molecules are depicted on the different graphics with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red.

are the OH addition and they showed that a shift appears around 450K, the hydrogen abstraction is becoming prominent. Moreover, they supposed that the hydrogen abstraction might occur at a lower temperature. To compare the values obtained by Bryukov et al.[244] with ReaxFF, the oxidation of the chlorobenzene molecule was investigated through reactive molecular dynamics simulations using the CHONCl-2022_weak force field.

In order to simulate atmospheric collisions, the chlorobenzene molecule was subjected to hydroxyl radical collisions coming from random positions sampled on a sphere centred on the centre of mass of the molecule. The initial velocity of the hydroxyl radical is oriented toward the centre of mass of the chlorobenzene molecule, and the velocity norm is sampled around the considered temperature. Then 8000 NVE simulations of 3 ps lengths were implemented to gather enough statistics using initial velocities at 80K, 250K, 300K, 350K and 800K. Most of the trajectories were inefficient and did not lead to any reaction, although, as expected, increasing the temperature leads to more reactions. Over the total of 8000 simulations for each temperature, branching ratio were computed over efficient trajectories, considering 108 reactions at 80K, 201 at 250K, 251 at 300K, 314 at 350K and 517 at 800K. Figure 3.33 presents an example of one trajectory along the abstraction of one hydrogen atom in ortho position. The results are presented in the table 3.15.

With ReaxFF, due to the lack of π arrangement description, the addition of the OH is rare. So, we compare the rate constant of the H-abstraction reactions and the chloride substitution computed by Bryukov et al.[244] with those obtained in our ReaxFF simulations. Bryukov et al.[244] computed the rate constant with different methods, but the IRTS (Isodesmic reactions for transition states) shows the best capacity to reproduce the experimental results[246,247].

The lowest temperature investigated by Bryukov et al.[244] is 251.68K. At this temperature, the rate con-

Products	80K	250K	300K	350K	800K
Ortho abstraction	32.4	38.3	41.4	42.7	39.5
Meta abstraction	49.1	46.3	34.3	31.9	25.7
Para abstraction	9.3	8.5	13.6	12.4	6.4
Cl Substitution	9.3	6.5	8.8	12.4	24.2
HOCl Formation	-	0.5	2.0	0.6	4.1
Para OH addition	-	-	-	-	0.2

Table 3.15: Branching ratio (in %) over the efficient trajectory of a hydroxyl radical collision with chlorobenzene at 80K, 250K, 300K, 350K and 800K, obtained with ReaxFF and the CHONCl-2022_weak force field.

stants scale is as follows: the smallest is the chlorine substitution; the H-abstraction in para and ortho positions are one order of magnitude larger, the H-abstraction in ortho being favoured; finally, the H-abstraction in meta position is the kinetic product with a rate constant one order of magnitude larger than the ortho and para cases. These results agree with the branching ratio presented in table 3.15. The H-abstraction in the meta position is the more probable, while the chlorine substitution is unlikely to occur. However, the efficiency of the H-abstraction in ortho and para positions is not well reproduced, although they are effectively less probable than the meta H-abstraction. These results might come from the geometry of the complex between the OH radical and the chlorobenzene molecule, far away from the transition state which are not included in the calculations done by Bryukov et al.[244]. From our QM calculations, the OH radical comes first between two carbon atoms before it reaches the transition state. Looking at the energy of the transition states on figure 3.32, the ortho position is favoured against the ipso position, the meta position is favoured against the para position, and the meta position is favoured against the ortho position. This is consistent with the branching ratio obtained in table 3.15. The calculations of the rate constants by including a two steps model with the approach of the hydroxyl radical may help to consider that point.

Increasing the temperature leads to a balance shift between the reaction pathway on positions ortho and meta, the para position remaining disfavored and the chloride atom substitution being more and more likely to occur. Bryukov et al.[244] showed that increasing the temperature reduces the difference between the rate constants of the chloride atom substitution and the three reaction pathways associated with the H-abstraction. The chloride atom substitution becomes prominent above $2343.75\,\mathrm{K}$[244]. Although our ReaxFF simulations with CHONCl-2022_weak slightly overestimate this substitution pathways they agree with the trend of this reaction pathway to become favoured by increasing the temperature. Moreover, these results are consistent with reaction pathways calculations, on figure 3.32, where the chloride atom substitution corresponds effectively to the thermodynamic product, whatever the QM or ReaxFF calculation level.

The approach presented here, averaging molecular collisions using ReaxFF simulations, may help investigate atmospheric reactions using a few calculation times. ReaxFF results being reliable, such collision analyses may help probe the possible reactive pathways quickly and gather the possible reaction products.

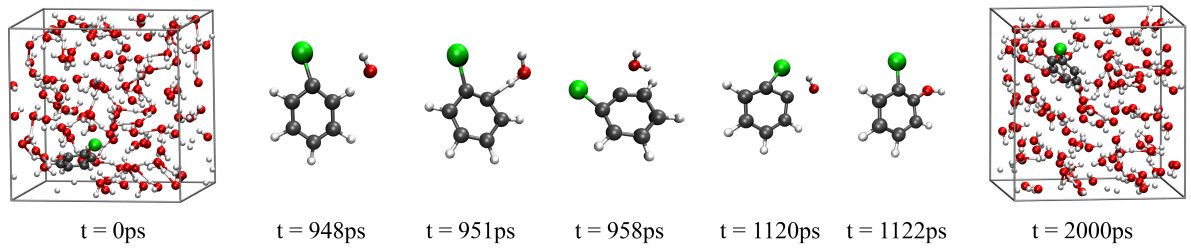


Figure 3.34: Initial and last simulation boxes of the plasma-induced reactivity of chlorobenzene in water along with intermediate key snapshots of the reaction pathway and their corresponding timestep.

3.6.3.4 Plasma-induced degradation of chlorobenzene in water

Finally, in order to demonstrate our new potential's capability, simulations corresponding to plasma-induced degradation of chlorobenzene in water were implemented with the CHONCl-2022_weak force field. To that ends, we considered a simulation box, of 15 length, containing one chlorobenzene molecule, 102 water molecules and 10 hydroxyl radicals to mimic the effect of the plasma. All compounds were randomly positioned in the box. The box was heated between 550 K and 750 K to fasten the reaction during 2 ns. A total of 35 simulations were made, and the results were compared to those obtained from the experimental investigation of Liu and Jiang[248]. Figure 3.34, presents an example of a simulation box at the first time step and at the last one, along with key snapshots illustrating the reactivity along one simulation.

Among the 35 simulations, the degradation of chlorobenzene only occur in two cases. One leads to the formation of a phenol molecule and the other one leads to a chlorophenol molecule, the hydroxyl radical being crafted on the ortho position. Figure 3.34 presents the formation of the chlorophenol molecule. These results agree with the conclusions of Liu and Jiang[248] who showed from plasma-induced experiments that the main intermediates products of the chlorobenzene decomposition are the three chlorophenols and the phenol compounds. This opens the field of application of CHONCl-2022_weak force field to the investigation of organochloride decomposition from plasma-induced degradation in water.

3.6.4 Conclusion

In this work, we developed a new reactive potential in the ReaxFF formalism called CHONCl-2022_weak force field. It provides the chlorine atom parameters for organochloride compounds to the CHON-2017_weak force field and focuses on conjugated and aromatic compounds. The force field parameters were fitted against high-level quantum mechanics calculations, including CASSCF/CASPT2 calculations for bond dissociation and DFT calculations for PESs along internal coordinates or reaction pathways. Its performance was evaluated against a validation set. We obtained an RMSD against QM energies of $0.38\,\mathrm{eV}$ ($8.91\,\mathrm{kcal\,mol^{-1}}$). Moreover, the RMSD over structural parameters of the new force field is 0.06Å for the bond lengths, $11.86°$ for the angles and $4.12°$ for the dihedral angles. The bond dissociation energies and PESs along bending and dihedral angles,

computed at a QM level, are well reproduced. Moreover, the central torsion angle of molecules from the PCB family agree well with reference values from a structural and energetic point of view which is of great interest as this torsion is a key factor in the toxicity of these compounds. Finally, the charge distribution over PCB molecules, computed with the new force field agrees with the one computed at the DFT level, allowing the identification of favourable zones for reactive nucleophilic addition on these compounds.

The atmospheric oxidation of the chlorobenzene molecule initiated by hydroxyl radical was investigated with the new force field and compared with DFT calculations. Two main reaction pathways were investigated: the addition of the OH radical on the chlorobenzene and the abstraction of one hydrogen atom or the substitution of the chlorine atom by the radical. Although ReaxFF tends to overestimate the energy barriers, the new force field reproduces the trends of QM results. The energy barriers of the two considered pathways are very closed. The reaction on the ipso position exhibits the largest energy barrier but leads to the lowest energy products. Considering the presence of a dixoygen molecule, the new force field reproduces the formation of chlorophenol and highly oxygenated molecules. To further increase the accuracy of the calculations, it could be needed to consider a better description of the π-stacking interaction in order to improve the interaction between the aromatic ring and the OH radical.

The regioselectivity associated with the reaction with an OH radical was also compared to rate constant calculations. Using an original approach, we simulated the atmospheric collisions between the chlorobenzene molecule and the OH radical and gathered statistical data to analyze the more probable reaction pathways. Again, the results obtained with the CHONCl-2022_weak force field agree with the literature, and we evidence the shift between the H-abstraction and the chloride atom substitution when increasing the temperature. This opens the possibility of probing reaction pathways and identifying reaction products quickly.

Finally, the degradation of chlorobenzene in water from the reaction with hydroxyl radicals was investigated using the CHONCl-2022_weak force field. The formation of phenol and chlorophenol molecules is obtained which agree with experimental investigation which identified these molecules as the major intermediates. This opens the filed of application of reactive molecular dynamic simulations using ReaxFF to the investigation of the degradation of organochloride compounds in water.

Acknowledgments

This work was achieved using HPC ressources from GENCI-CINES (grant 2021–A0100806920), MCIA (Mésocentre de Calcul Intensif Aquitain) and the Université de Pau et des Pays de l'Adour.

Supplementary Material

Supplementary material is available and provides (i) computational details about CASSCF/CASPT2 calculations (ii) density of several solvents computed with ReaxFF (iii) the structure and nomenclature of PCBs and (iv) reaction energies used to obtain the parameters. An archive is provided including data files, in the usual format for parameters optimization with AMS, containing the energies and the structures used to build the training set and the validation set and the force field parameters in a suitable format for LAMMPS or AMS.

Molecule	Bond	Active space
Dichlore	Cl-Cl	(12,10)
Chlorobenzene	C-Cl	(12,10)
Chloromethane	C-Cl	(10,9)
Chloromethanol	C-Cl	(14,14)
Chloroethan	C-Cl	(14,14)
Chloroethanol	C-Cl	(14,14)
Methyl hypochlorite	O-Cl	(14,14)
Hydrogen chloride	H-Cl	(12,10)

Table 3.16: This table presents the active spaces selected for the different CASSCF computations using the cc-pVTZ basis set. For the active space, the first number indicates the number of valence electrons distributed and the second the number of molecular orbitals.

Solvent	Density Literature	Density ReaxFF
Chlorobenzene	1.10	1.34
CH2Cl2	1.29	1.36
CH3Cl (249K)	1.00	1.28
CH2CHCl	0.91	1.09

Table 3.17: The densities (in g.cm^{-3}) of four different chlorine solvents obtained with the new reactive force field are compared with literature values. Except for the CH3Cl, which was computed at 249K, all the simulations were made at 300K.

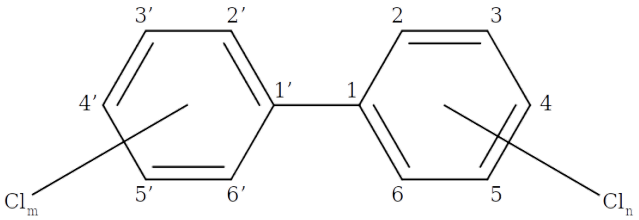


Figure 3.35: Here is a schematizing of a PCB, with the different positions indicated by numbers.

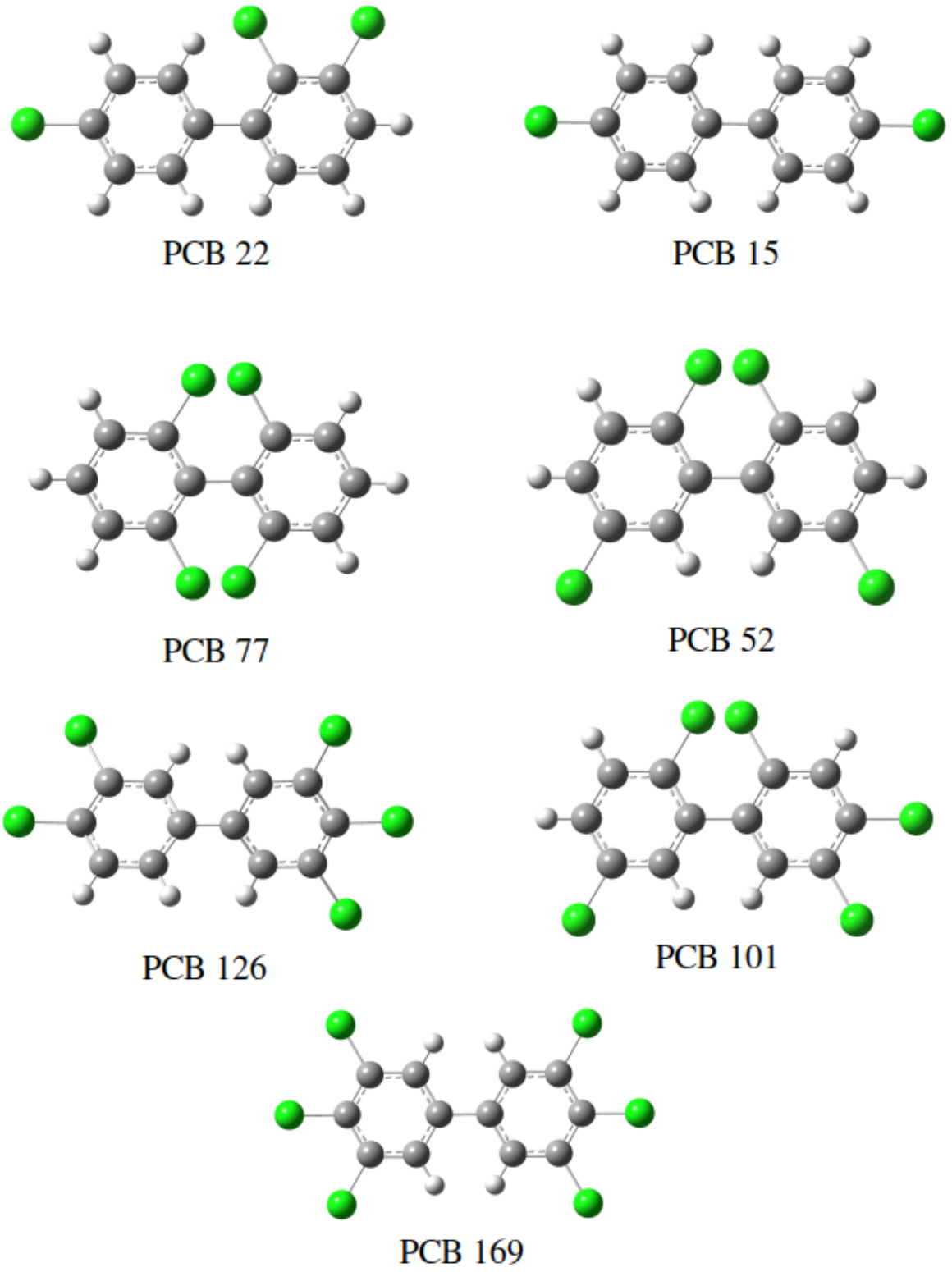


Figure 3.36: Representation of six different PCBs cited in the article. The PCB 15, 22, 52, 77, 126 and 169. The grey atoms are carbons, green chlorines, and white hydrogens.

Reaction	ReaxFF	QM
$CH_3Cl + H_2 \longrightarrow CH_4 + HCl$	-0.84	-0.13
$CH_2CHCl + O_2 \longrightarrow OOCHClCH_2$	+2.06	+1.28
$CH_2O_2 + HCl \longrightarrow H_2O + ClCOH$	+0.49	-0.08
$ClOH \longrightarrow OClH$	+2.61	+2.78
$HOCH_2Cl \longrightarrow HCl + OCH_2$	+0.39	+0.41
$C_6H_5Cl + H_2 \longrightarrow C_6H_6 + HCl$	+2.88	+0.66

Table 3.18: Comparison of reaction energies obtained from QM and ReaxFF (eV).

C-Cl bond dissociation	ReaxFF	QM
$C_6H_5Cl \longrightarrow C_6H_5 + Cl$	+4.25	+4.27
$C_2H_5ClO \longrightarrow C_2H_5O + Cl$	+2.38	+3.85
$C_2H_5Cl \longrightarrow C_2H_5 + Cl$	+2.19	+3.66
$CH_3Cl \longrightarrow CH_3 + Cl$	+3.36	+3.50
$CH_2OHCl \longrightarrow CH_2OH + Cl$	+3.47	+3.59
O-Cl bond dissociation	**ReaxFF**	**QM**
$CH_3OCl \longrightarrow CH_3O + Cl$	+1.67	+1.93
H-Cl bond dissociation	**ReaxFF**	**QM**
$H - Cl \longrightarrow H + Cl$	+4.63	+3.95
Cl-Cl bond dissociation	**ReaxFF**	**QM**
$Cl_2 \longrightarrow Cl + Cl$	+2.57	+2.81

Table 3.19: Comparison of bond dissociation energies obtained from QM and ReaxFF (eV).

dichlorobenzene	ReaxFF	QM
meta-dichlorobenzene	REF	REF
ortho-dichlorobenzene	+0.49	+0.08
para-dichlorobenzene	+0.48	0
dichloroethene	**ReaxFF**	**QM**
cis-1,2-Dichloroethene	REF	REF
trans-1,2-Dichloroethene	-0.06	+0.03
PCB	**ReaxFF**	**QM**
PCB1	REF	REF
PCB2	-0.08	-0.08
PCB3	-0.08	-0.08

Table 3.20: Comparison of energies difference between different isomers, the reference is indicated by REF (eV).

	Planar Structure		Vertical Structure	
Energy rotational barrier of PCBs	ReaxFF	QM	ReaxFF	QM
PCB15	+0.04	+0.09	+0.09	+0.09
PCB52	+1.55	+1.24	+0.04	+0.003
PCB77	+ 0.05	+0.08	+0.11	+0.11
PCB101	+ 0.49	+0.69	+0.06	+0.002
PCB126	+0.04	+0.08	+0.08	+0.09
PCB169	+0.08	+0.08	+0.08	+0.09

Table 3.21: Comparison of conformers energies of torsional barriers from different PCBs (eV).

OH addition on chlorobenzene	ReaxFF	QM
$C_6H_5Cl + OH \longrightarrow C_6H_5ClOH$ – ipso	-0.8	-0.82
$C_6H_5Cl + OH \longrightarrow C_6H_5ClOH$ – ortho	-0.77	-0.77
$C_6H_5Cl + OH \longrightarrow C_6H_5ClOH$ – meta	-1.04	-0.71
$C_6H_5Cl + OH \longrightarrow C_6H_5ClOH$ – para	-0.86	-0.76
O_2 addition on chlorobenzene	ReaxFF	QM
$C_6H_5Cl + OH + O_2 \longrightarrow C_6H_5ClO_2$ – ortho + H_2O	-2.45	-2.1
$C_6H_5Cl + OH + O_2 \longrightarrow C_6H_5ClO_2$ – meta + H_2O	-2.2	-2.38
$C_6H_5Cl + OH + O_2 \longrightarrow C_6H_5ClO_2$ – para + H_2O	-2.22	-2.31

Table 3.22: Comparison of energies reactions, with the OH addition on chlorobenzene, and the addition of O_2 on the chlorobenzene (eV).

Molecule	$\Delta_f H^{ReaxFF}$	QM	difference
Metazachlor	-199.90	-200.93	1.03
DDT	-157.71	-160.33	2.62
TCDD	-137.96	-139.19	1.23
Diuron	-133.43	-133.82	0.39
Oxazépam	-194.79	-192.85	1.94
PCB15	-132.47	-132.28	0.19
PCB52	-130.20	-130.06	0.15
PCB77	-130.13	-129.52	0.61
PCB101	-129.01	-128.78	0.23
PCB126	-128.92	-127.91	1.01
PCB169	-127.63	-126.24	1.38

Table 3.23: Comparison of Heat of formation computed from equation 14 of reference[5] and QM calculations.

3.7 Conclusion

Thus in the article, we make sure that the force field could reproduce correctly the key parameters identified in chapter II on the PCBs as the central torsion and the rotational barriers. Nevertheless, one point not mentioned in the article is the central bond distance between the two carbons atoms as a function of the chlorine substituted in the PCBs. This investigation was made at the CASSCF/CASPT2 level in comparison with the new ReaxFF force field, and it was observed that the chlorine position did not impact the central bond distance on the PCB, which stays around 1.48 Å.

Moreover, with the new force field, we successfully investigate the degradation in an atmospheric and aqueous environment for chlorobenzene. The force field's evaluation performance indicated that ReaxFF lacks in describing the π arrangement. Otherwise, it is able to describe all the other aspects of PCB reactivity. It allows us now to investigate the different organochloride degradation reactivities in different environments in the two following chapters. In the continuity of the thesis, the following chapter presents a pyrolysis investigation using the new force field. A dedicated results section is made on the pyrolysis of the chlorobenzene, completing the description of chlorobenzene degradation in various environments.

Chapter 4

Pyrolysis investigation using ReaxFF

4.1 Introduction

After developing the new aqueous force field CHONCl-2022_weak force field, this chapter aims to provide the first application of this force field to systems of interest. This article took place in the new trends of ReaxFF molecular dynamics simulations by investigating pyrolysis degradation. Indeed, to our knowledge, the usage of an aqueous reactive force field was never used to investigate pyrolysis simulations. That is why we decided to compare the efficiency of two ReaxFF force fields coming from the combustion against the new developed force field coming from the aqueous branch. Moreover, we discuss the necessity to implement the simulations by averaging the results of the pyrolysis over a set of reactive simulations. This work was done in collaboration between the two authors, and the first author was responsible for producing all the results and their analyses. The development of the computational details was done through two internships supervised by the first author.

Hence, three different plastics: PE, PP, and PS, in the presence of additives such as toluene and chlorobenzene, have been studied through reactive molecular dynamics.

In the article, an introduction is presented introducing the energetic and environmental contexts of plastics and why pyrolysis can help their recycling. Then, the computational details present a quick review of ReaxFF and the simulation details. The first results concern thermogravimetric analysis from the three plastics pyrolysis. We compare literature results with reactive simulations results obtained with the ReaxFF potentials from the combustion and the aqueous branches. Then, the different products of pyrolysis are investigated. The next results are the role of the two additives, toluene and chlorobenzene, using one combustion force field and the new aqueous we developed during this thesis and presented in chapter III. Finally, we analyzed the pyrolysis of chlorobenzene, and the pyrolysis products were compared to experimental results.

4.2 Article 2 : Reactive simulations of plastics pyrolysis with additives: a comparison of the aqueous and combustion branches of ReaxFF

Reactive simulations of plastics pyrolysis with additives: a comparison of the aqueous and combustion branches of ReaxFF

Matthieu Wolf, Germain Salvato Vallverdu

germain.vallverdu@univ-pau.fr

Universite de Pau et des Pays de l'Adour, E2S UPPA, CNRS, IPREM, UMR 5254, Pau France

Abstract

In this work, the efficiency of the combustion and aqueous branches of the reactive potential ReaxFF are compared through pyrolysis investigations from reactive molecular simulations of three polymers: polyethylene (PE), polystyrene (PS) and polypropylene (PP) with additive compounds. Thermogravimetric analyses were implemented, computing the weight loss percentage along with analyzes of the pyrolysis products. Three force fields were considered: the CHON-2019 and the CHOSFClN potentials coming from the combustion branch and the CHONCl-2022_weak potential coming from the aqueous branch. The results show that both CHON-2019 and CHONCl-2022_weak force fields give similar results in adequation with the experimental results and are better than the CHOSFClN force field from the combustion branch. They can correctly predict the kinetic order of polymer degradation and identify the species formed and their quantities. As a result, the latest versions of the ReaxFF potential, belonging to the aqueous branch, are well suited to investigate the pyrolysis processes and may give equal or better results than force fields from the combustion branch. Furthermore, force fields from the aqueous branch seem to be more reliable for the investigation of the effect of the additives' role on plastic pyrolysis. Finally, an important aspect presented in this article is the use of statistical simulations for the different investigations. An error of 24 K was computed on ten thermogravimetric analyses simulated with ReaxFF on the PP pyrolysis, emphasizing the necessity of statistical simulations.

4.2.1 Introduction

In the field of computational chemistry, quantum mechanics (QM) methodologies are the most accurate for investigating chemical systems and reactivity as they include an explicit description of the electronic density. However, the computational cost of fully dynamic simulations with those methodologies is often limiting for large systems. That is why classical molecular mechanics approaches are used to investigate the dynamics of complex systems like polymers. These approaches are well-suited for non-reactive simulations. Nevertheless, they require predefined connectivity of the atoms and the definition of molecules, which prevents reactions from occurring. When it comes to looking at the reactivity of complex systems, it is thus necessary to have a methodology able to describe the breaking and the formation of bonds. In this context, reactive potentials such as ReaxFF[5] which are based on the bond-order formalism bridge the gap between molecular mechanics and quantum chemistry, allowing the investigation of the reactivity of complex systems.

Since the first publication in 2001 by van Duin et al.[5], ReaxFF has proven its efficiency with numerous studies of complex systems, including a large chemical diversity[58]. Initially, it was designed to describe hydrocarbon compounds focusing on combustion reactions. The potential took its final formalism from the work of Strachan et al.[55] in 2005, leading to a complete C/H/O force field by Chenoweth et al.[249] designed to investigate combustion reactions, named hereafter CHO_comb_2008. This potential was well suited to investigate high-temperature reactions. Many variations and improvements of CHO_comb_2008 appeared[250,62], which are parts of a branch of ReaxFF potentials entitled "combustion branch". The ReaxFF formalism does not use atom types but has only one set of parameters describing an atom regardless of its chemical environment. CHO_comb_2008 describes well water in gas phases but not in the liquid phase. This led to the development in 2018 of an aqueous branch of the ReaxFF potentials, including a good description of weak interaction in the liquid phase[68]. These two main branches evolved separately with different goals[58], leading to a low transferability of the parameters inside each branch but a very weak transferability from one branch to another.

Starting from combustion reactions, reactive molecular simulations were widely applied to pyrolysis, with more than 180 articles published in the last five years. More specifically, numerous papers address the question of the pyrolysis of plastic compounds. In 2019, global plastics production reached almost 370 million tons, and predictions assume that the production might reach 12000 million tons in 2050[251]. However, in 2017, only 9% of the total production of plastics since 1950 was recycled[251]. It is leading to more than 79% plastic waste accumulated in landfills and different environments[252,253,254]. This significant accumulation of plastic waste in landfills is of great concern for the surrounding ecosystem and human health[254]. There are principally six kinds of typical plastic components in municipal solid wastes, which are polyethylene high and low density(HDPE and LDPE), polypropylene (PP), polystyrene (PS), polyvinyl chloride (PVC) and

polyethylene terephthalate (PET)[255]. Mechanical recycling, energy recovery or thermal recycling and chemical recycling or feedstock recycling are the three main methods of recycling tools of plastic wastes[256]. Due to the inhomogeneity of polymers, mechanical recycling is not an ideal choice for recycling plastics[257], but it is a better solution than thermal recycling because incineration releases undesirable pollutants into the atmosphere[256,251,258]. Many articles, highlighted that the most suitable approach for plastics wastes recycling is thus the chemical recycling from pyrolysis and gasification[259,260,261,262]. Indeed, in addition to the protection of the environment, chemical recycling can also lead to sustainable use of plastic waste by producing fuel[263,259,264,265]. The different pyrolysis mechanisms of the principal plastic wastes were investigated through experimental study[259,266,267,268,269] and more recently from computational methods[270,271,272,273] with some of them using ReaxFF potentials.

This paper addresses two main objectives. First of all, because of the development since 2019 of the new aqueous branch of the family of the ReaxFF potential, dedicated to the investigation of the reactivity in the condensed phase, we aim to compare the quality of reactive potentials from the combustion and the aqueous branches to describe pyrolysis. To that end, we decided to investigate the pyrolysis of different major plastic compounds (PE, PP and PS) and of the chlorobenzene molecule, which is well described in the literature[274]. Second, plastic generally includes molecular additives to fit the required properties for certain applications. These additives may lead to the formation of new pyrolysis products or alter the behaviour of the pure plastic phase along the pyrolysis process. In this work, The effect of common additives such as toluene and chlorobenzene was investigated through reactive molecular simulations. These two molecules are commonly used solvents during plastic synthesis or can be used as flame retardant[275,276].

4.2.2 Computational methods

4.2.2.1 ReaxFF: A reactive force field

The ReaxFF potential is based on two mains concept: (i) it uses the bond order (BO) paradigm to describe all the different intramolecular energetic contributions of the total energy[249]; (ii) there is only one unique atom type per chemical element. ReaxFF allows bonds to break and form along the simulations by including the BO in the potential formalism. The potential is defined through multiple different empirical parameters. The description of the ReaxFF potential energy terms, which compose the total energy, is presented in the equation 4.1.

$$E_{system} = \underline{E_{bond}} + \underline{E_{lp}} + \underline{E_{over}} + \underline{E_{under}} + \underline{E_{val}} + \underline{E_{pen}} + \underline{E_{coa}} + \underline{E_{C2}} + \underline{E_{triple}} + \underline{E_{tors}} + \underline{E_{conj}} + \underline{E_{Hbond}} + E_{vdW} + E_{Coulomb} \quad (4.1)$$

The system's total energy with the ReaxFF potential is divided into fourteen different contributions. All the underlined terms depend on the BO on the equation 1.14. All energy terms can be divided into two groups. The first one corresponds to intramolecular contributions and is composed of the first eleven terms, where the common factor is the BO dependency. In the order of the equation 1.14 we have the bond energy term, the lone pair term, both over coordination and undercoordination corrections, the valence angle contribution, another correction of double bonds angle valency issues, correction for NO_2 group stability, correction to destabilize triple bonds for allene, a particular correction for carbon monoxide, torsions and conjugation energies. The second group is composed of the last three terms in equation 1.14 and corresponds to the intermolecular interactions. This group includes a hydrogen bond term contribution depending on the BO, then Van der Waals interactions and Coulombic interactions. It is behind the scope of the article to present all the different equations corresponding to the different energetic contributions. However, we invite the reader to read the work of Chenoweth et al.[249], which proposes a complete description of the ReaxFF potential.

4.2.2.2 Computational details

Three polymers were investigated: polyethylene (PE), polypropylene (PP) and polystyrene (PS). For each polymer, reactive molecular dynamics simulations corresponding to the pyrolysis of the pure polymers and the polymers with additives were implemented. The pyrolysis of a pure chlorobenzene phase from reactive molecular dynamics was also investigated and compared to experimental data. In order to compare the aqueous and the combustion branches of the ReaxFF potential several potentials were considered, including the C/H/O/N elements and two of them with the Cl element. From the combustion branch the CHOSFClN potential, from the work of Wood et al.[60] and the CHON-2019 potential from the work of Kowalik et al.[62] were implemented. From the aqueous branch, the CHONCl-2022_weak potential from the work of Wolf et al.[277] was implemented as it includes parameters for the chloride atoms from the last ReaxFF potential in the aqueous branch CHON-2017_weak from Zhang and van Duin[68].

In order to investigate the pyrolysis from reactive molecular simulations, three different boxes were implemented. First, a box filled up with 540 chlorobenzene molecules was built to evaluate the quality of the force fields from the combustion or the aqueous branches on the pyrolysis of the chlorobenzene against experimental data. Second, one box composed of ten pure polymers with fifty monomers was built for each polymer (PE, PP and PS). Finally, a box with again ten polymers of fifty monomers with one additional toluene or chlorobenzene molecule was built for each polymer (PE, PP and PS). To obtain reliable results, averages computed from molecular simulations have to be independent of the initial conditions and, in particular, independent from the initial configuration. Considering only ten polymers chain or one additive molecule in the box leads to poor statistics. Consequently, ten boxes were built for each system corresponding to one polymer (PE, PP

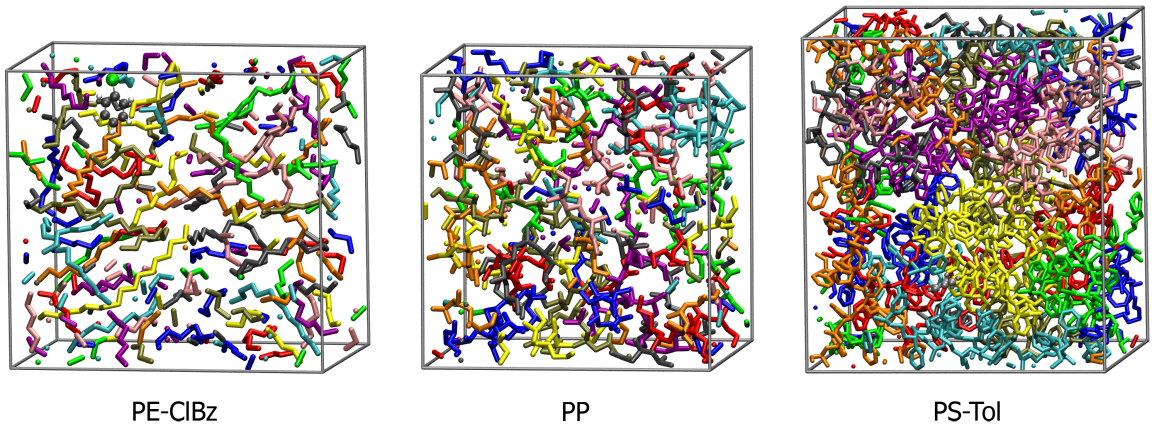


Figure 4.1: Three boxes of molecular dynamics simulation starting points. The first is a PE box with one chlorobenzene molecule, the second a PP box, and the final a box of PS with one toluene molecule. On each box, the different polymers are coloured.

and PS) and one additive (no-additive, toluene or chlorobenzene), and reactive molecular simulations corresponding to the pyrolysis were implemented on each of these ten boxes. The results and conclusions about each system were then produced based on the average over the ten simulations. To our knowledge, such an approach was never implemented in previous reactive molecular simulations of pyrolysis processes and will increase the quality and reliability of the results.

To create a realistic box of polymers, several simulations with the UFF classical force field[23] were implemented using the Force Field module of the AMS[90] software. We followed a similar methodology as previous works on polymer pyrolysis[271]. Starting from a large box with a cell size of 400Å, the ten polymers with fifty monomers were put randomly in the box. In the case of systems including an additive molecule, it was also added in the box randomly. The systems were then compressed using an NPT simulation implementing the Berendsen thermostat and barostat[278] at 300 K and 1 GPa, with a time step of 0.25 fs, during 50000 steps. The systems reaches an equilibrium density quite rapidly, and ten independent configurations were extracted from the simulations for each system. For each of these ten boxes and each system, they were homogenized by doing a second simulation step of 100 ps with a timestep of 1 fs in the NVT ensemble at 1000 K. The systems were then cooled at a rate of 1 K/ps from 1000 K to 300 K at 1 bar in the NPT ensemble. Finally, the last simulation step in the NPT ensemble at 300 K and 1 bar during 100 ps was implemented for all systems to reach an equilibrium at ambient temperature. These final configurations constitute the starting point of the reactive molecular dynamics simulations implemented to produce the pyrolysis process, and examples of three different polymer boxes are presented in figure 4.1.

For each system, considering the ten initial configurations, the pyrolysis process was simulated using the different reactive force fields. These simulations were implemented using the LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator)[89] software, with the version October 2021 and default cutoff

parameters, namely, a nbrhood_cutoff of 5 Å and a hbond_cutoff of 7.5 Å. With a temperature of 500 K, each system was heated progressively in the NVT ensemble, using the Berendsen thermostat, until it reaches a temperature of 3200 K for boxes with polymers and 2500 K in the case of chlorobenzene. A ramp of 10 K/ps was implemented with a timestep of 0.1 fs which ensure, in particular at high temperature, accurate integration of the equations of motion. The final temperature is far above common pyrolysis processes' experimental temperature. Experimentally, the time length of pyrolysis is in order of a few seconds, while the length of the reactive molecular dynamics simulations implemented here is a few hundred picoseconds. Using such a high temperature is thus a way to ensure that chemical reactions will occur during a short simulation time. This strategy is usual in ReaxFF simulations, and previous works already showed that although the temperature is high, the simulations capture the mechanism of the experimental pyrolysis and, in particular, reproduce the successive steps of the pyrolysis[271,270,279,273]. In this article, when comparing the experimental and computational results at a given temperature, the temperature from experimental results will be named the experimental temperature. In contrast, the temperature from reactive simulations will be named the ReaxFF temperature.

4.2.3 Results and discussions

4.2.3.1 Statistical average over a set of simulations

In the literature, ReaxFF simulations are commonly used to reproduce pyrolysis processes heating the system progressively using a temperature ramp. When considering polymers such as PE, PP and PS in this study, one of the first degradation steps will be the rupture of the polymer chain. For an obvious computational cost reduction, the number of polymer chains and the number of monomers is limited. Depending on the initial configuration and velocities, this first rupture of the polymer chain may occur in a wide temperature range and lead to sharp variations of the quantities followed along the pyrolysis process. For example, following the number of species in the simulation box, if one considers ten polymer chains when one chain is broken, this represents 10% of the total amount of polymer. In order to avoid these sharp variations, people usually include smoothing of the data. In this first subsection, we will highlight the results obtained after averaging over ten reactive simulations. We compare the thermogravimetric analysis (TG) computed as the weight loss percentage. To obtain the TG from our ReaxFF simulations, we followed the methodology already used in previous works[280,281,271] where the pyrolysis products with more than 40 carbons are classified into char, and then we followed the evolution of the char mass using the equation 4.2.

$$\text{Char mass} = \frac{\sum^{i} M_i \times n_i}{M_{total}} \tag{4.2}$$

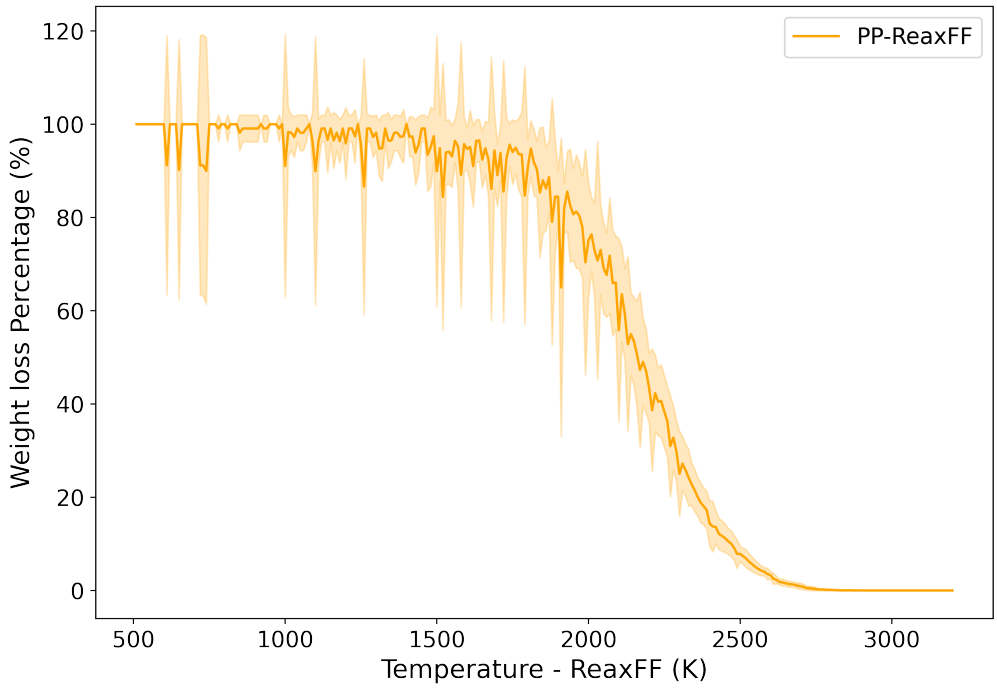


Figure 4.2: Thermogravimetric analyses obtained from 10 reactive simulations of PP pyrolysis. The dark yellow curve is the average over the 10 TG curves and the light envelope correspond to the standard deviation with a width of 2σ.

The char mass is obtained by computing the sum of each specie produced i of the molar mass M_i times the number of molecule specie n_i divided by the total molar mass M_{total} at the beginning of the simulation.

Figure 4.2 presents the average TG curve obtained over all the ten reactive simulations of PP pyrolysis and the envelope corresponds to the standard deviation (with a width of 2σ). The temperature for which the weight loss percentage reach 50% range from 2090 K to 2230 K, the average value is 2175 K with a standard deviation of 42 K. Assuming a rectangular error distribution (worst case), the associated uncertainties on this temperature is 24 K. Such analyses give two indications. First, it gives an estimation of the expected error on the TG curves computed hereafter. Second, thanks to the calculation of the dispersion from the standard deviation gives an overview of the envelope containing the expected TG curves. When two TG curves obtained from different conditions are closed, this will help to decide if the differences are relevant or not.

All the results presented hereafter in this paper are obtained from an average over 10 independent simulations. This was done both for the TG curve analyzes and the pyrolysis products.

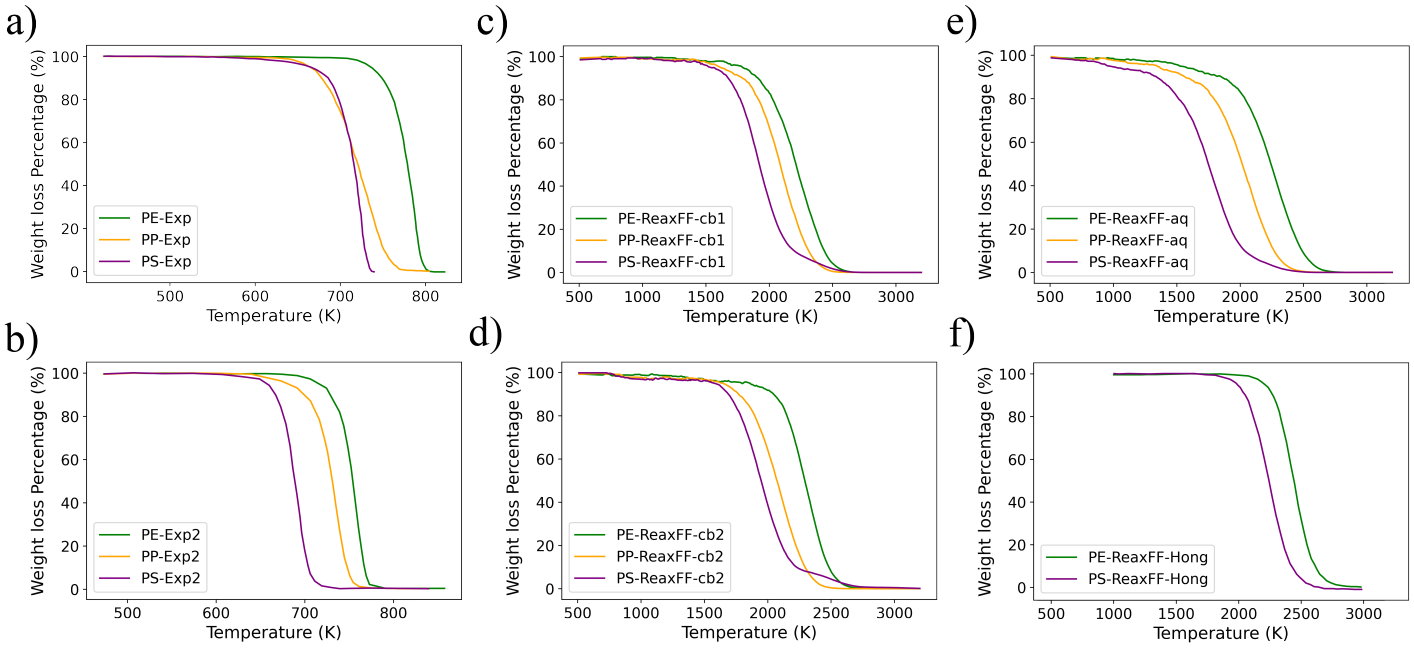


Figure 4.3: Thermogravimetric analyses obtained from experimental studies and ReaxFF simulations for pure PP, PE and PS polymers. a) and b) are experimental work of Abbas-Abadi[269] and Dümichen et al.[253] ; ReaxFF simulations of this work with c) CHOSFClN potential[60] ; d) CHON-2019 potential[62] ; e) CHONCl-2022_weak[277]. F) ReaxFF simulations from the work of Hong et al.[271]. c), d) and f) correspond to potential from the combustion branch and e) from the aqueous branch. Purple, green and orange curves correspond to the TG of PS, PE and PP respectively.

4.2.3.2 Thermogravimetric Analysis from plastic pyrolysis

This second subsection compares the TG analysis obtained from experimental studies against ReaxFF simulations using potentials from the combustion and the aqueous branches. The results are gathered in figure 4.3 for all polymers in the pure phase case from experimental or computational studies. The experimental results come from the literature and the work of Abbas-Abadi[269], figure 4.3 a), and Dümichen et al.[253], figure 4.3 b). The results from ReaxFF simulations using potentials of the combustion branch are presented figure 4.3 c), for the CHOSFClN force field[60], and figure 4.3 d), for the CHON-2019 force field[62]. Figure 4.3 e) presents the results obtained using the force field CHONCl-2022_weak[277] from the aqueous branch. The results in figure 4.3 f) come from the work of Hong et al.[271] implementing the initial C/H/O force field of the combustion branch and are presented for comparison. The weight loss percentage represented in the TG plots in figure 4.3 is defined as the proportion of molecules in the simulation box belonging to the char. This quantity is computed throughout the simulation during the heating procedure and plotted as a function of the temperature.

From figure 4.3, the two experimental investigations both obtained similar behaviour for the pyrolysis of the PE and the PS, but there is more uncertainty on the degradation of the PP. Such a shift in the temperature at which the degradation of PP occurs may come from the polymer density or the experimental setup. A previous work by Matsuzawa et al.[282], enforce the results of Dümichen et al.[253], figure 4.3 b), corresponding to

a temperature of degradation of PP closer to PE. Experimentally, the degradation behaviour of these polymers is thus the following: the degradation of PS occurs at a lower temperature than PP, and the degradation of PP occurs at a lower temperature than PE.

Focusing now on the results obtained from ReaxFF simulations, figure 4.3 c, d, e and f, the degradation behaviour of the three polymers agrees with the experimental results. The temperature at which the polymer starts to decompose increases in the following order: PS, PP and then PE. However, the temperature scale is not the same, and the temperature used in the simulation is higher. Such a high temperature is required when pyrolysis simulations are performed with ReaxFF to reduce the simulation time and permit the decomposition to occur[279].

Comparing now the results obtained with the combustion branch's potentials and the aqueous branch's potential, we also obtained the same behaviour and still the same order of the degradation temperature for the three polymers, whatever the ReaxFF branch of the potential. Looking closely at the temperature, using the potentials of the combustion branch, the degradation of PE starts at 1960 K for CHOSFClN, at 2100 K for CHON-2019 and 2250 K with the C/H/O potentials used by Hong et al.[271]. Using the CHONCl-2022_weak[277] potential of the aqueous branch, the degradation of PE starts at 1950 K, which is the lowest temperature. Based on these results, it is impossible to conclude about the best potential to investigate such pyrolysis because all potentials provide results that agree with experiments. Nevertheless, one can see a trend that the degradation temperature obtained from the aqueous force field is lower than the one obtained from combustion force fields.

The TG analysis provides an overview of the macroscopic behaviour of pyrolysis. In the following, in order to refine the comparison between aqueous and combustion potentials, we focus on the pyrolysis products predicted by the reactive simulations.

4.2.3.3 Pyrolysis products analyses

In this subsection, we compare the species produced from the reactive simulations of the pyrolysis process of the three polymers PE, PP and PS with experimental results. We used the same potentials as in the previous subsection from either the combustion or the aqueous branches. The TG analyses presented in figure 4.3 were used to scale the temperature and compare the pyrolysis products obtained from the simulations and those obtained experimentally at different temperatures Hong et al.[279].

PE pyrolysis products

Figure 4.4 presents the pyrolysis products of PE obtained from the experimental work of Honus et al.[268] and ReaxFF simulations with the two potentials from the combustion branch (CHOSFClN and CHON-2019)

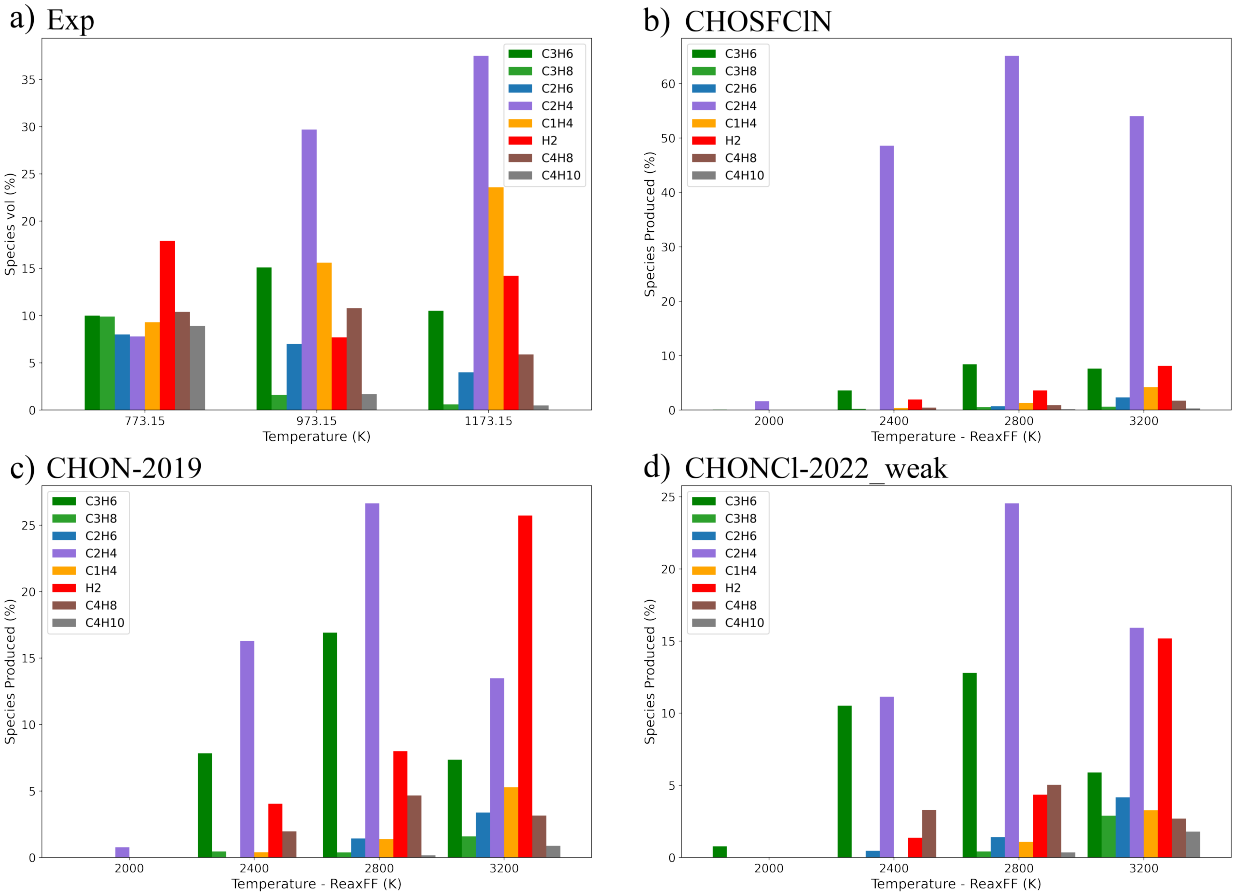


Figure 4.4: Amount of species produced from the pyrolysis of PE as a function of the temperature. a) experimental results from Honus et al.[268]. The amount of species is reported in the gas phase as the volumetric percentage. b) c) and d) present the reactive simulation results using the CHOSFClN, the CHON-2019 and the CHONCl-2022_weak force fields, respectively. The amount of each species is the percentage over the total number of molecules. The same species are present on all graphics.

and the potential from the aqueous branch (CHONCl-2022_weak). The amount of each species is presented at several temperatures along the pyrolysis process. Experimentally the species are presented at 773 K, 973 K and 1173 K, which correspond to the beginning of the PE degradation, almost the end of the degradation and a temperature above the end of the degradation of the polymer where only pyrolysis products exist, respectively. Following the same steps on the TG analysis of the ReaxFF simulations, the species obtained along the pyrolysis process were analysed at 2000 K, 2400 K, 2800 K and 3200 K corresponding to the beginning of the PE degradation, a weight loss percentage of 50%, the end of the degradation and the end of the pyrolysis simulation, respectively.

First, all simulations agree with the experimental results, the main pyrolysis products being C_2H_4, the initial monomer, followed by C_2H_6. All species identified experimentally were also present as pyrolysis products at the end of ReaxFF simulations. However, the amount of the H_2 molecule seems to be overestimated from the simulations. However, experimentally, the species abundances are measured as a volumetric percentage, while from the simulations, the abundances are computed as a number of species. This may explain the vari-

ation of the relative abundances between experimental and simulation results, in particular in the case of the H_2 molecule, which is the smallest one.

Comparing the reliability of the force fields considered in the reactive simulations, one can see in the figure 4.4 b) that at 2800 K and 3200 K, the CHOSFClN force field highly overestimates the production of C_2H_4 which corresponds to 60% of all the species produced. On the contrary, looking at figure 4.4 c) and d), the relative abundances of the species obtained from the simulations with the CHON-2019 and CHONCl-2022_weak force fields at 2800 K agree with the experimental results of Honus et al.[268] at 973 K except for methane, figure 4.4 a). These results also agree with the experimental results of Williams and Williams[255], who experimentally identified the gas compositions product formed from the pyrolysis of different plastics (weight percentage) presented in the table 4.1 in the supporting information, and in the case of the PE, they obtained, in terms of abundance, almost the same as the two results obtained with CHON-2019 and CHONCl-2022_weak. At the end of the pyrolysis process, at 3200 K in reactive simulations and 1173 K experimentally, there is a good agreement concerning the list of species and their relative abundance.

Both combustion and aqueous branches of ReaxFF potential, associated with the CHON2019 and CHONCl-2022_weak force fields, respectively, lead to a list of species and their relative abundance, which agrees with experimental results of the pyrolysis of PE. However, the CHOSFClN force field from the combustion branch largely overestimates the production of C_2H_4 at the expense of the other species.

PP pyrolysis products

The next polymer pyrolysis that is investigated is PP. The pyrolysis products obtained from experimental results and ReaxFF simulations are presented in figure 4.5 following the same approach as the previous subsection. The comparison of the results at a given experimental temperature was made on an equivalent temperature in reactive simulations corresponding to the same step on the TG analyses, figure 4.3. From experimental results, 773 K corresponds to the end of the polymer degradation, and then 973 K and 1173 K define several states post-polymer degradation. In reactive simulations, a temperature of 2400 K corresponds to the beginning of the polymer degradation. Then temperatures of 2800 K and 3200 K in the reactive simulations correspond to different states after the complete degradation of the polymer, and they can be linked to the states of the system just above 973 K and 1173 K, respectively, on the experimental results. Thus, the different experimental states at 773 K, 973 K and 1175 K are compared here with the state at 2400K, 2800K and 3200K, respectively, in the ReaxFF simulations.

The first point that can be noticed from the reactive simulation results is that they are all consistent with each other and in agreement with the experimental results, figure 4.5 a). They also agree with the work of Williams and Williams[255] presented in the table 4.1 in the SI. The species which is obtained with the largest abundance is C_3H_6. Moreover, looking at the evolution of the abundance over the temperature, all reactive

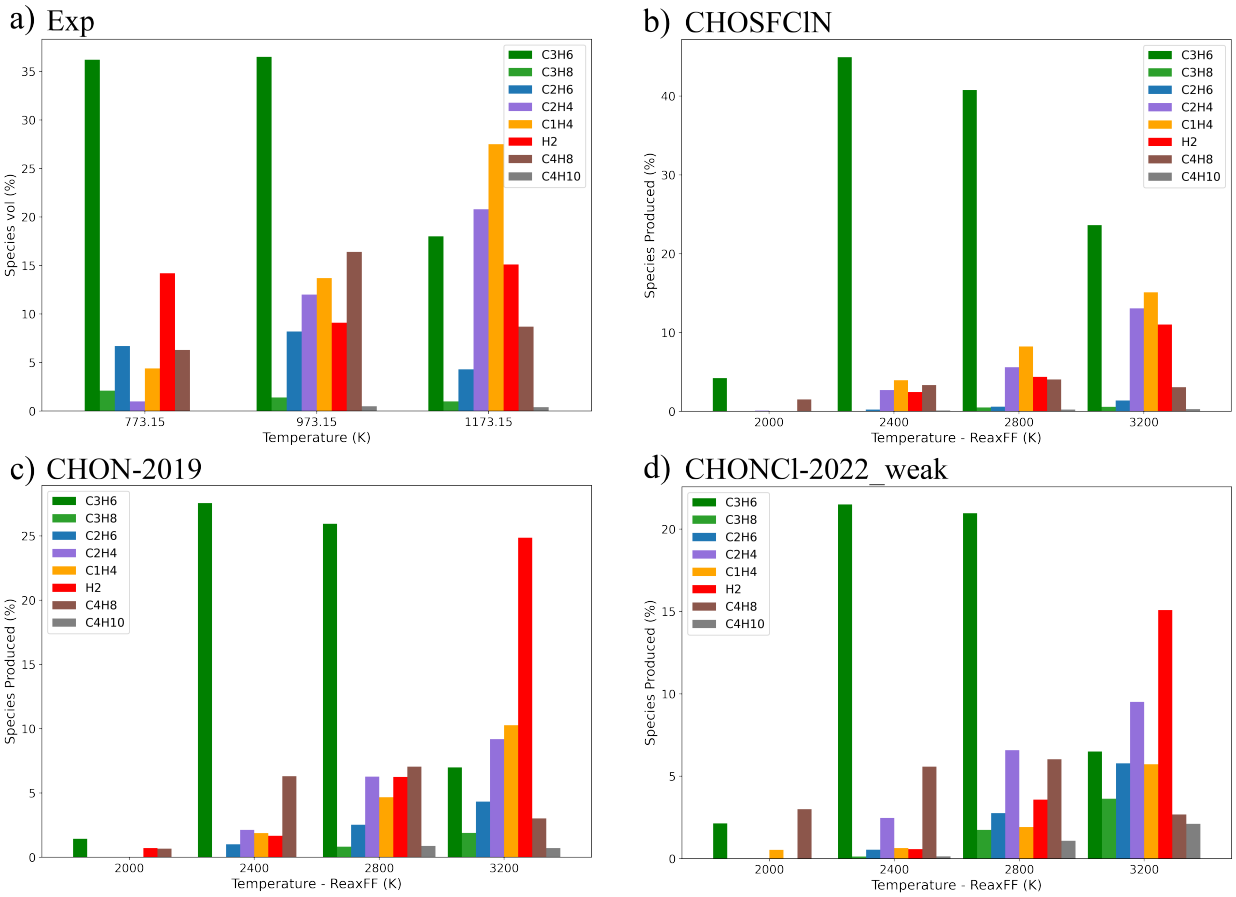


Figure 4.5: Amount of species produced from the pyrolysis of PP as a function of the temperature. a) experimental results from Honus et al.[268]. The amount of species is reported in the gas phase as the volumetric percentage. b) c) and d) present the reactive simulation results using the CHOSFClN, the CHON-2019 and the CHONCl-2022_weak force fields, respectively. The amount of each species is the percentage over the total number of molecules. The same species are present on all graphics.

simulation results agree with the experimental observations that the volumetric species percentage of C_2H_6 decreases when the pyrolysis continues.

For the other species, at 773 K experimentally and 2400K theoretically, the three force fields show the formation of all the different species detected experimentally but with few variations concerning the relative abundance, which could be attributed to the low abundance of these species. At 2800K and 3200 K, the results from ReaxFF simulations with the CHON2019 and the CHONCl-2022_weak force fields are in good agreement with the results of Williams and Williams [255] and Honus et al. [268] keeping almost the same species predominance. Concerning the results obtained with the CHOSFClN force field, they are consistent with the experimental results except for the abundance of C_4H_8, which is underestimated.

Finally, the pyrolysis products of PP from reactive molecular simulations agree with the experimental results along the pyrolysis process and after the complete degradation of the polymer.

PS pyrolysis products

In this last subsection of pyrolysis products analysis, we compare the species produced by the ReaxFF simulations of PS with the experimental results of Maafa [283]. The results are gathered in figure 4.6. Looking at figure 4.3 a) and b), the temperature of the beginning of degradation of PS is about 650 K and finishes at 750 K. From the experimental work of Maafa [283], the pyrolysis products of PS are available at 748 K, which thus corresponds to the end of the degradation. This step of the pyrolysis process will be compared with the system's state at 2400 K in the reactive molecular simulations. The pyrolysis products will then be compared after the complete degradation of the polymer, at 823 K on the experimental results and 2800 K from the reactive simulation results.

In figure 4.6, one can notice that the pyrolysis products are largely dominated by C_8H_8 corresponding to the initial styrene monomer. From figure 4.6 a), experimentally, the amount of styrene molecules slightly decreases at the end of the pyrolysis process. This decrease is obtained from reactive molecular simulations at 2800 K and agrees with the experimental results. Increasing the temperature in the reactive simulations again, at 3200 K, the end of the pyrolysis leads to the degradation of the styrene molecule itself in the case of the CHON-2019 and the CHONCl-2022_weak force fields, figure 4.6 b) and c), respectively. In the case of the CHOSFClN force field, a slight decrease is obtained at high temperatures.

Experimentally, figure 4.6 a), the minor pyrolysis products of PS are CH_4, C_6H_6 and C_7H_8. From reactive molecular simulations, these products are obtained only in the case of the CHON-2019 and the CHONCl-2022_weak force fields. C_6H_6 (benzene) and C_7H_8 (toluene) are obtained at 2800 K, but other species are only obtained at a higher temperature.

Finally, the three force fields show that at a higher temperature the degradation of styrene leads to the formation of lower-weight molecular compounds and releases a large quantity of H_2, it is at a higher temper-

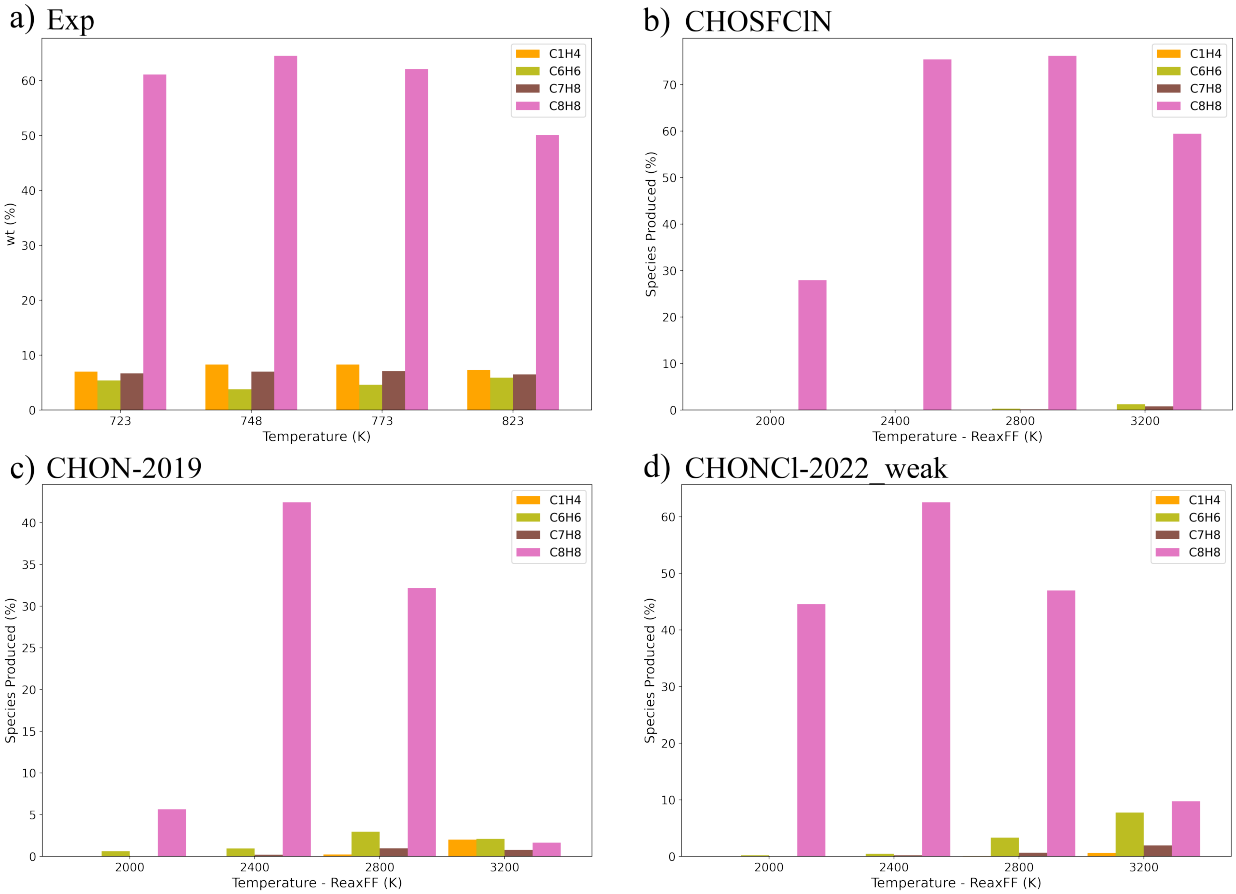


Figure 4.6: Amount of species produced from the pyrolysis of PS as a function of the temperature. a) experimental results from Maafa [283]. The amount of species is reported as the mass fraction in percentage. b) c) and d) present the reactive simulation results using the CHOSFClN, the CHON-2019 and the CHONCl-2022_weak force fields, respectively. The amount of each species is the percentage over the total number of molecules. The same species are present on all graphics.

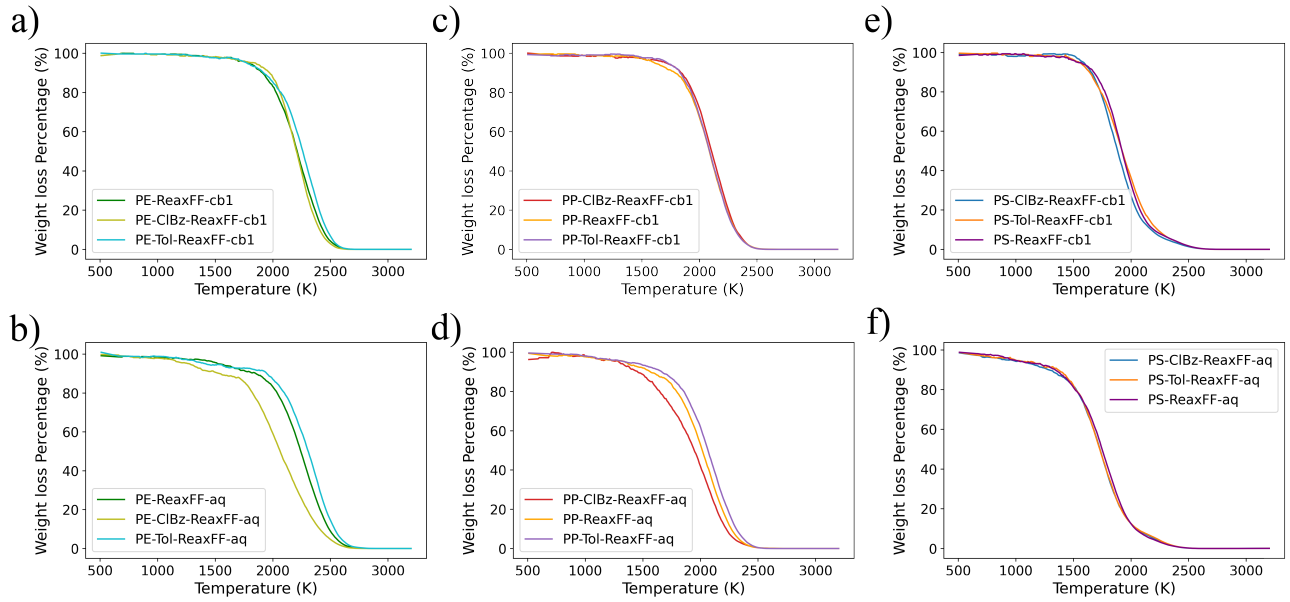


Figure 4.7: Thermogravimetric analyses obtained from ReaxFF simulations for PP, PE and PS polymers with no additive, one chlorobenzene (ClBz) molecule or toluene (tol) molecule. cb1, stands for combustion branch and corresponds to the CHOSFClN force field. aq, stands for aqueous branch and corresponds to the CHONCl-2022_weak force field.

ature than the two other force fields for the CHOSFClN. These results agree with the experimental results of Honus et al.[268], who measures at 1173 K a significant volumetric percentage of H_2.

The three force fields are able to reproduce from reactive molecular simulations the pyrolysis products of PS and the degradation products of styrene at high temperatures. However, CHON2019, from the combustion branch and the CHONCl-2022_weak force field, from the aqueous branch, are more efficient than CHOSFClN in matching experimental results.

4.2.3.4 Effect of additive during pyrolysis of plastic

This subsection presents an investigation of the potential effect of small quantities of additives on the behaviour of three plastics from reactive molecular simulations of the pyrolysis process. We investigate the presence of one molecule of chlorobenzene or one molecule of toluene for each of the three polymers: PE, PP and PS. These two molecules are used as solvents during the synthesis of these plastics[275,276] and might be found in small quantities inside. The chlorine atom is not parameterized in the CHON2019 force field, we only used the CHOSFClN and the CHONCl-2022_weak force fields in this part, which belong to the combustion and the aqueous branches, respectively. Figure 4.7 presents all the different thermogravimetric analyses obtained from reactive molecular simulations with the two reactive force fields in the presence of one chlorobenzene molecule or one toluene molecule.

First, we look at the results obtained with the CHOSFClN force field from the combustion branch, first line of figure 4.7. The presence of the additive molecules does not seem to impact the degradation of the

polymers. Considering now the results obtained with the CHONCl-2022_weak force field from the aqueous branch, in the case of PS, again, the additive molecules seem to have no impact on the pyrolysis process. In the case of PE and PP and the reactive simulations with the CHONCl-2022_weak force field, figure 4.7 b) and d), one can see that the toluene molecule tends to slow down the degradation process of the polymers while the chlorobenzene seems to speed up the degradation process.

Because of the small deviation of the TG curves associated to the presence of an additive molecule or not, the statistical averages performed over ten simulations for each system is of high importance. Figure 4.10 in the SI presents the average value of the TG curves surrounded by the standard deviation as a function of the temperature for PE and PP, with or without an additive molecule and with the CHONCl-2022_weak force field. The envelope representing the standard deviation gives an insight of the uncertainty of the TG curves. In the case of PP, one can see that the standard deviation of the three TG curves overlap. No conclusion can thus be done about the role of these additive molecules on the PP temperature degradation. In the case of PE, following the same arguments, the effect of the toluene additive is not relevant. However, in the case of a chlorobenzene additive, the envelope of the TG curve do not overlap with the TG curve without additive. The effect of chlorobenzene is thus relevant and tends to accelerate the degradation of PE.

The pyrolysis products obtained from the reactive molecular simulations were analyzed in the case of inserting a chlorobenzene molecule in the simulation boxes. Figure 4.11 in the supporting information presents the evolution of the chlorobenzene in the ten simulations of the PE with the additive using the CHONCl-2022_weak force field. One can see that in the case of the force field CHONCl-2022_weak, the chlorine atom of the chlorobenzene leads to the formation of an HCl molecule. However, for the combustion CHOSFClN force field, from the ten simulations, different chlorinated products are obtained as chloromethane, HCl, chloroethene and chlorobenzene. Moreover, following the percentage of the C_2H_4 molecules along the pyrolysis in the simulation with CHONCl-2022_weak force field, its formation is slightly accelerated with 13% per cent at 2400K, against 11% without chlorobenzene as presented in figure 4.4 d), which seems to speed up the degradation.

Due to this difference in the results obtained from the two force fields of the two branches of ReaxFF, we go further in the investigation of the chlorobenzene molecule. In order to get a better overview of the accuracy of the two force fields the pyrolysis of the chlorobenzene molecule was investigated from reactive molecular simulations and compared to experimental results.

4.2.3.5 Pyrolysis of chlorobenzene

In this subsection, we compare the results of ReaxFF simulations and experimental observation of the chlorobenzene pyrolysis. The reactive molecular simulations were implemented using both the CHOSFClN

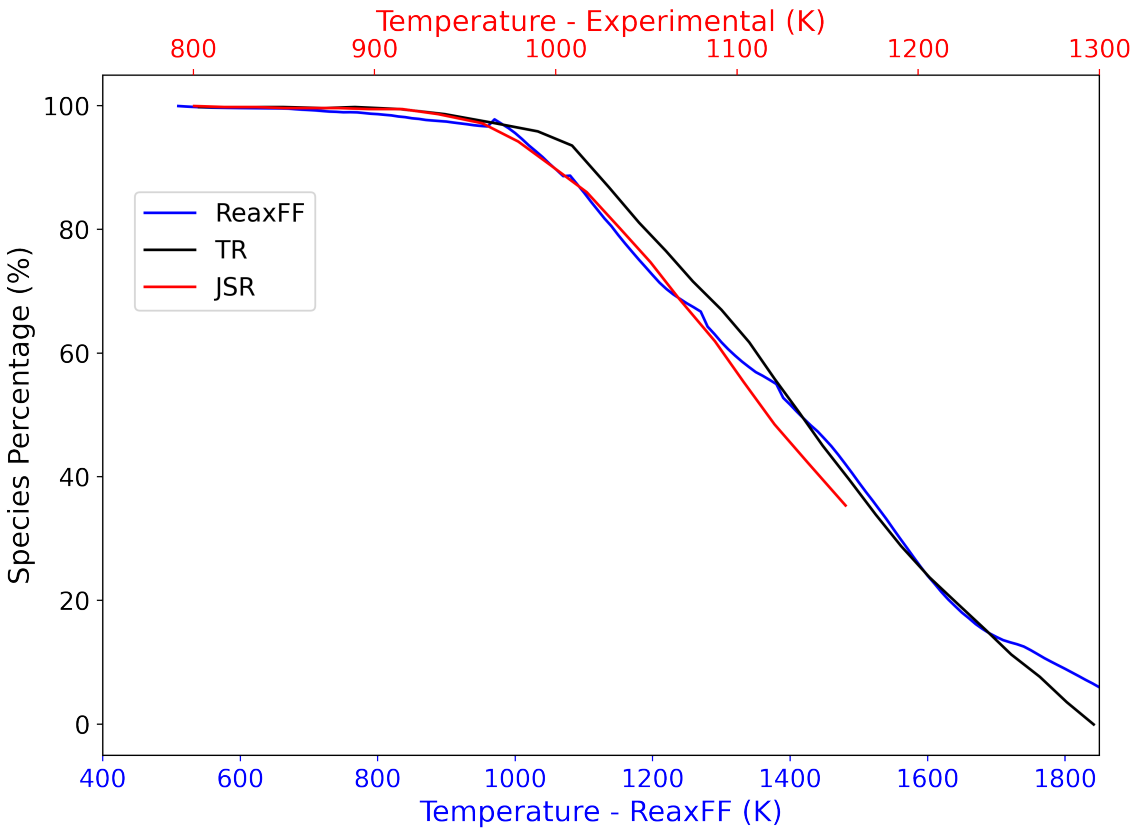


Figure 4.8: Percentage of the number of chlorobenzene molecules as a function of the temperature. In blue, and with the blue x-axis, from ReaxFF simulations with the CHONCl-2022_weak force field. In red and black, experimental results[274] from a jet-stirred-reactor (JSR) and a tubular reactor (TR), respectively. The two x-axis are scaled to match the experimental and simulation results.

force field from the combustion branch and the CHONCl-2022_weak force field from the aqueous branch. The percentage of chlorobenzene molecules that are converted during the pyrolysis was computed along with the evolution of the amount of the species produced from the pyrolysis.

First, the chlorobenzene molecules do not degrade during the reactive simulations when using the CHOSFClN force field. All chlorobenzene molecules were conserved throughout the simulations over the temperature range from 500 K to 2500 K. This reactive potential was successfully used to investigate the pyrolysis of PVC[279], but such polymer only contains an aliphatic chlorine group contrary to the aromatic character of chlorobenzene. The chlorobenzene is thus over-stabilized by the CHOSFClN force field, which might explain the different behaviour observed in the PE pyrolysis in the presence of a chlorobenzene molecule additive.

Figure 4.8 presents the conversion percentage of chlorobenzene from the experimental work of Vin et al.[274] and from the ReaxFF simulations with the CHONCl-2022_weak force field. Two kinds of reactors were used by Vin et al.[274]: a jet-stirred-reactor and a tubular reactor. The results from these two reactors will be labelled JSR and TR, respectively. One can see that the ReaxFF temperature was scaled to the same range as the experimental temperature to match the conversion curves. The experimental ramp between 750 K to 1300 K is thus

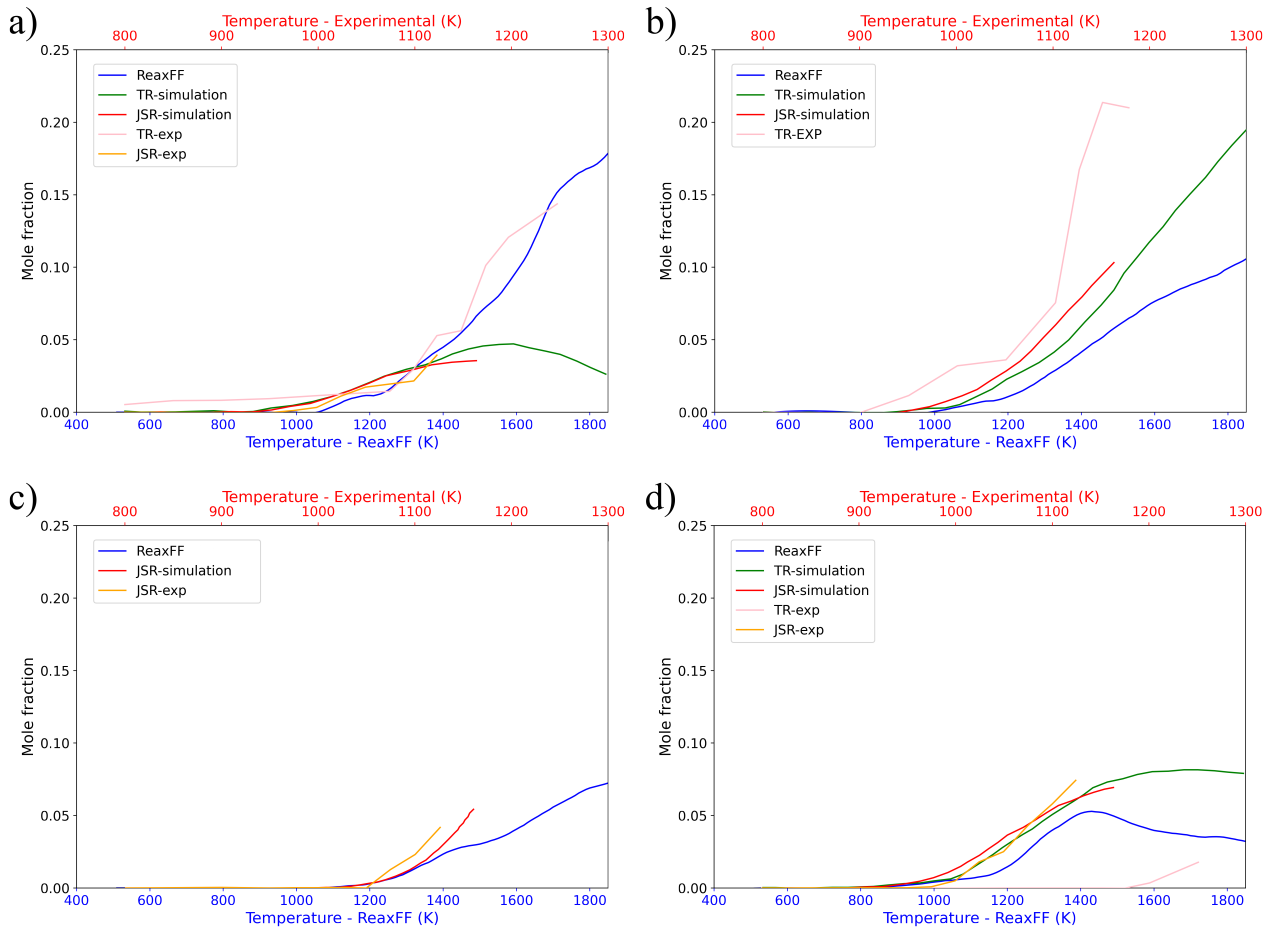


Figure 4.9: Mole fraction of a) benzene, b) HCl, c) biphenyl and d) PCB (PolyChlorinated Biphenyl) as a function of the temperature along the pyrolysis. Results from reactive simulations are in blue. Pink and Orange curves are experimental results from Vin et al.[274]. Red and Green curves are extrapolation from experimental investigation of Vin et al.[274]. On each plot, the two x-axis are scaled according to figure 4.8.

linked to the ramp from 400 K to 1825 K in ReaxFF simulations. Using the same temperature scales, figure 4.9 presents the amount of benzene, HCl, biphenyl, and Poly-Chlorinated Biphenyls (PCB) produced along the pyrolysis experimentally from the work of Vin et al.[274] and from the reactive molecular simulations.

A good agreement is obtained from the formation of benzene between ReaxFF simulations and TR or JCR experimental results, figure 4.9 a). The extrapolation of the TR experimental results[274] predicts a decrease in the amount of benzene molecule at a high temperature, which is not obtained from reactive simulations.

Figure 4.9 b) presents the evolution of the amount of HCl molecules during the pyrolysis. The experimental results came from the TR experiments only. ReaxFF simulations underestimated the formation of HCl, but the results are still consistent.

Figure 4.9 c) represents the formation of biphenyl molecules from reactive molecular simulations and JSR experiments. The ReaxFF simulation results agree with the JCR results. Until 1350 K (ReaXFF temperature), the simulation results are identical to the JSR experiments.

Figure 4.9 d) presents the evolution of all the different PCBs produced during the pyrolysis of chlorobenzene.
Until 1400 K (ReaxFF temperature), the ReaxFF simulation results are in good agreement with the JSR experiments and JSR simulations. At higher temperatures, above 1400 K, the mole fraction of PCB decreased in the ReaxFF simulation results, which was not observed experimentally when the amount of PCB reached a plateau. Nevertheless, it is observed from the TR simulation a decrease in the PCB mole fraction after 1250K in experimental temperature. Thus, ReaxFF might over-represent this behaviour of PCB degrading after a specific pyrolysis temperature might still occur.

Finally, the reactive molecular simulations produced with the CHONCl-2022_weak force field, from the aqueous branch, agree well with the behaviour of the chlorobenzene pyrolysis and reproduce both the conversion percentage of chlorobenzene and the amount of the main pyrolysis products.

4.2.4 Conclusion

This article focuses on the methodological aspect of producing reliable and accurate results on the investigation of pyrolysis processes from reactive molecular simulations. First, the efficiency of reactive force fields from different ReaxFF branches to reproduce the results of plastic pyrolysis is compared. Two force fields from the combustion branch were implemented, the CHOSFClN and the CHON-2019 force fields, along with the CHONCl-2022_weak force field from the aqueous branch. Secondly, we show how increasing the statistics over the results by producing a series of reactive molecular simulation improve confidence in the results. From the comparison with experimental results, we conclude that the three force fields are able to reproduce the kinetic order of degradation of the three polymers, PE, PP and PS. The analysis of the pyrolysis products

obtained from the reactive simulations, the CHON-2019 and the CHONCl-2022_weak force fields, give similar results in agreement with the experiments. The CHOSFClN force field gives coherent results but has a higher tendency to overestimate the formation of the main specie at the expense of the other species. Then we investigated the role of additive molecules with the example of chlorobenzene and toluene in the pyrolysis of the three polymers. Whatever the force fields, the toluene molecule does not perturb the behaviour of the plastics along the pyrolysis. The presence of one molecule of chlorobenzene speed up the degradation of the PE. To increase the knowledge of the reliability of the force fields, we investigated the pyrolysis of chlorobenzene and compared the results from reactive molecular simulations with experimental results. The CHOSFClN from the combustion branch over-stabilises the chlorobenzene molecule while the CHONCl-2022_weak force field from the aqueous branch predicts the chlorobenzene degradation during pyrolysis correctly.

Reactive potentials from the aqueous and combustion branches seem to both lead to reliable results. The force field from the aqueous branch might tend to produce better results concerning the small molecules and the pyrolysis products. Moreover, the quality of the results obtained from reactive molecular simulations is significantly improved when the results' uncertainties are evaluated by averaging the data over several independent simulations. In particular, using a set of independent simulations allows basing a comparison over averaged quantities considering the overlap of their standard deviations.

acknowledgement

This work was achieved using HPC ressources from GENCI-CINES (grant 2021-A0100806920), MCIA (Mésocentre de Calcul Intensif Aquitain) and the Université de Pau et des Pays de l'Adour.

Supplementary Material

Plastic	HDPE	PP
methane	1.90	0.93
ethane	2.21	1.45
ethene	6.08	3.52
propane	1.31	1.00
propene	4.56	3.53
butene	0.36	1.29

Table 4.1: Gas Composition from Fixed-Bed Pyrolysis of Individual Plastics (wt %) from the work of Williams and Williams [255].

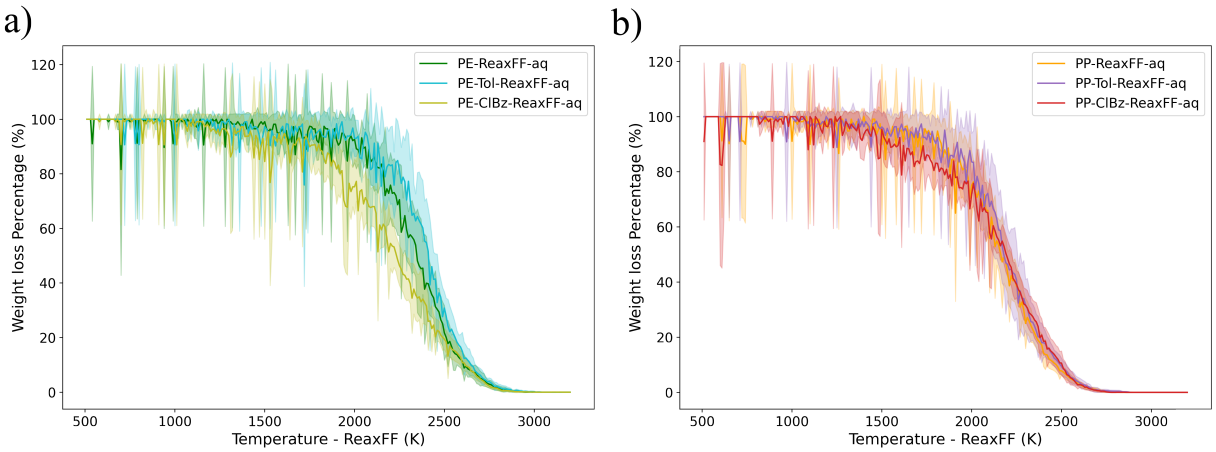


Figure 4.10: The average value of TG curves which are surrounded by the standard deviation as a function of the temperature for PE and PP pure or with one chlorobenzene (ClBz) molecule or toluene (tol) molecule. cb1, stands for combustion branch and corresponds to the CHOSFClN force field. aq, stands for aqueous branch and corresponds to the CHONCl-2022_weak force field.

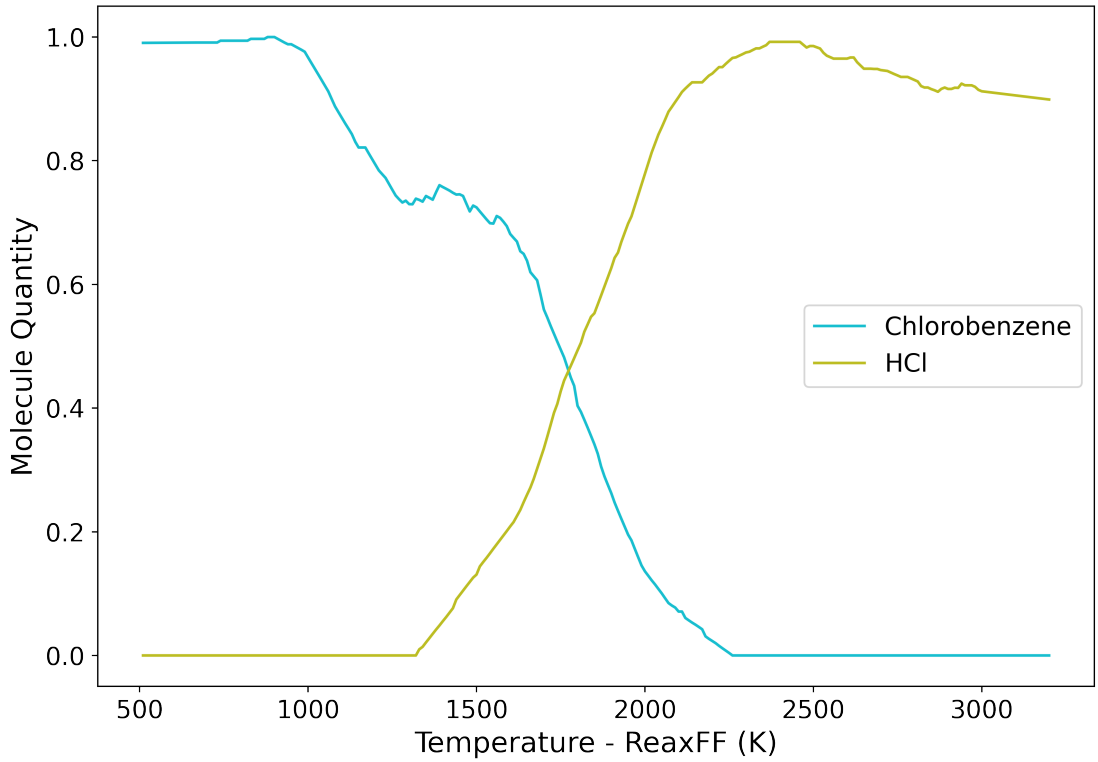


Figure 4.11: Evolution of molecule quantity of chlorobenzene and HCl species in the ten simulations of the PE pyrolysis with the CHONCl-2022_weak force field.

4.3 Conclusion

With this article, we successfully investigate through reactive molecular dynamics the pyrolysis degradation of different plastic polymers. We have shown the capacity of the new aqueous force field CHONCl-2022_weak force field in simulating these degradation processes, even better than the combustion force fields. Furthermore, statistical investigations make the results obtained more reliable than some other investigations made in the literature.

In the idea of exploring the reactivity of chlorinated molecules using the new force field, in this article, the pyrolysis of the chlorinated molecules, which is linked with the first article, brings a complete overview of the chlorobenzene degradation, in atmospheric conditions, in water conditions, and here in a combustion environment. The following chapter will bring more reactivity tracking of different chlorinated molecules through the creation of new software.

Chapter 5

Development of the new software SnookRMol

5.1 Introduction

This chapter presents a new software developed during this thesis. The software is a Python-based code named SnookRMol, created to track the reactivity of simple reactions using LAMMPS and, until now, reactive simulation that implements the ReaxFF potential. In terms of contribution to the article, the code was developed in cooperation with the two authors, with equal implications. The first author implemented the software on the platform of the university dedicated to scientific computing and was responsible for producing the results and their analysis.

To simply define SnookRMol, it generates a set of reactive simulations with target molecules on which probe molecules are thrown, such as projectiles and automatically identifies the reaction. In the first section of the publication, an introduction is given, followed by a complete description of the software, and finishing with the computational details. The first results presented are a comparison of the two types of surfaces used to generate the initial positions of the projectiles around the target molecules using the example of chlorobenzene as the target molecule with the hydroxyl radical as the probe molecule. Afterwards, with SnookRMol, the reactivity of four molecules of high environment interest, DDT, diclofenac, PCB 169 and TCDD are investigated by probing their reactivity with hydroxyl radicals. The results are then compared with the available data in the literature.

5.2 Article 3 : Introduction to SnookRMol, new Python software, using Lammps and ReaxFF for fast regioselectivity identification

Introduction to SnookRMol, new Python software, using Lammps and ReaxFF for fast regioselectivity identification

Matthieu Wolf, Germain Salvato Vallverdu

germain.vallverdu@univ-pau.fr

Universite de Pau et des Pays de l'Adour, E2S UPPA, CNRS, IPREM, UMR 5254, Pau France

abstract

SnookRMol is a Python-based software created to statistically probe the reactivity of a target molecule from a set of simple collisions with a probing one. This article presents the complete workflow of SnookRMol. It allows, with a low computational cost using LAMMPS and ReaxFF potentials, to statistically investigate the regioselectivity of the target molecule, and may be used to determine the most probable pathways of reactions. The initial points of the probe molecule are distributed on the surface and can be generated either from one sphere or to a set of fused spheres. The choice of the surface may depend on the shape of the molecule and its size. For each point of the grid, a reactive simulation is performed, and the software automatically identifies the reactions occurring, which allows data processing afterwards. First, the choice of the surface for the initial points of the projectile is discussed using hydroxyl radical as the projectile and chlorobenzene as the target molecule. Then, a comparison of the results produced with SnookRMol and obtained from the literature is made on four molecules with a high environmental impact, DDT, Diclofenac, PCB 169 and TCDD. The results show that SnookRMol produced consistent results on the regioselectivity and the first step of the reaction pathways. This highlight the capability of SnookRMol to give an overview of the most reactive sites of the target molecule, and it can be used as pre-treatment software for large molecular systems to filter the more plausible reactions and then use a higher level of theory to investigate them.

5.2.1 Introduction

In the field of theoretical chemistry, two main approaches coexist quantum chemistry and molecular mechanics calculations. Both allow us to compute molecular systems' energy using either approximations or empirical potential. When it comes to investigate a molecule's reactivity from theoretical methods, quantum chemistry calculations are highly relevant as they include the description of the electronic density at the origin of the chemical bonds[1,3,4]. Increasing the complexity and the size of molecules leads to numerous reactive pathways and reactive sites. Considering, in an exhaustive way, all the possibilities need a huge amount of computational resources in particular when high-level quantum chemistry methodologies are implemented.

On the other side, reactive molecular simulations, which are extensions of molecular mechanics methodologies are efficient tools for investigating the reactivity of molecular systems. Although the electronic density is not explicitly described, ReaxFF potentials developed initially by van Duin et al.[5] to investigate combustion reactions was extended to various chemical conditions and were used on systems with a high chemical diversity[56,58]. ReaxFF simulations are at least two orders of magnitudes faster than standard quantum chemistry calculations which open the investigation of the reactivity to complex molecules in the condensed phase, including thermodynamic conditions. Some examples of applications include the oxidation of hydrocarbons[56], evaluating the carbonation of polymers[62], DNA-damage by hydroxyl radicals[70], proteins or polymers conformation[71,72].

In this work, we present a new software named `SnookRMol`, taking benefit of the efficiency of the reactive potentials to ease the investigation of the reactive pathways of complex molecules. The approach consists of gathering statistical information over a set of reactive simulations which consist of collisions between a target molecule and a probe molecule. The analyses of the systems at the end of the trajectory helps to identify the highest reactive sites of the target molecule and the main products are resulting from the collisions. The new software is developed as a python library and relies on the LAMMPS code[89] to implement the molecular dynamics simulations. The LAMMPS functionalities are available from a python library, and a standard version of LAMMPS has been available using the conda package manager since 2018[284]. This leads to an easy way to distribute and use the software.

Similarly, a module in the AMS software[90] named molecular gun allows one to place molecules and atoms in a simulation box controlling velocity to simulate collisions and visualize them. `SnookRMol`, on the contrary, is based on a statistical probing of the molecule reactivity by analyzing the products obtained from a large set of trajectories.

This article is divided into three subsections. First, the software architecture is presented. Second, the computational details, including individual trajectories, are detailed. Finally, `SnookRMol` is used to investigate organochloride molecules' reactivity and the reaction pathways.

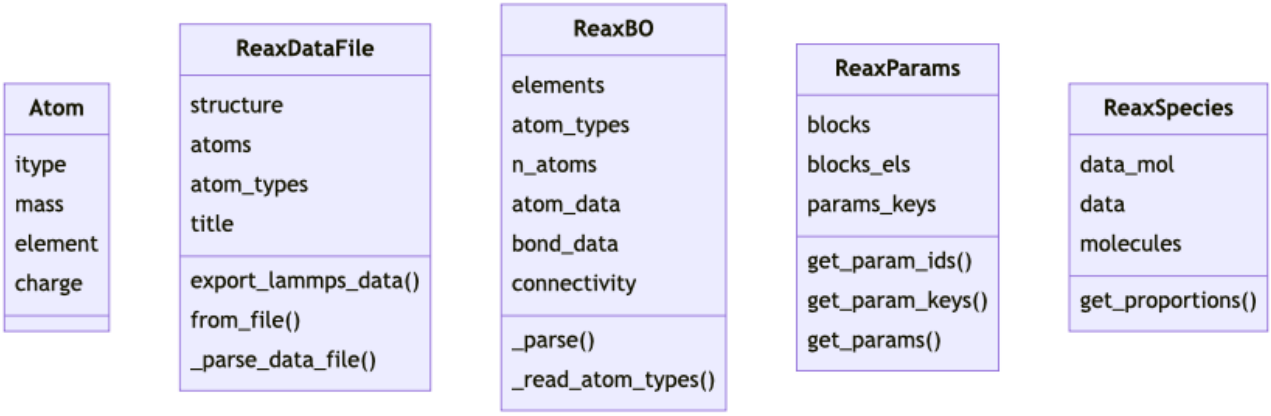


Figure 5.1: Structure of the PyReax classes in the Unified Modeling Language (UML).

5.2.2 Software description

The concept of this software is to fast determine the most favourable reactive site of a molecule and the underlying products. In that scope, it considers a projectile or a probe molecule, such as an hydroxyl radical and sends it numerously through the studied or target molecule and then perform statistical analyses. In order to probe all the possible sites, the initial positions of the probe molecule are distributed over the surface of a sphere centred on a given point or at the surface of fused spheres.

The software is available from a versioned repository on the development platform of UPPA and is part of the PyReax project: `https://git.univ-pau.fr/gvallver/pyreax`. In addition to `SnookRMol`, the PyReax library contains high-level python objects in order to perform pre or post-treatments of ReaxFF simulations. The structure of the classes available in PyReax is presented in figure 5.1. The class `ReaxParms` aims to explore and compare the parameters of ReaxFF force fields. The classes `ReaxDataFile`, `ReaxBO` and `ReaxSpecies` are used in `SnookRMol` to identify the reactions by extracting the bond order values, the connectivity and the chemical species from the output files of a ReaxFF simulation performed with LAMMPS. In order to make it run, `SnookRMol` relies on several python packages related to computational chemistry or material sciences. When installing the software, it is recommended to create a python environment including the different packages, which are all available from the conda package manager. The main ones are: Lammps[89], Pymatgen[205], Matplotlib[285], Numpy[286] and Pandas[287].

Now, we focus on the features of the `SnookRMol` software itself. A graphical representation of the software architecture is given in figure 5.2. The workflow is divided into two subsections corresponding to the production of the simulations, at the top of figure 5.2, and the analysis of the results, at the bottom of figure 5.2. The production of the simulations is performed with the `SnookRMol`.py script and the analysis with the Ana`SnookRMol`.py script.

The next subsections describe the main part of the software and the associated input parameters or use cases.

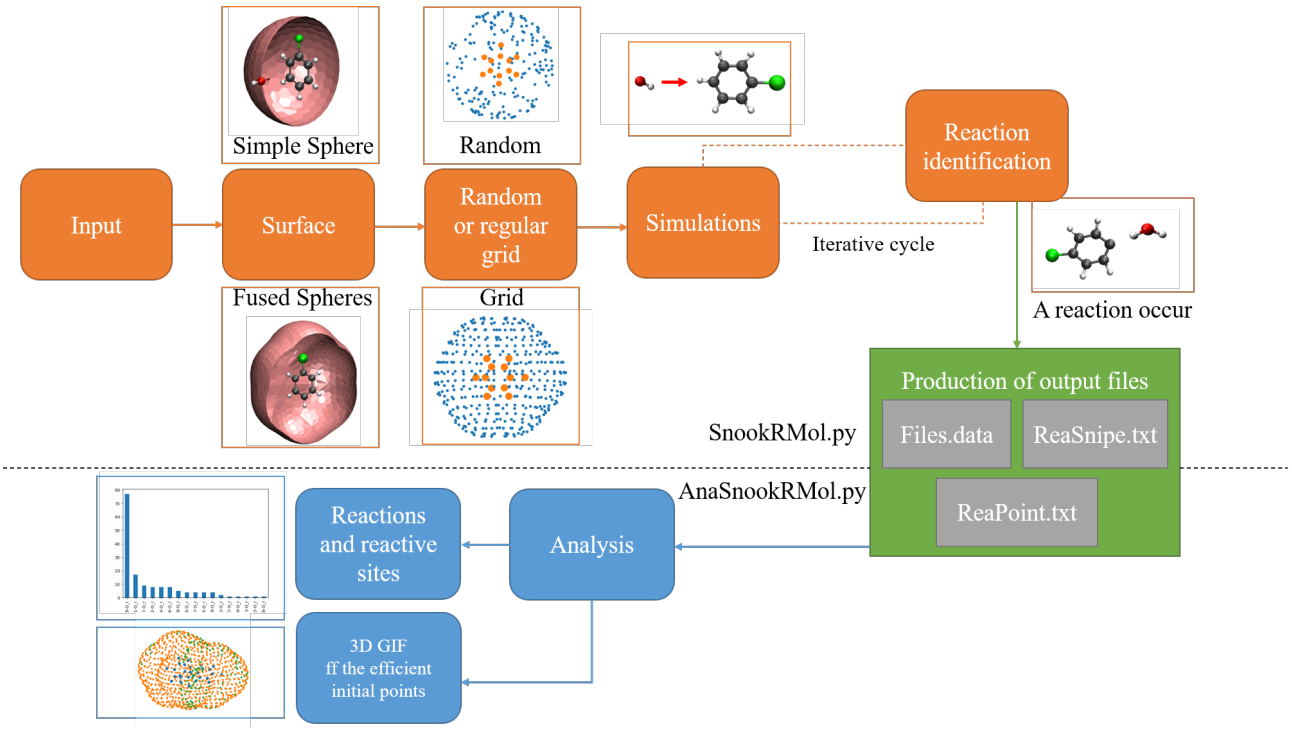


Figure 5.2: Schematizing of the workflow implemented in SnookRMol using the two programs: SnookRMol.py and AnaSnookRMol.py. (i) the top part, in orange, corresponds to the input files ; (ii) the green part, describes the output of the simulations ; (iii) the bottom part, in blue, describes the results of the analysis.

Input Data

The first step of the workflow is to define the different inputs needed. At first, one has to provide the structure of the target and the probe molecules from common molecule file formats such as xyz, mol or pdb files. In figure 5.2, a chlorobenzene molecule was chosen as the target molecule and the hydroxyl radical as the probe or projectile.

Then, as is presented in figure 5.2, the user needs to determine the surface on which the initial positions of the probe molecule will be generated. Two kinds of surfaces are available, and for each kind, the points can be distributed regularly over the surface or randomly. The simplest case corresponds to the surface of one sphere centred around the molecule's centre of mass or a given point defined by the user. In the case of a molecule with a non-globular structure, it might be better to define the surface as fused spheres centred on a series of a given set of points. It means that points included in a neighbour sphere are removed to keep only the surface. One obvious choice from a chemical point of view is to generate a fused sphere's surface using the atom positions as the centres of the spheres. The needed inputs are, thus, the number of points corresponding to the centres of the spheres and the radius of each sphere. After the merging procedure of the spherical grids originating from each sphere, the number of points is obviously reduced and will be smaller than the requested total number of points. Considering the case of regular grids, one must remember that depending on the radius, the number of points effectively generated may be slightly smaller than the requested number

of points due to the overlap between points. This is the case in particular small sphere radius.

Finally, for each point of the grid defined previously, one simulation will be implemented. The user can provide the computational details of these simulations to update predefined parameters, including default LAMMPS values, described in the next subsection. Two specific points are required to be considered carefully. First, obviously, the user has to choose the reactive force field or, more generally, the potential or the Hamiltonian used to compute the forces and the energies and perform the molecular dynamics simulation. Second, the temperature of the simulation has to be chosen globally, or a different temperature can be set for the probe molecule and the target molecule. The probe molecule's temperature will define the projectile's initial velocities, which will be projected in the direction of the target molecule. The temperature of the target molecule will correspond to vibrational motions (rotational and translational motions being removed) and may help to activate the bonds, speed up the reaction or investigate the possible products at different energy levels.

If the target molecule accepts one or several symmetric elements, using the surface obtained from a unique sphere, it is possible to limit the surface to a subpart by giving the boundaries of the θ and φ angles of the spheric coordinates.

Run Simulations

With all the input data being chosen, the simulations for each individual point of the grid can be run. The Software generates the velocity of the projectiles using a Boltzmann sampling to send them towards the molecule. In the case of a simple grid distributed over a sphere surface, the direction of the velocity is aligned with the right joining the point of the grid and the centre of the sphere. In the case of a simple grid distributed over a sphere surface, the direction of the velocity is aligned with the line joining the point of the grid and the centre of the sphere. If the fused spheres grid is made from spheres centred on each atom of the molecule, the initial velocity of the probe molecule is directed toward the atom at the centre of the sphere from which the grid point originates.

The number of simulations depends on the number of points on the defined grid. After each simulation, SnookRMol identified if a reaction occurred by analysing the connectivity between the atoms. These analyses can be done based on the bond order, which is available when using ReaxFF simulations or, more generally, using cutoff distances. Several output files are saved. First, the data file used as input of the LAMMPS simulations containing the initial positions and velocities is saved. This is important, particularly if one wants to reproduce the same trajectory. Then, in the file ReaSnookRMol.txt, all the different kinds of reactions identified over all the simulations are listed. On the file SnookRMolPoint.txt, all the grid points leading to an efficient simulation are listed.

Analysis of the results

At the end of the simulations, the analyses script returns an image of the distribution of all the different reactions identified. The reaction is sorted from the most probable to the lowest one. The nomenclature of the reaction is as follow :

$$\mathrm{At}_i - \mathrm{lab}\ \mathrm{At}_j - \mathrm{lab} \tag{5.1}$$

With At indicating the element of the atom concerned by the reaction, i or j indicates the number of the atom, lab is a label which can be either 'brk' indicating that this atom has one bond break, or 'form' indicating that the atom formed a bond.

Moreover, it generates 3D animated images in a GIF format, representing in orange all the considered grid points and, in green, the points for which an efficient trajectory was obtained, and a reaction occurred.

5.2.3 Computational details

In this subsection the general computational details, used as default parameters in the software are described along with a brief description of the ReaxFF potential. All this computational details can be modified when the software is used.

5.2.3.1 ReaxFF force field

ReaxFF is a force field based on the bond order (BO), which makes that quantity the key point that tunes many energetic terms. The main consequence is that ReaxFF is able to describe bond dissociation and enable the investigation of large reactive systems[211]. The following equation represents the complete expression of the potential with all energy terms which contribute to the total energy:

$$E_{system} = E_{bond} + E_{lp} + E_{over} + E_{under} + E_{val} + E_{pen} + E_{coa} + E_{C2} + E_{triple} + E_{tors} + E_{conj} + E_{Hbond} + E_{vdW} + E_{Coulomb} \tag{5.2}$$

The energetic terms involved in this equation are divided into two main groups: intramolecular interactions (11 first terms) and intermolecular interactions (3 last terms). The BO is used in all intramolecular terms. The intermolecular interactions take into account the Van der Waals, Coulombic and hydrogen-bond interactions. On the intramolecular interactions terms, following the order of the equation (5.2), they are terms to describe the bonds, lone pairs, overcoordination, undercoordination, valence angles, a correction of double bonds angle valency issues, a correction for NO_2 group stability, a correction to destabilize triple bond for carbon, a particular correction for carbon monoxide and finally torsions and conjugation energies. If the reader is

interested in a more detailed review of the ReaxFF potential we recommend to read the published article of Chenoweth et al.[211].

5.2.3.2 Simulations details

All the reactive molecular dynamics simulations using the ReaxFF potential were done using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) code version June 2022[89]. Defaults LAMMPS cutoff parameters were used in all ReaxFF calculations, in particular nbrhood_cutoff equal to 5Å and hbond_cutoff equal to 7.5Å. In the simulations, we used a timestep of 0.1 fs and the length of the trajectory is by default 30000 steps corresponding to 3 ps. The length of the simulations is a crucial parameter which depend on the chosen temperature, the target and the probe molecule and the generated grid. One has to check on a few number of simulations if the length is enough to make reactions occur.

Velocities of the target and the probe molecules were generated from the Boltzmann density assuming a temperature of 500 K. This high temperature is chosen with ReaxFF because it allows to speed up the reaction[279] by activating the bonds of the target molecule. The simulations were implemented in the NVE ensemble.

Use cases of the SnookRMol are presented considering molecules which raise awarness on the environmental point of view and belongs to the organochloride family. The reactive force field CHONCl-2022_weak[7] was thus used in the simulations as it includes the chlorine atom.

5.2.4 Results and discussions

In order to demonstrate the efficiency of this new software to probe at a low computational cost the reactive site of a target molecule, in this subsection are presented the results obtained with SnookRMol.

5.2.4.1 Choice of the surface

Firstly, the choice of the surface on which are distributed the initial position of the probe molecule is discussed considering the results on the chlorobenzene molecule. Recommendations on the choice of the surface are provided.

The SnookRMol software provides the possibility to distribute the initial position of the probe molecule on two different kinds of surfaces: a unique sphere surface or the surface obtained from fused spheres. Firstly, the difference between these two surfaces is their shape. While the first simple surface is a unique sphere, which can be centered on the center of mass of the molecule, the second surface is obtained from the merging of several adjacent spheres. This last surface may be well suited in the case of molecule with a complex shape such as a straight shape for the biphenyl molecules. The amount of spheres and their positions is one of the

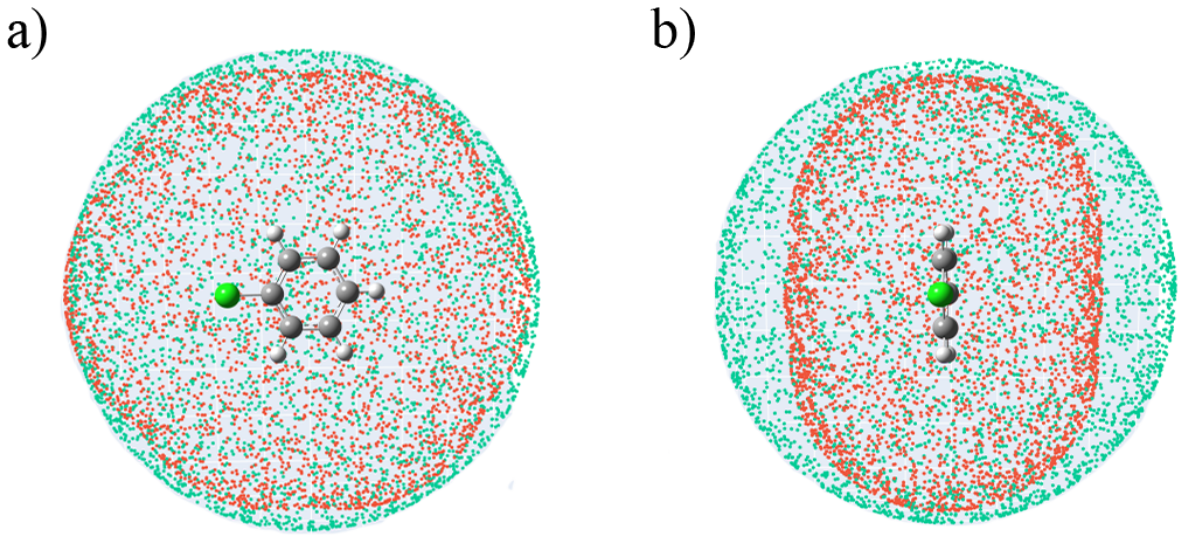


Figure 5.3: Superimposition of the surfaces obtained from a unique sphere of 8.3 Å and from fused spheres of a radius of 5.5 Å represented by cyan and red points, respectively. a) on the left, view in the direction perpendicular to the aromatic plane ; b) on the right view in a direction that contains the aromatic plane.

input of the software. One common approach is to center a sphere on each atom with a given radius and retained the surface uniting the whole spheres. Secondly, the vector direction of the projectile towards the target molecule is different and depend on the choice of the surface. Hence, for the unique sphere, the probe molecule is sent towards the centre of this sphere, whereas in the case of fused spheres, the direction is chosen towards the closest atom (or the center of the sphere to which the point belongs).

Because of these two differences we propose comparing the two surfaces' implementation, using, as an example, the hydroxyl radical as the probe molecule and the chlorobenzene as the target molecule. In order to make comparable results, we selected the radius of the spheres trying to obtain two surfaces that match each other as most as possible. For the fused spheres, we used a radius of 5.5 Å and we used a radius of 8.3 Å in the case of one unique sphere. It leads to the two surfaces depicted in figure 5.3, with the unique sphere and the surface obtained from fused spheres represented by cyan and red points, respectively.

Looking at the surfaces in the direction perpendicular to the aromatic plane, figure 5.3 a) the two surfaces match correctly. However, due to its construction, the fused spheres surface is flatter in the direction perpendicular to the aromatic plane, see figure 5.3 b). Initial velocities were generated in the both cases by sampling the Boltzmann distribution at a temperature of 500 K, and 3580 simulations were performed in each case. It took around twenty hours for both of the whole simulations. The main difference between the two cases is thus the initial direction of the probe molecule.

The results are gathered in figure 5.4 which presents a comparison of the percentage over the effective trajectories of each kind of identified reactions associated to the breaking or the formation of bonds. First

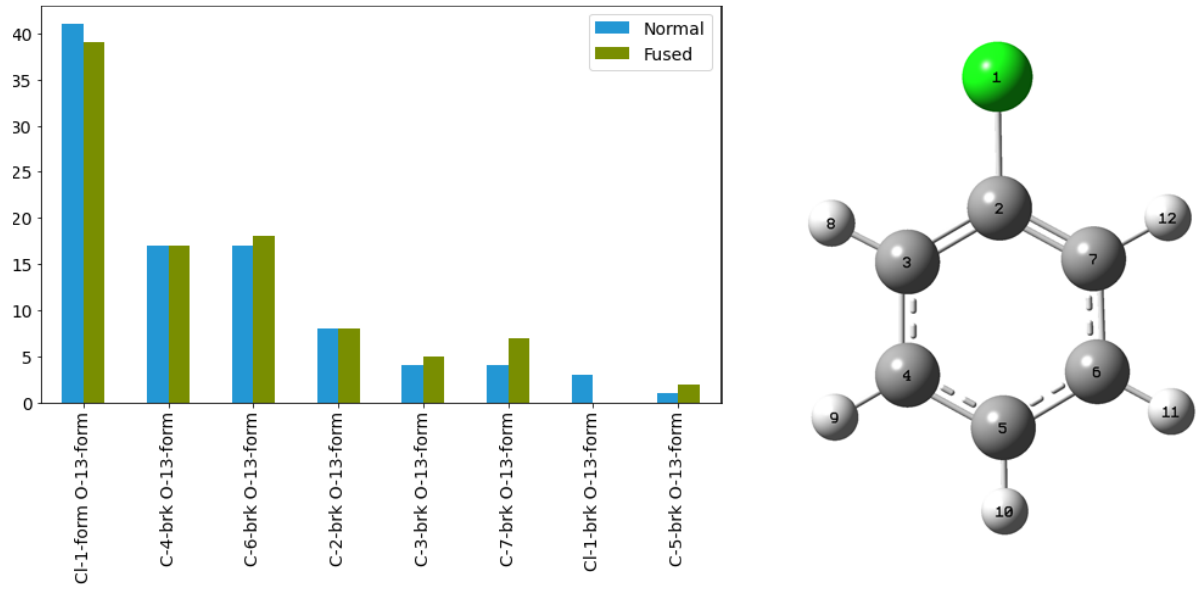


Figure 5.4: a) Percentage of each identified reaction over the total number of reactions when using a surface from a unique sphere (blue) or a fused spheres surface (green) b) Representation and labels of the chlorobenzene molecule. Reaction named include: form (bond-forming) and brk (bond-breaking) with the label of the atom. Carbon atoms are grey, hydrogen atoms are the chloride atom is green.

of all, one can see that the results are very similar in terms of the regioselectivity and the different kind of reactions. This result shows the reliability of the implemented approach to probe the reactivity.

In the both cases, the first major reaction is the abstraction of hydrogen atom in ortho positions (hydrogen 8 or 12) followed by a slight delocalisation of the chlorine atom creating a bond with the remaining carbon. This reaction represents 40% of the total number reactions at 500 K for both surfaces. The second major reaction, considering the symmetrical aspect, is the abstraction of the hydrogen atom in meta positions (corresponding to the hydrogen atom 9 and 11), with both positions around 15%, so in total, 30% of the total reactions. Then, in the both cases around 8% of the reactions correspond of the substitution of the chlorine atom by the oxygen atom leading to the formation of chloride acid. Finally the rarest reactions correspond to the hydrogen abstraction in the ortho (between 5% and 7%) and para position (around 3%). The only difference is that on the unique sphere a low percentage of substitution of the chlorine atom by the hydroxyl radical is obtained.

Comparing the absolute number of reactive simulations instead of the percentage over the number of reactions, 129 reactions against 69 were identified using a unique sphere or the fused spheres surface, respectively. Given the total number of simulations, the amount of effective reactions is very small. This is associated to the stability of the chlorobenzene molecule but enforce the use of cheap computational methodology to perform such analysis. Looking at the points on the surface from which the probe molecule originate, almost no reactions occur from point on the perpendicular axis of the chlorobenzene ring. In the case of a fused spheres surface, the shorter distance of the point in this direction leads to inefficient trajectories. The

hydroxyl radical being closer to the chlorobenzene, it does not have the time to arrange itself, arriving close to the molecule, and it is directly repulsed by the aromatic cycle. This behavior may explain why the amount of efficient trajectories is smaller in the case of the fused sphere surface.

Hence, both surfaces are useful and give similar results on chlorobenzene reactivity. However, in correlation with the previous paragraph, the repulsion observed with the fused spheres surface, which impacts the efficient reaction points obtained at the end of the simulations, might be more meaningful than those obtained with the surface sphere. Indeed, looking at at the surface obtained of the efficient points obtained after all the simulations with the two surfaces. In the surface sphere, the efficient points will be distributed all around the molecules thanks to the arrangement that the probe molecules made to be able to react, and would not bring information of the surface efficient points. Whereas the fused spheres surface obtained well-distributed points in coherence with the reactions probed. From this inspection of the efficient reactive points on the two surfaces, it can be concluded that the surface sphere obtained more statistical information on the reactions, the fused spheres surface can give more meaningful results on the efficient initial points.

Finally, a last comparison can be made between the use of the two surfaces. In the case of large molecules, the usage of the fused spheres surface might be recommended for the simulations. Indeed, let us take, for example, the investigation of a long thin molecule; if the user chooses the surface sphere, it should consider the number of simulation steps to allow the probe molecule to reach the centre of the molecule on the perpendicular axis. Whereas, with the fused spheres surface, if the user chooses a radius between 5.5 Å and 6.5 Å around each atom and a number of steps around 30000 with the default timestep of LAMMPS, the probe molecule will always have the time to reach the molecule and might react with it. That is why in the following subsections, for all the simulations, we used the fused spheres surface with a radius of 5.5 Å and 30000 steps.

5.2.4.2 Reactivity of Organochloride compounds

In this subsection, we present the reactive sites probe with `SnookRMol` of four different molecules: the DDT (dichlorodiphenyltrichloroethane), Diclofenac, the PCB 169 (polychlorinated biphenyl 169) and the TCDD (2,3,7,8-tetrachlorodibenzo-p-dioxin). These molecules were chosen due to their high environmental concern [108,134,288] and the availability of a force field able to describe organochlorides [7].

Reactivity of the DDT

For the investigation of the DDT molecule, the results obtained with `SnookRMol` and the force field CHONCl-2022_weak [7] are compared against the results of, on the one hand, Bai et al. [289] who published a theoretical work, at the MPWB1K/6−311+G(3df,2p)//MPWB1K/6−31+G(d,p) level, on the OH-initiated atmospheric reaction with DDT and, on the other hand, the experimental results of Ma et al. [290] who investigated

the product formed from plasma degradation of desorbed DDT. Bai et al.[289] investigated different kinds of reactions, including the addition of the hydroxyl radical on the DDT, or the abstraction of a hydrogen atom by the hydroxyl radical. However, they did not investigate the chlorine atom substitution. The structure of the DDT molecule is depicted in figure 5.5.

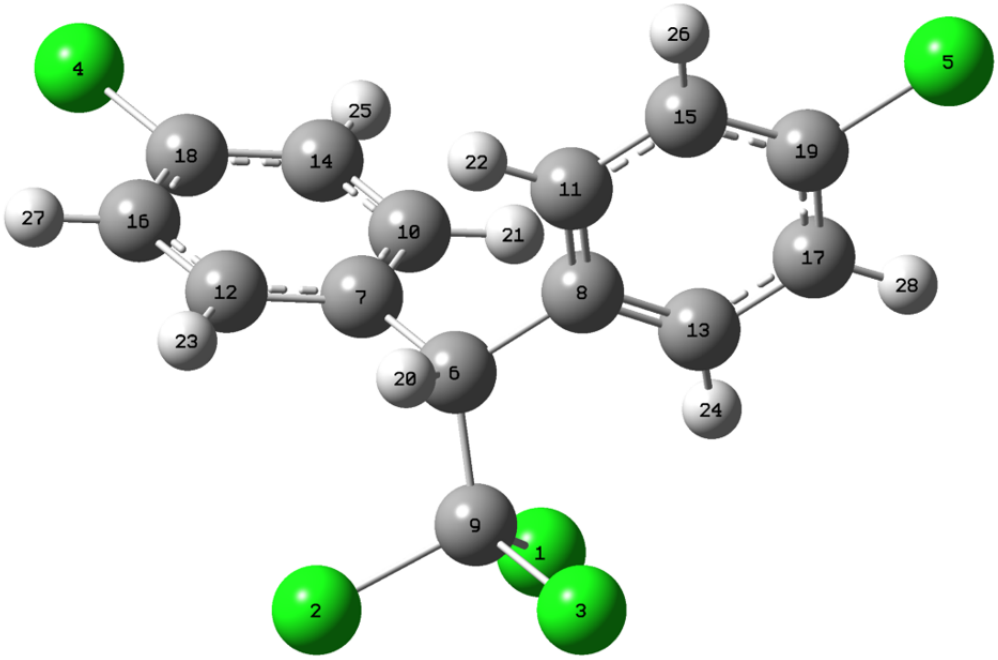


Figure 5.5: Representation of the DDT molecule with the atom labels used to describe the reactivity. Green, gray and white atoms correspond to the chlorine, carbon and hydrogen atoms, respectively.

Concerning addition reactions, it corresponds in the study of Bai et al.[289] to the addition of the OH radical on a carbon atom of one aromatic cycle. Due to the DDT symmetry, carbon atoms are equivalent on the two aromatic cycles. They showed that between 200 and 600 K, the most reactive site was, using the atom labels in figure 5.5, the addition of the C-13 or C-12 atoms, then C-16 or C-17, C-15 or C-14, C-8 or C-7, and finally, C-11 or C-10 atoms. Finally, when the temperature is higher than 600 K, the addition of the carbon atoms C-13 or C-12 becomes less prominent than the additions on the atoms C-17/C-18 and C-15/C-14.

Concerning the abstraction of one hydrogen atom, the major reaction, whatever the temperatures, is an abstraction of the central hydrogen atom, H20, figure 5.5. Then, there is a competition between the atoms H-23/24, H-27/28 and H-25/26, the less reactive being the abstraction of H-21/22 atoms.

With SnookRMol, 3500 simulations were implemented using a temperature for the initial velocities of 500 K. Two major reactions were obtained. Hence, 35% of the occurring reactions are the abstraction of the central hydrogen atom H20, which agrees with the results of Bai et al.[289]. The second prominent reaction is the substitution or abstraction of the chlorine atoms linked to the C-9 atom, with around 33% of the total

reactions. Then, around 3-4% of the reactions concern the chlorine atoms Cl-4 and Cl-5. Unfortunately, the chlorine atoms abstraction or substitution were not investigated by Bai et al.[289]. Nevertheless, in the experimental study of Ma et al.[290], they identified that 87% of chlorine atoms in the degraded DDTs was converted into chloride ions which can be correlated with the high reactions rate of the chlorine atoms' abstraction from our results obtained with SnookRMol.

It is important to note that Bai et al.[289] identified the addition reaction as more prominent than the abstraction. The fact that SnookRMol did not probe the addition of OH comes from the methodology and, more precisely, to the ReaxFF potential. Indeed, using a force field from the aqueous branch such as CHONCl-2022_weak[7], it was indicated that the potential tends to lack in describing accurately the reactivity of OH addition on aromatic rings, which is usually favoured by a π-stacking arrangement.

Reactivity of Diclofenac

In this subsection, we focus on the reaction of the Diclofenac molecule with a hydroxyl radical. The Diclofenac molecule is represented in figure 5.6 a). In the literature, Agopcan Cinar et al.[291] investigated from computational approaches (DFT B3LYP/6-31+G(d) and MPWB1K/6-311+G(3df,2p) level of calculation) the abstraction of hydrogen atoms and the OH additions on carbon atoms. They identified the abstraction of the H-20 or H-21 atom linked to the C-8 atom as the major abstraction reaction. In an experimental investigation Homlok et al.[292] proposed different reaction pathways of OH additions, with one being the addition of OH radical on the C-15 or C-14 atoms and C-10 or C-7 atoms. Finally, the formation of the Diclofenac amide by a new bond between C-16 and N-5 atoms was experimentally identified by Monteagudo et al.[293]. In that case, the O-3, H-30 and H-22 atoms are abstracted from the molecule.

2780 simulations were implemented using SnookRMol, and the results were compared with the literature. The most probable reaction, corresponding to 48% of the total amount of reactions, is the formation of the Diclofenac anionic structure from the abstraction of the H-30 atom belonging to the carboxylic group. This anionic structure was the one considered in the publication of Agopcan Cinar et al.[291] as it is the most probable chemical state in aqueous solution as considered by Homlok et al.[292]. Thus, it seems to be a possible reaction site. Then the two most favoured sites for hydrogen abstraction are the C-8 carbon atom with the H-20 of H-21 abstraction and N-5 with the abstraction of the H-22 atom. These results are in agreement with the work of Agopcan Cinar et al.[291], who determined the same hydrogen atom abstraction as the most probable. Moreover, the abstraction of the H-22 atom can be one of the preliminary steps of the formation of the diclofenac amide identified by Monteagudo et al.[293]. Finally, concerning the addition of hydroxyl radical, the most reactive sites probed by SnookRMol are the carbon atoms C-15 and C-14. The reaction rates are quite low but in agreement with some of the additions reactions described by Homlok et al.[292].

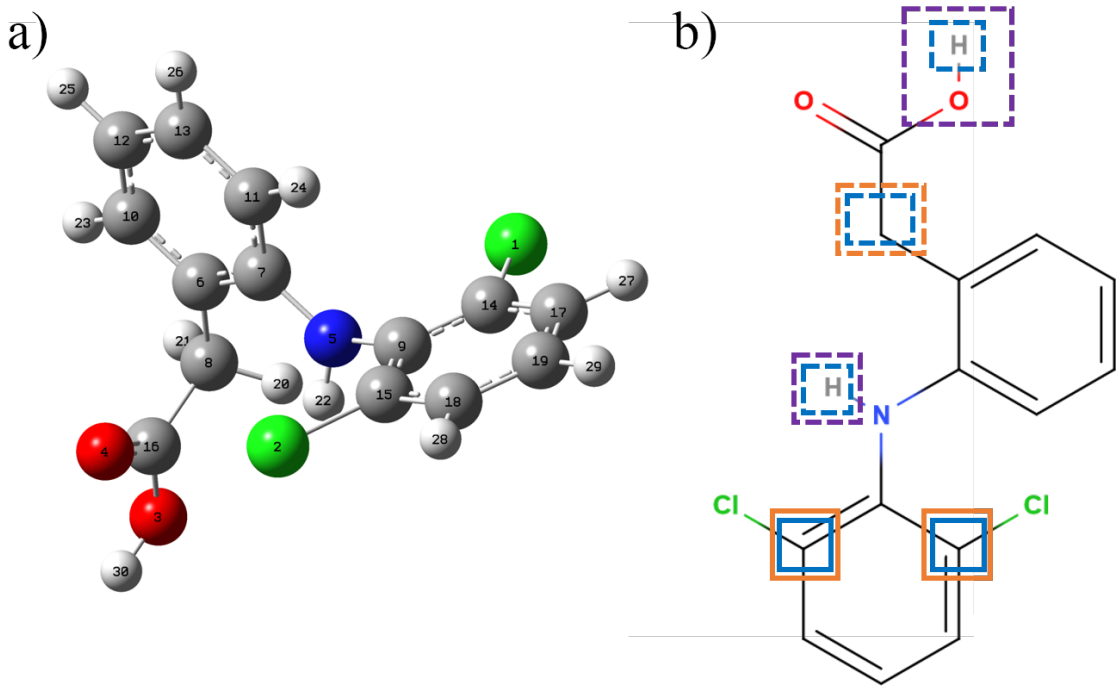


Figure 5.6: a) 3D Representation and labelling of the diclofenac molecule. b) superimposition of the identified reactive sites probe with SnookRMol, in blue, from the theoretical work of Agopcan Cinar et al.[291], in orange and from the experimental work of Monteagudo et al.[293], in violet. Solid line and dashed line are associated to OH addition or atom abstraction, respectively.

In conclusion, SnookRMol gives coherent results on the reactivity of Diclofenac, figure 5.6 b), schematizes the reactive sites of the Diclofenac molecule which are identified by blue, orange and violet rectangles from the results of SnookRMol, the works of Agopcan Cinar et al.[291] and Monteagudo et al.[293], respectively. The solid lines and the dashed lines correspond to the addition of an OH radical or the abstraction of a hydrogen atom, respectively. One can see that SnookRMol is able to identify all the reactive sites on the molecule. Even if the reaction rate is not very accurate, and this may help to filter the possible reaction pathways and select the ones that have to be investigated at a more accurate level of calculation.

PCB169

In this subsection, the reactivity of the PCB 169 is investigated. Only the reactions with a percentage higher than 3% are taken into account. The structure of the PCB 169 is presented in figure 5.7. Padmanabhan et al.[176] computed the local electrophilicity of the PCB 169 at the B3LYP/6-31G(d) level of calculation. They showed that for the C-6 and C-12 atoms, the electrophilicity was higher than any other carbon atoms of the molecule. This indicates a greater reactivity on these two sites. Only two kinds of reactions were identified using the SnookRMol software over 4700 simulations. They are the substitutions of the Cl-19 and Cl-20 atoms linked to the C-12 and C-6 atoms, respectively which is consistent with the electrophilicity calculations of Padmanabhan et al.[176].

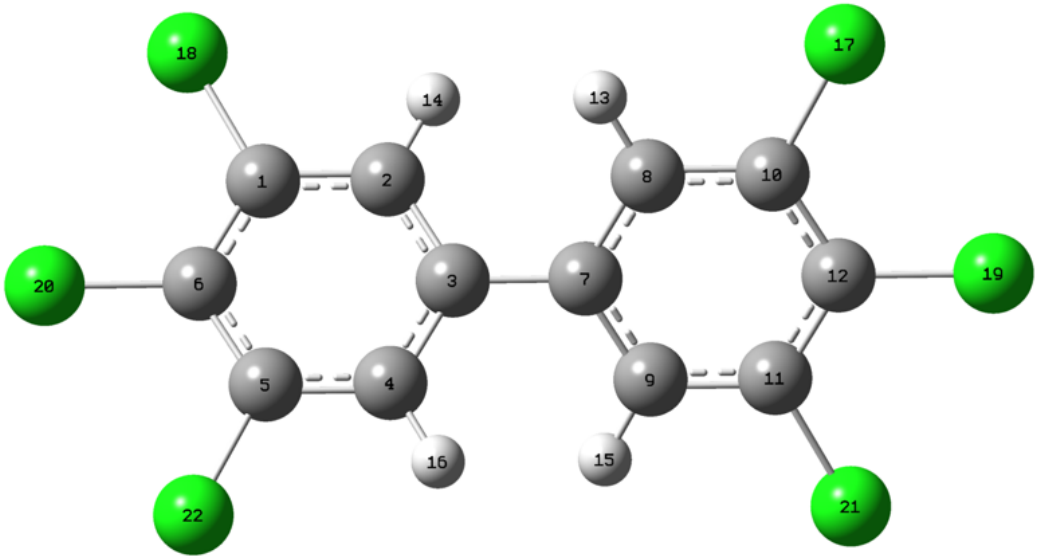


Figure 5.7: Representation of the PCB 169 molecule with the atom labels used to describe the reactivity. Green, gray and white atoms correspond to the chlorine, carbon and hydrogen atoms, respectively.

TCDD

The last molecule investigated in this work is using SnookRMol, concerned the TCDD molecule which is represented in figure 5.8. Using SnookRMol, 4500 simulations were implemented. All the reactions probed with the software are substitutions or abstractions of all the TCDD chlorine atoms with similar reaction rates fluctuating between 6 and 12% of the total number of reactions.

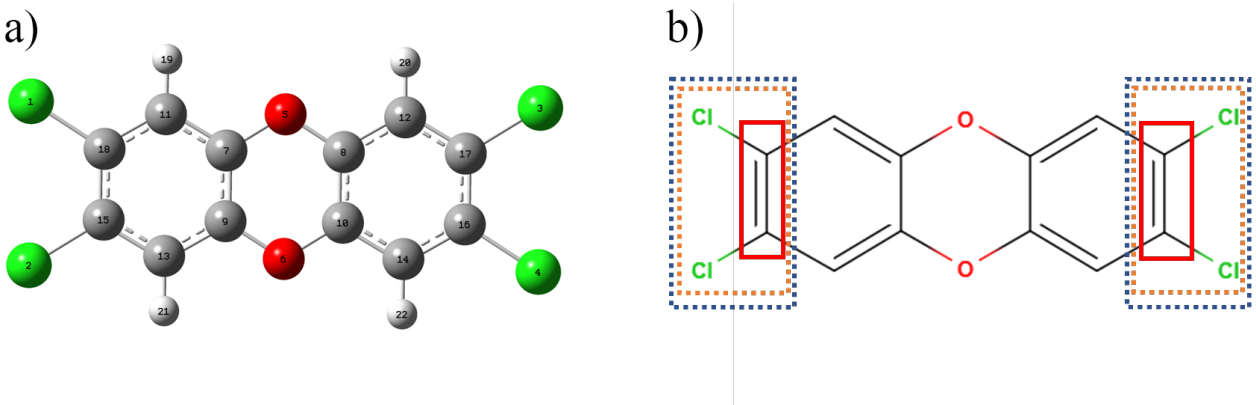


Figure 5.8: a) 3D Representation and labelling of the TCDD molecule. b) superimposition of the identified reactive sites probe with SnookRMol, in blue, from the calculation of electrophilicity by Arulmozhiraja et al. [294], in red and from the calulation of reaction pathways by Wen et al. [295], in orange.

Figure 5.8 b) schematizes the most probable reactive sites from the results obtained with SnookRMol, from the theoretical work of Arulmozhiraja et al. [294] and Wen et al. [295], with blue, red and orange rectangles. In

their work Arulmozhiraja et al. [294] computed the electrophilicity at the B3LYP/6-311G(d,p) level of calculation. They showed that the carbon atoms linked to the chlorine atoms are the most reactive sites. These results are consistent with the regioselectivity obtained from the SnookRMol results. The results also agree with the work of Wen et al. [295] who computed at the B3LYP/6-31G(d)//B3LYP/6-311++G(d,p) level of calculation, the reaction pathways of the substitution of the hydrogen and chloride atoms of the TCDD molecule. Indeed they concluded that the most substitution with the lowest energetical cost is linked to the chlorine atom of the molecule.

5.2.5 Conclusion

This publication presents a new Python software, called SnookRMol, which aims to fast probe the reactivity of a target molecule using a probe molecule as a projectile. Numerous simulations are implemented using the LAMMPS code and the reactive ReaxFF potential in order to gather enough statistical information about the regioselectivity and the first step of the most probable reactive pathways. SnookRMol is included in a versioned python package providing high-level objects for pre and post-treatments of ReaxFF simulations with LAMMPS. The initial positions of the projectile are distributed on the surface of a sphere or a set of fused spheres and the initial velocities are directed toward the atoms or the center of mass of the molecule. The software provides a tool to perform statistical analyses of the reactions occurring during the simulations based on the connectivity of the target molecule and the probe molecule after the collision.

We showed that the percentage of the identified reactions does not rely on the choice of the grid points used to generate the initial position of the probe molecule. This enforces the reliability of the methodology. However, fine-tuning the grid is necessary to save computational time. For small molecules, the sphere surface might gather more statistical information. However, the fused spheres surfaces give more meaningful results on the efficient initial points, and it is more easily applied to large molecules.

We implemented a series of calculations using SnookRMol in order to investigate the reactivity of four molecules which present a great interest due to their environmental impact: DDT, Diclofenac, PCB 169 and TCDD. By comparing our results to experimental or theoretical works available, we demonstrate the efficiency of SnookRMol to predict the main reactive pathways and the regioselectivity of these molecules.

The results provided by SnookRMol can be used in two ways. First, independently, the results quickly provide an overview of the most reactive sites or the main products. This can be implemented in large-scale screening of molecules of interest to investigate the reactivity of a molecule family. Second, SnookRMol being cheaper than quantum mechanics calculations, it may be considered as a pre-treatment software when one is interested in a large or complex molecule. The results obtained from SnookRMol can thus be used to filter the whole possible reaction and implement a more accurate methodology on the most relevant.

acknowledgement

This work was achieved using HPC ressources from GENCI-CINES (grant 2021–A0100806920), MCIA (Mésocentre de Calcul Intensif Aquitain) and the Université de Pau et des Pays de l'Adour.

5.3 Conclusion

This final chapter and the final article presented SnookRMol, the new Python-based software coupling Lammps and ReaxFF to quickly track the reactivity of simple reactions. In the context of this thesis, the software performance was made using chlorinated molecules with the chlorobenzene in the first results section, followed by DDT, diclofenac, PCB 169 and the TCDD reactivity investigations. The results show that, first, a great understanding of the ReaxFF force field will allow the user to extract the results correctly. Secondly, the usage of the normal sphere and the fused gives similar results in terms of reaction percentage distribution. However, even if they both can be used with adequate parameters, it is easier for complex molecules to use the fused sphere, even though it probes fewer reactions. Finally, the reactions tracked with SnookRMol of the four last molecules are coherent with the literature values, demonstrating its capability.

General conclusion

In this thesis, we presented the implementation of reactive molecular simulations to investigate the degradation of organic compounds, specifically organochloride compounds. To succeed in this work, we first developed a new reactive potential to include the chloride atom from the last version of ReaxFF suitable to investigate the chemical reactivity in water. Then, this new potential was validated and used in the simulations of plastic pyrolysis and the development of a new tool in order to probe the reactivity of organic compounds.

Optimizing a new set of parameters for a ReaxFF force field is a hard task because of the numerous parameters involved in all the energetic terms included in the analytical expression of the potential. This leads to optimization with highly undetermined parameters which may not have any physical meaning and strong correlations between some of them. We detailed in chapter III the procedure we followed to obtain our new potential. However, we want to highlight that this procedure might not be the optimal one, but we tried to respect validation checkpoints allowing us to tune the parameters slightly until we reach satisfying results. Among the different force fields published, there are almost none or only a few papers detailing the force field parameterization or giving hints to proceed. That is why we felt it was important to describe how we did. We hope that the procedure implemented for optimizing our new ReaxFF potential in this thesis might help develop new ReaxFF force fields.

Concerning the new force field that we obtained and published, we proved its reliability and identified the improvement that could be made. From this new force, named field CHONCl-2022_weak, several studies can be made on the degradation of pollutants in water with hydroxyl radicals, such as diclofenac, oxazepam, metazachlor, which have been identified in water and sometimes in high quantity. Thus, doing reactive molecular dynamics using the CHONCl-2022_weak force field will bring new insight into the degradation processes of these compounds by hydroxyl radicals in water. Nevertheless, the ReaxFF potential still fails to reproduce the interaction of molecules accurately through π arrangement. In order to improve this point, a new adjustment of the parameters is necessary but including the core part of the ReaxFF potential and C/H/O elements.

In chapter IV, a second article is presented concerning the pyrolysis of plastics and organic molecules. This study shows the great capacity of the new force field to reproduce chlorobenzene pyrolysis. That is why,

in the future, using the new CHONCl-2022_weak force field, an investigation of PCB pyrolysis should be made, as it is one of the major removal technology of the PCB, and it will give information on the secondary products that might be formed from it.

The last chapter of this thesis describes the first version of the SnookRMol software. The aim of SnookRMol is to provide an efficient tool able to probe the reactivity of molecules. This will help investigations in the field of chemical reactivity, providing either a methodology to implement an efficient screening of a whole molecule family or a fast approach in order to identify the first step of the most probable reaction pathways. Many extensions can be applied to the software. For example, in the article, we only used hydroxyl radicals to probe the reactivity, but the structure of the software includes the possibility to consider different probing molecules such as ozone, sulfur dioxide or sulfate. Furthermore, in the context of the thesis, we implemented only reactive simulations based on the ReaxFF force fields, but a good trade-off between accuracy and computational cost could be to implement simulations with semi-empirical Hamiltonian or tight-binding methods.

Appendix A

ReaxFF force field published

In the following table it is presented all the different force field published from the initial 2001 until 2021. They were used in the representation of all the force fields given in the chapter I.

Year	element	DOI	Branch
2001	(C/H)	https://doi.org/10.1021/jp004368u	NaN
2003	(Si/O)	https://doi.org/10.1021/jp0276303	NaN
2004	(Al/O)	https://doi.org/10.1103/PhysRevB.69.045423	NaN
2005	(H/B/N)	https://doi.org/10.1063/1.1999628	NaN
2005	(C/H/O/N/S/Si/Pt/Zr/Ni/Cu/Co)	http://dx.doi.org/10.1021/jp046244d	combustion
2005	(Mg/H)	https://doi.org/10.1021/jp0460184	NaN
2005	(C/H/O/Si)	http://dx.doi.org/10.1021/ja050980t	combustion
2005	(Ti/O)	https://doi.org/10.1021/jp054667p	NaN
2005	(Li/C/H)	https://doi.org/10.1021/jp051450m	NaN
2006	(C/H/O/N/S/Si)	https://doi.org/10.1103/PhysRevLett.96.095505	NaN
2008	(Na/H)	http://dx.doi.org/10.1063/1.2908737	NaN
2008	(C/H/Pt)	http://dx.doi.org/10.1021/jp074806y	NaN
2008	(V/O/C/H)	http://dx.doi.org/10.1021/jp802134x	combustion
2008	(Zn/O/H)	http://dx.doi.org/10.1016/j.susc.2007.12.023	water
2008	(C/H/O)	http://dx.doi.org/10.1021/jp709896w	combustion
2008	(C/H/O/Ba/Zr/Y)	http://dx.doi.org/10.1021/jp076775c	combustion
2009	(C/H/O/N/S/Si)	http://dx.doi.org/10.1021/jp901353a	combustion
2009	(Ni/C/H/O/N/S/F/Pt/Cl)	http://	combustion
2009	(C/H/O/N)	http://dx.doi.org/10.1021/jp900194d	combustion
2010	(Zr/Y/O/H)	https://doi.org/10.1016/j.susc.2010.05.006	NaN
2010	(C/H/O/N/S/Si/Pt/Zr /Ni/Cu/Co/He/Ne/Ar/Kr/Xe)	http://dx.doi.org/10.1021/jp1080302	combustion
2010	(H/O/N/B)	http://dx.doi.org/10.1021/jp100136c	combustion
2010	(Fe/O/C/H/Cl)	http://dx.doi.org/10.1021/jp101332k	water
2010	(Si/O/H)	http://dx.doi.org/10.1063/1.3407433	water
2010	(Ni/C/H)	http://dx.doi.org/10.1021/jp9035056	NaN
2010	(Au/C/S/O/H)	http://dx.doi.org/10.1103/PhysRevB.81.235404	water
2010	(H/C/O/N/S/B)	http://dx.doi.org/10.1021/jp100136c	combustion

2010	(Cu/Cl/H/O)	http://dx.doi.org/10.1021/jp9090415	water
2011	(Au/S/C/H)	http://dx.doi.org/10.1021/jp201496x	NaN
2011	(C/H/O/Fe/Al/Ni/Cu/S)	http://dx.doi.org/10.1021/jp204894m	water
2011	(C/H/O/N)	http://dx.doi.org/10.1021/jp108642r	water
2011	(C/H/O/N/S/Si)	http://dx.doi.org/10.1021/jp201599t	combustion
2011	(Al/H/O)	http://dx.doi.org/10.1016/j.ijhydene.2011.02.035	water
2012	(C/H/O/Ca/Si/X)	http://dx.doi.org/10.1021/la204338m	water
2012	(Si/C/O/H/N/S)	http://dx.doi.org/10.1021/jp306391p	combustion
2012	(C/H/O/N/S/Si/Na/Al)	http://dx.doi.org/10.1021/jp300221j	water
2012	(C/H/O/Fe/Cl/Si/Al/Ca)	http://dx.doi.org/10.1021/ja208894m	water
2012	(H/O/Si/Al/Li)	http://dx.doi.org/10.1088/0965-0393/20/1/015002	water
2012	(C/H/O/N/S/Mg/P/Na/Li)	http://dx.doi.org/10.1021/jp210345b	water
2012	(C/H/O/Fe/Cl/Si/Al)	http://dx.doi.org/10.1007/s11837-012-0463-5	water
2013	(C/H/O/N/S/Mg/P/Na/Ti/Cl/F)	http://dx.doi.org/10.1021/jp4031943	water
2013	(Pd/O)	http://dx.doi.org/10.1063/1.4815820	combustion
2013	(C/H/O/Fe/Cl/Si/Al)	http://dx.doi.org/10.1016/j.combustflame.2012.12.007	water
2013	(C/H/O/N/S/Si/Na/P)	http://dx.doi.org//10.1021/jp406360u	NaN
2013	(C/H/O/N/Si/S)	http://dx.doi.org/10.1021/jp3086649	combustion
2013	(H/O/S/Mg)	http://dx.doi.org/10.1002/jcc.23246	combustion
2013	(C/H/O/N/S/Mg/P/Na/Cu/Cl)	http://dx.doi.org/10.1039/C3CP51931G	water
2013	(C/H/O/N/S/Mg/P/Na/Ti/Cl/F)	http://dx.doi.org/10.1557/jmr.2012.367	water
2013	(C/H/O/N/S/Mg/P/Na/Cu)	http://dx.doi.org/10.1063/1.4774332	water
2013	(C/H/O/N/S/Mg/P/Na/Cu/Cl)	http://dx.doi.org/10.1021/jp312828d	water
2013	(Au/S/C/H)	http://dx.doi.org/10.1021/jp405992m	NaN
2014	(H/O/Ti/Pd)	http://dx.doi.org/10.1021/nn501817w	NaN
2014	(C/H/O/S/F/Cl/N)	http://dx.doi.org/10.1021/jp406248m	combustion
2014	(Pd/H)	http://dx.doi.org/10.1021/jp411015a	NaN
2014	(Co)	http://dx.doi.org/10.1021/jp500053u	combustion
2014	(C/H/O/S/Mo/Ni/Li/B/F/P/N)	http://dx.doi.org/10.1149/2.005408jes	combustion
2014	(C/H/O/N/S/Si/Na/F/Zr)	http://dx.doi.org/10.1021/jp4121029	water
2014	(C/H/O/N/S/Si/Pt /Ni/Cu/Co/Zr/Y/Ba)	http://dx.doi.org/10.1021/jz501891y	combustion
2014	(C/H/O/N/S/Si/Na/P)	http://dx.doi.org/10.1021/jp5054277	combustion
2014	(C/H/O/N/S/Mg/P/Na/Ti/Cl/F)	http://dx.doi.org/10.1016/j.commatsci.2014.07.056	water
2014	(O/Pt)	http://dx.doi.org/10.1039/C4CP03111C	combustion
2015	(Li/S)	http://dx.doi.org/10.1039/C4CP04532G	combustion
2015	(Ce/O)	https://doi.org/10.1021/acs.jpcc.5b01597	NaN
2015	(C/H/O/Li)	http://dx.doi.org/10.1021/ct501027v	water
2015	(C/H/O/Fe/Al/Ni/Cu/S/Cr)	http://dx.doi.org/10.1016/j.actamat.2014.09.047	water
2015	(C/H/O/N/S/Si/Ca /Cs/K/Sr/Na/Mg/Al/Cu)	http://dx.doi.org/10.1021/acs.jpcc.5b00699	water
2015	(C/H/O/Fe/Al/Ni/Cu/S/Cr)	http://dx.doi.org/10.1021/acs.jpcc.5b01275	water
2015	(C)	http://dx.doi.org/10.1021/jp510274e	combustion
2015	(Li/Si)	http://dx.doi.org/10.1039/C4CP05198J	combustion
2015	(C/H/O/Fe/Al/Ni/Cu/S/Cr)	http://dx.doi.org/10.1021/acscatal.5b01766	water
2015	(C/H/O/N/S/F/Pt/Cl/Ni)	http://dx.doi.org/10.1002/cphc.201500527	water

2016	(C/H/O/N/S/Mg/P/Na/Ti/Cl/F/X)	https://doi.org/10.1021/acs.jpcc.5b11332	NaN
2016	(C/H/B/N)	http://dx.doi.org/10.1039/C5CP05486A	NaN
2016	(Al/C/H/O)	http://dx.doi.org/10.1021/acs.jpcc.6b00786	water
2016	(C/H/O/S)	http://dx.doi.org/10.1021/acs.jctc.6b00461	combustion
2016	(C/H/O/N/S/Si/Ge)	https://doi.org/10.1016/j.apsusc.2016.08.082	water
2016	(C/H/Na)	http://dx.doi.org/10.1039/C6CP06774C	NaN
2016	(C/H/Fe)	https://doi.org/10.1039/C5CP06108C	NaN
2016	(C/H/O/N/Si/Cu/Ag/Zn)	http://dx.doi.org/10.1016/j.susc.2015.11.009	water
2016	(C/H/Ar/He/Ne/Kr)	http://dx.doi.org/10.1021/acsnano.6b03036	NaN
2016	(C/H/O/N/S/Si/Ge)	http://dx.doi.org/10.1016/j.susc.2015.08.019	combustion
2016	(C/H/O)	https://dx.doi.org/10.1021/acs.jpca.6b12429	combustion
2016	(C/H/O/N/S/Mg/P/Na/Ti/Cl/F/Au)	http://dx.doi.org/10.1021/acs.jpclett.5b02769	water
2016	(C/H/O/N/S/Mg/P/Na /F/B/Li/El/Ho)	http://dx.doi.org/10.1021/acs.jpcc.6b08688	water
2016	(H/O/Si/Al/Li)	http://dx.doi.org/10.1021/acs.jpca.5b11908	NaN
2017	(H/S/Mo)	http://dx.doi.org/10.1021/acs.jpclett.6b02902	NaN
2017	(H/O/X)	http://dx.doi.org/10.1021/acs.jpcb.7b02548	water
2017	(C/H/O/Fe/Al/Ni/Cu/S/Cr/Si/Ge)	http://dx.doi.org/10.1021/acsami.7b01618	NaN
2017	(C/H)	https://doi.org/10.1063/1.5000534	NaN
2017	(C/H/O)	http://dx.doi.org/10.1016/j.nimb.2016.11.007	combustion
2018	(Ru/N/H)	https://doi.org/10.1021/acsami.8b05070	NaN
2018	(C/H/O/Li/Al/Ti/P)	https://doi.org/10.1039/C8CP03586E	water
2018	(C/H/O/Si/Na)	https://dx.doi.org/10.1021/acs.jpcc.8b05852	water
2018	(C/H/O/N/S/Mg/P/Na/Cu/Cl/X)	http://dx.doi.org/10.1021/acs.jpcb.8b01127	water
2018	(C/H/O/N/S/Mg/P/Na/Cu/Cl/X)	https://doi.org/10.1021/acs.jpca.8b03826	water
2018	(C/H/O/S/Cu/Cl/X)	https://doi.org/10.1021/acs.jpcb.7b06976	NaN
2018	(C/H/O/N/S/Si)	http://dx.doi.org/10.1021/acs.jpcc.8b07075	NaN
2019	(Si/O/H)	https://doi.org/10.1021/acs.jpca.9b01481	combustion
2019	(Zr/Y/O/Ni/H)	https://doi.org/10.1038/s42004-019-0148-x	combustion
2019	(C/H/O/Ge)	http://dx.doi.org/10.1021/acs.jpcc.8b08862	combustion
2019	(C/H/O/Cs/K/Na/Cl/I/F/Li)	http://dx.doi.org/10.1021/acs.jpca.8b10453	water
2019	(C/H/O/N)	https://doi.org/10.1021/acs.jpcb.9b04298	combustion
2019	(Cu/Zr)	https://doi.org/10.1063/1.5112794	NaN
2020	(C/H/O/N/Si)	https://doi.org/10.1021/acs.jpcc.0c01645	water
2020	(Cu/C/O/H)	https://doi.org/10.1021/acs.jpcc.0c02573	combustion
2020	(Av/El)	https://doi.org/10.1063/5.0018971	NaN
2020	(C/H/O/N/S/Mg/P/Na/Ti/Cl/F/K/Li)	https://doi.org/10.1021/acsami.0c17536	water
2021	(Ni/Cr)	https://doi.org/10.1021/acs.jpcc.1c03605	NaN
2021	(W/S/H/Al/O)	https://doi.org/10.1021/acs.jpcc.1c03605	NaN
2021	(I/Pb/Cs/X)	https://doi.org/10.1021/acs.jpclett.1c01192	NaN
2021	(H/O/N/Si/F)	https://doi.org/10.1021/acsomega.1c01824	combustion
2021	(C/H/Ga)	https://doi.org/10.1021/acs.jpcc.1c01965	NaN
2021	(C/H/In)	https://doi.org/10.1021/acs.jpcc.1c01965	NaN
2021	(C/H/O/N/S/Zr)	https://doi.org/10.1021/acs.jpclett.0c02930	combustion
2021	(Si/Al/Mg/O)	https://doi.org/10.1021/acs.jpcc.1c01190	NaN
2021	(C/H/O/S/Mo/Ni/Li/B/F/P/N)	https://doi.org/10.1021/acs.jpclett.1c00279	NaN

Appendix B

Removal of organochlorines

In 2018, Jing et al.[296] published a review on the remediation method used to remove PCBs. This review is divided into three categories, the traditional, the different technologies, and the multi-technologies.

Traditional technologies include phytoremediation, microbial degradation, dehalogenation by chemical reagents and removal by activated carbon.

In phytoremediation technology[297], the biodegradation products of pollutants present in soils and sediments are absorbed by the plant roots and transformed by their enzymes or are directly volatilized into the atmosphere. It has demonstrated the ability to reduce the concentration of PCBs by 31.4% and 78.4% after the first year, and the second year of use. The advantage of this technology is that it is in line with environmental standards and has a low impact. In the short term, it is a technology that needs to be monitored over a long period and can lose its effectiveness due to weathering. The introduction of vegetation into ecosystems can also be detrimental[298]. This technology is also used with DDT[299] with demonstrated efficiency and TCDD[300].

Microbial degradation technology is a natural technology associated with any process where microorganisms degrade and transform contaminants. For PCBs, two routes are possible, anaerobic dehalogenation and aerobic degradation. The anaerobic degradation of highly chlorinated PCBs is mainly carried out by halorespiration (use of electron-accepting halogenated compounds in anaerobic respiration where hydrogen atoms replace chlorine atoms)[301]. The aerobic degradation process uses different enzymes that allow the aromatic ring to be destroyed with the least amount of substituted chlorine. The advantages of this technique is the low technical barriers required. The degraded PCBs are generally less toxic. It allows the dechlorination of highly chlorinated PCBs into conformers with 2 or 3 chlorines atomes making them less toxic. However, the problem with this technology is that only the sediment's upper layers are suitable for aerobic environments. A study on Vietnam soil and sediments also showed that the microbial activity seems to degrade the TCDD[115] slowly, and an efficient strategy using microbial degradation has been developed for DDT removal[302] with 53.6%, of removal in 18 months.

The technology of dehalogenation by chemical reagents consists of converting PCBs to less toxic products by progressively replacing the chlorine atoms under high temperature and high pressure. This technic is much faster than biological treatments. For example, it was observed that the efficiency of PCBs was reduced by 78 % in 2 hours at 100°C[303]. However, it is a difficult technology because of the extreme conditions of applications. For the DDT, using sandwiching iron strategy allows removal of (79 %) in one hour[304].

Finally, the last traditional technology presented here is the use of activated carbon which allows the adsorption of PCBs thanks to its high porosity. This is one of the most effective fundamental approaches to immobilise PCBs[305]. Furthermore, activated carbon can easily be used with other removal technologies.

Now we will look at more advanced technologies. These include supercritical water oxidation, ultrasonic radiation, catalytic hydrodehalogenation using bimetallic systems, and reductive dehalogenation using nanoparticles.

The oxidation of supercritical water is a clean technology that takes place at a temperature and pressure close to the critical point of water (647K and 22.064MPa)[306]. Under these supercritical conditions, water molecules lose their hydrogen bonds and change from a polar to a non-polar liquid. This leads to an increase in the solubility of the PCBs with the supercritical water, which then degrades under high temperature forming carbon dioxide, water and mineral acids[307]. This is a very efficient technology. It has been shown, for example, by Weber et al. in 2002[308], that under alkaline conditions, 99% of PCBs are destroyed. However, during oxidation in supercritical water, the chlorine atoms in the biphenyl rings can form hydrochloric acids, which cause corrosion. Salts can also precipitate and deposit on the surface of equipment requiring high maintenance costs[309]. This technology is also extremely efficient for the TCDD[310], and quite efficient for the DDT with 75% of removal[311].

Ultrasonic radiation is a promising technology for removing PCBs. It has shown an elimination efficiency of over 90%[312] and seems to work with the DDT[313]. The mechanism involved is that of cavitation. When the pressure becomes lower than the saturating vapour pressure, this leads to the formation of vapour bubbles. These bubbles will then release energy resulting in a high temperature and high pressure destroying the chemical bonds. Despite its good efficiency, low environmental impact and no by-product formation, this technology is expensive. It is not suitable for large-scale applications and requires a lot of energy[313].

Bimetallic systems consist of two metals sharing an interface or boundary between two separate phases. The alloy formed at the interface will be the active catalytic phase. These systems are generally made up of a zero-valent metal with a negative reduction potential and another, which would be a transition metal with a high reduction potential, the reducing catalyst[314]. The dehalogenation process will first involve the generation of hydrogen atoms by corrosion of the zero-valent metal with water at ambient pressure and temperature. Following this, the hydrogens will be adsorbed onto the catalyst's surface to form a hybrid metal to be the

substrate for dehalogenation[315]. The efficiency of this technology has been demonstrated to dechlorinate PCBs[316]. It also shown the same conclusion for the TCDD[317] with Ag and Fe, and DDT with Fe and Pd[318]. This technology shows many interests, but some metals are costly, and the phenomenon of passivation can slow down the reaction. The formation of less halogenated and more toxic products can occur.

Finally, the zero-valent iron nanoparticle technology (less than 100nm) starts with forming a hydroxide layer on the surface of the particles after reacting with water and oxygen in an aqueous medium. This layer allows electron transfer from the metal Fe^0 to the PCBs in an acidic medium, thus dehalogenating the PCBs by producing Fe^{2+}. This method has made it possible to eliminate 84% of the PCBs present in the Housatonic River[319]. Moreover, a study on DDT[320] demonstrate the removal of the DDT in the water of 99.2% in 4h. This technology has many advantages, notably its catalytic properties and high degradation rate. However, it remains controversial due to toxicity risks[321].

Finally, in the review by Jing et al.[296], three multi-technologies are presented for the removal of PCBs from the soil. Indeed, it is presented a technology combining biofilms covering activated carbon, an electro-kinetic technology, and finally, a technology linking non-valent iron nanoparticles with a second metal.

The technology combining biofilms and activated carbon is the method that, according to the criteria of Jing et al.[296], has the best performance score considering the cost, efficiency, and the time needed to degrade the PCBs and toxicity of the by-products obtained. As early as the 1970s, it was discovered that biofilms covered with activated carbon effectively eliminated organic pollutants[322]. Thus, to eliminate PCBs, the activated carbon will adsorb the PCBs and biofilms made up of micro-organisms will biodegrade them. However, during the growth and proliferation of certain micro-organisms that slowly dechlorinate PCBs, they can be affected by more easily biodegradable pollutants[323]. This implies a longer elimination time for PCBs.

Electrokinetic removal technology uses a low-level electric current as a cleaning agent to remove organic pollutants from soils[324]. Thus, using a direct external current with a cathode and an anode placed in saturated soil, ion migration and electrophoresis phenomena will transport the pollutants to the electrodes where the current is applied. A coupling with dechlorinating microorganisms has also been tested, giving an efficiency of between 40-60%[325]. A strategy using the Electrokinetic removal on DDT shows a capacity to remove the DDT at around 72%[326].

Finally, the last technology used to eliminate PCBs from the soil, presented here from the review by Jing et al.[296], is the combination of zero-valent iron nanoparticles with a second metal. Here, a second metal (Pt, Ag, Ni) will cover the zero-valent iron nanoparticles, which will reduce the activation energy of the PCBs in order to increase the dechlorination reaction rate[327]. In addition, this technology has shown faster dehalogenation by increasing the particle's surface area, and thus the surface area of activity[328]. This technology has shown efficiencies of 28% for the most chlorinated PCBs up to 99% for tri-chlorinated ones. Moreover, it seems to

be enhanced by high pH (more than 10) during the investigation of DDT dechlorination with zero-valent nanoparticles and Ni[329].

In order to treat the PCBs that are emitted into the atmosphere, the technologies used are closely related to the implementation of PCB treatment technology in incinerators[330]. For example, a dust collector can be used in incinerators to collect particle-bound dioxins. It is also possible to use electrostatic precipitators and scrubbers. The use of activated carbon filters has also been used to limit dioxin emissions.

List of Figures

1 Representation of the subject area of ReaxFF publications, from 4594 articles. xi

1.1 In this schematization in figure a), the blue and yellow points are inseparable in this two-dimensional space (x_1, x_2). However, it can be transformed into a higher dimensional feature space thanks to a suitable kernel function represented by the black arrow. Thus it leads to a linearly separable space (x_1, x_2, x_3) presented in figure b) where the yellow points and blue points can be linearly separated by a plane. 10

1.2 In this schematization of a simple neural network with one neuron, it is represented the different layer with the input X, the weight associated W, the biases term b, the artifical neuron composed of the weighted sum and the activation function (an example of activation function is represented here), and finaly the output predicted Y. 11

1.3 Representation of a neural network, with the initial input X being the initial layer, followed by several Hidden Layers, which the last one being the output layer predicting the output Y. The neurons are indicated in grey, similar to the figure 1.2 with the layer denoted as h_i, the synaptic connections in orange with the weight vector W associated. 12

1.4 Here it is represented a schematization of a combustion subgroup ReaxFF development from the RDX ReaxFF of 2005. Each box represents the year of publication with the element of the force field published and the letters for references with a[55], b[56], c[63], d[64], e[65], f[66], g[60], h[67], i[62]. 14

1.5 Here it is represented a schematization of an aqueous subgroup for biomolecules investigation, developed from the glycine ReaxFF of 2011[80]. Each box represents the year of publication with the element of the force field published and the letters for references with a)[80], b) named protein-2013[71], c) entitled CHON_2017_weak[68], and d) the CHON_2017_weak_bb[72]. 15

1.6 Here it is represented a schematization of the ReaxFF development from the 2001 initial ReaxFF. At the centre of the chart is represented the initial force field of 2001, then each extension of the circle corresponds to two more years, starting with 2003/2004, followed by 2005/2006..., except the last one which contains ReaxFF from 2019, 2020 and 2021. The ReaxFF in red parts belong to the combustion branch, those in blue to the aqueous branch, and in green, they are independent (usually, the O/H parameters are not included). The 116 ReaxFF references depicted here are available in the annexe. 16

1.7 Here, the evolution of the BO' is presented as a function of the interatomic distance r_{ij}. In Black, it is the BO', in red the sigma contribution, in blue the pie contribution and in green the piepie contribution. 18

1.8 Schematization of the Molecular Dynamics algorithm workflow. It starts with an initial phase point denoted as $P(t_0)$, and then the system's total energy is computed, the force is obtained, and with the equation of motion resolved, it gives a new configuration. The iterative cycle is repeated as a function of the number of steps chosen n. 31

1.9 Schematization of PBC in two dimensions with, in black, the simulation box and around the different images. The green particles represent particles going out of the boundary of the simulation box, and the images, re-appearing on the other side, are represented by lighter green particles. 34

1.10 Schematization of LEBC in two dimensions with, in black, the simulation box and around the different images. In these conditions, the system is defined in three layers; In the upper is moving periodic images with constant velocities U_v, and the lower layer with velocities -U_v. The last layer is the medium layer, one with the simulation box. The images on this layer are named stationary periodic images. L indicates the length on the y-axis of the simulation box. 35

1.11 Schematization of the Velocity Verlet algorithms. In red correspond the use of the total forces, in blue the velocity and green the computation of the new forces from the new positions. . . . 36

2.1 It schematizes the synthesis of the DDT ($C_{14}H_9Cl_5$) molecule represented on the right, from the chloral (C_2HCl_3O), the first molecule and the chlorobenzene (C_6H_5Cl), the second. The chlorine is represented in green and the oxygen in red. 40

2.2 Schematization of the formation of the TCDD (C14H9Cl5) molecule from the synthesis of chlorophenols (here with the example of the 3-chlorophenol) or the phenoxy acetic acid. The chlorine is represented in green, and the oxygen in red. 42

2.3 Representation of the PCBs structure using the IUPAC nomenclature. The positions ortho, meta and para are indicated. The letters m and n are associated with the chlorine atoms representing the number of substituted hydrogens. 43

2.4 Here is represented an example of the PCB-1 using the nomenclature of Ballschmiter and Zell. 43

2.5 Schematization of the diverse transport routes of the PCBs in the environment. On the right, it is indicated a legend of all the different ways of transport. In order to simplify the visualization of the representation, the bin symbolizes a landfill, a closed factory is represented in black and white, and a working factory is also represented, the transport is indicated by a car on the lithosphere and by a boat on the hydrosphere. The bioaccumulation is indicated along the trophic chain with the ant and the lion on the lithosphere and the fish with the shark on the hydrosphere. 47

2.6 Here it schematizes on the left the perpendicular structure of the biphenyl and on the right the planar structure. The carbon atoms are represented in grey and hydrogen white. 51

2.7 Representation of PCB-169 with the numbering of the atoms. The grey atoms are carbon, the white ones are hydrogen, and the green ones are chlorine. 55

3.1 Schematisation of a PCB, the number represents the position of the hydrogen or chlorine atoms. Position 6 or 2 are named ortho, 3 and 5 meta and 4 the para. 60

3.2 Schematization of all the different molecules directly used in the training set or the validation set. Chlorine atoms are represented in green, hydrogen grey, oxygen red, and carbon black. . . 61

3.3 We decided to add all the orbitals in the active space for the CASSCF at the equilibrium point of the PES. In the red rectangle, there are the six π orbitals of the benzene ring; in the green rectangle, the σ bonding and anti-bonding orbitals of the C-Cl bond; and in the blue one, the lone pair π orbitals of the chlorine atom. Bellow each orbital is indicated its position and, in parenthesis, its symmetry. 64

3.4 Comparison of the PES of the C-Cl bond dissociation of the chlorobenzene obtained from the CASSCF calculations in blue, with the value obtained from the CASSCF with the NEVPT2 correction in green and the value published by Liu et al.[194] with correction CASPT2 in red. The evolution of the Relative energy (eV) as a function of the distance C-Cl (in Angstrom). . . 65

3.5 Comparison of basis sets for the C-Cl bond dissociation of CH_3Cl. Relative energy (in kcal/mol) as a function of the distance C-Cl (in Angstrom). In red is a singlet PES scan with the B3LYP method and the 6-311g(d,p) basis set, in blue the 6-311+g(d,p) basis set, and in green the 6-31+g(d). 66

3.6 Here is the energetic profile following the reaction starting with HOCl (R1) to HClO (P1) with the transition state (TS). Above each level, the energy difference is indicated (in comparison with the R1 value) in kcal/mol and the image of the molecule (in green is the chlorine atom, in red is the oxygen, and white is the hydrogen). 69

3.7 Representation of two examples of selected points from two different scans. On the left is the choice from the angle Cl-C-C of the dichlorobenzene meta and on the right is the torsion Cl-C-C-H of the chloroethanol. The red curve represents the full PES, and the green is the selected curve. 70

3.8 This graph represents the evolution of the bond order as a function of the C-Cl bond from the combustion reactive force field CHOSFClN[60]. The equation plotted is $\exp\left(-0.0827 \cdot \left(\frac{x}{1.5598}\right)^{5.9023}\right)$. 73

3.9 Here is the 3D representation of the first iteration of the CMA-ES optimizer to obtain a new generation. The surface at the bottom of the 3D representation is a heat map schematizing the fitness evolution in the parameters space with the legend on the rigth side of the graph. The blue cross represents the first initial values of the two parameters in the equation 3.7, and the blue dotted circle is the covariance associated with these parameters. The red cross represents the mean distribution obtained from the equation 3.9, which will be the next initial step for the next iteration. Black points represent the force field generated at this step, with the green point representing the best force field with the lowest fitness. The yellow and blue curves correspond to the two distributions of parameters 1 and 2. The two green axes of the distributions correspond to the two values of these parameters, giving the best fitness among all the different force fields generated. The green point represents this best force field in the parameters space. 79

3.10 Here is the next step after the initial one presented in the figure 3.9, with the mean distribution of the parameters depicted by the red cross and the covariance matrix by the red dotted circle. 80

3.11 Here is the generation step of different force fields represented by black points. 80

3.12 In this step, the different force fields are ranked as a function of the fitness obtained from the equation 3.1. The whiter the point is, the lowest the fitness is. 80

3.13 In this step, the lowest fitness force fields are kept, and the others discarded. 80

3.14 In this step, the mean distribution of the parameters is updated as the covariance matrix associated. 80

3.15 It is shown a new step of generation of force fields with the new mean distribution and covariance matrix. 80

3.16 Following the same procedure as shown in the figure 3.12, 3.13 and 3.14 with the starting points of the figure 3.15, generation of new force fields with new mean distribution and new covariance matrix. 81

3.17 Here is a generation step of new force fields from the new mean distribution and covariance matrix obtained from the previous generation. 81

3.18 This step is the last generation step of force fields close to the global optimum. 81

3.19 10 evaluations of the convergence as a function of the step of the optimization. Each graph contains 10 runs of the optimization which are all denoted by a number. 82

3.20 Distribution of the parameter of the block bonds involving the atoms 10 (Chlorine) and 3 (Oxygen) named "pbo2". In red, the initial value, in green the best force field and in purple, the force field value with the highest fitness. 83

3.21 Distribution of the parameter of the block bonds involving the atoms 10 (Chlorine) and 10 (Chlorine) named "De sigma" (Sigma-bond dissociation energy). In red, the initial value, in green the best force field and in purple, the force field value with the highest fitness. 83

3.22 This heat map represents the correlation between all the parameters optimized from one hundred CMA-ES optimizations. The scale on the right indicates the correlation rate between the parameters, with 0.8 being the highest correlation. 84

3.23 This heat map represents the correlation between all the parameters optimized from one hundred CMA-ES optimizations. The scale on the right indicates the correlation rate between the parameters, with 0.8 being the highest correlation. 85

3.24 The distribution of three different parameter is presented in Figures a) b) and c). A blue curve indicates the KDE (Kernel density estimation) curve on each graph. The distributions of the parameters taking into account only force fields with high fitness (purple curve), low fitness (red curve) and super low fitness (green curve) are presented. The figure d), represents the distribution of the force field fitness, with super low fitness between 1300 and 2100, low fitness between 2100 and 3050 and high fitness more than 3050. 86

3.25 Here it is represented the values of three different parameters as a function of the fitness values. 87

3.26 Summary of the full method developed and used to obtain our new reactive force field. 89

3.27 Comparison between PES (in eV) along bond dissociations obtained with ReaxFF (blue circle) and at the QM CASSCF/NEVPT2 level (red square) as a function of the bond distance (in Å). a) Cl-Cl bond in dichloride molecule; b) Cl-C bond in chloromethanol; c) Cl-O bond in acetyl chloride and d) Cl-H bond in hydrogen chloride. The molecules are depicted on the different graphics with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red. 98

3.28 ReaxFF (blue circle) and QM-B3LYP/6-31+g(d)/D3BJ (red square) comparison of the relative energy (eV) as a function of the angle (in degrees) for the Cl-C-C angle of the chloroethane, the Cl-C-H of the chloromethane, the Cl-C-O of chloromethanol and the Cl-C-C of the ortho-dichlorobenzene. There is a representation of these molecules on the different graphics with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red. Black dotted points show the angle concerned by the PES. 99

3.29 CHONCl-2022_weak ReaxFF (blue circle) and QM-B3LYP/6-31+g(d)/D3BJ (red square) PES (in eV) along several dihedral angles (in degrees). a) the Cl-C-C-H torsion of the p-dichlorobenzene, b) the Cl-C-O-H of the chloromethanol, c) the C-C-C-C torsion around the central bond of the PCB101 and the PCB169 (d). There is a representation of these molecules on the different graphics with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red. Black arrows indicate the considered dihedral angle. 101

3.30 Carbon atomic charges of several PCBs, a) PCB 22, b) PCB 52, and c) PCB 101, computed with ReaxFF CHONCl-weak_2022 (in blue) and DFT from the reference 177 (in red). Blue rectangles indicate the favourable carbon atoms for electrophilic additions and in red for nucleophilic. . . 104

3.31 Reaction pathways obtained at the B3LYP/6-31+g(d)/D3BJ level (red) and with CHONCl-2022_weak on the QM geometries (black) for the reactions a) $CH_3Cl + H_2 \longrightarrow CH_4 + HCl$, b) $HCOOH + HCl \longrightarrow HCOCl + H_2O$, c) the recombination of HOCl in HClO and d) the dissociation $CH_2Cl-OH \longrightarrow H_2CO + HCl$. There is a representation of these molecules on the different diagrams with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red. 105

3.32 Reaction pathways obtained at the M06-2X/6-311++g(2df,2p) level (orange from reference 243 or this work) and using ReaxFF with CHONCl-2022_weak (black). For all schemes, the initial reactants are chlorobenzene and hydroxyl radical in the middle. The left side corresponds to the addition of the hydroxyl radical, while the right side corresponds to the abstraction of a hydrogen atom of the phenyl ring. The reaction is done on the substituent in position a) ipso, b) ortho, c) meta and d) para. The final products are obtained after the reaction with dioxygen. The energy barriers are in electron-volt and relative to the initial reactants except for the last products, after the addition of dioxygen materialized by the blue vertical dashed lines, where in these cases, the barrier is relative to the previous state. There is a representation of all the different molecules on the different diagrams with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red. 108

3.33 Representation of key snapshots along one reaction pathway leading to the abstraction of an hydrogen atom in ortho position. The molecules are depicted on the different graphics with the chlorine atom in green, carbon in grey, hydrogen in white and oxygen in red. 109
3.34 Initial and last simulation boxes of the plasma-induced reactivity of chlorobenzene in water along with intermediate key snapshots of the reaction pathway and their corresponding timestep. 111
3.35 Here is a schematizing of a PCB, with the different positions indicated by numbers. 114
3.36 Representation of six different PCBs cited in the article. The PCB 15, 22, 52, 77, 126 and 169. The grey atoms are carbons, green chlorines, and white hydrogens. 115

4.1 Three boxes of molecular dynamics simulation starting points. The first is a PE box with one chlorobenzene molecule, the second a PP box, and the final a box of PS with one toluene molecule. On each box, the different polymers are coloured. 124
4.2 Thermogravimetric analyses obtained from 10 reactive simulations of PP pyrolysis. The dark yellow curve is the average over the 10 TG curves and the light envelope correspond to the standard deviation with a width of 2σ. 126
4.3 Thermogravimetric analyses obtained from experimental studies and ReaxFF simulations for pure PP, PE and PS polymers. a) and b) are experimental work of Abbas-Abadi [269] and Dümichen et al. [253] ; ReaxFF simulations of this work with c) CHOSFClN potential [60] ; d) CHON-2019 potential [62] ; e) CHONCl-2022_weak [277]. F) ReaxFF simulations from the work of Hong et al. [271]. c), d) and f) correspond to potential from the combustion branch and e) from the aqueous branch. Purple, green and orange curves correspond to the TG of PS, PE and PP respectively. . 127
4.4 Amount of species produced from the pyrolysis of PE as a function of the temperature. a) experimental results from Honus et al. [268]. The amount of species is reported in the gas phase as the volumetric percentage. b) c) and d) present the reactive simulation results using the CHOSFClN, the CHON-2019 and the CHONCl-2022_weak force fields, respectively. The amount of each species is the percentage over the total number of molecules. The same species are present on all graphics. 129
4.5 Amount of species produced from the pyrolysis of PP as a function of the temperature. a) experimental results from Honus et al. [268]. The amount of species is reported in the gas phase as the volumetric percentage. b) c) and d) present the reactive simulation results using the CHOSFClN, the CHON-2019 and the CHONCl-2022_weak force fields, respectively. The amount of each species is the percentage over the total number of molecules. The same species are present on all graphics. 131

4.6 Amount of species produced from the pyrolysis of PS as a function of the temperature. a) experimental results from Maafa[283]. The amount of species is reported as the mass fraction in percentage. b) c) and d) present the reactive simulation results using the CHOSFClN, the CHON-2019 and the CHONCl-2022_weak force fields, respectively. The amount of each species is the percentage over the total number of molecules. The same species are present on all graphics. 133

4.7 Thermogravimetric analyses obtained from ReaxFF simulations for PP, PE and PS polymers with no additive, one chlorobenzene (ClBz) molecule or toluene (tol) molecule. cb1, stands for combustion branch and corresponds to the CHOSFClN force field. aq, stands for aqueous branch and corresponds to the CHONCl-2022_weak force field. 134

4.8 Percentage of the number of chlorobenzene molecules as a function of the temperature. In blue, and with the blue x-axis, from ReaxFF simulations with the CHONCl-2022_weak force field. In red and black, experimental results[274] from a jet-stirred-reactor (JSR) and a tubular reactor (TR), respectively. The two x-axis are scaled to match the experimental and simulation results. 136

4.9 Mole fraction of a) benzene, b) HCl, c) biphenyl and d) PCB (PolyChlorinated Biphenyl) as a function of the temperature along the pyrolysis. Results from reactive simulations are in blue. Pink and Orange curves are experimental results from Vin et al.[274]. Red and Green curves are extrapolation from experimental investigation of Vin et al.[274]. On each plot, the two x-axis are scaled according to figure 4.8. 137

4.10 The average value of TG curves which are surrounded by the standard deviation as a function of the temperature for PE and PP pure or with one chlorobenzene (ClBz) molecule or toluene (tol) molecule. cb1, stands for combustion branch and corresponds to the CHOSFClN force field. aq, stands for aqueous branch and corresponds to the CHONCl-2022_weak force field. . 140

4.11 Evolution of molecule quantity of chlorobenzene and HCl species in the ten simulations of the PE pyrolysis with the CHONCl-2022_weak force field. 140

5.1 Structure of the PyReax classes in the Unified Modeling Language (UML). 146

5.2 Schematizing of the workflow implemented in SnookRMol using the two programs: SnookRMol.py and AnaSnookRMol.py. (i) the top part, in orange, corresponds to the input files ; (ii) the green part, describes the output of the simulations ; (iii) the bottom part, in blue, describes the results of the analysis. 147

5.3 Superimposition of the surfaces obtained from a unique sphere of 8.3 Å and from fused spheres of a radius of 5.5 Å represented by cyan and red points, respectively. a) on the left, view in the direction perpendicular to the aromatic plane ; b) on the right view in a direction that contains the aromatic plane. 151

5.4 a) Percentage of each identified reaction over the total number of reactions when using a surface from a unique sphere (blue) or a fused spheres surface (green) b) Representation and labels of the chlorobenzene molecule. Reaction named include: form (bond-forming) and brk (bond-breaking) with the label of the atom. Carbon atoms are grey, hydrogen atoms are the chloride atom is green. 152

5.5 Representation of the DDT molecule with the atom labels used to describe the reactivity. Green, gray and white atoms correspond to the chlorine, carbon and hydrogen atoms, respectively. 154

5.6 a) 3D Representation and labelling of the diclofenac molecule. b) superimposition of the identified reactive sites probe with SnookRMol, in blue, from the theoretical work of Agopcan Cinar et al.[291], in orange and from the experimental work of Monteagudo et al.[293], in violet. Solid line and dashed line are associated to OH addition or atom abstraction, respectively. 156

5.7 Representation of the PCB 169 molecule with the atom labels used to describe the reactivity. Green, gray and white atoms correspond to the chlorine, carbon and hydrogen atoms, respectively. 157

5.8 a) 3D Representation and labelling of the TCDD molecule. b) superimposition of the identified reactive sites probe with SnookRMol, in blue, from the calculation of electrophilicity by Arulmozhiraja et al.[294], in red and from the calulation of reaction pathways by Wen et al.[295], in orange. 157

Bibliography

[1] Hehre, W. J. *A guide to molecular mechanics and quantum chemical calculations*; Wavefunction Irvine, CA, 2003; Vol. 2.

[2] Greenstein, G.; Zajonc, A. *The Quantum Challenge*; Jones and Bartlett Publishers, 2006.

[3] Deglmann, P.; Schäfer, A.; Lennartz, C. Application of quantum calculations in the chemical industry—An overview. *International Journal of Quantum Chemistry* **2015**, *115*, 107–136.

[4] Nalewajski, R. F. *Chemical Reactivity in Quantum Mechanics and Information Theory*; Elsevier: Walthm, MA, USA, 2022.

[5] van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A Reactive Force Field for Hydrocarbons. *The Journal of Physical Chemistry A* **2001**, *105*, 9396–9409.

[6] Breivik, K.; Sweetman, A.; Pacyna, J. M.; Jones, K. C. Towards a global historical emission inventory for selected PCB congeners — A mass balance approach: 3. An update. *Science of The Total Environment* **2007**, *377*, 296–307.

[7] Wolf, M.; Bégué, D.; Salvato Vallverdu, G. Development of a novel ReaxFF reactive potential for organochloride molecules. *J. Chem. Phys.* **2022**,

[8] Goldstine, H. H. *The Computer from Pascal to Von Neumann*; Princeton University Press, 1993.

[9] Eckhardt, R. Stan ulam, john von neumann, and the monte carlo method. *Los Alamos Science* **1987**, *15*, 30.

[10] Nascimento, M.; Peixoto, J. Monte Carlo Method and Their Codes Applied to Ionizing Radiation. *Journal of Physics: Conference Series* **2021**, *1826*, 012049.

[11] Haigh, T.; Priestley, M.; Rope, C. Los Alamos Bets on ENIAC: Nuclear Monte Carlo Simulations, 1947-1948. *IEEE Annals of the History of Computing* **2014**, *36*, 42–63.

[12] Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* **1953**, *21*, 1087–1092.

[13] Fermi, E.; Pasta, P.; Ulam, S.; Tsingou, M. STUDIES OF THE NONLINEAR PROBLEMS. **1955**,

[14] Alder, B. J.; Wainwright, T. E. Molecular dynamics by electronic computers. *Transport processes in statistical mechanics* **1958**, 97–131.

[15] Alder, Kawasaki Named This Year's Boltzmann Medalists. *Physics Today* **2001**, *54*, 73–73.

[16] Gibson, J. B.; Goland, A. N.; Milgram, M.; Vineyard, G. H. Dynamics of Radiation Damage. *Physical Review* **1960**, *120*, 1229–1253.

[17] Rahman, A. Correlations in the Motion of Atoms in Liquid Argon. *Physical Review* **1964**, *136*, A405–A411.

[18] McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of folded proteins. *Nature* **1977**, *267*, 585–590.

[19] González, M. Force fields and molecular dynamics simulations. *École thématique de la Société Française de la Neutronique* **2011**, *12*, 169–200.

[20] Morse, P. M. Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels. *Physical Review* **1929**, *34*, 57–64.

[21] Jones, J. E.; Chapman, S. On the determination of molecular fields. —II. From the equation of state of a gas. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1924**, *106*, 463–477.

[22] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* **1995**, *117*, 5179–5197.

[23] Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American Chemical Society* **1992**, *114*, 10024–10035.

[24] Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry* **1983**, *4*, 187–217.

[25] Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. Structure and energetics of ligand binding to proteins: Escherichia coli dihydrofolate reductase-trimethoprim, a drug-receptor system. *Proteins: Structure, Function, and Bioinformatics* **1988**, *4*, 31–47.

[26] Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry* **1996**, *17*, 490–519.

[27] Allinger, N. L.; Yuh, Y. H.; Lii, J. H. Molecular mechanics. The MM3 force field for hydrocarbons. 1. *Journal of the American Chemical Society* **1989**, *111*, 8551–8566.

[28] Sun, H. COMPASS: An ab Initio Force-Field Optimized for Condensed-Phase ApplicationsOverview with Details on Alkane and Benzene Compounds. *The Journal of Physical Chemistry B* **1998**, *102*, 7338–7364.

[29] van Gunsteren, W. F.; Billeter, S.; Eising, A.; Hünenberger, P.; Krüger, P.; Mark, A.; Scott, W.; Tironi, I. Biomolecular simulation: the GROMOS96 manual and user guide. *Vdf Hochschulverlag AG an der ETH Zürich, Zürich* **1996**, *86*, 1–1044.

[30] Jorgensen, W. L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society* **1988**, *110*, 1657–1666.

[31] Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology* **1976**, *103*, 227–249.

[32] Patel, S.; Brooks, C. L. Fluctuating charge force fields: recent developments and applications from small molecules to macromolecular biological systems. *Molecular Simulation* **2006**, *32*, 231–249.

[33] Dudek, M. J.; Ponder, J. W. Accurate modeling of the intramolecular electrostatic energy of proteins. *Journal of Computational Chemistry* **1995**, *16*, 791–816.

[34] Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *The Journal of Physical Chemistry B* **2010**, *114*, 2549–2564.

[35] Zhang, C.; Lu, C.; Jing, Z.; Wu, C.; Piquemal, J.-P.; Ponder, J. W.; Ren, P. AMOEBA Polarizable Atomic Multipole Force Field for Nucleic Acids. *J. Chem. Theory Comput.* **2018**, *14*, 2084–2108.

[36] Levitt, M.; Warshel, A. Computer simulation of protein folding. *Nature* **1975**, *253*, 694–698.

[37] Marrink, S. J.; de Vries, A. H.; Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *The Journal of Physical Chemistry B* **2004**, *108*, 750–760.

[38] Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K.-R. Machine Learning Force Fields. *Chem. Rev.* **2021**, *121*, 10142–10186.

[39] Benoit, M.; Amodeo, J.; Combettes, S.; Khaled, I.; Roux, A.; Lam, J. Measuring transferability issues in machine-learning force fields: The example of Gold-Iron interactions with linearized potentials. *Mach. Learn.: Sci. Technol.* **2020**,

[40] Rupp, M.; Tkatchenko, A.; Müller, K.-R.; von Lilienfeld, O. A. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. *Phys. Rev. Lett.* **2012**, *108*, 058301.

[41] Unke, O. T.; Chmiela, S.; Gastegger, M.; Schütt, K. T.; Sauceda, H. E.; Müller, K.-R. SpookyNet: Learning force fields with electronic degrees of freedom and nonlocal effects. *Nat. Commun.* **2021**, *12*, 1–14.

[42] Gong, Z.; Wu, Y.; Wu, L.; Sun, H. Predicting Thermodynamic Properties of Alkanes by High-Throughput Force Field Simulation and Machine Learning. *J. Chem. Inf. Model.* **2018**, *58*, 2502–2516.

[43] Montavon, G.; Rupp, M.; Gobre, V.; Vazquez-Mayagoitia, A.; Hansen, K.; Tkatchenko, A.; Müller, K.-R.; von Lilienfeld, O. A. Machine Learning of Molecular Electronic Properties in Chemical Compound Space. *arXiv* **2013**,

[44] Koner, D.; Unke, O. T.; Boe, K.; Bemish, R. J.; Meuwly, M. Exhaustive state-to-state cross sections for reactive molecular collisions from importance sampling simulation and a neural network representation. *J. Chem. Phys.* **2019**, *150*, 211101.

[45] Noé, F.; De Fabritiis, G.; Clementi, C. Machine learning for protein folding and dynamics. *Curr. Opin. Struct. Biol.* **2020**, *60*, 77–84.

[46] Sauceda, H. E.; Gálvez-González, L. E.; Chmiela, S.; Paz-Borbón, L. O.; Müller, K.-R.; Tkatchenko, A. BIGDML—Towards accurate quantum machine learning force fields for materials. *Nat. Commun.* **2022**, *13*, 1–16.

[47] Goldberg, Y.; Elhadad, M. splitSVM: fast, space-efficient, non-heuristic, polynomial kernel computation for NLP applications. Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Short Papers. 2008; pp 237–240.

[48] Jaakkola, T. S.; Diekhans, M.; Haussler, D., et al. Using the Fisher kernel method to detect remote protein homologies. ISMB. 1999; pp 149–158.

[49] Convolution Kernels on Discrete Structures. 1999.

[50] McCulloch, W. S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics* **1943**, *5*, 115–133.

[51] Yoo, P.; Sakano, M.; Desai, S.; Islam, M. M.; Liao, P.; Strachan, A. Neural network reactive force field for C, H, N, and O systems. *npj Comput. Mater.* **2021**, *7*, 1–10.

[52] Akkermans, R. L. C.; Toxvaerd, S.; Briels, W. J. Molecular dynamics of polymer growth. *The Journal of Chemical Physics* **1998**, *109*, 2929–2940.

[53] van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A Reactive Force Field for Hydrocarbons. *The Journal of Physical Chemistry A* **2001**, *105*, 9396–9409.

[54] van Duin, A. C. T.; Strachan, A.; Stewman, S.; Zhang, Q.; Xu, X.; Goddard, W. A. ReaxFFSiO Reactive Force Field for Silicon and Silicon Oxide Systems. *The Journal of Physical Chemistry A* **2003**, *107*, 3803–3811.

[55] Strachan, A.; Kober, E. M.; van Duin, A. C. T.; Oxgaard, J.; Goddard, W. A. Thermal decomposition of RDX from reactive molecular dynamics. *The Journal of Chemical Physics* **2005**, *122*, 054502.

[56] Chenoweth, K.; Van Duin, A. C.; Goddard, W. A. ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation. *The Journal of Physical Chemistry A* **2008**, *112*, 1040–1053.

[57] Raymand, D.; van Duin, A. C. T.; Baudin, M.; Hermansson, K. A reactive force field (ReaxFF) for zinc oxide. *Surface Science* **2008**, *602*, 1020–1031.

[58] Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M.; Verstraelen, T.; Grama, A.; Duin, A. C. T. v. The ReaxFF reactive force-field: development, applications and future directions. *npj Computational Materials* **2016**, *2*, 1–14.

[59] Weismiller, M. R.; van Duin, A. C. T.; Lee, J.; Yetter, R. A. ReaxFF Reactive Force Field Development and Applications for Molecular Dynamics Simulations of Ammonia Borane Dehydrogenation and Combustion. *The Journal of Physical Chemistry A* **2010**, *114*, 5485–5492.

[60] Wood, M. A.; van Duin, A. C. T.; Strachan, A. Coupled Thermal and Electromagnetic Induced Decomposition in the Molecular Explosive αHMX; A Reactive Molecular Dynamics Study. *The Journal of Physical Chemistry A* **2014**, *118*, 885–895.

[61] Merinov, B. V.; Mueller, J. E.; van Duin, A. C. T.; An, Q.; Goddard, W. A. ReaxFF Reactive Force-Field Modeling of the Triple-Phase Boundary in a Solid Oxide Fuel Cell. *The Journal of Physical Chemistry Letters* **2014**, *5*, 4039–4043.

[62] Kowalik, M.; Ashraf, C.; Damirchi, B.; Akbarian, D.; Rajabpour, S.; van Duin, A. C. T. Atomistic Scale Analysis of the Carbonization Process for C/H/O/N-Based Polymers with the ReaxFF Reactive Force Field. *The Journal of Physical Chemistry B* **2019**, *123*, 5357–5367.

[63] Chenoweth, K.; Van Duin, A. C.; Persson, P.; Cheng, M.-J.; Oxgaard, J.; Goddard Iii, W. A. Development and application of a ReaxFF reactive force field for oxidative dehydrogenation on vanadium oxide catalysts. *The Journal of Physical Chemistry C* **2008**, *112*, 14645–14654.

[64] Zhang, L.; Zybin, S. V.; Van Duin, A. C.; Dasgupta, S.; Goddard III, W. A.; Kober, E. M. Carbon cluster formation during thermal decomposition of octahydro-1, 3, 5, 7-tetranitro-1, 3, 5, 7-tetrazocine and 1, 3, 5-triamino-2, 4, 6-trinitrobenzene high explosives from ReaxFF reactive molecular dynamics simulations. *The Journal of Physical Chemistry A* **2009**, *113*, 10619–10640.

[65] Kamat, A. M.; Van Duin, A. C.; Yakovlev, A. Molecular dynamics simulations of laser-induced incandescence of soot using an extended ReaxFF reactive force field. *The Journal of Physical Chemistry A* **2010**, *114*, 12561–12572.

[66] Liu, L.; Liu, Y.; Zybin, S. V.; Sun, H.; Goddard, W. A. ReaxFF-lg: Correction of the ReaxFF Reactive Force Field for London Dispersion, with Applications to the Equations of State for Energetic Materials. *The Journal of Physical Chemistry A* **2011**, *115*, 11016–11022.

[67] Ashraf, C.; Van Duin, A. C. Extension of the ReaxFF combustion force field toward syngas combustion and initial oxidation kinetics. *The Journal of Physical Chemistry A* **2017**, *121*, 1051–1068.

[68] Zhang, W.; van Duin, A. C. T. Improvement of the ReaxFF Description for Functionalized Hydrocarbon/Water Weak Interactions in the Condensed Phase. *The Journal of Physical Chemistry B* **2018**, *122*, 4083–4092.

[69] Zhang, W.; van Duin, A. C. T. Second-Generation ReaxFF Water Force Field: Improvements in the Description of Water Density and OH-Anion Diffusion. *The Journal of Physical Chemistry B* **2017**, *121*, 6021–6032.

[70] Abolfath, R. M.; van Duin, A. C. T.; Brabec, T. Reactive Molecular Dynamics Study on the First Steps of DNA Damage by Free Hydroxyl Radicals. *The Journal of Physical Chemistry A* **2011**, *115*, 11045–11049.

[71] Monti, S.; Corozzi, A.; Fristrup, P.; Joshi, K. L.; Shin, Y. K.; Oelschlaeger, P.; Duin, A. C. T. v.; Barone, V. Exploring the conformational and reactive dynamics of biomolecules in solution using an extended version of the glycine reactive force field. *Physical Chemistry Chemical Physics* **2013**, *15*, 15062–15077.

[72] Vashisth, A.; Ashraf, C.; Zhang, W.; Bakis, C. E.; van Duin, A. C. T. Accelerated ReaxFF Simulations for Describing the Reactive Cross-Linking of Polymers. *The Journal of Physical Chemistry A* **2018**, *122*, 6633–6642.

[73] Aryanpour, M.; van Duin, A. C. T.; Kubicki, J. D. Development of a Reactive Force Field for Iron-Oxyhydroxide Systems. *The Journal of Physical Chemistry A* **2010**, *114*, 6298–6307.

[74] Zou, C.; Van Duin, A. Investigation of Complex Iron Surface Catalytic Chemistry Using the ReaxFF Reactive Force Field Method. *JOM* **2012**, *64*, 1426–1437.

[75] Fogarty, J. C.; Aktulga, H. M.; Grama, A. Y.; van Duin, A. C. T.; Pandit, S. A. A reactive molecular dynamics simulation of the silica-water interface. *The Journal of Chemical Physics* **2010**, *132*, 174704.

[76] Russo, M. F.; Li, R.; Mench, M.; van Duin, A. C. T. Molecular dynamic simulation of aluminum–water reactions using the ReaxFF reactive force field. *International Journal of Hydrogen Energy* **2011**, *36*, 5828–5835.

[77] Zou, C.; Shin, Y. K.; van Duin, A. C. T.; Fang, H.; Liu, Z.-K. Molecular dynamics simulations of the effects of vacancies on nickel self-diffusion, oxygen diffusion and oxidation initiation in nickel, using the ReaxFF reactive force field. *Acta Materialia* **2015**, *83*, 102–112.

[78] Kim, S.-Y.; van Duin, A. C.; Kubicki, J. D. Molecular dynamics simulations of the interactions between TiO2 nanoparticles and water with Na+ and Cl-, methanol, and formic acid using a reactive force field. *Journal of Materials Research* **2013**, *28*, 513–520.

[79] Bai, C.; Liu, L.; Sun, H. Molecular Dynamics Simulations of Methanol to Olefin Reactions in HZSM-5 Zeolite Using a ReaxFF Force Field. *The Journal of Physical Chemistry C* **2012**, *116*, 7029–7039.

[80] Rahaman, O.; van Duin, A. C. T.; Goddard, W. A.; Doren, D. J. Development of a ReaxFF Reactive Force Field for Glycine and Application to Solvent Effect and Tautomerization. *The Journal of Physical Chemistry B* **2011**, *115*, 249–261.

[81] Zhang, Q.; Çağın, T.; van Duin, A.; Goddard, W. A.; Qi, Y.; Hector, L. G. Adhesion and nonwetting-wetting transition in the Al / α - Al_2O_3. *Physical Review B* **2004**, *69*, 045423.

[82] Sanz-Navarro, C. F.; Åstrand, P.-O.; Chen, D.; Rønning, M.; van Duin, A. C. T.; Jacob, T.; Goddard, W. A. Molecular Dynamics Simulations of the Interactions between Platinum Clusters and Carbon Platelets. *The Journal of Physical Chemistry A* **2008**, *112*, 1392–1402.

[83] Mueller, J. E.; van Duin, A. C. T.; Goddard, W. A. Development and Validation of ReaxFF Reactive Force Field for Hydrocarbon Chemistry Catalyzed by Nickel. *The Journal of Physical Chemistry C* **2010**, *114*, 4939–4949.

[84] Cheung, S.; Deng, W.-Q.; van Duin, A. C. T.; Goddard, W. A. ReaxFFMgH Reactive Force Field for Magnesium Hydride Systems. *The Journal of Physical Chemistry A* **2005**, *109*, 851–859.

[85] Ojwang, J. G. O.; van Santen, R.; Kramer, G. J.; van Duin, A. C. T.; Goddard, W. A. Modeling the sorption dynamics of NaH using a reactive force field. *The Journal of Chemical Physics* **2008**, *128*, 164714.

[86] Senftle, T. P.; Janik, M. J.; van Duin, A. C. T. A ReaxFF Investigation of Hydride Formation in Palladium Nanoclusters via Monte Carlo and Molecular Dynamics Simulations. *The Journal of Physical Chemistry C* **2014**, *118*, 4967–4981.

[87] Koparde, V. N.; Cummings, P. T. Molecular Dynamics Simulation of Titanium Dioxide Nanoparticle Sintering. *The Journal of Physical Chemistry B* **2005**, *109*, 24280–24287.

[88] Keith, J. A.; Fantauzzi, D.; Jacob, T.; van Duin, A. C. T. Reactive forcefield for simulating gold surfaces and nanoparticles. *Physical Review B* **2010**, *81*, 235404.

[89] LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications* **2022**, *271*, 108171.

[90] te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.

[91] Mortier, W. J.; Ghosh, S. K.; Shankar, S. Electronegativity-equalization method for the calculation of atomic charges in molecules. *Journal of the American Chemical Society* **1986**, *108*, 4315–4320.

[92] Janssens, G. O.; Baekelandt, B. G.; Toufar, H.; Mortier, W. J.; Schoonheydt, R. A. Comparison of cluster and infinite crystal calculations on zeolites with the electronegativity equalization method (EEM). *The Journal of Physical Chemistry* **1995**, *99*, 3251–3258.

[93] Furman, D.; Wales, D. J. Transforming the Accuracy and Numerical Stability of ReaxFF Reactive Force Fields. *The Journal of Physical Chemistry Letters* **2019**, *10*, 7215–7223.

[94] Furman, D.; Wales, D. J. A well-behaved theoretical framework for ReaxFF reactive force fields. *The Journal of Chemical Physics* **2020**, *153*, 021102.

[95] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, 2017.

[96] Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Elsevier, 2001.

[97] Newton, I.; Chittenden, N. W. L. o. S. I. N.; Adee, D.; Motte, A.; Hill, T. P. E. A. m. b. C.-B. *Newton's Principia : the mathematical principles of natural philosophy*; New-York : Published by Daniel Adee, 1846.

[98] Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984**, *81*, 3684–3690.

[99] Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics* **1984**, *81*, 511–519.

[100] Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **1985**, *31*, 1695–1697.

[101] Sharma, S.; Kumar, P.; Chandra, R. In *Molecular Dynamics Simulation of Nanocomposites Using BIOVIA Materials Studio, Lammps and Gromacs*; Sharma, S., Ed.; Micro and Nano Technologies; Elsevier, 2019; pp 1–38.

[102] Lees, A. W.; Edwards, S. F. The computer study of transport processes under extreme conditions. *Journal of Physics C: Solid State Physics* **1972**, *5*, 1921–1928.

[103] Leonhardi Euleri, *Institutiones_calculi_integrals_volumen-Primum*; Lipsiae Et Berolini, 1768.

[104] Butcher, J. C. A history of Runge-Kutta methods. *Applied Numerical Mathematics* **1996**, *20*, 247–260.

[105] Verlet, L.; Weis, J.-J. Equilibrium Theory of Simple Liquids. *Physical Review A* **1972**, *5*, 939–952.

[106] Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *Journal of Chemical Physics* **1982**, *76*, 637–649.

[107] Birdsall, C.; Langdon, A.; Biskamp, D. Mirror-Based and Field-Reversed Approaches to Magnetic Fusion Course: 2 volumes; and Workshop: 1 volume. *Nuclear Fusion* **1985**, *25*, 643.

[108] Mansouri, A.; Cregut, M.; Abbes, C.; Durand, M.-J.; Landoulsi, A.; Thouand, G. The Environmental Issues of DDT Pollution and Bioremediation: a Multidisciplinary Review. *Appl. Biochem. Biotechnol.* **2017**, *181*, 309–339.

[109] *Zeitschrift des allgemeinen Oesterreichischen Apotheker-Vereines, Volume 39*; Im Selbstverlage des Vereines, 1901.

[110] Carson, R. *Thinking about the environment*; Routledge, 2015; pp 150–155.

[111] Programme, A. M. a. A., Jensson, H., Eds. *AMAP Assessment 2002: The influence of global change on contaminant pathways to, within, and from the Arctic*; AMAP Arctic Monitoring and Assessment Programme: Oslo, 2003.

[112] Mishra, K.; Sharma, R. C. Contamination of aquatic system by chlorinated pesticides and their spatial distribution over North-East India. *Toxicol. Environ. Health Sci.* **2011**, *3*, 144–155.

[113] Silva, V.; Mol, H. G. J.; Zomer, P.; Tienstra, M.; Ritsema, C. J.; Geissen, V. Pesticide residues in European agricultural soils – A hidden reality unfolded. *Sci. Total Environ.* **2019**, *653*, 1532–1545.

[114] Sadasivaiah, S.; Tozan, Y.; Breman, J. G. *Defining and Defeating the Intolerable Burden of Malaria III: Progress and Perspectives: Supplement to Volume 77(6) of American Journal of Tropical Medicine and Hygiene*; American Society of Tropical Medicine and Hygiene, 2007.

[115] Olson, K. R.; Morton, L. W. Long-Term Fate of Agent Orange and Dioxin TCDD Contaminated Soils and Sediments in Vietnam Hotspots. *Open Journal of Soil Science* **2019**, *9*, 1–34.

[116] Eskenazi, B.; Mocarelli, P.; Warner, M.; Needham, L.; Patterson, D. G., Jr.; Samuels, S.; Turner, W.; Gerthoux, P. M.; Brambilla, P. Relationship of serum TCDD concentrations and age at exposure of female residents of Seveso, Italy. *Environ. Health Perspect.* **2004**,

[117] IARC, *Polychlorinated Biphenyls and Polybrominated Biphenyls.*

[118] Ballschmiter, K.; Zell, M. Analysis of polychlorinated biphenyls (PCB) by glass capillary gas chromatography. *Fresenius' Zeitschrift für analytische Chemie* **1980**, *302*, 20–31.

[119] Novick, S. *Our world in peril an Environment review.*; Fawcett Publications: Greenwich Conn., 1971.

[120] Cairns, T.; Siegmund, E. G. PCBs. Regulatory history and analytical problems. *Analytical Chemistry* **1981**, *53*, 1183A–1193A.

[121] Robinson, G. K.; Lenn, M. J. The Bioremediation of Polychlorinated Biphenyls (PCBs): Problems and Perspectives. *Biotechnology and Genetic Engineering Reviews* **1994**, *12*, 139–188.

[122] Johnson, G. W.; Quensen, I., John F.; Chiarenzelli, J. R.; Hamilton, M. C. In *Environmental Forensics*; Morrison, R. D., Murphy, B. L., Eds.; Academic Press: Burlington, 2000; pp 187–225.

[123] Schwartz, L. Dermatitis from Synthetic Resins and Waxes *. *American Journal of Public Health and the Nations Health* **1936**, *26*, 586–592.

[124] Drinker, C. K.; Warren, M. F.; Bennett, G. A. The Problem of Possible Systemic Effects from Certain Chlorinated Hydrocarbons. *Journal of Industrial Hygiene and Toxicology* **1937**, *19*, 283–99.

[125] Jensen, S.; Johnels, A. G.; Olsson, M.; Otterlind, G. DDT and PCB in Marine Animals from Swedish Waters. *Nature* **1969**, *224*, 247–250.

[126] Fujiwara, K. Environmental and food contamination with PCB's in Japan. *Science of The Total Environment* **1975**, *4*, 219–247.

[127] Fischbein, A.; Wolff, M. S.; Lilis, R.; Thornton, J.; Selikoff, I. J. Clinical Findings Among Pcb-Exposed Capacitor Manufacturing Workers*. *Annals of the New York Academy of Sciences* **1979**, *320*.

[128] Warshaw, R.; Fischbein, A.; Thornton, J.; Miller, A.; Selikoff, I. J. Decrease in Vital Capacity in Pcb-Exposed Workers in a Capacitor Manufacturing Facility*. *Annals of the New York Academy of Sciences* **1979**, *320*.

[129] M.D, A. F.; Ph.D, M. S. W.; Bernstein, J.; M.D, I. J. S.; Ph.D, J. T. Dermatological Findings in Capacitor Manufacturing Workers Exposed to Dielectric Fluids Containing Polychlorinated Biphenyls (PCBs). *Archives of Environmental Health: An International Journal* **1982**, *37*, 69–74.

[130] Maroni, M.; Colombi, A.; Cantoni, S.; Ferioli, E.; Foa, V. Occupational exposure to polychlorinated biphenyls in electrical workers. I. Environmental and blood polychlorinated biphenyls concentrations. *Occupational and Environmental Medicine* **1981**, *38*, 49–54.

[131] Weltgesundheitsorganisation„ on Chemical Safety, I. P., for the Sound Management of Chemicals, I.-O. P., Eds. *Polychlorinated biphenyls: human health aspects*; Consise international chemical assessment document 55; World Health Organization: Geneva, 2003.

[132] Hornbuckle, K.; Robertson, L. Polychlorinated Biphenyls (PCBs): Sources, Exposures, Toxicities. *Environmental Science & Technology* **2010**, *44*, 2749–2751.

[133] Van den Berg M, et al. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. *Environmental Health Perspectives* **1998**, *106*, 775–792.

[134] Van den Berg, M. et al. The 2005 World Health Organization reevaluation of human and Mammalian toxic equivalency factors for dioxins and dioxin-like compounds. *Toxicological Sciences: An Official Journal of the Society of Toxicology* **2006**, *93*, 223–241.

[135] Hemming, H.; Bager, Y.; Flodström, S.; Nordgren, I.; Kronevi, T.; Ahlborg, U. G.; Wärngård, L. Liver tumour promoting activity of 3,4,5,3',4'-pentachlorobiphenyl and its interaction with 2,3,7,8-tetrachlorodibenzo-p-dioxin. *European Journal of Pharmacology: Environmental Toxicology and Pharmacology* **1995**, *292*, 241–249.

[136] Reddy, A. V. B.; Moniruzzaman, M.; Aminabhavi, T. M. Polychlorinated biphenyls (PCBs) in the environment: Recent updates on sampling, pretreatment, cleanup technologies and their analysis. *Chemical Engineering Journal* **2019**, *358*, 1186–1207.

[137] Kania-Korwel, I.; Wu, X.; Wang, K.; Lehmler, H.-J. Identification of lipidomic markers of chronic 3,3',4,4',5-pentachlorobiphenyl (PCB 126) exposure in the male rat liver. *Toxicology* **2017**, *390*, 124–134.

[138] Chen, S.-J.; Tian, M.; Zheng, J.; Zhu, Z.-C.; Luo, Y.; Luo, X.-J.; Mai, B.-X. Elevated Levels of Polychlorinated Biphenyls in Plants, Air, and Soils at an E-Waste Site in Southern China and Enantioselective Biotransformation of Chiral PCBs in Plants. *Environmental Science & Technology* **2014**, *48*, 3847–3855.

[139] Aken, B. V.; Correa, P. A.; Schnoor, J. L. Phytoremediation of Polychlorinated Biphenyls: New Trends and Promises. *Environmental Science & Technology* **2010**, *44*, 2767–2776.

[140] HellouJocelyne,; LebeufMichel,; RudiMarilynn, Review on DDT and metabolites in birds and mammals of aquatic ecosystems. *Environ. Rev.* **2012**,

[141] Eskenazi, B.; Chevrier, J.; Rosas, L. G.; Anderson, H. A.; Bornman, M. S.; Bouwman, H.; Chen, A.; Cohn, B. A.; de Jager, C.; Henshel, D. S.; Leipzig, F.; Leipzig, J. S.; Lorenz, E. C.; Snedeker, S. M.; Stapleton, D. The Pine River Statement: Human Health Consequences of DDT Use. *Environ. Health Perspect.* **2009**,

[142] Geyer, H. J.; Scheunert, I.; Rapp, K.; Gebefügi, I.; Steinberg, C.; Kettrup, A. The relevance of fat content in toxicity of lipophilic chemicals to terrestrial animals with special reference to dieldrin and 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). *Ecotoxicol. Environ. Saf.* **1993**, *26*, 45–60.

[143] Ododo, M. M.; Wabalo, B. K. Polychlorinated Biphenyls (PCBs) and Their Impacts on Human Health: A Review. *Journal of Environment Pollution and Human Health* **2019**, 6.

[144] Dvorská, A.; Šír, M.; Honzajková, Z.; Komprda, J.; Čupr, P.; Petrlík, J.; Anakhasyan, E.; Simonyan, L.; Kubal, M. Obsolete pesticide storage sites and their POP release into the environment—an Armenian case study. *Environ. Sci. Pollut. Res.* **2012**, *19*, 1944–1952.

[145] Lin, T.; Li, J.; Xu, Y.; Liu, X.; Luo, C.; Cheng, H.; Chen, Y.; Zhang, G. Organochlorine pesticides in seawater and the surrounding atmosphere of the marginal seas of China: Spatial distribution, sources and air–water exchange. *Sci. Total Environ.* **2012**, *435-436*, 244–252.

[146] Armitage, J. M.; Ginevan, M. E.; Hewitt, A.; Ross, J. H.; Watkins, D. K.; Solomon, K. R. Environmental fate and dietary exposures of humans to TCDD as a result of the spraying of Agent Orange in upland forests of Vietnam. *Sci. Total Environ.* **2015**, *506-507*, 621–630.

[147] Morton, L. W.; Culbertson, C. Persistence of Dioxin TCDD in Southern Vietnam Soil and Water Environments and Maternal Exposure Pathways with Potential Consequences on Congenital Heart Disease Prevalence in Vietnam. *Open Journal of Soil Science* **2022**, *12*, 119–150.

[148] ERICKSON, M. PCB properties, uses, occurrence, and regulatory history. *PCBs : Recent advances in environmental toxicology and health effects* **2001**,

[149] Hermanson, M. H.; Hites, R. A. Long-term measurements of atmospheric polychlorinated biphenyls in the vicinity of Superfund dumps. *Environmental Science & Technology* **1989**, *23*, 1253–1258.

[150] Shen, T.; Tofflemire, T. Air pollution aspects of land disposal of toxic wastes. *Journal of the Environmental Engineering Division, ASCE* **1980**, *106*, 211–226.

[151] Södergren, A.; Larsson, P.; Knulst, J.; Bergqvist, C. Transport of incinerated organochlorine compounds to air, water, microlayer, and organisms. *Marine Pollution Bulletin* **1990**, *21*, 18–24.

[152] Swackhamer, D. L.; Armstrong, D. E. Estimation of the atmospheric and nonatmospheric contributions and losses of polychlorinated biphenyls to Lake Michigan on the basis of sediment records of remote lakes. *Environmental Science & Technology* **1986**, *20*, 879–883.

[153] McLachlan, M.; MacKay, D.; Jones, P. H. A conceptual model of organic chemical volatilization at waterfalls. *Environmental Science & Technology* **1990**, *24*, 252–257.

[154] Verbrugge, D. A.; Giesy, J. P.; Mora, M. A.; Williams, L. L.; Rossmann, R.; Moll, R. A.; Tuchman, M. Concentrations of Dissolved and Particulate Polychlorinated Biphenyls in Water from the Saginaw River, Michigan. *Journal of Great Lakes Research* **1995**, *21*, 219–233.

[155] Eisenreich, S. J.; Looney, B. B.; Thornton, J. D. Airborne organic contaminants in the Great Lakes ecosystem. *Environmental Science & Technology* **1981**, *15*, 30–38.

[156] Leister, D. L.; Baker, J. E. Atmospheric deposition of organic contaminants to the chesapeake bay. *Atmospheric Environment* **1994**, *28*, 1499–1520.

[157] Atlas, E.; Giam, C. Ambient concentration and precipitation scavenging of atmospheric organic pollutants. *Water, Air, and Soil Pollution* **1988**, *38*, 19–36.

[158] Mackay, D. *Multimedia Environmental Models: The Fugacity Approach, Second Edition*, 2nd ed.; 2001.

[159] Bremle, G.; Okla, L.; Larsson, P. Uptake of PCBs in Fish in a Contaminated River System: Bioconcentration Factors Measured in the Field. *Environmental Science & Technology* **1995**, *29*, 2010–2015.

[160] Porte, C.; Albaigés, J. Bioaccumulation patterns of hydrocarbons and polychlorinated biphenyls in bivalves, crustaceans, and fishes. *Archives of Environmental Contamination and Toxicology* **1994**, *26*, 273–281.

[161] Saghir, A.; Koritz, G.; Hansen, L. Toxicokinetics of 2,2',4,4'- and 3,3',4,4'-Tetrachlorobiphenyl in House Flies Following Topical Administration. *Pesticide Biochemistry and Physiology* **1994**, *49*, 94–113.

[162] Göteborgs universitet, *The Nordic Expert Group for Criteria Documentation of Health Risks from Chemicals and The Dutch Expert Committee on Occupational Safety. 145, 145,*; Göteborgs universitet: Göteborg, 2011.

[163] Tang, A.; Wang, L.; Zhou, R. Gibbs energies of formation of chlorinated benzoic acids and benzoates and application to their reductive dechlorination. *Journal of Molecular Structure: THEOCHEM* **2010**, *960*, 31–39.

[164] Lohmann, R. The Emergence of Black Carbon as a Super-Sorbent in Environmental Chemistry: The End of Octanol? *Environmental Forensics* **2003**, *4*, 161–165.

[165] Arulmozhiraja, S.; Fujii, T. Torsional barrier, ionization potential, and electron affinity of biphenyl—A theoretical study. *The Journal of Chemical Physics* **2001**, *115*, 10589–10594.

[166] Katon, J.; Lippincott, E. The vibrational spectra and geometrical configuration of biphenyl. *Spectrochimica Acta* **1959**, *15*, 627–650.

[167] Carreira, L. A.; Towns, T. G. Raman spectra and barriers to internal rotation: biphenyl and nitrobenzene. *Journal of Molecular Structure* **1977**, *41*, 1–9.

[168] Almenningen, A.; Bastiansen, O.; Fernholt, L.; Cyvin, B. N.; Cyvin, S. J.; Samdal, S. Structure and barrier of internal rotation of biphenyl derivatives in the gaseous state: Part 1. The molecular structure

and normal coordinate analysis of normal biphenyl and pedeuterated biphenyl. *Journal of Molecular Structure* **1985**, *128*, 59–76.

[169] Bastiansen, O.; Samdal, S. Structure and barrier of internal rotation of biphenyl derivatives in the gaseous state: Part 4. Barrier of internal rotation in biphenyl, perdeuterated biphenyl and seven non-ortho-substituted halogen derivatives. *Journal of Molecular Structure* **1985**, *128*, 115–125.

[170] Rubio, M.; Merchán, M.; Ortí, E. The internal rotational barrier of biphenyl studied with multiconfigurational second-order perturbation theory (CASPT2). *Theoretica chimica acta* **1995**, *91*, 17–29.

[171] McKinney, J.; Singh, P. Structure-activity relationships in halogenated biphenyls: Unifying hypothesis for structural specificity. *Chemico-Biological Interactions* **1981**, *33*, 271–283.

[172] McKinney, J. D.; Gottschalk, K. E.; Pedersen, L. A theoretical investigation of the conformation of polychlorinated biphenyls (PCB's). *Journal of Molecular Structure: THEOCHEM* **1983**, *104*, 445–450.

[173] Arulmozhiraja, S.; Fujii, T.; Morita, M. Density Functional Theory Studies on Radical Ions of Selected Polychlorinated Biphenyls. *The Journal of Physical Chemistry A* **2002**, *106*, 10590–10595.

[174] Arulmozhiraja, S.; Selvin, P. C.; Fujii, T. Structures, Potential Energy Curves, and Torsional Barrier Heights for Selected Polychlorinated Biphenyls: A Density Functional Theory Study. *The Journal of Physical Chemistry A* **2002**, *106*, 1765–1769.

[175] Parthasarathi, R.; Padmanabhan, J.; Subramanian, V.; Maiti, B.; Chattaraj, P. K. Chemical Reactivity Profiles of Two Selected Polychlorinated Biphenyls. *The Journal of Physical Chemistry A* **2003**, *107*, 10346–10352.

[176] Padmanabhan, J.; Parthasarathi, R.; Subramanian, V.; Chattaraj, P. K. Chemical reactivity analysis on 33'44'55'-hexa chlorobiphenyl—A DFT approach. *Journal of Molecular Structure: THEOCHEM* **2005**, *730*, 221–226.

[177] Gorbunova, T. I.; Subbotina, J. O.; Saloutin, V. I.; Chupakhin, O. N. Reactivity of polychlorinated biphenyls in nucleophilic and electrophilic substitutions. *Journal of Hazardous Materials* **2014**, *278*, 491–499.

[178] Jensen, B.; Kuznetsova, T.; Kvamme, B.; Oterhals, Å. Molecular dynamics study of selective adsorption of PCB on activated carbon. *Fluid Phase Equilib.* **2011**, *307*, 58–65.

[179] Chana, A.; Concejero, M. A.; de Frutos, M.; González, M. J.; Herradón, B. Computational Studies on Biphenyl Derivatives. Analysis of the Conformational Mobility, Molecular Electrostatic Potential, and

Dipole Moment of Chlorinated Biphenyl: Searching for the Rationalization of the Selective Toxicity of Polychlorinated Biphenyls (PCBs). *Chem. Res. Toxicol.* **2002**, *15*, 1514–1526.

[180] Oturan, M. A.; Aaron, J.-J. Advanced Oxidation Processes in Water/Wastewater Treatment: Principles and Applications. A Review. *Critical Reviews in Environmental Science and Technology* **2014**, *44*, 2577–2641.

[181] Dominguez, C. M.; Oturan, N.; Romero, A.; Santos, A.; Oturan, M. A. Optimization of electro-Fenton process for effective degradation of organochlorine pesticide lindane. *Catal. Today* **2018**, *313*, 196–202.

[182] Brault, P.; Abraham, M.; Bensebaa, A.; Aubry, O.; Hong, D.; Rabat, H.; Magureanu, M. Insight into plasma degradation of paracetamol in water using a reactive molecular dynamics approach. *J. Appl. Phys.* **2021**, *129*, 183304.

[183] Wu, R.; Wang, S.; Wang, L. Atmospheric oxidation mechanism of chlorobenzene. *Chemosphere* **2014**, *111*, 537–544.

[184] Atkinson, R.; Aschmann, S. M.; Winer, A. M.; Pitts, J. N. Atmospheric gas phase loss processes for chlorobenzene, benzotrifluoride, and 4-chlorobenzotrifluoride, and generalization of predictive techniques for atmospheric lifetimes of aromatic compounds. *Arch. Environ. Contam. Toxicol.* **1985**, *14*, 417–425.

[185] Ajiboye, T. O.; Kuvarega, A. T.; Onwudiwe, D. C. Recent Strategies for Environmental Remediation of Organochlorine Pesticides. *Appl. Sci.* **2020**, *10*, 6286.

[186] Roos, B. O.; Taylor, P. R.; Sigbahn, P. E. M. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chemical Physics* **1980**, *48*, 157–173.

[187] Roos, B.; Lindh, R.; Malmqvist, P.; Veryazov, V.; Widmark, P.-O. *Multiconfigurational Quantum Chemistry*; Multiconfigurational Quantum Chemistry; 2016.

[188] Angeli, C.; Bories, B.; Cavallini, A.; Cimiraglia, R. Third-order multireference perturbation theory: The n-electron valence state perturbation-theory approach. *The Journal of Chemical Physics* **2006**, *124*, 054108.

[189] Neese, F. The ORCA program system. *WIREs Computational Molecular Science* **2012**, *2*, 73–78.

[190] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *The Journal of Chemical Physics* **1992**, *96*, 6796–6806.

[191] Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 energy evaluation by direct methods. *Chemical physics letters* **1988**, *153*, 503–506.

[192] Rubio-Pons, Ó.; Loboda, O.; Minaev, B.; Schimmelpfennig, B.; Vahtras, O.; Ågren, H. CASSCF calculations of triplet state properties: applications to benzene derivatives. *Mol. Phys.* **2003**, *101*, 2103–2114.

[193] Liu, Y.-J.; Persson, P.; Lunell, S. Theoretical Study of the Fast Photodissociation Channels of the Monohalobenzenes. *The Journal of Physical Chemistry A* **2004**, *108*, 2339–2345.

[194] Liu, Y.-J.; Persson, P.; Lunell, S. Multireference calculations of the phosphorescence and photodissociation of chlorobenzene. *The Journal of Chemical Physics* **2004**, *121*, 11000–11006.

[195] Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. The multi-state CASPT2 method. *Chemical physics letters* **1998**, *288*, 299–306.

[196] Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.

[197] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* **1988**, *37*, 785–789.

[198] Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics* **1980**, *58*, 1200–1211.

[199] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *The Journal of Physical Chemistry* **1994**, *98*, 11623–11627.

[200] Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li–F. *Journal of Computational Chemistry* **1983**, *4*, 294–301.

[201] Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. A complete basis set model chemistry. I. The total energies of closed-shell atoms and hydrides of the first-row elements. *The Journal of Chemical Physics* **1988**, *89*, 2193–2218.

[202] Petersson, G. A.; Al-Laham, M. A. A complete basis set model chemistry. II. Open-shell systems and the total energies of the first-row atoms. *The Journal of Chemical Physics* **1991**, *94*, 6081–6090.

[203] Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **2011**, *32*, 1456–1465.

[204] Rajzmann, M.; Wang, J.; Humbel, S. Metal-catalyzed rearrangement of allenylsulfides to furan: A theoretical mechanistic approach. *Molecular Catalysis* **2017**, *443*, 148–154.

[205] Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science* **2013**, *68*, 314–319.

[206] Shchygol, G.; Yakovlev, A.; Trnka, T.; van Duin, A. C. T.; Verstraelen, T. ReaxFF Parameter Optimization with Monte-Carlo and Evolutionary Algorithms: Guidelines and Insights. *J. Chem. Theory Comput.* **2019**, *15*, 6799–6812.

[207] Dittner, M.; Müller, J.; Aktulga, H. M.; Hartke, B. Efficient global optimization of reactive force-field parameters. *J. Comput. Chem.* **2015**, *36*, 1550–1561.

[208] Iype, E.; Hütter, M.; Jansen, A. P. J.; Nedea, S. V.; Rindt, C. C. M. Parameterization of a reactive force field using a Monte Carlo algorithm. *J. Comput. Chem.* **2013**, *34*, 1143–1154.

[209] Hansen, N.; Ostermeier, A. Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. Proceedings of IEEE international conference on evolutionary computation. 1996; pp 312–317.

[210] Hansen, N.; Müller, S. D.; Koumoutsakos, P. Reducing the Time Complexity of the Derandomized Evolution Strategy with Covariance Matrix Adaptation (CMA-ES). *Evolutionary Computation* **2003**, *11*, 1–18.

[211] Chenoweth, K.; van Duin, A. C. T.; Persson, P.; Cheng, M.-J.; Oxgaard, J.; Goddard, W. A. Development and Application of a ReaxFF Reactive Force Field for Oxidative Dehydrogenation on Vanadium Oxide Catalysts. *The Journal of Physical Chemistry C* **2008**, *112*, 14645–14654.

[212] Gittus, O. R.; Bresme, F. Thermophysical properties of water using reactive force fields. *The Journal of Chemical Physics* **2021**, *155*, 114501.

[213] Shi, L.; Ying, Z.; Xu, A.; Cheng, Y. Anomalous proton conduction behavior across a nanoporous two-dimensional conjugated aromatic polymer membrane. *Physical Chemistry Chemical Physics* **2020**, *22*, 2978–2985.

[214] Dong, D.; Zhang, W.; Barnett, A.; Lu, J.; Van Duin, A. C. T.; Molinero, V.; Bedrov, D. Multiscale Modeling of Structure, Transport and Reactivity in Alkaline Fuel Cell Membranes: Combined Coarse-Grained, Atomistic and Reactive Molecular Dynamics Simulations. *Polymers* **2018**, *10*, 1289.

[215] Ali, U.; Syed, J. H.; Malik, R. N.; Katsoyiannis, A.; Li, J.; Zhang, G.; Jones, K. C. Organochlorine pesticides (OCPs) in South Asian region: A review. *Science of The Total Environment* **2014**, *476-477*, 705–717.

[216] Zokm, G. M. E.; Ismail, M. M.; Okbah, M. A. E. Seaweed as bioindicators of organic micropollutants polycyclic aromatic hydrocarbons (PAHs) and organochlorine pesticides (OCPs). *Environmental Science and Pollution Research* **2022**,

[217] Okeagu, O. M.; Akinsanya, B.; Isibor, P. O.; Daniel-Rugu, J.; Onadeko, A. B.; Yalwaji, B.; Adekoya, K. O. Bioaccumulation of organochlorine pesticides in the parasite Cosmocerca sp. (Nematoda: Cosmocercidae) and the amphibian host Amietophrynus regularis (Reuss, 1833) within Lagos metropolis, Nigeria. *Toxicology Reports* **2022**, *9*, 136–146.

[218] Keswani, C.; Dilnashin, H.; Birla, H.; Roy, P.; Tyagi, R. K.; Singh, D.; Rajput, V. D.; Minkina, T.; Singh, S. P. Global footprints of organochlorine pesticides: a pan-global survey. *Environmental Geochemistry and Health* **2022**, *44*, 149–177.

[219] Ding, Y.; Huang, H.; Chen, W.; Zhang, Y.; Chen, W.; Xing, X.; Qi, S. Background levels of OCPs, PCBs, and PAHs in soils from the eastern Pamirs, China, an alpine region influenced by westerly atmospheric transport. *Journal of Environmental Sciences* **2022**, *115*, 453–464.

[220] Chen, W.; Peng, B.; Huang, H.; Kuang, Y.; Qian, Z.; Zhu, W.; Liu, W.; Zhang, Y.; Liao, Y.; Zhao, X.; Zhou, H.; Qi, S. Distribution and Potential Sources of OCPs and PAHs in Waters from the Danshui River Basin in Yichang, China. *International Journal of Environmental Research and Public Health* **2022**, *19*, 263.

[221] Veludo, A. F.; Martins Figueiredo, D.; Degrendele, C.; Masinyana, L.; Curchod, L.; Kohoutek, J.; Kukučka, P.; Martiník, J.; Přibylová, P.; Klánová, J.; Dalvie, M. A.; Röösli, M.; Fuhrimann, S. Seasonal variations in air concentrations of 27 organochlorine pesticides (OCPs) and 25 current-use pesticides (CUPs) across three agricultural areas of South Africa. *Chemosphere* **2022**, *289*, 133162.

[222] Idowu, G. A.; Aiyesanmi, A. F.; Oyegoke, F. O. Organochlorine pesticide residues in pods and beans of cocoa (Theobroma cacao L.) from Ondo State Central District, Nigeria. *Environmental Advances* **2022**, *7*, 100162.

[223] AMAP Assessment 2002: The Influence of Global Change on Contaminant Pathways to, within, and from the Arctic.

[224] UNEP, Report of the Conference of the Parties to the Stockholm Convention on Persistent Organic Pollutants on the work of its fifth meeting. **2011**,

[225] Singh, Z. Toxic Effects of Organochlorine Pesticides: A Review. *American Journal of BioScience* **2016**, *4*, 11.

[226] Jayaraj, R.; Megha, P.; Sreedev, P. Review Article. Organochlorine pesticides, their toxic effects on living organisms and their fate in the environment. *Interdisciplinary Toxicology* **2016**, *9*, 90–100.

[227] Chen, S.; Han, Z.; Zhang, D.; Zhan, J. Theoretical study of the inclusion complexation of TCDD with cucurbit[n]urils. *RSC Advances* **2014**, *4*, 52415–52422.

[228] Ali, I.; Alharbi, O. M. L.; Alothman, Z. A.; Alwarthan, A. Enantio-selective molecular dynamics of (±)-o,p-DDT uptake and degradation in water-sediment system. *Environmental Research* **2018**, *160*, 353–357.

[229] Bai, J.; Yang, W.; Zhang, C.; Zhao, Y.; Gong, C.; Sun, X.; Zhang, Q.; Wang, W. Theoretical study on the OH-initiated atmospheric reaction of 1,1,1-trichloro-2,2-bis(4-chlorophenyl) ethane (DDT). *Atmospheric Environment* **2013**, *67*, 177–183.

[230] Kao, C. M.; Chen, S. C.; Liu, J. K.; Wu, M. J. Evaluation of TCDD biodegradability under different redox conditions. *Chemosphere* **2001**, *44*, 1447–1454.

[231] Wang, W.; Wang, Y.; Feng, W.; Wang, W.; Li, P. Theoretical Investigations on the Reactivity of Hydrogen Peroxide toward 2,3,7,8-Tetrachlorodibenzo-p-dioxin. *Molecules* **2018**, *23*, 2826.

[232] Wen, Z.; Wang, Z.; Xu, J.; Liu, Y.; Cen, K. Quantum chemistry study on the destruction mechanism of 2,3,7,8-TCDD by OH and O3 radicals. *Chemosphere* **2013**, *92*, 293–298.

[233] Rashidian, N.; Zahedi, E.; Shiroudi, A. Kinetic and mechanistic insight into the OH–initiated atmospheric oxidation of 2,3,7,8-tetrachlorodibenzo-p-dioxin via OH–addition and hydrogen abstraction pathways: A theoretical investigation. *Science of The Total Environment* **2019**, *679*, 106–114.

[234] Liao, Z.; Zeng, M.; Wang, L. Atmospheric oxidation mechansim of polychlorinated biphenyls (PCBs) initiated by OH radicals. *Chemosphere* **2020**, *240*, 124756.

[235] Hur, J.; N. Abousleiman, Y.; L. Hull, K.; Qomi, M. J. A. Reactive force fields for modeling oxidative degradation of organic matter in geological formations. *RSC Advances* **2021**, *11*, 29298–29307.

[236] Frisch, M. J. et al. Gaussian~16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.

[237] Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 energy evaluation by direct methods. *Chemical Physics Letters* **1988**, *153*, 503–506.

[238] Sæbø, S.; Almlöf, J. Avoiding the integral storage bottleneck in LCAO calculations of electron correlation. *Chemical Physics Letters* **1989**, *154*, 83–89.

[239] Frisch, M. J.; Head-Gordon, M.; Pople, J. A. Semi-direct algorithms for the MP2 energy and gradient. *Chemical Physics Letters* **1990**, *166*, 281–289.

[240] Frisch, M. J.; Head-Gordon, M.; Pople, J. A. A direct MP2 gradient method. *Chemical Physics Letters* **1990**, *166*, 275–280.

[241] Head-Gordon, M.; Head-Gordon, T. Analytic MP2 frequencies without fifth-order storage. Theory and application to bifurcated hydrogen bonds in the water hexamer. *Chemical Physics Letters* **1994**, *220*, 122–128.

[242] Dorofeeva, O. V.; Novikov, V. P.; Moiseeva, N. F.; Yungman, V. S. Density functional theory study of conformations, barriers to internal rotations and torsional potentials of polychlorinated biphenyls. *Journal of Molecular Structure: THEOCHEM* **2003**, *637*, 137–153.

[243] Wu, R.; Wang, S.; Wang, L. Atmospheric oxidation mechanism of chlorobenzene. *Chemosphere* **2014**, *111*, 537–544.

[244] Bryukov, M. G.; Knyazev, V. D.; Gehling, W. M.; Dellinger, B. Kinetics of the Gas-Phase Reaction of OH with Chlorobenzene. *The Journal of Physical Chemistry A* **2009**, *113*, 10452–10459.

[245] Priestley, M. et al. Chemical characterisation of benzene oxidation products under high- and low-NO_x conditions using chemical ionisation mass spectrometry. *Atmospheric Chemistry and Physics* **2021**, *21*, 3473–3490.

[246] Knyazev, V. D. Computational Study of the Reactions of H Atoms with Chlorinated Alkanes. Isodesmic Reactions for Transition States. *The Journal of Physical Chemistry A* **2002**, *106*, 11603–11615.

[247] Knyazev, V. D. Isodesmic Reactions for Transition States: Reactions of Cl Atoms with Methane and Halogenated Methanes. *The Journal of Physical Chemistry A* **2003**, *107*, 11082–11091.

[248] Liu, Y.; Jiang, X. Plasma-induced Degradation of Chlorobenzene in Aqueous Solution. *Plasma Chem. Plasma Process.* **2008**, *28*, 15–24.

[249] Chenoweth, K.; van Duin, A. C. T.; Goddard, W. A. ReaxFF Reactive Force Field for Molecular Dynamics Simulations of Hydrocarbon Oxidation. *The Journal of Physical Chemistry A* **2008**, *112*, 1040–1053.

[250] Ashraf, C.; van Duin, A. C. Extension of the ReaxFF Combustion Force Field toward Syngas Combustion and Initial Oxidation Kinetics. *The Journal of Physical Chemistry A* **2017**, *121*, 1051–1068, Publisher: American Chemical Society.

[251] Geyer, R.; Jambeck, J. R.; Law, K. L. Production, use, and fate of all plastics ever made. *Science Advances* **2017**, *3*, e1700782, Publisher: American Association for the Advancement of Science.

[252] Qureshi, M. S.; Oasmaa, A.; Pihkola, H.; Deviatkin, I.; Tenhunen, A.; Mannila, J.; Minkkinen, H.; Pohjakallio, M.; Laine-Ylijoki, J. Pyrolysis of plastic waste: Opportunities and challenges. *Journal of Analytical and Applied Pyrolysis* **2020**, *152*, 104804.

[253] Dümichen, E.; Eisentraut, P.; Bannick, C. G.; Barthel, A.-K.; Senz, R.; Braun, U. Fast identification of microplastics in complex environmental samples by a thermal degradation method. *Chemosphere* **2017**, *174*, 572–584.

[254] Chae, Y.; An, Y.-J. Current research trends on plastic pollution and ecological impacts on the soil ecosystem: A review. *Environmental Pollution* **2018**, *240*, 387–395.

[255] Williams, P. T.; Williams, E. A. Interaction of Plastics in Mixed-Plastics Pyrolysis. *Energy & Fuels* **1999**, *13*, 188–196, Publisher: American Chemical Society.

[256] Dodbiba, G.; Takahashi, K.; Sadaki, J.; Fujita, T. The recycling of plastic wastes from discarded TV sets: comparing energy recovery with mechanical recycling in the context of life cycle assessment. *Journal of Cleaner Production* **2008**, *16*, 458–470.

[257] Sadat-Shojai, M.; Bakhshandeh, G.-R. Recycling of PVC wastes. *Polymer Degradation and Stability* **2011**, *96*, 404–415.

[258] Li, C. T.; Zhuang, H. K.; Hsieh, L. T.; Lee, W. J.; Tsao, M. C. PAH emission from the incineration of three plastic wastes. *Environment international* **2001**, *27*, 61–67.

[259] Onwudili, J. A.; Insura, N.; Williams, P. T. Composition of products from the pyrolysis of polyethylene and polystyrene in a closed batch reactor: Effects of temperature and residence time. *Journal of Analytical and Applied Pyrolysis* **2009**, *86*, 293–303.

[260] Kaminsky, W.; Hartmann, F. New Pathways in Plastics Recycling. 3.

[261] Sarker, M.; Rashid, M. M. Waste Plastics Mixture of Polystyrene and Polypropylene into Light Grade Fuel using Fe2O3 Catalyst. **2013**, *2*, 12.

[262] Bai, B.; Wang, W.; Jin, H. Experimental study on gasification performance of polypropylene (PP) plastics in supercritical water. *Energy* **2020**, *191*, 116527.

[263] Bhaskar, T.; Kaneko, J.; Muto, A.; Sakata, Y.; Jakab, E.; Matsui, T.; Uddin, M. Pyrolysis studies of PP/PE/PS/PVC/HIPS-Br plastics mixed with PET and dehalogenation (Br, Cl) of the liquid products. *Journal of Analytical and Applied Pyrolysis* **2004**, *72*, 27–33.

[264] Wong, S. L.; Ngadi, N.; Abdullah, T. A. T.; Inuwa, I. M. Current state and future prospects of plastic waste as source of fuel: A review. *Renewable and Sustainable Energy Reviews* **2015**, *50*, 1167–1180.

[265] Kunwar, B.; Cheng, H. N.; Chandrashekaran, S. R.; Sharma, B. K. Plastics to fuel: a review. *Renewable and Sustainable Energy Reviews* **2016**, *54*, 421–428.

[266] Gunasee, S. D.; Danon, B.; Görgens, J. F.; Mohee, R. Co-pyrolysis of LDPE and cellulose: Synergies during devolatilization and condensation. *Journal of Analytical and Applied Pyrolysis* **2017**, *126*, 307–314.

[267] Duemichen, E.; Eisentraut, P.; Celina, M.; Braun, U. Automated thermal extraction-desorption gas chromatography mass spectrometry: A multifunctional tool for comprehensive characterization of polymers and their degradation products. *Journal of Chromatography A* **2019**, *1592*, 133–142.

[268] Honus, S.; Kumagai, S.; Molnár, V.; Fedorko, G.; Yoshioka, T. Pyrolysis gases produced from individual and mixed PE, PP, PS, PVC, and PET—Part II: Fuel characteristics. *Fuel* **2018**, *221*, 361–373.

[269] Abbas-Abadi, M. S. The effect of process and structural parameters on the stability, thermo-mechanical and thermal degradation of polymers with hydrocarbon skeleton containing PE, PP, PS, PVC, NR, PBR and SBR. *Journal of Thermal Analysis and Calorimetry* **2021**, *143*, 2867–2882.

[270] Wang, Y.; Li, Y.; Zhang, C.; Yang, L.; Fan, X.; Chu, L. A study on co-pyrolysis mechanisms of biomass and polyethylene via ReaxFF molecular dynamic simulation and density functional theory. *Process Safety and Environmental Protection* **2021**, *150*, 22–35.

[271] Hong, D.; Li, P.; Si, T.; Guo, X. ReaxFF simulations of the synergistic effect mechanisms during co-pyrolysis of coal and polyethylene/polystyrene. *Energy* **2021**, *218*, 119553.

[272] He, X.-c.; Chen, D.-z. ReaxFF MD study on the early stage co-pyrolysis of mixed PE/PP/PS plastic waste. *Journal of Fuel Chemistry and Technology* **2022**, *50*, 346–356.

[273] Guo, W.; Fan, K.; Guo, G.; Wang, J. Atomic-scale insight into thermal decomposition behavior of polypropylene: A ReaxFF method. *Polymer Degradation and Stability* **2022**, *202*, 110038.

[274] Vin, N.; Battin-Leclerc, F.; Le Gall, H.; Sebbar, N.; Bockhorn, H.; Trimis, D.; Herbinet, O. A study of chlorobenzene pyrolysis. *Proceedings of the Combustion Institute* **2019**, *37*, 399–407.

[275] Peng, Z.; Shao, M.; Yu, M.; Lu, B.; Zhao, X.; Li, P.; Song, S.; Liu, Q.; Zhou, T.; Zhang, Q. Calculation and Experimental Validation of a Novel Approach Using Solubility Parameters as Indicators for the Extraction of Additives in Plastics. *Analytical Chemistry* **2021**, *93*, 14837–14843, Publisher: American Chemical Society.

[276] Prasad, M. N. V. In *Industrial and Municipal Sludge*; Prasad, M. N. V., de Campos Favas, P. J., Vithanage, M., Mohan, S. V., Eds.; Butterworth-Heinemann, 2019; pp xxv–xxxi.

[277] Wolf, M.; Bégué, D.; Salvato Vallverdu, G. Development of a novel ReaxFF reactive potential for organochloride molecules. *The Journal of Chemical Physics* **0**, *0*, null.

[278] Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

[279] Hong, D.; Gao, P.; Wang, C. A comprehensive understanding of the synergistic effect during co-pyrolysis of polyvinyl chloride (PVC) and coal. *Energy* **2022**, *239*, 122258.

[280] Zhong, Q.; Zhang, Y.; Shabnam, S.; Mao, Q.; Xiao, J.; van Duin, A. C.; Mathews, J. P. ReaxFF MD simulations of petroleum coke CO2 gasification examining the S/N removal mechanisms and CO/CO2 reactivity. *Fuel* **2019**, *257*.

[281] Zhang, Z.; Guo, L.; Zhang, H.; Zhan, J.-H. Comparing product distribution and desulfurization during direct pyrolysis and hydropyrolysis of Longkou oil shale kerogen using reactive MD simulations. *International Journal of Hydrogen Energy* **2019**, *44*, 25335–25346.

[282] Matsuzawa, Y.; Ayabe, M.; Nishino, J.; Kubota, N.; Motegi, M. Evaluation of char fuel ratio in municipal pyrolysis waste. *Fuel* **2004**, *83*, 1675–1687.

[283] Maafa, I. M. Pyrolysis of Polystyrene Waste: A Review. *Polymers* **2021**, *13*, 225, Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

[284] Anaconda Software Distribution. 2020; `https://docs.anaconda.com/`.

[285] Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* **2007**, *9*, 90–95.

[286] Harris, C. R. et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362.

[287] pandas development team, T. pandas-dev/pandas: Pandas. 2020; `https://doi.org/10.5281/zenodo.3509134`.

[288] Shipingana, L. N. N.; Shivaraju, H. P.; Yashas, S. R. Quantitative assessment of pharmaceutical drugs in a municipal wastewater and overview of associated risks. *Appl. Water Sci.* **2022**, *12*, 1–9.

[289] Bai, J.; Yang, W.; Zhang, C.; Zhao, Y.; Gong, C.; Sun, X.; Zhang, Q.; Wang, W. Theoretical study on the OH-initiated atmospheric reaction of 1,1,1-trichloro-2,2-bis(4-chlorophenyl) ethane (DDT). *Atmos. Environ.* **2013**, *67*, 177–183.

[290] Ma, F.; Zhu, Y.; Wu, B.; Zhang, Q.; Xu, D.; Xu, J.; Wang, B.; Gu, Q.; Li, F. Degradation of DDTs in thermal desorption off-gas by pulsed corona discharge plasma. *Chemosphere* **2019**, *233*, 913–919.

[291] Agopcan Cinar, S.; Ziylan-Yavaş, A.; Catak, S.; Ince, N. H.; Aviyente, V. Hydroxyl radical-mediated degradation of diclofenac revisited: a computational approach to assessment of reaction mechanisms and by-products. *Environ. Sci. Pollut. Res.* **2017**, *24*, 18458–18469.

[292] Homlok, R.; Takács, E.; Wojnárovits, L. Elimination of diclofenac from water using irradiation technology. *Chemosphere* **2011**, *85*, 603–608.

[293] Monteagudo, J. M.; El-taliawy, H.; Durán, A.; Caro, G.; Bester, K. Sono-activated persulfate oxidation of diclofenac: Degradation, kinetics, pathway and contribution of the different radicals involved. *Journal of Hazardous Materials* **2018**, *357*, 457–465.

[294] Arulmozhiraja, S.; Fujii, T.; Sato, G. Density functional theory-based reactivity descriptors for dioxins. *Mol. Phys.* **2002**, *100*, 423–431.

[295] Wen, Z.; Wang, Z.; Xu, J.; Liu, Y.; Cen, K. Quantum chemistry study on the destruction mechanism of 2,3,7,8-TCDD by OH and O3 radicals. *Chemosphere* **2013**, *92*, 293–298.

[296] Jing, R.; Fusi, S.; Kjellerup, B. V. Remediation of Polychlorinated Biphenyls (PCBs) in Contaminated Soils and Sediment: State of Knowledge and Perspectives. *Frontiers in Environmental Science* **2018**, *6*.

[297] Harvey, P. J.; Campanella, B. F.; Castro, P. M. L.; Harms, H.; Lichtfouse, E.; Schäffner, A. R.; Smrcek, S.; Werck-Reichhart, D. Phytoremediation of polyaromatic hydrocarbons, anilines and phenols. *Environmental Science and Pollution Research* **2002**, *9*, 29–47.

[298] Pivetz, B. E. *Phytoremediation of Contaminated Soil and Ground Water at Hazardous Waste Sites*; U.S. Environmental Protection Agency, Office of Research and Development, Office of Solid Waste and Emergency Response, 2001.

[299] Gao, J.; Garrison, A. W.; Hoehamer, C.; Mazur, C. S.; Wolfe, N. L. Uptake and Phytotransformation of o,p‘-DDT and p,p‘-DDT by Axenically Cultivated Aquatic Plants. *J. Agric. Food Chem.* **2000**, *48*, 6121–6127.

[300] Hanano, A.; Almousally, I.; Shaban, M.; Moursel, N.; Shahadeh, A.; Alhajji, E. Differential tissue accumulation of 2,3,7,8-Tetrachlorinated dibenzo-p-dioxin in Arabidopsis thaliana affects plant chronology, lipid metabolism and seed yield. *BMC Plant Biol.* **2015**, *15*, 1–13.

[301] Field, J. A.; Sierra-Alvarez, R. Microbial degradation of chlorinated dioxins. *Chemosphere* **2008**, *71*, 1005–1018.

[302] Zhu, Z.-q.; Yang, X.-e.; Wang, K.; Huang, H.-g.; Zhang, X.; Fang, H.; Li, T.-q.; Alva, A. K.; He, Z.-l. Bioremediation of Cd-DDT co-contaminated soil using the Cd-hyperaccumulator Sedum alfredii and DDT-degrading microbes. *J. Hazard. Mater.* **2012**, *235-236*, 144–151.

[303] Nah, I. W.; Hwang, K.-Y.; Shul, Y.-G. Effect of metal and glycol on mechanochemical dechlorination of polychlorinated biphenyls (PCBs). *Chemosphere* **2008**, *73*, 138–141.

[304] Abbas, T.; Wadhawan, T.; Khan, A.; McEvoy, J.; Khan, E. Iron turning waste: Low cost and sustainable permeable reactive barrier media for remediating dieldrin, endrin, DDT and lindane in groundwater. *Environ. Pollut.* **2021**, *289*, 117825.

[305] Choi, Y.; Cho, Y.-M.; Luthy, R. G. In Situ Sequestration of Hydrophobic Organic Contaminants in Sediments under Stagnant Contact with Activated Carbon. 1. Column Studies. *Environmental Science & Technology* **2014**, *48*, 1835–1842.

[306] Marulanda, V.; Bolaños, G. Supercritical water oxidation of a heavily PCB-contaminated mineral transformer oil: Laboratory-scale data and economic assessment. *The Journal of Supercritical Fluids* **2010**, *54*, 258–265.

[307] Pistone, L.; Trifirò, F.; Miertus, S. DESTRUCTION TECHNOLOGIES FOR POLYCHLORINATED BIPHENYLS. 2000.

[308] Weber, R.; Yoshida, S.; Miwa, K. PCB Destruction in Subcritical and Supercritical Water Evaluation of PCDF Formation and Initial Steps of Degradation Mechanisms. *Environmental Science & Technology* **2002**, *36*, 1839–1844.

[309] Marrone, P. A.; Hodes, M.; Smith, K. A.; Tester, J. W. Salt precipitation and scale control in supercritical water oxidation—part B: commercial/full-scale applications. *The Journal of Supercritical Fluids* **2004**, *29*, 289–312.

[310] Zainal, S.; Onwudili, J. A.; Williams, P. T. Supercritical water oxidation of dioxins and furans in waste incinerator fly ash, sewage sludge and industrial soil. *Environ. Technol.* **2014**, *35*, 1823–1830.

[311] de Souza, G. B. M.; Pereira, M. B.; Mourão, L. C.; dos Santos, M. P.; de Oliveira, J. A.; Garde, I. A. A.; Alonso, C. G.; Jegatheesan, V.; Cardozo-Filho, L. Supercritical water technology: an emerging treatment process for contaminated wastewaters and sludge. *Rev. Environ. Sci. Biotechnol.* **2022**, *21*, 75–104.

[312] Lu, J.-y.; Du, X.; Lipscomb, G. Cleaning membranes with focused ultrasound beams for drinking water treatment. 2009 IEEE International Ultrasonics Symposium. 2009; pp 1195–1198.

[313] Doosti, M.; Kargar, R.; Sayadi, M. Water treatment using ultrasonic assistance: A review. *Proceedings of the international academy of ecology and environmental sciences* **2012**, *2*, 96.

[314] Gomes, H. I.; Dias-Ferreira, C.; Ribeiro, A. B. Overview of in situ and ex situ remediation technologies for PCB-contaminated soils and sediments and obstacles for full-scale application. *Science of The Total Environment* **2013**, *445-446*, 237–260.

[315] Patel, U. D.; Suresh, S. Dechlorination of chlorophenols using magnesium–palladium bimetallic system. *Journal of Hazardous Materials* **2007**, *147*, 431–438.

[316] Wu, B.-Z.; Chen, H.-Y.; Wang, S. J.; Wai, C. M.; Liao, W.; Chiu, K. Reductive dechlorination for remediation of polychlorinated biphenyls. *Chemosphere* **2012**, *88*, 757–768.

[317] Wang, Z.; Huang, W.; Peng, P.; Fennell, D. E. Rapid dechlorination of 1,2,3,4-TCDD by Ag/Fe bimetallic particles. *Chem. Eng. J.* **2015**, *273*, 465–471.

[318] Ulucan-Altuntas, K.; Debik, E. Dechlorination of dichlorodiphenyltrichloroethane (DDT) by Fe/Pd bimetallic nanoparticles: Comparison with nZVI, degradation mechanism, and pathways. *Front. Environ. Sci. Eng.* **2020**, *14*, 1–13.

[319] Gardner, K. In-situ Dechlorination of Polychlorinated Biphenyls in Sediments Using Zero-Valent Iron. **2004**, 28.

[320] Poursaberi, T.; Konoz, E.; Sarrafi, A. M.; Hassanisadi, M.; Hajifathli, F. Application of nanoscale zero-valent iron in the remediation of DDT from contaminated water. *Chemical Science Transactions* **2012**, *1*, 658–668.

[321] Xue, W.; Huang, D.; Zeng, G.; Wan, J.; Cheng, M.; Zhang, C.; Hu, C.; Li, J. Performance and toxicity assessment of nanoscale zero valent iron particles in the remediation of contaminated soil: A review. *Chemosphere* **2018**, *210*, 1145–1156.

[322] Dussert, B. W.; Tramposch, W. G. Impact of Support Media on the Biological Treatment of Ozonated Drinking Water. *Ozone: Science & Engineering* **1997**, *19*, 97–108.

[323] Abromaitis, V.; Racys, V.; van der Marel, P.; Meulepas, R. J. W. Biodegradation of persistent organics can overcome adsorption–desorption hysteresis in biological activated carbon systems. *Chemosphere* **2016**, *149*, 183–189.

[324] Acar, Y. B.; Alshawabkeh, A. N. Principles of electrokinetic remediation. *Environmental Science & Technology* **1993**, *27*, 2638–2647.

[325] Chun, C. L.; Payne, R. B.; Sowers, K. R.; May, H. D. Electrical stimulation of microbial PCB degradation in sediment. *Water Research* **2013**, *47*, 141–152.

[326] Ni, M.; Tian, S.; Huang, Q.; Yang, Y. Electrokinetic-Fenton remediation of organochlorine pesticides from historically polluted soil. *Environ. Sci. Pollut. Res.* **2018**, *25*, 12159–12168.

[327] Liou, Y. H.; Lo, S.-L.; Lin, C.-J.; Kuan, W. H.; Weng, S. C. Chemical reduction of an unbuffered nitrate solution using catalyzed and uncatalyzed nanoscale iron particles. *Journal of Hazardous Materials* **2005**, *127*, 102–110.

[328] Clark, C. J.; Rao, P. S. C.; Annable, M. D. Degradation of perchloroethylene in cosolvent solutions by zero-valent iron. *Journal of Hazardous Materials* **2003**, *96*, 65–78.

[329] Tian, H.; Li, J.; Mu, Z.; Li, L.; Hao, Z. Effect of pH on DDT degradation in aqueous solution using bimetallic Ni/Fe nanoparticles. *Sep. Purif. Technol.* **2009**, *66*, 84–89.

[330] Kulkarni, P. S.; Crespo, J. G.; Afonso, C. A. M. Dioxins sources and current remediation technologies — A review. *Environment International* **2008**, *34*, 139–153.

Résumé

Cette thèse présente le développement d'un nouveau champ de force réactif pour étudier la réactivité de composés organochlorés dans divers environnements ou conditions tels que les solutions aqueuses, les conditions atmosphériques et la pyrolyse. Ces composés organochlorés sont d'un grand intérêt en raison de leurs longues demi-vies dans les milieux naturels et de leurs toxicités bien connues. Les champs de force réactifs sont une méthodologie intermédiaire entre la mécanique moléculaire et de la mécanique quantique. Ils permettent d'étudier la réactivité de systèmes complexes et de calculer des grandeurs macroscopiques en mettant en œuvre des simulations de dynamique moléculaire réactives. Parmi les potentiels réactifs existants, ReaxFF est aujourd'hui largement utilisé mais n'inclut pas l'atome de chlore. Après un premier chapitre présentant les concepts de base des simulations de dynamiques moléculaires, un deuxième chapitre présente les familles de composés organochlorés considérées dans cette thèse et un troisième chapitre présente la méthodologie utilisée pour développer le champ de force réactif pour les organochlorés, ainsi que l'évaluation de ses performances, en comparant les résultats obtenus avec des valeurs expérimentales et théoriques issues de cette thèse ou de la littérature. Ensuite, dans le chapitre 4, le nouveau champ de force a été appliqué à l'étude des simulations de pyrolyses par simulations moléculaires réactives de trois plastiques différents : PE, PP et PS, incluant du toluène et du chlorobenzène comme additifs. Enfin, le dernier chapitre concerne une librairie python, appelée SnookRMol, mettant en œuvre des simulations réactives de collisions entre des molécules d'intérêt telles que des molécules organochlorées avec de petites molécules telles que les radicaux hydroxyles. Les capacités du logiciel sont présentées et comment il peut être utilisé pour sonder la réactivité d'une molécule.

Abstract

This thesis presents the development of a novel reactive force field designed to study the reactivity of organochlorine compounds in various environments or conditions such as aqueous solutions, atmospheric reactivity and pyrolysis. Such organochloride compounds are highly interesting because of their long half-life in the environment and their well-known toxicity. Reactive force fields bridge the gap between molecular mechanics and quantum mechanics methodologies that allow the investigation of the reactivity of complex systems and compute macroscopic quantities by implementing reactive molecular dynamics simulations. Among the available reactive potentials, ReaxFF is nowadays widely used but lacks to include chlorine atom. After a first chapter that introduces the concepts of molecular dynamics simulations, a second chapter introduces

organochlorides compounds families considered in this thesis, and a third chapter presents the methodology used to develop this new reactive force field for organochlorides, and describing the new force field created along with its performance by comparing results obtained with the new force field to both experimental and theoretical values from this work or the literature. Then, in chapter 4, the new force field was applied to the investigation of pyrolysis simulations by reactive molecular simulations of three different plastic PE, PP and PS, including toluene and chlorobenzene as additives. Finally, the last chapter concerns a python framework named SnookRMol, implementing reactive simulations of collisions between molecules of interest, such as organochlorides, with small molecules such as hydroxyl radicals. The software's capabilities are presented, and how it can be used to probe the reactivity of a molecule.

ÉCOLE DOCTORALE :

École doctorale des sciences exactes et leurs applications

LABORATOIRE :

Institut des Sciences Analytiques et de Physico-Chimie Pour l'Environnement et les Matériaux

UMR 5254 — Équipe de Chimie Physique

Matthieu Wolf

matthieu-wolf@laposte.net

Université de Pau et des Pays de l'Adour

Avenue de l'Université

BP 576

64012 Pau Cedex

www.univ-pau.fr