August 2019

Modeling Chemical Reactivity in Aqueous and Organic Systems: From Electronic Structure Methods to Force Field Development

Caitlin Gibson Bresnahan
Louisiana State University and Agricultural and Mechanical College, caitlinbresnahan92@gmail.com

MODELING CHEMICAL REACTIVITY IN AQUEOUS AND ORGANIC SYSTEMS: FROM ELECTRONIC STRUCTURE METHODS TO FORCE FIELD DEVELOPMENT

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Chemistry

by
Caitlin Gibson Bresnahan
B.S., University of Wisconsin-Eau Claire, 2014
December 2019

Acknowledgements

To say that attending graduate school has been a wild journey that I had not anticipated would be a colossal understatement. Moving across the country, and the experiences granted me throughout my tenure at LSU has been incredibly exciting, and yet intimidating at times. I've learned so much through these experiences, both personally and academically.

My advisor, Dr. Revati Kumar, has pushed me out of my comfort zone and presented me with many opportunities and intellectual challenges including sending me to many different conferences, regions of the country, and even encouraging me to pursue international opportunities. Not only am I grateful for your role in my professional development, I am grateful to you for challenging me personally as well; from political views to insights on different cultures to offering a new viewpoint on my home-state.

I would like to thank Dr. Anne Milet who has been an exceptional sounding board when I've been frustrated at work, offering insight from another perspective. I am grateful that I was able to spend five months working under your guidance. Not only did I grow professionally, but I learned more than I ever anticipated about French culture; from the right way to pronounce your name, your arguably better version of king cake, how to properly enjoy a six course meal, and an extensive amount of the history of Grenoble. Thank you for your input, your help, and for laughing with me when it turned out that your assessment of my hiking shoes was correct; they weren't up to the job.

To my committee members, Dr. George Stanley and Dr. Bin Chen, thank you for joining my journey through my doctorate. I would like to thank Dr. Adam Bruner and Dr. Rolf David. While I'm sorry your time here never overlapped for more than a conference here and there, the two of you have offered an immense amount of help and guidance which I appreciate more than I

can articulate. I look forward to someday fulfilling our ACS-New Orleans bucket list. To Dr. Serge Nader, Dr. Suzanne Adam, and Jen Fry; I will forever be grateful for your guidance during my time in France, particularly the suggestions that ended up with us at the Christmas market or supporting regionally made products.

Last, and not least, I would like to thank my family. I didn't realize that when I decided to go to graduate school, you all would be attending in spirit with me. To my parents, I want to say thank you for every supportive phone call, video chat and visit, particularly in those months that Baton Rouge was determined to devour my car. I also really loved every terrible chemistry joke sent my way. Megan, the editor in chief, I thank you thirty-seven times over for the hours you've spent during my career doing the final proof read, and the reminder to just do it, Doug. To Erin, the constant unwavering support, bourbon chicken, and lately baby pictures have kept me grounded. To Jeremy, thank you for always being curious about my research and asking interesting questions, and to Crowley for keeping life amusing. Lastly, I would like to thank my Grandma, for always being one phone call away and reminding me that there is more to life than whatever I am currently focused on.

Table of Contents

Abstract

Modeling reactivity in chemical systems has evolved dramatically in line with the capabilities of modern computing. Despite the advances in computational ability, the level in which one can model a system depends on a number of factors including the region of reactivity, size of the system, level of sophistication required in the molecular description, and so on. Electronic structure methods allow for a detailed description of the potential energy surface and inherently include all essential physics required for reactivity to occur, however these methods are limited by their computational expense. On the other hand, force fields allow for an atomistic description of the interactions and drastically reduce the simulation time, yet typical force fields are dependent on a fixed bond topology, and as such, cannot model bond cleavage and formation.

This dissertation addresses modeling reactivity from electronic structure methods to force field development for reactive systems. The first section of the dissertation will focus on the hydrated HCl system. Accurately modeling covalent HCl, as well as ionization and subsequent proton shuttling, is essential in systems such as gas-liquid nucleation in the atmosphere, concentrated acid solutions, and HCl at the air-water interface. The amount of sampling required for gas-liquid nucleation pathways, or simulation time for large system sizes in the case of concentrated acid simulations necessitates an expedient description of the potential energy surface. To this end, a reactive force field has been developed. In order to determine the solvent environment factors required for an accurate force field description, ab initio molecular dynamics and metadynamics have been performed on $HCl(H_2O)_{n(n=1\text{-}22)}$. These simulations will be discussed in chapter two, while the development and performance of a reactive force field based on the multi-state empirical bond formalism will be described in chapter three. The second section of the dissertation will focus on modeling reactivity with electronic structure methods for two organic

systems. The systems range from determining the factors guiding the regioselectivity of silyloxyallyl cations by analyzing reaction profiles, SAPT energy decomposition, and molecular orbital analysis (chapter four), to the formation of an EDA complex and the corresponding charge transfer (chapter five).

Chapter 1. Introduction

1.1. Overview of Computational Simulation Methods

From electronic structure methods to force fields, a significant effort in computational chemistry has focused on how to efficiently model chemical systems at progressively longer length and time scales. Computational chemistry allows for exquisite insight into the atomistic or molecular description of chemical processes and can be a useful tool in solving chemical problems. Different techniques can be employed in the simulation of important processes in order to determine chemical and physical properties that are difficult to observe at the molecular level experimentally. For instance, spectroscopy still has difficulty probing transient states with short lifetimes although these spectroscopic techniques are quickly evolving.[1, 2] In particular, experiments and computations work in tandem to explain the atomistic insight into macroscopic properties. Computations can determine full reaction pathways,[3] transition states,[4] excited states, charge delocalization,[5] transport properties[6] and a multitude of other fundamental and physical properties.[7-9] Additionally, comparison between the atomistic level computations and experiments can take place by relating calculated spectra from simulations with experimental spectra from Raman,[10-12] IR,[12-14] and NMR[15, 16] spectroscopy, as well as other properties such as surface tension,[17, 18] conductivity,[19] etc.[20] Drug design often uses computational packages focused on docking to screen potential binding molecules.[21-23] Large proteins can be analyzed to determine large motions or broken down into reactive calculations of differing binding sites.[24-26] Computations can even be useful in helping direct the synthetic process. These are just a few examples of the power of computations, particularly in tandem with experiments. All in all, the field of computational chemistry has been growing and strong collaborations with experimentalists have allowed for the molecular level picture to be analyzed in greater detail.

Molecular systems can be simulated using two different methods. The first, molecular dynamics (MD) simulations, are essentially virtual experiments where the simulation imitates real systems providing the underlying intermolecular interactions are accurately modeled.[27-32] MD simulations provide the temporal evolution of the system at the atomistic level by integrating Newtonian's second law of motion, namely, $\boldsymbol{F}_i = m_i\ddot{\boldsymbol{r}}$**.** Here $\boldsymbol{F}_i$ is the force on particle i which is given by the negative gradient of the potential of interaction, $-\nabla_i U$, while m_i and $\ddot{\boldsymbol{r}}$ are the mass and acceleration of particle i. Monte Carlo (MC) based molecular simulations are a stochastic method based on equilibrium statistical mechanics.[33-37] Briefly, in canonical MC simulations, a new state of the system is generated by random displacements at each step and this new state is accepted or rejected based on a Boltzmann probability that includes the energy difference between the old and new states. The above canonical MC simulations can be extended to include non-Boltzmann transition probabilities giving rise to enhanced sampling.[38-41] While molecular dynamics can give both static and dynamical insight, MC based methods are restricted to static data. However, the new states generated during MC moves are not restricted the way they are in molecular dynamics simulations.

The accuracy of both MC and MD simulations strongly depends on the potential of interaction. In principle one should use the purely quantum mechanical energy (or the energy from electronic structure calculations), to describe the potential energy surface (PES) but given the expense, a cheaper alternative is to use force fields that are essentially a reduced representation of the key intermolecular interactions. The main focus of this work is to develop force fields from high level electronic structure calculations.

1.2. Ab Initio Description of the PES

In a truly quantum mechanical treatment, one would solve the time-dependent Schrödinger equation, given in equation 1.1. The Hamiltonian operator, $\hat{H}$, is given in equation 1.2.

$$i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi \tag{1.1}$$

$$\hat{H} = \frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + V(x, y, z) \tag{1.2}$$

However, this is computationally intractable for most chemical systems of interest. Instead, one can carry out Born-Oppenheimer Molecular Dynamics (BOMD), also often referred to as ab initio molecular dynamics (AIMD). The Born-Oppenheimer approximation maintains the underlying assumption that the motion of the nuclei is decoupled from that of electrons since the latter are so much lighter and have much more rapid motion than nuclear motion. Hence, in BOMD, the nuclei are propagated classically based on atomic forces determined by solving the electronic time-independent Schrodinger equation, equation 1.3, at each step. There are two main motifs of electronic structure methods: non-empirical and semi-empirical.

$$\hat{H}\psi = E\psi \tag{1.3}$$

Semi-empirical methodologies, while expedient, rely on a number of parameters which are obtained from experiment.[42-44] Semi-empirical methods work well on the systems they are parameterized for, resulting in good qualitative agreement; however, these methods struggle to accurately model other systems. For this reason, most of the electronic structure methods used in this work will rely on ab initio methods. Ab initio methods, resulting from first principle, do not rely on parameterization and contain all essential physics of the system, including nuclear quantum effects. These ab initio methods can be wave function based such as Hartree-Fock[45], Moller-Plesset Perturbation Theory,[46] etc., or electron density based such as Density Functional Theory

(DFT).[47-50] In this work, both wave function and density based methods for single point energies and optimizations from water clusters containing acid defects to organic reaction pathways, were utilized; however, for the AIMD simulations conducted, only DFT based methods are feasible. DFT includes electron-electron correlation and allows for relatively quick calculations of molecular properties. DFT methods are ideal for modeling reactive pathways in organic synthesis and have been shown repeatedly to perform well in these scenarios. While this methodology is ideal for modeling reactivity, it suffers from the drawback of computational expense when compared to force fields. Currently, ab initio molecular dynamics (AIMD) is limited to roughly a few hundred atoms with a sampling of approximately 100-200 ps.[51, 52]

Given the cost of ab initio methods, one often must decide at what level one can treat solvent effects. An explicit or implicit description of the solvent is readily available. In situations where the solvent plays a large role in the reactivity or movement of the system under study, an explicit solvent is required. These cases model the solvent in the same fashion as the other species in the system. However, as the number of electrons increases, so does the number of calculations; leading to longer simulation times. In cases where thousands of solvent molecules are required, the computational cost is once again intractable. Implicit solvation, or continuum solvents, have a uniform medium where the solvation is modeled as a polarizable dielectric continuum implicit solvent.[53-55] While this is sufficient in some cases, it is not in others, such as cavity systems where the solvent plays a large role in transport properties.

1.3. Classical Force Field Description of the PES

A more computationally efficient model is the use of force fields based on classical physics. In classical all atom force fields the electronic degrees of freedom are removed and the potential is based strictly on atom-atom interactions. Due to the removal of explicit electron treatment,

classical force fields are much more expedient in calculations, therefore allowing significantly longer simulation times and larger sampling of phase space. A plethora of classical force fields has been developed for different systems.[56] The most widely used force fields include the Assisted Model Building and Energy Refinement (AMBER)[57, 58] which was parameterized for studies of proteins and DNA, Chemistry at HARvard Molecular Mechanics (CHARMM)[59-61] built for small molecules and widely used on small and macromolecules, the GROningen Molecular Simulation (GROMOS)[62-64] parameterized for biomolecular systems, and the Optimized Potential for Liquid Simulations (OPLS),[65-67] which includes an all atom and united atom version, parameterized for ionic liquids. This list is by no means exhaustive, but gives an indication of the number of force fields present in the literature as well as the systems they have been parameterized for. Force fields determine the potential energy surface (PES) given a functional form dependent on the summation over bonded and nonbonded interactions, given in equation 1.4.

$$V_{total} = \sum V_{bonded} + \sum V_{nonbonded} \tag{1.4}$$

Classical non-polarizable force fields have similar potential energy functions with slightly different functional forms. As an example, the CHARMM functional form for the PES can be seen in equation 1.5.[59]

$$V(x) = \sum_{bonds} \frac{k_i}{2}(r_i - r_i^0)^2 + \sum_{angles} \frac{k_j}{2}\left(\theta_j - \theta_j^0\right)^2 + \sum_{dihedrals} \sum_{m=1}^{m_{max}} k_k^m cos(m\varphi_k - \gamma_k)$$
$$+ \sum_{impropers} \frac{k_l}{2}(\phi - \phi_l^0)^2 + \sum_{i<j} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} + 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{1.5}$$

Despite the increase of computational speed and ease of implementing an explicit solvent, force fields have a number of limitations. As the number of parameters used is quite extensive, accuracy is dependent on what systems the force field is parameterized to. Force fields can be fit to an ab

initio PES, or the function can be fit to reproduce experimental parameters such as rotational and vibrational spectra.[68] The parameters are only as accurate as the methods used to calculate the ab initio PES or the resolution of the experimental spectra. Additionally, many force fields, while covering a broad range of chemical structures, are not easily transferable to any type of system. There is currently no universal force field that would provide an accurate potential energy description over various drastically different chemical environments. Furthermore, most conventional classical force fields are hindered by a fixed bond topology; the bonds are not allowed to break or form throughout the simulation: a necessary component of chemical reactions.

1.4. Incorporating Reactivity in a Computationally Efficient Manner

One way to bypass the problem of modeling reactivity in MD is through the use of combined quantum mechanics molecular mechanics (QM/MM) simulations.[69-72] With a QM/MM description one can treat solvent effects explicitly, as well as large molecules like proteins, or really any non-reactive part of a system under study, with a molecular mechanics description, while using a quantum description for the reactive site. These methods have enhanced speed when compared to full ab initio simulations and still allow for the reactive center to be treated by first principles. However, QM/MM has a number of difficulties to manage as well.[52, 70-72] In order to effectively use QM/MM, previous knowledge of the location of the reactive sites is important. In addition, when looking at effects like proton shuttling, one would have to include a large number of solvation molecules within the QM sphere as proton diffusion rates are large. The larger number of atoms within the QM region, the slower the simulation occurs because QM/MM simulations are limited by the time taken to calculate the QM section as the MM section cannot evolve on differing time scales. Another issue arises when determining how to describe the boundary

between the QM and MM regions.[51, 71-73] Many studies have found that the manner in which the boundary region is treated can drastically change the energetics and dynamics of the simulation.

Due to the need for expedient reactive models, as well as models that do not necessarily restrict reactivity to one site of the system, reactive force fields have been developed. There are a number of reactive force fields that differ in model approach, computational expense, and performance. In general, there are two very distinct approaches to treat reactivity in a force field: 1) use of bond orders to allow for reactivity[74-79] 2) coupling of empirical potentials to achieve reactivity.[5, 14, 80-105]

One of the early bond order reactive force fields is the so-called Bond Energy Bond Order (BEBO) potential.[74] BEBO suffers from the lack of ability to describe the equilibrium bonds in more complex systems. This model is also limited by the assumption that the bond order is the same in both the reactant and product. The Reactive Bond Order Method (REBO) uses other existing models and added bond orders to allow for reactivity and can be seen in equation 1.6.[27, 75] In the original algorithm, partial charges and long range interactions were neglected. The attractive term is factored by a bond order term which determines the strength of the interactions, and b_{ij} is a function based on the angles and coordination of the atoms as well as conjugation and dihedrals.

$$E_{binding} = \sum_{i} \sum_{j>i} \left[V^{repulsive}(r_{ij}) - b_{ij} V^{attractive}(r_{ij}) \right] \tag{1.6}$$

The first iteration of the REBO model was made for large carbon crystal structures. The second generation of REBO modified the intermolecular potential, included Coulomb interactions, as well as parameterized to a larger number of structures. While the second generation greatly improved over the first iteration, the REBO force field still neglected a description of long term interactions. To this end, the Adaptive Intermolecular Reactive Bond Order (AIREBO) model was generated.[27] This improved over the other bond order models previously discussed as non-bonded interactions

were introduced. In order to accurately allow for reactivity and a proper description of non-bonded interactions, switching functions were added to turn off the long range interactions as the topology evolved.

Perhaps the most well-known reactive force field based off of a bond order scheme is ReaxFF.[27, 76-79] ReaxFF has shown significant and rigorous improvement over other bond order methods as it includes all typical interaction terms as seen in a classical force field, with other terms included to modulate the reactivity. The general reaction can be seen in equation 1.7.

$$E_{system} = E_{bond} + E_{over} + E_{under} + E_{val} + E_{pen} + E_{tors} + E_{conj} + E_{vdW} + E_{Coul} \quad (1.7)$$

In the spirit of the other bond order models, potential terms are tempered by a bond order term, which can get to be quite complicated. Torsion and valence angles are handled with switching functions so that as the bond breaks, the respective terms appropriately approach zero. Description for van der Waals interactions and Coulomb interactions are present. Additional correction terms were added to fix errors that may occur from the bond order description. There is a correction term that fixes any inaccuracies from over-bonding, or an atom participating in more bonds than it is physically able to. There is an undercoordination term that was introduced for when the valence terms are not fully filled. Additionally, there is a conjugated correction term which allows for that potential to approach zero as angles approach zero.

Use of ReaxFF has become more widespread. As studies continue using the formalism, parameter sets have been generated to allow for a more general use. However, despite the lofty goal of total transferability, ReaxFF parameters often require re-parameterization for new reactions.[51] This parameterization can be quite extensive as even systems containing only carbon, hydrogen, nitrogen and, oxygen have 611 parameters to fit.[68] The amount of parameterization makes general use of ReaxFF difficult and time-consuming.

Reactive methods based on a combination of empirical potentials exist. Perhaps one of the more primitive models is the adiabatic reactive molecular dynamics (ARMD) potential.[80] The assumption of ARMD is that two states, a reactant and a product are adiabatic; consequently, nuclear quantum effects and quantum tunneling, cannot be modeled with this potential. ARMD uses a potential for the reactant, V_r, the product, V_p, and adds a constant to the reactant such that the zero of V_r is the same as V_p. At each step, the change in potential is calculated by the potential of the reactant, V_r ,minus V_p plus the constant term. When the change in potential switches signs, a switching function, f(t), is used to turn off the reaction potential and turn on the product potential. The potential is given in equation 1.8. The drawback of ARMD is the requirement that the reactant and product states are adiabatically coupled, however, this is not always the case.

$$V = f(t)(V_r + \Delta) + \big(1 - f(t)\big)V_p \tag{1.8}$$

Another reactive force field used often in the literature is the Empirical Valence Bond (EVB) method[27, 82-88]. The mathematics and method behind EVB will be described in more detail in chapter 3, however, a quick overview is included in this section. In this formalism, the system is propagated as a linear combination of states that differ in their bonding topology, thereby allowing for a smooth transition from reactant to product. The EVB methodology requires prior knowledge of the reactions that will take place throughout the simulation, and generally has to be reparametrized for significantly different uses; however, it contains far fewer parameters in comparison to the bond order models, and is much faster than ab initio or QM/MM methods, allowing for a wide range of phase space to be modeled.

1.5. Purview of this Dissertation

The theme of this dissertation is the analysis of chemical reactivity pathways for different systems using both, and often, a combination of electronic structure methods to force field based

calculations. The layout is as follows. Chapter 2 introduces the problem of the solvated HCl system and describes both the ab initio molecular dynamics (AIMD) and ab initio metadynamics (MTD) simulations performed on HCl water clusters, and analyzes the different solvation environments that are seen. Chapter 3 details the development of a Multi-State Empirical Valence Bond (MS-EVB) based force field to accurately model the HCl water system in molecular simulations. Chapter 4 discusses electronic structure calculations and ab initio metadynamics simulations performed to determine the factors guiding regioselectivity experimentally seen in the addition of indole to unsymmetrical silyloxyallyl cations with an implicit solvent. Chapter 5 covers DFT investigations into charge transfer and the possible formation of an electron donor acceptor complex hypothesized in O-glycosylation experiments. Each chapter will contain background on simulations and previous literature for each project type. Finally, chapter 6 will discuss conclusions and future directions.

Chapter 2. Ab Initio Investigations of the Hydrated HCl System

2.1. Literature Review: An Overview of Previous HCl Computational Studies

Acid dissociation and the fundamental physical properties that determine the pathways of dissociation have been the question of numerous studies. In particular, the ionization of covalent HCl has warranted several molecular level computational simulations with both MD and MC approaches. Despite the volume of studies focused on this subject, many uncertainties arise due to the challenges faced when modeling HCl. Hydrogen chloride, while ionic in bulk solvation situations, has strong covalent character when the species exists in the gas phase. In fact, it is expected that a significant amount of covalent character is present in highly concentrated solutions.

There are many practical topics that involve covalent HCl. One such application involves chemical transformations of HCl in the atmosphere, particularly reactions which lead to the ozone layer depletion crisis. Gaseous HCl has been found to undergo reactions in the stratosphere creating the chloride ion, which undergoes subsequent chemical transformations to possible disastrous effects, as it reacts with other species and diminishes the amount of ozone in the atmosphere.[106-112]

It has been established that acidic species play a considerable role in atmospheric aerosols as a result of the high concentration of acids in the atmosphere.[113-116] HCl particles, notably, are expected to participate in atmospheric reactions.[103, 110, 111, 113, 117-119] Particles released in the atmosphere may undergo chemical transformations and can cause changes to the composition of the atmosphere.[112-115, 119] Moreover, these aerosols can potentially act as nuclei for ice or liquid particles which may result in cloud formation.[114, 118] Aerosols affect the scattering and absorption of incoming radiation, and, as such, have a large impact on the natural balance of the scattering vs absorption process. HCl exists naturally in marine environments[113, 118] and can be released into the

atmosphere through numerous processes such as combustion.[120] Recent studies have focused on HCl trapping by polycyclic aromatic hydrocarbons and the subsequent water addition due to the chlorine attached to the surface.[118] Due to this wide, although by no means exhaustive, list of covalent HCl either entering the atmosphere or participating in processes in the atmosphere; it is important to understand the chemical transport and transformations of HCl and other ubiquitous species present.

Many investigations of HCl systems focus on the fundamental physical proton transfer in small clusters containing water or other various atmospherically relevant species. For instance, much debate has existed over the number of water molecules required for HCl ionization. While it is widely accepted that HCl is not fully dissociated upon the addition of three water molecules,[112, 117, 121-123] many calculations have shown that HCl and four water molecules will lead to a dissociated, zwitterionic cluster.[107, 109, 124-127] Indeed, despite the presence of a minimum structure consisting of the solvent separated ion pair (SSIP), many of these calculations have found that there is a covalent minimum that is similar in energy.[107, 109, 125] Some ambiguity remains as other investigations claim that the dissociation process begins upon the addition of a fourth water molecule; however, the dissociation mechanism is not entirely accomplished until five water molecules are present.[117, 121] While these explorations have found undissociated and dissociated zwitterionic HCl water clusters with four water molecules, the question remains of how the system advances from covalent to the SSIP structural motif. It has been found that cyclic ring complexes allow for proton transfer processes;[114] however the ring $HCl(H_2O)_4$ structure is undissociated.[125] Metadynamics simulations were performed to resolve the pathway of ionization for clusters containing HCl and four water molecules.[125] This study determined the classical reaction pathway in which the structure began at the undissociated ring structure, traveled through a contact ion pair

(CIP), and finally resulted in the zwitterionic solvent separated ion pair (SSIP). While calculations were able to reproduce this reaction mechanism, there were large energy barriers between each step in the pathway yielding an improbable reaction process. Aggregation studies effectively circumvented the large activation barriers between isomers by allowing for the approach of a fourth water molecule to the center of the $HCl(H_2O)_3$ ring structure, supporting dissociation and solvent separation.[125] Studies on larger water structures have found that ionization is a stepwise reaction in which the HCl approaches the surface and dissociates with a low activation barrier followed by solvent reorganization to allow for an SSIP structure, that requires some activation.[128]

Based off the general form of the reaction pathway for ionization of HCl upon increasing hydration or accommodation to various surfaces; it is clear that studying new surfaces, extended surfaces, or systems containing covalent HCl, requires the ability to accurately and efficiently include explicit solvation as well as ionization and subsequent solvent rearrangement. Furthermore, reactions containing NO_x, HCl, and water have demonstrated a significant amount of charge transfer between HCl and its corresponding water as the HCl starts to dissociate,[119] indicating fluctuations between the undissociated HCl/H_2O and the Cl^-/H_3O^+ contact ion pair (CIP). One study attempted to model Cl^- utilizing a classical force field where parameters were tuned to model the CIP environment, however the previous investigations determining charge fluctuations as well as further ab initio studies comparing results of highly concentrated hydrochloric acid solutions to X-ray absorption fine structure (XAFS) experiments, it is indubitably clear that this method fails at capturing the real nature of the CIPs.[129]

As it turns out, accurately modeling CIP Zundel-like states is essential. Two studies have found nuclear quantum effects (NQE) dictate the character of CIPs.[124, 126] Essentially, the NQE results in similar energies between the CIP and SSIP of $HCl(H_2O)_4$ clusters. In fact, the quantum

effects are responsible for pushing the hydrogen back towards the chloride after dissociation.[124] The proton is essentially delocalized in the CIP Zundel-like form,[126, 130] allowing for shuttling back and forth between the chloride and oxygen. Numerous investigations have examined CIPs and the presence of the Zundel-like ion.[129-131] One such study explored high concentrations and determined ion pairing to be a prevalent interaction; their calculations agreeing with EXAFS data.[130] At 6 and 10 m concentrations, ions mainly exist in the CIP form.[129] This CIP interaction is Zundel-like[129] and contains a mixture of covalent and ionic character.[131]

Computational investigations at the molecular level of HCl ionization with increasing hydration, gas-liquid nucleation in the presence of HCl defects, etc., require significant sampling of phase space that is unachievable using ab initio/electronic structure based methods. While classical force fields are an attractive alternative for molecular simulations of these processes, these methods rely on a fixed bond topology and are therefore unreactive. The lack of description for reactivity is a major problem in these systems since they will be unable to model proton transfer/shuttling events, the so-called Grotthuss shuttling,[132, 133] of the acidic proton from the Cl to water and between water molecules. Additionally, to accurately model solvent rearrangement[89] and charge transfer, especially seen in CIPs, an accurate force field is required. A DFT description of the PES can be utilized to determine the number of water molecules required for HCl dissociation in small clusters; however, this sheds little insight on the dissociation pathways itself, particularly in large systems or extended surfaces. MC (using QM based energies) studies have been performed for the analysis of HCl ionization in water[128] or the pathway of ionization in the stratosphere upon addition to ice particles.[106] These different investigations were limited to a small number of water molecules and HCl. Combined quantum mechanics and molecular mechanics (QM/MM) simulations were also utilized and determined varying amounts of covalent and ionic

species throughout the ionization process.[108] However, QM/MM suffers from several drawbacks when modeling an excess proton, due to the subsequent transport mechanism upon the dissociation of hydrochloric acid in water.[72] The QM section must be large enough to facilitate an accurate description of the transport mechanism which essentially negates the increase of computational speed by using the QM/MM method, as the shuttling events span large amounts of water molecules. However, if the QM region is not appropriately spacious, the excess proton may shuttle until it reaches the MM region, where no QM oxygen is available for bonding, creating an artificial wall, essentially blocking the shuttling mechanism. This introduces unphysical interactions at the border region lending to an inaccurate description of energetics and dynamics.[73] Development of an "on the fly" partition process of the QM to MM regions has made progress in the past few years, however, these methodologies are still quite computationally expensive, and problems arise when determining how to identify the hydronium molecule.[70, 72, 73]

Due to this wide range of data showing differing solvation structures and ionization rates, it is clear that in order to accurately model HCl ionization, one must include an accurate description of both covalent and ionic HCl. An attempt to model these systems with reactive force fields has been attempted. Voth and co-workers used the reactive MS-EVB algorithm for an excess proton in water, with the addition of the chloride ion.[89] This work shows improvement over classical force fields as it models proton shuttling and allows for solvent rearrangement, however it does not include covalent HCl, and as discussed previously, simply tuning parameters for ion pairing does not allow for an accurate description of the surprisingly complex CIP environment. Wick generated an MS-EVB model with covalent HCl, however, his model was parameterized on low lying CIP and SSIP structures of $HCl(H_2O)_4$ as well as the HCl-H_2O dimer, and as such may not

accurately model covalent HCl and CIPs at small cluster sizes or higher energetic structures at any cluster size.[103]

A new reactive force field will be developed to include the complex nature of the covalent HCl and transformation to the CIP motif. However, in order to develop the force field, initial studies of HCl water clusters using higher level quantum (electronic structure) calculations must be performed to provide the data to develop the reactive HCl-water force fields. The reactive HCl force field developed, which will be described in the next chapter, is intended to be utilized on systems such as gas-liquid nucleation events in the atmosphere, HCl and the approach to the air-water interface, and concentrated acidic calculations. Previous studies have been limited to small clusters, or accommodation to an ice surface, which while revealing a path of ionization, does not show the HCl character at larger cluster sizes. Therefore, in this chapter, the electronic structure investigations on $HCl(H_2O)_n$ where n can range up to twenty-two water molecules are presented and includes the study of the solvation patterns, hydrogen bonding networks, and location of the ions in these clusters. A large variety of structures ranging in energetics are explored. These calculations will be used for the development of the search algorithm, parameterization, and validation of the model discussed in the next chapter.

2.2. Methodologies for the Determination of Relevant Hydrated HCl Structures

In order to fully understand the various solvation structures of $HCl(H_2O)_n$ cluster systems, ab initio calculations were performed. The structures determined in this chapter will be used in the development of a reactive force field for the study of these binary systems in the next chapter. The hydrogen bonding network at the cluster sizes analyzed in this work is very different in comparison to bulk water. This is due to the presence of a counterion at a much higher concentration as compared to typical aqueous HCl solutions, as well as finite size effects.

The different structures of smaller clusters, those ranging from $HCl(H_2O)_2$-$HCl(H_2O)_6$ were determined using manual construction followed by minimization of the structures. Optimizations were performed with the Gaussian 09 suite of programs.[134] All relevant structures of $HCl(H_2O)_2$ to $HCl(H_2O)_4$ found in previous studies were also identified and verified. MP2 optimization with the aug-cc-pvtz basis set was utilized for clusters up to $HCl(H_2O)_5$. Due to computational expense, $HCl(H_2O)_6$ clusters had a different optimization scheme, using B3LYP with the aug-cc-pvtz basis-set and Grimme D3 dispersion.[135, 136] The MP2 method with the aug-cc-pvtz basis set was utilized to determine single point energies of these clusters.

With increasing cluster size, the number of potential minima relevant to the systems of interest increases. To this end, ab initio metadynamics simulations (MTD)[137] were executed on clusters ranging from $HCl(H_2O)_{7\text{-}12}$ and $HCl(H_2O)_{20\text{-}22}$ in the CP2K package.[138, 139] MTD is a molecular dynamics simulation in which the Free energy potential is biased in order to explore different areas of phase space in a guided manner, based on a set of collective variables (CVs) . In this work, the simulations were driven by two CVs simultaneously in order to generate different solvation environments. One CV used to explore the solvation environments of the system was the distance between the chlorine atom and center of the cluster. The center of the cluster, also referred to as the cluster COM, is determined by the center of mass (COM) of the oxygens. The second CV included the coordination between hydrogen and chlorine atoms. In order to keep the clusters from fragmenting, a reflective wall was placed along the pathway of the first CV. These simulations were meant to generate a wide variety of solvation structures rather than to determine the Free energy surface, and so hills were added to the CV coordinates every 8 fs, with a width of 0.4 Bohr, and height of 0.7 kcal/mol. Simulations proceeded until analysis of the trajectories illustrated a number of different configurations and hydronium states. These calculations were done in the NVE

ensemble at 300 K and the potential was described with the BLYP functional with Grimme D3[135,136] dispersion and the DZVP-MOLOPT-SR-GTH basis set[140] and Goedecker-Teter-Hutter (GTH) pseudopotentials[141] with a timestep of 0.8 fs. The Gaussian Plane Wave framework was used and had a real space cutoff of 400 Ry.[139] A wavelet Poisson solver was used as the simulations were non-periodic.[142] Ab initio molecular dynamics (AIMD) simulations were also performed using the same description of interactions. Calculations proceeded in the NVT ensemble at 300 K with a timestep of 0.5 fs. Simulations were performed until several proton transfers were observed. The trajectories produced from the simulations were taken and analyzed with an in-house code which was used to identify the solvation type of the cluster [covalent HCl, contact ion pair (CIP), solvent separated ion pair (SSIP) or no ion pairing (NoIP)] as well as the distance between Cl^- and the center of the cluster. Configurations were then selected based on their relative energetics, solvation identity, and Cl^- cluster distance, and minimized in an effort to relax the structures into various minima on the Free energy surface. Optimizations of these initial coordinates taken from AIMD and MTD for the $HCl(H_2O)_7$-$HCl(H_2O)_{12}$ cluster sizes began with minimization using B3LYP with the 6-31+(d) basis set and Grimme D3 dispersion following the same protocol as used in the n=6 configurations for further optimization and stationary point calculations. Minimization of $HCl(H_2O)_{20}$-$HCl(H_2O)_{22}$ cluster sizes utilized B3LYP with the 6-31+(d) basis set and Grimme D3 dispersion. MP2 calculations for larger clusters requires an excessive amount of memory and computational time, and as such RI-MP2 calculations were used instead at the aug-cc-pvtz basis set. RI-MP2, or the so-called resolution of identity MP2, utilizes an auxiliary basis expansion scheme to reduce the computational expenditure associated with the electron correlation integrals.[143] This method of calculation has been used successfully for ionized

water clusters in the literature[144] and was tested on smaller clusters to ensure accuracy. RI-MP2 calculations were performed using the Orca package.[145]

2.3. $HCl(H_2O)_n$ Structural Evolution with Increasing Hydration

From here on out, all energy values refer to the potential energy and are given relative to the lowest energy structure at each cluster size.

2.3.1. $HCl(H_2O)_n$ where n=1-4

As these structures have been discussed in many other papers, a quick review of what is important pertaining to the solvation of these small clusters for future MS-EVB development is mentioned. The dihydrate structure yields one conformer in the shape of a triangular ring optimizing hydrogen bonding. The structure can be seen in appendix A.1. At the tri-hydrate level, two HCl configurations exist. The lowest energy structure is a square shaped ring with the HCl hydrogen bonding with two water molecules and can be seen in figure 2.1.a.

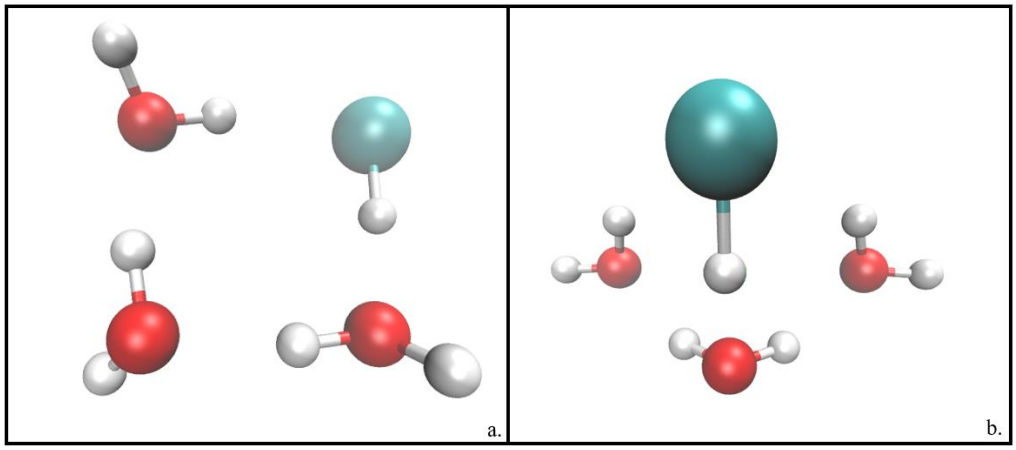

Figure 2.1. a. Lowest energy $HCl(H_2O)_3$ configuration b. second HCl $(H_2O)_3$ isomer, 3.8 kcal/mol higher than the lowest energy conformer

The second structure has all three waters coordinated to the hydrogen chloride, illustrated in figure 2.1.b, however, it is found to be 3.8 kcal/mol higher in energy than the ring configuration. The HCl bond length has increased to 1.39 Å in the higher energy structure from 1.34 Å in the ring

structure. This work also found the lowest energy conformation of $HCl(H_2O)_4$ to be the zwitterionic SSIP. A CIP was also found to be an energetic minimum, only 1.62 kcal/mol higher in energy than the SSIP structure. The HCl distance is 1.79 Å indicating both covalent and ionic character of the Zundel-like ion pair. The CIP and SSIP structures can be seen in appendix A.1. Two covalent HCl clusters were found at this system size, 3.4 and 4.6 kcal/mol higher than the SSIP energy minimum. The first HCl cluster is a pentamer-ring structure with two waters hydrogen bonding to the HCl. The second is a square ring with one outside water coordinated to two waters of the ring, creating a side triangular ring. These structures can be seen in figure 2.2.

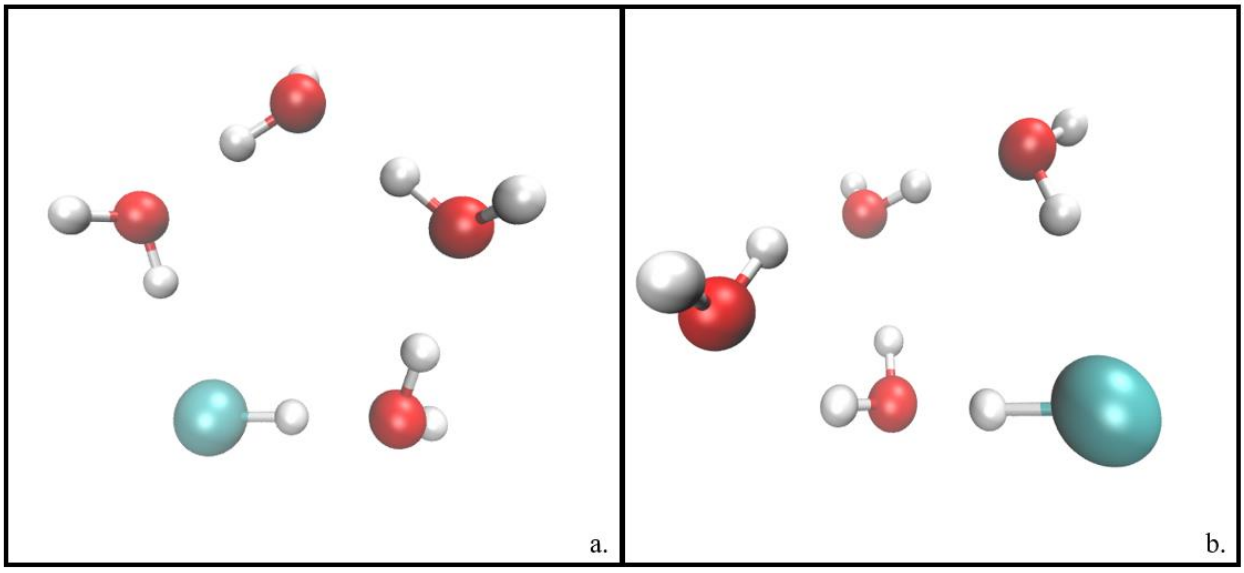

Figure 2.2. $HCl(H_2O)_4$ undissociated structures. a. is 3.4 kcal/mol higher in energy than the lowest energy structure b. has a relative energy of 4.6 kcal/mol

2.3.2. $HCl(H_2O)_5$ Clusters

Nine distinct structures were determined at this cluster size [1 HCl, 7 CIP, 1 SSIP, 0 NoIP]. As before the minimum energy structure is an SSIP configuration. It is similar to the $Cl^-(H_2O)_3H_3O^+$ zwitterionic SSIP found in the previous section with subtle differences. An Eigen cation $H_9O_4^+$ has formed between the hydronium and three water molecules, however, in this case,

only two of the waters participating in the Eigen cation are coordinated to the chloride ion, whereas the third Eigen water is hydrogen bonding to another water molecule, which subsequently coordinates to the chloride ion. This structure can be seen in appendix A.1. Seven CIP structures were found, two of which had a relative energy within 0.3 kcal/mol of the minimum structure. The two lowest energy CIP structures can be seen in appendix A.1. In all but one of the CIP structures, the hydronium and two water molecules were coordinated to the chloride ion. The HCl bond lengths ranged from 1.76-1.88 Å and the O-Cl bond lengths range from 2.80-2.87 Å. One undissociated $HCl(H_2O)_5$ cluster was found that is only 5.9 kcal/mol higher in energy than the SSIP configuration. This structure can be seen in figure 2.3. It has a book like structure, with two waters hydrogen bonding to the covalent HCl. The HCl bond length is 1.32 Å.

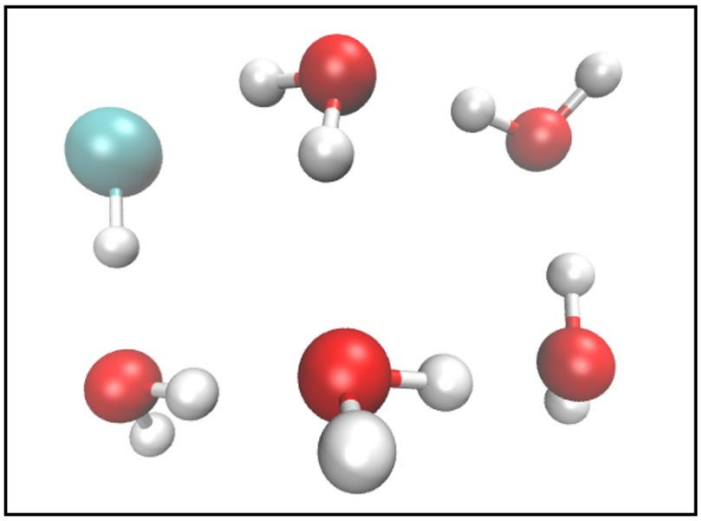

Figure 2.3. Undissociated HCl cluster, $HCl(H_2O)_5$ with a relative energy of 5.9 kcal/mol

2.3.3. $HCl(H_2O)_6$ and $HCl(H_2O)_7$

Eight clusters containing HCl and 6 water molecules were found [0 HCl, 4 CIP, 4 SSIP, 0 NoIP], 3 found upon the addition of a seventh water molecule [0 HCl, 0 CIP, 3 SSIP, 0 NoIP]. Of the eight structures at n=6, the lowest energy structure is SSIP and is 1.5 kcal/mol lower in energy than the next highest energy structure, which is a CIP. Both the energetic minimum SSIP

and lowest energy CIP can be seen in appendix A.1. Four structures are found to be SSIP with the second highest energy SSIP lying 2.4 kcal/mol above the minimum and the highest energy SSIP possessing a relative energy of 4.0 kcal/mol. Cl^- to the oxygen of hydronium distances range from 3.91-4.22 Å indicating how compact these small clusters are. The CIP motifs have relative energies ranging from 1.5-3.4 kcal/mol at this cluster size. H-Cl distances range from 1.86-2.03 Å and O-Cl from 2.88-2.99 Å indicating that these Zundel-like H_3O^+/Cl^- structures show both ionic and covalent character and are tightly bound.

All clusters at the n=7 level are CIPs and are within 3.6 kcal/mol of the lowest energy structure. HCl distances range from 1.99-2.02 Å and the O-Cl distance is 2.16-2.98 Å indicating these CIPs are less strongly bound than at the smaller clusters. It is possible that initial metadynamics simulations used to search for structures were trapped in a distinct energy well and that simulations were not run long enough to generate a larger variety of structures.

2.3.4. $HCl(H_2O)_9$ – $HCl(H_2O)_{12}$

Five clusters were found at n=9 resulting in SSIPs [0 HCl, 0 CIP, 9 SSIP, 0 NoIP]. The Cl^- to hydronium oxygen distance ranged from 4.32-4.46 Å. Two structures had five water molecules participating in hydrogen bonds to the chloride ion, while the other three structures had four water molecules bonding. The highest energy structure found had a relative energy of 9.1 kcal/mol; however, all the other structures fall within 4.1 kcal/mol of the minimum. The lowest energy structure can be found in appendix A.2.

Six clusters were determined for $HCl(H_2O)_{10}$ clusters [0 HCl, 1 CIP, 3 SSIP, 2 NoIP]. The lowest energy structure is of the SSIP motif with five water molecules coordinated to the chloride ion. The second lowest energy structure has no ion pairing (NoIP), seen in figure 2.4, with a relative energy of 1.0 kcal/mol, with a second NoIP structure found 3.1 kcal/mol higher than the

SSIP minimum. Both NoIP structures have three water molecules hydrogen bonding to the chloride ion. Two more SSIP structures are found, the first with a relative energy of 3.4 kcal/mol with five coordinating water molecules to Cl^-, and the other has a relative energy of 6.7 kcal/mol and three waters coordinating to the anion. One to two water molecules bridge the hydronium to the chloride ion, with the lowest energy structure having one bridging water molecule. One CIP is determined, with an HCl distance of 1.92Å. At this level of hydration, the CIP Zundel-like cation is perhaps more ionic than the smaller clusters, however, it is expected that some covalent character still exists. The lowest energy structure SSIP, NoIP, and CIP can be seen in appendix A.2.

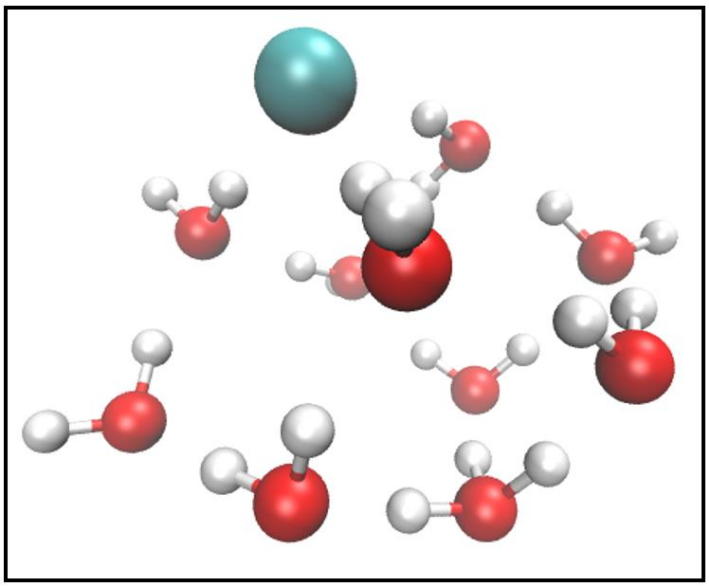

Figure 2.4. $HCl(H_2O)_{10}$ first no ion pairing structure determined with a relative energy of 1.0 kcal/mol

At the n=11 level, 16 structures were found [1 HCl, 6 CIP, 8 SSIP, 1 NoIP], and the lowest energy structure has no ion pairing. There are eight SSIP motifs with relative energetics ranging from 1.5 to 16.6 kcal/mol higher than the minimum structure. The lowest energy SSIP has six waters coordinated to the chloride ion, but the motifs range from three to six waters coordinated to the anion. All SSIPs have one or two waters bridging the hydronium to Cl^-. Six CIP structures

were found ranging from 4.2-18.4 kcal/mol higher than the NoIP minimum. H-Cl distances range from 1.7-1.97 Å and O-Cl distances from 2.82-2.97 Å for CIPs. One fully undissociated HCl structure is found at the $HCl(H_2O)_{11}$ cluster size. It is has a relative energy of 18.0 kcal/mol and only participates in a hydrogen bond with one water molecule. The oxygen of water is 3.06 Å from the chloride anion and the H-Cl bond distance is 1.32 Å. The lowest lying structure for each environment type (CIP, SSIP, NoIP) can be found in appendix A.2. The amount of hydration of the anion ranges extensively from one to six hydrogen bonds.

The lowest lying structure of $HCl(H_2O)_{12}$ clusters is a SSIP environment with one linking water between the hydronium and anion and can be seen in figure 2.5. Fifteen isomers were found [1 HCl, 3 CIP, 10 SSIP, 1 NoIP].

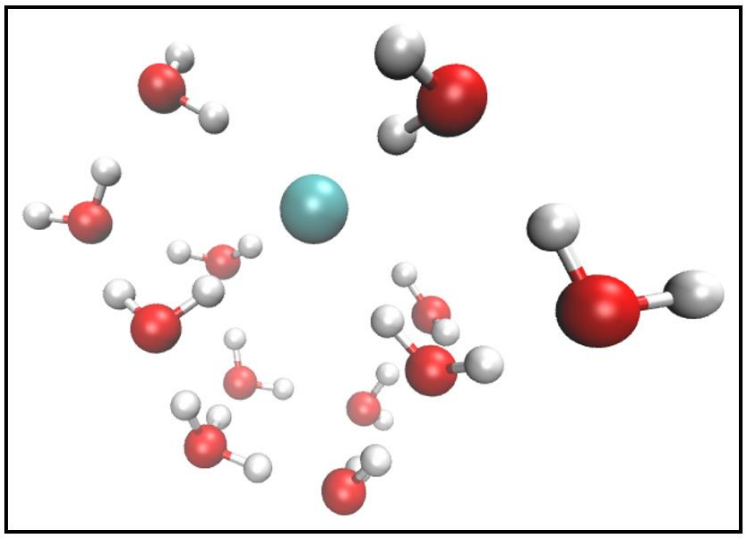

Figure 2.5. Lowest energy $HCl(H_2O)_{12}$ isomer, an SSIP motif

The highest energy SSIP configuration has a relative energy of 11.2 kcal/mol. Anywhere between one, to all three waters participating in the Eigen cation, coordinate to Cl^-. The SSIP structure where all three waters coordinating to the anion are also coordinated around the hydronium have a relative energy of 7.2 kcal/mol and can be seen in appendix A.2. It is essentially a hydrated

version of the n=4 SSIP zwitterion. The lowest lying CIP structure has a relative energy of 4.8 kcal/mol, the H-Cl distance of 2.00 Å and O-Cl at 2.99 Å indicating an evolution to slightly more ionic Cl^-/H_3O^+ Zundel-like states. However, the higher energy CIP isomers have similar O-Cl and H-Cl distances. The one covalent HCl found participates in three hydrogen bonds, has an H-Cl distance of 1.38 Å, an O-Cl distance of 2.89 Å, and is only 6.8 kcal/mol in relative energy. This structure can be seen in figure 2.6.

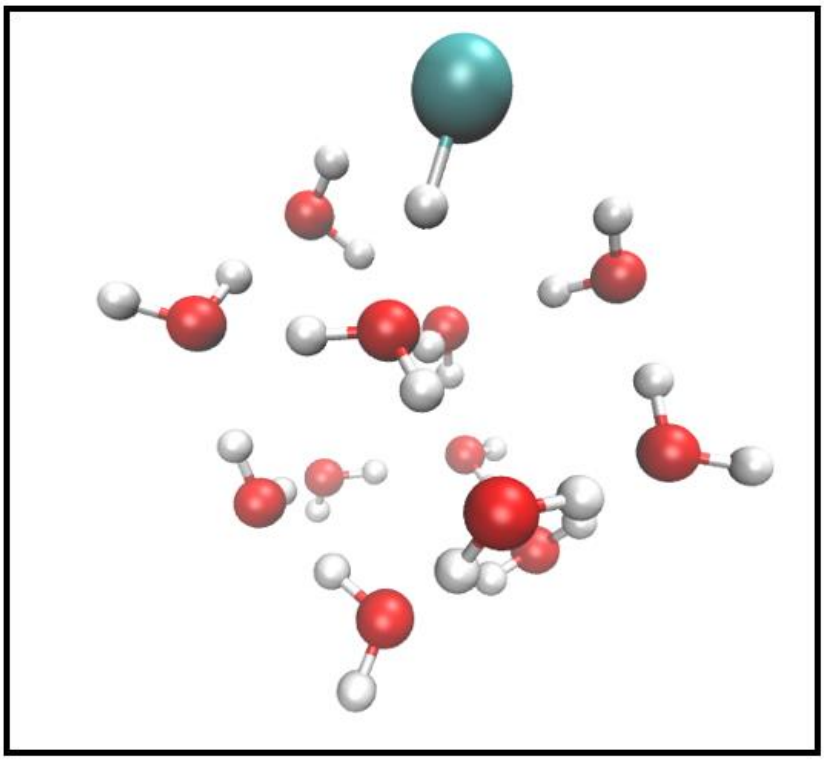

Figure 2.6. $HCl(H_2O)_{12}$ undissociated conformer

Only one NoIP configuration was found and is 7.2 kcal/mol in relative energy indicating that ion pairing is still favored at this cluster size. There are more water molecules coordinated to the water in the n=11 case, as three to six hydrogen bonds exist

2.3.5. $HCl(H_2O)_{20}$ – $HCl(H_2O)_{22}$

Thirteen structures were determined for $HCl(H_2O)_{20}$ clusters [0 HCl, 5 CIP, 4 SSIP, 4 NoIP]. The lowest energy structure was of SSIP configuration. Of the four SSIP structures found, all isomers only contained one water linking the hydronium to the chloride anion. Interestingly, one SSIP structure within 5.7 kcal/mol of the lowest energy configuration has the anion near the COM of the cluster and can be seen in figure 2.7.a. Five CIP structures were found with relative energies ranging from 4.4 to 19.5 kcal/mol higher than the minimum structure. HCl distances ranged from 1.97-2.06 Å and O-Cl from 2.99-3.04 Å like at the medium sized clusters. The highest energy CIP has the ion pair located nearly at the COM of the cluster and can be seen in figure 2.7.b.

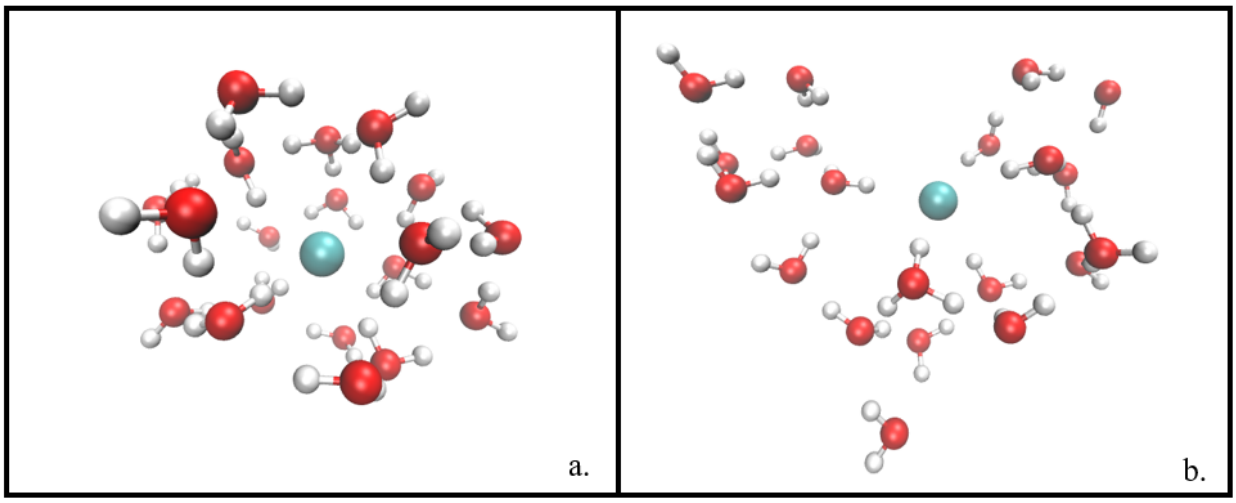

Figure 2.7. $HCl(H_2O)_{20}$ structures with chloride on the interior of the cluster a. SSIP motif with a relative energy of 5.7 kcal/mol b. CIP motif with relative energy of 19.5 kcal/mol

The comparison of the SSIP and CIP structures with the chloride anion towards the center of the cluster illustrates the lack of correlation between Cl^- location and energetics. Four NoIP structures were found with relative energies ranging from 4.5 to 8.3 kcal/mol. In the NoIP case, Cl^- tends to

stay at the exterior of the cluster. All solvation structures had the anion participating in four to five hydrogen bonds. The lowest energy structure of each solvation type is shown in appendix A.3.

Twenty $HCl(H_2O)_{21}$ isomers were found [1 HCl, 3 CIP, 7 SSIP, 9 NoIP], however, only four of these structures were within 5 kcal/mol of the lowest energy structure, which was found to have no ion pairing. In fact, all of the structures within 5 kcal/mol of the minimum energy structure consisted of the NoIP solvation type. Of the NoIP found at this cluster size, the location of Cl^- ranged from the center of the cluster to the outside edges, however, the Cl^- position does not determine the favorability as the location trends do not match the energetic trends. One undissociated structure was found, illustrated in figure 2.8. It was the highest energy isomer with a relative energy of 21.6 kcal/mol.

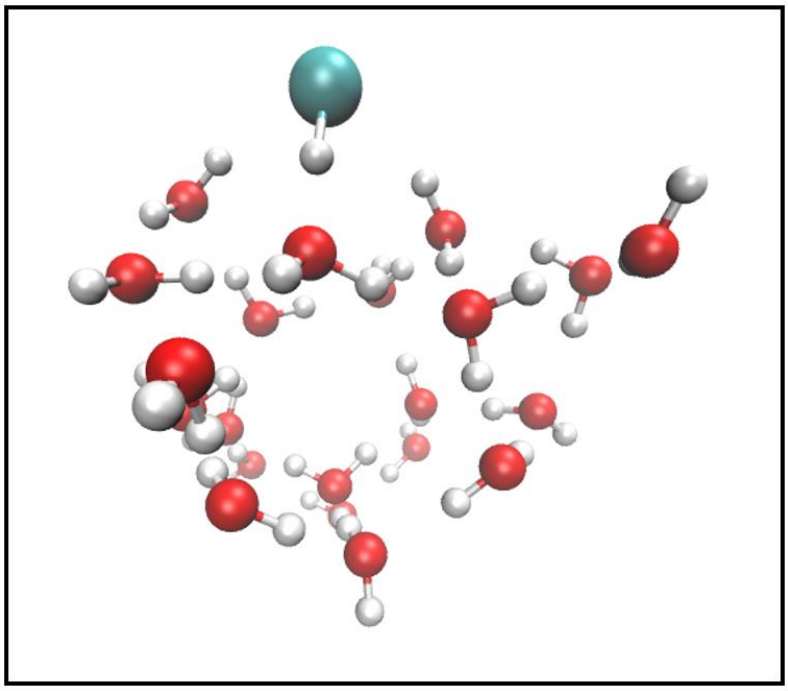

Figure 2.8. Undissociated $HCl(H_2O)_{21}$ cluster with a relative energy of 21.6 kcal/mol

The lowest energy SSIP has a relative energy of 6.3 kcal/mol, with the highest energy SSIP 14.7 kcal/mol higher than the minimum energy NoIP isomer. Three CIP structures were found, all were

high in relative energy ranging from 11.9-20.9 kcal/mol. It is located on the outside of the cluster and participates in three hydrogen bonds. The lowest energy structure of each solvation type not already pictured can be seen in appendix A.3.

The last structure size analyzed included 22 water molecules and HCl. Sixteen isomers were found [0 HCl, 1 CIP, 11 SSIP, 4 NoIP], however, no undissociated states were found at this cluster size. Interestingly, the lowest energy structure is SSIP where one water molecule links the hydronium and chloride ions. Relative energies of SSIP configurations ranged from the minimum energy isomer to 16.0 kcal/mol higher than the minimum structure. The majority of isomers found were SSIP. Only one CIP structure was found, however, it only has a relative energy of 8.05 kcal/mol with an H-Cl distance of 2.00 Å and O-Cl distance of 3.00 Å. Four NoIP configurations were found ranging in relative energies of 7.3 to 14.7 kcal/mol, however, the distance of Cl^- to the COM of oxygens varied and were not consistent with the energetic trends. The lowest energy structures of each solvation type can be seen in appendix A.3.

2.4. Solvation Motifs of $HCl(H_2O)_n$: From Covalent HCl to No Ion Pairing

Table 2.1 shows the number of distinct minima found at select cluster sizes where the size is given by n in $HCl(H_2O)_n$ as well as the percentage breakdown of each solvation type.

Table 2.1. Percentage breakdown of solvation patters that were seen at each cluster size. Cluster size is given by n of $HCl(H_2O)_n$ and number of minima refers to the number of distinct structures found at each cluster size

Cluster Size	Number of Minima	% HCl	% CIP	% SSIP	% NoIP
5	9	11.11	77.78	11.11	0.00
6	8	0.00	50.00	50.00	0.00
11	16	6.25	37.50	50.00	6.25
12	15	6.67	20.00	66.67	6.67
20	13	0.00	38.46	30.77	30.77
21	20	5.00	15.00	35.00	45.00
22	16	0.00	6.25	68.75	25.00

The cluster sizes were chosen for their relative novelty as well as representation of small to larger clusters. It is interesting to note that the NoIP configurations do not represent the majority of clusters found, even at n=22. This illustrates the requirement for model potentials to include accurate interactions of ion pairing. It also cannot be understated that a significant amount of CIP and SSIP configurations exist at every cluster size analyzed, with CIP configurations the majority solvation type at small clusters, and SSIP structures the majority at mid-sized clusters. In fact, roughly a third of all large cluster configurations found were SSIP.

While table 2.1 contained the percentage of each solvation type that was found at each cluster size, it is also important to determine which environment types dominate the low lying structures and which environments exist at high energetics. To this end, table 2.2 shows the energy ranges for each solvation type at the representative cluster sizes.

Table 2.2. Breakdown of the energetics that each solvation type spans at the given cluster size. Lower bounds give the lowest relative energy at the respective cluster and solvation types whereas the upper bound gives the highest relative energy. Cluster size is given by n of $HCl(H_2O)_n$ and number of minima refers to the number of distinct structures found at the given cluster size

Cluster Size	**Number of Minima**	**Energy Range HCl (kcal/mol)**	**Energy Range CIP (kcal/mol)**	**Energy Range SSIP (kcal/mol)**	**Energy Range NoIP (kcal/mol)**
5	9	5.9	0.3-5.1	0.00	-
6	8	-	1.5-3.4	0.00-4.0	-
11	16	18.0	4.2-18.4	1.5-16.5	0.00
12	15	6.8	4.8-17.9	0-11.2	7.17
20	13	-	4.4-19.5	0-8.1	4.5-8.3
21	20	21.6	11.9-20.9	6.1-14.7	0-20.0
22	16	-	8.1	0-16.0	7.3-14.5

Undissociated structures are still relatively low lying for both n=5 and n=12. The other HCl structures are higher in energy, however, whatever model potential chosen to model HCl-water

gas liquid nucleation, or HCl-water slabs at the air-water interface should still be able to accurately reproduce these values. CIP configurations have relatively low lying isomers present until 21 water molecules have hydrated the HCl. SSIP configurations are almost always the lowest energy structure and when they are not, they still span low lying isomers. The two cases where SSIP structures are not the lowest energy configurations are when NoIP structures are the most stable.

In order to fully illustrate the range of diverse structures found in terms of chemical environments in these clusters, table 2.3 shows a variety of properties at these cluster sizes.

Table 2.3. Range of solvation properties seen at each cluster size. Cluster size is given by n of $HCl(H_2O)_n$ and number of minima refers to the number of distinct structures found at the given cluster size. The third column lists the range of distances between the chloride anion and the COM of the cluster, defined in section 2.2. The fourth row tabulates the range of distances of the hydronium to COM of the cluster, and the fifth shows the ranges of free OH present at the given cluster size

Cluster Size	Number of Minima	Cl- to cluster COM (Å)	hydronium to oxygen COM (Å)	Number of Free OH
5	9	2.26-3.94	1.47-2.14	3 to 4
6	8	2.26-2.88	1.47-2.16	2 to 5
11	16	2.03-5.46	2.51-4.91	3 to 6
12	15	1.28-5.53	2.17-4.28	3 to 5
20	13	0.39-4.69	1.62-5.21	6 to 8
21	20	0.98-5.76	2.87-6.37	4 to 9
22	16	1.15-6.05	0.52-6.09	3 to 9

One property analyzed was the distance of the chlorine atom to the COM of the oxygens in the cluster and the second is the hydronium distance to the COM of the oxygens. The third property assessed is the amount of free OH groups in the cluster, or rather hydrogens attached to water molecules that are not involved in the hydrogen bonding network. This gives an idea of how tightly bound the clusters are. What can be determined from these structures is that a wide variety of configurations exist, and no distinct trend can be found. This in itself is important, because the

model developed in the next chapter must be able to reproduce energetics for these clusters which drastically vary in solvation type.

2.5. Conclusion of the Hydrated HCl System

The potential energy landscape for clusters ranging from $HCl(H_2O)_2$ to $HCl(H_2O)_{22}$ have been explored in detail. Structures up to n=6 were generated by hand, whereas MTD simulations were used to explore a wider range of phase space for larger cluster sizes. One hundred and nineteen structures in all were found. Of these, 9 undissociated HCl motifs were determined, 34 CIP, 55 SSIP, and 21 NoIP configurations were explored. This analysis shows the importance of a model potential to incorporate an accurate description of ion pairing as about 82% of the structures analyzed here have some amount of ion pairing.

Despite the differing cluster sizes, as well as a clear definition of ion pairing, it is difficult to determine the driving factors that govern the energetic favorability. The range of energetics at each cluster size shows a mixture of HCl, CIP, SSIP, and NoIP often relatively close in energy. The amount of hydrogen bonding to the chloride anion was determined for small and medium cluster sizes and showed no general trend corresponding to the increase in relative energy. Both Cl^- and H_3O^+ ion distances to the COM of the clusters were analyzed, and in this case, there was no general trend of distance related to the energetic scale. Lastly, the number of free OH groups were determined. One would expect that as energy increased, the amount of free OH would also increase as they are not participating in stabilizing effects of the hydrogen bond network. However, this trend is not present. Essentially, despite the seemingly relative simplicity of binary $HCl(H_2O)_n$ systems, the hydrogen bonding network is quite complex.

The next chapter focuses on the development of a reactive force field in order to study these complex water acid systems. The force field must accurately reproduce the energetics found

in these cluster configurations. A reactive force field based off an existing reactive hydronium model will be used. The fit set will contain the first 28 structures, ranging from n=2-8 where ion pairing, particularly CIP and undissociated HCl, are present at low relative energies, with the expectation that as the ion pairing decreases, the underlying hydronium model will take over. In order to validate the force field, the 91 structures ranging from n=9-22 will be used.

Chapter 3. Development of a Reactive Potential for HCl Water Systems

3.1. Introduction to the Problem of Modeling Acid Defects in Water Clusters

In the introduction to the previous chapter, the importance of modeling HCl, from its environmental impact within the atmosphere to concentrated acidic solutions, was discussed. Covalent HCl exists in the gas phase but can readily ionize upon the addition to a surface such as ice or polycyclic aromatic hydrocarbons, or in bulk water. Furthermore, once HCl ionizes in water it can undergo proton transport via the Grotthuss shuttling mechanism. In this mechanism, which has been the subject of many detailed investigations, an excess proton can traverse multiple water molecules in a very short amount of time with very little structural rearrangement of the solvation environment; essentially delocalizing the charge over multiple water molecules.[97, 132, 133] For any study focusing on concentrated acid solutions or atmospheric implications, such as gas-liquid nucleation or HCl at the air-water interface, it is necessary to model covalent HCl, the dissociation process, and subsequent shuttling mechanism/charge delocalization. Essentially, the species is reactive, and in order to properly model the physics of the system, a reactive description is required.

Ideally one would use ab initio methods in order to describe the potential energy of HCl water systems. Ab initio methods do not require further parameterization and contain all the physics required to model the reactivity. However, these methods require a substantial amount of computational time which makes their use for extended simulations where a large amount of sampling or long simulation times are required improbable. Classical force fields facilitate an efficient alternative to phase space sampling and obtain long simulation times, however, most conventional classical force fields are restricted to a fixed bond topology, and as such, cannot model the reactivity of these systems. The QM/MM description of the potential has the ability to

model one section of the simulation with high level ab initio descriptions, while the rest of the system is described with a classical force field; however these calculations are still relatively expensive, are limited in speed by the QM region, and require careful treatment of the QM/MM boundary section.[71-73] The mistreatment of the boundary can yield inaccurate results. Furthermore, the QM section for systems that have proton shuttling has to be prohibitively large, otherwise one risks the excess proton shuttling from the QM region to the MM region which will also result in errors in the potential energy description due to artifacts in the border region.[71, 72] To this end, reactive force fields allow for a more computationally efficient method of describing the reactive potential energy surface in which no treatment of boundary conditions are necessary. The goal of this work is to build a multi-state empirical valence bond (MS-EVB) algorithm that describes covalent HCl as well as the ionic Cl^-/H_3O^+ water systems.

3.2. The Two State Empirical Valence Bond Formalism

The Empirical Valence Bond (EVB) formalism was developed by Warshel in order to simulate covalent to ionic states in aqueous solutions with active enzymes.[82] In order to illustrate the method, an example relevant to the current research, the HCl-water dimer, is considered. Figure 3.1 shows the HCl water dimer in two states, one where HCl is covalent, and the other where the system is ionic. Both states have the same coordinates, but they vary in bonding topology. While both states are physically possible, with a classical force field only one of the states can be sampled during the simulation, either state $|1\rangle$ described by the H_{11} force field or state $|2\rangle$ described by the H_{22} force field.

The EVB methodology allows for a seamless transition between the two states due to the inclusion of the coupling term. In this formalism, the system Hamiltonian is written as a linear combination as seen in equation 3.1.

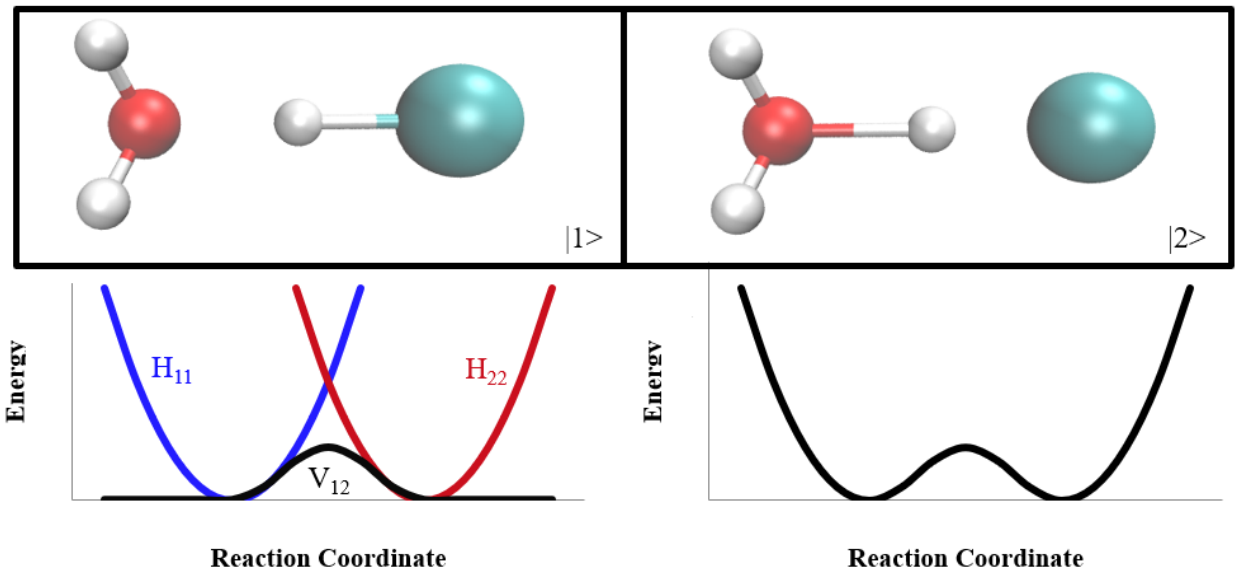

Figure 3.1. The top panel shows the two possible dimer states |1> the undissociated HCl and |2> the CIP dimer. The bottom left panel is an illustration of two potential energy wells described by a classical potential H_{11}, and H_{12}. With the classical potential, only one state will be sampled. The coupling term V_{12} couples the states so a smooth transition can occur as seen in the bottom right panel.

$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{12} & H_{22} \end{bmatrix} \tag{3.1}$$

The diagonal elements are described by a classical force field for each state as seen in equations 3.2 and 3.3.

$$H_{11} = V_{H_2O}^{intramolecular} + V_{HCl}^{intramolecular} + V_{H_2O-HCl}^{intermolecular} \tag{3.2}$$

$$H_{22} = V_{H_3O^+}^{intramolecular} + V_{H_3O^+-Cl^-}^{intermolecular} \tag{3.3}$$

The off-diagonal term is a function that couples the two states and can be parameterized to ab initio calculations. This is an eigenvalue problem by construction, as seen in equation 3.4, and so diagonalization of the matrix yields two eigenvalues and corresponding eigenvectors.

$$HC = EC \tag{3.4}$$

The lowest energy eigenvalue, E_0, corresponds to the ground state energy and the system is propagated along the ground state energy surface. From the corresponding ground state

eigenvector $C_0=[C_1,C_2]$ the contribution of each state, calculated by the square of each element, C_i^2, can be extracted.

The EVB methodology has been successfully used in many studies.[27, 82-88] A wide range of reactions have been investigated, including proton transfer in 2-pyridone to 2-hydroxypyridine,[84] Claisen rearrangement in allyl vinyl ethers,[84] charge delocalization in water wires,[83] and a benchmarking study of methane plus a chlorine atom.[86] While the work on the water wire system was able to determine the amount of delocalization of an excess proton between two water molecules, the EVB formalism is limited to two states; consequently, the description of shuttling is restricted to the two potential hydronium sites. Due to this limitation, the excess charge can only delocalize across two water molecules, when in reality more than two states may be accessible to the system.

3.3. Multi-State Empirical Valence Bond Formalism

Voth et. al. expanded the EVB methodology to a multi-state method.[14, 89-101] The extension to this multi-state interpretation is relatively straight-forward in theory but becomes quite complex in practice. To illustrate the multi-state version for this project, the $HCl(H_2O)_4$ cluster is used. Seven possible states exist, four hydronium states and three covalent HCl states, as seen in figure 3.2. As before, the coordinates in each state are the same, but the bonding topology is different. The Hamiltonian for this $HCl(H_2O)_4$ cluster can be written by the 7x7 matrix that is depicted in equation 3.5 where the diagonal terms are given by the classical force field and the off-diagonal terms are again used to couple the states which share a proton. However, unlike the Hamiltonian description of the two state EVB model where a singular reaction step is occurring, not all configurational states in the MS-EVB formalism share a proton, and therefore not every off-diagonal element has a contribution to the potential. For instance, state |1> and state |2> share a

bonded proton that transfers through a Zundel cation, however, states |1> and |7> do not share a proton; as such, H_{17} is simply zero in the matrix. This reduces the Hamiltonian matrix for this system to equation 3.6.

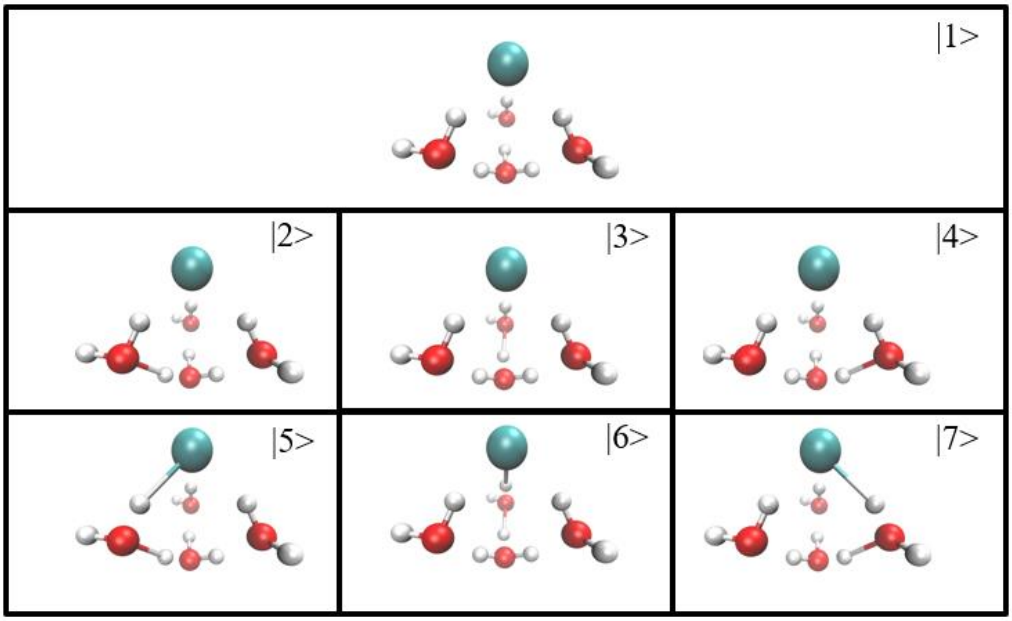

Figure 3.2. The seven possible configurational states in the SSIP $HCl(H_2O)_4$ isomer. States |1> through |4> show the possible hydronium states, while |5> through |7> depict possible covalent HCl states

$$H = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} & H_{15} & H_{16} & H_{17} \\ H_{12} & H_{22} & H_{23} & H_{24} & H_{25} & H_{26} & H_{27} \\ H_{13} & H_{23} & H_{33} & H_{34} & H_{35} & H_{36} & H_{37} \\ H_{14} & H_{24} & H_{34} & H_{44} & H_{45} & H_{46} & H_{47} \\ H_{15} & H_{25} & H_{35} & H_{45} & H_{55} & H_{56} & H_{57} \\ H_{16} & H_{26} & H_{36} & H_{46} & H_{56} & H_{66} & H_{67} \\ H_{17} & H_{27} & H_{37} & H_{47} & H_{57} & H_{67} & H_{77} \end{bmatrix} \tag{3.5}$$

$$H = \begin{bmatrix} H_{11} & H_{12} & H_{13} & H_{14} & 0 & 0 & 0 \\ H_{12} & H_{22} & 0 & 0 & H_{25} & 0 & 0 \\ H_{13} & 0 & H_{33} & 0 & 0 & H_{36} & 0 \\ H_{14} & 0 & 0 & H_{44} & 0 & 0 & H_{47} \\ 0 & H_{25} & 0 & 0 & H_{55} & 0 & 0 \\ 0 & 0 & H_{36} & 0 & 0 & H_{66} & 0 \\ 0 & 0 & 0 & H_{47} & 0 & 0 & H_{77} \end{bmatrix} \tag{3.6}$$

When the matrix is diagonalized, seven eigenvalues and eigenvectors will be determined and the lowest energy eigenvalue, E_0, is the total energy of the system. The contribution of each bonding configuration can yet again be calculated as the square of each element of the corresponding eigenvector $C_0=[C_1,C_2,C_3,C_4,C_5,C_6,C_7]$. Unlike a classical force field, the charge given by the ions are delocalized across the cluster, due to the inclusion of multiple states. In this case, a majority of the contribution to the energy is coming from the first state.

The MS-EVB formalism has been used to investigate proton transport in a number of different systems including hydronium and bulk water,[14, 81, 91-96, 102] hydronium and small water clusters,[93, 96, 97, 102] ions at interfaces,[5, 91, 93, 100, 103, 105] concentrated acid solutions,[89, 101] hydroxide,[105] ionizable amino acids in proteins,[98] water-filled channels,[105] polymer electrolyte membranes,[95] aqueous interfaces,[105] and other biological systems such as human carbonic anhydrase.[90] In fact, the multiconfigurational formalism has recently been extended into a coarse grained exploration of the folding of dodeca-alanine.[146]

A few MS-EVB studies have focused on hydrogen chloride. One such study analyzed concentrated acid solutions and did not include covalent HCl.[89] The underlying assumption in that investigation hinged on the idea that covalent HCl is negligible, even at high concentrations. However, as discussed in section 2.1, previous studies have determined that Lennard-Jones parameters of Cl^- and H_3O^+ in tandem with mixing rules do not suffice in the description of contact ion pairs (CIP). Another study, by Wick, developed an EVB approach to examine HCl accommodation to a bulk water slab.[103] There are a few notable differences between this work and the HCl model by Wick, which will be discussed in detail in the development section. Nevertheless, a general comment on the differences between Wicks model and the MS-EVB-HCl algorithm developed in this work, is that the formers model was developed for immediate

dissociation upon HCl addition to the bulk water slab, and is parameterized to only a few low lying cluster structures including the HCl-water dimer, the lowest energy SSIP $HCl(H_2O)_4$ zwitterionic structure, and the CIP $HCl(H_2O)_4$. Wick's model is also fit to the H-Cl PES of dissociation for the HCl-water dimer and the CIP $HCl(H_2O)_4$. The model in this work, MS-EVB-HCl, was developed with the intent to study HCl induced nucleation, concentrated solutions, and HCl at the air water interface where a significant amount of covalent HCl and H_3O^+/Cl^- CIP systems are to be expected, as well as distributions of cluster structures at the higher range of the energetic scale rather than simply global minima for a few small structures.

3.4. Development of MS-EVB Model Which Includes Covalent HCl

The recipe for an MS-EVB algorithm requires a search algorithm to determine all possible configurational states, inclusion of an accurate potential for the diagonal states, an algorithm to find any species pair in which a proton transfer reaction may occur, a potential form for the coupling term, and parameters to describe the extent of interactions, followed by diagonalization of the Hamiltonian matrix.

3.4.1. Diagonal State Search Algorithm

It is essential that all relevant bonding topology states are incorporated for a given configuration. This MS-EVB-HCl algorithm is based on the MS-EVB3.2 model.[91] The MS-EVB3.2 uses the same search algorithm as MS-EVB3,[96] which improves upon the MS-EVB2 algorithm.[92] The MS-EVB2 assigned hydronium states by determining the initial hydronium, or the pivot hydronium. The pivot hydronium is defined as the oxygen with three closest hydrogen molecules to it. From the pivot hydronium, the second closest oxygen to the protons is determined. The proton, second closest oxygen, and the hydrogens bound to that oxygen already are labeled as the new hydronium state. This process is iterative until three solvation shells are found. The MS-

EVB3 search algorithm was slightly altered in order to find every possible hydronium state within three solvation shells of the pivot hydronium, as some possible states were not found in the MS-EVB2 algorithm. Once the pivot hydronium is found, instead of finding just the second closest oxygen to the pivot hydronium proton, the algorithm searches for any oxygen within 2.5 Å. This allows for oxygens to participate in more than one hydronium state as the updated model finds any bifurcated oxygen.

Once all of the hydronium states are located, the covalent HCl states must be found. This search is slightly more complicated. The algorithm works as follows: 1) Any hydrogen within 2.3 Å of the chlorine atom is located. This hydrogen is labeled as the covalent hydrogen. From the covalent hydrogen, the topology of the water must be found 2) The closest oxygen to the covalent hydrogen is found. This is where the algorithm splits into two possible pathways where the closest oxygen to the covalent hydrogen a) only has a water identity from the hydronium search or only exists in one possible hydronium state, or b) participates in two hydronium states. Following the split through the a) pathway then the next step is 3.a) Here, the next two closest hydrogens that are not the covalent hydrogen are found, and together with the closest oxygen are labeled as a water molecule. 4.a) From this first water oxygen, the next closest oxygen is found, followed by the next two closest hydrogen atoms, and labeled as a second water molecule. 5.a) This process repeats until all water molecules have bonding type assigned so that there are no unintended free floating H^+ or OH^- groups.

If the closest oxygen has two hydronium states, the water topology search becomes slightly more complicated and computationally expensive. This is because, while one covalent hydrogen is found, the water topology can split in two different ways, as can be seen in figure 3.3.

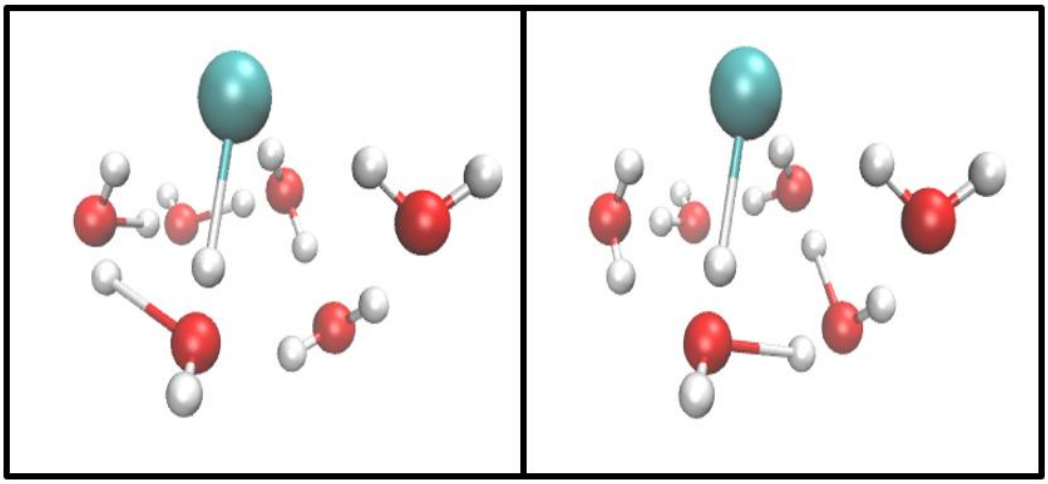

Figure 3.3. Two MS-EVB states for the system have the same covalent HCl, however, the water bonding topology is different.

Both water topologies must be accounted for, and as such this covalent HCl state will actually exist in two configurational states. In this way, starting from step 2, the determination that the closest oxygen to the covalent hydrogen participates in two hydronium states leads the algorithm to deviate towards path b) and then 3.b) The topology of the corresponding first hydronium is called, then 4.b) The covalent hydrogen is removed from bonding with the hydronium and is reassigned to the chlorine atom, and the rest of the water topology is taken from the corresponding first hydronium configurational state 5.b) The topology of the second hydronium for the closest oxygen is called next 6.b) The covalent hydrogen is reassigned from the second hydronium to the chloride ion. Followed by 7.b) the topology of the water is taken from the second hydronium configurational state. The HCl search slightly impedes the speed of the search, as there are many criteria that must be examined in this loop of code to ensure the water topology is found without splitting water. The reason behind the inclusion requirement of both configurational states differing in water topology has to do with the search for the off-diagonal states and will be discussed in the next section.

3.4.2. Off-Diagonal State Search

In order to incorporate reactivity, all configurational states that share a proton must be found. In this case, the topology of any species not participating in proton transfer must be the same. For instance, as illustrated in figure 3.4, it can be seen that the water topology of the two MS-EVB states are the same, with the only difference in topology arising from covalent HCl to CIP, in the first state the configuration is an H_3O^+/Cl^- ion pair, whereas in the second state the HCl is covalent and is hydrogen bonding with the water molecule which has hydronium identity in the first state.

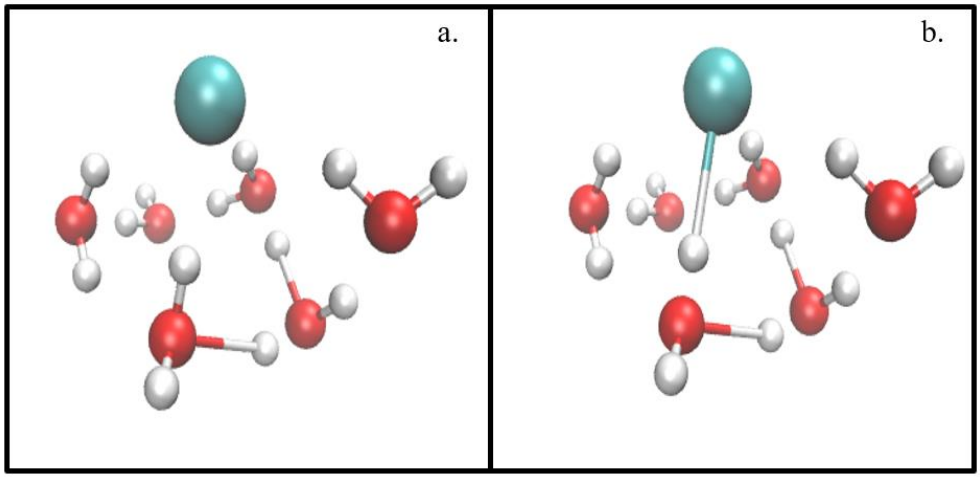

Figure 3.4. Illustrates the off-diagonal coupling between state a. which has the HCl configuration and state b. of the CIP motif. The excess proton is allowed to transfer from state a. to state b. due to the off-diagonal coupling term. The bonding topology in both states is exactly the same, except for the species in which the excess proton is bound to.

To find the coupling between two hydronium states, the search begins by identifying the common Zundel cation, $H_5O_2^+$. All the hydronium states are saved in topology arrays, and the search loops through the hydronium proton section of each state to identify shared protons between two hydronium oxygens. When this occurs, the Zundel state is identified. For states involving proton transfer between a hydronium and a chloride anion, the Zundel-like ion Cl-H-OH_2 is found. The

covalent hydrogen identity is saved in an array, and the search loops through the hydronium protons to find any state that shares the proton. The Zundel and Zundel-like ion are depicted in figure 3.5.

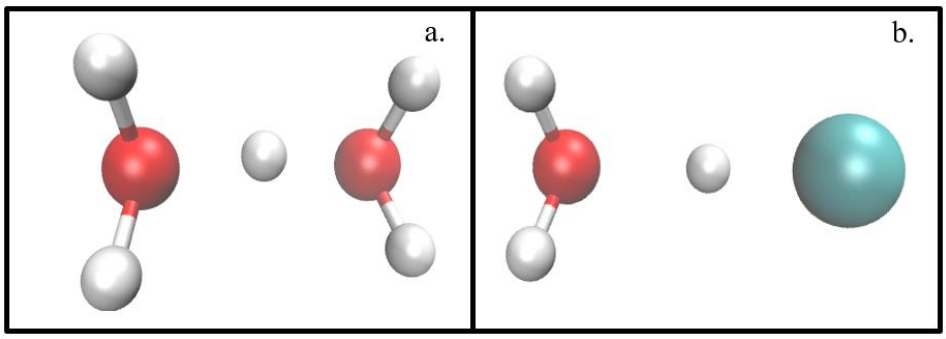

Figure 3.5. a. Zundel ion b. Zundel-like ion

3.4.3. Potential Energy Description

Once all states are found, the potential energy of every element that is non-zero in the Hamiltonian must be calculated. The form of the potential depends on what species exist. In the case of hydronium, chloride, and water, the potential can be seen in equation 3.7.

$$H_{ii}(H_3O^+, Cl^-, H_2O)$$

$$= V_{H_3O^+}^{intra} + \sum_{i}^{N_{H_2O}} V_{H_3O^+-H_2O_i}^{inter,i} + \sum_{i}^{N_{H_2O}} V_{H_2O_i}^{intra,i} + \sum_{i<j}^{N_{H_2O}} V_{H_2O_i-H_2O_j}^{inter,i,j} + \sum_{i}^{N_{H_2O}} V_{Cl^--H_2O_i}^{inter,i}$$

$$+ V_{Cl^--H_3O^+}^{inter} + \sum_{i}^{N_{H_2O}} \sum_{j}^{N_{H_2O}} V_{H_2O_i-Cl^--H_2O_i}^{three-body} + \sum_{i}^{N_{H_2O}} V_{H_3O^+-Cl^--H_2O_i}^{three-body} \quad (3.7)$$

The choice of water potential, particularly as the majority of interactions come from water interactions in water clusters, is essential. The TIP3P,[92] TIP4P,[102] SPC/Fw,[89,91,96] and qSPC/Fw[96] water molecules have all been used in previous hydronium-water MS-EVB algorithms. Polarizable MS-EVB models tend to parameterize their own flexible water models.[104] In this work, the SPC/Fw water model will be utilized.[147] The SPC/Fw is an excellent flexible water potential. In

addition, the MS-EVB3.2 Hydronium model,[91] which will be used for all hydronium-water interactions in this work, was parameterized with an SPC/Fw water potential. Parameters for the intermolecular interactions of Cl^- with water and hydronium were taken from previous work by Voth and co-workers on concentrated acid solutions.[100, 101] For all two-body parameters, unless specifically mentioned by the underlying potentials, all unlike atoms interacting used Lorentz-Berthelot mixing rules.

The system is highly polarizable once HCl ionizes forming the polarizable Cl^- ion, however, to include a fully polarizable treatment of the potential is quite computationally expensive. The dominant many-body effect of the ab initio energy is the three-body interaction. Rather than using a polarizable model in the MS-EVB-HCl, a computationally inexpensive geometric three-body potential was incorporated. Two, three-body terms were added, one describing the interactions between the heavy atoms of water-chloride-water (equation 3.8), and the second between the heavy atoms of hydronium-chloride-water (equation 3.9).

$$V_{H_2O-Cl^-H_2O}^{three-body} = \sum_{i}^{N_{H_2O}} \sum_{i<j}^{N_{H_2O}} \left(Ae^{-\alpha\left(r_{Cl^-O_i}-3.0\right)^2} e^{-\alpha\left(r_{Cl^-O_i}-3.0\right)^2} e^{-\beta r_{O_iO_j}} \right) sf_{Cl^-O_i} sf_{Cl^-O_j} sf_{O_iO_j} \quad (3.8)$$

$$V_{H_2O-Cl^--H_3O^+}^{three-body} = \sum_{i}^{N_{H_2O}} \left(Be^{-\gamma\left(r_{Cl^-O_i}-3.0\right)^2} e^{-\zeta(r_{Cl^-O}-2.8)^2} e^{-\kappa r_{O_iO}} \right) sf_{Cl^-O_i} sf_{Cl^-O} sf_{O_iO_j} \quad (3.9)$$

Switching functions are added in order to seamlessly cut the three-body potential off after the first solvation shell. The switching function, equation 3.10, is present for both chloride-oxygen distances, and the oxygen-oxygen distance, in order to prevent over solvation effects due to contributions from waters positioned on completely opposite sides of the chloride ion. The r_1-r_2 value is set at 3.2-3.3 Å for the oxygen-chloride distance, and 3.8-4.0 Å for the oxygen-oxygen distance.

$$sf = \begin{cases} 1 & r < r_1 \\ 1 - (r_2 - r_1)^{-3}(r - r_1)^2(3r_2 - r_1 - 2r) & r_1 \leq r \leq r_2 \\ 0 & r > r_2 \end{cases} \quad (3.10)$$

The potential for covalent HCl and water is seen in equation 3.11. As before, the SPC/Fw potential describes the water interactions. A Morse potential describes the HCl intramolecular potential. Equation 3.12 shows the Morse potential, where parameters for D and α are taken from Wick's work[103]; in this work, r_0 and U_0 were set to match the HCl-water dimer calculations from this study (table 3.1). The two body potential for the HCl-Water interactions uses a Lennard-Jones potential and Coulomb interaction instead of the Buckingham 6-exponential used in Wicks work, however covalent HCl σ, ε, and q parameters are taken from Wick's model. Lorentz-Berthelot combination rules were used for unlike atoms

$$H_{ii}(HCl, H_2O) = \sum_{i}^{N_{H_2O}} V_{H_2O_i}^{intra,i} + \sum_{i<j}^{N_{H_2O}} V_{H_2O_i - H_2O_j}^{inter,i,j} + \sum_{i}^{N_{H_2O}} V_{H_2O_i - HCl}^{inter} + V_{HCl}^{intra} \quad (3.11)$$

$$V_{HCl}^{intra} = D\left[1 - e^{-\alpha(r - r_0)}\right]^2 + U_0 \quad (3.12)$$

The off-diagonal potential is essential to accurately describe the proton transfer between two states. Many studies have focused on the best way to empirically describe the transfer between states, such as a Gaussian multiplied by a polynomial added across the potential energy surface,[84] or a less complex Gaussian function based on the reaction coordinates.[86] In this work, the potential describing the Zundel cation is a complex function following the MS-EVB3.2 algorithm.

For the HCl-water "Zundel-like" states, a simple Gaussian function will be used. As discussed in the previous chapter, the interplay between covalent HCl-water and CIP H_3O^+/Cl^- is relatively complex, however, two structural motifs exist. To this end, the empirical Gaussian description for the coupling states contains a term based on both the oxygen-chlorine and hydrogen-chlorine distance, shown in equation 3.13. The switching function in equation 3.10 was

utilized to swiftly cut the terms to zero with r_1-r_2 values of 1.7-2.3 Å for the hydrogen chloride distance and 3.0-3.3 Å for the oxygen chloride distance.

$$H_{ij} = \left(A_f e^{-\alpha_f (r_{O-Cl}-2.8)^2}\right) sf_{O-Cl} + \left(B_f e^{-\beta_f (r_{H-Cl}-1.293)^2}\right) sf_{H-Cl} \tag{3.13}$$

3.5. Parameterization of the Model

Eleven parameters, between the addition of the geometric three body potentials, and the Zundel-like off-diagonal potential, must be fit. Parameterizing reactive force fields is slightly more difficult than classical force fields as there are more parameters for each atom describing the interactions, and the parameters are more strongly coupled.[52] In this work, a two-step process in fitting was taken. In the first step, the four parameters required for the off-diagonal potential were fit to 14 dimer structures. Ten of the structures describe the Cl-H-OH_2 potential energy curve. In this potential energy curve, the chloride oxygen distance was held constant, and the reaction coordinate allowed to vary is the hydrogen-chloride distance. The computational methodology for generating the potential energy curve is detailed in appendix B. These structures fall mostly under the HCl-water motif, and so four dimers that had stronger H_3O^+/Cl^- CIP character were added. In the second step of the fitting procedure, the parameters for the geometric three-body potential were fit using the parameterization set of n=2-8 water molecules described in the previous chapter. The choice of fit set exemplifies the significant difference between Wicks model and the MS-EVB-HCl. Wicks model fits to the HCl-Water dimer potential energy surface, the solvent separated ion pair $H_3O^+(H_2O)_3Cl^-$, and the potential energy surface of the CIP $H_3O^+Cl^-(H_2O)_3$. The reaction coordinate for the potential energy surfaces is the hydrogen chloride distance. For the MS-EVB-HCl, the fit set for the three-body potential contains HCl, CIP, and SSIP structures including higher energy structures, with relative energies compared to the minimum energy structure at each cluster size as high as ~6 kcal/mol, rather than just the low-lying minima.

A genetic algorithm (GA) was utilized for the fitting process. The GA used is an in-house code that works based on the theory of natural selection.[52, 68] For each parameter, a range of potential values is input. From these ranges, a random initial population is generated. This population is tested to determine the strength of the set, in this case using least-squares fitting. Once the fit has been determined, a genetic operation such as cross-breeding or mutation is performed generating the new fit. The validity of this fit is quantified and this process is repeated until a specified termination point. The parameters determined for this model can be seen in 3.1.

Table 3.1. List of parameters used in the MS-EVB-HCl model

Parameter	Value	Parameter	Value
r_0	1.293Å	ζ	12.795 Å-2
U_0	-165.11 kcal/mol	κ	2.1097264 Å-1
A	-22.3110 kcal/mol	A_f	-4.7960 kcal/mol
α	14.998 Å-2	α_f	0.800 Å-2
β	0.5504 Å-1	B_f	-19.578 kcal/mol
B	-251.79 kcal/mol	β_f	0.800 Å-2
γ	16.539 Å-2		

3.6. Performance of MS-EVB-HCl

The MS-EVB-HCl accurately predicts the HCl potential energy curve and can be seen in figure 3.6. From the potential energy curve, it can be inferred that the simple off-diagonal potential of the MS-EVB-HCl model accurately describes the proton transfer between hydronium and chloride.

While the two body potential accurately models that transfer, it is important for the geometric three-body term to accurately model the interaction with water at all ion pairing situations. The fit set for the three body potential contained structures with covalent HCl, CIP, and SSIP motifs. Figure 3.7 shows the parameterization set and its performance compared to the MS-

EVB3.2 force field that was not designed to include covalent HCl character. All energies are given with respect to the lowest energy structure (as determined from ab initio calculations) at each particular cluster size in kcal/mol.

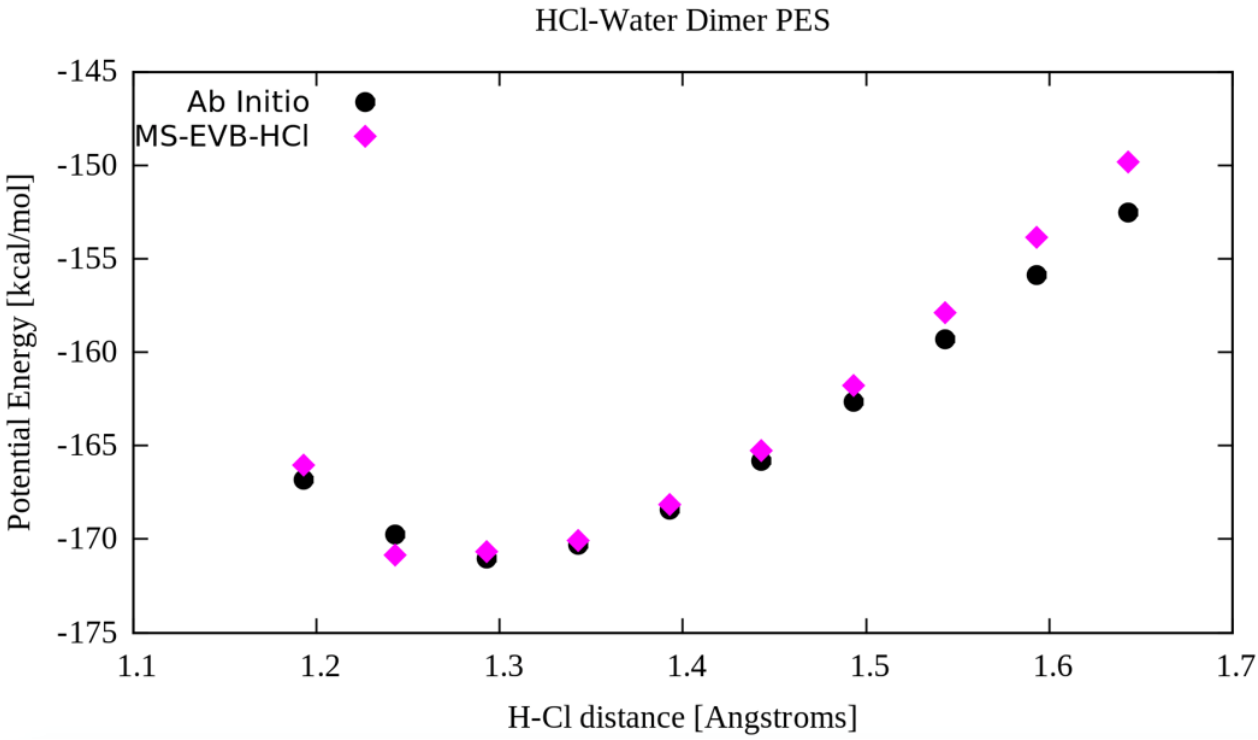

Figure 3.6. Performance of MS-EVB-HCl vs Ab Initio energy

As can be seen from figure 3.7, MS-EVB3.2 is unable to reproduce the relative energetics of a few key structures. The RMSD for the MS-EVB-HCl is 2.62 kcal/mol and 11.45 kcal/mol for the MS-EVB3.2 model. This shows the necessity of inclusion of covalent HCl character. The points that are drastically different from the ideal curve in the MS-EVB3.2 performance are covalent HCl states, but it is expected that the MS-EVB-HCl improves upon CIPs as well. This will be discussed later.

While the fit set of the MS-EVB-HCl is found to perform quite well in comparison to the ab initio energies, it is imperative that the model reproduces energetics of structures not used in fitting the parameters. The validation set includes 91 structures ranging from $HCl(H_2O)_9$ to

$HCl(H_2O)_{22}$ and contains a wide range of solvation structures with all four main motifs: covalent HCl, CIP, SSIP, and NoIP.

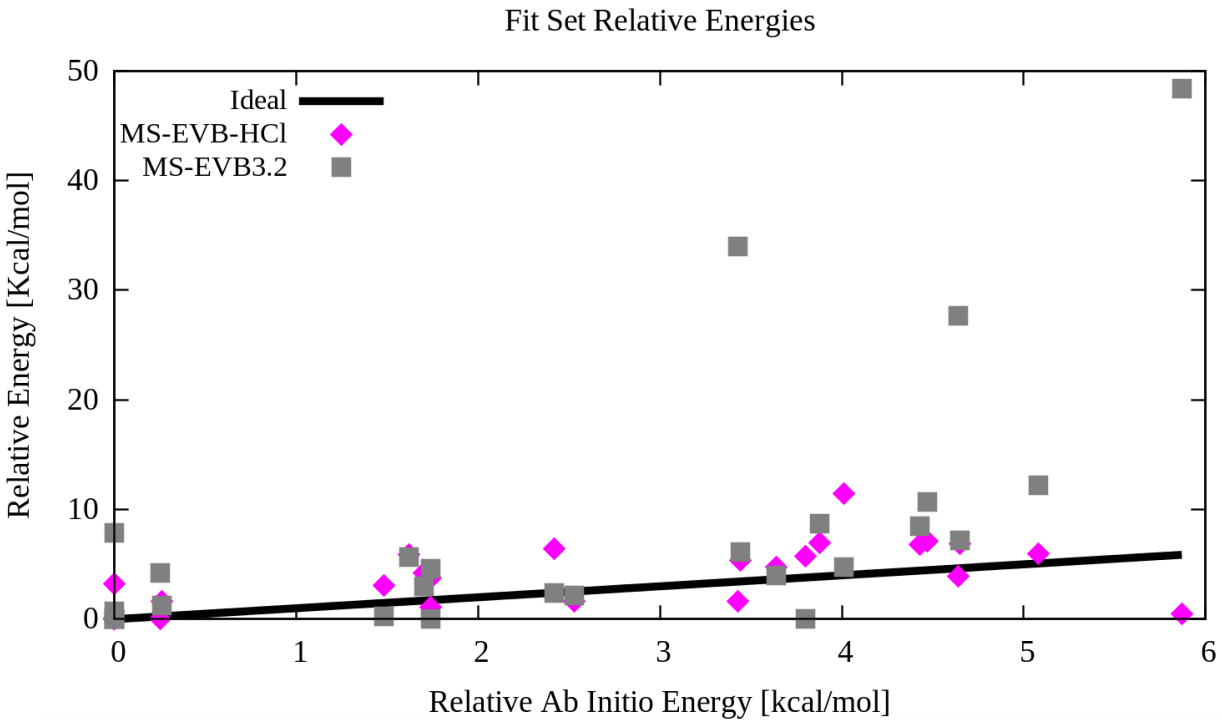

Figure 3.7. Comparison of the MS-EVB-HCl and MS-EVB3.2 performance on reproducing the relative energies of the fit set

The performance of MS-EVB3.2 and MS-EVB-HCl with respect to the ideal energies given by ab initio calculations can be seen in figure 3.8. The RMSD value for MS-EVB3.2 is 8.34 kcal/mol and is 2.76 kcal/mol for MS-EVB-HCl, indicating the necessity of inclusion of explicit covalent HCl configurational states as the inclusion improves upon the relative energetics. As before, all energies are given with respect to the minimum energy cluster at each cluster size, in kcal/mol. Keeping in line with trends from the fit set, MS-EVB3.2 has drastically high energies for covalent HCl clusters. Additionally, data from the figure shows improvement at more clusters than just the covalent HCl states with MS-EVB-HCl.

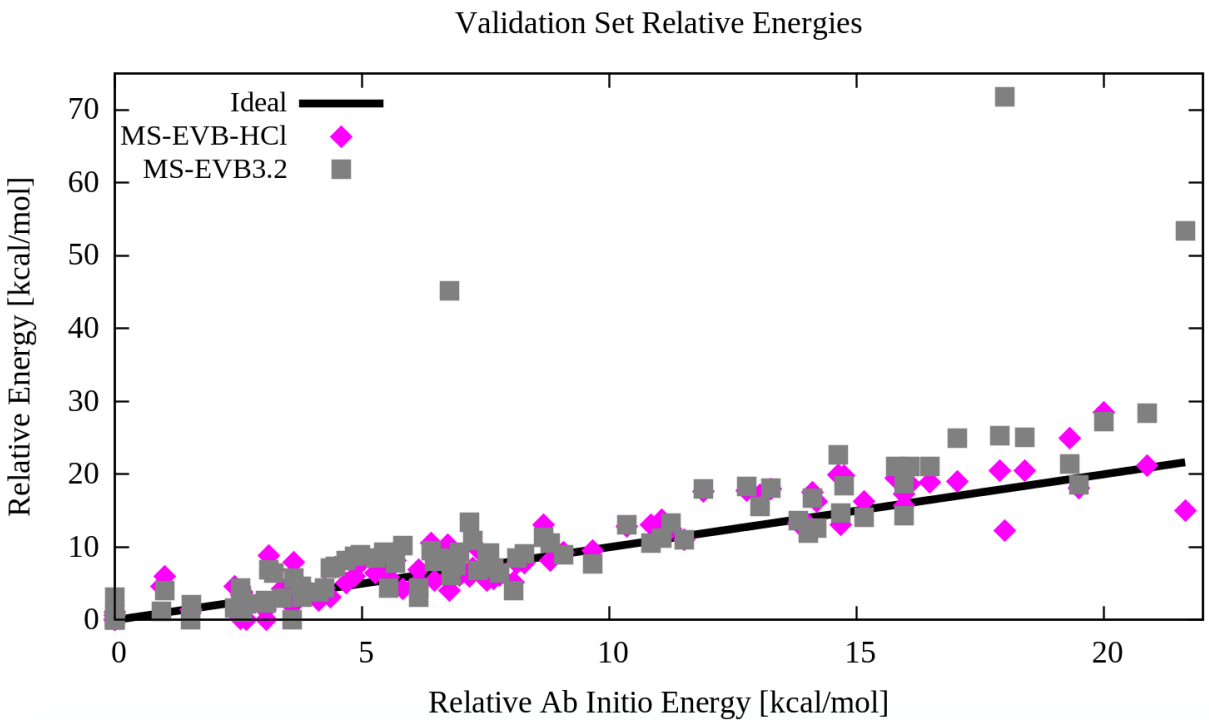

Figure 3.8. Comparison of the MS-EVB-HCl and MS-EVB3.2 performance on reproducing the relative energies of the validation set

In order to quantify this, the fit and test sets were broken up into ion pairing sets. In figure 3.9 the performance of MS-EVB3.2 and MS-EVB-HCl for only covalent HCl states is depicted. With this solvation motif, MS-EVB3.2 has an RMSD value of 33.04 kcal/mol and MS-EVB-HCl has an RMSD value of 3.89 kcal/mol. Essentially, the idea that neglecting covalent HCl in favor of tuning Lennard-Jones parameters to reproduce more contact ion pairing has been found to be insufficient in this work. The covalent character for water clusters, particularly smaller clusters, is significant. The MS-EVB3.2 model, while showing tremendous success in bulk water solutions, was not built to include covalent states and therefore cannot model these states with a degree of accuracy.

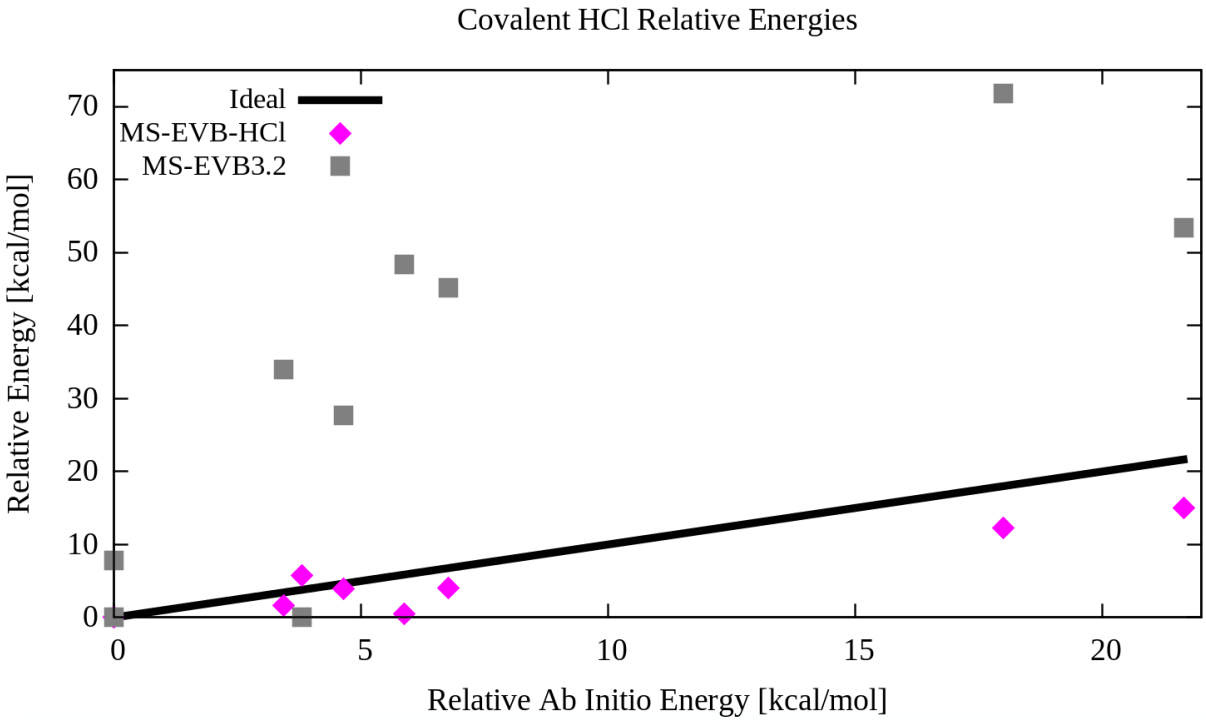

Figure 3.9. Comparison of the MS-EVB-HCl and MS-EVB3.2 performance on reproducing the relative energies of HCl configurations

CIP motifs were analyzed in order to determine the effect of including covalent terms on the improvement of the MS-EVB algorithm. Figure 3.10 shows the performance of MS-EVB3.2 and MS-EVB-HCl on water clusters of the CIP motif only. The RMSD value is reduced from 4.50 kcal/mol in the MS-EVB3.2 iteration to 2.60 kcal/mol with MS-EVB-HCl. The previous chapter discussed the different ranges for the closest hydrogen to chlorine distance at each solvation motif. The CIP had a range of 1.72 Å – 2.06 Å. The CIP structures that resulted in significant improvement with inclusion of covalent states in the MS-EVB-HCl model are CIP isomers with the closest hydrogen to chloride distances falling in the first half of that range, or under 1.9 Å, with a moderate improvement on a majority of the other structures with large hydrogen chloride distances (due to the inclusion of the three-body interaction between the Cl- ion and it solvation shell).

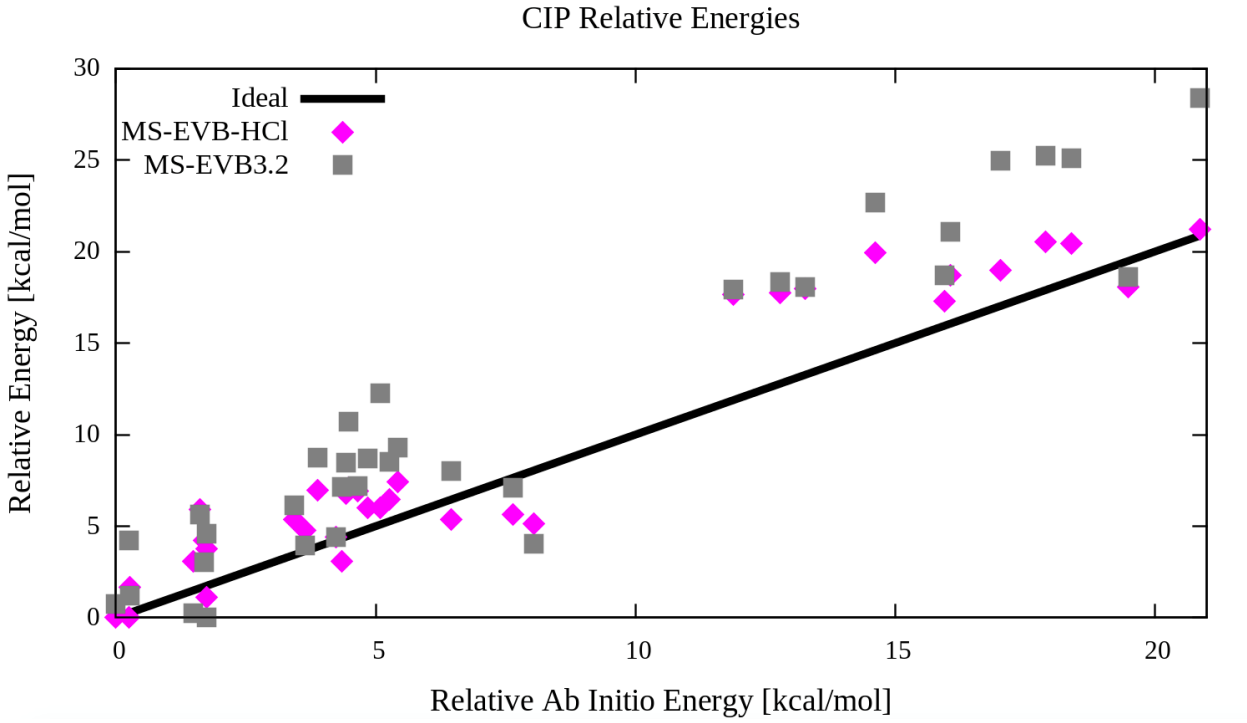

Figure 3.10. Comparison of the MS-EVB-HCl and MS-EVB3.2 performance on reproducing the relative energies of CIP configurations

Figure 3.11 shows the structural breakdown of performance for the MS-EVB3.2 and MS-EVB-HCl for the SSIP and NoIP motifs. It can be seen that the MS-EVB-HCl force field approaches the MS-EVB3.2 performance as the ion-pairing decreases, which is expected since at the lower ion pairing limit the MS-EVB3.2 model should accurately describe the interactions. Table 3.2 shows the ab initio energy range for each motif, similar to in the previous chapter, where each energy is with respect to the lowest energy structure at a given cluster size. The table has three additional columns: the first is the ab initio energy range, which shows the maximum energy cluster found at each cluster size, or the energy range of clusters found with the ab initio energy description, the second is the MS-EVB3.2 energy range with respect to the minimum energy cluster at that size, the third is the MS-EVB-HCl energy range. All energies are relative energies and are given in kcal/mol.

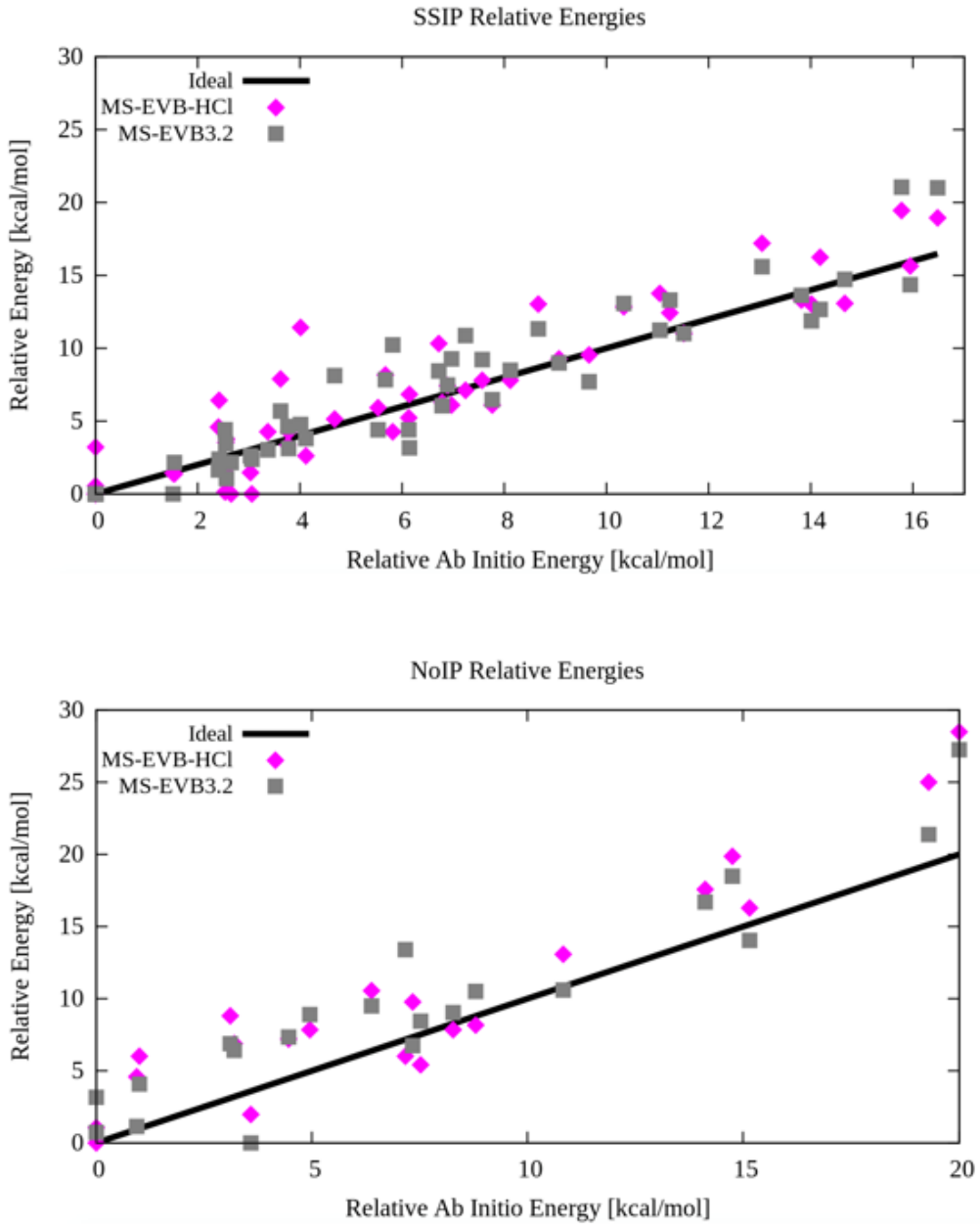

Figure 3.11. Comparison of the MS-EVB-HCl and MS-EVB3.2 performance on reproducing the relative energies of SSIP and NoIP motifs

From the graphs comparing the energies of each MS-EVB model to the ideal ab initio energy, as well as the tables, it can be seen that there was a significant improvement upon the addition of covalent HCl to the model.

Table 3.2. Energy ranges of all solvation types and, as well as the energy ranges for Ab Initio, MS-EVB3.2 and MS-EVB-HCl calculations. All relative energies are given in kcal/mol.

Cluster size	HCl Energy Range	CIP Energy Range	SSIP Energy Range	NoIP Energy Range	Ab Initio Energy Range	MS-EVB3.2 Energy Range	MS-EVB-HCl Energy Range
3	0-3.8	-	-	-	3.8	7.9	5.8
4	3.4-4.6	1.62	0	-	4.6	34.0	6.0
5	5.9	0.3-5.1	0	-	5.9	48.4	7.1
6	-	1.5-3.4	0.0-4.0	-	4.0	6.1	11.4
12	6.8	4.8-17.9	0-11.2	7.2	17.9	45.2	20.6
20	-	4.4-19.5	0-8.1	4.5-8.3	19.5	18.7	18.1
21	21.6	11.9-20.9	6.1-14.7	0-20.0	21.6	53.4	28.5
22	-	8.05	0-15.6	7.3-14.5	16.0	21.1	19.9

Figures 3.12 and 3.13 show five structures that have been improved with MS-EVB-HCl, three CIP structures (figure 3.12), two covalent HCl (figure 3.13), and one SSIP structure (figure 3.13), as well as the energies given by the ab initio, MS-EVB3.2, and MS-EVB-HCl description.

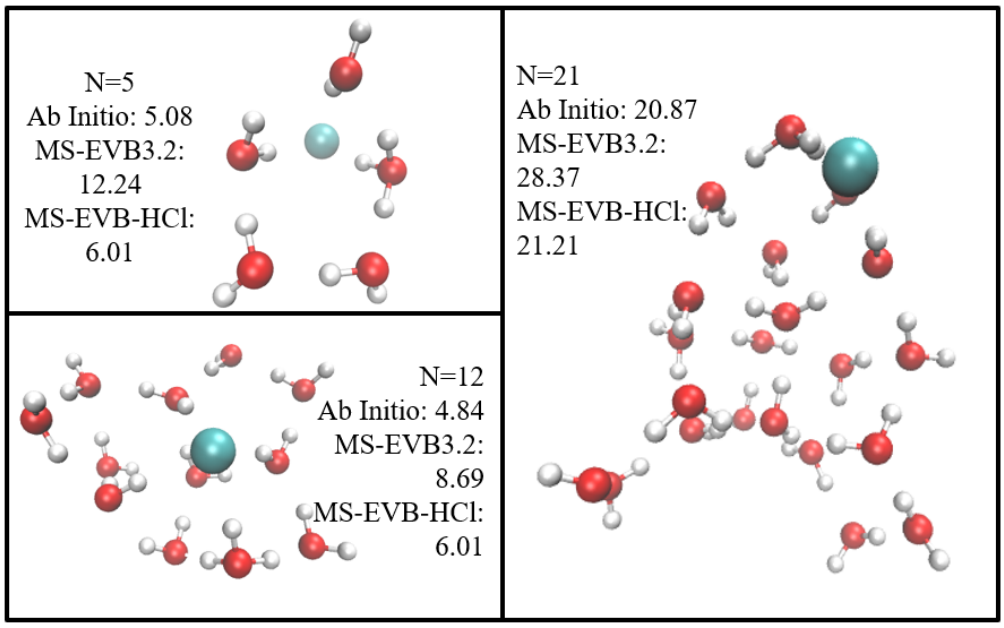

Figure 3.12. Comparison of performance of MS-EVB3.2 and MS-EVB-HCl at various representative clusters. All relative energies are given in kcal/mol.

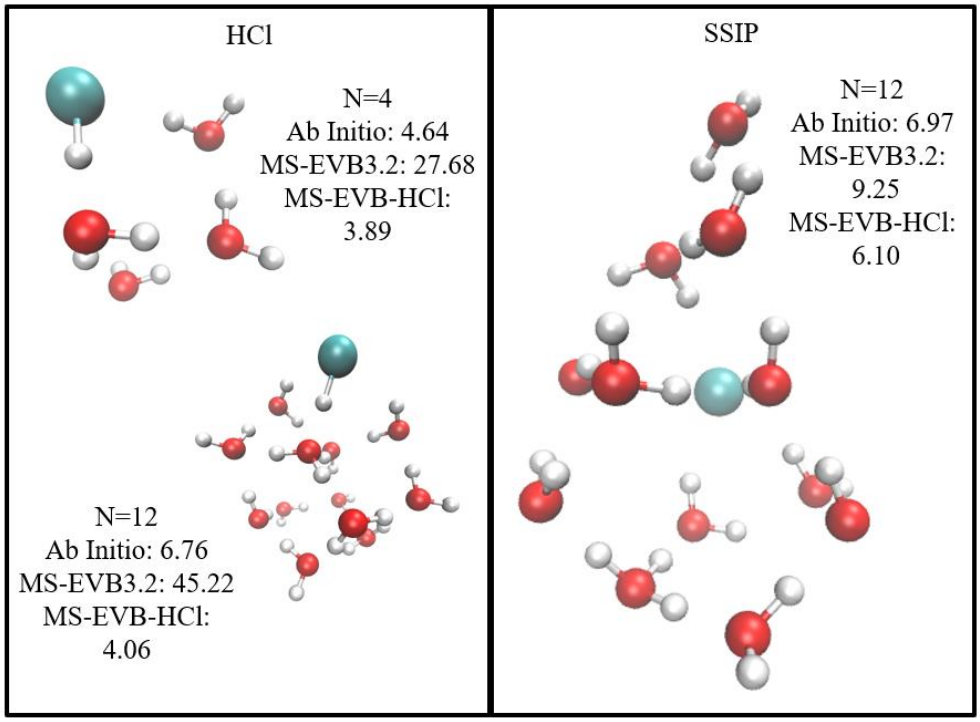

Figure 3.13. Comparison of performance of MS-EVB3.2 and MS-EVB-HCl at various representative clusters. All relative energies are given in kcal/mol.

3.7. Conclusion of the MS-EVB-HCl Development

A reactive force field, MS-EVB-HCl, based on the MS-EVB3.2 algorithm has been developed for the study of systems such as HCl-induced gas-liquid nucleation, HCl at the air water interface, and concentrated acidic solutions. Previous models have focused on parameterizing to the lowest energy structures at small cluster sizes, or have not included covalent states based on the assumption that the contribution of covalent HCl is negligible and simple Lennard-Jones parameters with mixing rules could accurately reproduce the energetics. The new MS-EVB-HCl algorithm includes covalent HCl and was parameterized to a wide range of structures including the H-Cl-OH_2 potential energy surface, and more CIP like H_3O^+/Cl^- dimer structures for the two body potential and clusters ranging in solvation type, cluster size [$HCl(H_2O)_{2\text{-}8}$], and relative energy range (0 to ~6 kcal/mol) for the three body potential.

The MS-EVB-HCl model shows improvement at all cluster sizes [$HCl(H_2O)_{2\text{-}22}$]. The most significant improvement in accuracy occurs for covalent HCl and CIP motifs, however, there is a slight improvement on some SSIP and NoIP configurations as well. The MS-EVB-HCl model proves the necessity of inclusion of covalent HCl states as well as three-body interactions between the Cl- ion and its solvation shell in order to accurately model concentrated solutions and small to medium sized clusters. The geometric three-body potential suffices in the description of water-chloride-water and hydronium-chloride-water interactions in a computationally efficient manner.

Chapter 4. DFT Investigations into the Regioselectivity of Unsymmetrical Silyloxyallyl Cations

4.1. Introduction: Experimental Motivation

Oxyallyl cations have been found to be extremely useful in organic synthesis. These cations are often produced via Narazarov electrocyclization,[148-150] and proceed to form carbon-carbon bonds via cycloaddition.[151-153] While oxyallyl cations allow for a significant amount of control over the synthetic process, regioselective addition of nucleophiles to unsymmetrical oxyallyl cations at the α vs α'-carbon has been difficult.[154, 155] To this end, the Kartika group has worked to develop a reaction method that allows for a tailored addition to the less substituted α'-carbon.[156-158] This mechanism begins with ionization of the hydroxyl group with a Bronsted acid to make a stable silyloxyallyl cation. From this cation, addition to the α'-carbon is possible. When putative silyoxyallyl cations, under the conditions specified by the Kartika methodology, react with indole it was found that a large amount of product formed where the carbon of indole attached to the less substituted α'-carbon of the cation; whereas in the methyloxyallyl case, no regioselectivity was determined experimentally.

In order to understand the full reaction mechanism, from the initial formation of a cation, to the intermediate indole addition, and final deprotonation of the cation, as well as the regioselectivity dictated in the silyloxyallyl case, Density Functional Theory (DFT) calculations were performed. These calculations determined whether the formation proceeded in an SN1 or SN2 fashion, and the reaction pathways for every conformer for the addition of indole to both the silyloxyallyl (TBS), and methyloxyallyl (Me) cations. The indole ring and trace waters ability to stabilize the cations have been determined and discussed and kinetic, thermodynamic, charge, and steric control of the selectivity has been analyzed. Ultimately the full reaction profile and its reversibility will be discussed. The work described herein has been published under the title

"Mechanistic Perspectives in the Regioselective Indole Addition to Unsymmetrical Silyloxyallyl Cations", in the Journal of Organic Chemistry.[8]

4.2. Computational Methodology

4.2.1. Ab Initio Metadynamics

In order to determine whether the formation of oxyallyl cations proceeds through an SN1 or an SN2 mechanism, ab initio metadynamics (MTD)[137] simulations were performed with the CP2K package.[138, 139] All MTD simulations utilized the BLYP functional and Grimme D3 dispersion,[135, 136] the DZVP-MOLOPT-GTH basis set,[140] Goedecker-Teter-Hutter (GTH) pseudopotential[141] within the Gaussian Plane Wave framework.[139] The real space cutoff was set to 400 Ry and a wavelet Poisson solver was used.[142] Simulations were executed at 300K, with a 0.5 fs time step. In order to maintain indole within reactive distance to the cation, a reflective wall was used. Two collective variables (CV's), described in more detail below, were set in order to accurately model dissociation of the hydroxyl group and addition of indole. Gaussians were added every 15 steps, a height of 0.8 kcal/mol and a width of 0.5 Bohr for CV1 and 0.8 Bohr for CV2.

4.2.2. Electronic Structure Methods

Electronic structure methods: The Gaussian 09[134] suite of programs was used to explore the pathways resulting in each possible configuration. Density Functional Theory (DFT) with the B3LYP functional, Grimme D3 dispersion corrections[135, 136] and the 6-31+G(d) basis set were used in all calculations. Potential Energy Surface (PES) scans were used to investigate the regioselectivity of indole addition to the α and α'-carbons of both TBS and Me containing cations. Potential reactant, transition, and product states were taken from these PES scans and further minimized. The Berny algorithm optimization scheme was used to locate transition states, which were further verified by the presence of one imaginary frequency.[159] In order to verify that the

correct transition state has been identified, the mode must correspond to the correct transition under investigation. Solvation was modeled with an implicit solvent. The SCRF (Polarizable Continuum Model using the integral equation formalism variant) keyword was used to set the implicit solvent of Toluene (ε=2.3741).[55] One explicit water molecule was added to stabilize the transition states. NBO analysis was utilized to calculate charges and visualize orbitals.[160] The Psi4 package[161] was used for SAPT energy decomposition.[162]

4.3. First Step: Cation Formation

4.3.1. Metadynamics Investigation into Mechanism Pathway

Figure 4.1 shows the schematic of the cations. MTD calculations were performed in order to determine if the reaction proceeded through an SN1 or SN2 mechanism. Calculations were

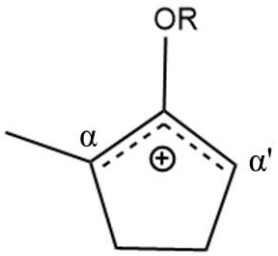

Figure 4.1. illustration of the cation. R can be TBS or Me. The α- carbon is the substituted addition site and α' is the unsubstituted addition site.

executed on the non-regioselective Me cation. The initial configuration began with the pyridinium in a suitable position to the hydroxyl leaving group; the proton was pointed towards the oxygen within a reactive distance. Two collective variables (CVs) were used to probe the formation of the cation as well as the addition of indole. The first CV explored the distance between the α'-carbon and the oxygen of the hydroxy group. This distance is chosen as it is the bond that is required to dissociate in order to form the cation. The second CV described the indole addition step and was defined as the distance between the nucleophilic carbon of indole and the α'-carbon. Figure 4.2

illustrates the CVs used in the simulation. This simulation saw the dissociation of the O-C bond upon the addition of 70 Gaussians, however, this occurred without an accompanying C-C bond formation indicating that the reaction proceeds through an SN1 mechanism. As the C-O bond started to dissociate, the hydrogen from the pyridinium started bonding with the oxygen, generating a water molecule and pyridine.

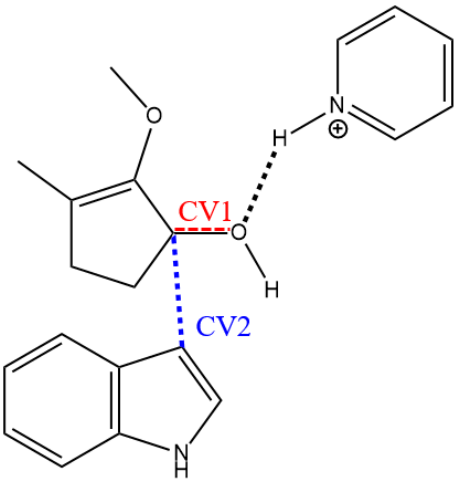

Figure 4.2. Depiction of the CVs used to determine mechanism pathway. CV1 is shown in red, CV2 in blue, and the black dotted line shows the hydrogen bond between the oxygen hydrogen, leads to bond formation at larger values of CV1

4.3.2. Effects of Indole on Cation Stabilization

Interestingly, the formation of the cation, 11a, would not occur in the metadynamics simulation without the presence of indole, despite its lack of reactivity in the first step. To investigate this phenomenon further, the cation formation step was calculated with (pre-11a to 11a) and without indole (pre-11a-NI to 11a-NI). A relaxed potential energy scan was performed along the oxygen of the leaving group to the hydrogen of pyridinium, which is also described by the same coordinate as CV2 in the MTD simulations. From this scan, a possible reactant, transition, and product states were extracted and minimized. The reaction profiles can be seen in figure 4.3. These calculations verified that the addition of a proton to the OH group and departure of the

leaving group from the oxyallyl system occurred in the same step, despite the use of two different coordinates (O-H bond distance in the PES, and C-O distance in metadynamics). Potential and Gibbs free energies, at standard temperature and pressure, were calculated. The format when both energies are shown has the Gibbs free energy given first followed by the potential energy (PE) in parenthesis.

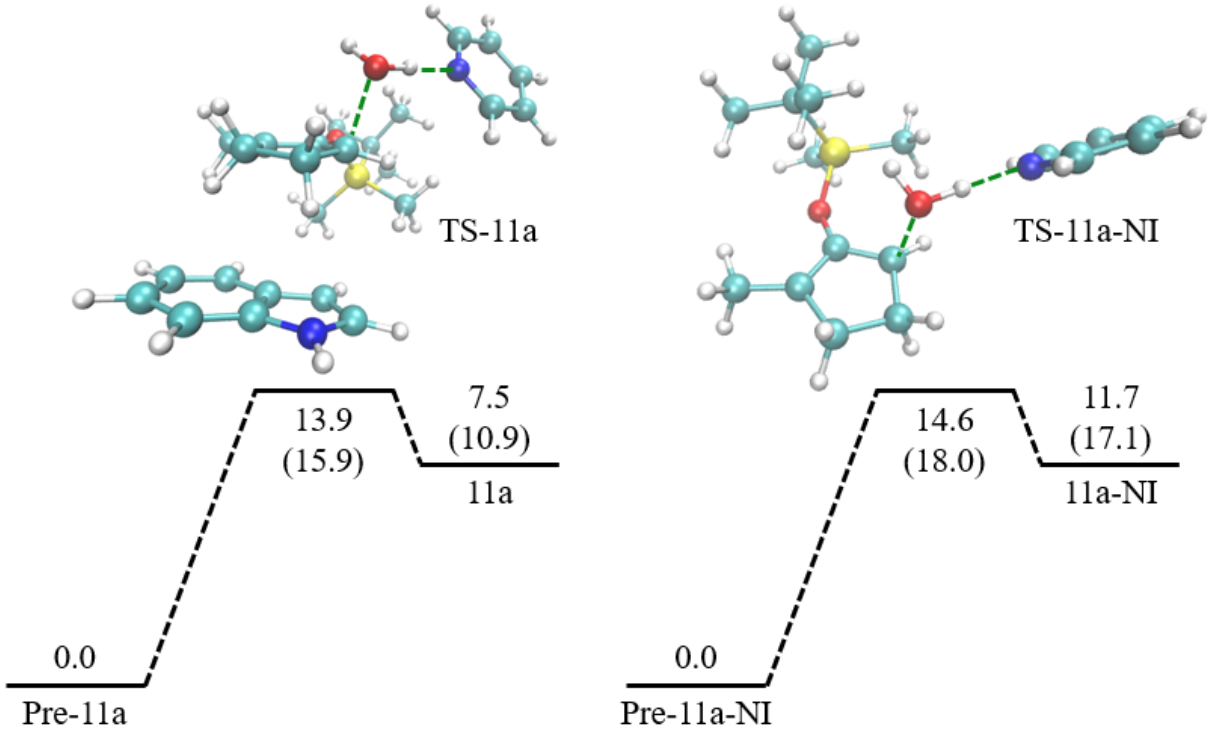

Figure 4.3. reaction mechanism for the formation of the TBS cation with (left) and without (right) the indole ring. Gibbs free energies are given first followed by potential energy in the parenthesis. All energies are given relative to the respective reactant and are given in kcal/mol.

Besides this notation, all energies are Gibbs free energies unless specifically stated. It is useful to show both energies because Gibbs free energy gives some indication of the entropic effects of the system which is important when modeling chemical reactivity. However, strictly speaking, the Gibbs free energy calculations carried out here assume that the particles are non-interacting and as such can only be accurately applied to an ideal gas. It is difficult to determine the extent this

assumption will have on the accuracy of the calculations, therefore the potential energy is included in order to establish general trends.

It was found that cation formation in the TBS case yields an activation barrier of 14.6 kcal/mol (18.0 kcal/mol PE) when no indole was present and 15.9 kcal/mol (20.3 kcal/mol PE) for the formation of Me case. As the reaction proceeds in an SN1 fashion, the first step is the rate determining step, which is unsurprising due to the presence the large activation barriers. The barrier for the formation of the TBS cation was decreased to 13.9 kcal/mol (15.9 kcal/mol PE) when indole was present, and the product was significantly more stable with indole to stabilize the cation, a product of 7.5 kcal/mol (10.9 kcal/mol PE), vs 11.7 kcal/mol (17.1 kcal/mol PE) without the indole present. Additionally, the barrier for the reverse reaction is significantly smaller leading to a more competitive reverse reaction without indole present. It is clear that the indole provides a significant amount of stabilization leading to the long lifetimes of the putative oxyallyl cations.

4.3.3. Investigation of π-π Interactions

In order to understand the origin of the cation stabilization by indole, the HOMO (highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbitals) states were extracted for analysis. Visualizing these molecular orbitals gives insight into the influence of π-π stacking often seen in mid-range complexation of ring systems. In order to preserve clarity in the visualization of the π-stacking, the indole complexed with the Me cation (11e) was used, as the TBS is quite bulky and may obstruct the view of the π-π stacking. Furthermore, as the Me cation allows for bonding at both the α and α'-carbon, a full picture of stacking will be available. The visualized molecular orbitals can be seen in figure 4.4. From the figure, it is clear that there is a significant amount of π-overlap between the indole, and the reactive carbons, i.e. the α and

α'-carbons. This π-stacking between the rings allows the positive charge of the cation to be more delocalized, and therefore more stable.

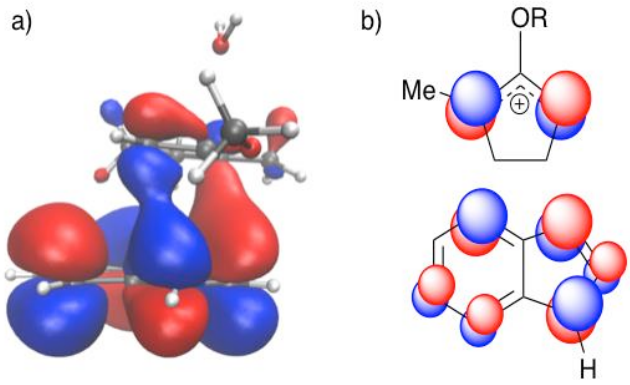

Figure 4.4. a. the HOMO-LUMO molecular orbitals are illustrated for structure 11e. b. a schematic of the orbitals that participate in π-bonding between the cation and indole. R=TBS or Me. The figure has been reprinted with permission from Bresnahan, C. G.; Taylor-Edinbyrd, K. A.; Cleveland, A. H.; Malone, J. A.; Dange, N. S.; Milet, A.; Kumar, R.; Kartika, R., Mechanistic Perspectives in the Regioselective Indole Addition to Unsymmetrical Silyloxyallyl Cations. *The Journal of Organic Chemistry* **2019**. ©2019 American Chemical Society

4.4. Intermediate Addition of Indole to the Cations

All possible pathways of indole addition to each cation were modeled to determine the kinetics of the formation, and gain insight into the causes of the regioselectivity. For each cation, this requires the addition to both the α and α'-carbon. However, each potential addition site has two possible isomers, depending on the orientation of the indole ring. Furthermore, indole can approach from both an outward or inward orientation. To this end, all sixteen possible reaction pathways, eight involving the Me cations, and eight involving the TBS cations were modeled.

4.4.1. Analysis of Intermediate Reaction Pathways for the TBS Cation

All TBS pathways are depicted in figure 4.5. Structures of motif-type 12 and 13 correspond to nucleophilic addition at the substituted α-carbon whereas addition at the unsubstituted α'-carbon will be designated as motif-type 14 and 15. As before, all Gibbs free energies are given followed

by potential energies in parenthesis. All energies are given with respect to the reactant of the given pathway.

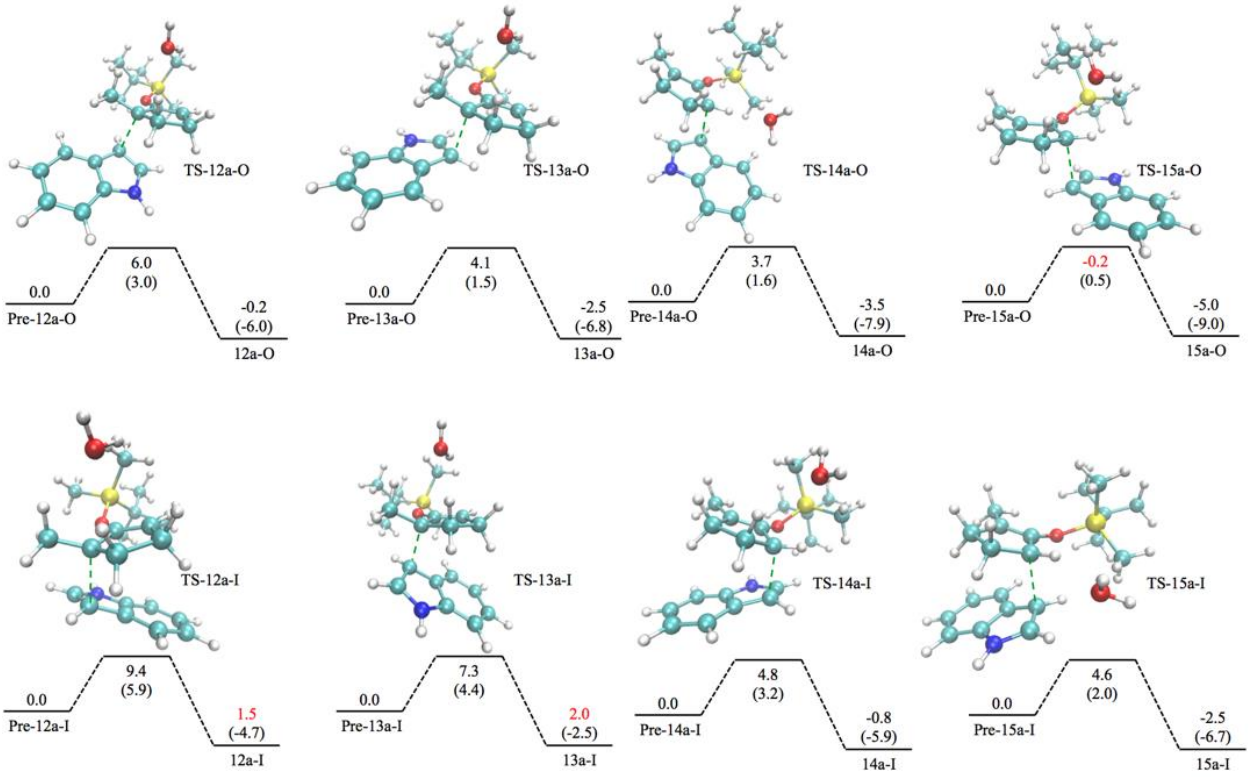

Figure 4.5. Reaction pathways for the addition of indole to the TBS cation. The first row depicts the outward approach of indole to the cation, the second row depicts the inward approach. A green dashed line shows where the bond is starting to form in each transition state. Gibbs energies are given first followed by PE in parenthesis. Energies are given with respect to each pathways reactant. Transition state energies given in red indicate a transition state with a barrier lower than the reactant, while red product energies indicate a product higher in energy than the reactant.

Table 4.1 lists the activations barriers for each isomer as well as the relative energies, potential and, free energies, with respect to the most stable structure. Boltzmann distributions of all configurations are tabulated as well; the equation for the distribution calculation is given in equation 4.1. R is the ideal gas constant and T is set to 300K. Figure 4.6 shows the reaction profile for the formation of the TBS cation and shows the relative energy levels of all isomeric products.

$$B = e^{\left(\frac{-\Delta\Delta G}{RT}\right)} \tag{4.1}$$

Table 4.1. Structures in the table refer to energies of TBS cations. Relative energies are given per state (reactant, transition, product) with respect to the most stable structure of that state (reactant, transition, product) [a] relative potential energy of each state [b] relative Gibbs free energy of each state [c] Boltzmann distribution of each state [d] activation barrier given as the difference between the transition state and reactant. Energies are given in kcal/mol. Terms in light blue indicate the lowest relative energy structure, while those in pink show the lowest relative energy addition at the other binding site.

Motif	Reactant			Transition State				Product		
	$\Delta\Delta E^{[a]}$	$\Delta\Delta G^{[b]}$	$B^{[c]}$	$\Delta G^{\ddagger[d]}$	$\Delta\Delta E^{[a]}$	$\Delta\Delta G^{[b]}$	$B^{[c]}$	$\Delta\Delta E^{[a]}$	$\Delta\Delta G^{[b]}$	$B^{[c]}$
12aO	5.4	5.4	0.0	6.0	6.4	6.8	0.0	6.0	7.7	0.0
13aO	4.7	6.1	0.0	4.1	4.3	5.6	0.0	4.6	6.1	0.0
14aO	3.6	4.8	0.0	3.7	3.2	3.9	0.0	2.4	3.9	0.0
15aO	5.4	7.5	0.0	-0.2	3.9	2.7	0.0	3.1	5.0	0.0
12aI	2.0	2.2	0.0	9.4	6.0	6.9	0.0	4.0	6.2	0.0
13aI	0.3	1.8	0.1	7.3	2.7	4.5	0.0	4.4	6.3	0.0
14aI	0.7	2.7	0.0	4.8	2.0	2.8	0.0	1.5	4.3	0.0
15aI	0.0	0.0	1.0	4.6	0.0	0.0	1.0	0.0	0.0	1.0

Table 4.1 shows the relative energies at each state. The outward configurations tend to be higher in energy than that of their corresponding inward trajectories. Calculated Boltzmann distributions show a ratio of zero outward isomers, compared to the inward configurations. Additionally, as illustrated in figure 4.6, all outward configurations are similar in energy to the transition state in the formation of the cation. This leads to an essentially barrierless transition for the reverse reaction. The outward trajectory is clearly unfavored thermodynamically. The stability of the inward trajectory can likely be attributed to the stabilizing effect of π-stacking with the indole, although one can hypothesize the steric hindrance from the TBS group plays a role. This will be discussed in further detail below. With this in mind, it is reasonable to assume that addition to the TBS cation has a less competitive reverse reactions when the indole has an inward approach to the cation and further analysis will focus on the inward approach.

The activation barrier for type 14 and 15, which allows for formation at the less substituted α'-carbon, are lower than those of type 12 and 13 where the bond forms at the substituted α-carbon indicating formation at the less substituted carbon is kinetically favored.

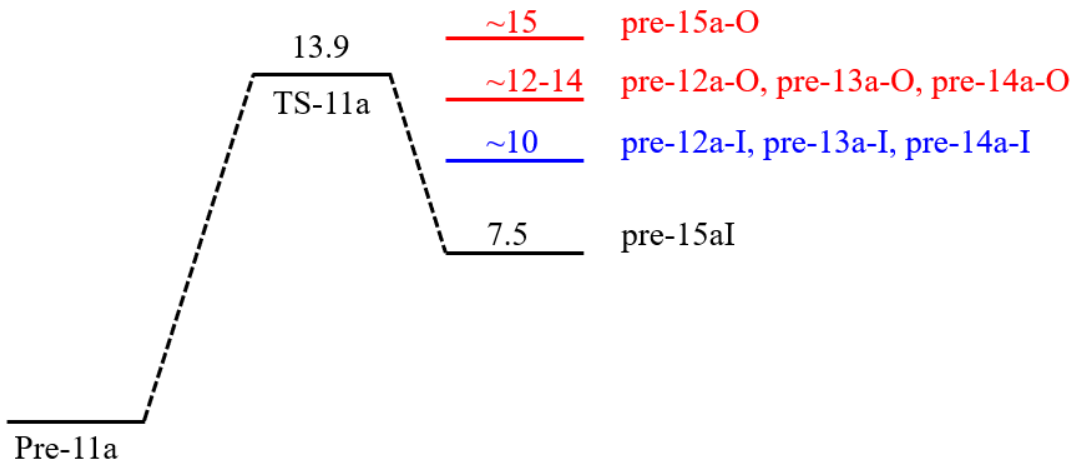

Figure 4.6. Free energy profile for the ionization of the TBS cation

In fact, 12aI and 13aI have products higher in energy than their corresponding reactants, and as such, have a lower barrier for the reverse reaction to take place. Table 4.1 shows that the relative energies and the transition and product states have a more favorable addition at the α'-carbon. The lowest energy α'-carbon addition is in blue, whereas the lowest energy α-addition is in pink. Gibbs free energy of the product favor 15a-I with an energy 6.2 kcal/mol lower in energy than 13a-I, the most stable α-addition product. Boltzmann distributions show a small ratio of pre-13a-I to pre-15aI, however, the ratio is negligible. From this analysis, nucleophilic addition to the α'-carbon has been found to be both thermodynamically favored and kinetically favored.

4.4.2. Analysis of Intermediate Reaction Pathways for the Me Cation

A similar analysis was performed on the reaction pathways in the Me case. The full reaction pathways are illustrated in figure 4.7. The activation barriers as well as ΔΔE, ΔΔG, and Boltzmann

distributions values in table 4.2. The reaction profile for the first step, formation of the cation, with the relative product energy for each isomer is given in figure 4.8.

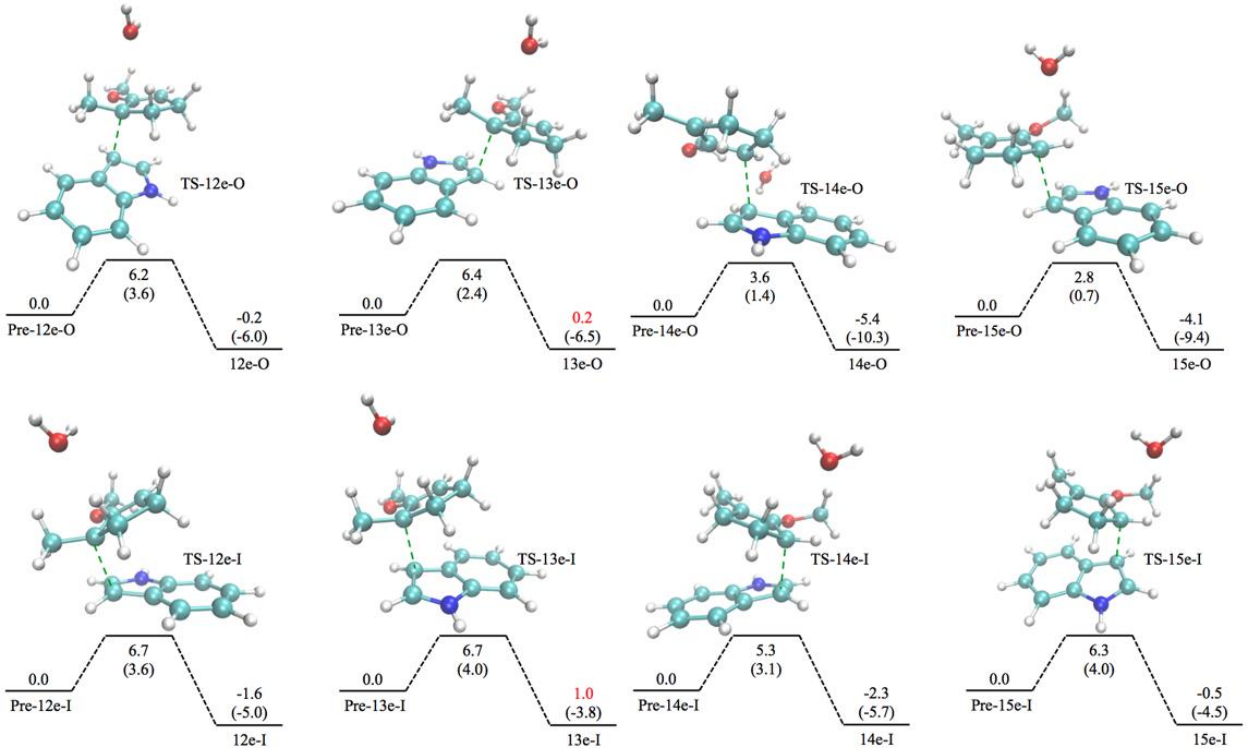

Figure 4.7. Reaction pathways for the addition of indole to the Me cation. The first row depicts the outward approach of indole to the cation, the second row depicts the inward approach. A green dashed line shows where the bond is starting to form in each transition state. Gibbs energies are given first followed by PE in parenthesis. Energies are given with respect to each pathways reactant. Red product energies indicate a product higher in energy than the reactant.

Experimentally, there is no preference for addition at either the α or α'-carbon. Thermodynamically speaking, the calculated ΔΔE and ΔΔG values have no appreciable difference between nucleophilic addition to the α and α'-carbons, matching experimental results. Furthermore, Boltzmann distributions at every state show a mixture of addition to both substituted and unsubstituted carbons. It is interesting to note that in this case, activation barriers are similar across all isomers (~5-6 kcal/mol) with the exception of the outward approach for TS-14e-O and TS-15e-O which require ~3kcal/mol for formation to occur.

Table 4.2. Structures in the table refer to energies of Me cations. Relative energies are given per state (reactant, transition, product) with respect to the most stable structure of that state (reactant, transition, product) [a] relative potential energy of each state [b] relative Gibbs free energy of each state [c] Boltzmann distribution of each state [d] activation barrier given as the difference between the transition state and reactant. Energies are given in kcal/mol.

	Reactant			Transition State				Product		
Motif	$\Delta\Delta E^{[a]}$	$\Delta\Delta G^{[b]}$	$B^{[c]}$	$\Delta G^{\ddagger[d]}$	$\Delta\Delta E^{[a]}$	$\Delta\Delta G^{[b]}$	$B^{[c]}$	$\Delta\Delta E^{[a]}$	$\Delta\Delta G^{[b]}$	$B^{[c]}$
12eO	3.1	2.2	0.0	6.2	2.7	2.1	0.0	1.3	3.1	0.0
13eO	2.8	1.1	0.2	6.4	1.2	1.2	0.1	0.4	2.3	0.0
14eO	6.2	4.3	0.0	3.6	3.7	1.7	0.1	0.0	0.0	1.0
15eO	6.2	4.2	0.0	2.8	3.0	0.7	0.3	0.9	1.2	0.1
12eI	1.3	1.5	0.1	6.7	1.0	2.0	0.0	0.4	1.0	0.2
13eI	0.0	0.1	0.8	6.7	0.0	0.6	0.4	0.2	2.2	0.0
14eI	2.2	1.7	0.1	5.3	1.3	0.8	0.3	0.6	0.5	0.4
15eI	0.4	0.0	1.0	6.3	0.4	0.0	1.0	0.0	0.6	0.4

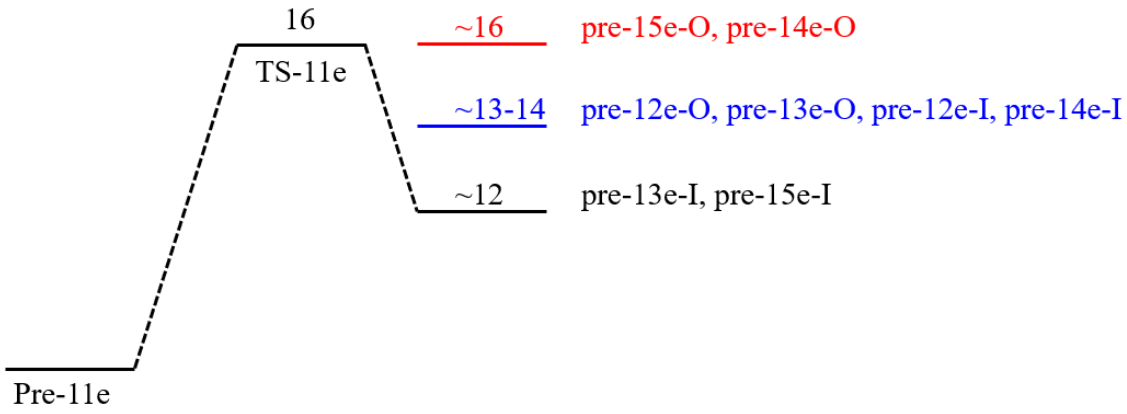

Figure 4.8 Illustrates the Free energy profile for the ionization of the TBS cation

However, despite the low activation barrier for indole addition, structures pre-15e-O and pre-14e-O, as seen in figure 4.8, are very similar in energy to their previous transition state leading towards cation formation. It would be expected that the reverse reaction would be highly competitive. It can be gathered from this figure that the main pathways for indole addition to the Me cation are formations of 15e-I (addition at the less substituted carbon) and 13e-I (addition at the substituted carbon), as the other isomers can undergo the reverse reaction. The activation

barriers for both configurations are similar, their respective reactants are essentially the same energy, and the products are within 2kcal/mol of each other. Formation of 15e-I and 13e-I are likely competing approaches, both thermodynamically and kinetically, which explains the lack of regioselectivity seen in experiment.

4.4.3. Effect of Water on Cation Stabilization

In order to stabilize the cation formation, it was necessary to include one water molecule. Without this water molecule present, a stable transition state could not be located in any optimizations or ab initio metadynamics simulations. Experiments have shown trace amounts of water can actually catalyze this reaction. Basis set superposition error (BSSE) calculations were performed to quantify the stabilizing effect of water bound to the TBS cation. Reactant, transition and product states for the formation of 15a-I were used because this is the most favorable pathway of the TBS configurations analyzed. The water was found to stabilize pre-15a-I by -7.7 kcal/mol, TS-15a-I by -9.8 kcal/mol, and 15a-I by -10.05 kcal/mol.

4.4.4. Steric Effects on Regioselectivity

Given the above discussion, it is hypothesized that the regioselectivity seen in the TBS-cation case is due to steric hindrance from the large protecting group. In order to gain a more quantitative view of the effect of steric hindrance, SAPT energy decomposition analysis was performed. Due to the complicated nature of transition states, the composition analysis was calculated at two levels of theory in order to check for consistency. SAPT energy decompositions were executed on α-carbon addition configuration 13 and α'-carbon configuration 15 for both Me and TBS variants. Decomposition analysis looked at the reactant, and transition state. The results calculated with jun-cc-pvdz are shown in table 4.3, and aug-cc-pvdz are shown in table 4.4.

Table 4.3. SAPT energy decomposition calculated with the jun-cc-pvdz basis set. All energies are given in kcal/mol

Configuration	Electrostatics	Exchange	Induction	Dispersion	Total
Pre-15a-I	-17.35	27.36	-10.31	-18.73	-19.03
Pre-13a-I	-16.24	25.13	-9.65	-17.97	-18.73
TS-15a-I	-41.57	79.07	-44.90	-27.58	-34.98
TS-13a-I	-56.12	105.54	-71.67	-33.24	-55.49
Pre-15e-I	-15.92	23.10	-9.67	-16.26	-18.75
Pre-3e-I	-15.62	23.16	-9.66	-16.17	-18.30
TS-15e-I	-43.39	81.67	-49.51	-26.89	-38.12
TS-13e-I	-44.22	82.09	-52.39	-27.63	-42.15

Table 4.4. SAPT energy decomposition calculated with the aug-cc-pvdz basis set. All energies are given in kcal/mol

Configuration	Electrostatics	Exchange	Induction	Dispersion	Total
Pre-15a-I	-17.01	25.68	-11.27	-20.78	-23.37
Pre-12a-I	-15.85	23.34	-10.48	-19.94	-22.92
TS-15a-I	-42.27	81.87	-49.35	-29.54	-39.29
TS-12a-I	-56.67	108.60	-76.11	-35.08	-59.26
Pre-15e-I	-15.46	21.27	-10.08	-18.19	-22.46
Pre-12e-I	-15.15	21.30	-10.06	-18.08	-21.99
TS-15e-I	-43.97	84.76	-53.92	-28.82	-41.95
TS-12e-I	-44.64	84.44	-55.91	-29.59	-45.70

Both jun-cc-pvdz and aug-cc-pvdz calculations show the same trends. In SAPT energy decomposition, the exchange term gives a quantitate view of steric hindrance. There is no strong difference between pre-15aI/pre-13aI or pre-15eI/pre-13eI indicating that the steric hindrance is similar for both configurations at the reactant step. Upon the formation of the transition state, the exchange term for TS-13e-I becomes much more unfavorable than that of TS-15e-I by more than 20 kcal/mol. Conversely TS-13eI and TS-15e-I have similar exchange terms. This is a strong indication that steric effects play a major role in the regioselective addition of indole to unsymmetrical TBS cations.

4.4.5. NBO Analysis of the Intermediates

NBO charge analysis was performed to see the effect of the lone pair (LP) electrons of the oxygen of the protecting group, as well as to determine if the addition of indole is charge driven as well as sterically driven. Amount of hybridization can be seen in table 4.5. Calculations were done for the product of the first step, with and without indole.

Table 4.5. Hybridization of oxygen lone pairs, with and without indole.

Group	Oxygen LP	% S	% P	% D
11a	LP1	18.30	81.64	0.06
	LP2	13.56	86.36	0.07
11a-NI	LP1	19.87	80.07	0.06
	LP2	11.32	88.61	0.07
11e	LP1	38.51	61.43	0.06
	LP2	0.07	99.83	0.10
11e-NI	LP1	38.58	61.36	0.06
	LP2	0.00	99.90	0.10

Molecular orbitals for the lone pairs of the oxygen of the protecting group were visualized and can be seen in figure 4.9. The second LP electron on each oxygen showed strong p-character, regardless of the protecting group. It was determined that despite the strong p-character, there was no overlap between the lone pair oxygen orbitals and the indole ring. Less hybridization occurs when the protecting group is Me.

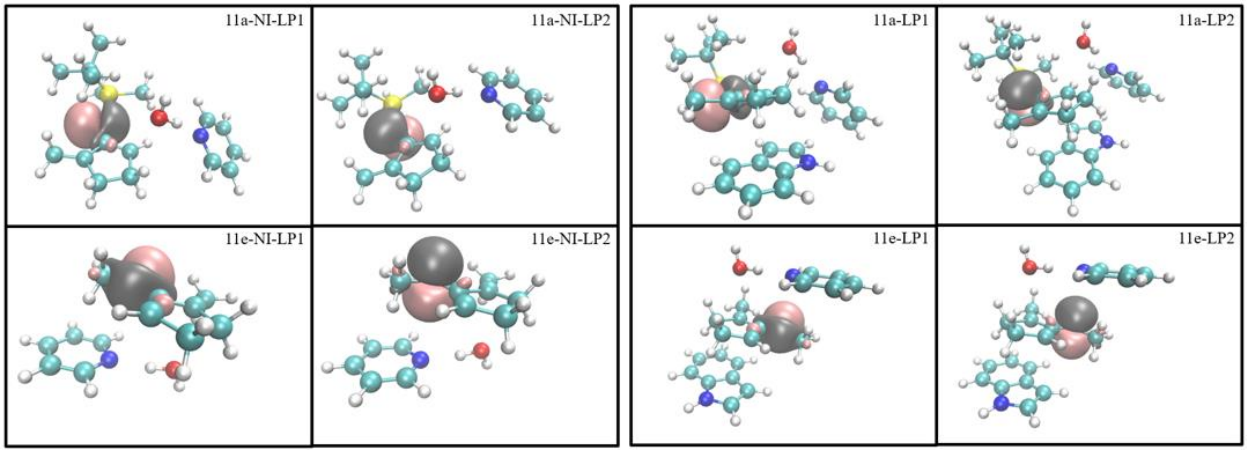

Figure 4.9. Molecular orbitals of the 11a and 11e oxygen lone pairs. NI indicates indole is not present and LP1 refers to the first oxygen lone pair, while LP2 refers to the second

NBO charges were determined for the α and α'-carbon as well as the carbon attached to the protecting group, defined as CR. These calculations were performed on the product of cation formation, the reactant for indole addition in both the 13 and 15 type isomers Me and TBS case. Calculations are done with and without water as water has been found to stabilize the reaction. Charges can be seen in table 4.6 below. From the charge analysis, one would assume that both the TBS cation and Me cation would see preferential binding at the unsubstituted carbon. Charges are consistent across all structures indicating that both cations should experience a regioselective addition. As the experiments have determined that the TBS cation reports regioselectivity whereas the Me cation does not, it has been determined that this reaction is not charge controlled. There is no significant change in charge when the water is present, indicating that the water has a larger effect on stabilizing the transition state for the addition of indole, rather than the cation itself, which likely has its largest stabilizing effect from the indole ring.

Table 4.6. NBO charge analysis for α, α', and CR carbons. All charges have units of elementary charge NW refers to charge analysis without the water present

Structure	α	α'	CR
11a	0.277	0.015	0.246
Pre-15a-I	0.219	-0.033	0.246
Pre-15a-I-NW	0.223	-0.044	0.251
Pre-13a-I	0.264	-0.041	0.242
Pre-13a-I-NW	0.231	-0.043	0.250
11e	0.301	-0.023	0.232
Pre-15e-I	0.230	-0.014	0.229
Pre-15e-I-NW	0.239	-0.053	0.234
Pre-13e-I	0.278	-0.059	0.229
Pre-13e-I-NW	0.245	-0.058	0.234

4.5. Last Step of the Reaction Mechanism and Overall Reaction Profile

The last step of the reaction, deprotonation, creates a very stable product, which can be seen in figure 4.10. The overall reaction profile for indole addition to the α'-carbon of the TBS cation is shown in figure 4.11, and the addition to the α'-carbon of the Me cation is shown in figure 4.12. Both the final product in the Me and TBS case are very stable. For the complete reverse reaction to occur, a reverse barrier of almost 24 kcal/mol would have to be crossed, which is unlikely to happen in standard condition.

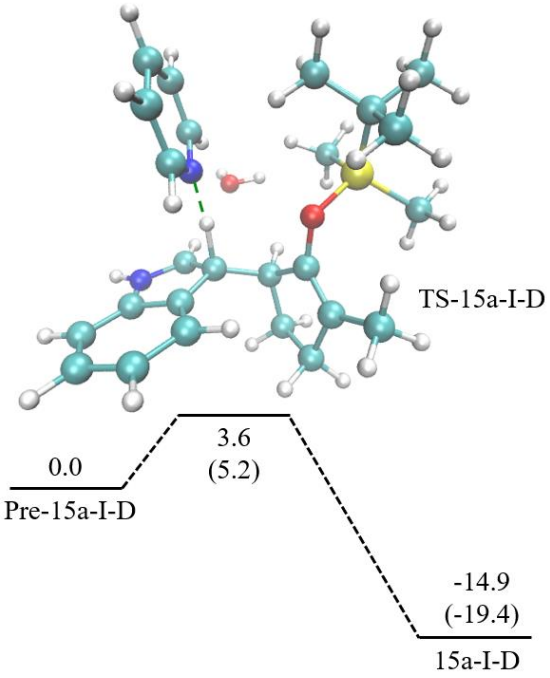

Figure 4.10. Final deprotonation of the TBS-indole complex

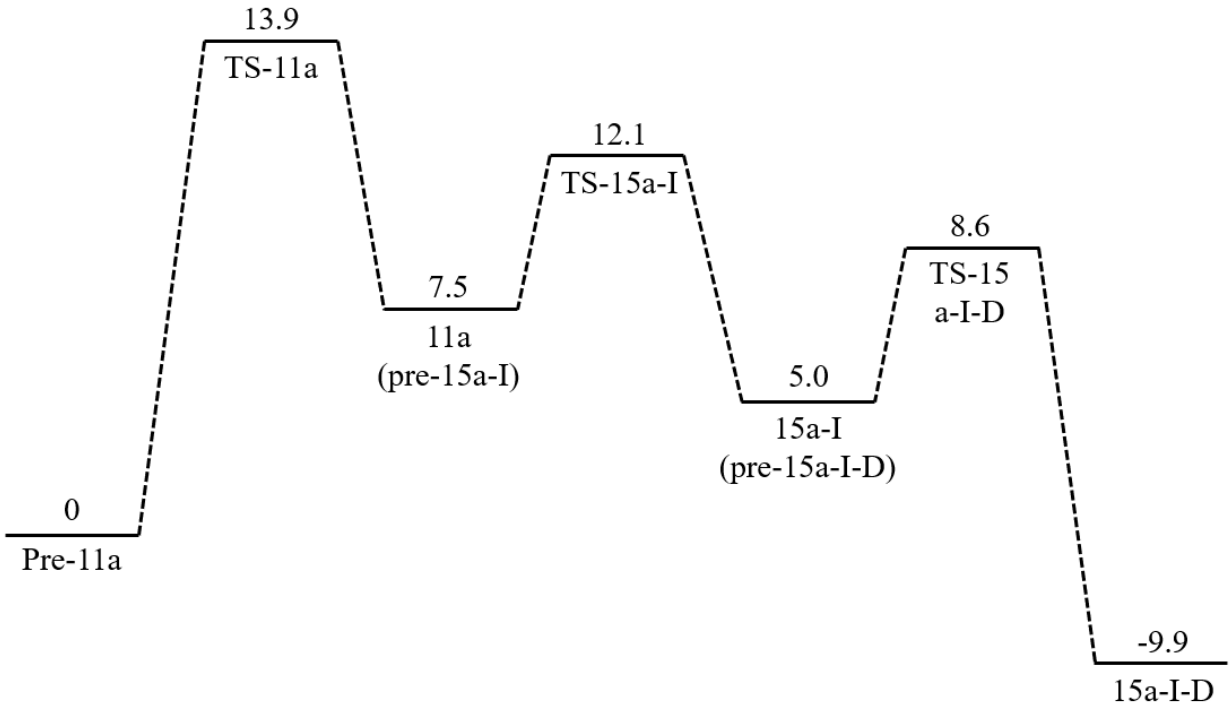

Figure 4.11. Overall reaction profile for the addition of indole to TBS cation

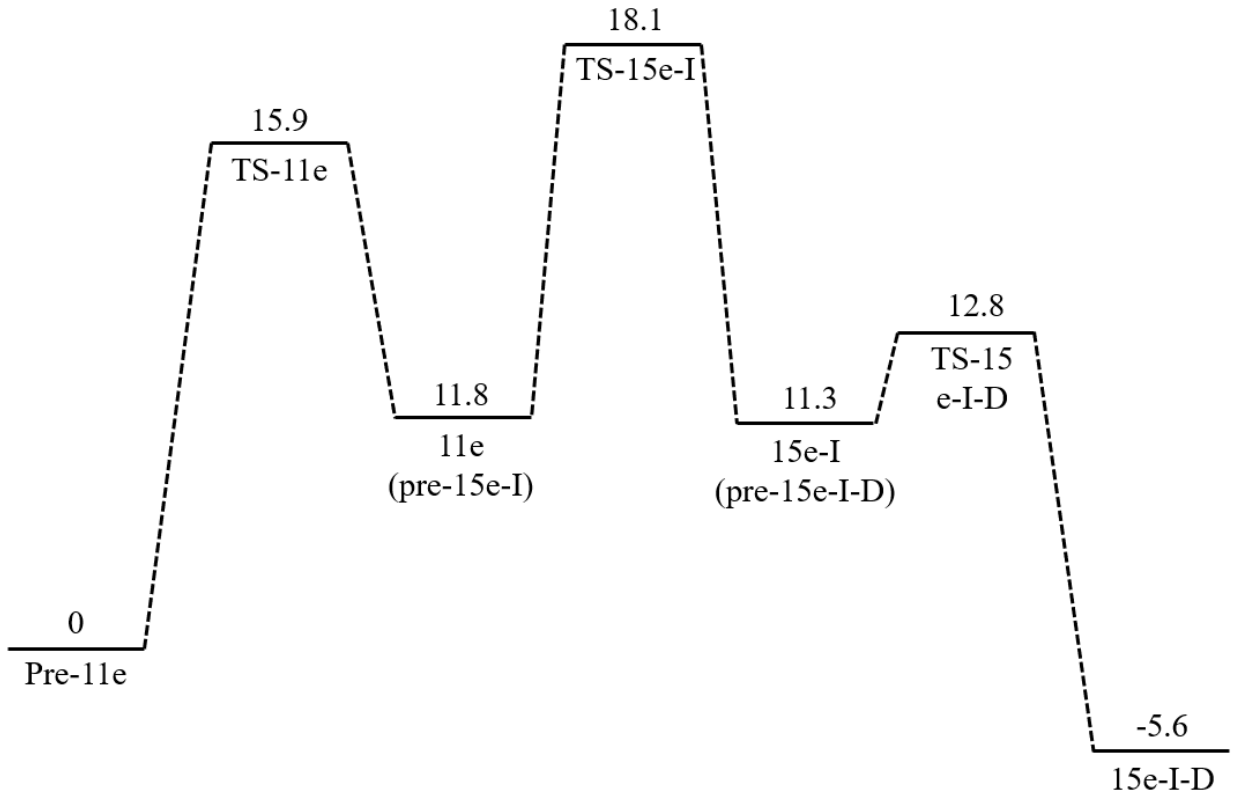

Figure 4.12. Overall reaction profile of indole addition to Me cation

4.6. Conclusion to the Regioselectivity Investigation

Regioselectivity for the addition of indole to the unsubstituted α'-carbon of TBS cations has been extensively studied in this work. Indole addition to the Me cation was used as a litmus test to discern the differences, as the Me cation does not undergo a regioselective addition. The mechanism proceeds through an SN1 formalism, where the first step of cation formation is the rate determining step. Indole was found to stabilize the cation due to π-stacking. For the formation of the TBS cation, the reverse reaction had a barrier of 6.5 kcal/mol when indole was present vs. only 3 kcal/mol without indole.

All possible reaction pathways were calculated. When the protecting group of the cation was TBS, it was found that the inward approach was favored over the outward, likely due to the π-stacking with indole. Nucleophilic addition via the α'-carbon of the TBS cation was found to be

both kinetically and thermodynamically favored over α-addition. Me cations have two main competitive pathways for indole addition. One has the C-C bond formation at the substituted α-carbon, and one with formation at the unsubstituted α'-carbon. Between these two pathways, there was not a discernable difference between their respective $\Delta G^{\ddagger}$, $\Delta\Delta E$, or $\Delta\Delta G$ values. Trace water was found to have a stabilizing effect on every step of these intermediate reaction.

NBO charge analysis determined that the reaction was not charge controlled, and the lone pairs on the oxygen of the protecting group had little effect on indole addition. SAPT energy decomposition confirmed that there was a large difference between the exchange term for the addition of indole to the two binding sites of the TBS cation. This indicates that steric effects strongly dominate the regioselectivity of the TBS cation.

The deprotonation of both the TBS-indole complex and Me-cation complex generates a very stable product. In fact, the reverse barrier is around 24 kcal/mol, and is improbable once the deprotonation of the complex is complete. Overall, many aspects of the reaction mechanism were probed, and the sources of regioselectivity were identified.

Chapter 5. DFT Analysis into the Formation of an EDA Complex Hypothesized from O-Glycosylation Experiments

5.1. Experimental Motivation: O-Glycosylation

The synthesis of oligosaccharides is of fundamental importance due to their critical role in biological functions. Oligosaccharides participate in mechanisms involving cell-cell interactions, immune response, etc.[163] In order to investigate biological systems, it is imperative to have a reliable synthetic process to generate these oligosaccharides. The O-glycosylation step is an important aspect in the synthetic process of oligosaccharides, and often use thioglycosides as the donor.[164-166] The Ragains group has focused on methodologies that utilize a stable donor in order to store chemicals for a longer period of time; procedures able to progress at ambient temperatures on a standard benchtop.[167, 168] The goal of the Ragains group is to proceed through the O-glycosylation step using a stable thioglycoside donor that is not easily activated, in mild environments in order to generate stereoselective O-glycosylations. Thioglycosides are frequently used in O-glycosylation due to their stability as well as other properties such as the ability to tune their reactive properties as well as synthetic ability. They were able to achieve their goal of activating 4-p-methoxyphenyl-3-butenylthioglycosides for use in O-glycosylation in mild reaction conditions. Initially, the activation occurred with the substrate in tandem with a photocatalyst and Umemoto's reagent. Their investigation took an interesting turn; while testing the control solutions, without the photocatalyst, it was observed that the solution of Umemoto's reagent and 4-p-methoxyphenyl-3-butenylthioglycoside resulted in a yellow solution. UV-vis spectra revealed absorbance in the visible region. In order to understand the structural motif responsible for the color change, spectrophotometry of 4-p-methoxyphenyl-3-butenylthioglycoside with Umemoto's reagent was compared to 4-methoxystyrene complexed with Umemoto's reagent, resulting in identical results. This proved the hypothesis that the electron rich styrene section of the

4-p-methoxyphenyl-3-butenylthioglycoside is the cause of the color change. It is expected that an electron-donor-acceptor (EDA) complex has been generated, and as such, Density Functional Theory (DFT) calculations were performed in order to determine the formation of an EDA complex, the extent of charge transfer in the complex, as well as simulated UV-Vis spectra with Time-Dependent DFT (TD-DFT). This work is published in Angewandte Chemie International Edition under the title "A Visible-Light-Promoted O-Glycosylation with a Thioglycoside Donor".[167]

5.2. Computational Methodology

5.2.1. Electronic Structure Methods

Rather than using the entire 4-p-methoxyphenyl-3-butenylthioglycosides substrate, a model system, p-methoxystyrene was utilized. Calculations were also performed with styrene, in which the experiments did not see any color changes, as a control for comparison. Structures of p-methoxystyrene, styrene, and the S-trifluoromethyldibenzothiophenium cation, which will be referred to as the Umemoto's reagent, were all optimized using the B3LYP functional 6-31+G(d) and Grimme D2 dispersion.[169] Calculations were also performed with the PBE0 functional in order to check for consistency in results and functional dependence. These calculations also proceeded with the 6-31+G(d) basis set and the Grimme D2 dispersion correction. All calculations were carried out with the Gaussian09[134] suite of programs, except when indicated. In order to mimic the experimental environment, an implicit solvent, using the Polarizable Continuum Model with the integral equation formalism, of acetonitrile was added with a dielectric constant value of ε=35.688.[55] The methoxystyrene complex has two isomers, a and b, which depend on the orientation of the oxygen group as can be seen in the first panel of figure 5.1. Once the initial minimized structures were determined, different configurational isomers of Umemoto's cation

with styrene, methoxystyrene-a, and methoxystyrene-b were generated and minimized to determine how Umemoto's cation orients over the differing styrene moieties. Once the configurational isomers were determined, any within 1.0 kcal/mol of the lowest energy configurational isomer was taken for further analysis.

5.2.2. Atomistic Charge Calculations

A lot of debate exists over the best way to calculate charge as it is not a direct quantum observable. In order to get a qualitative view of charge transfer in this system, two different methods are used for the charge calculation. The first is the commonly used Natural Bond Orbital (NBO) methodology.[160] NBO allows for analysis and property determination based on the input wavefunction. It essentially reconstructions the wavefunction into a set of bound, or unbound electrons, in order to recreate the Lewis structure approach. One drawback of the NBO methodology is that it is dependent on the description from the basis set, albeit less so than Mulliken charge analysis, as well as the ability to describe the system by a Lewis structure. NBO analysis should give a qualitative view of charge transfer, but in order to confirm the qualitative view of charge transfer, the Bader Atoms In Molecules (AIM)[170] approach was utilized as well. In theory, the AIM model should be more transferable across different basis sets and has an elegant and logical partitioning scheme. Like with NBO, the charge is built from the electronic density, however, the partitioning is different. AIM relies on the assumption that for any property of the system, the sum of the pieces should equal that of the whole property of that molecule, the physics of any given atom should be consistent across the atom type, and partitioning of the atoms occurs in such a way that the form of the atom is not significantly distorted. A grid is applied to the wavefunction and using the gradient, critical points are determined and surfaces diagrams are generated allowing for the partitioning of the molecule into atoms. A charge is assigned to each

atom after integrating over the density of the partitioned surface. In this way, the accuracy is determined on the grid rather than the method used. NBO analysis was performed with the Gaussian 09[134] suite of programs while the AIMAll package[171] was used for calculations with the Bader AIM approach. Both styrene and methoxystyrene are neutral molecules. In order to quantify the extent of charge transfer, the charges of each atom in styrene or methoxystyrene complexed with Umemoto's reagent were summed over, as seen in equation 5.1. In this way, the extent of charge transfer is clear as a charge larger than the neutral charge of the styrene moieties by themselves indicates that electrons have transferred to the electron deficient Umemoto's reagent.

$$q_{total} = \sum_{i=atom}^{molecule} q_i \tag{5.1}$$

Lastly, UV-Vis spectra of the first six excited states were calculated at the same level of theory as above using TD-DFT[172] in the Gaussian 09[134] suite of programs.

5.3. Charge Transfer and EDA Complex Formation

Methoxystyrene-a and methoxystyrene-b are similar in energy. All relevant isomers of the complexed structures, as well as the numbering scheme for the styrene motifs, can be seen in figure 5.1. Three configurational motifs for the complexation of the styrene moieties and Umemoto's reagent have been isolated. The first has the sulfur of Umemoto's reagent over the C1 carbon (see the top panel of figure 5.1 for the numbering scheme) of the styrene moiety. In the second motif, C1/C8-type, the sulfur complexes over the u-bend created by C1-C2-C3-C8; the sulfur is above the center of space created by the dihedral. The third configurational motif has sulfur complexing over the C4 carbon of the styrene moiety.

All data shown has been calculated with the B3LYP functional. Calculations performed with the PBE0 functional resulted in similar structures while the charge analysis showed the same

trends. Table 5.1 contains the bond distances between the sulfur of Umemoto's reagent and the carbon on the styrene or p-methoxystyrene. It is followed by the charge on each structure resulting from complexation with Umemoto's reagent calculated by NBO charge analysis and the AIM approach. The charge is given in units of elementary charge.

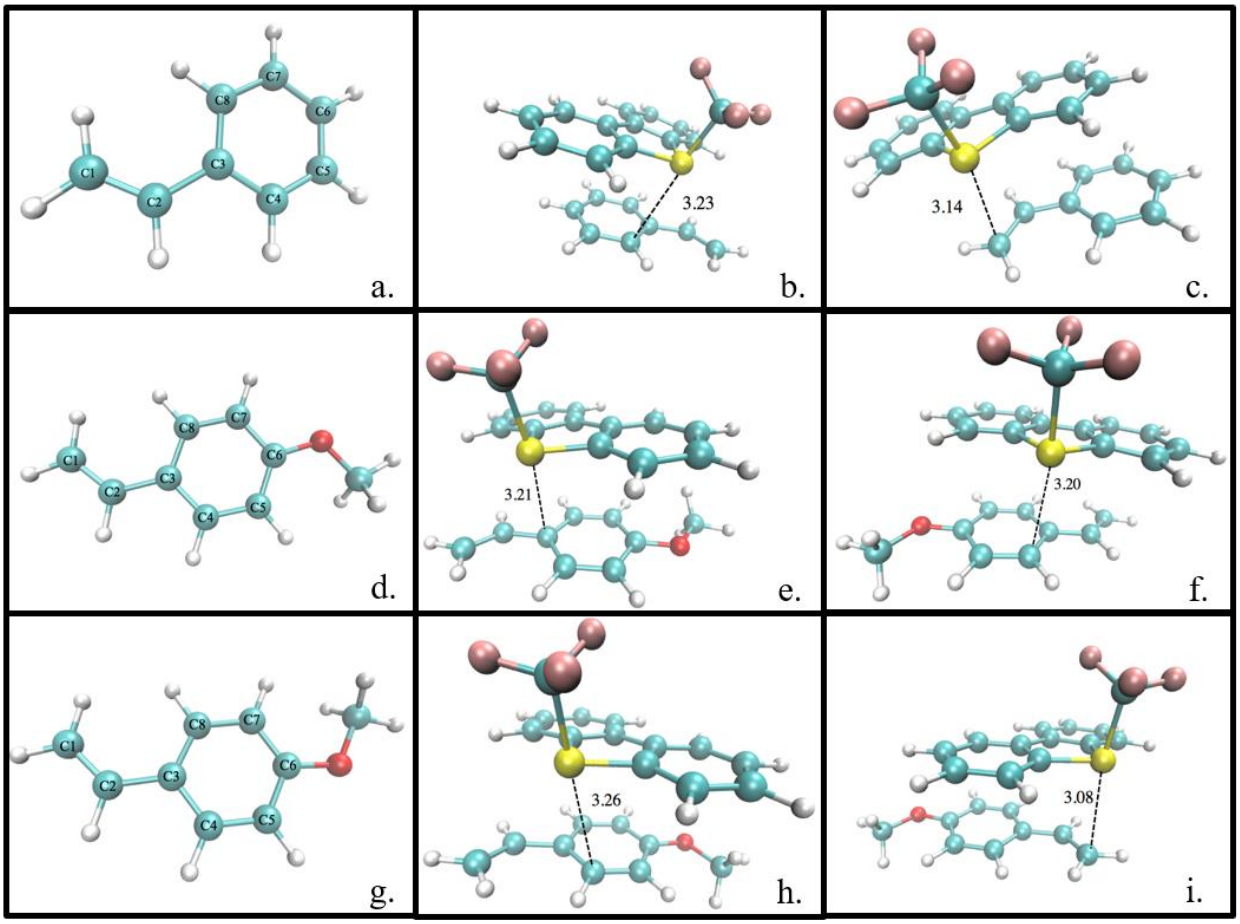

Figure 5.1. The first column contains the numbering scheme for the styrene moieties, the second contains all C1/C8 configurations and the third illustrates the C1 or C4 configurations. The distance between the sulfur of Umemoto's reagent to the closet carbon of the styrene moiety is shown and all distances are given in angstroms. a. styrene numbering scheme b. C1/C8 complex of styrene and Umemoto's reagent. c. C1 styrene-Umemoto's complex d. the numbering scheme for methoxystyrene-a. e. C1/C8 methoxystyrene-a-Umemoto's complex f. methoxystyrene-a C4 type motif with Umemoto's reagent g. methoxystyrene-b numbering scheme h. methoxystyrene-b-Umemoto's complex in the C1/C8 configuration. i. Umemoto's reagent and methoxystyrene-b complexed in the C1 configuration

No C1 type was found for methoxystyrene-a, although, a motif of C4 was found, which did not exist within 1.0 kcal/mol of the lowest isomer for the Umemoto's complex with methoxystyrene-b or styrene. Styrene was found to have minimal charge transfer as expected, nevertheless, the C1 configuration saw more charge transfer than the C1/C8 type.

Table 5.1. Lists all structures complexed with Umemoto's reagent, followed by the configurational motif. Bond distance of the sulfur of Umemoto's reagent to the closest carbon of the styrene moiety is listed in Å, followed by charges determined by NBO and AIMAll analysis, given in units of elementary charge.

Structure Complexed with Umemoto's Reagent	Conformation Type	Bond Distance	NBO charge on Structure	AIMAll charge on Structure
Styrene	C1/C8	3.23	0.023	0.057
Styrene	C1	3.14	0.038	0.065
Methoxystyrene-a	C1/C8	3.21	0.035	0.061
Methoxystyrene-a	C4	3.20	0.021	0.046
Methoxystyrene-b	C1/C8	3.26	0.036	0.59
Methoxystyrene-b	C1	3.08	0.065	0.096

The methoxystyrene-a case has less charge transfer than the methoxystyrene-b configuration. There is less charge transfer seen when Umemoto's reagent is coordinated to methoxystyrene-a in the C4 configuration, in comparison to the C1/C8 case. However, when methoxystyrene-b complexes with Umemoto's reagent there is significantly more charge transfer in the C1 motif vs the C1/C8 motif. This is likely partially explained by the proximity of methoxystyrene-b to Umemoto's reagent in the C1 configuration. This structure has a shorter interplanar distance than any other configuration analyzed, perfect for a charge transfer complex. When comparing the C1 and C1/C8 configurations between the methoxystyrene-b and styrene complex it is quite clear that there is significantly more charge transfer in the methoxystyrene-b case. Furthermore, the interplanar distance shows that these complexes are ideally suited for as an EDA complex.

UV-Vis spectra were calculated for the first six excited states of every configurational motif at both the B3LYP and PBE0 level. The spectra are depicted in figure 5.2 (styrene and Umemoto's reagent), and 5.3 (methoxystyrene-b and Umemoto's reagent), and are representative of the trends seen and are calculated with the B3LYP functional for the C1 Umemoto's reagent/methoxystyrene-b structure that showed the most charge transfer, and its corresponding C1 styrene Umemoto's reagent counterpart. It can be seen that methoxystyrene-b has a transition in the visible range, in the representative case at around 500 nm, nonetheless, all calculations consistently show a transition ranging from 485-515 nm.

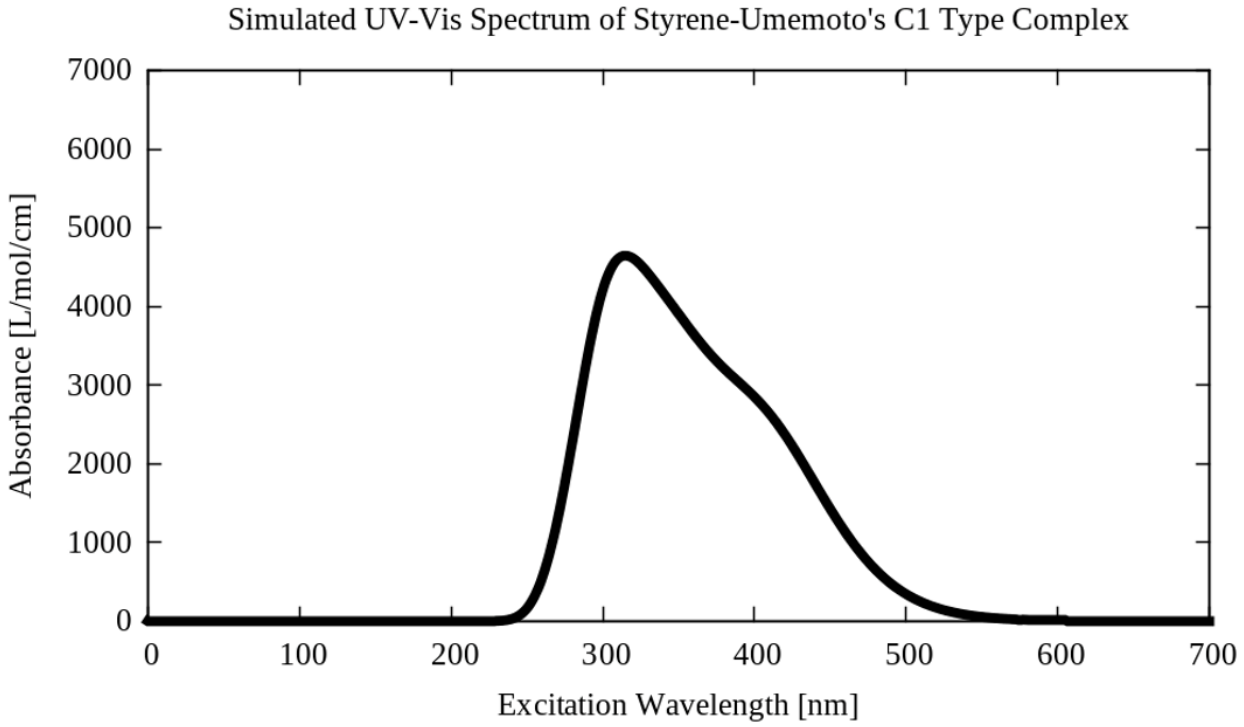

Figure 5.2. Simulated UV-Vis spectrum of styrene and Umemoto's reagent, C1 type motif.

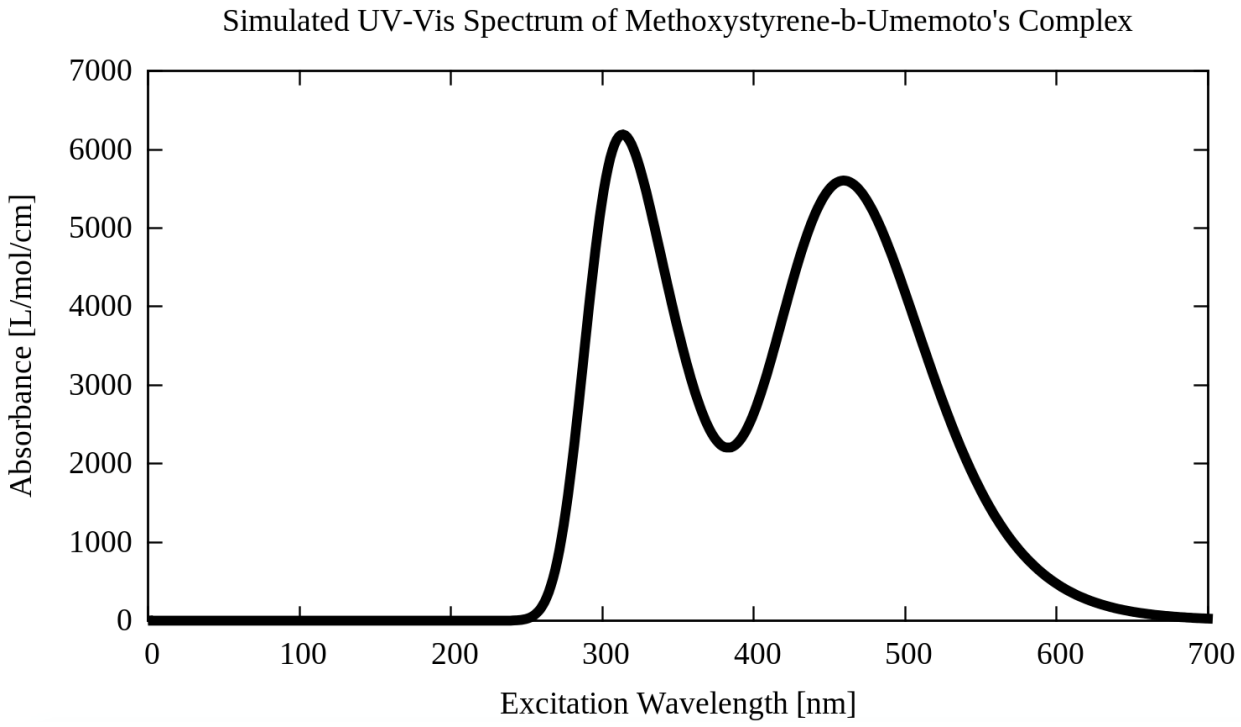

Figure 5.3. Simulated UV-Vis spectrum of methoxystyrene-b and Umemoto's reagent, C1 type

5.4. Conclusion

DFT calculations have been performed in order to determine the extent of charge transfer between three complexes 1) Styrene and Umemoto's reagent 2) Methoxystyrene-a and Umemoto's reagent and 3) methoxystyrene-b and Umemoto's reagent. It was determined that methoxystyrene-b complexed with Umemoto's reagent in a C1 configuration showed the most significant amount of charge transfer, significantly more than the corresponding C1-styrene/Umemoto's reagent complex. Further investigation into the interplanar distance indicates that the complexes are at an ideal distance for an EDA complex. UV-Vis spectra confirm the results of the experiments done by the Ragains group, with a transition in the visible region once p-methoxystyrene is complexed by Umemoto's reagent.

Chapter 6. Conclusion and Outlook

6.1. Conclusions

The theme of this dissertation has spanned reactivity from electronic structure methods to force field development. Modeling chemical reactivity is difficult due to many factors such as the system size, description of molecular interactions, computational resources, etc. Three distinct systems have been analyzed in this work with electronic structure methods; hydrated HCl, indole addition to methyloxyallyl and silyloxyallyl cations, and the formation of an EDA complex theorized in the O-glycosylation processes. The data gathered from electronic structure calculations of hydrated HCl clusters were used in the development of a reactive force field for HCl-water systems.

In chapter two, AIMD and MTD simulations were performed in order to determine the solvation structure of HCl water clusters ranging in composition from $HCl(H_2O)_2$ to $HCl(H_2O)_{22}$. In all, one hundred and nineteen structures were established. These structures consisted of four solvation types, covalent HCl and water, contact ion pairs, solvent separated ion pairs, and no ion pairing. Additionally, structures spanned a wide range of energetics rather than focusing on the minimum energy structure at each cluster size. No trend between solvation motif and energetic scale was found, so other configurational properties such as the number of free OH groups, which indicates the extent of hydrogen bonding, the distance between Cl^- and the center of the cluster, and distance of H_3O^+ to the center of the cluster were examined. No distinct correlation between any one of these properties and solvation type, or energetic trends, were identified. Indeed, the hydrogen bonding network of these cluster systems is found to be quite complex. These studies clearly underlined the importance of developing accurate force fields, given that very different

solvation environments are found to be of comparable energy thereby requiring the incorporation of the essential physics to accurately model these systems.

Chapter three describes the development of a reactive force field which incorporates covalent HCl. The model is based off the MS-EVB3.2 formalism which was developed for an excess proton in water. The coupling term is a combination of Gaussians dependent on the H-Cl and O-Cl distance of the complex Cl-H-OH_2 Zundel-like state. Chloride is a highly polarizable ion; rather than integrate a complicated iterative polarization scheme, a simple short range geometric three-body potential dependent on the heavy atoms of the water-chloride-water and water-hydronium-chloride interactions was employed. Eleven parameters in total were fit, the off-diagonal terms parameterized to the Cl-H-OH_2 potential energy surface, along with four structures that were stronger CIP rather than HCl/H_2O character, and the three-body terms were parameterized to the potential energies of twenty-eight structures that encompassed HCl, CIP, and SSIP solvation motifs as well as relative energies as high as approximately 6 kcal/mol. The new force field, MS-EVB-HCl, accurately reproduced the Cl-H-OH_2 potential energy curve as well as the relative energetics of structures at all cluster sizes explored in the previous chapter. MS-EVB-HCl showed improvement over the MS-EVB3.2 for structures containing covalent HCl and strong ion pairing. The MS-EVB-HCl model will be incorporated into enhanced sampling schemes to determine the pathway of HCl-induced gas liquid nucleation. The model potential will also be used to analyze HCl at the air-water interface. With this model, the effects of the counterion, Grotthuss proton shuttling, presence of Zundel and Eigen states (and potential disruption due to the counterion), and the solvation environments present in these systems can be determined.

Chapter four used DFT methods to investigate the regioselectivity seen in the addition of indole to methyloxyallyl and silyloxyallyl cations. The reaction pathway for the first step of cation formation, the intermediate addition of indole to the cation, which resulted in eight different possibilities (addition to the α or α'-carbon, for which two possible isomers were available at each binding site, from both an inward and an outward approach of the indole ring) per cation, and final deprotonation of the complex were examined. Relative energetics at each state (reactant, transition, and product) were determined as well as the Boltzmann distribution. Addition to the α'-carbon was determined to be both kinetically and thermodynamically favored for the silyloxyallyl cation, while kinetic and thermodynamic analysis resulted in two favorable additions for the methyloxyallyl cation, one at the α-carbon, the other at the α'-carbon. SAPT energy decomposition was utilized to quantify the amount of steric hindrance, which indicated that the regioselectivity seen in the addition of indole to the silyloxyallyl cation is in fact, steric controlled. Molecular orbital analysis determined that the indole plays a stabilizing role on the cation, while the lone pairs of the oxygen of the protecting group do not play a role in π-bonding to the indole ring. NBO analysis proved that charge is not a controlling factor in the regioselectivity.

Chapter five used a model complex to determine the properties seen by styrene moieties in the O-glycosylation step of oligosaccharide synthesis. Experiments found that the addition of Umemoto's reagent to 4-p-methoxyphenyl-3-butenylthioglycosides lead to a color change. The addition of methoxystyrene to Umemoto's reagent saw similar color changes, whereas the styrene-Umemoto's reagent solutions experienced no color changes. To this end, a configurational search for the different complexes of Umemoto' reagent to of methoxystyrene and styrene was performed. Three main configurations were determined. The amount of charge transfer was found to be dependent both on the configuration, as well as the species. Methoxystyrene Umemoto's reagent

complexes generally experience more charge transfer than styrene and Umemoto's reagent complexes. Distance between the sulfur of Umemoto's reagent and the closest carbon indicate an EDA complex has formed. Additionally, the simulated UV-Vis spectrum for the methoxystyrene case had a strong transition in the visible region.

6.2. Future Directions

The main focus of the thesis is the development of reactive models to study the effect of acidic defects in clusters from counterion effects to gas-liquid nucleation as well as counterion effects at aqueous solution-hydrophobic interfaces (such as air-aqueous solution interfaces). While this thesis focused on the development of the reactive model for HCl-water system, the next set of studies will concentrate on the effect of HCl at the air-aqueous interface as well as in gas-liquid nucleation. Previous studies in the literature did not allow for the acidic proton to hop to the Cl^- anion, but the electronic structure studies revealed the importance of these structures in small to medium sized clusters. While in the bulk phase HCl is completely dissociated, the air-water interface presents a very different environment wherein CIP and undissociated HCl can become important, especially in the case of concentrated acidic solutions. The next step will be the development of an MS-EVB formalism to analyze H_2SO_4 water cluster systems, as sulfuric acid has been found to be a major driver of gas liquid nucleation. This model is more complex as a larger variety of states can exist (H_2SO_4, HSO_4^-, SO_4^{-2}, and up to two H_3O^+) and an iterative scheme is required to handle the presence of multiple excess protons, which is also the case for concentrated HCl solutions. The development of a reactive scheme, from data generation to force fields, for these two acids will form the basis of a generalized platform that can be used for other acidic species. The codes developed during the course of this work will be released on an open source platform thereby contributing to progress in the field of reactive simulations.

Appendix A. $HCl(H_2O)_{2\text{-}22}$ Solvation Structures

All energetics in appendix A are given relative to the lowest energy structure at the specified cluster size and have units of kcal/mol.

A.1. Solvation Patterns of $HCl(H_2O)_{2\text{-}6}$

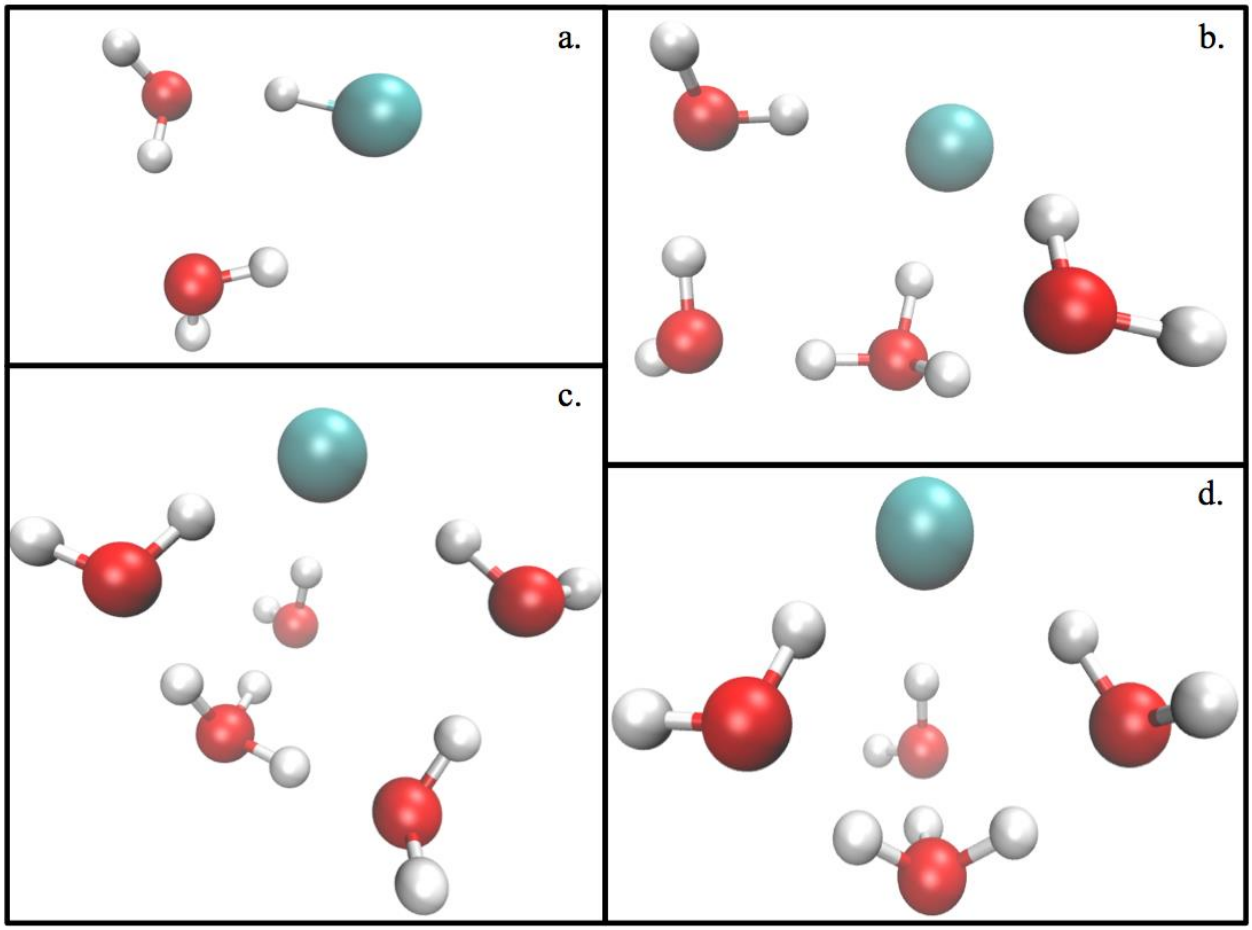

Figure A.1.a. Triangular ring structure of $HCl(H_2O)_2$ b. CIP solvation structure of $HCl(H_2O)_4$ with a relative energy of 1.6 kcal/mol c. The lowest energy $HCl(H_2O)_5$ cluster which is similar to zwitterionic $HCl(H_2O)_4$ SSIP structure d. $HCl(H_2O)_4$ zwitterionic SSIP structure

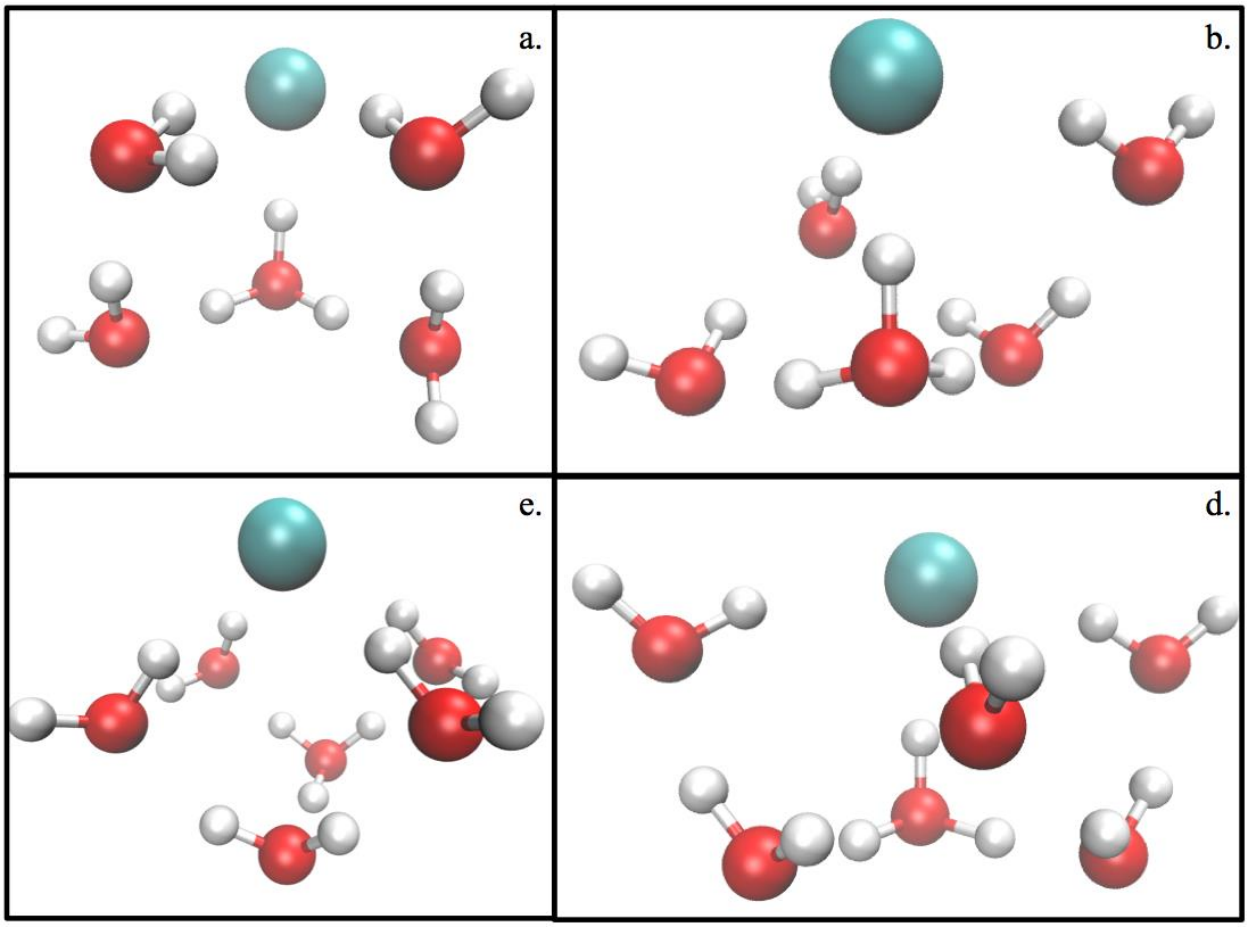

Figure A.2.a. lowest energy CIP for $HCl(H_2O)_5$ with a relative energy of 0.25 kcal/mol. b. second lowest energy CIP at $HCl(H_2O)_5$ e. overall minimum energy structure for $HCl(H_2O)_6$. It is a SSIP and is similar to the zwitterionic n=4 structure d. CIP minimum of $HCl(H_2O)_6$ with a relative energy of 1.5 kcal/mol

A.2. Solvation Patterns of $HCl(H_2O)_{9\text{-}12}$

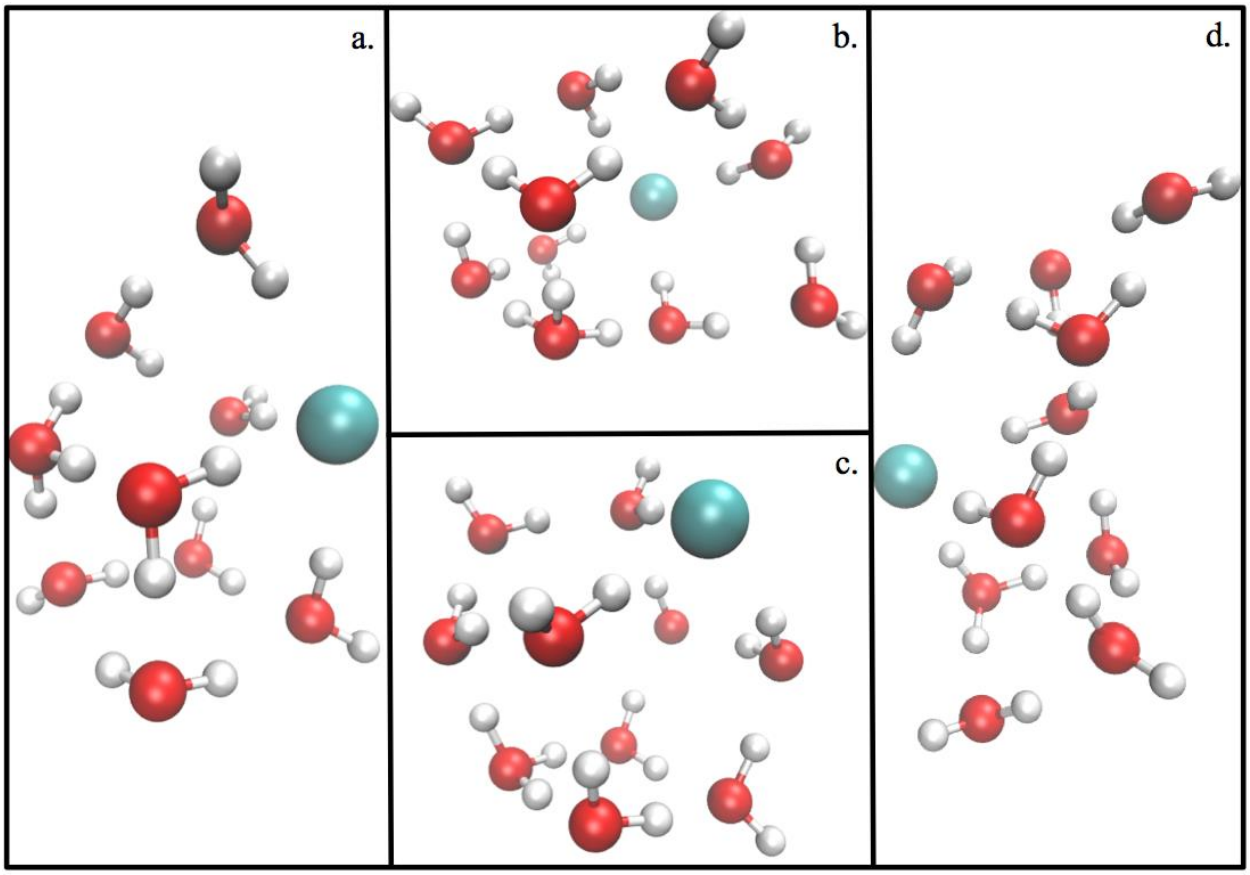

Figure A.3.a. $HCl(H_2O)_9$ SSIP minimum structure b. $HCl(H_2O)_{10}$ lowest energy structure which is SSIP. c. lowest energy $HCl(H_2O)_{10}$ NoIP solvation environment with relative energy 1.0 kcal/mol d. $HCl(H_2O)_{10}$ minimum energy CIP with relative energy of 5.3 kcal/mol

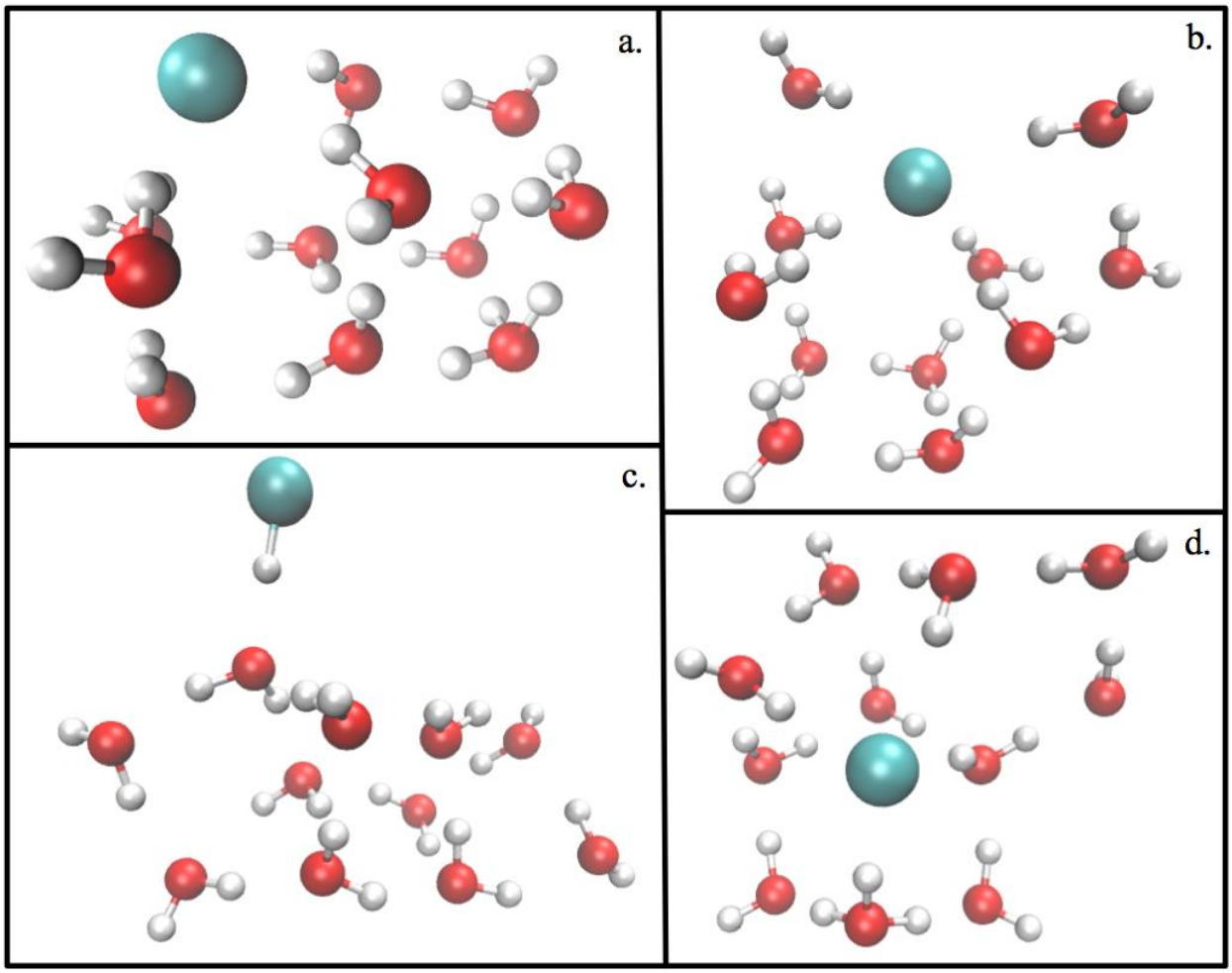

Figure A.4. all structures in this figure are $HCl(H_2O)_{11}$ cluster size 1.NoIP minimum energy structure b. lowest energy SSIP with a relative energy of 1.5 kcal/mol c. HCl minimum, 18.0 kcal/mol d. CIP minimum, 4.23 kcal/mol

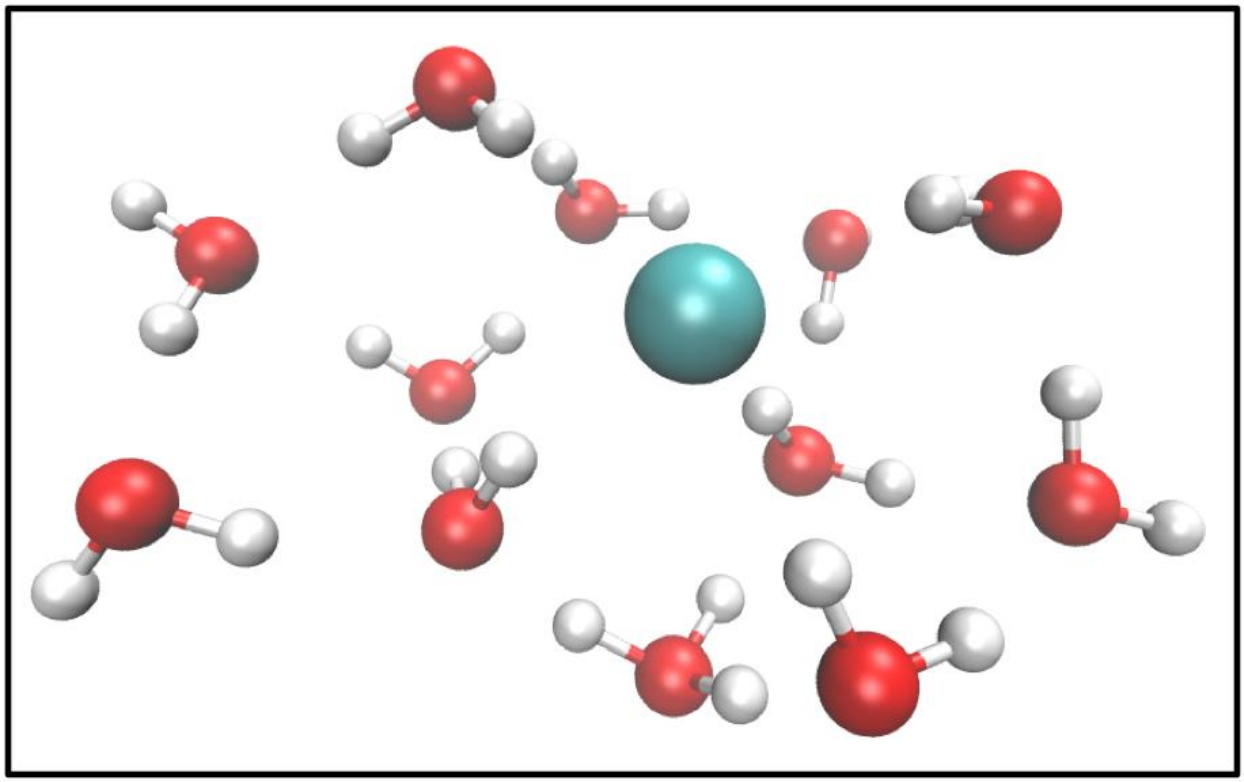

Figure A.5. $HCl(H_2O)_{12}$ SSIP configuration where all three hydrogens of the hydronium coordinate to a water that hydrogen bonds to the chloride anion. This structure has relative energy of 7.3 kcal/mol

A.3. Solvation Patterns of $HCl(H_2O)_{20-22}$

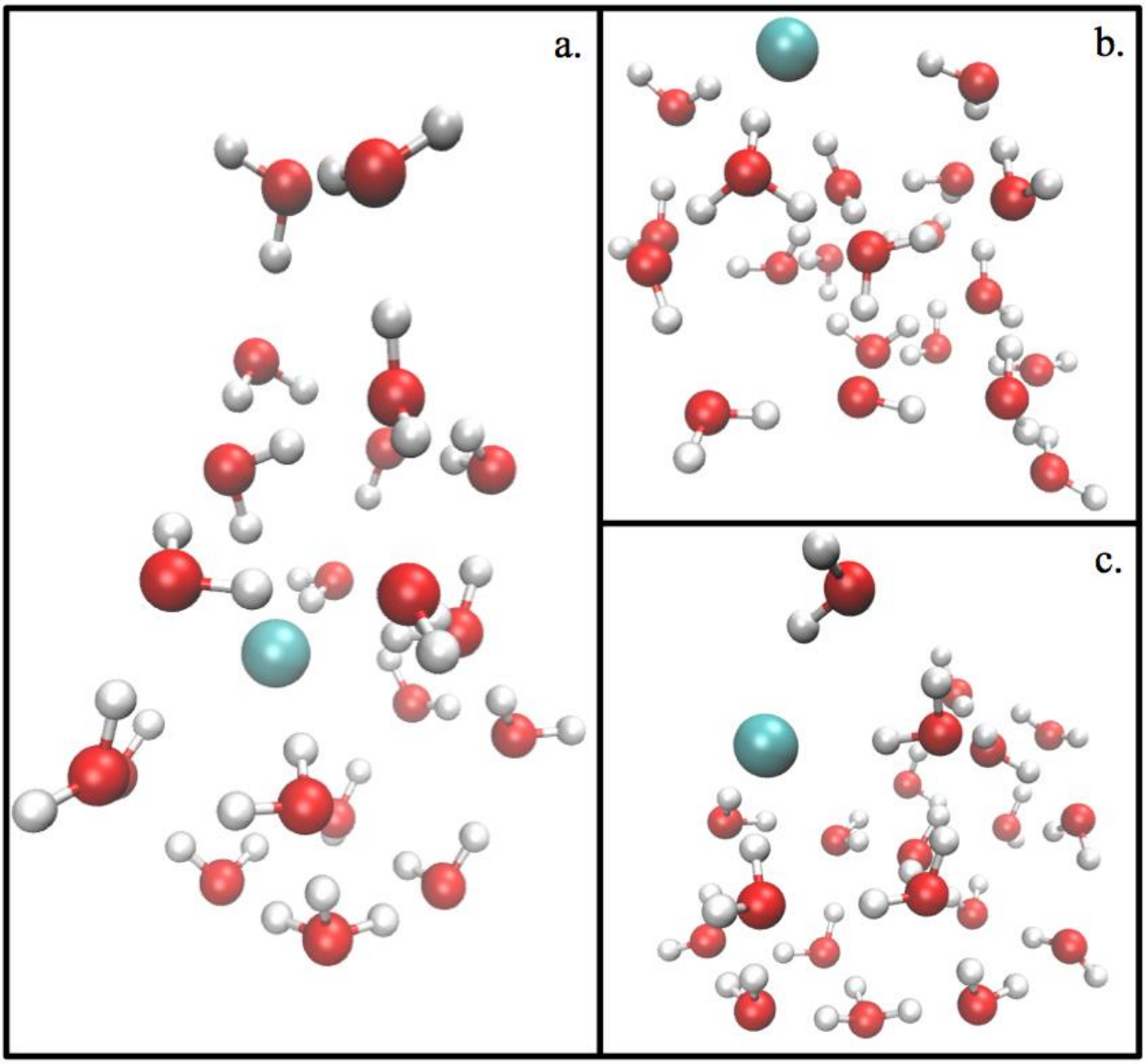

Figure A.6. Lowest energy structures of each solvation type found in clusters containing $HCl(H_2O)_{20}$ a. The SSIP structure is the minimum energy structure found at this cluster size. b. CIP with a relative energy of 4.4 kcal/mol c. NoIP with a relative energy of 4.5 kcal/mol

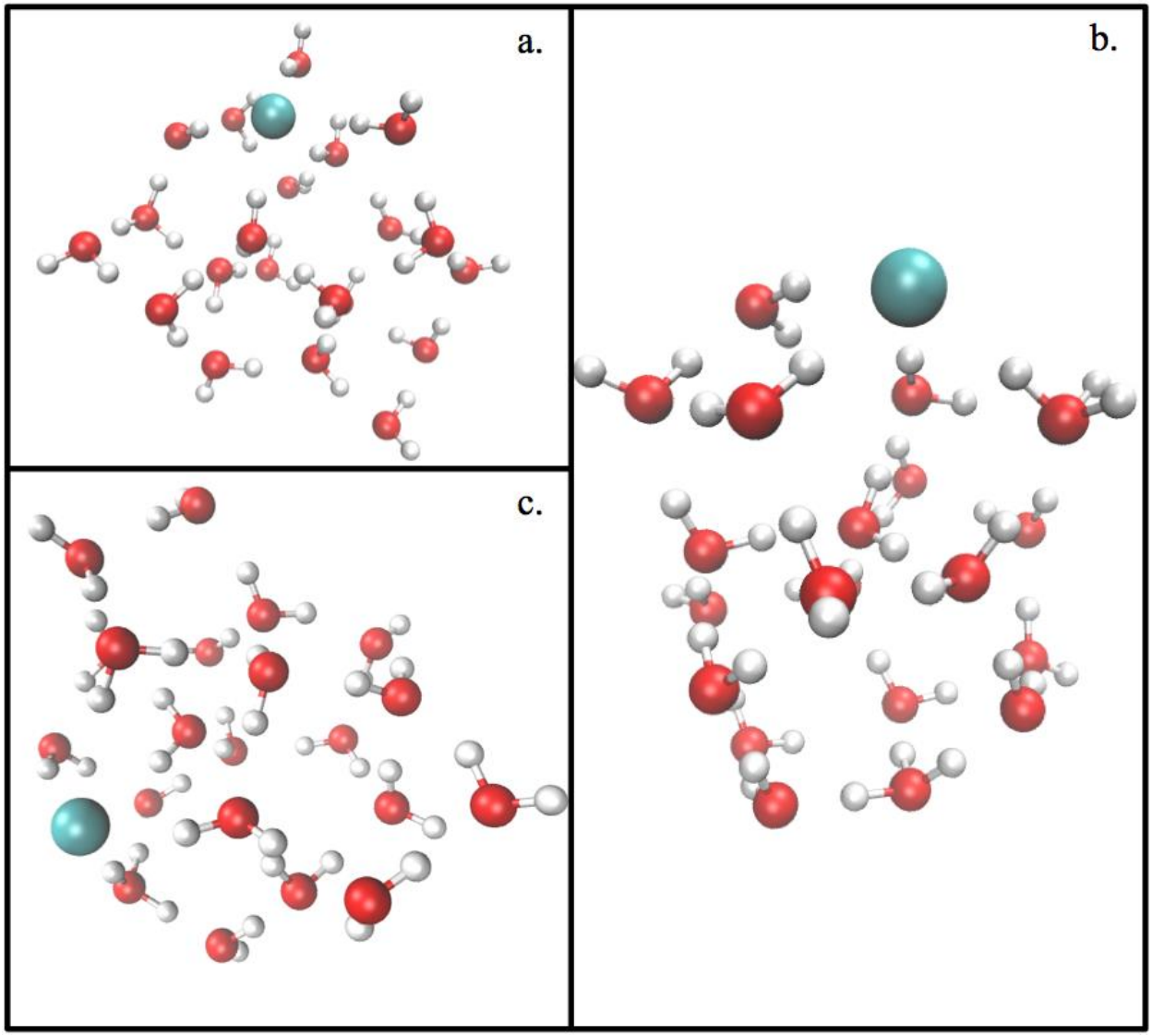

Figure A.7. Lowest energy structures of each solvation type found in clusters containing $HCl(H_2O)_{21}$ a. SSIP configuration with a relative energy of 6.1 kcal/mol b. lowest energy structure found at this cluster size and has no ion pairing c. CIP with relative energy of 12.0 kcal/mol

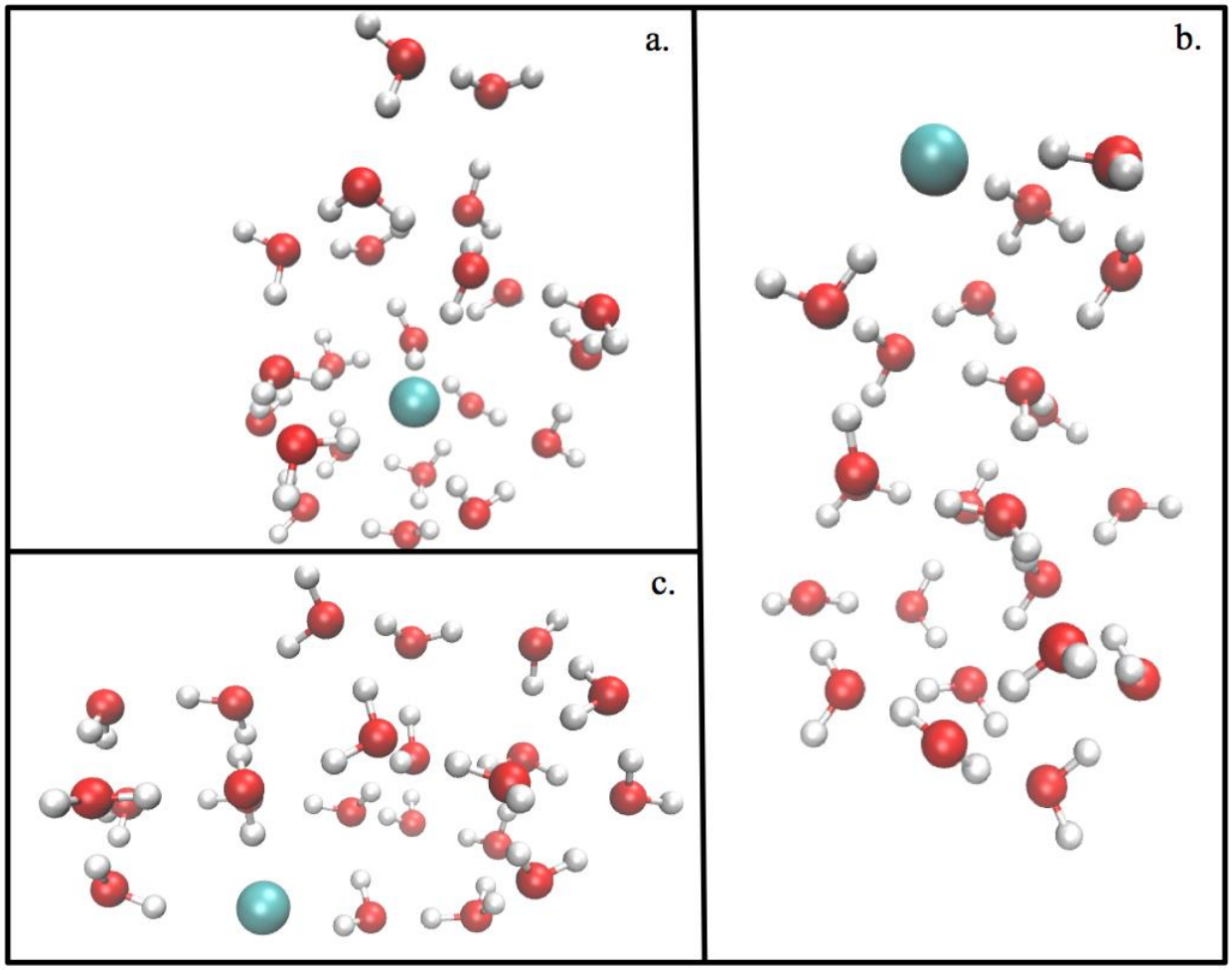

Figure A.8. Lowest energy structures of each solvation type found in clusters containing $HCl(H_2O)_{22}$ a. NoIP structure with a relative energy of 7.3 kcal/mol b. CIP configuration with a relative energy of 8.2 kcal/mol c. Minimum energy structure at this cluster size, with the SSIP solvation type

Appendix B. Generation of HCl-Water PES

The potential energy surface (PES) of Cl-H-OH_2 was calculated for use in the fitting of two body parameters. This work was done in the Gaussian 09 suite of programs.[134] Initially, the HCl-Water dimer was optimized with the MP2 basis set at the aug-cc-pvtz level of theory in order to generate the starting geometry. From this initial configuration, the oxygen to chloride distance was frozen, while the H-Cl distance was varied, allowing the hydrogens on the water molecule to relax at each state. The distance was modified by 0.05 angstroms, and ten steps along the potential energy surface were recorded.

Appendix C. Publication Agreements and Permissions

Copyright Clearance Center RightsLink®

Home | Create Account | Help

ACS Publications
Most Trusted. Most Cited. Most Read.

Title: Mechanistic Perspectives in the Regioselective Indole Addition to Unsymmetrical Silyloxyallyl Cations

Author: Caitlin G. Bresnahan, Kiara A. Taylor-Edinbyrd, Alexander H. Cleveland, et al

Publication: The Journal of Organic Chemistry

Publisher: American Chemical Society

Date: May 1, 2019

Copyright © 2019, American Chemical Society

LOGIN

If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials.

Already a RightsLink user or want to learn more?

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.

BACK | CLOSE WINDOW

Works Cited

1. Elsaesser, T., Introduction: Ultrafast Processes in Chemistry. *Chemical Reviews* **2017,** *117* (16), 10621-10622.

2. Ghosh, R.; Mora, A. K.; Nath, S., Disentangling Time Scales of Vibrational Cooling, Solvation, and Hydrogen Bond Reorganization Dynamics Using Ultrafast Transient Infrared Spectroscopy of Formylperylene. *The Journal of Physical Chemistry B* **2019,** *123* (20), 4408-4414.

3. Maeda, S.; Harabuchi, Y., On Benchmarking of Automated Methods for Performing Exhaustive Reaction Path Search. *Journal of Chemical Theory and Computation* **2019,** *15* (4), 2111-2115.

4. Debnath, J.; Invernizzi, M.; Parrinello, M., Enhanced Sampling of Transition States. *Journal of Chemical Theory and Computation* **2019,** *15* (4), 2454-2459.

5. Kumar, R.; Knight, C.; Wick, C. D.; Chen, B., Bringing Reactivity to the Aggregation-Volume-Bias Monte Carlo Based Simulation Framework: Water Nucleation Induced by a Reactive Proton. *The Journal of Physical Chemistry B* **2015,** *119* (29), 9068-9075.

6. Taylor, J.; Guo, H.; Wang, J., Ab initio modeling of quantum transport properties of molecular electronic devices. *Physical Review B* **2001,** *63* (24), 245407.

7. Kümmel, S., Charge-Transfer Excitations: A Challenge for Time-Dependent Density Functional Theory That Has Been Met. *Advanced Energy Materials* **2017,** *7* (16), 1700440.

8. Bresnahan, C. G.; Taylor-Edinbyrd, K. A.; Cleveland, A. H.; Malone, J. A.; Dange, N. S.; Milet, A.; Kumar, R.; Kartika, R., Mechanistic Perspectives in the Regioselective Indole Addition to Unsymmetrical Silyloxyallyl Cations. *The Journal of Organic Chemistry* **2019**.

9. Fuller, J.; Wilson, T. R.; Eberhart, M. E.; Alexandrova, A. N., Charge Density in Enzyme Active Site as a Descriptor of Electrostatic Preorganization. *Journal of Chemical Information and Modeling* **2019,** *59* (5), 2367-2373.

10. Kudin, K. N.; Ozbas, B.; Schniepp, H. C.; Prud'homme, R. K.; Aksay, I. A.; Car, R., Raman Spectra of Graphite Oxide and Functionalized Graphene Sheets. *Nano Letters* **2008,** *8* (1), 36-41.

11. Michalska, D.; Wysokiński, R., The prediction of Raman spectra of platinum(II) anticancer drugs by density functional theory. *Chemical Physics Letters* **2005,** *403* (1), 211-217.

12. Malik, M.; Wysokiński, R.; Zierkiewicz, W.; Helios, K.; Michalska, D., Raman and Infrared Spectroscopy, DFT Calculations, and Vibrational Assignment of the Anticancer Agent Picoplatin: Performance of Long-Range Corrected/Hybrid Functionals for a Platinum(II) Complex. *The Journal of Physical Chemistry A* **2014,** *118* (34), 6922-6934.

13. Flynn, S. D.; Skvortsov, D.; Morrison, A. M.; Liang, T.; Choi, M. Y.; Douberly, G. E.; Vilesov, A. F., Infrared Spectra of HCl−H2O Clusters in Helium Nanodroplets. *The Journal of Physical Chemistry Letters* **2010,** *1* (15), 2233-2238.

14. Biswas, R.; Carpenter, W.; Voth, G. A.; Tokmakoff, A., Molecular modeling and assignment of IR spectra of the hydrated excess proton in isotopically dilute water. *The Journal of Chemical Physics* **2016,** *145* (15), 154504.

15. Grimblat, N.; Sarotti, A. M., Computational Chemistry to the Rescue: Modern Toolboxes for the Assignment of Complex Molecules by GIAO NMR Calculations. *Chemistry – A European Journal* **2016,** *22* (35), 12246-12261.

16. Rychnovsky, S. D., Predicting NMR Spectra by Computational Methods: Structure Revision of Hexacyclinol. *Organic Letters* **2006,** *8* (13), 2895-2898.

17. Dreher, T.; Lemarchand, C.; Soulard, L.; Bourasseau, E.; Malfreyt, P.; Pineau, N., Calculation of a solid/liquid surface tension: A methodological study. *The Journal of Chemical Physics* **2018,** *148* (3), 034702.

18. Ghoufi, A.; Malfreyt, P.; Tildesley, D. J., Computer modelling of the surface tension of the gas–liquid and liquid–liquid interface. *Chemical Society Reviews* **2016,** *45* (5), 1387-1409.

19. Yoshida, H.; Inagaki, T.; Miura, K.; Inaba, M.; Ogumi, Z., Density functional theory calculation on the effect of local structure of doped ceria on ionic conductivity. *Solid State Ionics* **2003,** *160* (1), 109-116.

20. Gennarini, F.; David, R.; López, I.; Le Mest, Y.; Réglier, M.; Belle, C.; Thibon-Pourret, A.; Jamet, H.; Le Poul, N., Influence of Asymmetry on the Redox Properties of Phenoxo- and Hydroxo-Bridged Dicopper Complexes: Spectroelectrochemical and Theoretical Studies. *Inorganic Chemistry* **2017,** *56* (14), 7707-7719.

21. Amaro, R. E.; Baudry, J.; Chodera, J.; Demir, Ö.; McCammon, J. A.; Miao, Y.; Smith, J. C., Ensemble Docking in Drug Discovery. *Biophysical Journal* **2018,** *114* (10), 2271-2278.

22. Śledź, P.; Caflisch, A., Protein structure-based drug design: from docking to molecular dynamics. *Current Opinion in Structural Biology* **2018,** *48*, 93-102.

23. Alonso, H.; Bliznyuk, A. A.; Gready, J. E., Combining docking and molecular dynamic simulations in drug design. *Medicinal Research Reviews* **2006,** *26* (5), 531-568.

24. Lalaoui, N.; David, R.; Jamet, H.; Holzinger, M.; Le Goff, A.; Cosnier, S., Hosting Adamantane in the Substrate Pocket of Laccase: Direct Bioelectrocatalytic Reduction of O2 on Functionalized Carbon Nanotubes. *ACS Catalysis* **2016,** *6* (7), 4259-4264.

25. Cole, D. J.; Cabeza de Vaca, I.; Jorgensen, W. L., Computation of protein–ligand binding free energies using quantum mechanical bespoke force fields. *MedChemComm* **2019**.

26. Špačková, N. a.; Trošanová, Z.; Šebesta, F.; Jansen, S.; Burda, J. V.; Srb, P.; Zachrdla, M.; Žídek, L.; Kozelka, J., Protein environment affects the water–tryptophan binding mode. MD, QM/MM, and NMR studies of engrailed homeodomain mutants. *Physical Chemistry Chemical Physics* **2018,** *20* (18), 12664-12677.

27. Farah, K.; Müller-Plathe, F.; Böhm, M. C., Classical Reactive Molecular Dynamics Implementations: State of the Art. *ChemPhysChem* **2012,** *13* (5), 1127-1151.

28. Alder, B. J.; Wainwright, T. E., Studies in Molecular Dynamics. I. General Method. *The Journal of Chemical Physics* **1959,** *31* (2), 459-466.

29. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R., Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984,** *81* (8), 3684-3690.

30. van Gunsteren, W. F.; Berendsen, H. J. C., Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry. *Angewandte Chemie International Edition in English* **1990,** *29* (9), 992-1023.

31. Andersen, H. C., Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics* **1980,** *72* (4), 2384-2393.

32. Maginn, E. J.; Elliott, J. R., Historical Perspective and Current Outlook for Molecular Dynamics As a Chemical Engineering Tool. *Industrial & Engineering Chemistry Research* **2010,** *49* (7), 3059-3078.

33. Leontidis, E.; Forrest, B. M.; Widmann, A. H.; Suter, U. W., Monte Carlo algorithms for the atomistic simulation of condensed polymer phases. *Journal of the Chemical Society, Faraday Transactions* **1995,** *91* (16), 2355-2368.

34. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E., Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* **1953,** *21* (6), 1087-1092.

35. Theodorou, D. N., Progress and Outlook in Monte Carlo Simulations. *Industrial & Engineering Chemistry Research* **2010,** *49* (7), 3047-3058.

36. Valleau, J. P.; Cohen, L. K., Primitive model electrolytes. I. Grand canonical Monte Carlo computations. *The Journal of Chemical Physics* **1980,** *72* (11), 5935-5941.

37. Smith, G. R.; Bruce, A. D., A study of the multi-canonical Monte Carlo method. *Journal of Physics A: Mathematical and General* **1995,** *28* (23), 6623-6643.

38. Chen, B.; Siepmann, J. I.; Oh, K. J.; Klein, M. L., Aggregation-volume-bias Monte Carlo simulations of vapor-liquid nucleation barriers for Lennard-Jonesium. *The Journal of Chemical Physics* **2001,** *115* (23), 10903-10913.

39. Chen, B.; Kim, H.; Keasler, S. J.; Nellas, R. B., An Aggregation-Volume-Bias Monte Carlo Investigation on the Condensation of a Lennard-Jones Vapor below the Triple Point and Crystal Nucleation in Cluster Systems: An In-Depth Evaluation of the Classical Nucleation Theory. *The Journal of Physical Chemistry B* **2008,** *112* (13), 4067-4078.

40. Chen, B.; Siepmann, J. I., Improving the Efficiency of the Aggregation−Volume−Bias Monte Carlo Algorithm. *The Journal of Physical Chemistry B* **2001,** *105* (45), 11275-11282.

41. Torrie, G. M.; Valleau, J. P., Monte Carlo study of a phase-separating liquid mixture by umbrella sampling. *The Journal of Chemical Physics* **1977,** *66* (4), 1402-1408.

42. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P., Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *Journal of the American Chemical Society* **1985,** *107* (13), 3902-3909.

43. Stewart, J. J. P., Optimization of parameters for semiempirical methods I. Method. *Journal of Computational Chemistry* **1989,** *10* (2), 209-220.

44. Stewart, J. J. P., Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *Journal of Molecular Modeling* **2007,** *13* (12), 1173-1213.

45. Roothaan, C. C. J., New Developments in Molecular Orbital Theory. *Reviews of Modern Physics* **1951,** *23* (2), 69-89.

46. Møller, C.; Plesset, M. S., Note on an Approximation Treatment for Many-Electron Systems. *Physical Review* **1934,** *46* (7), 618-622.

47. Hohenberg, P.; Kohn, W., Inhomogeneous Electron Gas. *Physical Review* **1964,** *136* (3B), B864-B871.

48. Kohn, W.; Sham, L. J., Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review* **1965,** *140* (4A), A1133-A1138.

49. Andzelm, J.; Wimmer, E., Density functional Gaussian-type-orbital approach to molecular geometries, vibrations, and reaction energies. *The Journal of Chemical Physics* **1992,** *96* (2), 1280-1303.

50. Becke, A. D., Density-functional thermochemistry. I. The effect of the exchange-only gradient correction. *The Journal of Chemical Physics* **1992,** *96* (3), 2155-2160.

51. Nakata, H.; Bai, S., Development of a new parameter optimization scheme for a reactive force field based on a machine learning approach. *Journal of Computational Chemistry* **2019,** *0* (0).

52. Dittner, M.; Müller, J.; Aktulga, H. M.; Hartke, B., Efficient global optimization of reactive force-field parameters. *Journal of Computational Chemistry* **2015,** *36* (20), 1550-1561.

53. Miertuš, S.; Scrocco, E.; Tomasi, J., Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chemical Physics* **1981,** *55* (1), 117-129.

54. Miertuš, S.; Tomasi, J., Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes. *Chemical Physics* **1982,** *65* (2), 239-245.

55. Tomasi, J.; Mennucci, B.; Cammi, R., Quantum Mechanical Continuum Solvation Models. *Chemical Reviews* **2005,** *105* (8), 2999-3094.

56. Leonard, A. N.; Wang, E.; Monje-Galvan, V.; Klauda, J. B., Developing and Testing of Lipid Force Fields with Applications to Modeling Cellular Membranes. *Chemical Reviews* **2019,** *119* (9), 6227-6269.

57. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation* **2015,** *11* (8), 3696-3713.

58. Cheatham Iii, T. E.; Case, D. A., Twenty-five years of nucleic acid simulations. *Biopolymers* **2013,** *99* (12), 969-977.

59. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr, A. D., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry* **2010,** *31* (4), 671-690.

60. Mackerell Jr, A. D., Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry* **2004,** *25* (13), 1584-1604.

61. Soteras Gutiérrez, I.; Lin, F.-Y.; Vanommeslaeghe, K.; Lemkul, J. A.; Armacost, K. A.; Brooks, C. L.; MacKerell, A. D., Parametrization of halogen bonds in the CHARMM general force field: Improved treatment of ligand–protein interactions. *Bioorganic & Medicinal Chemistry* **2016,** *24* (20), 4812-4825.

62. Reif, M. M.; Hünenberger, P. H.; Oostenbrink, C., New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field. *Journal of Chemical Theory and Computation* **2012,** *8* (10), 3705-3723.

63. Schuler, L. D.; Daura, X.; van Gunsteren, W. F., An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *Journal of Computational Chemistry* **2001,** *22* (11), 1205-1218.

64. Soares, T. A.; Hünenberger, P. H.; Kastenholz, M. A.; Kräutler, V.; Lenz, T.; Lins, R. D.; Oostenbrink, C.; van Gunsteren, W. F., An improved nucleic acid parameter set for the GROMOS force field. *Journal of Computational Chemistry* **2005,** *26* (7), 725-737.

65. Jorgensen, W. L.; Tirado-Rives, J., The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society* **1988,** *110* (6), 1657-1666.

66. Doherty, B.; Zhong, X.; Gathiaka, S.; Li, B.; Acevedo, O., Revisiting OPLS Force Field Parameters for Ionic Liquid Simulations. *Journal of Chemical Theory and Computation* **2017,** *13* (12), 6131-6145.

67. Sambasivarao, S. V.; Acevedo, O., Development of OPLS-AA Force Field Parameters for 68 Unique Ionic Liquids. *Journal of Chemical Theory and Computation* **2009,** *5* (4), 1038-1050.

68. Pahari, P.; Chaturvedi, S., Determination of best-fit potential parameters for a reactive force field using a genetic algorithm. *Journal of Molecular Modeling* **2012,** *18* (3), 1049-1061.

69. David, R.; Jamet, H.; Nivière, V.; Moreau, Y.; Milet, A., Iron Hydroperoxide Intermediate in Superoxide Reductase: Protonation or Dissociation First? MM Dynamics and QM/MM Metadynamics Study. *Journal of Chemical Theory and Computation* **2017,** *13* (6), 2987-3004.

70. Zheng, M.; Waller, M. P., Adaptive quantum mechanics/molecular mechanics methods. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2016,** *6* (4), 369-385.

71. Brunk, E.; Rothlisberger, U., Mixed Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulations of Biological Systems in Ground and Electronically Excited States. *Chemical Reviews* **2015,** *115* (12), 6217-6263.

72. Pezeshki, S.; Lin, H., Adaptive-Partitioning QM/MM for Molecular Dynamics Simulations: 4. Proton Hopping in Bulk Water. *Journal of Chemical Theory and Computation* **2015,** *11* (6), 2398-2411.

73. Jiang, T.; Boereboom, J. M.; Michel, C.; Fleurat-Lessard, P.; Bulo, R. E., Proton Transfer in Aqueous Solution: Exploring the Boundaries of Adaptive QM/MM. In *Quantum Modeling of Complex Molecular Systems*, Rivail, J.-L.; Ruiz-Lopez, M.; Assfeld, X., Eds. Springer International Publishing: Cham, 2015; pp 51-91.

74. Arthur, N. L.; McDonell, J. A., BEBO Calculations. I. Activation Energies and Kinetic Isotope Effects for the Reactions of CH3 and CF3 Radicals with HCl and H2S. *The Journal of Chemical Physics* **1972,** *56* (6), 3100-3110.

75. Hur, J.; Stuart, S. J., Modified reactive empirical bond-order potential for heterogeneous bonding environments. *The Journal of Chemical Physics* **2012,** *137* (5), 054102.

76. van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A., ReaxFF: A Reactive Force Field for Hydrocarbons. *The Journal of Physical Chemistry A* **2001,** *105* (41), 9396-9409.

77. Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M.; Verstraelen, T.; Grama, A.; van Duin, A. C. T., The ReaxFF reactive force-field: development, applications and future directions. *Npj Computational Materials* **2016,** *2*, 15011.

78. Chenoweth, K.; van Duin, A. C. T.; Goddard, W. A., ReaxFF Reactive Force Field for Molecular Dynamics Simulations of Hydrocarbon Oxidation. *The Journal of Physical Chemistry A* **2008,** *112* (5), 1040-1053.

79. Nielson, K. D.; van Duin, A. C. T.; Oxgaard, J.; Deng, W.-Q.; Goddard, W. A., Development of the ReaxFF Reactive Force Field for Describing Transition Metal Catalyzed Reactions, with Application to the Initial Stages of the Catalytic Formation of Carbon Nanotubes. *The Journal of Physical Chemistry A* **2005,** *109* (3), 493-499.

80. Danielsson, J.; Meuwly, M., Atomistic Simulation of Adiabatic Reactive Processes Based on Multi-State Potential Energy Surfaces. *Journal of Chemical Theory and Computation* **2008,** *4* (7), 1083-1093.

81. Brancato, G.; Tuckerman, M. E., A polarizable multistate empirical valence bond model for proton transport in aqueous solution. *The Journal of Chemical Physics* **2005,** *122* (22), 224507.

82. Warshel, A.; Weiss, R. M., An empirical valence bond approach for comparing reactions in solutions and in enzymes. *Journal of the American Chemical Society* **1980,** *102* (20), 6218-6226.

83. Mei, H. S.; Tuckerman, M. E.; Sagnella, D. E.; Klein, M. L., Quantum Nuclear ab Initio Molecular Dynamics Study of Water Wires. *The Journal of Physical Chemistry B* **1998,** *102* (50), 10446-10458.

84. Sonnenberg, J. L.; Schlegel, H. B., Empirical valence bond models for reactive potential energy surfaces. II. Intramolecular proton transfer in pyridone and the Claisen reaction of allyl vinyl ether. *Molecular Physics* **2007,** *105* (19-22), 2719-2729.

85. Schlegel, H. B.; Sonnenberg, J. L., Empirical Valence-Bond Models for Reactive Potential Energy Surfaces Using Distributed Gaussians. *Journal of Chemical Theory and Computation* **2006,** *2* (4), 905-911.

86. Hornung, B.; Harvey, J. N.; Preston, T. J.; Dunning, G. T.; Orr-Ewing, A. J., Empirical Valence Bond Theory Studies of the CH4 + Cl → CH3 + HCl Reaction. *The Journal of Physical Chemistry A* **2015,** *119* (37), 9590-9598.

87. Hartke, B.; Grimme, S., Reactive force fields made simple. *Physical Chemistry Chemical Physics* **2015,** *17* (26), 16715-16718.

88. Kim, Y.; Corchado, J. C.; Villà, J.; Xing, J.; Truhlar, D. G., Multiconfiguration molecular mechanics algorithm for potential energy surfaces of chemical reactions. *The Journal of Chemical Physics* **2000,** *112* (6), 2718-2735.

89. Wang, F.; Izvekov, S.; Voth, G. A., Unusual "Amphiphilic" Association of Hydrated Protons in Strong Acid Solution. *Journal of the American Chemical Society* **2008,** *130* (10), 3120-3126.

90. Taraphder, S.; Maupin, C. M.; Swanson, J. M. J.; Voth, G. A., Coupling Protein Dynamics with Proton Transport in Human Carbonic Anhydrase II. *The Journal of Physical Chemistry B* **2016,** *120* (33), 8389-8404.

91. Biswas, R.; Tse, Y.-L. S.; Tokmakoff, A.; Voth, G. A., Role of Presolvation and Anharmonicity in Aqueous Phase Hydrated Proton Solvation and Transport. *The Journal of Physical Chemistry B* **2016,** *120* (8), 1793-1804.

92. Day, T. J. F.; Soudackov, A. V.; Čuma, M.; Schmitt, U. W.; Voth, G. A., A second generation multistate empirical valence bond model for proton transport in aqueous systems. *The Journal of Chemical Physics* **2002,** *117* (12), 5839-5849.

93. Swanson, J. M. J.; Maupin, C. M.; Chen, H.; Petersen, M. K.; Xu, J.; Wu, Y.; Voth, G. A., Proton Solvation and Transport in Aqueous and Biomolecular Systems: Insights from Computer Simulations. *The Journal of Physical Chemistry B* **2007,** *111* (17), 4300-4314.

94. Schmitt, U. W.; Voth, G. A., Multistate Empirical Valence Bond Model for Proton Transport in Water. *The Journal of Physical Chemistry B* **1998,** *102* (29), 5547-5551.

95. Voth, G. A., Computer Simulation of Proton Solvation and Transport in Aqueous and Biomolecular Systems. *Accounts of Chemical Research* **2006,** *39* (2), 143-150.

96. Wu, Y.; Chen, H.; Wang, F.; Paesani, F.; Voth, G. A., An Improved Multistate Empirical Valence Bond Model for Aqueous Proton Solvation and Transport. *The Journal of Physical Chemistry B* **2008,** *112* (2), 467-482.

97. Knight, C.; Voth, G. A., The Curious Case of the Hydrated Proton. *Accounts of Chemical Research* **2012,** *45* (1), 101-109.

98. Lee, S.; Liang, R.; Voth, G. A.; Swanson, J. M. J., Computationally Efficient Multiscale Reactive Molecular Dynamics to Describe Amino Acid Deprotonation in Proteins. *Journal of Chemical Theory and Computation* **2016,** *12* (2), 879-891.

99. Savage, J.; Voth, G. A., Proton Solvation and Transport in Realistic Proton Exchange Membrane Morphologies. *The Journal of Physical Chemistry C* **2016,** *120* (6), 3176-3186.

100. Iuchi, S.; Chen, H.; Paesani, F.; Voth, G. A., Hydrated Excess Proton at Water−Hydrophobic Interfaces. *The Journal of Physical Chemistry B* **2009,** *113* (13), 4017-4030.

101. Wang, F.; Voth, G. A., A linear-scaling self-consistent generalization of the multistate empirical valence bond method for multiple excess protons in aqueous systems. *The Journal of Chemical Physics* **2005,** *122* (14), 144105.

102. Kumar, R.; Christie, R. A.; Jordan, K. D., A Modified MSEVB Force Field for Protonated Water Clusters. *The Journal of Physical Chemistry B* **2009,** *113* (13), 4111-4118.

103. Wick, C. D., HCl Accommodation, Dissociation, and Propensity for the Surface of Water. *The Journal of Physical Chemistry A* **2013,** *117* (47), 12459-12467.

104. Wick, C. D., Hydronium Behavior at the Air–Water Interface with a Polarizable Multistate Empirical Valence Bond Model. *The Journal of Physical Chemistry C* **2012,** *116* (6), 4026-4038.

105. Wick, C. D., Comparing hydroxide and hydronium at the instantaneous air-water interface using polarizable multi-state empirical valence bond models. *Computational and Theoretical Chemistry* **2017,** *1116*, 64-72.

106. Buch, V.; Sadlej, J.; Aytemiz-Uras, N.; Devlin, J. P., Solvation and Ionization Stages of HCl on Ice Nanocrystals. *The Journal of Physical Chemistry A* **2002,** *106* (41), 9374-9389.

107. Bacelo, D. E.; Binning, R. C.; Ishikawa, Y., Ab Initio Monte Carlo Simulated Annealing Study of HCl(H2O)n (n = 3, 4) Clusters. *The Journal of Physical Chemistry A* **1999,** *103* (24), 4631-4640.

108. Svanberg, M.; Pettersson, J. B. C.; Bolton, K., Coupled QM/MM Molecular Dynamics Simulations of HCl Interacting with Ice Surfaces and Water Clusters − Evidence of Rapid Ionization. *The Journal of Physical Chemistry A* **2000,** *104* (24), 5787-5798.

109. Sobolewski, A. L.; Domcke, W., Photochemistry of HCl(H2O)4: Cluster Model of the Photodetachment of the Chloride Anion in Water. *The Journal of Physical Chemistry A* **2003,** *107* (10), 1557-1562.

110. Molina, M. J.; Tso, T.-L.; Molina, L. T.; Wang, F. C. Y., Antarctic Stratospheric Chemistry of Chlorine Nitrate, Hydrogen Chloride, and Ice: Release of Active Chlorine. *Science* **1987,** *238* (4831), 1253.

111. Tolbert, M. A.; Rossi, M. J.; Malhotra, R.; Golden, D. M., Reaction of Chlorine Nitrate with Hydrogen Chloride and Water at Antarctic Stratospheric Temperatures. *Science* **1987,** *238* (4831), 1258.

112. Zuraski, K.; Kwasniewski, D.; Samanta, A. K.; Reisler, H., Vibrational Predissociation of the HCl–(H2O)3 Tetramer. *The Journal of Physical Chemistry Letters* **2016,** *7* (21), 4243-4247.

113. Chang, C.-T.; Liu, T.-H.; Jeng, F.-T., Atmospheric concentrations of the Cl atom, ClO radical, and HO radical in the coastal marine boundary layer. *Environmental Research* **2004,** *94* (1), 67-74.

114. Kumar, M.; Sinha, A.; Francisco, J. S., Role of Double Hydrogen Atom Transfer Reactions in Atmospheric Chemistry. *Accounts of Chemical Research* **2016,** *49* (5), 877-883.

115. Pöschl, U., Atmospheric Aerosols: Composition, Transformation, Climate and Health Effects. *Angewandte Chemie International Edition* **2005,** *44* (46), 7520-7540.

116. Jungwirth, P.; Tobias, D. J., Specific Ion Effects at the Air/Water Interface. *Chemical Reviews* **2006,** *106* (4), 1259-1281.

117. Ndongmouo, U. F. T.; Lee, M. S.; Rousseau, R.; Baletto, F.; Scandolo, S., Finite-Temperature Effects on the Stability and Infrared Spectra of HCl(H2O)6 Clusters. *The Journal of Physical Chemistry A* **2007,** *111* (49), 12810-12815.

118. Radola, B.; Martin-Gondre, L.; Picaud, S.; Rayez, M.-T.; Rayez, J.-C., Hydrogen chloride adsorption on large defective PAHs modeling soot surfaces and influence on water trapping: A DFT and AIMD study. *Chemical Physics* **2019,** *523*, 18-27.

119. Hammerich, A. D.; Finlayson-Pitts, B. J.; Gerber, R. B., NOx Reactions on Aqueous Surfaces with Gaseous HCl: Formation of a Potential Precursor to Atmospheric Cl Atoms. *The Journal of Physical Chemistry Letters* **2012,** *3* (23), 3405-3410.

120. Ouf, F. X.; Mocho, V. M.; Pontreau, S.; Wang, Z.; Ferry, D.; Yon, J., Physicochemical properties of aerosol released in the case of a fire involving materials used in the nuclear industry. *Journal of Hazardous Materials* **2015,** *283*, 340-349.

121. Chaban, G. M.; Gerber, R. B.; Janda, K. C., Transition from Hydrogen Bonding to Ionization in (HCl)n(NH3)n and (HCl)n(H2O)n Clusters: Consequences for Anharmonic Vibrational Spectroscopy. *The Journal of Physical Chemistry A* **2001,** *105* (36), 8323-8332.

122. Botti, A.; Bruni, F.; Imberti, S.; Ricci, M. A.; Soper, A. K., Ions in water: The microscopic structure of a concentrated HCl solution. *The Journal of Chemical Physics* **2004,** *121* (16), 7840-7848.

123. Packer, M. J.; Clary, D. C., Interaction of HCl with water clusters: (H2O)nHCl, n = 1-3. *The Journal of Physical Chemistry* **1995,** *99* (39), 14323-14333.

124. Walewski, Ł.; Forbert, H.; Marx, D., Quantum Induced Bond Centering in Microsolvated HCl: Solvent Separated versus Contact Ion Pairs. *The Journal of Physical Chemistry Letters* **2011,** *2* (24), 3069-3074.

125. Forbert, H.; Masia, M.; Kaczmarek-Kedziera, A.; Nair, N. N.; Marx, D., Aggregation-Induced Chemical Reactions: Acid Dissociation in Growing Water Clusters. *Journal of the American Chemical Society* **2011,** *133* (11), 4062-4072.

126. Hassanali, A. A.; Cuny, J.; Ceriotti, M.; Pickard, C. J.; Parrinello, M., The Fuzzy Quantum Proton in the Hydrogen Chloride Hydrates. *Journal of the American Chemical Society* **2012,** *134* (20), 8557-8569.

127. Milet, A.; Struniewicz, C.; Moszynski, R.; Wormer, P. E. S., Theoretical study of the protolytic dissociation of HCl in water clusters. *The Journal of Chemical Physics* **2001,** *115* (1), 349-356.

128. Ando, K.; Hynes, J. T., Molecular Mechanism of HCl Acid Ionization in Water: Ab Initio Potential Energy Surfaces and Monte Carlo Simulations. *The Journal of Physical Chemistry B* **1997,** *101* (49), 10464-10478.

129. Fulton, J. L.; Balasubramanian, M., Structure of Hydronium (H3O+)/Chloride (Cl−) Contact Ion Pairs in Aqueous Hydrochloric Acid Solution: A Zundel-like Local Configuration. *Journal of the American Chemical Society* **2010,** *132* (36), 12597-12604.

130. Baer, M. D.; Fulton, J. L.; Balasubramanian, M.; Schenter, G. K.; Mundy, C. J., Persistent Ion Pairing in Aqueous Hydrochloric Acid. *The Journal of Physical Chemistry B* **2014,** *118* (26), 7211-7220.

131. Chialvo, A. A.; Cummings, P. T.; Simonson, J. M., H3O+/Cl− ion-pair formation in high-temperature aqueous solutions. *The Journal of Chemical Physics* **2000,** *113* (18), 8093-8100.

132. Cukierman, S., Et tu, Grotthuss! and other unfinished stories. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2006,** *1757* (8), 876-885.

133. Agmon, N., The Grotthuss mechanism. *Chemical Physics Letters* **1995,** *244* (5), 456-462.

134. 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, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian 09 (Gaussian, Inc., Wallingford CT, 2009).

135. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics* **2010,** *132* (15), 154104.

136. Grimme, S.; Ehrlich, S.; Goerigk, L., Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **2011,** *32* (7), 1456-1465.

137. Barducci, A.; Bonomi, M.; Parrinello, M., Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011,** *1* (5), 826-843.

138. Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J., cp2k: atomistic simulations of condensed matter systems. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2014,** *4* (1), 15-25.

139. VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J., Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Computer Physics Communications* **2005,** *167* (2), 103-128.

140. VandeVondele, J.; Hutter, J., Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *The Journal of Chemical Physics* **2007,** *127* (11), 114105.

141. Krack, M., Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functionals. *Theoretical Chemistry Accounts* **2005,** *114* (1), 145-152.

142. Genovese, L.; Deutsch, T.; Goedecker, S., Efficient and accurate three-dimensional Poisson solver for surface problems. *The Journal of Chemical Physics* **2007,** *127* (5), 054704.

143. Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R., RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chemical Physics Letters* **1998,** *294* (1), 143-152.

144. Herr, J. D.; Talbot, J.; Steele, R. P., Structural Progression in Clusters of Ionized Water, (H2O)n=1–5+. *The Journal of Physical Chemistry A* **2015,** *119* (4), 752-766.

145. Neese, F., The ORCA program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012,** *2* (1), 73-78.

146. Sharp, M. E.; Vázquez, F. X.; Wagner, J. W.; Dannenhoffer-Lafage, T.; Voth, G. A., Multiconfigurational Coarse-Grained Molecular Dynamics. *Journal of Chemical Theory and Computation* **2019,** *15* (5), 3306-3315.

147. Wu, Y.; Tepper, H. L.; Voth, G. A., Flexible simple point-charge water model with improved liquid-state properties. *The Journal of Chemical Physics* **2006,** *124* (2), 024503.

148. Grant, T. N.; Rieder, C. J.; West, F. G., Interrupting the Nazarov reaction: domino and cascade processes utilizing cyclopentenyl cations. *Chemical Communications* **2009,** (38), 5676-5688.

149. Shimada, N.; Stewart, C.; Tius, M. A., Asymmetric Nazarov Cyclizations. *Tetrahedron* **2011,** *67* (33), 5851-5870.

150. Vaidya, T.; Eisenberg, R.; Frontier, A. J., Catalytic Nazarov Cyclization: The State of the Art. *ChemCatChem* **2011,** *3* (10), 1531-1548.

151. Harmata, M., The (4+3)-cycloaddition reaction: heteroatom-substituted allylic cations as dienophiles. *Chemical Communications* **2010,** *46* (47), 8904-8922.

152. Lohse, A. G.; Hsung, R. P., (4+3) Cycloaddition Reactions of Nitrogen-Stabilized Oxyallyl Cations. *Chemistry – A European Journal* **2011,** *17* (14), 3812-3822.

153. Battiste, M. A.; Pelphrey, P. M.; Wright, D. L., The Cycloaddition Strategy for the Synthesis of Natural Products Containing Carbocyclic Seven-Membered Rings. *Chemistry – A European Journal* **2006,** *12* (13), 3438-3447.

154. Tang, Q.; Chen, X.; Tiwari, B.; Chi, Y. R., Addition of Indoles to Oxyallyl Cations for Facile Access to α-Indole Carbonyl Compounds. *Organic Letters* **2012,** *14* (7), 1922-1925.

155. Vander Wal, M. N.; Dilger, A. K.; MacMillan, D. W. C., Development of a generic activation mode: nucleophilic α-substitution of ketones via oxy-allyl cations. *Chemical Science* **2013,** *4* (8), 3075-3079.

156. Ayala, C. E.; Dange, N. S.; Fronczek, F. R.; Kartika, R., Brønsted Acid Catalyzed α′-Functionalization of Silylenol Ethers with Indoles. *Angewandte Chemie International Edition* **2015,** *54* (15), 4641-4645.

157. Malone, J. A.; Cleveland, A. H.; Fronczek, F. R.; Kartika, R., Effects of Solvent and Residual Water on Enhancing the Reactivity of Six-Membered Silyloxyallyl Cations toward Nucleophilic Addition. *Organic Letters* **2016,** *18* (17), 4408-4411.

158. Dange, N. S.; Stepherson, J. R.; Ayala, C. E.; Fronczek, F. R.; Kartika, R., Cooperative benzylic–oxyallylic stabilized cations: regioselective construction of α-quaternary centers in ketone-derived compounds. *Chemical Science* **2015,** *6* (11), 6312-6319.

159. Li, X.; Frisch, M. J., Energy-Represented Direct Inversion in the Iterative Subspace within a Hybrid Geometry Optimization Method. *Journal of Chemical Theory and Computation* **2006,** *2* (3), 835-839.

160. Foster, J. P.; Weinhold, F., Natural hybrid orbitals. *Journal of the American Chemical Society* **1980,** *102* (24), 7211-7218.

161. Parrish, R. M.; Burns, L. A.; Smith, D. G. A.; Simmonett, A. C.; DePrince, A. E.; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M.; Gonthier, J. F.; James, A. M.; McAlexander, H. R.; Kumar, A.; Saitow, M.; Wang, X.; Pritchard, B. P.; Verma, P.; Schaefer, H. F.; Patkowski, K.; King, R. A.; Valeev, E. F.; Evangelista, F. A.; Turney, J. M.; Crawford, T. D.; Sherrill, C. D., Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *Journal of Chemical Theory and Computation* **2017,** *13* (7), 3185-3197.

162. Hohenstein, E. G.; Sherrill, C. D., Density fitting and Cholesky decomposition approximations in symmetry-adapted perturbation theory: Implementation and application to

probe the nature of π-π interactions in linear acenes. *The Journal of Chemical Physics* **2010,** *132* (18), 184111.

163. *Essentials of Glycobiology*, 2nd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor NY, 2009.

164. Codée, J. D. C.; Litjens, R. E. J. N.; van den Bos, L. J.; Overkleeft, H. S.; van der Marel, G. A., Thioglycosides in sequential glycosylation strategies. *Chemical Society Reviews* **2005,** *34* (9), 769-782.

165. Shiao, T. C.; Roy, R., "Active–Latent" Thioglycosyl Donors and Acceptors in Oligosaccharide Syntheses. In *Reactivity Tuning in Oligosaccharide Assembly*, Fraser-Reid, B.; Cristóbal López, J., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; pp 69-108.

166. Koeller, K. M.; Wong, C.-H., Synthesis of Complex Carbohydrates and Glycoconjugates: Enzyme-Based and Programmable One-Pot Strategies. *Chemical Reviews* **2000,** *100* (12), 4465-4494.

167. Spell, M. L.; Deveaux, K.; Bresnahan, C. G.; Bernard, B. L.; Sheffield, W.; Kumar, R.; Ragains, J. R., A Visible-Light-Promoted O-Glycosylation with a Thioglycoside Donor. *Angewandte Chemie International Edition* **2016,** *55* (22), 6515-6519.

168. Spell, M. L.; Deveaux, K.; Bresnahan, C. G.; Ragains, J. R., O-Glycosylation Enabled by Remote Activation. *Synlett* **2017,** *28* (07), 751-761.

169. Grimme, S., Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry* **2006,** *27* (15), 1787-1799.

170. Bader, R. F. W., Atoms in molecules. *Accounts of Chemical Research* **1985,** *18* (1), 9-15.

171. AIMAll (Version 15.05.18), Todd A. Keith, TK Gristmill Software, Overland Park KS, USA, 2015 (aim.tkgristmill.com).

172. Bauernschmitt, R.; Ahlrichs, R., Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chemical Physics Letters* **1996,** *256* (4), 454-464.

Vita

Caitlin G. Bresnahan majored in chemistry at the University of Wisconsin-Eau Claire, receiving her B.S. in 2014. While at Eau Claire, Caitlin participated in computational research under the guidance of Dr. Sudeep Bhattacharyay, investigating π-stacking interactions of various aromatics to lumiflavin. She is pursuing a Ph.D. in physical chemistry at Louisiana State University under the guidance of Dr. Revati Kumar. During her time at Louisiana State University, Caitlin was awarded a Paula Barbra Scholarship to attend a TSRC meeting in Telluride, Co. Additionally, in 2018, Caitlin was awarded a Chateaubriand Fellowship of the Office for Science & Technology of France in the United States to perform research under the guidance of professor Anne Milet at the University of Grenoble-Alpes, in Grenoble France, for five months. Her research has focused on a broad range of topics from detailed investigations at the molecular level of organic systems in collaboration with experimentalists, to reactive force field development for HCl-water clusters.