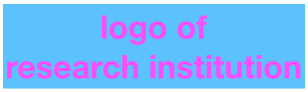
Structure, Solvation, Thermodynamic and Fragmentation of Molecular Clusters



Linjie Zheng

Supervised by
Jérôme Cuny & Mathias Rapacioli

Laboratoire de Chimie et Physique Quantiques

Université de Toulouse 3 Paul Sabatier

A thesis submitted for the degree of

PhD

month 2021

1. Reviewer: Name

2. Reviewer:

Day of the defense:

Signature from head of PhD committee:

•

Abstract

This thesis aims at studying in details the behavior of complex molecular clusters and focuses on two main aspects. First, the description of low-energy isomers of ammonium/ammonia water clusters and protonated uracil water clusters through an extensive exploration of potential energy surfaces (PES) using a combination of global and local optimization schemes. Structural, solvation and thermodynamics properties of the newly identified low-energy isomers were characterized. Second, the dynamical simulations of collision-induced dissociation of protonated uracil water clusters and pyrene dimer cation were carried out to explore collision trajectories, dissociation mechanism, energy partition, mass spectra, and collision cross-sections to complement experimental measurements conducted on these species.

Global optimization of $(H_2O)_{1-10}NH_4^+$ and $(H_2O)_{1-10}NH_3$ clusters is conducted at the self-consistent-charge density-functional based tight-binding (SCC-DFTB) level of theory, for which improved N-H parameters are proposed, in combination with the parallel-tempering molecular dynamics (PTMD) approach. Low-energy isomers of $(H_2O)_{1-10}NH_4^+$ and $(H_2O)_{1-10}NH_3$ are further optimized at MP2 level in order to evaluate the reliability of our modified N-H parameters. Both structures and binding energies obtained with SCC-DFTB are in line with the results at MP2/Def2TZVP level, which demonstrates the ability of SCC-DFTB to describe the PES of molecular species and represents a first step towards the modeling of complex aggregates of atmospheric interest.

Focus on protonated uracil water clusters aims at providing a detailed description of recent collision-induced dissociation (CID) experiments. First, stable isomers of $(H_2O)_{1-7,11,12}UH^+$ are calculated using the same methodology as described above. Then, dynamical simulations of the collisions between various $(H_2O)_{1-7,11,12}UH^+$ isomers and argon is conducted at a constant collision energy at the SCC-DFTB level. Simulated proportion of formed neutral vs. protonated

uracil containing clusters, fragmentation cross-section as well as mass spectra are consistent with the experimental data which highlights the accuracy of our simulations. They allow to probe which fragments are formed on the short time scale and rationalize the location of the excess proton on these fragments.This latter property is highly influenced by the nature of the aggregate undergoing the collision. Analyses of proportion of time-dependent fragments and mass spectra demonstrate that, up to 7 water molecules, a shattering mechanism occurs after collision whereas for n=11,12 a statistical mechanism is more likely to participate. These simulations appear as a useful tool to complement CID experiments of hydrated molecular species.

Dynamical simulation of CID experiments of pyrene dimer cation for different collision energies between 2.5 and 30 eV is also presented. The dynamical simulations allow to understand the dissociation processes. The agreement between the simulated and measured mass spectra suggests that the main processes are captured by this approach. It appears that most of the dissociation occurs on a short timescale (less than 3 ps). At low collision energies, the dissociation cross-section increases with collision energies whereas it remains almost constant for collision energies greater than 10 to 15 eV. Analysis of the kinetic energy partition is used to get insights into the collision/dissociation processes at the atomic scale. The simulated time of flight mass spectra of the parent and dissociated products are obtained from the combination of molecular dynamics simulations and phase space theory to address the short and long timescales dissociation, respectively.

Keywords: SCC-DFTB, CID, molecular dynamics, ammonium/ammonia water clusters, uracil water clusters,

Résumé de la thèse

Cette thèse vise à étudier en détails le comportement d'agrégats moléculaires complexes et se concentre sur deux aspects principaux. Tout d'abord, la description des isomères de faible énergie des clusters d'ammonium et ammoniac et des clusters d'uracile/eau protonés à travers l'exploration des surfaces d'énergie potentielle (PES) en utilisant une combinaison d'approches d'optimisation globales et locales. Les propriétés structurelles, de solvatation et thermodynamiques des isomères de basse énergie nouvellement identifiés ont été caractérisées. Par la suite, des simulations dynamiques de la dissociation induite par collision des agrégats d'uracile/eau protonés et du dimère de pyrène ont été réalisées et analysées en termes de : mécanisme de dissociation, répartition d'énergie, spectres de masse et sections efficaces de collision pour complémenter des mesures expérimentales récentes menées sur ces espèces.

L'optimisation globale des clusters $(H_2O)_{1-10}NH_4^+$ et $(H_2O)_{1-10}NH_3$ a été réalisée au niveau de théorie SCC-DFTB (pour self-consistent-charge density-functional based tight-binding), pour laquelle des paramètres NH améliorés ont été proposés, en combinaison avec l'approche d'exploration PTMD (pour parallel-tempering molecular dynamics). Les isomères de basse énergie nouvellement déterminés ont été optimisés au niveau MP2 afin d'évaluer la fiabilité de nos paramètres N-H modifiés. Les structures et les énergies de liaison obtenues avec la méthode SCC-DFTB sont en très bon accord avec les résultats de niveau MP2/Def2TZVP, ce qui démontre la capacité de l'approche SCC-DFTB à décrire la PES de ces espèces moléculaires et représente ainsi une première étape vers la modélisation d'agrégats complexes d'intérêt atmosphérique.

L'intérêt porté aux agrégats uracile/eau protonés vise à fournir une description détaillée d'expériences récentes de dissociation induite par collision (CID). Premièrement, les isomères stables des agrégats $(H_2O)_{1-7,11,12}UH^+$ sont calculés en utilisant la même méthodologie que celle décrite ci-dessus. Ensuite, des simulations dynamiques des collisions entre divers isomères $(H_2O)_{1-7,11,12}UH^+$ et un atome d'argon sont réalisées à énergie de collision constante au niveau SCC-DFTB. La proportion simulée de d'agrégats neutres contenant l'uracile par rapport à celle d'agrégats chargés contenant l'uracile, la section efficace de fragmentation ainsi

que les spectres de masse sont cohérents avec les données expérimentales ce qui met en évidence la précision de nos simulations. Ces dernières permettent de sonder en details les fragments qui se forment aux temps courts et de rationaliser la localisation du proton en excès sur ces fragments. Cette dernière propriété est fortement influencée par la nature de l'agrégat soumis à la collision. L'analyses de la proportion des fragments en fonction du temps et des spectres de masse démontrent que, jusqu'à 7 molécules d'eau, un mécanisme de dissociation direct en mis en jeu après la collision alors que pour 11 et 12 molécules, un mécanisme statistique est plus susceptible d'intervenir. Ces simulations, uniques jusqu'à présent, apparaissent comme un outil indispensable pour comprendre et interpréter les expériences CID d'espèces moléculaires hydratées.

Enfin, des simulations d'expériences CID du dimère de pyrène à différentes énergies de collision, entre 2,5 et 30 eV, sont également présentées. Les simulations permettent de comprendre les processus de dissociation mis en jeu. L'accord entre les spectres de masse simulés et mesurés suggère que les principaux processus sont bien pris en compte par cette approche. Il semble que la majeure partie de la dissociation se produise sur une courte échelle de temps (moins de 3 ps). Aux faibles énergies de collision, la section efficace de dissociation augmente avec les énergies de collision alors qu'elle reste presque constante pour des énergies de collision comprises entre 10 et 15 eV. L'analyse de la répartition d'énergie cinétique est utilisée pour obtenir des informations sur les processus de collision/dissociation à l'échelle atomique. Les spectres de masse simulés des clusters parents et dissociés sont obtenus à partir en combinant simulations de dynamique moléculaire et théorie de l'espace des phases pour traiter respectivement la dissociation aux courtes et longues échelles de temps.

Mots clés : SCC-DFTB, CID, dynamique moléculaire, agrégats d'ammonium/ammoniac, agrégats d'uracile eau protonés

To ...

Acknowledgements

I would like to acknowledge ....

Contents

Glossary

BIRD blackbody infrared radiative dissociation

BO Born-Oppenheimer

BSSE basis set superposition errors

CAD collisionally activated dissociation

CC coupled cluster

CI configuration interaction

CID collision-induced dissociation

CM3 class IV - charge model 3

CSF configuration state function

DFT density functional theory

DFTB density-functional based tight-binding

DFTB3 third-order extension of DFTB

FF force field

GGA generalized gradient approximation

HF Hartree-Fock

HK Hohenberg–Kohn

KS Kohn-Sham

LCAO linear combination of atomic orbitals

LDA local density approximation

LEP location of excess proton

LOC line of center model

MCPs micro-channel plates

MCTDH multi-configuration time-dependent Hartree

MD molecular dynamics

MM molecular mechanics

MP2 Møller-Plesset perturbation theory at second-order

NBO natural bond order

NUL neutral uracil loss

PAH polycyclic aromatic hydrocarbons

PES potential energy surface

PST phase space theory

PTMD parallel-tempering molecular dynamics

QM quantum mechanics

RRKM Rice-Ramsperger-Kassel-Marcus

SCC-DFTB self-consistent-charge density-functional based tight-binding

SCF self-consistent field

TCID threshold collision-induced dissociation

TOF time-of-flight

TOFMS time-of-flight mass spectrometry

VV velocity Verlet algorithm

WF wavefunction

ZPVE zero-point vibrational energy

1

General Introduction

The term *cluster* was coined by F. A. Cotton in the early 1960s to refer to compounds containing metal-metal bonds such as $[Re_2Cl_8]^{2-}$ and $[Re_2Br_8]^{2-}$.[1] He defined metal atom cluster compounds as "*those containing a finite group of metal atoms which are held together entirely, mainly, or at least to a significant extent, by bonds directly between the metal atoms even though some non-metal atoms may be associated intimately with the cluster*". Subsequently, the study of clusters, also referred to as aggregates, has greatly diversified and the definition of the term *cluster* has evolved considerably from that given by Cotton. Indeed, in chemistry, the term cluster now refers to an ensemble of bound atoms or molecules that can be isolated or incorporated within a larger chemical compounds, for instance within a solid-state compounds. A cluster is intermediate in size between a single molecule or atom and a nanoparticle. A hundred billion *particles* (here the term particle referred to the constituents of the cluster, which can be either atoms, ions, molecules or a mix) held together behave in most ways like bulk matter whereas small clusters contain no more than a few hundred or a thousand particles and a large cluster designates something containing about a few thousands of particles.[2] Clusters are also intermediate in terms of properties between a single molecule or atom and the corresponding bulk compound.

Cluster chemistry developed contemporaneously along several independent research lines and several families of compounds can be referred to as clusters. Among them, one can mention **naked metal clusters**,[3, 4] metal cluster compounds such as **transition metal carbonyl clusters**,[5] **transition metal halide clusters**,[6] **transition metal organic carbon clusters** (organometallic),[7] **metalloid clusters**,[8] **intermetalloid clusters**,[9, 10] as well as **atomic clusters** composed of non-metal atoms [11, 12, 13, 14] and **molecular clusters**.[15, 16]

1. GENERAL INTRODUCTION

Naked metal clusters encompass only metal atoms that are held together by metallic bond for instance Rh_{13} and Au_{13}.[3] **Transition metal carbonyl clusters** are compounds containing a core that consists of two or more metal atoms linked in part by metal-metal bonds and embraced by carbon monoxide (CO) ligand groups exclusively or predominantly. Similarly, **transition metal halide clusters** are compounds that contains two or more metal atoms (prevalent for heavy metals) linked in part by metal-metal bonds and embraced by halide ligands. Some representative species for transition metal carbonyl and halide clusters are $Mn_2(CO)_{10}$,[5] $Fe_2(CO)_9$,[17] $[Rh_6(CO)_{15}]^{2-}$,[17] $Re_3Cl_{12}^{3-}$,[18] $(Mo_6Cl_8)Cl_4$,[6] Nb_3Cl_8,[19] **Organometallic clusters** contain metal-metal bonds as well as at least one organic ligand directly bonded to a metal atom. It can be neutral or ionic. One example of organometallic cluster is $[Co_3(CCH_3)(CO)_9]$.[7] **Metalloid clusters** are ligand-stabilized clusters that metal atoms possess more direct element-element contacts than element-ligand contacts such as $[Al_{69}(N(SiMe_3)_2)_{18}]^{3-}$ and $[Ga_{84}(N(SiMe_3)_2)_{20}]^{4-}$.[8] The suffix "oid" highlights that such clusters possess atom arrangements that appear in bulk intermetallic compounds with high coordination numbers of the atoms at a molecular scale. **Intermetalloid** clusters consist in at least two different (semi) metallic elements, and possesses more direct metal-metal contacts than metal-ligand contacts. This kind of cluster often appears as discrete units in intermetallic compounds separated from each other by electropositive atoms for instance $[Sn@Cu_{12}@Sn_{20}]^{12-}$.[9, 10] Finally, **clusters composed of non-metal atoms or molecules** are usually found in gas-phase for instance **fullerenes**, [12] **rare-gas clusters**,[11, 14] **water clusters**,[20, 21] and **PAHs (Polycyclic aromatic hydrocarbons) clusters**.[15, 16]

These various kinds of clusters, which list has no mean to be exhaustive, can be differentiated by the bounding mode, *i.e.* the nature of the interaction, between the cluster particles. They can be of different natures:

- **Van der Waals interactions**, which is the main interaction in the rare-gas clusters such as argon clusters.[22]
- **Hydrogen-bond interaction**, which is of paramount importance in a variety of molecular clusters, in particular those containing water molecules.
- **Covalent bond**, as found in fullerenes,[12] or more generally pure carbonaceous aggregates, and other atomic aggregates made of non-metallic atoms.
- **Metallic bond**, as found in Cu, Ag, and Au clusters.[4]

- **Ionic bond**, which exists in ionic clusters such as NaCl[23] or NaF clusters.[24]

Properties of clusters stem from both their size and composition. Clusters can thus exhibit very specific physical and chemical properties that are strongly influences by their structures, which themselves are strongly determined by the number of atoms or molecules they are made of. Furthermore, when a given cluster of a well defined composition switches between different stable configurations, chemical and physical properties can also be strongly impacted. For instance, for different conformational isomers of small Ni and Fe clusters, compact structures are more stable than open structures and the photoabsorption spectra of two isomers of Ni_4 are different.[25] This becomes all the more true as the chemical complexity of the cluster increases, *i.e.* when it is constituted of more than one chemical element, for instance several types of molecules for a molecular cluster or different atoms for an atomic cluster. Depending on the cluster type, see above, intermolecular interactions can be rather weak.[26] This is true for atomic or molecular clusters which cohesion is governed by Van der Waals and/or hydrogen-bond interactions. In that case, the potential energy surface (PES), or energy landscape, can be extremely complex and a large variety of local minima displaying equivalent stabilities exist. The properties of a given cluster can significantly differ from the properties of the corresponding bulk material. For instance, the magnetic moment of small iron particles at room temperature is smaller than that of the bulk.[27] However, a gradual transition occurs between the properties of the clusters and those of the corresponding bulk as cluster size increases.[28] This transition can be rough or continuous depending on the considered species and properties. For instance, U. Landman *et al.* reported that anionic gold clusters favor planar structures up to $\sim$13 atoms.[4] J.-M. L'Hermite *et al.* also reported that the transition temperature extracted from he heat capacity curve of protonated water clusters $(H_2O)_nH^+$ has a strong size dependence as seen in Figure 1.1.[29] Consequently, the study of clusters allow to bridge the gap between single molecule or atom properties and bulk materials, which can be of help to reveal microscopic aspects which are hardly observable in the bulk only.

The field of cluster research can be traced back to 1857 when M. Faraday gave his lecture entitled "*Experimental Relation of (Colloidal) Gold to Light*" which paved the way for modern work on both metal clusters and the interaction of photons with clusters.[32] Cluster research have since drawn a lot of interest and the field has undergone a dramatic growth which can be explained by two main reasons. The first one is the **development of efficient and accurate characterization techniques**. Indeed, experimental techniques now enable the investigation of

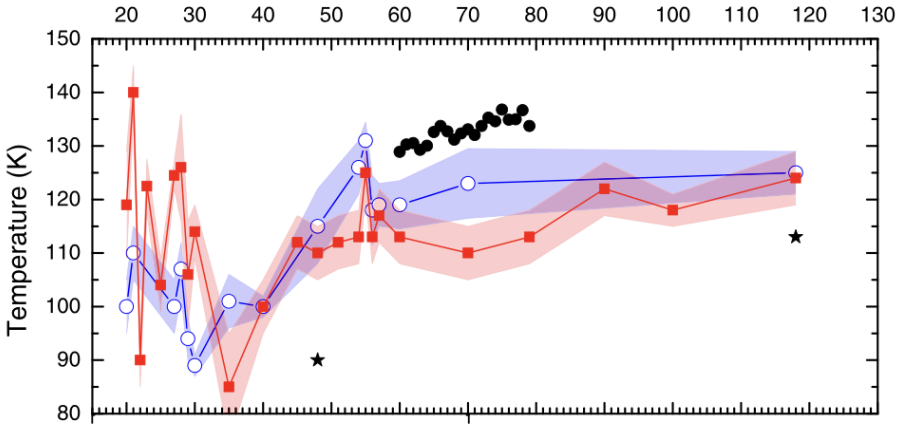


Figure 1.1: Transition temperature of $(H_2O)_nH^+$ clusters (red squares) and $(H_2O)_{n-1}OH^-$ (blue circles) as a function of n. The results obtained by M. Schmidt *et al.* on $(H_2O)_nH^+$ are also presented (black circles)[30] as well as those by C. Hock *et al.* on $(H_2O)_n^-$ clusters (black stars).[31] Figure extracted from reference [29].

clusters of interest in several scientific domains such as astrophysics and astrochemistry,[16] atmospheric physico-chemistry,[33] biochemistry,[34] and environmental science.[35] With the help of mass spectrometer, well-defined cluster sizes can now be isolated and observed.[36] The advent of the laser technology also provides a new dimension to the field as it enables detailed spectroscopic observations.[37] The second reason is related to **application of clusters**. Indeed, clusters may offer ways to develop new kinds of materials,[38] to carry out chemical reactions in new ways,[39] and to get new kinds of understanding of bulk matter by learning how the bulk properties emerge from properties of clusters as the cluster grows larger and larger.[28] For instance, the study of clusters has provided new insights into phase transition, e.g. condensation of gas mixtures,[40] evaporation,[41] precipitation,[42] solidification of liquid mixtures[43] and melting of solids.[44] The study of clusters also helps to understand nucleation phenomena, for instance the formation of nanoscale materials and aerocolloids, as well as ultrafine particles.[45, 46, 47, 48] Study of clusters in gas phase can provide detailed structural, energetic, and spectroscopic information which are hardly accessible from measurements on the bulk.[49, 50, 51, 52] Finally, clusters containing organic/inorganic molecules or ions and water molecules can be viewed as intermediates between a dilute gas phase and a solution. Consequently, their study allows to explore the effects of solvation on the chemistry of gas-phase molecules and ions.[53, 54, 55, 56, 57]

Although it is possible to experimentally probe a large range of properties of clusters, one difficulty is to extract all the chemical and physical information provided by these experiments. Indeed, in the "simplest case", a property determined experimentally can result from a unique isomer of the probed species. A first major task is thus to determine the nature of this lowest energy isomer which is not straightforward. This is where theoretical calculations come in. Indeed, a vast majority of experiments requires the contribution of theoretical calculations in order to determine the lowest energy isomer of a given cluster. For instance, a vast amount of theoretical calculations have been conducted to determine the low energy structures of $(H_2O)_n$ and $(H_2O)_nH^+$ aggregates. Among them, we can mention the studies performed by D. Wales and co-workers using the basin-hopping algorithm.[58, 59, 60, 61] In more difficult cases, the probe properties result from the contribution of several isomers which have to be taken into account. When considering finite-temperature properties, an ergodic exploration of the PES also need to be performed. For instance, J. Boulon *et al.* reported heat capacity curves as a function of temperature of mass selected protonated water clusters and highlighted a stronger steepness of the curve of $(H_2O)_{21}H^+$ as compared to adjacent sizes.[29] Theoretical simulations latter provided explanations for this peculiar behavior.[62] When considering dissociation of clusters, which can be a non-equilibrium process, theoretical calculations allow to understand dissociation mechanisms and energy partition which are not accessible from the experiment.[63, 64, 65, 66] It is worth noting that theoretical calculations can also be useful to make predictions when the experiments are restricted by cost or other conditions.[67]

Among these variety of systems and properties, the present thesis has focused on the study of two kinds of molecular clusters: **water clusters containing an impurity** and **polycyclic aromatic hydrocarbon clusters** with a focus on the **exploration of PES** and the modelling of **collision induced dissociation** processes. In the following, I briefly introduce these different aspects.

Water clusters. Water is ubiquitous in our environment. In view of the importance of water to life and its complex properties, a significant amount of experimental [68, 69, 70, 71, 72, 73, 74, 75] and theoretical [76, 77, 78, 79] studies have been devoted to this fundamental substance since the first realistic interaction potential of water was proposed in 1933.[80, 81] Water clusters are intermediate species between gas and condensed phases, their study is therefore of fundamental importance to understand properties of liquid water and ice. They also offer the opportunity to understand how the properties of liquid water and ice emerge from the assembling of an increasing number of water molecules.[82] They also allow to study

at the molecular scale proton transfer processes,[83, 84] finite-temperature effects as well as nuclear quantum effects. Molecular clusters with a controlled number of solvent molecules are also ideal model systems for providing a fundamental understanding of solute-solvent and solvent-solvent interactions at the molecular level.[85] From a more applicative point of view, they play a significant role in atmospheric sciences where the physical and chemical properties of aerosols are strongly impacted by the properties of the water clusters they are made of.[86, 87, 88, 89, 90, 91, 92] In particular, water clusters can absorb a significant amounts of radiative energy,[93] and therefore they have to be included in climate models.[94] This is not actually the case due to the lack of data about their formation. They can also play a role in astrochemistry where water ice can act as a catalyst for the formation of a large range of chemical species. [95, 96, 97, 98]

From a theoretical point of view, the study of water clusters is not straightforward as water clusters display **two major difficulties**:

- As stated above, the PES of aggregates can display a large number of local minima, *i.e.* stable configurations, and energy barriers. Determination of low-energy structures or ergodic exploration of PES is thus not straightforward. This is all the more true that, for molecular aggregates, the range of considered temperatures often results in a low diffusion of molecules which makes possible for a given aggregate to be trapped in a local minimum of the PES. One textbook case for the complexity of water clusters is $(H_2O)_6$. Despite the apparent simplicity of $(H_2O)_6$, which is the smallest neutral water cluster displaying a tridimensional structure, the nature of its lowest energy isomer has been a subject of debate for several years. It is only in 2012 that C. Pérez *et al.* published an experimental paper in Science in which the authors unambiguously identified three of its isomers: cage, prism and book and concluded that the most stable isomer is the cage.[99] The theoretical description of water clusters thus requires simulation tools specifically devoted to the exploration of complex PES such as **molecular dynamics** or **Monte-Carlo simulations** in combination with efficient **enhanced sampling methods**.

- Molecular scale modelling of water is also made difficult as there is no potential, *ab initio* or empirical, that makes possible to reproduce all the properties of the different phases of water, that is applicable to large systems and that is easily transferable. It is therefore often necessary to make a choice between computational efficiency, transferability, and accuracy. This balance determines the nature of the questions that can be

addressed. Furthermore, the aforementioned **enhanced sampling methods** generally require to repeat a large amount of calculations. Therefore, they need to be combined with computationally efficient approaches to compute the PES. As presented in chapter 2, the method I use within this thesis is the **self-consistent-charge density-functional based tight-binding** (SCC-DFTB) method.

Water clusters are usually combined with other inorganic/organic ions or molecules that make them relevant to astrochemistry, atmospheric chemistry and biological sciences. Therefore, it is of paramount importance to investigate **water clusters containing an impurity**, whether it is experimentally or theoretically. And indeed, in parallel to the study of pure water clusters, such studies have drastically grown over the last years. For instance, one can mention studies devoted to $Cl^-(H_2O)_n$,[100] $Na^+(H_2O)_n$, $H_2PO_4^-(H_2O)_n$,[101] $NH_4^+(H_2O)_n$, $NH_3(H_2O)_n$, $C_6H_6O(H_2O)_n$,[20] $H_2SO_4(H_2O)_n$,[102, 103] $SO_4^{2-}(H_2O)_n$,[103] $(CO)_m(H_2O)_n$, $((CH_3)_2NH_2^+)_m(HSO_4^-)_m(H_2O)_n$,[35] $C_4H_5N_2O_2^+(H_2O)_n$,[104] and $(C_5H_5N)_mH^+(H_2O)_n$.[105] In the domain of astrochemistry, the growth of atmospheric particles can initiate the process of acid cloud formation and participates in reactions leading to the destruction of the ozone layers in polar regions.[106, 107] More detailed studies of atmospheric particles demonstrated the existence of charged molecular aggregates in the stratosphere,[108, 109] in particular sulfate containing aggregates,[103] and ammonium/ammonia containing aggregates.[20, 110] In the latter case, **ammonia has been highlighted as an important component of atmospheric nucleation** together with water and sulphuric acid.[111, 112, 113] This important role of ammonia and ammonium water clusters, and the lake of theoretical studies devoted to these species, motivated a thorough benchmark of the SCC-DFTB approach to model these systems which is presented in chapter 3. In parallel, understanding the **properties of the proton** and how it can impact the solvation properties of molecules of biological interest is of paramount importance for understanding fundamental processes in biology and chemistry. In particular, uracil, one of the nucleobases, plays a key role in the encoding and expression of genetic information in living organisms. The study of **water clusters containing uracil** is therefore a good playground to probe how uracil properties vary from isolated gas-phase to hydrated species and how this is impacted by protonation. Chapters 3 and 4 try to address these questions.

Polycyclic aromatic hydrocarbon clusters. Polycyclic aromatic hydrocarbons (PAHs) are a family of organic molecules made up of two or more aromatic carbon rings containing peripheral hydrogen atoms. These hydrocarbon molecules have aromatic behavior resulting

from the presence of sp^2 carbon atoms. Several examples of PAHs molecules are presented in Figure 1.2.

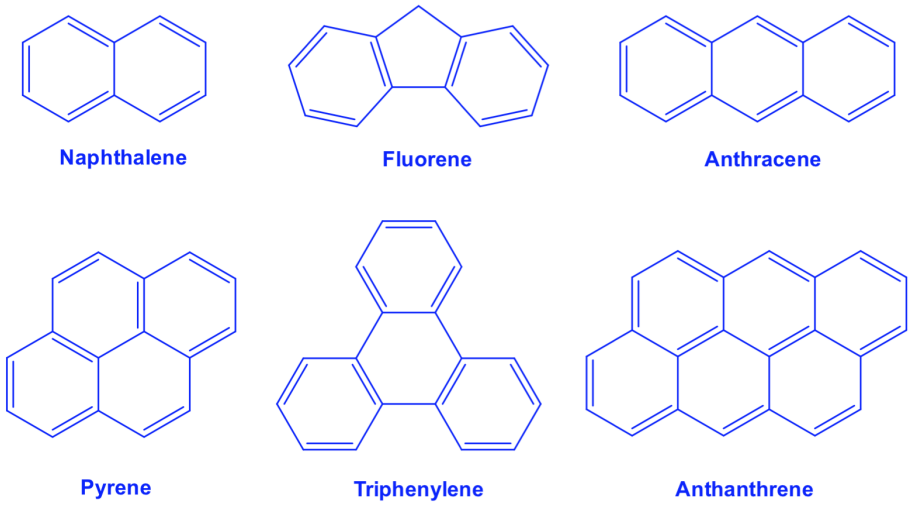


Figure 1.2: Examples of several PAH molecules.

PAHs have been investigated in various scientific fields, both experimentally and theoretically, for instance in astrophysics and astrochemistry, environmental science, combustion science, or the search for new organic solar cell devices.

The presence of PAHs in the interstellar medium was proposed in the middle of the 80s,[114, 115] and they have since played an important role in the astrophysical context. In particular, the so-called unidentified infrared bands in the gas phase of the interstellar medium are thought to be partially attributable to emission by PAHs.[114, 116, 117] They have been proposed to be present in the form of a mixture of neutral, ionised, and partly dehydrogenated molecules and to account for $\sim$10 - 20% of the total carbon in the interstellar medium.[117, 118] In addition, cationic PAH clusters are expected to be abundant in photo-dissociation regions[119, 120] since the ionization energy of the clusters is lower than that of neutral PAHs and decreases with the cluster size,[121, 122] leading to the efficient formation of cationic clusters. These charged species are expected to survive longer than their neutral counterparts due to higher dissociation energies, as predicted by calculations.[121] PAHs are also found in the atmosphere as highly toxic molecules. Their significant abundance arises from their efficient formation as by-products of natural processes, biomass burning, or human activities such as combustion of fossil fuels.[123] In the atmosphere, PAHs with more than three rings can be adsorbed

by various particles, for instance carbonaceous aerosols, ferric oxides, and icy particles.[124] The role of **PAH clusters** in the process of soot nucleation is a major topic in the context of combustion and leads to consider the competition between clustering, evaporation, and oligomerization.[125, 126] Finally, PAH stacks provide possible compounds to define new organic solar cell junctions.[127, 128] Due to the importance of PAHs as mentioned above, the stability of PAH clusters have been extensively studied experimentally and theoretically. [16, 65, 119, 129, 130, 131] IN particular, their evolution following absorption of photons, collision with high or low energetic particles as well as their behaviour in very high pressure environments has been thoroughly studied.[122, 129, 132, 133, 134, 135] Chapter 4 provides thorough theoretical analysis of the **collision-induced dissociation** of the simplest pyrene cluster, *i.e.* the **pyrene dimer cation** in order to complement recent experiments.

Collision-induced dissociation of molecular clusters. The structure, energetics and reactivity of a variety of molecular clusters can be explored by collision-induced dissociation.[136, 137, 138, 139, 140, 141] By colliding a molecule or a molecular cluster with a non-reactive noble gas atom or a small molecule such as N_2, it is possible to monitor the parent ions and collision products by means of mass spectrometry that can provide a wealth of structural information from which one can infer, for instance, dissociation mechanisms,[142, 143] or bond and hydration enthalpies.[144] Collision-induced dissociation has also been used to understand the impact of high-energy radiations on living cells and DNA or RNA,[139, 145, 146] as well as the impact of low-energy collisions on biological molecules.[147, 148]

Extracting energetics or collision process from collision-induced dissociation is not straightforward and it often needs to be **complemented by theoretical calculations**. Two main methodologies can be conducted. The first one is to make an exhaustive description of the PES connecting both parent ions and products. Energetic information on both minima and transition states can then be introduced in Rice-Ramsperger-Kassel-Marcus [149, 150] and/or Kinetic Monte Carlo simulations.[151, 152] The second approach is to perform **molecular dynamics simulations to explicitly model the collision trajectory** of the target ion and the projectile, the energy redistribution, the subsequent reorganizations and fragmentations. A potential is needed to describe the PES of the system and its reactivity in both methodology. For the latter one, the potential needs to reach a very good balance between accuracy and computational efficiency as this methodology requires the propagation of tens, hundreds or even thousands of trajectories. With this in view, it appears that wave-function based methods do not allow to reach a sufficient amount of simulations to describe dynamical behavior at finite temperature.

Unfortunately, the same is true for density-functional theory (DFT). Force-field approaches can easily handle molecular dynamics simulations of system with hundred of atoms for several hundred nanoseconds, but they can poorly describe formation or breaking of covalent bonds and they are poorly transferable. In between DFT and force-field methods, semi-empirical approaches provide interesting alternatives. In particular, the **SCC-DFTB** method allows to perform molecular dynamical simulations of systems containing several tens or hundreds of atoms for simulation time of several hundred picoseconds. This approach has thus been used in the present thesis to model collision-induced dissociation experiments.

To summarize, the goal of this thesis is to go a step further into the theoretical description of the properties of molecular clusters in the view to complement complex experimental measurements. It has focused on two different types of molecular clusters. First, I focused on water clusters containing an impurity, *i.e.* an additional ion or molecule. I have first focused my studies on **ammonium and ammonia water clusters** in order to thoroughly explore their PES to characterize in details low-energy isomers for various cluster sizes. Then I tackle the study of **protonated uracil water clusters** through two aspects: characterize low-energy isomers and model collision-induced dissociation experiments to probe dissociation mechanism in relation with recent experimental measurements. Finally, I address the study of the **pyrene dimer cation** to explore collision trajectories, dissociation mechanism, energy partition, mass spectra, and cross-section.

To introduce, develop, and conclude on these different subjects, this manuscript is organised as follow:

- The **first chapter** introduces the objectives of this thesis. Generalities about clusters, in particular molecular clusters, and collision-induced dissociation are provided.

- The **second chapter** is devoted to the introduction of the fundamental concepts used in theoretical chemistry to solve the electronic structure problem and to explore the PES. It describes the main approaches used along this thesis and their foundations. The **SCC-DFTB** approach, which is the main method used along this thesis, is described in details as well as the **parallel-tempering molecular dynamics** approach to explore PES.

- The **third chapter** focuses on the thorough exploration of the PES of ammonium and ammonia water clusters, as well as protonated uracil water clusters, in the view to discuss their structural and energetic properties. Along this chapter, the results obtained at the

SCC-DFTB level are compared to MP2 results and discuss in the light of the actual literature.

- The **fourth chapter** presents molecular dynamics simulations of collision-induced dissociation of protonated uracil water clusters and pyrene dimer cation. In the former case, the theoretical proportion of formed neutral uracil aggregates *vs.* protonated water cluster as well as total fragmentation cross sections are compared to the experimental results by S. Zamith and J.-M. L'Hermite. The molecular dynamics simulations allow to probe the nature of the formed fragments one the short time scale and to rationalize the location of the excess proton on these fragments. The simulation of the collision-induced dissociation of the pyrene dimer cation at different collision energies is then addressed in this chapter.

- Finally, the conclusions of this thesis as well as a number of perspectives are presented in the **fifth chapter**.

2

Computational Methods

Chemistry is a central scientific discipline dealing with the construction, transformation and properties of molecules. It provides a foundation for understanding both basic and applied scientific disciplines at a fundamental level.[153] Chemistry started out as an experimental science. So, historically, all the chemical processes studied and the theories developed have been through experience. However, with the development of theoretical concepts dealing with the structure of matter and with recent developments in computer technology, a new trend in chemical research has developed: **theoretical chemistry**. **Computational chemistry**, as a subfield of theoretical chemistry, combines fundamental laws with mathematical methods to solve chemical problems through calculations or simulations. Nowadays, modelling of physico-chemical properties has become a important part of research in science of matter and the role of computational chemistry has become fundamental. However, when resorting to computational chemistry, one always has to choose **a balance between accuracy and computational efficiency**.

There exist many methods to model the interactions between different particles in atomic or molecular aggregates. Some of them are depicted in Figure 2.1 and are classified in terms of system sizes (y-axis) and simulation times (x-axis) they can tackle. If focusing on the y-axis, the closer a given method is from the origin of the graph, the more accurate it is. The farther from the origin a method is, the larger are the system sizes it can simulate. For instance, **force fields** (FF), also referred to as **molecular mechanics** (MM) methods, describe the interactions between particles by empirical interatomic potentials and electrons are not treated explicitly. FF methods thus bypass the solving of the **Schrödinger equation** and the quantum aspects of nuclear motion are neglected which allow them to model very large systems. On the opposite,

electronic structures methods such as **full configuration interaction** (Full CI) make it possible to describe electron distributions with a very high accuracy but only for very small systems.

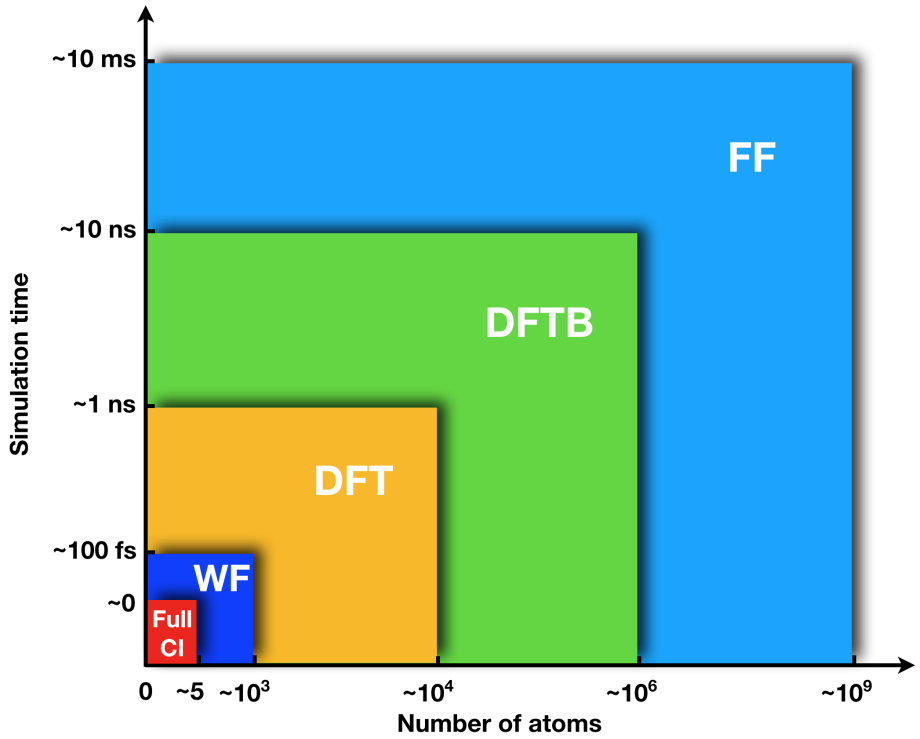


Figure 2.1: Comparison of the computational efficiency, *i.e.* system sizes and simulation times, of various computational chemistry methods. The *y*-axis indicates the length of time accessible from classical molecular simulations for average system sizes tackle by each method. The *x*-axis indicates the approximative maximum system size tractable by each method in a single-point energy calculation.

In parallel to the accuracy, the ability of a given method to efficiently and accurately explore PES is also of paramount importance to model a variety of properties. Here also, different methods allow for different possibilities. For instance, the multi-configuration time-dependent Hartree (MCTDH) method allows for a full quantum treatment of the nuclear degrees of freedom as long as only a limited number of them is taken into account.[154] For large systems, to access time-dependent information, one has to resort to **molecular dynamics simulations** which is dealt with **classical mechanics**, *i.e.* Newton's second law.

This chapter gives a brief description of the theoretical foundations of the computational methods that were used during this thesis. Two main aspects are developed. First, I de-

scribe the main methods to compute PES, *i.e.* the electronic energy, through the solving of the Schrödinger equation. Second, I present the main computational tools that I have used to explore PES.

2.1 Schrödinger Equation

The discovery of the Schrödinger equation by E. Schrödinger in 1925 is an extremely significant landmark in the development of the quantum mechanics. The behavior of a molecular system can be described by the Schrödinger equation (a linear partial differential equation), which describes the wavefunction or state function of a quantum-mechanical system.[155] In quantum mechanics, the concept of wavefunction Ψ is a fundamental postulate which defines the state of a system at each spatial position and time. The Schrödinger equation governs the evolution of the wavefunction Ψ of particles in an atomic or molecular system. In the case of a system composed of M nuclei and N electrons, the Schrödinger equation for the wavefunction in position space $\Psi(\mathbf{R}_\alpha, \mathbf{r}_j, t)$ can be written as:

$$i\hbar \frac{\partial \Psi(\mathbf{R}_\alpha, \mathbf{r}_j, t)}{\partial t} = \hat{H}(\mathbf{R}_\alpha, \mathbf{r}_j, t)\Psi(\mathbf{R}_\alpha, \mathbf{r}_j, t), \quad \alpha = 1, 2, ..., M; j = 1, 2, ..., N. \tag{2.1}$$

The vectorial variables are in bold in this manuscript. This equation describes the evolution of the wavefunction in space and time, where i is the imaginary unit. $\hbar$=$h/2\pi$ (1.054572×10^{-34} J·s) is the reduced Planck constant and t is the time. $\mathbf{R}_\alpha$ and $\mathbf{r}_j$ refer to the coordinates of nucleus α and electron j, respectively. $\hat{H}$ is the so called Hamiltonian operator corresponding to the total energy of the system. When the Hamiltonian itself is explicitly independent on time (the total wavefunction still has a time dependency), it is possible to decompose the space variables and the time variable to write the time-independent Schrödinger equation:

$$\hat{H}\Psi_k(\mathbf{R}_\alpha, \mathbf{r}_j) = E_k \Psi_k(\mathbf{R}_\alpha, \mathbf{r}_j) \tag{2.2}$$

where E_k is the total energy of the system associated with the eigenstate Ψ_k. According to this, the evolution of the wavefunction becomes:

$$\Psi(\mathbf{R}_\alpha, \mathbf{r}_j, t) = \sum_k c_k \Psi_k(\mathbf{R}_\alpha, \mathbf{r}_j) e^{-iE_k t/\hbar} \tag{2.3}$$

where c_k is a coefficient. The eigenstates obtained from equation 2.2 are the stationary states of the system and form a complete basis of orthonormal vectors. The lowest energy eigenstate is called the ground state usually denoted as Ψ_0 and E_0 is the corresponding energy.

In a system made up of M nuclei and N electrons, the Hamiltonian operator (in a non-relativistic framework) is written as following:

$$\begin{aligned}\hat{H} &= \hat{T} + \hat{V} \\ &= \hat{T}_n + \hat{T}_e + \hat{V}_{nn} + \hat{V}_{ee} + \hat{V}_{ne} \\ &= -\frac{1}{2}\sum_{\alpha=1}^{M}\frac{1}{M_\alpha}\nabla_\alpha^2 - \frac{1}{2}\sum_{j=1}^{N}\nabla_j^2 + \sum_{\alpha=1}^{M}\sum_{\beta>\alpha}^{M}\frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} \\ &\quad + \sum_{j=1}^{N}\sum_{i>j}^{N}\frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} - \sum_{\alpha=1}^{M}\sum_{j=1}^{N}\frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{r}_j|}\end{aligned} \tag{2.4}$$

in which the atomic unit system is used. The energy is thus expressed in Hartree. The mass is in unit of the mass of electron and the length is in Bohr. M_α refers to the mass of nucleus α (in atomic unit) and Z_α is the atomic number. $\hat{T}_n$ and $\hat{T}_e$ are the kinetic energy operators of nuclei and electrons, respectively. $\hat{V}_{nn}$, $\hat{V}_{ee}$ and $\hat{V}_{ne}$ denote the potential energy operators of the repulsion between the nuclei, repulsion between the electrons and electrostatic attraction between nuclei and electrons, respectively. ∇^2 is the Laplace operator. For nucleus α in three dimensions using Cartesian coordinates, its position vector is $\mathbf{R}_\alpha$=(X_α, Y_α, Z_α) and ∇^2 is expressed as following:

$$\nabla^2 = \frac{\partial^2}{\partial X_\alpha^2} + \frac{\partial^2}{\partial Y_\alpha^2} + \frac{\partial^2}{\partial Z_\alpha^2} \tag{2.5}$$

2.2 Born-Oppenheimer Approximation

Electrons are very light particles which can not be described correctly even qualitatively through classical mechanics. If we want to describe the electron distribution in details, quantum mechanics must be applied, *i.e.* solving the Schrödinger equation. Usually in atomic and molecular systems, it is very hard, if not impossible, to obtain the exact solutions of Schrödinger equation. It is only possible in the system containing one nucleus and one electron, *i.e.*, the hydrogen atom or hydrogenic ions. For molecular species, the mathematical complexity to solve the Schrödinger equation will increase with the number of degrees of freedom of the system. Thus, it is necessary to resort to approximations in almost all cases. The **Born-Oppenheimer (BO) approximation** or adiabatic approximation is a cornerstone in real-life quantum analysis of atoms and molecules, which helps to solve the Schrödinger equation.

BO approximation is based on the large difference of mass between nuclei and electron and correspondingly the time scales of their motions. Indeed, the mass of proton is much higher

than that of electron ($m_p/m_e \approx 1836$). With the same amount of kinetic energy, the electrons move much faster than the nuclei.

Then, it is considered that the electrons move in the field of the fixed atomic nuclei, *i.e.*, the electrons adapt instantly to the displacement of the nuclei while remaining in their ground state.[156]

BO approximation consists of expressing the total wavefunction of a molecule as the product of a nuclear (vibrational, rotational) wavefunction and an electronic wavefunction, which enables a separation of the Hamiltonian operator into the fast electronic term and the usually much slower nuclear term, where the coupling between electrons and nuclei is neglected so that the two smaller and non-coupled systems can be solved more efficiently. In mathematical terms, the total wavefunction Ψ_{tot} of a molecule can be expressed as an expansion in the complete set of electronic wavefunctions ψ_k^e with the expansion coefficients being parametric functions of the nuclei coordinates $\mathbf{R}_\alpha$:

$$\Psi_{\text{tot}}(\mathbf{R}_\alpha, \mathbf{r}_j) = \sum_{k=1}^{\infty} \psi_k^e(\mathbf{r}_j; \mathbf{R}_\alpha)\psi_k^n(\mathbf{R}_\alpha) \tag{2.6}$$

where the semicolon symbolizes the positions of the nuclei as parameters and not the variables of electronic wavefunction. This indicates that, although ψ_k^e is a real-valued function of $\mathbf{r}_j$, its functional form depends on $\mathbf{R}_\alpha$.

Two smaller, consecutive steps can be used when using the BO approximation. In the first step, the nuclei is treated as stationary. The corresponding operator $\hat{T}_n$ is subtracted from the total molecular Hamiltonian operator $\hat{H}$ leading to the electronic Hamiltonian without considering nuclear kinetic energy:

$$\hat{H}_e = \hat{T}_e + \hat{V}_{nn} + \hat{V}_{ee} + \hat{V}_{ne} \tag{2.7}$$

The major computational work is in solving the electronic Schrödinger equation for a given set of nuclear coordinates.

$$\hat{H}_e(\mathbf{R}_\alpha, \mathbf{r}_j)\psi_k^e(\mathbf{R}_\alpha, \mathbf{r}_j) = E_k^e \psi_k^e(\mathbf{R}_\alpha, \mathbf{r}_j) \tag{2.8}$$

where the eigenvalue E_k^e, electronic energy, depends on the chosen positions $\mathbf{R}_\alpha$ of the nuclei. By varying these positions $\mathbf{R}_a$ in small steps and repeatedly to solve the electronic Schrödinger equation, so one can obtain E_k^e as a function of $\mathbf{R}_\alpha$, which generates the PES.

In the second step, the nuclear kinetic energy T_n is reintroduced, and the Schrödinger equation for the nuclear motion is:

$$(\hat{T}_n + E_k^e(\mathbf{R}_\alpha) + \langle\Psi_k|\nabla_n^2|\Psi_k\rangle)\psi_k^n(\mathbf{R}_\alpha) = E_k^{\text{tot}}\psi_k^n(\mathbf{R}_\alpha) \tag{2.9}$$

The eigenvalue E_k^{tot} is the total energy of the molecule, which includes the overall rotation translation of the molecule, contributions from electrons, and nuclear vibrations. This second step involves a separation of vibrational, translational, and rotational motions. M. Born and R. Oppenheimer assumed that the integral $\langle\Psi_k|\nabla_n^2|\Psi_k\rangle$ (diagonal correction) weakly depends on the nuclear coordinates, so that it can be ignored. [156] Therefore, the Born-Oppenheimer approximation allows to describe the movement of nuclei in the corresponding potential to an adiabatic electronic state by the following equation:

$$(\hat{T}_n + E_k^e(\mathbf{R}_\alpha))\psi_k^n(\mathbf{R}_\alpha) = E_0^{\text{tot}}\psi_k^n(\mathbf{R}_\alpha) \tag{2.10}$$

In this thesis, I will assume electrons adapt fast to reach their electronic ground state. The potential energy E_0^e thus equals the ground state electronic energy E_0 and the total energy is then $E_0^{\text{tot}} = T_{\text{n}} + E_0$. The Schrödinger equation for $\psi_0^n(\mathbf{R}_\alpha)$ can therefore be written as:

$$(\hat{T}_n + E_0)\psi_0^n(\mathbf{R}_\alpha) = E_0^{\text{tot}}\psi_0^n(\mathbf{R}_\alpha) \tag{2.11}$$

The next step is usually to consider the nuclei can be described classically. One can then consider that they evolve classically *i.e.* following new Newton's equation, on a PES defined by the ground state electronic energy. The calculation of the ground state electronic energy is discussed in section 2.3 The BO approximation only introduces very small errors in most systems, which explains it is widely applied in quantum chemistry to speed up the computation of molecular wavefunctions of large molecules. However, the BO approximation is only valid when the electronic state is sufficiently separated from other electronic states. That is to say that, when two or more solutions of the electronic Schrödinger equation come close to each other energetically, the approximation loses validity (usually called "break down") but it can be used as a starting point for more refined methods.[157, 158]

2.3 Computation of Electronic Energy

Experimental spectroscopic investigations help in understanding the electronic structure of molecules, for instance measurements of absorption, emission and scattering. These measurements can often provide a detailed picture of molecular systems but sometimes they are

difficult to interpret. In the last few decades, molecular electronic-structure theory has developed to a stage where it can provide invaluable assistance in the interpretation of experimental measurements of a wide range of important properties of molecules in rotational and vibrational spectroscopies, magnetic-resonance spectroscopies, ultraviolet/visible spectroscopies, and others.[159, 160, 161, 162, 163, 164]

Electronic wavefunction of systems including three or more interacting particles can not be obtained analytically, so approximations must be applied. Many approximations have been proposed to obtain approximate solutions of the exact electronic wavefunction. Each one of them is usually the basis of one or more calculation approaches, which have their own advantages and disadvantages. When solutions of the Schrödinger equation are obtained without reference to experimental data, the methods are usually referred to as *ab initio* compared to semi-empirical models. The Hartree-Fock (HF) model takes all interactions between electrons into account except for the correlation between electrons that is neglected. Post-HF theory usually generates more accurate results by considering the electronic correlation.[165] DFT methods in the Kohn-Sham formulation can be regarded as an improvement over the HF theory as it considers approximated electronic correlation. Density-functional based tight-binding (DFTB) formalism is an approximated DFT method that involves additional approximations.[166, 167, 168, 169] Compared with quantum chemical (QM) methods that require considerable computer resources, molecular mechanics (MM) calculations are much cheaper but present severe limitations in the treatment of chemical reactivity. QM/MM techniques that combine QM for the reactive region and MM for the remainder are very promising, especially for large systems.[170, 171, 172]

The following section focuses on the description of wavefunction based methods, DFT method, and the DFTB methods used to solve Schrödinger equation and the electronic structure problem.

2.3.1 Wavefunction based Methods

Hartree-Fock theory. Hartree-Fock theory is one of the earliest wavefunction based approximation methods, which is the foundation for a large part of the computational work on the electronic structure of atoms and molecules.[173, 174, 175] The HF approximation transforms the many-body Schrödinger equation into many coupled single-particle equations, which often assumes the N-body electronic wavefunction of a system is approximated by a single Slater determinant and every electron is considered to be independent. Hartree proposed that the

electronic wavefunction could be approximated by assuming that in addition to the nuclei, the individual electrons could be separated as well. Therefore, the many-electron wavefunction would be a product of one-electron wavefunctions ψ_j:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)...\psi_n(\mathbf{r}_n) \tag{2.12}$$

The HF theory assumes that every electron moves in an average field of all the other electrons and the nuclei in the molecule, which is an example of a mean-field approximation. The HF equations for an individual electron j moving in the mean field V_i^{HF}, can be expressed as:

$$\left(-\frac{1}{2}\nabla_j^2 + V_\sigma(\mathbf{r}_j) + V_{\mathrm{H}}(\mathbf{r}_j)\right)\psi_\sigma^{\mathrm{HF}}(\mathbf{r}_j) - \sum_{k=1}^{N_\sigma}\int d^3r' \frac{\psi_{k\sigma}^{\mathrm{HF}*}(\mathbf{r}')\psi_{k\sigma}^{\mathrm{HF}}(\mathbf{r}_j)}{\mid \mathbf{r}_j - \mathbf{r}' \mid}\psi_\sigma^{\mathrm{HF}}(\mathbf{r}') = E_\sigma^{\mathrm{HF}}\psi_\sigma^{\mathrm{HF}}(\mathbf{r}_j) \tag{2.13}$$

where ∇_j^2 depends on the jth electron coordinates. $V_\sigma(\mathbf{r}_j)$ refers to the external potential. σ is the spin. The last term on the left-hand side is the HF exchange potential. Using the electronic density, we can obtain:

$$V_{\mathrm{H}}(\mathbf{r}_j) = \int d^3r_j \frac{\rho(\mathbf{r}')}{\mid \mathbf{r}_j - \mathbf{r}' \mid} \tag{2.14}$$

$V_H(\mathbf{r}_j)$ is the Hartree potential:

$$\rho(\mathbf{r}_j) = \sum_\sigma\sum_j^{N_\sigma}\rho_\sigma(\mathbf{r}_j) = \sum_\sigma\sum_j^{N_\sigma} \mid \psi_\sigma^{\mathrm{H}}(\mathbf{r}_j) \mid^2 \tag{2.15}$$

r' refers to the positions of all other electrons except electron j.

HF or self-consistent field (SCF) theory can be derived by invoking the variational principle in a restricted space of wavefunctions, which was introduced in many books.[176, 177, 178, 179] C. Roothaan then put forward the approximation of molecular orbitals (MOs) as a linear combination of atomic orbitals (LCAO) in 1951, namely, a linear combination of atomic basis functions to solve the HF equations of molecules.[180] Thus was a significant improvement in the practical solution of the HF equations. Roothaan equations allow to transform the HF problem into a linear algebra problem, for which algebraic equations are particularly suitable for modern computers.[181] The vast majority of computational approaches, whether *ab initio*, semi-empirical, or even some sophisticated force-fields are based on the concept that the molecular orbitals of a given molecule can be built from the atomic orbitals of its constituent atoms.[182]

The HF theory showed some success.[183] However, it does not contain the electron correlation beyond the minimum required to satisfy the antisymmetry for electronic wavefunctions. The resulting approximated electronic energies are therefore not accurate enough for most practical applications in chemistry. Good HF results can account for over 99% of the true total energy of the system. Due to the variational principle, the HF wavefunction is always too high in energy. So, the remaining 1% error with respect to the true total energy, which is defined as the correlation energy ($E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}$) is essential to account for the chemical properties of atoms and molecules. Electron correlation results from the correlated behavior of electrons, and the failure of HF theory to describe it originates from that the mean-field approximation can not treat electron-electron interactions properly.[184]

Post Hartree-Fock methods. Post Hartree-Fock methods provide improvements to HF theory by adding the electron correlation.[165, 185] The calculation of the correlation energy is then the objective of several post-HF methods (for instance, the configuration interaction (CI) method,[186, 187] Møller-Plesset perturbation theory (MP2, MP3 and MP4) methods,[188, 189, 190] coupled cluster (CC) method,[191, 192, 193, 194] quantum chemistry composite (G2, G3, and T1) methods,[195, 196, 197] and so on). Post-Hartree-Fock methods usually give more accurate results than Hartree-Fock calculations, [186, 198, 199] but the additional accuracy comes to the price of a higher computational cost. Among the aforementioned post-HF methods, the Møller-Plesset perturbation theory at second-order (MP2) method was used along this thesis.

HF theory variationally optimize the orbitals of a single configuration state function (CSF). In the case of a closed-shell singlet and some open-shell cases, this would be a single Slater determinant. To include the electron correlation in wavefunction theory, a superposition of CSFs should be considered. This superposition, referred to as a multiconfigurational wavefunction, is a linear combination of two or more CSFs. Standard multiconfigurational approaches are the CI, CC, and perturbation theory methods.

2.3.2 Density Functional Theory

For a long time, approximations based on the wavefunction were systematically applied to solve the Schrödinger equation. However, it is usually impractical to perform a wavefunction based calculation with chemical accuracy for complex or large systems. Density functional theory is based on the electron density rather than the electronic wavefunction.[200, 201] Because DFT displays a more favourable scaling of computational resources with respect to system

size, DFT is nowadays the most widely used method available in computational chemistry, computational physics, and condensed-matter physics for ground state calculation of large and complex systems.

Although DFT has a history almost as old as the Schrödinger equation, the modern form dates back to the paper published by P. Hohenberg and W. Kohn [202] that introduced the two Hohenberg–Kohn (HK) theorems in 1964 and the extension by M. Levy in 1979.[203] The theory is usually applied in the form latter suggested by W. Kohn and L. J. Sham in 1965.[200] DFT makes it possible to transform the problem of electrons interacting and evolving in a nuclear potential to a problem of independent electrons evolving in an effective potential. The electron density $\rho(\mathbf{r})$ corresponds to the number of electrons per unit volume in a given state.

$$\rho(\mathbf{r}) = N \int \Psi^*(\mathbf{r}, \mathbf{r}_2, \mathbf{r}_3, ..., \mathbf{r}_N)\Psi(\mathbf{r}, \mathbf{r}_2, \mathbf{r}_3, ..., \mathbf{r}_N) d\mathbf{r}_2 d\mathbf{r}_3 ... d\mathbf{r}_N \tag{2.16}$$

The central idea of DFT is to promote $\rho(\mathbf{r})$ (function which only depends on three spatial coordinates) as the key variable in the determination of the electronic energy of a system.

This idea originates from the the model of the uniform electron gas in the phase space around an atom developed in 1927 by L. H. Thomas [204] and E. Fermi [205], which is the predecessor to density functional theory. Nevertheless, the Thomas-Fermi model is unable to correctly describe molecular bonds because it does not take into account the exchange and correlation energies.

The first HK theorem shows that, for a many-electron system in its ground state, the energy is uniquely determined by the electron density $\rho(\mathbf{r})$. In other words, the first HK theorem shows that it is not necessary to know the wavefunction of the system to obtain its energy and that the knowledge of the electron density alone is sufficient. It sets down the foundation for reducing the many-body problem of N electrons with $3N$ spatial coordinates to three spatial coordinates, through using functionals of the electron density. From this theorem, it follows that $\rho(\mathbf{r})$ determines the external potential $V_{\text{ext}}(\mathbf{r})$ and N can be obtained via the normalization of $\rho(\mathbf{r})$:

$$\int \rho(\mathbf{r}) d\mathbf{r} = N \tag{2.17}$$

N and $V_{\text{ext}}(\mathbf{r})$ determine the electronic Hamiltonian. $\rho(\mathbf{r})$ determines the energy and all other ground state electronic properties of a system. This is clearly shown in Figure 2.2.

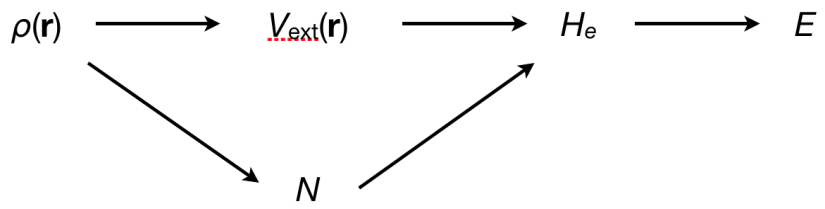


Figure 2.2: Interdependence of basic variables in the Hohenberg-Kohn theorem.

The second HK theorem is a variational electron density theorem which defines an energy functional of the electronic density for a system. For a given external potential $V_{\text{ext}}(\mathbf{r})$, the ground state energy E_0 of the system is obtained by minimizing this energy function with respect to the electronic density, the corresponding density is exactly the ground state electronic density $\rho(\mathbf{r})$. From the HK theorems, we can write the functional of the total energy of the system as a sum of the kinetic energy of th electrons $T_{\text{e}}[\rho(\mathbf{r})]$, and the electronic interaction energy $E_{\text{ee}}[\rho(\mathbf{r})]$:

$$E[\rho(\mathbf{r})] = T_{\text{e}}[\rho(\mathbf{r})] + V_{\text{ee}}[\rho(\mathbf{r})] + \int V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})d(\mathbf{r}) \tag{2.18}$$

the last term represents the interaction between the lectron density and the external potential, *i.e.* the nuclei in the case of atoms and molecules. To obtain the energy of the ground state from this equation, the variation principle can be applied with respect to $\rho(\mathbf{r})$. To do this, the form of $T_{\text{e}}[\rho(\mathbf{r})]$ and $V_{\text{ee}}[\rho(\mathbf{r})]$ should be known.

The HK theorems do not provide mathematical expressions for $T_{\text{e}}[\rho(\mathbf{r})]$ and $V_{\text{ee}}[\rho]$ for a system of interacting particles. To solve this problem, Kohn and Sham first proposed to treat the electrons as non-interacting particles subject to the $V_{\text{ext}}(\mathbf{r})$ potential only.[200] The idea was to work with a fictitious system of N non-interacting electrons evolving in an effective potential and having the exact electronic density of the system. The resolution of equation 2.18 for a system of non-interacting electrons is known exactly, and if the correct electron density is reproduced, then the exact electronic energy of the system can be calculated. The total energy of the real system is composed as follows:

$$E_{\text{DFT}}[\rho(\mathbf{r})] = T_{\text{no}}[\rho(\mathbf{r})] + \underbrace{E_{\text{H}}[\rho(\mathbf{r})] + \int V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})d(\mathbf{r}) + E_{\text{xc}}[\rho(\mathbf{r})]}_{E_{\text{eff}}[\rho(\mathbf{r})]} \tag{2.19}$$

where the sum of last three terms on the right-hand side is the effective energy $E_{\text{eff}}[\rho(\mathbf{r})]$. $T_{\text{no}}[\rho(\mathbf{r})]$ is the kinetic energy of a system of non-interacting electrons:

$$T_{\text{no}}[\rho(\mathbf{r})] = \sum_{i}^{N} \left\langle \Psi_i \left| -\frac{1}{2}\nabla^2 \right| \Psi_i \right\rangle \tag{2.20}$$

$E_H[\rho(\mathbf{r})]$ represents the Hartree energy which corresponds to the interaction energy of a classical charge distribution of density $\rho(\mathbf{r})$:

$$E_{\text{H}}[\rho(\mathbf{r})] = \frac{1}{2}\int\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \tag{2.21}$$

$V_{\text{ext}}(\mathbf{r})$ is the external potential. The remaining energy components are assembled in the exchange-correlation energy $E_{\text{xc}}[\rho(\mathbf{r})]$ functional containing the difference between the kinetic energy of the real system $T[\rho(\mathbf{r})]$ and that of the non-interacting system $T_n[\rho(\mathbf{r})]$ and the non-classical part of $E_{\text{ee}}[\rho]$, adding the difference between $V_{\text{ee}}[\rho(\mathbf{r})$ and $E_{\text{H}}[\rho(\mathbf{r})]$. The $E_{\text{xc}}[\rho(\mathbf{r})]$ functional can thus be expressed as:

$$E_{\text{xc}}[\rho(\mathbf{r})] = (T[\rho(\mathbf{r})] - T_{\text{no}}[\rho(\mathbf{r})]) + (V_{\text{ee}}[\rho(\mathbf{r})] - E_{\text{H}}[\rho(\mathbf{r})]) \tag{2.22}$$

In practice, to minimize the energy $E[\rho(\mathbf{r})]$ with respect to $\rho(\mathbf{r})$ applying by the variational principle while considering the constraints of orbital orthogonality, one performs an optimization under the constraints using Lagrange multipliers.

Combining eqs 2.19, 2.20, and 2.21, the effective potential $V_{\text{eff}}[\rho(\mathbf{r})]$ can be introduced as:

$$\begin{aligned} V_{\text{eff}}[\rho(\mathbf{r})] &= V_{\text{ext}}[\rho(\mathbf{r})] + V_{\text{H}}[\rho(\mathbf{r})] + V_{\text{xc}}[\rho(\mathbf{r})] \\ &= V_{\text{ext}}[\rho(\mathbf{r})] + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{\text{xc}}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \qquad (2.23) \\ &= \frac{\delta E_{\text{eff}}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \qquad (2.24) \end{aligned}$$

where $V_H[\rho(\mathbf{r})]$ refers to the Hartree potential and $V_{xc}[\rho(\mathbf{r})]$ is the exchange-correlation potential. Therefore, for a given $V_{\text{eff}}[\rho(\mathbf{r})]$, one can obtain $\rho(\mathbf{r})$ by making the right-hand side of equation **??** independent of $\mathbf{r}$ which is done by introducing the molecular orbitals $\psi_i(\mathbf{r})$ such that:

$$\rho(\mathbf{r}) = \sum_{i}^{N} |\psi_i(\mathbf{r})|^2 \tag{2.25}$$

The molecular orbitals $\psi_i(\mathbf{r})$ should satisfy the one-electron KS equations:

$$\underbrace{\left(-\frac{1}{2}\nabla_i^2 + V_{\text{eff}}[\rho(\mathbf{r})]\right)}_{\text{KS operator}} \psi_i(\mathbf{r}) = E_i \psi_i(\mathbf{r}) \tag{2.26}$$

This result can be reobtained within a variational context when looking for those orbitals minimizing the energy functional of equation 2.18, subject to orthonormality conditions:

$$\int \psi_i^*(\mathbf{r})\psi_j(\mathbf{r})d\mathbf{r} = \delta_{ij} \tag{2.27}$$

The one-electron KS equation of equation 2.26, just as the HF equations, needs to be solved iteratively. The computational cost for the incorporation of electron correlation is the one necessary to calculate $V_{xc}[\rho(\mathbf{r})]$.

DFT is in principle correct if one knows the "exact" exchange-correlation functional. However, despite a lot of work aimed at determining this exact functional, its form is still unknown and a systematic strategy for improvement is not available. Therefore, it is always necessar to use an approximate functional that is characterized by more or less important artefacts in effective calculations. Over the last decades, many exchange-correlation functionals have been proposed. Although they are different, it is possible to classify them into families according to some common characters. **Local Density Approximation** (LDA), **Generalized Gradient Approximation** (GGA), **meta-GGA**, and **hydrid functionals** (containing to some extent a contribution of HF exact-exchange) are some of the most widely used approximations. **LDA** is the first approximation of $E_{xc}[\rho(\mathbf{r})]$ proposed by W. Kohn and L. J. Sham,[200] which is based on the description of the homogeneous electron gas. For atomic or molecular systems, when densities vary rapidly in space, the assumption of a uniform electron density is not correct and the LDA approximation is not applicable any more. For example, binding energies in molecules are usually overestimated by the LDA approximation. To solve this problem, a new family of functionals, **GGA functional**, which includes a contribution of the electron density gradient was developed.[200, 206, 207] **Hybrid functionals** were first proposed by Becke in 1993.[208] The main idea is that for an uncorrelated system, the HF energy is exact while for a highly correlated system LDA or GGA energy is more appropriate. These two states are then connected by a continuum of partially correlated real systems of identical density. This connection is described by the adiabatic connection formula:

$$E_{\text{xc}}[\rho(\mathbf{r})] = \int_0^1 E_{\text{xc}}[\rho(\mathbf{r})]\lambda d\lambda \tag{2.28}$$

where λ=0 corresponds to the uncorrelated case and λ=1 corresponds to the highly correlated case. Hybrid functionals usually provide higher quality results than those provided by LDA and GGA functionals for the study of molecular properties. Hybrid functionals have been very successful for ground-state properties. B3LYP (Becke, 3-parameter, Lee-Yang-Parr),[208, 209, 210, 211] PBE with one empirical parameter (PBE0, also called PBE1PBE),[212] Heyd-Scuseria-Ernzerhof (HSE06),[213, 214], and M06-2X [215] are some of the most popularly hybrid functionals. Approximate functionals suffer from some generic problems, many of which can be traced to the so-called delocalization error [216] or the closely related problem of self-interaction error. Nevertheless, it should be emphasized that today DFT, cast in the Kohn-Sham formalism, provides a computational tool with a remarkable quality and computationally less expensive than wavefunction based methods.

2.3.3 Density Functional based Tight-Binding Theory

DFT is computationally too expensive for systems with more than hundreds of atoms especially when one needs to perform global optimization or to perform molecular dynamics (MD) simulations with sufficient statistical sampling of the initial conditions or to perform fairly long trajectories. It is therefore necessary to further simplify the method in order to reduce the computational cost. To do so, W. Foulkes and R. Haydock showed that tight-binding models can be derived from the DFT.[217] Later, DFTB was proposed by D. Porezag *et al*..[218]

Non-self-consistent DFTB scheme is suitable to study systems in which polyatomic electronic density is well described by a sum of atom-like densities. This is the case ighly ionic and homonuclear covalent systems. However, uncertainty rises in the non-self-consistent DFTB scheme whem chemical bonds are controlled by a subtler charge balance between atoms especially for polar, semi-conductor and heteronuclear molecules. The self-consistent-charge extension of DFTB, SCC-DFTB, was born as an improvement of standard DFTB to provide a better description of electronic systems in which long-range Coulomb interactions are significant.[168, 168, 218, 219, 220, 221] However, this method is not fully suitable for the calculation of molecular aggregates because of the dispersion and charge resonance present in such systems. Further corrections were made to the model in order to address these problems. I know describe in more details how SCC-DFTB can be derived from DFT.

General principles of DFTB. DFTB is derived from DFT based on the following approximations:

- Only valence electrons are treated explicitly.
- Molecular orbitals are developed on atomic valence orbitals.
- A Taylor expansion of the total energy around a reference density is realized.
- Integrals involving more than two centers are neglected.

First-order DFTB (historically referred to as zeroth-order DFTB) takes into account the first term of Taylor's expansion and which is equivalence with the other tight-binding model. **Second-order DFTB** (historically the SCC-DFTB) introduces a self-consistent procedure on atomic charges. There is also a more recent **third-order extension of DFTB** (referred to as DFTB3).[222] DFTB3 was not used in this thesis so the third-order expansion term will not be shown in the following equations. According to equation 2.19, $E_{\mathrm{DFT}}[\rho(\mathbf{r})]$ can be written as:

$$E_{\mathrm{DFT}}[\rho(\mathbf{r})] = \langle\psi_i|\hat{T}_{\mathrm{no}}[\rho(\mathbf{r})]|\psi_i\rangle + E_{\mathrm{eff}}[\rho(\mathbf{r})] \tag{2.29}$$

The SCC-DFTB scheme is built on the second-order Taylor series development around a reference electronic density $\rho_0(\mathbf{r})$, so $\rho(\mathbf{r})$=$\rho_0(\mathbf{r})$+$\delta\rho(\mathbf{r})$. In practice, $\rho_0(\mathbf{r})$ is taken as the superimposition of the densities of isolated atoms. Then $E_{\mathrm{DFTB}}[\rho_0(\mathbf{r})+\delta\rho(\mathbf{r})]$ can be written as follows:

$$\begin{aligned} E_{\mathrm{DFTB}}[\rho_0(\mathbf{r})+\delta\rho(\mathbf{r})] = &\langle\psi_i|\hat{T}_{\mathrm{no}}[\rho(\mathbf{r})]|\psi_i\rangle + E_{eff}[\rho_0(\mathbf{r})] + \int \frac{\delta E_{eff}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}\Big|_{\rho_0}\delta\rho(\mathbf{r})d\mathbf{r} \\ &+ \frac{1}{2}\iint \frac{\delta^2 E_{eff}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\Big|_{\rho_0,\,\rho_0'}\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \end{aligned} \tag{2.30}$$

with $\delta\rho(\mathbf{r})$ = $\rho(\mathbf{r})$-$\rho_0(\mathbf{r})$. Equation 2.30 can be rewritten as:

$$\begin{aligned} E_{\mathrm{DFTB}}[\rho_0(\mathbf{r})+\delta\rho(\mathbf{r})] = &\overbrace{E_{\mathrm{eff}}[\rho_0(\mathbf{r})] - \int \frac{\delta E_{\mathrm{eff}}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}\Big|_{\rho_0}\rho_0(\mathbf{r})d\mathbf{r}}^{E_{\mathrm{rep}}} \\ &\underbrace{+\langle\psi_i|\hat{T}_{\mathrm{no}}[\rho(\mathbf{r})]|\psi_i\rangle + \int \frac{\delta E_{\mathrm{eff}}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}\Big|_{\rho_0}\rho(\mathbf{r})d\mathbf{r}}_{E_{\mathrm{band}}} \\ &\underbrace{+\frac{1}{2}\iint \frac{\delta^2 E_{\mathrm{eff}}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\Big|_{\rho_0,\,\rho_0'}\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')d\mathbf{r}d\mathbf{r}'}_{E_{\mathrm{2nd}}} \end{aligned} \tag{2.31}$$

From the first line of equation 2.31, we can define that the reference Hamiltonian $\hat{H}^0$ only depends on the reference electron density ρ_0:

$$\hat{H}^0 = -\frac{1}{2}\nabla^2 + \underbrace{V_{\text{ext}}[\rho(\mathbf{r})] + \int \frac{\rho_0'(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + V_{\text{xc}}[\rho_0(\mathbf{r})]}_{V_{\text{eff}}([\rho_0(\mathbf{r})])} \tag{2.32}$$

where we combined the last three terms in the operator as $V_{\text{eff}}([\rho_0(\mathbf{r})])$.

The right-hand side terms in the first line of equation 2.31 vary linearly with $\rho_0(\mathbf{r})$, and correspond to a repulsive contribution E_{rep}. The sum of the terms in the second line is the so-called band energy E_{band}. The third line is the second-order energy $E_{2\text{nd}}$. Equation 2.31 can be rewritten as follows:

$$E_{\text{DFTB}}[\rho_0(\mathbf{r}) + \delta\rho(\mathbf{r})] = E_{\text{rep}}[\rho_0(\mathbf{r})] + \overbrace{\sum_i^{\text{occ}} n_i \left\langle \psi_i \left| \hat{H}^0 \right| \psi_i \right\rangle}^{E_{\text{band}}} + E_{2\text{nd}}\left[\rho_0(\mathbf{r}), (\delta\rho(\mathbf{r}))^2\right] \tag{2.33}$$

Band energy term. In DFTB, one relies on the use of LCAO for the description of the Kohn-Sham molecular orbitals $\psi_i(\mathbf{r})$. Here the atomic orbitals are limited to the valence orbitals of atoms:

$$\psi_i(\mathbf{r}) = \sum_\nu C_{i\nu}\varphi_\nu(\mathbf{r} - \mathbf{R}_\alpha) \tag{2.34}$$

E_{band} is the sum over the energies of all occupied orbitals obtained by diagonalization of the parameterized Hamiltonian matrix. For atoms α and β, the corresponding atomic orbitals are φ_μ and φ_ν. Then E_{band} can be rewritten with equations 2.24 and 2.34:

$$\sum_i^{\text{occ}} n_i \left\langle \psi_i \left| \hat{H}^0 \right| \psi_i \right\rangle = \sum_i^{\text{occ}} n_i \sum_\mu^{\text{occ}} \sum_\nu^{\text{occ}} C_{i\mu} C_{i\nu} \underbrace{\left\langle \varphi_\mu \left| \hat{T}_e[\rho(\mathbf{r})] + V_{\text{eff}}[\rho_0(\mathbf{r})] \right| \varphi_\nu \right\rangle}_{H^0_{\mu\nu}} \tag{2.35}$$

$\forall \mu \in \alpha, \nu \in \beta$. The Hamiltonian matrix element $H^0_{\mu\nu}$ is defined as:

$$H^0_{\mu\nu} = \left\langle \varphi_\mu \left| \hat{H}^0 \right| \varphi_\nu \right\rangle, \tag{2.36}$$

The effective potential $V_{\text{eff}}[\rho_0(\mathbf{r})]$ is defined as the sum of potentials $V_\alpha(\mathbf{r})$ centered on the atoms:

$$V_{\text{eff}}[\rho_0(\mathbf{r})] = \sum_\alpha V_\alpha(\mathbf{r}_\alpha) \tag{2.37}$$

where $\mathbf{r}_\alpha = \mathbf{r} - \mathbf{R}_\alpha$. The Hamiltonian matrix elements can be written as follows:

$$H^0_{\mu\nu} = \left\langle \varphi_\mu \left| -\frac{1}{2}\nabla^2_\nu + V_\alpha + (1-\delta_{\alpha\beta})V_\beta \right| \varphi_\nu \right\rangle, \text{with } \mu \in \alpha, \nu \in \beta \tag{2.38}$$

where $\delta_{\alpha\beta}$ is the Kronecker's delta.

For diagonal elements, the energy level in the free atom is chosen, which ensures the correct dissociation limits. The interatomic blocks are computed as given in equation 2.38, depending on the choice of potential generation. Because of the orthogonality of the basis functions, the off-diagonal elements of the intraatomic blocks are exactly zero. To summarize, within the electronic density superposition approach, the $H^0_{\mu\nu}$ elements can be unfolded as:

$$H^0_{\mu\nu} = \begin{cases} \varepsilon_\mu^{\text{free atom}}, & \mu = \nu \\ \left\langle \varphi_\mu \left| -\frac{1}{2}\nabla^2_\nu + V_\alpha + (1-\delta_{\alpha\beta})V_\beta \right| \varphi_\nu \right\rangle, & \mu \in \alpha, \nu \in \beta, \alpha \neq \beta \\ 0, & \text{otherwise} \end{cases} \tag{2.39}$$

It should be noted that the $H^0_{\mu\nu}$ elements only depend on atoms α and β. Therefore only the two-center matrix elements and the two-center elements of the overlap matrix can be explicitly calculated, in other words, interactions at three or more centers are neglected as stated above.

Repulsive energy term. E_{rep} is a repulsive contribution obtained from the sum of atomic-pair terms, which only depend on the reference electronic density $\rho_0(\mathbf{r})$.

Due to the screening of terms of more than two centers, one can assume that the two-center contributions are short ranged. But E_{rep} doesn't decay to zero for long interatomic distances. Instead, it decays to a constant given by the atomic contributions:

$$\lim_{R_{\alpha\beta}\to\infty} E_{\text{rep}}[\rho_0(\mathbf{r})] = \sum_\alpha^N E_{\text{rep}}[\rho_0^\alpha(\mathbf{r}_\alpha)] \tag{2.40}$$

The right side of equation 2.40 is assumed to make E_{rep} only rely on the two-center contributions:

$$E_{\text{rep}}[\rho_0(\mathbf{r})] \approx \frac{1}{2}\sum_\alpha^N \sum_\beta^N V(\mathbf{R}_\alpha - \mathbf{R}_\beta) \tag{2.41}$$

In practice, it it possible to calculate E_{rep} with known values of $\rho_0(\mathbf{r})$, but it's more convenient to adjust the expression of E_{rep} to *ab initio* calculations. Therefore, E_{rep} is determined by comparing the difference between the DFT energy E_{DFT} and E_{band}+E_{2nd} as a function of the interatomic distance $R_{\alpha\beta}$:

$$E_{\text{rep}}[\rho_0(\mathbf{r})] \equiv E_{\text{rep}}(R_{\alpha\beta}) = E_{\text{DFT}}(R_{\alpha\beta}) - E_{\text{band}}(R_{\alpha\beta}) - E_{\text{2nd}}(R_{\alpha\beta}) \tag{2.42}$$

Second-order term. In SCC-DFTB, the electronic density is corrected by including the second-order contribution $E_{2\text{nd}}$ in equation 2.33, which is ignored in first-order DFTB.

To include the density fluctuations in a simple but efficient way according to tight-binding method, $\delta\rho$ can be written as a superposition of atom-like contributions $\delta\rho_\alpha$, which has a fast decrease with the increase of the distance from the corresponding atomic center:

$$\delta\rho(\mathbf{r}) = \sum_{\alpha}^{N} \delta\rho_\alpha(\mathbf{r}) \tag{2.43}$$

where $\delta\rho_\alpha$ can be simplified with the monopole approximation as follows:

$$\delta\rho_\alpha(\mathbf{r}) = \Delta q_\alpha F_0(\mathbf{r} - \mathbf{R}_\alpha) \tag{2.44}$$

where the atomic charge fluctuation Δq_α (difference between the Mulliken population q_α [223] of atomic α and the number of valence electrons of the atom at infinity) is estimated by the Mulliken expression. F_0 represents the normalized radial dependence of the electronic density fluctuation in atom α. This means the effects of charge transfer are included, however, the changes in the shape of the electronic density are ignored. $E_{2\text{nd}}$ can be rewritten with equations 2.31 and 2.33 as follows:

$$\begin{aligned} E_{2\text{nd}} \approx & \frac{1}{2}\sum_{\alpha}^{N}\sum_{\beta}^{N}\Delta q_\alpha \Delta q_\beta \overbrace{\iint \left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta^2 E_{xc}[\rho_0(\mathbf{r})]}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\Big|_{\rho_0,\,\rho_0'}\right) F(\alpha,\beta) d\mathbf{r} d\mathbf{r}'}^{\gamma_{\alpha\beta}} \\ = & \frac{1}{2}\sum_{\alpha}^{N}\sum_{\beta}^{N}\Delta q_\alpha \Delta q_\beta \gamma_{\alpha\beta} \end{aligned} \tag{2.45}$$

$$F(\alpha,\beta) = F_0(\mathbf{r} - \mathbf{R}_\alpha) \times F_0(\mathbf{r}' - \mathbf{R}_\beta) \tag{2.46}$$

where the two-electron integrals $\gamma_{\alpha\beta}$ is introduced for convenience.

To calculate equation 2.46, $\gamma_{\alpha\beta}$ must be analyzed. In the limit case, the interatomic distance is very large, $|\mathbf{R}_\alpha - \mathbf{R}_\beta| = |\mathbf{r} - \mathbf{r}'| \to \infty$ with GGA-DFT, the exchange-correlation term tends to zero. $\gamma_{\alpha\beta}$ that describes the interaction of two normalized spherical electronic densities reduces to $1/|\mathbf{R}_\alpha - \mathbf{R}_\beta|$, so $E_{2\text{nd}}$ can be expressed as follows:

$$E_{2\text{nd}} = \frac{1}{2}\sum_{\alpha}^{N}\sum_{\beta}^{N}\frac{\Delta q_\alpha \Delta q_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} \tag{2.47}$$

It is worth noting that the electronic density $\rho(\mathbf{r})$ influences explicitly the calculation of the electrostatic energy in DFT. In the context of DFTB, point charges are used and the electronic density around the atom is condensed at a point. In practice, Mulliken's definition of charge is often used,[223] which is defined as:

$$q_\alpha^{\text{Mull}} = \frac{1}{2}\sum_\alpha^{\text{occ}} n_i \sum_{\mu\in\alpha}\sum_\nu \left(C_{i\mu}^* C_{i\nu} S_{\mu\nu} + C_{i\nu}^* C_{i\mu} S_{\nu\mu}\right) \tag{2.48}$$

where μ denotes the orbitals belonging to the atom α. This definition does not allow the bond between two different atoms to polarize.

Total energy. The total energy in SCC-DFTB can be written from the previous different contributions as follows:

$$\begin{aligned} E_{\text{SCC}} = &\sum_i^{\text{occ}} n_i \sum_\mu^{\text{occ}}\sum_\nu^{\text{occ}} C_{i\mu} C_{i\nu} \left\langle \psi_i \middle| \hat{H}^0 \middle| \psi_i \right\rangle + \frac{1}{2}\sum_\alpha^N\sum_\beta^N V(\mathbf{R}_\alpha - \mathbf{R}_\beta) \\ &+ \frac{1}{2}\sum_\alpha^N\sum_\beta^N \Delta q_\alpha \Delta q_\beta \gamma_{\alpha\beta} \end{aligned} \tag{2.49}$$

Here the contribution from H^0 is exactly the same with the one in standard DFTB scheme and the molecular orbitals ψ_i are expanded in a LCAO model (equation 2.34).

Secular equations. From this LCAO model, we can get the secular equations:

$$\sum_\nu C_{i\nu}\left(H_{\mu\nu}^0 - \varepsilon_i S_{\mu\nu}\right) = 0,\ \forall \mu,\ \nu \tag{2.50}$$

where $H_{\mu\nu}^0$ are the Hamiltonian matrix elements and $S_{\mu\nu}$ are the overlap matrix elements.

The secular equations can be rewritten with modified Hamiltonian matrix elements, $H_{\mu\nu}$, defined by:

$$H_{\mu\nu} = H_{\mu\nu}^0 + \frac{1}{2} S_{\mu\nu} \sum_\zeta (\gamma_{\alpha\zeta} + \gamma_{\beta\zeta}) \Delta q_\zeta \tag{2.51}$$

The $H_{\mu\nu}^0$ and $S_{\mu\nu}$ matrix elements are the same than the ones defined in the standard DFTB method (equation 2.36). The $H_{\mu\nu}$ elements rely on the atomic charges explicitly, and the atomic charges depend on the molecular orbitals (see equation 2.48). Then the resolution can be achieved in a self-consistent way. First, from an initial set of charges the $H_{\mu\nu}$ elements which depend on these charges can be computed. The KS equations (equations 2.50) are then solved which gives the energy of the KS orbitals and the corresponding eigenvectors. The corresponding coefficients allow to compute a new set of charges which will be used in the calculation of new $H_{\mu\nu}$ elements. This procedure is repeated until the atomic charges are converged.

DFTB is derived from DFT, it therefore inherits the specific problems of DFT. For instance, the traditional DFT functionals can not describe properly dispersion interaction and charge resonance phenomena in charged aggregates. DFTB also display some specific problems because of its own approximations such as the use of Mulliken atomic charges, the absence of atomic polarization, the absence of coupling between atomic orbitals located on the same atom. This differs from DFT that explicitly considers atomic polarization.

Atomic charges. As presented above, DFTB was initially developed with Mulliken charges; However, other definitions of atomic charges are possible such as Natural Bond Order (NBO)[224, 225, 226, 227] and Electrostatic Potential Fitting (EPF) charges.[228, 229] EPF has a fairly good representation of the electrostatic term of a molecule dominated by the Van der Waals interactions. CM3 (Class IV / Charge Model 3) charges were proposed by J. Li *etal*. in 1998 and they have been considered in DFTB. They give good results for the description of the electric dipole and the electrostatic potential, partial atomic charges in molecules, and Coulombic intermolecular potential of polycyclic aromatic hydrocarbon clusters.[230, 231, 232, 233] The CM3 charges are defined as:

$$q_\alpha^{\mathrm{CM3}} = q_\alpha^{\mathrm{Mull}} + \sum_{\beta \neq \alpha} \left(C_{t_\alpha t_\beta} K_{\alpha\beta} + D_{t_\alpha t_\beta} K_{\alpha\beta}^2 \right) \tag{2.52}$$

where $K_{\alpha\beta}$ is the Mayer bond order,[234, 235, 236] which depends on the density matrix of the orbitals belonging to each of the two atoms α and β. $C_{t_\alpha t\beta}$ and $D_{t_\alpha t\beta}$ are empirical parameters which are related to the nature of atoms α and β.

In this thesis, in practice for the calculation of electronic energy, the definition of CM3 charges is simplified as:

$$q_\alpha^{\mathrm{CM3}} = q_\alpha^{\mathrm{Mull}} + \sum_{\beta \neq \alpha} C_{t_\alpha t_\beta} K_{\alpha\beta} \tag{2.53}$$

Dispersion energy. In order to correctly describe the energies of molecular systems, it is necessary to take into account the Van der Waals interactions. London dispersion interaction acts between atoms and molecules and represents a large part of the Van der Waals interactions. London dispersion interaction arises from the interactions between fluctuating dipoles. One of the major drawbacks within DFTB is that it does not take dispersion interactions into account. This is also true for DFT, when using LDA or GGA functionals. To overcome this limitation, semi-empirical energy corrections can be applied in DFT and DFTB calculations, which usually gives good results.[237, 238, 239, 240, 241, 242, 243, 244] Moreover, a semi-empirical

correction to the dispersion in the DFTB energy leaves the freedom to use the already existing DFTB parameters withjout any need of re-parametrization. For the studies presented in this thesis, the dispersion energy plays a fairly important role. The correction we apply for the dispersion energy is of the following form:

$$E_{\text{disp}}(\mathbf{R}_{\alpha\beta}) = -\sum_{\alpha}\sum_{\beta\neq\alpha} f(\mathbf{R}_{\alpha\beta})\frac{C^6_{\alpha\beta}}{\mathbf{R}^6_{\alpha\beta}} \tag{2.54}$$

where $f(\mathbf{R}_{\alpha\beta})$ is a cutoff function, which allows to avoid the divergence of this term at a short distance. $C^6_{\alpha\beta}$ is an empirical coefficient calculated for each pair of atoms.

2.3.4 Force Field Methods

Force field is a computational method utilized to estimate the forces between particles, in other words it is the functional form and parameter sets applied to calculate the potential energy of a system. FF is interatomic potential and use the same concept with force field in classical physics, a vector field which describes a non-contact force acting on a particle at different positions in space. The acting force on each particle is derived as a gradient of the potential energy with respect to the particle positions.[245] In such case, the interactions in a system are determined from parameterized potentials in which the electronic structure can not be described explicitly because each particle is treated as a material point. The particles interact with each other through the FF and the integration algorithm is applied to the particles. In most cases, this leads to a big decrease of the precision level in the description of the system but it can reduce the calculation cost drastically, which allows to model systems containing several thousands of particles.

Different potentials have been proposed,[246, 247, 248, 249, 250, 251] which can be classified two main groups: pair potentials and multi-body potentials. For pair potentials, harmonic interaction is the most basic form:[252]

$$V(R_{\alpha\beta}) = k(R_{\alpha\beta} - R_{eq})^2 \tag{2.55}$$

where $R_{\alpha\beta}$ is the distance between two interacting particles α and β. k is the harmonic force constant. R_{eq} is the equilibrium distance where the force of repulsion equals to the one of attraction. This potential is sufficient for systems only deviating very slightly from the bond distance at equilibrium and interactions reasonably limited to adjacent pairs of particles. However, for systems with large deviations, other potential forms must be used, for instance, the

Morse potential which describes the potential energy of a diatomic molecule:[248, 253]

$$V(R_{\alpha\beta}) = D_{eq}\left(1 - e^{-a(R_{\alpha\beta}-R_{eq})}\right)^2 \tag{2.56}$$

where D_{eq} is the depth of the Morse potential well. Parameter a determines the width of the potential, the smaller a the larger the well. The force constant of the bond can be found via the Taylor expansion of $V(R_{\alpha\beta})$ around $R_{\alpha\beta} = R_{eq}$ to the second derivative of the potential energy function, from which it can obtain $a = (k_{eq}/2D_{eq})^{\frac{1}{2}}$ in which k_{eq} is the force constant of the minimum well.

Lennard-Jones potential as known as LJ potential or 12-6 potential is a pair potential, which is proposed by J. Lennard-Jones in 1924.[246, 247] It models soft repulsive and attractive interactions, therefore, the LJ potential describes electronically neutral atoms or molecules. Because of its simple mathematical form, it is one of the most widely used intermolecular potentials especially to describe the interaction within noble gas molecules. The total energy can be written as the sum of the interaction energy of all atomic pairs, which is defined as follows:[254]

$$V_{\mathrm{LJ}}(R_{\alpha\beta}) = 4\varepsilon_0\left[\left(\frac{\sigma}{R_{\alpha\beta}}\right)^{12} - \left(\frac{\sigma}{R_{\alpha\beta}}\right)^{6}\right] = \varepsilon_0\left[\left(\frac{R_{eq}}{R_{\alpha\beta}}\right)^{2n} - 2\left(\frac{R_{eq}}{R_{\alpha\beta}}\right)^{n}\right] \tag{2.57}$$

ε_0 denotes the depth of the potential well that usually refers to the dispersion energy. σ is the interparticle distance at which the potential energy is zero. $n = 6$ and ε_0 denotes the bonding energy, the energy required to separate the atoms. And the LJ potential has its minimum $(-\varepsilon_0)$ at a distance of $R_{\alpha\beta} = R_{eq} = 2^{\frac{1}{6}}\sigma$.

It becomes more complicated for molecular systems in which the modes of intermolecular and intramolecular interactions are very different. So it requires to develop force fields including several kinds of potentials. The expression of the potential energy for a molecular system which is the most frequently used for simple organic molecules and biological macromolecules is written as follows:[255]

$$\begin{aligned} V_{\mathrm{total}}(R) &= \overbrace{V_{\mathrm{bond}} + V_{\mathrm{angle}} + V_{\mathrm{dihedral}}}^{V_{\mathrm{bonded}}} + \overbrace{V_{\mathrm{VW}} + V_{\mathrm{Coulomb}}}^{V_{\mathrm{nonbonded}}} \\ &= \sum_{\mathrm{bond}} \frac{k_{\alpha\beta}}{2}(R_{\alpha\beta} - R_{eq})^2 + \sum_{\mathrm{angle}} \frac{k_\theta}{2}(\theta - \theta_{eq})^2 + \sum_{\mathrm{dihedral}} \frac{k_\phi}{2}(1 + \cos(n\phi - \phi_{eq}))^2 \\ &\quad + \sum_{\mathrm{VW}} 4\varepsilon_0\left[\left(\frac{\sigma}{R_{\alpha\beta}}\right)^{12} - \left(\frac{\sigma}{R_{\alpha\beta}}\right)^{6}\right] + \sum_{\mathrm{Coulomb}} \frac{1}{4\pi\varepsilon_0}\frac{q_\alpha q_\beta}{R_{\alpha\beta}} \end{aligned} \tag{2.58}$$

Molecular interactions can determine the macroscopic properties of matter. Van der Waals interaction is an important force between atoms and molecules but it is extremely short ranged. Van der Waals interaction energy is also termed London dispersion energy. This repulsive distance dependence is usually modeled as a term that scales with $1/R_{\alpha\beta}^{12}$ although there is no absolute physical reason for it. The van der Waals interaction energy is usually approximated by a Lennard-Jones potential.

Now we know how to calculate the potential energy of nuclei (electronic energy). Many thermal dynamical, chemical, and physical properties require the association of PES. Different methods will be shown for the exploration of the PES in the next section.

2.4 Exploration of PES

The PES is a function giving the energy of a system according to one or more nuclear coordinates. If there is only one coordinate, the PES is called the energy profile or potential energy curve. Figure 2.3 is a model of PES in two dimension. The analogy of a hilly landscape with peaks, valleys, mountain passes helps to understand the PES. Stable molecular structures correspond to the minima in the valleys on a PES. The shape of the valley around a minimum determines the vibrational spectrum. The key points on a PES can be classified according to the first and all second derivatives of the energy with respect nuclear coordinates, which correspond to the gradients and the curvatures, respectively. When points have a zero gradient (stationary points) and their second derivatives are positive, this corresponds to a **local minimum** (physically stable structure). Among these local minima, the lowest energy minima is called the **global minimum**. When at least one of the second derivatives is negative, the point is a transition state (saddle point).

Molecular structure, properties, chemical reactivity, dynamics, and vibrational spectra of molecules can be readily understood in terms of PES. Only very simple PES can be obtained from experiment whereas computational chemistry has developed different kinds of methods to efficiently explore PES. To survey the PES, the choice of an exploration method can be guided by the shape of the PES (the statistical set that one wishes to study) and the temporal aspect.

Monte Carlo and classical **MD** simulations are widely recognized approaches for the exploration of PES of systems containing a large number of degrees of freedom such as molecular aggregates. Monte Carlo methods allow the sampling of a PES by performing random shifts

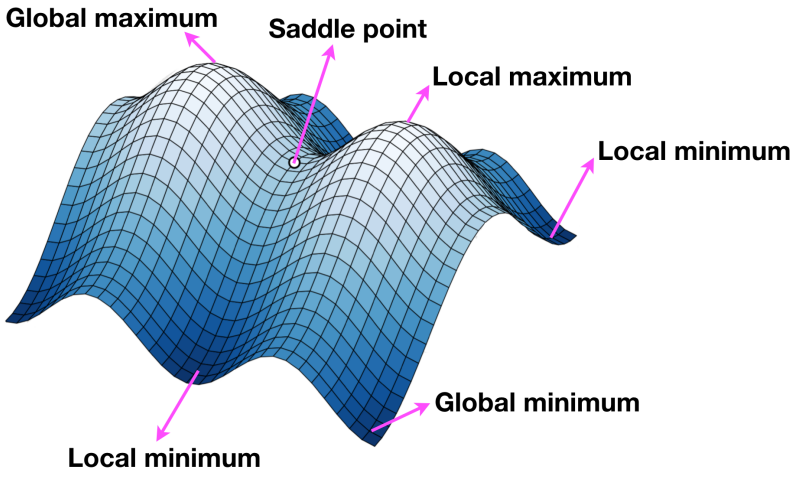


Figure 2.3: Schematic representation of some key points on a model potential energy surface.

in order to correctly reproduce the probability distribution of the configurations which are accessible from the phase space of a system as a function of state variables such as temperature, energy, or number of particles. One of the main advantages of Monte Carlo methods is that they do not require the calculation of gradients. In MD simulations, one needs to numerically integrate Newton's equations of motion of the interacting particles. One can then obtain statistical properties by performing time averaging along the resulting deterministic trajectory in phase space. **Parallel-tempering molecular dynamics** (PTMD) is an improvement of classical MD that allows for a more efficient sampling of the PES. The Monte Carlo, classical MD and PTMD methods are briefly described in the following sections.

2.4.1 Monte Carlo Simulations

The term Monte Carlo denotes a class of algorithmic methods that aims at a probabilistic description, relying on the use of random numbers. The name Monte Carlo alludes to the games of chance taking place at the Monte Carlo casino. The Monte Carlo method was introduced in 1947 by N. Metropolis,[256] and was first published in 1949 by N. Metropolis in collaboration with S. Ulam.[151] Monte Carlo methods have been widely applied in computational physics, computational statistics, biomedicine, machine learning, industrial engineering, economics and finance, and other fields.[257] **Monte Carlo methods can** generally be roughly **divided into two categories**. The first category is applied to problems that have inherent randomness, and the computing power of the computer can directly simulate this random process. For example,

in nuclear physics, analysis of the transmission process of neutrons in a reactor. The second category applies to problems that can be transformed into randomly distributed characteristic numbers, such as the probability of a random event. Through the random sampling method, the probability of random events is estimated by the frequency of occurrence, and this is used as the solution to the problem. This method is usually used to solve complicated multi-dimensional integration problems.

Monte Carlo methods are mainly used in the field of optimization, numerical integration, and generating draws from a probability distribution. In physics related problems, Monte Carlo methods are applied to simulat systems diplaying many coupled degrees of freedom, such as fluids, strongly coupled solids, disordered materials, and cellular structures. In statistical physics not related to thermodynamics, Monte Carlo molecular simulation is an alternative to MD simulations. Monte Carlo methods can be used to compute statistical field theories of simple particles and polymer systems.[258, 259, 260] When solving practical problems using a Monte Carlo methods, it includes two main parts. First, random variables with various probability distributions need to be generated to simulate a certain process. Second, statistical methods need to be used to estimate the numerical characteristics of the model to obtain a numerical solutions to the actual problem. Monte Carlo methods make it possible for the sampling of a PES by performing random shifts in order to correctly reproduce the probability distribution of the configurations. One example for the calculation procedure of a molecular simulation is as follows:

1. A random molecular configuration is generated using a random number generator.

2. Random changes are made to the particle coordinates of this molecular configuration, resulting in a new molecular configuration.

3. Calculate the energy of the new molecular configuration.

4. Compare the energy change between the new molecular configuration and the former one to determine whether to accept the configuration change or not.

    - If the energy of the new molecular configuration is lower than that of the original one, the new configuration change is accepted, and the new configuration is used for the next iteration.

- If not, the Boltzmann factor is calculated and a random number is generated. If this random number is greater than the calculated Boltzmann factor, then the new configuration is discarded. If not, the new configuration is conserved and is used for the next iteration.

5. If a new iteration is required, the process is repeated from step 2.

General principles. Monte Carlo calculations that lead to quantitative results may be regarded as attempts to estimate the value of a multiple integral. This is particularly true for the applications in equilibrium statistical thermodynamics, where one hopes to calculate the thermal average $\langle A \rangle_T$ of an observable $A(\mathbf{X})$ as an integral over phase space Ω, where $\mathbf{X}$ is a point in Ω:

$$\langle A \rangle_T = \frac{1}{Z} \int_\Omega d\mathbf{X} A(\mathbf{X}) e^{-H(\mathbf{X})/k_B T} \tag{2.59}$$

in which Z is the partition function, and k_B is the Boltzmann constant. T refers to the temperature, and $H(\mathbf{X})$ denotes the Hamiltonian of the system. To illustrate the general application of Monte Carlo techniques, here the standard example of the one-dimensional integral I over integration space Ω is taken:

$$I = \int_\Omega f(\mathbf{x}) d\mathbf{x} \tag{2.60}$$

Such integral can be evaluated more efficiently using conventional numerical means than using Monte Carlo methods. However, the extension to higher dimensions (number of degrees of freedom greater than 3) is always difficult in practice with conventional numerical integration. It is then possible to use stochastic approaches that one explores the configuration space randomly and the computation of the integral is estimated in the form of an average value:

$$I_N = \frac{V}{N} \sum_i^N f(\mathbf{x}_i) \tag{2.61}$$

where V is the volume of the integration space and N is the number of points drawn randomly. Within the limit of numbers ($N \to \infty$), the computation of the integral is exact as follows:

$$I = \lim_{N \to \infty} I_N \tag{2.62}$$

In addition, according to the central limit theorem, we can obtain a Gaussian distribution of the integral result and the statistical error is independent of the dimension of the problem to be solved.

In practice, a completely random exploration is inefficient if $f(\mathbf{x})$ is located in a specific space region. It is then interesting to constrain the exploration of space by introducing a distribution $\rho(\mathbf{x})$ to preferentially visit the regions where the function has the maximum influence in the integral. The integral can be rewritten as:

$$I = \int_{\Omega} \frac{f(\mathbf{x})}{\rho(\mathbf{x})} \rho(\mathbf{x}) d\mathbf{x} \tag{2.63}$$

then the integral is estimated with the mean value of f/p of the points explored:

$$I_N = \frac{V_p}{N} \sum_{i}^{N} \frac{f(\mathbf{x}_i)}{\rho(\mathbf{x}_i)} \tag{2.64}$$

where V_p is the volume of the integration space weighted by the distribution $\rho(\mathbf{x})$. In order to explore the surface as efficiently as possible in calculating the mean value, the best choice for the points distribution is $\rho(\mathbf{x}) = f(\mathbf{x})$.

In physical system, it usually needs to consider the weighting of different possible configurations to calculate the mean values. By definition, the computation of an average value of an observable A is the integral of A over the whole phase space weighted by $\rho(\mathbf{x})$ divided by the volume of this weighted space and is written as follows:

$$\langle A \rangle = \frac{\int A(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x}}{\int \rho(\mathbf{x}) d\mathbf{x}} \tag{2.65}$$

Then the problem is to generate a distribution of configurations according to the law $\rho(\mathbf{x})$. The solution used by the Monte Carlo methods is to generate these points using a Markov chain, in other words, a sequential series of configurations where each configuration belongs to the state space and only depends on the previous point. The properties of Markov processes have the following consequences:

$$\sum_{x_{i+1}} p(\mathbf{x}_i \to \mathbf{x}_{i+1}) = 1 \tag{2.66}$$

$$\sum_{\mathbf{x}_i} \rho(\mathbf{x}_i) p(\mathbf{x}_i \to \mathbf{x}_{i+1}) = \rho(\mathbf{x}_{i+1}) \tag{2.67}$$

the transition probability from configuration i to configuration j is $p(\mathbf{x}_i \to \mathbf{x}_j)$. To ensure the validity of equation 2.67, the condition of microreversibility is sufficient that the movements in one direction are exactly compensated by the reverse movements. This condition is verified by the detailed balance sheet equation:

$$\rho(\mathbf{x}_i) p(\mathbf{x}_i \to x_j) = \rho(\mathbf{x}_j) p(\mathbf{x}_j \to \mathbf{x}_i) \tag{2.68}$$

The transition probability $p(\mathbf{x}_i \rightarrow \mathbf{x}_j)$ can be defined as the product of the probability of attempting a transition $\eta(\mathbf{x}_i \rightarrow \mathbf{x}_j)$ with the probability of accepting this same transition $\sigma(\mathbf{x}_i \rightarrow \mathbf{x}_j)$. In the case of symmetrical movements, $\eta(\mathbf{x}_i \rightarrow \mathbf{x}_j) = \eta(\mathbf{x}_j \rightarrow \mathbf{x}_i)$ and the probability of accepting a relocation must satisfy the condition of the detailed assessment:

$$\frac{\sigma(\mathbf{x}_i \rightarrow \mathbf{x}_{i+1})}{\sigma(\mathbf{x}_{i+1} \rightarrow \mathbf{x}_i)} = \frac{\rho(\mathbf{x}_{i+1})}{\rho(\mathbf{x}_i)} \tag{2.69}$$

2.4.2 Classical Molecular Dynamics

MD is a powerful tool for analyzing the physical movements of atoms and molecules of many-body systems. MD was originally developed following the earlier successes of Monte Carlo simulations. The first work about MD was published in 1957 by B. Alder *et al.* which focused on the integration of classical equations of motion for a system of hard spheres.[261] Before long, radiation damage at low and moderate energies were studied using MD in 1960 and MD was also applied to simulate liquid argon in 1964.[262, 263] MD experienced an extremely rapid development in the years that followed. MD simulations have been applied in chemistry, biochemistry, physics, biophysics, materials science, and branches of engineering, which is often coupled with experimental measurements to facilitate interpretation. MD has a strong predictive potential thus making it possible to motivate the implementation of new experiments. The diversity, broadness, and sophistication level of MD techniques have been continuously reported.[245, 264, 265, 266, 267, 268, 269, 270, 271, 272] The range of applications of MD simulations is extremely wide, for instance, the study of structure,[269, 273, 274, 275] thermodynamic,[276, 277, 278, 279, 280] diffusion,[281, 282, 283, 284] viscosity,[282, 285, 286, 287] and spectroscopies.[288, 289, 290, 291] Moreover, MD simulations are not only limited to the study of homogeneous systems but also allow for the description of phase equilibria, the relaxation of metastable states and the dynamics of processes at interfaces. In addition, MD can also model chemical reactions in complex environments.

MD simulations allow to model real time evolution of particles, one can then access time-dependent properties. If the time evolution is obtained by **integrating Newton's equations of motion** for a system of interacting particles, it is referred to as **classical MD**. A classical MD simulation needs the definition of a potential describing the interaction between particles in order to calculate the PES. Potentials can be of different levels of accuracy as described in the previous section. The most commonly used potentials in chemistry are **force fields**. In that case, one refers to **molecular mechanics**, which embodies a classical mechanics treatment of

the interactions between particles. In classical molecular dynamics, electrons and nuclei are not distinguished and one refers only to particles. As already mentioned, the main drawback of force fields is that they usually can not model chemical reactions. If the potential comes from a quantum chemical treatment of the electrons only, the nuclei being treated as interacting point charge particles, on refers to **ab initio molecular dynamics**. **Quantum dynamics** differs from classical molecular dynamics as the temporal evolution of a system is described by the time-dependent Schrödinger equation. For instance, the MCTDH approach[292, 293] and discrete-variable representations,[294, 295] are particularly accurate but are limited in terms of number of degrees of freedom. The reduction from a full quantum description of all particles, electrons and nuclei, to a classical treatment involves two main approximations. The first one is the BO approximation as described in section 2.2 which allows to treat separately electrons and nuclei. The second one treats the nuclei (much heavier than electrons) as point charge particles that follow classical Newtonian dynamics. In this thesis, **classical molecular dynamics was used** to perform simulations and use the term MD to denote classical molecular dynamics only.

Principles. In the classical formulation of MD, each particle or nucleus in the system is represented by a material point which interacts with all other particles via a potential defined by their positions. The principle of the dynamics of N atoms is to determine the forces $\mathbf{F}_\alpha$ acting on each of the particles in a given geometry and then calculate the accelerations and the velocities of the particles from these forces using Newton's second law:

$$\mathbf{F}_\alpha = m_\alpha \mathbf{a}_\alpha, \quad \alpha = 1,2...,N \tag{2.70}$$

where m_α is the mass of the atom α, $\mathbf{a}_\alpha$ being its acceleration and $\mathbf{F}_\alpha$ is the total force exerted on α. $\mathbf{F}_\alpha$ is defined as the derivative of the potential energy V of the system with respect to the corresponding position ($\mathbf{R}_\alpha$) of α:

$$\mathbf{F}_\alpha = -\frac{\partial V}{\partial \mathbf{R}_\alpha} = m_\alpha \frac{d^2\mathbf{R}_\alpha}{dt^2} \tag{2.71}$$

where V depends on the positions of all atoms or particles. This leads to a system of $f \times N$ second-order differential equations where f is the dimension of space. In our case, the degree of freedom f is equal to 3. According to the known initial positions of particles, the potential energy can be obtained. Then, a numerical resolution of the partial derivative equations provided in equation 2.71 can be obtained using a suitable **integration algorithm**. The

integration algorithm gives access to the positions and velocities of atoms or particles and to the forces acting on these atoms or particles over time. Here, the time is discretized in regular intervals and calculations are repeated at each time interval referred to as **time step**. Many high order integration algorithms have been proposed depending on the desired accuracy: Euler algorithm,[296, 297] Verlet algorithms, predictor-corrector algorithm,[298, 299] and Runge-Kutta algorithm.[300, 301, 302, 303] The Verlet algorithms include the Simple Verlet (SV),[304] the Leapfrog Verlet (LFV),[305] and the Velocity Verlet (VV).[306] The **Velocity Verlet algorithm** is the most widely used in many MD codes owing to its numerical stability and implementation simplicity. Furthermore, movement constants are very well preserved over time. We briefly describe the basis of the Velocity Verlet algorithm below.

The local error is quantified by inserting the exact values $\mathbf{R}_\alpha(t_{n-1})$, $\mathbf{R}_\alpha(t_n)$, and $\mathbf{R}_\alpha(t_{n+1})$ into the iteration and calculating the Taylor expansions at time $t = t_n$ of the position vector $\mathbf{R}_\alpha(t \pm \delta t)$ in different time directions:

$$\mathbf{R}_\alpha(t+\delta t) = \mathbf{R}_\alpha(t) + \mathbf{v}_\alpha(t)\delta t + \frac{\mathbf{a}_\alpha(t)\delta t^2}{2} + \frac{\mathbf{b}_\alpha(t)\delta t^3}{6} + O(\delta t^4) \tag{2.72}$$

$$\mathbf{R}_\alpha(t-\delta t) = \mathbf{R}_\alpha(t) - \mathbf{v}_\alpha(t)\delta t + \frac{\mathbf{a}_\alpha(t)\delta t^2}{2} - \frac{\mathbf{b}_\alpha(t)\delta t^3}{6} + O(\delta t^4) \tag{2.73}$$

where $\mathbf{v}_\alpha(t)$ is the velocity of α and $\mathbf{b}_\alpha(t)$ is the derivative of $\mathbf{a}_\alpha(t)$ with respect to the time. Through summing equations 2.72 and 2.73, we can get the Verlet integrator:

$$\mathbf{R}_\alpha(t+\delta t) = 2\mathbf{R}_\alpha(t) - \mathbf{R}_\alpha(t-\delta t) + \mathbf{a}_\alpha(t)\delta t^2 + O(\delta t^4) \tag{2.74}$$

We can notice that the first-order and third-order terms cancel out from the Taylor expansion, which makes the Verlet integrator more accurate than the integration by a Taylor expansion only.

We can see in equation 2.74 that the position propagation equation does not involve the velocities. They can be computed by the following finite difference:

$$\mathbf{v}_\alpha(t) = \frac{R_\alpha(t+\delta t) - R_\alpha(t-\delta t)}{2\delta t} + O(\delta t^2) \tag{2.75}$$

This provides velocities at time t and not at time $t + \delta t$, which means the velocity term is a step behind the position term. The use of equation 2.75 has the advantage of low data storage,

i.e. less memory is required, but a problem emerges in the calculation of the kinetic energy at time $t+\delta t$.

The VV algorithm is often applied to solve this problem as it allows velocities and positions to be computed simultaneously:

$$\mathbf{R}_\alpha(t+\delta t) = \mathbf{R}_\alpha(t) + \mathbf{v}_\alpha(t)\delta t + \frac{\mathbf{a}_\alpha(t)\delta t^2}{2} \tag{2.76}$$

$$\mathbf{v}_\alpha(t+\delta t) = \mathbf{v}_\alpha(t) + \frac{\mathbf{a}_\alpha(t)+\mathbf{a}_\alpha(t+\delta t)}{2}\delta t \tag{2.77}$$

The standard implementation of the VV algorithm is a four steps scheme: firstly to calculate equation 2.78, secondly to calculate equation 2.79, thirdly to derive $\mathbf{a}_\alpha(t+\delta t)$ from the interaction potential with $\mathbf{R}_\alpha(t+\delta t)$, finally to compute equation 2.80.

$$\mathbf{v}_\alpha(t+\frac{1}{2}\delta t) = \mathbf{v}_\alpha(t) + \frac{1}{2}\mathbf{a}_\alpha(t)\delta t \tag{2.78}$$

$$\mathbf{R}_\alpha(t+\delta t) = \mathbf{R}_\alpha(t) + \mathbf{v}_\alpha(t+\frac{1}{2}\delta t)\delta t \tag{2.79}$$

$$\mathbf{v}_\alpha(t+\delta t) = \mathbf{v}_\alpha(t+\frac{1}{2}\delta t) + \frac{1}{2}\mathbf{a}_\alpha(t+\delta t)\delta t \tag{2.80}$$

¡

Temperature is not a state variable in the simulations, an average kinetic temperature can be calculated from the average value of the kinetic energy.

At the end of each integration step, the VV algorithm gives directly access to $\mathbf{R}_\alpha(t+\delta t)$, $\mathbf{v}_\alpha(t+\delta t)$, and $\mathbf{F}_\alpha(t+\delta t)$. The VV algorithm needs to ensure two intrinsic properties of the classical equations of motion. One is temporal reversibility, the invariance of the trajectories at t and $-t$. This symmetry leads to the independence of he dynamics from the direction of time. The other property is the conservation of the Hamiltonian function over time. Because of the discretization of trajectories, this conservation can not be insured. A stable integration algorithm must impose this conservation for long enough time steps (δt) to allow for sufficiently long simulation times. The VV algorithm is able to do this due to its sufficient numerical stability. According to the conservation of energy, the natural ensemble corresponding to such dynamics is the microcanonical ensemble (N,V,E).[307, 308] N is the number of particles. V

denotes the volume and E is the energy of the system which is conserved. It is possible to extend the MD to other statistical ensembles by modifying the Hamiltonian of the system. For instance, in the canonical ensemble (N,V,T), a thermostat is added to the system allowing for the control of the temperature T.[309, 310, 311, 312] The temperature is not a state variable in the simulations, an average kinetic temperature can be calculated from the average value of the kinetic energy.

The force, *i.e.* acceleration, calculation in the MD simulations is an extremely important part as it determines both the accuracy of the potential energy and the computational cost. Different methods were presented in section 2.3 to calculate the forces. **In this thesis, the SCC-DFTB method has been applied to compute the forces in all the MD simulations.**

2.4.3 Parallel-Tempering Molecular Dynamics

In a number of cases, it is necessary to explore the PES as thoroughly as possible for the study of dynamical, thermodynamical, and structural properties of a given system. For atomic and molecular clusters, PES usually presents a large number of stable configurations linked together by energy barriers. Unfortunately, MD simulations cannot overcome these energy barriers in a reasonable simulation time even within the canonical ensemble (N,V,T). This leads to non-ergodic simulations that cannot be used to extract meaningful statistical averages. Actually, when a simulation explores the well of a PES, it may often be blocked in this well at low temperature because the energy barriers are too high to be crossed. In this case, if E_b refers to the energy barrier and T is the temperature of the simulation, one can consider that $E_b >> k_B T$ where k_B is the Boltzmann constant. At intermediate temperature, the possibility of crossing the energy barriers during a simulation increases, but this can not guarantee the PES to be explored exhaustively. For high temperatures, one has a high probability to cross the energy barriers whereas the bottoms of the wells can not be explored comprehensively. Therefore, it is not possible to both cross the energy barriers and thoroughly explore the bottom of the wells using a unique MD simulation at a given temperature. Many methods have been proposed to solve this question and are referred to as **enhanced sampling methods** They are classified into two groups: **biased methods** and **non-biased methods**. In **biased methods**, the dynamics of the system is influenced by a external factor, usually a non-physical force, which makes it possible to push the system outside of the wells even at low T.[313, 314, 315, 316, 317] For instance, Metadynamics is a biased method.[318, 319, 320] In **non-biased methods**, the dynamics of the system is not modified directly. Examples are simulated annealing,[321, 322]

and multi-replica approaches such as the **parallel-tempering molecular dynamics** approach, which has been used in this thesis.

The replica exchange approach also termed parallel-tempering was originally devised by R. H. Swendsen *et al*. in 1986 [323] then extended by Geyer and coauthor in 1991 [324] and was further developed by Hukushima and Nemoto,[325] M. Falcioni and M. W. Deem,[326] D. J. Earl and coworker.[327] Y. Sugita and Y. Okamoto formulated a MD version of parallel tempering to enhance conformational sampling.[328] PTMD is a method which aims at enhancing the ergodicity of MD simulations. The principle of PTMD is shown in Figure 2.4.

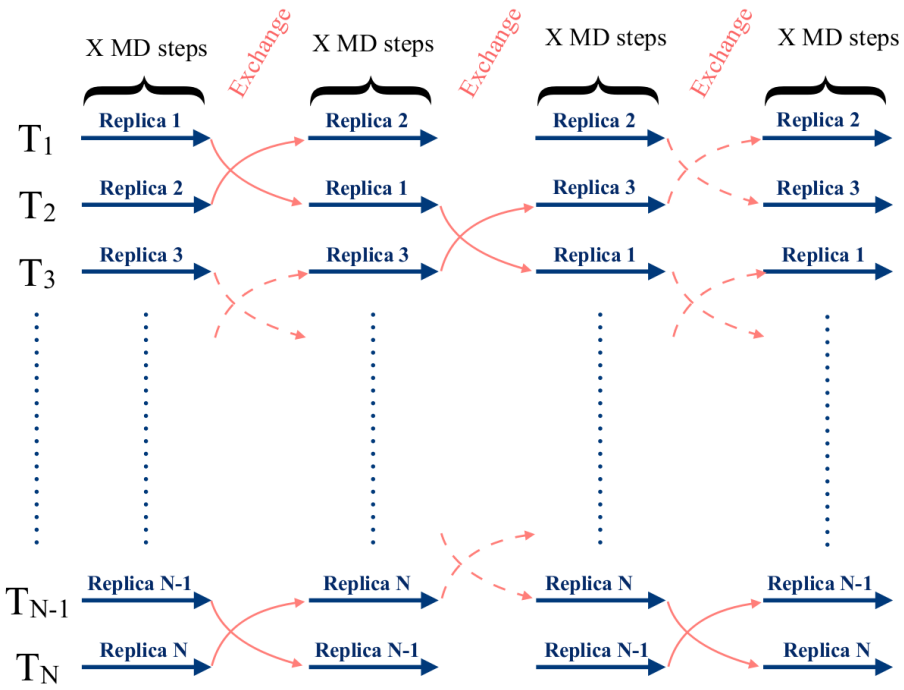


Figure 2.4: Schematics of the PTMD algorithm in its synchronous version. Replicas of the same system, numbered from C_1 to C_N, are simulated subject to different temperatures (from T_1 to T_N). Once X MD steps (straight solid arrows) have been performed by each replica, configuration exchanges are attempted between neighbouring simulations according to the Metropolis criterion. Some of them undergo successful swapping (solid curved arrows) while other not (dashed curved arrows). MD simulations then proceed for X additional MD steps before new attempts of exchange.

N replicas ($C_i, i = 1,2,...,N$) of the same system are simulated in parallel each one at a given temperature T_i, (i = 1, 2, ..., N) in the canonical ensemble. The time evolution of each replica is independent with each other but exchanges of configurations between adjacent repli-

cas C_i and C_j, where $T_i < T_j$ and $i = j - 1$ are permitted at regular time intervals. The choice of the extreme temperatures T_1 and T_N is very important for the algorithm to be optimal. The lowest temperature (T_1) should be the one at which usual simulations are blocked and the highest temperature (T_N) should be chosen so that all significant energy barriers can be overcome during the simulation. Moreover, the temperatures between T_1 and T_N must be chosen to lead to sufficient overlap between the density of states of the adjacent replicas. Indeed, if this overlap is too small, the probability of exchange is very low, which makes the PTMD simulations inefficient and leads to a bad exploration of the PES. In contrast, if the overlap is too large, a large amount of redundant information will be produced, which will cost unnecessary computational resources. Configurations between two neighbouring replicas at different T are exchanged based on the Metropolis–Hastings criterion with probability:

$$\rho(C_i \Leftrightarrow C_j) = \min\left(1, e^{(V_i - V_j)\left(\frac{1}{kT_i} - \frac{1}{kT_j}\right)}\right) = \min\left(1, e^{m}\right) \tag{2.81}$$

where V_i and V_j are the potential energies of replicas i and j, respectively. T_i and T_j are the temperatures of replicas i and j, respectively. If the energy of replica configuration C_j at high temperature T_j is lower than the energy of replica configuration C_i at low temperature T_i, which means the exponent (m) of e is positive, then the exchange is allowed. If m is negative, the exchange between neighbouring replicas is only allowed when e^m is greater than ω (a random value between 0 and 1). To accelerate the equilibration of the system after the exchange, the velocities of all particles can be renormalized as follows:

$$\mathbf{v}_\alpha^{\text{new}} = \left(\frac{T_{\text{new}}}{T_{\text{old}}}\right)^{\frac{1}{2}} \mathbf{v}_\alpha^{\text{old}} \tag{2.82}$$

2.4.4 Global Optimization

Global Optimization refers to the determination of the lowest energy point on a PES, *i.e.* the global minimum. As this latter usually includes a large number of stationary points, it is not straightforward to find the global minimum. Local optimization methods do not make it possible to cross the energy barriers between local minima. Therefore, a global optimization scheme such as MD or Monte Carlo simulations is needed to perform a more exhaustive exploration of the PES to get to the lowest energy minimum.

There exits a vast amount of methods to perform global optimization and each one has its strength and weaknesses. For instance, the Basin-Hopping method is a particular useful global optimization technique in high-dimensional landscapes that iterates by performing a random

perturbation of coordinates, making a local optimization, and rejecting or accepting new coordinates based on a minimized function value.[58, 60] Genetic algorithms are also among the most used methods to find a global minimum.[329, 330, 331, 332] A genetic algorithm is inspired by the process of natural selection. Genetic algorithms are usually applied to generate high-quality solutions of optimization. The ergodicity problem appears in all of these global optimization methods. In principle, one can only be sure of having found the real global minimum after an infinite number of iterations.

In order to avoid ergodicity problems, it is interesting to combine global and local optimization methods. A very popular combination is the simulated annealing method combined with local optimizations. The PES of molecular aggregates display many degrees of freedom and contains a large number of low-energy isomers. The PTMD algorithm coupled with a great number of local optimizations is a good choice to search for low-energy structures in this kind of system. Local optimizations are performed many times from initial conditions structures which are extracted from all PTMD trajectories, whether it be low or high temperature, in order to maximize sampling. **This approach, in combination with SCC-DFTB, has been conducted along this thesis to perform global Optimization.**

3

Exploration of Structural and Energetic Properties

This **third chapter** of my thesis merges two independent studies dealing with the determination of the low-energy isomers of ammonium/ammonia water clusters, $(H_2O)_nNH_4^+$ and $(H_2O)_nNH_3$, and protonated uracil water clusters, $(H_2O)_nUH^+$. As highlighted in the general introduction of this thesis and in chapter 2, performing global optimization of molecular clusters is not straightforward. The two studies presented in this chapter thus share a main common methodology which is the combination of the **self-consistent-charge density functional based tight-binding** (SCC-DFTB) method for the efficient calculation of the potential energy surfaces (PES) and the **parallel-tempering molecular dynamics** (PTMD) approach for their exploration. All low-energy isomers reported in this chapter are discussed in terms of structure, relative energy and binding energy which are compared to the literature when available. Calculations at higher level of theory are also performed to refine the results obtained at the SCC-DFTB level or to validate the results it provides. In particular, in this chapter, an improve set of parameters is proposed to describe sp^3 nitrogen containing compounds at the SCC-DFTB level. Our results are also used to complement collision-induced dissociation experiments performed by S. Zamith and J.-M. L'Hermite at the *Laboratoire Collisions Agrégats Réactivité* (LCAR).

3.1 Computational Details

3.1.1 SCC-DFTB Potential

SCC-DFTB electronic structure calculations presented in this chapter were all performed with the deMonNano code.[333] The details of the method are presented in section 2.3.3 of chapter 2. The mio-set for the Slater-Koster tables of integrals was used.[168] However, it has been shown that these integrals do not properly described sp^3 hybridized nitrogen, in particular, proton affinity.[334] Consequently, in order to avoid spurious deprotonation of the sp^3 hybridized nitrogen in NH_4^+ and to correctly reproduce binding energies calculated at the MP2/Def2TZVP level, I propose to modify the original mio-set for Slater-Koster tables of N-H integrals by applying them a multiplying factor. Several of them were tested and I present here the results obtained for two of them: 1.16 and 1.28. For the sp^2 nitrogen of uracil, the original integrals of the mio-set were used. To improve description of the intermolecular interactions, the original Mulliken charges were replaced by the CM3 charges,[230, 233, 335] (see equation 2.52 in section 2.3.3) and an empirical correction term (see equation 2.54 in section 2.3.3) was used to describe dispersion interactions.[233, 239, 336] A. Simon *et al.* developed a SCC-DFTB potential that leads to geometries, frequencies, and relative energies close to the corresponding experimental and CCSD(T)/aug-cc-pVTZ results.[337, 338] The corresponding D_{OH} parameter, *i.e.* 0.129, is retained in the studies presented in thus chapter. D_{NH} is tested in the study of ammonium/ammonia water clusters and two values were retained and thoroughly tested: 0.12 and 0.14. D_{NO} is set to zero.

3.1.2 SCC-DFTB Exploration of PES

To determine the lowest-energy isomers of $(H_2O)_{1-10,20}NH_4^+$, $(H_2O)_{1-10}NH_3$ and $(H_2O)_{1-7,11,12}UH^+$ clusters, their PES were thoroughly explored using PTMD [327, 328, 339] simulations combined with a SCC-DFTB [168] description of the energies and gradients. I describe below the **detailed parameters used for all the simulations conducted within this chapter**.

Detailed parameters for PTMD simulations of $(H_2O)_{1-10,20}NH_4^+$ and $(H_2O)_{1-10}NH_3$ clusters are as follows. For $(H_2O)_{1-3}NH_4^+$ and $(H_2O)_{1-3}NH_3$ clusters, 16 replicas were used with a linear distribution of temperatures with a 15 K step ranging from 10 to 250 K. 40 replicas with a 6 K step ranging from 10 to 250 K were considered for $(H_2O)_{4-10,20}NH_4^+$ and $(H_2O)_{4-10}NH_3$ species. All trajectories were 5 ns long and a time step of 0.5 fs was used

to integrate the equations of motion. A Nosé-Hoover chain of 5 thermostats was employed for all the simulations to achieve simulations in the canonical ensemble.[310, 311] Thermostat frequencies were fixed at 400 cm^{-1}. To identify low-energy isomers of $(H_2O)_{1-3}NH_4{}^+$ and $(H_2O)_{1-3}NH_3$ clusters, 303 geometries were periodically selected from each replicas and further optimized at the SCC-DFTB level, which produced 4848 optimized geometries per cluster. For $(H_2O)_{4-10,20}NH_4{}^+$ and $(H_2O)_{4-10}NH_3$ clusters, 500 geometries were periodically selected from each replicas leading to 20000 optimized geometries per cluster. For $(H_2O)_{20}NH_4{}^+$, the initial structure used for the global optimization process was the lowest-energy structure reported by J. Douady *et al.*.[340] The five lowest-energy isomers among the 4848 or 20000 optimized geometries were further optimized using the MP2/Def2TZVP method. See below for the details on MP2/Def2TZVP calculations.

Detailed parameters for PTMD simulations of $(H_2O)_{1-7,11,12}UH^+$ clusters are as follows. 40 replicas with temperatures rnaging linearly from 50 to 350 K were used. Each trajectory was 4 ns long, and the integration time step was 0.5 fs. A reasonable time interval for the PT exchanges was 2.5 ps. A Nosé-Hoover chain of five thermostats with frequencies of 800 cm^{-1} was applied to achieve an exploration in the canonical ensemble.[310, 311] To avoid any spurious influence of the initial geometry on the PES exploration, three distinct PTMD simulations were carried out with distinct initial proton location: on the uracil in two cases and on a water molecule in the other one. In the former cases, I used two isomers u178 and u138 of UH^+ shown in Figure 3.1 as the initial geometries.[341, 342] 600 geometries per temperature were linearly selected along each PTMD simulation for subsequent geometry optimization leading to 72000 structures optimized at SCC-DFTB level. These structures were sorted in ascending energy order and checked for redundancy. 9, 23, 46, 31, 38, 45, 63, 20, and 29 structures were then selected for $(H_2O)UH^+$, $(H_2O)_2UH^+$, $(H_2O)_3UH^+$, $(H_2O)_4UH^+$, $(H_2O)_5UH^+$, $(H_2O)_6UH^+$, $(H_2O)_7UH^+$, $(H_2O)_{11}UH^+$ and $(H_2O)_{12}UH^+$ respectively, to perform geometry optimizations at the MP2/Def2TZVP level. See below for the details on MP2/Def2TZVP calculations.

3.1.3 MP2 Geometry Optimizations, Relative and Binding Energies

Some low-energy isomers obtained at the SCC-DFTB level were further optimized at the MP2 level of theory in combinationwith an all electron Def2TZVP basis-set.[343, 344] All calculations used a a tight criteria for geometry convergence and an ultrafine grid for the numerical integration. All MP2 calculations were performed with the Gaussian 09 package.[345]

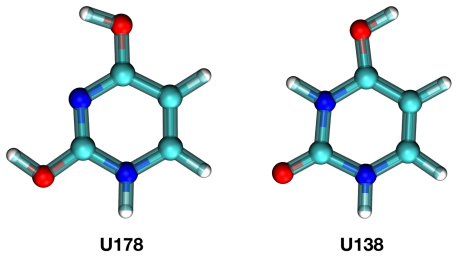


Figure 3.1: Structures of the two protonated uracil isomers, u178 (keto-enol form) and u138 (di-keto form), used as initial conditions in the PTMD simulations.

Detailed parameters for $(H_2O)_{1-10,20}NH_4^+$ and $(H_2O)_{1-10}NH_3$ clusters. Following SCC-DFTB optimizations, the five lowest-energy isomers of $(H_2O)_{1-10}NH_4^+$ and $(H_2O)_{1-10}NH_3$ clusters were further optimized at the MP2/Def2TZVP level of theory. In section 3.2, relatives energies with respect to the lowest-energy isomer of each cluster are reported. Impact of zero-point vibrational energy (ZPVE) corrections on relative energies were evaluated at MP2/Def2TZVP level. To evaluate the strength of water-ammonium and water-ammonia interactions and to assess the accuracy of the SCC-DFTB method, binding energies are also reported. Two distinct approaches were used to calculate binding energies. The first one considers only the binding energy between the water cluster as a whole and the impurity, NH_4^+ or NH_3, while the second one considers the binding energy between all the molecules of the cluster. In both cases, the geometry of the molecules is the one found in the optimized cluster. Using these two methods, relative binding energies ($E_{bind.}$(SCC-DFTB)-$E_{bind.}$(MP2/Def2TZVP)) $\Delta E_{bind.}^{whole}$ and $\Delta E_{bind.}^{sep.}$ were obtained. For all binding energies of $(H_2O)_{1-10}NH_4^+$ and $(H_2O)_{1-10}NH_3$ clusters calculated at MP2/Def2TZVP level, basis set superposition errors (BSSE) correction was considered by using the counterpoise method of Boys and Bernardi.[346]

Detailed parameters for $(H_2O)_{1-7,11,12}UH^+$ clusters. Following SCC-DFTB optimizations, the six lowest-energy isomers of $(H_2O)_{1-7,11,12}UH^+$ clusters were further optimized at the MP2/Def2TZVP level of theory. The binding and relative energies calculated at MP2/Def2TZVP level without BSSE correction of clusters $(H_2O)_{2-7,11,12}UH^+$ are discussed in section 3.3.

3.1.4 Structure Classification

For clusters $(H_2O)_{1-10}NH_4^+$ and $(H_2O)_{1-10}NH_3$, "n-x"and "n′-x"labels are used to distinguish between the $(H_2O)_nNH_4^+$ and $(H_2O)_nNH_3$ reported isomers, respectively, obtained at the SCC-DFTB level. For comparison, "n-x"and "n′-x"isomers are also optimized at the

MP2/Def2TZVP level. In that case, the resulting structures are referred to as "n-x*"and "n′-x*" to distinguish them more easily although they display the same general topology as "n-x"and "n′-x"isomers. In these notations, n and n′ denote the number of water molecules in the ammonium and ammonia water clusters, respectively. x is an alphabetic character going from a to e that differentiates between the five low-energy isomers reported for each cluster in ascending energy order, *i.e.* a designates the lowest-energy isomer.

3.2 Structural and Energetic Properties of Ammonium/Ammonia including Water Clusters

3.2.1 General introduction

Water clusters play an important role in various areas such as atmospheric and astrochemical science, chemistry and biology.[118, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359] They are involved into the critical stages of nucleation and growth of water-containing droplets in the atmosphere thus contributing to the physical and chemical properties of this medium.[360] In many cases, the presence of chemical impurities interacting with water aggregates strongly affect their properties. For instance, ammonia is an important compound commonly found in the atmosphere and which displays a key role in aerosol chemistry.[361] Its high basicity makes it a potential proton sink that can form a ionic center for nucleation.[362, 363] E. Dunne and co-workers also reported that most nucleation occurring in the atmosphere involves ammonia or biogenic organic compounds, in addition to sulfuric acid.[113] J. Kirkby *et al.* also found that even a small amount of atmospherically relevant ammonia can increase the nucleation rate of sulphuric acid particles by several orders of magnitude.[112] The significance of ammonium and ammonia water clusters have thus motivated a large amount of experimental and theoretical studies during the past decades.[340, 362, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374]

As a few examples, in 1984, $(H_2O)_2NH_4^+$ was identified using mass spectrometry by M. D. Perkin *et al.*[362] In 1997, Stenhagen and co-workers studied the $(H_2O)_{20}H_3O^+$ and $(H_2O)_{20}NH_4^+$ clusters and found that both species display a similar structure.[366] P. Hvelplund *et al.* later reported a combined experimental and theoretical study devoted to protonated mixed ammonia/water which highlighted the idea that small protonated mixed clusters of water and ammonia contain a central NH_4^+ core.[370] Theoretical calculations devoted to ammonium and ammonia water clusters have also been extensively conducted.[368, 369, 370,

373, 374, 375, 376, 377, 378] Among them, J. Novoa *et al.* studied the $(H_2O)_4NH_3$ aggregate and found the existence of a minimum in its potential energy surface corresponding to a $(H_2O)_3\cdots NH_4^+\cdots OH^-$ structure, resulting from one proton transfer from a water molecule to the ammonia molecule.[375] D. Bacelo later reported a number of low-energy minima for $(H_2O)_{3-4}NH_3$ clusters obtained from *ab initio* calculation and a Monte Carlo exploration of the PES.[373] More recently, J. Douady *et al.* performed a global optimization of $(H_2O)_nNH_4^+$ (n = 1-24) clusters again using a Monte Carlo procedure in combination with a Kozack and Jordan empirical force field.[371, 379] In this study, the finite temperature properties as well as vibrational signature of several clusters thus highlighting the key contribution of simulations in understanding such species. Morrell and Shields also studied the $(H_2O)_nNH_4^+$ (n = 1-10) aggregates *via* a mixed molecular dynamics and quantum mechanics methodology to calculate energies and free energies of formations which were in good agreement with previous experimental and theoretical results.[372] More recently, S. Pei *et al.* determined that $(H_2O)_nNH_4^+$ clusters start to adopt a closed-cage geometry at n=8.[380] Finally, W. Walters and collaborators determined the geometry of $(H_2O)_{16}NH_3$ and $(H_2O)_{16}NH_4^+$ at the HF/6-31G(d) level, and observed strong hydrogen bonding between water and the lone pair of NH_3 and bewteen NH_4^+ and the four adjacent water molecules.[381]

As for the study of other molecular clusters, the range of applicability of theoretical simulations to describe ammonium and ammonia water clusters is dictated by the balance between accuracy, transferability and computational efficiency. While *ab-initio* methods can accurately model small aggregates, their application to large species is more difficult, in particular when an exhaustive exploration of the PES is required. In contrast, force-field potentials are computationally extremely efficient and can be coupled to global optimization methods but their transferability is limited. The SCC-DFTB approach can be seen as an intermediate approach which combines the strengths of both *ab-initio* and force-field methods. Indeed, it can be as accurate as DFT while computationally more efficient and is more transferable than force fields (see chapter 2) In recent years, SCC-DFTB has been successfully applied to the study of various molecular clusters: pure, protonated, and de-protonated water clusters,[62, 382, 383, 384] water clusters on PAHs,[337, 385] sulfate-containing water clusters,[103] water clusters in an argon matrix,[130] whether it is for global optimization or for the study of finite-temperature properties. However, in its original formulation, SCC-DFTB does not provide good results for the description of ammonia and ammonium as nitrogen hybridization seems to be a problem for minimal basis-set methods like SCC-DFTB.[386] M. Elstner and coworkers found consistent

errors (about 14.0 kcal.mol^{-1}) for deprotonation energies of sp^3 hybridized nitrogen containing systems, whereas sp^1 and sp^2 systems display much smaller errors.[334]

In this section, I first propose an improvement of the SCC-DFTB scheme to describe ammonium and ammonia water clusters by modifying both Hamiltonian and overlap N-H integrals and introducing optimized atomic charges.[121, 335] By combining this improved SCC-DFTB scheme with PTMD simulations, global optimization of the $(H_2O)_{1-10}NH_4^+$ and $(H_2O)_{1-10}NH_3$ clusters is then performed which allows to report a number of low-energy isomers for these species. Among them, a selected number of structures are further optimized at the MP2/Def2TZVP level of theory to confirm they are low-energy structures of the PES and to rationalize the difference in relative energy between both methods. A detailed description of the reported low-energy isomers is then provided as well as comparisons with the literature. The heat capacity curve of $(H_2O)_{20}NH_4^+$ is also obtained at the SCC-DFTB level and compared to previously published simulations. Some conclusions are finally presented. A very small part of this work has been published in 2019 in a review in*Molecular Simulation*.[384] A full paper devoted to this work is in preparation.

3.2.2 Results and Discussion

3.2.2.1 Dissociation Curves and SCC-DFTB Potential

In order to define the best SCC-DFTB parameter to model ammonia and ammonium water clusters, I have tested various sets of corrections. Each correction involves two modifications of the potential, the first one is the CM3 charge parameter D_{NH} and the second one is the multiplying factor, noted xNH, applied to the NH integrals in the Slater-Koster tables. So a given set is noted D_{NH}/xNH. Two sets of corrections have provided satisfactory results, 0.12/1.16 and 0.14/1.28. Figure 3.2 and 3.3 present dissociation curves obtained at the MP2/Def2TZVP, MP2/Def2TZVP with BSSE correction, original SCC-DFTB, SCC-DFTB 0.14/1.28 and SCC-DFTB 0.12/1.16 levels of theory. These curves are obtained using the same set of geometries regardless of the method applied to calculate the binding energies. They are obtained from the MP2/Def2TZVP optimized structures in which the distance between the water and the ammonium/ammonia was shifted along the N–O vector, all other geometrical parameters being kept fixed.

From Figure 3.2, the five curves display the same trends with a minimum located at almost the same N—O distance. At the curve minimum, binding energies vary between -25.57

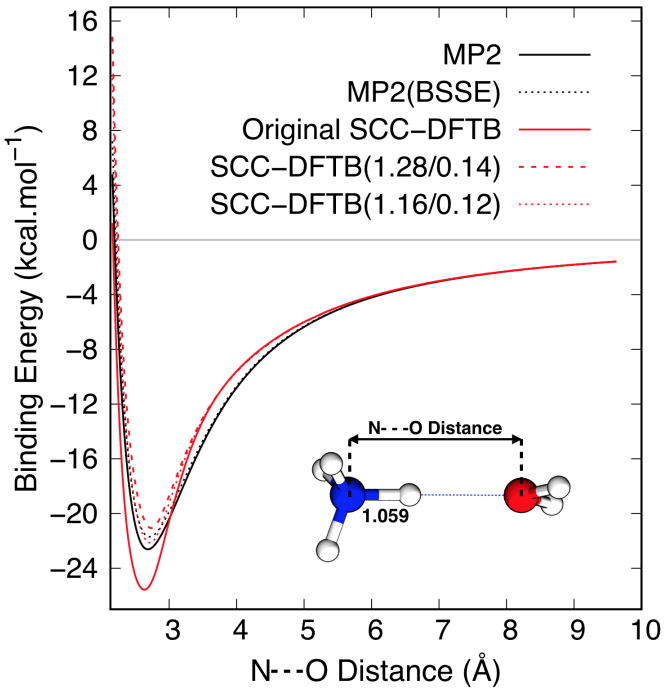


Figure 3.2: Binding energies of $(H_2O)NH_4^+$ as a function of the N—O distance at MP2/Def2TZVP (plain black), MP2/Def2TZVP with BSSE correction (dotted black), original SCC-DFTB (plain red), SCC-DFTB (0.14/1.28) (dotted red) and SCC-DFTB (0.12/1.16) (dashed red) levels of theory.

and -21,07 kcal.mol^{-1} at the original SCC-DFTB and SCC-DFTB 0.14/1.28 levels, respectively. The binding energy obtained at the SCC-DFTB 0.12/1.16 level is the closest to that obtained at MP2/Def2TZVP level with BSSE correction with a binding energy difference of only 0.47 kcal.mol^{-1}. The SCC-DFTB 0.14/1.28 curve is also very close with a difference in binding energy only 0.16 kcal.mol^{-1} higher. It is worth mentioning that both sets of corrections lead to improved results as compared to the original SCC-DFTB parameters which leads to a too low binding energy as compared to MP2/Def2TZVP level with BSSE correction. Also the position of the minimum is more shifted at the original SCC-DFTB level (2.64 Å) than with corrections (2.73 Å). So from structural and energetic point of views, both sets of corrections are satisfactory.

From Figure 3.3, the five curves display significant differences. This effect is accentuated by smaller binding energy values: they vary from -3.82 to -7,39 kcal.mol^{-1} at the original SCC-DFTB and MP2/Def2TZVP levels, respectively, at the minimum of the curves. The binding energy obtained at the SCC-DFTB 0.12/1.16 level is the closest to that obtained at

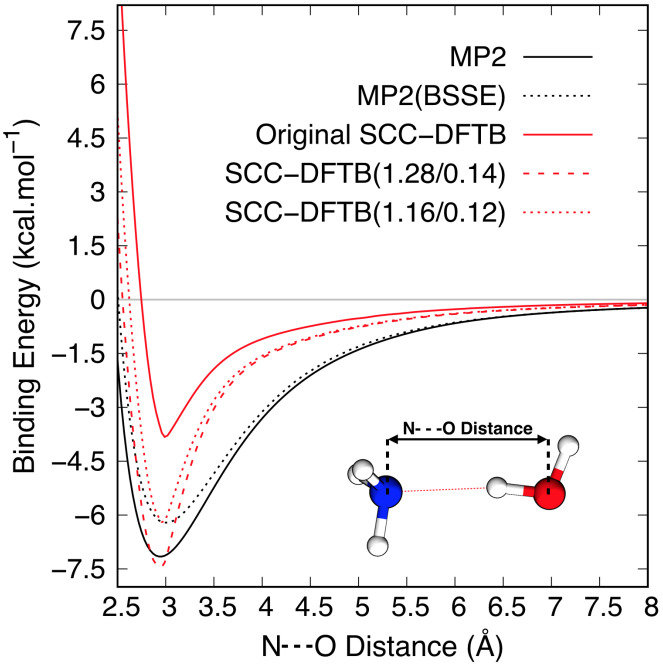


Figure 3.3: Binding energies of $(H_2O)NH_3$ as a function of the N—O distance at MP2/Def2TZVP (plain black), MP2/Def2TZVP with BSSE correction (dotted black), original SCC-DFTB (plain red), SCC-DFTB (0.14/1.28) (dotted red) and SCC-DFTB (0.12/1.16) (dashed red) levels of theory.

MP2/Def2TZVP level with BSSE correction with a binding energy difference of only 0.01 kcal.mol^{-1}. The SCC-DFTB 0.14/1.28 curve is also rather close with a difference in binding energy only 1.3 kcal.mol^{-1} higher. Here also, both sets of corrections lead to improved results as compared to the original SCC-DFTB parameters. The position of the minimum is also well reproduced by the corrected potentials. In contrast to $(H_2O)NH_4^+$, the shape of the curves for $(H_2O)NH_3$ obtained at the SCC-DFTB level differs significantly from those obtained at MP2 level. Vibrational frequencies calculated at the SCC-DFTB level for this systems are therefore expected to be inacurate. It is worth mentioning that the large difference in binding energy between $(H_2O)NH_4^+$ and $(H_2O)NH_3$ was expected owing to a stronger electrostatic contribution of NH_4^+ to the binding energy.

Another very important point when comparing the original SCC-DFTB potential and the corrected potentials, is the structure obtained for the $(H_2O)NH_4^+$ dimer. Figure 3.4 compares the structure obtained from geometry optimization at the SCC-DFTB 0.14/1.28 and original SCC-DFTB levels. The N-H covalent bond involved in the hydrogen bond is longer with the original potential while the N—O distance is smaller by 0.14 Å. This is reminiscent of the too

low proton affinity of NH_4^+ predicted by the original SCC-DFTB potential. This discrepancy has been previously highlighted in other studies,[334, 386] and makes this potential unusable in any realistic molecular dynamics simulation as it leads to a spurious deprotonation. Both sets of corrections are free of this error.

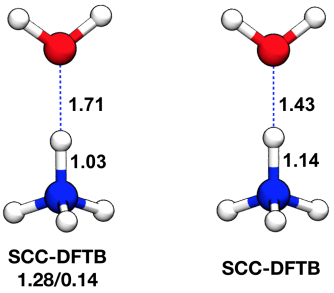


Figure 3.4: Structure of $(H_2O)NH_4^+$ obtained from geometry optimization at the SCC-DFTB 0.14/1.28 (right) and original SCC-DFTB (left) levels.

Figures 3.2 and 3.3 show that SCC-DFTB 0.12/1.16 better describe both $(H_2O)NH_3$ and $(H_2O)NH_4^+$ dissociation curves. Furthermore, as $(H_2O)NH_3$ is characterized by a much lower binding energy than $(H_2O)NH_4^+$, an error of the order of $\sim$1.0 kcal.mol^{-1} is more likely to play a significant role for ammonia than ammonium containing species. All the following discussion therefore involve the SCC-DFTB 0.12/1.16 potential.

3.2.2.2 Small Species: $(H_2O)_{1-3}NH_4^+$ and $(H_2O)_{1-3}NH_3$

As a first test case for the application of the SCC-DFTB 0.14/1.28 potential is the study of small ammonium and ammonia water clusters: $(H_2O)_{1-3}NH_4^+$ and $(H_2O)_{1-3}NH_3$. Due to the limited number of low-energy isomers for these species, we only consider the lowest-energy isomer of $(H_2O)_{1-2}NH_4^+$ and $(H_2O)_{1-3}NH_3$ and the two lowest-energy isomers for $(H_2O)_3NH_4^+$. As displayed in Figure 3.5 and3.6, the reported low-energy isomers 1-a, 1′-a, 2-a, 2′-a, 3-a, 3-b, and 3′ display a structure very similar to those obtained at the MP2/Def2TZVP level (1-a*, 1′-a*, 2-a*, 2′-a*, 3-a*, 3-b* and 3′-a*). Indeed, although differences in bond lengths are observed, they are rather small. In terms of energetics,

From en energetic point of view, it is interesting to first look at the relative energy between the two reported isomers of $(H_2O)_3NH_4^+$. Isomer 3-b is 2.12 kcal·mol^{-1} higher than 3-a at the SCC-DFTB level. At the MP2/Def2TZVP level, 3-b is 0.30 kcal·mol^{-1} lower than 3-a when ZPVE is not considered while it is 1.21 kcal·mol^{-1} higher when it is considered. In comparison,

in the experimental results by H. Chang and co-workers, 3-a is more stable than 3-b.[367, 369] The authors also complemented their measurements by theoretical calculations that show that at the B3LYP/6-31+G(d) level, 3-a is higher than 3-b but. In contrast, at the MP2/6-31+G(d) level corrected with ZPVE, the energy of 3-a is lower than that of 3-b while it is inverted if ZPVE is taken into account.[367, 369] Additionally, F. Spiegelman and co-workers, conducted a global Monte Carlo optimizations with an intermolecular polarizable potential that lead to 3-a as lowest-energy isomer.[371] All these results show that for the specific question of lowest-energy isomer of $(H_2O)_3NH_4^+$, SCC-DFTB has an accuracy close to other *ab initio* methods which confirms its applicability.

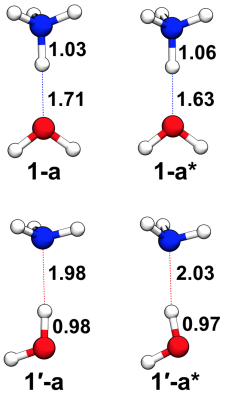


Figure 3.5: Structure of 1-a and 1′-a isomers obtained at the SCC-DFTB level and corresponding structures obtained at MP2/Def2TZVP level (1-a* and 1′-a* isomers). Selected bond lengths are in Å.

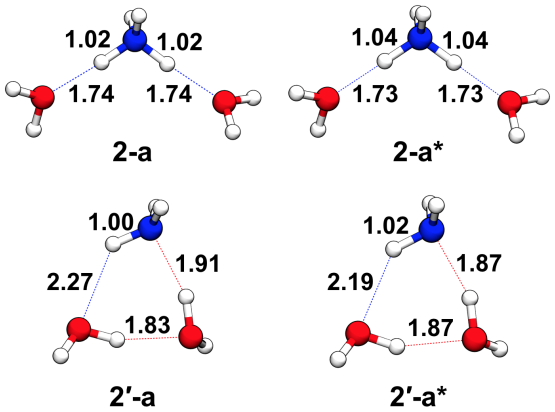


Figure 3.6: Structure of 2-a and 2′-a isomers obtained at the SCC-DFTB level and corresponding structures obtained at MP2/Def2TZVP level (2-a*, 2′-a* isomers). Selected bond lengths are in Å.

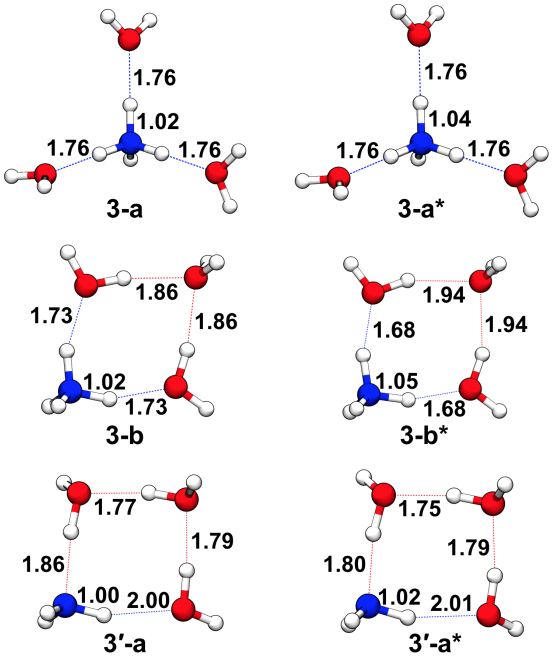


Figure 3.7: Structure of 3-a, 3-b and 3′-a isomers obtained at the SCC-DFTB level and corresponding structures obtained at MP2/Def2TZVP level (3-a*, 3-b* and 3′-a* isomers). Selected bond lengths are in Å.

As listed in Table 3.1, the relative binding energies $\Delta E_{bind.}^{whole}$ or $\Delta E_{bind.}^{sep.}$ of $(H_2O)NH_4^+$ and $(H_2O)NH_3$ are 1.21 and -1.17 kcal·mol^{-1}, respectively, which again highlights that SCC-DFTB is in agreement with MP2/Def2TZVP. For $(H_2O)NH_3$, the negative value show that MP2/Def2TZVP binding energy is smaller than the SCC-DFTB value. This is inverse to what is shown in Figure 3.3 and results from structural reorganization after optimization. All other values of Table 3.1 are equal of smaller than these values, whether considering $\Delta E_{bind.}^{whole}$ or $\Delta E_{bind.}^{sep.}$, which again demonstrates that the presently proposed SCC-DFTB potential provides results in line with reference MP2/Def2TZVP calculations.

$(H_2O)_{4-10}NH_4^+$ clusters have been studied by molecular dynamic and Monte Carlo simulations in combination with DFT and MP2 approaches although these latter are computationally expensive.[340, 367, 369, 371, 372, 387] In contrast, to the best of our knowledge, no theoretical calculation about $(H_2O)_{5-10}NH_3$ clusters have been conducted. The low computational cost of SCC-DFTB and its seemingly good performances on small clusters provide an appealing opportunity to thoroughly explore the PES of both large ammonium and ammonia containing water clusters. In the following section, the five lowest-energy isomers of clusters

Table 3.1: Relative binding energies $\Delta E_{bind.}^{whole}$ and $\Delta E_{bind.}^{sep.}$ of the low-energy isomers of $(H_2O)_{1-3}NH_4^+$ and $(H_2O)_{1-3}NH_3$ clusters. Values are given in kcal.mol^{-1}.

$(H_2O)_nNH_4^+$	$\Delta E_{bind.}^{whole}$	$\Delta E_{bind.}^{sep.}$	$(H_2O)_nNH_3$	$\Delta E_{bind.}^{whole}$	$\Delta E_{bind.}^{sep.}$
1-a	1.21	1.21	1′-a	-1.17	-1.17
2-a	0.82	0.91	2′-a	0.57	0.28
3-a	-0.25	0.11	3′-a	0.91	0.01
3-b	1.21	-0.15	-	-	-

$(H_2O)_{4-10}NH_4^+$ and $(H_2O)_{4-10}NH_3$ are presented and discussed in details.

3.2.2.3 Properties of $(H_2O)_{4-10}NH_4^+$ Clusters

The five lowest-energy isomers of $(H_2O)_4NH_4^+$ are depicted in Figure 3.8. 4-a is the lowest-energy isomer obtained from the global SCC-DFTB optimization and also the lowest-energy configuration after optimization at MP2/Def2TZVP level with ZPVE corrections. This result is consistent with previous computational studies[367, 369, 371, 387, 388] and the experimental studies by H. Chang and co-workers.[367, 368] Isomer 4-a displays four hydrogen bonds around the ionic center which lead to no dangling N-H bonds. Other isomers of comparable stability are displayed in Figure 3.8 The energy ordering of 4-a to 4-e at SCC-DFTB level is consistent with that at MP2/Def2TZVP level with ZPVE correction, although they are slightly higher by$\sim$2.0 kcal.mol^{-1}. Isomer 4-c was not reported in H. Chang's study,[369] and the corresponding energy ordering of the five lowest-energy isomers was the same as ours which certainly results from the use of a different basis set.

The relative binding energy of SCC-DFTB method to MP2/Def2TZVP method with BSSE correction for isomers 4-a to 4-e are listed in Table 3.2. When the four water molecules are considered as a whole part to calculate the binding energy, the relative binding energy of isomers 4-a to 4-e are -1.67, 0.00, 0.77, 0.77 and -4.04 kcal·mol^{-1}. As shown in Table 3.2, for isomers 4-a to 4-e, when the four water molecules are separately considered using the geometry in the cluster to calculate the binding energy, the biggest absolute value of the relative binding energy is 0.87 kcal·mol^{-1}. This shows the results of SCC-DFTB are in good agreement with those of MP2/Def2TZVP with BSSE correction for $(H_2O)_4NH_4^+$. From the relative binding energy of $(H_2O)_4NH_4^+$, it indicates that all the water molecules considered as a whole part or separately

Table 3.2: Relative binding energies $\Delta E_{bind.}^{whole}$ and $\Delta E_{bind.}^{sep.}$ of the five lowest-energy isomers of $(H_2O)_{4-10}NH_4^+$ and $(H_2O)_{4-10}NH_3$. Binding energies are given in kcal·mol^{-1}.

$(H_2O)_nNH_4^+$	$\Delta E_{bind.}^{whole}$	$\Delta E_{bind.}^{sep.}$	$(H_2O)_nNH_3$	$\Delta E_{bind.}^{whole}$	$\Delta E_{bind.}^{sep.}$
4-a	-1.67	-0.87	4′-a	-1.11	-1.76
4-b	0.00	0.61	4′-b	-0.29	-1.62
4-c	0.77	0.44	4′-c	-0.29	-1.38
4-d	0.77	0.42	4′-d	1.08	-0.49
4-e	-4.04	0.69	4′-e	1.02	-1.07
5-a	-1.62	0.56	5′-a	0.82	-1.78
5-b	0.72	0.48	5′-b	-0.23	-2.26
5-c	0.69	0.55	5′-c	-0.34	-2.50
5-d	-1.08	-0.78	5′-d	-0.59	-1.84
5-e	-2.08	0.88	5′-e	-0.38	-2.60
6-a	-1.71	-0.38	6′-a	-0.27	-3.05
6-b	-1.14	-0.76	6′-b	-0.31	-3.55
6-c	-2.06	0.27	6′-c	-1.11	-4.67
6-d	-2.90	-1.06	6′-d	-0.05	-4.44
6-e	-1.18	-0.60	6′-e	0.55	-1.96
7-a	-2.95	-0.39	7′-a	1.09	-2.02
7-b	-2.92	-0.38	7′-b	-0.02	-4.07
7-c	-2.17	0.09	7′-c	-0.40	-4.15
7-d	-1.28	-1.35	7′-d	-0.14	-3.10
7-e	-3.22	-2.27	7′-e	-1.11	-4.32
8-a	-2.20	-1.63	8′-a	-1.12	-4.41
8-b	-1.61	-2.01	8′-b	-0.10	-3.04
8-c	-3.71	-1.17	8′-c	-0.41	-4.46
8-d	-2.43	-0.36	8′-d	0.20	-3.68
8-e	-0.55	0.35	8′-e	-1.28	-4.75
9-a	-2.02	-1.39	9′-a	-0.15	-4.47
9-b	0.51	-0.84	9′-b	-1.01	-4.45
9-c	-3.31	-0.85	9′-c	-1.04	-4.42
9-d	-1.58	-1.78	9′-d	-1.09	-5.14
9-e	-2.39	-0.91	9′-e	0.41	-2.57
10-a	-2.64	-1.94	10′-a	-0.03	-4.80
10-b	-5.79	-4.35	10′-b	0.13	-5.61
10-c	-1.26	-2.36	10′-c	-0.62	-6.50
10-d	-1.98	-1.42	10′-d	-1.10	-6.30
10-e	-7.17	-1.54	10′-e	0.23	-8.36

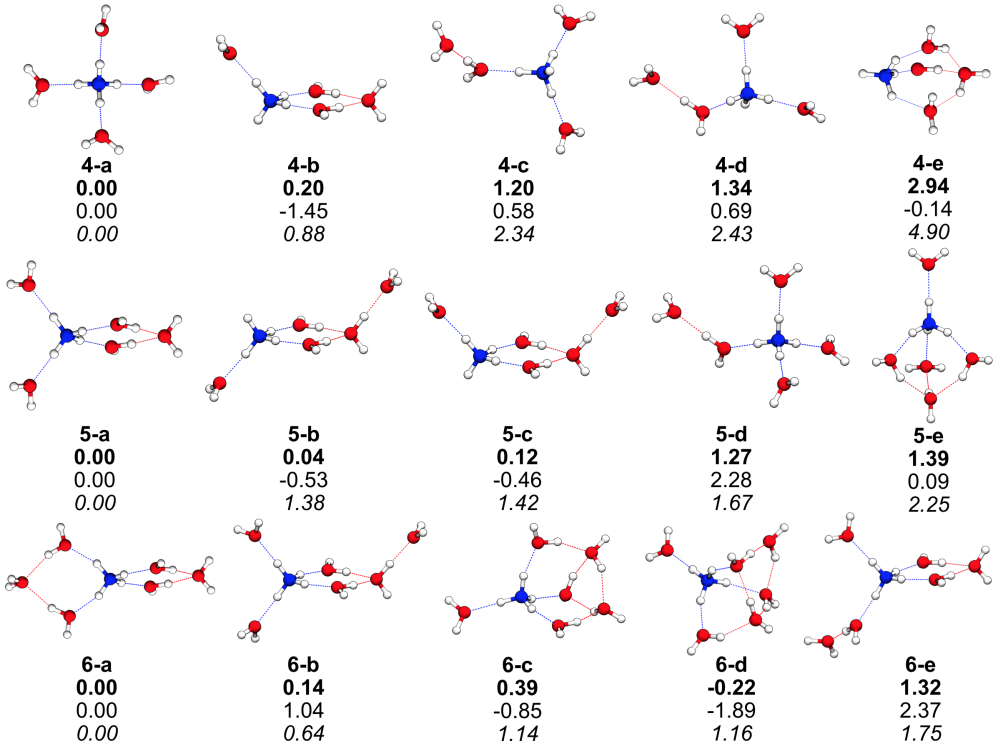


Figure 3.8: Five lowest-energy isomers of $(H_2O)_{4-6}NH_4^+$ and corresponding relative energies at MP2/Def2TZVP level with (bold) and without ZPVE (roman) correction and SCC-DFTB level (italic). Relative energies are given in kcal·mol^{-1}.

has an effect on the relative binding energy for the cluster $(H_2O)_4NH_4^+$ and the overall $\Delta E_{bind.}^{whole}$ are bigger than $\Delta E_{bind.}^{sep.}$.

For cluster $(H_2O)_5NH_4^+$, the five low-energy isomers are illustrated in Figure 3.8. The isomer 5-a is the most stable one, which is consistent with F. Spiegelman's result using the global Monte Carlo optimization and G. Shields's results obtained with a mixed molecular dynamics/quantum mechanics moldel.[371, 372] The energy order of 5-a to 5-e at SCC-DFTB level is consistent with that at MP2/Def2TZVP level with ZPVE correction. 5-a, 5-d and 5-e have a complete solvation shell while one dangling N-H bond is exposed in 5-b and 5-c. For the five low-energy isomers, the energy order of our results are not exactly the same with H. Chang's calculation results at MP2/6-31+G(d)level with ZPVE correction.[369] In H. Chang's results, 5-d is the low-energy isomer and 5-a is the second low-energy isomer. They didn't find isomers 5-b and 5-c. From the comparison, it implies the combination of SCC-DFTB and PTMD is good enough to find the low-energy isomer and the basis set can affect the energy

order when using the MP2 approach.

When all the water molecules are considered as a whole part, the obtained binding energy has a deviation due to the interaction of water molecules. As listed in Table 3.2, for isomers 5-a to 5-e, the relative binding energy $\Delta E_{bind.}^{whole}$ are -1.62, 0.72, 0.69, -1.08 and -2.08 kcal·mol^{-1} and $\Delta E_{bind.}^{sep.}$ are -0.56, 0.48, 0.55, -0.78 and 0.88 kcal·mol^{-1}, respectively. The $\Delta E_{bind.}^{whole}$ is bigger than corresponding $\Delta E_{bind.}^{sep.}$, which indicates it is better to calculate the binding energy with considering the water molecules separately. The $\Delta E_{bind.}^{sep.}$ is less than 1.00 kcal·mol^{-1} for the five low-energy isomers of cluster $(H_2O)_5NH_4^+$, so the SCC-DFTB method is good enough compared to MP2/Def2TZVP with BSSE correction for cluster $(H_2O)_5NH_4^+$.

For cluster $(H_2O)_6NH_4^+$, no N-H bond is exposed in the five low-energy isomers displayed in Figure 3.8. 6-a is the first low-energy isomer at SCC-DFTB level, which is a symmetric double-ring species connected together by eight hydrogen bonds making it a robust structure. 6-a is also the first low-energy isomer obtained using the Monte Carlo optimizations with the intermolecular polarizable potential.[371] 6-d is the first low-energy isomer at MP2/Def2TZVP level with ZPVE correction but it is only 0.22 kcal·mol^{-1} lower than 6-a. In Shields's results, 6-d is also the first low-energy isomer at MP2/aug-cc-pVDZ level.[372] In H. Chang's study, 6-b with a three-coordinated H2O molecule is the first low-energy isomer for cluster $(H_2O)_6NH_4^+$ at B3LYP/6-31+G(d) level.[367] 6-b is also the first low-energy isomer at B3LYP/6-31++G(d,p) level including the harmonic ZPE contribution.[371] The energy of 6-b is only 0.14 kcal·mol^{-1} higher than that of 6-a at MP2/Def2TZVP level with ZPVE correction. The energies of 6-a, 6-b and 6-d are very close at both SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels, which implies it is easy to have a transformation among 6-a, 6-b and 6-d. It shows SCC-DFTB is good to find the low-energy isomers of cluster $(H_2O)_6NH_4^+$ compared to MP2 and B3LYP methods.

As shown in Table 3.2, for isomers 6-a to 6-e, the relative binding energy $\Delta E_{bind.}^{whole}$ are -1.71, -1.14, -2.06, -2.90 and -1.18 kcal·mol^{-1} and the $\Delta E_{bind.}^{sep.}$ are -0.38, -0.76, 0.27, -1.06 and -0.60 kcal·mol^{-1}, respectively. It indicates the binding energy are very close at SCC-DFTB and MP2/Def2TZVP with BSSE correction levels when water molecules are calculated separately. The $\Delta E_{bind.}^{whole}$ is bigger than corresponding $\Delta E_{bind.}^{sep.}$ because of the interaction of water molecules when all the water molecules are considered as a whole part.

For cluster $(H_2O)_7NH_4^+$, the five low-energy isomers are shown in Figure 3.9. The ion core NH_4^+ has a complete solvation shell in isomers 7-a to 7-e. 7-a and 7-b with three three-coordinated H_2O molecules are the first low-energy isomers at SCC-DFTB level. In

F. Spiegelman's study, 7-a is also the first low-energy isomer using the Monte Carlo optimizations with the intermolecular polarizable potential.[371] 7-c is the first low-energy isomer at MP2/Def2TZVP with ZPVE correction level including three three-coordinated water molecules. 7-c is also the first low-energy isomer at B3LYP/6-31++G(d,p) level including the harmonic ZPE contribution.[371] 7-e is the first low-energy isomer with three three-coordinated H_2O molecules at MP2/aug-cc-pVDZ level in G. Shields's study.[372] As illustrated in Figure 3.9, the energy difference between 7-a, 7-c and 7-e at SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels are less than 0.61 kcal·mol^{-1} so it is possible that the first low-energy iosmer is different when different method are applied. The energy of 7-a and 7-b are the same at both SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels and their structures are similar, which indicates it is easy for them to transform to each other. The results for cluster $(H_2O)_7NH_4^+$ verify the accuracy of SCC-DFTB approach.

As shown in Table 3.2, for isomers 7-a to 7-e, the relative binding energy $\Delta E_{bind.}^{whole}$ are -2.95, -2.92, -2.17, -1.28 and -3.22 kcal·mol^{-1} and the $\Delta E_{bind.}^{sep.}$ are only -0.39, -0.38, 0.09, -1.35 and -2.27 kcal·mol^{-1}, respectively. It indicates the binding energies of 7-a to 7-e at SCC-DFTB agree well especially for 7-a to 7-d with those at MP2/Def2TZVP with BSSE correction level when water molecules are calculated separately. When all the water molecules are regarded as a whole part, the results of SCC-DFTB are not as good as those of the MP2 with BSSE method.

For cluster $(H_2O)_8NH_4^+$, 8-a to 8-e are the five low-energy isomers displayed in Figure 3.9. In 8-a to 8-d, the ion core NH_4^+ has a complete solvation shell. 8-a is the first low-energy isomer in our calculation at SCC-DFTB level. In F. Spiegelman's study, 8-b is the first low-energy isomer at B3LYP/6-31++G(d,p) level including the harmonic ZPE contribution.[371] The structures of 8-a and 8-b are very similar and the energy differences are only 0.09 and 0.18 kcal·mol^{-1} at SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels, respectively. 8-d with seven three-coordinated H_2O molecules in the cube frame is the first low-energy isomer in our calculation at MP2/Def2TZVP with ZPVE correction level, which is consistent with F. Spiegelman's results obtained using Monte Carlo optimizations.[371] In 8-e, NH_4^+ has an exposed N-H bond and it also has seven three-coordinated H_2O molecules in its cage frame. The energies of isomers 8-a to 8-e are very close calculated using SCC-DFTB and MP2 methods, so it's possible that the energy order will change when different methods or basis sets are applied. The results certificate the SCC-DFTB is good enough to find the low-energy isomers for cluster $(H_2O)_8NH_4^+$.

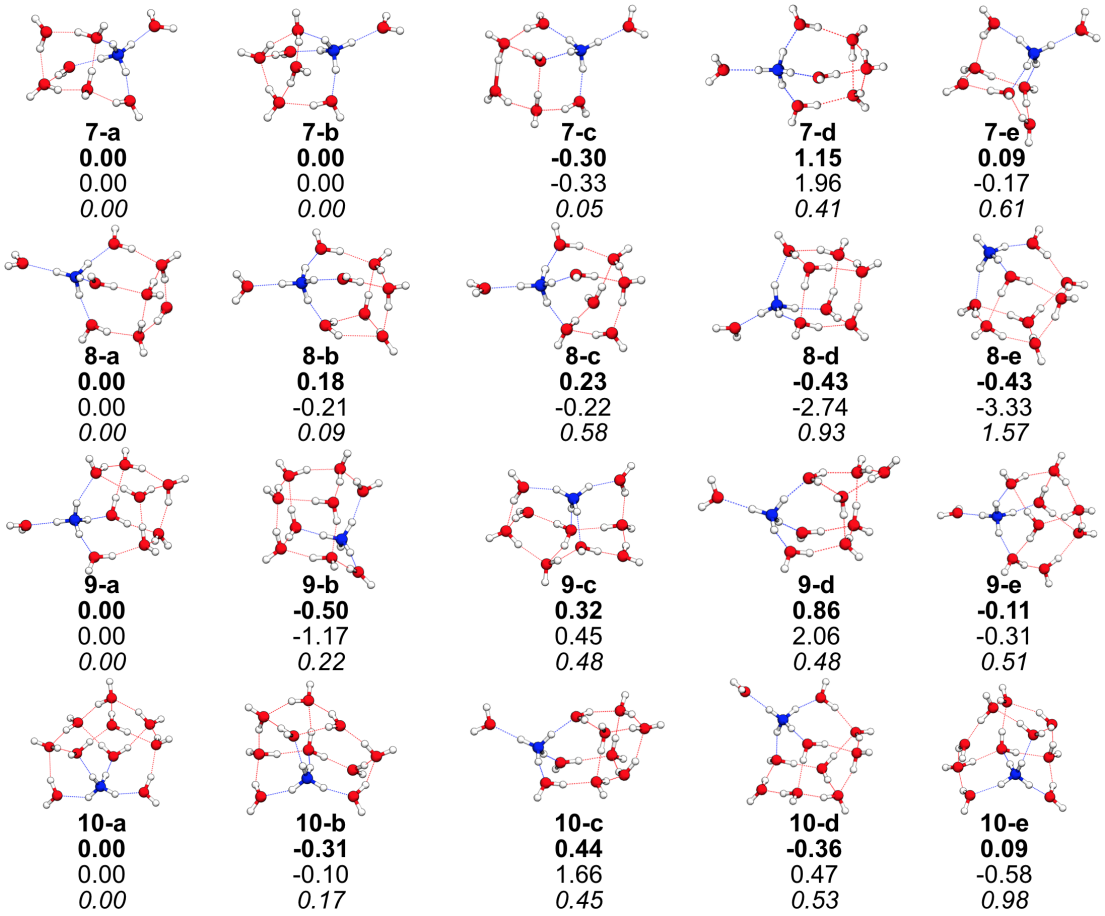


Figure 3.9: The five low-energy isomers of clusters $(H_2O)_{7-10}NH_4^+$ and the associated relative energies (in kcal·mol^{-1}) at MP2/Def2TZVP level with (bold) and without ZPVE correction and SCC-DFTB level (italic).

As shown in Table 3.2, for isomers 8-a to 8-e, the relative binding energy $\Delta E_{bind.}^{whole}$ are -2.20, -1.61, -3.71, -2.43 and -0.55 kcal·mol^{-1}, respectively and the biggest $\Delta E_{bind.}^{sep.}$ is -2.01 kcal·mol^{-1}. It shows the binding energies at SCC-DFTB level agree well with those at MP2/Def2TZVP with BSSE correction level when water molecules are calculated separately. From these results, when all the water molecules are considered as a whole part, the results of SCC-DFTB didn't agree well with those of the MP2 with BSSE correction method.

For cluster $(H_2O)_9NH_4^+$, the five low-energy structures of $(H_2O)_9NH_4^+$ are illustrated in Figure 3.9. 9-a with seven three-coordinated H_2O molecules in the cage frame is the first low-energy isomer at SCC-DFTB level. 9-a is also the first low-energy structure at B3LYP/6-31++G(d,p) level including the harmonic ZPE contribution in F. Spiegelman's study.[371] 9-b

with one N-H bond exposed in NH_4^+ is the second low-energy isomer whose energy is only 0.22 kcal·mol^{-1} higher than that of 9-a in the results of SCC-DFTB calculation. 9-b is the first low-energy isomer at MP2/Def2TZVP with ZPVE correction level in our calculation and it is also the first low-energy isomer at B3LYP/6-31++G(d,p) level in F. Spiegelman's study.[371] 9-c, 9-d and 9-e have a complete solvation shell. All the water molecules are connected together in the structure of 9-c. The structures of 9-a and 9-e are very similar and their energy difference is only 0.11 kcal·mol^{-1} at MP2/Def2TZVP with ZPVE correction level. The energy difference of isomers 9-a to 9-e is less than 0.51 kcal·mol^{-1} at SCC-DFTB and less than 0.86 kcal·mol^{-1} at MP2/Def2TZVP with ZPVE correction, so it's easy for them to transform to each other making it possible for the variation of the energy order. The results certificate the SCC-DFTB is good enough to find the low-energy isomers for cluster $(H_2O)_9NH_4^+$.

As shown in Table 3.2, for isomers 9-a to 9-e, the relative binding energy $\Delta E_{bind.}^{whole}$ are -2.20, -1.61, -3.71, -2.43 and -0.55 kcal·mol^{-1} and the relative binding energy $\Delta E_{bind.}^{sep.}$ are -1.39, -0.84, -0.85, -1.78, and -0.91 kcal·mol^{-1}, respectively. It is obvious that the absolute values of $\Delta E_{bind.}^{whole}$ are bigger than the corresponding $\Delta E_{bind.}^{sep.}$. It shows the binding energies at SCC-DFTB level agree well with those at MP2/Def2TZVP with BSSE correction level when water molecules are calculated separately. According to the results, When all the water molecules are considered as a whole part, the results of SCC-DFTB didn't agree well with those of the MP2 with BSSE correction method.

For cluster $(H_2O)_{10}NH_4^+$, 10-a to 10-e are the five low-energy isomers in which the ion core NH_4^+ has a complete solvation shell shown in Figure 3.9. 10-a with eight three-coordinated H2O molecules in its big cage structure is the first low-energy isomer calculated using the SCC-DFTB approach. 10-a is also the first low-energy structure at B3LYP/6-31++G(d,p) level including the harmonic ZPE contribution in F. Spiegelman's study.[371] In 10-b and 10-e, there is a four-coordinated H_2O molecule in their cage structures. 10-d is the first low-energy structure in our calculation results using MP2/Def2TZVP with ZPVE correction, which is also the first low-energy isomer at B3LYP/6-31++G(d,p) level in F. Spiegelman's study.[371] The energy of 10-b is only 0.17 kcal·mol^{-1} higher than that of 10-a at SCC-DFTB level, and it is only 0.31 kcal·mol^{-1} lower than that of 10-a at MP2/Def2TZVP with ZPVE correction level. The energy of isomers 10-a to 10-e are very close at both SCC-DFTB and MP2/Def2TZVP levels, which indicates the results with SCC-DFTB agree well with those using MP2/Def2TZVP method for cluster $(H_2O)_{10}NH_4^+$.

As shown in Table 3.2, for isomers 10-a to 10-e, the relative binding energies $\Delta E_{bind.}^{whole}$ and $\Delta E_{bind.}^{sep.}$ are not as small as the corresponding ones of clusters $(H_2O)_{1-9}NH_4{}^+$, which implies the error of the relative binding energy increases with the number of water molecules in the cluster. The whole results of $\Delta E_{bind.}^{whole}$ are still bigger than those of $\Delta E_{bind.}^{sep.}$ for isomers 10-a to 10-e.

3.2.2.4 Properties of $(H_2O)_{4-10}NH_3$ Clusters

For cluster $(H_2O)_4NH_3$, the five low-energy structures 4′-a to 4′-e are displayed in Figure 3.10. 4′-a with three N-H bonds exposed is the first low-energy isomer at SCC-DFTB level. 4′-b with two N-H bonds exposed is the second low-energy isomer at SCC-DFTB level but it is the first low-energy isomer at MP2/Def2TZVP with ZPVE correction level. The energy differences between 4′-a to 4′-b are only 0.20 and 0.07 kcal·mol^{-1} at SCC-DFTB and MP2/Def2TZVP with ZPVE correction level, respectively. The energy difference of isomers 4′-a to 4′-e is less than 0.75 kcal·mol^{-1} at MP2/Def2TZVP with ZPVE correction, so it's possible for the variation of the energy order when different methods or basis sets are used. 4′-d with a nearly planar pentagonal structure with nitrogen atom and the four oxygen atoms at the apexes is the first low-energy isomer at MP2/6-31+G(d,p) studied by J. Novoa et al[375] 4′-d is also the first low-energy isomer in D. Bacelo's study using QCISD(T) for a single-point energy calculation based on the MP2/6-311++G(d,p) results.[373] In addition, 4′-a to 4′-e are also the five low-energy isomers in D. Bacelo's study even the energy order is different.[373] The results show the SCC- DFTB is good enough to find the low-energy isomers isomers for cluster $(H_2O)_4NH_3$.

The relative binding energies of isomers 4′-a to 4′-e are shown in Table 3.2. Except 4′-d, the values of $\Delta E_{bind.}^{whole}$ for 4′-a to 4′-e are smaller than the corresponding values of $\Delta E_{bind.}^{sep.}$. The $\Delta E_{bind.}^{sep.}$ of 4′-d is smaller than those of other isomers. 4′-d has a nearly planar pentagonal structure that only contains three O-H···O hydrogen bonds among the four water molecules while other isomers contain four O-H···O hydrogen bonds among the four water molecules. So the intermolecular interaction of the four water molecules in 4′-d is not as strong as it in other isomers, this may explain the $\Delta E_{bind.}^{sep.}$ of 4′-d is smaller than those of other isomers. In general, both relative binding energies $\Delta E_{bind.}^{sep.}$ and $\Delta E_{bind.}^{sep.}$ are not big that indicates SCC-DFTB performs well compared to the MP2 method with BSSE correction for calculating the binding energy of cluster $(H_2O)_4NH_3$.

For cluster $(H_2O)_5NH_3$, 5′-a to 5′-e are the five low-energy isomers shown in Figure 3.10. 5′-a with four three-coordinated water molecules is the first low-energy structure at both SCC-

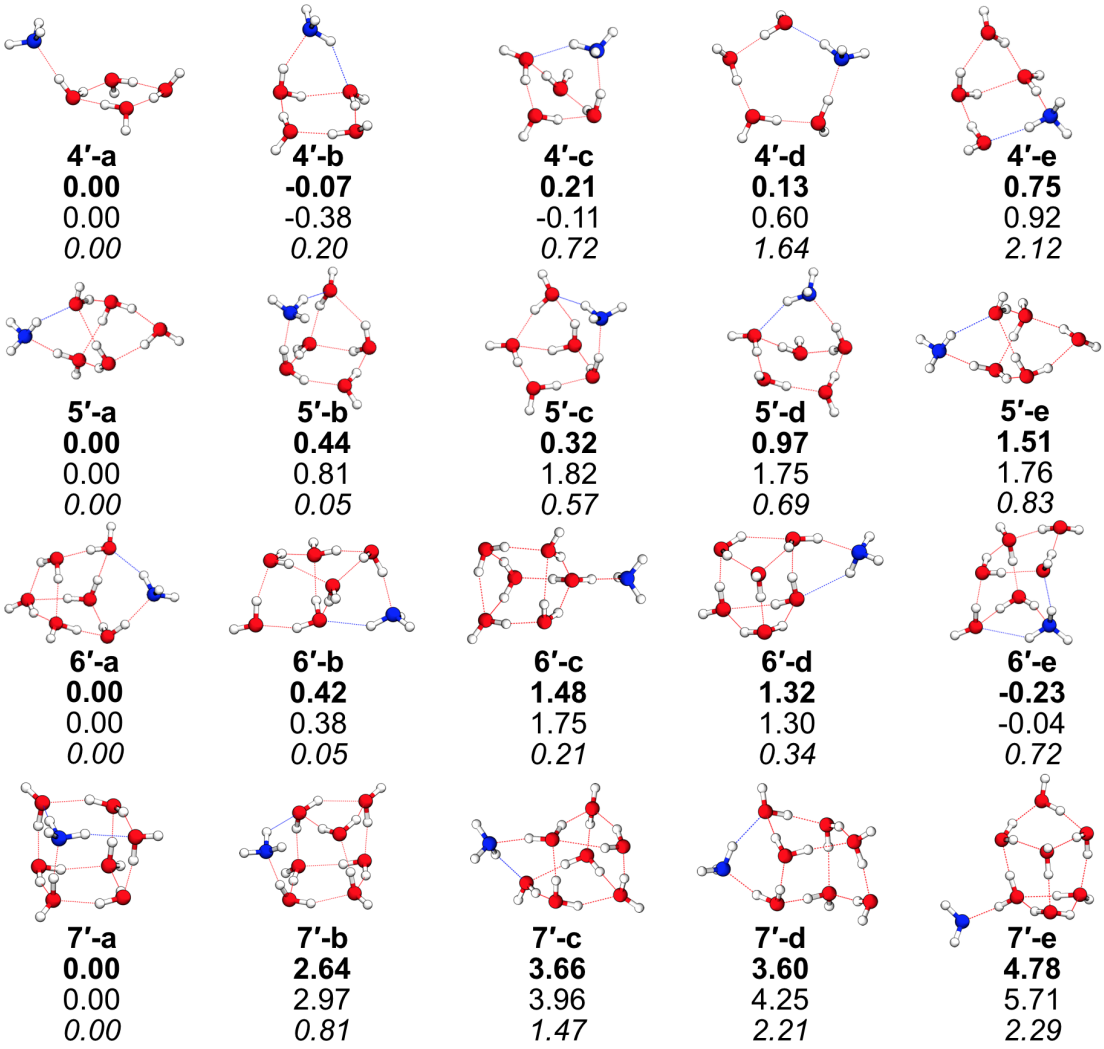


Figure 3.10: The five low-energy isomers of cluster $(H_2O)_{4-7}NH_3$ and the associated relative energies (in kcal·mol^{-1}) at MP2/Def2TZVP level with (bold) and without ZPVE correction and SCC-DFTB level (italic).

DFTB and MP2/Def2TZVP with ZPVE correction levels. 5′-b and 5′-c are the second and third isomers at SCC-DFTB level and they are the third and second isomers at MP2/Def2TZVP level with ZPVE. The energy difference between 5′-b and 5′-c is only 0.05 and 0.44 kcal·mol^{-1} at SCC-DFTB level and MP2/Def2TZVP with ZPVE correction level, respectively. In addition, the structures of 5′-b and 5′-c are very similar so it is possible for them to transform to each other. 5′-d with two three-coordinated water molecules is the fourth low-energy structure at both SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels. 5′-e with four

three-coordinated water molecules is the fifth low-energy structure at both SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels. The frames of 5′-a and 5′-e are almost the same but the water molecule who offers the hydrogen or oxygen to form the O-H···O hydrogen bonds has a small difference. The energy of 5′-e is 1.51 kcal·mol^{-1} higher than that of 5′-a at MP2/Def2TZVP with ZPVE correction level, which implies the intermolecular connection mode has an influence on the stability of the isomers. The results show the SCC-DFTB approach performs well to find the low-energy isomers for cluster $(H_2O)_5NH_3$ compared with MP2/Def2TZVP with ZPVE correction method.

The relative binding energies of isomers 5′-a to 5′-e are shown in Table 3.2. The values of $\Delta E_{bind.}^{whole}$ are less than 0.82 kcal·mol^{-1} for 5′-a to 5′-e. The values of $\Delta E_{bind.}^{sep.}$ are bigger than the corresponding values of $\Delta E_{bind.}^{whole}$. It indicates SCC-DFTB agrees better with MP2/Def2TZVP $\Delta E_{bind.}^{whole}$ when all the water molecules are regarded as a whole part than considering separately for calculating the binding energy of cluster $(H_2O)_5NH_3$.

For cluster $(H_2O)_6NH_3$, the five low-energy structures 6′-a to 6′-e are displayed in Figure 3.10. 6′-a is the first low-energy structure at SCC-DFTB level. All water molecules in 6′-a are three-coordinated. 6′-b is the second low-energy isomer at SCC-DFTB level and it's only 0.05 and 0.42 kcal·mol^{-1} higher than the ones of 6′-a at SCC-DFTB level and MP2/Def2TZVP with ZPVE correction level, respectively. 6′-c to 6′-d are the third and fourth low-energy isomers in which the six water molecules form a triangular prism structure and there are one and two four-coordinated water molecules in 6′-c to 6′-d, respectively. 6′-e is the fifth low-energy structure at SCC-DFTB level but it's the first low-energy isomer at MP2/Def2TZVP with ZPVE correction level. The energy of 6′-a to 6′-e are very close at both SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels that it is difficult to keep the energy order when different methods or basis sets are applied. This also shown the SCC-DFTB method used is efficient to find the low-energy isomers of cluster $(H_2O)_6NH_3$.

The relative binding energies of isomers 6′-a to 6′-e are listed in Table 3.2. The smallest and the biggest values of $\Delta E_{bind.}^{whole}$ are -0.05 and -1.11 kcal·mol^{-1}, respectively. The smallest absolute value of $\Delta E_{bind.}^{sep.}$ is 1.96 kcal·mol^{-1}. The binding energies calculated with SCC-DFTB agree well with those calculated at MP2/Def2TZVP level for cluster $(H_2O)_6NH_3$ when all the water molecules are considered as a whole part.

For cluster $(H_2O)_7NH_3$, the five low-energy isomers 7′-a to 7′-e are illustrated in Figure 3.10. 7′-a with a cubic structure is the first low-lying energy structure at both SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels. 7′-b is the second low-energy structure at both

SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels. 7′-b has a similar structure with 7′-a but the NH_3 in it has two exposed N-H bonds. 7′-c and 7′-d have similar structures and they are the third and fourth low-lying energy isomers at SCC-DFTB level and their energy difference is only 0.74 kcal·mol^{-1}. 7′-e with three exposed N-H bonds is the fifth low-energy isomer at both SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels. The results of SCC-DFTB method agree well with those of MP2/Def2TZVP with ZPVE correction for the five low-energy isomers of cluster $(H_2O)_7NH_3$.

The smallest and the biggest values of $\Delta E_{bind.}^{whole}$ of isomers 7′-a to 7′-e are -0.02 and -1.11 kcal·mol^{-1}, respectively and the smallest absolute value of $\Delta E_{bind.}^{sep.}$ is 2.02 kcal·mol^{-1} shown in Table 3.2. The binding energies calculated with SCC-DFTB agree well with those obtained using MP2/Def2TZVP for cluster $(H_2O)_7NH_3$ when all the water molecules are considered as a whole part.

For cluster $(H_2O)_8NH_3$, 8′-a to 8′-e are the five low-energy structures shown in Figure 3.11. 8′-a in which eight water molecules constitute a cube is the first low-lying energy structure in SCC-DFTB calculation results. 8′-b also with a water-cube structure is the second low-energy structure at SCC-DFTB level and it is the first low-energy isomer at MP2/Def2TZVP with ZPVE correction level. The energy differences between 8′-a and 8′-b are only 0.93 an 0.30 kcal·mol^{-1} at SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels. From Figure 3.11, the fifth low-energy isomer 8′-e includes less number of hydrogen bonds than other isomers and its energy has a clearly increase compared to other isomers. The results show the SCC-DFTB method performs well to obtain the low-energy isomers of cluster $(H_2O)_8NH_3$.

The smallest and the biggest values of $\Delta E_{bind.}^{whole}$ of isomers 8′-a to 8′-e are -0.1 and -1.28 kcal·mol^{-1}, respectively while the smallest absolute value of $\Delta E_{bind.}^{sep.}$ is 3.04 kcal·mol^{-1} shown in Table 3.2. The binding energies calculated with SCC-DFTB agree better with those obtained at MP2/Def2TZVP level when all the water molecules are considered as a whole part in cluster $(H_2O)_8NH_3$ than the ones when water molecules calculated separately.

For cluster $(H_2O)_9NH_3$, 9′-a to 9′-e are the five low-lying energy structures displayed in Figure 3.11. 9′-a with a "chair" structure is the first low-energy structure at SCC-DFTB level. 9′-b, 9′-c and 9′-d in which the nine water molecules have the similar configuration are the second, third and fourth isomers. In 9′-b and 9′-c, the NH_3 has three exposed N-H bonds and the energies of 9′-b and 9-c are very close at both SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels. The NH_3 has two exposed N-H bonds in 9′-d. 9′-e is the fifth low-energy

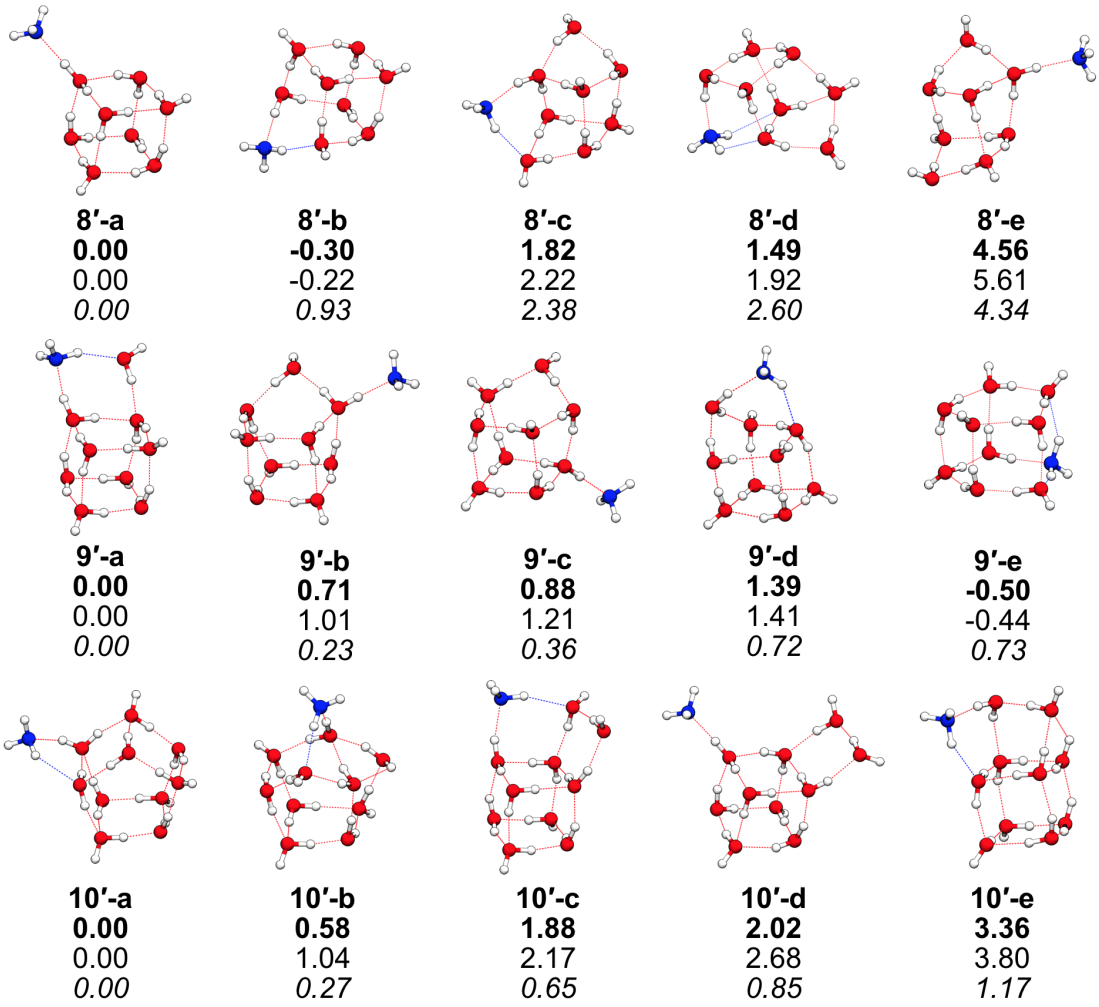


Figure 3.11: The five low-energy isomers of clusters $(H_2O)_{8-10}NH_3$ and the associated relative energies (in kcal·mol^{-1}) at MP2/Def2TZVP level with (bold) and without ZPVE correction and SCC-DFTB level (italic).

isomer in the SCC-DFTB calculation results but it is the first low-energy isomer in the calculation results of MP2/Def2TZVP with ZPVE correction. 9′-e has a pentagonal prism structure and all the water molecules in it are three-coordinated. The relative energy for each isomer between SCC-DFTB level and MP2/Def2TZVP with ZPVE correction level is less than 1.23 kcal·mol^{-1}. This shows our SCC-DFTB calculation results are consistent with the calculation results of MP2/Def2TZVP with ZPVE correction for low-energy isomers optimization of cluster $(H_2O)_9NH_3$.

The relative binding energies of isomers 9′-a to 9′-e are shown in Table 3.2. The absolute

values of $\Delta E_{bind.}^{whole}$ are less than 1.09 kcal·mol^{-1} while the smallest absolute value of $\Delta E_{bind.}^{sep.}$ is 2.57 kcal·mol^{-1}. The binding energies calculated with SCC-DFTB agree well with those acquired at MP2/Def2TZVP level when all the water molecules are considered as a whole part for cluster $(H_2O)_9NH_3$.

For cluster $(H_2O)_{10}NH_3$, 10′-a to 10′-e are the five low-energy structures illustrated in Figure 3.11. The energy order for the five low-energy structures is the same at SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels. 10′-a and 10′-b are the first and second low-energy isomer in which the ten water molecules constitute the pentagonal prism. The energy differences of 10′-a and 10′-b are only 0.27 and 0.58 kcal·mol^{-1} at SCC-DFTB and MP2/Def2TZVP with ZPVE correction levels. 10′-c and 10′-d are the third and fourth low-energy isomers in which eight water molecules constitute a cube and the energy differences between 10′-c and 10′-d are very small calculated with SCC-DFTB or MP2/Def2TZVP with ZPVE correction. 10′-e is the fifth low-energy structure in which eight water molecules also constitute a cube but its energy is obviously higher than those of 10′-c and 10′-d. The calculation results of SCC-DFTB are consistent with those of MP2/Def2TZ for the optimization of the low-energy isomers of cluster $(H_2O)_{10}NH_3$. According to the structures of the five low-energy isomers of clusters $(H_2O)_{1-10}NH_3$, in most cases, the NH_3 usually contains two or three exposed N-H bonds.

The smallest and biggest values of $\Delta E_{bind.}^{whole}$ of isomers 10′-a to 10′-e are -0.03 and -1.10 kcal·mol^{-1} while the smallest absolute value of $\Delta E_{bind.}^{sep.}$ is 4.80 kcal·mol^{-1} shown in Table 3.2. The values of $\Delta E_{bind.}^{whole}$ implies that SCC-DFTB agree very well with MP2/Def2TZVP for cluster $(H_2O)_{10}NH_3$ when all the water molecules are regarded as a whole part.

3.2.2.5 Properties of $(H_2O)_{20}NH_4^+$ Cluster

For cluster $(H_2O)_{20}NH_4^+$, the lowest-energy structure shown in Figure 3.12 (a) was obtained with the combination of SCC-DFTB and PTMD which is consistent with that of previous study.[340, 389, 390] Microcanonical and canonical caloric curves were obtained using exchange Monte Carlo simulations by F. Spiegelman's group.[340] I also calculated the canonical heat capacities of cluster $(H_2O)_{20}NH_4^+$ using the combination of SCC-DFTB and PTMD depicted in Figure.

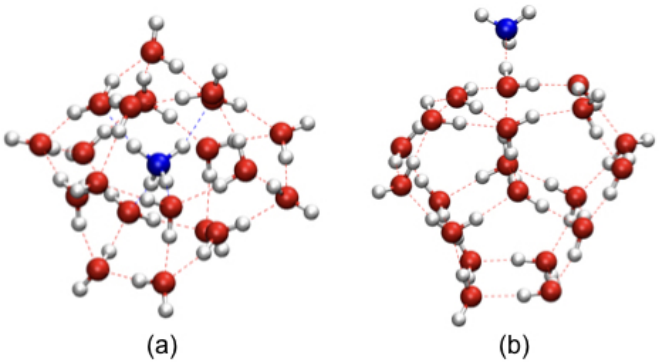


Figure 3.12: The five low-energy isomers of cluster $(H_2O)_{20}NH_4^+$ (a) and $(H_2O)_{20}NH_3$ (b) at SCC-DFTB level.

3.2.3 Conclusions for Ammonium/Ammonia Including Water Clusters

The low-energy isomers reported for $(H_2O)_{1-10,20}NH_4^+$ and $(H_2O)_{1-10,20}NH_3$ clusters are obtained with a combination of SCC-DFTB (0.12/1.16) and PTMD. Binding energies as a function of the N—O distance in $(H_2O)NH_4^+$ and $(H_2O)NH_3$ demonstrate that the improve parameters I proposed are in much better agreement with reference calculations than the original SCC-DFTB parameters. The low-energy isomers of clusters $(H_2O)_{1-10}NH_4^+$ and $(H_2O)_{1-10}NH_3$ at the SCC-DFTB (0.12/1.16) level agree well with those at MP2/Def2TZVP level and the corresponding results in the literature. The SCC-DFTB binding energies also agree well with those calculated with MP2/Def2TZVP method with BSSE correction. This demonstrate that SCC-DFTB (0.12/1.16) approach is good enough to model ammonium and ammonia containing water clusters.

Among the five lowest-energy structures of $(H_2O)_4NH_4^+$, four of them display a dangling N-H bond. Among the five lowest-energy structures of $(H_2O)_5NH_4^+$, only two structures display a dangling N-H bond. Among the five lowest-energy isomers of $(H_2O)_{6-10}NH_4^+$, all the structures, except 8-e, display a ion core NH_4^+ that has a complete solvation shell but it is not located at the center of the water cluster. In the most stable structures of $(H_2O)_{20}NH_4^+$, the ion core NH_4^+ has a complete solvation shell and it is in the center of the water cluster. In contrast, in the low-energy isomers of $(H_2O)_{1-10}NH_3$ clusters, NH_3 is never fully solvated by the water cluster.

The present study demonstrate thet ability of SCC-DFTB to model small size ammonium and ammonia containing water clusters, which is less expensive than *ab initio* methods. It

is possible for SCC-DFTB to describe the larger scaled ammonium and ammonia containing water clusters.

3.3 Structural and Energetic Properties of Protonated Uracil Water Clusters

3.3.1 General introduction

Gas phase investigations of molecules help to understand the intrinsic properties of molecules that are free from the effects of solvents. The gas phase study needs to be extended towards more realistic biomolecular systems, to reveal how the intrinsic molecular properties are affected by the surrounding medium when the biomolecules are in a natural environment.[147, 391, 392, 393] The hydration study of biomolecules is of paramount importance to get insights into their behavior in aqueous medium, especially the effects on their structure, stability and dynamics.

The nucleobases in DNA and RNA play a significant role in the encoding and expression of genetic information in living systems while water is a natural medium of many reactions in living organisms. The study of the interaction between nucleobase molecules and aqueous environment has attracted a lot of interests among biologists and chemists. Exploring the clusters composed of nucleobase molecules with water is a good workbench to observe how the properties of nucleobase molecules change when going from isolated gas-phase to hydrated species.

The radiation can cause damages on RNA and DNA molecules, which is proficiently applied in radiotherapy for cancer treatment. The major drawback in radiotherapy is the unselective damage in both healthy and tumor cells, which has a big side effect. This makes it particularly important to explore the radiation fragments. Uracil , $C_4H_4N_2O_2$, is one of the four nucleobases of RNA, has been paid attention concerning radiation damage. Protonated uracil UH^+ can be generated by radiation damages.[394] The reasons for such degradation can be due to the interaction with slow electrons, as shown by the work of B. Boudaiffa *et al.* [395] Several studies have been devoted to the effect of hydration on the electron affinity of DNA nucleobases. [396, 397, 398] For instance, A. Rasmussen *et al.* found that a water molecule is more likely to interact with a charged species than with a neutral one though the study of hydration effects on the lowest triplet states of cytosine, uracil, and thymine by including one

or two water molecules explicitly, [399] However, a lot of work is still needed to be performed to understand the role of aqueous environment on charged nucleobases of DNA and RNA.

Collision experiments is a useful tool that can be applied to understand the reactivity of molecules and provide access to structural information.[141] Fragmentation of the bare protonated U has already been performed under collision-induced dissociation (CID) with tandem mass spectrometry,[142, 400, 401] however, there are only few studies available concerning the effect of hydration on such process. Infrared photodissociation spectroscopy of singly hydrated protonated uracil shows that the most stable tautomeric form of the neutral uracil (diketo) differs from the most stable one for bare protonated uracil (keto-enol).[402] However, fragmentation studies of such species under CID conditions have not been performed. **S. Zamith and J.-M. L'Hermite conducted such CID experiments on protonated uracil water species $(H_2O)_{1-15}UH^+$ during my thesis and I collaborated with them in order to provide a theoretical support to their measurements**.

Theoretical studies have already been devoted to mixed uracil-water clusters and intended to describe the lowest energy structures. However, only neutral species ($(H_2O)_nU$) were considered.[403, 404, 405, 406, 407, 408] Those studies showed that for sizes up to with $n = 3$, the water molecules arrange in monomers or dimers in the plane of the uracil molecule [404, 405, 406, 407, 408] with no trimer formation. But for $n > 3$, very different structures were predicted depending on the considered study. For instance, M. Ghomi predicted that for $n = 7$,[406] water molecules arrange in dimers and trimers in the plane of the uracil molecule, whereas for $n = 11$, water molecules form locked chains.[403] 3D configurations were also proposed. For instance, all water molecules lie above the uracil plane for $n = 4, 5$ reported by F. Calvo *et al.*.[408] Similarly, for $n = 11$, V. Danilov *et al.* also obtained a structure that consists of a water cluster above the uracil molecule.[407] Such structures are predicted to start with 4 water molecules reported by F. Calvo and collaborator [408] or with 6 water molecules (though 5 have not been calculated) reported by S. Gadre *et al.*.[404] Those studies may suggest that for few water molecules (up to two), the proton should be located on the uracil molecule, whereas when a large number of water molecules surround the uracil, the charge is expected to be located on the water molecules. Of course, the excess proton is expected to strongly influence the structure of the lowest energy isomers of each species, as observed for pure water clusters, so the size at which the proton is transferred from uracil to water cannot be deduced from the aforementioned studies. Moreover, all those theoretical studies do not lead to the same low-energy structures as highlighted by V. Danilov and F. Calvo.[407, 408] Consequently, although

it is instructive from a qualitative point of view, the analysis of the experimental data by S. Zamith and J.-M. L'Hermite cannot be based on those studies. I have therefore undertaken a theoretical simulation of hydrated protonated uracil clusters $(H_2O)_{1-7,11,12}UH^+$ to determine their lowest-energy structures to complete the experiments by S. Zamith and J.-M. L'Hermite at the *Laboratoire Collisions Agrégats Réactivité* (LCAR). This work has been published in 2019 in the *The Journal of Chemical Physics*.[104]

3.3.2 Results and Discussion

In the following section, section 3.3.2.1, I present in details the results obtained from the CID experiments of S. Zamith and J.-M. L'Hermite and the main concepts used to interpret the data. The following section, section 3.3.2.2, is devoted to the theoretical determination of the low-energy isomers of the $(H_2O)_{1-7,11,12}UH^+$ clusters. A more detailed presentation of CID experiments is also provided in section 4.1 of chapter 4, where these details are important to explicitly model CID experiments.

3.3.2.1 Experimental Results

Time-of-flight of mass spectrum. A typical fragmentation mass spectrum obtained by colliding $(H_2O)_7UH^+$ with neon at a center of mass collision energy of 7.2 eV is shown in Figure 3.13. The more intense peak on the right comes from the parent cluster $(H_2O)_7UH^+$, the next 7 peaks at the left of the parent peak correspond to the loss of 1-7 water molecules of parent cluster, and the next 5 peaks to the left results from the evaporation of the uracil molecule and several water molecules from parent cluster. This mass spectrum is obtained at the highest pressure explored in the present experiments. This is still true for the largest size investigated here, namely, $(H_2O)_{15}UH^+$. From the result of the fragmentation mass spectrum displayed in Figure 3.13, it indicates multiple collisions are possible, which allows the evaporation of all water molecules. Moreover, the intensity of evaporation of water molecules is bigger than the one of evaporation of U. In the study, mainly focus on two specific channels. Channel 1 corresponds to the loss of only neutral water molecules, whereas channel 2 corresponds to the loss of neutral uracil and one or several water molecules,

$$\text{Channel 1, } (H_2O)_nUH^+ \rightarrow (H_2O)_{n-x}UH^+ + xH_2O \tag{3.1}$$

$$\text{Channel 2, } (H_2O)_nUH^+ \rightarrow (H_2O)_{n-x}H^+ + xH_2O + U \tag{3.2}$$

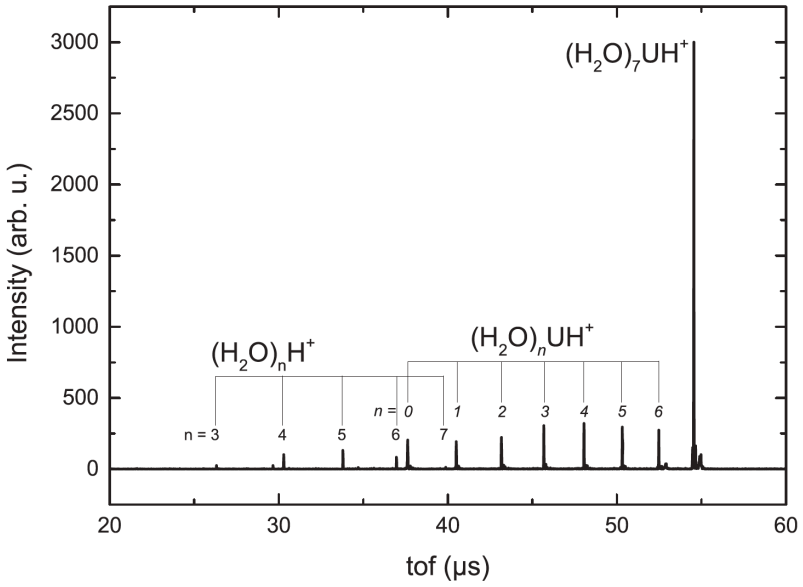


Figure 3.13: Time-of-flight of mass spectrum obtained by colliding $(H_2O)_7UH^+$ with Ne at 7.2 eV center of mass collision energy (93.5 eV in the laboratory frame).

Fragmentation cross section. The total fragmentation cross sections of clusters $(H_2O)_{n-1}UH^+$, pure water clusters
$(H_2O)_{2-6}H^+$,[409] and deuterated water clusters $(D_2O)_{5,10}H^+$ [410] are plotted in Figure 3.14 as a function of the cluster size n. Here n stands for the total number of molecules when the cluster includes uracil molecule. Different target atoms and molecules were used in these experiments: Water molecules or neon atoms in our experiments, xenon atoms in Dalleska's experiments. These experimental data are compared to the geometrical (*i.e.*, hard sphere) cross sections given by:

$$\sigma_{geo} = \pi \left(\left[n_w \times r_w^3 + n_U r_U^3 \right]^{1/3} + r_T \right)^2 \tag{3.3}$$

where n_w is the number of water molecules, and n_U is the number of uracil molecules (n_U = 0 or 1 in the present study). r_w, r_U, and r_T refer to the molecular radii of water, uracil, and the target atom or molecule, respectively. The molecular radii are deduced from macroscopic densities that gives r_U = 3.2 Å[411] and r_w = 1.98 Å. The radii of rare gas target atoms are taken as their Van der Waals radii r_{Ne} = 1.54 Åand r_{Xe} = 2.16 Å.

The main differences between the curves in Figure 3.14 can be rationalized as follows: The larger the size of the target atom (or molecule) is, the bigger the fragmentation cross section will

be. The experimental fragmentation cross sections of clusters $(H_2O)_{n-1}H^+$ colliding with water molecules are larger than the values obtained for collisions with Ne atoms. In the same vein, for a given number of molecules in the cluster, the cross section is larger for clusters containing uracil. The overall trend of all curves in Figure 3.14 is the same: The fragmentation cross sections increase with the size and seem to tend toward the geometrical one. The cross sections measured for clusters containing uracil colliding with water molecules (black squares) are of the same magnitude as the ones previously obtained for deuterated pure water clusters (green full circles) at a similar collision energy.[410] For clusters containing uracil, fragmentation cross sections are systematically larger than the one for pure water clusters by an amount of the same magnitude as the one predicted by the geometrical cross sections. For instance, the difference between red squares and blue stars, and the difference between red full line and blue dashed line has the same magnitude.

The fragmentation cross sections obtained by N. Dalleska and coworkers [409] for protonated water clusters are within our error bars for $n = 5$, 6 and about a factor of 2 lower for $n = 3$, 4. However their cross section is notably lower for $(H_2O)_2H^+$ as compared to our measurement for $(H_2O)UH^+$. This difference may be explained by the fact that UH^+ forms a weaker bond with water than H_2OH^+ does. Indeed the dissociation energy $D[H_2OH^+–H_2O]$ is 1.35 eV [409, 412] whereas the value for $D[UH^+–H_2O]$ is estimated between 0.54 [394] and 0.73 eV. [402] The same behavior is observed for $n = 3$, and the dissociation energy $D[(H_2O)_2H^+–H_2O] = 0.86$ eV [409, 412] is greater than the dissociation energy $D[U(H_2O)H^+–H_2O] = 0.49$ eV.[394] Hence the dissociation of water molecules is more favored in the protonated uracil cluster than in the pure water clusters.

Intermolecular fragmentation. Figure 3.15 displays the percentage of the fragments that have lost a neutral uracil molecule over all the fragments, plotted as a function of the number of water molecules in the parent cluster $(H_2O)_nUH^+$. It shows that for the cluster $(H_2O)_nUH^+$ with a small number of water molecules, almost no neutral uracil is evaporated. From $n = 5$ and more clearly from $n = 6$, the loss of neutral uracil molecule increases up to about 20% for $(H_2O)_9UH^+$.

The fragmentation can arise from two distinct mechanisms (direct and statistical fragmentation processes) depending on the life time of the collision complex. On the one hand, if the fragmentation occurs in a very short time after collision, the dissociation is impulsive (direct). In this case, I thus assume that the nature of the collision products is partly determined by the nature of the lowest-energy isomers of parent clusters and especially by the location of the

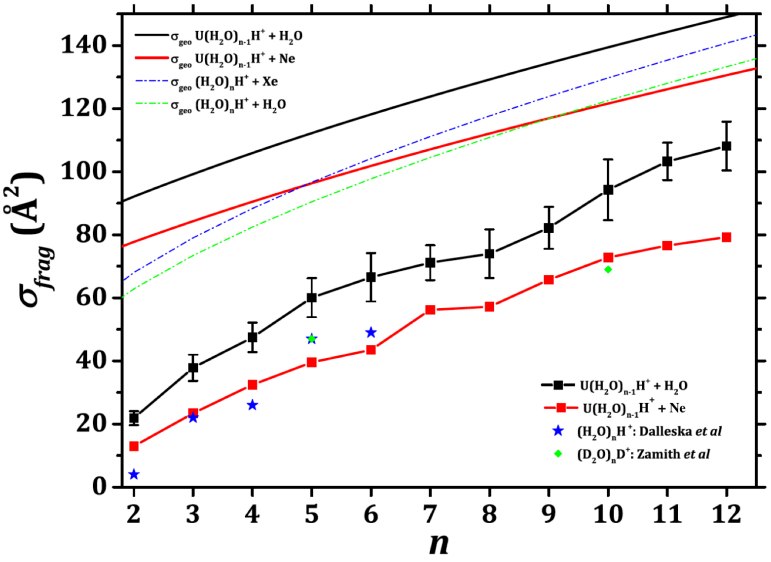


Figure 3.14: Fragmentation cross sections of clusters $(H_2O)_{n-1}UH^+$ at a collision energy of 7.2 eV plotted as a function of the total number n of molecules in the clusters. The experimental results and geometrical cross sections are shown for collision with H_2O and Ne. The results from N. Dalleska et al.[409] using Xe as target atoms on pure protonated water clusters $(H_2O)_{2-6}H^+$ and from S. Zamith *et al.* [410] using water as target molecules on deuterated water clusters $(D_2O)_{5,10}H^+$ are also shown. The geometrical collision cross sections of water clusters in collision with Xe atoms and water molecules are also plotted. Error bars represent one standard deviation.

excess proton in the structure. In other words, the lowest-energy isomer of the parent cluster obviously plays a major role in determining the fragmentation channels. On the other hand, in the case of long-lived collision complexes, collision energy is transferred to the parent cluster and is redistributed among all degrees of freedom. This is a slow process, and the structures involved during the fragmentation are no longer the lowest-energy isomer, *i.e.*, the structure of the cluster can undergo structural reorganizations before evaporation. Furthermore, the excess proton can also diffuse in the structure and for instance, recombine with the uracil. Then the role of the initial structure of the parent clusters is strongly reduced in determining the fragmentation channels.

In Figure 3.15, I focus on the loss of the neutral uracil molecule in the detected fragments since it indicates where the proton lies after collision, namely, on the uracil or on a water cluster. A transition in the nature of fragmentation product is clearly seen from n = 5-6. To account for this transition, the evaporation originates from a direct fragmentation process is

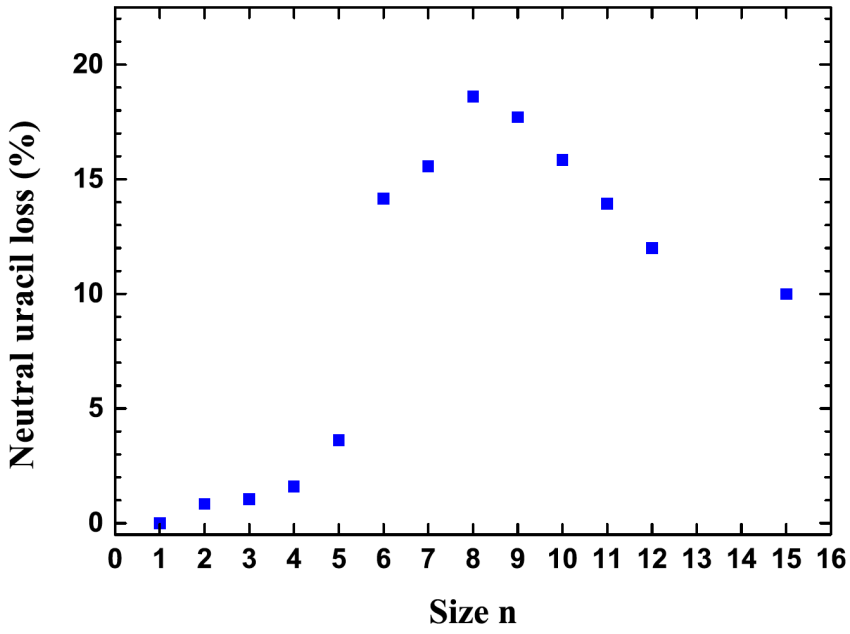


Figure 3.15: Proportion of neutral uracil molecule loss plotted as a function of the number of water molecules n in the parent cluster $(H_2O)_nUH^+$. Results obtained for collisions with Ne atoms at 7.2 eV center of mass collision energy.

considered. A short discussion about the implications of possible structural rearrangement prior to dissociation, which occurs in a statistical process, will be provided in section 3.3.2.2.

The relative proton affinities of each component of the mixed clusters gives a first estimate of which molecule, uracil or water, is more likely to carry the positive charge prior to collisions. Experimentally, the gas phase proton affinity of uracil is bracketed to 9 ± 0.12 eV.[413] For the proton affinity of water molecule, an experimental value is reported at 7.31 eV [414] and a theoretical one at 7.5 eV.[415] In the work of H. Cheng, it shows that the proton affinity of water clusters increases with their size.[415] The proton affinities extracted from the different studies for the uracil molecule and for water clusters as a function of the number of water molecules are displayed in Figure 3.16.

It clearly shows that the proton affinity of uracil, PA[U], is larger than the one of water monomer PA[H_2O]. Thus, for the mono-hydrated uracil, from the energetic point of view, the proton is on the uracil molecule and the only observed fragments are indeed protonated uracil molecules. Moreover, an experimental work [402] confirms that there is no proton transfer from the uracil to the water molecule in mono-hydrated clusters. Proton affinity of the uracil

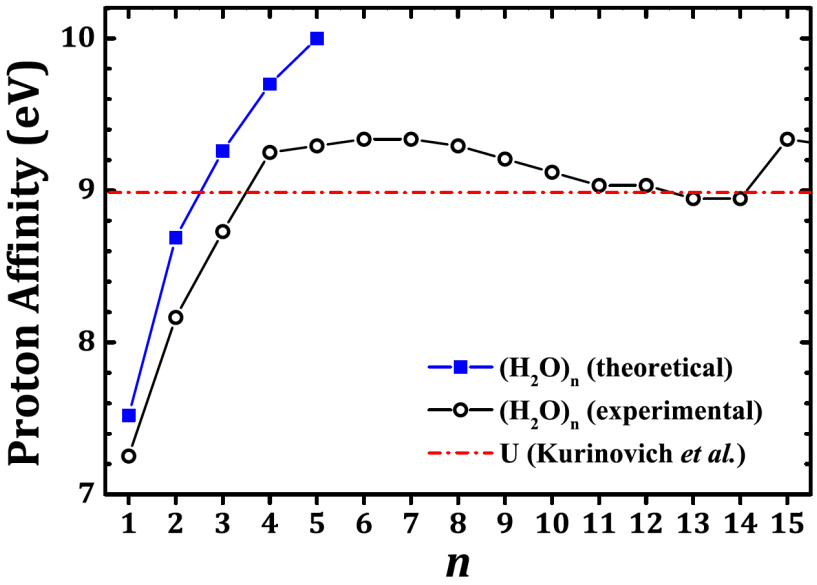


Figure 3.16: The proton affinities of water clusters as a function of the number of water molecules n, which are taken from the work of T. Magnera (black circles) [414] and from the work of Cheng (blue squares).[415] The value of the proton affinity of uracil (red dotted dashed line) is also plotted.[413]

molecule is also larger than that of the water dimer, or even the trimer: PA[U] $>$ PA[$(H_2O)_n$], n = 2 or 3 depending on the considered data for water. This is still consistent with our experimental observation of no neutral uracil molecule loss for n = 2 and 3. However from the PA values, one would predict that the appearance of neutral uracil should occur for n $\approx$ 3-4. For instance, for n = 4, assuming a statistical fragmentation for which the energies of final products are expected to be of relevance, the channel U + $(H_2O)_4H^+$ is energetically favorable. If one now assumes a direct dissociation, where the parent protonation state remains unchanged, one also expects that neutral uracil evaporates. However, experimentally, for n = 4, no neutral uracil evaporation is observed. The loss of neutral uracil starts at n = 5 and becomes significant only at n = 6.

This analysis based on PA is however quite crude. Indeed, it assumes that the protonated uracil cluster would be composed of a uracil molecule attached to an intact water cluster. However, one expects that the hydration of uracil may be more complicated than this simple picture. Therefore, the uracil hydration is explored theoretically in the next Section, Section 3.3.2.2, in order to determine the proton location more realistically.

Table 3.3: Binding energy of two $(H_2O)U$ isomers at MP2/Def2TZVP and SCC-DFTB levels of theory.

isomer	$E_{bind_{MP2}}$	$E_{bind_{DFTB}}$ $D_{NH_{0.0}}$	E_{Re} $D_{NH_{0.0}}$	$E_{bind_{DFTB}}$ $D_{NH_{0.12}}$	E_{Re} $D_{NH_{0.12}}$	$E_{bind_{DFTB}}$ $D_{NH_{0.14}}$	E_{Re} $D_{NH_{0.14}}$
a	-8.3	-8.6	-0.3	-9.8	-1.5	-10.0	-1.7
b	-6.9	-6.6	0.3	-6.9	0.0	-6.9	0.0

3.3.2.2 Calculated Structures of Protonated Uracil Water Clusters

As discussed in section 3.2, I have proposed a modified set of NH parameters to describe sp^3 nitrogen atoms. For, sp^2 nitrogen atoms there is no need to modified the integral parameters as SCC-DFTB describe them rather correctly. Consequently, only the D_{NH} parameter needs to be defined for the present calculations. Table 3.3 present the binding energy of the two $(H_2O)U$ isomers represented in Figure 3.17 at MP2/Def2TZVP and SCC-DFTB levels of theory. Both $D_{\mathrm{NH}} = 0.12$ and $D_{\mathrm{NH}} = 0.14$ lead to consistent binding energies. So, to be consistent with the work performed in the previous section, we have used $D_{\mathrm{NH}} = 0.12$ in the following.

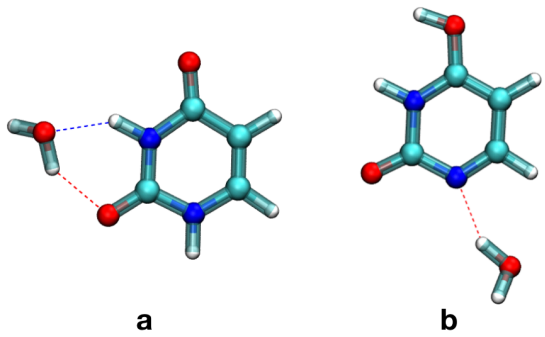


Figure 3.17: Structure of two $(H_2O)U$ isomers used for binding energy calculations.

The lowest-energy isomers determined theoretically for hydrated uracil protonated clusters $(H_2O)_{1-7,11,12}UH^+$ are shown in Figures 3.18-3.27. In the experiments, clusters are produced at a temperature of about 25 K, so only a very few isomers are likely to be populated. Indeed, the clusters are produced in the canonical ensemble at the temperature $T_c \approx 25$ K, so only isomers for which the Boltzmann factor $\exp(-\Delta E k_B T_c)$ is larger than 10^{-7} are considered here. In this formula, ΔE represents the relative energy of a considered isomer with

respect to the lowest-energy one. Thus for each isomer, only the six lowest-energy structures of $U(H_2O)_{1-7,11,12}UH^+$ obtained from the PES exploration will be discussed.

Figure 3.18 displays the six lowest-energy isomers obtained for $(H_2O)UH^+$. Two (1a and 1b) of them contain the u138-like isomer of U (each one with a different orientation of the hydroxyl hydrogen). Three of them (1c, 1d, and 1e) contain the u178 isomer and 1f contains the u137[341] isomer with a reverse orientation of the hydroxyl hydrogen. From those isomers, different sites are possible for the water molecule attachment which leads to variety of isomers even for such small size system. To the best of our knowledge, $(H_2O)UH^+$ is the most studied protonated uracil water cluster and our results are consistent with previous published studies. Indeed, S. Pedersen and co-workers [342] conducted ultraviolet action spectroscopy on $(H_2O)UH^+$ and discussed their measurements in the light of theoretical calculations performed on two isomers: ur138w8 (1a in the present study) and ur178w7 (1c).[342] Their energy ordering at 0 K is the same whatever the computational method they used: B3LYP/6-311++G(3df,2p), M06-2X/6-311++G(3df,2p), MP2/6-311++G(3df,2p), CCSD(T)/6-311++G(3df,2p), and CCSD(T)/augcc-pVTZ and is similar to what I obtained. Similarly, J. Bakker and co-workers[402] considered three isomers: $U(DK)H_W^+$ (1a), $U(KE)H_{Wa}^+$ (1c), and $U(KE)H_{Wb}^+$ (1e) at the B3LYP/6-311++G(3df,2p) level of theory and obtained the same energy ordering as I did. Our methodology has thus allowed us to retrieve those isomers and to locate two new low-energy structures (1b and 1d). 1f is too high in energy to be considered in low-temperature experiments that are in the same range of relative energies but have never been discussed. To ensure that they are not artificially favored in our computational method, calculations were also performed at the B3LYP/6-311++G(3df,2p) level of theory. The results are presented in Figure 3.19, which are consistent with the MP2/Def2TZVP ones. This makes us confident in the ability of the present methodology to locate meaningful low energy structures. Importantly, no isomer with the proton on the water molecule was obtained, neither at the DFTB or MP2 levels.

Figures 3.20 and 3.21 display the six lowest-energy isomers obtained for $(H_2O)_2UH^+$ and $(H_2O)_3UH^+$, respectively. For $(H_2O)_2UH^+$, the lowest energy structure, 2a contains the u138 isomer of uracil. 2b, 2d, and 2e contain u178 and 2c contains u138 with reverse orientation of the hydroxyl hydrogen. 2f contains u178 with reverse orientation of the hydroxyl hydrogen. This demonstrates that, similarly to $(H_2O)UH^+$, a diversity of uracil isomers are present in the low-energy structures of $(H_2O)_2UH^+$ which makes an exhaustive exploration of its PES more difficult. The same behavior is observed for $(H_2O)_3UH^+$. The configuration of u138 does not

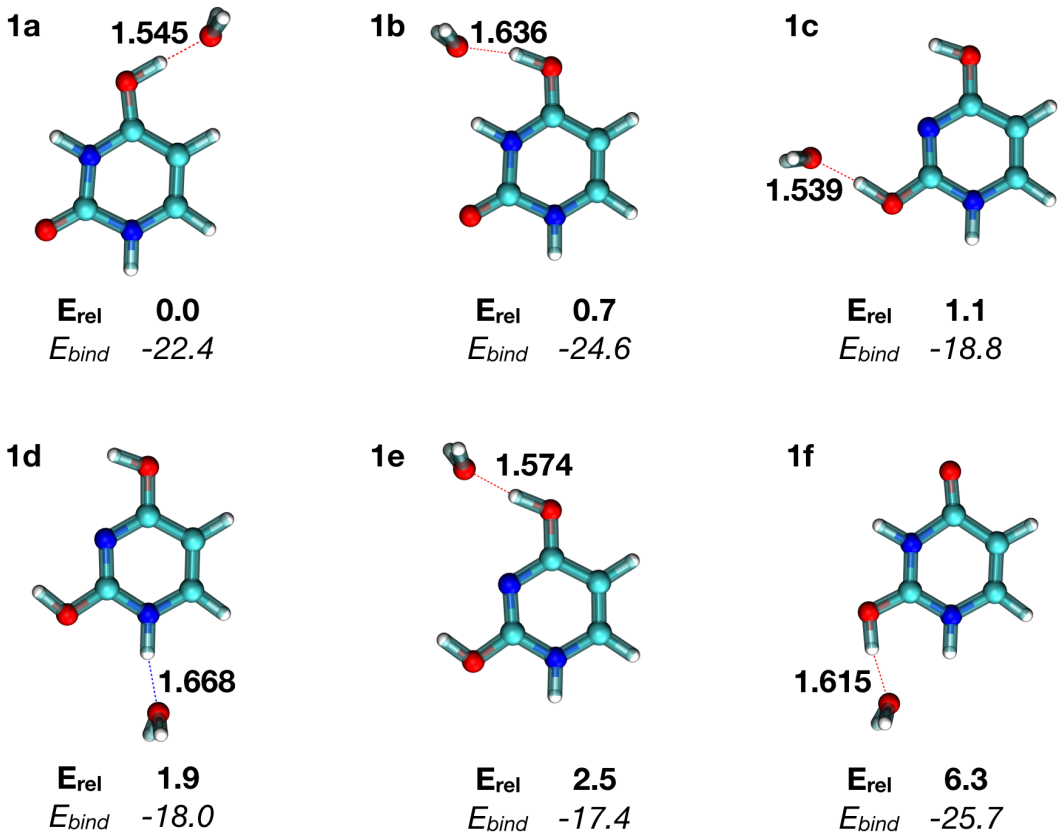


Figure 3.18: Lowest-energy structures of $(H_2O)UH^+$ obtained at the MP2/Def2TZVP level of theory. Relative (E_{rel}) and binding energies (E_{bind}) are given in kcal.mol^{-1}. Important hydrogen-bond distances are indicated in bold and are given in Å.

allow for the formation of a water dimer which leads to two unbound water molecules in 2a. By contrast, a water-water hydrogen bond is observed for 2b and 2c. The existence of a water dimer was not encountered in the low-energy isomers of the unprotonated $(H_2O)_2U$ species due to the absence of the hydroxyl group on U. It is worth pointing out that 2a, 2b, 2c, and 2d are very close in energy which makes their exact energy ordering difficult to determine. However, no isomer displaying an unprotonated uracil in the low-energy isomers of $(H_2O)_2UH^+$ was located. The lowest-energy structure of $(H_2O)_3UH^+$, 3a, is characterized by two water-water hydrogen bond that forms a linear water trimer. Higher energy isomers display only one (3b, 3d, and 3e) or zero (3c and 3f) water-water bond (see Figure 3.21). Similarly to $(H_2O)_2UH^+$, no isomer displaying an unprotonated uracil was located for $(H_2O)_3UH^+$.

The six lowest-energy isomers obtained for $(H_2O)_4UH^+$ and $(H_2O)_5UH^+$ are displayed in

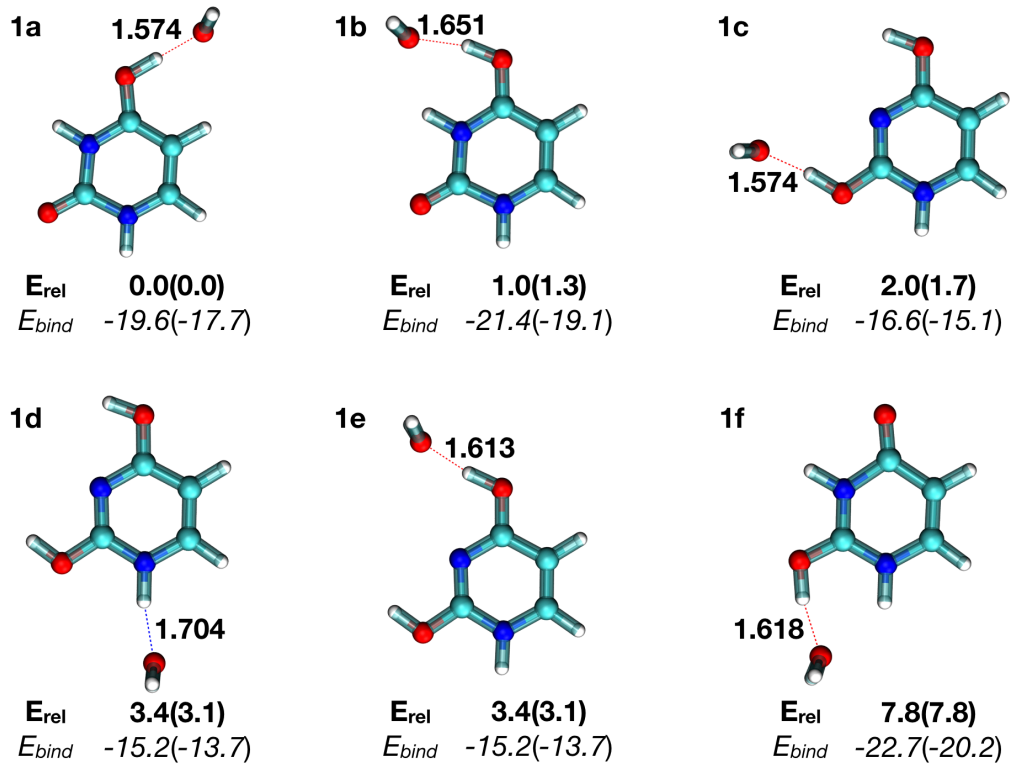


Figure 3.19: Lowest-energy structures of $(H_2O)UH^+$ obtained at the B3LYP/6-311++G(3df,2p) level of theory. Relative (E_{rel}) and binding energies (E_{bind}) are given in kcal.mol^{-1}. The corresponding values with ZPVE corrections are provided in brackets. Important hydrogen-bond distances are indicated in bold and are given in Å.

Figures 3.22 and 3.23, which constitute a transition in the behavior of the proton. Indeed, in $(H_2O)_4UH^+$, two kind of low-lying energy structures appear: (i) structures composed of UH^+, one water trimer, and one isolated water molecule (4b, 4d, 4e, and 4f); (ii) structures composed of U and a protonated water tetramer (4a and 4c). In the latter case, the hydronium ion is always bounded to an uracil oxygen atom. The U–H_2OH^+ bond is always rather strong as compared to U–H_2O bonds as highlighted by the corresponding short oxygen-hydrogen distance. Furthermore, speaking of distances, the difference between the U–H_2OH^+ and UH^+–H_2O forms is rather fuzzy and might be sensitive to computational parameters and also to quantum fluctuations of the hydrogen. This suggests that collision with $(H_2O)_4UH^+$ is more likely to induce evaporation of H_2O rather than H_2OH^+ or a protonated water cluster. The picture is significantly different in $(H_2O)_5UH^+$ where the lowest-energy structure displays a hydronium ion

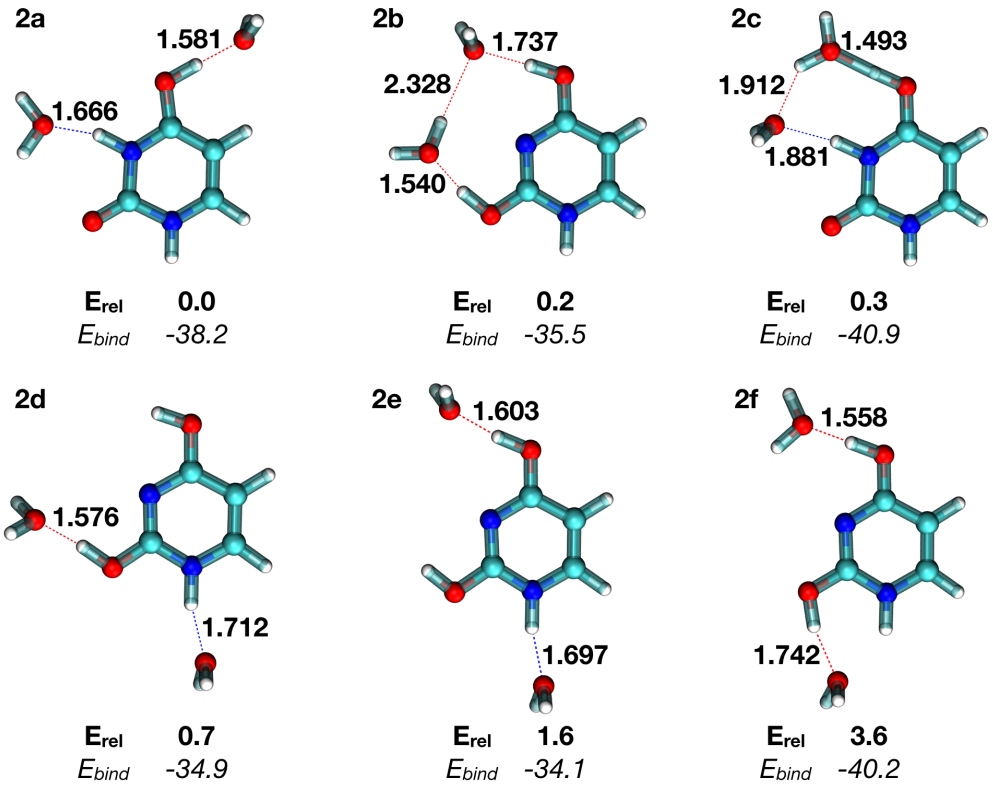


Figure 3.20: Lowest-energy structures of $(H_2O)_2UH^+$ obtained at the MP2/Def2TZVP level of theory. Relative (E_{rel}) and binding energies (E_{bind}) are given in kcal.mol^{-1}. Important hydrogen-bond distances are indicated in bold and are given in Å.

separated by one water molecule from U. Such structures do not appear in $(H_2O)_4UH^+$ due to the limited number of water molecules available to separate H_2OH^+ from U. Such separation suggests that, if considering a direct dissociation process, evaporation of neutral uracil can now occurs in agreement with the experimental observations (see discussion above). One can see that 5b, which is only 0.3 kcal.mol^{-1} higher in energy than 5a, still displays a U–H_2OH^+ link. This is in line with the low amount of neutral uracil that is evaporated in the experiment (see Figure 3.15).

Figures 3.24 and 3.25 display the six lowest-energy isomers obtained for $(H_2O)_6UH^+$ and $(H_2O)_7UH^+$. Similarly to $(H_2O)_5UH^+$, the first lowest-energy structure, 6a and 7a, for both species $(H_2O)_6UH^+$ and $(H_2O)_7UH^+$ have the excess proton on a water molecule that is separated by one water molecule from the uracil. This appears to be common to the clusters with

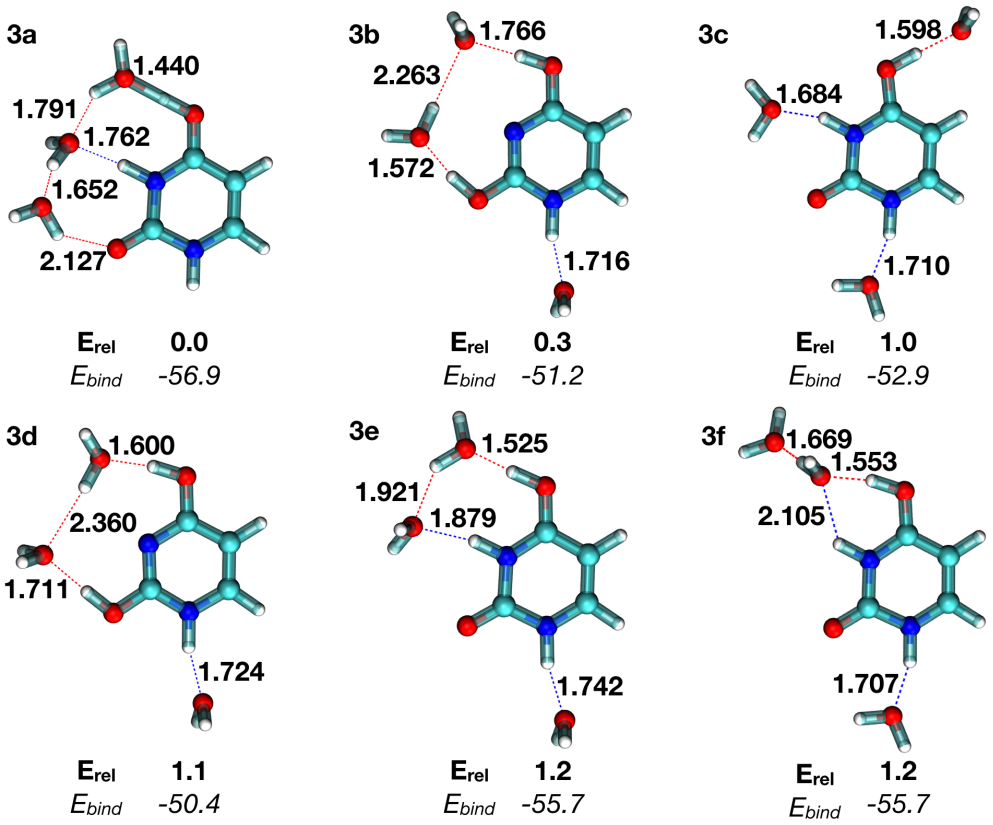


Figure 3.21: $(H_2O)_3UH^+$ lowest-energy structures obtained at the MP2/Def2TZVP level of theory. Relative (E_{rel}) and binding energies (E_{bind}) are given in kcal.mol^{-1}. Important hydrogen-bond distances are indicated in bold and are given in Å.

at least 5 water molecules. This is also observed for higher-energy isomers (6c, 6d, 7c, 7e, and 7f). Other characteristics of the proton are also observed: proton in a similar Zundel form [416] bounded to the uracil (6b, 6e, and 7d) or H_2OH^+ still bounded to uracil (6f and 7b).

Finally, due to the neutral uracil loss proportion starts to decrease from n=9 (see Figure 3.15), which attracted us to perform the optimization of big cluster $(H_2O)_{11,12}UH^+$ as examples to explore why it has this change. The six low-lying energy isomers obtained for cluster $(H_2O)_{11,12}UH^+$ are shown in Figures 3.26 and 3.27.

In all isomers (11a to 11f) of cluster $(H_2O)_{11}UH^+$, the excess is on the water cluster and was separated by water molecule from uracil. For 12a, 12b, 12c, and 12d, it is obvious that the excess proton is not directly bounded to the uracil. The uracil in 12a and 12d belongs to the di-keto form (there is a hydrogen atom on each nitrogen of uracil), and the excess proton

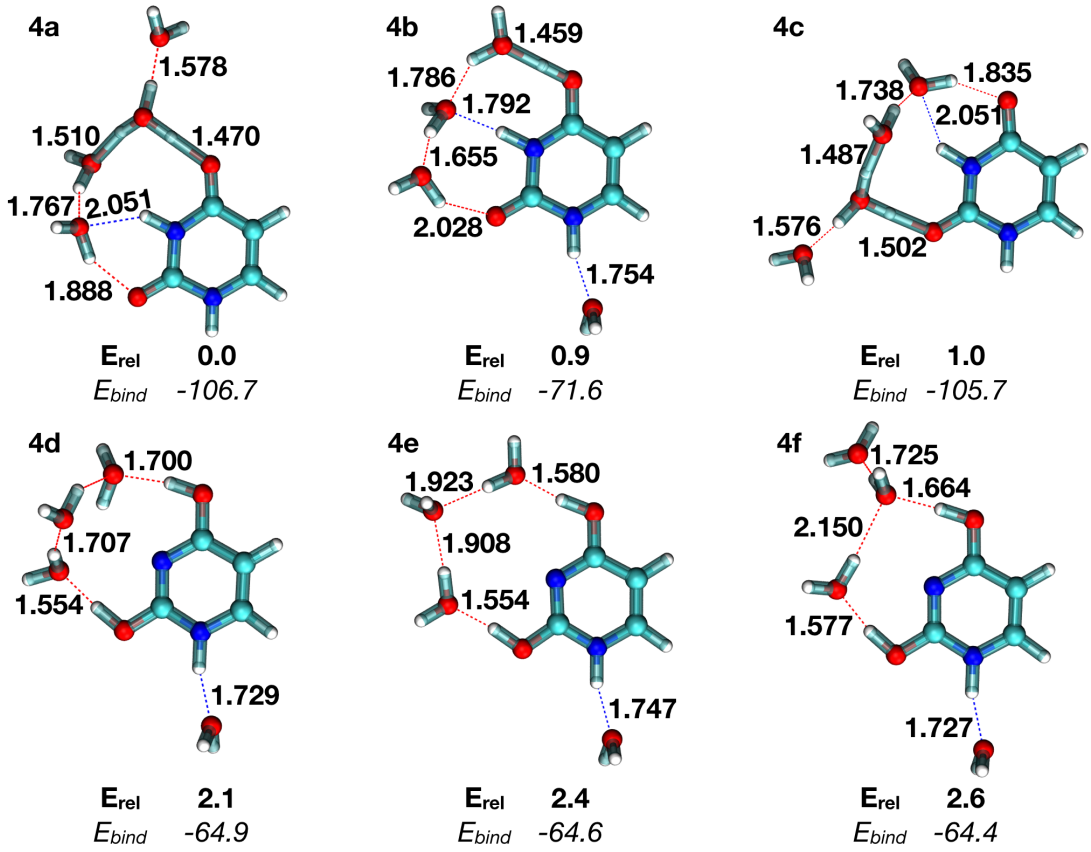


Figure 3.22: Lowest-energy structures of $(H_2O)_4UH^+$ obtained at the MP2/Def2TZVP level of theory. Relative (E_{rel}) and binding energies (E_{bind}) are given in kcal.mol^{-1}. Important hydrogen-bond distances are indicated in bold and are given in Å.

was separated by one water molecule from uracil, additionally, the uracil is surrounded by the water cluster, all of these may lead the excess proton to go to the near oxygen atom of uracil. For 12b, the excess proton is on the water cluster and is very far from the uracil. For 12c, the excess proton was separately by one water molecule from uracil. For isomers 12e and 12f, the excess proton is between the uracil and a water molecule. The uracil is surrounded by the water cluster in 12e but it is not in 12f. Of course, for $(H_2O)_{11}UH^+$, $(H_2O)_{12}UH^+$, $(H_2O)_6UH^+$ and $(H_2O)_7UH^+$ and also $(H_2O)_4UH^+$ and $(H_2O)_5UH^+$, the amount of low-energy isomers is expected to be very large and do not intended to find them all. Furthermore, due to the limited number of MP2 geometry optimization I performed, there are few chances that I located the global energy minima for $(H_2O)_6UH^+$, $(H_2O)_7UH^+$, $(H_2O)_{11}UH^+$ and $(H_2O)_{12}UH^+$. However, the general picture I am able to draw from the present discussed structures fully supports

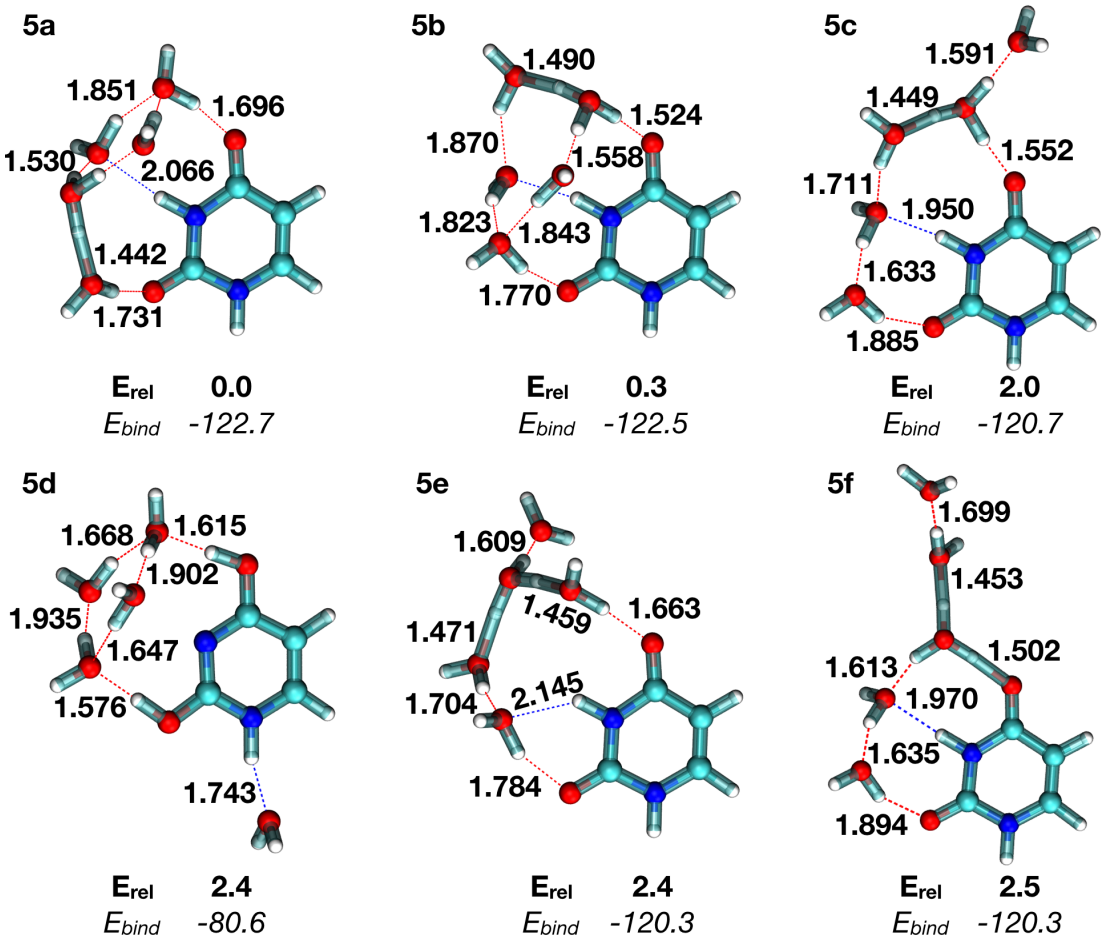


Figure 3.23: Lowest-energy structures of $(H_2O)_5UH^+$ obtained at the MP2/Def2TZVP level of theory. Relative (E_{rel}) and binding energies (E_{bind}) are given in kcal.mol^{-1}. Important hydrogen-bond distances are indicated in bold and are given in Å.

the experimental results: from $(H_2O)_5UH^+$, it exists low-energy structures populated at very low temperature in which the excess proton is not directly bound to the uracil molecule. Upon fragmentation, this allows the proton to remain bounded to the water molecules.

All the aforementioned low-lying energy structures are relevant to describe the $(H_2O)_{1-7,11,12}UH^+$ species at low temperature and to understand the relation between the parent cluster size and the amount of evaporated neutral uracil in the case of direct dissociation. However, as already stated, one has to keep in mind that upon collision statistical dissociation can also occur. In that case, structural rearrangements are expected to occur which are important to understand each individual mass spectra of the $(H_2O)_{1-15}UH^+$ clusters and the origin

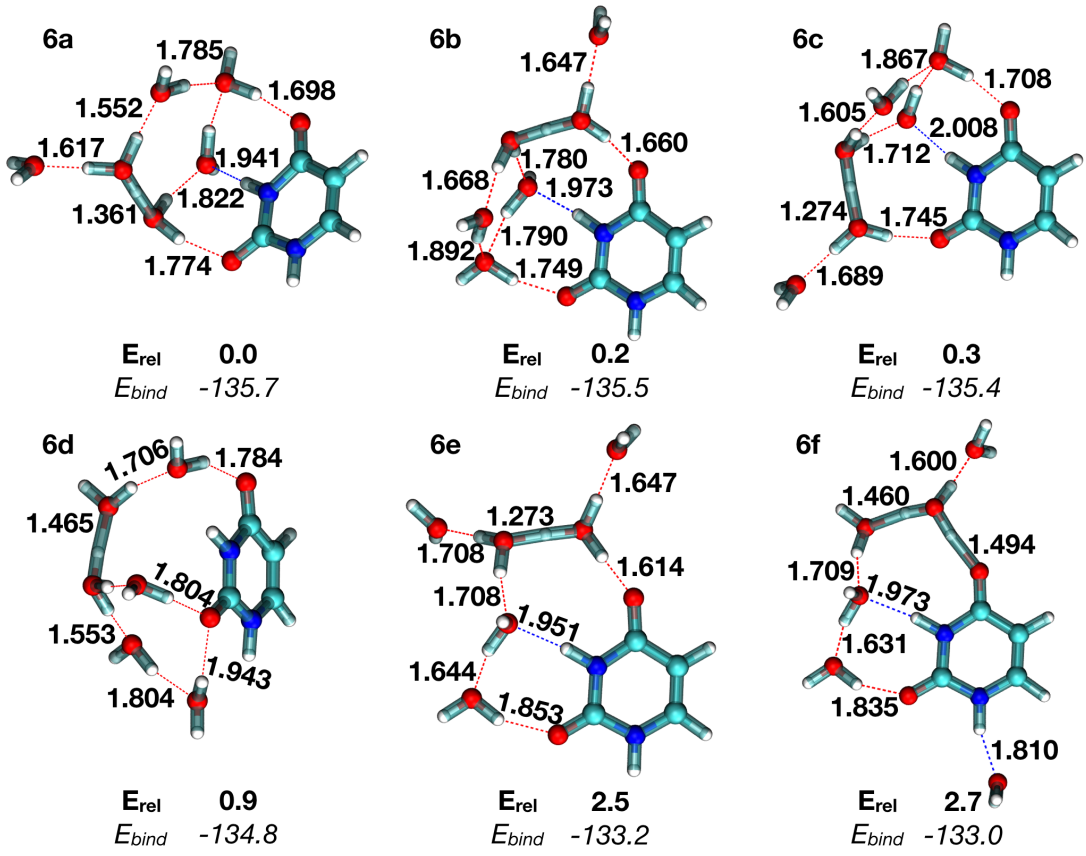


Figure 3.24: Lowest-energy structures of $(H_2O)_6UH^+$ obtained at the MP2/Def2TZVP level of theory. Relative (E_{rel}) and binding energies (E_{bind}) are given in kcal.mol^{-1}. Important hydrogen-bond distances are indicated in bold and are given in Å.

of each collision product. For instance, the fragment UH^+ is detected for all cluster sizes in experiment. This means that for the largest sizes, for which I have shown from the calculation that the proton is located away from the uracil, proton transfer does occur prior to dissociation. One possible scenario is that after collision, water molecules sequentially evaporates. When the number of water molecules is small enough, the proton affinity of uracil gets larger than the one of the remaining attached water cluster. Proton transfer is then likely and therefore protonated uracil can be obtained at the end.

If one turns to the neutral uracil evaporation channel, it appears that the smaller clusters H_2OH^+ and $(H_2O)_2H^+$ are not present in the time-of-flight mass spectra. This absence might have two origins. First, the dissociation energies of the protonated water monomers and dimers

are substantially higher than larger sizes, and they are therefore less prone to evaporation. Second, as already mentioned, for such small sizes, the proton affinity of uracil gets larger than the one of the water dimer or trimer and proton transfer to the uracil is likely to occur.

In order to confirm the above scenarios, simulations and/or evaporation rate calculation would have to be conducted to describe the fragmentation channels in details. MD simulations of protonated uracil have already been performed by R. Spezia and co-workers to understand the collision-induced dissociation.[143, 401] Although, in the present case, the initial position of the excess proton appears as a key parameter to explain the evaporation of neutral uracil, such MD simulations could be additionally conducted to provide a clearer picture on the various evaporation pathways, which will be shown in section 4.3.

3.3.3 Conclusions on $(H_2O)_nUH^+$ clusters

The work in this section presents the collision-induced dissociation of hydrated protonated uracil $(H_2O)_{1-15}UH^+$ clusters and their experimental absolute fragmentation cross sections. The experiments demonstrate that the evaporation channels evolve with size: Below $n = 5$, the observed charged fragments always contain the uracil molecule, whereas from $n = 5$, the loss of a neutral uracil molecule becomes significant. To understand this transition, I conducted an exhaustive exploration of the potential energy surface of $(H_2O)_{1-7,11,12}UH^+$ clusters combining a rough exploration at the SCC-DFTB level with fine geometry optimizations at the MP2 level of theory. Those calculations show that below $n = 5$, the excess proton is either on the uracil or on a water molecule directly bound to uracil, *i.e.*, forming a strongly bound U–H_2OH^+ complex. From $n = 5$ and above, clusters contain enough water molecules to allow for a net separation between uracil and the excess proton: The latter is often found bound to a water molecule which is separated from uracil by at least one other water molecule. Upon direct dissociation, the excess proton and the uracil can thus belong to different fragments. This study demonstrates that combination of collision-induced dissociation experiments and theoretical calculation allows to probe the solvation and protonation properties of organic molecules such as nucleobases. This is a step toward a better understanding of the role of water in the chemistry of in vivo DNA and RNA bases. However, the knowledge of the lowest-energy isomers of the species involved in CID experiments is not enough to understand all the collision process. To get a deeper understanding of the collision mechanism, an explicit modelling of the collision is needed. This question is addressed in the next chapter of this thesis.

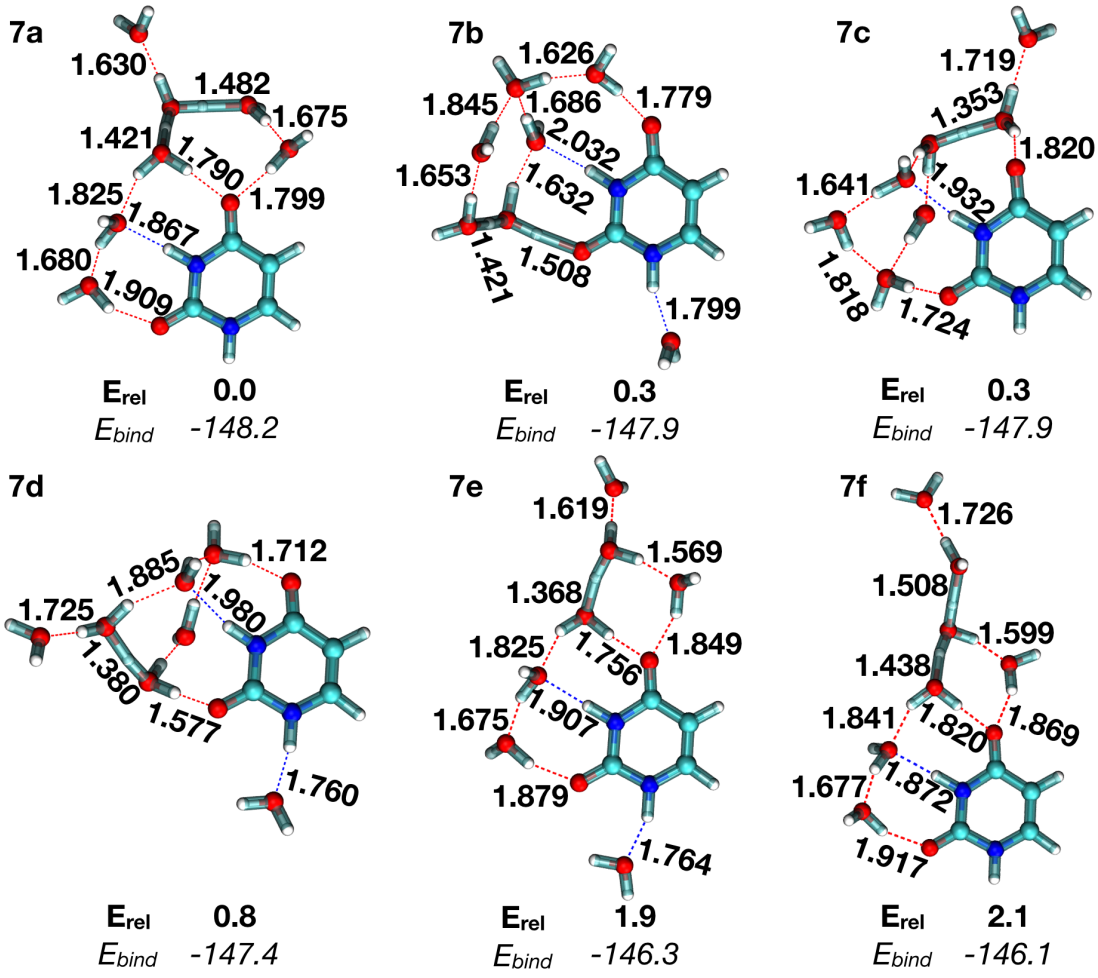


Figure 3.25: Lowest-energy structures of $(H_2O)_7UH^+$ obtained at the MP2/Def2TZVP level of theory. Relative (E_{rel}) and binding energies (E_{bind}) are given in kcal.mol^{-1}. Important hydrogen-bond distances are indicated in bold and are given in Å.

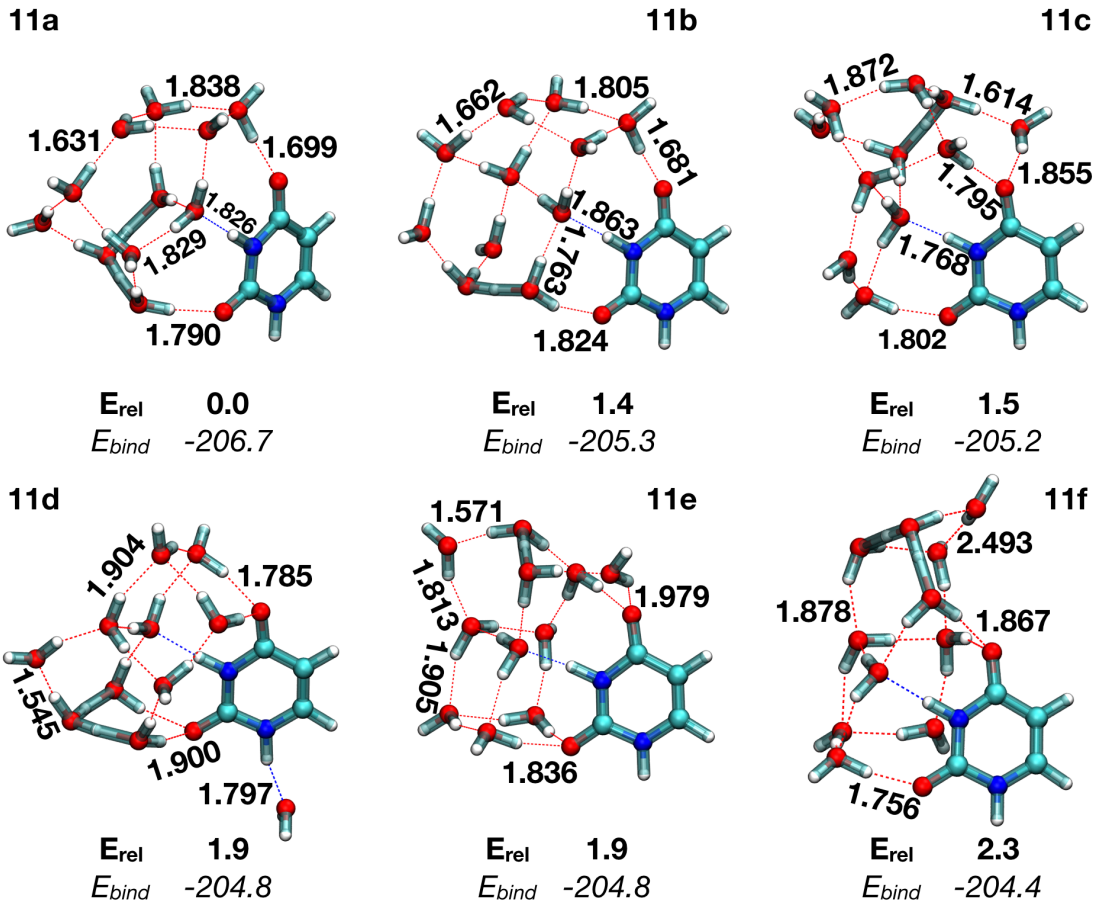


Figure 3.26: Lowest-energy structures of $(H_2O)_{11}UH^+$ obtained at the MP2/Def2TZVP level of theory. Relative (E_{rel}) and binding energies (E_{bind}) are given in kcal.mol^{-1}. Important hydrogen-bond distances are indicated in bold and are given in Å.

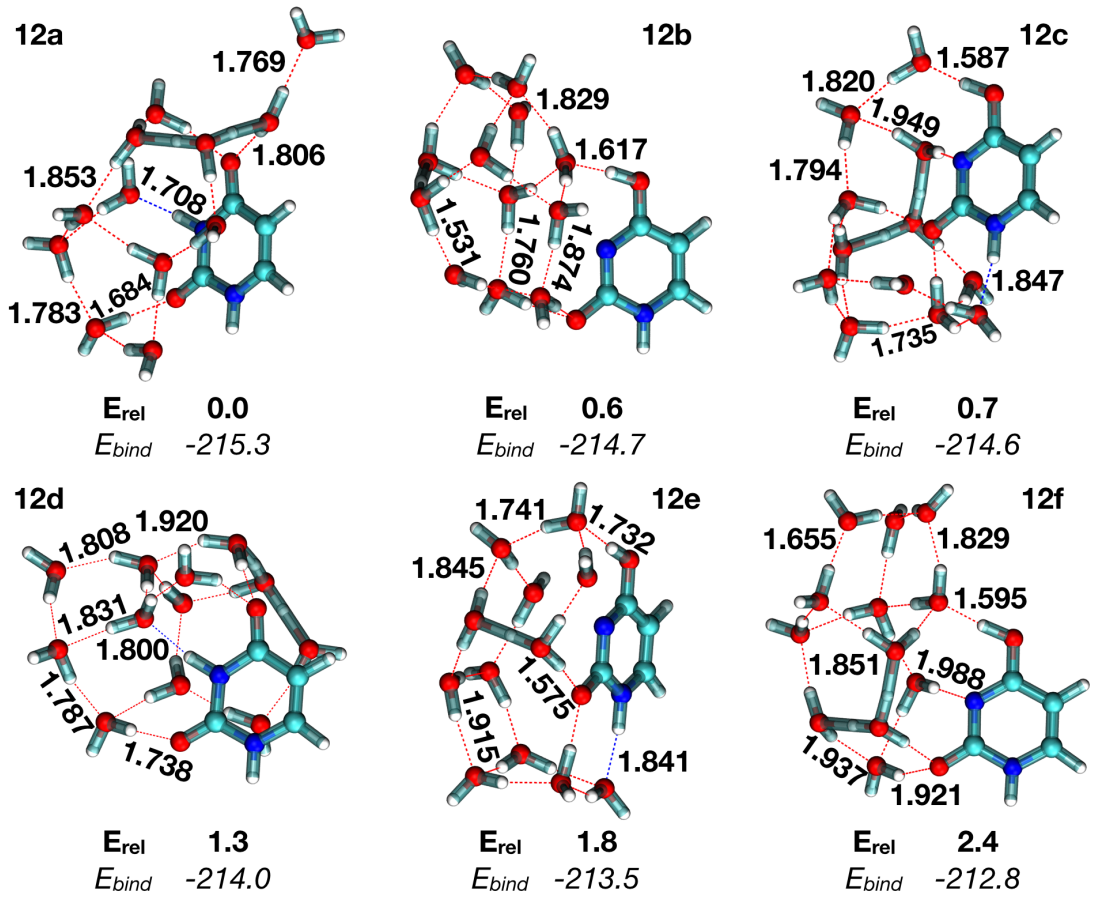


Figure 3.27: Lowest-energy structures of $(H_2O)_{12}UH^+$ obtained at the MP2/Def2TZVP level of theory. Relative (E_{rel}) and binding energies (E_{bind}) are given in kcal.mol^{-1}. Important hydrogen-bond distances are indicated in bold and are given in Å.

4

Dynamical Simulation of Collision-Induced Dissociation

This **fourth chapter of this thesis** merges two independent studies relating to the dynamical simulation of the collision-induced dissociation of $(H_2O)_nUH^+$ clusters and pyrene dimer cation Py_2^+. The two studies in this chapter share the same methodology to generate the collision trajectories for the collision of argon with $(H_2O)_nUH^+$ and Py_2^+. The collision process of the two studies involves dynamical simulations carried out at a QM/MM level where argon is treated as a polarisable MM particle and the lowest-energy targets $(H_2O)_nUH^+$ and Py_2^+ are treated as SCC-DFTB level. The dynamical simulations performed in these two studies allow to visualise the collision trajectories from which it is possible to analyse in details a number of properties. The theoretical results are compared with the CID experimental results conducted on the same systems, *i.e.* $(H_2O)_nUH^+$ and Py_2^+, by S. Zamith and J.-M. l'Hermite, which facilitates their interpretation and complete the CID experiments.

4.1 Experimental Methods

The stability of cluster can be investigated from dissociation experiments. Clusters can be dissociated in electric field, magnetic field, high pressure environment, or by heating (such as absorption of photons) or colliding with energetic particles and so on. For instance, the sodium cluster ions and lithium cluster cation were dissociated with a pulsed UV laser source.[417, 418] Gaseous hydrated trivalent metal ions were dissociated using blackbody infrared radiative dissociation (BIRD).[419, 420] The collision between cluster and high or low energetic par-

ticles at different pressure also have been investigated. Collisions between the high energetic projectile ions (such as 3 keV Ar^+, 22.5 keV He^{2+}) and neutral targets were investigated by Gatchell and A. Holm.[129, 134**?**] Collisions between clusters and projectile have been also explored at low collision energy, which allows for the derivation of dissociation energies and the thermal evaporation and stability of clusters. [135, 421, 422]

By colliding a molecule, or a molecular aggregate, with a non-reactive rare gas atom (neon, argon) or a small molecule such as H_2O or N_2, it is possible to monitor the parent ions and collision products by use, for instance, of tandem mass spectrometry (MS/MS).[423, 424] The resulting mass spectra provide a wealth of information about the structure of the parent and product ions from which one can infer, for instance, dissociation mechanisms [142, 143] or bond and hydration enthalpies [144].

The overall process of collisional activation followed by dissociation/fragmentation is commonly referred to as the collision-induced dissociation (CID) that is also named collisionally activated dissociation (CAD). CID is a mass spectrometry technique to induce dissociation/fragmentation of selected ions in the gas phase, which is one of standard methods for the determination of dissociation/fragmentation pathways. [422, 425] The CID technique consists of accelerating a given ion into a collision gas thereby the ion getting energy and inducing fragmentation. The produced ionic fragments are then mass analyzed, yielding essentially a mass spectrum.[426] The CID technique has been applied in different context. Higher-energy C-trap dissociation is a CID technique specific to the orbitrap mass spectrometer in which dissociation/fragmentation occurs outside the trap [427, 428] Sustained off-resonance irradiation collision-induced dissociation (SORI-CID) is a CID technique used in Fourier transform ion cyclotron resonance mass spectrometry which involves accelerating the ions in cyclotron motion, in a circle inside of an ion trap, in the presence of a collision gas.[429, 430]

CID and the dissociated/fragmented ions produced by CID are used for several purposes: First, partial or complete structural determination can be achieved. Second, CID can simply achieve more sensitive and specific detection. By detecting a unique dissociated/fragmented ion, the precursor ion can be detected in the presence of other ions of the same m/z value, mass-to-charge ratio, which reduces the background and increases the limit of detection.

CID has been applied to a variety of systems, in particular hydrated atomic ions [141, 431, 432, 433, 434] and molecular ions [137, 138, 140, 435]. In the second case, it has been used to understand the impact of high-energy radiations on living cells and DNA or RNA [139, 145, 146], as well as low-energy collisions on molecules of biological interest [147, 148].

Theoretical and experimental studies devoted to fragmentation of hydrated molecular aggregates are scarce, [139, 147, 393, 402, 436, 437] although CID has been applied to water clusters containing an atomic ion [141, 432, 433] and on charged water clusters [136, 402, 410, 431]. This is a real lack as understanding hydration of molecules and biomolecules is of paramount importance to get insights into their structure, stability, dynamics and reactivity in aqueous medium. In that respect, CID investigations could play an important role in understanding those properties in a environment free from long-range solvent effects but also for different hydration degrees or protonation states. This can be evidenced by the experimental study of B. Liu *et al.* on the fragmentation of the singly-charged adenosine 5'-monophosphate (AMP^-) which shows two different fragmentation channel depending on the solvation state of AMP^-.[139] However, to the best of our knowledge, no modelling was performed to complement these experiments except for a few static calculations.[141, 432, 433].

Threshold collision-induced dissociation (TCID) method has also been used, for instance to study the fragmentation patterns and to measure the dissociation energies of clusters.[438, 439] S. Zamith *et al.* did a CID study of the mass-selected protonated uracil water clusters with water molecules and noble gases, respectively.[104] In addition, they also reported the TCID study of pyrene cluster cations. [65] For these two projects, the single collision event is the predominant process. In this chapter, MD simulations based on a quantum chemical formalism are able to model such complex dissociation mechanism to provide an atomic-scale description for these collisions to explain and complete these experiments.

4.1.1 Principle of TCID

In usual TCID setups, experiments are done in ion guides, allowing to perform collisions with large mass atoms such as xenon without losing ions by deflection due to the collision. In order to unambiguously determine dissociation energies, one has to take care of a number of experimental parameters. First, the number of collisions should be as low as possible in order to insure single collision conditions. This can be achieved by performing experiments at various pressures and extrapolating results to zero pressure. Second, one has to consider the possible so-called kinetic shifts that can alter the dissociation energy measurement. Indeed, at threshold collision energy, the system under study might not dissociate during the timescale of the experiment. The apparent threshold has, therefore, to be corrected. This is usually done by extrapolating the experimental values using Rice-Ramsperger-Kassel-Marcus (RRKM) dissociation rates.[149, 150] Third, the initial thermal energy distribution has to be taken into

account. Finally, TCID experimental results are usually fitted assuming a given form for the CID cross section, which can be expressed as [439]:

$$\sigma(E_{col}) = (\sigma_0 n/E_{col}) \sum_i g_i \int_{E_0-E_i}^{E_{col}} [1 - e^{-k(\varepsilon+E_i)\tau}] \times (E_{col} - \varepsilon)^{n-1} d\varepsilon \tag{4.1}$$

where σ_0 is the collision cross section, n is the energy dependence of the reaction cross section, and E_{col} is the collision energy. The populations g_i of rovibrational states with energies E_i are used to carry out the thermal averaging. The dissociation rate k is usually calculated using RRKM type theories, and τ is the typical experimental time between the collision and detection. For comparison with experimental curves, eq 4.1 is further convolved with the kinetic energy distributions of both the ion and neutral reactants. If one needs to incorporate sequential fragmentation and/or competitive channels, these can also be included.[440, 441]

In this method, ion guides are not used. Therefore, it needs to simulate the full ion trajectories in order to ensure that ion losses are correctly taken into account. Collisions are, thus, described with a microscopic model rather than with the average curve given by eq 4.1. This approach allows to quite naturally include sequential dissociation and to potentially test energy transfer models. One advantage of the setup resides in the fact that the systems under study are thermalized at low temperature prior to collisions. This implies that averaging over thermal energies of the parent ion plays a minor role, thus leading to minor uncertainties.

4.1.2 Experimental Setup

The experimental setup of Laboratoire Collisions Agrégats Réactivié (LCAR) by S. Zamith *e*t al. for the collision of protonated uracil water clusters or pyrene dimer cation with noble gas atoms is set up as follows:

Clusters are produced in a gas aggregation source [442](a) and then thermalized (b) at 25 K through thousands of collisions with helium. The experimental setup can be used in two modes. In the first mode, only the first Wiley-McLaren acceleration stage (c) is used to work together with the reflectron (i). Clusters are detected using dual micro-channel plates (MCPs) (j) biased at-10 kV. This allows to perform regular Time of Flight Mass Spectrometry (TOFMS) and to optimize the cluster production. In this mode, the mass filter (d), the electrodes for energy focusing (e) and deceleration (f), and the second Wiley-McLaren acceleration stage (h), are grounded.

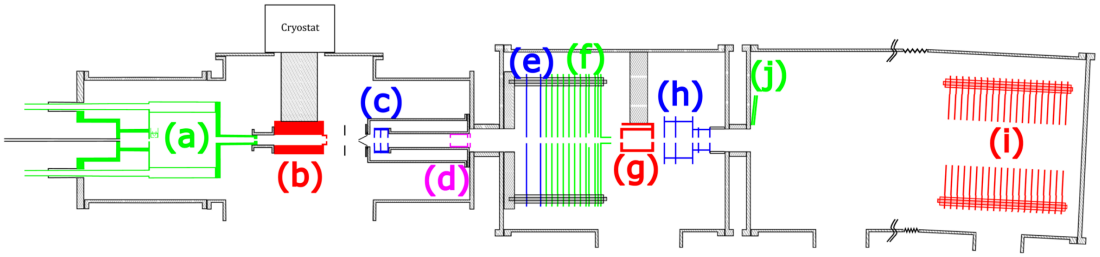


Figure 4.1: Schematic view of the experimental setup. (a) Cluster gas aggregation source. (b) Thermalization chamber. (c) First Wiley–McLaren acceleration stage. (d) Massfilter. (e) Energy focusing. (f) Deceleration. (g) Collision cell. (h) Second Wiley–McLaren acceleration stage. (i) Reflectron. (j) Micro-channel plate detector.

In the second mode, all the electrodes were used to mass-select the clusters. In order to perform collisions between the mass-selected clusters and the rare gas atoms, precisely delayed high voltage pulses were applied to electrodes (c)-(f). Pulsed high voltages applied to the first Wiley–McLaren electrodes (c) accelerate all the clusters, giving them an average kinetic energy of 622 eV. The applied voltages and the spacing between the electrodes of the Wiley-McLaren are chosen such that, 26 cm downstream, there is a linear relation (to first order) between the position of clusters and their kinetic energy. Using a pulsed high voltage, an electric field is created in this region (e) that compensates this linear kinetic energy dispersion, and all clusters finally have the same kinetic energy within a few electron volts. The time at which this pulsed high voltage is applied determines which cluster size is correctly energy focused. After this kinetic energy focusing, ions are decelerated by a potential barrier (f). At the end of the potential barrier, the potential is shut down in a field free zone and the mass-selected clusters then fly freely through the collision cell (g) up to the second Wile-McLaren acceleration stage (h). Clusters are then mass-analyzed using the reflectron (i) and the MCP detector (j). High voltage is applied on the mass filter (d) when the mass of interest enters the cylinder and shut down before it comes out. This allows us to eliminate part of the neighboring masses. In the experiments of pyrene clusters, the kinetic energy of the clusters in the laboratory frame is varied between 5 eV and 200 eV. For the experiments of protonated uracil water clusters, the kinetic energy of the clusters in the laboratory frame is 100 eV.

Kinetic energies of the ions can be easily deduced from experimental parameters. Indeed, since the distances in the apparatus are well-known, measuring, for instance, the time the ions take to travel from the end of the slowing down stage to the second acceleration stage gives the speed of the ions. More precise kinetic energy calibration is obtained by recording the

signal of the ions as a function of delays and/or voltages. These curves are then reproduced by simulations to obtain the kinetic energy distribution of the ions. [443]

The simplified experiment setup is shown in Figure 4.2. Clusters are produced in a gas aggregation source and thermalized at a temperature of 25 K. Clusters are then mass-selected with a chosen kinetic energy, which collide with argon atoms in a collision cell. The collision products are then analysed by TOFMS.

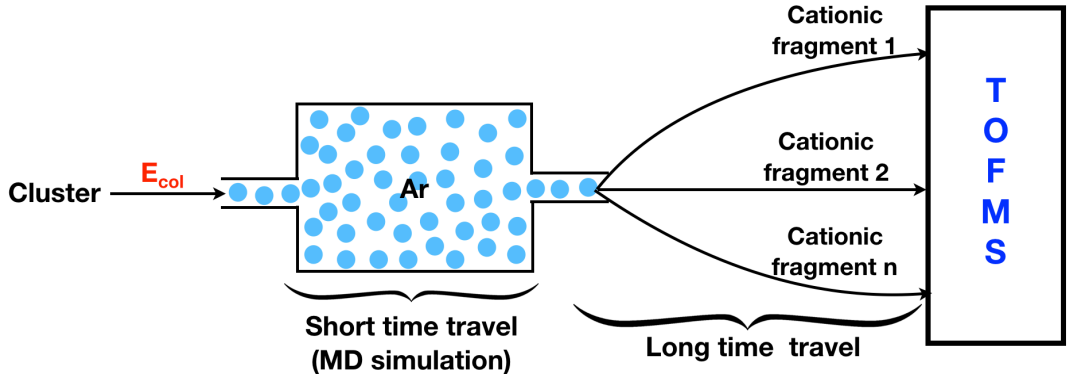


Figure 4.2: Schematic of the simplified experimental setup.

4.2 Computational Details

4.2.1 SCC-DFTB Potential

For the work in this chapter, the SCC-DFTB in combination with the mio-set for the Slater-Koster tables of integrals is applied. [168, 169, 218, 219, 444, 445] The SCC-DFTB potential for protonated uracil water clusters is shown in section 3.1.

DFTB is an efficient tool to perform MD simulations, in particular addressing the evaporation/dissociation processes in various chemical systems. [62, 446, 447, 448] The dynamics simulations of the collision process were performed with a QM/MM scheme [449] whose details can be found in the original paper[450, 451]. The projectile argon is treated as a polarisable MM particle interacting with the target (protonated uracil water cluster or pyrene dimer cation Py_2^+), the latter being treated at the DFTB level. The oscillation problem often appears for dissociated or close to dissociation systems. For the collision trajectories described below, a Fermi distribution (Fermi temperature 2000 K) was applied to avoid oscillation issues during the self-consistent procedure.[452] The Fermi distribution allows to recover the continuity in

energy and gradients in the case of level crossing. [452] It should be mentioned that, in order to keep a low computational cost, no correction has been used to improve the DFTB charge resonance description. However, this charge delocalization issue has been specifically addressed in the case of PAH cation dissociation and it was shown to have a minor effect on the final computed mass spectra [446]. I also mention that the collision energy is in principle high enough to have electronic excitation in the system, which is taken into account at a crude level by the use of a Fermi temperature. Finally, nuclear quantum effects are not taken into account. Although this may affect the results at very low collision energies, the effect is expected to be small for the experimental collision energies of 7.2 eV and 17.5 eV. Although all these limits should be kept in mind, I would like to emphasize that, recently, the dissociation of PAH molecules has been simulated and a good agreement with experimental results was obtained despite similar crude approximations, namely neglect of non-adiabatic and nuclear quantum effects, improper treatment of charge delocalization and use of a Fermi temperature [446, 448, 453].

4.2.2 Collision Trajectories

The preparation for the collisional trajectories for the collision of protonated uracil water clusters or Py_2^+ with Ar is the same. The schematic example for the collision of Py_2^+ with Ar is shown in Figure 4.3. Starting from the optimized Py_2^+ geometry [454], a preliminary thermalization run of 200 fs at 25 K (maintained by a Nosé-Hoover chain thermostat [309, 311]) is performed. Then, the argon atom projectile is introduced in the simulation with a velocity determined to reproduce a given collision energy. The target dimer Py_2^+ was positioned at the origin of the simulation referential and randomly rotated to allow all possible impact points on the dimer. The argon atom is initially positioned at x=10, y=b and z=0 Å, with b being the impact parameter. At each center of mass collision energy E_{col}, a series of 300 collision trajectories were conducted (the center of mass of the aggregate was kept at position (0, 0, 0)) for each of the 13 b values which are evenly distributed (interval being 0.5 Å) between 0 and R+0.5 Å. R refers to the radius of Py_2^+. 600 collision trajectories were performed per isomer of protonated uracil water clusters. Trajectory calculations have been performed with a time step of 0.5 fs and a total duration of 15 ps and 3 ps for the collision of argon with protonated uracil water clusters and Py_2^+, respectively. For the collision of Py_2^+ with argon, I have checked that for high collision energies such as 20 and 25 eV, a time step of 0.1 fs does not change significantly our numerical results, which will be shown in section 4.4.4.2.

It should be noted that the quaternion was used for rotation process in the generation of initial inputs. This approximation allows us to go from $3n-6$ degrees of freedom to $6N-6$, where n and N are the number of atoms and the number of molecules (3 degrees of translation and 3 degrees of rotation) in a system, respectively. The complex quaternion formalism ($\boldsymbol{q}$ = q_0, q_1, q_2, q_3) was used o describe the orientation of a solid body with respect to Euler angles ((θ, ϕ, ψ)) formalism. The quaternions involve an additional degree of freedom (similar to a homothety), which can be offset by using a normalization constraint on the vector $\boldsymbol{q}$:

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1 \tag{4.2}$$

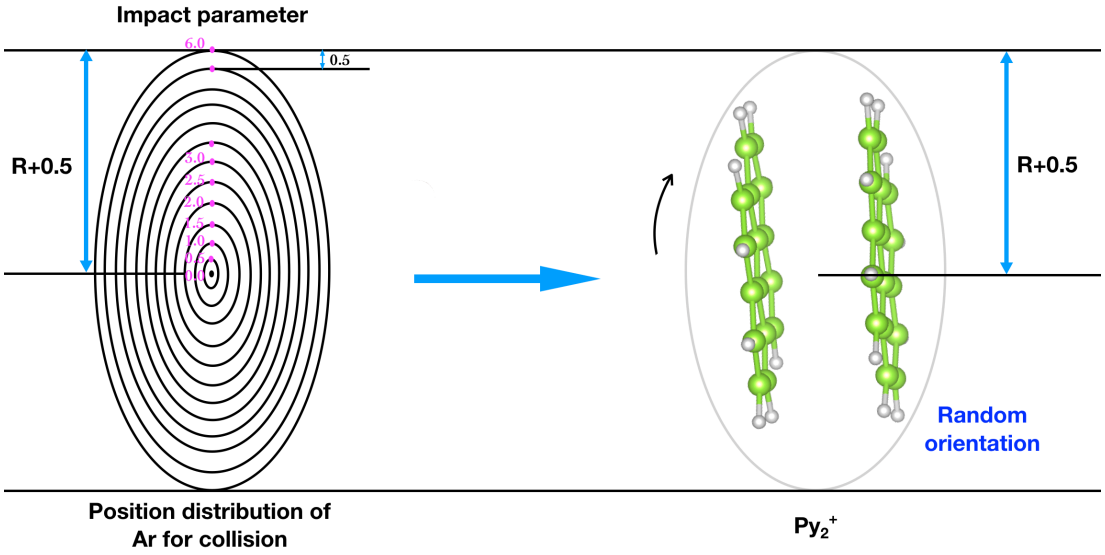


Figure 4.3: Schematic of the generation of the initial inputs.

4.2.3 Trajectory Analysis

During the results collection, the final snapshot was extracted for each trajectory. For collision of between Py_2^+ and argon, when the center of mass of the two pyrene monomer is more than 10 Å, a dissociation is defined. For the dissociaton definition of protonated uracil water clusters, it is a little more complicated than Py_2^+. A fragment is defined as a group of atoms in which the distance of any pair of adjacent atoms is less than 5.0 Å. The number of hydrogen, nitrogen and oxygen atoms in one fragment is denoted by k, l and m, respectively. For instance, a fragment

characterised by $l = 0$ and $k = 2m + 1$ is a pure water cluster containing the excess proton. Identifying such a fragment at the end of the trajectory means that a neutral uracil fragment exists, otherwise the excess proton is located on a uracil containing fragment. In practice, at each time step, the fragments are identified on the basis of their k, l, m values, allowing to record their time-dependent evolution. The mass spectrum is built retaining only the fragments containing the excess proton, as only charged fragments are detected in the experiment.

The opacity $P(b, E_{col})$, *i*.e. the dissociation probability as a function of impact parameter at a given collision energy is computed by averaging the results over the simulations corresponding to these conditions. The cross sections are then derived from the following formula:

$$\begin{aligned} \sigma_{frag}(E_{col}) &= \int_0^{b_{max}} 2\pi P(b, E_{col}) b db \\ &\simeq \sum_{i=0}^{b_{max}} \frac{P(b_i, E_{col}) + P(b_{i+1}, E_{col})}{2} \pi (b_{i+1}^2 - b_i^2) \end{aligned} \tag{4.3}$$

Mean values are computed using the same approach, followed by a division by πb_{max}^2. When mean values are restricted to trajectories leading to dissociation (noted $-d$) or not (noted $-ud$), additional normalisation by the total number of dissociated or undissociated trajectories is also necessary.

4.3 Dynamical Simulation of Collision-Induced Dissociation of Protonated Uracil Water Clusters

4.3.1 Introduction

Motivated by the recent CID experiments conducted by I. Braud *et al.* consisting in $(H_2O)_{1-15}UH^+$ clusters colliding with an impacting atom or molecule M (M = H_2O, D_2O, neon, and argon) at a constant center of mass collision energy of 7.2 eV,[104] the dynamical simulations of the collision between the protonated uracil water clusters $(H_2O)_{1-7,11,12}UH^+$ and an argon atom were performed. The low collision energy (7.2 eV) only leads to intermolecular bond breaking, without any electronic excitation, rather than intramolecular bond breaking. The branching ratios for different charged fragments were determined in experiment, which allows to deduce the fragmentation cross section for all $(H_2O)_{1-15}UH^+$ species and the location of the excess proton after collision: on a uracil containing cluster or on a pure water cluster. This allows to determine the proportion of neutral uracil loss (corresponding to cases where the excess proton is located on pure water clusters) as a function of the number n of water molecules. A

sharp increase of neutral uracil loss was observed for n = 5-6 (2.8% and 25.0% for n = 4 and 7, respectively). Those experiment were complemented by theoretical calculations that aim at characterizing the lowest-energy isomers of $(H_2O)_nUH^+$ (n = 1-7, 11, 12) clusters (see section 3.3), which shows that the location of the proton after collision recorded in the CID experiment is determined by its position in the lowest-energy parent isomer. In other words, a shattering mechanism occurs after collision. Despite these findings, static calculations can not provide a full picture for the fragmentation process and some issues are still not properly understood: (i) What is the main path of the fragmentation mechanisms? (ii) What are the fragments after collision? (iii) How does the proportion of fragments change according to time? (iv) Is the proportion of neutral uracil molecules loss only determined by the nature of the lowest-energy isomers?

To answer these questions, this simulations present a complete MD study of the fragmentation process for $(H_2O)_{1-7,11,12}UH^+$ aggregates colliding with an argon atom. Section 4.3.2 discusses the statistical convergence of collision trajectories, theoretical time-dependent proportion of fragments, proportion of neutral uracil loss, total fragmentation cross sections and mass spectra of fragments bearing the excess proton. These data are compared to available experimental results in order to discuss in details dissociation mechanism as a function of n. The main outcomes are summarized in section 4.3.7.

4.3.2 Results and Discussion

4.3.2.1 Statistical Convergence

In order to ensure that the statistical convergence is reached in the collision trajectories, initial conditions have to reproduce all possible collision orientation with good statistics. The procedure to generate a set of collision trajectories is described in section 4.2.2. As a visual proof, pictures a, b and c in Figure 4.4 represent 200, 400 and 600 random argon orientations with impact parameter being 0 for cluster $(H_2O)_3UH^+$, respectively. In these pictures, $(H_2O)_3UH^+$ is fixed and all initial positions for argon are orientated which leads to distribution maps of the initial positions of argon with respect to fixed $(H_2O)_3UH^+$. It is worth pointing out that in the collision trajectories, argon is fixed and uracil is rotated. Picture d in Figure 4.4 presents 200 random argon orientations with impact parameter being 0.0 and 6.0, respectively. The similar pictures for cluster $(H_2O)_{12}UH^+$ are displayed in Figure 4.5. These Figures demonstrate that

the more collision trajectories are performed, the more colliding opportunities of argon at all possible orientations are obtained.

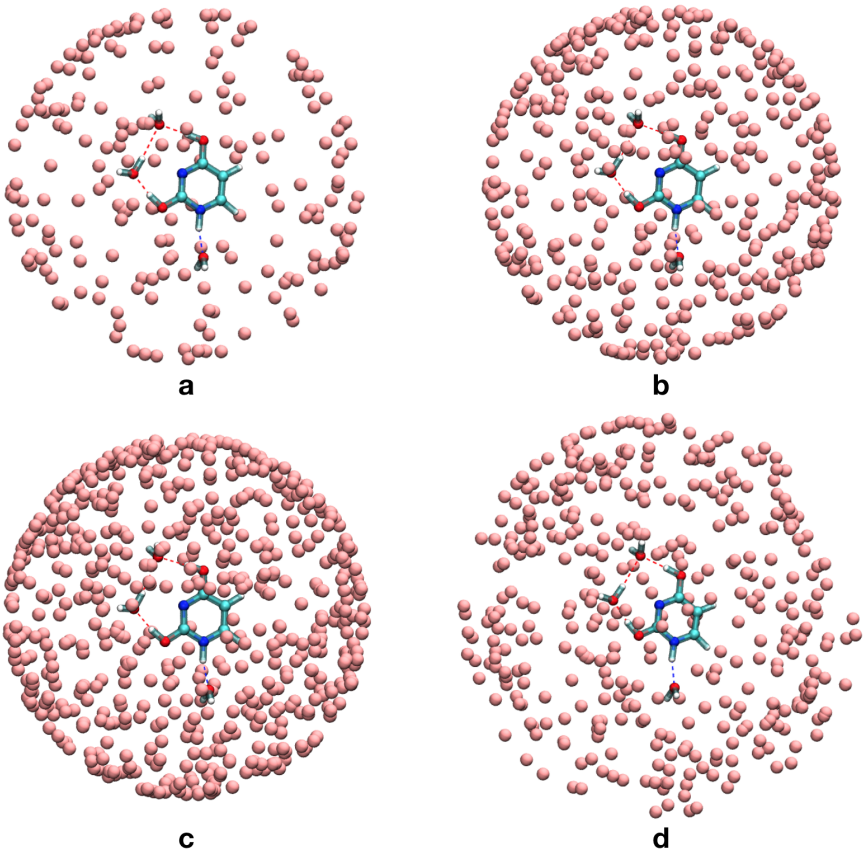


Figure 4.4: Schematic representation of random argon orientations for the collision with the second lowest-energy isomer of cluster $(H_2O)_3UH^+$. 200 (a), 400 (b) and 600 (c) random argon orientations are generated with impact parameter being 0. 200 orientations are generated with impact parameter being 0 and 6 (d), respectively.

In addition, to confirm that statistical convergence is reached for the properties discussed in sections 4.3.3, 4.3.4, 4.3.5, and 4.3.6, Tables 4.1 and 4.2 present the P_{NUL} involved from the

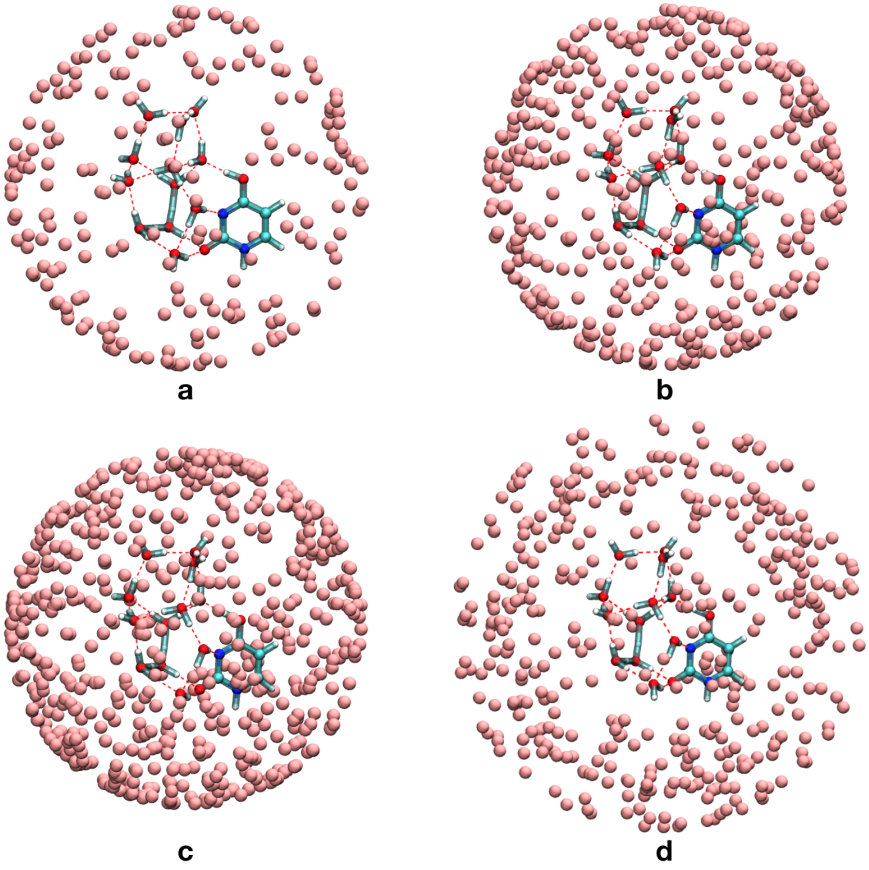


Figure 4.5: Schematic representation of random argon orientations for the collision with the second lowest-energy isomer of cluster $(H_2O)_{12}UH^+$. 200 (a), 400 (b) and 600 (c) random argon orientations are generated with impact parameter being 0. 200 orientations are generated with impact parameter value being 0 and 7 (d), respectively.

following formula 4.4

$$\begin{aligned} P_{NUL}(E_{col}) &= \int_0^{b_{max}} N_{NUL}(b,E_{col})2\pi b db \Big/ \int_0^{b_{max}} N_{frag}(b,E_{col})2\pi b db \\ &\simeq \frac{\sum_{i=0}^{b_{max}} \frac{1}{2}(N_{NUL}(b_i,E_{col})+N_{NUL}(b_{i+1},E_{col}))\pi(b_{i+1}^2-b_i^2)}{\sum_{i=0}^{b_{max}} \frac{1}{2}(N_{frag}(b_i,E_{col})+N_{frag}(b_{i+1},E_{col}))\pi(b_{i+1}^2-b_i^2)} \end{aligned} \tag{4.4}$$

and σ_{frag} of two isomers (the first lowest energy isomer and the one whose P_{NUL} fits best to the experiment results (in bold)) of each cluster $(H_2O)_{1-7,11,12}UH^+$ obtained from 200, 400, and 600 random argon orientations per impact parameter value.

Whatever the considered isomer, the three P_{NUL} and σ_{frag} values from 200, 400, and 600 random argon orientations are very close. Indeed, the largest difference is observed for isomer 7a which has P_{NUL} values of 29.5 and 31.3 % for 200 and 600 random orientations, respectively. This demonstrate that even for 200 initial random orientations, simulation are close to statistical convergence. In the present study, all results discussed in the main text were obtained with 600 initial random argon orientations per impact parameter value which ensures statistical convergence of the results independently of cluster size.

4.3.3 Time-Dependent Proportion of Fragments

The time-dependent proportion of each fragment was extracted from collision trajectories. To illustrate the change in behavior resulting from the difference in cluster size, Figures 4.6-4.11 display the time-dependent proportion of fragments obtained from the dissociation of the low-lying energy isomers of $(H_2O)_{1-7,11,12}UH^+$.

I will discuss the time-dependent proportion of one small cluster (7a) and one big cluster (12a) in detail as an example. For the sake of clarity, only the fragments displaying significant proportion, higher than 0.035 and 0.015 for 7a and 12a, respectively are considered in Figure 4.10. This corresponds to the eight and ten most prominent fragments for 7a and 12a, respectively. From Figure 4.10, it is clear that for both aggregates, the proportion of H_2O has the sharpest increase after collision and then stay almost constant as a function of time. For 7a, $\sim$3 ps after collision, the proportion of almost all fragments does not change any more. Before that, the proportion of the $(H_2O)_6UH^+$ fragment increases first and then decreases, which indicates a sequential dissociation of water molecules. For 12a, proportion of $(H_2O)_{11}UH^+$ and $(H_2O)_{10}UH^+$ fragments displays a sharp increase quickly after collision which is then followed by a fast decrease, and finally it keeps a minute decrease up to the end of the simulations. The decrease of proportion of $(H_2O)_{10}UH^+$ and $(H_2O)_{11}UH^+$, and the increase of proportion of $(H_2O)_6UH^+$, $(H_2O)_7UH^+$ and $(H_2O)_8UH^+$ indicate sequential dissociation after collision is occurring. It is worth noting that, in contrast to 7a, the proportions of the main fragments of 12a do not tend to be a constant at the end of the simulations. This implies that, for this large aggregate, structural rearrangements are more likely to occur prior to complete dissociation. Proportions of the main fragments of clusters 7d and 12c shown in Figure 4.11 display similar behavior as for 7a and 12a. As a first conclusion, Figure 4.10 suggests that clusters with 7 water molecules experience a direct dissociation mechanism as was hypothesised by I. Braud *et al.*.[104] A similar conclusion can be drawn for smaller cluster sizes as supported by Figures

Table 4.1: The proportions of P_{NUL} and σ_{frag} of first lowest-energy isomer and the isomer whose P_{NUL} fits the experiment (in bold) of $(H_2O)_{1-5}UH^+$ with simulations of 200, 400, and 600 as initial conditions.

Cluster	Simu	P_{NUL} (%)	σ_{frag} (Å^2)
1a	200	0.1	28.4
1a	400	0.1	28.3
1a	600	0.2	28.9
1b	200	0.2	26.3
1b	400	0.1	25.7
1b	600	0.1	25.9
2a	200	0.0	35.9
2a	400	0.0	36.5
2a	600	0.0	36.3
2b	200	0.0	34.7
2b	400	0.1	34.8
2b	600	0.1	34.9
3a	200	5.4	37.4
3a	400	5.2	36.2
3a	600	5.7	36.3
3b	200	0.0	41.2
3b	400	0.0	41.5
3b	600	0.0	41.9
4a	200	26.9	40.1
4a	400	28.2	40.3
4a	600	29.4	40.1
4b	200	2.7	45.3
4b	400	2.6	45.6
4b	600	2.6	45.2
5a	200	37.2	45.7
5a	400	37.8	46.1
5a	600	38.2	46.6
5d	200	0.1	47.3
5d	400	0.1	47.3
5d	600	0.1	47.5

Table 4.2: The proportions of P_{NUL} and σ_{frag} of first lowest-energy isomer and the isomer whose P_{NUL} fits the experiment (in bold) of $(H_2O)_{6,7,11,12}UH^+$ with simulations of 200, 400, and 600 as initial conditions.

Cluster	Simu	P_{NUL} (%)	σ_{frag} (Å^2)
6a	200	38.0	46.6
6a	400	38.0	45.6
6a	600	39.3	45.8
6f	200	18.9	54.2
6f	400	19.0	55.2
6f	600	18.5	55.0
7a	200	29.5	54.8
7a	400	31.3	53.4
7a	600	31.3	53.4
7d	200	22.6	55.3
7d	400	22.9	54.3
7d	600	23.0	54.0
11a	200	26.7	53.8
11a	400	28.2	53.5
11a	600	28.3	52.9
11d	200	14.5	55.2
11d	400	15.4	56.1
11d	600	15.6	56.5
12a	200	8.0	59.2
12a	400	7.5	60.5
12a	600	7.6	60.2
12c	200	10.4	55.3
12c	400	10.8	55.8
12c	600	10.8	55.4

4.6-4.8. In contrast, cluster with 11 (see Figure 4.9) and 12 water molecules shows a behavior compatible with a certain amount of statistical dissociation, namely a long-time evolution that allows structural rearrangements. These important observations can now be refined by looking at more detailed properties.

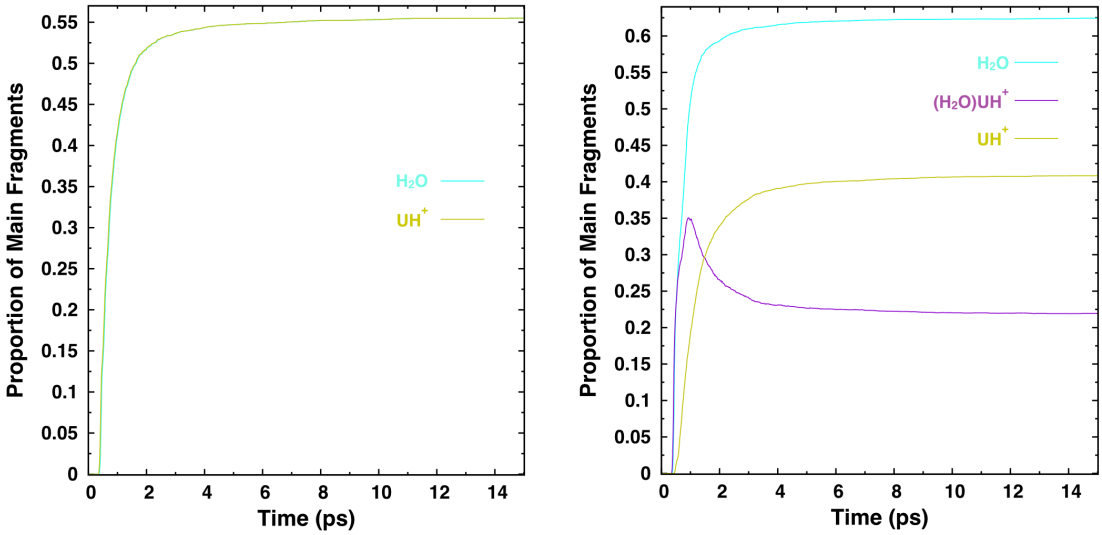


Figure 4.6: Time-dependent proportions of the main fragments obtained from the dissociation of the lowest-energy isomers of $(H_2O)_1UH^+$ (left) and $(H_2O)_2UH^+$ (right).

4.3.4 Proportion of Neutral Uracil Loss and Total Fragmentation Cross Sections for Small Clusters

In order to get more insights in the fragmentation, molecular dynamics trajectories were analysed in terms of proportion of neutral uracil loss (P_{NUL}) defined in section 4.3.2.1 and total fragmentation cross sections (σ_{frag}) defined in section 4.2.3. These two properties are also accessible from experiments. Another property extracted from the MD simulations, but not accessible from experiment, is the proportion of protonated uracil (P_{PU}) which is equal to the ratio of the number of simulations leading to a protonated uracil molecule over the number of simulations leading to a fragment containing the uracil and the excess proton. In order to correlate the outcome of the collision and the structure of the aggregate undergoing the collision, all considered low-energy isomers are characterized by there relative energies ($E_{rel.}$) and the location of the excess proton (LEP). For the latter, three distinct configurations were considered: The excess proton is bounded to the uracil molecule (noted U-H); The excess proton is

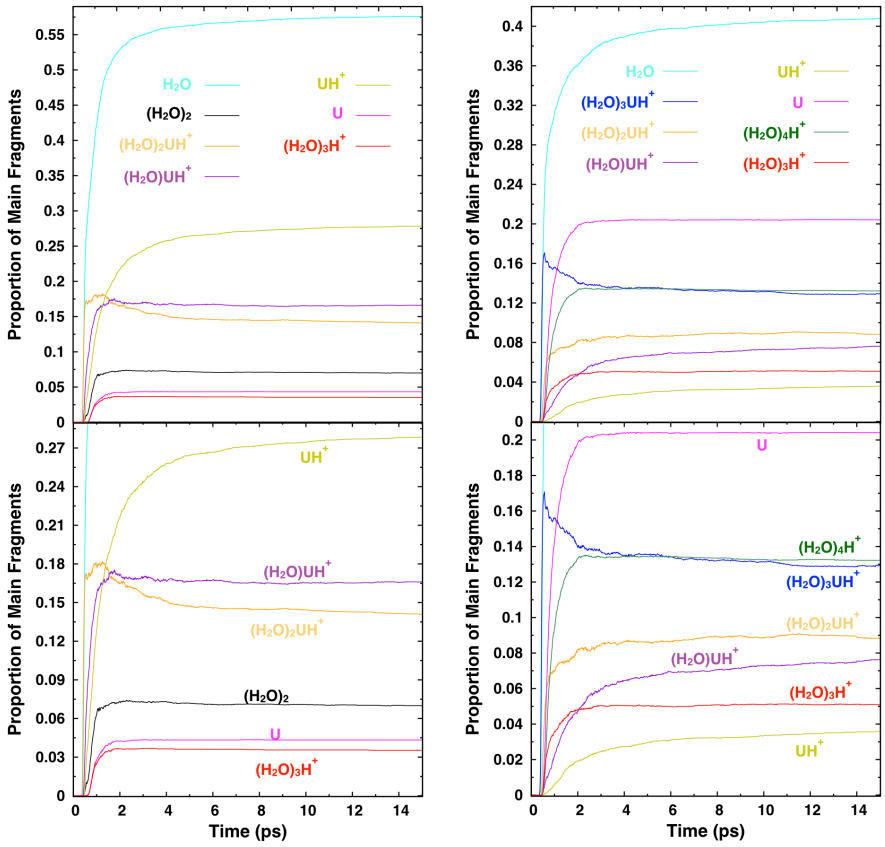


Figure 4.7: Time-dependent proportions of the main fragments obtained from the dissociation of the lowest-energy isomers of $(H_2O)_3UH^+$ (left) and $(H_2O)_4UH^+$ (right). Bottom panels correspond to a zoom over the lower proportions.

bounded to a water molecule that is adjacent to an oxygen atom of the uracil molecule (noted W-H-U); The excess proton is bounded to a water molecule that is separated by at least one other water molecule from the uracil molecule (noted W-H). All these data are gathered in Table 4.3 and will first discuss the behavior of the small species $(H_2O)_{1-7}UH^+$.

Various information can be inferred from these properties. Firstly, one observes a general increase of σ_{frag} as a function of cluster size with values ranging from 25.9 Å^2 for isomer 1b to 60.2 Å^2 for isomer 12a. Interestingly, only slight variations of σ_{frag} are observed for different isomers of the same aggregate. In contrast, P_{NUL} is much more sensitive to the nature of the considered isomers, in particular when these isomers display different LEP values. For instance, P_{NUL} is 46.6 % for 5a (W-H) while it is only 0.1 % for 5d (U-H). More interestingly, there seems to exist a strong correlation between P_{NUL} and LEP. Indeed, P_{NUL} values

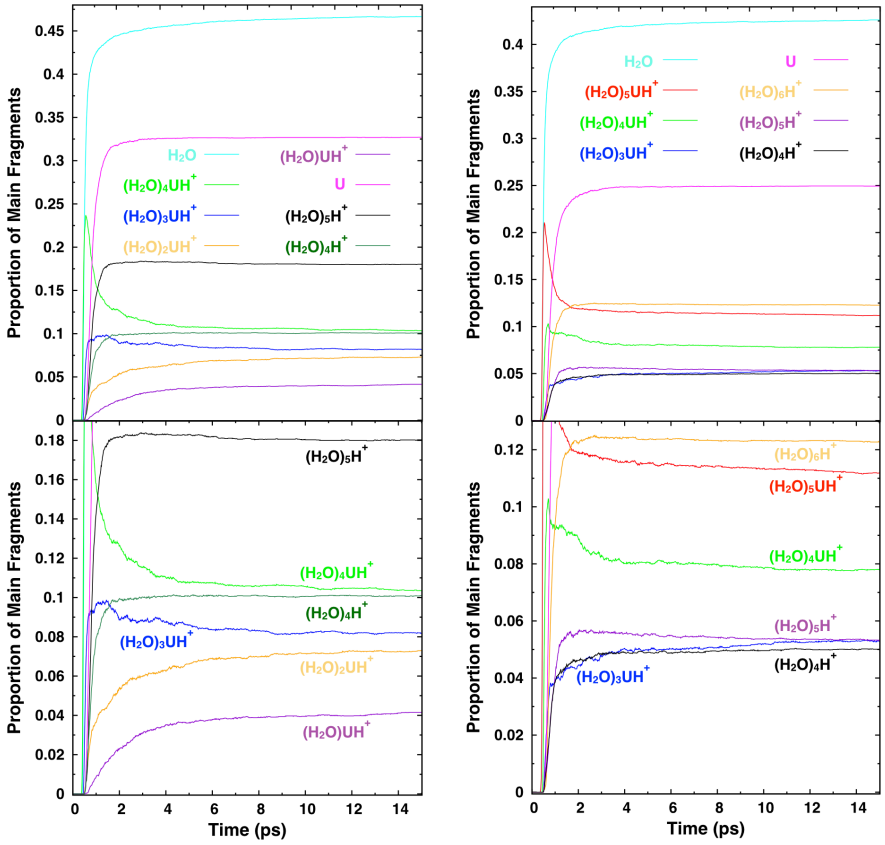


Figure 4.8: Time-dependent proportions of the main fragments obtained from the dissociation of the lowest-energy isomers of $(H_2O)_5UH^+$ (left) and $(H_2O)_6UH^+$ (right). Bottom panels correspond to a zoom over the lower proportions.

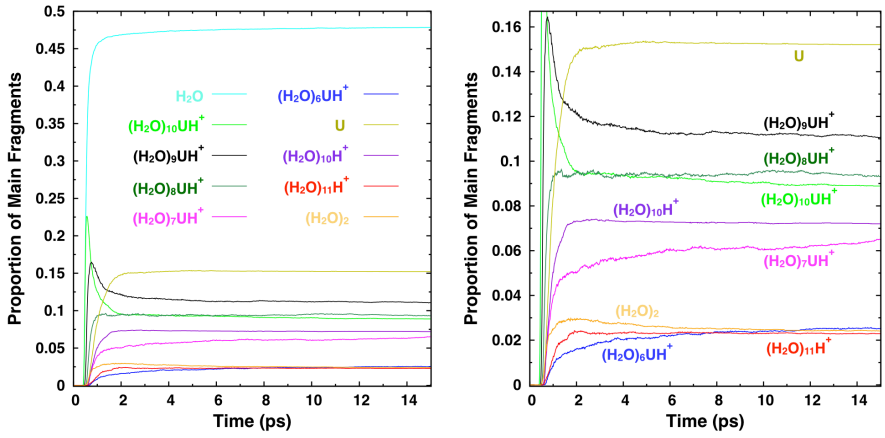


Figure 4.9: Time-dependent proportions of the main fragments obtained from the dissociation of the lowest-energy isomer of $(H_2O)_{11}UH^+$. Right panel corresponds to a zoom over the lower proportions.

Table 4.3: Relative energy $E_{rel.}$ (in kcal.mol^{-1}) at the MP2/Def2TZVP level, LEP, P_{PU} (in %), P_{NUL} (in %), σ_{frag} (in Å^2) of the considered low-energy isomers of $(H_2O)_{1-7,11,12}UH^+$ clusters. Isomers which P_{NUL} fit best to the experimental value are indicated in bold. $P_{NUL_{exp}}$ and $\sigma_{frag_{exp}}$ are the experimental values for P_{NUL} and σ_{frag}, respectively. For $(H_2O)_{12}UH^+$, experimental values were obtained for collision with Ne, whereas all other theoretical and experimental data are for collision with Ar.

Isomers	$E_{rel.}$	LEP	P_{PU}	P_{NUL}	$P_{NUL_{exp}}$	σ_{frag}	$\sigma_{frag_{exp}}$
1a	0.0	U-H	100	0.2	0.9	28.9	12.3
1b	0.7	U-H	100	0.1		25.9	
2a	0.0	U-H	100	0.0	0.4	36.3	22.8
2b	0.2	U-H	100	0.1		34.9	
3a	0.0	U-H	100	5.7	1.7	36.3	31.2
3b	0.3	U-H	100	0.0		41.9	
4a	0.0	W-H-U	98.0	29.4	2.8	40.1	43.4
4b	0.9	U-H	99.7	2.6		45.2	
5a	0.0	W-H	78.5	46.6	7.5	38.2	48.0
5b	0.3	W-H-U	89.0	28.5		38.7	
5c	2.0	W-H-U	87.8	27.1		44.6	
5d	2.4	U-H	100	0.1		47.5	
6a	0.0	W-H	44.1	39.3	18.0	45.8	54.3
6b	0.2	W-H	43.5	33.8		58.6	
6c	0.3	W-H	46.4	36.6		46.1	
6d	0.9	W-H	64.6	34.7		42.6	
6e	2.5	W-H	45.9	34.9		50.5	
6f	2.7	W-H-U	76.2	18.5		55.0	
7a	0.0	W-H	28.2	31.3	25.0	53.4	59.7
7b	0.3	W-H-U	52.4	21.4		51.7	
7c	0.3	W-H	41.3	31.1		49.5	
7d	0.8	W-H-U	40.9	23.0		54.0	
11a	0.0	W-H	4.6	28.3	11.8	52.9	63.8
11b	1.4	W-H	3.2	28.5		54.7	
11c	1.5	W-H	4.2	22.8		55.2	
11d	1.9	W-H	6.8	15.6		56.5	
11e	1.9	W-H	5.4	22.7		52.6	
11f	2.3	W-H	7.9	24.3		52.0	
12a	0.0	W-H	6.7	7.6	12.2	60.2	77.0
12b	0.6	W-H	34.0	22.4		52.2	
12c	0.7	W-H	48.7	10.8		55.4	
12d	1.3	W-H-U	5.4	9.7		54.3	
12e	1.8	W-H-U	67.5	6.0		54.2	
12f	2.4	W-H-U	55.0	17.1		54.1	

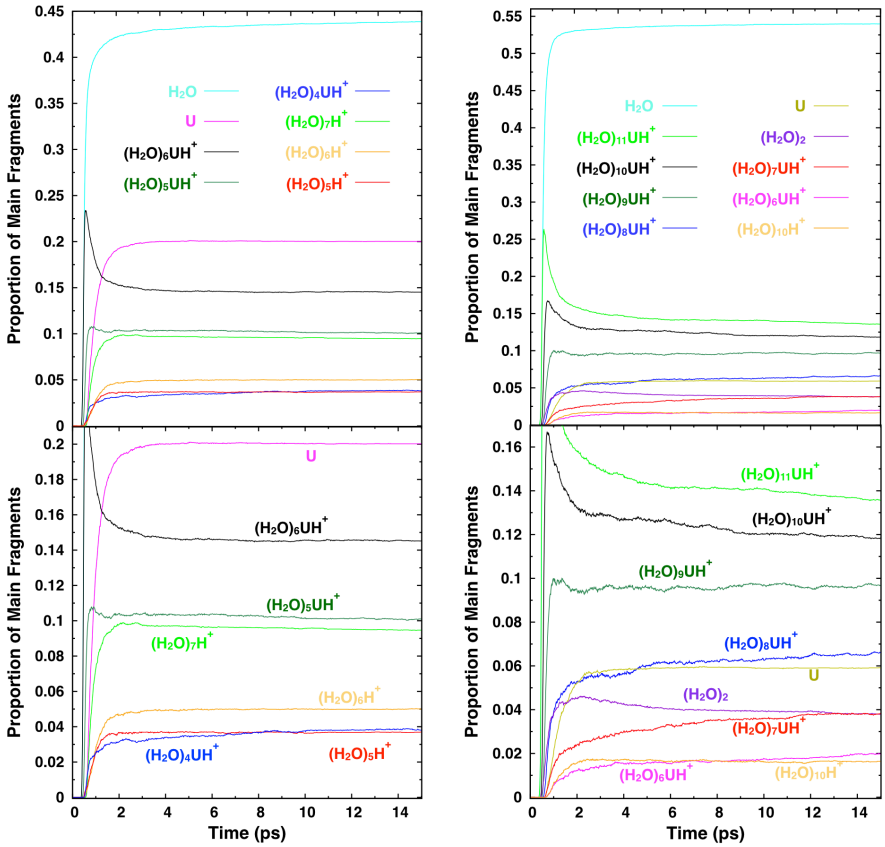


Figure 4.10: Time-dependent proportions of the main fragments obtained from the dissociation of the lowest-energy isomers of $(H_2O)_7UH^+$ (left) and $(H_2O)_{12}UH^+$ (right). Bottom panels correspond to a zoom over the lower proportions.

below 1.0 % are characterized by an excess proton initially bounded to uracil (U-H type). This suggests that when uracil is protonated, probability for deprotonation after collision is very low and thus P_{NUL} is close to 0%. P_{NUL} values between 9.7 and 29.4 % are obtained from W-H-U configurations while larger P_{NUL} values, above 31.1 %, arise from W-H configurations in clusters $(H_2O)_{5-7}UH^+$. This demonstrates that, from the excess proton point of view, the outcome of the collision is highly sensitive to the nature of the isomer undergoing the collision as hypothesised by I. Braud *et al.* [104] This important finding can be of help to determine which isomer, or set of isomers, is likely to undergo collision by comparing experimental and theoretical P_{NUL} as this is not necessarily the lowest-energy isomer as discussed below.

For $(H_2O)_{1-2}UH^+$, the theoretical and experimental P_{NUL} values, close to zero, are in good agreement regardless of the considered isomer. For $(H_2O)_3UH^+$, the experimental P_{NUL}

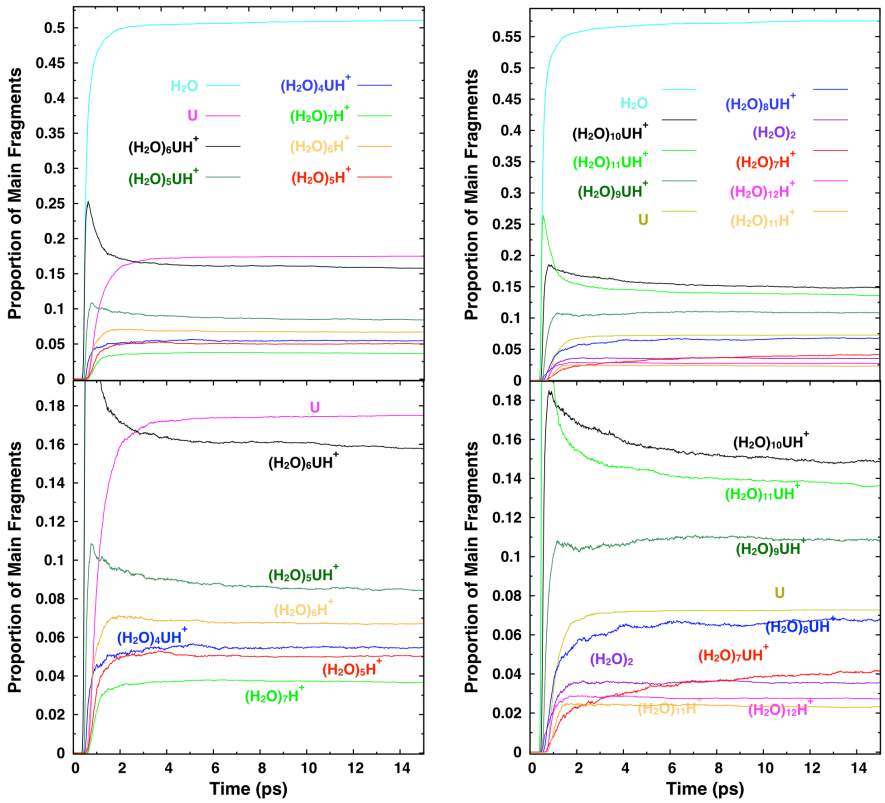


Figure 4.11: Time-dependent proportions of the main fragments obtained from the dissociation of the the third lowest-energy isomer of $(H_2O)_7UH^+$ (left) and the third lowest-energy isomer $(H_2O)_{12}UH^+$ (right). Bottom panels correspond to a zoom over the lower proportions.

is 1.7 % which is well reproduced by both isomers 3a and 3b although 3b is the one closer to the experimental value, 0.0 % against 5.7 % for 3a. This was expected as they are very close in energy, only 0.3 kcal.mol^{-1} difference, and in structure, as displayed in Figure 4.12, both being of U-H type structure. Consequently, in the experiment, each one of them could be at the origin of the experimental signal. $(H_2O)_4UH^+$ behaves differently. The two low-energy isomers, 4a and 4b, display very different P_{NUL} values, 29.4 and 2.6 %, respectively. The experimental value is 2.8 % which suggests that 4b, although slightly higher in energy by 0.9 kcal.mol^{-1}, is the isomer prevailing during the collision process. The difference in behavior can be explained by the U-H configuration of 4b, in which the excess proton is bounded to the uracil, whereas in 4a, it is bounded to a water molecule adjacent to uracil (see Figure 4.13). The case of $(H_2O)_5UH^+$ is more complex as this is the first species displaying the three types of LEP configuration among its four lowest-energy isomers as can be seen on Figure 4.13. This implies

very different P_{NUL} values: 46.6 % for 5a, 28.5 and 27.1 % for 5b and 5c, respectively, while it is only 0.1 % for 5d. The experimental P_{NUL} value for $(H_2O)_5UH^+$ is still relatively low, 7.5 %, which suggests that a U-H type structure prevails during the collision process. Although 5d is 2.4 kcal.mol^{-1} higher in energy than 5a, this isomer is thus expected to undergo the collision.

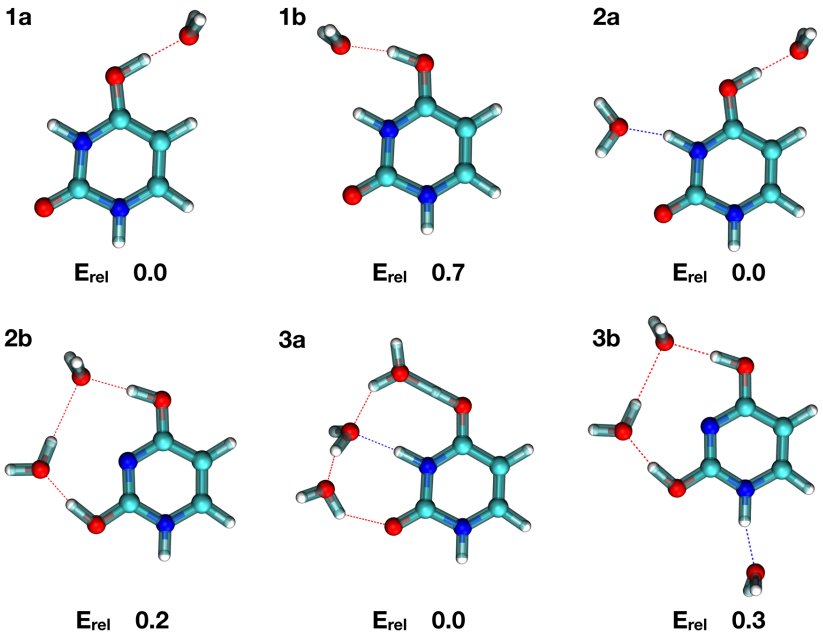


Figure 4.12: Selected low-energy configurations of $(H_2O)_{1-3}UH^+$. Relative energies at the MP2/Def2TZVP level are in kcal.mol^{-1}.

$(H_2O)_6UH^+$ and $(H_2O)_7UH^+$ are the first two aggregates for which no low-energy isomer belongs to the U-H type structure. As a consequence, in contrast to smaller species, the theoretical P_{NUL} values are all higher than 15 %. This is in line with the experimental values which display a net increase at $n = 6$. Isomers 6a, 6b, 6c, 6d, and 6e (see Figure 4.14) are all W-H type structures which leads to P_{NUL} values almost twice higher than the experimental one. Consequently, as for $(H_2O)_5UH^+$, one can assume that the isomer of $(H_2O)_6UH^+$ undergoing the collision is more likely to be a W-H-U type structure although it is higher in relative energy. Isomer 6f can be such a candidate as it displays of P_{NUL} value of 18.5% which is in agreement with the experimental value, 18.0%. Due to its increasing size, $(H_2O)_6UH^+$ displays W-H configurations with the excess proton at various distances from the recombining oxygen. Indeed, in 6a, 6c and 6d this distance is 1.774, 1.745 and 1.804 Å, while in 6b, 6e and 6f, it

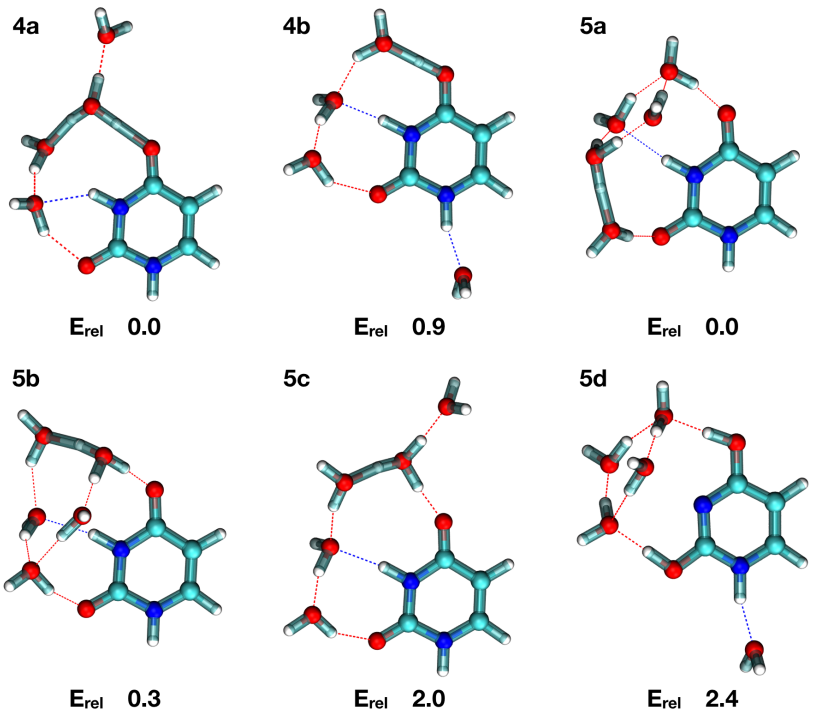


Figure 4.13: Selected low-energy configurations of $(H_2O)_{4-5}UH^+$. Relative energies at the MP2/Def2TZVP level are in kcal.mol^{-1}.

is shorter: 1.660, 1.614, and 1.494 Å, respectively. However, no net correlation is observed between this distance and the value of P_{NUL}: 39.3, 33.8, 36.6, 34.7, 34.9 and 18.5% for 6a, 6b, 6c, 6d, 6e and 6f, respectively. In particular, the behavior of 6e is striking. It has almost the same relative energy as 6f and they are structurally similar (see Figure 4.14) but display different P_{NUL} values. This suggests that, for n larger than 5, the ability of the water molecule network to stabilise the excess proton, *i.e* to promote or prevent its diffusion toward the uracil molecule, starts to be competitive with the configuration type of the isomer. In 6e, the excess proton is in a configuration close to the Zundel ion which may explain its high P_{NUL} value as compared to 6f. For $(H_2O)_7UH^+$, a W-U-H type configuration is also expected to fit best the experimental result. And indeed 7d, a W-H-U type structure, which is only 0.8 kcal.mol^{-1} above the lowest-energy isomer (see Figure 4.15), has a P_{NUL} value of 22.9 % as compared to 25.0 % experimentally. Isomers 7a and 7c have a W-H configuration and their P_{NUL} values (31.3 and 31.1 %, respectively) are higher than the ones of 7b and 7d which have a W-H-U configuration.

Finally, it is worth noting that even when the excess proton is initially bounded to a water molecule, *i.e.* when a W-H type structure is considered, the maximum P_{NUL} that has been obtained is only 46.6 %. This demonstrates that for small aggregates such as $(H_2O)_{5-7}UH^+$ ($(H_2O)_{1-4}UH^+$ do not display low-energy W-H type structures), dissociation mainly lead to protonated uracil containing fragments. This is in line with the experimental results. Analysis of P_{PU} values also show that uracil is protonated in a significant amount of these protonated uracil containing fragments. P_{PU} has a clear tendency to decrease with cluster size, but can be quite high even for W-H type structures, for instance 5a, 6d and 7c in Table 4.3. This demonstrates that upon collision, the excess proton is likely to transfer to uracil on a rather short time scale.

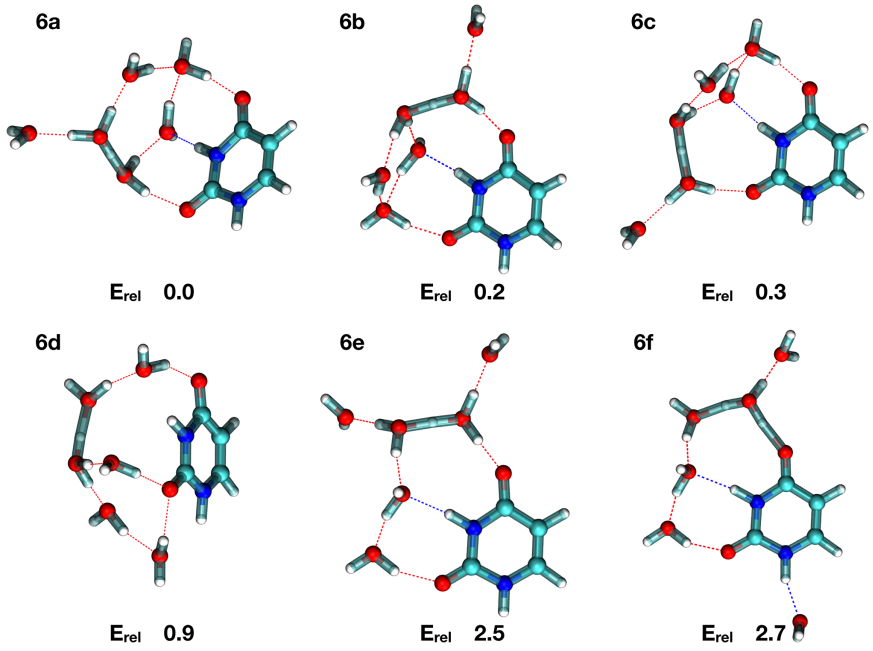


Figure 4.14: Selected low-energy configurations of $(H_2O)_6UH^+$. Relative energies at the MP2/Def2TZVP level are in kcal.mol^{-1}.

The clusters discussed above are characterized by complex potential energy surfaces characterized by several low-energy isomers, with relative energies that can be lower than 1 kcal.mol^{-1}, and which get more complex as the number of water molecules increases. Consequently, the exact energetic ordering between the low-energy isomers can not be precisely known as this is below chemical accuracy and thus can not claim here to have found the lowest-energy structure of each aggregate, or the isomer undergoing the collision. Nevertheless, what I show is that

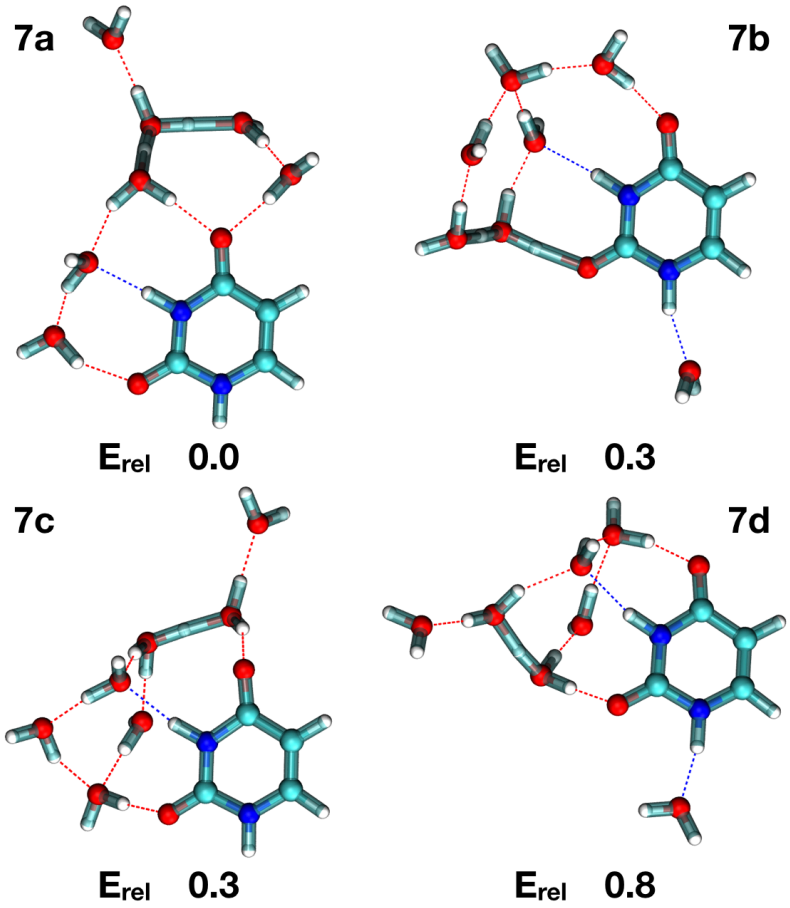


Figure 4.15: Selected low-energy configurations of $(H_2O)_7UH^+$. Relative energies at the MP2/Def2TZVP level are in kcal.mol^{-1}.

P_{NUL} is mainly determined by the initial position of the proton in the isomer undergoing the collision. Consequently, for the collision energy and the range of cluster size I have considered, the structure of the aggregate undergoing the collision plays a key role in determining the dissociation process and collision outcomes much more than energetics. This is consistent with the analysis of the time-dependent proportion of fragments which suggests a direct dissociation mechanism. This is further highlighted on Figure 4.16, which presents the experimental P_{NUL} for collision with Ar and Ne, respectively as a function of n and the corresponding theoretical values obtained from the lowest-energy isomers as well as from the isomers for which P_{NUL} matches best to the experimental data. As can be seen, a very good agreement can be obtained with the experimental data by considering a specific set of isomers. Interestingly, if a similar plot is drawn for σ_{frag} considering the same isomers (see Figure 4.17), a good agreement with

the experimental data and much better than σ_{geo} (calculated from formula 3.3)is also obtained with the two sets of isomers which confirms the weaker dependence upon isomer of σ_{frag}.

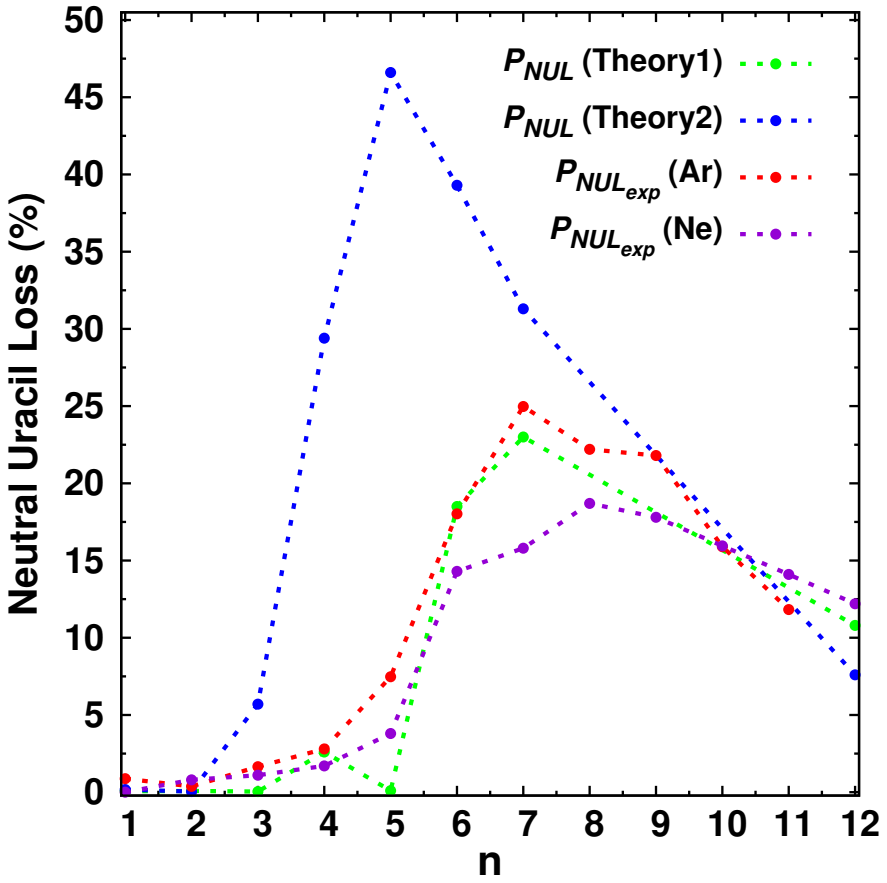


Figure 4.16: Theoretical (green and blue lines) and experimental (red line) P_{NUL} values for the $(H_2O)_{1-7,11,12}UH^+$ clusters. Theory 1 (green line) is obtained from the isomers which P_{NUL} matches best to the experimental data while Theory 2 (blue line) is obtained from lowest-energy isomers.

4.3.5 Behaviour at Larger Sizes, the Cases of $(H_2O)_{11,12}UH^+$

In the experiments conducted by I. Braud *et al.*,[104] P_{NUL} starts to decrease at n=8. This decrease is not consistent with the above argument of a direct dissociation mechanism and larger species more likely characterized by W-H and W-H-U type structures. This apparent discrepancy motivated us to extend the present study to a larger cluster, namely $(H_2O)_{11,12}UH^+$. For $(H_2O)_{12}UH^+$, the only available experimental data is for collisions with Ne instead of Ar, al-

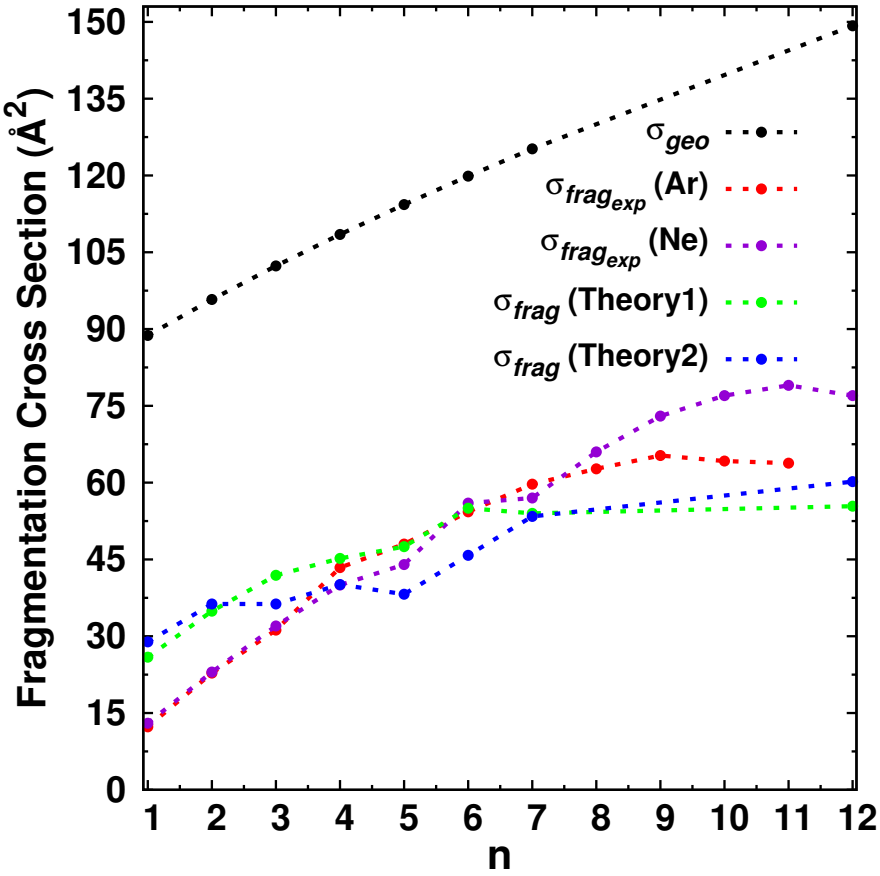


Figure 4.17: Theoretical (green and blue lines) and experimental (red line) σ_{frag} values for the $(H_2O)_{1-7,11,12}UH^+$ clusters. Theory 1 (green line) is obtained from the isomers which P_{NUL} matches best to the experimental data while Theory 2 (blue line) is obtained from lowest-energy isomers.

though for the same center of mass collision energy. As shown in Figure 4.16, experimental P_{NUL} values for Ne or Ar, although not equal, display similar trend. In the following, I thus discuss the experimental data of $(H_2O)_{12}UH^+$ colliding with Ne. For cluster $(H_2O)_{1-7,11}UH^+$, keep discussing the experimental data from colliding with Ar.

The behaviour for $(H_2O)_{11}UH^+$ and $(H_2O)_{12}UH^+$ is rather different when looking at detailed properties. Indeed, for $(H_2O)_{11}UH^+$, P_{NUL} values for isomers 11a, 11b, 11c, 11e, and 11f are very similar as they range from 22.7 to 29.8 %. For 11d, P_{NUL} equal 15.6 % which fits best to the experiment. These P_{NUL} values are lower than those of $(H_2O)_6UH^+$, as observed experimentally, and in the same range as $(H_2O)_7UH^+$. All isomers display a W-H type configuration as seen in Figure 4.18. P_{PU} is very small for all $(H_2O)_{11}UH^+$ isomers which shows

that on the time scale of the simulations, protonation of uracil hardly occurs.

For $(H_2O)_{12}UH^+$, 12c isomer, which has a W-H type configuration (see Figure 4.19), has a P_{NUL} value which fits best to the experiment, 10.8 % against 12.2 %, while isomer 12a, also a W-H type configuration (see Figure 4.19), has a P_{NUL} value equal to 7.6 %. Overall, P_{NUL} values calculated for $(H_2O)_{12}UH^+$ isomers are lower than those of $(H_2O)_6UH^+$ and $(H_2O)_7UH^+$, which is in line with the experiment. The main difference with the $(H_2O)_{1-7}UH^+$ aggregates is that no clear relation exist between the P_{NUL} value and the initial localisation of the excess proton. Indeed, 12a, 12b and 12c are all W-H type configurations but with P_{NUL} values ranging from 7.6 to 22.4 %. The same is observed for 12d, 12e and 12f although they are all W-H-U type configurations. Similarly, no difference in behaviour is obtained between W-H and W-H-U type configurations. This can be explained by assuming that the dissociation mechanism in $(H_2O)_{12}UH^+$ involves some amount of structural rearrangement that softens the impact of the isomer undergoing the collision. Indeed, as $(H_2O)_{12}UH^+$ has more degrees of freedom, it can more easily accommodate the kinetic energy transferred by the colliding atom prior to dissociation which thus takes place on a longer time scale. This excess of internal energy thus fosters structural rearrangements, in particular proton transfers toward the uracil, explaining the smaller P_{NUL} value for $(H_2O)_{12}UH^+$. This is in full agreement with the conclusions obtained in section 4.3.3 from Figures 4.9,4.10 and 4.11. To further support this conclusion, I conducted 200 MD simulations in the micro-canonical ensemble in which the whole kinetic energy of Ar was randomly distributed in all the vibrational modes of isomer 12c by drawing initial velocities in a 1185 K Boltzmann distribution. Among them, 166 simulations display dissociation with one or two water molecules dissociating from the main cluster. No neutral uracil loss is observed. To conclude, although the present simulations are too short to assert that $(H_2O)_{12}UH^+$ undergoes a statistical dissociation mechanism, they clearly show that a direct mechanism is not sufficient to account for the theoretical and experimental results. Consequently, structural rearrangements are very likely to occur prior to dissociation and the experimental results for P_{NUL} and σ_{frag} values can not result from a single $(H_2O)_{12}UH^+$ isomer. In contrast, similarities in both P_{NUL} and P_{PU} values for all considered $(H_2O)_{11}UH^+$ isomers, as well as P_{NUL} values close to $(H_2O)_7UH^+$ ones, do not evidence structural rearrangements in this species although they could be present.

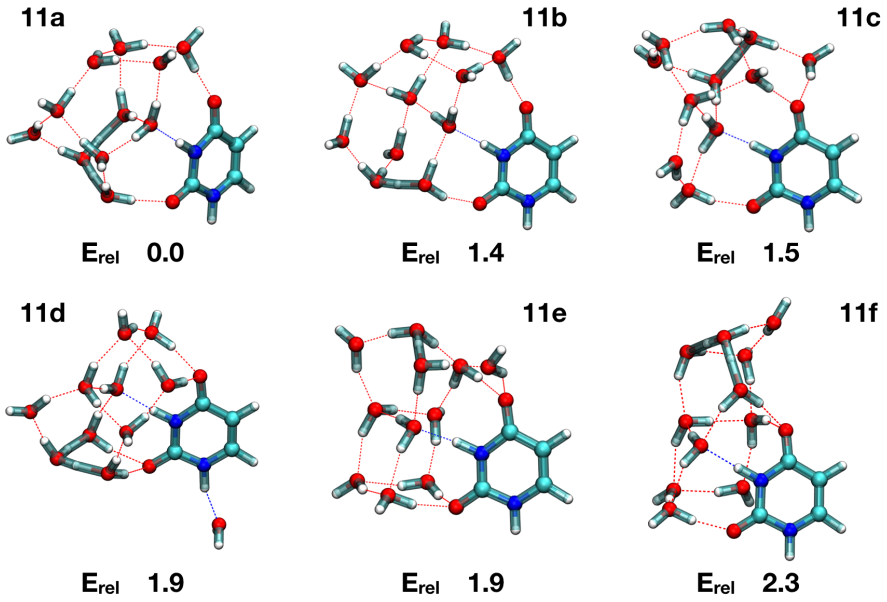


Figure 4.18: Selected low-energy configurations of $(H_2O)_{11}UH^+$. Relative energies at the MP2/Def2TZVP level are in kcal.mol^{-1}.

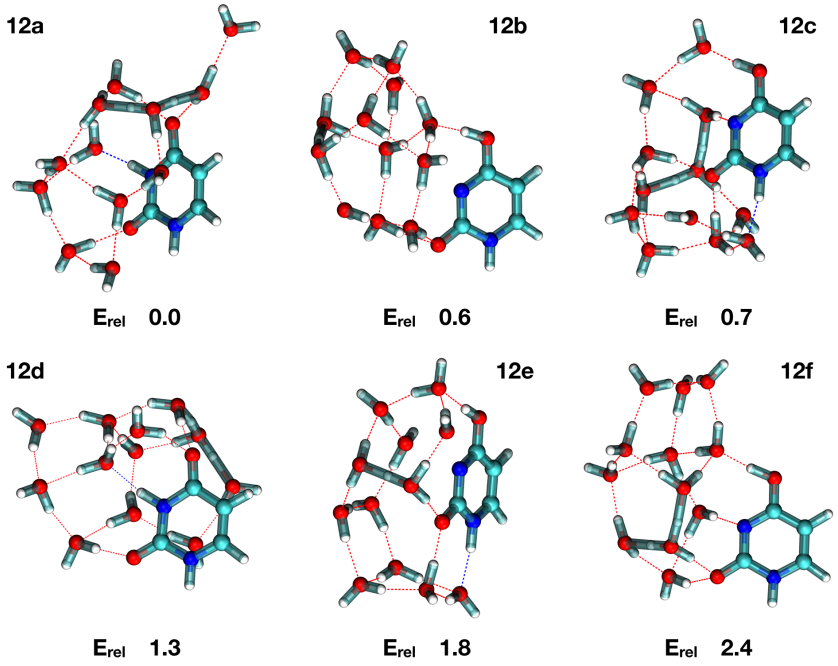


Figure 4.19: Selected low-energy configurations of $(H_2O)_{12}UH^+$. Relative energies at the MP2/Def2TZVP level are in kcal.mol^{-1}.

4.3.6 Mass Spectra of Fragments with Excess Proton

In this section, in order to analyse collision products in more details, the branching ratios of the different fragments containing the excess proton were extracted from the collision simulations of clusters $(H_2O)_{1-7,11,12}UH^+$ and compared with the experimental ones shaped as mass spectra.[104] For each cluster size, only simulations corresponding to the isomer which P_{NUL} value fits best to the experiment were considered (1a, 2b, 3b, 4b, 5d, 6f, 7d, 11d, 12c). The results are presented in Figures 4.20, 4.21 and 4.22. For cluster $(H_2O)_{12}UH^+$, there is no experimental data for collision with argon. For $(H_2O)_{12}UH^+$, experimental results were obtained for collision with neon. From the experimental results for argon and neon for 1a, 2b, 3b, 4b, 5d, 6f, 7d, and 11d, it shows the branch ratios for collision with argon and neon are close and have the same trend. So it should be reasonable to compare the simulated branch ratios of 12c with the ones of experimental data from the collision of $(H_2O)_{12}UH^+$ with argon.

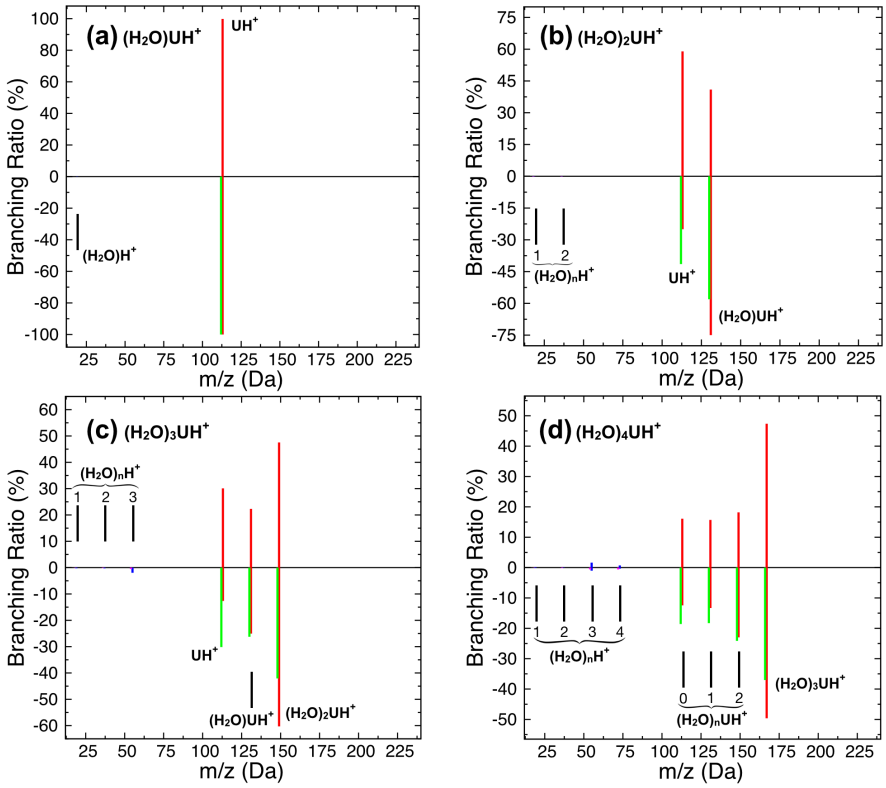


Figure 4.20: Simulated mass spectra (positive area) of the charged fragments after 15 ps simulation time (fragments $(H_2O)_nH^+$ in red and $(H_2O)_nUH^+$ in blue for argon; $(H_2O)_nH^+$ in pink and $(H_2O)_nUH^+$ in green for neon) from isomers (a) 1a, (b) 2b, (c) 3b, (d) 4b. The counterparts in experiment are plotted (negative area).

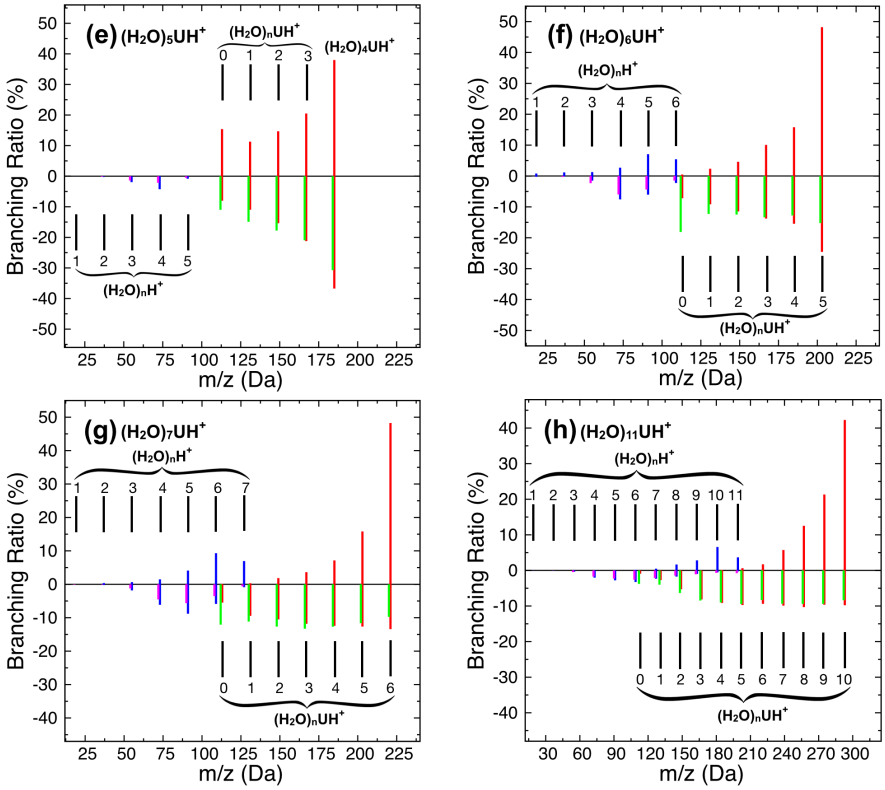


Figure 4.21: Simulated mass spectra (positive area) of the charged fragments after 15 ps simulation time (fragments $(H_2O)_nH^+$ in red and $(H_2O)_nUH^+$ in blue) from isomers (e) 5d, (f) 6f, (g) 7d, and (h) 11d. The counterparts in experiment are plotted (negative area).

Overall, the experimental and theoretical spectra present the same general trends: (i) The mass spectra present a broad distribution of sizes, without prominence of a particular peak; (ii) All the spectra are dominated by the heaviest protonated uracil containing fragment (loss of a single water molecule) with the exception of the simulated mass spectrum for 2b; (iii) Fragments containing protonated uracil prevail over pure protonated water fragments, as already observed from the P_{NUL} values provided in Table 4.3; (iv) Pure protonated water fragments only appear for the largest cluster sizes. Indeed, although very minor contributions are observed in both the simulated and experimental spectra for parent clusters with n=3-5, significant contributions of these species only appear when the parent cluster contains at least 6 water molecules.

A more detailed discussion of the simulated and experimental mass spectra will be made as follows. For $(H_2O)_2UH^+$, fragments $(H_2O)UH^+$ and UH^+ are observed in both experiment and theory although their relative ratio is different. For $(H_2O)_3UH^+$, the simulated and exper-

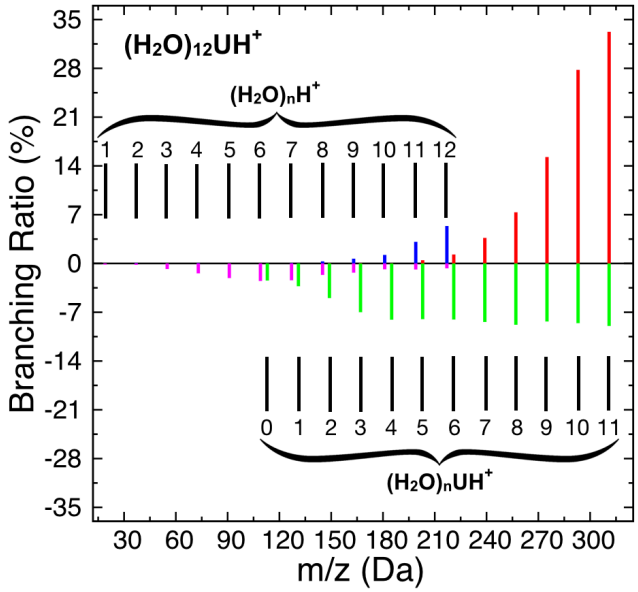


Figure 4.22: Simulated mass spectra (positive area) of the charged fragments after 15 ps simulation time (fragments $(H_2O)_nH^+$ in red and $(H_2O)_nUH^+$ in blue) from isomers 12c. The counterparts in experiment obtained for collision with neon are plotted in negative area $(H_2O)_nH^+$ in pink and $(H_2O)_nUH^+$ in green).

imental spectra agree quite well with a dominant peak for $(H_2O)_2UH^+$. For $(H_2O)_4UH^+$, in both experimental and theoretical spectra, the peak intensity of the fragments containing protonated uracil increases with the number of water molecules. For $(H_2O)_5UH^+$ and $(H_2O)_6UH^+$, this is also the case except for the UH+ fragment which is overestimated when compared to the experimental result in Figure 4.21 (e). For $(H_2O)_{7,11,12}UH^+$, the intensities for the heaviest fragments are overestimated.

From Figures 4.20, 4.21 and 4.22, it is clear that the smaller the cluster (except Figure 4.20 (b)) is, the better the agreement between the simulated and experimental branching ratios is. This trend indicates that for small clusters, *i.e.* for $n = 1-6$, short simulation time is enough to capture the full dissociation pattern, in other words, the dissociation mechanism is direct with no noticeable contribution of long term evolution. However, for larger clusters, starting at n=7, owing to the larger number of degrees of freedom, short simulation time does not capture the full dissociation pattern, *i.e.* long term statistical dissociation is more likely to play a noticeable role. This is fully in line with the conclusions obtained in section 4.3.5 for $(H_2O)_{12}UH^+$ and refine the interpretation given in section 4.3.4 for $(H_2O)_7UH^+$. This also shows that although the data presented in section 4.3.5 for $(H_2O)_{11}UH^+$ do no evidence the

Table 4.4: Energies of different $(H_2O)_6UH^+$ fragments selected from the dissociation of 7d at SCC-DFTB level, and the lowest energies $(H_2O)_5UH^+$ and (H_2O) at SCC-DFTB level. The relative energy $\Delta E = E_{(H_2O)_6UH^+}$ -($E_{(H_2O)_5UH^+}$ +E_{H_2O}). All energies here are given in eV.

$E_{(H_2O)_6UH^+}$	$E_{(H_2O)_5UH^+}$	E_{H_2O}	ΔE
-44.310	-40.312	-4.057	1.605
-44.322	-40.312	-4.057	1.279
-44.307	-40.312	-4.057	1.687
-44.344	-40.312	-4.057	0.680
-44.373	-40.312	-4.057	-0.109

contribution of structural re-arrangements on the short time scale, they are very likely to occur as in $(H_2O)_{12}UH^+$.

One has to keep in mind that modeling the complete duration of the experiment (up to μs) is out of reach with MD/SCC-DFTB simulations. In this work, the simulation time was 15 ps, for all cluster sizes. Large fragments such as $(H_2O)_{6-12}UH^+$ may lose more water molecules if long enough simulation time were available, as suggested from the time dependent evolution of selected trajectories in section 4.3.3. To certify this, the total energy of $(H_2O)_6UH^+$ fragments at SCC-DFTB level is calculated originating from the dissociation of $(H_2O)_7UH^+$ (7d) from all the 1421 trajectories producing fragment $(H_2O)_6UH^+$ over the total 600 $\times$ 15 trajectories. Then the energies of the lowest-energy isomer of $(H_2O)_5H^+$ and H_2O at SCC-DFTB level are subtracted. The deduced relative energies ΔE are reported in Table 4.4 for five cases. When ΔE is greater than zero, it is possible for the $(H_2O)_6UH^+$ fragment to lose a water molecule. The percentage of ΔE being positive in all the trajectories leading to fragment $(H_2O)_6UH^+$ is 53.0 %, which indicates that many $(H_2O)_6UH^+$ fragments have still the potential to lose one more water molecule after the end of the simulation.

4.3.7 Conclusions about CID of $(H_2O)_nUH^+$

Collision-induced dissociation of protonated uracil water clusters $(H_2O)_{1-7,11,12}UH^+$ at constant center of mass collision energy has been investigated by molecular dynamics simulations using the SCC-DFTB method. The very good agreement between the simulated and measured P_{NUL} and σ_{frag} as well as branching ratios indicate that the essence of the dissociation induced by collisions is well captured by the simulations.

The P_{NUL} values from the different isomers of the $(H_2O)_{1-7}UH^+$ cluster show that the localization of the excess proton after dissociation is strongly determined by the initial configuration of the isomer undergoing the collision. This suggests that $(H_2O)_{1-7}UH^+$ aggregates primarily engage a direct dissociation path after collision that takes place on a very short time scale, *i.e.* lower than 15 ps. More strikingly, in most cases, the proposed lowest-energy isomer does not lead to the best fit to the experiment. However, the relative energy between the lowest-energy isomers and the isomers best fitting to the experiment is less than 1.0 kcal.mol^{-1} for $(H_2O)_{1-4,7}UH^+$ clusters and less than 2.7 kcal.mol^{-1} for $(H_2O)_{5,6}UH^+$ clusters. This is in line with the strong sensitivity of the collision outcome with the nature of the isomer undergoing the collision. This even suggests that the LEP can help in determining the main characteristic of the isomer involved in the collision. For $(H_2O)_{11,12}UH^+$, these conclusions do not apply any more which shows that significant structural rearrangements occur after collision. This is confirmed by the time-dependent proportion of fragments which continue to vary even at 15 ps for $(H_2O)_{11,12}UH^+$ whereas it is almost flat for $(H_2O)_{1-7}UH^+$. Analysis of the fragment branching ratios helps in clarifying these points. Indeed, for the smallest clusters, $(H_2O)_{1-5}UH^+$, the short simulation time well reproduces the corresponding experimental results which is in line with a direct mechanism. In contrast, for $(H_2O)_{6-7}UH^+$, although P_{NUL} is well reproduced by the simulations, the experimental and theoretical branching ratios differ which show that more time is needed to properly describe the dissociation. For $(H_2O)_{11,12}UH^+$, neither theoretical nor experimental branching ratios and P_{NUL} are in agreement which is a strong indication that a significant contribution of structural rearrangements occur; This suggests that a contribution of a statistical mechanism is more likely to occur for larger species such as $(H_2O)_{11,12}UH^+$.

This work demonstrates that explicit molecular dynamics simulations performed at a quantum chemical level can provide a wealth of information about collision-induced mechanism in molecular clusters, in particular, hydrated molecular species. Such simulations thus represent a key tool to complement CID experiments and hope the present study will motivate similar computational studies on future CID experiments of hydrated molecular aggregates.

4.4 Dynamical Simulation of Collision-Induced Dissociation for Pyrene Dimer Cation

4.4.1 Introduction

PAH clusters have been investigated in several scientific fields. In combustion science, the role of PAH clusters in combustion processes is still under debate, in particular they might or not be the intermediate systems in the growth of soot particles. [125, 455, 456, 457, 458] In atmospheric and environmental science, PAHs are known as the pollutants, which is harmful to human health. For instance, the carcinogenic PAHs associated with particulate matter in air pollution has showed clear evidence of genotoxic effects, such as DNA adduct, chromosome aberrations. [459, 460] In new energy resources field, for the understanding of the properties of organic crystal or the design of new organic solar cell devices, PAH stacks are investigated as the prototypes.[461] In astrophysics, PAHs species are believed to be ubiquitous and abundant in the interstellar medium because of their compact and stable structure. [117] The PAH clusters are important contributors to the diffuse interstellar bands and UV-visible absorption bands. PAH clusters have been proposed to be the origin of a series of infrared emission bands, which are ubiquitous in the Universe. [114, 115] The broadening of these bands in regions protected from the star's UV flux suggests the following scenario: PAHs are trapped in clusters in UV-protected regions and photo-evaporated by star's UV photons in the so-called photodissociated region. [462, 463] For all these topics, it is necessary to make a better understanding of the fundamental properties of PAH clusters. The crucial quantities are the stability, molecular growth processes, dissociation energies and their evolution with PAH charge, species, cluster size.

The investigation of PAH clusters has been performed in experiment. Many studies focused on the investigation of structural properties of these clusters at the most stable geometrical configurations [132, 458, 464, 465, 466]. Their energetic properties such as ionisation potentials have been recorded [122] as well as their spectral properties [467, 468]. clusters may evaporate, breaking the PAH units themselves or leading to chemical reactivity between the different units, which shows the role of PAH clusters in the growth of PAHs themselves. If free flying PAHs are possibly from the evaporation of larger clusters, then this calls for more experimental data on their dissociation properties. The evolution of PAH clusters has been explored from experiments following the evaporation after absorption of UV photons, collision with low or high energetic particles or in a high-pressure environment. [122, 129, 132, 133, 134, 135] The

range of collision energies considered experimentally is quite large, ranging from eV to high energy collision at a few keV. Low energy collision experiments allow for the derivation of dissociation energies [135] whereas the oligomerization of PAHs within the cluster induced by high energy collisions [469] or photoabsorption [16] suggests the possible role of clusters in the interstellar PAHs growth process [131].

The quantitative data from experiments of PAH clusters are still rather limited, which motivates the modeling studies of them. In the calculation of PAH clusters, the size of the systems limits the use of *ab initio* wave function methods to the investigation of properties of the smallest clusters, namely dimers [470, 471], whereas larger clusters can be addressed either at the DFT level or with more semi-empirical schemes [233, 457, 472, 473]. Many of these studies, focused on structural properties, evidence a stacking growth process in agreement with experimental results. In addition, IR properties were also reported at the DFT level [474]. Most of the theoretical studies involve neutral clusters, mostly due to the fact that treating charge resonance process in ions is a challenging task for DFT based methods [475]. The singly charged PAH clusters are more stable than their neutral counterparts due to charge resonance stabilization.[121] Cationic PAH clusters are expected to be abundant in the photo-dissociation regions because the ionization energy of the PAH cluster is lower than that of the isolated PAH, which leads to the efficient formation of cationic PAH clusters. In addition, the ionized PAH clusters are easier to control, so it is more important to study them. It should be mentioned the recent studies computing ionisation potentials [122] as well as structural [454] and spectral (electronic [476] and vibrational [477]) properties of cations, performed with an original model combining DFTB [168, 218, 219, 445] with a configuration interaction scheme[478].

With respect to these studies, very few is known about the dynamical aspects of PAH clusters carrying internal energy. High energy collisions of PAH clusters with energetic ions have been simulated by M. Gatchel *e*t al. [479, 480] at the semi-empirical and DFTB levels. Recently experiments at lower collision energies were performed by S. Zamith *et al.* [65] (the principle of this experiment and the experimental setup were shown in sections 4.1.1 and 4.1.2), which were analysed by treating statistically the dissociation after collision energy deposition. Namely, the dissociation rate of pyrene clusters has been computed using phase space theory (PST)[135]. A fair agreement with experimental results was obtained concerning the collision energy dependence of the dissociation cross section. However, the employed model failed at reproducing in details the shape of the peaks in the time-of-flight (TOF) spectra. In this section, it is aimed at extending the description of such low energy collision processes (less than several

tens of eV) combining a dynamical simulations to describe the fast processes in addition to the statistical theory to address dissociation at longer timescales. With this approach, (i) good agreement between simulated and experimental mass spectra will be shown, thus validating the model, (ii) dissociation cross sections as a function of the collision energy is derived, (iii) the kinetic energy partition between dissociative and non-dissociative modes will be discussed and (iv) the energy transfer efficiency between intra and intermolecular modes will also be discussed.

4.4.2 Calculation of Energies

In the analysis, I will discuss the kinetic energy contributors, applying the following decomposition of the total kinetic energy E^k_{tot} of the dimer:

$$\begin{aligned}
E^k_{tot} &= E^k_{Ar} + E^k_{td} + E^k_{Py^1} + E^k_{Py^2} + E^k_{Re} \\
E^k_{tot} &= \frac{1}{2}\sum_{i=1}^{52} m_i(\vec{v}_i)^2 \\
E^k_{Ar} &= \frac{1}{2} m_{Ar}\vec{v}^2_{Ar} \\
E^k_{td} &= \frac{1}{2} m_{Py_2}\vec{v}^2_t(Py_2) \\
E^k_{Re} &= \frac{1}{2}\frac{m_{Py^1}m_{Py^2}}{m_{Py^1}+m_{Py^2}}(\vec{v}_t(Py^2) - \vec{v}_t(Py^1))^2 \\
E^k_{Py^n} &= \frac{1}{2}\sum_{i=1}^{26} m^n_i(\vec{v}^n_i - \vec{v}_t(Py^n))^2
\end{aligned} \tag{4.5}$$

In these equations and in the following, Py_2 refers to the pyrene dimer (possibly dissociated) whereas Py^1 and Py^2 refer to the first and second monomers, respectively. E^k_{tot} can be also calculated from the masses m^n_i and velocities $\vec{v}_i$ of its atoms. E^k_{Ar} refers to the kinetic energy of the argon (with mass m_{Ar} and velocity $\vec{v}_{Ar}$). E^k_{td} is the translation kinetic energy of the dimer (with mass m_{Py_2} and velocity $\vec{v}_t(Py_2)$). E^k_{Re} is the relative kinetic energy of the two pyrene monomers, computed from their masses of $m_{Py^1} = m_{Py^2}$ and monomer global translation velocities $\vec{v}_t(Py^{1,2})$. $E^k_{Py^n}$ is the rovibrational kinetic energy of the monomer n computed from the masses and velocities of its atoms (m^n_i and $\vec{v}^n_i$, respectively).

The intramolecular vibrational kinetic energy ($E^k_{intra^n}$) of monomer n obtained after removing the contributions associated to the monomer translation and rotation modes is calculated as

follows:

$$E^k_{intra^n} = \frac{1}{2}\sum_{i=1}^{26} m^n_i(\vec{v}^n_i - \vec{v}_t(Py^n) - \vec{v}^n_{ir})^2 \tag{4.6}$$

where $\vec{v}^n_{ir}$ is the velocity of atom i associated to the monomer global rotation. In addition, the dimer intermolecular kinetic energy (E^k_{inter}) is calculates as follows:

$$E^k_{inter} = E^k_{tot} - E^k_{Ar} - E^k_{td} - E^k_r - E^k_{intra^1} - E^k_{intra^2} \tag{4.7}$$

where E^k_r refers to the rotation kinetic energy of the dimer. $\vec{v}^n_{ir}$ and E^k_r are calculated using the following formulas.

$$\begin{aligned}
\vec{L}(Py^n) &= \sum_{i=1}^{26} m^n_i(\vec{r}^n_i - \vec{r}_{CM}(Py^n)) \times (\vec{v}^n_i - \vec{v}_t(Py^n)) \\
\vec{L}(Py_2) &= \sum_{i=1}^{52} m_i(\vec{r}_i - \vec{r}_{CM}(Py_2)) \times (\vec{v}_i - \vec{v}_t(Py_2)) \\
I &= mr^2 \\
\vec{\omega} &= [I]^{-1} \times \vec{L} \\
\vec{v}^n_{ir} &= \vec{\omega}(Py^n) \times (\vec{r}^n_i - \vec{r}_{CM}(Py^n)) \\
E^k_{tot^n} &= E^k_{r^n} + E^k_{td^n} + E^k_{intra^n} \\
E^k_{tot^n} &= \frac{1}{2}\sum_{i=1}^{26} m^n_i(\vec{v}^n_i)^2 \\
E^k_{td^n} &= \frac{1}{2} m_{Py^n}\vec{v}^2_t(Py^n) \\
E^k_r &= \frac{1}{2}\vec{\omega}(Py_2) \times [I](Py_2) \times \vec{\omega}(Py_2) = \frac{1}{2}\vec{L}(Py_2) \times [I]^{-1}(Py_2) \times \vec{L}(Py_2)
\end{aligned} \tag{4.8}$$

(4.9)

$E^k_{tot^n}$ is the total kinetic energy of monomer n. $E^k_{td^n}$ is the translation kinetic energy of pyrene monomer n. $\vec{L}(Py^n)$ is the angular momentum of pyrene monomer n. $\vec{L}(Py_2)$ is the angular momentum of the dimer. $[I]$ refers to the moment of inertia tensor. $[I]^{-1}$ is the inverse of $[I]$. $\vec{\omega}$ is the angular velocity. $\vec{r}^n_i$ and $\vec{r}_{CM}(Py^n)$ denote the coordinates of atom i and center of mass of dimer of monomer n, respectively. $\vec{r}_i$ and $\vec{r}_{CM}(Py_2)$ and denote the coordinates of atom i and center of mass of dimer, respectively.

From the endpoint of the simulation, the total energy transferred towards internal rovibrational modes of the pyrene dimer cal also be computed as:

$$\Delta E^{Py_2}_{int} = E^{k,0}_{Ar} - E^k_{Ar} - E^k_{td} \tag{4.10}$$

where $E_{Ar}^{k,0}$ is the initial argon kinetic energy whereas E_{Ar}^{k} and E_{td}^{k} correspond to kinetic energies at the end of the MD simulation. In the case of dissociated dimers at the end of the simulations, the energy deposited in the rovibrational modes of the monomers can be deduced as:

$$\Delta E_{int}^{Py^1+Py^2} = E_{Ar}^{k,0} - E_{Ar}^{k} - E_{td}^{k} - E_{Re}^{k} \tag{4.11}$$

4.4.3 Simulation of the Experimental TOFMS

The experimental TOFMS are reproduced by simulating the ion trajectories through the experimental setup in the presence of the electric fields. These are calculated by solving numerically the Laplace equation. Equations of motion are integrated using the fourth order Runge-Kutta method with adaptive step size. The occurrence of collision or dissociation is decided at each time step of the ion trajectory based on the collision and dissociation probabilities.

In the work of S. Zamith *et* al., [65] the energy transfer was treated upon collision by using the Line of Center model (LOC) [481]. In the LOC model, the transferred energy is the kinetic energy along the line of centers. Evaporation rates were then estimated using PST, in which only statistical dissociation to be possible after energy deposition in the cluster by collision was conside. Although this approach, which will be referred to as PST in the following, has been proved to be able to satisfactorily reproduce CID cross section experiments,[65] it fails to reproduce in details the shape and position of the fragment peaks in the TOFMS, as will be shown in section 4.4.4.1.

In order to better reproduce the position and peak shapes, the MD and PST methods were combined. The outputs of the MD simulations were used to treat the collisions in the ion trajectories. At each time step the probability for a collision is evaluated. The principle of MD+PST is displayed in Figure 4.23.

One MD trajectory (with proper weighting of the b values) was randomly picked from all outputs of MD simulations at a given collision energy. Then two cases have to be considered. First, if the dissociation occurred during the picked MD calculation (short time dissociation), then the MD final velocities of the fragments are used to further calculate the ion trajectories. On the other hand, if the pyrene dimer is still intact at the end of the picked MD calculation, then the dimer velocity is updated and use the collision energy transfer $\Delta E_{int}^{Py_2}$ deduced from the MD calculation to increase the internal energy of the cluster. The dissociation rate resulting from this new internal energy is then evaluated using PST. In the latter case, if dissociation occurs (long time dissociation), the relative velocities of the fragment are evaluated using the PST

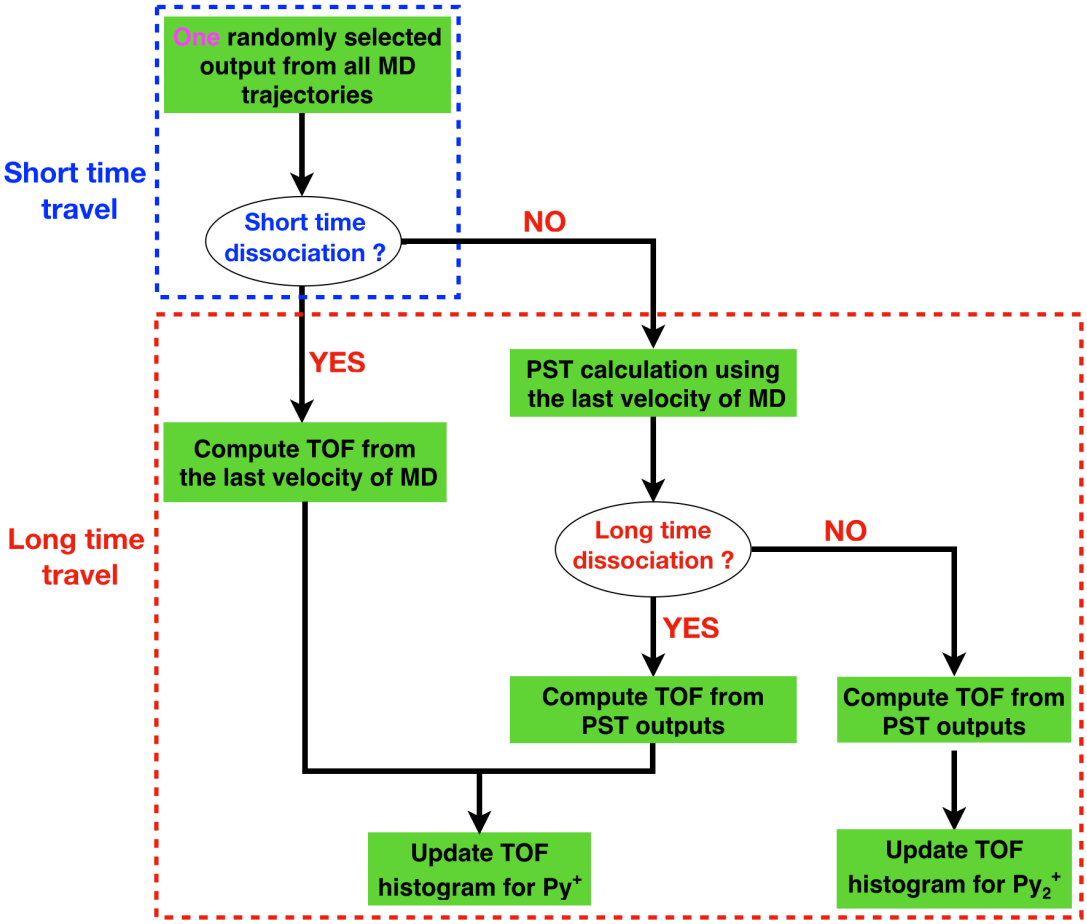


Figure 4.23: Principle of MD+PST.

outcome. The whole process of MD+PST is performed many times for different trajectories to ensure the reliability of the final obtained data. For each time, the TOFMS of Py^+ or Py_2^+ is updated.

Here I emphasize that, due to the short time scale of the MD calculations (3 ps), only direct dissociation can be captured by the MD simulations. Therefore, one has to evaluate the probability of dissociation at longer time scales after the energy deposition by collision. This is done here by considering that at longer time scales, dissociation occurs statistically and is treated by using PST.

4.4.4 Results and Discussion

In the following, I will discuss the dissociation at short, experimental and infinite timescales. The first two ones correspond to dissociation occurring during the MD simulation only or with the MD+PST model. The dissociation at infinite time accounts for all MD trajectories where the amount of energy transferred to the internal dimer rovibrational modes $\Delta E_{int}^{Py_2}$ is larger than the dissociation energy of 1.08 eV (value from references [65, 454]). It can be regarded as the dissociation occurring after an infinite time neglecting any cooling processes, such as thermal collisions or photon emissions.

4.4.4.1 TOFMS Comparison

An example of TOFMS is given in Figure 4.24. Figure 4.24(a) is centered around the intact parent mass (Py_2^+) whereas in (b) is displayed the region around the fragment peak (Py^+).

In Figure 4.24(a) are displayed three curves corresponding to the experimental one and the results of the two simulations (PST and MD+PST) for the parent ion. One can see that the peak shape and position are properly reproduced using the simulations; therefore, the essential of the ion propagation is captured by the simulations. Although some of the detected parent ions have undergone a collision without dissociation, no difference is seen in the parent peak since the collision rate is kept very small.

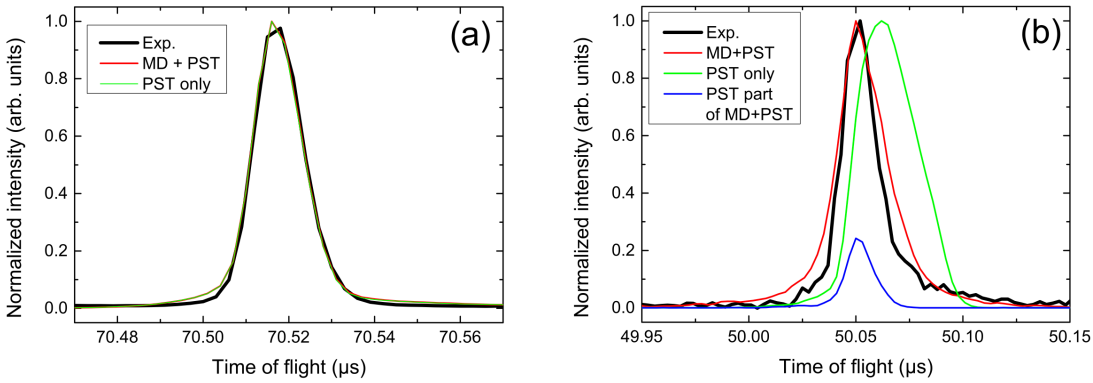


Figure 4.24: Normalized time of flight mass spectra of the parent pyrene dimer cation (a), and the pyrene fragment Py^+ (b) resulting from the collision of Py_2^+ with argon at a center of mass collision energy of 17.5 eV. The black line is for the experimental result whereas red and green curves are the MD+PST and PST model results. The blue curve is the PST subcontribution of the MD+PST model.

In Figure 4.24(b), the experimental result is compared to the PST and MD+PST simulations. Clearly, the PST based simulation fails to reproduce both the position and the shape of the peak. On the other hand, a much better agreement is found when using the output of the MD+PST simulations. This agreement is a good indication that this scheme captures the essence of the pyrene dimer cation dissociation induced by argon collisions at this collision energy. Actually, in this scheme, the largest contribution to the TOFMS results from dimers dissociating on short timescales, *i.e.* during the MD simulation. The remaining contribution, *i.e.* resulting from dimers dissociating at longer timescales and computed from the second step PST calculation, is minor and represented in blue in Figure 4.24(b).

4.4.4.2 Molecular Dynamics Analysis

Description of selected trajectories

A first qualitative description of the collision processes can be obtained from the analysis of some arbitrarily selected MD trajectories. Figure 4.25 (top and bottom) reports some snapshots extracted from two trajectories with the same collision energy (17.5 eV) and impact parameter (3.5 Å). Only the top one leads to the Py_2^+ dissociation. During the results collection, the final snapshot for each trajectory is extracted and the dimer is dissociated is considered if the distance between the two monomers molecular mass centers is larger than 10 Å. Figures 4.25-1/1* represent the system after its preliminary thermalization, when the argon atom introduced in the simulation with its initial velocity. Figures 4.25-2/2* and 3/3* represent the beginning and end of the collision. From these points, the two trajectories show different behaviors. For the top trajectory in Figure 4.25, snapshot 5 corresponds to the step where the two pyrene monomers start to go away from each other. After this, the intermolecular distance continues to increase further in snapshot 6. For the bottom trajectory in Figure 4.25, Figures 4.25-5* and 6* correspond to the middle and ending snapshots of the simulation, respectively. The snapshots 4*, 5* and 6* show the process of energy redistribution within the clusters. In particular, the soft modes associated to global deformation of the molecular planes appear to be excited.

From these two particular cases, it can be seen that the evolution of the trajectory either toward a dissociation or a redistribution of the transferred energy strongly depends on the process of energy transfer during the collision. In the top trajectory in Figure 4.25, the argon atom is pushing the two monomers far away from each other, *i.e.* the transferred energy is mostly localised in an intermolecular dissociative mode. On the opposite, in the bottom trajectory in

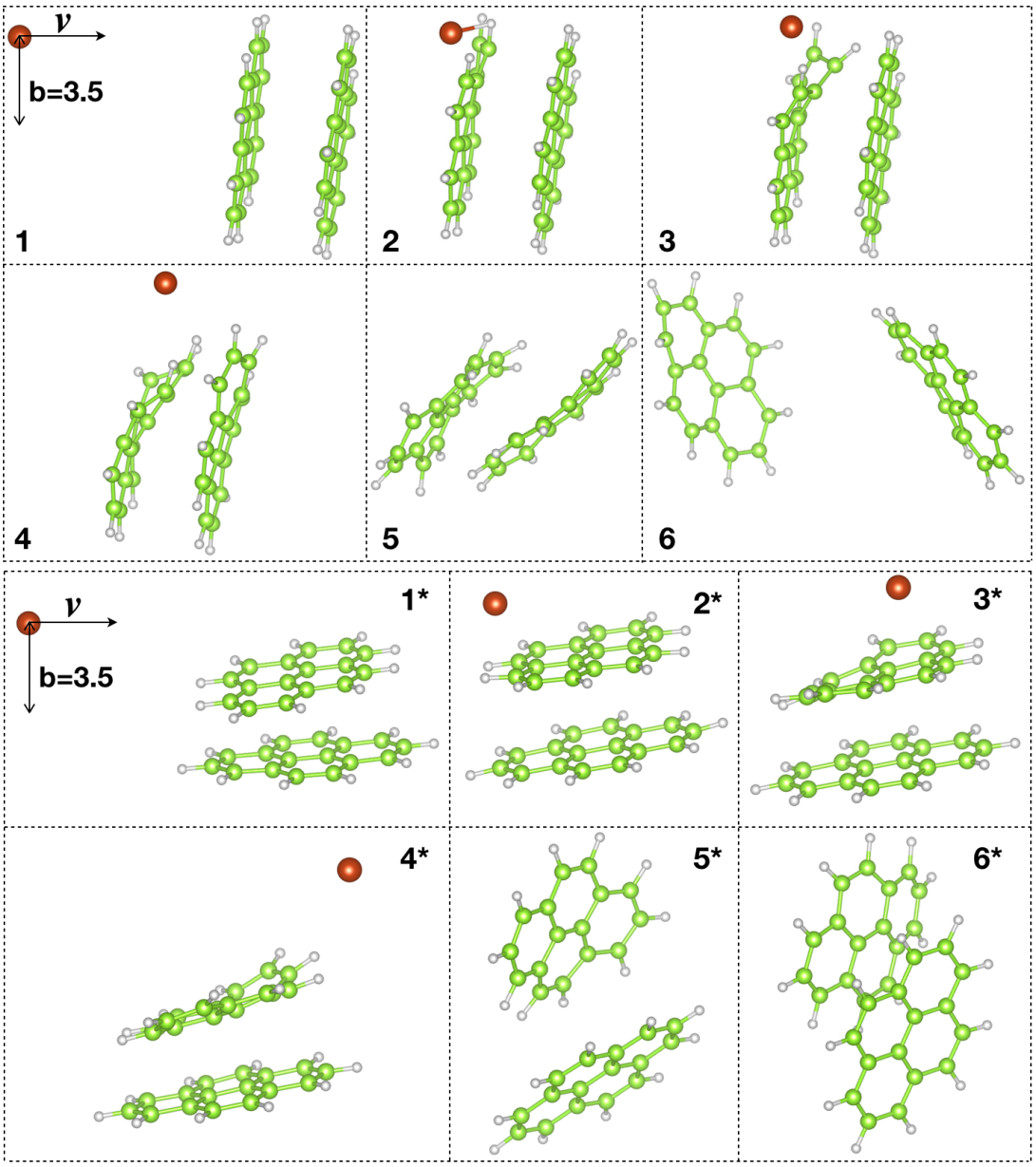


Figure 4.25: Snapshots for two different molecular dynamics trajectories. Top and bottom: trajectories with impact parameter of 3.5 Å and a collision energy of 17.5 eV, leading to dissociation and non-dissociation (top and bottom, respectively).

Figure 4.25, the collision mostly involves an intramolecular soft vibrational mode. The transferred energy is then redistributed over all the other modes. The statistical distribution of this energy is then hardly favorable to the dissociation due to the large number of intramolecular modes (72 per pyrene) with respect to the 6 intermolecular modes, only 3 of them (1 breathing and 2 parallel displaced modes) being dissociative modes.

The amount of transferred energy is also a major ingredient for the fate of the cluster. Depending on the details of the collision such as impact parameter or cluster orientation, very different amounts of energy can be transferred. This is illustrated in Figure 4.26 where the distribution of transferred energy $\Delta E_{int}^{Py_2}$ restricted to trajectories that would dissociate after infinite time, is plotted for simulations at the experimental collision energy of 17.5 eV. This distribution could hardly be guessed without a dynamical description of the collision at the atomic level. Indeed, a simpler model such as the LOC model (used in the pure PST approach) would lead to a constant distribution between the binding energy and the maximum collision energy as shown in Figure 4.26. In the distribution resulting from MD simulations, lower transferred energies are favored with respect to the distribution extracted from the LOC model. All these effects are intrinsically taken into account in the MD simulations on the opposite to the pure PST model, explaining the better agreement of the MD+PST scheme with the experimental results.

Finally, I note that the pyrene monomers remained intact (no fragmentation) up to collision energies of 25 eV. The snapshots of a fragmentation trajectory at collision energy of 27.5 eV are shown in Figure 4.27. It can be seen that the pyrene molecule impacted by the argon undergoes an opening of an aromatic cycle and the loss of two hydrogen atoms, leaving as a H_2 molecule. As the study of monomer's fragmentation is beyond the scope of the present paper, I will focus on trajectories with collision energies below this fragmentation threshold energy in the following.

Dissociation cross section

The opacity curves are presented in Figure 4.28 for various collision energies. At low impact parameters, the dissociation is very efficient even at low collision energy. At the lowest collision energy of 2.5 eV, the opacity curve presents a smooth decrease from 2 to 5 Å, whereas for collision energies larger than 10 eV, all curves are very similar. These high energy curves show high dissociation probability below 3.5 Å, reach 50% at about 4.5 Å and drop to zero for values larger than 5.5 Å. These values can be compared to the van der Waals radius of argon (1.88 Å) plus half of (i) the distance between the two molecules centers of masses (3.04 Å),

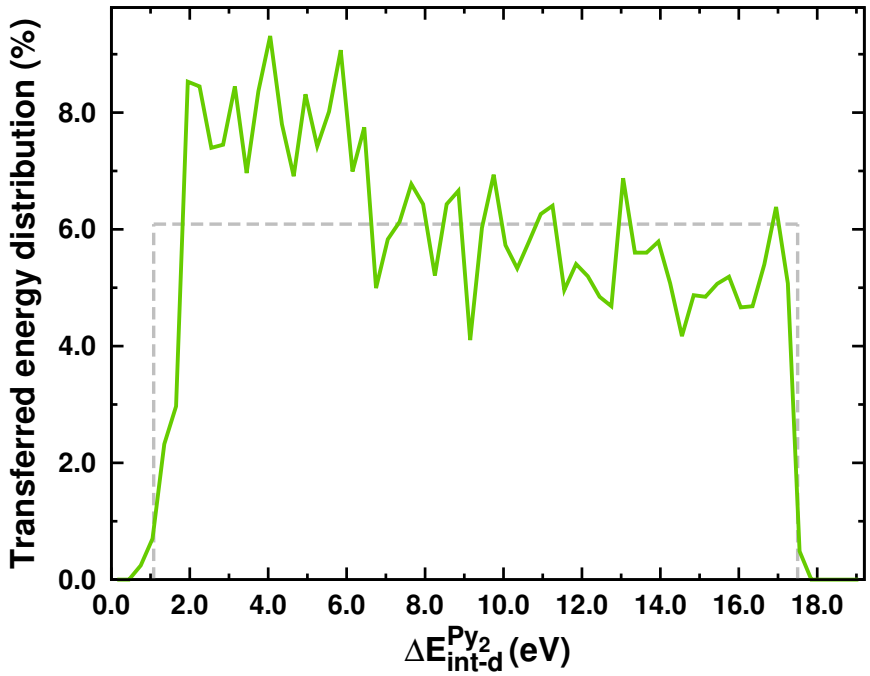


Figure 4.26: Distribution of transferred energy in rovibrational modes $\Delta E_{int}^{Py_2}$ for trajectories leading to dissociation at the end of MD (center of mass collision energy of 17.5 eV). The dashed line shows the distribution of transferred energy used in the LOC model.

(ii) the smallest (6.82 Å) or (iii) largest pyrene axes (9.25 Å) leading to distances of 3.40, 5.29 and 6.50 Å, respectively. Below 3.40 Å, all trajectories involve a frontal impact of the argon on the dimer carbonaceous system and almost all of them lead to dissociation. Unexpectedly, the opacity curve drops to zero at 5.5 Å which is lower than the largest computed value of 6.5 Å. Interestingly, taking the largest distance between carbon atoms in pyrene (7.0 Å) instead of that between hydrogen atoms (9.25 Å) leads to a value of 5.4 Å which is in line with the opacity curves. This suggests that the dissociation is efficient only if the carbonaceous skeleton area is impacted, the impact in the region of external hydrogen atoms resulting mostly in an intramolecular C-H mode excitation at the expense of dissociative modes. As a conclusion, it seems that for energies larger than 10 eV, the opacity curves are similar as they are driven by simple geometric rules, in other words, if the dimer receives a direct impact of the argon on the carbonaceous skeleton area, it will dissociate. Interestingly, this seems to be in agreement with previous works [480, 482] which also pointed out the efficient nuclear stopping power of carbon atoms in a very different context (higher energy collisions leading to knock-out process).

The blue curve in Figure 4.29 shows the MD dissociation cross sections of pyrene dimers

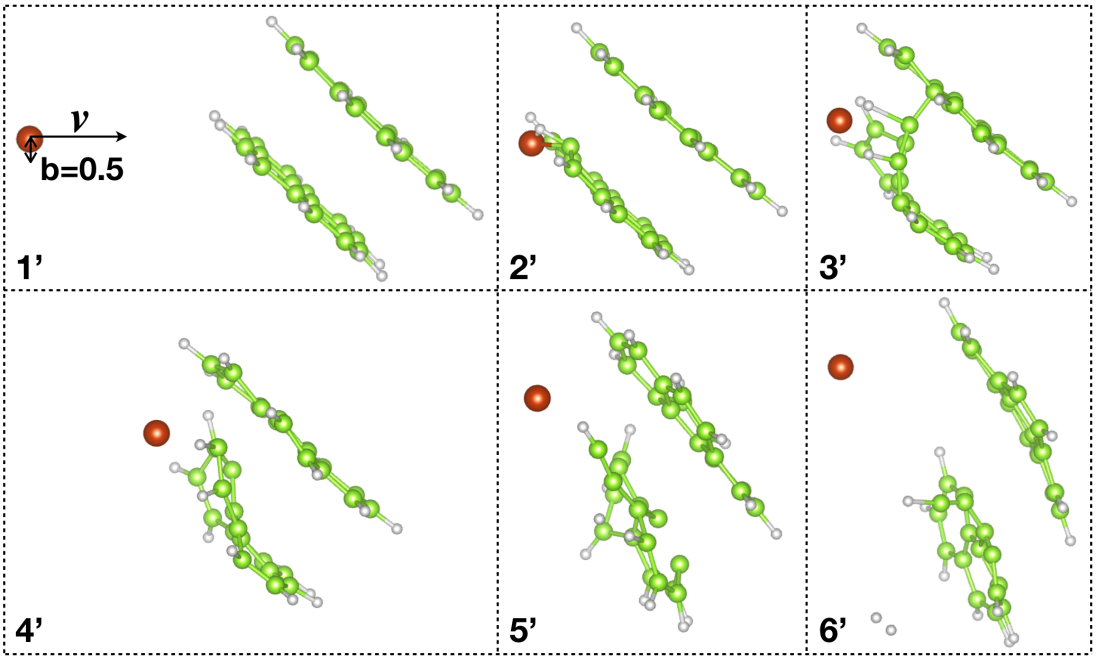


Figure 4.27: Snapshots for molecular dynamics trajectory with impact parameter of 0.5 Å and a collision energy of 27.5 eV leading to intramolecular fragmentation.

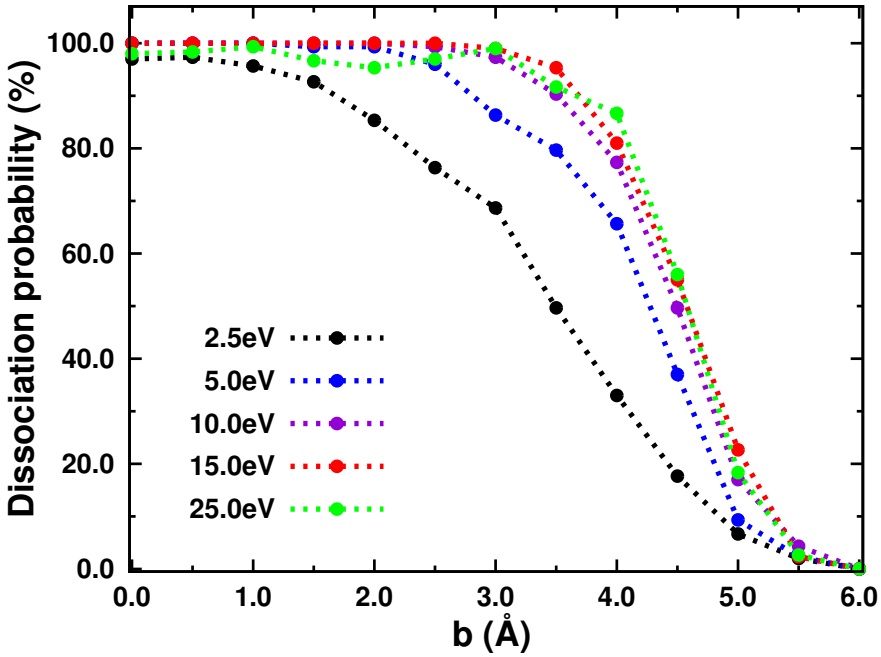


Figure 4.28: Opacity curves as a function of the impact parameter *b* for several selected center of mass collision energies.

obtained from the opacity curves following eq. 4.3. It presents a steep increase for energies bellow 7.5 eV before remaining almost constant around 65 Å^2 for collision energies greater than 10-15 eV. This is a direct consequence from the already discussed similarity of opacity curves for the high collision energies. The purple curve corresponds to dissociation at infinite timescales. Figure 4.29 also reports the cross sections computed from the MD+PST model. It can be seen that, for low collision energies, the MD and MD+PST cross sections are very close, indicating that most of the dissociations occur on the short timescales. On the opposite, at high collision energies, a non-negligible fraction of the dimers, which are not dissociated at the end of the MD simulation, carry enough energy to evaporate on the experimental timescales. At the experimental center of mass collision energy of 17.5 eV, the MD+PST cross section (about 70 Å^2) is slightly above the pure MD dissociation ratio, which indicates that the dissociation at long timescales represents a small fraction of the dissociated pyrene dimers as already seen from the TOF spectra analysis (see Figure 4.24).

I have also plotted in Figure 4.29 the model cross section σ_∞ that successfully reproduced the threshold collision-induced dissociation experimental results [65]. This model cross section is obtained by considering that the collision energy transfer is given by the LOC model and the expression for the cross section is given by:

$$\sigma_{LOC}(E_{col}) = \sigma_0(E_{col} - D)/(E_{col}). \tag{4.12}$$

where $D = 1.08$ eV is the dissociation energy [65, 454] and $\sigma_0 = 63$ Å^2 is a scaling factor usually thought as the geometrical cross section. This model cross section is usually further convolved with dissociation rates, collision energy distributions and internal energy distributions in order to be compared with experimental results. However, since here for the theoretical calculations there is no collision energy distribution, this curve could in principle be directly compared with the purple one in Figure 4.29, namely the cross section for infinite time. One can see that the MD, MD+PST results and the model cross section have similar collision energy dependence. The magnitude of the two cross sections is rather different at high collision energy, with about 60 Å^2 and 74 Å^2 for the model and infinite timescale cross sections respectively. Nevertheless, this difference is probably within the error bars of the experimental cross section measurement.

The dissociation cross sections for MD timescales with time step being 0.1 fs at collision energy of 20 and 25 eV ($\sigma_{\mathrm{MD}}(0.1)$) in Figure 4.29 are close to the ones of time step being

0.5 fs (σ_{MD}), which indicates a time step of 0.1 fs used in the MD simulation does not change significantly the corresponding dissociation cross section.

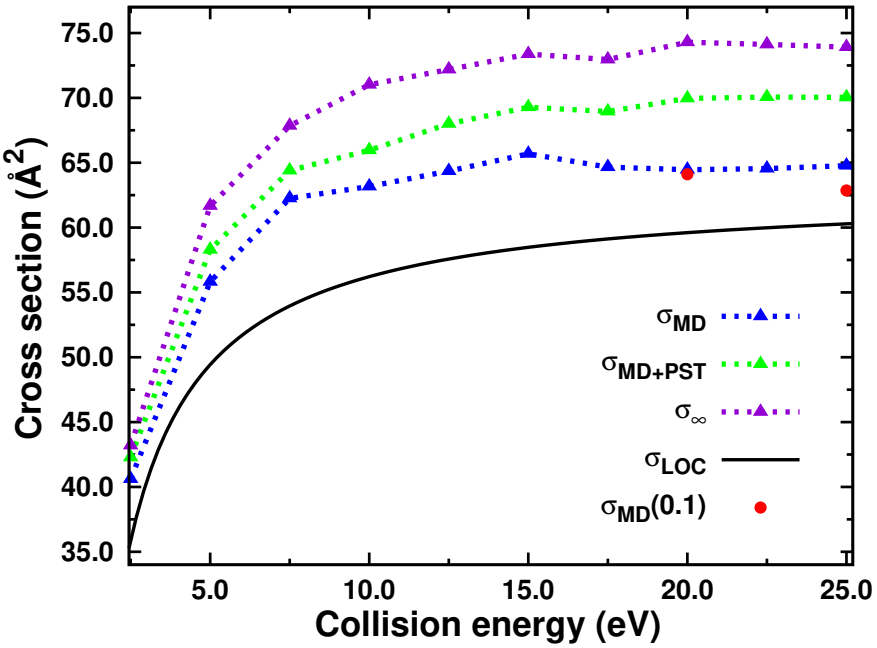


Figure 4.29: Dissociation cross sections of Py_2^+ after collision with argon as a function of center of mass collision energy for the short (MD), experimental (MD+PST) and infinite timescales. Cross sections resulting from the LOC model are also plotted. σ_{MD} (0.1) denotes the dissociation cross section for short (MD) timescale with a time step of 0.1 fs.

Energy partition

The mean value obtained for the transferred energy after removing the translation kinetic energy of the dimer, namely $\Delta E_{int}^{Py_2}$, is plotted in Figure 4.30 as a function of the collision energy. Although this quantity evolves almost linearly with the collision energy, the curves are different when one considers only the trajectories leading to dissociation or non-dissociation. For trajectories where the dimer does not dissociate, $\Delta E_{int-ud}^{Py_2}$ remains small for all collision energies below 20 eV and shows a very slight increase for collision energies larger than 20 eV. For trajectories leading to dissociation, $E_{int-d}^{Py_2}$ grows almost linearly, but above 10-15 eV most of the absorbed energy is actually used to heat the individual monomers (the green curve) whereas the energy given in the dissociative mode (difference between the blue and green curves) remains almost constant. I note that, despite the trends of the mean energy values derived from all simulations or restricted to the undissociated cases are interesting, their absolute values have

small meaning as they depend on the arbitrarily chosen b_{max} value, *i.e.* increasing b_{max} would result in more undissociated trajectories with less and less energy transfer. On the opposite, absolute values of mean energies for the dissociation trajectories are relevant, as increasing the b_{max} value would not result in new dissociation trajectories.

For MD simulations with time step being 0.1 fs at collision energy of 20 and 25 eV, the corresponding energies in Figure 4.30 are close to the ones of time step being 0.5 fs, which indicates a time step of 0.1 fs used in the MD simulation does not change significantly the corresponding deposition of the total transferred energy.

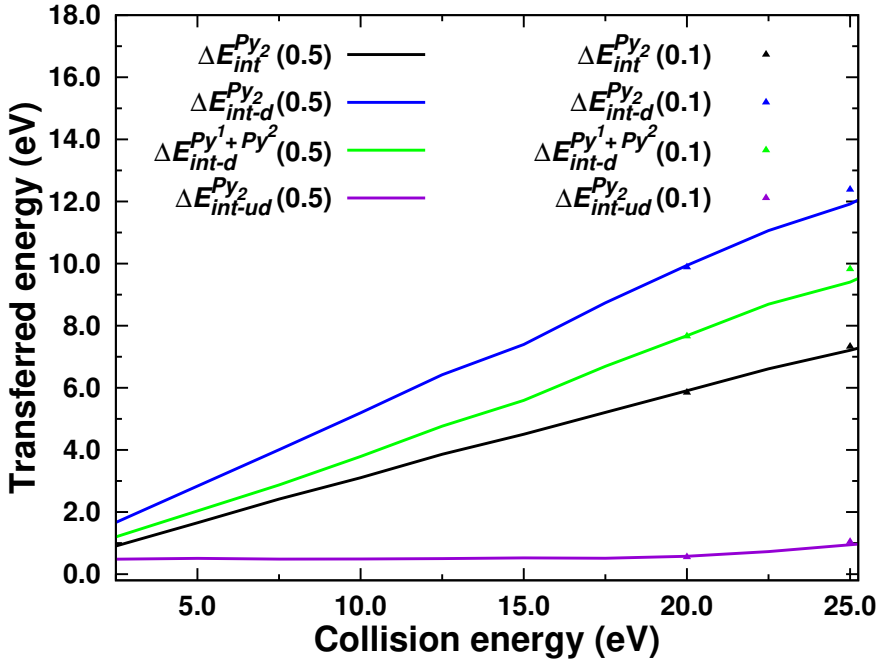


Figure 4.30: At the end of the MD collision simulations with a time step of 0.1 and 0.5 fs, the total transferred energy $\Delta E_{int}^{Py_2}$ to the rovibrational modes or restricted to the sole dissociated ($\Delta E_{int-d}^{Py_2}$) or undissociated ($\Delta E_{int-ud}^{Py_2}$) pyrene dimers as a function of collision energy. The transeferred energy to the monomers rovibrational modes for the dissociated dimers $\Delta E_{int-d}^{Py^1+Py^2}$ is also plotted.

It is also interesting to focus on the kinetic energy partition, in particular because its decomposition in sub-contributions (dissociative vs non-dissociative modes) is easier (see eqs. 4.5) than that of the potential (and consequently total) energy. For each simulated collision energy, the values for the kinetic energy sub-contributions (eqs. 4.5) are averaged over all the trajectories and reported in Table 4.5 and Figure 4.31. In addition, the averaged kinetic energy sub-contributions are also calculated over the dissociated and undissociated trajectories

Table 4.5: The kinetic energy partition after the collision of pyrene dimer with argon at different collision energies E_{col}. All energies are in eV.

E_{col}	E^k_{td}	E^k_{Ar}	$E^k_{Py^1}$	$E^k_{Py^2}$	E^k_{Re}
2.5	0.17	1.67	0.25	0.25	0.19
5.0	0.30	3.52	0.40	0.41	0.41
7.5	0.42	5.38	0.55	0.56	0.64
10.0	0.53	7.32	0.71	0.70	0.80
12.5	0.65	9.20	0.89	0.88	0.95
15.0	0.74	11.20	1.03	1.03	1.06
17.5	0.82	13.16	1.16	1.22	1.18
20.0	0.89	15.12	1.37	1.32	1.30
22.5	0.96	17.09	1.52	1.51	1.36
25.0	1.01	19.18	1.61	1.69	1.45

separately for each simulated collision energy. Then sum the contributions of dissociated and undissociated trajectories (black curves in Figure 4.31) calculated by the following eqs 4.13,

$$\begin{aligned} E^k_1 &= E^k_{td-d} * P + E^k_{td-ud} * (1-P) \\ E^k_2 &= E^k_{Re-d} * P + E^k_{Re-ud} * (1-P) \\ E^k_3 &= (E^k_{Py^1} + E^k_{Py^2})_{-d} * P + (E^k_{Py^1} + E^k_{Py^2})_{-ud} * (1-P) \end{aligned} \tag{4.13}$$

where P refers to the dissociation probability at a given center of mass collision energy. The results of sum the contributions of dissociated and undissociated trajectories are the same with the ones over all the trajectories (see Figure 4.31). This ensures our calculations for the mean kinetic energy sub-contributions are right.

The comparison of these kinetic energy sub-contributions between the time step being 0.1 and 0.5 fs used in the MD simulations is shown in Table 4.6 and Figure 4.32, which indicate a time step of 0.1 fs almost didn't affect the results of kinetic energy sub-contributions.

In Figure 4.33 are reported the ratios of the pyrene dimer translational kinetic energy E^k_{td}, relative kinetic energy E^k_{Re} and monomers rovibrational kinetic energies $E^k_{Py^1} + E^k_{Py^2}$ over the total pyrene dimer kinetic energy $E^k_{tot} - E^k_{Ar}$. It clearly appears that, whereas the contribution of the dimer translation kinetic energy (E^k_{td}) remains almost constant (very slight decrease from

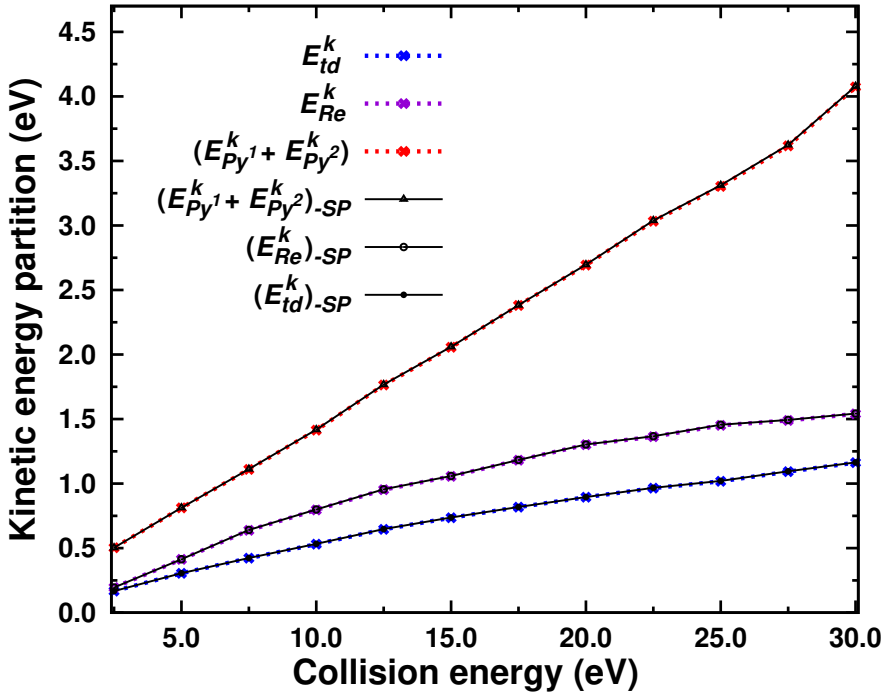


Figure 4.31: Mean kinetic energy partition at the end of the MD simulations.

about 20% to 18% of the dimer kinetic energy), this is not the case for the other two contributions. For collision energies below 7.5 eV, the proportion of the kinetic energy associated to the center of mass relative velocities increases whereas the opposite is observed for the monomers rovibrational kinetic energy. These two trends are reversed above 7.5 eV.

Again, it is convenient to analyse separately the kinetic energy partition for trajectories leading to dissociation or not as done in Figure 4.34 for E^k_{td}, E^k_{Re} and $E^k_{Py^1} + E^k_{Py^2}$. For both dissociated and undissociated trajectories, the total energy in the system computed from the initial energy at 25 K (0.32 eV) plus the transferred energy is twice the final kinetic energy computed from velocities shown in Figure 4.34 (black curves). This is exactly what one should expect from the Virial theorem.

In the absence of dissociation, the transferred energy is either small or redistributed over all the vibrational modes of the dimer, leading to small values for E^k_{Re-ud} (mean value always below 0.04 eV). The monomers rovibrational kinetic energies remain constant with an increase for collision energies above 20 eV, indicating that the slight increase of transferred energy results in a heating of the monomers, as already inferred from Figure 4.30. Once a dimer dissociates, the two pyrene molecules relative kinetic energy E^k_{Re-d} can not be transferred back

Table 4.6: The kinetic energy partition and cross section at the end of MD simulations with time step being 0.1 and 0.5 at different collision energies of 20 and 25. All energies are in eV. Time step ($Tstep$) is in fs. Cross section σ_{MD} is in Å.

E_{col}	$Tstep$	E^k_{td}	E^k_{Ar}	$E^k_{Py^1}$	$E^k_{Py^2}$	E^k_{Re}	σ_{MD}
20.0	0.1	0.89	15.18	1.34	1.35	1.28	64.11
20.0	0.5	0.89	15.12	1.37	1.32	1.30	64.45
25.0	0.1	1.03	19.04	1.71	1.67	1.44	62.86
25.0	0.5	1.01	19.18	1.61	1.69	1.45	64.77

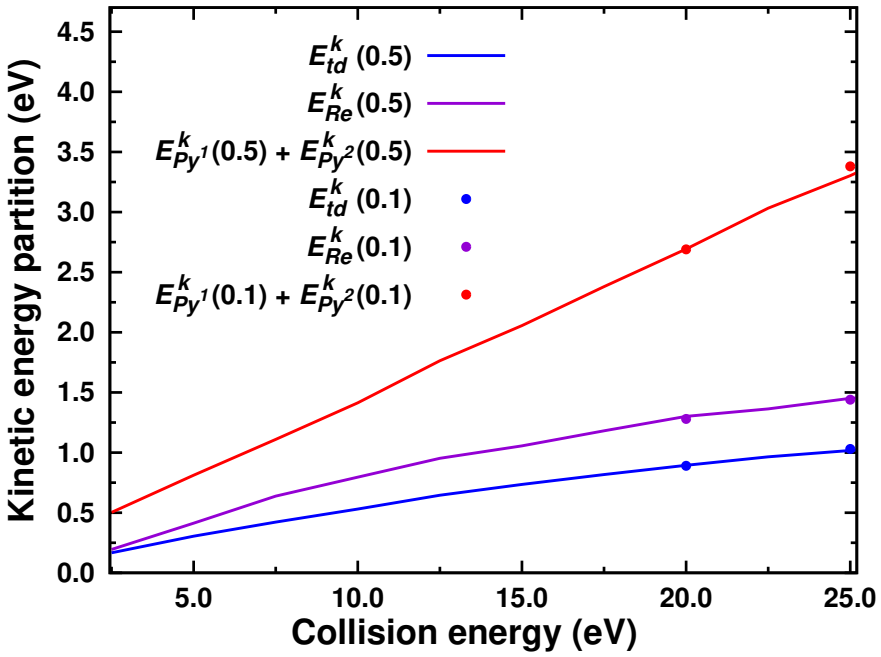


Figure 4.32: Mean kinetic energy partition at the end of the MD simulations with time step being 0.5 fs at the center of mass collision energy from 2.5 to 25 eV. The mean kinetic energy partition with time step being 0.1 fs at center of mass collision energies of 20 and 25 eV are plotted with filled round circles.

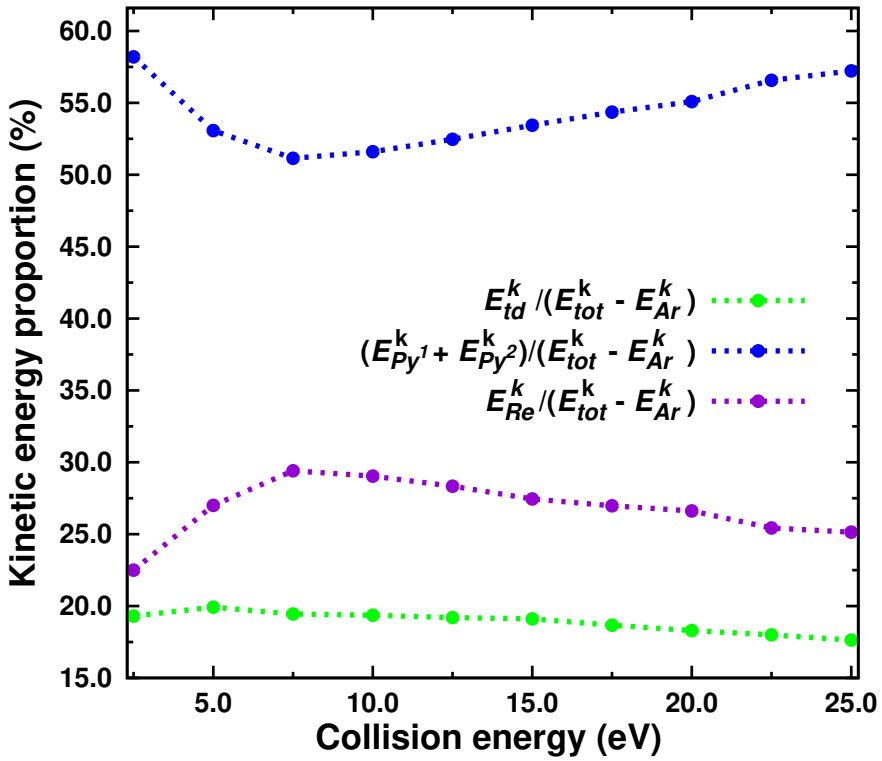


Figure 4.33: Kinetic energy proportion after collision of Py_2^+ with argon as a function of collision energy.

to the intramolecular modes and its mean value is never negligible with respect to the monomers rovibrational kinetic energies $(E^k_{Py^1} + E^k_{Py^2})_{-d}$. However, although the slope of $(E^k_{Py^1} + E^k_{Py^2})_{-d}$ remains constant with collision energies, that of E^k_{Re-d} decreases clearly. This is in line with the analysis of Figure 4.30, which shows the amount of energy transferred to the dissociative modes remains constant for high collision energies whereas the monomers are getting more internal energy.

Finally, some characteristic timescales are computed, which are presented in Figure 4.35. They correspond to the timescales for the argon with its initial velocity to travel across some characteristic distances, namely, a C-H (1.10 Å) or a C-C bond (1.40 Å) and the largest molecular axis (9.25 Å). These timescales can be compared with those of the pyrene dimer vibrational modes as an efficient energy transfer would be favored by similar orders of magnitudes. The intermolecular dimer modes possibly mixed with very soft folding modes are lying within the 70-120 cm^{-1} spectral range [477] with corresponding half-periods of 130-240 fs. These timescales are of the same order of magnitude as the time for the argon to travel across the

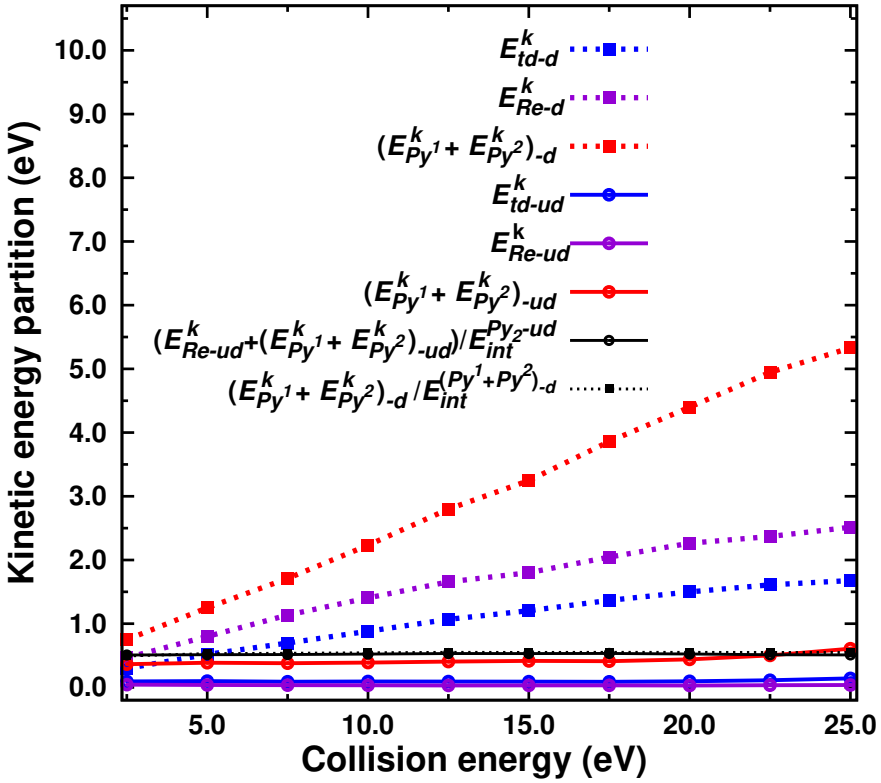


Figure 4.34: Kinetic energy partition for dissociated (-d) and undissociated (-ud) trajectories at the end of the MD simulation as a function of collision energy.

largest pyrene axis for collision energies below 10 eV. Typical frequencies for intramolecular non-soft modes are lying from 500 cm^{-1} to 3000 cm^{-1} (C-H stretching modes), leading to half-periods of 5-33 fs. For all the simulated collision energies, the characteristic times required for the argon to travel across typical C-H or C-C bond distances belong to the same order of magnitude as some of the intramolecular hard modes. Therefore, it appears from this qualitative description that the collision energy transfer toward the intermolecular modes is easier at collision energies lower than 10 eV whereas the transfer toward intramolecular modes is efficient for all the simulated collision energies. This is actually in line with the fact that the part of the absorbed collision energy taken by the non-soft intramocular modes is increasing with the collision energy at the expense of that taken by the intermolecular and soft intramolecular modes, which is in agreement with the previous energy analysis (Figures 4.30, 4.33 and 4.34).

Efficiency of energy transfer within the dimer

In this section, I address how the energy is shared inside the dimer after the collision. In par-

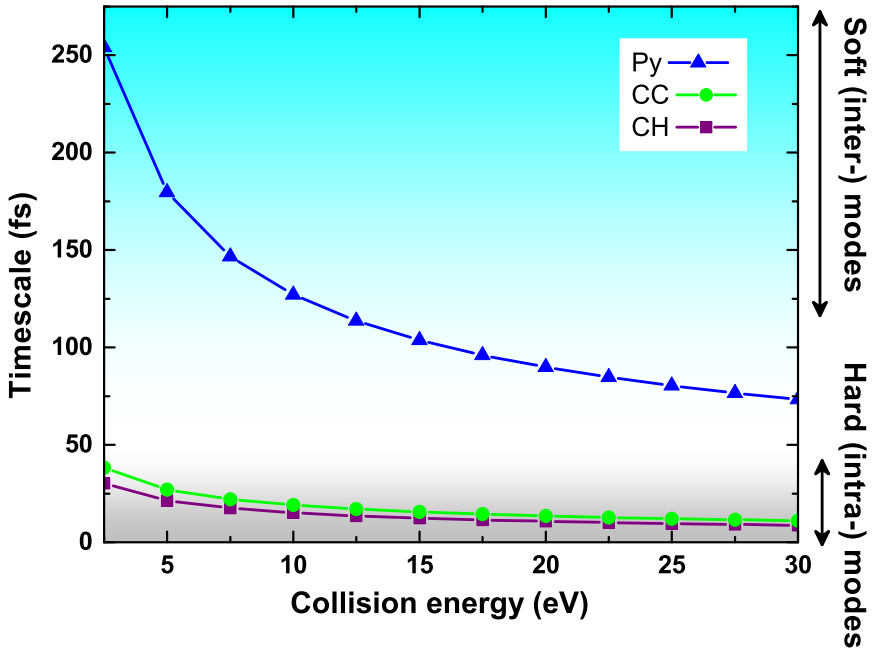


Figure 4.35: Timescales, as a function of center of mass collision energy, for argon to travel across some typical distances: a carbon-carbon bond (green), a carbon-hydrogen bond (purple) or the largest axis of the pyrene molecule (blue).

ticular, I look at the efficiency of energy transfer between the intramolecular modes of each unit and the intermolecular modes. The amount of deposited energy as well as its partition between the intramolecular modes of each molecule and the intermolecular modes is strongly dependent on the collision condition: the impact parameter, the orientation of the dimer, whether a head on collision occurs with one of the dimer atoms (and its nature, carbon or hydrogen). This results in very different evolutions of the subsequent energy flows for which precise values concerning timescales can hardly be derived. Nevertheless, the analysis of the trajectories allows to identify some characteristic behaviors. In order to estimate the thermalization process efficiency, the instantaneous intra and intermolecular kinetic temperatures are evaluated using the following formula:

$$T^k = 2\frac{< E^k >}{nk_b} \tag{4.14}$$

where k_b refers to the Boltzmann constant. n is the number of involved modes and E^k is the kinetic energy for the intra or intermolecular modes (see eqs. 4.5). T^k is plotted in Figure 4.36 for some selected trajectories obtained for collision energies of 22.5 eV and various impact pa-

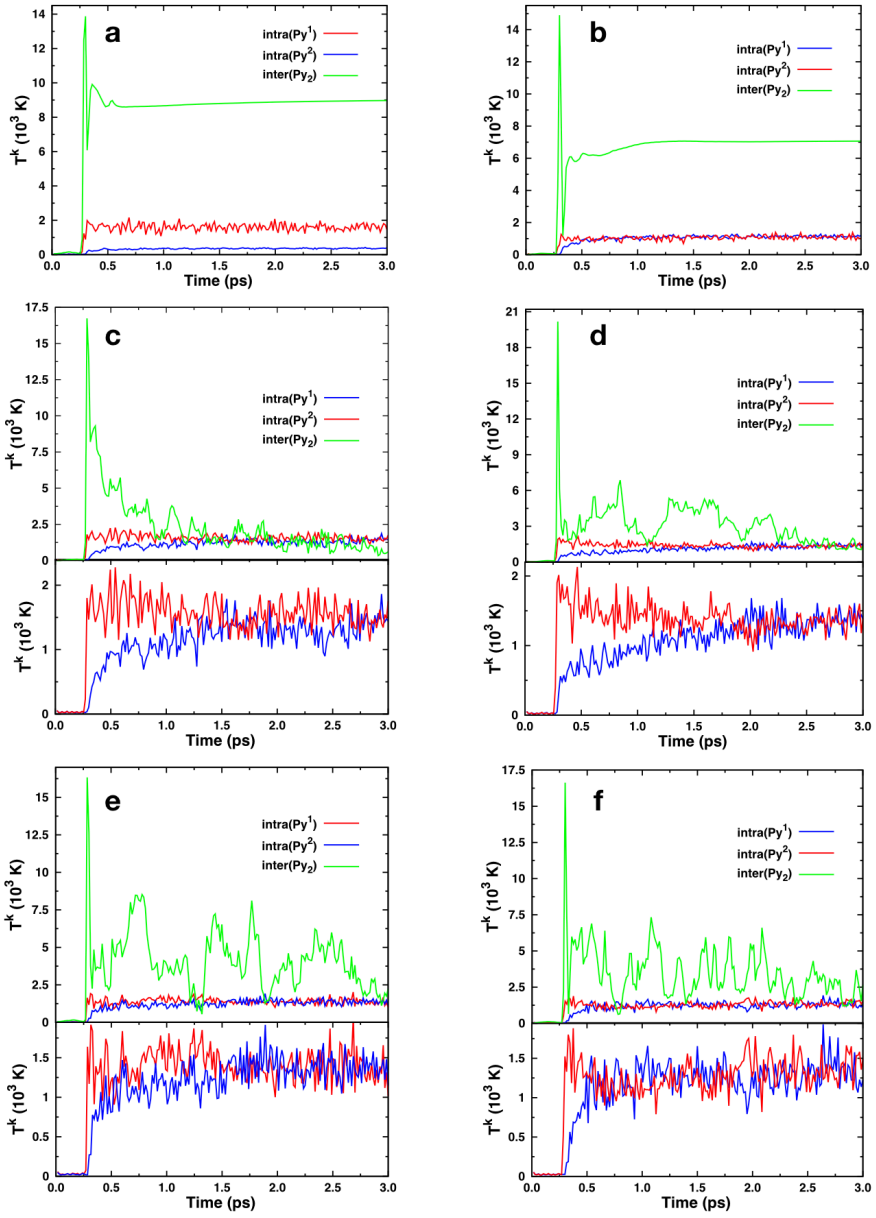


Figure 4.36: Instantaneous kinetic temperatures as a function of time for intra and intermolecular modes of the pyrene dimer at a collision energy of 22.5 eV. Impact parameters *b* are (a) 2, (b) 3, (c) 0, (d) 2.5, (e) 2, and (f) 2 Å. In cases (a) and (b) dissociation takes place whereas in the other cases the dimer remains undissociated at the end of the MD simulation. In (c) to (f) the lower panel is a vertical zoom of the corresponding intramolecular parts in upper panel.

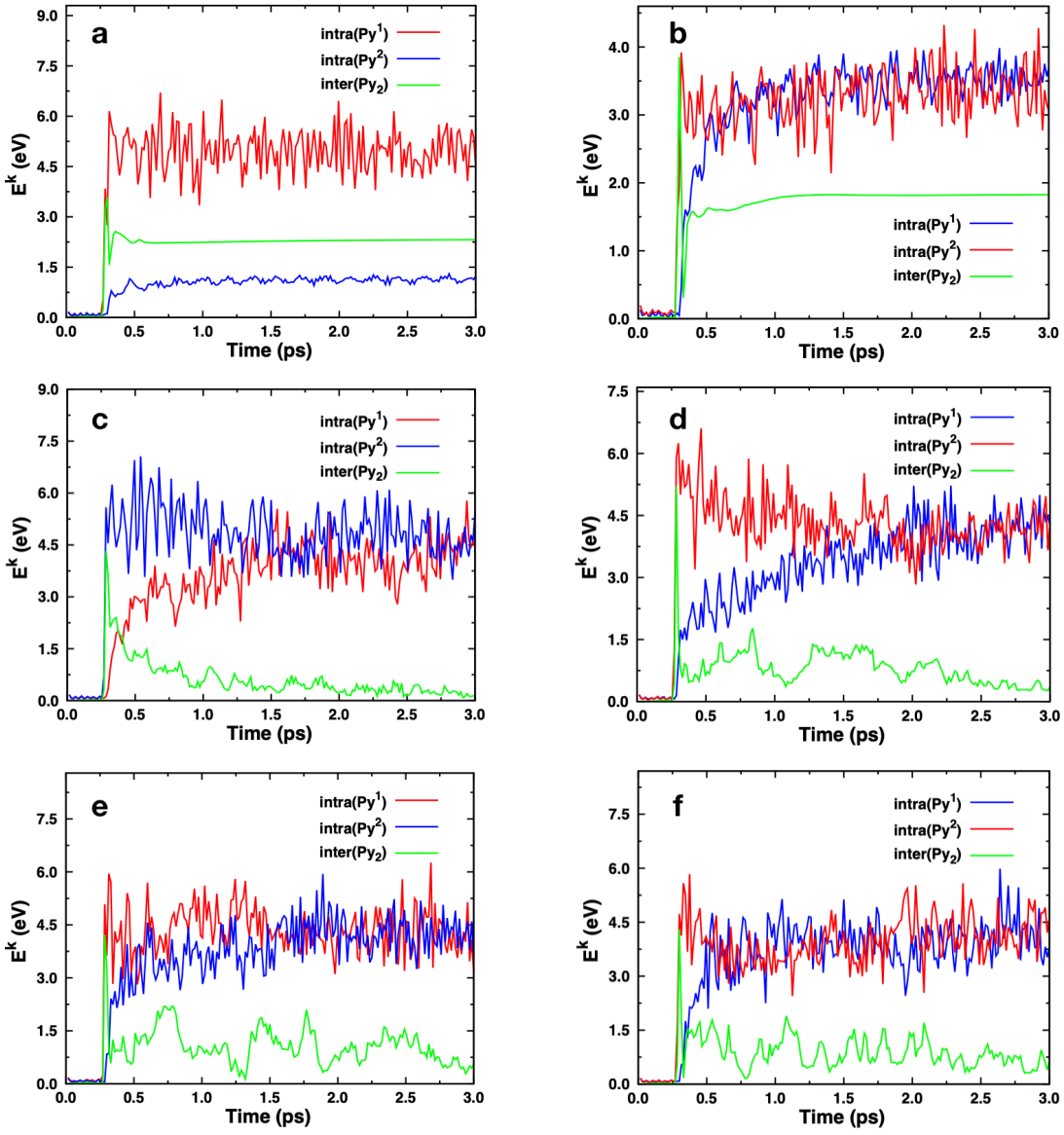


Figure 4.37: Instantaneous kinetic energies as a function of time for intra and intermolecular modes of the pyrene dimer at a collision energy of 22.5 eV. Impact parameters *b* are (a) 2, (b) 3, (c) 0, (d) 2.5, (e) 2, and (f) 2 Å. In cases (a) and (b), dissociation takes place whereas in the other cases the dimer remains undissociated at the end of the MD simulation.

rameters. The evolution of the corresponding energies ($E^k_{intra^1}$, $E^k_{intra^2}$ and E^k_{inter}) are presented in Figure 4.37. In Figures 4.36 and 4.37, simulations (a) and (b) correspond to trajectories for which dissociation occurred, whereas the dimer remained intact in the other simulations. In the simulation (a), a larger amount of energy is deposited in the first monomer with respect to the second one. The dissociation occurs before an efficient energy transfer takes place between the two monomers, leading to one hot monomer and one cold monomer at the end of the simulation. The situation is slightly different in the dissociation trajectory (b): there is a much smaller difference between the energies received during the collission by each of the monomers. One can observe that the equilibration of the two monomers intramolecular energies can take place before dissociation, leaving the two monomers with similar energies/kinetic temperatures. In the other four pictures (c, d, e, and f), corresponding to undissociated trajectories, one can see that the thermalization between the two monomers intramolecular modes occurs with timescales from 0.2 to 1.5 ps shown in Figure 4.36. On the other side, the energy equilibration between intra and intermolecular modes takes more time. Indeed, the thermalization is almost achieved in simulations (c) and (d) at 1.5 and 2.5 ps, respectively, but would take more than the simulated duration 3 ps for trajectories (e) and (f) displayed in Figure 4.36. As a conclusion of these trajectories analyses, it seems that the thermalization between intramolecular modes of the two monomers is relatively efficient (on the order of ps). On the other hand, the thermalization with the intermolecular modes is less efficient and sometimes is not observed during the simulated time of 3 ps. The direct dissociation of the dimer is a fast process (on the order of a few tenths of ps) which may prevent the thermalization taking place, leading to monomer temperatures reflecting the initial energy deposition.

4.4.5 Conclusions about CID of Py_2^+

A QM/MM dynamics study of the collision of Py_2^+ with argon at various collision energies were carried out. Argon was treated as a polarisable MM particle and Py_2^+ was treated using the SCC-DFTB method. In the dynamical simulations, a time step of 0.5 fs is proper even for high collision energies 25 eV. The TOF mass spectra of parent Py_2^+ and dissociation product Py^+ were simulated by the PST using the MD outputs at a centre of mass collision energy of 17.5 eV. With respect to TOF mass spectra extracted from pure PST simulations, considering non-statistical dissociation processes that take place before the energy redistribution from MD simulations improves the match between experimental and theoretical TOF spectra. The agreement between the measured and simulated mass spectra peak shapes and positions shows

that the essence of the collision-induced dissociation is captured by the simulation. It appears that the TOF spectra mostly result from dimers dissociating on short timescales (during the MD simulation) and the remaining minor contribution is from dimers dissociating at longer timescales (the second step, during PST calculation). This indicates that Py_2^+ primarily engages a direct dissociation path after collision.

The extraction of snapshots from the MD simulations allows to visualize the collision processes. It shows that the evolution of the trajectories either toward a dissociation or a redistribution of the transferred energy strongly depends on the initial collision conditions. Intramolecular fragmentation of the monomers occurred only for collision energies above 25 eV. The dissociation cross sections show a steep increase for collision energies below 7.5 eV and remain almost constant for collision energies greater than 10 eV. The dissociation cross section of Py_2^+ increases when dissociation occurring on longer timescale is included. As such, the dissociation cross section computed from the MD+PST model at the centre of mass collision energy of 17.5 eV is slightly higher than the value derived from pure MD simulations. The analysis of the partition of the final kinetic energy as a function of the collision energy shows how the absorbed energy is shared between the dissociative modes and the heating of individual monomers. It shows that above 7.5 eV, increasing the collision energy mostly results in an increase of the intramolecular energy. The qualitative analysis of the different timescales involved in the collision further supports the kinetic energy partition analysis. Finally, the analysis of energy transfer efficiency within the dimer suggests that direct dissociation is too fast to allow significant thermalization of the system. On the other hand, when there is no dissociation, thermalization can occur with a faster equilibration between the intramolecular modes of the two units than with the intermolecular modes.

The present results can be compared with experimental and theoretical works discussing the direct and indirect fragmentation of PAH and PAH clusters submitted to higher energy collisions [480, 482]. These authors showed that the nuclear stopping power dominates over the electronic one below 1 keV, giving a justification to our approach based on classical MD and PST. They also showed that the direct non-statistical PAH fragmentation (knock-out) is an efficient process above 20 eV. This is in line with the fact that monomer fragmentation was only observed in our MD simulations above 25 eV. Our work shows that, for PAH clusters, a regime exists below this collision energy where the dimer dissociation is governed by non-statistical processes.

4. DYNAMICAL SIMULATION OF COLLISION-INDUCED DISSOCIATION

In this study, the collision process, dissociation path, energy partition and distribution, and the efficiency of energy transfer were deeply explored for the Py_2^+ system, which can provide valuable reference for the CID study of larger PAH cation clusters.

5

General Conclusions and Perspectives

5.1 General Conclusions

As stated in the general introduction, the goal of this thesis was to go a step further into the theoretical description of properties of molecular clusters in the view to complement complex experimental measurements. It has focused on two different types of molecular clusters. First, we have focused on water clusters containing an impurity, *i.e.* an additional ion or molecule. We have first focused our studies on ammonium and ammonia water clusters in order to thoroughly explore their PES to characterize in details low-energy isomers for various cluster sizes. We have then tackled the study of protonated uracil water clusters through two aspects: characterize low-energy isomers and model collision-induced dissociation experiments to probe dissociation mechanism in relation with recent experimental measurements by S. Zamith and J.-M. l'Hermite. Finally, we addressed the study of the pyrene dimer cation to explore collision trajectories, dissociation mechanism, energy partition, mass spectra, and cross-section. These four studies have been organized in two chapters, each one gathering two studies involving similar computational tools. Below are gathered the main conclusions obtained along this thesis.

Structural and energetic properties. The structures and binding energies of the lowest-energy isomers of $(H_2O)_{1-10}NH_4^+$ and $(H_2O)_{1-10}NH_3$ clusters were obtained through a synergistic use of SCC-DFTB and PTMD. The reported low energy isomers were further optimized at the MP2/Def2TZVP level of theory. In order to improve the description of sp^3 nitrogen, we have proposed a modified set of N-H parameters. Through comparing the configurations and binding energies of the lowest-energy isomers obtained at SCC-DFTB an MP2/Def2TZVP

levels and by comparing the corresponding results to the literature, we demonstrate that this modified set of NH parameters is accurate enough to model both ammonia and ammonium water clusters. This work has thus allowed to report a number of new low-energy isomers for the studied species. Finally, PTMD simulation of $(H_2O)_{20}NH_4^+$ was conducted and the heat capacity curve of this aggregate was obtained. It is in agreement with previous results reported in the literature.

A similar exploration of the PES of $(H_2O)_{1-7,11,12}UH^+$ clusters was also performed. The reported low-energy isomers for these systems are all new and therefore constitute new data set to discuss and analyse the hydration properties of nucleobases found in RNA. The complement available structures already reported for the non-protonated $(H_2O)_nUH^+$ species. These structures have also helped use to understand recent collision-induced measurements performed by S. Zamith and J.-M. l'Hermite.

The theoretical results show that when there are 1 or 2 water molecules, the proton located is on the uracil. When there are 3 or 4 water molecules, the proton is still on the uracil but it has a tendency to be transferred to the water molecule which is directly bounded to uracil *i.e.*, forming a strongly bound U–H_2OH^+ complex. From n = 5 and above, clusters contain enough water molecules to allow for a net separation between uracil and the excess proton: The latter is often bound to a water molecule which is separated from uracil by at least one other water molecule. The localization of the excess proton in different clusters $(H_2O)_{n=1-7,11,12}UH^+$ helps to understand the evaporation channels of clusters after collision.

Overall, SCC-DFTB super powerfull continue to demonsdtret blablabl

Collision-induced dissociation study

The QM/MM dynamical simulations using SCC-DFTB method for collision-induced dissociation of low-energy protonated uracil water clusters $(H_2O)_{1-7,11,12}UH^+$ and pyrene dimer cation were performed, which provides a wealth of important information for recent experimental CID measurements.

For the explicit dynamical collision simulations of $(H_2O)_{1-7,11,12}UH^+$ at constant center of mass collision energy, the theoretical proportion of formed neutral *vs.* protonated uracil containing clusters, total fragmentation cross sections as well as the mass spectra of charged fragments are consistent with the experimental data which highlights the accuracy of the simulations. They allow to probe which fragments are formed on the short time scale and rationalize the location of the excess proton on these fragments. The results show that this latter property is highly influenced by the nature of the aggregate undergoing the collision. Analyses of the

time evolution of the fragments populations and theoretical and experimental branching ratios of $(H_2O)_{1-7,11,12}UH^+$ indicate that $(H_2O)_{1-7}UH^+$ engage a direct/shattering mechanism (dissociation on a very short time scale) after collision whereas for $(H_2O)_{11-12}UH^+$ a significant contribution of structural rearrangements occur. This suggests that a contribution of a statistical mechanism is more likely to occur for larger species such as $(H_2O)_{11-12}UH^+$. This study demonstrates that explicit molecular dynamics simulations appear as a useful tool to complement collision-induced dissociation experiments of hydrated molecular clusters.

For the dynamical simulations of collision between Py_2^+ and argon at different center of mass collision energies between 2.5 and 30 eV, the collision process, dissociation path, energy partition and distribution, and the efficiency of energy transfer were deeply explored for the Py_2^+ system, which can provide valuable reference for the CID study of larger PAH cation clusters. The simulated TOFMS of parent and dissociated products are obtained from the combination of MD simulations and PST to address the short and long timescales dissociation, respectively. The agreement between the simulated and measured mass spectra suggests that the main processes are captured by this approach. It appears that the TOFMS spectra mostly result from dimers dissociating on short timescales (during the MD simulation) and the remaining minor contribution is from dimers dissociating at longer timescales (the second step, during PST calculation). This indicates that Py_2^+ primarily engages a direct dissociation path after collision. The dynamical simulations allow to visualise the dissociation processes. It shows that the evolution of the trajectories either toward a dissociation or a redistribution of the transferred energy strongly depends on the initial collision conditions. Intramolecular fragmentation of the monomers occurred only for collision energies above 25 eV. At low collision energies, the dissociation cross section increases with collision energies whereas it remains almost constant for collision energies greater than 10-15 eV. The analysis of the kinetic energy partition as a function of the collision energy shows the absorbed energy is shared between the dissociative modes and the heating of individual monomers. It shows that above 7.5 eV, increasing the collision energy mostly results in an increase of the intramolecular energy. Finally, the analysis of energy transfer efficiency within the dimer suggests that direct dissociation is too fast to allow significant thermalization of the system. On the other hand, when there is no dissociation, thermalization can occur with a faster equilibration between the intramolecular modes of the two units than with the intermolecular modes.

5.2 Perspectives

Based on the work of this thesis, several perspectives can be implemented in the future:

- The proposed N-H integral parameter and bond parameter in SCC-DFTB can be used to explore the low-energy isomers and binding energies of clusters $(NH_3)_m$, $(NH_3)_mH^+$, $(H_2O)_n(NH_3)_m$, $(H_2O)_n(NH_3)_mH^+$ and mixed sulfate ammonia/ammonium water clusters and compare their results with the ones at MP2 level to see if these proposed parameters are proper for the calculation of these clusters.

- It would be of great interest to pursue dynamical simulations of protonated uracil water clusters by looking at the influence of collision energy, both lower or higher, on the dissociation mechanism as a function of the cluster size. Furthermore, inclusion of nuclear quantum effects in the simulations could also help to increase the accuracy of the model and improve the comparison with the experiments.

- The dynamical simulations for collision-induced dissociation of pyrene dimer cation have been verified successfully. It is possible to do the dynamical simulations for collision-induced dissociation of PAHs water clusters to compare with the experimental results and to explain and complete the experiments.

- All the simulations of water clusters in this thesis were performed in the electronic ground state, it would be interesting to investigate solvation effect on organic/inorganic molecule in both electronic ground and excited states using TD-DFTB method. It would be wonderful to calculate both absorption spectra from electronic ground state and emission spectra from electronic excited state of organic/inorganic molecule containing water clusters.

References

[1] F A Cotton, NF Curtis, CB Harris, BFG Johnson, SJ Lippard, JT Mague, WR Robinson, and JS Wood. **Mononuclear and polynuclear chemistry of rhenium (III): its pronounced homophilicity**. *Science*, **145**(3638):1305–1307, 1964. 1

[2] Hellmut Haberland. *Clusters of atoms and molecules: theory, experiment, and clusters of atoms*, **52**. Springer Science & Business Media, 2013. 1

[3] Günter Schmid. **Metal clusters and cluster metals**. *Polyhedron*, **7**(22-23):2321–2329, 1988. 1, 2

[4] Hannu Häkkinen, Michael Moseler, and Uzi Landman. **Bonding in Cu, Ag, and Au clusters: relativistic effects, trends, and surprises**. *Phys. Rev. Lett.*, **89**(3):033401, 2002. 1, 2, 3

[5] L. F. Dahl and R. E. Rundle. **The crystal structure of dimanganese decacarbonyl $Mn_2(CO)_{10}$**. *Acta Crystallogr.*, **16**(5):419–426, 1963. 1, 2

[6] L Fucugauchi, S Millan, A Mondragon, and M Solache-Rios. **Chemical effects of (n, γ) nuclear reaction on $(Mo_6Cl_8)Cl_4$**. *J. Radioanal. Nucl. Chem.*, **178**(2):437–442, 1994. 1, 2

[7] Paul W. Sutton and Lawrence F. Dahl. **The Molecular structure of $Co_3(CO)_9CCH_3$. A tricyclic organocobalt complex containing a metal-coordinated triply bridging aliphatic carbon atom**. *J. Am. Chem. Soc.*, **89**(2):261–268, 1967. PMID: 22175962. 1, 2

[8] Andreas Schnepf and Hansgeorg Schnöckel. **Metalloid aluminum and gallium clusters: element modifications on the molecular scale?** *Angew. Chem. Int. Ed.*, **41**(19):3532–3554, 2002. 1, 2

[9] Thomas F. Fässler and Stephan D. Hoffmann. **Endohedral zintl ions: intermetalloid clusters**. *Angew. Chem. Int. Ed.*, **43**(46):6242–6247, 2004. 1, 2

[10] Saskia Stegmaier and Thomas F Fässler. **A Bronze Matryoshka: The Discrete Intermetalloid Cluster $[Sn@Cu_{12}@Sn_{20}]^{-}_{12}$ - in the Ternary Phases $A_{12}Cu_{12}Sn_{21}$ (A= Na, K)**. *J. Am. Chem. Soc.*, **133**(49):19758–19768, 2011. 1, 2

[11] J Farges, MF De Feraudy, B Raoult, and G Torchet. **Structure and temperature of rare gas clusters in a supersonic expansion**. *Surf. Sci.*, **106**(1-3):95–100, 1981. 1, 2

[12] Harry W Kroto, AW Allaf, and SP Balm. **C_{60}: Buckminsterfullerene**. *Chem. Rev.*, **91**(6):1213–1235, 1991. 1, 2

[13] X Blase, A San Miguel, A Perez, et al. **Cluster assembled silicon networks**. In *Nanosilicon*, pages 79–113. Elsevier, 2008. 1

[14] Christian Siedschlag and Jan-Michael Rost. **Small rare-gas clusters in soft x-ray pulses**. *Phys. Rev. Lett.*, **93**(4):043402, 2004. 1, 2

[15] M Rapacioli, Florent Calvo, Fernand Spiegelman, C Joblin, and DJ Wales. **Stacked clusters of polycyclic aromatic hydrocarbon molecules**. *J. Phys. Chem. A*, **109**(11):2487–2497, 2005. 1, 2

[16] Junfeng Zhen, Tao Chen, and Alexander GGM Tielens. **Laboratory photochemistry of pyrene clusters: An efficient way to form large PAHs**. *Astrophys. J.*, **863**(2):128, 2018. 1, 2, 4, 9, 132

[17] Paul J Dyson and J Scott McIndoe. *Transition metal carbonyl cluster chemistry*, **2**. CRC Press, 2000. 2

[18] R Colton and R. L. Martin. **Magnetic evidence for trinuclear clusters in rhenium (IV) chloride**. *Nature*, **205**(4968):239–240, 1965. 2

[19] Jiho Yoon, Edouard Lesne, Kornelia Sklarek, John Sheckelton, Chris Pasco, Stuart S. P. Parkin, Tyrel M. McQueen, and Mazhar N. Ali. **Anomalous thickness-dependent electrical conductivity in van der Waals layered transition metal halide, Nb_3Cl_8**. *J. Phys. Condens. Matter*, **32**(30), JUL 15 2020. 2

[20] Giel Berden, W Leo Meerts, Michael Schmitt, and Karl Kleinermanns. **High resolution UV spectroscopy of phenol and the hydrogen bonded phenol-water cluster**. *J. Chem. Phys.*, **104**(3):972–982, 1996. 2, 7

[21] Udo Buck and Friedrich Huisken. **Infrared spectroscopy of size-selected water and methanol clusters**. *Chem. Rev.*, **100**(11):3863–3890, 2000. 2

[22] IA Harris, RS Kidwell, and JA Northby. **Structure of charged argon clusters formed in a free jet expansion**. *Phys. Rev. Lett.*, **53**(25):2390, 1984. 2

[23] A Ayuela, JM López, JA Alonso, and V Luana. **Theoretical study of NaCl clusters**. *Z. Phys. D*, **26**(1):213–215, 1993. 3

[24] Florent Calvo. **Melting and structural transitions**. *Frontiers of Nanoscience*, **12**:295–331, 2018. 3

[25] JA Alonso. **Electronic and atomic structure, and magnetism of transition-metal clusters**. *Chem. Rev.*, **100**(2):637–678, 2000. 3

[26] Henry Margenau and Neil R Kestner. *Theory of Intermolecular Forces: International Series of Monographs in Natural Philosophy*, **18**. Elsevier, 2013. 3

[27] Keisaku Kimura. **Magnetic properties of iron: from clusters to bulk**. *Physics Letters A*, **158**(1-2):85–88, 1991. 3

[28] Joshua Jortner. **Cluster size effects**. *Z. Phys. D*, **24**(3):247–275, 1992. 3, 4

[29] Julien Boulon, Isabelle Braud, Sébastien Zamith, Pierre Labastie, and Jean-Marc L'Hermite. **Experimental nanocalorimetry of protonated and deprotonated water clusters**. *J. Chem. Phys.*, **140**(16):164305, 2014. 3, 4, 5

[30] M Schmidt and B von Issendorff. **Gas-phase calorimetry of protonated water clusters**. *J. Chem. Phys.*, **136**(16):164307, 2012. 4

[31] C Hock, M Schmidt, R Kuhnen, C Bartels, L Ma, H Haberland, and B v Issendorff. **Calorimetric observation of the melting of free water nanoparticles at cryogenic temperatures**. *Phys. Rev. Lett.*, **103**(7):073401, 2009. 4

[32] Michael Faraday. **The Bakerian Lecture.—Experimental relations of gold (and other metals) to light**. *Philosophical Transactions of the Royal Society of London*, **147**:145–181, 1857. 3

[33] Markku Kulmala, Liisa Pirjola, and Jyrki M Mäkelä. **Stable sulphate clusters as a source of new atmospheric particles**. *Nature*, **404**(6773):66–69, 2000. 4

REFERENCES

[34] HUI WANG, NED S WINGREEN, AND RANJAN MUKHOPADHYAY. **Self-organized periodicity of protein clusters in growing bacteria**. *Phys. Rev. Lett.*, **101**(21):218101, 2008. 4

[35] JOSEPH W DEPALMA, DOUGLAS J DOREN, AND MURRAY V JOHNSTON. **Formation and growth of molecular clusters containing sulfuric acid, water, ammonia, and dimethylamine**. *J. Phys. Chem. A*, **118**(29):5464–5473, 2014. 4, 7

[36] I KATAKUSE, T ICHIHARA, Y FUJITA, T MATSUO, T SAKURAI, AND H MATSUDA. **Mass distributions of copper, silver and gold clusters and electronic shell structure**. *Int. J. Mass Spectrom. Ion Processes*, **67**(2):229–236, 1985. 4

[37] JAN POSTHUMUS. *Molecules and clusters in intense laser fields*. Cambridge University Press, 2009. 4

[38] AW CASTLEMAN JR AND SN KHANNA. **Clusters, superatoms, and building blocks of new materials**. *J. Phys. Chem. C*, **113**(7):2664–2675, 2009. 4

[39] ARNIM HENGLEIN. **Non-metallic silver clusters in aqueous solution: stabilization and chemical reactions**. *Chem. Phys. Lett.*, **154**(5):473–476, 1989. 4

[40] NG KOROBEISHCHIKOV, AE ZARVIN, V ZH MADIRBAEV, AND RG SHARAFUTDINOV. **Condensation of argon, monosilane and their mixtures in a pulse free jet**. *Plasma Chem. Plasma Process.*, **25**(4):319–349, 2005. 4

[41] QIAOFEI XU, BAIJIE XU, HUI KONG, PEIPEI HE, JIAWEI WANG, THIRUMURTHY KANNAN, PENGTAO MA, JINGPING WANG, AND JINGYANG NIU. **Synthesis and Characterization of a Crown-Shaped 36-Molybdate Cluster and Application in Catalyzing Knoevenagel Condensation**. *Inorg. Chem.*, **59**(15):10665–10672, AUG 3 2020. 4

[42] TIAN XIAO-LIN, ZHAO YU-HONG, TIAN JIN-ZHONG, AND HOU HUA. **Effects of interatomic potential on precipitation sequences of medium Al concentration in Ni75AlxV25-x alloys**. *Acta Phys. Sin.*, **67**(23), DEC 5 2018. 4

[43] Y. H. DENG, D. D. WEN, Y. LI, J. LIU, AND P. PENG. **Local atomic structures of amorphous $Pd_{80}Si_{20}$ alloys and their configuration heredity in the rapid solidification**. *Philos. Mag.*, **98**(31, A):2861–2877, NOV 2 2018. 4

[44] MATHIAS RAPACIOLI, FERNAND SPIEGELMAN, AND NATHALIE TARRAT. **Evidencing the relationship between isomer spectra and melting: the 20-and 55-atom silver and gold cluster cases**. *Phys. Chem. Chem. Phys.*, **21**(45):24857–24866, DEC 7 2019. 4

[45] A. W CASTLEMAN AND P. M. HOLLAND. **Molecular-properties of small clusters and their relationship to understanding nucleation phenomena**. In *Abstracts of papers of the american chemical society*, **175**, pages 74–74. Am. Chem. Soc. 1155 16TH ST, NW, WASHINGTON, DC 20036, 1978. 4

[46] AW CASTLEMAN JR, PAUL M HOLLAND, AND ROBERT G KEESEE. **The properties of ion clusters and their relationship to heteromolecular nucleation**. *J. Chem. Phys.*, **68**(4):1760–1767, 1978. 4

[47] Q ZHONG AND AWA CASTLEMAN. **An ultrafast glimpse of cluster solvation effects on reaction dynamics**. *Chem. Rev.*, **100**(11):4039–4058, 2000. 4

[48] ANDREW PINKARD, ANOUCK M CHAMPSAUR, AND XAVIER ROY. **Molecular clusters: nanoscale building blocks for solid-state materials**. *Acc. Chem. Res.*, **51**(4):919–929, 2018. 4

[49] ASUKA FUJII AND KENTA MIZUSE. **Infrared spectroscopic studies on hydrogen-bonded water networks in gas phase clusters**. *Int. Rev. Phys. Chem.*, **32**(2):266–307, 2013. 4

[50] ZHIXUN LUO, A. W. CASTLEMAN, AND SHIV N. KHANNA. **Reactivity of metal clusters**. *Chem. Rev.*, **116**(23):14456–14492, 2016. PMID: 27960263. 4

[51] LAI-SHENG WANG. **Photoelectron spectroscopy of size-selected boron clusters: from planar structures to borophenes and borospherenes**. *Int. Rev. Phys. Chem.*, **35**(1):69–142, 2016. 4

[52] SHUAI JIANG, XIAOLI ZHANG, XIUCHENG ZHENG, AND XIANGTAO KONG. **Structures, energetics, and infrared spectra of the cationic monomethylamine clusters**. *Chemical Physics*, **523**:7–11, 2019. 4

[53] MICHAEL MEOT-NER. **The ionic hydrogen bond and ion solvation. 2. Solvation of onium ions by one to seven H2O molecules. Relations between monomolecular, specific, and bulk hydration**. *J. Am. Chem. Soc*, **106**(5):1265–1272, 1984. 4

[54] A. W. CASTLEMAN AND S. WEI. **Cluster Reactions**. *Annual Review of Physical Chemistry*, **45**(1):685–719, 1994. 4

[55] A. W. CASTLEMAN AND K. H. BOWEN. **Clusters:FIXME Structure, energetics, and dynamics of intermediate states of matter**. *J. Phys. Chem.*, **100**(31):12911–12944, 1996. 4

[56] JAMES M FARRAR, WILLIAM HUNDLEY SAUNDERS, ET AL. *Techniques for the study of ion-molecule reactions*. John Wiley & Sons, 1988. 4

[57] PAUL M MAYER. **Ion rearrangement at the beginning of cluster formation: isomerization pathways and dissociation kinetics for the ionized dimethylamine dimer**. *International J. Mass Spectrom.*, **218**(1):1–10, 2002. 4

[58] DAVID J WALES AND JONATHAN PK DOYE. **Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms**. *J. Phys. Chem. A*, **101**(28):5111–5116, 1997. 5, 47

[59] DAVID J WALES AND MATTHEW P HODGES. **Global minima of water clusters $(H_2O)_n$, $n \leqslant 21$, described by an empirical potential**. *Chem. Phys. Lett.*, **286**(1-2):65–72, 1998. 5

[60] DAVID J WALES AND HAROLD A SCHERAGA. **Global optimization of clusters, crystals, and biomolecules**. *Science*, **285**(5432):1368–1372, 1999. 5, 47

[61] TIM JAMES, DAVID J WALES, AND JAVIER HERNÁNDEZ-ROJAS. **Global minima for water clusters $(H_2O)_n$, $n \leqslant 21$, described by a five-site empirical potential**. *Chem. Phys. Lett.*, **415**(4-6):302–307, 2005. 5

[62] KSENIIA KORCHAGINA, AUDE SIMON, MATHIAS RAPACIOLI, FERNAND SPIEGELMAN, JEAN-MARC L'HERMITE, ISABELLE BRAUD, SÉBASTIEN ZAMITH, AND JÉRÔME CUNY. **Theoretical investigation of the solid–liquid phase transition in protonated water clusters**. *Phys. Chem. Chem. Phys.*, **19**(40):27288–27298, 2017. 5, 54, 102

[63] YASUHIRO HADA. **A molecular dynamics simulation of cluster dissociation process under cluster ion implantation**. *Phys. Rev. B Condens. Matter*, **340**:1036–1040, 2003. 5

[64] DEBDUTTA CHAKRABORTY, HANS LISCHKA, AND WILLIAM L HASE. **Dynamics of Pyrene-Dimer Association and Ensuing Pyrene-Dimer Dissociation**. *J. Phys. Chem. A*, **124**(43):8907–8917, 2020. 5

[65] SÉBASTIEN ZAMITH, JEAN-MARC L'HERMITE, LÉO DONTOT, LINJIE ZHENG, MATHIAS RAPACIOLI, FERNAND SPIEGELMAN, AND CHRISTINE JOBLIN. **Threshold collision induced dissociation of pyrene cluster cations**. *J. Chem. Phys.*, **153**(5):054311, 2020. 5, 9, 99, 132, 135, 137, 143

[66] LINJIE ZHENG, SÉBASTIEN ZAMITH, AND MATHIAS RAPACIOLI. **Dynamical simulation of collision-induced dissociation of pyrene dimer cation**. *Theor. Chem. Acc.*, **140**(2):1–14, 2021. 5

[67] ROBERT TIBSHIRANI AND GUENTHER WALTHER. **Cluster validation by prediction strength**. *J. Comput. Graph. Stat.*, **14**(3):511–528, 2005. 5

[68] AS WOUTERSEN, U EMMERICHS, AND HJ BAKKER. **Femtosecond mid-IR pump-probe spectroscopy of liquid water: Evidence for a two-component structure**. *Science*, **278**(5338):658–660, 1997. 5

[69] CHONG-YU RUAN, VLADIMIR A LOBASTOV, FRANCO VIGLIOTTI, SONGYE CHEN, AND AHMED H ZEWAIL. **Ultrafast electron crystallography of interfacial water**. *Science*, **304**(5667):80–84, 2004. 5

[70] J-B BRUBACH, AL MERMET, A FILABOZZI, A GERSCHEL, AND P ROY. **Signatures of the hydrogen bonding in the infrared bands of water**. *J. Chem. Phys.*, **122**(18):184509, 2005. 5

[71] UWE BERGMANN, DENNIS NORDLUND, PH WERNET, MICHAEL ODELIUS, LARS GM PETTERSSON, AND ANDERS NILSSON. **Isotope effects in liquid water probed by x-ray Raman spectroscopy**. *Phys. Rev. B*, **76**(2):024202, 2007. 5

[72] WANDARED POKAPANICH, HENRIK BERGERSEN, IOANA L BRADEANU, RICARDO RT MARINHO, ANDREAS LINDBLAD, SEBASTIEN LEGENDRE, ALDANA ROSSO, SVANTE SVENSSON, MAXIM TCHAPLYGUINE, NIKOLAI V KRYZHEVOI, ET AL. **Auger electron spectroscopy as a probe of the solution of aqueous ions**. *J. Am. Chem. Soc.*, **131**(21):7264–7271, 2009. 5

[73] QIANG SUN. **The single donator-single acceptor hydrogen bonding structure in water probed by Raman spectroscopy**. *J. Chem. Phys.*, **132**(5):054507, 2010. 5

[74] Y HARADA, J MIYAWAKI, H NIWA, K YAMAZOE, LARS GM PETTERSSON, AND ANDERS NILSSON. **Probing the OH stretch in different local environments in liquid water**. *J. Phys. Chem. Lett.*, **8**(22):5487–5491, 2017. 5

[75] KOSUKE YAMAZOE, JUN MIYAWAKI, HIDEHARU NIWA, ANDERS NILSSON, AND YOSHIHISA HARADA. **Measurements of ultrafast dissociation in resonant inelastic x-ray scattering of water**. *J. Chem. Phys.*, **150**(20):204201, 2019. 5

[76] PIER LUIGI SILVESTRELLI AND MICHELE PARRINELLO. **Structural, electronic, and bonding properties of liquid water from first principles**. *J. Chem. Phys.*, **111**(8):3572–3580, 1999. 5

[77] DAMIEN LAAGE AND JAMES T HYNES. **A molecular jump mechanism of water reorientation**. *Science*, **311**(5762):832–835, 2006. 5

[78] VYACHESLAV S BRYANTSEV, MAMADOU S DIALLO, ADRI CT VAN DUIN, AND WILLIAM A GODDARD III. **Evaluation of B3LYP, X3LYP, and M06-class density functionals for predicting the binding energies of neutral, protonated, and deprotonated water clusters**. *J. Chem. Theory Comput.*, **5**(4):1016–1026, 2009. 5

[79] PIER LUIGI SILVESTRELLI. **Hydrogen bonding characterization in water and small molecules**. *J. Chem. Phys.*, **146**(24):244315, 2017. 5

[80] JOHN D BERNAL AND RALPH H FOWLER. **A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions**. *J. Chem. Phys.*, **1**(8):515–548, 1933. 5

[81] ROBERT M. SHIELDS, BERHANE TEMELSO, KAYE A. ARCHER, THOMAS E. MORRELL, AND GEORGE C. SHIELDS. **Accurate predictions of water cluster formation, $(H_2O)_{n=2-10}$, journal = J. Phys. Chem. A, volume = 114, number = 43, pages = 11725-11737, year = 2010**. 5

[82] JONATHON K GREGORY AND DAVID C CLARY. **Structure of water clusters. The contribution of many-body forces, monomer relaxation, and vibrational zero-point energy**. *J. Phys. Chem.*, **100**(46):18014–18022, 1996. 5

[83] MARINUS KUNST AND JOHN M WARMAN. **Proton mobility in ice**. *Nature*, **288**(5790):465–467, 1980. 6

[84] MIQUEL TORRENT-SUCARRAT AND JOSEP M ANGLADA. **Anharmonicity and the Eigen-Zundel dilemma in the IR spectrum of the protonated 21 water cluster**. *J. Chem. Theory Comput.*, **7**(2):467–472, 2011. 6

[85] XUE-BIN WANG, KAROL KOWALSKI, LAI-SHENG WANG, AND SOTIRIS S XANTHEAS. **Stepwise hydration of the cyanide anion: A temperatur-controlled photoelectron spectroscopy and ab initio computational study of CN-$(H_2O)_n$, n = 2-5**. *J. Chem. Phys.*, **132**(12):124306, 2010. 6

[86] E. K. BIGG. **Stratospheric particles**. *J. Atmos. Sci.*, **32**(5):910–917, 1975. 6

[87] VERONICA VAIDA AND JILL E HEADRICK. **Physicochemical properties of hydrated complexes in the Earth's atmosphere**. *J. Phys. Chem. A*, **104**(23):5401–5412, 2000. 6

[88] SIMONE ALOISIO AND JOSEPH S FRANCISCO. **Radical-water complexes in earth's atmosphere**. *Acc. Chem. Res.*, **33**(12):825–830, 2000. 6

[89] VCPJ RAMANATHAN, PJ CRUTZEN, JT KIEHL, AND DANIEL ROSENFELD. **Aerosols, climate, and the hydrological cycle**. *science*, **294**(5549):2119–2124, 2001. 6

[90] PATRICK R MCCURDY, WAYNE P HESS, AND SOTIRIS S XANTHEAS. **Nitric acid- water complexes: Theoretical calculations and comparison to experiment**. *J. Phys. Chem. A*, **106**(33):7628–7635, 2002. 6

[91] GREGORY M HARTT, GEORGE C SHIELDS, AND KARL N KIRSCHNER. **Hydration of OCS with one to four water molecules in atmospheric and laboratory conditions**. *J. Phys. Chem. A*, **112**(19):4490–4495, 2008. 6

[92] VERONICA VAIDA. **Perspective: Water cluster mediated atmospheric chemistry**. *J. Chem. Phys.*, **135**(2):020901, 2011. 6

[93] HENRIK G KJAERGAARD, TIMOTHY W ROBINSON, DARYL L HOWARD, JOHN S DANIEL, JILL E HEADRICK, AND VERONICA VAIDA. **Complexes of importance to the absorption of solar radiation**. *J. Phys. Chem. A*, **107**(49):10680–10686, 2003. 6

[94] VERONICA VAIDA, HENRIK G KJAERGAARD, AND KARL J FEIERABEND. **Hydrated complexes: Relevance to atmospheric chemistry and climate**. *Int. Rev. Phys. Chem.*, **22**(1):203–219, 2003. 6

[95] PETR KLÁN, DAVID DEL FAVERO, ALENA ANSORGOVÁ, JANA KLÁNOVÁ, AND IVAN HOLOUBEK. **Photodegradation of halobenzenes in water ice**. *Environ. Sci. Pollut. Res.*, **8**(3):195–200, 2001. 6

[96] L AMIAUD, F DULIEU, J-H FILLION, A MOMENI, AND JL LEMAIRE. **Interaction of atomic and molecular deuterium with a nonporous amorphous water ice surface between 8 and 30 K**. *J. Chem. Phys.*, **127**(14):144709, 2007. 6

[97] TF KAHAN, R ZHAO, AND DJ DONALDSON. **Hydroxyl radical reactivity at the air-ice interface**. *Atmos. Chem. Phys.*, **10**(2):843–854, 2010. 6

REFERENCES

[98] M Minissale, T Nguyen, and F Dulieu. **Experimental study of the penetration of oxygen and deuterium atoms into porous water ice**. *Astron. Astrophys.*, **622**:A148, 2019. 6

[99] Cristóbal Pérez, Matt T. Muckle, Daniel P. Zaleski, Nathan A. Seifert, Berhane Temelso, George C. Shields, Zbigniew Kisiel, and Brooks H. Pate. **Structures of cage, prism, and book isomers of water hexamer from broadband rotational spectroscopy**. *Science*, **336**(6083):897–901, 2012. 6

[100] Alex J Huneycutt, Ross J Stickland, Fredrik Hellberg, and Richard J Saykally. **Infrared cavity ringdown spectroscopy of acid–water clusters: HCl–H_2O, DCl–D_2O, and DCl–$(D_2O)_2$**. *J. Chem. Phys.*, **118**(3):1221–1229, 2003. 7

[101] Carl Caleman and David van der Spoel. **Evaporation from water clusters containing singly charged ions**. *Phys. Chem. Chem. Phys.*, **9**(37):5105–5111, 2007. 7

[102] M Rozenberg and A Loewenschuss. **Matrix isolation infrared spectrum of the sulfuric acid- monohydrate complex: new assignments and resolution of the "missing h-bonded ν (oh) band" issue**. *J. Phys. Chem. A*, **113**(17):4963–4971, 2009. 7

[103] Kseniia A. Korchagina, Aude Simon, Mathias Rapacioli, Fernand Spiegelman, and Jérôme Cuny. **Structural characterization of sulfur-containing water clusters using a density-functional based tight-binding approach**. *J. Phys. Chem. A*, **120**(45):9089–9100, 2016. 7, 54

[104] Isabelle Braud, Sébastien Zamith, Jérôme Cuny, Linjie Zheng, and Jean-Marc L'Hermite. **Size-dependent proton localization in hydrated uracil clusters: A joint experimental and theoretical study**. *J. Chem. Phys.*, **150**(1):014303, 2019. 7, 77, 99, 105, 109, 116, 122, 126

[105] Mauritz Johan Ryding, Alexey S Zatula, Patrik Urban Andersson, and Einar Uggerud. **Isotope exchange in reactions between D 2 O and size-selected ionic water clusters containing pyridine, H+(pyridine) m (H 2 O) n**. *Phys. Chem. Chem. Phys.*, **13**(4):1356–1367, 2011. 7

[106] Thomas Koop and Kenneth S Carslaw. **Melting of H2SO4· 4H2O particles upon cooling: Implications for polar stratospheric clouds**. *Science*, **272**(5268):1638–1641, 1996. 7

[107] Kenneth S Carslaw, Thomas Peter, and Simon L Clegg. **Modeling the composition of liquid stratospheric aerosols**. *Rev. Geophys.*, **35**(2):125–154, 1997. 7

[108] F Arnold, D Krankowsky, and KH Marien. **First mass spectrometric measurements of positive ions in the stratosphere**. *Nature*, **267**(5606):30–32, 1977. 7

[109] F Arnold, AA Viggiano, and H Schlager. **Implications for trace gases and aerosols of large negative ion clusters in the stratosphere**. *Nature*, **297**(5865):371–376, 1982. 7

[110] JD Payzant, AJ Cunningham, and P Kebarle. **Gas phase solvation of the ammonium ion by NH_3 and H_2O and stabilities of mixed clusters $nh_4^+(nh_3)_n(H_2O)_w$**. *Can. J. Chem.*, **51**(19):3242–3249, 1973. 7

[111] M Kulmala, H Vehkamäki, T Vesala, JC Barrett, and CF Clement. **Aerosol formation in diffusive boundary layer: Binary homogeneous nucleation of ammonia and water vapours**. *J. Aerosol Sci.*, **26**(4):547–558, 1995. 7

[112] Jasper Kirkby, Joachim Curtius, João Almeida, Eimear Dunne, Jonathan Duplissy, Sebastian Ehrhart, Alessandro Franchin, Stéphanie Gagné, Luisa Ickes, Andreas Kürten, et al. **Role of sulphuric acid, ammonia and galactic cosmic rays in atmospheric aerosol nucleation**. *Nature*, **476**(7361):429–433, 2011. 7, 53

[113] Eimear M Dunne, Hamish Gordon, Andreas Kürten, João Almeida, Jonathan Duplissy, Christina Williamson, Ismael K Ortega, Kirsty J Pringle, Alexey Adamov, Urs Baltensperger, et al. **Global atmospheric particle formation from CERN CLOUD measurements**. *Science*, **354**(6316):1119–1124, 2016. 7, 53

[114] A. Léger and J. L. Puget. **Identification of the 'unidentified' IR emission features of interstellar dust?** *Astron. Astrophys.*, **137**:L5–L8, 1984. 8, 131

[115] L J Allamandola, A G G M Tielens, and JRz Barker. **Polycyclic aromatic hydrocarbons and the unidentified infrared emission bands-Auto exhaust along the Milky Way**. *Astrophys. J.*, **290**:L25–L28, 1985. 8, 131

[116] JL Puget and A Léger. **A new component of the interstellar matter: Small grains and large aromatic molecules**. *Annu. Rev. Astron. Astrophys.*, **27**(1):161–198, 1989. 8

[117] Alexander G G M Tielens. **Interstellar polycyclic aromatic hydrocarbon molecules**. *Annu. Rev. Astron. Astrophys.*, **46**:289–337, 2008. 8, 131

[118] Alexander G G M Tielens. *The physics and chemistry of the interstellar medium*. Cambridge University Press, 2005. 8, 53

[119] M Rapacioli, F Calvo, C Joblin, P Parneix, D Toublanc, and F Spiegelman. **Formation and destruction of polycyclic aromatic hydrocarbon clusters in the interstellar medium**. *Astron. Astrophys.*, **460**(2):519–531, 2006. 8, 9

[120] Julien Montillaud and Christine Joblin. **Absolute evaporation rates of non-rotating neutral polycyclic aromatic hydrocarbon clusters**. *Astron. Astrophys.*, **567**:A45, 2014. 8

[121] Mathias Rapacioli and Fernand Spiegelman. **Modelling singly ionized coronene clusters**. *The European Physical Journal D*, **52**(1):55–58, 2009. 8, 55, 132

[122] Christine Joblin, Léo Dontot, GA Garcia, Fernand Spiegelman, Mathias Rapacioli, Laurent Nahon, Pascal Parneix, T Pino, and P Bréchignac. **Size effect in the ionization energy of PAH clusters**. *J. Phys. Chem. Lett.*, **8**(15):3697–3702, 2017. 8, 9, 131, 132

[123] Barbara J Finlayson-Pitts and James N Pitts Jr. **Atmospheric chemistry. Fundamentals and experimental techniques**. 1986. 8

[124] MS Callén, MT De la Cruz, JM López, R Murillo, MV Navarro, and AM Mastral. **Some inferences on the mechanism of atmospheric gas/particle partitioning of polycyclic aromatic hydrocarbons (PAH) at Zaragoza (Spain)**. *Chemosphere*, **73**(8):1357–1365, 2008. 9

[125] NA Eaves, SB Dworkin, and MJ Thomson. **The importance of reversibility in modeling soot nucleation and condensation processes**. *Proc. Combust. Inst.*, **35**(2):1787–1794, 2015. 9, 131

[126] Angela Violi and Sergei Izvekov. **Soot primary particle formation from multiscale coarse-grained molecular dynamics simulation**. *Proc. Combust. Inst.*, **31**(1):529–537, 2007. 9

[127] Reinhard Scholz, Regina Luschtinetz, Gotthard Seifert, Till Jägeler-Hoheisel, Christian Körner, Karl Leo, and Mathias Rapacioli. **Quantifying charge transfer energies at donor–acceptor interfaces in small-molecule solar cells with constrained DFTB and spectroscopic methods**. *J. Condens. Matter Phys.*, **25**(47):473201, 2013. 9

[128] Ala Aldin MHM Darghouth, Mark E Casida, Walid Taouali, Kamel Alimi, Mathias P Ljungberg, Peter Koval, Daniel Sánchez-Portal, and Dietrich Foerster. **Assessment of density-functional tight-binding ionization potentials and electron affinities of molecules of interest for organic solar cells against first-principles GW calculations**. *Computation*, **3**(4):616–656, 2015. 9

[129] Anne I S Holm, Henning Zettergren, Henrik AB Johansson, Fabian Seitz, Stefan Rosen, Henning T Schmidt, A Ławicki, Jimmy Rangama, Patrick Rousseau, Michaël Capron, et al. **Ions colliding with cold polycyclic aromatic hydrocarbon clusters**. *Phys. Rev. Lett.*, **105**(21):213401, 2010. 9, 98, 131

[130] Aude Simon, J A Noble, Guillaume Rouaut, Audrey Moudens, C Aupetit, Christophe Iftner, and Joëlle Mascetti. **Formation of coronene: water complexes: FTIR study in argon matrices and theoretical characterisation**. *Phys. Chem. Chem. Phys.*, **19**(12):8516–8529, 2017. 9, 54

[131] Tao Chen. **Formation of covalently bonded polycyclic aromatic hydrocarbons in the interstellar medium**. *Astrophys. J.*, **866**(2):113, 2018. 9, 132

[132] Martin Schmidt, Albert Masson, and Catherine Bréchignac. **Coronene cluster experiments: Stability and thermodynamics**. *International J. Mass Spectrom.*, **252**(2):173–179, 2006. 9, 131

[133] Michael Gatchell, Mark H Stockett, Nathalie de Ruette, Tao Chen, Linda Giacomozzi, Rodrigo F Nascimento, Michael Wolf, Emma K Anderson, Rudy Delaunay, Violaine Vizcaino, et al. **Failure of hydrogenation in protecting polycyclic aromatic hydrocarbons from fragmentation**. *Phys. Rev. A*, **92**(5):050702, 2015. 9, 131

[134] Michael Gatchell, Rudy Delaunay, Giovanna D'Angelo, Arkadiusz Mika, Kostiantyn Kulyk, Alicja Domaracka, Patrick Rousseau, Henning Zettergren, Bernd A Huber, and Henrik Cederquist. **Ion-induced molecular growth in clusters of small hydrocarbon chains**. *Phys. Chem. Chem. Phys.*, **19**(30):19665–19672, 2017. 9, 98, 131

[135] Sébastien Zamith, Ming-Chao Ji, Jean-Marc l'Hermite, Christine Joblin, Léo Dontot, Mathias Rapacioli, and Fernand Spiegelman. **Thermal evaporation of pyrene clusters**. *J. Chem. Phys.*, **151**(19):194303, 2019. 9, 98, 131, 132

[136] PH Dawson. **The collision-induced dissociation of protonated water clusters studied using a triple quadrupole**. *Int. J. Mass Spectrom.*, **43**(2-3):195–209, 1982. 9, 99

[137] Susan T Graul and Robert R Squires. **Energy-Resolved Collision-Induced Dissociation of Proton-Bound Cluster Ions as a Structural Probe: The Acetonitrile-Water System**. *Int. J. Mass Spectrom. Ion Process*, **94**(1-2):41–61, 1989. 9, 98

[138] S Wei, WB Tzeng, RG Keesee, and AW Castleman Jr. **Metastable Unimolecular and Collision-Induced Dissociation of Hydrogen-Bonded Clusters: Evidence for Intracluster Molecular Rearrangement and the Structure of Solvated Protonated Complexes**. *J. Am. Chem. Soc.*, **113**(6):1960–1969, 1991. 9, 98

[139] B. Liu, S. Brondsted Nielsen, P. Hvelplund, H. Zettergren, H. Cederquist, B. Manil, and B. A. Huber. **Collision-Induced Dissociation of Hydrated Adenosine Monophosphate Nucleotide Ions: Protection of the Ion in Water Nanoclusters**. *Phys. Rev. Lett.*, **97**(13):133401, SEP 29 2006. 9, 98, 99

[140] Daniel J Goebbert, Hao Chen, and Paul G Wenthold. **Collision-induced dissociation studies of protonated ether-$(H_2O)_n$ (n= 1–3) clusters**. *J. Mass Spectrom.*, **41**(2):242–247, 2006. 9, 98

[141] Rebecca A Coates and PB Armentrout. **Binding energies of hydrated cobalt (ii) by collision-induced dissociation and theoretical studies: evidence for a new critical size**. *Phys. Chem. Chem. Phys.*, **20**(2):802–818, 2018. 9, 76, 98, 99

[142] Chad C Nelson and James A McCloskey. **Collision-induced dissociation of uracil and its derivatives**. *J. Am. Soc. Mass Spectrom.*, **5**(5):339–349, 1994. 9, 76, 98

[143] Estefanía Rossich Molina, Daniel Ortiz, Jean-Yves Salpin, and Riccardo Spezia. **Elucidating collision induced dissociation products and reaction mechanisms of protonated uracil by coupling chemical dynamics simulations with tandem mass spectrometry experiments**. *J. Mass Spectrom.*, **50**(12):1340–1351, 2015. 9, 92, 98

[144] Damon R. Carl, Robert M. Moision, and P. B. Armentrout. **Binding energies for the inner hydration shells of Ca^{2+}: An experimental and theoretical investigation of $Ca^{2+}(H_2O)_x$ complexes (x=5-9)**. *Int. J. Mass Spectrom.*, **265**(2-3):308–325, SEP 1 2007. 9, 98

[145] Viet Hung Nguyen, Carlos Afonso, and Jean-Claude Tabet. **Concomitant EDD and EID of DNA evidenced by MSn and double resonance experiments**. *Int. J. Mass Spectrom.*, **301**(1-3, SI):224–233, MAR 30 2011. 9, 98

[146] Sarah C. Shuck, Kristie L. Rose, and Lawrence J. Marnett. **Mass spectrometric methods for the analysis of nucleoside–protein cross-links: application to oxopropenyl-deoxyadenosine**. *Chem. Res. Toxicol.*, **27**(1):136–146, 2014. 9, 98

[147] MC Castrovilli, P Markush, P Bolognesi, P Rousseau, S Maclot, A Cartoni, R Delaunay, A Domaracka, J Kočišek, BA Huber, and L Avaldi. **Fragmentation of pure and hydrated clusters of 5br-uracil by low energy carbon ions: observation of hydrated fragments**. *Phys. Chem. Chem. Phys.*, **19**(30):19807–19814, 2017. 9, 75, 98, 99

[148] Anupam Bera, Bruno Concina, Baptiste Schindler, Gina Predelus Renois, Gabriel Karras, Isabelle Compagnon, Atanu Bhattacharya, and Franck Lépine. **Collision induced dissociation of positive ions of dimethylnitramine, a model system for nitramine energetic molecules**. *Int. J. Mass Spectrom.*, **431**:15–21, 2018. 9, 98

[149] Stephen J Klippenstein. **Variational optimizations in the rice–ramsperger–kassel–marcus theory calculations for unimolecular dissociations with no reverse barrier**. *J. Chem. Phys.*, **96**(1):367–371, 1992. 9, 99

[150] Tomas Baer and William L Hase. *Unimolecular Reaction Dynamics: Theory and Experiments*, **31**. Oxford University Press on Demand, 1996. 9, 99

[151] Nicholas Metropolis and Stanislaw Ulam. **The Monte Carlo method**. *J. Am. Stat. Assoc.*, **44**(247):335–341, 1949. 9, 36

[152] Arthur F Voter. **Introduction to the kinetic Monte Carlo method**. In *Radiation Effects in Solids*, pages 1–23. Springer, 2007. 9

[153] Theodore Lawrence Brown. *Chemistry: the central science*. Pearson Education, 2009. 13

REFERENCES

[154] M.H. Beck, A. Jäckle, G.A. Worth, and H.-D. Meyer. **The multiconfiguration time-dependent hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets**. *Phys. Rep.*, **324**(1):1–105, 2000. 14

[155] David J Griffiths and Darrell F Schroeter. *Introduction to quantum mechanics*. Cambridge University Press, 2018. 15

[156] Max Born and Robert Oppenheimer. **Zur quantentheorie der molekeln**. *Ann. Phys.*, **389**(20):457–484, 1927. 17, 18

[157] Saul T Epstein. **Ground-state energy of a molecule in the adiabatic approximation**. *J. Chem. Phys.*, **44**(2):836–837, 1966. 18

[158] Laurie J Butler. **Chemical reaction dynamics beyond the Born-Oppenheimer approximation**. *Ann. Rev. Phys. Chem.*, **49**(1):125–171, 1998. 18

[159] Cristina Puzzarini, John F Stanton, and Juergen Gauss. **Quantum-chemical calculation of spectroscopic parameters for rotational spectroscopy**. *Int. Rev. Phys. Chem.*, **29**(2):273–367, 2010. 19

[160] Trygve Helgaker, Michał Jaszuński, and Kenneth Ruud. **Ab initio methods for the calculation of NMR shielding and indirect spin-spin coupling constants**. *Chem. Rev.*, **99**(1):293–352, 1999. 19

[161] Alfonso Pedone, Malgorzata Biczysko, and Vincenzo Barone. **Environmental effects in computational spectroscopy: Accuracy and interpretation**. *ChemPhysChem*, **11**(9):1812–1832, 2010. 19

[162] Graham Fleming. **Chemical applications of ultrafast spectroscopy**. 1986. 19

[163] Jaan Laane, Hiroaki Takahashi, and Andre Bandrauk. *Structure and Dynamics of Electronic Excited States*. Springer Science & Business Media, 1999. 19

[164] Trygve Helgaker, Sonia Coriani, Poul Jørgensen, Kasper Kristensen, Jeppe Olsen, and Kenneth Ruud. **Recent advances in wave function-based methods of molecular-property calculations**. *Chem. Rev.*, **112**(1):543–631, 2012. 19

[165] Frank Jensen. *Introduction to computational chemistry*. John wiley & sons, 2017. 19, 21

[166] Dirk Porezag, Th Frauenheim, Th Köhler, Gotthard Seifert, and R Kaschner. **Construction of tight-binding-like potentials on the basis of density-functional theory: application to carbon**. *Phys. Rev. B*, **51**(19):12947–12958, 1995. 19

[167] G Seifert, D Porezag, and Th Frauenheim. **Calculations of molecules, clusters, and solids with a simplified LCAO-DFT-LDA scheme**. *Int. J. Quantum Chem.*, **58**(2):185–192, 1996. 19

[168] Marcus Elstner, Dirk Porezag, G Jungnickel, J Elsner, M Haugk, Th Frauenheim, Sandor Suhai, and Gotthard Seifert. **Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties**. *Phys. Rev. B*, **58**(11):7260, 1998. 19, 26, 50, 102, 132

[169] Marcus Elstner and Gotthard Seifert. **Density functional tight binding**. *Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci.*, **372**(2011):20120483, 2014. 19, 102

[170] U Chandra Singh and Peter A Kollman. **A combined ab initio quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the $CH_3Cl + Cl^-$ exchange reaction and gas phase protonation of polyethers**. *J. Comput. Chem*, **7**(6):718–730, 1986. 19

[171] Jiali Gao. **Methods and applications of combined quantum mechanical and molecular mechanical potentials**. *Rev. Comput. Chem.*, **7**:119–186, 1996. 19

[172] Tiziana Z Mordasini and Walter Thiel. **Combined quantum mechanical and molecular mechanical approaches**. *Chimia*, **52**(6):288–291, 1998. 19

[173] D R Hartree. **The waves mechanics of an atom with non-coulombic central field: parts I, II, III**. In *Proc. Cambridge Phil. Soc*, **24**, pages 89–132, 1928. 19

[174] D R Hartree. **The calculation of atomic structures**. *Rep. Prog. Phys.*, **11**:113–143, 1947. 19

[175] John Clarke Slater. **Quantum theory of matter**. 1968. 19

[176] Roy McWeeny. *Methods of molecular quantum mechanics*. Academic press, 1992. 20

[177] A Szabo and N S Ostlund. **Introduction to advanced electronic structure theory**. *Modern Quantum Chemistry*, 1996. 20

[178] Jorge Kohanoff. *Electronic structure calculations for solids and molecules: theory and computational methods*. Cambridge University Press, 2006. 20

[179] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2014. 20

[180] Clemens Carel Johannes Roothaan. **New developments in molecular orbital theory**. *Rev. Mod. Phys.*, **23**(2):69–89, 1951. 20

[181] Seymour Michael Blinder and James E House. *Mathematical Physics in Theoretical Chemistry*. Elsevier, 2018. 20

[182] R Koch and T Clark. *The chemist's electronic book of orbitals*. Springer, 1999. 20

[183] JW Moskowitz and MC Harrison. **Gaussian Wavefunctions for the 10-Electron Systems. III. OH^-, H_2O, H_3O^+**. *J. Chem. Phys.*, **43**(10):3550–3555, 1965. 21

[184] Per-Olov Löwdin. **The historical development of the electron correlation problem**. *Int. J. Quantum Chem.*, **55**(2):77–102, 1995. 21

[185] Christopher J Cramer and FM Bickelhaupt. **Essentials of computational chemistry**. *Angew. Chem. Int. Ed.*, **42**(4):381–381, 2003. 21

[186] Martin Head-Gordon, Rudolph J Rico, Manabu Oumi, and Timothy J Lee. **A doubles correction to electronic excited states from configuration interaction in the space of single substitutions**. *Chem. Phys. Lett.*, **219**(1-2):21–29, 1994. 21

[187] David Maurice and Martin Head-Gordon. **Analytical second derivatives for excited electronic states using the single excitation configuration interaction method: theory and application to benzo [a] pyrene and chalcone**. *Mol. Phys.*, **96**(10):1533–1541, 1999. 21

[188] Chr Møller and Milton S Plesset. **Note on an approximation treatment for many-electron systems**. *Phys. Rev.*, **46**(7):618, 1934. 21

[189] John A Pople, J Stephen Binkley, and Rolf Seeger. **Theoretical models incorporating electron correlation**. *Int. J. Quantum Chem.*, **10**(S10):1–19, 1976. 21

[190] R Krishnan and John A Pople. **Approximate fourth-order perturbation theory of the electron correlation energy**. *Int. J. Quantum Chem.*, **14**(1):91–100, 1978. 21

[191] JIŘÍ ČÍŽEK. **On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods**. *J. Chem. Phys*, **45**(11):4256–4266, 1966. 21

[192] GEORGE D PURVIS III AND RODNEY J BARTLETT. **A full coupled-cluster singles and doubles model: The inclusion of disconnected triples**. *J. Chem. Phys*, **76**(4):1910–1918, 1982. 21

[193] KRISHNAN RAGHAVACHARI, GARY W TRUCKS, JOHN A POPLE, AND MARTIN HEAD-GORDON. **A fifth-order perturbation comparison of electron correlation theories**. *Chem. Phys. Lett.*, **157**(6):479–483, 1989. 21

[194] TROY VAN VOORHIS AND MARTIN HEAD-GORDON. **Two-body coupled cluster expansions**. *J. Chem. Phys*, **115**(11):5033–5040, 2001. 21

[195] LARRY A CURTISS, KRISHNAN RAGHAVACHARI, GARY W TRUCKS, AND JOHN A POPLE. **Gaussian-2 theory for molecular energies of first- and second-row compounds**. *J. Chem. Phys*, **94**(11):7221–7230, 1991. 21

[196] LARRY A CURTISS, KRISHNAN RAGHAVACHARI, PAUL C REDFERN, VITALY RASSOLOV, AND JOHN A POPLE. **Gaussian-3 (G3) theory for molecules containing first and second-row atoms**. *J. Chem. Phys*, **109**(18):7764–7776, 1998. 21

[197] WILLIAM S OHLINGER, PHILIP E KLUNZINGER, BERNARD J DEPPMEIER, AND WARREN J HEHRE. **Efficient calculation of heats of formation**. *J. Phys. Chem. A*, **113**(10):2165–2175, 2009. 21

[198] ROY MCWEENY. *Post-Hartree-Fock Methods*, chapter 8, pages 133–139. John Wiley & Sons,Ltd. 21

[199] HERBERT DACOSTA AND MAOHONG FAN. *Rate constant calculation for thermal reactions*. Wiley Online Library, 2011. 21

[200] WALTER KOHN AND LU JEU SHAM. **Self-consistent equations including exchange and correlation effects**. *Physical review*, **140**(4A):A1133, 1965. 21, 22, 23, 25

[201] WALTER KOHN. **Nobel Lecture: Electronic structure of matter—wave functions and density functionals**. *Rev. Mod. Phys.*, **71**(5):1253, 1999. 21

[202] PIERRE HOHENBERG AND WALTER KOHN. **Inhomogeneous electron gas**. *Physical Review*, **136**(3B):B864, 1964. 22

[203] MEL LEVY. **Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem**. *Proc. Natl. Acad. Sci.*, **76**(12):6062–6065, 1979. 22

[204] LLEWELLYN H THOMAS. **The calculation of atomic fields**. *Proc. Camb. Philos. Soc.*, **23**(5):542–548, 1927. 22

[205] ENRICO FERMI. **Eine statistische methode zur bestimmung einiger eigenschaften des atoms und ihre anwendung auf die theorie des periodischen systems der elemente**. *Zeitschrift für Physik*, **48**(1-2):73–79, 1928. 22

[206] FRANK HERMAN, JOHN P VAN DYKE, AND IRENE B ORTENBURGER. **Improved statistical exchange approximation for inhomogeneous many-electron systems**. *Phys. Rev. Lett.*, **22**(16):807–811, 1969. 25

[207] JOHN P PERDEW AND KIERON BURKE. **Comparison shopping for a gradient-corrected density functional**. *Int. J. Quantum Chem.*, **57**(3):309–319, 1996. 25

[208] AXEL D BECKE. **Density functional thermochemistry. III. The role of exact exchange**. *J. Chem. Phys.*, **98**(492):5648–5652, 1993. 25, 26

[209] AXEL D BECKE. **Density-functional exchange-energy approximation with correct asymptotic behavior**. *Phys. Rev. A*, **38**(6):3098, 1988. 26

[210] CHENGTEH LEE, WEITAO YANG, AND ROBERT G PARR. **Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density**. *Phys. Rev. B*, **37**(2):785–789, 1988. 26

[211] PHILIP J STEPHENS, FRANK J DEVLIN, CARY F CHABALOWSKI, AND MICHAEL J FRISCH. **Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields**. *J. Phys. Chem.*, **98**(45):11623–11627, 1994. 26

[212] CARLO ADAMO AND VINCENZO BARONE. **Toward reliable density functional methods without adjustable parameters: The PBE0 model**. *J. Chem. Phys.*, **110**(13):6158–6170, 1999. 26

[213] THOMAS M HENDERSON, ARTUR F IZMAYLOV, GIOVANNI SCALMANI, AND GUSTAVO E SCUSERIA. **Can short-range hybrids describe long-range-dependent properties?** *J. Chem. Phys.*, **131**(4):044108, 2009. 26

[214] ALIAKSANDR V KRUKAU, OLEG A VYDROV, ARTUR F IZMAYLOV, AND GUSTAVO E SCUSERIA. **Influence of the exchange screening parameter on the performance of screened hybrid functionals**. *J. Chem. Phys.*, **125**(22):224106, 2006. 26

[215] YAN ZHAO AND DONALD G TRUHLAR. **The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals**. *Theor. Chem. Acc.*, **120**(1-3):215–241, 2008. 26

[216] CHEN LI, XIAO ZHENG, NEIL QIANG SU, AND WEITAO YANG. **Localized orbital scaling correction for systematic elimination of delocalization error in density functional approximations**. *Natl. Sci. Rev.*, **5**(2):203–215, 2018. 26

[217] W MATTHEW C FOULKES AND ROGER HAYDOCK. **Tight-binding models and density-functional theory**. *Phys. Rev. B*, **39**(17):12520, 1989. 26

[218] DIRK POREZAG, TH FRAUENHEIM, TH KÖHLER, GOTTHARD SEIFERT, AND R KASCHNER. **Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon**. *Phys. Rev. B*, **51**(19):12947, 1995. 26, 102, 132

[219] G SEIFERT, D POREZAG, AND TH FRAUENHEIM. **Calculations of molecules, clusters, and solids with a simplified LCAO-DFT-LDA scheme**. *Int. J. Quantum Chem.*, **58**(2):185–192, 1996. 26, 102, 132

[220] MARCUS ELSTNER, T FRAUENHEIM, J MCKELVEY, AND G SEIFERT. **Density functional tight binding: Contributions from the american chemical society symposium**. *J. Phys. Chem. A*, **111**(26):5607–5608, 2007. 26

[221] AUGUSTO F OLIVEIRA, GOTTHARD SEIFERT, THOMAS HEINE, AND HÉLIO A DUARTE. **Density-functional based tight-binding: an approximate DFT method**. *J. Braz. Chem. Soc., volume=20, number=7, pages=1193–1205, year=2009, publisher=SciELO Brasil*. 26

[222] MICHAEL GAUS, QIANG CUI, AND MARCUS ELSTNER. **DFTB3: extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB)**. *J. Chem. Theory Comput.*, **7**(4):931–948, 2011. 27

[223] ROBERT S MULLIKEN. **Electronic population analysis on LCAO–MO molecular wave functions. I**. *J. Chem. Phys.*, **23**(10):1833–1840, 1955. 30, 31

REFERENCES

[224] RICHARD F. W. BADER. **Atoms in molecules**. *Acc. Chem. Res.*, **18**(1):9–15, 1985. 32

[225] RICHARD F W BADER. *Atoms in molecules: a quantum theory*. Oxford: Clarendon press, 1990. 32

[226] ERIC D GLENDENING AND FRANK WEINHOLD. **Natural resonance theory: II. Natural bond order and valency**. *J. Comput. Chem*, **19**(6):610–627, 1998. 32

[227] ERIC D GLENDENING, CLARK R LANDIS, AND FRANK WEINHOLD. **Natural bond orbital methods**. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2**(1):1–42, 2012. 32

[228] U CHANDRA SINGH AND PETER A KOLLMAN. **An approach to computing electrostatic charges for molecules**. *J. Comput. Chem*, **5**(2):129–145, 1984. 32

[229] BRENT H BESLER, KENNETH M MERZ JR, AND PETER A KOLLMAN. **Atomic charges derived from semiempirical methods**. *J. Comput. Chem*, **11**(4):431–439, 1990. 32

[230] JIABO LI, TIANHAI ZHU, CHRISTOPHER J CRAMER, AND DONALD G TRUHLAR. **New class IV charge model for extracting accurate partial charges from wave functions**. *J. Phys. Chem. A*, **102**(10):1820–1831, 1998. 32, 50

[231] JAROSLAW A KALINOWSKI, BOGDAN LESYNG, JASON D THOMPSON, CHRISTOPHER J CRAMER, AND DONALD G TRUHLAR. **Class IV charge model for the self-consistent charge density-functional tight-binding method**. *J. Phys. Chem. A*, **108**(13):2545–2549, 2004. 32

[232] CASEY P KELLY, CHRISTOPHER J CRAMER, AND DONALD G TRUHLAR. **Accurate partial atomic charges for high-energy molecules using class IV charge models with the MIDI! basis set**. *Theor. Chem. Acc.*, **113**(3):133–151, 2005. 32

[233] MATHIAS RAPACIOLI, FERNAND SPIEGELMAN, DAHBIA TALBI, TZONKA MINEVA, ANNICK GOURSOT, THOMAS HEINE, AND GOTTHARD SEIFERT. **Correction for dispersion and Coulombic interactions in molecular clusters with density functional derived methods: Application to polycyclic aromatic hydrocarbon clusters**. *J. Chem. Phys.*, **130**(24):244304, 2009. 32, 50, 132

[234] ISTVAN MAYER. **Charge, bond order and valence in the AB initio SCF theory**. *Chem. Phys. Lett.*, **97**(3):270–274, 1983. 32

[235] E CANCES, BENEDETTA MENNUCCI, AND J TOMASI. **A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics**. *J. Chem. Phys.*, **107**(8):3032–3041, 1997. 32

[236] TIAN LU AND FEIWU CHEN. **Multiwfn: a multifunctional wavefunction analyzer**. *J. Comput. Chem*, **33**(5):580–592, 2012. 32

[237] JAMES P LEWIS AND OTTO F SANKEY. **Geometry and energetics of DNA basepairs and triplets from first principles quantum molecular relaxations**. *Biophys. J.*, **69**(3):1068–1076, 1995. 32

[238] FA GIANTURCO, F PAESANI, MF LARANJEIRA, V VASSILENKO, AND MA CUNHA. **Intermolecular forces from density functional theory. III. A multiproperty analysis for the Ar(^{1}S)-CO($^1\Sigma$) interaction**. *J. Chem. Phys.*, **110**(16):7832–7845, 1999. 32

[239] MARCUS ELSTNER, PAVEL HOBZA, THOMAS FRAUENHEIM, SÁNDOR SUHAI, AND EFTHIMIOS KAXIRAS. **Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density-functional-theory based treatment**. *J. Chem. Phys.*, **114**(12):5149–5155, 2001. 32, 50

[240] QIN WU AND WEITAO YANG. **Empirical correction to density functional theory for van der Waals interactions**. *J. Chem. Phys.*, **116**(2):515–524, 2002. 32

[241] STEFAN GRIMME. **Accurate description of van der Waals complexes by density functional theory including empirical corrections**. *J. Comput. Chem*, **25**(12):1463–1473, 2004. 32

[242] URS ZIMMERLI, MICHELE PARRINELLO, AND PETROS KOUMOUTSAKOS. **Dispersion corrections to density functionals for water aromatic interactions**. *J. Chem. Phys.*, **120**(6):2693–2699, 2004. 32

[243] MARCUS A NEUMANN AND MARC-ANTOINE PERRIN. **Energy ranking of molecular crystals using density functional theory calculations and an empirical van der Waals correction**. *J. Phys. Chem. B*, **109**(32):15531–15541, 2005. 32

[244] ANNICK GOURSOT, TZONKA MINEVA, RUSLAN KEVORKYANTS, AND DAHBIA TALBI. **Interaction between n-alkane chains: Applicability of the empirically corrected density functional theory for van der Waals complexes**. *J. Chem. Theory Comput.*, **3**(3):755–763, 2007. 32

[245] DAAN FRENKEL AND BEREND SMIT. *Understanding molecular simulation: from algorithms to applications*, **1**. Elsevier, 2001. 33, 40

[246] JOHN EDWARD JONES. **On the determination of molecular fields.—I. From the variation of the viscosity of a gas with temperature**. *Proc. Math. Phys. Eng. Sci.*, **106**(738):441–462, 1924. 33, 34

[247] JOHN EDWARD LENNARD-JONES. **On the determination of molecular fields. II. From the equation of state of gas**. *Proc. Roy. Soc. A*, **106**:463–477, 1924. 33, 34

[248] PHILIP M MORSE. **Diatomic molecules according to the wave mechanics. II. Vibrational levels**. *Physical review*, **34**(1):57–64, 1929. 33, 34

[249] MAX BORN AND JOSEPH E MAYER. **Zur gittertheorie der ionenkristalle**. *Zeitschrift für Physik*, **75**(1-2):1–18, 1932. 33

[250] FRANK H STILLINGER AND THOMAS A WEBER. **Computer simulation of local order in condensed phases of silicon**. *Phys. Rev. B*, **31**(8):5262–5271, 1985. 33

[251] ILYA G KAPLAN. *Intermolecular interactions: physical picture, computational methods and model potentials*. John Wiley & Sons, 2006. 33

[252] IVET BAHAR, ALI RANA ATILGAN, AND BURAK ERMAN. **Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential**. *Fold Des.*, **2**(3):173–181, 1997. 33

[253] LOUIS A GIRIFALCO AND VICTOR G WEIZER. **Application of the Morse potential function to cubic metals**. *Physical Review*, **114**(3):687–690, 1959. 34

[254] JOHN EDWARD LENNARD-JONES. **Cohesion**. *Proceedings of the Physical Society*, **43**(5):461–482, 1931. 34

[255] LUCA MONTICELLI AND D PETER TIELEMAN. **Force fields for classical molecular dynamics**. In *Biomolecular simulations*, pages 197–213. Springer, 2013. 34

[256] N METROPOLIS. **The beginning of the Monte Carlo method**. *Los Alamos Science*, **15**:125–130, 1987. 36

[257] DIRK P KROESE, TIM BRERETON, THOMAS TAIMRE, AND ZDRAVKO I BOTEV. **Why the Monte Carlo method is so important today**. *Wiley Interdisciplinary Reviews: Computational Statistics*, **6**(6):386–392, 2014. 36

[258] MARSHALL N ROSENBLUTH AND ARIANNA W ROSENBLUTH. **Monte Carlo calculation of the average extension of molecular chains**. *J. Chem. Phys.*, **23**(2):356–359, 1955. 37

[259] KURT BINDER, DIETER HEERMANN, LYLE ROELOFS, A JOHN MALLINCKRODT, AND SUSAN MCKAY. **Monte Carlo simulation in statistical physics**. *Comput. Phys.*, **7**(2):156–157, 1993. 37

[260] STEPHAN A BAEURLE. **Multiscale modeling of polymer materials using field-theoretic methodologies: a survey about recent developments**. *J. Math. Chem.*, **46**(2):363–426, 2009. 37

[261] BERNI JULIAN ALDER AND THOMAS EVERETT WAINWRIGHT. **Phase transition for a hard sphere system**. *J. Chem. Phys.*, **27**(5):1208–1209, 1957. 40

[262] JB GIBSON, AN N GOLAND, M MILGRAM, AND GOHO VINEYARD. **Dynamics of radiation damage**. *Physical Review*, **120**(4):1229–1253, 1960. 40

[263] ANEESUR RAHMAN. **Correlations in the motion of atoms in liquid argon**. *Physical review*, **136**(2A):A405–A411, 1964. 40

[264] MICHAEL P ALLEN AND DOMINIC J TILDESLEY. *Computer simulation of liquids*. Oxford university press, 2017. 40

[265] G CICCCOTTI, D FRENKEL, AND IR MCDONALD. **Simulations of liquids and solids–molecular dynamics and monte carlo methods in statistical mechanics**, 1987. 40

[266] WILFRED F VAN GUNSTEREN AND HERMAN JC BERENDSEN. **Computer simulation of molecular dynamics: methodology, applications, and perspectives in chemistry**. *Angew. Chem. Int. Ed.*, **29**(9):992–1023, 1990. 40

[267] BRUCE J BERNE, GIOVANNI CICCOTTI, AND DAVID F COKER. *Classical and Quantum Dynamics in Condensed Phase Simulations: Proceedings of the International School of Physics*. World Scientific, 1998. 40

[268] MARK E TUCKERMAN AND GLENN J MARTYNA. **Understanding modern molecular dynamics: techniques and applications**, 2000. 40

[269] MARTIN KARPLUS AND J ANDREW MCCAMMON. **Molecular dynamics simulations of biomolecules**. *Nat. Struct. Biol.*, **9**(9):646–652, 2002. 40

[270] DENNIS C RAPAPORT. *The art of molecular dynamics simulation*. Cambridge university press, 2004. 40

[271] PAUL TAVAN, HEIKO CARSTENS, AND GERALD MATHIAS. **Molecular dynamics simulations of proteins and peptides: Problems, achievements, and perspectives**. *Protein Folding Handbook*, pages 1170–1195, 2005. 40

[272] WILFRED F VAN GUNSTEREN, DIRK BAKOWIES, RICCARDO BARON, INDIRA CHANDRASEKHAR, MARKUS CHRISTEN, XAVIER DAURA, PETER GEE, DAAN P GEERKE, ALICE GLÄTTLI, PHILIPPE H HÜNENBERGER, ET AL. **Biomolecular modeling: goals, problems, perspectives**. *Angew. Chem. Int. Ed.*, **45**(25):4064–4092, 2006. 40

[273] MICHELE PARRINELLO AND ANEESUR RAHMAN. **Crystal structure and pair potentials: A molecular-dynamics study**. *Phys. Rev. Lett.*, **45**(14):1196–1199, 1980. 40

[274] PIETRO BALLONE, WANDA ANDREONI, ROBERTO CAR, AND MICHELE PARRINELLO. **Equilibrium structures and finite temperature properties of silicon microclusters from ab initio molecular-dynamics calculations**. *Phys. Rev. Lett.*, **60**(4):271–274, 1988. 40

[275] RAFAEL GUTIERREZ, THOMAS FRAUENHEIM, THOMAS KÖHLER, AND GOTHARD SEIFERT. **Stability of silicon carbide structures: from clusters to solid surfaces**. *J. Mater. Chem. A .*, **6**(10):1657–1663, 1996. 40

[276] CL BRIANT AND JJ BURTON. **Molecular dynamics study of the structure and thermodynamic properties of argon microclusters**. *J. Chem. Phys.*, **63**(5):2045–2058, 1975. 40

[277] JOHAN PM POSTMA, HERMAN JC BERENDSEN, AND JAN R HAAK. **Thermodynamics of cavity formation in water. A molecular dynamics study**. In *Faraday Symposia of the Chemical Society*, **17**, pages 55–67. Royal Society of Chemistry, 1982. 40

[278] J DANA HONEYCUTT AND HANS C ANDERSEN. **Molecular dynamics study of melting and freezing of small Lennard-Jones clusters**. *Journal of Physical Chemistry*, **91**(19):4950–4963, 1987. 40

[279] LIEM X DANG AND TSUN-MEI CHANG. **Molecular dynamics study of water clusters, liquid, and liquid–vapor interface of water with many-body potentials**. *J. Chem. Phys.*, **106**(19):8149–8159, 1997. 40

[280] PETER V COVENEY AND SHUNZHOU WAN. **On the calculation of equilibrium thermodynamic properties from molecular dynamics**. *Phys. Chem. Chem. Phys.*, **18**(44):30236–30240, 2016. 40

[281] SG CHARATI AND SA STERN. **Diffusion of gases in silicone polymers: molecular dynamics simulations**. *Macromolecules*, **31**(16):5529–5535, 1998. 40

[282] IN-CHUL YEH AND GERHARD HUMMER. **System-size dependence of diffusion coefficients and viscosities from molecular dynamics simulations with periodic boundary conditions**. *J. Phys. Chem. B*, **108**(40):15873–15879, 2004. 40

[283] CARLOS BRAGA, AMPARO GALINDO, AND ERICH A MÜLLER. **Nonequilibrium molecular dynamics simulation of diffusion at the liquid-liquid interface**. *J. Chem. Phys.*, **141**(15):154101, 2014. 40

[284] GAURAV PRANAMI AND MONICA H LAMM. **Estimating error in diffusion coefficients derived from molecular dynamics simulations**. *J. Chem. Theory Comput.*, **11**(10):4586–4592, 2015. 40

[285] MAURIZIO MONDELLO AND GARY S GREST. **Viscosity calculations of n-alkanes by equilibrium molecular dynamics**. *J. Chem. Phys.*, **106**(22):9327–9336, 1997. 40

[286] BERK HESS. **Determining the shear viscosity of model liquids from molecular dynamics simulations**. *J. Chem. Phys.*, **116**(1):209–217, 2002. 40

[287] YONG ZHANG, AKIHITO OTANI, AND EDWARD J MAGINN. **Reliable viscosity calculation from equilibrium molecular dynamics simulations: A time decomposition method**. *J. Chem. Theory Comput.*, **11**(8):3537–3546, 2015. 40

[288] SR ARAGON AND R PECORA. **Fluorescence correlation spectroscopy as a probe of molecular dynamics**. *J. Chem. Phys.*, **64**(4):1791–1803, 1976. 40

[289] RODOLPHE VUILLEUMIER AND DANIEL BORGIS. **Transport and spectroscopy of the hydrated proton: a molecular dynamics study**. *J. Chem. Phys.*, **111**(9):4251–4266, 1999. 40

[290] MICHELE PAVONE, GIUSEPPE BRANCATO, GIOVANNI MORELLI, AND VINCENZO BARONE. **Spectroscopic properties in the liquid phase: combining high-level ab initio calculations and classical molecular dynamics**. *ChemPhysChem*, **7**(1):148–156, 2006. 40

[291] GIUSEPPE BRANCATO, NADIA REGA, AND VINCENZO BARONE. **A quantum mechanical/molecular dynamics/mean field study of acrolein in aqueous solution: analysis of H bonding and bulk effects on spectroscopic properties**. *J. Chem. Phys.*, **125**(16):164515, 2006. 40

[292] MICHAEL H BECK, ANDREAS JÄCKLE, GRAHAM A WORTH, AND H-D MEYER. **The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets**. *Phys. Rep.*, **324**(1):1–105, 2000. 41

[293] HAOBIN WANG AND MICHAEL THOSS. **Multilayer formulation of the multiconfiguration time-dependent Hartree theory**. *J. Chem. Phys.*, **119**(3):1289–1299, 2003. 41

[294] HUA WEI AND TUCKER CARRINGTON JR. **Discrete variable representations of complicated kinetic energy operators**. *J. Chem. Phys.*, **101**(2):1343–1360, 1994. 41

[295] JOHN C LIGHT AND TUCKER CARRINGTON JR. **Discrete-variable representations and their utilization**. *Advances in Chemical Physics*, **114**:263–310, 2000. 41

[296] JOHN T BATINA. **Unsteady Euler algorithm with unstructured dynamic mesh for complex-aircraft aerodynamic analysis**. *AIAA J.*, **29**(3):327–333, 1991. 42

[297] JOHN CHARLES BUTCHER AND NICOLETTE GOODWIN. *Numerical methods for ordinary differential equations*, **2**. Wiley Online Library, 2008. 42

[298] CHARLES WILLIAM GEAR. **The numerical integration of ordinary differential equations of various orders**. Technical report, Argonne National Lab., Ill., 1966. 42

[299] KAI DIETHELM, NEVILLE J FORD, AND ALAN D FREED. **A predictor-corrector approach for the numerical solution of fractional differential equations**. *Nonlinear Dyn.*, **29**(1-4):3–22, 2002. 42

[300] CARL RUNGE. **Über die numerische Auflösung von Differentialgleichungen**. *Math. Ann*, **46**(2):167–178, 1895. 42

[301] WILHELM KUTTA. **Beitrag zur naherungsweisen Integration totaler Differentialgleichungen**. *Z. Math. Phys.*, **46**:435–453, 1901. 42

[302] JOHN CHARLES BUTCHER. **A history of Runge-Kutta methods**. *Appl. Numer. Math.*, **20**(3):247–260, 1996. 42

[303] JC BUTCHER. **Runge–Kutta methods for ordinary differential equations**. In *Numerical Analysis and Optimization*, pages 37–58. Springer, 2015. 42

[304] LOUP VERLET. **Computer" experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules**. *Physical review*, **159**(1):98–103, 1967. 42

[305] FY FRAIGE AND PA LANGSTON. **Integration schemes and damping algorithms in distinct element models**. *Advanced Powder Technology*, **15**(2):227–245, 2004. 42

[306] WILLIAM C SWOPE, HANS C ANDERSEN, PETER H BERENS, AND KENT R WILSON. **A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters**. *J. Chem. Phys.*, **76**(1):637–649, 1982. 42

[307] JOHN R RAY AND HW GRABEN. **Direct calculation of fluctuation formulae in the microcanonical ensemble**. *Mol. Phys.*, **43**(6):1293–1297, 1981. 43

[308] JOHN R RAY AND HONGWEI ZHANG. **Correct microcanonical ensemble in molecular dynamics**. *Phys. Rev. E*, **59**(5):4781–4785, 1999. 43

[309] SHUICHI NOSÉ. **A unified formulation of the constant temperature molecular dynamics methods**. *J. Chem. Phys.*, **81**(1):511–519, 1984. 44, 103

[310] SHŪICHI NOSÉ. **A molecular dynamics method for simulations in the canonical ensemble**. *Mol. Phys.*, **52**(2):255–268, 1984. 44, 51

[311] WILLIAM G HOOVER. **Canonical dynamics: Equilibrium phase-space distributions**. *Phys. Rev. A*, **31**(3):1695–1697, 1985. 44, 51, 103

[312] PHILIPPE H HÜNENBERGER. **Thermostat algorithms for molecular dynamics simulations**. In *Advanced computer simulation*, pages 105–149. Springer, 2005. 44

[313] GLENN M TORRIE AND JOHN P VALLEAU. **Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling**. *J. Comput. Phys.*, **23**(2):187–199, 1977. 44

[314] ULRICH HE HANSMANN AND YUKO OKAMOTO. **Prediction of peptide conformation by multicanonical algorithm: New approach to the multiple-minima problem**. *J. Comput. Chem*, **14**(11):1333–1338, 1993. 44

[315] MASSIMO MARCHI AND PIETRO BALLONE. **Adiabatic bias molecular dynamics: a method to navigate the conformational space of complex molecular systems**. *J. Chem. Phys.*, **110**(8):3697–3702, 1999. 44

[316] CHRISTIAN BARTELS. **Analyzing biased Monte Carlo and molecular dynamics simulations**. *Chem. Phys. Lett.*, **331**(5-6):446–454, 2000. 44

[317] ERIC DARVE AND ANDREW POHORILLE. **Calculating free energies using average force**. *J. Chem. Phys.*, **115**(20):9169–9183, 2001. 44

[318] ALESSANDRO LAIO AND MICHELE PARRINELLO. **Escaping free-energy minima**. *Proc. Natl. Acad. Sci.*, **99**(20):12562–12566, 2002. 44

[319] MARCELLA IANNUZZI, ALESSANDRO LAIO, AND MICHELE PARRINELLO. **Efficient exploration of reactive potential energy surfaces using Car-Parrinello molecular dynamics**. *Phys. Rev. Lett.*, **90**(23):238302, 2003. 44

[320] ALESSANDRO BARDUCCI, MASSIMILIANO BONOMI, AND MICHELE PARRINELLO. **Metadynamics**. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **1**(5):826–843, 2011. 44

[321] SCOTT KIRKPATRICK, C DANIEL GELATT, AND MARIO P VECCHI. **Optimization by simulated annealing**. *science*, **220**(4598):671–680, 1983. 44

[322] PETER JM VAN LAARHOVEN AND EMILE HL AARTS. **Simulated annealing**. In *Simulated annealing: Theory and applications*, pages 7–15. Springer, 1987. 44

[323] ROBERT H SWENDSEN AND JIAN-SHENG WANG. **Replica Monte Carlo simulation of spin-glasses**. *Phys. Rev. Lett.*, **57**(21):2607, 1986. 45

[324] CHARLES J GEYER ET AL. **Computing science and statistics: Proceedings of the 23rd Symposium on the Interface**. *American Statistical Association, New York*, **156**, 1991. 45

[325] KOJI HUKUSHIMA AND KOJI NEMOTO. **Exchange Monte Carlo method and application to spin glass simulations**. *J. Phys. Soc. Jpn.*, **65**(6):1604–1608, 1996. 45

[326] MARCO FALCIONI AND MICHAEL W DEEM. **A biased Monte Carlo scheme for zeolite structure solution**. *J. Chem. Phys.*, **110**(3):1754–1766, 1999. 45

[327] DAVID J EARL AND MICHAEL W DEEM. **Parallel tempering: Theory, applications, and new perspectives**. *Phys. Chem. Chem. Phys.*, **7**(23):3910–3916, 2005. 45, 50

[328] YUJI SUGITA AND YUKO OKAMOTO. **Replica-exchange molecular dynamics method for protein folding**. *Chem. Phys. Lett.*, **314**(1-2):141–151, 1999. 45, 50

[329] BERND HARTKE. **Global geometry optimization of clusters using genetic algorithms**. *J. Phys. Chem.*, **97**(39):9973–9976, 1993. 47

[330] RON UNGER AND JOHN MOULT. **Genetic algorithms for protein folding simulations**. *J. Mol. Biol., volume=231, number=1, pages=75–81, year=1993, publisher=Elsevier*. 47

[331] SN SIVANANDAM AND SN DEEPA. **Genetic algorithms**. In *Introduction to genetic algorithms*, pages 15–37. Springer, 2008. 47

[332] CLAUDIO FABIANO MOTTA TOLEDO, L OLIVEIRA, AND PAULO MORELATO FRANÇA. **Global optimization using a genetic algorithm with hierarchically structured population**. *J. Comput. Appl. Math.*, **261**:341–351, 2014. 47

[333] T HEINE, M RAPACIOLI, S PATCHKOVSKII, J FRENZEL, A KOSTER, P CALAMINICI, HA DUARTE, S ESCALANTE, R FLORES-MORENO, AND A GOURSOT. **deMonNano**, 2009. 50

[334] MICHAEL GAUS, ALBRECHT GOEZ, AND MARCUS ELSTNER. **Parametrization and benchmark of DFTB3 for organic molecules**. *J. Chem. Theory Comput.*, **9**(1):338–354, 2013. 50, 55, 58

[335] JASON D THOMPSON, CHRISTOPHER J CRAMER, AND DONALD G TRUHLAR. **Parameterization of charge model 3 for AM1, PM3, BLYP, and B3LYP**. *J. Comput. Chem.*, **24**(11):1291–1304, 2003. 50, 55

[336] LYUBEN ZHECHKOV, THOMAS HEINE, SERGUEI PATCHKOVSKII, GOTTHARD SEIFERT, AND HELIO A DUARTE. **An efficient a posteriori treatment for dispersion interaction in density-functional-based tight binding**. *J. Chem. Theory Comput.*, **1**(5):841–847, 2005. 50

[337] AUDE SIMON, MATHIAS RAPACIOLI, JOËLLE MASCETTI, AND FERNAND SPIEGELMAN. **Vibrational Spectroscopy and Molecular Dynamics of Water Monomers and Dimers Adsorbed on Polycyclic Aromatic Hydrocarbons**. *Phys. Chem. Chem. Phys.*, **14**(19):6771–6786, 2012. 50, 54

[338] JA ODUTOLA AND TR DYKE. **Partially deuterated water dimers: Microwave spectra and structure**. *J. Chem. Phys.*, **72**(9):5062–5070, 1980. 50

[339] YUJI SUGITA AND YUKO OKAMOTO. **Replica-exchange multicanonical algorithm and multicanonical replica-exchange method for simulating systems with rough energy landscape**. *Chem. Phys. Lett.*, **329**(3-4):261–270, 2000. 50

[340] JULIE DOUADY, FLORENT CALVO, AND FERNAND SPIEGELMAN. **Effect of an ionic impurity on the caloric curves of water clusters**. *Eur. Phys. J. D*, **52**(1-3):47–50, 2009. 51, 53, 60, 73

[341] JILL K WOLKEN AND FRANTIŠEK TUREČEK. **Proton affinity of uracil. A computational study of protonation sites**. *J. Am. Soc. Mass Spectrom.*, **11**(12):1065–1071, 2000. 51, 84

[342] SARA ØVAD PEDERSEN, CAMILLA SKINNERUP BYSKOV, FRANTISEK TURECEK, AND STEEN BRØNDSTED NIELSEN. **Structures of protonated thymine and uracil and their monohydrated gas-phase ions from ultraviolet action spectroscopy and theory**. *J. Phys. Chem. A*, **118**(24):4256–4265, 2014. 51, 84

[343] FLORIAN WEIGEND AND REINHART AHLRICHS. **Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy**. *Phys. Chem. Chem. Phys.*, **7**(18):3297–3305, 2005. 51

[344] FLORIAN WEIGEND. **Accurate Coulomb-fitting basis sets for H to Rn**. *Phys. Chem. Chem. Phys.*, **8**(9):1057–1065, 2006. 51

[345] M. J. FRISCH, G. W. TRUCKS, H. B. SCHLEGEL, G. E. SCUSERIA, M. A. ROBB, J. R. CHEESEMAN, G. SCALMANI, V. BARONE, B. MENNUCCI, G. A. PETERSSON, AND *et. al.* 51

[346] SF BOYS AND F BERNARDI. **The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors**. *Mol. Phys.*, **100**(1):65–73, 2002. 52

[347] R. G. KEESEE. **Nucleation and particle formation in the upper atmosphere**. *J. Geophys. Res. Atmos.*, **94**(D12):14683–14692, 1989. 53

[348] J J GILLIGAN, DJ MOODY, AND A W CASTLEMAN. **Reactions of protonated water clusters with chlorine nitrate revisited**. *Z. Phys. Chem.*, **214**(10):1383–1391, 2000. 53

[349] PETR G SENNIKOV, STANISLAV K IGNATOV, AND OTTO SCHREMS. **Complexes and clusters of water relevant to atmospheric chemistry: H2O complexes with oxidants**. *ChemPhysChem*, **6**(3):392–412, 2005. 53

[350] ALBA VARGAS-CAAMAL, JOSE LUIS CABELLOS, FILIBERTO ORTIZ-CHI, HENRY S. RZEPA, ALBEIRO RESTREPO, AND GABRIEL MERINO. **How Many Water Molecules Does it Take to Dissociate HCl?** *Chem. Eur. J.*, **22**(8):2812–2818, 2016. 53

[351] ESAM A ORABI AND GUILLAUME LAMOUREUX. **Polarizable interaction model for liquid, supercritical, and aqueous ammonia**. *J. Chem. Theory Comput.*, **9**(4):2035–2051, 2013. 53

[352] MARTIN BOMMER, ANA-NICOLETA BONDAR, ATHINA ZOUNI, HOLGER DOBBEK, AND HOLGER DAU. **Crystallographic and computational analysis of the barrel part of the PsbO protein of photosystem II: Carboxylate–water clusters as putative proton transfer relays and structural switches**. *Biochemistry*, **55**(33):4626–4635, 2016. 53

[353] SD RODGERS AND SB CHARNLEY. **Chemical evolution in protostellar envelopes: Cocoon chemistry**. *Astrophys. J.*, **585**(1):355, 2003. 53

[354] EWINE F VAN DISHOECK. **ISO spectroscopy of gas and dust: from molecular clouds to protoplanetary disks**. *Annu. Rev. Astrophys. J.*, **42**, 2004. 53

[355] EL GIBB, DCB WHITTET, ACA BOOGERT, AND AGGM TIELENS. **Interstellar ice: the infrared space observatory legacy**. *Astrophys. J., Suppl. Ser.*, **151**(1):35, 2004. 53

[356] B PARISE, C CECCARELLI, AND S MARET. **Theoretical HDO emission from low-mass protostellar envelopes**. *Astron. Astrophys.*, **441**(1):171–179, 2005. 53

[357] AC ADWIN BOOGERT, PERRY A GERAKINES, AND DOUGLAS CB WHITTET. **Observations of the icy universe**. *Annu. Rev. Astron. Astrophys.*, **53**, 2015. 53

[358] F DULIEU, L AMIAUD, E CONGIU, J-H FILLION, E MATAR, A MOMENI, V PIRRONELLO, AND JL LEMAIRE. **Experimental evidence for water formation on interstellar dust grains by hydrogen and oxygen atoms**. *Astron. Astrophys.*, **512**:A30, 2010. 53

[359] ERIC MICHOULIER, NADIA BEN AMOR, MATHIAS RAPACIOLI, JENNIFER A NOBLE, JOËLLE MASCETTI, CÉLINE TOUBIN, AND AUDE SIMON. **Theoretical determination of adsorption and ionisation energies of polycyclic aromatic hydrocarbons on water ice**. *Phys. Chem. Chem. Phys.*, **20**(17):11941–11953, 2018. 53

[360] M. KULMALA, H. VEHKAMÄKI, T. PETÄJÄ, M. DAL MASO, A. LAURI, V.-M. KERMINEN, W. BIRMILI, AND P.H. MCMURRY. **Formation and growth rates of ultrafine atmospheric particles: A review of observations**. *J. Aerosol Sci.*, **35**(2):143–176, 2004. 53

REFERENCES

[361] H Ziereis and F Arnold. **Gaseous ammonia and ammonium ions in the free troposphere**. *Nature*, **321**(6069):503–505, 1986. 53

[362] Mark D Perkins and Fred L Eisele. **First mass spectrometric measurements of atmospheric ions at ground level**. *J. Geophys. Res. Atmos.*, **89**(D6):9649–9657, 1984. 53

[363] F Arnold, KH Wohlfrom, J Schneider, M Klemm, T Stilp, and F Grimm. **Atmospheric measurements of aerosol precursor gases in the upper troposphere and lower stratosphere**. *J. Aerosol Sci.*, **1001**(28):S65–S66, 1997. 53

[364] P Herbine and TR Dyke. **Rotational spectra and structure of the ammonia–water complex**. *J. Chem. Phys.*, **83**(8):3768–3774, 1985. 53

[365] Paul A Stockman, Roger E Bumgarner, Sakae Suzuki, and Geoffrey A Blake. **Microwave and tunable far-infrared laser spectroscopy of the ammonia-water dimer**. *J. Chem. Phys.*, **96**(4):2496–2510, 1992. 53

[366] Gustaf Hulthe, Gunnar Stenhagen, Olof Wennerström, and Carl-Henrik Ottosson. **Water clusters studied by electrospray mass spectrometry**. *J. Chromatogr. A*, **777**(1):155–165, 1997. 53

[367] Yih-Sheng Wang, Hai-Chou Chang, Jyh-Chiang Jiang, Sheng H. Lin, Yuan T. Lee, and Huan-Cheng Chang. **Structures and isomeric transitions of $NH_4^+(H_2O)_{3-6}$: from single to double rings**. *J. Am. Chem. Soc.*, **120**(34):8777–8788, 1998. 53, 59, 60, 61, 64

[368] Hai-Chou Chang, Yih-Sheng Wang, Yuan T Lee, and Huan-Cheng Chang. **Studying protonated ion hydration by infrared spectroscopy of size-selected $NH_4^+(H_2O)_n$ clusters in a free jet expansion**. *Int. J. Mass Spectrom.*, **179**:91–102, 1998. 53, 61

[369] J C Jiang, H-C Chang, YT Lee, and SH Lin. **Ab initio studies of $NH_4^+(H_2O)_{1-5}$ and the influence of hydrogen-bonding nonadditivity on geometries and vibrations**. *J. Phys. Chem. A*, **103**(16):3123–3135, 1999. 53, 59, 60, 61, 63

[370] Preben Hvelplund, Theo Kurten, Kristian Støchkel, Mauritz Johan Ryding, Steen Brøndsted Nielsen, and Einar Uggerud. **Stability and structure of protonated clusters of ammonia and water, $H^+(NH_3)_m(H_2O)_n$**. *J. Phys. Chem. A*, **114**(27):7301–7310, 2010. 53

[371] Julie Douady, Florent Calvo, and Fernand Spiegelman. **Structure, stability, and infrared spectroscopy of $(H_2O)_nNH_4^+$ clusters: A theoretical study at zero and finite temperature**. *J. Chem. Phys.*, **129**(15):154305, 2008. 53, 54, 59, 60, 61, 63, 64, 65, 66, 67

[372] Thomas E Morrell and George C Shields. **Atmospheric implications for formation of clusters of ammonium and 1-10 water molecules**. *J. Phys. Chem. A*, **114**(12):4266–4271, 2010. 53, 54, 60, 63, 64, 65

[373] Daniel E Bacelo. **Theoretical study of microscopic solvation of ammonia in water clusters: $NH_3(H_2O)_n$, $n = 3, 4$**. *J. Phys. Chem. A*, **106**(46):11190–11196, 2002. 53, 54, 68

[374] AE Galashev. **A computer study of ammonium adsorption on water clusters**. *Russ. J. Phys. Chem. B*, **7**(4):502–508, 2013. 53, 54

[375] Chengteh Lee, George Fitzgerald, Marc Planas, and Juan J Novoa. **Ionization of bases in water: Structure and stability of the $NH_4^+ \cdots OH^-$ ionic forms in ammonia-water clusters**. *J. Phys. Chem.*, **100**(18):7398–7404, 1996. 54, 68

[376] Piotr Skurski and Maciej Gutowski. **Theoretical study of the dipole-bound anion $(H_2O \ldots NH_3)^-$**. *J. Chem. Phys.*, **108**(15):6303–6311, 1998. 54

[377] DJ Donaldson. **Adsorption of atmospheric gases at the air- water interface. I. NH_3**. *J. Phys. Chem. A*, **103**(1):62–70, 1999. 54

[378] J Sadlej, R Moszynski, J Cz Dobrowolski, and AP Mazurek. **Structure and energetics of the weakly bound $NH_3 \cdots H_2O$ complex**. *J. Phys. Chem. A*, **103**(42):8528–8536, 1999. 54

[379] RE Kozack and PC Jordan. **Polarizability effects in a four-charge model for water**. *J. Chem. Phys.*, **96**(4):3120–3130, 1992. 54

[380] Shi-Tu Pei, Shuai Jiang, Yi-Rong Liu, Teng Huang, Kang-Ming Xu, Hui Wen, Yu-Peng Zhu, and Wei Huang. **Properties of ammonium ion–water clusters: analyses of structure evolution, noncovalent interactions, and temperature and humidity effects**. *J. Phys. Chem. A*, **119**(12):3035–3047, 2015. 54

[381] Wendell W Walters, Jiajue Chai, and Meredith G Hastings. **Theoretical phase resolved ammonia–ammonium nitrogen equilibrium isotope exchange fractionations: Applications for tracking atmospheric ammonia gas-to-particle conversion**. *ACS Earth Space Chem.*, **3**(1):79–89, 2018. 54

[382] Tae Hoon Choi and Kenneth D Jordan. **Application of the SCC-DFTB method to $H^+(H_2O)_6$, $H^+(H_2O)_{21}$, and $H^+(H_2O)_{22}$**. *J. Phys. Chem. B*, **114**(20):6932–6936, 2010. 54

[383] Tae Hoon Choi, Ruibin Liang, C Mark Maupin, and Gregory A Voth. **Application of the SCC-DFTB method to hydroxide water clusters and aqueous hydroxide solutions**. *J. Phys. Chem. B*, **117**(17):5165–5179, 2013. 54

[384] Aude Simon, Mathias Rapacioli, Eric Michoulier, Linjie Zheng, Kseniia Korchagina, and Jérôme Cuny. **Contribution of the density-functional-based tight-binding scheme to the description of water clusters: methods, applications and extension to bulk systems**. *Mol. Simul.*, **45**(4-5):249–268, 2019. 54, 55

[385] Aude Simon and Fernand Spiegelman. **Water clusters adsorbed on polycyclic aromatic hydrocarbons: Energetics and conformational dynamics**. *J. Chem. Phys.*, **138**(19):194309, 2013. 54

[386] Paul Winget, Cenk Selcuki, Anselm HC Horn, Bodo Martin, and Timothy Clark. **Towards a "next generation"neglect of diatomic differential overlap based semiempirical molecular orbital technique**. *Theor. Chem. Acc.*, **110**(4):254–266, 2003. 54, 58

[387] Han Myoung Lee, P Tarakeshwar, Jungwon Park, Maciej Roman Kołaski, Yeo Jin Yoon, Hai-Bo Yi, Woo Youn Kim, and Kwang S Kim. **Insights into the structures, energetics, and vibrations of monovalent cation-(water) 1-6 clusters**. *J. Phys. Chem. A*, **108**(15):2949–2958, 2004. 60, 61

[388] Frank C Pickard IV, Meghan E Dunn, and George C Shields. **Comparison of model chemistry and density functional theory thermochemical predictions with experiment for formation of ionic clusters of the ammonium cation complexed with water and ammonia; atmospheric implications**. *J. Phys. Chem. A*, **109**(22):4905–4910, 2005. 61

[389] Jan K Kazimirski and Victoria Buch. **Search for low energy structures of water clusters $(H_2O)_n$, n= 20- 22, 48, 123, and 293**. *J. Phys. Chem. A*, **107**(46):9762–9775, 2003. 73

[390] Bernhard Bandow and Bernd Hartke. **Larger water clusters with edges and corners on their way to ice: Structural trends elucidated with an improved parallel evolutionary algorithm**. *J. Phys. Chem. A*, **110**(17):5809–5822, 2006. 73

[391] SYLVAIN MACLOT, MICHAEL CAPRON, RÉMI MAISONNY, ARKADIUSZ ŁAWICKI, ALAIN MÉRY, JIMMY RANGAMA, JEAN-YVES CHESNEL, SADIA BARI, RONNIE HOEKSTRA, THOMAS SCHLATHÖLTER, B MANIL, L ADOUI, P ROUSSEAU, AND BA HUBER. **Ion-induced fragmentation of amino acids: effect of the environment**. *ChemPhysChem*, **12**(5):930–936, 2011. 75

[392] ALICJA DOMARACKA, MICHAEL CAPRON, SYLVAIN MACLOT, JEAN-YVES CHESNEL, ALAIN MÉRY, JEAN-CHRISTOPHE POULLY, JIMMY RANGAMA, LAMRI ADOUI, PATRICK ROUSSEAU, AND BERND A HUBER. **Ion interaction with biomolecular systems and the effect of the environment**. *J. Phys. Conf. Ser.*, **373**:012005, 2012. 75

[393] PAL MARKUSH, PAOLA BOLOGNESI, ANTONELLA CARTONI, PATRICK ROUSSEAU, SYLVAIN MACLOT, RUDY DELAUNAY, ALICJA DOMARACKA, JAROSLAV KOCISEK, MATTEA C CASTROVILLI, BERND A HUBER, AND LORENZO AVALDI. **The role of the environment in the ion induced fragmentation of Uracil**. *Phys. Chem. Chem. Phys.*, **18**(25):16721–16729, 2016. 75, 99

[394] HENRYK WINCEL. **Microhydration of protonated nucleic acid bases and protonated nucleosides in the gas phase**. *J. Am. Soc. Mass Spectrom.*, **20**(10):1900–1905, 2009. 75, 79

[395] BADIA BOUDAÏFFA, PIERRE CLOUTIER, DAREL HUNTING, MICHAEL A HUELS, AND LÉON SANCHE. **Resonant formation of DNA strand breaks by low-energy (3 to 20 eV) electrons**. *Science*, **287**(5458):1658–1660, 2000. 75

[396] MAEVE SMYTH AND JORGE KOHANOFF. **Excess electron localization in solvated DNA bases**. *Phys. Rev. Lett.*, **106**(23):238108, 2011. 75

[397] KATRIN R SIEFERMANN AND BERND ABEL. **The hydrated electron: a seemingly familiar chemical and biological transient**. *Angew. Chem. Int. Ed.*, **50**(23):5264–5272, 2011. 75

[398] ELAHE ALIZADEH, ANA G SANZ, GUSTAVO GARCIA, AND LEON SANCHE. **Radiation damage to DNA: The indirect effect of low-energy electrons**. *J. Phys. Chem. Lett.*, **4**(5):820–825, 2013. 75

[399] ANDREW M RASMUSSEN, MARIA C LIND, SUNGHWAN KIM, AND HENRY F SCHAEFER III. **Hydration of the lowest triplet states of the DNA/RNA pyrimidines**. *J. Chem. Theory Comput.*, **6**(3):930–939, 2010. 76

[400] LEILA SADR-ARANI, PIERRE MIGNON, HENRY CHERMETTE, AND THIERRY DOUKI. **Theoretical and experimental study of the fragmentation of protonated uracil**. *Chem. Phys. Lett.*, **605**:108–114, 2014. 76

[401] ESTEFANÍA ROSSICH MOLINA, JEAN-YVES SALPIN, RICCARDO SPEZIA, AND EMILIO MARTINEZ-NUNEZ. **On the gas phase fragmentation of protonated uracil: A statistical perspective**. *Phys. Chem. Chem. Phys.*, **18**(22):14980–14990, 2016. 76, 92

[402] JOOST M BAKKER, RAJEEV K SINHA, THIERRY BESSON, MAURIZIO BRUGNARA, PAOLO TOSI, JEAN-YVES SALPIN, AND PHILIPPE MAÎTRE. **Tautomerism of uracil probed via infrared spectroscopy of singly hydrated protonated uracil**. *J. Phys. Chem. A*, **112**(48):12393–12400, 2008. 76, 79, 81, 84, 99

[403] OLEG V SHISHKIN, LEONID GORB, AND JERZY LESZCZYNSKI. **Modeling of the hydration shell of uracil and thymine**. *Int. J. Mol. Sci.*, **1**(2):17–27, 2000. 76

[404] SHRIDHAR R GADRE, K BABU, AND ALISTAIR P RENDELL. **Electrostatics for exploring hydration patterns of molecules. 3. Uracil**. *J. Phys. Chem. A*, **104**(39):8976–8982, 2000. 76

[405] TANJA VAN MOURIK, SARAH L PRICE, AND DAVID C CLARY. **Diffusion Monte Carlo simulations on uracil–water using an anisotropic atom–atom potential model**. *Faraday discussions*, **118**:95–108, 2001. 76

[406] MARIE-PIERRE GAIGEOT AND MAHMOUD GHOMI. **Geometrical and Vibrational Properties of Nucleic Acid Constituents Interacting with Explicit Water Molecules as Analyzed by Density Functional Theory Calculations. 1. Uracil + $n_W H_2O$ (n_W= 1, ..., 7)**. *J. Phys. Chem. B*, **105**(21):5007–5017, 2001. 76

[407] VICTOR I DANILOV, TANJA VAN MOURIK, AND VALERY I POLTEV. **Modeling of the 'hydration shell' of uracil and thymine in small water clusters by DFT and MP2 methods**. *Chem. Phys. Lett.*, **429**(1-3):255–260, 2006. 76

[408] MARIE-CHRISTINE BACCHUS-MONTABONEL AND FLORENT CALVO. **Nanohydration of uracil: emergence of three-dimensional structures and proton-induced charge transfer**. *Phys. Chem. Chem. Phys.*, **17**(15):9629–9633, 2015. 76

[409] NF DALLESKA, KENJI HONMA, AND PB ARMENTROUT. **Stepwise solvation enthalpies of protonated water clusters: Collision-induced dissociation as an alternative to equilibrium studies**. *J. Am. Chem. Soc.*, **115**(25):12125–12131, 1993. 78, 79, 80

[410] SÉBASTIEN ZAMITH, PIERRE LABASTIE, AND JEAN-MARC L'HERMITE. **Fragmentation cross sections of protonated water clusters**. *J. Chem. Phys.*, **136**(21):214301, 2012. 78, 79, 80, 99

[411] RICHARD L MYERS. *The 100 most important chemical compounds: a reference guide*. ABC-CLIO, 2007. 78

[412] K HANSEN, PU ANDERSSON, AND E UGGERUD. **Activation energies for evaporation from protonated and deprotonated water clusters from mass spectra**. *J. Chem. Phys.*, **131**(12):124303, 2009. 79

[413] MARY ANN KURINOVICH, LINDA M PHILLIPS, SEEMA SHARMA, AND JEEHIUN K LEE. **The gas phase proton affinity of uracil: measuring multiple basic sites and implications for the enzyme mechanism of orotidine 5'-monophosphate decarboxylase**. *Chemical communications*, **21**(20):2354–2355, 2002. 81, 82

[414] THOMAS F MAGNERA, DONALD E DAVID, AND JOSEF MICHL. **The first twenty-eight gas-phase proton hydration energies**. *Chem. Phys. Lett.*, **182**(3-4):363–370, 1991. 81, 82

[415] HAI-PING CHENG. **Water clusters: fascinating hydrogen-bonding networks, solvation shell structures, and proton motion**. *J. Phys. Chem. A*, **102**(31):6201–6204, 1998. 81, 82

[416] G. ZUNDEL AND H. METZGER. **Energiebänder der tunnelnden Überschuß-Protonen in flüssigen Säuren. Eine IR-spektroskopische Untersuchung der Natur der Gruppierungen $H_5O_2^+$**. *Proc. Z. Phys. Chem.*, **58**(5_6):225–245, 1968. 88

[417] C BRÉCHIGNAC, PH CAHUZAC, J LEYGNIER, AND J WEINER. **Dynamics of unimolecular dissociation of sodium cluster ions**. *J. Chem. Phys.*, **90**(3):1492–1498, 1989. 97

[418] C BRÉCHIGNAC, H BUSCH, PH CAHUZAC, AND J LEYGNIER. **Dissociation pathways and binding energies of lithium clusters from evaporation experiments**. *J. Chem. Phys.*, **101**(8):6992–7002, 1994. 97

[419] RICHARD L WONG, KOLJA PAECH, AND EVAN R WILLIAMS. **Blackbody infrared radiative dissociation at low temperature: Hydration of $X^{2+}(H_2O)_n$, for X= Mg, Ca**. *International J. Mass Spectrom.*, **232**(1):59–66, 2004. 97

[420] MATTHEW F BUSH, RICHARD J SAYKALLY, AND EVAN R WILLIAMS. **Reactivity and infrared spectroscopy of gaseous hydrated trivalent metal ions**. *J. Am. Chem. Soc.*, **130**(28):9122–9128, 2008. 97

REFERENCES

[421] KRISTIE A BOERING, JOSEPH ROLFE, AND JOHN I BRAUMAN. **Low energy collision induced dissociation: phase-shifting excitation control of ion kinetic energy in ion cyclotron resonance spectrometry**. *Int. J. Mass Spectrom. Ion Processes*, **117**:357–386, 1992. 98

[422] J MITCHELL WELLS AND SCOTT A MCLUCKEY. **Collision-induced dissociation (CID) of peptides and proteins**. *Methods Enzymol.*, **402**:148–185, 2005. 98

[423] YL MA, QM LI, H VAN DEN HEUVEL, AND M CLAEYS. **Characterization of flavone and flavonol aglycones by collision-induced dissociation tandem mass spectrometry**. *Rapid Commun. Mass*, **11**(12):1357–1364, 1997. 98

[424] SAIFUL M CHOWDHURY, XIUXIA DU, NIKOLA TOLIC, SI WU, RONALD J MOORE, M ULJANA MAYER, RICHARD D SMITH, AND JOSHUA N ADKINS. **Identification of cross-linked peptides after click-based enrichment using sequential collision-induced dissociation and electron transfer dissociation tandem mass spectrometry**. *Anal. Chem.*, **81**(13):5524–5532, 2009. 98

[425] LEKHA SLENO AND DIETRICH A VOLMER. **Ion activation methods for tandem mass spectrometry**. *J. Mass Spectrom.*, **39**(10):1091–1112, 2004. 98

[426] RB CODY, RC BURNIER, AND BS FREISER. **Collision-induced dissociation with Fourier transform mass spectrometry**. *Anal. Chem.*, **54**(1):96–101, 1982. 98

[427] JESPER V OLSEN, BORIS MACEK, OLIVER LANGE, ALEXANDER MAKAROV, STEVAN HORNING, AND MATTHIAS MANN. **Higher-energy C-trap dissociation for peptide modification analysis**. *Nat. Methods.*, **4**(9):709–712, 2007. 98

[428] GENE HART-SMITH AND MARK J RAFTERY. **Detection and characterization of low abundance glycopeptides via higher-energy C-trap dissociation and orbitrap mass analysis**. *J. Am. Soc. Mass Spectrom.*, **23**(1):124–140, 2011. 98

[429] JW GAUTHIER, TR TRAUTMAN, AND DB JACOBSON. **Sustained off-resonance irradiation for collision-activated dissociation involving Fourier transform mass spectrometry. Collision-activated dissociation technique that emulates infrared multiphoton dissociation**. *Anal. Chim. Acta*, **246**(1):211–225, 1991. 98

[430] JULIA LASKIN AND JEAN H FUTRELL. **Activation of large Ions in FT-ICR mass spectrometry**. *Mass Spectrom. Rev.*, **24**(2):135–167, 2005. 98

[431] KERI MCQUINN, FRASER HOF, AND J SCOTT MCINDOE. **Collision-induced dissociation of protonated nanodroplets**. *Int. J. Mass Spectrom.*, **279**(1):32–36, 2009. 98, 99

[432] DAMON R CARL AND PETER B ARMENTROUT. **Hreshold collision-induced dissociation of hydrated magnesium: experimental and theoretical investigation of the binding energies for $Mg^{2+}(H_2O)_x$ complexes (x= 2–10)**. *ChemPhysChem*, **14**(4):681–697, 2013. 98, 99

[433] THERESA E. HOFSTETTER AND P. B. ARMENTROUT. **Threshold collision-induced dissociation and theoretical studies of hydrated Fe(II): binding energies and coulombic barrier heights**. *J. Phys. Chem. A*, **117**(6):1110–1123, 2013. 98, 99

[434] REBECCA A COATES AND PB ARMENTROUT. **Thermochemical investigations of hydrated nickel dication complexes by threshold collision-induced dissociation and theory**. *J. Phys. Chem. A*, **121**(19):3629–3646, 2017. 98

[435] N HAAG, B LIU, S BRØNDSTED NIELSEN, H ZETTERGREN, P HVELPLUND, B MANIL, B A HUBER, H A B JOHANSSON, H T SCHMIDT, AND H CEDERQUIST. **Collisions with biomolecules embedded in small water clusters**. *J. Phys. Conf. Ser.*, **194**(1):012053, nov 2009. 98

[436] SHIJIAN LI AND ELLIOT R BERNSTEIN. **Toluene-water clusters: Ion fragmentation and chemistry**. *J. Chem. Phys.*, **97**(2):792–803, 1992. 99

[437] C BOBBERT, S SCHÜTTE, C STEINBACH, AND U BUCK. **Fragmentation and reliable size distributions of large ammonia and water clusters**. *Eur. Phys. J. D*, **19**(2):183–192, 2002. 99

[438] VASSIL A SPASOV, TAECK-HONG LEE, AND KENT M ERVIN. **Threshold collision-induced dissociation of anionic copper clusters and copper cluster monocarbonyls**. *J. Chem. Phys.*, **112**(4):1713–1720, 2000. 99

[439] PB ARMENTROUT, KENT M ERVIN, AND MT RODGERS. **Statistical rate theory and kinetic energy-resolved ion chemistry: theory and applications**. *J. Phys. Chem. A*, **112**(41):10071–10085, 2008. 99, 100

[440] MT RODGERS AND PB ARMENTROUT. **Statistical modeling of competitive threshold collision-induced dissociation**. *J. Chem. Phys.*, **109**(5):1787–1800, 1998. 100

[441] PB ARMENTROUT. **Statistical modeling of sequential collision-induced dissociation thresholds**. *J. Chem. Phys.*, **126**(23):234302, 2007. 100

[442] ISABELLE BRAUD, SÉBASTIEN ZAMITH, AND J-M L'HERMITE. **A gas aggregation source for the production of heterogeneous molecular clusters**. *Rev. Sci. Instrum.*, **88**(4):043102, 2017. 100

[443] FABIEN CHIROT, SÉBASTIEN ZAMITH, PIERRE LABASTIE, AND J-M L'HERMITE. **New device to study unimolecular cluster nucleation**. *Rev. Sci. Instrum.*, **77**(6):063108, 2006. 102

[444] J FRENZEL, AF OLIVEIRA, N JARDILLIER, T HEINE, AND G SEIFERT. **Semi-relativistic, self-consistent charge Slater-Koster tables for density-functional based tight-binding (DFTB) for materials science simulations**. *Zeolites*, **2**(3):7, 2004. 102

[445] FERNAND SPIEGELMAN, NATHALIE TARRAT, JÉRÔME CUNY, LEO DONTOT, EVGENY POSENITSKIY, CARLES MARTÍ, AUDE SIMON, AND MATHIAS RAPACIOLI. **Density-functional tight-binding: basic concepts and applications to molecules and clusters**. *Adv. Phys. -X*, **5**(1):1710252, 2020. 102, 132

[446] AUDE SIMON, MATHIAS RAPACIOLI, GUILLAUME ROUAUT, GEORGES TRINQUIER, AND FX GADÉA. **Dissociation of polycyclic aromatic hydrocarbons: molecular dynamics studies**. *Philos. Trans. R. Soc. A*, **375**(2092):20160195, 2017. 102, 103

[447] MATHIAS RAPACIOLI, NATHALIE TARRAT, AND FERNAND SPIEGELMAN. **Melting of the Au_{20} gold cluster: Does charge matter?** *J. Phys. Chem. A*, **122**(16):4092–4098, 2018. 102

[448] AUDE SIMON, JEAN-PHILIPPE CHAMPEAUX, MATHIAS RAPACIOLI, PATRICK MORETTO-CAPELLE, X. GADÉA, FLORENT, AND MARTINE SENCE. **Dissociation of polycyclic aromatic hydrocarbons at high energy: MD/DFTB simulations versus collision experiments**. *Theor. Chem. Acc.*, **137**(7):106, 2018. 102, 103

[449] ARIEH WARSHEL AND MICHAEL LEVITT. **Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme**. *J. Mol. Biol.*, **103**(2):227–249, 1976. 102

[450] QIANG CUI, MARCUS ELSTNER, EFTHIMIOS KAXIRAS, THOMAS FRAUENHEIM, AND MARTIN KARPLUS. **A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method**. *J. Phys. Chem. B*, **105**(2):569–585, 2001. 102

[451] Christophe Iftner, Aude Simon, Kseniia Korchagina, Mathias Rapacioli, and Fernand Spiegelman. **A density functional tight binding/force field approach to the interaction of molecules with rare gas clusters: Application to $(C_6H_6)^{+/0}Ar_n$ clusters**. *J. Chem. Phys.*, **140**(3):034301, 2014. 102

[452] Edwin Kukk, DT Ha, Y Wang, Dariusz Grzegorz Piekarski, Sergio Diaz-Tendero, Kuno Kooser, Eero Itälä, Helena Levola, Manuel Alcamí, Elisabeth Rachlew, et al. **Internal energy dependence in X-ray-induced molecular fragmentation: an experimental and theoretical study of thiophene**. *Phys. Rev. A*, **91**(4):043417, 2015. 102, 103

[453] Mathias Rapacioli, Stéphanie Cazaux, Nolan Foley, Aude Simon, Ronnie Hoekstra, and Thomas Schlathölter. **Atomic hydrogen interactions with gas-phase coronene cations: hydrogenation versus fragmentation**. *Phys. Chem. Chem. Phys.*, **20**(35):22427–22438, 2018. 103

[454] Léo Dontot, Fernand Spiegelman, and Mathias Rapacioli. **Structures and energetics of neutral and cationic pyrene clusters**. *J. Phys. Chem. A*, **123**(44):9531–9543, 2019. 103, 132, 137, 143

[455] Seung-Hyun Chung and Angela Violi. **Peri-condensed aromatics with aliphatic chains as key intermediates for the nucleation of aromatic hydrocarbons**. *Proc. Combust. Inst.*, **33**(1):693–700, 2011. 131

[456] Chiara Saggese, Sara Ferrario, Joaquin Camacho, Alberto Cuoci, Alessio Frassoldati, Eliseo Ranzi, Hai Wang, and Tiziano Faravelli. **Kinetic modeling of particle size distribution of soot in a premixed burner-stabilized stagnation ethylene flame**. *Combust. Flame*, **162**(9):3356–3369, 2015. 131

[457] Qian Mao, Adri CT Van Duin, and KH Luo. **Formation of incipient soot particles from polycyclic aromatic hydrocarbons: a ReaxFF molecular dynamics study**. *Carbon*, **121**:380–388, 2017. 131, 132

[458] Chen S Wang, Norman C Bartelt, Regina Ragan, and Konrad Thürmer. **Revealing the molecular structure of soot precursors**. *Carbon*, **129**:537–542, 2018. 131

[459] SA Kyrtopoulos, Panagiotis Georgiadis, H Autrup, N Demopoulos, P Farmer, A Haugen, K Katsouyanni, B Lambert, S Ovrebo, R Sram, et al. **Biomarkers of genotoxicity of urban air pollution: overview and descriptive data from a molecular epidemiology study on populations exposed to moderate-to-low levels of polycyclic aromatic hydrocarbons: the AULIS project**. *Mutat. Res. Genet. Toxicol. Environ.*, **496**(1-2):207–228, 2001. 131

[460] Peter B Farmer, Rajinder Singh, Balvinder Kaur, Radim J Sram, Blanka Binkova, Ivan Kalina, Todor A Popov, Seymour Garte, Emanuela Taioli, Alena Gabelova, et al. **Molecular epidemiology studies of carcinogenic environmental pollutants: effects of polycyclic aromatic hydrocarbons (PAHs) in environmental pollution on exogenous and oxidative DNA damage**. *Mutat. Res.-Rev. Mutat. Res.*, **544**(2-3):397–402, 2003. 131

[461] Cyril Aumaitre and Jean-François Morin. **Polycyclic aromatic hydrocarbons as potential building blocks for organic solar cells**. *Chem. Rec.*, **19**(6):1142–1154, 2019. 131

[462] M Rapacioli, C Joblin, and P Boissel. **Spectroscopy of polycyclic aromatic hydrocarbons and very small grains in photodissociation regions**. *Astron. Astrophys.*, **429**(1):193–204, 2005. 131

[463] O Berné, C Joblin, M Rapacioli, J Thomas, J-C Cuillandre, and Y Deville. **Extended red emission and the evolution of carbonaceous nanograins in NGC 7023**. *Astron. Astrophys.*, **479**(3):L41–L44, 2008. 131

[464] Annette Eschenbach, Reinhard Wienberg, and Bernd Mahro. **Fate and stability of nonextractable residues of [14C] PAH in contaminated soils under environmental stress conditions**. *Environ. Sci. Technol.*, **32**(17):2585–2590, 1998. 131

[465] Marcelo Goulart, Martin Kuhn, Bilal Rasul, Johannes Postler, Michael Gatchell, Henning Zettergren, Paul Scheier, and Olof Echt. **The structure of coronene cluster ions inferred from H 2 uptake in the gas phase**. *Phys. Chem. Chem. Phys.*, **19**(41):27968–27973, 2017. 131

[466] Lei Lei, Yuzhong Yao, Jie Zhang, Dale Tronrud, Wei Kong, Chengzhu Zhang, Lan Xue, Léo Dontot, and Mathias Rapacioli. **Electron diffraction of pyrene nanoclusters embedded in superfluid helium droplets**. *J. Phys. Chem. Lett.*, **11**(3):724–729, 2019. 131

[467] JE Roser and A Ricca. **Polycyclic aromatic hydrocarbon clusters as sources of interstellar infrared emission**. *Astrophys. J.*, **801**(2):108, 2015. 131

[468] Alexander K Lemmens, Sébastien Gruet, Amanda L Steber, Jens Antony, Stefan Grimme, Melanie Schnell, and Anouk M Rijs. **Far-IR and UV spectral signatures of controlled complexation and microhydration of the polycyclic aromatic hydrocarbon acenaphthene**. *Phys. Chem. Chem. Phys.*, **21**(7):3414–3422, 2019. 131

[469] Rudy Delaunay, Michael Gatchell, Patrick Rousseau, Alicja Domaracka, Sylvain Maclot, Yang Wang, Mark H Stockett, Tao Chen, Lamri Adoui, Manuel Alcamí, et al. **Molecular growth inside of polycyclic aromatic hydrocarbon clusters induced by ion collisions**. *J. Phys. Chem. Lett.*, **6**(9):1536–1542, 2015. 132

[470] Manuel Piacenza and Stefan Grimme. **Van der Waals complexes of polar aromatic molecules: Unexpected structures for dimers of azulene**. *J. Am. Chem. Soc.*, **127**(42):14841–14848, 2005. 132

[471] Ö Birer and E Yurtsever. **Dimer formation of perylene: An ultracold spectroscopic and computational study**. *J. Mol. Struct.*, **1097**:29–36, 2015. 132

[472] Y. Zhao and Donald G. Truhlar. **A Prototype for graphene material simulation; structures and interaction potentials of coronene dimers**. *J. Phys. Chem. C*, **112**:4061–4067, 2008. 132

[473] Kimberly Bowal, Jacob W Martin, and Markus Kraft. **Partitioning of polycyclic aromatic hydrocarbons in heterogeneous clusters**. *Carbon*, **143**:247–256, 2019. 132

[474] Alessandra Ricca, Charles W Bauschlicher Jr, and Louis J Allamandola. **The infrared spectroscopy of neutral polycyclic aromatic hydrocarbon clusters**. *Astrophys. J.*, **776**(1):31, 2013. 132

[475] Jürgen Gräfenstein and Dieter Cremer. **The self-interaction error and the description of non-dynamic electron correlation in density functional theory**. *Theor. Chem. Acc.*, **123**(3):171–182, 2009. 132

[476] Léo Dontot, Nicolas Suaud, Mathias Rapacioli, and Fernand Spiegelman. **An extended DFTB-CI model for charge-transfer excited states in cationic molecular clusters: model studies versus ab initio calculations in small PAH clusters**. *Phys. Chem. Chem. Phys.*, **18**(5):3545–3557, 2016. 132

[477] Léo Dontot, Fernand Spiegelman, Sébastien Zamith, and Mathias Rapacioli. **Dependence upon charge of the vibrational spectra of small Polycyclic Aromatic Hydrocarbon clusters: the example of pyrene**. *Eur. Phys. J. D*, **74**(11):1–9, 2020. 132, 149

[478] Mathias Rapacioli, Fernand Spiegelman, Anthony Scemama, and André Mirtschink. **Modeling charge resonance in cationic molecular clusters: combining DFT-tight binding with configuration interaction**. *J. Chem. Theory Comput.*, **7**(1):44–55, 2011. 132

[479] MICHAEL GATCHELL. *Molecular Hole Punching: Impulse Driven Reactions in Molecules and Molecular Clusters*. PhD thesis, Department of Physics, Stockholm University, 2016. 132

[480] MICHAEL GATCHELL AND HENNING ZETTERGREN. **Knockout driven reactions in complex molecules and their clusters**. *J. Phys. B*, **49**(16):162001, 2016. 132, 141, 155

[481] RAPHAEL D LEVINE, RICHARD BARRY BERNSTEIN, ET AL. *Molecular reaction dynamics and chemical reactivity*. Oxford University Press, USA, 1987. 135

[482] TAO CHEN, MICHAEL GATCHELL, MARK H STOCKETT, JOHN D ALEXANDER, Y ZHANG, PATRICK ROUSSEAU, A DOMARACKA, S MACLOT, R DELAUNAY, L ADOUI, ET AL. **Absolute fragmentation cross sections in atom-molecule collisions: scaling laws for non-statistical fragmentation of polycyclic aromatic hydrocarbon molecules**. *J. Chem. Phys.*, **140**(22):224306, 2014. 141, 155

/

Declaration

I herewith declare that I have produced this paper without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This paper has not previously been presented in identical or similar form to any other examination board.

The thesis work was conducted from 2017 to 2021 under the supervision of Jérôme Cuny and Mathias Rapacioli in Laboratoire de Chimie et Physique Quantiques.