Atomistic Simulations of Molecule Formation and Decomposition: From Astrophysics to Atmospheric Chemistry

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel
von
Meenu Upadhyay

Basel, 2024

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Erstbetreuer: Prof. Dr. Markus Meuwly

Zweitbetreuer: Prof. Dr. O. Anatole von Lilienfeld

Externer Experte: Prof. Dr. Andrew J. Orr-Ewing FRS MAE FRSC

Basel, den 25. Juni 2024

Prof. Dr. Marcel Mayor

Dekan

Abstract

Atomistic simulations, with their increasing accuracy and predictive capabilities, have become indispensable tools for understanding molecular behavior at the atomic level. Employing reactive molecular dynamic (MD) simulations with accurate interaction potentials offers valuable insights into both atmospheric and interstellar chemistry. This thesis explores the reaction dynamics of small molecules in gas and condensed phases under controlled conditions.

In the gas phase, photoinduced dynamics of *syn*-acetaldehyde oxide Criegee intermediate leading to the formation of vinoxy and hydroxyl (OH) radicals, as well as glycolaldehyde is explored. Molecular dynamics simulations are performed, utilizing classical force fields and neural network based potential energy surfaces at the MP2 and CASPT2 levels of theory. By comparing the computed final OH translational and rotational state distributions with experimental data, the O–O bond dissociation energy (D_e^{OO}) of vinyl hydroperoxide is estimated to be between 22 and 25 kcal/mol. The computed OH elimination rates are consistent with experiments and show limited sensitivity to D_e^{OO} due to the non-equilibrium nature of the process. Additionally, a new OH roaming pathway is observed following O–O scission, leading to glycolaldehyde formation on the picosecond time scale. This, along with recent work on methyl-ethyl-substituted Criegee intermediate, highlights the significance of OH roaming pathways, which should be incorporated into molecular-level atmospheric models.

In the condensed phase, atom-diatom collisions "in" and "on" amorphous solid water (ASW) are examined under conditions typical in cold molecular clouds. Reactive MD simulations, combined with machine-learned energy functions and fluctuating charge models, are employed to characterize the ground-state formation of CO_2 and NO_2 from O + CO and O + NO collisions. The simulations reveal that surface diffusion allows the reactants to recombine on the nanosecond timescale, even with initial separations up to 20 Å at 50 K. After recombination, the newly formed internally excited molecules relax by transferring energy to internal water modes and surface phonon modes of ASW on the picosecond timescale. Furthermore, surface diffusion

and desorption energetics are explored through extended and non-equilibrium MD simulations and are compared with literature data.

By combining computational results with experimental observables, this thesis enhances the understanding of molecular processes in diverse environments.
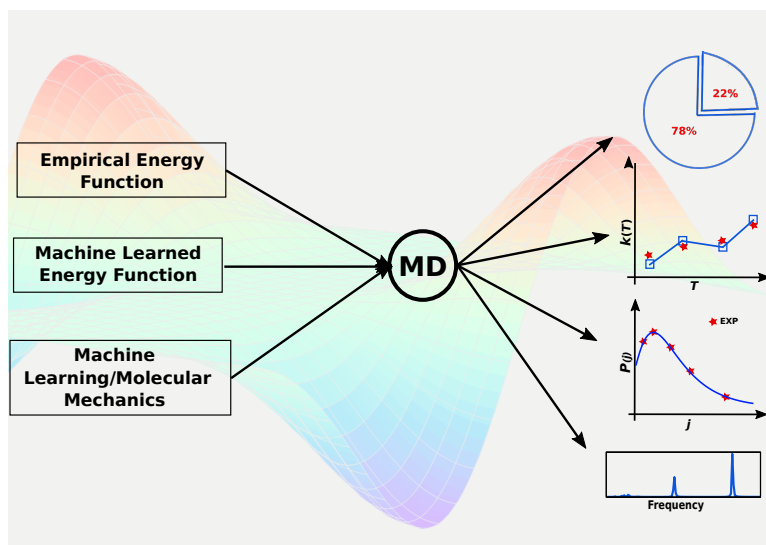
Acknowledgement

First and foremost, I would like to express my sincere gratitude to Prof. Dr. Markus Meuwly for giving me the opportunity to pursue my doctoral studies in his research group. I appreciate his unwavering support, guidance and mentorship throughout my PhD journey.

I am deeply grateful to Prof. Dr. O. Anatole von Lilienfeld for graciously agreeing to serve as my second supervisor, and to Prof. Dr. Andrew J. Orr-Ewing for kindly accepting the role of my external examiner.

Furthermore, I extend my thanks to our collaborators Prof. Dr. Edvardas Narevicius and Prof. Dr. Christiane Koch for their insightful contributions and enriching scientific exchanges. Additionally, I express gratitude to Prof. Dr. Joel Bowman and Prof. Dr. Marsha I. Lester for their insightful discussions and guidance.

Thanks to all past and present group members without whom the PhD experience would lose its flavor: Dr. Michael Devereux, Dr. Juan Carlos, Prof. Dr. Debasish Koner, Dr. Sarbani Patra, Dr. Sebastian Brickel, Dr. Sugata Goswami, Dr. Marco Pezzella, Dr. Maryam Salehi, Dr. Taylan Turan, Dr. Kai Töpfer, Dr. Silvan Käser, Eric D. Boittier, Dr. Luis Itzá Vázquez Salazar and Kaisheng Song. Special thanks to Dr. Kai Töpfer and Eric D. Boittier for proofreading my thesis and providing helpful suggestions.

I am also thankful to Sourabh for his encouragement and unwavering support. Additionally, I extend my heartfelt gratitude to Mahesh and Mohini for their support and strength.

Finally, I would like to thank my wonderful parents and siblings, Rashmi and Neeraj, for their enduring love, support and encouragement. Their constant support has been invaluable, providing me with the emotional and moral strength needed to persevere.

Contents

Abbreviations

AA	**A**cetaldehyde
AIMD	**A**b **I**nitio **M**olecular **D**ynamics
AO	**A**tomic **O**rbital
ASE	**A**tomic **S**imulation **E**nvironment
ASW	**A**morphous **S**olid **W**ater
BO	**B**orn-**O**ppenheimer
CI	**C**riegee **I**ntermediate
DZ	**D**ouble **Z**eta
FF	**F**orce **F**ield
FPC	**F**luctuating **P**oint **C**harge
FWHM	**F**ull **W**idth at **H**alf **M**aximum
GAPO	**GA**ussian and **PO**lynomial
GA	**G**lycolaldehyde
HDNN	**H**igh-**D**imensional **N**eural **N**etwork
HT	**H**ydrogen **T**ransfer
IRC	**I**ntrinsic **R**eaction **C**oordinate
ISM	**I**nterstellar **M**edium
IVR	**I**nternal **V**ibrational energy **R**edistribution
KDE	**K**ernel **D**ensity **E**stimation
KIE	**K**inetic **I**sotope **E**ffect
KKY	**K**umagai, **K**awamura, **Y**okokawa
LJ	**L**ennard-**J**ones
MAE	**M**ean **A**bsolute **E**rror

MD	Molecular Dynamics
MDCM	Minimal Distributed Charge Model
MEP	Minimum Energy Path
ML	Machine Learning
MO	Molecular Orbital
MS-ARMD	Multi-Surface Adiabatic Reactive Molecular Dynamics
MSE	Mean Squared Error
NN	Neural Network
NBO	Natural Bond Orbital
OH	Hydroxyl
PES	Potential Energy Surface
PIP	Permutationally Invariant Polynomial
POZ	Primary OZonide
QCT	QuasiClassical Trajectory
QM	Quantum Mechanics
QZ	Quadruple Zeta
RKHS	Reproducing Kernel Hilbert Space
RMSE	Root Mean Square Error
RRKM	Rice-Ramsperger-Kassel-Marcus
SCF	Self-Consistent Field
SE	Schrödinger Equation
SOA	Secondary Organic Aerosol
TKER	Total Kinetic Energy Release
TL	Transfer Learning
TZ	Triple Zeta
vDOS	vibrational Density of States
VA	Vinylalcohol
VHP	Vinyl HydroPeroxide
VOC	Volatile Organic Compound
ZPVE	Zero-Point Vibrational Energy
5Z	Quintuple Zeta

Chapter 1

Introduction

Chemical processes are everywhere, from the evolution of stars to our everyday experiences of sight, smell, taste, and touch. These processes involve intricate series of chemical reactions and interactions, prompting scientists to ponder: what is happening, why and how? Can they occur more efficiently or in a novel way? Can we exert selective control over them? How do they respond to disturbances and different conditions? And many more. To address such questions, understanding the behavior of molecules at the atomic level is crucial. As Richard Feynman famously stated, the most powerful assumption guiding our understanding of life is that all things are made of atoms, and everything can be understood in terms of the jigglings and wigglings of atoms.[1] Over the past five decades, much of the scientific community has dedicated itself to unraveling the nature of these atomic movements. However, capturing the dynamic essence of molecules experimentally, particularly at the atomic level and in a selective and controlled manner, is challenging. This is where computer simulations offer a compelling addition.[2–4] Molecular dynamics (MD) simulations have emerged as invaluable tools, providing microscopic insights into molecular dynamics and behavior.[5–7] Their effectiveness is strengthened by the advancement in computational power, algorithms, and techniques, enabling precise control over simulation conditions and providing a virtual platform to explore the complexities of atoms and molecules.[8–11]

Atomistic simulations not only aid in interpreting experimental and laboratory measurements but also play a vital role in predicting molecular properties, reaction

pathways, and products across diverse conditions.[12–14] These conditions can exhibit vast variability, spanning from the frigid temperatures of a few kelvins in the interstellar depths[15] to several hundred kelvins (ranging from 180 to 340 K) in the atmosphere.[16–18] However, beneath this diversity lies a common complexity: the intricate network of reactions within multiphase environments.[19–21] Researchers in both fields continuously discover new molecules and pathways, endeavoring to unravel the mechanisms governing molecular evolution and their potential impacts.[22–25]

In the realm of lower atmospheric chemistry, the reactions of ozone with volatile organic compounds (VOCs) have been a subject of interest for over a century.[26–30] These reactions significantly influence the concentrations of OH radicals, NO_x, ozone and secondary organic aerosol (SOA) precursors, thereby impacting both global warming and air pollution. Detailed chemical models are used to understand and predict their tropospheric concentrations, crucial for making environmental policy decisions.

The ozonolysis reaction proceeds through a 1,3-cycloaddition of ozone across the C=C double bond to form a primary ozonide (POZ) that subsequently decomposes into carbonyl compounds and energized carbonyl oxides, known as Criegee Intermediates (CIs).[27] These highly energized intermediates rapidly undergo either unimolecular decay to OH radicals or collisional stabilization.[31,32] Stabilized CIs can isomerize and decompose into products including the OH radicals or undergo bimolecular reactions with water vapor, SO_2, NO_2, and acids.[28,33] These intermediates are a nonphotolytic source of atmospheric (detergent) OH radicals during the day and account for nearly all of them at nighttime.[34,35]

Acetaldehyde oxide (CH_3CHOO) is one of the smallest CIs that can undergo unimolecular decay to OH radical.[36] The decomposition process of stabilized *syn*-CH_3CHOO proceeds through a five-membered transition state with an energy barrier of ~17 kcal/mol,[36–38] following 1,4-hydrogen transfer to form vinyl hydroperoxide (VHP). Subsequent homolytic cleavage of the O–OH bond leads to OH and vinoxy radicals.[39] Conversely, starting from the *anti*-CH_3CHOO isomer leads to methyldioxirane as the main product.[40] The conversion of the *syn*- to the *anti*-isomer

involves a barrier of ~42 kcal/mol.[41]

Experimentally, stabilized CIs have been extensively investigated by the Lester group through the use of lasers to photolyze precursors and pump-probe methods to unravel their decomposition into OH radicals.[36,42] For *syn*-CH₃CHOO, OH formation rates and products final state distributions were reported as a function of excitation in the CH overtone region.[42,43] Tunnelling through the barrier was also observed in the decomposition reaction.[44] On the theoretical front, electronic structure calculations were used to determine the energies, optimized geometries, vibrational frequencies and intrinsic reaction coordinates.[37,45] For rate computations, statistical Rice-Ramsperger-Kassel-Marcus (RRKM) theory along with tunneling corrections, was applied.[42,46] However, the statistical approach assumes rapid internal vibrational energy redistribution (IVR), meaning the OH-elimination step is fast compared to the 1,4-hydrogen transfer, and thus lacks a detailed understanding of the dynamics involved. To gain a dynamic perspective, quasiclassical trajectory (QCT) calculations were conducted on a full-dimensional potential energy surface (PES) based on permutationally invariant polynomials (PIPs)[47] and final state energy distributions were compared with experiments.[43] The QCT calculations were initiated at the transition state separating the *syn*-CH₃CHOO and VHP or followed the homolytic cleavage from VHP to dissociated products, instead of the *syn* minimum. This approach was chosen due to the long lifetime of the energized *syn*-CH₃CHOO and the large computational effort required to propagate the trajectories. However, the non-equilibrium nature of such processes was neglected in the calculations and the entire *syn*-CH₃CHOO $\longrightarrow$ CH₂CHOOH $\longrightarrow$ CH₂CHO+OH reaction pathway was not followed.

Chapter 3.1 explores in detail the decomposition dynamics of thermally and vibrationally excited *syn*-CH₃CHOO including all degrees of freedom and relevant states along the reaction pathway (see Figure 1.1).[48] Utilizing nanosecond timescale reactive MD simulations and PESs at the Møller-Plesset second-order (MP2) level of theory, the reactivity is studied. The vibrationally excited non-equilibrium state is prepared by scaling the instantaneous velocity vector along the methyl-CH stretch mode to match experimental excitation energies. Additionally, thermal simulations

reflecting the initial conditions post-ozonolysis of 2-butene are explored. By computing reaction time distributions and energy-dependent dissociation rates, insights into the underlying dynamics are obtained. A critical unknown parameter governing OH elimination is the O–O bond strength (D_e^{OO}) in VHP. Upon homolysis of the O–O bond, singlet VHP transforms into two doublet radicals, posing challenges associated with spin contamination necessitating high-level multireference calculations. Depending on the basis set and active space used for *ab initio* calculations, the energy required for O–O scission falls within the range of 18.8 - 31.5 kcal/mol.[43,45] Finally, the influence of D_e^{OO} on rates is also examined.
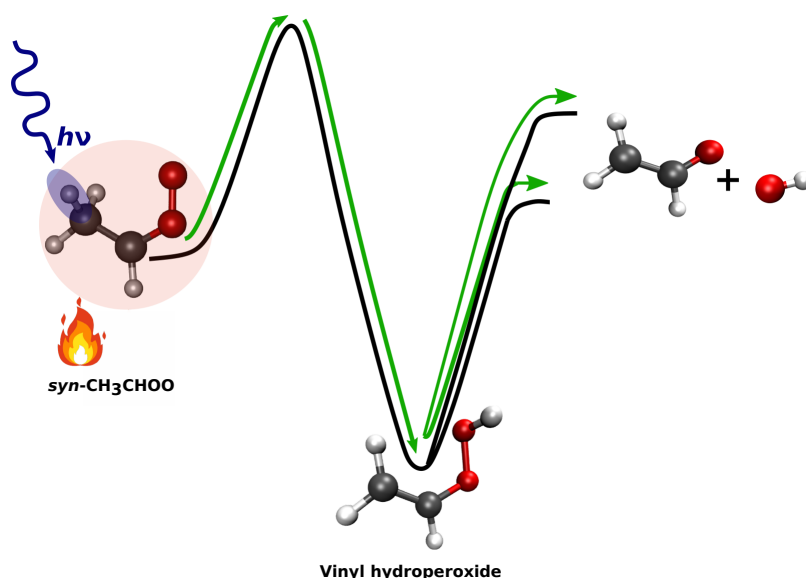
Figure 1.1: Schematic of PES relating the species involved in Chapter 3.1.

In Chapter 3.2, the dissociation of O–O bond in VHP is precisely addressed by constructing a full-dimensional PES using the complete active space second-order perturbation theory (CASPT2) method.[14] A reliable range for D_e^{OO} is determined by varying O–O bond strengths (highlighted part of the PES in Figure 1.2) and comparing computational final state analysis with experimental observables. Furthermore, an alternative reaction pathway involving OH roaming, which leads to the formation of glycolaldehyde (GA), is explored alongside the CH_2CHO + OH product channel. GA, an atmospheric VOC, is known to be produced from isoprene,[49] ethene[50] and biomass burning.[51] Experimental validation of GA formation follow-

ing the ozonolysis of trans-2-butene has been reported using molecular-beam mass spectrometry.[52] Additionally, recent experiments, RRKM and minimum energy path calculations have implicated a similar roaming behavior in the formation of hydroxybutanone from the unimolecular decay of the methyl-ethyl substituted CI.[53]
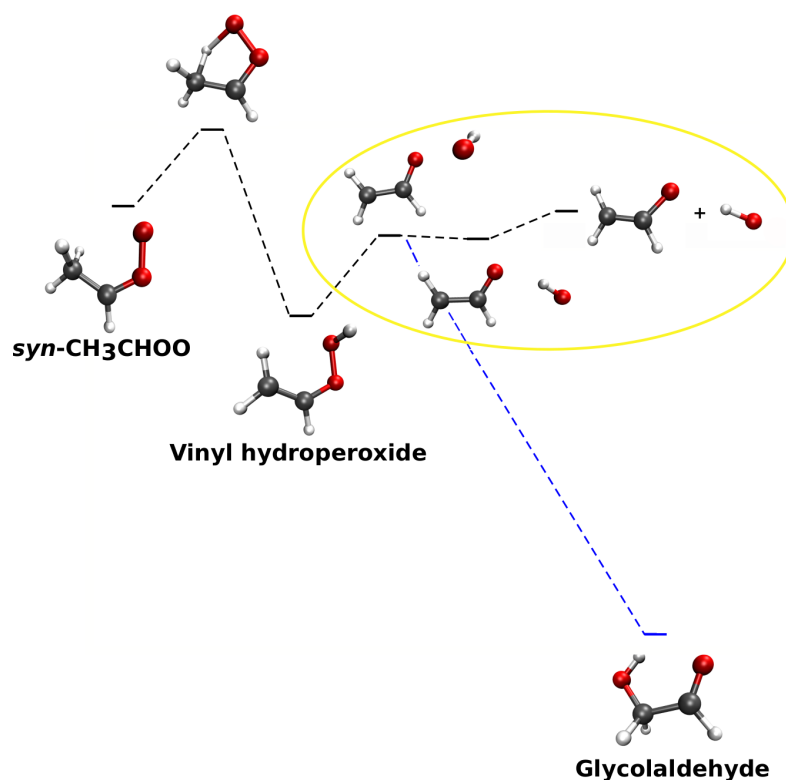
Figure 1.2: Schematic of PES relating the species involved in Chapter 3.2.

Interestingly, GA has also been detected in the interstellar medium[54] (ISM) and is believed to play a key role in the origin of life within the "RNA world hypothesis".[55,56] ISM is the diffuse material present in space that is not part of stars or other large celestial bodies, accounting for about 10 - 15% of our galaxy's total mass.[57] This mass is primarily composed of 99% gas (in ionic, atomic and molecular form) and 1% solids (dust or ice particles), with hydrogen atoms making up 89% of the gas, helium atoms 9% and the remaining 2% being heavier elements. Interstellar dust predominantly consists of heavy elements particles.[58] Despite its diffuse nature, discrete clouds can form in the ISM. These clouds typically exhibit translucent and diffuse characteristics at higher temperatures (T ~50 K), while dense molecular clouds are prevalent at temperatures around T ~10 K. Till now, more than 300

molecules have been detected in ISM (according to Cologne Database for Molecular Spectroscopy), with only about ten are cyclic, such as c-SiC$_2$, c-C$_3$H, c-C$_3$H$_2$, c-H$_2$C$_3$O, c-C$_2$H$_4$O, 2H-azirine, benzene, C$_{60}$, C$_{60}$ and C$_{60}^+$, while the rest are straight chain molecules. Molecules in ISM are detected through their unique spectroscopic signatures.[59,60] Over 80% of all the known species have been detected via their rotational spectral features, as transitions between rotational states are generally the least energetic compared to transitions between electronic and vibrational states.

The chemistry in interstellar clouds involves both gas-phase processes and reactions occurring on and within the surfaces of dust grains.[61,62] These surfaces primarily consist of silicates, carbonaceous species (graphite, amorphous carbon, polyaromatic hydrocarbons) or water.[63] In cold molecular clouds, cosmic dust grains are typically covered with water ice, which can be polycrystalline or amorphous. At low temperatures, the amorphous phase dominates over the polycrystalline phase.[64] Typically, bulk water in these environments is present in the form of ASW. ASWs are porous structures with surface roughness and internal cavities of varying sizes, capable of retaining molecular or atomic guests.[65] The high porosity facilitates the adsorption of chemical reagents either on the surface or within ASW, making it a suitable catalyst for gas-surface reactions involving oxygen,[66–73] hydrogen,[74] carbonaceous[75–78] and nitrogen-containing[79] species.

To gain insights into species abundances and their evolution, chemical models based on rate equations are commonly used.[62,80] These models rely on kinetic data sourced from databases[81,82] for species adsorption, desorption, diffusion, and reaction as input. However, a significant challenge arises from the lack of appropriate laboratory data on binding energies and desorption efficiencies of molecular ices of astrophysical interest, as well as the energy barriers and branching ratios for surface reactions. While large models can handle up to a few hundred different surface species, only a limited number of reaction systems and molecules have been theoretically or experimentally studied. To address these limitations and gain dynamical insights, computationally expensive Monte Carlo methods, *ab initio* molecular dynamics (AIMD), and classical MD approaches are utilized.[83–86] Chapter 4 uses a mixed Machine Learning/Molecular Mechanics (ML/MM) approach to model atom-diatom

reactions on and within ASW. In this approach, a small part of the system (specifically species involving bond breaking or forming) is described by accurate high-level machine-learned PESs and the remainder of the system is described by a classical force fields (FFs).[72,87–89]

In Chapters 4.1 and 4.2, the formation of ground-state CO_2 on and within ASW under conditions typical of cold molecular clouds is characterized using reactive MD simulations.[87,89] In these clouds, CO serves as a main carbon reservoir[61] and its reaction with atomic oxygen $O(^1D)$ is a barrierless pathway for CO_2 formation. To describe the atom-diatom energetics, model-based and coupled cluster (CC) level PESs are employed and reactivity is quantitatively probed using fixed point charges for CO and O. Here, fixed point charges refer to the fact that the charges for O ($q_O = -0.3$ e) and CO ($q_{CO} = +0.3$ e) remain fixed regardless of whether they are recombined or far from each other. Different aspects of the reactions are taken into account: the efficiency of the reaction, the relaxation mechanism and the possibility that desorption mechanism could happen in ISM condition. Given the highly exothermic nature of the recombination reaction, the relaxation of CO_2 on and within the water surface is also monitored.

The suppressed diffusion of reactants on the water surface is observed in Chapters 4.1 and 4.2 due to the fixed point charges used. At low temperatures (~50 K) in the interstellar medium, diffusion plays a pivotal role in driving chemical reactions. To ensure an accurate depiction of diffusion, the fluctuating point charge (FPC) model is employed for reacting species in Chapter 4.3. In the FPC model, atomic oxygen and diatoms (CO and NO) behave as neutrals on the water surface at larger distances, while at closer proximity, the charges on atoms (O, C and N) change as a function of the O–CO and O–NO separations.[88] Alongside CO_2, the NO_2 formation from nitric oxide (NO) and atomic oxygen is also explored.[90,91] The diffusion and desorption energies of the involved species are determined and compared with experiments wherever possible. Additionally, energy relaxation to the ASW and into the internal degrees of freedom of water molecules is investigated by analyzing the vibrational density of states (vDOS).

This dissertation is structured as follows: Chapter 2 provides an overview of the basic quantum mechanical concepts and computational methods used. Chapters 3 and 4 present the scientific publications comprising my doctoral research. Lastly, Chapter 5 summarizes the key findings of the publications and suggests potential future directions.

Chapter 2

Theoretical Background

This thesis mainly deals with the motion of atoms in both gaseous and condensed phases, utilizing MD simulations to propagate the system's state over the PES with time. To achieve this, a range of methodologies are employed, including molecular mechanics (MM) and machine learning (ML) approaches, to access the energy and gradient of the system. Quantum mechanics calculations play a pivotal role in providing the essential database for ML training and in parameterizing the FFs used. This chapter provides a concise overview of the methods used, setting the stage for their application in the subsequent chapters.

2.1 Potential Energy Surface

In quantum mechanics, the state of a system is completely described by its wavefunction Ψ. Certain quantum-mechanical operators are applied to extract quantifiable information from Ψ, some of which result in eigenvalue-eigenstate pairs. The most notable pair is the time-independent Schrödinger equation (SE)[92,93]

$$\hat{H}|\Psi\rangle = E|\Psi\rangle, \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian and E denotes the energy eigenvalue. For N nuclei and n electrons system, the Hamiltonian is expressed as

$$\hat{H}(\boldsymbol{r}, \boldsymbol{R}) = \hat{T}_{\mathrm{n}}(\boldsymbol{R}) + \hat{T}_{\mathrm{e}}(\boldsymbol{r}) + \hat{V}_{\mathrm{n-e}}(\boldsymbol{r}, \boldsymbol{R}) + \hat{V}_{\mathrm{e-e}}(\boldsymbol{r}) + \hat{V}_{\mathrm{n-n}}(\boldsymbol{R}), \tag{2.2}$$

where $\boldsymbol{R} = \{\boldsymbol{R}_A\}_{A=1}^{N}$ and $\boldsymbol{r} = \{\boldsymbol{r}_i\}_{i=1}^{n}$ represent the nuclear and electronic coordinates respectively. The first two terms describe the kinetic energy of nuclei $\hat{T}_n(\boldsymbol{R})$ and electrons $\hat{T}_e(\boldsymbol{r})$, while the subsequent terms describe the potential energy arising from nuclei-electron $\hat{V}_{n-e}(\boldsymbol{r}, \boldsymbol{R})$, electron-electron $\hat{V}_{e-e}(\boldsymbol{r})$ and nuclei-nuclei $\hat{V}_{n-n}(\boldsymbol{R})$ interactions.[94] The kinetic and potential energy operators are given as

$$\hat{T}_n(\boldsymbol{R}) = -\sum_{A}^{N} \frac{1}{2M_A} \nabla_A^2, \tag{2.3}$$

$$\hat{T}_e(\boldsymbol{r}) = -\sum_{i}^{n} \frac{\nabla_i^2}{2}, \tag{2.4}$$

$$\hat{V}_{n-e}(\boldsymbol{r}, \boldsymbol{R}) = -\sum_{A}^{N} \sum_{i}^{n} \frac{Z_A}{\bar{r}_{Ai}}, \tag{2.5}$$

$$\hat{V}_{e-e}(\boldsymbol{r}) = +\sum_{i}^{n} \sum_{i<j}^{n} \frac{1}{\bar{r}_{ij}}, \tag{2.6}$$

$$\hat{V}_{n-n}(\boldsymbol{R}) = +\sum_{A}^{N} \sum_{A<B}^{N} \frac{Z_A Z_B}{\bar{R}_{AB}}, \tag{2.7}$$

where M_A and Z_A represent the mass and charge of the nuclei, respectively, and $\bar{R}_{AB}$, $\bar{r}_{Ai}$ and $\bar{r}_{ij}$ correspond to the distances between the respective nuclei A, B and electrons i, j. Note that the equations are given in atomic units with electron mass m_e, reduced Planck constant $\hbar$ and elementary charge e equal to 1.

Unfortunately, the exact solutions of the SE are limited to single-electron systems, such as the hydrogen atom or one-electron ions. Consequently, for multi-electron systems, approximations become necessary. One of the foremost is the Born-Oppenheimer (BO) approximation,[95] which assumes that electronic and nuclear motions are independent due to their different masses. Electrons being much lighter than nuclei, move significantly faster. The approximation allows the molecular wavefunction to be expressed as the product of a nuclear and an electronic wavefunction: $\Psi(\boldsymbol{r}, \boldsymbol{R}) = \Psi_n(\boldsymbol{R}) \cdot \Psi_e(\boldsymbol{r}; \boldsymbol{R})$. Here, the electronic wavefunction depends on electronic coordinates while being parametrically dependent on nuclear coordinates. At any given time, the state of electrons is determined by the instantaneous positions of nuclei, resulting in nuclei to appear at rest from the electron's perspective. This fur-

ther leads to the neglection of nuclear kinetic energy, while nuclear-nuclear repulsion is treated as constant. Within the BO approximation, the electronic Hamiltonian is represented as:

$$\hat{H}_e(\boldsymbol{r}; \boldsymbol{R}) = \hat{T}_{\mathrm{e}}(\boldsymbol{r}) + \hat{V}_{\mathrm{n-e}}(\boldsymbol{r}; \boldsymbol{R}) + \hat{V}_{\mathrm{e-e}}(\boldsymbol{r}) + \underbrace{\hat{V}_{\mathrm{n-n}}(\boldsymbol{R})}_{\text{constant}} \qquad (2.8)$$

The eigenvalue of the electronic SE corresponds to the potential energy, thereby introducing the concept of a PES. The PES represents the energy of the molecule (or a collection of molecules) as a function of its nuclear coordinates.[94] It governs molecular dynamics, akin to how the motion of a ball rolling through a hilly landscape is dictated by the surface's topology. An illustrative example of a PES for the Ne–H_2^+ system is shown in Figure 2.1. In the following sections (2.2 and 2.3), a brief overview of the methods used to solve (or approximate) the electronic SE is provided.
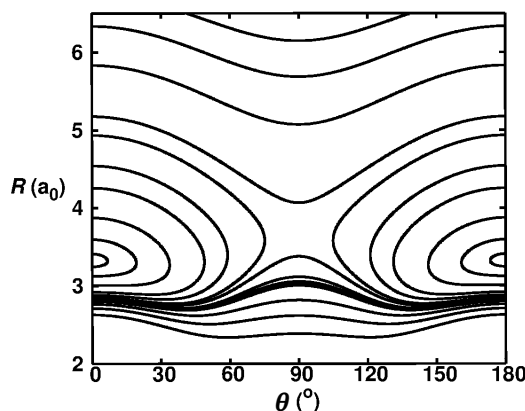
Figure 2.1: Contour plot of the Ne–H_2^+ PES at CCSD(T)/aug-cc-pV5z level of theory, with a fixed $r_{H_2^+} = 2.0$ a$_0$. Here, R represents the distance between the neon atom and center of mass of H_2^+, and θ denotes the angle between $\vec{R}$ and $\vec{r}_{H_2^+}$.

2.2 Basis Sets

Another approximation inherent in essentially all *ab initio* methods, crucial for solving the electronic SE, is the utilization of atom-centered basis set.[94] These are a set of known mathematical functions used to represent the unknown (single or many-) electron wavefunction of atoms and molecules. Within this framework, a set of N_{basis} known basis functions χ (also termed atomic orbitals) is considered, and

the unknown molecular orbital ϕ (MO) is expressed as a linear combination of these functions

$$\phi_i = \sum_{\alpha}^{N_{\text{basis}}} c_{\alpha i} \chi_{\alpha}, \tag{2.9}$$

where $c_{\alpha i}$ is the coefficient for a given χ_{α}. This approach is known as the linear combination of atomic orbitals (LCAO). While such expansion is not an approximation if the basis set is complete, achieving completeness requires an infinite number of functions, making it impractical. Therefore, the size and type of the basis set are critical factors influencing results accuracy. The goal is to strike a balance: using the smallest possible basis set without compromising accuracy.

The basis function used in computational chemistry can be classified into two main types: Slater-type orbitals (STOs) and Gaussian-type orbitals (GTOs).[94,96] They are expressed mathematically in terms of polar coordinates as:

$$\text{STO:} \quad \chi_{\zeta,n,l,m}(r,\Theta,\varphi) = N Y_{l,m}(\Theta,\varphi) r^{n-1} e^{-\zeta r} \tag{2.10}$$

$$\text{GTO:} \quad \chi_{\zeta,n,l,m}(r,\Theta,\varphi) = N Y_{l,m}(\Theta,\varphi) r^{2n-2-l} e^{-\zeta r^2} \tag{2.11}$$

where N and $Y_{l,m}$ correspond to normalization constant and spherical harmonic function, and r is the distance between electron and nucleus. Further, ζ is the orbital exponent, and n, l and m are principal, angular momentum, and magnetic quantum numbers, respectively. The r^2 dependence in the exponent makes the GTOs inferior to the STOs. While STOs closely resemble hydrogen-like orbital, GTOs offer the advantage of being less computationally demanding. However, a single Gaussian function cannot accurately represent an atomic orbital. Therefore, a combination of GTOs are used to construct an approximate Slater-type function. As a rough guideline, achieving a desired level of accuracy typically requires approximately three times as many GTOs as STOs.

The primary consideration in selecting a basis set is the number of functions to be utilized. The smallest number of functions possible is a minimal basis set (STO-3G,

i.e., 3 GTOs per STO), which employs a single basis function for each atomic orbital (AO). This means a single *s*-function for H and He atoms. For the first row elements of the periodic table, it employs two *s*- and one set of *p*-functions (1s/2s and $2p_x/2p_y/2p_z$). However, this basis set encounters challenges in accurately describing the expansion or contraction of electronic density when transitioning from an isolated atom to an atom within a molecule.[94]

So, the advancement involves doubling all basis functions known as double zeta (DZ) type basis. For first-row elements, a DZ basis set employs two s-functions for H, enhancing the description of the electron distribution. Similarly, the triple zeta (TZ), quadruple zeta (QZ), and quintuple zeta (5Z) employ three, four, and five times more basis functions, respectively, compared to the minimum basis set. To reduce the computational cost of DZ basis, split-valence basis sets are used. This approach leverages the concept that flexibility in valence orbitals is more crucial than in core orbitals, owing to the fact that core orbitals are usually less influenced by bond formation. It uses a different number of Gaussians to describe the core and the valence. Among the most commonly used split-valence basis sets are Pople basis sets.[97] The notation typically follows the format K-LMG, where K indicates the number of functions used for the core, L and M describe the valence, and the letter "G" indicates that these are Gaussian functions. For example, the 6-31G basis set uses 6 functions for the core electrons and four functions for the valence electrons: 3 contracted ones and 1 that is diffuse.

Further improvements can be achieved by incorporating the polarization effect using polarized basis functions (denoted by *). Additionally, diffuse basis functions, which exhibit slow decay at large distances, can be included for a better description of anions and van der Waal complexes (denoted by +). Moreover, correlation-consistent (cc) basis sets are often used to recover the correlation energy of the valence electrons by including functions that contribute similar amounts of correlation energy at the same stage, independent of the function type. For example, the first *d*-function provides a significant decrease in energy, and the (small) contribution from a second *d*-function is akin to that from the first *f*-function. The energy reduction from a third *d*-function is similar to that from the second *f*-function and the first *g*-function.

The addition of polarization functions therefore follows the order: 1d, 2d1f and 3d2f1g. These basis sets are denoted as cc-pVNZ where N = D, T, Q, 5 (D= double, T= triple, etc). The "cc-p" stands for "correlation-consistent polarized" while "V" denotes a split-valence basis. The inclusion of diffuse functions further improves the flexibility of the outer valence region, leading to the augmented correlation-consistent basis sets (aug-cc-pVNZ).[98] Explicitly correlated basis set (F12) can also be used for fast convergence of correlation energies.[99]

2.3 Ab Initio Methods

Ab initio methods, also known as the first principles methods, are fundamental in solving the electronic SE. Unlike semi-empirical methods, which rely on experimental data, *ab initio* methods derive solutions solely from theoretical principles. In the following, an introduction to the Hartree-Fock theory,[100,101] configuration interaction method,[102] Møller–Plesset perturbation theory,[103] coupled cluster theory[104], and multireference approaches[105–108] is given.

2.3.1 Hartree-Fock Theory

Hartree-Fock (HF) theory is fundamental to much of electronic structure theory, serving as the basis for MO theory. It assumes that the motion of each electron can be described by a single-particle function (orbital), independent of the instantaneous motions of other electrons ($\hat{V}_{e-e}(r) = 0$). Thus, it represents the *n*-electron wavefunction as the product of the single-electron wavefunctions, known as the Hartree product:[94]

$$\Psi_{\text{HP}}(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_n) = \Phi_1(\boldsymbol{r}_1)\Phi_2(\boldsymbol{r}_2)\cdots\Phi_n(\boldsymbol{r}_n). \tag{2.12}$$

However, this formulation fails to satisfy the Pauli exclusion principle,[109] which states that the wavefunction must change sign upon the interchange of any two electrons. This underscores the necessity of extending the electrons description

beyond spatial coordinates alone to include an intrinsic spin coordinate (denoted by α or β), resulting in a set of spin-orbitals $\chi(\boldsymbol{x}) = \Phi(\boldsymbol{r})\alpha$ or $\Phi(\boldsymbol{r})\beta$. However, this replacement alone does not yield an antisymmetric wavefunction. Therefore, to fulfill this requirement, the wavefunction is constructed as a Slater determinant:

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \chi_1(\boldsymbol{x}_1) & \chi_2(\boldsymbol{x}_1) & \cdots & \chi_n(\boldsymbol{x}_1) \\ \chi_1(\boldsymbol{x}_2) & \chi_2(\boldsymbol{x}_2) & \cdots & \chi_n(\boldsymbol{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\boldsymbol{x}_n) & \chi_2(\boldsymbol{x}_n) & \cdots & \chi_n(\boldsymbol{x}_n) \end{vmatrix}. \tag{2.13}$$

A single Slater determinant is used as an approximation to the wavefunction, and the MOs are expressed as linear combinations of atomic orbitals (χ^{AO}) (one-electron functions or basis functions):

$$\chi_i = \sum_{k=1}^{N_{\mathrm{basis}}} c_k^i \chi_k^{\mathrm{AO}}. \tag{2.14}$$

Here, c_k^i are the expansion coefficients and χ_k^{AO} are the AOs or basis functions. Then the variational principle,[94,110] which states that the best spin orbitals for a given Slater determinant are those that minimize the electronic energy, is employed to determine the optimal c_k^i.

$$E_{\mathrm{e}} \leq \langle \Psi_{\mathrm{test}} | \hat{H}_{\mathrm{e}} | \Psi_{\mathrm{test}} \rangle \qquad \Psi \text{ normalized} \tag{2.15}$$

The electronic energy minimization employs an iterative procedure called self-consistent field (SCF) method,[94] resulting in the Hartree-Fock equation

$$\hat{f}(i)\chi_i(\boldsymbol{x}_i) = \epsilon_i \chi_i(\boldsymbol{x}_i), \tag{2.16}$$

$$\hat{f}(i) = \hat{h}(i) + \sum_{j \neq i} \left[\mathcal{J}_j(i) - \mathcal{K}_j(i) \right], \tag{2.17}$$

where $\hat{f}(i)$ is the Fock operator and ϵ_i represents the energy of the i-th orbital. The terms $\hat{h}$, $\mathcal{J}$ and $\mathcal{K}$ are the one-electron, Coulomb, and exchange operators, respectively, given by

$$\mathcal{J}_j(i)\chi_i(\boldsymbol{x}_i) = \left[\int d\boldsymbol{x}_j\chi_j^*(\boldsymbol{x}_j)\frac{1}{r_{ij}}\chi_j(\boldsymbol{x}_j)\right]\chi_i(\boldsymbol{x}_i), \tag{2.18}$$

$$\mathcal{K}_j(i)\chi_i(\boldsymbol{x}_i) = \left[\int d\boldsymbol{x}_j\chi_j^*(\boldsymbol{x}_j)\frac{1}{r_{ij}}\chi_i(\boldsymbol{x}_j)\right]\chi_j(\boldsymbol{x}_i). \tag{2.19}$$

When summed over all orbitals, the Coulomb term describes the interaction of an electron in spin-orbital χ_i with the average charge distribution of the other electrons, illustrating the "mean field" character. The exchange operator is similar to the coulomb term but exchanges the spin orbitals. The HF wavefunction is the best non-interacting solution for an interacting multi-electron problem.

In HF approximation, electron interactions are treated solely at the mean-field level, where each electron experiences a spatially averaged repulsion from all others. Despite recovering over 99% of the total energy, the HF approach is not entirely exact.[111] This lack of accuracy arises because the single Slater determinant wavefunction includes no terms that account for the relative positions of two electrons and mean-field approximation. The remaining 'correlation' energy, denoted as E_{corr}, represents the difference between the exact energy and the HF energy.[112]

$$E_{\mathrm{corr}} = E_{\mathrm{exact}} - E_{\mathrm{HF}} \tag{2.20}$$

Beyond the HF description, electron correlation is broadly categorized into dynamic and static components. Dynamic correlation corrects the limitations of the mean-field approximation as electrons move synchronously to maximize their separation from each other.[113] However, the mean-field approximation assumes independent electron motion, leading to a systematic underestimation of electron distances and consequently an overestimation of electron repulsion energies. This coordinated electron movement primarily affects electrons within the same orbital but occurs among all electrons in a system.

Conversely, static correlation is associated with the qualitative breakdown of the single slater determinant wavefunction.[113] This occurs in cases of nearly degenerate occupied and unoccupied orbitals, where a model based on a single electronic configuration becomes inadequate. Static correlation is commonly quantified as the energy difference between a wavefunction incorporating all possible electron arrangements within the valence orbitals and the HF wavefunction.

Despite its limitations, HF provides a good starting point for more elaborate theoretical methods that address electron correlation effects, such as configuration interaction, many-body perturbation theory, and coupled cluster.

2.3.2 Configuration Interaction Theory

The configuration interaction (CI) method improves the HF solution by increasing the space from a single Slater determinant to a set of Slater determinants.[102] In this method, the wavefunction is expressed as a linear combination of *N*-electron configurations

$$|\Psi_{\text{CI}}\rangle = c_0 |\Psi_0\rangle + \underbrace{\sum_{a,r} c_a^r |\Psi_a^r\rangle}_{\text{single}} + \underbrace{\sum_{a<b,r<s} c_{ab}^{rs} |\Psi_{ab}^{rs}\rangle}_{\text{double}} + \underbrace{\sum_{a<b<c,r<s<t} c_{abc}^{rst} |\Psi_{abc}^{rst}\rangle}_{\text{triple}} + ..., \qquad (2.21)$$

where Ψ_0 is the HF determinant, *a*, *b*, *c* are occupied MOs, *r*, *s*, *t* are virtual (unoccupied) MOs, and $|\Psi_a^r\rangle$, $|\Psi_{ab}^{rs}\rangle$, $|\Psi_{abc}^{rst}\rangle$ are possible determinants after single, double and triple excitations. The coefficients of the different configurations are optimized using the variational principle. When all possible excited Slater determinants are included the method is called Full CI (FCI). Such calculation is even unfeasible for small molecules because the number of determinants that need to be handled increases rapidly with the system size and basis set. Therefore, a truncation of CI expansion is performed, in which the excitation operator is truncated to include only specific excitation levels to lower the cost of calculations. Among the widely used truncated methods are those that only include double excitations (CID) or

include both single and double excitations (CISD). [94,114]

Unfortunately, any truncated CI method is not size consistent in general, i.e. the energy of two non-interacting (infinitely separated) molecules is not simply the sum of their individual energies, as calculated within the same approximation. This discrepancy occurs because the same configurations are not taken into account for the separate fragments and the combined system. As the system's size increases, it becomes necessary to include higher excitations.

2.3.3 Møller-Plesset Perturbation Theory

Perturbation theory is a commonly utilized method to improve results obtained at a particular lower level of theory, such as HF. The theory states that the approximate solution deviates from the exact solution by a small correction term. It assumes that a part of the full Hamiltonian, denoted as $\hat{H}^0$, along with its solution E_0 is known. The full Hamiltonian is then expressed as the sum of this known component and a small perturbation for which solution is unknown, $\hat{H} = \hat{H}^0 + \lambda \hat{H}'$. The prefactor λ determines the magnitude of the perturbation and is intentionally chosen to be small, ensuring that the eigenfunctions of the full Hamiltonian remain closely aligned with those of the known Hamiltonian. [94,110]

When the Hamiltonian is chosen to be the HF Hamiltonian, one arrives at the Møller-Plesset (MP) perturbation method. [103] For a system of n electrons, the unperturbed HF Hamiltonian $\hat{H}^0$ is expressed as the sum over the Fock operators

$$\hat{H}^0 = \sum_{i=1}^{n} \hat{f}_i = \sum_{i=1}^{n} \left(\hat{h}_i + \sum_{j=1}^{n} \mathcal{J}_j - \mathcal{K}_j \right) \tag{2.22}$$

$$= \sum_{i=1}^{n} \hat{h}_i + \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \hat{v}_{ij} \rangle = \sum_{i=1}^{n} \hat{h}_i + 2\langle \hat{V}_{e-e} \rangle \tag{2.23}$$

which double counts the average electron-electron repulsion term. Then the perturbed Hamiltonian $\hat{H}'$ is given by:

$$\hat{H}' = \hat{H} - \hat{H}^0 = \sum_{i=1}^{n} \sum_{j=1}^{n} \hat{v}_{ij} - \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \hat{v}_{ij} \rangle = \hat{V}_{\text{e-e}} - 2\langle \hat{V}_{\text{e-e}} \rangle. \tag{2.24}$$

The HF determinant (zeroth-order wavefunction) yields the zeroth-order energy

$$E(\text{MP0}) = E_0 = \langle \Psi_0 | \hat{H}^0 | \Psi_0 \rangle = \langle \Psi_0 | \sum_{i=1}^{n} \hat{f}_i | \Psi_0 \rangle = \sum_{i=1}^{n} \epsilon_i \tag{2.25}$$

which corresponds to the sum of MO energies ϵ_i. Each MO energy represents the energy of an electron interacting with all the nuclei and its repulsion to all other electrons. Therefore, the electron-electron repulsion is counted twice. The first-order energy correction $E(\text{MP1})$ and the total electronic energy E_{tot} is given by

$$E(\text{MP1}) = E_1 = \langle \Psi_0 | \hat{H}' | \Psi_0 \rangle = \langle \hat{V}_{\text{e-e}} \rangle - 2\langle \hat{V}_{\text{e-e}} \rangle = -\langle \hat{V}_{\text{e-e}} \rangle, \tag{2.26}$$

$$E_{\text{tot}} = E(\text{MP0}) + E(\text{MP1}) = E_0 + E_1 = E(\text{HF}). \tag{2.27}$$

Therefore, MP1 compensates for the over-counting of electron-electron repulsion and matches the HF energy. Corrections addressing electron correlation begin from the second-order correction onward. The second-order energy correction requires summing over doubly excited determinants, expressed as

$$E(\text{MP2}) = \sum_{a<b}^{\text{occ.}} \sum_{r<s}^{\text{virt.}} \frac{[\langle \phi_a \phi_b | \hat{v} | \phi_r \phi_s \rangle - \langle \phi_a \phi_b | \hat{v} | \phi_s \phi_r \rangle]^2}{(\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s)}. \tag{2.28}$$

Here, a and b are occupied orbitals and r and s are unoccupied (virtual) orbitals. The terms ϵ_a, ϵ_b, ϵ_r and ϵ_s are the corresponding orbital energies. The E_{tot} is given by

$$E_{\text{tot}} = E(\text{MP0}) + E(\text{MP1}) + E(\text{MP2}). \tag{2.29}$$

As a perturbative approach, unlike variational, MP2 might yield energies lower than the exact ground state energy. However, its accuracy strongly depends on the choice of reference wavefunction.[115,116] The MP2 approximation fails when the HF determinant doesn't serve as a suitable reference wavefunction, particularly in cases

where bond breaking introduces a multireference character.

2.3.4 Coupled Cluster Theory

While perturbation methods incorporate all types of corrections (S, D, T, ...) to the reference wavefunction to a given order (first, second, third, ...), coupled cluster (CC) methods include all corrections of a certain type (S, D, T, ...) to infinite order.[104,117] In CC theory, the wavefunction is expressed using an exponential ansatz:

$$|\Psi\rangle = e^{\hat{T}}|\Phi_0\rangle \tag{2.30}$$

Here, $|\Phi_0\rangle$ denotes the initial Slater determinant typically constructed from HF molecular orbitals, and $\hat{T}$ represents the cluster operator, expressed as:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + ..., \tag{2.31}$$

where $\hat{T}_1$ represents the operator for all single excitations, $\hat{T}_2$ for all double excitations, and so forth. These excitation operators are often defined in terms of creation $\hat{a}^\dagger$ and annihilation operators $\hat{a}$ as follows:

$$\hat{T}_1 = \sum_a \sum_r t_a^r \hat{a}_a \hat{a}_r^\dagger, \tag{2.32}$$

$$\hat{T}_2 = \frac{1}{4} \sum_{ab} \sum_{rs} t_{ab}^{rs} \hat{a}_a \hat{a}_b \hat{a}_r^\dagger \hat{a}_s^\dagger. \tag{2.33}$$

Here, a and b (r and s) denote occupied (unoccupied) orbitals, with the coefficients of excitation operators denoted by t_a^r and t_{ab}^{rs} respectively. The exponential operator $e^{\hat{T}}$ is expressed by a Taylor expansion as

$$e^{\hat{T}} = 1 + \hat{T} + \frac{\hat{T}^2}{2!} + \frac{\hat{T}^3}{3!} + ... = 1 + \hat{T}_1 + \hat{T}_2 + \frac{\hat{T}_1^2}{2!} + \hat{T}_1\hat{T}_2 + \frac{\hat{T}_2^2}{2!} + ... \tag{2.34}$$

If all cluster operators up to $\hat{T}_n$ are included, the CC wavefunction becomes equivalent to FCI. However, for practical application, the cluster operator is usually truncated at a certain excitation level. For instance, when $\hat{T}$ includes only $\hat{T}_1$ and $\hat{T}_2$, it represents the CCSD method;[118] including $\hat{T}_1$, $\hat{T}_2$, and $\hat{T}_3$ yields the CCSDT

method. Additionally, when triples are derived based on perturbation theory, it leads to the CCSD(T) method,[119] often known as the current "gold standard method" in electronic structure theory. Even when excitations are included only up to $\hat{T}_2$, the wavefunction still contains contributions from Slater determinants of higher ($n > 2$) excitations. This is due to terms like $\hat{T}_1\hat{T}_2$ and $\frac{\hat{T}_2^2}{2!}$, representing triple and quadruple excitations, which contribute to recovering more correlation energy.[94]

While CCSD(T) is widely regarded as a top-tier method, it is important to note that it remains a single-reference approach and often encounters limitations in regions exhibiting multireference character or state degeneracies. Handling such challenges requires the use of computationally expensive multireference approaches.

2.3.5 Multireference Approaches

Single-reference methods like truncated CI, truncated CC and MP2 theory often fail in cases where orbital near-degeneracy problem arises. The emergence of multiple dominant electronic configurations indicates that the reference HF determinant inadequately approximates the exact wavefunction, highlighting the presence of strong static correlation effects. To tackle this limitation, a common approach involves employing a multireference wavefunction using the multiconfigurational self-consistent field (MCSCF)[105] framework. However, the computationally feasible MCSCF method typically provides only qualitative accuracy. Additional CI, CC, or perturbation approaches are often necessary to compute the remaining dynamic electron correlation.

Multiconfigurational Self-Consistent Field

When multiple electronic configurations dominate (or contribute more evenly) the exact wavefunction, a multiconfigurational wavefunction can be effectively represented as the truncated CI expansion:

$$|\Psi\rangle_{\text{MCSCF}} = \sum_I c_I |\Psi_I\rangle. \tag{2.35}$$

The transition to MCSCF method occurs when not only the CI expansion coefficients (c_I), but also the MOs used to construct the determinants are optimized through the variational principle.[105,113] By including all configurations in the MCSCF expansion, one can achieve the exact FCI energy and wavefunction. In practice, the set of configurations necessary to describe the multireference character of a system are chosen *a priori.*

The most common method for generating the configurations is the complete active space self-consistent field (CASSCF) method,[106,107] wherein the MOs are divided into sets of inactive, active and virtual orbitals (see Figure 2.2). The inactive orbitals, also known as core orbitals, remain doubly occupied, while the virtual orbitals clearly remain unoccupied. The active orbitals, constituting the active space, accommodate the 'problematic' *n* electrons and *m* MOs, conventionally denoted as CASSCF(*n,m*). All possible arrangements of these *n* electrons into the *m* MOs are considered within the active space. Consequently, the size of the active space determines the total number of Slater determinants in the MCSCF wavefunction. As the FCI expansion operates within the active space, the number of determinants scales factorially with the number of active electrons and orbitals, limiting the number of orbitals/electrons that can be included in the active space. The selection of an appropriate active space significantly influences the quality of results and is often the most challenging aspect of such calculations.
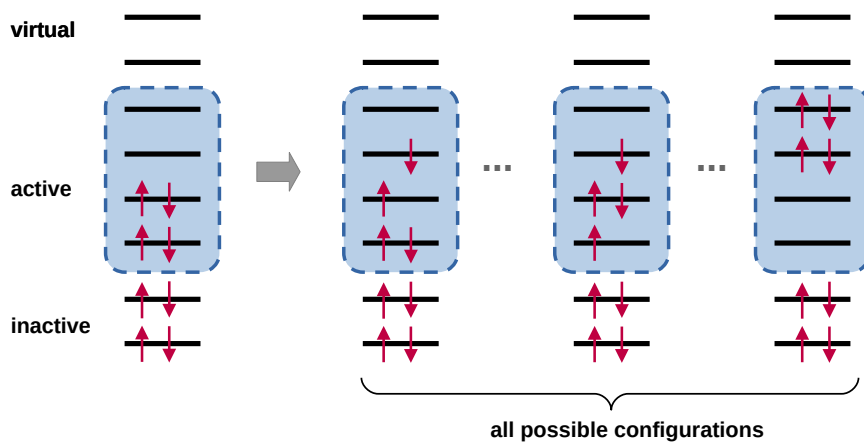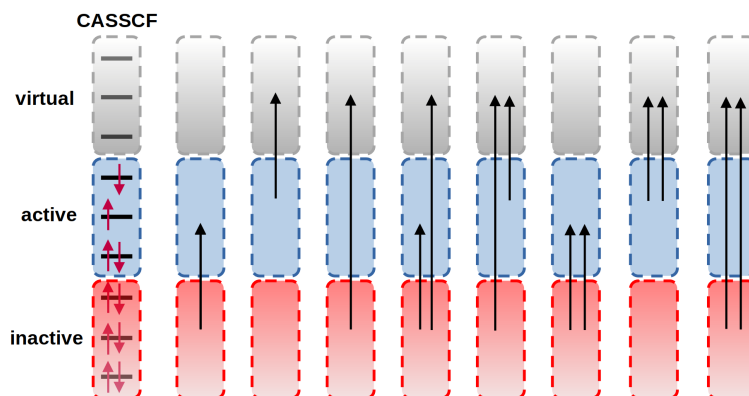

Figure 2.2: A schematic representation of the orbitals classification in the CASSCF method.

Another limitation of the method is its inability to fully capture the dynamic correlation effects. While electron correlation within the active space is covered, no electron correlation energy is recovered from the inactive space, and between the inactive, active, and virtual spaces. The remaining electron correlation, associated with the instantaneous short-range electron-electron interaction (dynamic correlation), can be recovered by using complete active space second-order perturbation theory (CASPT2)[108] and multireference configuration interaction (MRCI)[120] methods.

Complete Active Space Second-Order Perturbation Theory

The CASPT2 method is similar to the MP2 method discussed in section 2.3.3, but a CASSCF wavefunction is used as the reference (zero-order) wavefunction.[108,121] In this approach, more determinants are included by considering single and double replacements out of the inactive orbitals and into the virtual orbitals. The resultant excited determinants include, for instance, excitations from the inactive to the active orbitals, from the active to the virtual orbitals, from the inactive to virtual orbitals (see Figure 2.3). Being a perturbative approach the method is highly sensitive to the choice of reference orbitals. Nonetheless, it remains a widely employed perturbation approach, particularly in the context of excited-state calculations.

Figure 2.3: A schematic representation of the excited determinants included in the CASPT2 method.

At the end of this section, it should be underlined that the selection of a method depends on the system being studied. Often the most accurate methods are computa-

tionally expensive, necessitating a balance between results quality and computational cost. Additionally, their applicability is confined to relatively small systems. For larger systems, a simplified approach FF or MM is commonly employed due to their computational efficiency. [122,123]

2.4 Molecular Mechanics

In MM, atoms are represented as spheres with associated mass and charges. These atoms are then assembled into molecules, with the bonds between them modeled as springs. A set of parameters and equations are assigned to each atom pair to describe the potential energy of the system. The mathematical expression, consisting of these parameters and equations that describe the system's energy dependence on the coordinates of its atoms is termed a FF. The current generation of force fields offers a well-balanced compromise between accuracy and computational efficiency. Extensive testing has confirmed their capability to reproduce experimentally measured physical properties accurately. The most commonly used FFs include Chemistry at HARvard Molecular Mechanics (CHARMM), [124,125] Optimized Potential for Liquid Simulations (OPLS) [126] and Assisted Model Building with Energy Refinement (AMBER). [127]

The CHARMM FF is widely used to study small molecules and biomolecular systems, like peptides, proteins, nucleic acids, and lipids. [124] The general form of the potential energy function based on fixed point charges is calculated as a sum of bonded and non-bonded energy terms (Eq. 2.36). The bonded term, U_{bonded} includes contribution from bond stretching, angle bending and dihedral rotation as explained in Figure 2.4. The non-bonded term, $U_{\text{non-bonded}}$ accounts for interactions between non-bonded atoms i.e. atoms separated by more than three bonds. [128]

$$U(\vec{R}) = \underbrace{U_{\text{bonds}} + U_{\text{angles}} + U_{\text{dihedrals}} + U_{\text{impropers}}}_{U_{\text{bonded}}} + \underbrace{U_{\text{electrostatic}} + U_{\text{vdW}}}_{U_{\text{non-bonded}}} \qquad (2.36)$$

The potential energy associated with bonds, denoted as U_{bonds}, is characterized by a summation of harmonic potentials acting upon specified atom pairs identified as

24
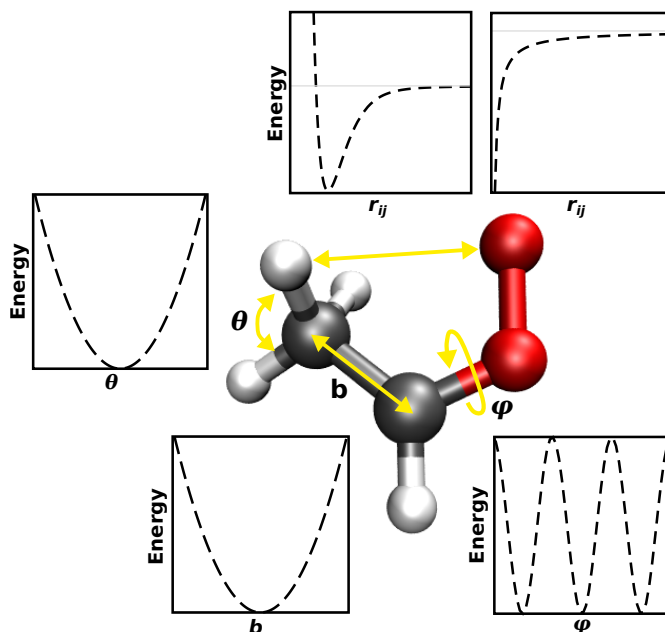
Figure 2.4: Schematic representation of the FF terms together with their potentials.

being bonded. The term accounts for the energy associated with bond length, b, as it varies from the equilibrium bond length, b_{eq}. The parameter k_b is the force constant which governs the strength of the bond. Both k_b and b_{eq} parameters are specific to the chemical type of the bonded pairs.

$$U_{\text{bonds}} = \sum_{\text{bonds}} \frac{1}{2} k_b (b - b_{\text{eq}})^2 \tag{2.37}$$

The harmonic potential used to describe the bonds is only valid when the bonds are close to their equilibrium configuration. In certain cases where bond breaking needs to be modeled, a Morse potential[129] is used to describe the anharmonicity.

$$U_{\text{morse}} = D_{\text{e}} \big[1 - e^{(-\beta(b - b_{\text{eq}}))} \big]^2 \tag{2.38}$$

Here, D_{e} is the dissociation energy, b_{eq} is the equilibrium bond length and β controls the width of the potential well.

Similar to the bond energy term, U_{angles} describes the deviations of the bond angles, θ from their equilibrium values θ_0 with a harmonic potential. The force constant, k_θ governs the strength of the bond angle.

$$U_{\text{angles}} = \sum_{\text{angles}} \frac{1}{2}k_\theta(\theta - \theta_0)^2 \tag{2.39}$$

In cases where four atoms are bonded, torsional terms are used to describe the rotational barriers between them. There are two types of torsional terms: proper and improper dihedrals. The proper torsional term, denoted as $U_{\text{dihedrals}}$, is defined for four bonded atoms in a row and associated with the dihedral angle (φ) between two planes sharing the common atoms i.e. the plane defined by atoms C, C, and O and the plane defined by atoms C, O, and O (see Figure 2.4).[128]

$$U_{\text{dihedrals}} = \sum_{\text{dihedrals}} \frac{1}{2}k_\varphi(1 + cos(n\varphi - \gamma)) \tag{2.40}$$

Here, k_φ denotes the strength parameter, n is the integer determining the number of minima and maxima the cosine function exhibits within the angular range of 0 to 2π and γ represents the phase shift.

On the other hand, the $U_{\text{impropers}}$ term is used to describe a four-atom system when three of the atoms are connected to a central atom that lies out of the plane. It is described by a harmonic function and required to maintain the chirality of a tetrahedral heavy atom and to maintain planarity about certain atoms.

$$U_{\text{impropers}} = \sum_{\text{impropers}} k_\omega(\omega - \omega_0)^2 \tag{2.41}$$

Here, k_ω represents the strength parameter, while ω is the angle associated with the out-of-plane motion of an atom C from the plane defined by three hydrogen atoms of -CH_3 group (see Figure 2.4).

The non-bonded term, denoted as $U_{\text{non−bonded}}$, includes both van der Waals and electrostatic contributions. The van der Waals term, U_{vdW} is commonly represented by a Lennard-Jones 12-6 potential[130] to account for the repulsion and attraction between two non-bonded atoms. The attractive component exhibits r^{-6} dependence

which arises due to interactions at long range (dispersion force), while the repulsive component has r^{-12} dependence, which becomes predominant at short distances.

$$U_{\text{vdW}} = \sum_{\text{vdW}} \varepsilon_{ij} \left[\left(\frac{r_{\min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{\min,ij}}{r_{ij}} \right)^{6} \right] \tag{2.42}$$

In the LJ potential, r_{ij} is the distance between non-bonded atoms where i and j are the indices of the interacting non-bonded atoms, ε_{ij} is the well depth and $r_{\min,ij}$ is the value of r at which LJ potential is minimum. Typically in CHARMM FF, the ε_{ij} and $r_{\min,ij}$ values for each pair are obtained by a combination rule i.e. ε_{ij} via geometric mean ($\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$) and $r_{\min,ij}$ via arithmetic mean ($r_{\min,ij} = (r_i + r_j)/2$).

The remaining contribution is the electrostatic interaction, denoted as $U_{\text{electrostatic}}$, between atom-centered point charges, represented by the Coulomb potential.

$$U_{\text{electrostatic}} = \sum_{\text{electrostatic}} \frac{q_i q_j}{4 \pi \varepsilon_0 r_{ij}} \tag{2.43}$$

Here, q_i and q_j are the partial charges of non-bonded atoms and ϵ_0 is the effective dielectric constant.

Conventional FFs typically assume fixed atomic charges and connectivity between atoms, as they were designed for low-amplitude vibrations and conformational motions. However, to model chemical reactions involving bond breaking or formation, several MD methods have been emerged. A key feature of these methods is the incorporation of a switching function between diabatic states. One such method is the Multi-Surface Adiabatic Reactive Molecular Dynamics (MS-ARMD).

2.4.1 Multi-Surface Adiabatic Reactive Molecular Dynamics

MS-ARMD[131] is a broadly applicable, energy conserving surface crossing algorithm based on the potential energy of the system under investigation. The method utilizes parameterized FFs for both reactant and product states, referred to as surfaces and GAussian and POlynomial (GAPOs) functions to describe the surface crossing regions (area near the transition state). MS-ARMD offers flexibility by allowing

the replacement of harmonic bonds with morse potentials, facilitating the bond formation and breaking. Additionally, to improve the description of van der Waals interactions along the reaction path, Generalized Lennard Jones (GLJ) replaces the traditional 6-12 Lennard Jones potentials.

$$V_{\text{GLJ}}(r; p, q, \epsilon, r_{\min}) = \frac{p\epsilon}{q - p} \left[\left(\frac{r_{\min}}{r} \right)^q - \frac{q}{p} \left(\frac{r_{\min}}{r} \right)^p \right] \tag{2.44}$$

Here, ϵ represents the well depth of the GLJ potential and $r_{\min}$ is the separation at the minimum energy.

In order to combine the individual surfaces of reactant and product states, coordinate dependent weights $w_i(\mathbf{x})$ are assigned for each surface $V_i(\mathbf{x})$. According to MS-ARMD, the effective surface is always the lowest-energy surface, except for geometries where other surfaces approach similar energies. In such instances, the algorithm smoothly transitions among them by adjusting their weights. The effective potential energy is defined as the linear combination of n PESs with weights $w_i(\mathbf{x})$.

$$V_{\text{MS−ARMD}}(\mathbf{x}) = \sum_{i=1}^{n} w_i(\mathbf{x}) V_i(\mathbf{x}) \tag{2.45}$$

In Eq. 2.45, $w_i(\mathbf{x})$ is obtained by renormalizing the raw weights $w_{i,0}(\mathbf{x})$ and an exponential decay function is used to calculate $w_{i,0}(\mathbf{x})$

$$w_i(\mathbf{x}) = \frac{w_{i,0}(\mathbf{x})}{\sum_{j=1}^{n} w_{j,0}(\mathbf{x})}, \tag{2.46}$$

$$w_{i,0}(\mathbf{x}) = \exp\left(-\frac{V_i(\mathbf{x}) - V_{\min}(\mathbf{x})}{\Delta V} \right), \tag{2.47}$$

where $V_i(\mathbf{x})$ and $V_{\min}$ is the potential energy of the i-th and the minimum energy surface, respectively. ΔV is a user defined switching parameter determined during the parametrization of the global PES by fitting it to *ab initio* data, such as the minimum energy path. To better characterize the barrier region GAPOs are used to match the electronic structure calculations energies,

$$\Delta V_{\text{GAPO},k}^{ij}(\mathbf{x}) = \exp\left(-\frac{(\Delta V_{ij}(\mathbf{x}) - V_{ij,k}^0)^2}{2\sigma_{ij,k}^2} \right) \sum_{l=0}^{m_{ij,k}} a_{ij,kl}(\Delta V_{ij}(\mathbf{x}) - V_{ij,k}^0)^l, \tag{2.48}$$

where $V_{ij,k}^0$ and $\sigma_{ij,k}$ are the center and standard deviation of the Gaussian function, and $a_{ij,k}$ is the coefficient of the polynomial of order $m_{ij,k}$. Finally, the resulting global PES is then given by

$$V_{\text{MS−ARMD}}(\boldsymbol{x}) = \sum_{i=1}^{n} w_i(\boldsymbol{x}) V_i(\boldsymbol{x}) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[w_i(\boldsymbol{x}) + w_j(\boldsymbol{x}) \right] \sum_{k=1}^{n_{ij}} \Delta V_{\text{GAPO},k}^{ij}(\boldsymbol{x}). \qquad (2.49)$$

The parametrization strategy used is described in the next section.

2.4.2 Force Field Parametrization

First, the initial parameters for reactant and product FFs were taken from Swiss-Param.[132] Representative structures were then sampled from MD simulations at the required temperature, and their *ab initio* energies were determined at the desired level of theory using programs such as Gaussian09[133] or Molpro.[134] Subsequently, the FFs for the reactant and product were fitted separately to *ab initio* energies using a downhill simplex[135] algorithm. This improved set of parameters were then used to generate new structures, whose energies were combined with reference energies and used for further refinement of the FF parameters. Several rounds of refinements were done until the root mean squared deviation (RMSD) for the final set between the target (*ab initio*) and the fitted FF energies were within chemical accuracy of $\approx$ 1 kcal/mol.

The downhill simplex algorithm, also known as the Nelder-Mead method,[135] is a robust technique used for the minimization of a function by rolling a polyhedron downhill to its lowest possible value. The term 'simplex' refers to a geometrical body with $N + 1$ vertices in an N-dimensional space. This method iteratively replaces the vertex of the simplex with the highest function value by a new point using reflection. Reflection involves mirroring the vertex along the line connecting it to the centroid of the simplex. If the new point achieves the lowest function value, an expansion operation is executed, pushing the point further along that direction. Conversely, if the newly computed point retains the highest function value, a contraction occurs, moving the point in the opposite direction. If the new point has a function value neither the highest nor the lowest, it is accepted, and the algorithm selects the

current vertex with the highest value to proceed with the process. The algorithm terminates if no further improvement can be found within a predefined tolerance value. Given that the construction of the initial simplex heavily relies on the quality of the initial parameters, it is important to provide a good estimation to ensure the algorithm's effectiveness. This procedure leads to the convergence of the FF energies with respect to *ab initio* energies within a few iterations.

The two FFs (here reactant and product) are then connected using GAPO functions to ensure a continuous connection along the reaction path. This process involves determining the intrinsic reaction coordinate (IRC) between the reactant and product, which is used in the fitting process. A genetic algorithm is employed to parametrize the GAPO functions and replicate the energies along the reaction path. The genetic algorithm operates on the principle of natural selection, where the best solutions survive until convergence is achieved.[136] It begins with an initial population of individuals (possible solutions), represented by genes or variables. Selection based on a fitness function is then performed to identify the high-fitness individuals, often determined by how well their parameters fit the target energy profile. Through genetic operations such as crossover and mutation, the algorithm generates new offspring by combining and modifying the parameters of selected individuals. Crossover involves combining the genetic information of two parents to explore the solution space further, while mutation introduces random changes to explore a wider range of the parameter space. Individuals with higher fitness are more likely to be chosen as parents for the next generation. The algorithm iterates until the newly generated offspring fail to improve upon the parent generation, indicating convergence. The applied algorithm is based on differential evolution, a method that systematically explores the parameter space, with a focus on regions where the function value is low.[136,137]

Employing empirical FFs for modeling PESs offers notable advantages, particularly in terms of computational efficiency and the ability to handle systems with thousands of atoms. However, the straightforward functional form of FFs, while enhancing computational efficiency, comes with inherent limitations. These limitations include compromised accuracy and a coarse representation of stretch/bend-couplings.[138] To

generate accurate PESs for small systems, it is often preferable to directly interpolate a set of *ab initio* energies to construct a continuous functional form. One such interpolation technique is the reproducing kernel Hilbert Space method (RKHS).[139]

2.5 Reproducing Kernel Hilbert Space Method

In recent years, advancements in the efficiency and accuracy of high-level electronic structure calculations have facilitated high-quality *ab initio* calculations for small polyatomic systems. However, maintaining the precision of these calculations requires representations capable of preserving accuracy. RKHS[139] emerges as a robust technique for constructing smooth and precise analytical global PES representations for small molecules. This method utilizes gridded *ab initio* reference data to accurately reproduce calculated on-grid energies and effectively captures long-range interactions when appropriate kernel polynomials are selected.

Within the RKHS framework, potential energies for a system can be expressed as a linear combination of reproducing kernel functions, using a set of known energies $V(x)$ at different configurations x. The represener theorem[140] states that for a given set of N training points x_i, along with the corresponding values $f_i = f(x_i)$ of a function $f(x)$, the function $f(x)$ can always be approximated as a linear combination of kernel products

$$f(x) = y = \sum_{i=1}^{N} c_i K(x, x_i), \tag{2.50}$$

where $K(x, x_i)$ is the reproducing kernel and c_i are coefficients that satisfy the linear relation

$$f_j = \sum_{i=1}^{N} c_i K_{ij} \tag{2.51}$$

with $K_{ij} = K(x_i, x_j)$. These coefficients are computed from the known values f_i in the training set using Cholesky decomposition[141]

$$c = K^{-1} y, \tag{2.52}$$

where $c = [c_1, ..., c_N]$ is the vector of coefficients, K is a $N \times N$ kernel matrix and $y = [y_1, ..., y_N]^T$ is the vector containing the calculated dataset. Once the coefficients have been determined, the unknown values at any arbitrary positions x can be interpolated using eq. 2.50.

Since the method reproduces the reference data exactly, it is prone to overfit the noise potentially present in the training data set. To mitigate this, a small positive constant λ is added to the diagonal of K to regularize the solution[142]

$$c = (K + \lambda I)^{-1} y, \tag{2.53}$$

where I is the identity matrix. When applied to represent discrete data for N configuration energies, the PES can be expressed as

$$V(x) = \sum_{i=1}^{N} c_i K(x, x_i), \tag{2.54}$$

where x_i represents the training set geometries for which *ab-initio* energies have been determined. The coefficients are then determined from the known *ab initio* energies for N configurations by solving the following matrix equation

$$\begin{vmatrix} K(x_1, x_1) & K(x_1, x_2) & \cdots & K(x_1, x_N) \\ K(x_2, x_1) & K(x_2, x_2) & \cdots & K(x_2, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ K(x_N, x_1) & K(x_N, x_2) & \cdots & K(x_N, x_N) \end{vmatrix} \begin{vmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{vmatrix} = \begin{vmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \end{vmatrix}. \tag{2.55}$$

For the multidimensional (M) case, the kernel can be approximated by a product of one-dimensional kernels $K_j(x, x')$

$$K(x, x') = \prod_{j=1}^{M} c_i K_j(x, x'). \tag{2.56}$$

To generate PES for triatomic systems, the grids for *ab initio* energy calculations are set up in Jacobi coordinates (R, r, θ), where, r is the diatomic bond length, R denotes the distance between the atom and the center of mass of the diatomic,

and θ represents the angle between r and R. For the radial dimensions (r and R) a reciprocal power decay kernel is used

$$k(x, x') = \frac{2}{15} \frac{1}{x_>^5} - \frac{2}{21} \frac{x_<}{x_>^6},$$

(2.57)

where $x_>$ and $x_<$ are the larger and smaller value of x and x'. The used kernel smoothly decays to zero according to x^{-4} as the leading term in the asymptotic region, thereby ensuring the accurate description of long-range behavior in atom-diatom interactions. For the angular dimension, a Taylor spline kernel is used

$$k(z, z') = 1 + z_< z_> + 2z_<^2 z_> - \frac{2}{3}z_<^3,$$

(2.58)

where $z_>$ and $z_<$ are similar to $x_>$ and $x_<$. The variable z is defined as $z = \frac{1-cos\theta}{2}$ to ensure that z remains within the interval $[0, 1]$. Finally, the 3-dimensional kernel is the product of $k(R, R')$, $k(r, r')$ and $k(z, z')$. To compute the derivatives of $f(x)$ analytically, the kernel function $K(x, x')$ is replaced with its corresponding derivative. When RKHS accurately represents the PES, its derivative also provides a good approximation of the gradients.

Despite its effectiveness, RKHS encounters scalability challenges with larger training sets and is primarily suited for constructing PESs for small systems (triatomics), where systematically generating reference data in grid form is feasible. In contrast, for medium-sized systems ($8 - 10$ atoms), Neural Network (NN) potential provides a balanced solution, offering both accuracy and flexibility in the PES construction. [143,144]

2.6 Neural Network

In high-dimensional configuration spaces where interpolation techniques such as splines or polynomial expansions do not work well, NN-based approaches emerge as powerful tools, effectively bridging the gap between precision and computational efficiency. Unlike traditional methods, NN-based PESs operate non-parametrically and are not bound by predetermined functional forms. Essentially, these methods

learn the functional relationship between inputs (chemical descriptors) and outputs (properties) from patterns or structures in the data.

The fundamental building blocks of fully connected NNs are known as dense layers,[145,146] which transform an input vector $\mathbf{x}$ to an output vector $\mathbf{y}$ according to the equation:

$$y = Wx + b, \tag{2.59}$$

where the weights $\boldsymbol{W}$ and biases $\boldsymbol{b}$ are learnable parameters. However, a single dense layer can only depict linear relationships. In order to model a non-linear relationship, a minimum of two dense layers, combined with a non-linear activation function σ (e.g. a sigmoid function), is necessary.

$$h = \sigma\left(W_1 x + b_1\right) \tag{2.60}$$

$$y = \left(W_2 h + b_2\right) \tag{2.61}$$

Equations 2.60 and 2.61 are general function approximators, capable of accurately modeling any relationship between input $\boldsymbol{x}$ and output $\boldsymbol{y}$ to arbitrary precision, provided that the dimensionality of the 'hidden layer' $\boldsymbol{h}$ is sufficiently large.[147] In practice, deep NNs with multiple hidden layers are more parameter-efficient. Such networks stack l hidden layers on top of each other:

$$h_1 = \sigma(W_1 x + b_1)$$
$$h_2 = \sigma(W_2 h_1 + b_2)$$
$$\vdots$$
$$h_l = \sigma(W_l h_{l-1} + b_l) \tag{2.62}$$
$$y = W_{l+1} h_l + b_{l+1}$$

These equations map the input $\boldsymbol{x}$ to increasingly complex feature spaces until the features $\boldsymbol{h_l}$ in the final layer are linearly related to the output $\boldsymbol{y}$. The parameters of the NN, including the entries in the matrices $\boldsymbol{W_l}$ and vectors $\boldsymbol{b_l}$, are randomly initial-

ized and then optimized to minimize a loss function that quantifies the difference between the NN's output and a given set of training data. The mean squared error (MSE) loss function is a common choice for regression tasks.

Many different flavors of NN-based PESs exist, tailored to different applications. For large chemical systems, a common strategy involves decomposing the total energy of an N atomic system into individual atomic contributions,[144]

$$E = \sum_{i=1}^{N} E_i. \tag{2.63}$$

Subsequently, the energies E_i of each atom are then predicted using single or multiple NNs. This method, first proposed by Behler and Parrinello and known as high-dimensional neural networks (HDNNs), relies on the chemically intuitive assumption that an atom's contribution to the total energy primarily depends on its local chemical environment.[144] Message passing is one variant of graph neural networks (GNNs), introduced by Schütt and coworkers.[148,149] It utilizes nuclear charges and cartesian coordinates as input to learn a meaningful description of the local chemical environment through exchanging "messages". This method also serves as the foundation for PhysNet,[143] a model utilized in this thesis.

PhysNet is a message-passing HDNN that predicts atomic energy contributions and partial charges from feature vectors that encode information about each atom's local chemical environment. These feature vectors are iteratively constructed from Cartesian coordinates $\mathbf{R}_i$ and nuclear charges Z_i of all atoms i by passing messages between atoms within a cutoff distance r_{cut}. Initially, the features of atom i rely solely on Z_i, and they are refined using learnable functions by coupling with all feature vectors of atoms j within r_{cut}. The total potential energy E of a system is expressed as

$$E = \sum_{i=1}^{N} E_i + k_e \sum_{i=1}^{N} \sum_{j>i}^{N} \tilde{q}_i \tilde{q}_j \chi(R_{ij}) + E_{\mathrm{D3}}. \tag{2.64}$$

Here, E_i represents the predicted atomic energy contribution, while q_i and q_j are corrected partial charges of atoms i and j. R_{ij} denotes the distance between atoms i and j, k_e is the Coulomb's constant and E_{D3} refers to the D3 dispersion correction.[150] The Coulomb term is modified and damped for small distances (R_{ij}), employing $\chi(R_{ij})$ to prevent numerical instabilities arising from the singularity at $R_{ij} = 0$.

$$\chi(R_{ij}) = \phi(2R_{ij})\frac{1}{\sqrt{R_{ij}^2 + 1}} + (1 - \phi(2R_{ij}))\frac{1}{R_{ij}} \tag{2.65}$$

Here, ϕ is a smooth cut-off function. A correction scheme, such as Equation 2.66, is employed to ensure charge conservation such that the sum of all predicted atomic partial charges q_i equals the total charge Q.

$$\tilde{q}_i = q_i - \frac{1}{N}\left(\sum_{j=1}^{N} q_j - Q\right) \tag{2.66}$$

The forces $\mathbf{F}_i$ acting on each atom i is obtained analytically by reverse mode automatic differentiation[151] and the total dipole moment is computed from the partial charges q_i (see Equation 2.67).

$$\boldsymbol{\mu} = \sum_{i=1}^{N} q_i \mathbf{R}_i \tag{2.67}$$

During training, all parameters are fitted to reference *ab initio* energies, forces and dipole moments using the procedure described in the Ref.[143] The parameters of the NN are optimized, minimizing a loss function

$$
\begin{aligned}
L = w_E|E - E^{\text{ref}}| &+ \frac{w_F}{3N}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\left|-\frac{\partial E}{\partial R_{i,\alpha}} - F_{i,\alpha}^{\text{ref}}\right| + w_Q\left|\sum_{i=1}^{N} q_i - Q^{\text{ref}}\right| \\
&+ \frac{w_p}{3}\sum_{\alpha=1}^{3}\left|\sum_{i=1}^{N} q_i R_{i,\alpha} - p_\alpha^{\text{ref}}\right| + L_{\text{nh}},
\end{aligned} \tag{2.68}
$$

where E^{ref}, Q^{ref}, $\boldsymbol{\mu}^{\text{ref}}$ and $\mathbf{F}_i^{\text{ref}}$ are the reference energy, total charge, dipole moment and force respectively. w_E, w_Q, w_p and w_F are the hyperparameters, determining the relative contribution of the respective errors to the loss function. The L_{nh} term is a

nonhierarchicality penalty. If the weights of lower-lying modules (building blocks of the PhysNet) do not decrease, the prediction is penalized. This is because lower-lying modules are expected to capture higher-order interactions, which typically decay fast in magnitude.

Constructing a NN-based PES typically requires extensive *ab initio* calculations, ranging from tens to hundreds of thousands. This often limits accuracy to low-level theories like MP2 due to computational constraints, especially for medium-sized molecules. To address this challenge, the transfer learning (TL) approach has gained significant attention in recent years.[152–154]

TL improves the global PES from a lower to a higher level of theory, providing a data-efficient alternative that requires only a fraction of the high-level data. This approach bypasses the need for extensive high-level *ab initio* datasets by leveraging the observation that molecular PES topographies at different theory levels often exhibit similarities. The TL process involves two stages: Initially a low-level PES is constructed using a large dataset of low-level (less computationally expensive) reference data to obtain an optimal set of weights W_i^{LL} and biases b_i^{LL}. This process minimizes the difference between the lower-level properties (energies and forces) and the NN predictions for a given set of structures. Next, the (extensive) low-level reference data is replaced with limited (≈ 10 times less) high-level reference data, and the weights W_i^{LL} and biases b_i^{LL} of the NN model are retrained. As a result, a high-level PES is achieved using significantly fewer reference data points. This approach effectively reduces the training and computation times while maintaining accuracy.

Finally, these accurate PESs are used to run quantitative atomistic MD simulations, providing a detailed description of the underlying intermolecular interactions.

2.7 Molecular Dynamics Simulations

MD is a numerical technique used to determine the evolution of a system over time by solving equations of motion. It requires defining a set of initial conditions (positions and velocities of each particle), employing a robust model to represent the forces acting between the particles (either empirical FFs or ML-potentials), and setting appropriate boundary conditions.[124,155] Subsequently, the following equations are solved

$$\text{Newton's second law:} \qquad \boldsymbol{f}_i = m_i \boldsymbol{a}_i = m_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2}. \qquad (2.69)$$

Here, $\boldsymbol{f}_i$ is the exerted force, m_i , $\boldsymbol{a}_i$, $\boldsymbol{r}_i$ are the mass, acceleration, and position of atom i at time t, respectively. The force $\boldsymbol{f}_i$ is determined by the gradient of the potential energy function, $U(\boldsymbol{r}^N)$:

$$\boldsymbol{f}_i = -\frac{\partial U}{\partial \boldsymbol{r}_i} \qquad (2.70)$$

The trajectory is obtained by solving the differential equation, adhering to specific criteria such as energy conservation, time-reversibility, and computational efficiency. Commonly used integration algorithms include velocity Verlet[156] and leapfrog.[157]

In the velocity Verlet algorithm, the new positions and velocities of each atom at time $t + \delta t$ are computed using the positions and velocities at time t:

$$\boldsymbol{r}(t + \delta t) = \boldsymbol{r}(t) + \boldsymbol{v}(t)\delta t + \frac{1}{2}\boldsymbol{a}(t)\delta t^2 \qquad (2.71)$$

$$\boldsymbol{v}(t + \delta t) = \boldsymbol{v}(t) + \frac{1}{2}[\boldsymbol{a}(t) + \boldsymbol{a}(t + \delta t)]\delta t \qquad (2.72)$$

The propagation of coordinates and velocities (using Equation 2.71 and 2.72) enables the recording of a trajectory. The simultaneous calculation of positions and velocities in the velocity Verlet algorithm offers a key advantage over the leapfrog algorithm, where positions and velocities are updated half a time step apart.

In the leapfrog algorithm, velocities at time $t + \frac{1}{2}\delta t$ are initially calculated using velocities at time $t - \frac{1}{2}\delta t$ and the acceleration at time t. These velocities at time $t + \frac{1}{2}\delta t$ are then used to compute positions at time $t + \delta t$ as follows:

$$\boldsymbol{r}(t + \delta t) = \boldsymbol{r}(t) + \boldsymbol{v}(t + \frac{1}{2}\delta t)\delta t \tag{2.73}$$

$$\boldsymbol{v}(t + \frac{1}{2}\delta t) = \boldsymbol{v}(t - \frac{1}{2}\delta t) + \boldsymbol{a}(t)\delta t \tag{2.74}$$

MD simulations are typically performed in ensembles where the number of atoms N and two other properties remain constant. Keeping energy E and volume V constant results in a microcanonical ensemble (NVE). Additionally, when temperature T is kept constant alongside N and V, one achieves a canonical ensemble (NVT). In the isothermal-isobaric ensemble (NpT), temperature and pressure p are maintained constant, along with N.[124]

The following chapters will illustrate the utilization of the discussed theoretical tools in exploring a wide array of processes, ranging from molecule formation to decomposition in astrophysics and atmospheric chemistry fields.
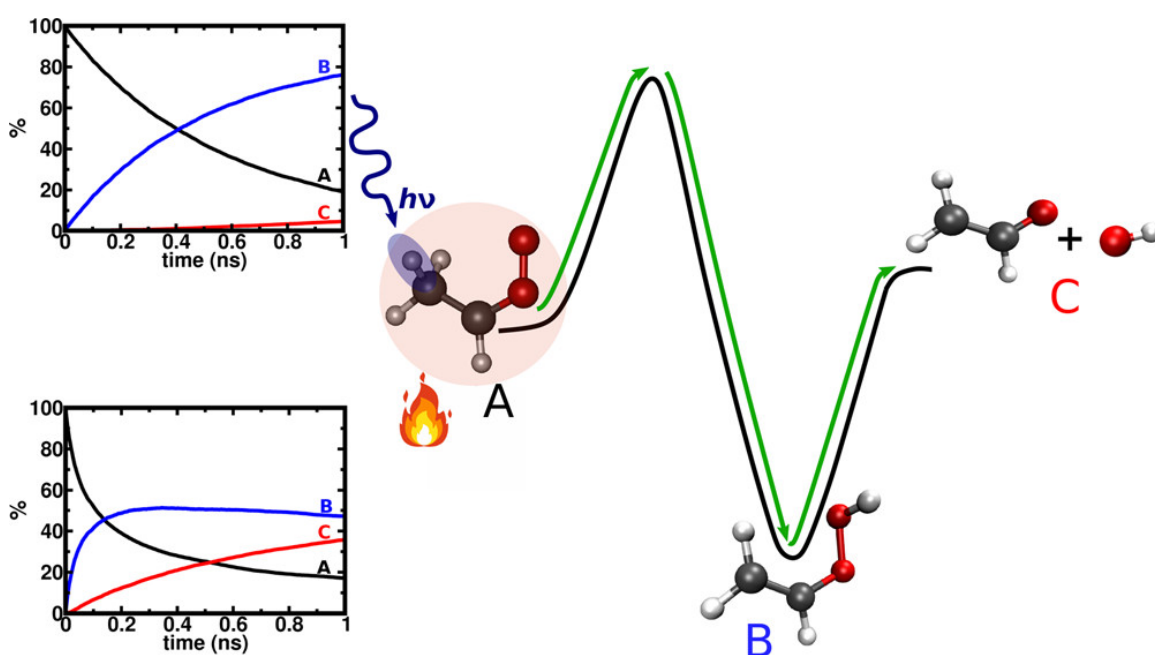
Chapter 3

Unimolecular Decomposition of *syn*-CH$_3$CHOO Criegee Intermediate

3.1 Publication: Thermal and Vibrationally Activated Decomposition of the *syn*-CH₃CHOO Criegee Intermediate

The results presented in this chapter have been previously published in:

ACS Earth Space Chem. **2021**, *5(12)*, 3396–3406.

3.1.1 Abstract

The full reaction pathway between the *syn*-CH$_3$CHOO CI via VHP to CH$_2$COH+OH is followed for vibrationally excited and thermally prepared reactants. Reactivity along the entire pathway was characterized from an aggregate of more than 10 μs of reactive MD simulations using energy functions with accuracies at the Møller-Plesset second order level of theory. Reaction times for OH elimination are on the nanosecond time scale, and the energy dependence of the rates is consistent with experiments in the jet. The actual rates depend on the O–O dissociation energy (D_e^{OO} = 31.5 kcal/mol at the MP2/aug-cc-pVTZ level or D_e^{OO} = 23.5 kcal/mol closer to earlier CASPT2 calculations). Also, the initial preparation of the reactants influences both, the VHP formation/OH elimination rates and the OH yields. For most conditions with initial vibrational excitation 80 % or more of *syn*-CH$_3$CHOO progress to OH elimination on the 5 ns time scale. However, for internally cold conformational ensembles generated at low temperature (50 K) only 20 % to 30 % of VHP eliminate OH on the 5 ns time scale which increases to 55 % to 67 % depending on excitation energy from simulations on the 15 ns time scale. For thermal preparation of *syn*-CH$_3$CHOO, which is relevant in the atmosphere, 35 % of the trajectories lead to OH-elimination within 1 ns. This compares with experimentally reported yields of 24 % to 64 % in a collisional environment. The estimated thermal rate at 300 K is 10^3 s^{-1}, extrapolated from results at 500 K to 900 K, is consistent with the experimentally measured rate of 182 ± 66 s^{-1}.

3.1.2 Introduction

The hydroxyl radical, one of the most powerful oxidizing agents, plays an important role in the chemical evolution of the atmosphere.[158] OH, also referred to as the "detergent of the troposphere",[159,160] triggers degradation of many pollutants including VOCs and is an important chain initiator in most oxidation processes in the atmosphere. The amount of OH generated from alkene ozonolysis is an important determinant required for chemical models of the lower atmosphere. Field studies have suggested that ozonolysis of alkenes is responsible for the production of about one third of the atmospheric OH radicals during daytime, and is the predominant source of OH radicals at night.[35,161] Alkene ozonolysis proceeds through a 1,3-cycloaddition of ozone across the C=C bond to form a primary ozonide which then decomposes into carbonyl compounds and energized carbonyl oxides, known as CIs.[27] These highly energized intermediates rapidly undergo either unimolecular decay to hydroxyl radicals[31] or collisional stabilization.[32] Stabilized CIs can isomerize and decompose into products including the OH radical or undergo bimolecular reactions with water vapor, SO$_2$, NO$_2$ and acids.[28,33] The high energy and short lifetime of these zwitterionic species complicates their direct experimental characterization.

One of the smallest CIs that can either follow unimolecular decay to generate OH or bimolecular reactions under atmospheric conditions is the acetaldehyde oxide (CH$_3$CHOO). This species is generated from ozonolysis of trans-2-butene.[162–164] Unimolecular decomposition of stabilized *syn*-CH$_3$CHOO proceeds through a five membered transition state with an energy barrier of ~17 kcal/mol,[36–38] following 1,4-hydrogen transfer to form VHP. Subsequent homolytic cleavage of the O$_A$–O$_B$H bond (see Figure 3.1A) leads to OH and vinoxy radical.[39] Conversely, starting from the *anti*-CH$_3$CHOO isomer, the main product is methyl-dioxirane which proceeds through a ring-closure step.[40] The conversion of the *syn*- to the *anti*-isomer involves a barrier of ~42 kcal/mol which makes such a reorganization highly unlikely.[41]

In a comprehensive joint theory/experiment study[43] a PES based on PIPs was reported and used in QCT calculations. This PES was fit to ~157000 electronic energies (a mixture of CCSD(T) and CASPT2) and covered the *syn*-CH$_3$CHOO,
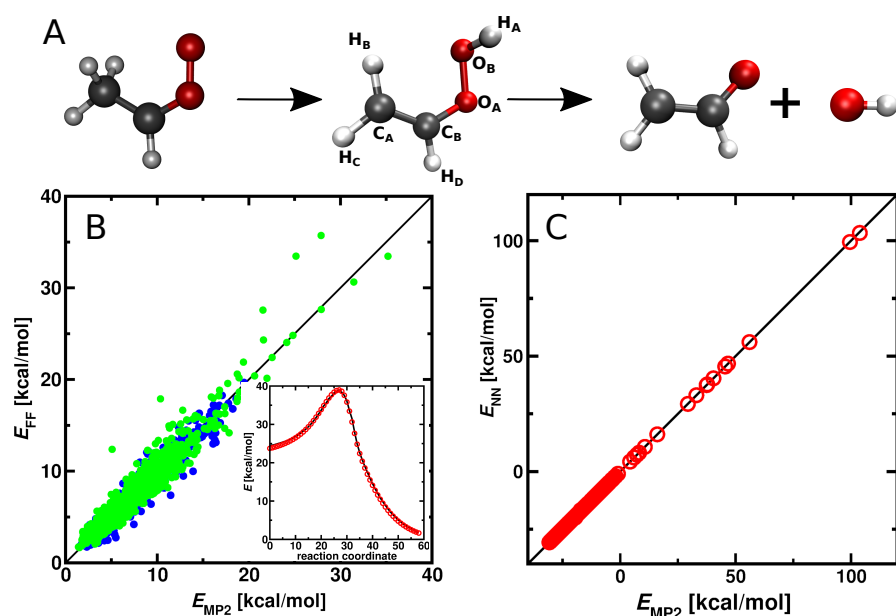
Figure 3.1: Quality of the PESs. Panel A: Schematic representation of OH formation starting from the *syn*-CH$_3$CHOO CI (left) via VHP (middle) through a 1,4-hydrogen shift reaction and subsequent OH-elimination to yield vinoxy radical + OH (right). Panel B: Correlation of 1612 *ab initio* reference structures and the fitted force field for reactant (blue) and product (green) with RMSE value of 1.1 and 1.2 kcal/mol, respectively. Inset: *ab initio* IRC (red circles) and fitted MS-ARMD (black curve). Panel C: Correlation between reference (MP2) and predicted (PhysNet) energies. The 10540 test set structures are predicted with MAE of 0.02 kcal/mol and RMSE of 0.19 kcal/mol.

the TS to VHP, several exit channel wells and the OH+vinoxy products. The QCT calculations were initiated at the TS separating the *syn*-CH_3CHOO and VHP wells instead of the *syn* minimum. This was done owing to the long lifetime of the energized *syn*-CH_3CHOO and the large computational effort needed to propagate the trajectories. Since that is exactly the focus of the present work this PES was not used but instead a computationally more efficient representation based on MS-ARMD was used. The computational effort for MS-ARMD is similar to evaluating an empirical force field and is thus suitable for running statistically meaningful numbers of trajectories ($\sim 10^4$ and more) for long times (tens to hundreds of nanoseconds).[131] In addition, an accurate NN-based representation following the PhysNet architecture is used.[143] Such PESs have been previously used, e.g., to investigate isomerization of acetaldehyde on the 0.5 μs time scale[165] and are therefore also suitable for exhaustive sampling and sufficiently long-time simulations.

Direct time-domain experimental rates, for appearance of OH from unimolecular dissociation of *syn*-CH_3CHOO under collision free conditions were obtained by vibrationally activating the molecules at specific energies in the vicinity of[42] and below[44] the transition state barrier of 1,4-hydrogen transfer. Statistical RRKM[166] rates with tunneling and zero point energy corrections agreed with experimentally determined OH formation rates at energies close to the barrier (~ 6000 cm⁻¹). It is interesting to note that at the highest excitation energies the corrections due to tunneling are only a factor of 3. Towards lower energies the RRKM rates without tunneling corrections differ more substantially because the number of open channels is small. For the present case of a two-step reaction the RRKM calculations need to assume that the OH-elimination step is fast compared with the 1,4 hydrogen transfer for direct comparison with experiment. Also, RRKM simulations provide rates but an understanding of the underlying dynamics involved is not possible.

The second step for OH formation is O–O cleavage starting from VHP which converts singlet VHP into two doublet radicals (OH and CH_2COH).[45] Electronic energy differences at the MP2 level of theory between VHP and the dissociation products are 31.5 and 35.8 kcal/mol using the 6-31G(d) and aug-cc-pVTZ basis sets, respectively.[45] Calculations at the multireference configuration interaction singles and

doubles (MRCISD) level of theory report dissociation energies between 14.3 and 17.8 kcal/mol depending on the basis set used. However, as the asymptotic energy differences appear to depend both, on the size of the active space and the basis set used, a definitive dissociation energy for this process is currently not available. In another, more recent effort [43] the PES for the OH elimination reaction was determined at the CASPT2(12,10)/cc-pVDZ level of theory which find a first submerged barrier, 23.0 kcal/mol above VHP, and the second, "positive barrier", 29.3 kcal/mol above VHP. Both of them connect to the same asymptotic state ($CH_2COH+OH$) which is 27.2 kcal/mol above the VHP energy.

Experiments on deuterated *syn*-CD_3CHOO reported a large kinetic isotope effect of ~50, indicative of tunneling, for unimolecular decay at energies near the barrier. [167] Similarly, experiments for *syn*-CH_3CHOO which only excited one quantum in the CH stretch (~3000 cm^{-1}, i.e., below the barrier for the 1,4-hydrogen shift reaction) also reported OH formation which is only possible through tunneling. [168] Computational work reported the possibility for rearrangement of the vinoxy and hydroxyl radical to form hydroxyacetaldehyde instead of O–O bond homolysis in VHP. [169] The dynamics of energized *syn*-CH_3CHOO from simulations initiated at the TS towards VHP reported prompt OH-dissociation without visiting VHP along the pathway. [170] Finally, thermal unimolecular decay is also a relevant pathway for *syn*-CH_3CHOO loss in the atmosphere with rates at room temperature of 100 to 300 s^{-1}. [171,172]

The previous high-level computational efforts either started the dynamics at the TS separating *syn*-CH_3CHOO and VHP or followed the homolytic cleavage from VHP to dissociated products using an accurate full-dimensional PES [43] or the step between *syn*-CH_3CHOO and VHP involving the 1,4-hydrogen shift reaction was investigated based on statistical RRKM theory including tunneling corrections. [42] Both approaches do not provide rates for the overall *syn*-$CH_3CHOO \rightarrow CH_2CHOOH \rightarrow CH_2CHO+OH$ reaction which is the topic of the present work. For a more atomistically resolved understanding of the entire reaction pathway *syn*-$CH_3CHOO \rightarrow CH_2CHOOH \rightarrow CH_2CHO+OH$, a statistically significant number of reactive trajectories based on full-dimensional reactive potential energy surfaces is required

and analyzed which is done in the present work. This is the first such attempt to determine decomposition rates of CH_3CHOO including all degrees of freedom and all relevant states along the reaction pathway. To follow experimental preparations, energy to the system is provided by vibrationally exciting CH_3CHOO. In addition, thermal dissociation is investigated as this is the relevant process in the atmosphere.

The work is structured as follows. First, the computational methods including the construction of the PESs is described. This is followed by validation of the PESs and the reactive MD simulations to determine energy-dependent dissociation rates and the reaction time distributions. Next, the results are discussed in the context of experiments and previous work on the reaction rates, and finally, conclusions are drawn.

3.1.3 Computational Methods

This section presents the different computational methods employed. Two representations of the full-dimensional, reactive PES were developed based on two quantum chemical treatments. They included MS-ARMD[131] and PhysNet[143] for the representation and the MP2/6-311++G(2d,2p) and MP2/aug-cc-pVTZ levels of theory for the reference calculations which were carried out using Gaussian 09[133] and MOLPRO.[134] This allows validation of the representations and direct assessment of the level of theory. Also, the MD simulations are described.

Reactive Force Fields

MS-ARMD is a computationally efficient, energy-conserving surface crossing algorithm to investigate chemical reactions based on empirical force fields.[131] It uses parameterized force fields for the reactant and product state and GAPOs to describe the surface crossing region, i.e., the region around the transition state (TS). The parametrized force fields for the reactant and product complexes are iteratively fit to reference data from electronic structure calculations.

The initial parameters for the reactant (*syn*-CH_3CHOO, methyl-substituted CI) and product (CH_2CHOOH, VHP) were taken from SwissParam.[132] First, representative

structures for CH$_3$CHOO (1612) and CH$_2$CHOOH (1650) were sampled from 500 ps MD simulations at 300 K. For these structures, energies were determined at the MP2/6-311++G(2d,2p) level of theory. These were fit separately to parametrized force fields for the reactant and the product using a downhill simplex[135] algorithm. Parametrization starts with a set of 100 structures for CH$_3$CHOO and CH$_2$CHOOH and first fitting for the two states were carried out which was then followed by further MD simulations using this improved set of parameters from which another 200 structures was extracted and included in the fit. Several rounds of refinements were done until the root mean squared deviation for the final set between the target (*ab initio*) and the fitted energies for CH$_3$CHOO and CH$_2$CHOOH reached 1.1 kcal/mol and 1.2 kcal/mol, respectively. The bonds involved in bond breaking and bond formation (the three C–H, the O–O, and the O–H bonds) were described by Morse potentials and the charges from *ab initio* calculations at the MP2/6-311++G(2d,2p) level of theory. Generalized Lennard-Jones potential is included and fitted, to represent the van der Waals interaction between atoms(between H$_A$ and O$_B$ in reactant FF and between H$_A$ and C$_A$ in product FF) in the reactive region.

The reactant and product force fields are connected by using GAPO-functions to yield a continuous connection along the reaction path. For this, the IRC between reactant and product was also determined and included in the fitting. To parametrize the adiabatic barrier, genetic algorithm was used to parametrize the GAPO[131] functions and to reproduce the energies along the reaction path. The TS barrier for *syn*-CH$_3$CHOO is 16.02 kcal/mol at MP2/6-311++G(2d,2p) level of theory.

Neural Network

As an alternative to MS-ARMD, a NN based on the PhysNet architecture[143] was trained, which is designed to construct PES after learning molecular properties like energy, forces, charges and dipole moments from *ab initio* reference data. Here, MP2/aug-cc-pVTZ level of theory[103] is used for generation of *ab initio* reference data and the energies, forces and dipole moments that are calculated using MOLPRO software package.
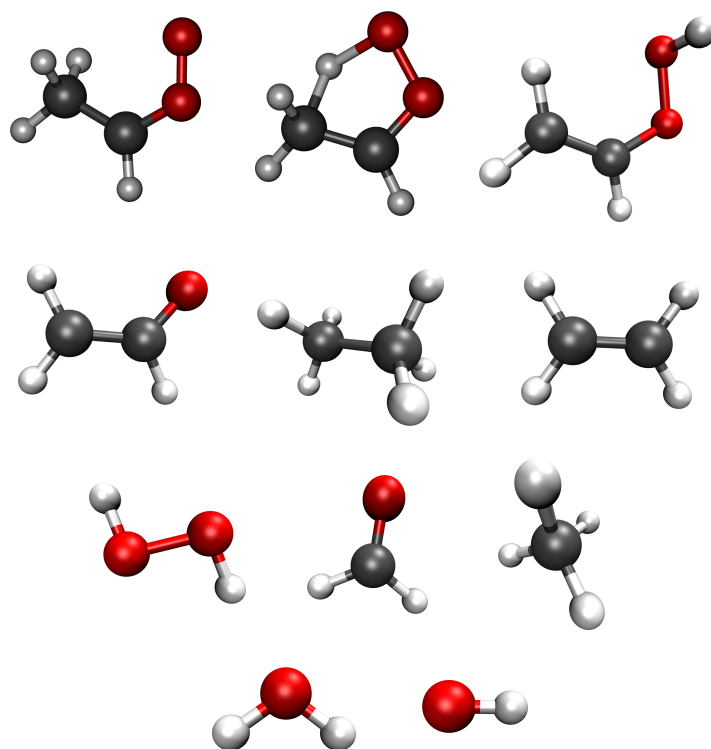
Figure 3.2: Different structures and amons[173] used for training the neural network representation. The first row contains the reactant, transition, product structures and subsequent rows contain their fragmented structures.

Following the "amons" approach[173], a set of molecules (Figure 3.2) are included covering a range of fragmentation reactant, products, stable intermediates and van der Waals complexes in the dataset. To obtain a broad range of molecular geometries, MD simulations were started from optimized geometries which are propagated using Langevin dynamics at 1000 K with a time step of 0.1 fs. Then, the data set is extended based on adaptive sampling and normal mode sampling.[174] The final dataset used for training contains 105403 structures, which was then split into training (84322), validation (10540) and test set (10541). The TS barrier for *syn*-CH_3CHOO is 14.9 kcal/mol at MP2/aug-cc-pVTZ level of theory. It is noted that MS-ARMD requires close to two orders of magnitude fewer points than the NN-based PES.

Molecular Dynamics Simulations

The MD simulations for MS-ARMD were carried out with a suitably modified version of CHARMM.[124] Simulations based on PhysNet were run with the Atomic Simulation Environment (ASE).[175] All MD simulations were carried out in the *NVE* ensemble (i.e., without a thermal bath) and in the gas phase (i.e., in a collision-free environment), and were started from the minimized geometry of *syn*-CH_3CHOO using the respective PES, i.e., the MS-ARMD PES for reactive MD simulations and the PhysNet PES for all NN-based simulations. The time step in all simulations was $\Delta t = 0.1$ fs and most trajectories were propagated for 1 ns except for several 10^3 long-time trajectories which were run for 5 ns and 15 ns, respectively.

For the thermal simulations with MS-ARMD *syn*-CH_3CHOO was heated to the desired temperature (including 865, 877, 886, 902, 925, 932, 956 K) and equilibrated for 50 ps followed by free *NVE* dynamics. The simulations with ASE were initialized with momenta assigned from a Maxwell-Boltzmann distribution, followed by Langevin dynamics for 100 ps with a small friction term (0.01 atomic units corresponding to 2.41×10^{-7} ps) to keep the energy of the system constant and equilibration of temperature for long time during the dynamics. In addition, *NVE* simulations were also carried out with ASE for direct comparison.

The vibrationally excited (vibEX) simulations using MS-ARMD were heated to 300 K and equilibrated for 50 ps followed by free *NVE* dynamics for 1 ns. From 7 independent simulations coordinates and velocities were saved regularly to obtain 70000 initial conditions for each of the excitation energies. Then a non-equilibrium state is prepared by scaling the instantaneous velocity vector along the CH stretch mode. 10000 independent trajectories for each excitation energy were run with a simulation time of 1 ns for D_e^{OO} = 31.5 kcal/mol and 100 ps for D_e^{OO} = 23.5 kcal/mol. For the vibEX simulations with ASE, the momenta were assigned from a Maxwell-Boltzmann distribution at 300 K to the geometry optimized structure of CH₃CHOO, followed by 50 ps of free dynamics. The vibrationally excited, non-equilibrium state is prepared by scaling the instantaneous velocity vector along the CH mode to provide energies commensurate with the excitation energies used in the experiments,[42] i.e., for excitation to [5603, 5709, 5748, 5818, 5951, 5984, 6082] cm⁻¹, corresponding to energies of 16.0 to 17.3 kcal/mol. For each excitation energy 10000 independent trajectories were propagated for 100 ps.

The type of initial conditions from which the present simulations were started do not include quantization and have been used earlier[43] for studying the decomposition of VHP to CH₂CHO+OH and also in previous work on the decomposition reactions involving sulfuric acid and halogenated variants of it.[176–178] This differs somewhat from the initial conditions for the simulations that were initialized from the transition state separating *syn*-CH₃CHOO and VHP for which normal mode sampling at the TS was used together with translational energy along the imaginary mode.[43] Most rigorous among the different possibilities are quasiclassical trajectory simulations which are started from semiclassical quantized initial conditions[179] and often used for bimolecular reactions such as A+BC→AB+C.[180,181] However, their use for unimolecular reactions is problematic because the zero point vibrational energy of the reactant state is larger than the barrier for formation of VHP which leads to unrealistic rates, as was also recently discussed.[182]

In addition to the explicit MD simulations, the minimum energy path was also determined.[183] This path connecting reactant and product geometry, passing through the exact transition state with zero excess energy was calculated by assigning momenta

along the normal mode vector. Then, the MD simulation from transition state to both reactant and product channel using the PhysNet PES.

3.1.4 Results

Quality of the Potential Energy Surfaces

The quality of the MS-ARMD and PhysNet representations of the reactive PESs is reported in Figures 3.1B, C. For MS-ARMD the fitted PESs for the reactant (blue, CH_3CHOO) and the product (green, VHP) states have overall root mean squared errors (RMSEs) of 1.1 and 1.2 kcal/mol, typical for such an approach.[177,184,185] The IRC closely follows the reference MP2 calculations, see inset Figure 3.1B, which underlines the quality of the reactive PES. For OH-elimination a conventional Morse fit to the reference MP2 calculations was used with $r_e^{OO} = 1.45$ Å, $D_e^{OO} = 31.5$ kcal/mol, and $\beta = 2.3$ Å^{-1}. However, because O–O bond breaking may involve multireference character[43,45] simulations with a value of $D_e^{OO} = 23.5$ kcal/mol were also carried out, consistent with findings from CCSD(T)-F12b/CASPT2 calculations.[43]

The quality of the PhysNet representation of the global reactive PES is reported in Figure 3.1C. Here, the mean average error on the test set (~10000 structures not used for training or validation) is 0.02 kcal/mol with a RMSE of 0.19 kcal/-mol and $R^2 = 1$–10^{-7} . Again, this performance is in line with previous work.[165,186,187]

Reaction Rates from VibEX and Thermal Simulations

As a first exploration of the PES the minimum dynamic path for the 1,4 hydrogen shift reaction was determined on the PhysNet PES, see Figure 3.3. Starting from the initial structure "A", this reaction passes through a five-membered ring (structure "C") before formation of VHP (structure "E").

An illustrative example for a reactive trajectory from MS-ARMD simulations is reported in Figure 3.5A. Here, the C_AH_A, O_BH_A, and O_AO_B time series are shown (for labeling see Figure 3.1A) which are directly relevant to the reaction. Initially,
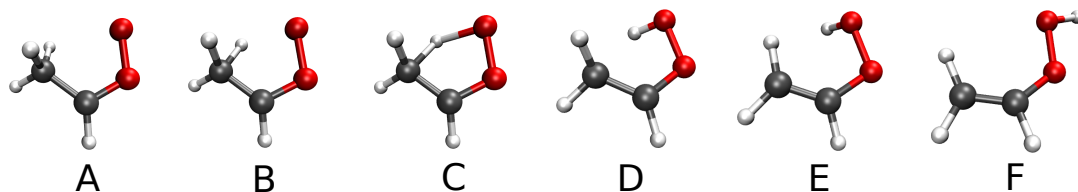
Figure 3.3: Minimum Dynamic Path[183] on the PhysNet PES for the 1,4-hydrogen shift from carbon to oxygen in CI. The reaction proceeds through a five membered transition state for which r_{CH} = 1.30 Å and r_{OH} = 1.41 Å. From A→ F the C–C distance change is reported in Figure 3.4. The C–O distance changes from 1.27 to 1.37 Å and the O–O distance changes from 1.32 to 1.45 Å.
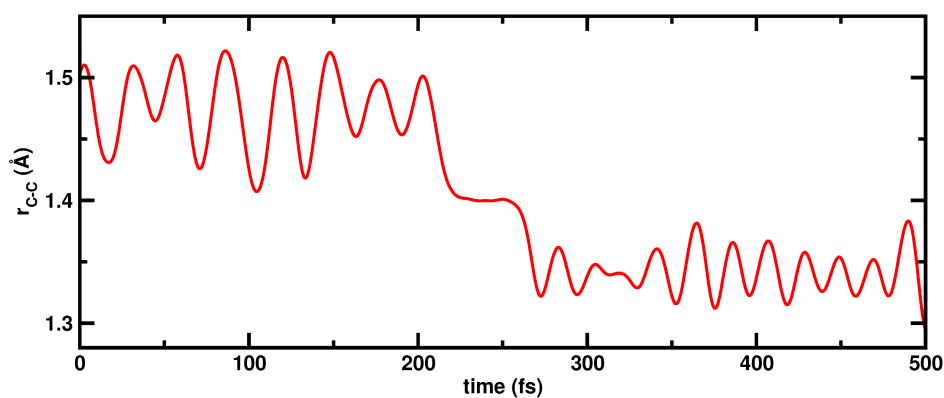

Figure 3.4: C–C distance along the minimum dynamic path as a function of time from CH₃CHOO to CH₂CHOOH.

the C_AH_A separation (black trace in Figure 3.5A) fluctuates around 1.12 Å which is close to the equilibrium bond length, and the O_BH_A separation is large (ranging from 2 to 4 Å) as the system is in its reactant state. Such large variations are elicited by the CH_3 rotation. At $t \approx 0.5$ ns the 1,4-hydrogen shift reaction occurs which increases the C_AH_A and O_AO_B separations and shortens the O_BH_A bond due to bond formation. For the next ~0.3 ns the system is in its VHP state before OH elimination takes place at $t \approx 0.79$ ns following breaking of the O_AO_B bond.
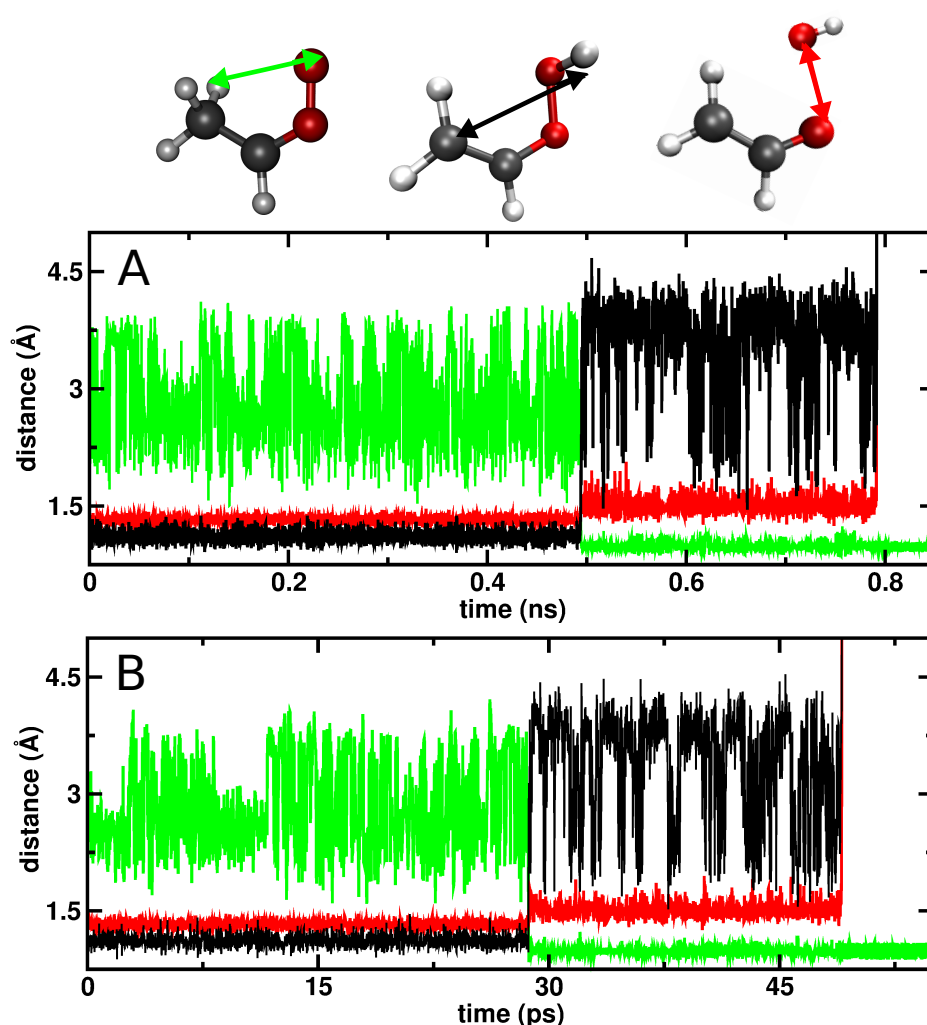
Figure 3.5: Time series for different distances for the reaction syn-CH₃CHOO → VHP → CH₂CHO+OH. Panel A: using MS-ARMD simulations at $t \approx 0.5$ ns VHP forms and at $t \approx 0.79$ ns OH elimination takes place. Panel B: using PhysNet at $t \approx 28$ ps VHP forms and at $t \approx 47$ ps OH elimination takes place. Black, red, and green solid lines correspond to the C_AH_A, O_AO_B, and O_BH_A distances, respectively.
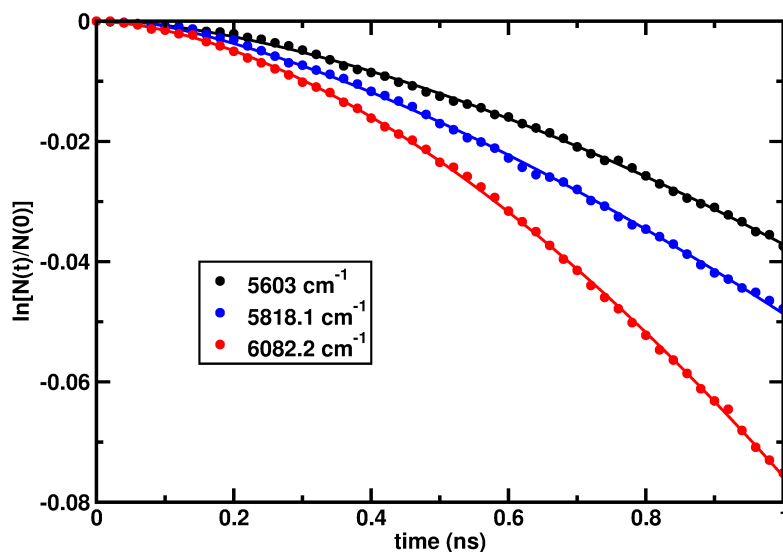
Figure 3.6: Fit of $\ln[N(t)/N(0)]$ for $D_e = 31.5$ kcal/mol from vibEX simulations at energies of 5603 cm^{-1}, 5818.1 cm^{-1} and 6082.2 cm^{-1} to a stretched exponential decay.

A similar trajectory, run with ASE and the PhysNet representation, is shown in Figure 3.5B. Again, CH₃ rotation is found. As for the MS-ARMD simulations, for VHP (between $t \approx 28$ and 47 ps) the position of the OH-group switches between *syn-* and *anti-*conformers, respectively. These lead to long and short C_A–$H_A O_B$ separations (black trace).

In laboratory-based experiments[42,43] the reaction path following *syn-*CH₃CHOO $\rightarrow$ VHP $\rightarrow$ CH₂CHO+OH is initiated by excitation of the CH stretch vibration of the terminal CH₃ group with energies ranging from ~5600 cm^{-1} to ~6000 cm^{-1}. This corresponds to about 2 quanta in the methyl-CH stretch mode. In the simulations, excitation of this mode was accomplished by scaling the velocities[176,177] along the CH-local mode and the starting ensemble was generated from equilibrium simulations of *syn-*CH₃CHOO at specified temperatures. It should, however, be noted that "temperature" as determined from the equivalence of kinetic energy and $3/2Nk_BT$, as is usually done in MD simulations,[188] should not be directly compared with experimentally reported temperatures (e.g., 10 K rotational temperature in Ref.[42]).

Simulations were run using excitation energies as those reported from experiments, i.e. [5603, 5709, 5748, 5818, 5951, 5984, 6082] cm^{-1}. For each energy $N_{tot} = 10000$ individual simulations were carried out. The ensuing rates were determined from
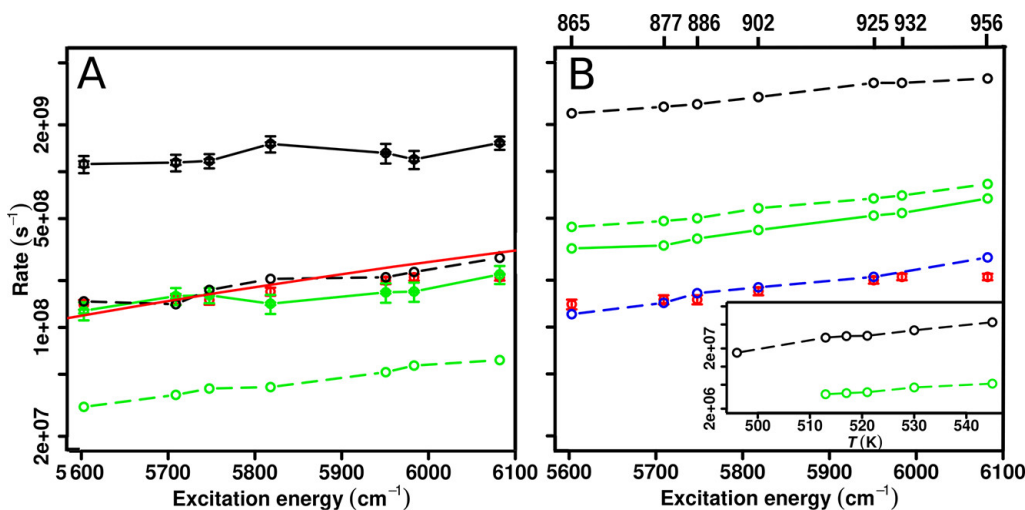
Figure 3.7: OH formation rates from vibEX (panel A) and thermal (panel B) simulations. Panel A: Rates from MS-ARMD simulations following excitation of the C_AH_A stretch vibration (vibEX) at different excitation frequencies. The green and black symbols are for $D_e^{OO} = 31.5$ kcal/mol (consistent with the present and earlier[45] MP2 calculations), and $D_e^{OO} = 23.5$ kcal/mol, respectively. Filled symbols are rates based on fitting $N(t)/N_{tot}$ to stretched exponentials and open symbols from using a single exponential, respectively, with separate error bars from bootstrapping and for the parameters. The red points with error bars are the experimental results and the solid red line is the rate from RRKM calculations including tunneling. Panel B: Rates from thermal simulations using MS-ARMD (identical color code as in panel A), and from PhysNet (dashed blue line). The top y axis gives the temperature T as determined from $T = 2E_{kin}/(3Nk_B)$. The inset reports results from simulations at somewhat lower temperatures. Reproduced from Ref.[42]

following the number $N(t) = 1 - N_{OH}$, i.e., counting those that had not reached the OH-elimination product, see Table 3.1 and 3.2 and fitting $N(t)$ to either a single ($\sim \exp(-kt)$) or a stretched-exponential ($\exp(-kt)^\gamma$) dependence (see Figure 3.6).[189] The rates k as a function of excitation energy are reported in Figure 3.7A together with those from experiment (red symbols).[42]

MS-ARMD simulations from an ensemble generated at 300 K were carried out for two values of the O–O dissociation energy, see above. Rates from simulations with $D_e^{OO} = 31.5$ kcal/mol (green) and $D_e^{OO} = 23.5$ kcal/mol (black) are shown in Figure 3.7A. Depending on the value of D_e^{OO} the rates differ by a factor of ~10 but following a similar energy dependence, consistent with that observed in the experiments. Bootstrapping is used to determine statistical errors by randomly sampling 8000
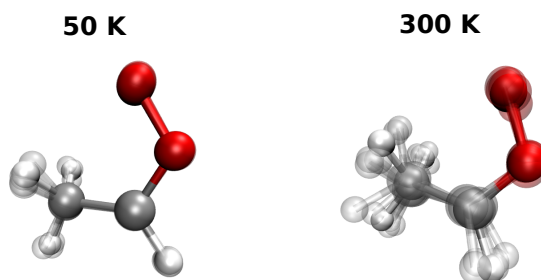
Figure 3.8: Superposition of the structural ensemble generated at 50 K (left) and 300 K (right).

trajectories 30 times from all 10000 trajectories at each energy. The error bars due to parameter uncertainties in the fitting of $\ln N(t)/N(0)$ vs time are a factor of 4 to 10 smaller than those from bootstrapping.

To assess the sensitivity of the results to simulation length and the initial preparation of the system, additional simulations were carried out. For Set1, several 1000 vibEX simulations, each 5 ns in length at different excitation energies were run by sampling from ensembles generated at 50 K and 300 K using MS-ARMD and $D_e^{OO} = 23.5$ kcal/mol, respectively. A direct comparison of the structural ensemble generated at the two temperatures is reported in Figure 3.8. At 50 K there are 15 out of 2000 trajectories that lead to VHP from excitation with 5603 cm^{-1} out of which 7 lead to OH-production, i.e., 15/7/2000 for VHP/OH/N_{tot}. This changes to 23/4/1000 and 6/3/400 for excitation at 5747 and 5818 cm^{-1}, respectively (not shown graphically). At 300 K the 5 ns simulations yield 1994/1982/2000, 997/991/1000, and 399/395/400, i.e., most of the simulations show formation of VHP and OH elimination, see Figure 3.9. Hence, with the energy content at 300 K most trajectories form VHP and undergo OH elimination on the 5 ns time scale. At 50 K only half of the 5 ns trajectories that form VHP also lead to OH formation and the reaction probability is reduced by about 2 orders of magnitude.

Up to this point, the internal energy of the reactant was that of the ensemble at which the initial conditions were generated, i.e., at 50 K or 300 K, respectively. To probe the influence of the internal energy on 1,4-hydrogen shift and OH-formation, Set2 (internally "cold") was generated: after vibrational excitation, the velocities of all atoms except for those involved in the CH-stretch were explicitly set to zero
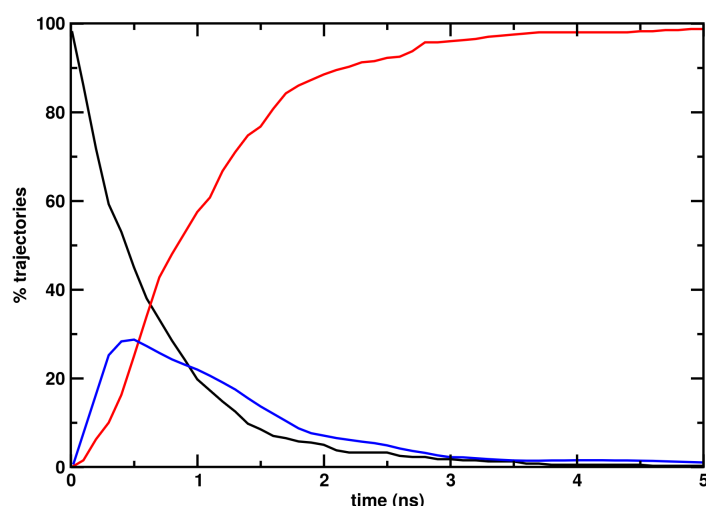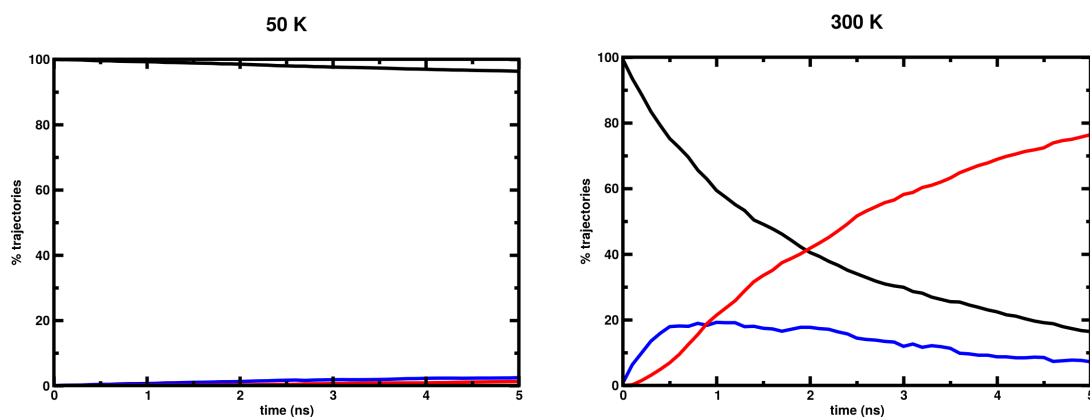
Figure 3.9: Change in concentration (in %) for vibEX CIs (black), VHP (blue) and CH₂COH+OH (red) as a function of time using MS-ARMD and D_e^{OO} = 23.5 kcal/mol from 5 ns simulations at 300 K (Set1).

and 5 ns simulations were carried out. For conformations generated at 300 K there are 655/551/1000 and 836/764/1000 H-transfers and OH-production trajectories for excitation with 5603 and 6082 cm⁻¹ which changes to 47/10/5000 and 186/62/5000 at 50 K, see Figure 3.10. Set 2 establishes that for structures sampled at 50 K only about 20% to 30% of the trajectories that exhibit H-transfer also proceed to OH elimination on the 5 ns time scale, which is the time resolution of the experiment.[43] Hence, for internally cold reactants VHP appears to become a bottleneck for OH-formation. Increasing the simulation time to 15 ns leads to 27/15/1200 and 202/136/2000 events for excitation at the two energies. Hence, with a simulation time of 15 ns a considerably larger percentage of VHP (50 % to 70 %) shows OH-elimination even for samples from cold initial distributions. Therefore, reactivity on the nanosecond time scale is also found for considerably lower internal energies (only CH-stretch excited), which is consistent with experiment.[42] For converged rates an even larger number of trajectories would be required, though.

In the atmosphere, vibrational excitation is not the likely primary mechanism by which *syn*-CH₃CHOO is energized. Rather, the ozonolysis reaction is expected[190] to lead to an excited parent molecule (*syn*-CH₃CHOO) which subsequently decays following the 1,4-hydrogen shift and O–O dissociation reactions.[191–193] Hence, "thermal" simulations were also carried out by heating the reactant to average

Figure 3.10: Change in concentration (in %) for vibEX CIs (black), VHP (blue) and CH₂COH+OH (red) as a function of time using MS-ARMD and D_e^{OO} = 23.5 kcal/mol from 5 ns simulations at 50 K (left) and 300 K (right) for "cold" preparation (Set2). Here the velocities of all atoms except those vibrationally excited were set to zero explicitly to mimic an internally cold, CH-vibrationally excited reactant.

temperatures between 865 and 956 K, commensurate with internal energies of ~5500 to 6000 cm⁻¹. Again, the equivalence between "temperature" and "kinetic energy" is rather qualitative. For this reason, additional simulations at somewhat lower temperatures (500 to 550 K) were also carried out and analyzed.

The rates were determined by following $N(t)$, see Tables 3.3, 3.4 and 3.5. For these "thermal" simulations with D_e^{OO} = 31.5 kcal/mol, $N(t)$ is well represented by a single exponential decay for times $t < 200$ ps whereas a stretched exponential is again required for longer simulation times, see Figure 3.11. Such information is of particular relevance in the context of *ab initio* MD simulations which can usually only be carried out on the multiple ~10 ps time scale for sufficiently high-level treatments of the electronic structure.[194]

Thermal rates from ~9000 independent simulations using the MS-ARMD (solid green and black lines) and PhysNet (dashed blue line) representations are reported in Figure 3.7B. MS-ARMD simulations were run for D_e^{OO} = 31.5 and 23.5 kcal/mol and yielded rates ranging from 3×10^8 to 5×10^9 s⁻¹. The temperature dependence is moderate and follows that found from experiments using vibrational excitation although the magnitude of the computed rate is consistently higher by a factor of 5
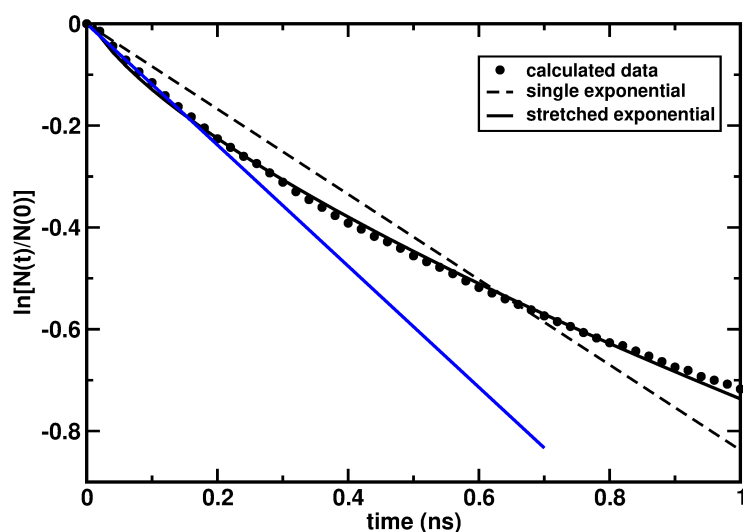
Figure 3.11: Fit of $\ln[N(t)/N(0)]$ for D_e = 31.5 kcal/mol from thermal simulations at 956 K where blue solid line is to show that till 200 ps the data follows a linear fit.

to 10. Thermal rates from simulations using PhysNet show the same temperature dependence as those observed in experiments and from MS-ARMD simulations but with an amplitude that is closer to that from experiment ($\sim 10^8$ s^{-1}). Exploring effects due to a change in D_e^{OO} is not easily accomplished using PhysNet because the underlying data set used in the fitting would need to be modified. An additional 1000 simulations in the *NVE* ensemble were carried out after heating to 865 K for which 10 trajectories show OH elimination which compares with 11 trajectories from the Langevin simulations at the same temperature. This is consistent with the expectation that Langevin simulations with a small friction coefficient are close to microcanonical simulations.

MS-ARMD simulations were also carried out for ensembles generated at lower temperatures, see inset Figure 3.7. As expected, the rates decrease by about an order of magnitude but remain consistent with experiment, exhibiting a comparable temperature dependence.

3.1.5 Discussion and Conclusions

The present work follows OH-elimination starting from conformational ensembles at 50 K and 300 K of *syn*-CH_3CHOO. The energy to induce hydrogen transfer is provided by vibrationally exciting the CH-stretch as is done in laser experiments.[43] From a statistically significant number of trajectories, the rates for OH-elimination are determined. It is found that the energy-dependence of the rates follows that reported from the experiments (see Figure 3.7), and by providing two quanta in the CH-stretch tunneling is not necessary for H-transfer to occur. This does, however, not preclude that tunneling happens and further accelerates the reaction.

With a statistically significant number of reactive trajectories ($\sim 10^4$) it is also possible to carry out additional analyses. Distributions of H-transfer and OH-formation times from vibEX and thermal MS-ARMD simulations for $D_e^{OO} = 31.5$ kcal/mol are reported in Figure 3.12. Reactions initiated from vibrational excitation have a comparatively flat reaction time distribution up to ~ 0.2 ns for the 1,4 H-shift reaction, red histogram in Figure 3.12A. The reaction time is defined by the time interval between the point of vibrational excitation and VHP formation for which a geometrical criterion was used (O–H separation < 1.6 Å and O–O separation < 2.8 Å). For longer simulation times (> 0.2 ns) the reaction probability decays towards zero. Decay to $CH_2COH+OH$ in the vibEX simulations is delayed and the number of trajectories reaching the final state compared with those that complete the first step (1,4 H-shift) depends on time. For lower excitation energy (top panel in Figure 3.12A) appreciable OH elimination only starts after ~ 0.2 ns and does not reach more than 10 % of the amount of VHP formed on the 1 ns time scale. For higher excitation energies, the onset of OH elimination shifts to shorter times and the amount of OH formed increases in proportion. The fraction of VHP intermediates formed ranges from 74 % to 87 % compared with a fraction of 3 % and 7 % for the amount of OH formed from vibEX simulations (5603 cm^{-1} to 6082.2 cm^{-1}).

The situation changes considerably for the thermal reaction, see Figure 3.12B. Here, the reaction time distributions for the 1,4 H-shift reaction approach an exponential decay with simulation time for all temperatures considered. Similarly, the amount
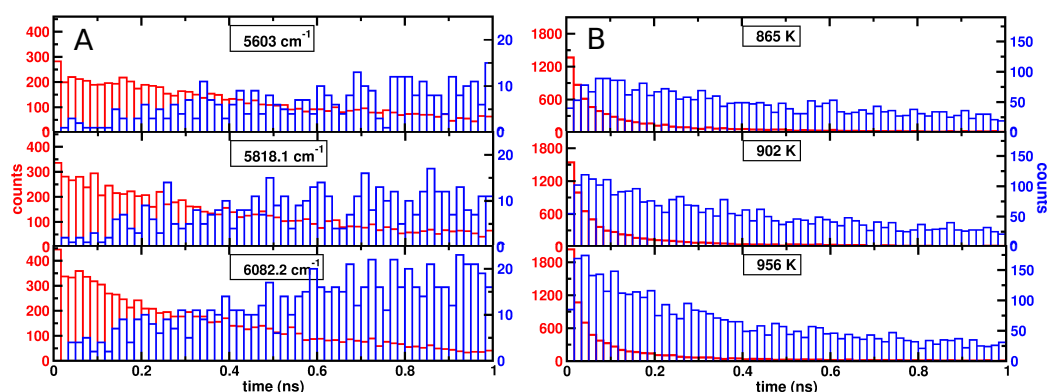
Figure 3.12: Distribution of reaction times for 1,4 H-shift [CH₃CHOO → VHP, red] and OH-formation [CH₃CHOO → CH₂COH+OH, blue] from vibEX (panel A) and thermal (panel B) simulations at different energies and temperatures using MS-ARMD with D_e^{OO} = 31.5 kcal/mol. Note that scales for the 1,4 H-shift distribution times are along the left-hand y axis and those for the total reaction time are along the right-hand y axis and that the magnitudes differ between panels A and B.

of OH formed follows this. With increasing temperature (internal kinetic energy) the amount of VHP formed increases by about 50 % between 865 and 956 K and the amount of OH formed by almost a factor of 2. The fraction of VHP intermediates formed ranges from 73 % to 76 % compared with a fraction of 28 % and 38 % for the amount of OH formed.

The results from Figure 3.12 indicate that the VHP intermediate has a broad distribution of lifetimes, see Figure 3.13. As the probability to form VHP in the vibEX simulations has not reached zero within 1 ns (red distribution Figure 3.12A), VHP can also be generated on longer time scales from trajectories that have undergone more extensive internal vibrational energy redistribution (IVR) which is consistent with the results presented earlier for "Set2". For the thermal trajectories the probability to form VHP has decayed to very low levels within 1 ns (Figure 3.12B). Thus, the VHP lifetime distributions in Figure 3.13 are expected to be close-to-converged.

Formation of the VHP intermediate follows different time traces depending on (a) the dissociation energy D_e^{OO} used in the simulations and (b) whether energy to *syn*-CH₃CHOO was provided by methyl-CH vibrational excitation or by thermal preparation. Figure 3.14A shows that for D_e^{OO} = 31.5 kcal/mol VHP concentration
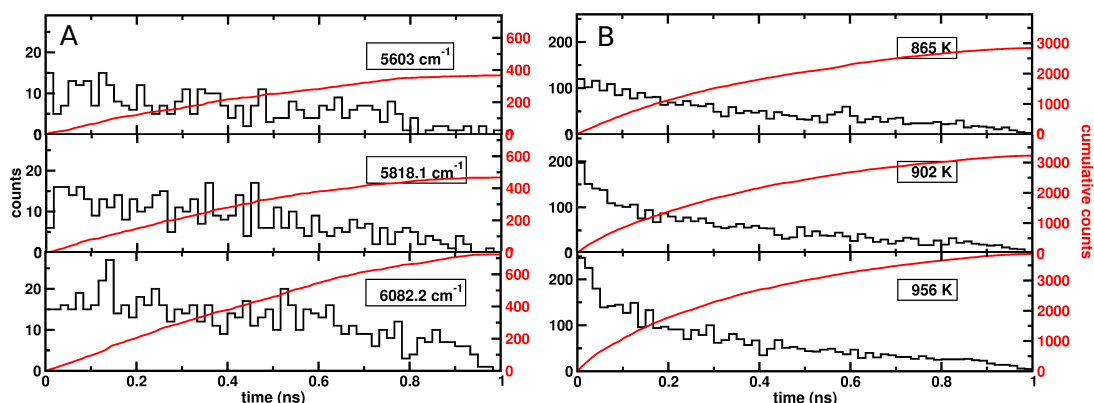
Figure 3.13: Distribution of VHP lifetimes before OH formation from vibEX (panel A) and thermal (panel B) using MS-ARMD and D_e^{OO} = 31.5 kcal/mol. Histograms (black) and cumulative (red line) events are reported in the same panel with corresponding labels along the $y-$axes. Note the different scaling of the $y-$axes in panels A and B. The kinetic temperatures used to label the thermal simulations (panel B) are close to the excitation energies used in the vibEX simulations (panel A).
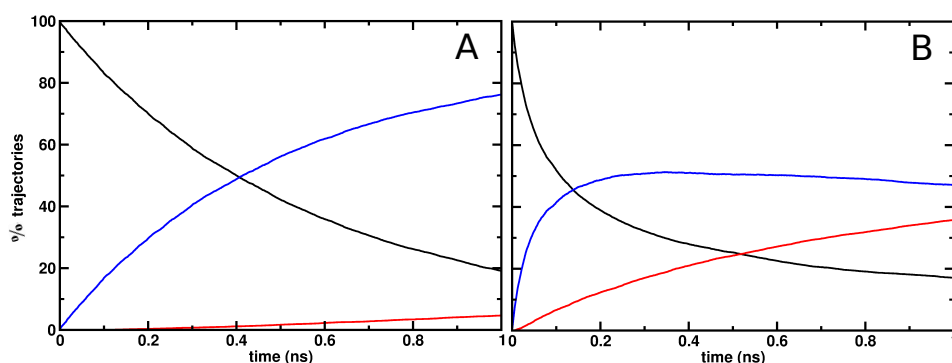
Figure 3.14: Change in concentration (in %) of CI, VHP and CH₂COH+OH as a function of time using MS-ARMD and D_e^{OO} = 31.5 kcal/mol. Panel A: from vibEX simulations at 5818.1 cm^{-1}. Panel B: from thermal simulations at 902 K for $N_{Criegee}$ (black), N_{VHP} (blue), and N_{OH} (red), respectively. For vibEX the rate limiting step is the OH-elimination step whereas for thermal excitation none of the two steps is a clear reaction bottleneck. With comparable energy content in the reactant, thermal simulations yield close to one order of magnitude more OH product than vibrational excitation of the methyl-CH-stretch. For simulations with D_e^{OO} = 23.5 kcal/mol see Figure 3.9.
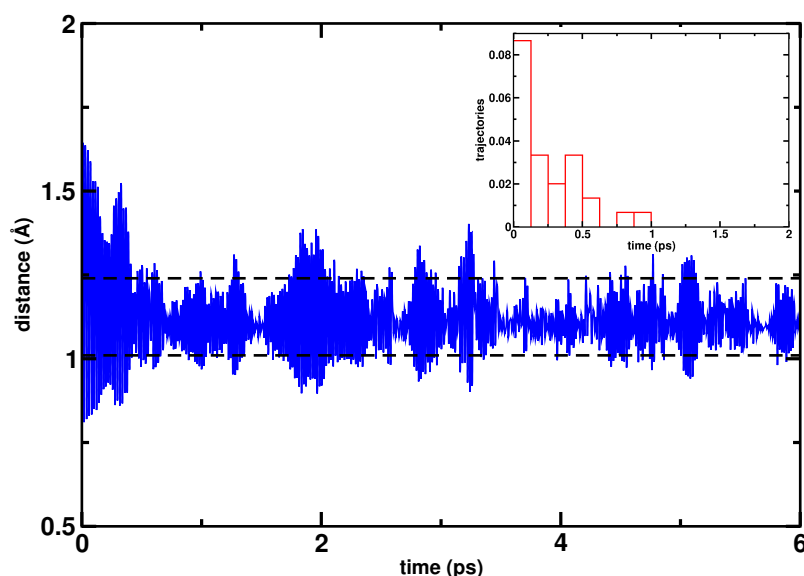
**Figure 3.15:** Time series of the CH bond for *syn*-CH$_3$CHOO from one of the vibEX simulations. The energy along the CH bond is found to remain for less than 1 ps in this mode. Inset: Distribution of CH-relaxation times from 30 trajectories. The excited CH stretch is considered "relaxed" when the CH bond length reaches 1.64 Å for the first time. The two dashed lines at 1.01 and 1.24 Å are the turning points of the CH bond at $v = 0$.

from vibEX monotonically increases (blue trace) on the 1 ns time scale and reaches close to 80 % but only for 5 % OH elimination has occurred. For vibEX simulations with $D_e^{OO} = 23.5$ kcal/mol close to 100 % of the trajectories lead to OH-elimination on the 3 ns time scale, see Figure 3.9. Thermal excitation, representative of the initial state of the reactant after ozonolysis of 2-butene, with $D_e^{OO} = 31.5$ kcal/mol leads to ~50% population of VHP within 0.3 ns. Over the same time close to 20% of the trajectories already show OH-elimination and within 1 ns more than 30 % of *syn*-CH$_3$CHOO have completed OH-elimination. OH production on the 1 ns time scale is about one order of magnitude larger with thermal compared to vibEX preparation of the reactant. This difference between thermal and vibEX simulations is explained by rapid energy redistribution from the excited CH stretch (see Figure 3.15) into the remaining vibrational modes which requires time to channel the energy back into the progression coordinate. For thermal simulations with $D_e^{OO} = 23.5$ kcal/mol a considerably higher yield than 30 % on the 1 ns time scale is expected.

Conversely, in thermal simulations all modes are already similarly energized which apparently speeds up hydrogen transfer. Experimentally, OH yields from 2-butene ozonolysis cover a broad range from 24 % to 64 %, depending on the experimental and detection method used.[39,163,195–200] Here, it should be noted that ozonolysis of 2-butene yields both, *syn-* and *anti*-CH₃CHOO of which, however, only the *syn*-isomer has been found to lead to OH-elimination whereas the simulations report results for *syn*-CH₃CHOO only. Hence, the difference with a yield of ~30 % from the simulations is also in the fact that within 1 ns of simulation time decomposition to products is not complete. Thus, the yield found from the present work are a lower limit to the true value. For a lower O–O dissociation energy of $D_e^{OO} = 23.5$ kcal/mol OH-production is almost complete within 5 ns at 300 K, see Figure 3.9, whereas for 50 K and 300 K and "cold" initial conditions OH production starts but is clearly slower (Figure 3.10) and increases with longer simulation times of 15 ns. From experiments in the jet no yields are available to the best of our knowledge.
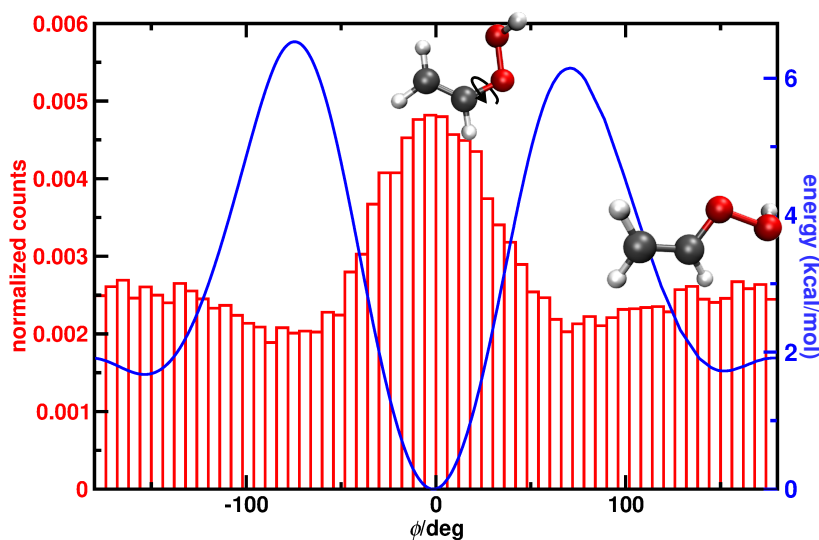
Figure 3.16: The CCOO dihedral angle distribution $P(\phi)$ for CH₂CHOOH (VHP) from thermal simulations at 902 K using ASE. The structure with $\phi = 0°$ corresponds to the -OH group in its *syn-* conformation whereas $\phi = \pm180°$ is the *anti* conformer. The barriers between 0° and 180° and 0° to −180° are 6.1 kcal/mol and 6.5 kcal/mol, respectively, at the MP2/aug-cc-pVTZ level of theory. The ϕ population (red) from the simulations is consistent with the PES as *syn*-VHP is more stable than *anti*-VHP.
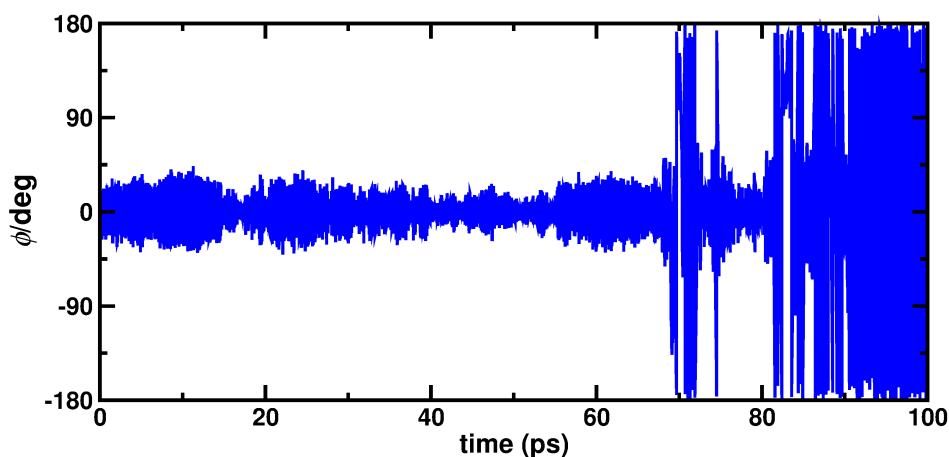
Figure 3.17: Time series $\phi(t)$ from a thermal trajectory at 902 K using ASE. Until $t =$ 67 ps the system is in the *syn*-CH$_3$CHOO state with $\phi \sim 0$. Between $t =$ 67 and 90 ps it is in VHP. Here, $\phi \in [-180 \ldots + 180]$, but most time is spent in the syn-form $\phi \in [-60 \ldots + 60]$. At $t = 90$ ps OH dissociation takes place and ϕ looses its meaning.

The extended lifetimes of VHP prior to OH elimination can also be rationalized from considering the CCOO dihedral angle ϕ, see Figure 3.16. Coming from CH$_3$CHOO the VHP intermediate is formed with a considerable amount of internal energy which is sufficient to overcome the *syn/anti* barrier height of ~6 kcal/mol, see also Figure 3.17. The population distribution $P(\phi)$ for VHP from simulations with PhysNet simulations at 902 K sample all values of ϕ, consistent with the topology of the potential energy curve (blue trace in Figure 3.16). Analysis of the simulations reveals that OH-elimination occurs from both, *syn-* and *anti*-VHP. This is consistent with the energetic preference for *syn*-VHP but at variance with findings from statistical modeling which reported that most of the VHP decomposes via the *anti*-pathway even if the *syn* form is thermodynamically more stable.[169]

Finally, bond length distributions (Figures 3.18 and 3.19) show that there is little difference between vibEX and thermal simulations (blue compared with red traces). Thermal distributions from MS-ARMD simulations using only the reactant force field at 300 K (black) are considerably more peaked around the equilibrium values compared with those at 760 K (red). If the 1,4 H-shift reaction is possible (by running the MS-ARMD simulations with the reactive force field), a few bond length distributions change appreciably (green traces in Figure 3.18). Excitation of the
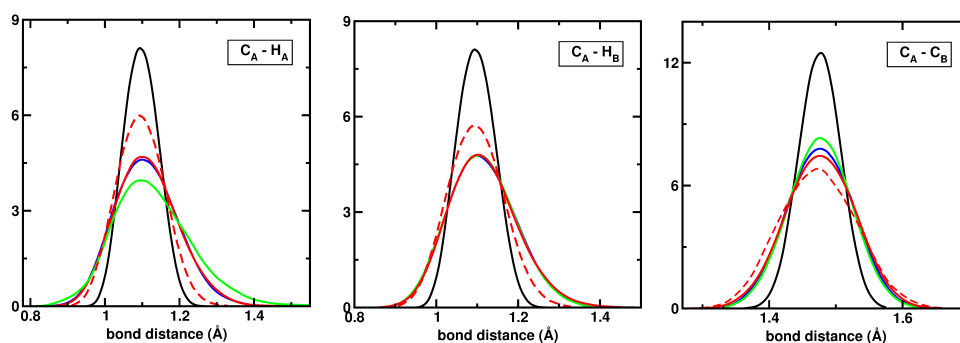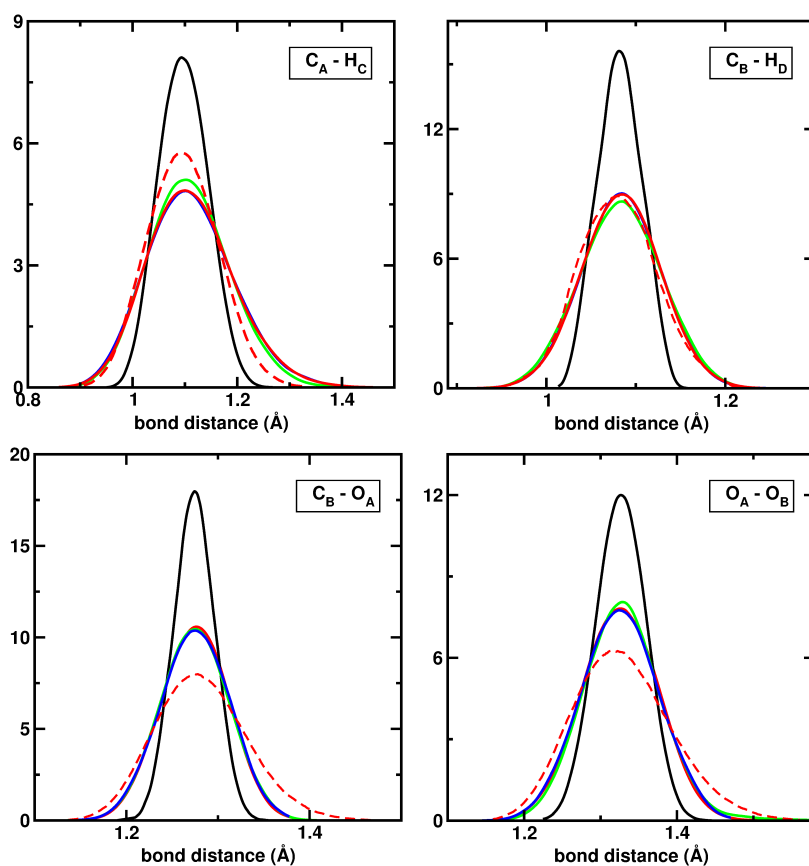
Figure 3.18: Bond distance distributions of C_AH_A, C_AH_B and C_AC_B for *syn*-CH₃CHOO at 300 K (black) 250 ps, vibEX at 5983 cm^{-1} from 250 ps long simulation (blue) using the reactant FF, thermal simulations at 760 K from 250 ps long simulation (red) using the reactant FF and vibEX at 5983 cm^{-1} from reactive simulation showing H-transfer within 10 ps (green) using MS-ARMD. Bond distance distribution from thermal simulations (red dashed) at 932 K using ASE.

C_AH_A bond leads to a considerably flatter, more extended distribution, similarly to the C_BH_D distribution. Conversely, the C_AC_B distribution function slightly narrows with an increased peak height at the minimum energy geometry. This suggests that compression of the C-C bond facilitates the 1,4 H-shift reaction.

Experimentally, vibrational excitation of the CH stretch introduces energies between 5603 and 6082 cm^{-1} (16.0 to 17.3 kcal/mol) compared with a height of the first transition state of 16.0 kcal/mol (14.0 kcal/mol zero-point corrected) and 18.6 kcal/mol (16.2 kcal/mol) from the higher-level CCSD(T) calculations[43] separating *syn*-CH₃CHOO and VHP. The more relevant of the two computed barrier heights is that corrected for ZPE. Hence *syn*-CH₃CHOO excited with 16.0 kcal/mol energizes the system right to the TS whereas with 17.3 kcal/mol it is above the TS energy by about 1 kcal/mol. Assuming that tunneling effects will be particularly prevalent for excitation below the TS energy and less relevant for excitation above the TS energy it is expected that tunneling effects will become progressively smaller as the photon energy increases. Given the qualitative character of this discussion it should also be noted that direct experimental proof for tunneling to be operative can be extremely challenging. Because the first reaction step involves a 1,4 H-shift it is quite likely that tunneling will speed up the rate for OH elimination. However, from the present simulations with excitation of two quanta in the CH-stretch vibration tunneling is

Figure 3.19: Bond distance distribution for *syn*-CH$_3$CHOO at 300 K (black solid line), Vibrationally activated at 5983 cm^{-1} from 250 ps long simulation (blue solid line) using reactant FF, thermally activated at 760 K from 250 ps long simulation (red solid line) using reactant FF and Vibrationally activated at 5983 cm^{-1} from reactive simulation showing H-transfer in 10 ps (green solid line) using MS-ARMD. Bond distance distribution from thermal simulations (red dashed) at 932 K using ASE.

not the determining factor for reactivity.

To further probe this, additional MS-ARMD simulations with D_e^{OO} = 31.5 kcal/mol were carried out. Starting from samples at 50, 100, 200, and 300 K with 4000 cm^{-1} of excess energy in the CH stretch, OH-elimination occurs. At an excitation energy of 4000 cm^{-1}, out of a total 100 simulations, each 5 ns long (i.e., a total of 500 ns), there are 0, 1, 2 and 62 H-transfers for equilibrated trajectories at the four temperatures. One caveat in using MD simulations for reactions is the fact that the energy content in particular modes in a polyatomic molecule can drop below its zero-point level. This difficulty has been recognized and various remedies have been proposed[201–203] but none of them solves it in general. For final state analysis, as was the focus in Ref.[43], it is possible to exclude trajectories for which the energy content in important modes drops below the ZPE from the analysis. One other possibility is to run classical MD on a full-dimensional, ZPE-corrected PES as has been done for malonaldehyde[204] and to compare the results from path integral MD simulations on the ZPE-uncorrected PES. For high temperatures (~1000 K) the H/D kinetic isotope effects (KIE) differ by a few percent whereas for room temperature and below they differ by a factor of two to three. Detailed analysis showed that for proton transfer in malonaldehyde at temperatures higher than 250 K, zero point effects are more important than tunneling.

The energy dependence of the experimentally measured rates and those from simulations and RRKM calculations are consistent, see Figure 3.7. However, there are several points of note. MS-ARMD simulations with a realistic dissociation energy for the O–O bond (black traces in Figure 3.7A) are close to experiment if $N(t)/N_{tot}$ is analyzed with a single exponential. With the more realistic stretched exponential model for $N(t)/N_{tot}$ (see Figure 3.11) the rates are about a factor of 5 higher. On the other hand, the barrier for the 1,4-hydrogen shift from MS-ARMD is lower by about 2 kcal/mol than the best estimates. Thus, rates from using stretched exponentials in the analysis will be reduced by about one order of magnitude if the higher barrier is used which brings them again close to experiment. Similarly, for the thermal simulations based on the PhysNet representation the barrier height for the 1,4-hydrogen shift reaction and the energy required for breaking the O–O bond are

under- and overestimated, respectively. The energy dependence of these rates are consistent with experiments in the jet but the absolute values depend on the value of D_e^{OO} and the initial ensemble. Finally, the RRKM calculations including tunneling assume that the second step (OH elimination from VHP) is rapid compared with the first step. This, however, is not guaranteed as earlier[191,205] and the present work (see Figure 3.14) show. Thus, the rates from RRKM plus tunneling calculations should not necessarily be close to those from experiment if the dynamics of the second step is included.

The PESs used here allow us to run a statistically significant number of reactive trajectories on the nanosecond time scale with qualities approaching the MP2 level of theory at the cost of an empirical force field. The two representations have their particular advantages and shortcomings. The MS-ARMD PES has an overall accuracy of ~1 kcal/mol which is certainly sufficient for qualitative studies. Such a parametrization allows exploration of parameter space as illustrated by the variation of the well depth D_e^{OO}. On the other hand, the PhysNet representation is highly accurate with respect to the reference points. Simulations with this PES are about two orders of magnitude slower which limits broad exploration of initial conditions. One example concerns vibEX simulations with PhysNet. These needed to be carried out with excitation energies ranging from 6500 cm^{-1} to 9000 cm^{-1} in order to observe reactive trajectories on the nanosecond time scale due to the O–O dissociation energy of $D_e^{OO} = 35.7$ kcal/mol. With 9000 cm^{-1} excess energy 1 out of 10 trajectories show OH-elimination on the 1 ns time scale whereas with 8000 cm^{-1}, 1 out of 5 trajectories progressed to product on the 5 ns time scale. Excitation with 6500 cm^{-1} does not lead to OH-elimination on the 25 ns time scale. Exploration of the influence of D_e^{OO} within PhysNet is not easily possible without dedicated modification of the underlying data set and retraining the NN.

VibEX simulations with $D_e^{OO} = 31.5$ kcal/mol indicate that the 1,4 H-shift reaction yields ~80 % VHP on the 1 ns time but only 5 % react further to the product, see Figure 3.14A. This changes to prompt OH-loss and a yield of close to 100 % within ~3 ns for $D_e^{OO} = 23.5$ kcal/mol, see Figure 3.9. Previous electronic structure calculations at the CASPT2 level[43] reported dissociation energies 23.0 kcal/mol

(submerged) and 29.3 kcal/mol (the "positive barrier"). Therefore, depending on the initial preparation and the O–O dissociation energy (D_e^{OO} = 31.5 or 23.5 kcal/mol) the second step (VHP → CH$_2$COH+OH) is or is not a bottleneck for OH generation following vibrational excitation. For D_e^{OO} = 31.5 kcal/mol rapid IVR after the 1,4 H-shift reaction delays OH-elimination. This can be seen, e.g., in the high excitation of the CCOO-dihedral motion after formation of VHP (Figure 3.17). Similar observations were made for the isomerization of acetaldehyde (AA) to vinylalcohol (VA).[165] Excitation of AA with an actinic photon (~94 kcal/mol) is not sufficient to trigger isomerization to VA on the 500 ns time scale although the AA → VA barrier height is only 68 kcal/mol.

Conversely, the thermal simulations which are more representative of initial conditions following ozonolysis of 2-butene find that on the 1 ns time scale similar amounts of VHP and OH-elimination products are formed with only ~20 % of reactant remaining, see Figure 3.14B. Hence, on this time scale no clear bottleneck can be identified. How much VHP finally reacts to form OH product also depends on the collisional quenching time which is between 1 ns and 10 ns.[177,206] Hence, if VHP does not form OH product on that time scale it is more likely to lose energy in collisions with the environment which limits OH production from this pathway. It is also possible to estimate the thermal rate at 300 K from the MS-ARMD simulations. This is done from the simulations with D_e^{OO} = 23.5 kcal/mol as the more likely dissociation energy from the present work, see open black symbols in Figure 3.7B. Linear extrapolation to 300 K yields a rate of 1.2×10^5 s^{-1}. Accounting for the underestimation of the barrier height for the 1,4 hydrogen transfer step (16.0 vs 18.6 kcal/mol, see Table 3.6) the estimated rate at 300 K of ~10^3 s^{-1} is consistent with an experimentally determined value of 182 ± 66 s^{-1}.[172]

OH-elimination from *syn*-CH$_3$CHOO following vibrational excitation or thermal preparation yields rates consistent with experiments using full dimensional MS-ARMD and NN-based PESs. The classical MD simulations do not include tunneling which is expected to further speed up the first step and to become important for excitation closer to the transition state energy. Following vibrational excitation of the CI, VHP can readily progress to OH-elimination or accumulate depending on

the O–O dissociation energy and the initial preparation. This is different for thermal preparation of *syn*-CH₃CHOO. One of the important "unknowns" remains the dissociation energy D_e^{OO} which would require large-scale multireference configuration interaction calculations. Overall, the present work provides molecular-level insights for an important reaction in atmospheric chemistry. The approaches used here are generic and expected to be applicable to a range of other reactions.

3.1.6 Supporting Information

IR excitation (cm^{-1})	t (ns)	H transfer	OH formed	IR excitation (cm^{-1})	t (ns)	H transfer	OH formed
5603	0.1	1307	8	5950.9	0.1	1860	12
	0.2	2479	21		0.2	3286	42
	0.3	3498	48		0.3	4484	83
	0.4	4378	85		0.4	5447	147
	0.5	5104	124		0.5	6272	209
	0.6	5679	158		0.6	6874	267
	0.7	6212	207		0.7	7344	342
	0.8	6700	254		0.8	7763	422
	0.9	7120	305		0.9	8089	506
	1.0	7452	367		1.0	8413	605
5709	0.1	1462	8	5983.5	0.1	1873	15
	0.2	2641	24		0.2	3371	50
	0.3	3709	49		0.3	4551	98
	0.4	4568	81		0.4	5487	161
	0.5	5325	134		0.5	6255	226
	0.6	5938	188		0.6	6932	307
	0.7	6474	253		0.7	7440	388
	0.8	6953	316		0.8	7896	460
	0.9	7346	376		0.9	8228	555
	1.0	7712	431		1.0	8528	647
5747.6	0.1	1544	10	6082.2	0.1	2128	15
	0.2	2761	35		0.2	3664	50
	0.3	3786	71		0.3	4825	101
	0.4	4761	109		0.4	5853	160
	0.5	5510	154		0.5	6661	232
	0.6	6200	208		0.6	7295	311
	0.7	6697	264		0.7	7788	406
	0.8	7142	334		0.8	8225	509
	0.9	7526	401		0.9	8547	612
	1.0	7873	485		1.0	8774	724
5818.1	0.1	1695	9				
	0.2	2996	31				
	0.3	4119	73				
	0.4	4999	116				
	0.5	5782	169				
	0.6	6415	225				
	0.7	6943	276				
	0.8	7381	340				
	0.9	7747	410				
	1.0	8097	467				

Table 3.1: Total number of H-transfer and OH-elimination reactions from MS-ARMD vibEX simulations with D_e^{OO} = 31.5 kcal/mol. The total simulation time is 1 ns for each trajectory and 10000 independent simulations are carried out. The aggregated number of reactions up to time t is reported.

IR excitation (cm^{-1})	t (ps)	H transfer	OH formed	IR excitation (cm^{-1})	t (ps)	H transfer	OH formed
5603	10	242	2	5950.9	10	267	3
	20	396	7		20	460	17
	30	534	16		30	644	27
	40	674	28		40	829	52
	50	795	50		50	1043	74
	60	935	68		60	1242	105
	70	1084	101		70	1409	133
	80	1215	128		80	1579	174
	90	1327	153		90	1749	215
	100	1440	178		100	1927	257
5709	10	215	3	5983.5	10	311	5
	20	361	12		20	513	23
	30	512	19		30	706	37
	40	655	27		40	881	58
	50	814	54		50	1053	84
	60	942	74		60	1264	118
	70	1090	90		70	1440	154
	80	1232	116		80	1599	181
	90	1351	144		90	1803	227
	100	1467	178		100	1956	275
5747.6	10	238	3	6082.2	10	337	4
	20	409	11		20	556	18
	30	585	21		30	766	38
	40	753	41		40	978	67
	50	912	62		50	1195	108
	60	1060	90		60	1414	140
	70	1187	110		70	1628	182
	80	1328	145		80	1816	226
	90	1458	178		90	1990	288
	100	1578	212		100	2176	344
5818.1	10	262	3				
	20	450	12				
	30	607	28				
	40	791	48				
	50	946	68				
	60	1104	95				
	70	1244	132				
	80	1390	172				
	90	1562	218				
	100	1711	253				

Table 3.2: Total number of H-transfer and OH-elimination reactions from MS-ARMD vibEX simulations with D_e^{OO} = 23.5 kcal/mol. The total simulation time is 100 ps for each trajectory and 10000 independent simulations are carried out. The aggregated number of reactions up to time t is reported.

T (K)	Total	t (ns)	H transfer	OH formed	T (K)	Total	t (ns)	H transfer	OH formed
865	9292	0.1	4034	416	925	8894	0.1	4648	657
		0.2	5272	861			0.2	5816	1230
		0.3	5905	1236			0.3	6395	1668
		0.4	6328	1561			0.4	6764	2094
		0.5	6642	1831			0.5	7006	2414
		0.6	6858	2076			0.6	7193	2685
		0.7	7039	2295			0.7	7337	2897
		0.8	7172	2501			0.8	7466	3117
		0.9	7268	2680			0.9	7551	3294
		1.0	7377	2839			1.0	7621	3449
877	9227	0.1	4228	510	932	8848	0.1	4663	690
		0.2	5351	971			0.2	5777	1311
		0.3	5959	1336			0.3	6358	1798
		0.4	6356	1684			0.4	6687	2178
		0.5	6640	1953			0.5	6960	2472
		0.6	6875	2181			0.6	7138	2746
		0.7	7048	2442			0.7	7305	2983
		0.8	7184	2610			0.8	7406	3165
		0.9	7308	2771			0.9	7520	3364
		1.0	7419	2924			1.0	7590	3525
886	9165	0.1	4269	490	956	8789	0.1	4982	827
		0.2	5415	974			0.2	6076	1532
		0.3	6054	1374			0.3	6620	2026
		0.4	6428	1717			0.4	6948	2455
		0.5	6708	1981			0.5	7152	2774
		0.6	6922	2273			0.6	7319	3063
		0.7	7069	2477			0.7	7435	3308
		0.8	7207	2701			0.8	7535	3528
		0.9	7339	2871			0.9	7621	3717
		1.0	7449	3039			1.0	7683	3880
902	9032	0.1	4347	601					
		0.2	5514	1115					
		0.3	6123	1539					
		0.4	6503	1895					
		0.5	6748	2191					
		0.6	6993	2454					
		0.7	7179	2684					
		0.8	7307	2876					
		0.9	7392	3068					
		1.0	7493	3228					

Table 3.3: Total number of H-transfer and OH-elimination reactions from MS-ARMD thermal simulations with D_e^{OO} = 31.5 kcal/mol. The total simulation time is 1 ns for each trajectory and 10000 independent simulations are carried out. The aggregated number of reactions up to time t is reported. Some simulations shown H-transfer in heating step which are excluded from the total.

T (K)	Total	t (ps)	H transfer	OH formed	T (K)	Total	t (ps)	H transfer	OH formed
865	9338	10	827	93	925	9000	10	1239	204
		20	1496	286			20	2103	542
		30	2022	524			30	2682	905
		40	2426	768			40	3143	1249
		50	2765	998			50	3538	1533
		60	3081	1224			60	3846	1818
		70	3332	1440			70	4139	2095
		80	3554	1631			80	4370	2342
		90	3767	1827			90	4574	2555
		100	3970	2023			100	4752	2768
877	9324	10	954	146	932	9000	10	1231	170
		20	1651	385			20	2112	472
		30	2196	615			30	2721	835
		40	2604	860			40	3203	1196
		50	2989	1107			50	3599	1525
		60	3279	1352			60	3911	1823
		70	3560	1575			70	4167	2100
		80	3787	1788			80	4396	2349
		90	4016	1973			90	4563	2550
		100	4188	2173			100	4750	2770
886	9257	1037	34	133	956	8917	10	1356	216
		20	1768	401			20	2268	616
		30	2283	650			30	2899	997
		40	2698	922			40	3382	1357
		50	3103	1174			50	3753	1695
		60	3411	1406			60	4085	2011
		70	3687	1642			70	4373	2277
		80	3900	1850			80	4623	2560
		90	4130	2066			90	4794	2795
		100	4332	2258			100	4989	3014
902	9167	10	1065	148					
		20	1822	410					
		30	2427	681					
		40	2876	986					
		50	3277	1232					
		60	3601	1520					
		70	3881	1766					
		80	4113	2001					
		90	4334	2232					
		100	4538	2407					

Table 3.4: Total number of H-transfer and OH-elimination reactions from MS-ARMD thermal simulations with $D_e^{OO} = 23.5$ kcal/mol. The total simulation time is 100 ps for each trajectory and 10000 independent simulations are carried out. The aggregated number of reactions up to time t is reported. Some simulations shown H-transfer in heating step which are excluded from the total.

78

T (K)	Total	t (ps)	H transfer	OH formed	T (K)	Total	t (ps)	H transfer	OH formed
865	9980	10	15	12	902	9891	10	30	24
		20	35	29			20	48	43
		30	44	43			30	68	62
		40	56	54			40	90	84
		50	69	66			50	108	103
		60	78	73			60	121	116
		70	90	88			70	134	131
		80	99	98			80	144	143
		90	107	106			90	161	153
		100	112	110			100	171	169
877	9978	10	24	19	925	9970	10	40	30
		20	38	34			20	58	53
		30	51	48			30	75	70
		40	62	62			40	100	92
		50	74	74			50	117	113
		60	90	83			60	141	136
		70	102	97			70	165	159
		80	120	114			80	175	172
		90	131	130			90	195	188
		100	141	139			100	211	208
886	9956	10	26	21	956	9961	10	41	34
		20	48	42			20	77	69
		30	62	55			30	108	102
		40	85	75			40	144	133
		50	99	98			50	164	162
		60	112	110			60	192	180
		70	126	125			70	208	207
		80	143	139			80	227	225
		90	153	149			90	242	239
		100	164	162			100	260	260

Table 3.5: Total number of H-transfer and OH-elimination reactions from PhysNet thermal simulations. The total simulation time is 100 ps for each trajectory and 10000 independent simulations are carried out. The aggregated number of reactions up to time t is reported. Some simulations shown H-transfer in heating step which are excluded from the total.

	MS-ARMD $D_{\mathrm{e}}^{\mathrm{OO}} = 31.5$	MS-ARMD $D_{\mathrm{e}}^{\mathrm{OO}} = 23.5$	PhysNet	CCSD(T)-F12b/ CASPT2[43]
syn-CH₃CHOO	0.00	0.00	0.00	0.00
TS	16.02 (14.02)	16.02 (14.02)	14.90 (12.90)	18.6 (16.2)
CH₂CHOOH (VHP)	-22.95	-22.95	-22.51	-19.2
CH₂CHO + OH	8.55	0.55	13.27	6.8

Table 3.6: Stationary points on the MS-ARMD and the NN-based PESs. Energies (in kcal/mol) are given with respect to *syn*-CH₃CHOO and the energy differences including zero point energy corrections are given in brackets. The total ZPE is 37.7 kcal/mol for *syn*-CH₃CHOO and 35.6 kcal/mol for the TS towards VHP.

3.2 Publication: Molecular Simulation for Atmospheric Reactions: Non-Equilibrium Dynamics, Roaming and Glycolaldehyde Formation Following Photo-Induced Decomposition of *syn*-Acetaldehyde Oxide
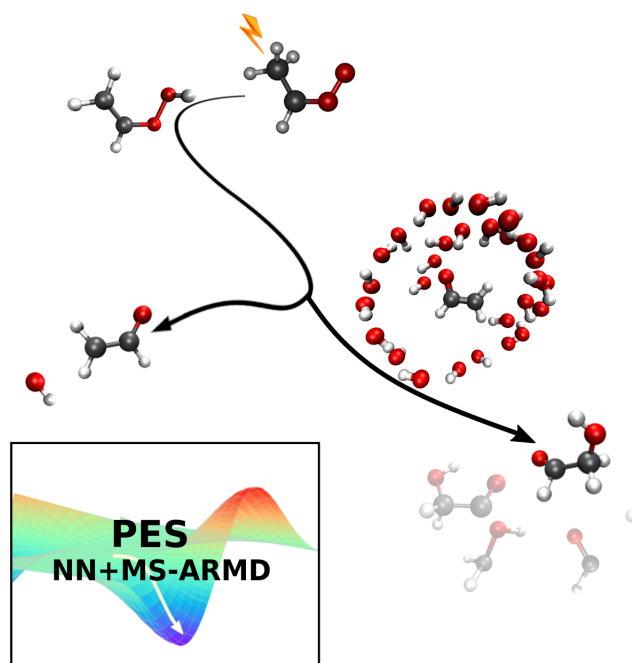
3.2.1 Abstract

The decomposition dynamics of vibrationally excited *syn*-CH$_3$CHOO to form either vinoxy + hydroxyl (CH$_2$CHO + OH) radicals or to recombine to form glycolaldehyde (CH$_2$OHCHO) are characterized using statistically significant numbers of molecular dynamics simulations using a full-dimensional neural-network-based potential energy surface at the CASPT2 level of theory. The computed final OH-translational and rotational state distributions agree well with experiments and probe the still unknown O–O bond strength D_e^{OO} for which best values from 22 to 25 kcal/mol are found. OH-elimination rates are consistent with experiments and do not vary appreciably with D_e^{OO} due to the non-equilibrium nature of the process. In addition to the OH-elimination pathway, OH-roaming is observed following O–O scission, which leads to glycolaldehyde formation on the picosecond time scale. Together with recent work involving the methyl-ethyl-substituted CI, we conclude that OH roaming is a general pathway to be included in molecular-level modeling of atmospheric processes. This work demonstrates that atomistic simulations with machine-learned energy functions provide a viable route to exploring the chemistry and reaction dynamics for atmospheric reactions.

3.2.2 Introduction

Chemical processing and the evolution of molecular materials in the atmosphere are primarily driven by photodissociation reactions. Sunlight photoexcites the molecules in the different layers of Earth's atmosphere and triggers chemical reactions. The photoproducts are then reactants for downstream reactions from which entire reaction networks emerge.[19] Within the extensive array of chemical reactions occurring in the biosphere, CIs are some of the eminent reactive species that have captured particular attention.[207] CIs make up an important class of molecules generated in the atmosphere from ozonolysis of alkenes, which proceeds through a 1,3-cycloaddition of ozone across the C=C bond. The resulting POZ then decomposes into a carbonyl oxide, also known as CI, and one carbonyl (aldehyde or ketone).[27] Highly energized CIs rapidly undergo either unimolecular decay to OH radicals[31] or collisional stabilization.[32] Stabilized CIs can isomerize and decompose into products, including the OH radical, or engage in bimolecular reactions with water vapor, SO_2, NO_2 and acids.[28,33]

In the laboratory, CIs in the gas phase were generated from iodinated precursors.[43,208] Such studies allowed experimental characterization of the photodissociation dynamics of *syn*-CH_3CHOO using laser spectroscopy[36,42,43] and provided important information about the reactivity and decomposition dynamics of CIs, including final state distributions of the OH product. Computationally, a range of methods, including RRKM theory[42] and MD simulations from the saddle point separating *syn*-CH_3CHOO and CH_2CHOOH,[43] were employed to study the dynamics and decomposition of vibrationally excited *syn*-CH_3CHOO. However, it was not until recently that the entire reaction pathway from energized *syn*-CH_3CHOO to $OH(X^2\Pi)$ elimination was followed using NN representations of the PESs.[48]

The intense interest in understanding the photodissociation dynamics of CIs, and in particular *syn*-CH_3CHOO, is due to the fact that one of its decomposition products is the OH radical, also termed the "detergent of the troposphere".[159,160] OH is one of the most powerful oxidizing agents and plays an important role in the chemical evolution of the atmosphere, triggering the degradation of many pollutants, including
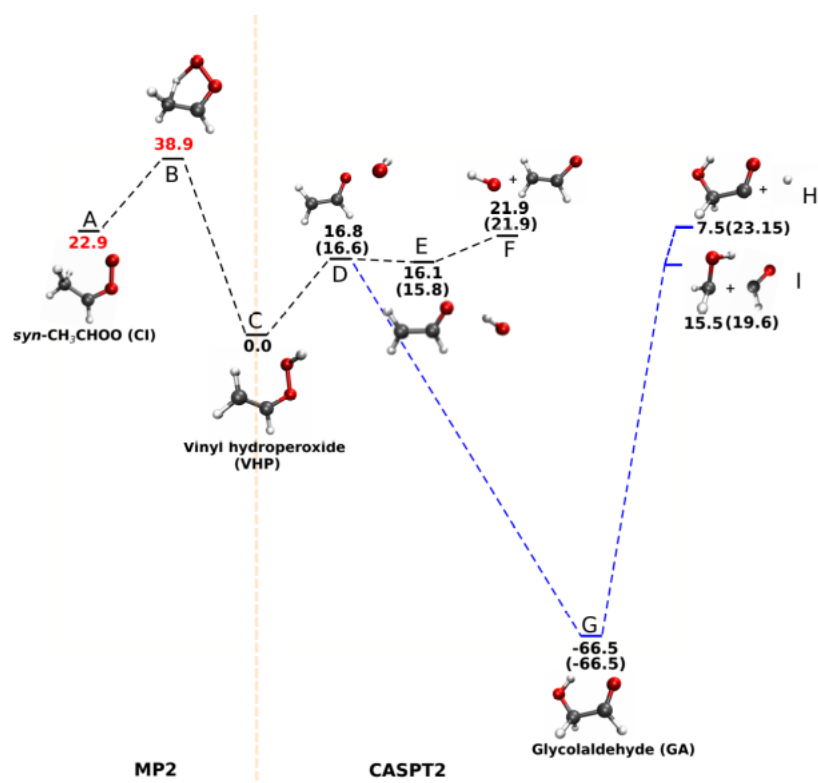
Figure 3.20: Schematic of the PES relating the species involved. Energies (in kcal/-mol) for optimized geometries with VHP (structure C) as the zero of energy. For the 1,4 H transfer (A to C), energies are reported from MP2/6-31G(d)-level PES and upon VHP decomposition (C – I), energies from the CASPT2(12,10)/cc-pVDZ level PES are shown. The *ab initio* reference energies are in parentheses and those from the NN-PES are without parentheses. The orange vertical line indicates the two PESs that were used in the hybrid simulations.
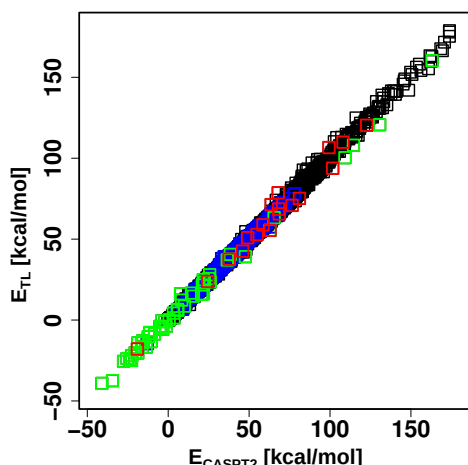
Figure 3.21: Correlation between the *ab initio* CASPT2 energies and predicted TL energies for the 2750 test set structures from the entire reaction pathway with overall RMSE of 1.56 kcal/mol and $r^2 = 0.99$. Color code: GA (G; green), $CH_2CHO + OH$ (D, E, F; blue), VHP and Criegee (A, B, C; black), $HCO + CH_2OH$ (I; red). Reference data for $CH_2OHCO + H$ was not included in the training as this decay channel was not considered.

VOCs.[158] Field studies have suggested that ozonolysis of alkenes is responsible for the production of approximately one-third of the atmospheric OH during daytime and is the predominant source at night.[35,161] A critical parameter that drives OH elimination is the O–O bond strength (D_e^{OO}) of VHP which is still unknown. One purpose of this work is to provide a reliable range for D_e^{OO} by using PESs with different O–O bond strengths and comparing results from MD simulations with experimental observables.

The dynamics studies employ an NN-based PES that is suitable for following chemical reactions. Except for the $CH_2CHO + OH$ product channel (states D – F in Figure 3.20), it was also found that GA (state G in Figure 3.20) is produced. GA is an atmospheric volatile organic compound and can be generated from isoprene[49], ethene[50] and biomass burning[51]. Using molecular-beam mass spectrometry, the formation of GA following ozonolysis of *trans*-2-butene was recently observed.[52] To the best of our knowledge, this is the only experimental work that reported the formation of GA from *syn*-CH₃CHOO so far.

This work uses atomistic simulations together with validated machine-learned PESs based on the PhysNet NN architecture.[143] An overview of the species considered is
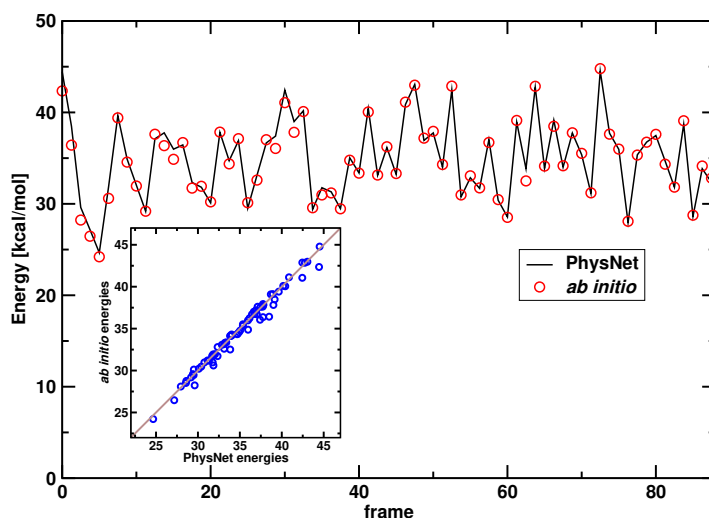
Figure 3.22: Validation of the NN-PES in the OH roaming region: Correlation between energies from the NN-PES (black line) and CASPT2 calculations (open circles) for a roaming trajectory. The frames are separated by 10 fs. Importantly, these structures and energies were never used in constructing the NN-PES. The RMSE between the NN-PES and the quantum reference calculations is 0.6 kcal/mol with $r^2 = 0.98$ (see inset).

shown in Figure 3.20. The excellent performance of the transfer-learned NN-PES (see Methods in the Supporting Information) is reported in Figures 3.21 and 3.22, which compare CASPT2/cc-pVDZ energies with the trained model for structures that were not used in training the NN. As will be demonstrated, the combined use of computed and experimental observables allows bracketing of the elusive O–O scission energy for OH elimination. To this end, the final state total kinetic energy release (TKER) and rotational state distributions, $P(\text{TKER})$ and $P(N)$, respectively, where N is the OH rotational quantum number, were determined from a statistically significant number of reactive MD simulations and compared with experiment. The simulations were initialized from an ensemble of low-temperature *syn*-CH₃CHOO structures, akin to those prepared experimentally. This was followed by vibrationally exciting the alkyl CH stretching vibration of *syn*-CH₃CHOO with energies commensurate with the experimental procedures.[43] Simulations and analyses were based on classical MD simulations without and with zero-point vibrational energy (ZPVE) in the reactant and products included, (see the Supporting Information for details).
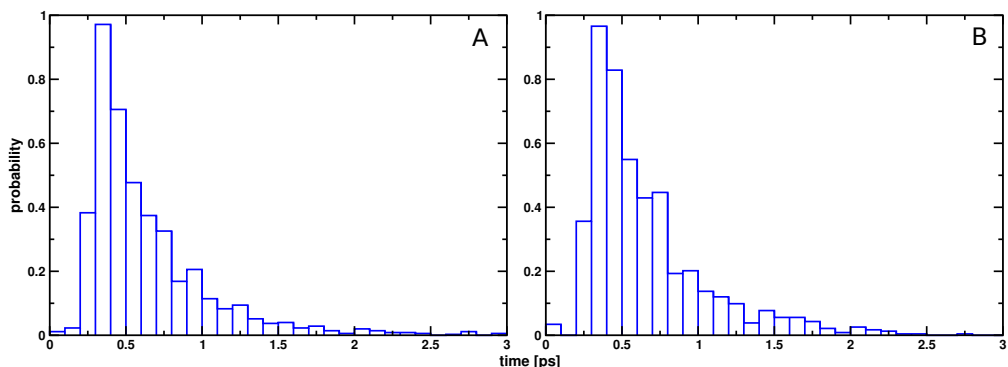
Figure 3.23: Distribution of OH dissociation times from O–O separation 3 Å to 10 Å for trajectories with CH$_2$CHO and OH as the reaction products (CH-stretch excitation with 5988 cm^{-1}). The time at which the O–O separation reaches 3 Å is set to be the zero of time. Panels A and B are for $D_{\mathrm{e}}^{\mathrm{OO}}$ = 22 and $D_{\mathrm{e}}^{\mathrm{OO}}$ = 25 kcal/mol, respectively.

3.2.3 Results

First, the TKER as the sum of the translational energies of the CH$_2$CHO and OH products from energized *syn*-CH$_3$CHOO was analyzed. In trajectories leading to OH as the final product, two different types of OH-elimination pathways were observed. The first is referred to as "direct" elimination, whereas the second is classified as "roaming" elimination. A semiquantitative criterion[209] based on the distance traveled by the OH fragment was used to distinguish the two types of processes. For this, the moment along the dynamics at which the level of the O–O separation exceeds 3 Å was set as the zero of time. Then, the time t_e required for the O–O distance to increase from 3 Å to 10 Å was recorded. When t_{e} < 0.8 ps, a trajectory is considered as direct elimination, whereas for all other cases (t_{e} ≥0.8 ps), it is a roaming elimination (see the Supporting Information and Figures 3.23 – 3.25).

Panels A and B of Figure 3.26 compare computed and experimental P(TKER) to form OH(X^{2}Π, v = 0, N = 3) following excitation of the CH-stretch vibration at 5988 cm^{-1}. MD simulations using the NN-PES with $D_{\mathrm{e}}^{\mathrm{OO}}$ values of 22 and 25 kcal/mol yield P(TKER) via direct elimination (dashed line), which agree favourably with experiment. In all simulations, the OH radical was in the v = 0 vibrational ground
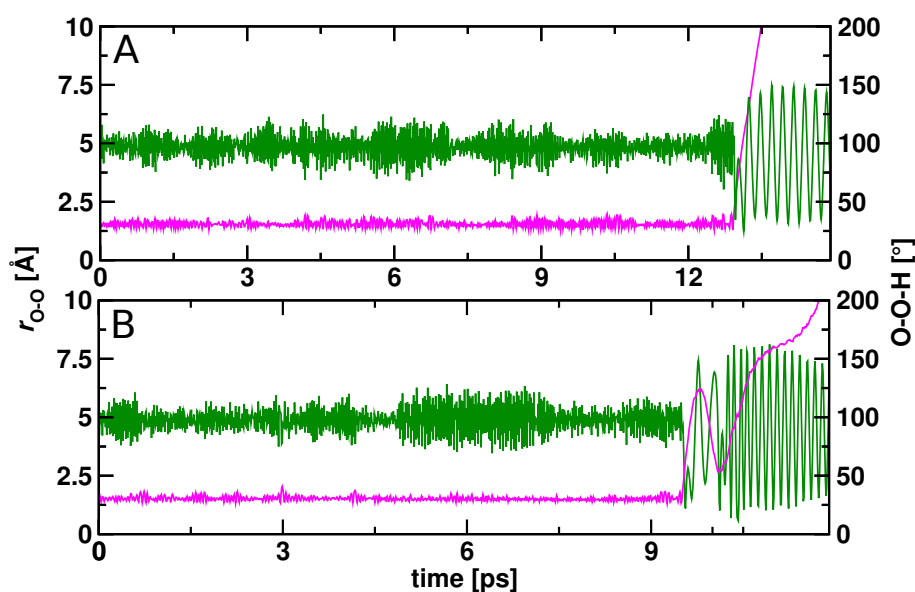
Figure 3.24: Time series for the O–O separation and the O–(OH) angle from a trajectory showing direct (Panel A) and roaming OH dissociation (Panel B) using the NN-PES with D_e^{OO} = 22 kcal/mol (CH-stretch excitation with 5988 cm^{-1}).
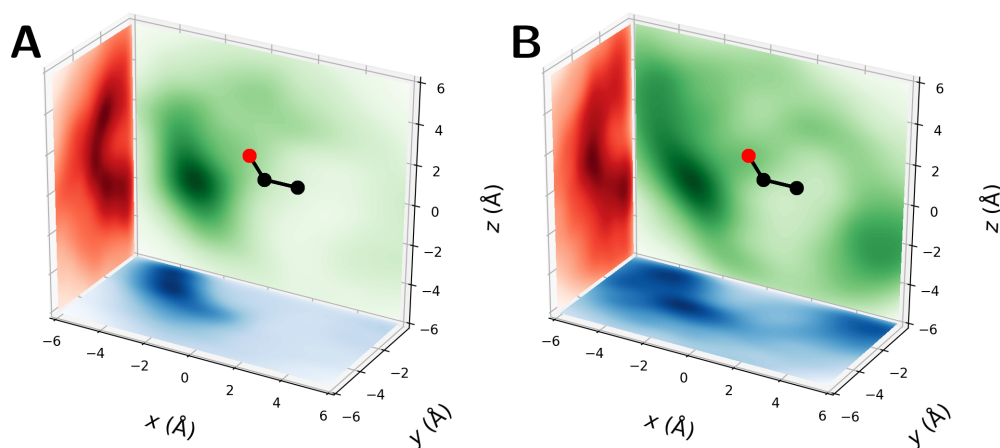
Figure 3.25: Probability densities of the OH radical around the CH₂CHO group (with C_A–C_B–O_A in the xz plane and the center of mass of CH₂CHO in the origin) from ~500 trajectories each leading to A: direct or B: roaming OH dissociation on the NN-PES with D_e^{OO} = 22 kcal/mol. Trajectories in panel B access the half planes (x, y) and (x, z) which are not visited in panel A for directly dissociating trajectories.
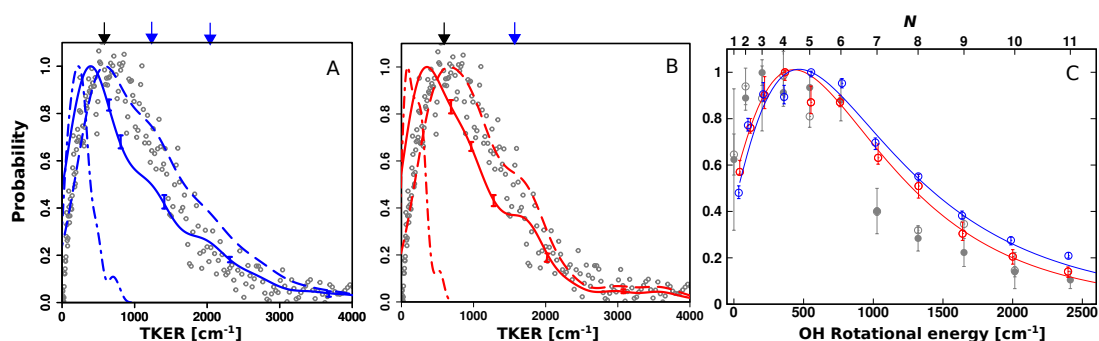
Figure 3.26: OH final state distributions from simulations with CH-stretch excitation at 5988 cm^{-1}. Panels A and B show P(TKER) values for the final OH(X$^2\Pi$, $v = 0$, $N = 3$) with D_e^{OO} values of 22 and 25 kcal/mol, respectively, and panel C shows $P(N)$ values for the two values of D_e^{OO}. Panels A and B report P(TKER) for direct elimination with roaming (solid lines), direct elimination (dashed lines), and roaming (dotted–dashed) together with experiment (gray circles).[43] The black arrow indicates the position of the maximum from experiment at 600 cm^{-1}.[43] Blue arrows point toward the local maxima in P(TKER), termed "waviness" in the text. For simulations forming OH(X$^2\Pi$, $v = 0$, all N) and those including ZPVE in the reactant and two products, see Figures 3.27 and 3.28. Panel C shows $P(N)$ from simulations (direct elimination with roaming) compared with experiments (gray circles with error bars).[43] Error bars are from bootstrapping.[210] The solid lines are Gamma distributions fit to the computed data. Positions of maxima and fwhms of all distributions are summarized in Table 3.7.
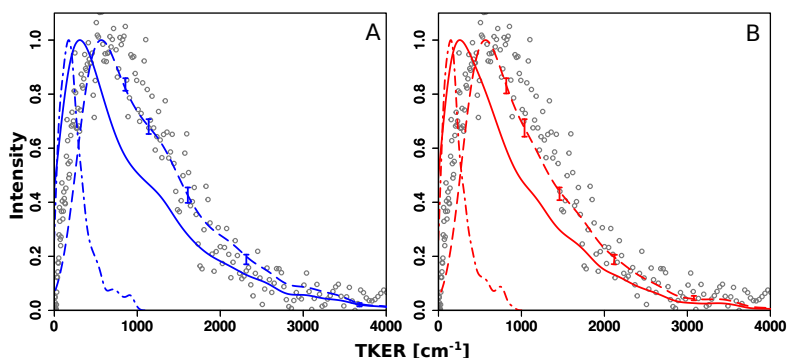
Figure 3.27: TKER distributions for final OH(X$^2\Pi$, $v = 0$, all N) from simulations and experiment. The computed distributions (CH-stretch excitation with 5988 cm^{-1}) compared with those reported from experiments[43] (grey circles, (at 6081 cm^{-1}). Panel A: $D_e^{OO} = 22$ kcal/mol; Panel B: $D_e^{OO} = 25$ kcal/mol for all (direct elimination plus roaming, solid lines), roaming (dotted-dashed) and direct elimination (dashed) trajectories.

state. The P(TKER) to yield OH(X$^2\Pi$, $v = 0$, all N) is reported in Figure 3.27. In Figure 3.26A ($D_e^{OO} = 22$ kcal/mol), the peak maxima for directly eliminating and all trajectories are at 588 and 390 cm^{-1} (600 cm^{-1} from experiment), which shift to 666 and 360 cm^{-1}, respectively, for a D_e^{OO} of 25 kcal/mol (panel B). In Figure 3.26A, the full width at half maximum (fwhm) decreases from 1522 to 1220 cm^{-1} if roaming trajectories are included, whereas in panel B, they are 1544 and 1170 cm^{-1} compared with 1340 cm^{-1} from experiment. One reason for the rather moderate change in the position of the peak maximum with D_e^{OO} is that independent of the precise value of D_e^{OO} the van der Waals complex in the product channel (CH$_2$CHO—HO) is always stabilized by ~6 kcal/mol (structures E and F in Figure 3.20) relative to the separated CH$_2$CHO + OH fragments.

The OH rotational distributions P(N) from simulations following CH-stretch excitation at 5988 cm^{-1} and for a D_e^{OO} value of 22 kcal/mol (blue) or 25 kcal/mol (red) are shown in Figure 3.26C and realistically describe the experimentally reported P_{exp}(N) (gray symbols). The position of the maximum (at $N_{max} = 4$) of the calculated P(N) and its fwhm of 1260 cm^{-1} from simulations using a D_e^{OO} of 25 kcal/mol agree well with the experimental P_{exp}(N), whereas with a D_e^{OO} of 22 kcal/mol, $N_{max} = 5$ and the fwhm increases to 1390 cm^{-1}. We note that the computed P(N) can be represented as a Gamma distribution (solid lines), which is also employed to model waiting
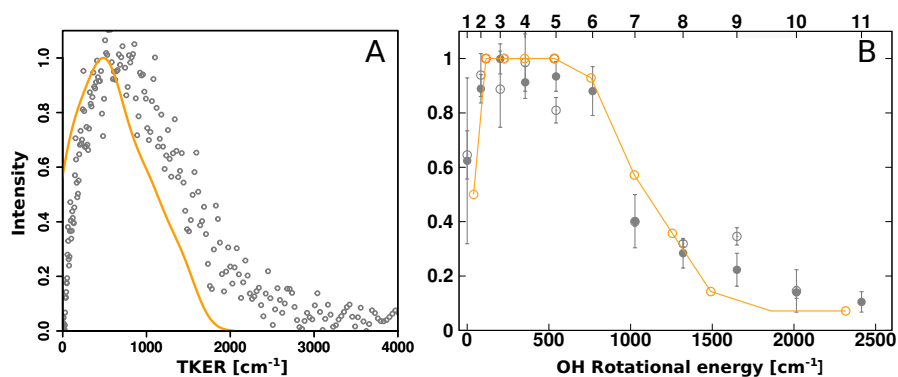
Figure 3.28: *P*(TKER) for all (direct and roaming) trajectories forming OH(X²Π, *v* = 0, *N* = 3) (Panel A) and *P*(*N*) (Panel B) of OH formed from simulations including ZPVE and using the NN-PES with D_e^{OO} = 22 kcal/mol (CH-stretch excitation with 5988 cm⁻¹) compared with experiments (grey open circles). The simulations were carried out by including ZPVE in the reactant and by analyzing only trajectories in which both products contained at least ZPVE of 1851 cm⁻¹ for OH and 9388 cm⁻¹ for CH₂COH, respectively.

time distributions[211] and successfully used to study photodissociation of H₂SO₄ following excitation of the OH stretch.[177,212]

An additional 6000 trajectories that account for ZPVE in the reactant and the CH₂CHO + OH products formed were run. In these simulations, only trajectories were analyzed for which OH and CH₂CHO contained at least the respective ZPVEs, which are 1851 cm⁻¹ and 9388 cm⁻¹, respectively. The number of trajectories satisfying this criterion was ~5 %, which evidently decreases the statistical significance of the final *P*(TKER) and *P*(*N*). The *P*(TKER) that includes directly eliminating and roaming trajectories to form OH(X²Π, *v* = 0, all *N* = 3) has its maximum at 475 cm⁻¹, and the FWHM is 1125 cm⁻¹, see Figure 3.28A. The rotational distribution *P*(*N*) (Figure 3.28B) agrees rather well with experiments for all values of *N* but specifically for the plateau region between 3 ≤ *N* ≤ 6. The question of whether to include ZPVE in such simulations is intensely debated,[213] primarily because vibrational energy can leak from one mode to other modes on the time scale of the simulation, which may affect computed observables. However, currently, no satisfactory and universally accepted treatment of this issue exists. Overall, it is found that except for the plateau region between (3 ≤ *N* ≤ 6) for *P*(*N*) the final state distributions change little for simulations including or excluding ZPVE effects.

Previous computational studies[43] reported $P(TKER)$ and $P(N)$ from simulations starting either from the TS of the 1,4 H-transfer or the submerged saddle point configuration preceding OH elimination (step D, Figure 3.20) as opposed to following the entire reaction from reactant to products. Starting at the H-transfer TS yielded $P(TKER)$ for $OH(X^2\Pi, v = 0, \text{all } N)$ with the maximum peak at 1550 cm⁻¹ compared with ~600 cm⁻¹ from experiment and a fwhm larger by a factor of ~2 compared with that observed experimentally (see Table 3.7). The position N_{max} of the maximum and width for $P(N)$ agree reasonably well with experiment. Initializing trajectories at the submerged barrier improve the position of the maximum for $P(TKER)$ to 880 cm⁻¹, but the fwhm remains too large; on the contrary, $P(N)$ narrows and the maximum shifts to a smaller N_{max}.

Energy-dependent rates were determined for three experimentally used excitation energies [$E = \{5603, 5818, 5988\}$ cm⁻¹ (see Figure 3.29)]. These rates were thoroughly investigated in earlier work[48] and are only considered briefly here for the new transfer-learned NN-PES. Reaction rates were determined from fitting the number $n(t)$ of nonreactive trajectories by time t to single- or stretched-exponential decays, which remarkably well represents the data with stretch exponents ranging from 1.1 to 1.6 (see Figure 3.30). Such values suggest that kinetic intermediates are visited that lead to distributed barriers.[189,214] The rates $k(E)$ for the two values of D_e^{OO} and from using the two fitting functions are reported in Figure 3.29 (red and blue traces) together with the experimentally determined rates (gray symbols). Importantly, the slope of $k(E)$ is correctly captured by the simulations, which has been reported previously.[48] On the contrary, the absolute rates from simulations not accounting for ZPVE are too large by 1 to 2 orders of magnitude. Accounting for ZPVE brings the rates into yet closer agreement with experiment (orange symbols in Figure 3.29).

Taken together, the results for final state distributions and energy-dependent rates validate and establish the quality of the transfer-learned PES. With a D_e^{OO} of 25 kcal/mol, $P(TKER)$ for $OH(X^2\Pi, v = 0, N = 3)$ as the final photoproduct (Figure 3.26B) and $P(N)$ in Figure 3.26C agree with experiment particularly favourably. The findings thus far also support results from the present and previous CASPT2 calculations that reported dissociation energies of ≤ 26 kcal/mol. Furthermore,
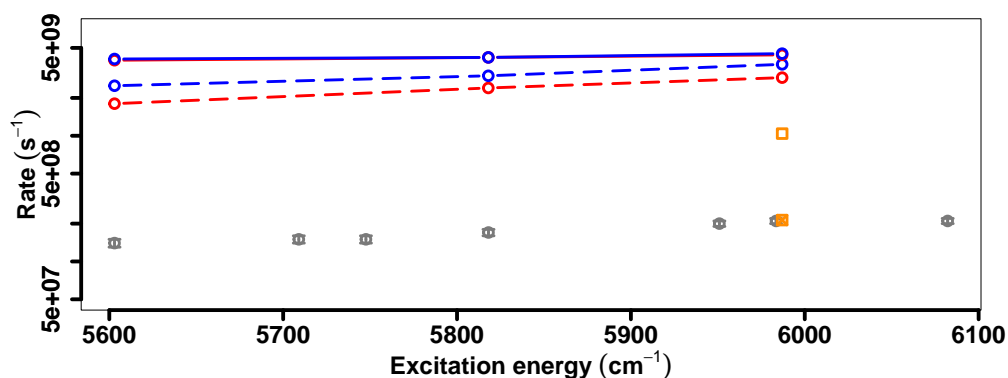
Figure 3.29: OH formation rates from simulations using the NN-PES compared with the experiments. The grey points with error bars are the experimental results.[42] The blue and red symbols are for D_e^{OO} = 22 and 25 kcal/mol, respectively. Dashed and solid lines are rates based on fitting $n(t)$ to single and stretched exponential decays, respectively, see Figure 3.30. Experimental rates are 1.4 to 2.1 ×10^8 s^{-1} for excitation energies between 5603 and 6082 cm^{-1}, compared with 1.8 to 2.9 ×10^9 with D_e^{OO} = 25 kcal/mol and 2.5 ×10^9 to 3.6 ×10^9 with D_e^{OO} = 22 kcal/mol using a single exponential function. The golden symbols are the rates from simulations accounting for ZPVE using the NN-PES with D_e^{OO} = 22 kcal/mol, fitted to single (open square) and stretched exponential (closed square).
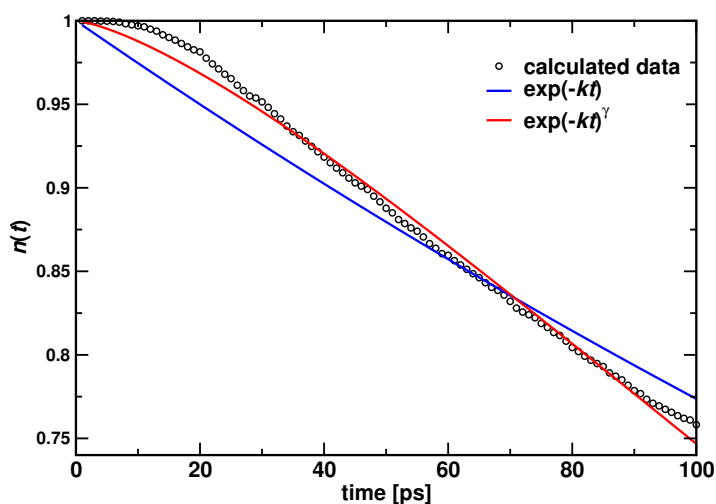
Figure 3.30: Fit of $n(t)$ (black circles) to single (blue) and stretched (red) exponential for simulations using the NN-PES with D_e^{OO} = 22 kcal/mol and vibrational excitation at 5603 cm^{-1}.
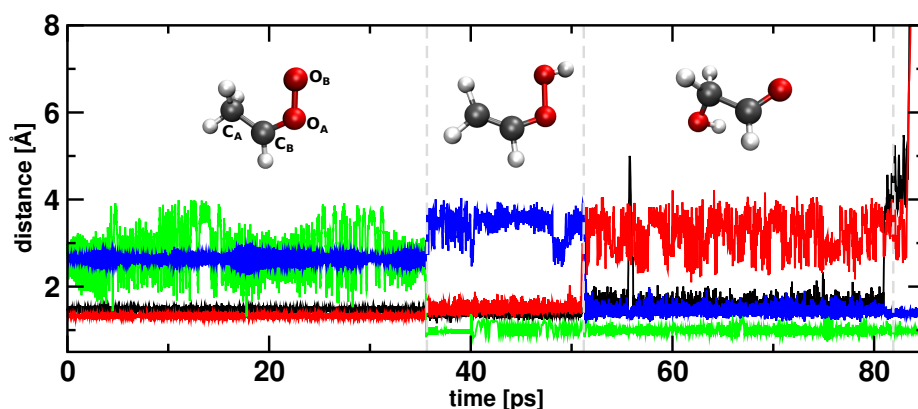
Figure 3.31: Time series for C_A–C_B (black), O_A–O_B (red), O_B–H (green) and C_A–O_B (blue) separations for a trajectory leading to HCO + CH₂OH radical formation using the NN-PES PES with a D_e^{OO} of 22 kcal/mol at 5988 cm⁻¹. At 36 ps, VHP is formed followed by GA formation and equilibration for the next 30 ps, and at 81 ps, the C–C bond dissociates to form HCO + CH₂OH.

they clarify that values of 31.5 [6-31G(d)] and 35.7 kcal/mol (aug-cc-pVTZ) from (single-reference) MP2 calculations are too high.[43]

Following the dynamics of OH elimination leads to products beyond the CH₂CHO + OH channel. Roaming of OH was found to generate GA through recombination with the carbon atom of the CH₂ group of CH₂CHO. GA is stabilized by 66.5 kcal/mol relative to VHP (see Figure 3.20G) and can further decay to products such as CH₂OH + HCO or CH₂OHCO + H (see Figure 3.20). The previous experiments[43] did not report OH roaming following photodissociation of *syn*-CH₃CHOO, whereas more recent work involving *anti*-methyl-ethyl substituted CI[53] shows that roaming is possible.

Explicit time series of important bond lengths involved in the formation of GA are reported in Figure 3.31. Starting from *syn*-CH₃CHOO at time zero, vibrational excitation of the CH-stretch leads to VHP formation after 36 ps. The system dwells in this state for ~15 ps until OH elimination and transfer to the CH₂ group yields GA with a residence time of ~30 ps followed by HCO elimination at t ~81 ps. The trajectory involves direct OH-transfer from the carbonyl oxygen to CH₂ group of VHP, without OH roaming.

94
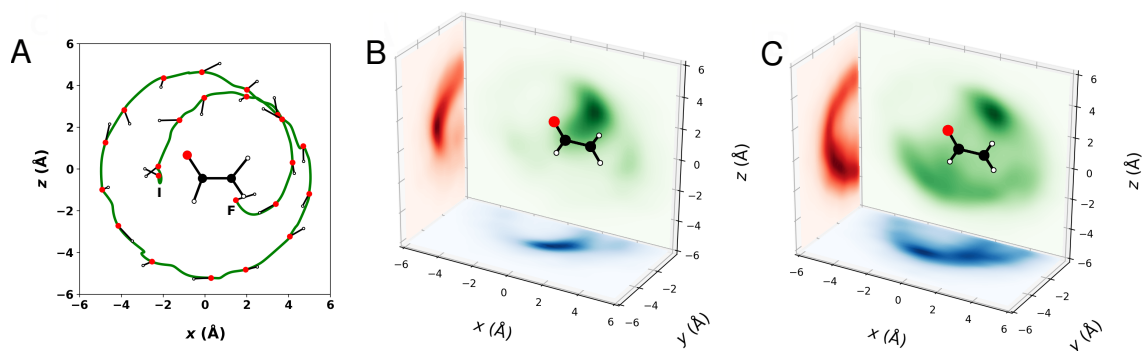
Figure 3.32: OH roaming around CH_2CHO before GA formation. (A) OH roaming trajectory projected on *x-z* plane around CH_2CHO before GA formation using the NN-PES with a D_e^{OO} of 22 kcal/mol at 5988 cm^{-1} for 1 ps. Here, **I** denotes the initial position of OH at which O–O bond of VHP dissociates and **F** denotes the final position of OH before GA-formation. (B and C) Probability densities of the OH radical around the CH_2CHO group (with C_A–C_B–O_A in the *x-z* plane and the center of mass of CH_2CHO in the origin) from ~500 trajectories leading to direct and roaming GA formation, respectively, on the NN-PES with D_e^{OO} of 22 kcal/mol. Trajectories in panel C access the half-planes $(-x, y)$ and $(-x, z)$, which are not visited in panel B for direct transfer trajectories.

Figure 3.32A shows an example of the OH roaming dynamics around the CH_2CHO radical before GA formation. This backside attack is a hallmark of roaming.[12] Probability densities for the OH radical moving around CH_2CHO before GA formation from 500 independent trajectories are shown in panels B and C of Figure 3.32. In both, direct transfer and roaming trajectories, the OH radical moves out-of-plane to attack the CH_2 group. A similar out-of-plane roaming was reported from experiments on formaldehyde.[215] In direct transfer trajectories, OH follows a semicircular path around the CH_2CHO group whereas pronounced roaming is followed for the other trajectories (see Figure 3.32C). Two-dimensional projections of the OH positions around the CH_2CHO radical as shown in panels B and C of Figure 3.32 provide an impression of the spatial range sampled by the OH radical before the formation of GA. Typical excursions involve separations of 3 to 4 Å from the center of mass of CH_2CHO.

Formation times of GA after dissociation of OH from VHP extend to ~2 ps (see Figure 3.33). The initial, pronounced peak at 0.1 ps is due to direct transfer between
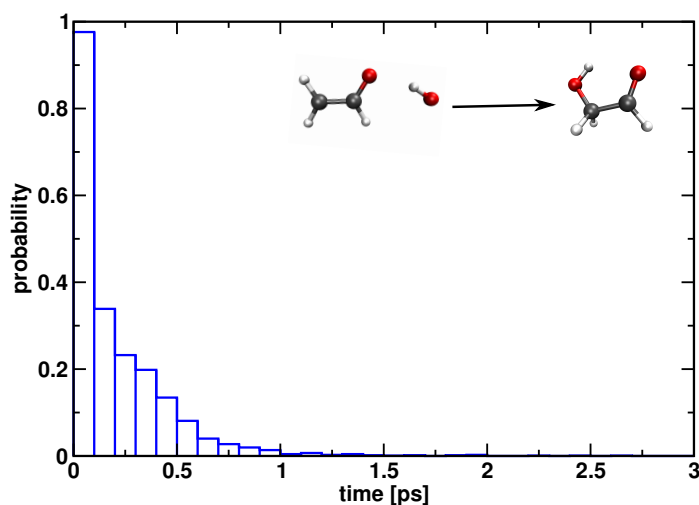
Figure 3.33: Distribution of reaction times for $CH_2CHO + OH \longrightarrow$ GA using the NN-PES with $D_e^{OO} = 22$ kcal/mol (CH-stretch excitation with 5988 cm^{-1}). Here the time at which the O–O separation reaches 3 Å is set as zero. The pronounced peak at short reaction time ($\tau_r \leq 0.1$ ps) involves direct OH transfer to the CH$_2$ group whereas the remainder of the distribution is for OH roaming trajectories, see Figure 3.32A for a trajectory with $\tau_r = 1$ ps.

VHP and GA, whereas the remainder of the exponentially decaying distribution involves OH roaming trajectories with maximum roaming times of 2.6 ps. Of the total amount of GA formed, 12 % (7 % of total VHP) shows C-C bond cleavage to form $CH_2OH + HCO$ radicals. The corresponding lifetime distribution of GA before dissociation, shown in Figure 3.34, features lifetimes of $\lesssim 10$ ps, with the most probable being ~1 ps.

The decay channels of internally hot VHP are summarized in Figure 3.35. Because the reaction is initiated by exciting the CH-stretch vibration, OH elimination from VHP is a non-equilibrium process because thermalization of the initial vibrational energy is not complete within the lifetime of VHP [~20 ps (see Figure 3.36)]. For each pathway, the formation probabilities of the final products on the 200 ps time scale are reported from using the PES with $D_e^{OO} = 22$ kcal/mol. Of all of the trajectories analyzed, 26 % lead to direct elimination of OH and 47 % roam around CH_2CHO, out of which 25 % (i.e., 12 % of the total) lead to OH elimination and 75 % (i.e., 35 % of total) form GA. The remaining 22 % directly yield GA, and
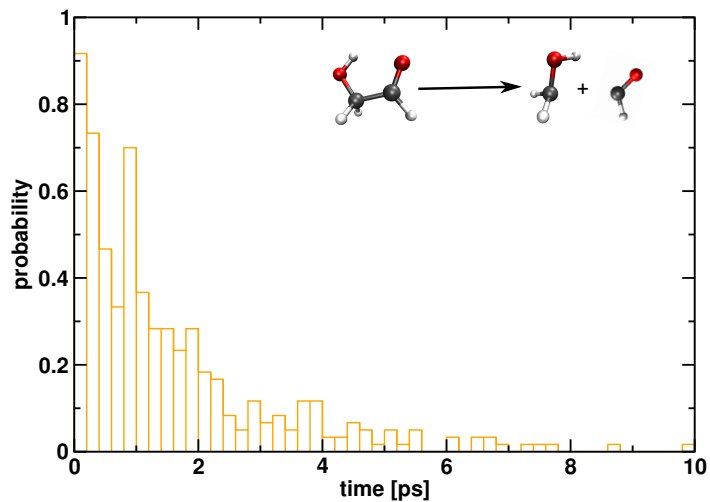
Figure 3.34: Lifetime distribution of glycolaldehyde before dissociating to CH_2OH + HCO using the NN-PES with D_e^{OO} = 22 kcal/mol (CH-stretch excitation with 5988 cm^{-1}).
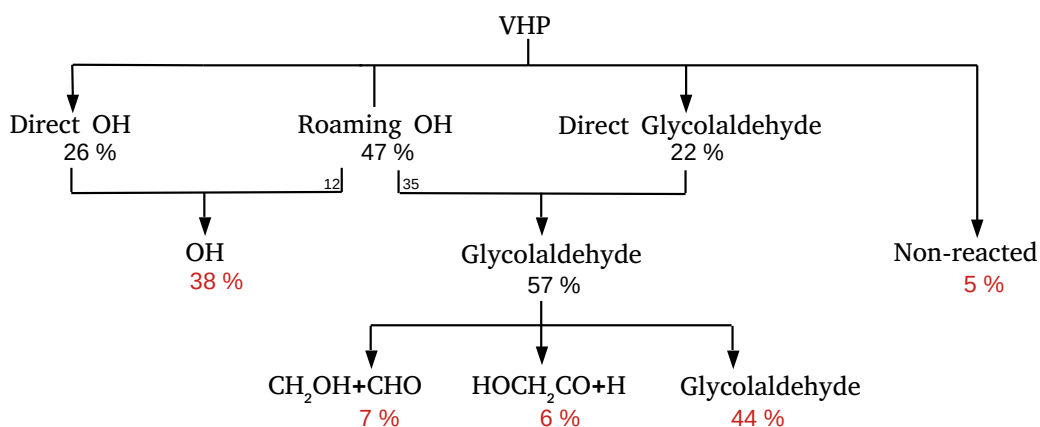
Figure 3.35: Branching ratios of CH_2CHO, OH, CH_2OH, HCO radicals and GA products observed from VHP using the NN-PES with a D_e^{OO} of 22 kcal/mol at 5988 cm^{-1} on the 200 ps time scale. The OH radical is formed from both direct and roaming trajectories, and CH_2OH + CHO is observed from GA.

97
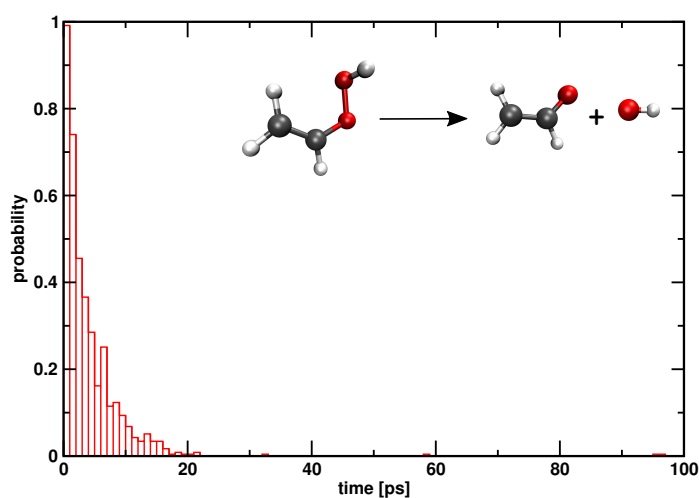
Figure 3.36: Lifetime distribution for VHP before dissociating to CH₂CHO + OH from simulations using the NN-PES with D_e^{OO} = 22 kcal/mol (CH-stretch excitation with 5988 cm^{-1}).

5 % remain in VHP on the 200 ps time scale but may undergo further chemical processing on longer time scales. From the total GA population formed (57 % of all trajectories), 7 % decay to CH₂OH + HCO, 6 % end up in CH₂OHCO + H, and the remainder end up in GA. It should be noted that branching ratios depend on the total simulation time considered. For example, it is expected that all GA formed in the gas phase will eventually decay, because sufficient energy is available.

3.2.4 Discussion and Conclusions

This work establishes that photodissociation of the energized *syn*-CH₃CHOO can lead to additional products besides the known CH₂CHO + OH fragments. This is possible because after O–O scission the OH radical can roam around the CH₂CHO radical to yield GA, from which other dissociation channels become accessible, e.g., leading to the CH₂OH + HCO or CH₂OHCO + H fragments. This discovery was possible because the computationally efficient NN-PES makes statistically significant numbers of sufficiently long reactive MD simulations possible, which are unfeasible
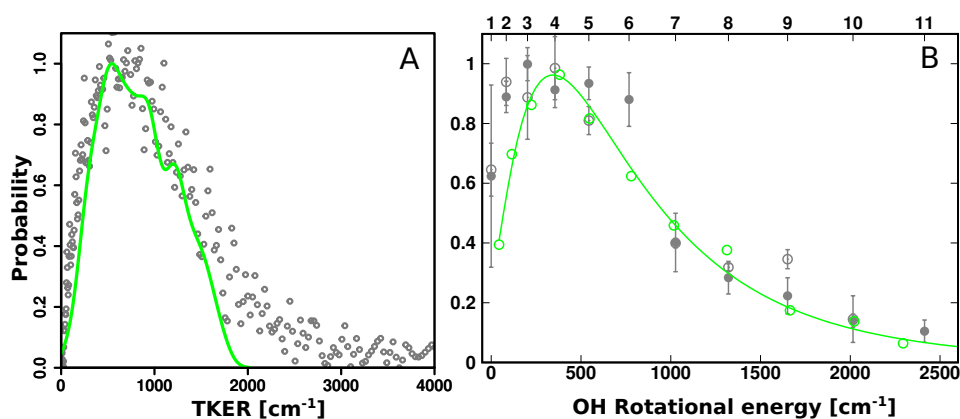
Figure 3.37: TKER (Panel A) and rotational state distribution (Panel B) for direct OH formed (CH-stretch excitation with 5988 cm^{-1}) using the NN-PES with D_e^{OO} = 28 kcal/mol compared with experiments.

with *ab initio* MD simulations at the required level of theory (CASPT2).

The translational kinetic energy and rotational state distributions, *P*(TKER) and *P*(*N*), respectively, of the OH radical from these simulations agree well with experiment. This finding differs from earlier efforts that initiated the dynamics at transition states and found only qualitative agreement, in particular for *P*(TKER).[43] On the basis of the final state distributions, O–O scission energies ranging from 22 to 25 kcal/mol are recommended from this work, which compares with the value of ~25 kcal/mol from electronic structure calculations at the CASPT2 level of theory.[43] With yet larger values for the O–O scission energy [e.g., D_e^{OO} = 28 kcal/mol (see Figure 3.37)] in particular *P*(TKER) deviates appreciably from experiment.

The waves superimposed on *P*(TKER) are spaced by ~1000 cm^{-1}. These could be related to the C–C stretch mode in CH₂CHO that is reported at 917 cm^{-1}.[216] Such an assignment is further corroborated by separately analyzing the C–C distribution functions for ~100 trajectories with the lowest and highest TKER values from Figure 3.26B. For a large kinetic energy release, less energy is available to distribute over the internal degrees of freedom, which leads to a smaller average C–C separation and a narrower distribution (see brown and cyan distributions in Figure 3.38).

Roaming times of OH before recombination to form GA are in the range of picoseconds, consistent with earlier reports on roaming in nitromethane.[217] GA formation
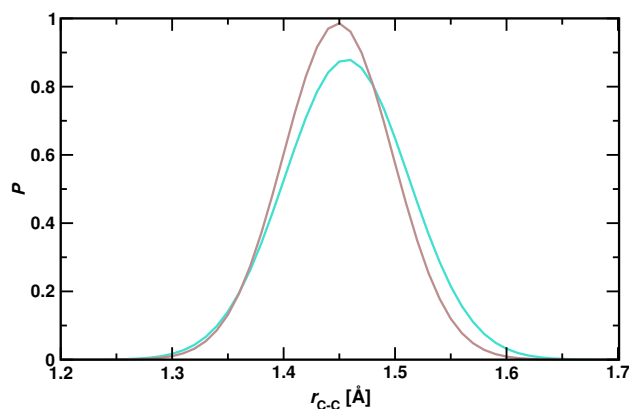
Figure 3.38: C-C bond distance distributions for the H$_2$C-CHO product for trajectories with low-(cyan) and high-(brown) TKER from analyzing 140 trajectories using the NN-PES with D_e^{OO} = 22 kcal/mol (CH-stretch excitation with 5988 cm^{-1}).

from photodissociation of *syn*-CH$_3$CHOO is therefore another case for "molecular roamers" (here OH).[217] As reported earlier, OH elimination under atmospheric conditions takes place on longer (microsecond) time scales.[42] Because OH roaming to form GA occurs on considerably shorter (picosecond) time scales, the processes considered in this work are also likely to be relevant to the chemical evolution of the atmosphere, depending on the environmental conditions (local density), and will lead to a rich and largely unexplored chemistry.

Another notable finding of this work is that the maxima and widths of P(TKER) and $P(N)$ are sensitive to the O–O scission energy whereas energy-dependent rates $k(E)$ are not. Assuming that $k(E)$ follows an equilibrium Arrhenius-type expression would lead to an expected increase in the rate by a factor of ~150 at room temperature, i.e., by ≈2 orders of magnitude, upon reducing the scission energy by 3 kcal/mol. Rather, and in agreement with experiment, these simulations lead to a change in the rate by a factor of only 2 – 5 (see Figure 3.29). Therefore, elimination of OH from VHP is a non-equilibrium process. Because the reaction is initiated through vibrational excitation and intramolecular vibrational energy redistribution occurs on considerably longer time scales across all internal degrees of freedom than when the reaction takes place (Figure 3.36), $k(E)$ is not particularly sensitive to the O–O scission energy. This also suggests the entire pathway from the initial reactant (here, *syn*-CH$_3$CHOO) to the products (here CH$_2$CHO + OH, CH$_2$OHCO + H, and

$CH_2OH + HCO$) should be followed.

Very recently, roaming has been implicated in the formation of hydroxybutanone following H transfer and unimolecular decay of the methyl-ethyl-substituted Criegee intermediate (MECI) from experiments, RRKM, and minimum energy path ("$T = 0$") calculations.[53] The roaming pathway was estimated to be populated in 7 % of the cases compared with OH elimination to yield 2-butanone-3-yl + OH for the remainder. This work also employed multireference electronic structure calculations as was done here. We also found OH elimination and formation of GA via OH roaming and recombination in our study. This confirms that OH roaming and formation of hydroxycarbonyl products are general results of unimolecular decay of CIs. Contrary to the work on MECI, our findings are based on extensive and statistically significant numbers of atomistic simulations using a full-dimensional PES that also provide molecular-level insights into the underlying dynamics. In this case, roaming is found for ~30 % of the GA-forming trajectories, which may, however, still change if the quality of the PES can be further improved.

In summary, this work provides novel molecular-level insights into the non-equilibrium reaction dynamics of OH elimination and roaming and formation of GA following vibrational excitation of *syn*-CH_3CHOO from statistically significant numbers of reactive trajectories on a full-dimensional, reactive PES at the CASPT2 level of theory. Running even a single *ab initio* MD simulation at the CASPT2 level of theory is computationally infeasible as it amounts to $\sim 10^6$ hours of CPU time for one 200 ps run. The final TKER and OH rotational distributions agree well with experiment and bracket the O–O scission energy between 22 and 25 kcal/mol. The widths and position of the maxima of the final state distributions depend on the value of D_e^{OO}, whereas the energy dependent rate $k(E)$ does not. Roaming of the OH product can yield GA on the picosecond time scale, which makes this a relevant alternative reaction pathway to be considered in atmospheric chemistry. Capitalizing on the advances to generate accurate, full-dimensional reactive PESs[218,219] paired with transfer-learning strategies[152,220] and integration into broadly applicable molecular simulation software[221] with techniques to enumerate all possible fragmentation products of a molecule,[222] the routine exploration and discovery

of the decomposition dynamics of small to medium-sized molecules becomes possible.

3.2.5 Supporting Information

Reactive PESs

The transition state for H-transfer connecting *syn*-CH_3CHOO and CH_2CHOOH (see Figure 3.20) was found at 16.0/18.6/19.8 kcal/mol using MP2[48]/CCSD(T)[43]/CASPT2 levels of theory, respectively, compared with an estimated barrier from experiment of ~18.0 kcal/mol based on the lowest photon energy leading to OH products.[36] For the O–O scission energy the computed values are 31.5/18.8/26.2 kcal/mol at the MP2/MRCISD(4,4)/CASPT2 levels of theory.[43,45,48] However, no direct information on the O–O scission energy and OH-roaming is available from experiments.

The analysis of the NN-PES available from previous work[48] motivated the following specific improvements. The NN-PES, based on the PhysNet architecture,[143] was transfer learned[152,220,223,224] to the CASPT2(12,10)/cc-pVDZ level of theory based on reference energies, forces and dipole moments for 26,500 structures computed by MOLPRO 2019.[134] Transfer-learning[224] starts from an available NN (base model) which was trained to MP2/aug-cc-pVTZ data in the present case.[48] Using the weights and biases of this base model a new NN-PES was trained using new and physically more meaningful data including multireference effects, here CASPT2. The new reference data set contains 16000 structures randomly chosen between CH_3CHOO and OH elimination from the MP2 data set and additional 10500 OH roaming, GA and CH_2OH + CHO structures. Out of a total of 26500 structures, 21000 structures were used for training and 2750 each for validation and testing. After training, the quality of the transfer-learned PES to the CASPT2(12,10)/cc-pVDZ level of theory is shown in Figure 3.21. Over a range of more than 200 kcal/mol the root mean squared error for 2750 test structures is 1.56 kcal/mol for all species involved, which establishes the high quality of the transfer-learned PES. A second validation is provided in Figure 3.22 which reports CASPT2 and NN-PES energies along part of a roaming trajectory. Again, the NN-PES accurately captures the energies from *ab initio* calculations with an RMSE of 0.6 kcal/mol and $r^2 = 0.98$

between the two data sets.

Evaluating the stationary points on the transfer-learned NN-PES reproduces energies from electronic structure calculations to within ≤ 0.3 kcal/mol except for the fragmentation products $CH_2OHCO + H$ and $CH_2OH + HCO$ for which only ~100 structures were included. Both channels were not of primary interest in the present work. The dissociation barrier for the O–O bond at the CASPT2(12,10)/cc-pVDZ level of theory starting from the VHP equilibrium to a separated vinoxy and OH radical complex is ~22 kcal/mol which is consistent with earlier work.[43]

In order to assess the sensitivity of the final observables such as $P(\text{TKER})$ or $P(N)$ on D_e^{OO}, the O–O dissociation energy of the NN-PES was modified by adding an energy term $f(r_{OO}) = \frac{\Delta V_s}{1+e^{-b(r_{OO}-r_e)}}$ to the transfer-learned NN-PES. This increases the dissociation energy by ΔV_s. The sigmoidal function depends on the O–O separation r_{OO} and the parameters $\Delta V_s = 3$ kcal/mol, $b = 14.0$ Å^{-1}, $r_e = 1.95$ Å increase the dissociation energy and influence the slope and the turning point of the curve, respectively. The energy correction $f(r_{OO})$ is close to 0 for *syn*-CH₃HCOO and VHP conformations but increases to $\Delta V_s = 3$ as the OH radical dissociates.

Molecular Dynamics Simulations

MD simulations using the NN-trained PES were carried out with pyCHARMM linked to c47a2 to make NN-energies and forces available in the MD code.[221,225] The implementation was validated for energy conservation in *NVE* simulations.[221]

All production MD simulations, each 200 ps in length, were carried out in the *NVE* ensemble with a time step of $\Delta t = 0.1$ fs to conserve total energy. For each variant of the NN-PESs distinguished by the value for D_e^{OO} at least 6000 independent trajectories were run. To generate a thermal ensemble reflecting experimental conditions[43] (low-temperature ensemble in its ground electronic and vibrational state), agnostic to any mode-specific vibrational excitation, initial conditions were generated from *NVT* simulations at 50 K from which structures and velocities were extracted at regular intervals, similar to MD simulations for the same system in previous work.[43]

Subsequently, the velocities of the atoms involved in the C–H bond closest to the accepting oxygen O_B were scaled along the CH-local mode commensurate with energies ranging from 5603 to 6082 cm^{-1}. This corresponds to ~2 quanta in the CH-stretch vibration and is consistent with the experiment.[43]

Depending on the energy used to initiate the nonequilibrium dynamics the fraction of *syn*-Criegee that forms VHP increases from 85 % at 5603 cm^{-1} to 99.5 % at 5987 cm^{-1}. To provide more insight into the distribution from which the initially prepared ensemble develops, the reaction time and corresponding CH bond length distributions $P(r_{CH})$ for the 1,4 H-shift at 5603 cm^{-1} are reported in Figure 3.39. The maximum probability for H-transfer is at ~50 ps after vibrational excitation (Figure 3.39A) and extends out to 500 ps. Figure 3.39B reports $P(r_{CH})$ for thermal equilibrium at 900 K (black) which is the averaged temperature from 1000 simulations with CH-excitation at 5603 cm^{-1}. This is compared with $P(r_{CH})$ for trajectories that feature H-transfer within 3 ps (blue), 25 ps (orange), and 135 ps (red) after vibrational excitation and no H-transfer within the maximum simulation time of 500 ps (green trace). For early reaction times (3 ps, 25 ps) the distributions $P(r_{CH})$ are noticeably broader than that from thermalized simulations. Hence, the system develops from a nonequilibrium initial state. As the reaction time increases to 135 ps and for simulations without H-transfer within 500 ps, $P(r_{CH})$ approaches that of thermal equilibrium. Thus, from the perspective of the excited CH-stretch excitation the initial ensemble develops from a mixture of [non-equilibrium / equilibrium] initial conditions of [60/40]% assuming a time t^* =100 ps for separating non-equilibrium from equilibrium distributions. This changes to [30/70]% for $t^* = 50$ ps which is the maximum of the reaction time distribution in Figure 3.39A.

To account for ZPVE in the reactant, additional simulations with excitation energy including the difference ΔE in ZPVE between *syn*-CH₃CHOO and the transition state for H-transfer were carried out. At the MP2/6-31G(d,p) level of theory $\Delta E =$ 2.1 kcal/mol.

The simulations used a hybrid protocol. For the H-transfer to form VHP from *syn*-CH₃CHOO the earlier MS-ARMD representation[48] featuring the correct 16.0
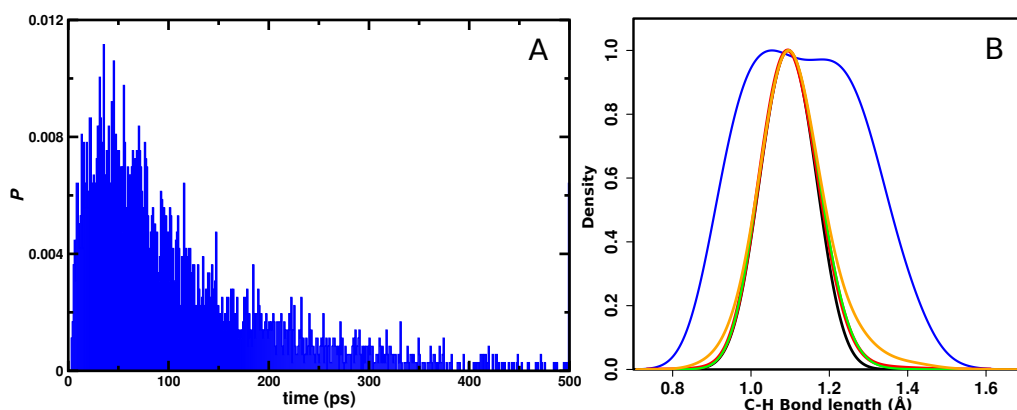
Figure 3.39: Normalized reaction time distribution (panel A) for 1,4 H-shift at 5603 cm^{-1} and C–H bond distance distributions (panel B) for *syn*-CH$_3$CHOO. The $P(r_{CH})$ in panel B were obtained from thermal simulations at 900 K (black) and are compared with simulations following excitation of the CH-stretch in which H-transfer occurs at 3 ps (blue), 25 ps (orange), 135 ps (red), and no transfer within 500 ps (green) after excitation.

kcal/mol barrier was used. Then, the positions $\vec{x}$ and momenta $\vec{p}$ of all atoms were stored and the simulation was restarted using the PhysNet representation. For consistency, a scaling λ for the momenta $\vec{p}$ of all atoms was determined from

$$T(\vec{p}) + \Delta V_{\text{MS–ARMD}}(\vec{x}) = T(\lambda \cdot \vec{p}) + \Delta V_{\text{PhysNet}}(\vec{x}) \tag{3.1}$$

to match the sum of the kinetic energy T and the potential energy difference ΔV between the potential at atom positions $\vec{x}$ and the VHP equilibrium structure for both PESs.

Final State Analysis

For the final state analysis, the total energy of the separated system was decomposed into fragment translational (E_{trans}), rotational (E_{rot}), and vibrational (E_{vib}) energy components. The measured TKER is the sum of the OH and vinoxy radical fragments translational energy contributions $E_{\text{trans},\alpha} = \vec{p}_{\text{CM},\alpha}^{\,2}/2M_\alpha$ where M_α is the mass and $\vec{p}_{\text{CM},\alpha}$ the center of mass momentum of fragment α obtained as the sum of the respective atom momenta. The rotational energy E_{rot} of the dissociating OH radical is $E_{\text{rot}} = L^2/2I$ with the angular momentum vector $\vec{L} = \sum_{i=1}^{2}(\vec{r}_i - \vec{r}_{\text{CM}}) \times \vec{p}_i$, where $\vec{r}_i - \vec{r}_{\text{CM}}$ is the atom position with respect to the center of mass position, $\vec{p}_i$ are

	Note	P(TKER)		P(N)	
		max	FWHM	max	FWHM
$D_e = 22$ kcal/mol	all	390	1220	$N = 5$	1390
	direct	588	1522		
	ZPVE (all)	475	1125	$N \in [3, 5]$	1090
$D_e = 25$ kcal/mol	all	360	1170	$N = 4$	1260
	direct	666	1544		
H-TS[43]		1550	2595	$N = 3$	860
Submerged-TS[43]		880	2605	$N = 3$	630
Expt.[43]		600	1340	$N = 3$	1055

Table 3.7: Comparison of the position of the maxima ("max") and full widths at half maximum (FWHM) for *P*(TKER) and *P*(*N*) from the present simulations, from experiment[43] and previous simulations starting from transition states along the pathways as indicated.[43] Widths and peak maxima from Ref.[43] were determined by extracting numerical data using g3data.[226] Notes "all" and "direct" refer to analyzing directly dissociating + roaming trajectories and only directly dissociating trajectories, respectively. "ZPVE" indicates analysis of trajectories including zero-point vibrational energy restrictions in reactant and products.

the atoms momenta and *I* is the OH moment of inertia. As the exchange of energy between vibrational and rotational degrees of freedom persists in the presence of non-zero internal angular momentum, E_{rot} is averaged over the final 25 ps after dissociation. The TKER and average E_{rot} become constant after both fragments have moved further than the interaction cutoff of 6 Å from each other. The vibrational energy of the OH fragment was then computed according to $E_{\text{vib}} = E_{\text{OH}} - E_{\text{trans}} - E_{\text{rot}}$ with $E_{\text{OH}} = T + V(r_{\text{OH}})$ which was invariably close to the quantum mechanical ground state vibrational energy for the respective rotational state of OH. Hence, final OH products are always in $v_{\text{OH}} = 0$.

Motivating the Value for t_e

To motivate the value of $t_e = 0.8$ ps, the distribution of dissociation times using NN-PESs with $D_e^{\text{OO}} = 22$ and 25 kcal/mol are reported in Figure 3.23. For both distributions, a step at $t_e > 0.8$ ps is found. Furthermore, a monotonic increase in the O–O separation with time is observed in the trajectories with $t_e < 0.8$ ps, see Figure 3.24A. On the other hand, for a "roaming elimination" shown in Figure 3.24B

the OH fragment first reaches an O–O separation of ~6 Å and re-approaches to 2.5 Å before finally dissociating from CH_2CHO. It is noted that there is in general some ambiguity in using geometrical criteria for differentiating between different types of trajectories and other definitions would be possible than the one used here. For a positionally resolved picture of the different types of OH dissociation, density maps of OH radical around CH_2CHO are shown in Figure 3.25. Roaming trajectories (Figure 3.25B) differ from direct dissociation (Figure 3.25A) by the coverage of the surrounding space by the OH radical around the vinoxy group. Unlike direct dissociation, roaming OH explores the vicinity of the vinoxy radical before complete fragmentation, see Figure 3.25.

Rates

To determine reaction rates, for each energy 6000 individual simulations were carried out. The rates were determined from fitting the number $n(t)$ of trajectories that had not reacted by time t to single ($\sim \exp(-kt)$) and stretched exponential decays $\exp(-kt)^\gamma$, see Figure 3.29 for the quality of the two fits.
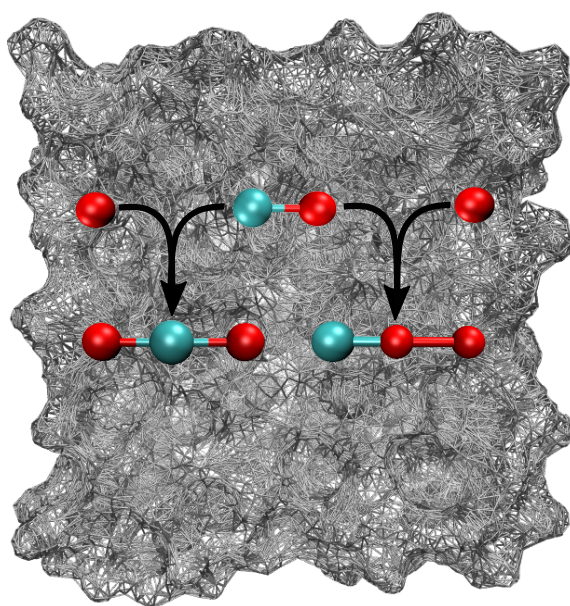
Chapter 4

Atom–Diatom Collisions on Cold Amorphous Solid Water

4.1 Publication: Genesis of Polyatomic Molecules in Dark Clouds: CO_2 Formation on Cold Amorphous Solid Water

4.1.1 Abstract

Understanding the formation of molecules under conditions relevant to interstellar chemistry is fundamental to characterize the chemical evolution of the universe. Using reactive MD simulations with model-based or high-quality PESs provides a means to specifically and quantitatively probe individual reaction channels at a molecular level. The formation of CO_2 from collision of $CO(^1\Sigma)$ and $O(^1D)$ is characterized on ASW under conditions typical in cold molecular clouds. Recombination takes place on the subnanosecond time scale and internal energy redistribution leads to stabilization of the product with CO_2 remaining adsorbed on the ASW on extended time scales. Using a high-level, reproducing kernel-based PES for CO_2, formation into and stabilization of CO_2 and COO is observed.

4.1.2 Introduction

The formation and chemical evolution of stars and galaxies is intimately linked to the presence of molecules.[61] This "Molecular Universe"[61,227] requires molecules to be generated "bottom-up" but also causes them to be destroyed "top-down" by reaction channels that are potentially specific to the conditions in the ISM. For the formation of neutrals, grain-surface chemistry involving interstellar ices inside dense molecular clouds is of particular importance.[61] Reactions on icy surfaces occur at the diffusion limit, and their efficiency depends primarily on the diffusion of one or several of the reacting species on the surface and whether or not particular reaction channels involve entrance barriers. In dense cold clouds, CO is the main carbon reservoir[61] and the second most abundant molecule in molecular clouds with an important role as a molecular tracer for probing and characterizing the chemical and physical conditions of the environment.[228]

In the interstellar medium, water ice grains were found in 1973[229] and direct detection of CO_2 was reported 30 years ago.[230] More recent observations indicate that H_2O and CO_2 are also sufficiently abundant on grains.[231,232] The chemical precursor for formation of CO_2 is believed to be carbon monoxide and the CO + O reaction has been proposed as a nonenergetic pathway, close to conditions in interstellar environments, for CO_2 formation 20 years ago.[233] Thermoluminescence experiments of added CO to photolyzed N_2O in an Argon matrix at 7 K suggested that the $O(^3P)$ + $CO(^1\Sigma^+)$ reaction yields excited CO_2^*, which, after emission of a photon, leads to formation of CO_2.[234] Such a process has also been proposed to occur on interstellar grains[63] and has been confirmed experimentally[75] with an estimated entrance barrier of 0.014 eV to 0.103 eV for the process on ASW, compared with a value of 0.3 eV from high-level electronic structure calculations.[235] The surrounding water matrix provides the necessary coupling[233] to facilitate relaxation of the $^3A'$ or $^3A''$ states of CO_2 to the $^1A'$ ground state (correlating with linear $^1\Sigma_g^+$). On the other hand, the presence of an entrance barrier for the $O(^3P)+CO(^1\Sigma^+)$ reaction has led to the consideration of the alternative CO+OH pathway for CO_2 formation[236] which was, however, reconsidered to yield the HOCO intermediate in such environments in more recent experiments.[237] Also, the reaction products of the CO+OH reaction

have been found to depend on the experimental conditions.[238,239] As the CO+OH reaction appears to have a barrier in the entrance channel,[239] attention has recently shifted to the HOCO+H reaction for CO_2 formation.[237]

For atomic oxygen, the possibility for diffusion on and within ASW was demonstrated down to temperatures of 10 K and 50 K from both experiments and simulations.[69,71,240] Typical temperatures of dense cold clouds have been estimated at ~15 K.[241] The oxygen mobility found experimentally even at such low temperatures[69] opens up ways to form small molecules, including O_2, O_3, or CO_2 through atom/atom or atom/molecule recombination. In this context, it is also of interest to note that experiments reported a higher mobility for excited $O(^1D)$ compared to the $O(^3P)$ ground state.[242,243] For molecular oxygen, atomistic simulations showed that both ground and excited state O_2 can be stabilized on ASW.[72,244]

Direct, barrierless recombination into the ground state of CO_2 is possible from $O(^1D) + CO(^1\Sigma^+)$. Excited $O(^1D)$ can be formed from photolysis of H_2O,[245] and the radiative lifetime is 110 minutes.[246] Also, experiments using electron-induced neutral dissociation of water into H_2 and $O(^1D)$ in the presence of CO have recently reported the formation of CO_2 in cryogenic CO/H_2O films.[247] Here, the reaction $O(^1D) + CO(^1\Sigma^+) \rightarrow CO_2(^1\Sigma_g^+)$ on the surface of ASW is investigated at a molecular level. Two PESs are used to describe the energetics of $CO + O \rightarrow CO_2$ formation (see Supporting Information for the energy functions used). One is a Morse-Morse-Harmonic (MMH) parametrization based on multireference CI calculations for CO-dissociation, fitted to a Morse functional form together with an empirical parametrization for the OCO angle θ. The second PES is a 3-dimensional RKHS representation of CCSD(T)-F12 calculations which (see Figure 4.1 for the quality of the representation) is more accurate than the MMH PES but also considerably more computationally expensive to evaluate. Energy conservation for MD simulations using both PESs is reported in Figure 4.2.

The 3-dimensional MMH-PES for the ground state of CO_2, fitted to MRCI/aug-cc-pVTZ data, has a linear equilibrium geometry with C-O distances of 1.1644 Å and bending, symmetric and asymmetric stretching frequencies at 645, 1226, and 2450
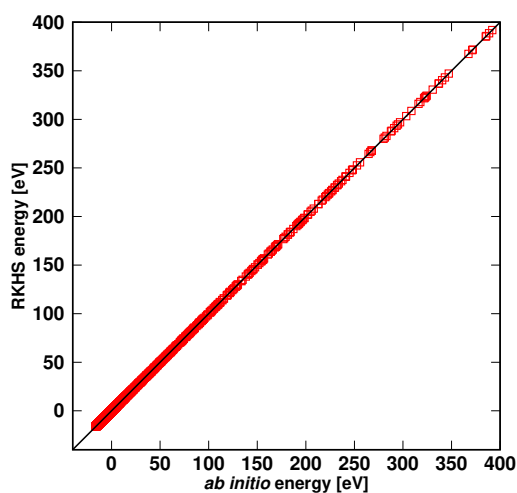
Figure 4.1: Quality of the RKHS representation for the CCSD(T)-F12 points from the electron structure calculations. The RMSE is 0.0026 eV (0.06 kcal/mol) and the $R^2 = 1.0$.
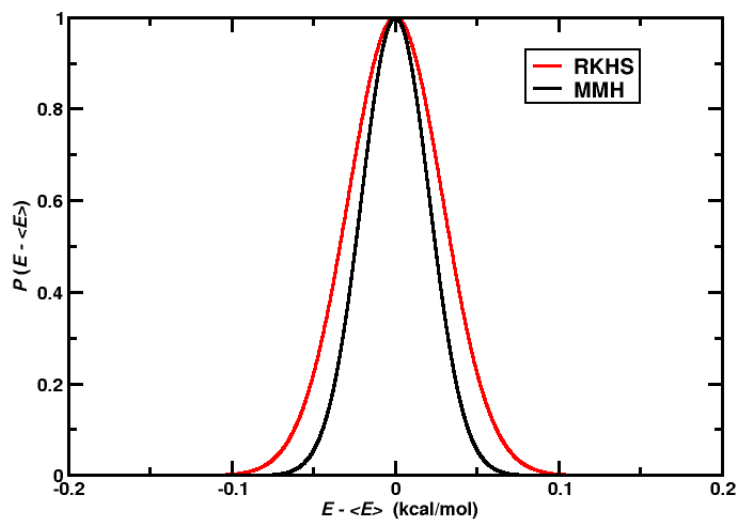
Figure 4.2: Energy conservation for a reactive simulation using MMH (black) and the RKHS representation of the CCSD(T)-F12 calculations (red). The Gaussian shape of the curve confirms the energy conservation.
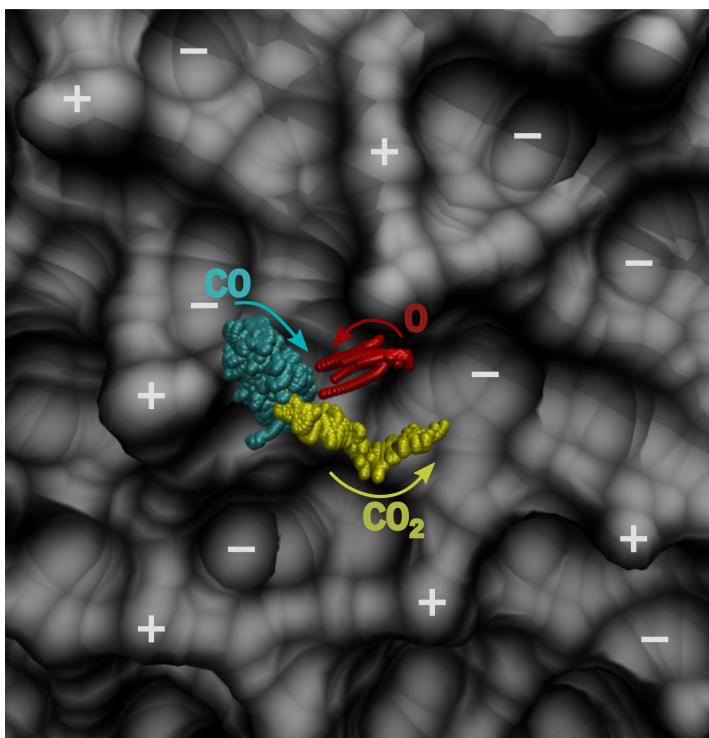
Figure 4.3: Trajectories leading to CO_2 recombination. The O-atom (red) and CO molecule (cyan) from eight different trajectories are shown to form CO_2 (yellow). Before and after recombination, all species diffuse on the ASW (gray colors) surface. The "plus" and "minus" signs indicate protuberances and indentations of the ASW surface.

cm^{-1}, respectively. This compares with an equilibrium separation of 1.1621 Å and frequencies at 667 cm^{-1} for bend, 1333 cm^{-1} for symmetric stretch and 2349 cm^{-1} for asymmetric stretch from experiments.[248] Thus, the MMH-PES is qualitatively correct but not spectroscopically accurate which is, however, not necessary for the present study. The differences in the frequencies are related to the level of theory at which the reference energies were calculated.

First, recombination simulations were run with the MMH model. This provides a qualitatively correct description of the CO_2 formation dynamics at reduced computational cost. Typical recombination trajectories projected onto the ASW surface are shown in Figure 4.3. Before recombination, CO (cyan) and O (red) diffuse separately in their respective adsorption sites on the ASW. After recombination, CO_2 (yellow) continues to diffuse on the ASW.
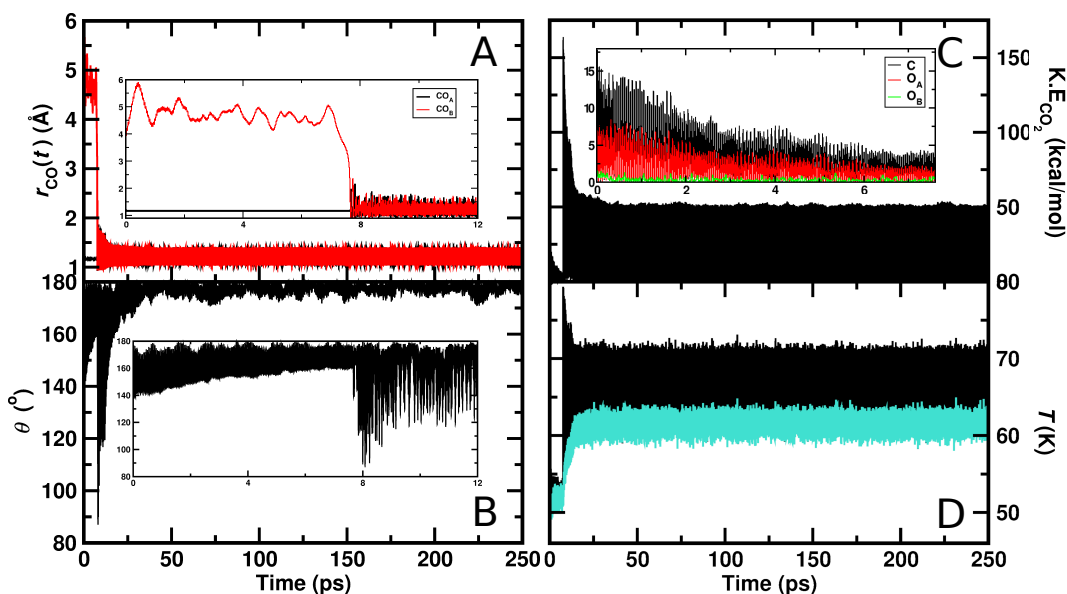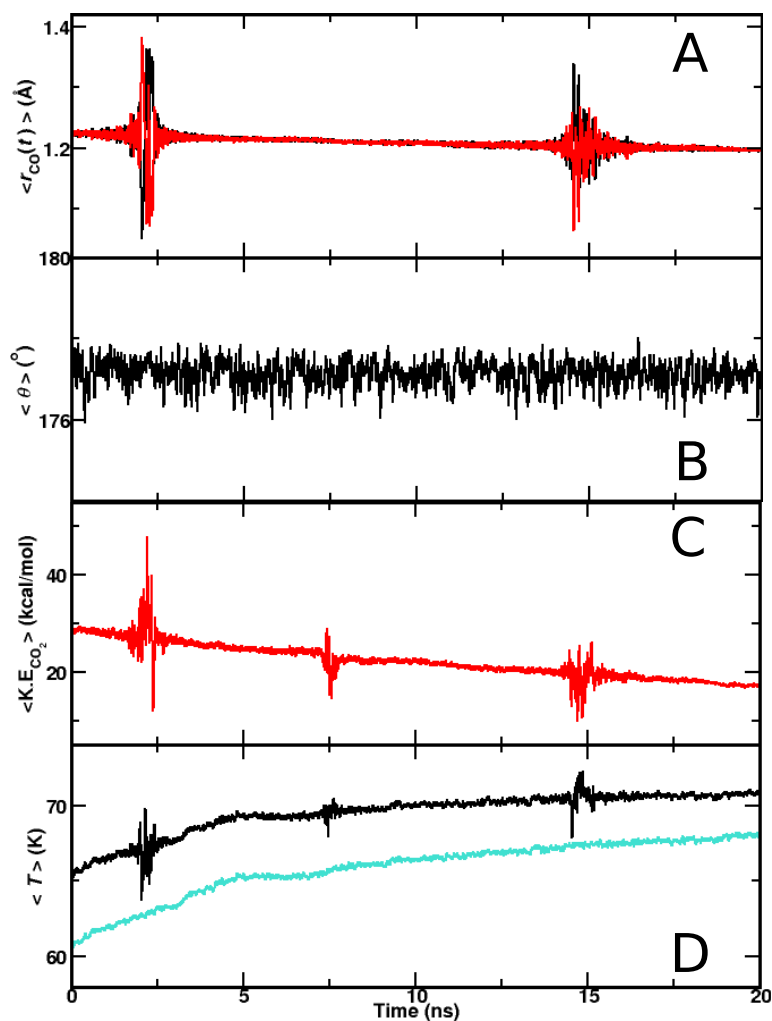
Figure 4.4: Interatomic distances CO_A and CO_B (panel A) and the OCO angle θ (panel B, running average) from a 250 ps simulation using the MMH PES. Initially, $R = 4.66$ Å and $\theta = 135°$. The diffusive motion of O_B can be seen in the inset before recombination takes place after ~7 ps. After recombination, both CO stretches are equally highly excited. Panel C: Kinetic energy of CO_2 along with the contribution from each of the atoms. Panel D: Average temperature of the ASW (cyan) and the full system (black) before and after CO_2 formation. After recombination there is a steep increase in temperature of the entire system (CO_2 plus ASW) whereas warming of the ASW is more gradual. This suggests that the underlying process is vibrational relaxation of hot CO_2 on a cool ASW surface which gradually warms and assumes a new thermal equilibrium.

117

A representative time series for $CO_A + O_B$ to form CO_2 is reported in Figure 4.4. From an initial separation of $R = 4.66$ Å and $\theta = 135°$, CO_2 is formed within ~8 ps (inset of Figure 4.4A). Upon recombination, both CO stretch coordinates are highly excited and the OCO angle fluctuates around $\theta = 180°$ with an amplitude of ~10° (Figures 4.4, parts A and B). For the remaining 250 ps, CO_2 relaxes slowly and remains as a diffusing product on the ASW (see also Figure 4.3).

It is also of interest to follow the kinetic energy and the temperature of the ASW before and after recombination, see Figures 4.4C and D. Before $CO_A + O_B$ recombination the average temperature of the water molecules and the entire system are close to one another and fluctuate around 50 K. Upon recombination, the temperatures as determined from the kinetic energies of the oxygen and carbon atoms increases considerably by about 10 K on average and assume different values as the system now consists of a "hot" CO_2 molecule adsorbed on a cool ASW surface. The temperature of the full system (recombined CO_2 and the ASW) first shows a prominent peak reaching 75 K right after recombination and relaxing subsequently. Contrary to that, the temperature of the water molecules does not show such a spike but rather increases gradually from 50 K to 60 K following relaxation of the CO_2 molecule (Figure 4.4D).

To better characterize relaxation of the internal energy, longer (20 ns) simulations were also run. Figure 4.5 reports the CO bond lengths (panel A), the OCO angle (panel B), the kinetic energy of CO_2 (panel C), and the temperatures of the total system (panel D, black) and the ASW (panel D, cyan). During the entire 20 ns, slow relaxation of the CO stretch amplitudes takes place, interrupted by occasional large amplitude motions, caused by scattering of the CO_2 from collisions with the water molecules. Such collisions are accompanied by increase of the kinetic energy of the entire CO_2 molecule. The raw data for $\theta(t)$ is reported in Figure 4.6. Energy exchange between the vibrationally excited CO_2 molecule and the ASW surface continues out to 20 ns and beyond which is reflected in the continued warming of the water molecules. However, the slope of the cyan curve in Figure 4.5D is steeper during the first 5 ns and then flattens out for the remainder of this trajectory.

Figure 4.5: Interatomic distances CO_A and CO_B (panel A), OCO angle (panel B), kinetic energy of CO_2 (panel C) and ASW (cyan) and overall system (black)(panel D) from 20 ns simulation. The spikes are due to collisions of CO_2 with the surrounding water matrix.
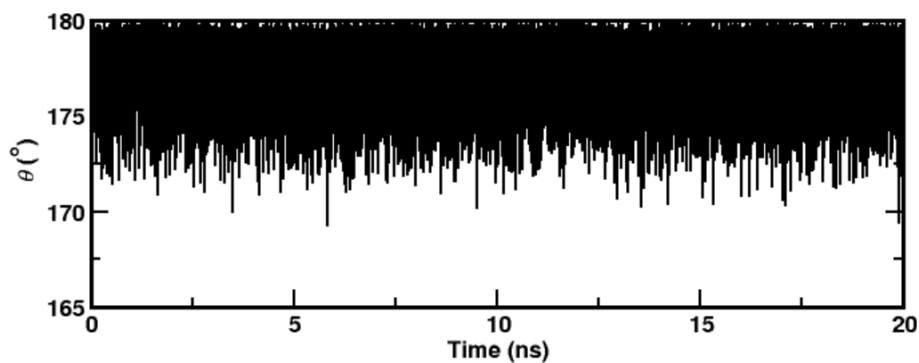
Figure 4.6: Time series for $\theta(t)$ from 20 ns simulation, reporting the raw data of Figure 4.5B

.
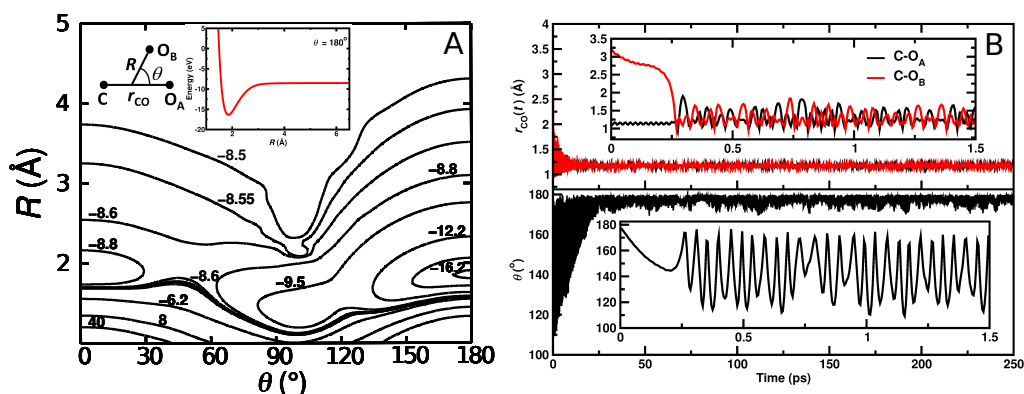
Figure 4.7: Panel A: Two-dimensional cut through the 3d RKHS PES at the CCSD(T)-F12 level of theory for the CO + O channel at fixed r_{CO_A} = 1.21 Å. The inset reports the 1d cut along θ = 180° for r_{CO_A} = 1.21 Å. All energies in eV and the zero of energy at the dissociation into atomic fragments. Panel B: Interatomic distances CO_A and CO_B (top panel, for labels see panel A) and the OCO angle (bottom panel) from a 250 ps simulation with R = 3.9 Å and θ = 180° using the RKHS PES. Initially, the CO_A bond (black) is fluctuating around its thermal (50 K) equilibrium separation but the amplitude increases considerably after recombination due to internal vibrational relaxation. The OCO angle relaxes to a quasi-linear structure on the 25 ps time scale.

The MMH simulations discussed so far demonstrate that CO + O collisions lead to formation of CO_2 which relaxes on time scales longer than nanoseconds and does not desorb from the ASW surface.

Corresponding simulations for recombination into ground state CO_2 were carried out with the more realistic but computationally more expensive RKHS[139] representation of the CCSD(T)-F12 PES (see Supporting Information). Figure 4.7 reports the two dimensional PES for r_{CO_A} = 1.21 Å and varying R and θ and shows a deep minimum for the linear OCO structure (θ = 180°) together with the high-energy, metastable COO structure (θ = 0°), 167.7 kcal/mol (7.29 eV) higher in energy than the global minimum. Such a COO intermediate has been proposed from the interpretation of the C+O_2 reaction[249] and was also found in calculations.[235,250] The two structures are separated by a barrier of 9.55 kcal/mol (0.41 eV) in going from the COO to the OCO conformation.

A typical reactive trajectory to form CO_2 using the RKHS PES is reported in Figure 4.7B. Again, recombination leads to a highly vibrationally excited CO_2 molecule
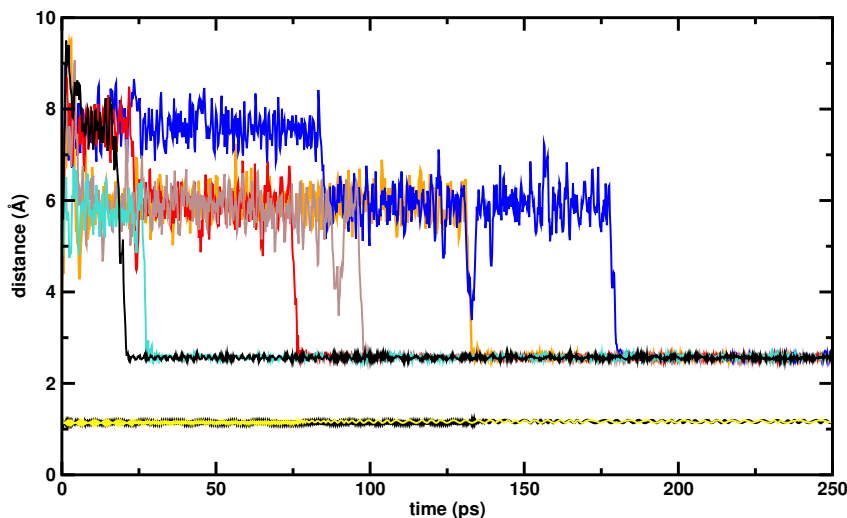
Figure 4.8: Time series for rebinding into the COO conformation from simulations using the RKHS PES.

with rapidly decaying CO stretch amplitude for the CO-bond that is newly formed (red trace) whereas the CO_A vibrates around 1 Å initially and extends to ~1.2 Å upon recombination. The OCO angle relaxes to a quasi-linear structure on the 25 ps time scale, see lower panel in Figure 4.7B. As this 3-dimensional PES also supports a COO conformation, it is of interest to start simulations from initial orientations $\theta = 0°$. Indeed, COO formation and stabilization on the subnanosecond time scale is observed, see Figure 4.8. Such an intermediate is of potential interest as it can lead to C + O_2 formation.

The possible final states after collision include (I) $CO_A + O_B \longrightarrow CO_A + O_B$ (flyby), (II) $CO_A + O_B \longrightarrow CO_B + O_A$ (atom exchange), (III) $CO_A + O_B \longrightarrow CO_2$ (CO_2 formation), (IV) CO_A remains on the surface and O_B desorbs (elastic scattering (ES1)), (V) CO_A desorbs and O_B remains on the surface (ES2), and (VI) both, CO and O desorb from the ASW surface. For each of the initial conditions between 100 and 1000 independent trajectories were run; see Tables 4.1, 4.2, 4.3 and 4.4. In total, 32500 simulations were run to determine the probability and rate for CO_2 formation depending on the initial conditions. All processes I to VI were observed in the simulations described in the following. One of the possible processes that was not
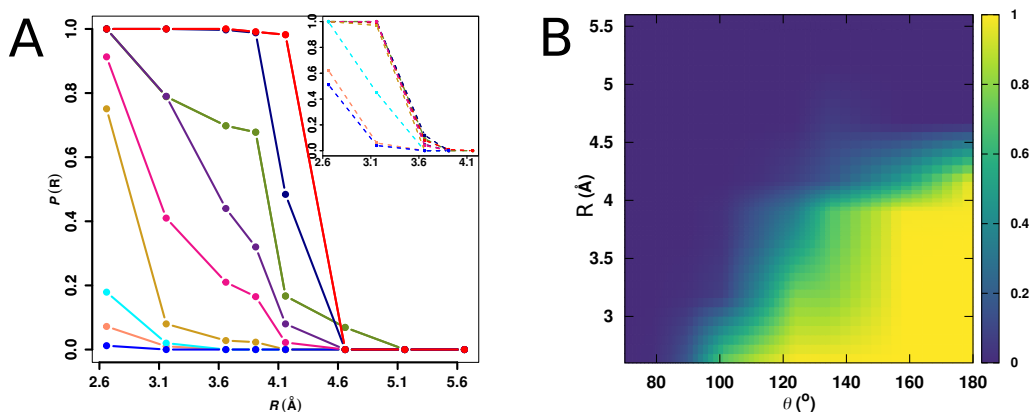
Figure 4.9: Panel A: The CO_2 formation probability as a function of R and θ normalized to the number of trajectories with Morse potential (main panel) and RKHS (inset) where different colors refer to probability for different angles. Color code: darkblue: 78.75°; salmon: 84.375°; skyblue: 90.0°, gold: 101.25°, pink: 112.5°, darkorchid: 123.75°, olive: 135.0°, navy: 157.5°, red: 180.0°. Panel B: Plot for the CO_2 formation probability as a function of R and θ using MMH PES. For a smoother shape or $P(R,\theta)$ kernel density estimation (KDE)[254] was used.

observed on the nanosecond time scale is CO_2 formation with subsequent desorption of CO_2. This is consistent with experiments[251] together with recent computer modeling which found a low (< 0.05) probability for O_2 and CO_2 desorption after the O+O and O+CO reaction, respectively.[72,252,253]

As the main focus of the present work concerns a) whether or not CO_2 is formed, b) whether CO_2 stabilizes and c) what other processes follow the recombination reaction, the initial separations between the two reaction partners to explore these questions was limited to $R \sim 6$ Å. As already found for the $O(^3P)$–$O(^3P)$ reaction even initial separations of ~10 Å can lead to recombination; see Figure 4.8. However, to maintain simulation times manageable for this two-dimensional problem, the initial separations were limited to a somewhat shorter value of R.

Figure 4.9A reports the reaction probability for CO_2 formation depending on the initial reaction geometry (R,θ) for the MMH (main panel) and the RKHS PES (inset). Simulations starting from $\theta = 180°$, i.e. along the $O_A C$–O_B approach have unit reaction probability with initial separations ranging from $R = 2.6$ Å up to $R = 4.5$ Å. Such separations cover both situations, the reacting partners within the same
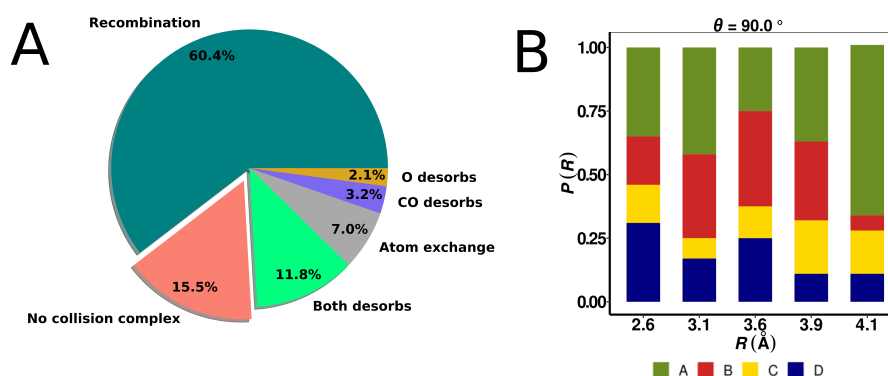
Figure 4.10: Panel A: Classification of outcomes from all the > 30000 simulations using the MMH PES. In more than half the cases CO_2 is formed whereas for 15 % of the cases the two collision partners do not meet. The remaining 25 % include atom exchange reactions, elastic scattering or desorption of both collision partners. Panel B: After atom exchange, there are 4 further possibilities, i.e, (A) both CO and O remain on the water surface, (B) both CO and O desorb, (C) O remains and CO desorbs, and (D) CO remains and O desorbs from the water surface for $\theta = 90°$. See Figure 4.11 for different angles.

and in two different wells. Moving along θ, it is found that on the MMH PES the reaction probabilities rapidly decrease whereas on the RKHS PES they remain close to 1 up to $\theta \approx 100°$ after which they decrease (inset Figure 4.9A). These differences are due to the different topographies of the two PESs.

For the MMH model there is a sufficient number of trajectories on a fine and extensive grid of initial collision geometries to also construct a 2-dimensional probability function for the reaction probability, see Figure 4.9B. This was done by using kernel density estimation[254] (KDE) applied to the raw data from Figure 4.9A. It is found that, as soon as the oxygen atom diffuses within a range of ~4 Å of the center of mass of CO molecule, formation of CO_2 is highly probable.

With respect to the distribution of the final state channels, more than 50 % of the trajectories run on the MMH PES lead to CO_2 and remain in this state, see Figure 4.10A. The multiple long-time (20 ns) simulations indicate that once formed, CO_2 is expected to remain adsorbed on the ASW and vibrationally cools on considerably longer time scales (microseconds to milliseconds). Among the other, minor, channels, the one not forming a collision complex is most probable (16 %), followed by
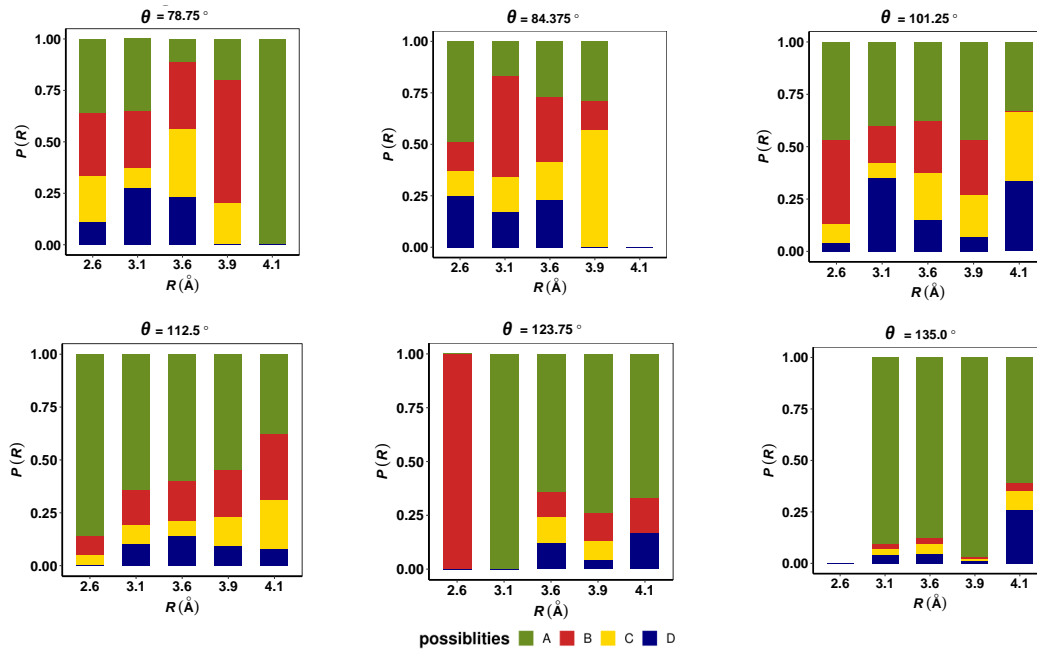
Figure 4.11: After atom exchange, there are 4 further possibilities i.e both CO and O remains on the water surface (A), both CO and O desorbs (B), O remains and CO desorbs (C), CO remains and O desorbs (D) from water surface. Plots for atom exchange probability for each of the grid points using MMH.

desorption of both reaction partners (12 %) and the atom exchange reaction (7 %), whereas desorption of either CO or O occur rarely. As CO_2 does not desorb following reactive CO + O collisions and the desorption energy of CO_2 (~2500 K)[255] is larger than that of CO or O_2, CO_2 may accumulate on ASW. On the other hand, the CO + O reaction also competes with, e.g., the O + O reaction,[75] which limits the amount of CO_2 generated, depending on the relative concentrations (i.e., surface coverage) of CO and O on the grain surface.

The atom exchange reaction is of particular interest because it provides another recombination channel, but probably on longer simulation times, if both reaction partners remain on the ASW. Figure 4.10B shows that CO_B and O_A remaining on the ASW after atom exchange (green) is most probable irrespective of the initial separation between CO_A and O_B. This is followed by desorption of both, whereas desorption of either CO_B or O_A is less likely on average. Thus, even if the initial collision process leads to atom exchange, there is still probability for subsequent

CO_2 formation because both reaction partners remain on the ASW.

4.1.3 Discussion and Conclusions

The present work demonstrates that CO_2 formation and stabilization from collision of CO and $O(^1D)$ on ASW is possible and occurs on the subnanosecond time scale for the reaction partners within 6 Å. This compares with a radiative lifetime of 110 minutes[246] of $O(^1D)$ which can be formed from photolysis of H_2O[245]. Relaxation of the highly excited CO_2 product occurs on much longer time scales which is not covered here. From previous work on $O(^3P) + O(^3P)$ to form O_2 it is known that full vibrational relaxation occurs on the microsecond time scale[72,256] and that the probability for chemical desorption of O_2 on ASW is low.[257] Similarly, it is expected that CO_2 vibrational relaxation extends out to considerably longer time scales than those covered in the present work (tens of nanoseconds).

The time scale for vibrational relaxation also depends on the coupling between the vibrationally excited CO_2 and the water substrate. Here, a rigid water model was used which only offers the phonon bath to couple to. However, previous work for O_2 on ASW showed that vibrational relaxation speeds up by ~30 % if a flexible KKY water model is used.[72,258] The typical water modes that couple in the ~ 2000 cm^{-1} region are the water bend (1600 cm^{-1}) and the framework rotation (600 cm^{-1}) as was also found for the vibrational relaxation of cyanide in water.[259] For $O(^3P)$ + $O(^3P)$ recombination to form O_2 it was found that energy is dissipated into the ASW bulk on the 10 ps time scale within a range of 8 Å around the reaction site.[72] Future simulations for the CO + $O(^1D)$ reaction using a flexible water model, such as KKY, will also provide information about energy migration on and within ASW following CO_2 formation.

Effects due to quantum nuclear dynamics are not included here but are expected to be small given the large mass of the particles involved. In fact, earlier work[71] has demonstrated that the experimentally[69] determined T−dependent mobility of $O(^3P)$ is a consequence of ASW surface roughness rather than tunneling at low temperature. Also, the validity of QCT-based simulations has been explicitly validated for

the [CNO] reactive system by comparing final vibrational state distributions with time-independent quantum simulations.[180] The efficiency of nonadiabatic effects due to $O(^1D)$ to $O(^3P)$ conversion have been estimated[235] to be 10 % at 300 K based on earlier measurements.[260] This may change at lower temperature, though, and explicit inclusion of nonadiabatic processes for the $O(^1D)$ + CO association reaction to form CO_2 appears to be of interest in the future.

The simulations on the RKHS PES support the possibility to form the COO isomer. Early emission spectra found that the $C(^3P)$ + $O_2(^3\Sigma_g^-)$ reaction generates CO with $v' = 17$ and that the transition state has the COO configuration rather than OCO.[261] As both the high-level CCSD(T)-F12 PES and a recently determined MRCISD+Q/aug-cc-pVTZ ground state PES[235] support stabilization of the COO structure, the present results also indicate that association of CO with $O(^1D)$ can lead to formation and stabilization of COO.

Complementary to the $O(^1D)$ + $CO(^1\Sigma^+)$ reaction considered here, the $O(^3P)$ + $CO(^1\Sigma^+)$ process to yield triplet (excited) CO_2 is also feasible. For this, nonenergetic experimental preparations have been reported to yield CO_2.[75,233,262] The $O(^3P)$+$CO(^1\Sigma^+)$ asymptotic state connects to the excited $^3A'$ and $^3A''$ states of CO_2 both exhibiting entrance barriers of 0.2 eV and 0.3 eV in the gas phase, respectively, at the CASSCF-MP2 level of theory[263] which change to 0.3 eV and 0.4 eV from MRCI+Q calculations.[235] Temperature-programmed desorption experiments on ASW indicate that the entrance barrier reduces to between 0.014 and 0.067 eV.[75] Furthermore, symmetry breaking due to the water environment[233,234] makes the transition from triplet to singlet (ground state) CO_2 possible. For the $O(^3P)$ + $CO(^1\Sigma^+)$ reaction equally accurate PESs are available[235] which allow to follow this pathway in molecular detail.

This work demonstrates that the genesis of small polyatomic molecules can be investigated from computations using high-level electronic structure and molecular dynamics methods. For the specific case of CO_2 formation along the $O(^1D)$ + $CO(^1\Sigma^+)$ asymptote the product stabilizes through intermolecular vibrational relaxation and does not desorb from the ASW on the < 100 ns time scale. The fact

that O(^{1}D) with a radiative lifetime[246] of 110 min is considerably more mobile than ground state O(^{3}P)[242,243] makes this process to generate CO_2 feasible. Furthermore, formation of the COO isomer is consistent with early gas-phase experiments. With recent advances in machine learning of reactive force fields for polyatomic molecules,[143,218,264,265] explicit characterization of a wide range of molecule formation processes at low temperatures and under realistic conditions in interstellar space become possible. Such studies will ideally complement laboratory-based and observational astrochemistry efforts to better understand molecule abundances and the chemical evolution of molecular clouds.

4.1.4 Supporting Information

Methods: The Energy Functions

Reactive MD simulations require PESs that allow bond formation and bond breaking.[131] Two such PESs were considered in the present work. One was based on fitting Morse potentials $V(r) = D_e \left(1 - \exp\left[-\beta(r - r_0)^2 \right] \right)$ to the C–O interaction based on *ab initio* calculations using the MOLPRO[134] suite of programs at the MRCI/aug-cc-pVTZ[98,120] level of theory. The fitting parameters are r_0 =1.164 Å, D_0 =173.50 kcal/mol, and β =2.53 Å^{-1}. For the angular potential (the OCO bend) the parameters were those from the CHARMM36 force field[266] and this PES is referred to as MMH. Such an approach is similar to that used previously for oxygen-oxygen recombination on ASW.[72,244]

The second ground state PESs is based on high-level CCSD(T)-F12[267] calculations with the aug-cc-pVTZ-F12[268] basis set using MOLPRO.[134] For the geometries, a grid of r, R, and θ values including $R \in [0.9, 5.9]$Å, $r \in [0.8, 2.1]$Å, and $\theta \in [0, 180]°$ was used. In total, 7800 reference energies were determined and represented as a RKHS.[139] The root mean squared difference between the reference energies and the RKHS representation is 0.06 kcal/mol with $R^2 = 1.00$, see Figure 4.1.

For water the TIP3P model[269] was used and for CO_2 the partial charges were q_O = –0.3e and q_C = 0.6e with standard van der Waals parameters from CHARMM. These

charges are consistent with those obtained from B3LYP/6-31G(d,p) calculations snapshots from the MD simulations with CO_2 adsorbed to the 10 nearest H_2O molecules which yield $q_C = 0.73e$ and $q_O = -0.35e$. Simulations starting from the reactant state (CO+O) were initialized with the same partial charges. This compares with charges of $q_C = 0.22e$ and $q_O = -0.21e$ for the CO molecule and $q_O = -0.1e$ for an oxygen atom adsorbed to the 10 nearest H_2O molecules. To assess the dependence of the results on the partial charges used, additional reactive MD simulations were carried out with $q_O = -0.1e$ and $q_C = 0.2e$ (i.e. $q_{CO} = 0.1e$) and with $q_O = -0.2e$ and $q_C = 0.4e$ (i.e. $q_{CO} = 0.2e$). In all cases, recombination was found to speed up considerably compared with $q_{CO} = 0.3e$ and $q_O = -0.3e$ due to the higher mobility of the CO molecule and the O atom on the ASW when reduced partial charges are used.

Although there are considerably more sophisticated models for water,[270,271] simulations with the TIP3P model and point charges for atomic and molecular oxygen have yielded satisfactory agreement for diffusion coefficients for atomic oxygen on ASW with experiments.[71,72,240,244] Similarly, adsorption of CO to the ASW as reported[272] from simulations with a quadrupolar model for CO together with TIP4P is captured with the present parametrization. For temperatures up to $T = 100$ K, a) CO with charges $q_C = 0.22e$ and $q_O = -0.21e$ and b) $q_{CO} = 0.3e$ and $q_O = -0.3e$ remain physisorbed to the ASW surface on the nanosecond time scale. Also, the more elaborate parametrizations, in particular for water, are computationally considerably more expensive and do not allow routine, multiple > 10 ns simulations to be carried out. As the main focus of the present work is on the CO+O recombination reaction the nonbonded models used here were deemed sufficient.

Molecular Dynamics Simulations

In the following, the coordinates are the CO stretch r, the separation R between the center of mass of CO_A and O_B and θ is the $O_A CO_B$ angle. Hence $\theta = 0$ corresponds to the COO structure whereas $\theta = 180°$ is the OCO conformation. Initial conditions were generated for a grid of angles $\theta \in [67.5, 78.75, 84.375, 90.0, 101.25, 112.5, 123.75, 135.0, 157.5, 180.0]°$ and separations $R \in [2.66, 3.16, 3.66, 3.91, 4.16, 4.66, 5.16, 5.66]$ Å. A total of 80 initial simulations were carried out to obtain initial coordinates and velocities

128

for each of the grid points. With constrained CO and O position first 750 steps of steepest descent and 100 steps Adopted Basis Newton-Raphson minimization was carried out, followed by 50 ps heating dynamics to 50 K. Then, 100 ps equilibration dynamics was carried out for 100 ps. From each of the runs coordinates and velocities were saved regularly to obtain 1000 initial conditions for each combination of angle and distance, i.e. in total 80000 initial conditions. Production simulations 250 ps in length were then run from saved coordinates and velocities with a time step of $\Delta t = 0.2$ fs in the *NVE* ensemble. Energy conservation demonstrating the correct implementation of the energies and forces for the reactive PESs is reported in Figure 4.2 for both, simulations with MMH and RKHS.

MD simulations have demonstrated that oxygen atom diffusion and desorption are consistent with experimental findings and the topography and binding site distributions and cavity volumes in the computational model are realistic.[240,273] Specifically, for atomic oxygen[71] the diffusional barriers from simulations range from $E_{dif} = 0.2$ kcal/mol to 2 kcal/mol (100 K to 1000 K) compared with $E_{dif} = 990^{+530}_{-360}$ K from experiments[274] and the computed average desorption energy is $E_{des} = 1082$ K compared with $E_{des} = 1410^{+290}_{-160}$ K from experiments[274] which yields a ratio $E_{dif}/E_{des} \approx 0.3$ compared with 0.3 to 0.55 typically used.[274,275] For CO an average over 6 independent simulations in which the velocity orthogonal to the ASW surface was scaled until desorption occurred (see also Ref.[72]) yield desorption energies between 3.1 and 4.0 kcal/mol (1560 K to 2012 K or 130 meV to 170 meV), compared with 120 meV from experiments[276] and 70 meV to 250 meV from previous simulations.[276] On non-porous and crystalline water surfaces the submonolayer desorption energies for CO are 1307 K and 1330 K (~115 meV), respectively.[255]

Table 4.1: List of observations from the simulations with the MMH potential for combination of angles and distances where R is the distance between center of mass ((CoM)(CO)) and O and θ is the angle between them.

$\theta(°)$	$R(\text{Å})$	Total	CO_2 formation	Atom exchange	No collision complex	ES1	ES2	Both desorb
67.5	2.66	100	0	0	1	2	3	94
78.75	2.66	500	6	20,17,12,6	107	96	23	213
78.75	3.16	500	0	22,18,6,17	9	8	14	406
78.75	3.66	100	0	1,3,3,2	2	8	0	81
84.375	2.66	500	36	89,25,21,45	47	37	44	156
84.375	3.16	100	1	2,6,2,2	7	2	6	72
84.375	3.66	100	0	6,7,4,5	6	6	1	65
90.0	2.66	1000	179	73,40,32,64	44	60	48	460
90.0	3.16	100	2	10,8,2,4	8	3	1	62
90.0	3.66	100	0	4,6,2,4	8	7	0	69
90.0	3.91	100	0	7,6,4,2	22	6	2	51
90.0	5.66	500	0	0	0	500	0	0
101.25	2.66	1000	751	44,37,9,4	0	21	5	129
101.25	3.16	100	8	11,5,2,9	20	6	3	36
101.25	3.66	500	14	41,27,24,16	66	52	19	241
101.25	3.91	300	7	23,13,10,3	97	33	8	106
101.25	4.16	100	0	1,0,1,1	88	3	2	4
101.25	4.66	100	0	0	99	0	0	1
112.5	2.66	1000	913	19,2,1,0	46	9	4	6
112.5	3.16	500	205	76,20,10,12	73	30	12	62
112.5	3.66	200	42	34,11,4,8	18	15	6	62
112.5	3.91	200	33	24,10,6,4	44	16	8	55
112.5	4.16	500	11	5,4,3,1	441	1	11	23
123.75	2.66	1000	998	0,1,0,0	0	0	0	1
123.75	3.16	100	79	9,0,0,0	7	2	1	2
123.75	3.66	100	44	16,3,3,3	7	5	3	16
123.75	3.91	100	32	17,3,2,1	14	3	11	17
123.75	4.16	100	8	4,1,0,1	84	0	0	2
123.75	4.66	100	0	0	100	0	0	0

Table 4.2: Continuation of Table 4.1.

$\theta(°)$	$R(\text{Å})$	Total	CO_2 formation	Atom exchange	No collision complex	ES1	ES2	Both desorb
135.0	2.66	1000	1000	0	0	0	0	0
135.0	3.16	1000	789	169,4,5,8	15	6	2	2
135.0	3.66	500	349	97,3,5,5	20	11	3	7
135.0	3.91	500	339	73,1,1,1	22	4	46	13
135.0	4.16	1000	167	14,1,2,6	253	35	378	144
135.0	4.66	1000	69	0	901	0	27	3
135.0	5.16	1000	0	0	1000	0	0	0
135.0	5.66	1000	0	0	1000	0	0	0
157.5	2.66	1000	1000	0	0	0	0	0
157.5	3.16	500	500	0	0	0	0	0
157.5	3.66	1000	997	0	1	0	1	1
157.5	3.91	500	494	0	3	0	2	1
157.5	4.16	500	242	0	233	0	24	1
157.5	4.66	500	0	0	500	0	0	0
157.5	5.66	100	0	0	100	0	0	0
180.0	2.66	1000	1000	0	0	0	0	0
180.0	3.16	1000	1000	0	0	0	0	0
180.0	3.66	1000	1000	0	0	0	0	0
180.0	3.91	1000	991	0	9	0	0	0
180.0	4.16	500	491	0	5	3	0	1
180.0	4.66	100	0	0	100	0	0	0
180.0	5.66	1000	0	0	1000	0	0	0

In the atom exchange column, labels (a, b, c, and d) refer to trajectories with atom exchange followed by a: CO and O remain on the surface, b: CO and O desorb, c: O remains on surface and CO desorbs, and d: CO remains on surface and O desorbs.

Table 4.3: List of observations from the simulations using the 3d RKHS PES where R is the distance between CoM(CO) and O and θ is the angle between them.

$\theta(°)$	$R(\text{Å})$	Total	CO_2 formation	Atom exchange	No collision complex	ES1	ES2	Both desorb	COO formation
67.5	2.66	100	17	0	0	0	1	82	0
78.75	2.66	100	51	0	3	1	0	45	0
78.75	3.16	100	4	0	0	0	2	2	94
78.75	3.66	100	0	0	89	1	0	0	10
84.375	2.66	100	62	0	2	5	3	27	1
84.375	3.16	100	6	0	0	0	0	0	94
84.375	3.66	100	0	0	96	2	0	0	2
90.0	2.66	100	97	0	3	0	0	0	0
90.0	3.16	100	45	0	5	0	0	0	50
90.0	3.66	100	0	0	0	97	0	0	3
101.25	2.66	100	100	0	0	0	0	0	0
101.25	3.16	100	97	0	3	0	0	0	0
101.25	3.66	100	0	0	96	3	0	0	1
112.5	2.66	100	100	0	0	0	0	0	0
112.5	3.16	100	100	0	0	0	0	0	0
112.5	3.66	100	4	0	93	0	1	2	0
112.5	3.91	100	0	0	94	4	1	0	1
123.75	2.66	100	100	0	0	0	0	0	0
123.75	3.16	100	100	0	0	0	0	0	0
123.75	3.66	100	5	0	93	0	0	2	0
123.75	3.91	100	0	0	93	4	1	0	2
135.0	2.66	100	100	0	0	0	0	0	0
135.0	3.16	100	100	0	0	0	0	0	0
135.0	3.66	500	46	0	393	0	56	0	5
135.0	3.91	1000	1	0	731	9	242	0	17
135.0	4.16	100	0	0	97	0	0	0	3

Table 4.4: Continuation of Table 4.3

$\theta(°)$	$R(\text{Å})$	Total	CO_2 formation	Atom exchange	No collision complex	ES1	ES2	Both desorb	COO formation
157.5	2.66	100	100	0	0	0	0	0	0
157.5	3.16	100	100	0	0	0	0	0	0
157.5	3.66	100	12	0	59	0	28	0	1
157.5	3.91	100	0	0	76	0	19	0	5
180.0	2.66	1000	1000	0	0	0	0	0	0
180.0	3.16	100	99	0	0	0	0	1	0
180.0	3.66	500	53	0	292	2	135	8	10
180.0	3.91	100	1	0	99	0	0	0	0
180.0	4.16	100	0	0	93	0	5	1	1

4.2 Publication: Energy Redistribution following CO$_2$ Formation on Cold Amorphous Solid Water

4.2.1 Abstract

The formation of molecules in and on ASW as it occurs in interstellar space releases appreciable amounts of energy that need to be dissipated to the environment. Here, energy transfer between CO$_2$ formed within and on the surface of ASW and the surrounding water is studied. Following $CO(^1\Sigma^+) + O(^1D)$ recombination the average translational and internal energy of the water molecules increases on the ~10 ps time scale by 15 % – 25 % depending on whether the reaction takes place on the surface or in an internal cavity of ASW. Due to tight coupling between CO$_2$ and the surrounding water molecules the internal energy exhibits a peak at early times which is present for recombination on the surface but absent for the process inside ASW. Energy transfer to the water molecules is characterized by a rapid ~10 ps and a considerably slower ~1 ns component. Within 50 ps a mostly uniform temperature increase of the ASW across the entire surface is found. The results suggest that energy transfer between a molecule formed on and within ASW is efficient and helps to stabilize the reaction products generated.

4.2.2 Introduction

The motion of adsorbates in and on ASW is essential for chemistry at astrophysical conditions. Typically, bulk water is present in the form of ASW which is the main component of interstellar ices.[277] The structure of ASW is usually probed by spectroscopic measurements[277,278] although interference-based methods have also been employed.[279] ASWs are porous structures characterized by surface roughness and internal cavities of different sizes which can retain molecular or atomic guests.[65] Under laboratory conditions the water ices have been reported to be porous[280,281] or non-porous[280–282] ASW whereas the morphology of ices in the interstellar medium is more debated.[283,284]

The high porosity of ASW[285–287] makes it a good catalyst for gas-surface reactions involving oxygen[66–73], hydrogen[74], carbonaceous[75–78] and nitrogen-containing[79] species and helps maintaining those species on[288] or inside ASW.[274,289] This increases the probability for the reaction partners to diffuse to locations for collisions and association reactions to occur. As the diffusivity of individual atoms and small molecules has been established from both, experiments and simulations,[71,75,240] this is a likely scenario for formation of molecules on and within ASW.

Earlier thermoluminescence experiments suggested that the $O(^3P)+CO(^1\Sigma^+)$ reaction with both reaction partners in their electronic ground state yields excited CO_2^* which, after emission of a photon, leads to formation of CO_2.[234] Such a process has also been proposed to occur on interstellar grains[63] and has been confirmed experimentally[75] with an estimated entrance barrier of 0.014 eV – 0.103 eV for the process on ASW, compared with a value of 0.3 eV from high-level electronic structure calculations.[235] The surrounding water matrix provides the necessary coupling[233] to facilitate relaxation of the $^3A'$ or $^3A''$ states of CO_2 to the $^1A'$ ground state (correlating with linear $^1\Sigma_g^+$). The presence of an entrance barrier for the $O(^3P)+CO(^1\Sigma^+)$ reaction has led to the consideration of the alternative CO+OH pathway for CO_2 formation.[236,290] This was, however, reconsidered to yield the HOCO intermediate in such environments in more recent experiments.[237] Also, the reaction products of the CO+OH reaction have been found to depend on the

experimental conditions.[238,239] As the CO+OH reaction also appears to have a barrier in the entrance channel,[239] attention has recently shifted to the HOCO+H reaction for CO_2 formation.[237]

For adsorbed species to react on ASW they need to be able to diffuse. This has been demonstrated from MD simulations with diffusion coefficients and desorption energies consistent with experiments.[240,273] Atomic oxygen[71] on ASW experiences diffusional barriers between E_{dif} = 0.2 kcal/mol and 2 kcal/mol (100 K to 1000 K) compared with values of $E_{dif} = 990^{+530}_{-360}$ K determined from experiments.[274] For CO, MD simulations reported[72] desorption energies between 3.1 and 4.0 kcal/mol (1560 K to 2012 K or 130 meV to 170 meV), compared with 120 meV from experiments.[276] It was also found that the CO desorption energy from ASW depends on CO coverage with ranges from E_{des} = 1700 K for low to 1000 K for high coverage[280] which is consistent with the simulations.[72] On non-porous and crystalline water surfaces submonolayer desorption energies for CO are 1307 and 1330 K (~115 meV), respectively.[255] Experimental diffusional barriers range from 350 ± 50 K[281] to 490 ± 12 K.[291]

As such association reactions are in general exothermic, the energy released needs to be transferred to environmental degrees of freedom for the reaction products to stabilize. This is the quest of the present work which investigates the time scale and degrees of freedom to receive the energy liberated for the $O(^1D)+CO(^1\Sigma^+)$ reaction to form ground state $CO_2(^1\Sigma_g^+)$. The chemical precursors for formation of CO_2 are believed to be carbon monoxide and atomic oxygen and the CO+O reaction has been proposed as a non-energetic pathway, close to conditions in interstellar environments, for CO_2 formation 20 years ago from experiments involving a water-ice cap on top of CO and O deposited on a copper surface.[233] Formation of $CO_2(^1\Sigma_g^+)$ from ground state $CO(^1\Sigma^+)$ and electronically excited $O(^1D)$ is barrierless. The excited atomic oxygen species can, for example, be generated from photolysis of H_2O[245] which has a radiative lifetime of 110 minutes.[246] An alternative pathway proceeds via electron-induced neutral dissociation of water into H_2 + $O(^1D)$.[247] In the presence of CO formation of CO_2 in cryogenic CO/H_2O films was observed.[247]
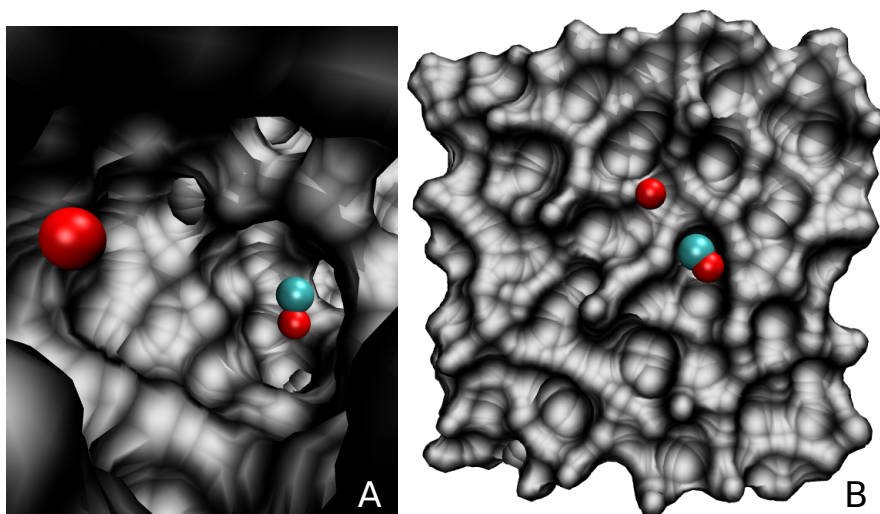
Figure 4.12: The simulation system for studying the O(^{1}D)+CO($^1\Sigma^+$)→ CO$_2$($^1\Sigma_g^+$) recombination reaction. Panel A: CO and O trapped inside a cavity of ASW; panel B: CO and O on the top of the ASW surface.

After recombination O(^{1}D)+CO($^1\Sigma^+$)→ CO$_2$($^1\Sigma_g^+$) the product is in a highly vibrationally excited state. For it to stabilize, excess internal energy needs to be channeled into the environment which is the ASW. The present work characterizes and quantifies energy relaxation of the CO$_2$($^1\Sigma_g^+$) product into internal and translational degrees of freedom of the surrounding water matrix. First, the methods used are described. Then, results are presented and discussed. Finally, conclusions are drawn.

4.2.3 Computational Methods

All MD simulations were carried out using the CHARMM suite of programs.[124] The simulation system, see Figure 4.12, consisted of an equilibrated cubic box of ASW with dimension 31 × 31 × 31 Å^3 containing 1000 water molecules. As all bonds and angles are flexible, the simulations were run with a time step of $\Delta t = 0.1$ fs and the non-bonded cutoff was at 13 Å. Simulations were started from an existing, equilibrated ASW structure[71,72,87] by adding CO$_A$ and O$_B$ inside (Figure 4.12A) or on top of (Figure 4.12B) ASW.

In the following, the coordinates are the CO stretch r, the separation R between the center of mass of CO_A and O_B and θ is the O_ACO_B angle. In addition, the C–O_B separation will be considered where appropriate. Initial conditions were generated for a grid of angles θ and separations R and simulations were carried out to obtain initial coordinates and velocities for each of the grid points. With constrained CO and O position, first 750 steps of steepest descent and 100 steps Adopted Basis Newton-Raphson minimization were carried out, followed by 50 ps heating dynamics to 50 K. Then, 100 ps equilibration dynamics was carried out. From each of the runs coordinates and velocities were saved regularly to obtain initial conditions for each combination of angle and distance. Production simulations 500 ps or 6 ns in length were then run from saved coordinates and velocities in the *NVE* ensemble. Data (energies, coordinates and velocities) were saved every 1000 steps for subsequent analysis.

Water was described by a reparametrized,[258,292] flexible KKY (Kumagai, Kawamura, Yokokawa) model.[293] The typical water modes that couple in the ~2000 cm^{-1} region relevant in the present work are the water bend (1600 cm^{-1}) and the framework rotation (600 cm^{-1}) as was also found for the vibrational relaxation of cyanide in water.[259] To describe CO_A+O_B recombination to form CO_2 the MMH parametrization was employed.[87] This model treats the two CO bonds with a Morse potential and the OCO bend as a harmonic function. MMH is a computationally efficient model (fitted to MRCI/aug-cc-pVTZ data), which yields results for recombination probabilities on ASW comparable to a more elaborate RKHS representation with an exothermicity of –7.27 eV.[87,235]

For CO_2, the partial charges were $q_O = -0.3$e and $q_C = 0.6$e with standard van der Waals parameters from CHARMM. These charges are consistent with those obtained from B3LYP/6-31G(d,p) calculations snapshots from the MD simulations with CO_2 adsorbed to a small water cluster $(H_2O)_{10}$ which yield $q_C = 0.73$e and $q_O = -0.35$e. This compares with charges of $q_C = 0.22e$ and $q_O = -0.21$e for the CO molecule and $q_O = -0.1$e for an oxygen atom adsorbed to $(H_2O)_{10}$. To assess the dependence of the results on the partial charges used, additional reactive MD simulations using the MMH parametrization were carried out with $q_O = -0.1$e and $q_C = 0.2$e (i.e., $q_{CO} =$

0.1e) and with q_O = –0.2e and q_C = 0.4e (i.e., q_{CO} = 0.2e). In all cases, recombination was found to speed up compared with q_{CO} = 0.3e and q_O = –0.3e due to the increased mobility of the CO molecule and the O atom on the ASW when reduced partial charges are used.

The main focus of the present work is to study energy redistribution within the system following recombination of $CO_A + O_B$ to form CO_2. For this, the average total, translational and internal energy of the water molecules is analyzed for recombination on top of and inside ASW. Both, the time scale and amount of energy dissipated into translational and internal degrees of freedom was determined. The translational energy for the water molecules at each timestep was determined by first calculating the magnitude of the linear momentum of each water molecule from the stored velocities. From this the translational energy contribution is calculated for each water molecule considered and the total translational energy is accumulated. The internal energy for each water molecule is determined from the difference of the total kinetic energy and the translational energy.

4.2.4 Results

In the following, the energy distribution in the water matrix of the ASW is separately discussed on the ~100 ps and on the nanosecond time scale. Next, the energy flow away from the recombination site is analyzed and, finally, the energy redistribution to neighboring water molecules surrounding the recombination site is considered.

Recombination on the 100 ps time scale

A typical trajectory for $CO_A + O_B$ recombination inside the ASW cavity is shown in Figure 4.13 (left column). Initially, the C–O_B separation is ~6 Å (Figure 4.13A). Within 150 ps recombination takes place and angular distortions lead to exploration of angles θ_{OCO} ~90° (Figure 4.13B). Relaxation of the angle occurs within the following 50 ps and the CO_2 molecule remains in an internally excited state on much longer time scales, see Figure 4.13C.[87] Concomitantly, the average internal
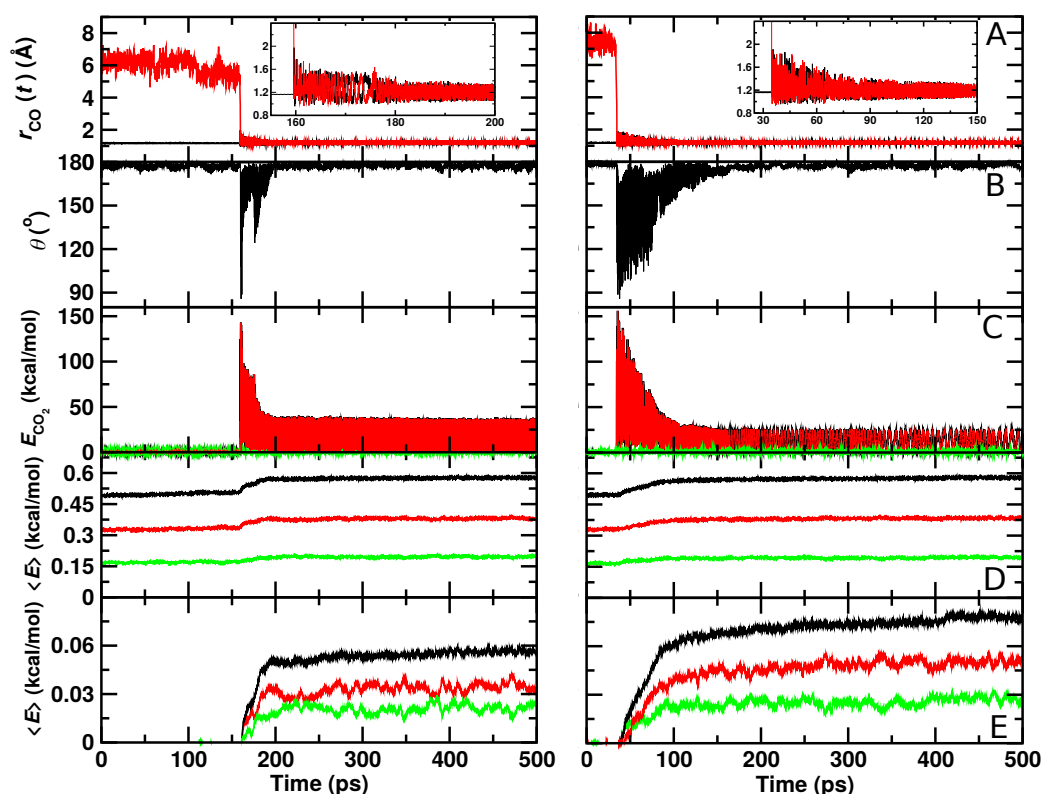
Figure 4.13: Recombination of O(^{1}D)+CO($^1\Sigma^+$) to form ground state CO$_2$($^1\Sigma_g$) in the ASW cavity (left column) and on top of the ASW surface (right column). Initially, $R = 6.0$ Å and $\theta = 180°$. (A): O$_B$–C$_{CO}$ separation (red) and CO$_A$ separation (black); (B): the O–C–O angle θ; (C – E): the average total (black), internal (red), and translational (green) energies for the CO$_2$ molecule (C), the average per water molecule (D), and the magnitude of the average per water molecule relative to the energy before recombination (E).

energy of the surrounding water molecules increases by about 10 %, see black, red and green traces in Figure 4.13D, E. The translational (phononic) modes (green) acquire approximately 1/3 of the additional energy whereas the internal energy (red) increases by the remaining 2/3.

Figure 4.13 (right column) reports a recombination trajectory on top of ASW. In this case, recombination takes place after ~35 ps and wide angular excursions extend out to 100 ps. The amount of energy picked up by the water matrix is larger compared to recombination inside ASW (Figure 4.13D, E). The average total
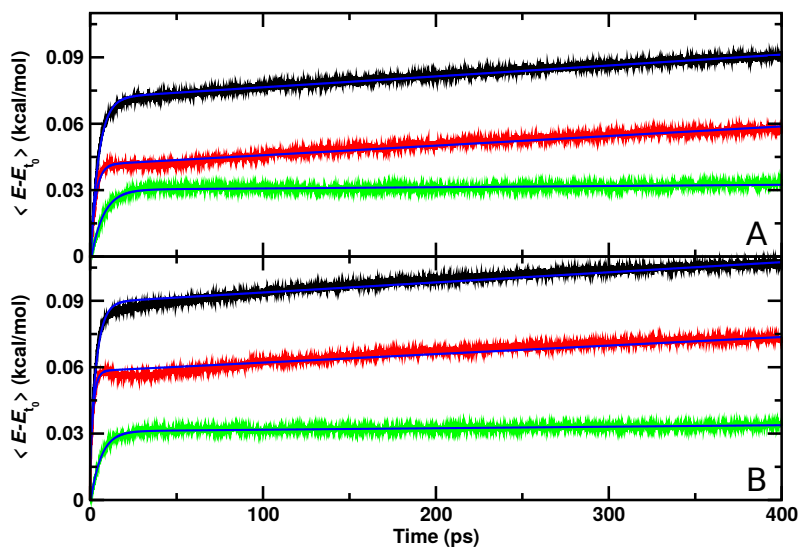
Figure 4.14: Average total (black), internal (red) and translational (green) energies for water over 70 independent runs relative to the average before recombination. The time of reaction for all trajectories is shifted to $t = 0$ and defined by the first instance at which $r_{C-O_B} < 1.6$ Å. (A): recombination within the ASW cavity. (B): recombination for $CO_A + O_B$ on the top of the ASW surface. The blue solid line is a fit to an empirical expression $\epsilon = a_0 e^{-t/a_1} + a_2 t + a_3$, see text.

energy per water molecule increases by close to 20 % and the amount that goes into internal degrees of freedom is considerably larger. For the translational modes, the energy after recombination is comparable to that for recombination within the cavity.

From a set of 70 recombination trajectories for the reaction within the cavity and on top of the ASW surface, the averaged energy contents in translational, internal and all degrees of freedom of the water molecules were determined (see Figure 4.14). For this analysis, the time of reaction was set to zero ($t = 0$) to align all reactive trajectories and all energies are reported relative to the averages before recombination. The translational contribution for recombination within and on top of ASW re-equilibrates on the ~25 ps time scale after which no change in the phononic degrees of freedom is observed. Contrary to that, the internal degrees of freedom (red traces) show temporal evolution on two time scales: a rapid phase on

the picosecond time scale, followed by a slow, long increase in the internal energies. This is also reflected in the averaged total energy (black).

As for the single trajectories, the amount of energy released from the recombination reaction into the translational degrees of freedom is similar for the reaction inside the cavity and on top of the ASW surface. For the internal degrees of freedom, however, recombination on top of the ASW surface leads on an average increase per water molecule by 0.075 kcal/mol within 400 ps (Figure 4.14B) compared with 0.06 kcal/mol for the process inside the cavity. Also, there is a characteristic decrease in the internal contribution for recombination on the surface after 15 ps which is even present when averaging over 70 independent runs. This feature is not found for recombination within ASW.

To estimate approximate time scales for the different processes involved, the average energies were fitted to an empirical expression $\epsilon = a_0 e^{-t/a_1} + a_2 t + a_3$ where ϵ is any of the energies considered. Such a functional form was chosen after inspection of the data in Figure 4.14 and accounts for the rapid initial increase in the three energies together with the slow variation of the internal energy on longer times. This parametrization is not able to model the dip around 15 ps for recombination on top of the surface, though. The time scales a_1 for total, internal, and translational energies are (4.8, 2.9, 7.1) ps for recombination inside the cavity and speed up to (3.9, 1.9, 6.1) ps for the process on the ASW surface. It is of interest to note that the rapid time scale for the internal energy is considerably faster than the kinetics of the translational degrees of freedom for both types of recombinations. The parameter a_2 which describes the slow increase of internal energy has a value of $a_2 = 4.3 \times 10^{-2}$ (kcal/mol)/ns for recombination in the cavity and $a_2 = 3.8 \times 10^{-2}$ (kcal/mol)/ns for the reaction on the surface, and is vanishingly small for the translational energy.

Average internal energies from representative independent runs for recombination inside the cavity and on top of the ASW surface are shown in Figures 4.15 and 4.16. For recombination inside the cavity (Figure 4.15) the results confirm that the energy content in the internal degrees of freedom increases considerably faster than for the translation. Also, it is found that the amount of energy transferred to translation
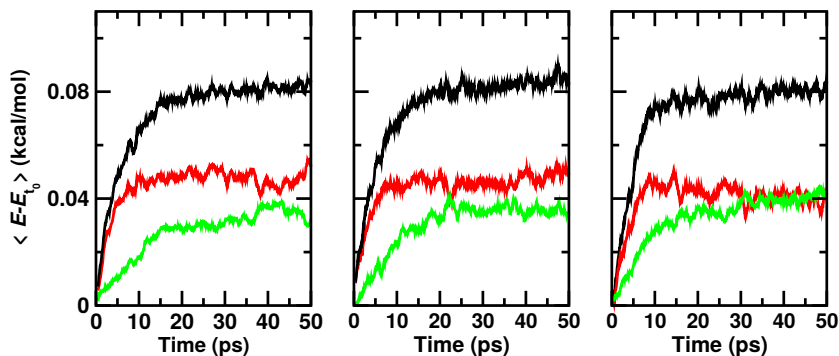
Figure 4.15: Total (black), internal (red), and translational (green) energies for water molecules from 3 independent simulations inside the ASW cavity. The time of reaction for all is shifted to $t = 0$ and defined by the first instance at which $r_{\mathrm{C-O_B}} < 1.6$.
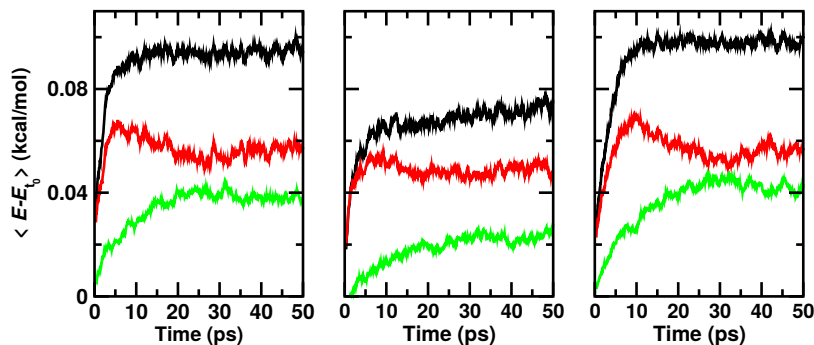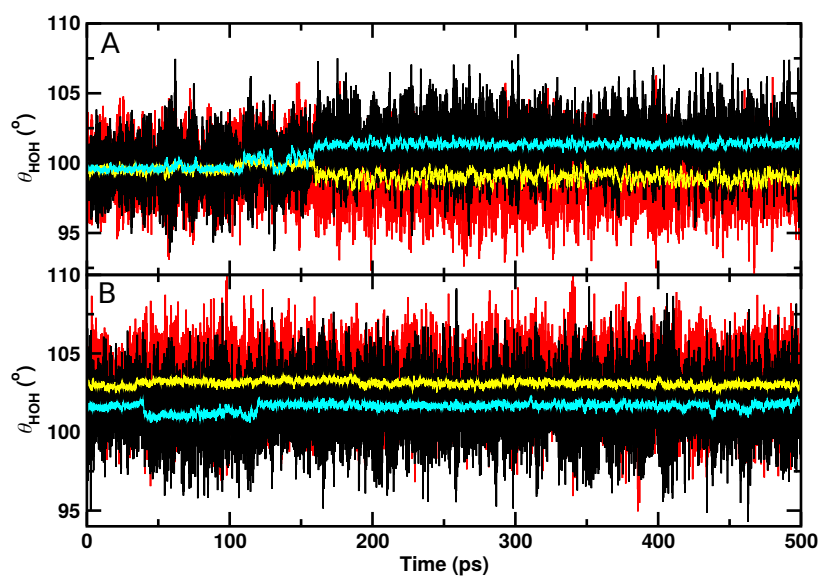

Figure 4.16: Total (black), internal (red), and translational (green) energies for water molecules from 3 independent simulations on the top of the ASW surface. The time of reaction for all is shifted to $t = 0$ and defined by the first instance at which $r_{\mathrm{C-O_B}} < 1.6$.

after recombination is smaller than that partitioned into internal degrees of freedom. For recombination on the ASW surface the same observations are made. In addition, the pronounced maximum after ~5 ps is present in all examples shown in Figure 4.16. To provide a molecularly resolved interpretation of this feature the HOH angle time series $\theta(t)$ was analyzed for a trajectory in which CO+O recombination occurred after 35 ps, see Figure 4.17B. At the time of reaction the water bending angle decreases from its average equilibrium value by $\langle \Delta\theta \rangle \sim 1°$ over the next 70 ps after which it relaxes back to the original value. The signature in the internal energy extends over ~30 ps, see Figure 4.16. Hence, it is possible that changes in the average water geometry following CO+O recombination are responsible for the overshooting and subsequent relaxation of the internal energy for the reaction on the surface. The HOH angle for a simulation within ASW in Figure 4.17A also shows a slight adjustment of the valence angle after recombination. Contrary to the situation on the surface, the average angle does not relax to the value before recombination, though.

Recombination Dynamics on Longer Time Scales

It is also of interest to analyze the energy redistribution on the multi-nanosecond time scale. Figure 4.18A demonstrates that the average total kinetic energy per water molecule continuously increases even on the nanosecond time scale. Most of this increase is due to the internal degrees of freedom although the translational component also shows a continuous slow increase on the nanosecond time scale.

The relaxation of the CO_2 internal energy is reported in Figure 4.18B. Within the first few picoseconds (inset) the internal energy is quenched to ~10 kcal/mol after which two relaxations are observed. A first phase during 1 nanosecond following recombination and a second, slower phase extending out to 6 ns and beyond. By the end of the simulation the average internal energy of the CO_2 molecule has decreased to ~2.5 kcal/mol on average. Hence, it is expected that energy transfer to the surrounding water continues but slows down considerably on the 10 ns time scale and longer.

Figure 4.17: The bending angle time series for two water molecules closest to the recombined CO_2. Panel A: recombination inside the ASW cavity at 160 ps. Panel B: recombination on the water surface at 35 ps. Black/red lines are the raw data and cyan/yellow lines are the respective average over 20 neighboring points.
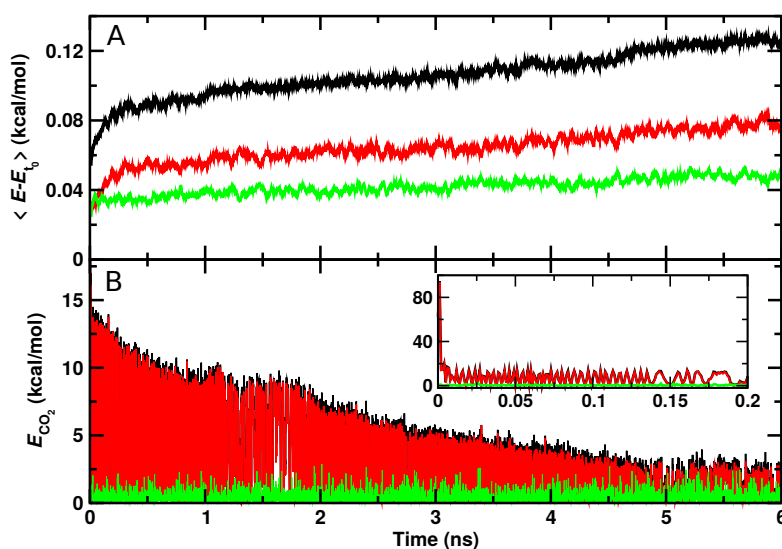
Figure 4.18: Total (black), internal (red), and translational (green) energies for water molecules (A) and for the recombined CO_2 molecule (B) from a 6 ns rebinding trajectory on the top of the ASW surface. The time of reaction is shifted to $t = 0$ and defined by the first instance at which $r_{C-O_B} < 1.6$ Å. CO_2 continues to relax and the energy in the ASW further increases beyond the maximum simulation time of 6 ns after CO_2 recombination.

Energy Migration around the Recombination Site

For a positionally resolved picture of energy flow the simulation system was separated in voxels with dimension $31 \times 1 \times 1$ Å^3. The kinetic energy of all water molecules within one such voxel was averaged along the trajectory and projected onto the (y, z)–plane. Which water molecules belong to a particular voxel was decided based on the water-oxygen atom coordinates. Figure 4.19A reports the distribution of total kinetic energy distribution before recombination. The recombination site is at $(y = 2, z = 2)$ Å and marked as a large cross. Within the first 5 ps after recombination the kinetic energy of water molecules within ~10 Å of the recombination site increases considerably, by up to a factor of 4. Following this, energy redistributes continuously across the entire surface on the 200 ps time scale, see panels C to F.

Certain regions that are initially "cold" (blue) – e.g., the region labelled "I" at $(y = 5, z = -5)$ Å in Figure 4.19 – warm up as energy transfer from CO_2 to the water molecules occurs. Conversely, other regions remain "cool", such as region "II" around $(y = 5, z = 5)$ Å for which the color code remains blue until 200 ps. Yet for other regions, such as "III", the total kinetic energy oscillates between cooler and warmer. It is also instructive to include only the first few ASW layers in this analysis which was done in Figure 4.20. Here, the voxels have sizes $10 \times 1 \times 1$ Å. For one, the cool regions are more extended before recombination. After recombination energy transfer occurs in a similar fashion as for the full system. However, the warm regions are less extended. This suggests that energy transfer also occurs to a considerable extent *into* the bulk rather than across the surface of the ASW even for recombination on top of ASW.

Energy Flow to Nearby Water Molecules

Finally, individual water molecules in immediate proximity of the recombination site are analyzed. For one trajectory with recombination on the ASW surface the average total, internal, and translational energies for the 5 water molecules closest to the recombination site are reported in Figure 4.21. During the first 10 ps after recombination the average total kinetic energy increases by up to 0.6 kcal/mol per
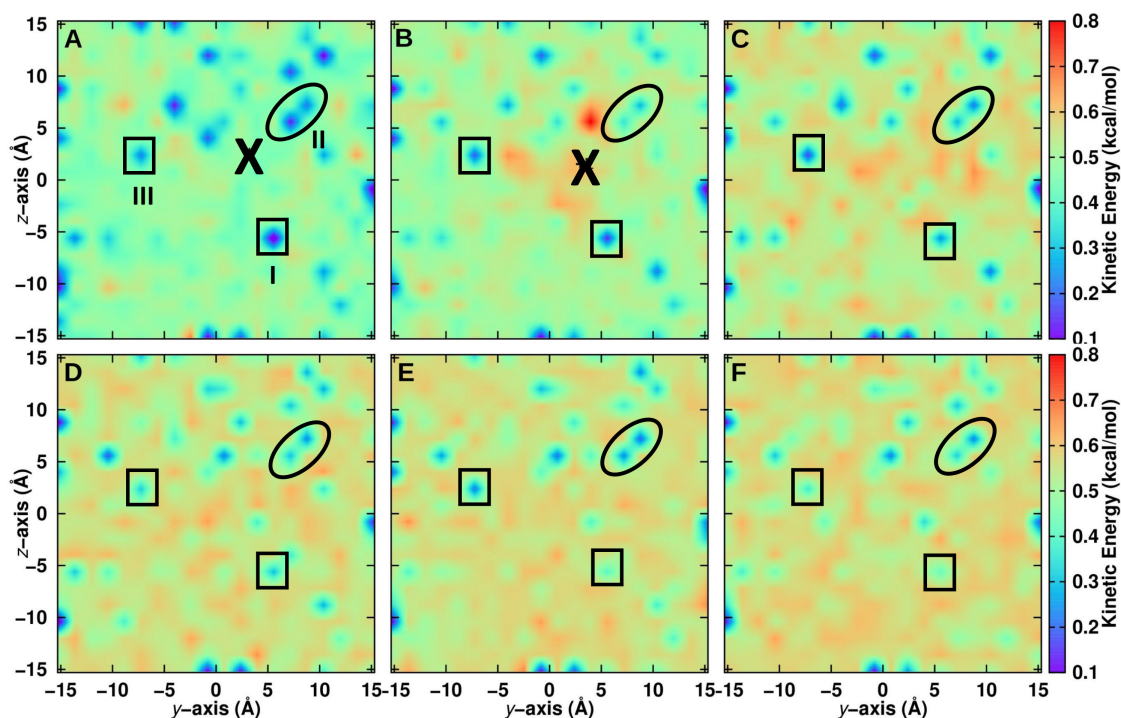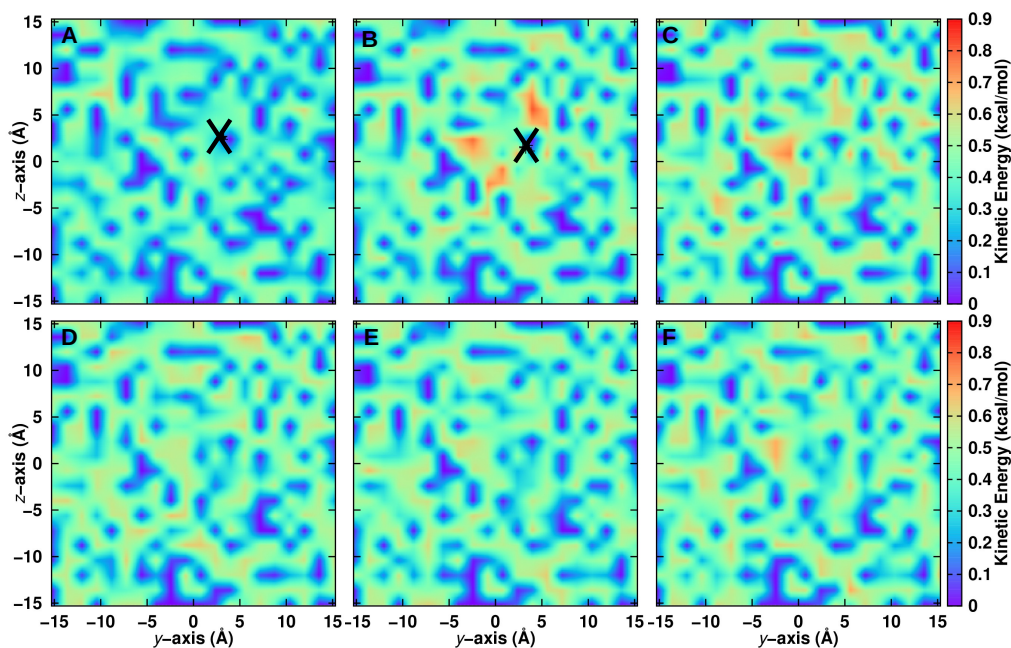
Figure 4.19: Kinetic energy of water molecules projected onto the (y, z)–plane averaged over 7 independent simulations on the top layer of the ASW surface. Recombination of CO and O takes place at the location labelled with "X". Before recombination (A): $-5 \leq t \leq 0$ ps) and after recombination (B): $0 \leq t \leq 5$ ps, (C): $5 \leq t \leq 10$ ps, (D): $10 \leq t \leq 50$ ps, (E): $50 \leq t \leq 100$ ps and (F): $100 \leq t \leq 200$ ps). The average position of CO is indicated by a large black cross. Noteworthy regions are labelled I to III and surrounded by solid lines. Region I is cool at early times and gradually warms up. Region II remains cool for most of the simulation time and region III alternates between cool and warm. For results within 10 Å of the surface, see Figure 4.20.

water molecule. Conversely, the translational energy contribution fluctuates around zero which indicates that the local structure of ASW remains intact and most of the energy flows into internal degrees of freedom.

After this initial increase, cooling of these few nearby water molecules takes place with a long-time average of -0.1 kcal/mol per water molecule in the internal degrees of freedom. On the 500 ps no noticeable changes in the translational energy content is observed. For the CO_2 molecule (see Figure 4.21B) the translational energy remains small throughout the trajectory whereas the internal energy decreases rapidly within the first 5 ps following recombination. Subsequently, slow gradual cooling on the 100 ps to nanosecond time scale takes place as was already found earlier, see Figure 4.18.

4.2.5 Discussion and Conclusions

The present work reports on the energy redistribution across ASW following $O(^1D)+CO(^1\Sigma^+)$ recombination to form $CO_2(^1\Sigma_g^+)$ on the surface and in a cavity. It is found that energy distribution occurs in two phases, one on the picosecond and one on the nanosecond time scale for both locations. Although the time dependence of the processes is similar for the two different recombination sites (inside vs. on top), the dynamics differs in a number of ways. Firstly, recombination on the surface leads to excess internal energy on the picosecond time scale which subsequently relaxes and additional energy transfer into water modes occurs on longer time scales. Secondly, recombination within the cavity considered here leads to smaller magnitude (~15 %) of energy transferred per water molecule compared with the process on the surface (~25 %). A possible reason for this is that within a sufficiently large cavity the recombined CO_2 molecule exchanges energy with the surrounding through direct collision whereas on the surface CO_2 is always in contact with the ASW. In other words, the coupling between CO_2 and water differs for recombination within ASW and on top of it. Finally, heating of the water molecules occurs on the 10 ps time scale following the recombination reaction. Consistent with earlier work,[87,294] no CO_2 desorption is found from the simulations carried

Figure 4.20: Kinetic energy of water molecules projected onto the (y, z)−plane averaged over 7 independent simulations on the top layer of ASW surface. The kinetic energy is averaged over voxels $10 \times 1 \times 1$ Å^3, i.e. only within 10 Å of the surface. Here, $t = 0$ is defined as the first instance at which $r_{C\text{-}O_B} < 1.6$ and all times are aligned with respect to this reference. Before recombination (Panel A: $-5 \leq t \leq 0$ ps) and after recombination (Panel B: $0 \leq t \leq 5$ ps, Panel C: $5 \leq t \leq 10$ ps, Panel D: $10 \leq t \leq 50$ ps, Panel E: $50 \leq t \leq 100$ ps and Panel F: $100 \leq t \leq 200$ ps). The average position of CO is indicated by black filled circles.
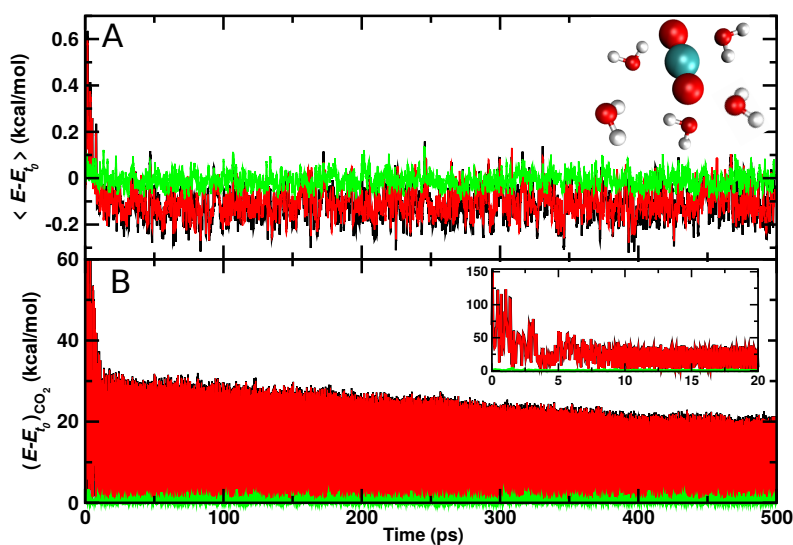
Figure 4.21: Average total (black), internal (red), and translational (green) energies for 5 water molecules (A) closest to the CO_2, and the CO_2 molecule (B) formed from a recombination trajectory on the ASW surface. The time of reaction is shifted to $t = 0$ and defined by the first instance at which $r_{C-O_B} < 1.6$. The initial time scale for energy redistribution from the relaxing CO_2 molecule to internal and translational degrees of freedom of the ASW occurs on the 10 ps time scale with slow gradual relaxation on the ~100 ps to ns time scale. On the 500 ps time scale the 5 water molecules slightly cool compared with the kinetic energy before recombination.

out here.

It is of interest to note that – ultimately – energy redistribution in such systems follows quantum mechanical principles. The present results suggest that the local energy generated from CO+O recombination is probably sufficient to excite internal modes of individual water molecules surrounding the recombination site. Hence, after CO+O recombination the ASW will be in a state characterized by a few internally and vibrationally excited water molecules embedded into a matrix of water molecules in the ground state. Earlier work on a related problem – the vibrational relaxation of a quantum oscillator coupled to oscillators of a biomolecule[295] – found that using classical mechanics leads to qualitatively correct results compared with a full quantum treatment. For the relaxation times a moderate factor of 2 for the difference between classical and rigorous quantum simulations was reported. Hence, for the present problem it is also expected that similar conclusions apply and that the nonequilibrium relaxation dynamics of individual vibrationally excited water molecules surrounded by vibrationally cold water molecules can be captured qualitatively from using classical dynamics.

In summary, the present work demonstrates that $O(^1D)+CO(^1\Sigma^+)$ recombination to form $CO_2(^1\Sigma_g^+)$ leads to excitation of both, phononic and internal modes of the water molecules that constitute the ASW. The time scales for this are on the pico- and nano-second and lead to warming the water matrix. Water molecules in direct proximity of the recombination site may become vibrationally excited and the time scale for their relaxation back to the ground state will depend on the coupling to the immediate environment. Full relaxation of the CO_2 molecule is expected to require several $10 - 100$ nanoseconds.

4.3 Publication: CO$_2$ and NO$_2$ Formation on Amorphous Solid Water

4.3.1 Abstract

Context. The dynamics of molecule formation, relaxation, diffusion, and desorption on amorphous solid water (ASW) is studied in a quantitative fashion.

Aims. The formation probability, stabilization, energy relaxation, and diffusion dynamics of CO$_2$ and NO$_2$ on cold ASW following atom+diatom recombination reactions are characterized quantitatively.

Methods. Accurate machine-learned energy functions combined with fluctuating charge models were used to investigate the diffusion, interactions, and recombination dynamics of atomic oxygen with CO and NO on ASW. Energy relaxation to the ASW and into water internal degrees of freedom were determined from the analysis of the vibrational density of states. The surface diffusion and desorption energetics were investigated with extended and nonequilibrium MD simulations.

Results. The reaction probability is determined quantitatively and it is demonstrated that surface diffusion of the reactants on the nanosecond time scale leads to recombination for initial separations of up to 20 Å. After recombination, both CO$_2$ and NO$_2$ stabilize by energy transfer to water internal and surface phonon modes on the picosecond timescale. The average diffusion barriers and desorption energies agree with those reported from experiments, which validates the energy functions. After recombination, the triatomic products diffuse easily, which contrasts with the equilibrium situation, in which both CO$_2$ and NO$_2$ are stationary on the multinanosecond timescale.

4.3.2 Introduction

Surface processes are of paramount importance for the genesis of molecules in the Universe. In the interstellar medium, nitrogen emerges as one of the chemically dynamic species, following hydrogen, oxygen, and carbon. Within the realm of prebiotic molecules and simple amino acids, including CH_2NH, CH_3NH_2, NH_2CH_2COOH, NH_2CHO, and HNCO, nitrogen stands out as the common element, underscoring the necessity to explore the reactivity of nitrogen-containing molecules. This is because three-body collisions in the gas phase are highly inefficient for molecule formation from atomic constituents. Astrophysically relevant surfaces consist of silicates, carbonaceous species (graphite, amorphous carbon, polyaromatic hydrocarbons), or water.[63] In cold molecular clouds, cosmic dust grains are typically covered by water ice, which can be polycrystalline or amorphous which is also referred to as amorphous solid water (ASW). At low temperatures, ASW dominates over the polycrystalline phase.[64] Typically, bulk water is present in the form of ASW, which is the main component of interstellar ices.[277] The structure of ASW is usually probed by spectroscopic measurements,[277,278] although interference-based methods have also been employed.[279] ASWs are porous structures characterized by surface roughness and internal cavities of different sizes, which can retain molecular or atomic guests.[65] Under laboratory conditions, water ices have been found to be porous[280,281] or nonporous,[280–282] whereas the morphology of ices in the interstellar medium is more heavily debated.[283,284]

The high porosity of ASW[285–287] makes it a suitable catalyst for gas–surface reactions involving oxygen,[66–73] hydrogen,[74] carbonaceous,[75–78] and nitrogen-containing[79] species. The surface morphology and chemistry help to adsorb chemical reagents on top of[288] or inside ASW.[274,289] This increases the probability that the reaction partners will diffuse to locations for collisions and association reactions to occur. As the diffusivity of individual atoms and small molecules has been established from both experiments and simulations,[71,75,240] this is a likely scenario for the formation of molecules on and within ASW.

The chemical precursors for the formation of CO_2 are believed to be carbon monoxide and atomic oxygen. In the past, the CO+O reaction was proposed as a nonenergetic pathway for CO_2 formation based on experiments involving a copper surface covered by a water-ice cap with CO and O deposited on it.[233] The formation of $CO_2({}^1\Sigma_g^+)$ from ground-state $CO({}^1\Sigma^+)$ and electronically excited $O({}^1D)$ is barrierless.[235,296] In addition to CO_2, nitrogen-containing species, such as nitric oxide (NO),[297–301] nitrous oxide (N_2O),[298,302] and nitrosyl hydride (HNO)[303] have been detected in the interstellar medium. However, their interstellar chemistry has been little explored so far. Specifically, NO is believed to be critical for the overall nitrogen chemistry of the interstellar medium.[304,305]

The aim of the present work is to investigate the energetics and dynamics of the CO and NO oxygenation reactions $O({}^1D)+CO({}^1\Sigma^+)$[235] and $O({}^3P)+NO(X^2\Pi)$[90,91] on ASW to form ground-state $CO_2({}^1\Sigma_g^+)$ and $NO_2({}^2A')$, respectively, at conditions representative of interstellar environments. Electronically excited atomic oxygen species can, for example, be generated from photolysis of H_2O,[245] which has a radiative lifetime of 110 minutes.[246] An alternative pathway proceeds via electron-induced neutral dissociation of water into $H_2 + O({}^1D)$.[247] In the presence of CO, the formation of CO_2 in cryogenic CO/H_2O films was observed.[247] Ground state ${}^3P$ atomic oxygen was observed from photolysis of O_2 using far-ultraviolet light.[306] These are also formation routes of $O({}^3P/{}^1D)$ that can occur in interstellar, photon-dominated regions.

At the low temperatures prevalent in the interstellar medium, surface diffusion of reaction partners is a major driving force for chemical reactions. Earlier work on the $CO + O \rightarrow CO_2$ recombination reaction[87] assumed that the CO and O moieties retain their total charges upon dissociation of $OCO \rightarrow O + CO$. This, however, is unrealistic as a CO molecule and an oxygen atom are electrically neutral for large separations between them. To provide a physically meaningful description, a fluctuating point charge (FPC) model is developed in the present work. Such a model describes each of the reactant species as a neutral atom (oxygen) or molecule (CO or NO), respectively, while leading to physically meaningful charges for the

reaction products (CO_2 and NO_2).

The methods used are described below, before presenting the CO_2 and NO_2 formation dynamics in their electronic ground states. Subsequently, the desorption dynamics of the species involved is described and the energy redistribution is analyzed. Conclusions are drawn in the final section of this work.

4.3.3 Methods

Intermolecular Interactions

Reactive molecular dynamics (MD) simulations require potential energy surfaces (PESs) that allow bond formation and bond breaking. Two such PESs were considered in the present work for the [COO] and [NOO] reactive collision systems. First, for both systems, global and reactive PESs are available, which are based on reproducing kernel Hilbert space (RKHS)[139] representations.[90,235] For the [COO] system, PES was based on extensive CCSD(T)-F12/aug-cc-pVTZ-F12 reference data,[87] whereas for [NOO], the MRCI+Q/aug-cc-pVTZ level of theory was used.[90] For both systems, all asymptotic channels are correctly described, including the atomization reaction. Second, a simplified parametrization was used that is based on Morse oscillators for the bond-stretch degrees of freedom and a harmonic potential for the bending vibration. This Morse–Morse–harmonic model is referred to as MMH. Such a simplified model is considered because of its much improved computational efficiency and because it offers the opportunity to explore the use of such simplified models for investigating larger molecules, for which high-level electronic structure calculations are problematic or unfeasible. It is stressed that both representations, RKHS and MMH, are capable of describing bond formation and bond breaking; that is, these are reactive energy functions for the [COO] and [NOO] systems.[90,235]

The use of RKHS representations has been validated in gas-phase experiments and they showed excellent performance.[87,90,235] For the MMH parametrizations, Morse potentials $V(r) = D_e(1 - \exp\{ [-\beta(r - r_0)^2] \})$ were fit to reference calculations at the MRCI/aug-cc-pVTZ level of theory, which was carried out using the MOL-

PRO[134] suite of programs.[134] To this end, the O_B–CO_A separation (see Figure 4.22 for labeling) was scanned on a regular grid of $R-$ values between 0.9 and 6.0 Å, where R is the separation between O_B and the center of mass of CO_A. The same procedure was followed for O_B–NO_A; see Figure 4.22. For the bending degree of freedom, the equilibrium angles were 180° and 134.2° for CO_2 and NO_2, respectively, and the force constants were 45.0 and 93.0 kcal/mol/radian². This yields normal mode bending frequencies of 660 cm^{-1} and 750 cm^{-1} , compared with 667 cm^{-1} and 750 cm^{-1} from experiments.[248,307] Such an approach is similar to that previously taken to explore oxygen–oxygen recombination on amorphous solid water.[72,244]

The fluctuating point charge model provides a means to correctly describe the asymptotic charge distributions in the reactant and product channels and to smoothly connect the two. Starting from separated neutral reactant species ($CO+O$ and $NO+O$), upon recombination the charges on all atoms change as a function of the O_B–CO_A and O_B–NO_A separations; see Figure 4.22. In addition, to account for the influence of the surface water molecules on the charges of the adsorbates (O atom, CO, NO, CO_2, and NO_2), electronic structure calculations for all species including the ten nearest water molecules —corresponding to a first hydration layer— were carried out. This was done at the M062x/aug-cc-pVTZ level of theory using the Gaussian program from which Mulliken charges were extracted;[308] charges based on a natural bond orbital (NBO) analysis[309] are discussed in the Section 3. After inspection of the charge variations depending on geometry, a sigmoidal function $y(\rho) = a + \frac{b-a}{1+e^{\frac{c-\rho}{d}}}$ was used to represent charges depending on the separation ρ between the oxygen atom O_B and the carbon atom of CO_A or the nitrogen atom of NO_A, respectively; see Figure 4.22. As can be seen in Figures 4.22A and B, the total charge of CO_2 and NO_2 is approximately zero, which is consistent with earlier work.[296] In other words, there is almost no charge transfer between the adsorbates and ASW.

Molecular Dynamics Simulations

All MD simulations were carried out using the CHARMM suite of programs[124] with provisions for bond forming reactions through MMH or RKHS representations.[139]
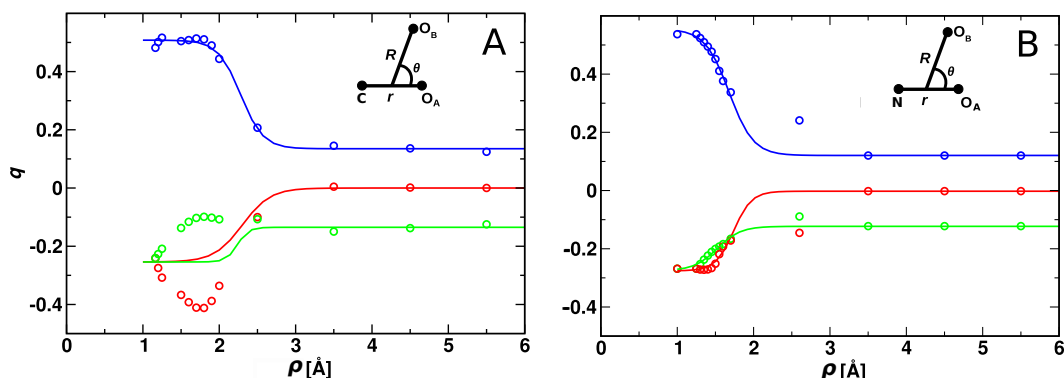
Figure 4.22: Change in charges of the atoms adsorbed to the 10 nearest H_2O molecules for formation of CO_2 (panel A) and NO_2 (panel B) as a function of ρ (C–O_B or N–O_B) distance. At low temperatures, the reaction probability completely depends on the diffusion of species for which charges fit well at asymptotic. At close range ($\rho \sim 2$ Å and shorter) bonded interactions dominate. Finally, at equilibrium, correct charges are achieved by the fitted function. Filled circles and solid line refers to *ab initio* determined and fitted charges, respectively. Color code: C and N atoms (blue) and atomic oxygen O_B (red).

The simulation system (see Figure 4.23) consists of an equilibrated cubic box of ASW containing 1000 water molecules with dimensions of $31\times31\times31$ Å³. Simulations were started from an existing, equilibrated, and validated ASW structure with a realistic distribution of internal cavity volumes.[71,72,87,240] For water, the TIP3P model,[310] as implemented in the CGenFF force field,[311] was used for consistency with the van der Waals parameters $[\epsilon, \sigma]$ for C: [–0.11 kcal/mol, 2.1 Å]; O: [–0.12 kcal/mol, 1.7Å]; N: [–0.2 kcal/mol, 1.85 Å], respectively.

Simulations for the formation of triatomics (CO_2 and NO_2) used a two-dimensional grid (R, θ) with $R \in [2.0 - 16]$ Å and $\theta \in [50 - 180]°$. For each combination of (R, θ), the reactants ([O_B, CO_A], and [O_B, NO_A]) were placed randomly on the ASW to broadly sample the surface. First, the structure of the system was optimized (750 steps of steepest descent and 100 steps of adopted basis Newton-Raphson minimization), with R and θ constrained (frozen) to each of the values on the grid. Next, Boltzmann-distributed velocities at 50 K were assigned to all atoms, followed by 50 ps of heating simulations (*NVT*) to 50 K and 100 ps of equilibration at 50 K, with (R, θ) still constrained. After releasing the constraints, a 500 ps equilibrium trajectory followed, during which coordinates and velocities were saved regularly to
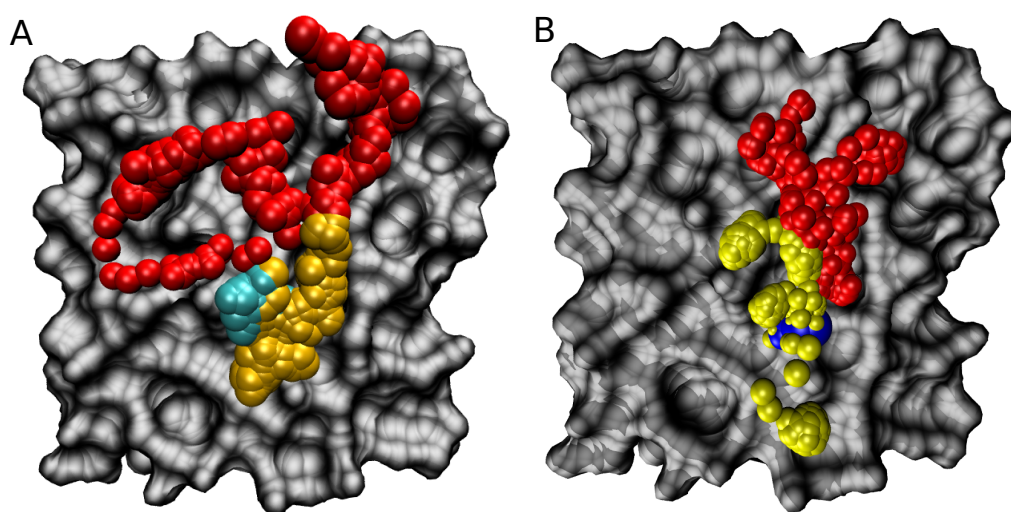
Figure 4.23: The simulation system for studying the CO_2 (Panel A) and NO_2 (Panel B) recombination reaction. Panel A color code: oxygen (red), CO (cyan), CO_2 (orange); Panel B color code: oxygen (red), NO (blue), NO_2 (yellow). Before and after recombination all species diffuse on the water surface in this case.

obtain 1000 initial conditions for $R > 4$ Å and 500 for $R < 4$ Å, accounting for the fact that the reaction probability is ~ 1 for short initial R. This yielded a statistically significant number of initial conditions from which CO_2 and NO_2 recombination trajectories were run. The subsequent production simulations were each 500 ps in length and were carried out in the NVE ensemble to ensure total energy conservation and to be able to analyze energy flow. For analysis, the total energies, coordinates, and velocities were saved every 0.5 ps.

Analysis

To characterize energy flow between the activated internal modes of newly formed species and the surface water molecules, the vibrational density of states (vDOS) was analyzed.[86,312] For this, the Fourier transform of the hydrogen atoms normalized velocity autocorrelation function was determined[313] according to

$$I_{\text{vDOS}}(\omega) = \int_0^{T_c} \frac{\langle v(0) \cdot v(t) \rangle}{\langle v(0) \cdot v(0) \rangle} e^{-i2\pi\omega t} dt, \tag{4.1}$$

where v denotes the velocity vector of the hydrogen atom. The value $T_c = 1$ ps was chosen because vibrational energy flow is expected to occur on such timescales[314] and because sufficient data for meaningful analysis can be accumulated from the simulations while still covering the tens-of-picoseconds timescale on which the equilibration of the newly formed triatomic molecules takes place. The formation probability of products was determined from the ratio of the number of successful recombination reactions to the total number of reactive collisions, as is usual from rate calculations in quasiclassical trajectory simulations.[315]

4.3.4 Results

Exploratory Simulations

A representative 500 ps trajectory for $CO_A + O_B$ recombination is shown in Figure 4.24 (left column), where O_A and O_B label the two oxygen atoms to distinguish them; see Figure 4.22 for labeling. Initially, the $C \cdots O_B$ separation was ~ 12 Å (see panel
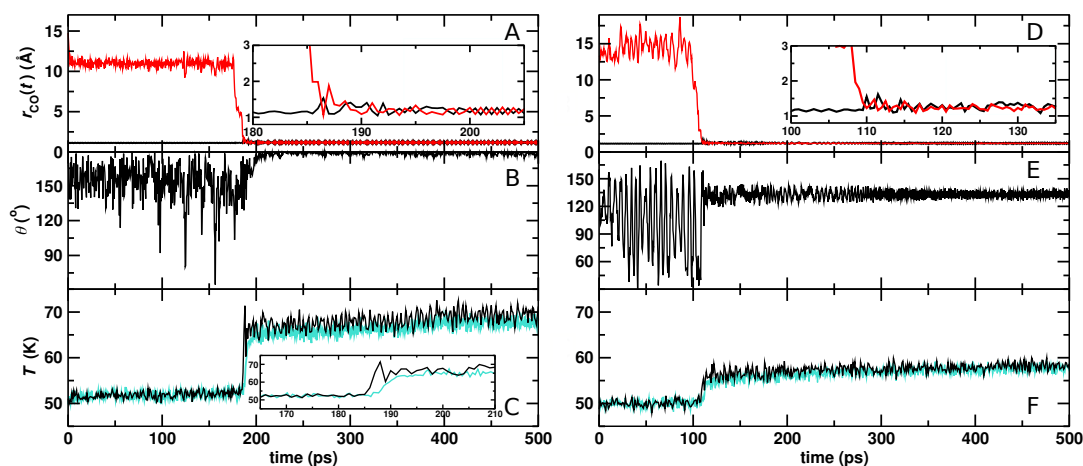
Figure 4.24: Recombination of CO + O to form CO_2 (left column) and NO + O to
NO_2 (right column) in ground state. Panel A and D: O_B–C/N (red) and
O_A–C/N (black) separation; Panel B and E: O–C/N–O angle; Panel C and
F: Temperature of the ASW (cyan) and the full system (black) before and
after recombination. Upon recombination both O–C/N stretches exhibit
equally pronounced excitation (shown in the insets).

A), which is in the asymptotic, noninteracting regime of the CO_A–O_B PES. For the
first 190 ps, oxygen diffusion on the water surface can be seen. Upon recombination
at $t \sim 190$ ps, the CO stretch and the OCO (panel B) angle are highly excited and
relax during the following few picoseconds to average values of around the CO_2
equilibrium geometry. The CO_2 product remains in an internally excited state for a
considerably longer timescale; see panel B. The temperatures of the ASW (cyan)
and the full system (black) are determined by the kinetic energies of the molecules.
A prominent peak in the temperature of the full system is observed immediately
after recombination, whereas the warming of cool water surface is gradual (panel C
inset).

Figure 4.24 (right column) shows a NO_A + O_B recombination trajectory. The ground
state $NO_2(^2A')$ has a nonlinear geometry with N-O distances of 1.2 Å and an O-N-O
angle of $134°$. In this case, recombination takes place after ~ 110 ps and the amount
of energy released is half compared to the CO_2. This is due to the different stabi-
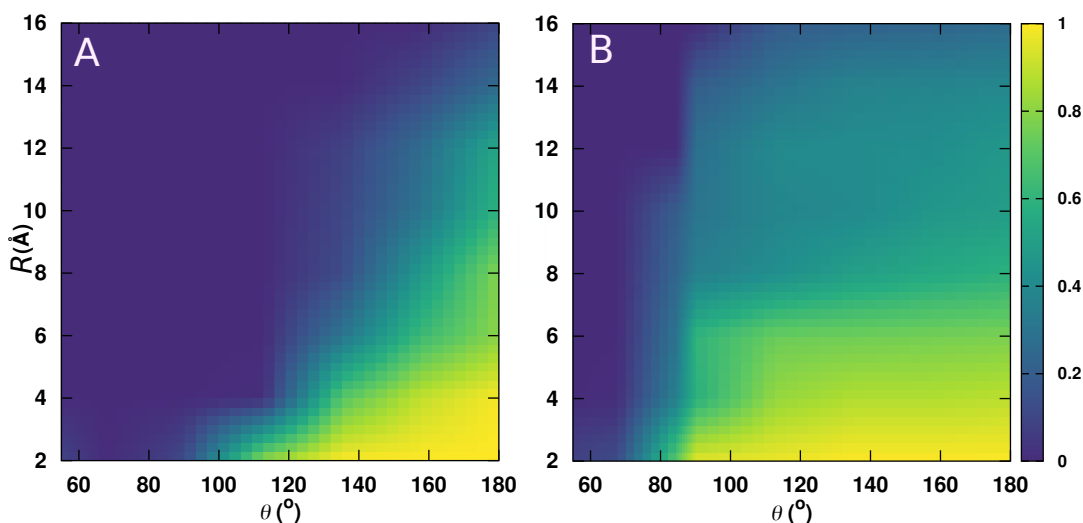
Figure 4.25: KDE density map for the average CO_2 formation probability depending on initial (R, θ) using the MMH (Panel A) and RKHS (Panel B) PES. The color palette indicates the probability of the reaction, with yellow representing $P = 1$ and blue representing $P = 0$. The COO formation probability using RKHS PES is shown in Figure 4.27.

lization energies of the triatomics with respect to the CO/NO+O asymptote, which are 7.71 eV for CO_2 (for $\theta = 180°$) and 3.24 eV for NO_2 (for $\theta = 135°$), respectively.

Reaction Probabilities for CO_2 and NO_2 Formation

The formation dynamics of CO_2 and NO_2 was followed from initial conditions on a (R, θ) grid. For each initial R and θ configuration, 250 or 500 trajectories at 50 K were run for 500 ps. Depending on the initial (R, θ) values, heat maps of the formation probability were determined and a kernel density estimate (KDE) was used for smoothing; see Figures 4.25 and 4.26.

Figure 4.25 reports probability heat maps for the likelihood to form CO_2 depending on initial (R, θ) using both MMH (Panel A) and RKHS (Panel B) PESs from simulations of 500 ps in length. With the MMH PES (Panel A) for $R \sim 2$ to 4.5 Å, the recombination probability is $P \sim 1$ near $180°$ and vanishes for $\theta \sim 90°$. For $\theta = 180°$, P changes from unity for $R = 4$ to $P \sim 0$ for 14 Å on the 500 ps timescale. For $R > 4$ Å and $\theta \leq 120°$, the formation probability is $P = 0$.
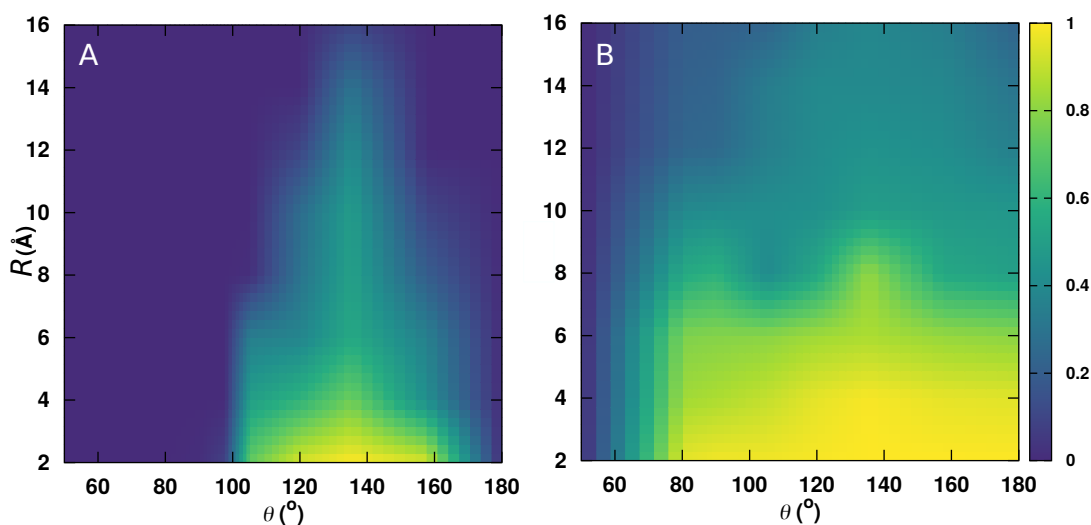
Figure 4.26: The NO_2 formation probability as a function of R and θ using MMH (Panel A) and RKHS (Panel B) PES. The color palette indicates the probability of the reaction, with yellow representing $P = 1$ and blue representing $P = 0$.

Using the RKHS PES, which is a considerably more accurate representation of the O+CO interaction than the MMH parametrization, the reaction probabilities depending on initial (R, θ) are larger throughout; see Figure 4.25 B. The probability for formation of CO_2 is particularly increased along the angular coordinate θ. These differences are due to the different topographies of the two PESs. Also, the angular dependence is observed for $R < 7\text{Å}$, and for larger distances, $R > 7\text{Å}$, a plateau is observed. This plateau implies that the reaction probability is primarily governed by diffusion for large distances and does not depend on the initial angular term on the 500 ps timescale.

Due to the more realistic and improved charge model for the reactants, their diffusivity increases. For simulations on the 500 ps, this leads to a nonvanishing reaction probability for initial separations of $R = 16$ Å. This contrasts remarkably with earlier simulations[87] that maintained the charges on the oxygen atom and CO molecule fixed at ± 0.3e (corresponding to their charges in CO_2), which suppressed the reactants' mobility and reaction probability for $R > 5$ Å for simulations on the same timescale. Overall, using the more realistic charge models from the present work increases the recombination probability of the species involved.

The recombination probability for NO_2 formation is reported in Figure 4.26. Panel A shows the 2D recombination probability using the MMH PES. For NO_2, the minimum energy structure has $\theta = 134°$, which leads to $P \sim 1$ for $\theta \in [125, 145]°$ and $R \leq 4$ Å, which gradually decreases to $P \sim 0$ as R increases. Using the considerably more accurate RKHS representation, the recombination probability is larger throughout; see Figure 4.26B. A formation probability of $P \sim 1$ is found for $\theta \in [90, 180]°$ and $R \leq 8$ Å. These differences are again due to the different topographies of the two PESs. Similar to CO_2, the recombination probability is not equal to zero at $R = 16$ Å on the 500 ps timescale and a plateau is observed for $R > 8$Å.

Although the structures of CO_2 and NO_2 differ (linear versus bent), the shape of the underlying PES is reflected for both in the geometry dependence of the rebinding probability for shorter R. For larger initial separations R, the rebinding probability is purely governed by diffusion and is independent of θ. Taking the results from simulations using the RKHS representation as the reference, the MMH PES yields qualitatively comparable results. However, in particular for NO_2 the reaction funnel is far too narrow. On the other hand, the MMH-PES is orders of magnitude more computationally efficient and improvements with techniques such as multistate adiabatic reactive MD (MS-ARMD) may yield considerable improvements over MMH.[48,316]

The RKHS PES also supports the COO conformation and results in the formation and stabilization of COO. This intermediate holds significant interest as it can decay to $C + O_2$. The probability heat map for the likelihood to form COO depending on R and θ is shown in Figure 4.27. Finally, the atom exchange reaction is found for 7 % of all simulations using the MMH PES for NO/CO + O. Atom exchange with RKHS is not observed for CO+O, whereas 0.4 % of the simulations show atom exchange for NO + O collision.

**Energy Dissipation upon Triatomic Product Relaxation**

Both association reactions considered here are exothermic. Upon recombination, products are formed in a highly excited internal state (vibration and rotation) and
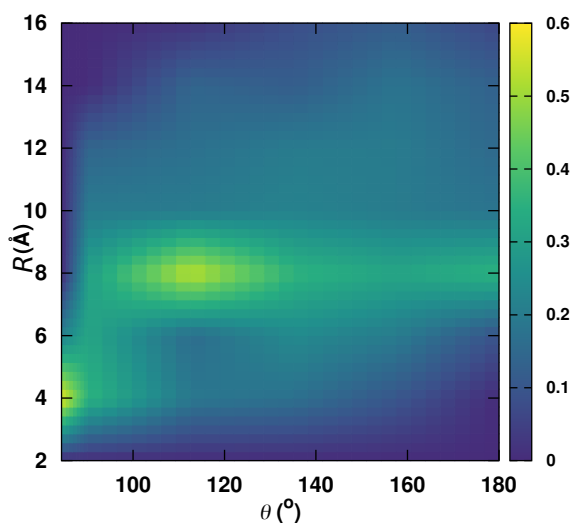
Figure 4.27: KDE density map for the average COO formation probability depending on initial (R, θ) using RKHS PES.

their relaxation depends on the coupling between excited stretching and bending modes of the newly formed species and phonon and internal modes of the underlying water ice surface.[86] To allow energy transfer between the adsorbate and the internal and phonon modes of ASW, the reparametrized[258,292] flexible Kumagai, Kawamura, and Yokokawa (KKY) water model[293] was used in these simulations. This model has been successfully used to study the spectroscopy of CO on and in ices,[258] the vibrational relaxation of solvated cyanide,[259] and the relaxation of O_2 formed on ASW following $O(^3P) + O(^3P)$ recombination and vibrational relaxation through coupling to the water-phonon modes.[72] Within 1 μs, O_2 relaxes to the $v = 2$ vibrational state, which is consistent with experimental findings[317] and validates the use of the KKY model to study energy transfer processes involving ASW.

To study energy dissipation following recombination, several hundred $O+CO \rightarrow CO_2$ and $O+NO \rightarrow NO_2$ trajectories were run for 500 ps. Subsequently, 280 recombination trajectories were analyzed for CO_2 and NO_2, respectively. First, the normalized final kinetic energy distribution $P(E_{kin}^{fin})$ at 450 ps after recombination was determined for the products; see Figure 4.28. The $P(E_{kin}^{fin})$ for CO_2 and NO_2 extend from low-E_{kin}^{fin} (~ 1 kcal/mol) to 60 kcal/mol. For NO_2, more than 80 % of the products contain $E_{kin}^{fin} < 5$ kcal/mol after 450 ps, whereas for CO_2 this fraction is considerably smaller. In other words, relaxation of the highly excited internal modes is more effective for
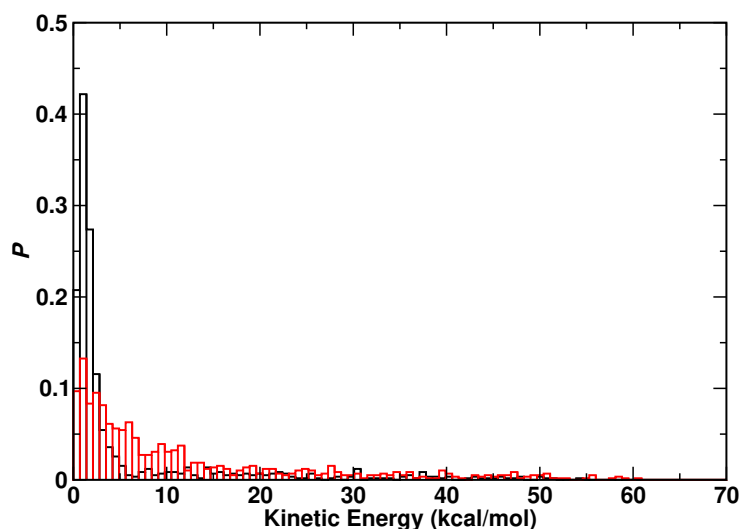
Figure 4.28: Normalized kinetic energy distribution for CO_2 (red) and NO_2 (black) from 280 reactive trajectories after 450 ps of recombination using RKHS PES. For NO_2 ~80 % of trajectories relaxed down to ~5 kcal/mol whereas this fraction is only 40 % for CO_2. In other words, upon recombination NO_2 relaxes more effectively than CO_2 on the sub-ns time scale.

NO_2 compared with CO_2. Possible reasons for such rapid relaxation include but are not limited to the shallow minimum for NO_2 compared to CO_2 and/or more efficient coupling of the vibrational modes of NO_2 to the surrounding water molecules.

Next, for one CO_2-forming trajectory, the kinetic energy distribution and vDOS of two water molecules W_A and W_B and the CO_2 molecule were determined for windows of 1 ps in length; see Figure 4.29. For the analysis, the time of recombination was set to zero and water molecules W_A and W_B were chosen such that, at $t = 0$ W_A was the water molecule nearest to the newly formed CO_2 and W_B was closest to CO_2 after 3.0 ps due to diffusion of the recombination product. As a reference point for the analysis, the averaged E_{kin} (red distributions) and vDOS (black lines) for ten water molecules away from the recombination site was determined for the 0.5 ps prior to CO_2-formation; see the top row in Figure 4.29.

For water molecule W_A, the maximum of $P(E_{kin})$ (red shaded area) shifts to higher E_{kin} and its width increases considerably for the first 15 ps after recombination.
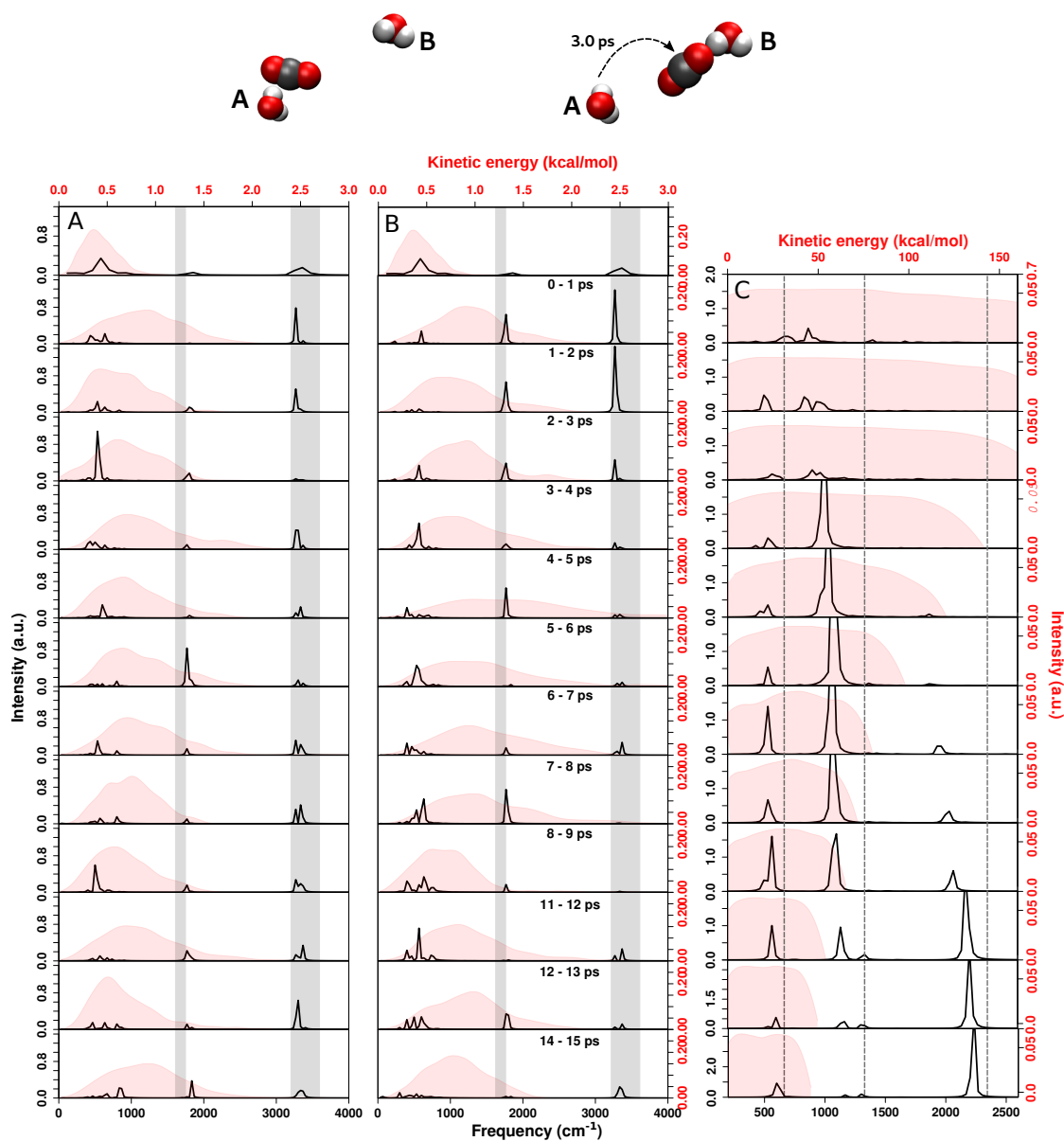
Figure 4.29: Vibrational density of states power spectrum (in black) and $P(E_{kin})$ (red) for water molecules A and B, W_A (panel A) and W_B (panel B), on the ASW surface at several delay times after the formation of CO_2 on the RKHS-PES. The IR bands at $1600 - 1750$ and $3200 - 3600$ cm^{-1} correspond to bending and OH-stretching modes.[277,318,319] Panel C reports the kinetic energy distribution (red) and the vDOS of the CO_2 molecule. Top panel: Averaged vDOS spectrum and kinetic energy distribution from 10 randomly chosen surface water molecules from data before recombination takes place. Subsequent panels are labeled with the time interval after CO_2 recombination. For an analysis of a trajectory using the MMH PES, see Figure 4.30.
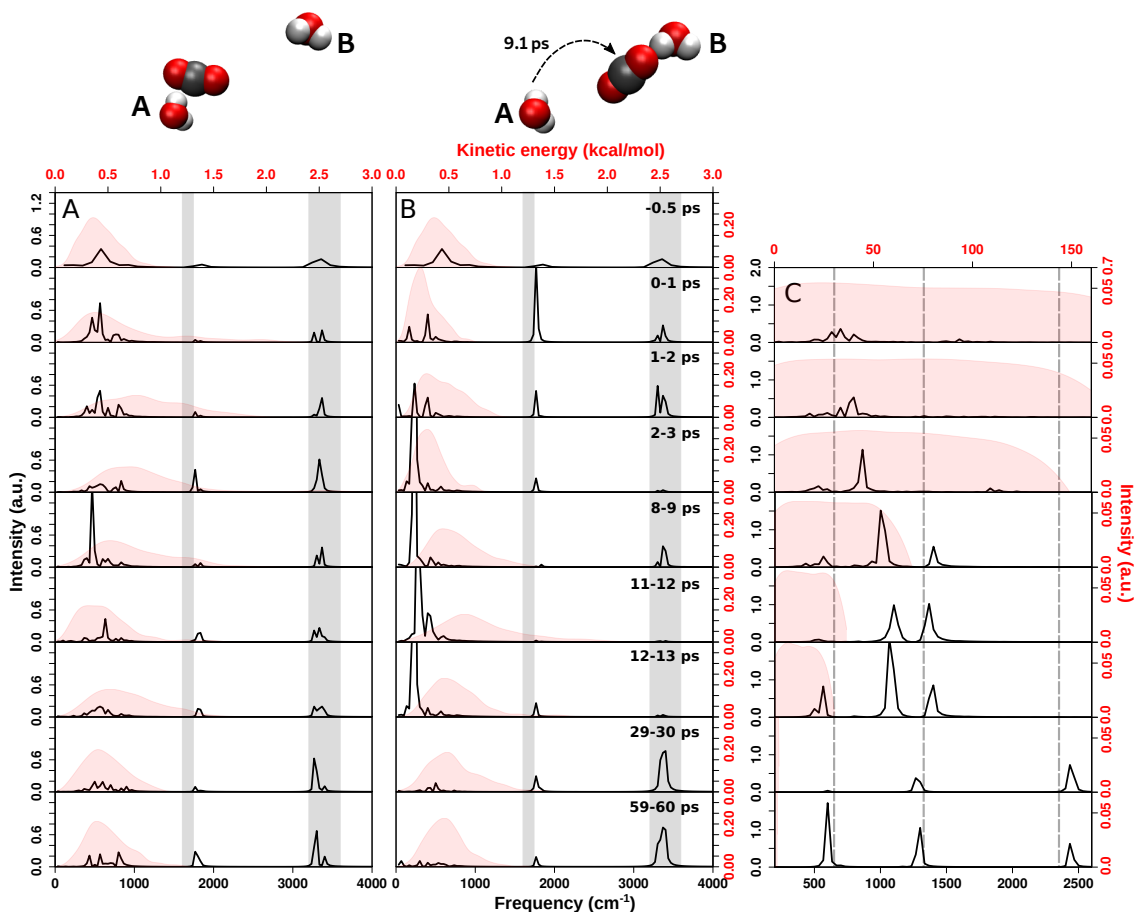
Figure 4.30: Vibrational density of states power spectrum (in black) and $P(E_{\text{kin}})$ (red) for water molecules A and B, $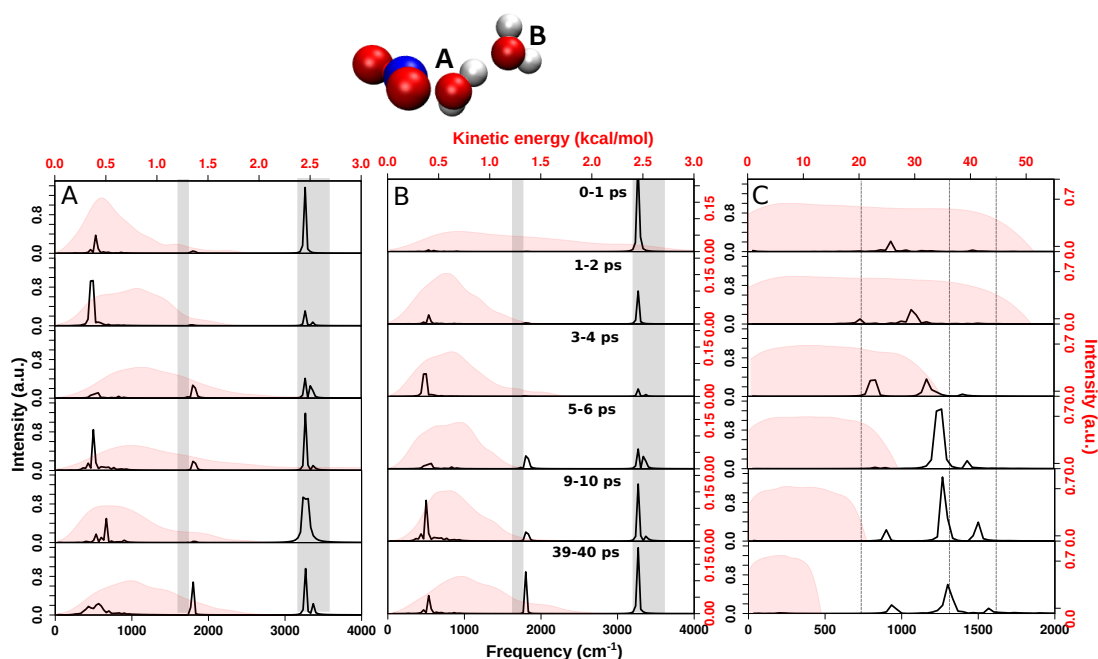W_A$ (panel A) and W_B (panel B), on the ASW surface at several delay times after the formation of CO_2 on the MMH-PES. Panel C reports the kinetic energy distribution (red) and the vDOS of the CO_2 molecule.

Energy flow is clearly taking place as can be seen from the varying widths and locations of the maxima depending on time after recombination. The same is observed for water molecule W_B (middle column). As judged from the vDOS (black lines), the water stretch, bend, and libration modes for W_A and W_B are populated for all time intervals. The two internal modes do not shift in frequency relative to the thermal equilibrium and their frequency is covered by the experimentally determined frequency range.[277,318,319]

The CO_2 molecule formed is initially highly excited, which leads to a broad, unstructured $P(E_{kin})$; see the top panel in Figure 4.29C. Over the next 15 ps, the width of $P(E_{kin})$ rapidly decreases due to energy exchange with the internal modes of the surrounding water molecules and the phonon modes of the ASW. The reaction product does not further cool until 60 ps after recombination. A picture emerges for the vDOS that is rather different than that for water. The vibrational frequencies for CO_2 are at 667 (bend), 1333 (symmetric stretch), and 2349 (antisymmetric stretch) cm^{-1} from experiments[248] (see vertical lines) and 647, 1374, and 2353 cm^{-1} from calculations using the reactive PES, in reasonably good agreement. For times up to 5 ps after recombination, only frequencies below ∼ 1000 cm^{-1} appear in the vDOS. Specifically, the antisymmetric stretch mode only manifests itself after ∼ 6 ps, because the reactive PES for CO_2 is fully anharmonic and for the first few picoseconds after recombination the highly excited stretching motions sample the narrowly spaced energies close to dissociation. For times between 6 ps and 14 ps after recombination, the asymmetric stretch shifts from ∼ 1800 cm^{-1} to 2250 cm^{-1} due to relaxation from higher to lower vibrationally excited states. The CO_2 bending vibration together with its overtone at 1130 cm^{-1} are prominently populated after 4 ps until 12 ps, after which the symmetric stretch appears and the overtone disappears.

A slightly different approach was used for NO_2 recombination in order to study the energy transfer between neighboring water molecules on the ASW. Again, $P(E_{kin})$ and the vDOS of water molecules W_A and W_B and NO_2 were determined for one NO_2−forming trajectory. In this case, W_A is the water molecule closest to NO_2, while W_B is hydrogen-bonded to W_A; see Figure 4.31. For the first 1 ps after recombination, the widths of $P(E_{kin})$ for W_A and W_B are considerably broader than the thermal

Figure 4.31: Vibrational density of states power spectrum (in black) and $P(E_{kin})$ (red) for water molecules A and B, W_A (panel A) and W_B (panel B), on the ASW surface at several delay times after the formation of NO_2 on the RKHS-PES. The IR bands at $1600 - 1750$ and $3200 - 3600$ cm^{-1} correspond to bending and OH-stretching modes.[277,318,319] Panel C reports the kinetic energy distribution (red) and the vDOS of the NO_2 molecule. For an analysis of a trajectory using the MMH PES, see Figure 4.32.

distribution; see the top panel of Figure 4.31A. Energy exchange between the two water molecules is readily observed for the following 5 ps with maxima and widths of $P(E_{kin})$ increasing and decreasing. The distributions for 10 ps and 40 ps no longer change appreciably and the two water molecules do not fully thermalize on that timescale. The kinetic energy distribution for NO_2 is broad at the beginning (50 kcal/mol) and transfers more than half of this energy to the environment within 10 ps of recombination (Figure 4.31C).

The experimentally measured[307] vibrational frequencies for gas phase NO_2 are at 750 (bend), 1318 (symmetric stretch), and 1618 (antisymmetric stretch) cm^{-1}. Up to 4 ps after recombination, the vDOS is rather sparse. Regarding CO_2, the asymmetric stretch mode is not populated at all at early times, whereas the bend and symmetric stretch manifest themselves by 6 ps. At 10 and 40 ps after recombination, the NO_2 vibrational spectrum can be recognized from the vDOS. For the water modes, the
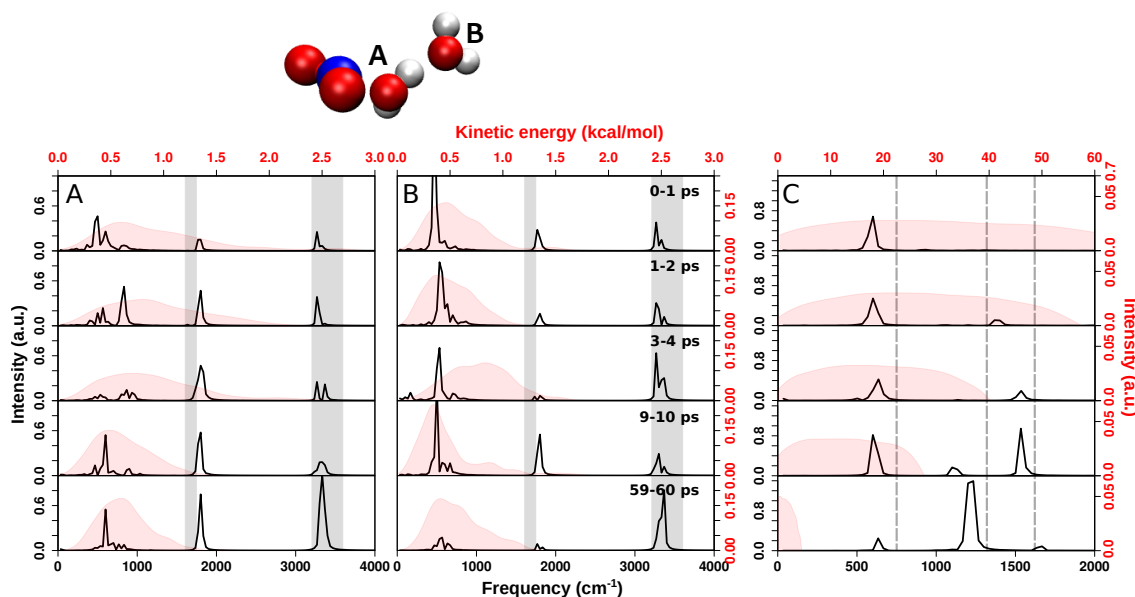
Figure 4.32: Vibrational density of states power spectrum (in black) and $P(E_{\text{kin}})$ (red) for water molecules A and B, W_A (panel A) and W_B (panel B), on the ASW surface at several delay times after the formation of NO_2 on MMH-PES. Panel C reports the kinetic energy distribution (red) and the vDOS of the NO_2 molecule.

librational and stretching modes are populated throughout, whereas the bending mode is present occasionally and clearly visible after 40 ps.

Specifically comparing the time evolution of $P(E_{\text{kin}})$ for CO_2 and NO_2, the two products are found to differ in the efficiency and speed of cooling. For NO_2, the fraction of products formed with $E_{\text{kin}} \sim 5$ kcal/mol within 450 ps is twice as large as it is for CO_2; see Figure 4.28. This may be related to the lower energy of formation (3.24 eV vs. 7.71 eV) but may also be due to the overall lower vibrational energies of the fundamentals for NO_2 compared with CO_2, which facilitates energy relaxation.

For a mode-specific analysis of energy transfer between the formed CO_2 and the surrounding water molecules, the time evolution of the integrated vDOS in three frequency ranges was considered: water libration: $(0-1000 \text{ cm}^{-1})$, bending $(1500-2000 \text{ cm}^{-1})$, and stretching $(2800-3800 \text{ cm}^{-1})$. For this analysis, water molecules within 10 Å in the $x-$direction and 5 Å in the $y/z-$ direction around the recombination site were used; see Figure 4.33. The vDOS was then computed for each water molecule
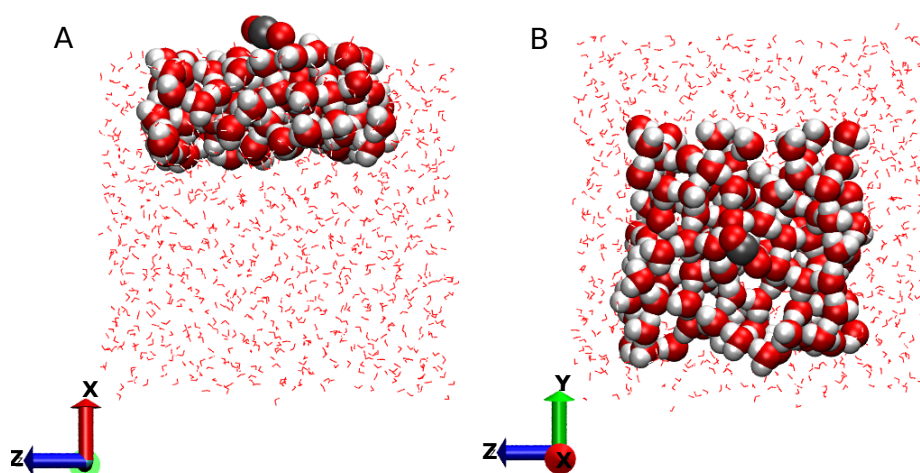
Figure 4.33: Side (left) and top (right) view of the ASW surface with CO_2 on the top. The water molecules within 10 Å in the $x-$direction and 5 Å in the $y-/z-$directions are highlighted.

separately, and the resulting peak areas were averaged over 100 water molecules located in the region defined above. Figure 4.34A reports the difference between the equilibrium averaged peak area and the area at time t after recombination. Alternatively, the difference between peak area immediately after excitation and the area at time t is shown in panel B. The integrated intensity of the low-frequency libration mode (black line) initially increases, followed by a decrease after $\sim$ 10 ps, before eventually reaching a plateau after 25 ps (panel A). Meanwhile, the intensity of the stretching modes (green) displays a progressive rise over the first 25 ps, followed by fluctuations around a specific value, and then a slow relaxation over about 100 ps. In comparison, the intensity of bending modes (red traces) remains comparable throughout the observed time frame (a similar observation is shown in Figure 4.35). The time series show that the libration mode first acquires and stores energy and subsequently releases it, whereas the stretching modes pick up energy over a timscale of approximately 100 ps and retain it beyond that. This was also observed for NH_3 formation on water ice.[86]

Species Desorption on ASW

To explicitly probe desorption, velocities for the adsorbants (CO, NO, CO_2, and NO_2) were drawn from a Maxwell-Boltzmann distribution and scaled along the
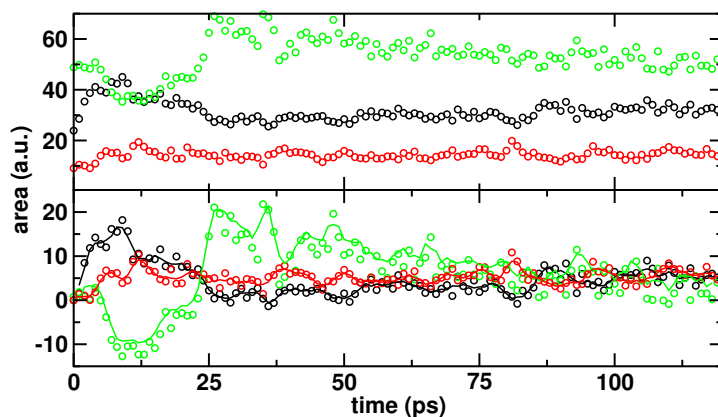
Figure 4.34: Time evolution of the integrated change of vDOS curves averaged over 100 water molecules for libration (black), bending (red), and stretching modes (green) of water using the RKHS-PES. In Panel A, the equilibrium, thermally averaged area before recombination is the reference for calculating the difference at time t, whereas for Panel B the peak area right after excitation ($t = 1$ps) is the reference. Here, solid lines represent the 4-point moving average.

x−direction, perpendicular to the water surface.[244] For the water molecules, the velocities were those from an equilibrium simulation at 50 K and the dynamics of the system was followed for 100 ps. If within this timescale the adsorbant remained on the ASW, it was considered physisorbed. However, for sufficiently large scaling of the velocity vector, the adsorbant leaves the ASW from which the desorption energy can be estimated.[244] Average desorption energies were determined for initiating the dynamics for different initial positions of the adsorbant on the ASW.

Computed desorption energies (E_{des}) for CO and NO from the ASW surface were 3.5 ± 0.7 kcal/mol (1812 K or 156 meV) and 3.1 ± 0.5 kcal/mol (1534 K or 132 meV), respectively. Earlier MD simulations reported[87] CO-desorption energies of between 3.1 and 4.0 kcal/mol (1560 K to 2012 K or 130 meV to 170 meV), compared with 120 meV (2.8 kcal/mol or 1392 K) from experiments.[276] It was also found that the desorption energy of CO from ASW depends on CO coverage with ranges from E_{des} = 1700 K for low to E_{des} = 1000 K for high coverage, respectively,[280] which supports the simulation results.[87] On nonporous and crystalline water surfaces, submonolayer desorption energies for CO are 1307 K and 1330 K

($\sim$ 115 meV), respectively.[255] For NO, experiments reveal a desorption energy of 1300 K (2.6 kcal/mol or 112 meV) when using temperature programmed desorption measurements[320] after exposure of NO/H_2O ice to atomic oxygen. Thus, for both diatomics, the present findings are consistent with experiments and earlier simulations.

For the triatomics, computed average desorption energies for CO_2 and NO_2 from ASW were 9.0 ± 1.9 kcal/mol (4529 K or 390 meV) and 13.1 ± 3.4 kcal/mol (6592 K and 568 meV). For CO_2, one experiment reported $E_{des} = 2490 \pm 240$ K[321] (4.9 ± 0.47 kcal/mol or 214 ± 20 meV) whereas a compilation of literature data reported desorption energies ranging from 4.5 to 6.0 kcal/mol,[322] all of which are considerably lower than the computed values. This indicates that the electrostatic interaction between CO_2 and the surface is somewhat too strong. Yet, recomputing CO_2 point charges using an alternative and potentially more physical NBO[309] analysis yields increased polarity and interaction strengths with ASW for CO_2 (see Figure 4.36), which is inconsistent with experimental findings. On the other hand, no experimentally measured data are available for NO_2;[322] however, the desorption temperatures of NO_2 and H_2O from the same ice surface were found to be comparable, which indicates that NO_2 interacts strongly with the surface.[323] Overall, for all four species relevant to the present work, the computed E_{des} compare reasonably (CO_2) to favorably (CO, NO) with reference values from experiments and previous calculations, particularly in light of the fact that appreciable variations in experimentally reported values exist.[322]

Species Diffusion on ASW

To determine typical diffusion barrier heights, longer simulations for adsorbants CO, NO, CO_2, and NO_2 on water surface were run with velocities from a Maxwell-Boltzmann distribution at 50 K. Guided by earlier work[71] and to broadly probe the ASW surface, the total simulation time considered was 30 ns. For each frame, the interaction energy between adsorbants and the ASW surface was determined and mapped onto the ASW surface to give a two-dimensional map; see Figures 4.37 and 4.38. Figure 4.37A shows the 2D projection of CO interaction energy with ASW along the entire trajectory. Close to 20 minima are probed, which suggests that the
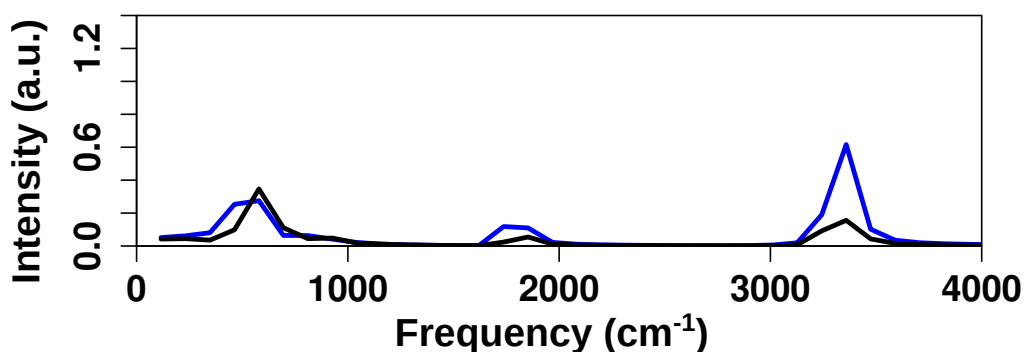
Figure 4.35: The vDOS spectrum for 10 randomly chosen surface water molecules (including W_A and W_B) before formation of CO_2 (black, thermal equilibrium) and 120 ps after CO+O recombination (blue) on the MMH-PES. The spectrum at t = 120 ps reproduces all features of the equilibrium vDOS, i.e. the system has returned to equilibrium on the 120 ps time scale. The intensities are "populations of modes within a given frequency interval".

simulations are sufficiently long to enable a representative exploration of the surface. The black solid line traces the path followed by the CO molecule on top of the ASW surface. Figure 4.37B shows the total path length with several minima enumerated and the interaction energy between CO and the ASW. The diffusion barrier heights vary between minima and are a consequence of ASW surface roughness.[71] For CO, the average activation barrier for diffusion between neighboring minima is 1.15 kcal/mol, with the lowest and highest barriers being 1.0 kcal/mol and 2.7 kcal/mol, respectively. Experimentally, a range of average diffusion barriers was found to extend from $E_b \sim 120 \pm 180$ K to 490 ± 12 K.[281,324] Another study obtained E_b separately for weakly and strongly bound sites, reporting ~ 350 K and ~ 930 K, respectively.[275] Hence, experimentally the average diffusion barriers range from 0.25 kcal/mol to 1.85 kcal/mol. The present simulations support a distribution of barriers, and these are consistent with the experimentally reported values.

The calculated average activation barrier for diffusion of NO (Figure4.38A) between neighboring minima is 0.91 kcal/mol with minimum and maximum barrier heights of 0.5 kcal/mol and 5.1 kcal/mol, respectively. From equilibrium trajectories at 50 K, no diffusion for CO_2 and NO_2 is observed on the 75 ns timescale. For CO_2, one experiment reported an average diffusion barrier of 2150 K (4.3 kcal/mol).[291] As a comparison, for CO the average diffusion barrier height is ~ 1 kcal/mol, which leads to a transition time between neighboring minima of ~ 1 ns. Within an Arrhenius
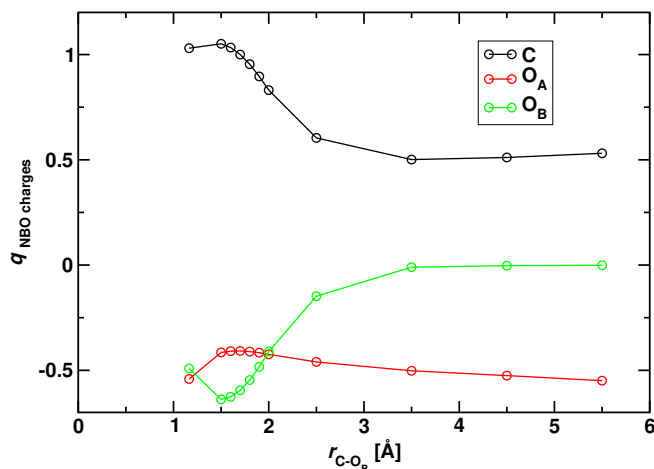
Figure 4.36: Fluctuating point charges from a natural bond order (NBO) analysis.[309] Calculations were carried out for the same snapshots as those in Figure 4.22. However, NBO charges are yet larger in magnitude and increase the interaction between CO_2 and the ASW which is inconsistent with experiments.[321,322]

picture for surface diffusion —which may not be entirely appropriate at sufficiently low temperatures[71]— a barrier of ∼ 4 kcal/mol corresponds to a diffusion timescale on the order of several hundred nanoseconds. This is consistent with the present findings that CO_2 does not diffuse on ASW on timescales of less than 100 ns. For comparison, the diffusion of CO_2 and NO_2 on the water surface after recombination is shown in Figure 4.39.

Earlier work already demonstrated that computed oxygen diffusion coefficients are consistent with measurements. Atomic oxygen[71] on ASW experiences diffusional barriers ranging from E_{dif} = 0.2 kcal/mol to 2 kcal/mol (100 K to 1000 K), compared with the average value of E_{dif} = 990^{+530}_{-360} K determined from experiments.[274]

4.3.5 Discussion and Conclusion

The findings of the present work support the conclusion that larger molecules can form and stabilize on ASW from atomic and molecular constituents. The rate-limiting step for CO_2 and NO_2 generation is the diffusion of the reactant species.
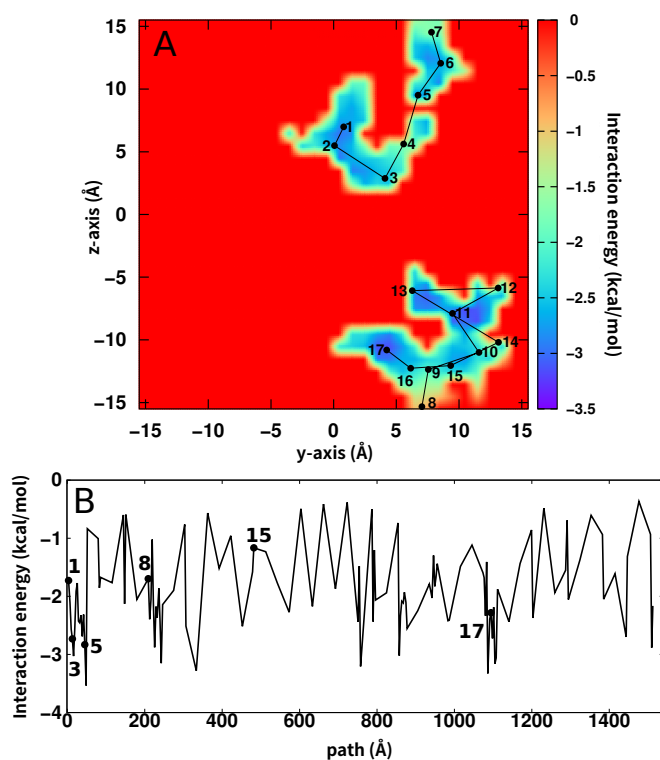
Figure 4.37: Diffusion of CO on ASW. Panel A: 2D projection of CO interaction energy and diffusion path onto the ASW $y - z$ plane from a 30 ns long simulation at 50 K. Panel B: 1D projection of the path followed with some of the minima visited in panel A labeled with corresponding numbers. Between minima 15 and 17 the adsorbant repeatedly visits minima 8 to 17. The average diffusion barrier height is 1.15 kcal/mol compared with a range from 0.25 kcal/mol to 1.85 kcal/mol from experiments.[275,324]
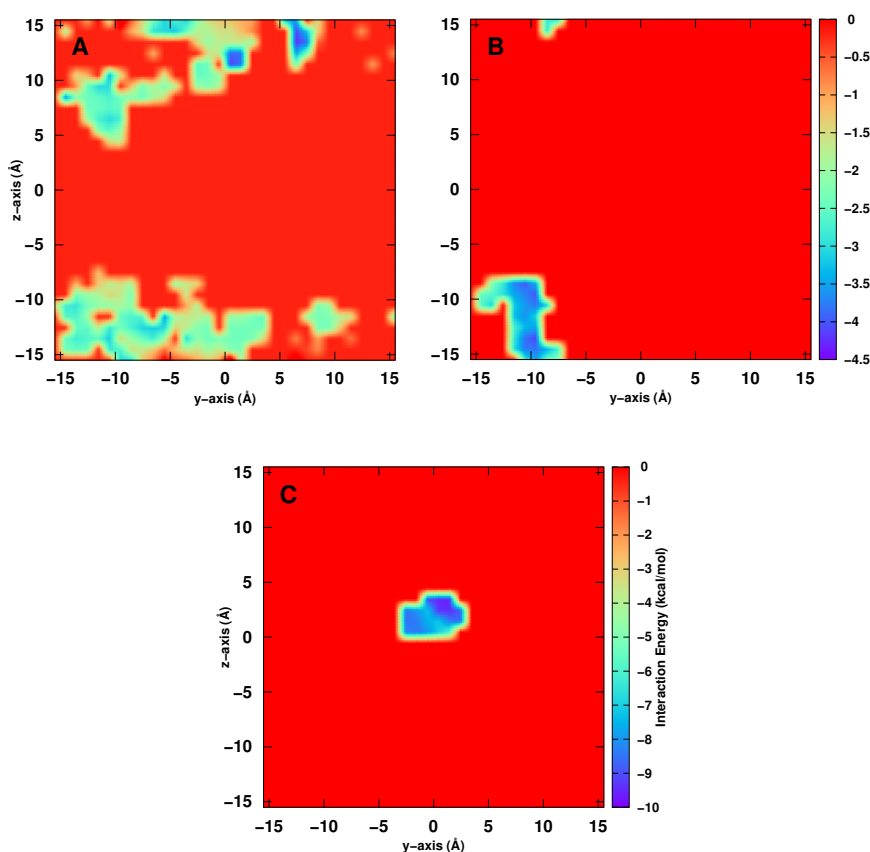
Figure 4.38: Interaction energy between the respective adsorbate and the ASW projected onto the y/z−plane from a 30 ns long simulation for NO (A), CO_2 (B), and NO_2 (C) on ASW at 50 K. Note that the color codes in panels A and B are the same but differ from that in panel C.
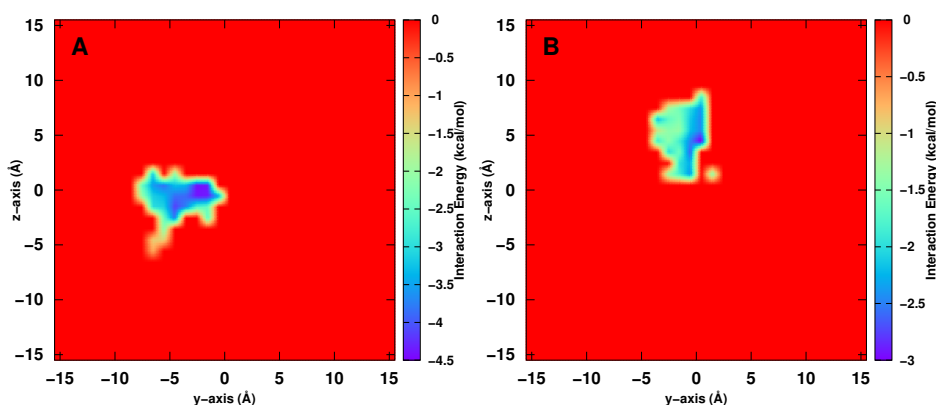
Figure 4.39: Diffusion of CO_2 (Panel A) and NO_2 (Panel B) on the water surface after recombination ($XO+O \longrightarrow XO_2$) was observed in simulations running for 10 ns and 6 ns, respectively. This allows for comparison with the diffusion at 50 K, as shown in Figure 4.38.

After recombination, both products efficiently cool on timescales of approximately 10 ps, although a smaller fraction still remains in an internally hot state. Use of a fluctuating charge model clarifies that, on the nanosecond timescale, recombination can occur from atom–molecule separations of up to at least ~ 20 Å, which is much larger than the average separation between neighboring adsorption sites. This finding is in stark contrast to earlier work, which reported nonvanishing recombination probabilities only over a range of ~ 5 Å due to using a simplified electrostatic model. Hence, it is concluded that recombination reactions for barrierless processes on ASW can occur on nanosecond timescales. It is also found that simplified models for the intramolecular energies, such as MMH, can yield qualitatively recombination probabilities, but are probably less suitable for studying energy relaxation because of the simplifications made. This is relevant when extending such studies to investigations of the formation of larger molecules for which construction of global and accurate machine-learning based PESs is still very challenging or even impossible. [143,325] Improved models, such as those based on MS-ARMD, provide viable future alternatives. [326]

Given the much longer timescales available in interstellar chemistry, the areas covered by the reactants will increase in proportion. The energy liberated upon recombination is efficiently transferred into water-internal modes and lattice vibrations (phonons) on picosecond timescales. Of particular note is the vanishingly small equilibrium diffusivity of CO_2 and NO_2 at low temperatures (here 50 K), which is consistent with experiments, whereas immediately after recombination the triatomics diffuse and roam wider parts of the ASW. For NO_2, 11 % of the recombined trajectories show desorption in simulations of 1 ns in length, whereas no desorption is observed for CO_2.

Astrophysical implications of the present work include physical and chemical aspects alike. Recombination reactions liberate energy to drive restructuring of the underlying surface. As a consequence of recombination, local heating of the ASW can take place, which also potentially increases the local diffusivity and leads to desorption of reactive species, such as H or OH. This is particularly relevant as the cold average temperatures (~ 10 K) restrict translational motion to the lightest particles. [327] Because the morphology of a surface in part determines its chemi-

cal activity, reconfiguration of the solid support is an important determinant for downstream reactivity. Our simulations also show that following recombination, products undergo nonthermal surface diffusion (see Figure 4.39), which leads to mass transport across the surface to initiate chemistry at other locations. This is particularly relevant because products can remain in an internally excited state, which can lead to breakup and radical formation.

In summary, the present work establishes that triatomics can be formed and stabilized following atom+diatom recombination reactions. The kernel-based potential energy surfaces combined with physically meaningful electrostatic models allow us to run statistically significant numbers of trajectories and provide molecular-level insight into energy relaxation and redistribution. This work provides a basis for broader reaction exploration on ASW, in particular in light of recent progress in representing high-dimensional, reactive PESs based on neural networks.[326]

Chapter 5

Conclusion and Outlook

In this thesis, reactive MD simulations combined with accurate representations of PESs were used to gain molecular-level insights into gas and condensed phase reactions. The methods for describing intermolecular interaction potentials varied with system complexity: RKHS interpolation of energies for small systems (triatomics), NN potentials and fitted FFs for medium-sized molecules (up to 10 atoms), and an ML/MM approach for condensed phase systems. The published papers presented in the previous chapters are summarized below.

Chapter 3.1 examined the unimolecular decomposition of syn-CH$_3$CHOO into CH$_2$CHO and OH radicals following thermal and vibrational excitation. To perform the dynamics, two representations of PESs at the MP2 level were used: a NN-based approach using the PhysNet architecture and a more empirical MS-ARMD method. The calculated reaction times for OH elimination were on the nanosecond time scale and the energy dependence of the rates was consistent with the experimental data. Results showed that the initial preparation of the reactants influence the VHP formation, OH elimination and OH yields. OH production was higher and faster with thermal excitation due to the uniform energy distribution across all modes. In contrast, vibrational excitation required time for energy redistribution to channel energy back into the progression coordinate, resulting in a non-zero probability of forming VHP even after 1 ns of excitation. The dependence of OH formation rates on D_e^{OO} (31.5 kcal/mol at MP2 level and 23.5 kcal/mol, closer to earlier CASPT2 calculations[43]) was also determined from a significant number of
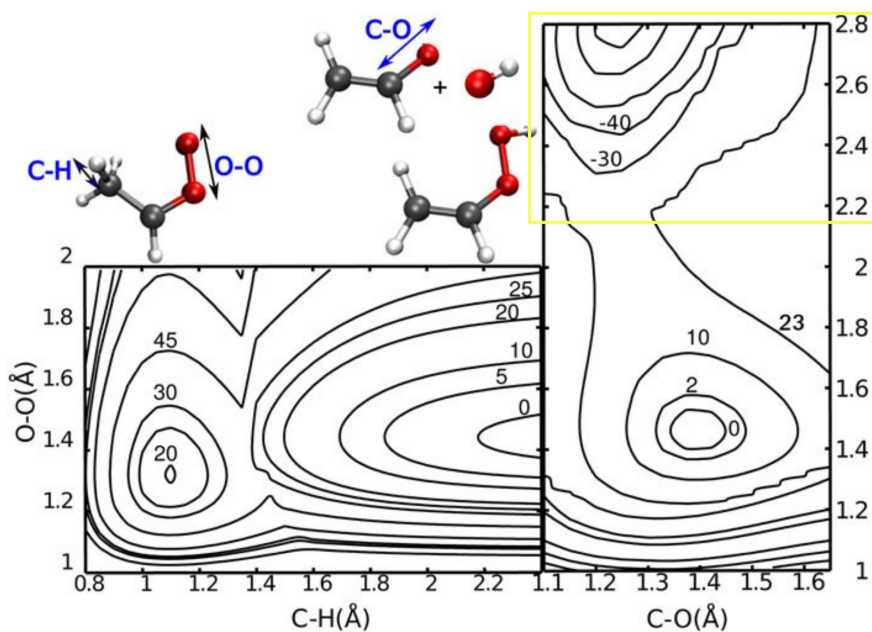
Figure 5.1: Two-dimensional cuts through the full-dimensional MP2 level PES for the *syn*-CH$_3$CHOO $\longrightarrow$ VHP $\longrightarrow$ CH$_2$CHO + OH decomposition channel. The PES is relaxed on each grid point. For the first step (left panel), the C–H (proton transfer) and O–O coordinates are the driving coordinates whereas for the second OH elimination step (right panel), the C–O and O–O separation are used, respectively. All energies are given in kcal/mol and the minimum of the VHP state is the zero of energy. The figure is used by Joel Bowman in ref.[328]

trajectories. However, the unsuitable features for large O–O separation made PESs inadequate for final state analysis. As highlighted in Figure 5.1, the MP2 energies exhibited incorrect asymptotic behavior for large O–O separations and failed to describe the electrostatics accurately.

In Chapter 3.2, the long-range behavior for OH elimination was correctly captured by transfer learning the MP2 level PES to the CASPT2 level of theory. The computed final OH-translational and rotational state distributions were compared with experiments to determine the unknown O–O bond strength D$_e^{OO}$. The best values found for D$_e^{OO}$ ranged from 22 to 25 kcal/mol. OH-elimination rates were consistent with experiments and remained largely unaffected by variations in D$_e^{OO}$ due to the non-equilibrium nature of the process. In addition to the OH-elimination pathway, OH roaming was observed following O–O scission, leading to GA formation on the picosecond time scale. This observation, along with recent experimental findings of

similar OH roaming behavior in the unimolecular decay of methyl-ethyl substituted CI, suggests that OH roaming is a general pathway to be included in molecular-level modeling of atmospheric processes. Following GA decomposition, other atmospheric species[52,329] such as CH_2OH, CHO, $HOCH_2CO$ and H radicals were also observed. The relevance of OH roaming in GA formation and the rapid generation of other radicals on picosecond timescales is likely crucial for understanding the complex and largely unexplored chemistry of the atmosphere.

The ability to explore the unimolecular decomposition of *syn*-CH_3CHOO using atomistic simulations and high-level ML potential has substantially advanced our understanding. Of course, the direct comparison with experimental observables has allowed an assessment of the methods and validation of models. Performing a single *ab initio* MD (AIMD) simulation at the CASPT2 level of theory is currently computationally infeasible, requiring ~106 hours of CPU time for a 200 ps run. In contrast, the approach used in this thesis took ~6 hours, providing a balance between accuracy and speed.[138,152,220] Nevertheless, generating a dataset that thoroughly covers the configurational and conformational space remains labour intensive. While empirical FFs (here MS-ARMD) allow for efficient and extended timescale simulations in mere minutes for a 1 ns simulation, but fall short in accuracy. Future research utilizing the ML/MM approach to explore the reactivity of CIs in aqueous environments will further deepen our understanding of atmospheric processes. Another significant improvement would be the inclusion of nonadiabatic effects to explore the fate of electronically excited CIs, which presents a considerable challenge. Emerging ML-based techniques[330–332] are beginning to address these issues, offering promising solutions.

Transitioning to condensed phase systems, Chapter 4.1 characterized the formation of CO_2 from the barrierless recombination of $CO(^1\Sigma)$ and $O(^1D)$ "on" ASW under conditions typical in cold molecular clouds. Employing an ML/MM approach, the dynamics were modeled using a rigid TIP3P water model combined with either a CCSD(T) level RKHS PES or a Morse representation of the reactive O–CO potential fitted at the MRCI level of theory. The simulations were initiated from different separation distances (R) and angular (θ) orientations (see inset in Figure 5.2) using

a fixed point charge model for CO and O. Results showed that CO_2 formation and stabilization from collision of CO and $O(^1D)$ on ASW is possible and occurs on the subnanosecond time scale for the reaction partners within close range ($\approx$6 Å). The effective redistribution of the released energy to the water surface played a pivotal role in stabilizing the resulting products. Furthermore, the formation and stabilization of COO intermediate supported the eventual generation of C + O_2.
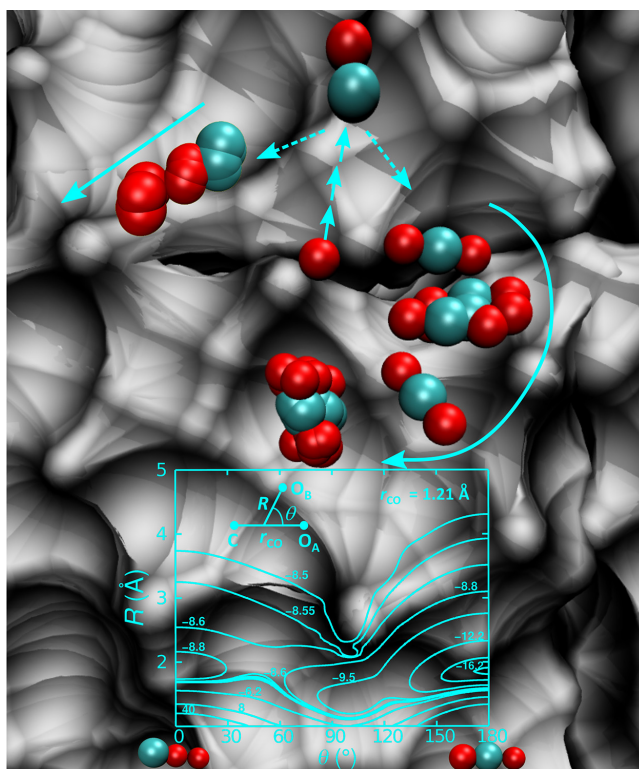
Figure 5.2: CO + O recombination on ASW surface. The ground-state PES supports both the global OCO and the local COO minimum energy structure. Both are formed and stabilized from reactive MD simulations using the RKHS interpolated CCSD(T) reference energies.

Taking a step further, Chapter 4.2 explored the energy distribution across ASW following CO_2 formation "on" the surface and "in" a cavity using a flexible KKY water model. It was found that energy distribution occurs in two phases, one on the picosecond and one on the nanosecond time scale for both locations. Although the time dependence of the processes is similar, the coupling between CO_2 and water differs for recombination within ASW and on its surface. Recombination within the cavity considered led to a smaller magnitude (~15%) of energy transferred per

water molecule compared with the process on the surface (~25%). This disparity was attributed to the reason that within a sufficiently large cavity, the recombined CO_2 molecule exchanges energy with the surrounding environment through direct collision, whereas on the surface, CO_2 is consistently in contact with the ASW. Mostly, within 50 ps, a uniform temperature increase of the ASW across the entire surface was observed.

With fixed point charges in Chapters 4.1 and 4.2, the reactants interacted more strongly with the water hydrogen atoms and only reaction partners within close range (~6 Å) recombined to CO_2. In Chapter 4.3, the FPC model replaced the fixed point charge model. Together with CO_2, NO_2 formation from NO + O is also investigated using MRCI level RKHS PES. After the replacement, recombination was even observed from $R = 20$ Å for both CO_2 and NO_2 on a 500 ps timescale, underlining that the rate-limiting step for CO_2 and NO_2 generation is the diffusion of the reactant species. After recombination, both CO_2 and NO_2 stabilized by energy transfer to water internal and surface phonon modes on the picosecond time scale. Furthermore, the average diffusion barriers and desorption energies for the involved species agree with those reported from experiments.

The work establishes that triatomics can form and stabilize on the ASW following atom + diatom recombination reactions. The kernel-based PESs combined with physically meaningful electrostatic models enable to run statistically significant numbers of trajectories and provide molecular-level insights into energy relaxation and redistribution in cold surfaces.

Moving forward, there are several exciting avenues for further investigation. Firstly, the exploration of the competition between CO + O and NO + O recombination processes on the same surface under similar conditions would be valuable. Additionally, while the present work employed rigid TIP3P and flexible KKY water models involving fixed point charges, adopting polarizable water models[333] or ML potentials can provide valuable insights. This would allow for explicit probing of the coupling modes between newly formed vibrationally excited molecules and surface water molecules, together with vDOS analysis for more precise understanding of the

energy dissipation mechanism.

In summary, increasing computational power, when paired with appropriate tools, has opened up new avenues for the exploration of chemical systems at the atomistic level on a significant time scale. The approaches used here are generic and expected to be applicable to a range of other reactions. Through rigorous exploration, this thesis has thoroughly investigated the complex dynamics of chemical systems, uncovering pathways of decomposition and the formation of atmospheric and astrophysical relevant species. These advancements in methodology and chemical insights may contribute to future advancements in the field.

Bibliography

[1] Feynman, R. P. *QED: The strange theory of light and matter.* Princeton University Press: **1985**.

[2] Levitt, M.; Lifson, S. *J. Mol. Biol.* **1969**, *46*, 269–279.

[3] Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002**, *9*(*9*), 646–652.

[4] Hollingsworth, S. A.; Dror, R. O. *Neuron* **2018**, *99*(*6*), 1129–1143.

[5] De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. *J. Med. Chem.* **2016**, *59*(*9*), 4035–4061.

[6] Meuwly, M. *J. Phys. Chem. B* **2022**, *126*(*11*), 2155–2167.

[7] Töpfer, K.; Upadhyay, M.; Meuwly, M. *Phys. Chem. Chem. Phys.* **2022**, *24*(*21*), 12767–12786.

[8] Karplus, M.; Sharma, R. D.; Porter, R. N. *J. Chem. Phys.* **1964**, *40*(*7*), 2033–2034.

[9] Koner, D.; Salehi, S. M.; Mondal, P.; Meuwly, M. *J. Chem. Phys.* **2020**, *153*(*1*), 010901.

[10] Dommer, A.; Casalino, L.; Kearns, F.; Rosenfeld, M.; Wauer, N.; Ahn, S.-H.; Russo, J.; Oliveira, S.; Morris, C.; Bogetti, A.; et al. *Int. J. High Perform. Comput. Appl.* **2023**, *37*(*1*), 28–44.

[11] Jung, J.; Kobayashi, C.; Kasahara, K.; Tan, C.; Kuroda, A.; Minami, K.; Ishiduki, S.; Nishiki, T.; Inoue, H.; Ishikawa, Y.; et al. *J. Comput. Chem.* **2021**, *42(4)*, 231–241.

[12] Bowman, J. M.; Shepler, B. C. *Ann. Rev. Phys. Chem.* **2011**, *62*, 531–553.

[13] Townsend, D.; Lahankar, S. A.; Lee, S. K.; Chambreau, S. D.; Suits, A. G.; Zhang, X.; Rheinecker, J.; Harding, L.; Bowman, J. M. *Science* **2004**, *306(5699)*, 1158–1161.

[14] Upadhyay, M.; Töpfer, K.; Meuwly, M. *J. Phys. Chem. Lett.* **2023**, *15(1)*, 90–96.

[15] Herbst, E. *Chem. Soc. Rev.* **2001**, *30(3)*, 168–176.

[16] Champion, K.; Cole, A.; Kantor, A. *Handbook of geophysics and the space environment.* Air Force Geophysics Laboratory: **1985**.

[17] Seinfeld, J. H.; Pandis, S. N. *Atmospheric chemistry and physics: from air pollution to climate change.* John Wiley & Sons: **2016**.

[18] Lydolph, P. E. *The climate of the earth.* Government Institutes: **1985**.

[19] Vereecken, L.; Aumont, B.; Barnes, I.; Bozzelli, J.; Goldman, M.; Green, W.; Madronich, S.; Mcgillen, M.; Mellouki, A.; Orlando, J.; et al. *Int. J. Chem. Kin.* **2018**, *50(6)*, 435–469.

[20] Linnartz, H.; Ioppolo, S.; Fedoseev, G. *Int. Rev. Phys. Chem.* **2015**, *34(2)*, 205–237.

[21] Abbatt, J.; Ravishankara, A. R. *Atmos. Chem. Phys.* **2023**, *23(17)*, 9765–9785.

[22] Berndt, T.; Chen, J.; Kjærgaard, E. R.; Møller, K. H.; Tilgner, A.; Hoffmann, E. H.; Herrmann, H.; Crounse, J. D.; Wennberg, P. O.; Kjaergaard, H. G. *Science* **2022**, *376(6596)*, 979–982.

[23] Fried, Z. T.; El-Abd, S. J.; Hays, B. M.; Wenzel, G.; Byrne, A. N.; Margulès, L.; Motiyenko, R. A.; Shipman, S. T.; Horne, M. P.; Jørgensen, J. K.; et al. *Astrophys. J. Lett.* **2024**, *965(2)*, L23.

[24] Sanz-Novo, M.; Rivilla, V. M.; Müller, H. S.; Jiménez-Serra, I.; Martín-Pintado, J.; Colzi, L.; Zeng, S.; Megías, A.; López-Gallifa, Á.; Martínez-Henares, A.; et al. *Astrophys. J. Lett.* **2024**, *965(2)*, L26.

[25] Fuentetaja, R.; Agúndez, M.; Cabezas, C.; Tercero, B.; Marcelino, N.; Pardo, J.; Vicente, P.; Cernicharo, J. *Astron. Astrophys.* **2022**, *667*, L4.

[26] Harries, C. *Justus Liebigs Ann. Chem.* **1905**, *343(2-3)*, 311–344.

[27] Criegee, R.; Wenner, G. *Justus Liebigs Ann. Chem.* **1949**, *564(1)*, 9–15.

[28] Taatjes, C. A. *Ann. Rev. Phys. Chem.* **2017**, *68*, 183–207.

[29] Orr-Ewing, A. J.; Osborn, D. L. *J. Phys. Chem. A* **2024**, *128(15)*, 2909–2911.

[30] Vereecken, L.; Glowacki, D. R.; Pilling, M. J. *Chem. Rev.* **2015**, *115(10)*, 4063–4114.

[31] Alam, M. S.; Camredon, M.; Rickard, A. R.; Carr, T.; Wyche, K. P.; Hornsby, K. E.; Monks, P. S.; Bloss, W. J. *Phys. Chem. Chem. Phys.* **2011**, *13(23)*, 11002–11015.

[32] Novelli, A.; Vereecken, L.; Lelieveld, J.; Harder, H. *Phys. Chem. Chem. Phys.* **2014**, *16(37)*, 19941–19951.

[33] Mauldin Iii, R.; Berndt, T.; Sipilä, M.; Paasonen, P.; Petäjä, T.; Kim, S.; Kurtén, T.; Stratmann, F.; Kerminen, V.-M.; Kulmala, M. *Nature* **2012**, *488(7410)*, 193–196.

[34] Emmerson, K. M.; Carslaw, N.; Carslaw, D.; Lee, J. D.; McFiggans, G.; Bloss, W. J.; Gravestock, T.; Heard, D. E.; Hopkins, J.; Ingham, T.; et al. *Atmos. Chem. Phys.* **2007**, *7(1)*, 167–181.

[35] Emmerson, K. M.; Carslaw, N. *Atmos. Environ.* **2009**, *43(20)*, 3220–3226.

[36] Liu, F.; Beames, J. M.; Petit, A. S.; McCoy, A. B.; Lester, M. I. *Science* **2014**, *345(6204)*, 1596–1598.

[37] Kuwata, K. T.; Hermes, M. R.; Carlson, M. J.; Zogg, C. K. *J. Phys. Chem. A* **2010**, *114(34)*, 9192–9204.

[38] Vereecken, L.; Novelli, A.; Taraborrelli, D. *Phys. Chem. Chem. Phys.* **2017**, *19(47)*, 31599–31612.

[39] Gutbrod, R.; Kraka, E.; Schindler, R. N.; Cremer, D. *J. Am. Chem. Soc.* **1997**, *119(31)*, 7330–7342.

[40] Long, B.; Bao, J. L.; Truhlar, D. G. *J. Am. Chem. Soc.* **2016**, *138(43)*, 14409–14422.

[41] Yin, C.; Takahashi, K. *Phys. Chem. Chem. Phys.* **2017**, *19(19)*, 12075–12084.

[42] Fang, Y.; Liu, F.; Barber, V. P.; Klippenstein, S. J.; McCoy, A. B.; Lester, M. I. *J. Chem. Phys.* **2016**, *144(6)*, 061102.

[43] Kidwell, N. M.; Li, H.; Wang, X.; Bowman, J. M.; Lester, M. I. *Nat. Chem.* **2016**, *8(5)*, 509–514.

[44] Fang, Y.; Liu, F.; Barber, V. P.; Klippenstein, S. J.; McCoy, A. B.; Lester, M. I. *J. Chem. Phys.* **2016**, *145(23)*, 234308.

[45] Kurtén, T.; Donahue, N. M. *J. Phys. Chem. A* **2012**, *116(25)*, 6823–6830.

[46] Klippenstein, S.J.; Elliott, S. N. *J. Phys. Chem. A* **2023**, *127(50)*, 10647–10662.

[47] Braams, B. J.; Bowman, J. M. *Intern. Rev. Phys. Chem.* **2009**, *28(4)*, 577–606.

[48] Upadhyay, M.; Meuwly, M. *ACS Earth Space Chem.* **2021**, *5(12)*, 3396–3406.

[49] Lee, A.; Goldstein, A. H.; Kroll, J. H.; Ng, N. L.; Varutbangkul, V.; Flagan, R. C.; Seinfeld, J. H. *J. Geophys. Res. Atmos.* **2006**, *111(D17)*, D17305.

[50] Niki, H.; Maker, P.; Savage, C.; Breitenbach, L. *Chem. Phys. Lett.* **1981**, *80(3)*, 499–503.

[51] Bertschi, I.; Yokelson, R. J.; Ward, D. E.; Babbitt, R. E.; Susott, R. A.; Goode, J. G.; Hao, W. M. *J. Geophys. Res. Atmos.* **2003**, *108(D13)*, 8472.

[52] Conrad, A. R.; Hansen, N.; Jasper, A. W.; Thomason, N. K.; Hidaldo-Rodrigues, L.; Treshock, S. P.; Popolan-Vaida, D. M. *Phys. Chem. Chem. Phys.* **2021**, *23(41)*, 23554–23566.

[53] Liu, T.; Elliott, S. N.; Zou, M.; Vansco, M. F.; Sojdak, C. A.; Markus, C. R.; Almeida, R.; Au, K.; Sheps, L.; Osborn, D. L.; et al. *J. Am. Chem. Soc.* **2023**, *145(35)*, 19405–19420.

[54] Hollis, J. M.; Lovas, F. J.; Jewell, P. R. *Astrophys. J.* **2000**, *540(2)*, L107.

[55] Powner, M. W.; Gerland, B.; Sutherland, J. D. *Nature* **2009**, *459(7244)*, 239–242.

[56] Robertson, M. P.; Joyce, G. F. *Cold Spring Harb. Perspect. Biol.* **2012**, *4(5)*, a003608.

[57] Ferrière, K. M. *Rev. Mod. Phys.* **2001**, *73(4)*, 1031–1066.

[58] Yamada, K. M. T.; Winnewisser, G. *Interstellar molecules: their laboratory and interstellar habitat* volume 241. Springer: **2011**.

[59] Tennyson, J. *Astronomical Spectroscopy: An Introduction to the Atomic and Molecular Physics of Astronomical Spectroscopy.* World Scientific: **2019**.

[60] McGuire, B. A. *Astrophys. J. Suppl. Ser.* **2022**, *259(2)*, 30.

[61] Tielens, A. G. G. M. *Rev. Mod. Phys.* **2013**, *85(3)*, 1021–1081.

[62] Wakelam, V.; Smith, I.; Herbst, E.; Troe, J.; Geppert, W.; Linnartz, H.; Öberg, K.; Roueff, E.; Agúndez, M.; Pernot, P. e.; et al. *Space Sci. Rev.* **2010**, *156*, 13–72.

[63] Ruffle, D. P.; Herbst, E. *Mon. Not. R. Astron. Soc.* **2001**, *324(4)*, 1054–1062.

[64] Wakelam, V.; Bron, E.; Cazaux, S.; Dulieu, F.; Gry, C.; Guillard, P.; Habart, E.; Hornekær, L.; Morisset, S.; Nyman, G. *Mol. Astrophys.* **2017**, *9*, 1–36.

[65] Bar-Nun, A.; Dror, J.; Kochavi, E.; Laufer, D. *Phys. Rev. B* **1987**, *35(5)*, 2427–2435.

[66] Ioppolo, S.; Cuppen, H. M.; Linnartz, H. *Rend. Lincei.* **2011**, *22*(*3*), 211–224.

[67] Romanzin, C.; Ioppolo, S.; Cuppen, H. M.; Dishoeck, E. F.; Linnartz, H. *J. Chem. Phys.* **2011**, *134*(*8*), 084504.

[68] Chaabouni, H.; Minissale, M.; Manicò, G.; Congiu, E.; Noble, J. A.; Baouche, S.; Accolla, M.; Lemaire, J. L.; Pirronello, V.; Dulieu, F. *J. Chem. Phys.* **2012**, *137*(*23*), 234706.

[69] Minissale, M.; Congiu, E.; Baouche, S.; Chaabouni, H.; Moudens, A.; Dulieu, F.; Accolla, M.; Cazaux, S.; Manico, G.; Pirronello, V. *Phys. Rev. Lett.* **2013**, *111*(*5*), 053201.

[70] Dulieu, F.; Minissale, M.; Bockelée-Morvan, D. *Astron. Astrophys.* **2017**, *597*, A56.

[71] Pezzella, M.; Unke, O. T.; Meuwly, M. *J. Phys. Chem. Lett.* **2018**, *9*(*8*), 1822–1826.

[72] Pezzella, M.; Meuwly, M. *Phys. Chem. Chem. Phys.* **2019**, *21*(*11*), 6247–6255.

[73] Christianson, D. A.; Garrod, R. T. *Front. Astron. Space Sci.* **2021**, *8*, 643297.

[74] Hama, T.; Watanabe, N. *Chem. Rev.* **2013**, *113*(*12*), 8783–8839.

[75] Minissale, M.; Congiu, E.; Manicò, G.; Pirronello, V.; Dulieu, F. *Astron. Astrophys.* **2013**, *559*, A49.

[76] Minissale, M.; Moudens, A.; Baouche, S.; Chaabouni, H.; Dulieu, F. *Mon. Not. R. Astron. Soc.* **2016**, *458*(*3*), 2953–2961.

[77] Qasim, D.; Fedoseev, G.; Chuang, K.-J.; He, J.; Ioppolo, S.; Dishoeck, E.; Linnartz, H. *Nat. Astron.* **2020**, *4*(*8*), 781–785.

[78] Molpeceres, G.; K, J.; Fedoseev, G.; Qasim, D.; Schömig, R.; Linnartz, H.; Lamberts, T. *J. Phys. Chem. Lett.* **2021**, *12*, 10854–10860.

[79] Minissale, M.; Fedoseev, G.; Congiu, E.; Ioppolo, S.; Dulieu, F.; Linnartz, H. *Phys. Chem. Chem. Phys.* **2014**, *16*, 8257–8269.

[80] Wakelam, V.; Cuppen, H. M.; Herbst, E. *Astrochemistry: synthesis and modelling. In Astrochemistry and Astrobiology.* 115–143, Springer: **2013**.

[81] Wakelam, V.; Herbst, E.; Loison, J.-C.; Smith, I.; Chandrasekaran, V.; Pavone, B.; Adams, N.; Bacchus-Montabonel, M.-C.; Bergeat, A.; Béroff, K.; et al. *Astrophys. J. Suppl. Ser.* **2012**, *199(1)*, 21.

[82] Millar, T.; Walsh, C.; Sande, M.; Markwick, A. *Astron. Astrophys.* **2024**, *682*, A109.

[83] Cuppen, H.; Karssemeijer, L.; Lamberts, T. *Chem. Rev.* **2013**, *113(12)*, 8840–8871.

[84] Chang, Q.; Herbst, E. *Astrophys. J.* **2014**, *787(2)*, 135.

[85] Arasa, C.; Hemert, M. C.; Dishoeck, E. F.; Kroes, G.-J. *J. Phys. Chem. A* **2013**, *117(32)*, 7064–7074.

[86] Ferrero, S.; Pantaleone, S.; Ceccarelli, C.; Ugliengo, P.; Sodupe, M.; Rimola, A. *Astrophys. J.* **2023**, *944(2)*, 142.

[87] Upadhyay, M.; Pezzella, M.; Meuwly, M. *J. Phys. Chem. Lett.* **2021**, *12*, 6781–6787.

[88] Upadhyay, M.; Meuwly, M. *arXiv:2403.15141* **2024**.

[89] Upadhyay, M.; Meuwly, M. *Front. Chem.* **2022**, *9*, 827085.

[90] Veliz, J. C. S. V.; Koner, D.; Schwilk, M.; Bemish, R. J.; Meuwly, M. *Phys. Chem. Chem. Phys.* **2020**, *22(7)*, 3927–3939.

[91] Ndengué, S.; Quintas-Sánchez, E.; Dawes, R.; Osborn, D. *J. Phys. Chem. A* **2021**, *125(25)*, 5519–5533.

[92] Schrödinger, E. *Phys. Rev.* **1926**, *28(6)*, 1049–1070.

[93] Dirac, P. A. M. *Proc. R. Soc. London, Ser. A* **1929**, *123(792)*, 714–733.

[94] Jensen, F. *Introduction to computational chemistry.* John wiley & sons: **2017**.

[95] Born, M.; Oppenheimer, R. *Ann. Phys.* **1927**, *389*, 457–484.

[96] Magalhaes, A. L. *J. Chem. Educ.* **2014**, *91(12)*, 2124–2127.

[97] Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56(5)*, 2257–2261.

[98] Dunning Jr, T. H. *J. Chem. Phys.* **1989**, *90(2)*, 1007–1023.

[99] Klopper, W.; Manby, F. R.; Ten-No, S.; Valeev, E. F. *Int. Rev. Phys. Chem.* **2006**, *25(3)*, 427–468.

[100] Hartree, D. R. *Mathematical Proceedings of the Cambridge Philosophical Society.* **1928**.

[101] Fock, V. *Zeitschrift für Physik* **1930**, *61*, 126–148.

[102] Nesbet, R. *Proc. R. Soc. London, Ser. A* **1955**, *230(1182)*, 312–321.

[103] Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46(7)*, 618–622.

[104] Kümmel, H. G. *Int. J. Mod. Phys. B* **2003**, *17(28)*, 5311–5325.

[105] Roos, B. O. *Multiconfigurational quantum chemistry. In Theory and Applications of Computational Chemistry.* 725–764, Elsevier: **2005**.

[106] Roos, B. O.; Taylor, P. R.; Sigbahn, P. E. *Chem. Phys.* **1980**, *48(2)*, 157–173.

[107] Siegbahn, P. E.; Almlöf, J.; Heiberg, A.; Roos, B. O. *J. Chem. Phys.* **1981**, *74(4)*, 2384–2396.

[108] Roos, B. O.; Linse, P.; Siegbahn, P. E.; Blomberg, M. R. *Chem. Phys.* **1982**, *66*, 197–207.

[109] Pauli, W. *Zeitschrift für Physik* **1925**, *31*, 765–783.

[110] Griffiths, D. J.; Schroeter, D. F. *Introduction to quantum mechanics.* Cambridge university press: **2018**.

[111] Szabo, A.; Ostlund, N. S. *Modern quantum chemistry: introduction to advanced electronic structure theory.* Courier Corporation: **2012**.

[112] Löwdin, P.-O. *Phys. Rev.* **1955**, *97(6)*, 1509.

[113] Helgaker, T.; Jorgensen, P.; Olsen, J. *Molecular electronic-structure theory*. John Wiley & Sons: **2013**.

[114] Bunge, C. F. *Phys. Rev.* **1968**, *168*(*1*), 92.

[115] Sergeev, A. V.; Goodson, D. Z.; Wheeler, S. E.; Allen, W. D. *J. Chem. Phys.* **2005**, *123*(*6*), 064105.

[116] Gill, P. M.; Pople, J. A.; Radom, L.; Nobes, R. H. *J. Chem. Phys.* **1988**, *89*(*12*), 7307–7314.

[117] Bartlett, R. J.; Musiał, M. *Rev. Mod. Phys.* **2007**, *79*(*1*), 291.

[118] Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*(*4*), 1910–1918.

[119] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*(*6*), 479–483.

[120] Werner, H.-J.; Knowles, P. J. *J. Chem. Phys.* **1988**, *89*(*9*), 5803–5814.

[121] Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys* **1992**, *96*(*2*), 1218–1226.

[122] MacKerell Jr, A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. *J. Phys. Chem. B* **1998**, *102*(*18*), 3586–3616.

[123] Das, A. K.; Meuwly, M. *Methods Enzymol.* **2016**, *577*, 31–55.

[124] Brooks, B.; III, C. B.; Jr, A. M.; Nilsson, L.; Petrella, R.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R.; Post, C.; Pu, J.; Schaefer, M.; Tidor, B.; Venable, R.; Woodcock, H.; Wu, X.; Yang, W.; York, D.; Karplus, M. *J. Comput. Chem.* **2009**, *30*(*10*), 1545–1614.

[125] MacKerell Jr, A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. *J. Phys. Chem. B* **1998**, *102*(*18*), 3586–3616.

[126] Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*(*6*), 1657–1666.

[127] Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham III, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*(*1-3*), 1–41.

[128] González, M. A. *École thématique de la Société Française de la Neutronique* **2011**, *12*, 169–200.

[129] Morse, P. M. *Phys. Rev.* **1929**, *34*(*1*), 57.

[130] Jones, J. E. *Proc. R. Soc. London, Ser. A* **1924**, *106*(*738*), 463–477.

[131] Nagy, T.; Yosa Reyes, J.; Meuwly, M. *J. Chem. Theory Comput.* **2014**, *10*, 1366–1375.

[132] Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. *J. Comput. Chem.* **2011**, *32*(*11*), 2359–2368.

[133] Gaussian09, R. A. *Inc., Wallingford CT* **2009**, *121*, 150–166.

[134] Werner, H.-J.; Knowles, P. J.; Manby, F. R.; Black, J. A.; Doll, K.; Heßelmann, A.; Kats, D.; Köhn, A.; Korona, T.; Kreplin, D. A.; et al. *J. Chem. Phys.* **2020**, *152*(*14*), 144107.

[135] Nelder, J. A.; Mead, R. *Comput. J.* **1965**, *7*(*4*), 308–313.

[136] Unke, O. T.; Devereux, M.; Meuwly, M. *J. Chem. Phys.* **2017**, *147*(*16*), 161712.

[137] Storn, R.; Price, K. *J. Global Optim.* **1997**, *11*(*4*), 341–359.

[138] Song, K.; Upadhyay, M.; Meuwly, M. *Phys. Chem. Chem. Phys.* **2024**.

[139] Unke, O. T.; Meuwly, M. *J. Chem. Inf. and Mod.* **2017**, *57*(*8*), 1923–1931.

[140] Schölkopf, B.; Herbrich, R.; Smola, A. J. *International conference on computational learning theory.* Springer: **2001**.

[141] Golub, G. H.; Van Loan, C. *Matrix Computations.* JHU Press: **2013**.

[142] Tikhonov, A. N.; Arsenin, V. *Solutions of ill-posed problems.* Winston: Washington, DC: **1977**.

[143] Unke, O. T.; Meuwly, M. *J. Chem. Theory Comput.* **2019**, *15*(*6*), 3678–3693.

[144] Behler, J.; Parrinello, M. *Phys. Rev. Lett.* **2007**, *98*(*14*), 146401.

[145] Haykin, S. *Neural networks and learning machines.* London: Pearson Education: **2009**.

[146] Ripley, B. D. *Pattern recognition and neural networks.* Cambridge university press: **2007**.

[147] Hornik, K. *Neural networks* **1991**, *4*(*2*), 251–257.

[148] Gilmer, J.; Schoenholz, S. S.; Riley, P. F.; Vinyals, O.; Dahl, G. E. *Int. Conf. Mach. Learn.* **2017**, *70*, 1263–1272.

[149] Schütt, K.; Kindermans, P.-J.; Sauceda Felix, H. E.; Chmiela, S.; Tkatchenko, A.; Müller, K.-R. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 992–1002.

[150] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*(*15*), 154104.

[151] Baydin, A. G.; Pearlmutter, B. A.; Radul, A. A.; Siskind, J. M. *J. Mach. Learn. Res.* **2018**, *18*(*153*), 1–43.

[152] Käser, S.; Richardson, J. O.; Meuwly, M. *J. Chem. Theory Comput.* **2022**, *18*(*11*), 6840–6850.

[153] Käser, S.; Meuwly, M. *Phys. Chem. Chem. Phys.* **2022**, *24*(*9*), 5269–5281.

[154] Smith, J. S.; Isayev, O.; Roitberg, A. E. *Sci. Data* **2017**, *4*(*1*), 1–8.

[155] Leach, A. R. *Molecular modelling: principles and applications.* Pearson education: **2001**.

[156] Verlet, L. *Phys. Rev.* **1967**, *159*(*1*), 98.

[157] Ruth, R. D. *IEEE Trans. Nucl. Sci.* **1983**, *30*, 2669–2671.

[158] Stone, D.; Whalley, L. K.; Heard, D. E. *Chem. Soc. Rev.* **2012**, *41*(*19*), 6348–6404.

[159] Gligorovski, S.; Strekowski, R.; Barbati, S.; Vione, D. *Chem. Rev.* **2015**, *115*(*24*), 13051–13092.

[160] Levy, H. *Science* **1971**, *173*(*3992*), 141–143.

[161] Khan, M.; Percival, C.; Caravan, R.; Taatjes, C.; Shallcross, D. *Environ. Sci.: Process. Impacts* **2018**, *20*(*3*), 437–453.

[162] Hatakeyama, S.; Kobayashi, H.; Akimoto, H. *J. Phys. Chem.* **1984**, *88*(*20*), 4736–4739.

[163] Atkinson, R.; Aschmann, S. M. *Env. Sci. Tech.* **1993**, *27*(*7*), 1357–1363.

[164] Campos-Pineda, M.; Zhang, J. *Chem. Phys. Lett.* **2017**, *683*, 647–652.

[165] Käser, S.; Unke, O. T.; Meuwly, M. *J. Chem. Phys.* **2020**, *152*(*21*), 214304.

[166] Baer, T.; Hase, W. L. *Unimolecular reaction dynamics: theory and experiments.* Oxford University Press: **1996**.

[167] Green, A. M.; Barber, V. P.; Fang, Y.; Klippenstein, S. J.; Lester, M. I. *Proc. Natl. Acad. Sci. USA* **2017**, *114*(*47*), 12372–12377.

[168] Barber, V. P.; Pandit, S.; Esposito, V. J.; McCoy, A. B.; Lester, M. I. *J. Phys. Chem. A* **2019**, *123*(*13*), 2559–2569.

[169] Kuwata, K. T.; Luu, L.; Weberg, A. B.; Huang, K.; Parsons, A. J.; Peebles, L. A.; Rackstraw, N. B.; Kim, M. J. *J. Phys. Chem. A* **2018**, *122*(*9*), 2485–2502.

[170] Wang, X.; Bowman, J. M. *J. Phys. Chem. Lett.* **2016**, *7*(*17*), 3359–3364.

[171] Nguyen, T. L.; McCaslin, L.; McCarthy, M. C.; Stanton, J. F. *J. Chem. Phys.* **2016**, *145*, 131102.

[172] Zhou, X.; Liu, Y.; Dong, W.; Yang, X. *J. Phys. Chem. Lett.* **2019**, *10*(*17*), 4817–4821.

[173] Huang, B.; Lilienfeld, O. A. *Nat. Chem.* **2020**, *12*(*10*), 945–951.

[174] Behler, J. *J. Phys. Condens. Matter* **2014**, *26*(*18*), 183001.

[175] Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christense n, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; s. *J. Phys. Condens. Matter* **2017**, *29*(*27*), 273002.

[176] Yosa Reyes, J.; Brickel, S.; Unke, Oliver T; Meuwly, M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 6780–6788.

[177] Yosa Reyes, J.; Nagy, T.; Meuwly, M. *Phys. Chem. Chem. Phys.* **2014**, *16*, 18533–18544.

[178] Brickel, S.; Meuwly, M. *J. Phys. Chem. A* **2017**, *121*, 5079–5087.

[179] Karplus, M.; Porter, R. N.; Sharma, R. D. *J. Chem. Phys.* **1965**, *43*(*9*), 3259–3287.

[180] Koner, D.; Bemish, R. J.; Meuwly, M. *J. Chem. Phys.* **2018**, *149*(*9*), 094305.

[181] Koner, D.; Bemish, R. J.; Meuwly, M. *J. Phys. Chem. A* **2020**, *124*(*31*), 6255–6269.

[182] Qu, C.; Nandi, A.; Bowman, J. M. *Adv. Chem. Phys.* **2018**, *163*, 151–166.

[183] Unke, O. T.; Brickel, S.; Meuwly, M. *J. Chem. Phys.* **2019**, *150*(*7*), 074107.

[184] Tong, X.; Nagy, T.; Reyes, J. Y.; Germann, M.; Meuwly, M.; Willitsch, S. *Chem. Phys. Lett.* **2012**, *547*, 1–8.

[185] Sweeny, B. C.; Pan, H.; Kassem, A.; Sawyer, J. C.; Ard, S. G.; Shuman, N. S.; Viggiano, A. A.; Brickel, S.; Unke, O. T.; Upadhyay, M.; et al. *Phys. Chem. Chem. Phys.* **2020**, *22*(*16*), 8913–8923.

[186] Käser, S.; Unke, O. T.; Meuwly, M. *New J. Phys.* **2020**, *22*(*5*), 055002.

[187] Rivero, U.; Unke, O. T.; Meuwly, M.; Willitsch, S. *J. Chem. Phys.* **2019**, *151*(*10*), 104301.

[188] Boltachev, G. S.; Schmelzer, J. W. *J. Chem. Phys.* **2010**, *133*(*13*), 134509.

[189] Soloviov, M.; Das, A. K.; Meuwly, M. *Angew. Chem. Int. Ed.* **2016**, *55*(*34*), 10126–10130.

[190] Kroll, J. H.; Clarke, J. S.; Donahue, N. M.; Anderson, J. G. *J. Phys. Chem. A* **2001**, *105*(*9*), 1554–1560.

[191] Donahue, N. M.; Drozd, G. T.; Epstein, S. A.; Presto, A. A.; Kroll, J. H. *Phys. Chem. Chem. Phys.* **2011**, *13*(*23*), 10848–10857.

[192] Osborn, D. L.; Taatjes, C. A. *Intern. Rev. Phys. Chem.* **2015**, *34*(*3*), 309–360.

[193] Wang, Z.; Dyakov, Y. A.; Bu, Y. *J. Phys. Chem. A* **2019**, *123*(*5*), 1085–1090.

[194] Kalinowski, J.; Räsänen, M.; Heinonen, P.; Kilpeläinen, I.; Gerber, R. B. *Angew. Chem. Int. Ed.* **2014**, *126*(*1*), 269–272.

[195] Rickard, A. R.; Johnson, D.; McGill, C. D.; Marston, G. *J. Phys. Chem. A* **1999**, *103*(*38*), 7656–7664.

[196] McGill, C.; Rickard, A.; Johnson, D.; Marston, G. *Chemosphere* **1999**, *38*(*6*), 1205–1212.

[197] Aschmann, S. M.; Arey, J.; Atkinson, R. *Atm. Envir.* **1996**, *30*(*17*), 2939–2943.

[198] Atkinson, R.; Aschmann, S. M.; Arey, J.; Shorees, B. *J. Geophys. Res.: Atm.* **1992**, *97*(*D5*), 6065–6073.

[199] Orzechowska, G. E; Paulson, S. E. *Atmosph. Env.* **2002**, *36*(*3*), 571–581.

[200] Alam, M. S.; Rickard, A. R.; Camredon, M.; Wyche, K. P.; Carr, T.; Hornsby, K. E.; Monks, P. S.; Bloss, W. J. *J. Phys. Chem. A* **2013**, *117*(*47*), 12468–12483.

[201] Bowman, J. M.; Gazdy, B.; Sun, Q. *J. Chem. Phys.* **1989**, *91*(*5*), 2859–2862.

[202] Ben-Nun, M; Levine, R. *J. Chem. Phys.* **1994**, *101*(*10*), 8768–8783.

[203] Brieuc, F.; Bronstein, Y.; Dammak, H.; Depondt, P.; Finocchi, F.; Hayoun, M. *J. Chem. Theory Comput.* **2016**, *12*(*12*), 5688–5697.

[204] Karandashev, K.; Xu, Z.-H.; Meuwly, M.; Vaníček, J.; Richardson, J. O. *Struct. Dyn.* **2017**, *4(6)*, 061501.

[205] Drozd, G. T.; Kroll, J.; Donahue, N. M. *J. Phys. Chem. A* **2011**, *115(2)*, 161–166.

[206] Drozd, G. T.; Kurtén, T.; Donahue, N. M.; Lester, M. I. *J. Phys. Chem. A* **2017**, *121(32)*, 6036–6045.

[207] Rousso, A. C.; Hansen, N.; Jasper, A. W.; Ju, Y. *Phys. Chem. Chem. Phys.* **2019**, *21(14)*, 7341–7357.

[208] Welz, O.; Savee, J. D.; Osborn, D. L.; Vasu, S. S.; Percival, C. J.; Shallcross, D. E.; Taatjes, C. A. *Science* **2012**, *335(6065)*, 204–207.

[209] Meuwly, M.; Karplus, M. *J. Chem. Phys.* **2002**, *116(6)*, 2572–2585.

[210] Mooney, C. Z.; Duval, R. D.; Duvall, R. *Bootstrapping: A nonparametric approach to statistical inference.* Sage: **1993**.

[211] Olguín-Arias, V.; Davis, S.; Gutiérrez, G. *Physica A* **2021**, *561*, 125198.

[212] Juvenal, Y. R.; Meuwly, M. *J. Phys. Chem. A* **2011**, *115(50)*, 14350–14360.

[213] Mukherjee, S.; Pinheiro Jr, M.; Demoulin, B.; Barbatti, M. *Philos. Trans. Royal Soc. A* **2022**, *380(2223)*, 20200382.

[214] Austin, R.; Beeson, K.; Eisenstein, L.; Frauenfelder, H.; Gunsalus, I. *Biochemistry* **1975**, *14(24)*, 5355–5373.

[215] Foley, C. D.; Xie, C.; Guo, H.; Suits, A. G. *Science* **2021**, *374(6571)*, 1122–1127.

[216] Linstrom, PJ; Mallard, W. *National Institute of Standards and Technology Standard Reference Database* **2001**.

[217] Bowman, J. M. *Mol. Phys.* **2014**, *112(19)*, 2516–2528.

[218] Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K.-R. *Chem. Rev.* **2021**, *121(16)*, 10142–10186.

[219] Käser, S.; Vazquez-Salazar, L. I.; Meuwly, M.; Töpfer, K. *Digital Discovery* **2023**, *2(1)*, 28–58.

[220] Käser, S.; Meuwly, M. *J. Chem. Phys.* **2023**, *158(21)*, 214301.

[221] Song, K.; Käser, S.; Töpfer, K.; Vazquez-Salazar, L. I.; Meuwly, M. *J. Chem. Phys.* **2023**, *159*, 024125.

[222] Désesquelles, P.; Van-Oanh, N.-T.; Thomas, S.; Domin, D. *Phys. Chem. Chem. Phys.* **2020**, *22(6)*, 3160–3172.

[223] Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. *Nat. Comm.* **2019**, *10(1)*, 2903.

[224] Taylor, M. E.; Stone, P. *J. Mach. Learn. Res.* **2009**, *10(7)*, 1633–1685.

[225] Buckner, J.; Liu, X.; Chakravorty, A.; Wu, Y.; Cervantes, L. F.; Lai, T. T.; Brooks III, C. L. *J. Chem. Theory Comput.* **2023**.

[226] Bauer, B.; Reynolds, M. *Behav. Res. Meth.* **2008**, *40(3)*, 858–868.

[227] Tielens, A. G. G. M. Hagen, W. *Astron. Astrophys.* **1982**, *114*, 245–260.

[228] Neininger, N.; Guélin, M.; Ungerechts, H.; Lucas, R.; Wielebinski, R. *Nature* **1998**, *395(6705)*, 871–873.

[229] Gillett, F. C.; Forrest, W. *Astrophys. J.* **1973**, *179*, 483–491.

[230] d'Hendecourt, L. B.; Jourdain de Muizon, M. *Astron. Astrophys.* **1989**, *223*, L5–L8.

[231] Whittet, D.; Gerakines, P.; Hough, J.; Shenoy, S. *Astrophys. J.* **2001**, *547(2)*, 872.

[232] Garrod, R. T.; Pauly, T. *Astrophys. J.* **2011**, *735(1)*, 15.

[233] Roser, J. E.; Vidali, G.; Manicò, G.; Pirronello, V. *Astrophys. J. Lett.* **2001**, *555(1)*, L61.

[234] Fournier, J.; Deson, J.; Vermeil, C.; Pimentel, G. *J. Chem. Phys.* **1979**, *70(12)*, 5726–5730.

[235] Veliz, J. C. S. V.; Koner, D.; Schwilk, M.; Bemish, R. J.; Meuwly, M. *Phys. Chem. Chem. Phys.* **2021**, *23(19)*, 11251–11263.

[236] Ioppolo, S.; Van Boheemen, Y.; Cuppen, H.; Van Dishoeck, E.; Linnartz, H. *Mon. Not. R. Astron. Soc.* **2011**, *413(3)*, 2281–2287.

[237] Qasim, D.; Lamberts, T.; He, J.; Chuang, K.-J.; Fedoseev, G.; Ioppolo, S.; Boogert, A.; Linnartz, H. *Astron. Astrophys.* **2019**, *626*, A118.

[238] Oba, Y.; Watanabe, N.; Kouchi, A.; Hama, T.; Pirronello, V. *Astrophys. J.* **2010**, *712(2)*, L174.

[239] Noble, J. A.; Dulieu, F.; Congiu, E.; Fraser, H. J. *Astrophys. J.* **2011**, *735(2)*, 121.

[240] Lee, M. W.; Meuwly, M. *Faraday Discuss.* **2014**, *168*, 205–222.

[241] Bergin, E. A.; Tafalla, M. *Annu. Rev. Astron. Astrophys.* **2007**, *45*, 339–396.

[242] Danilychev, A. V.; Apkarian, V. *J. Chem. Phys.* **1993**, *99(11)*, 8617–8627.

[243] Apkarian, V. A.; Schwentner, N. *Chem. Rev.* **1999**, *99(6)*, 1481–1514.

[244] Pezzella, M.; Koner, D.; Meuwly, M. *J. Phys. Chem. Lett.* **2020**, *11(6)*, 2171–2176.

[245] Stief, L. J.; Payne, W. A.; Klemm, R. B. *J. Chem. Phys.* **1975**, *62(10)*, 4000–4008.

[246] Garstang, R. *Mon. Not. R. Astron. Soc.* **1951**, *111(1)*, 115–124.

[247] Schmidt, F.; Swiderek, P.; Bredehöft, J. H. *ACS Earth Space Chem.* **2019**, *3(9)*, 1974–1986.

[248] Shimanouchi, T.; Matsuura, H.; Ogawa, Y.; Harada, I. *J. Phys. Chem. Ref. Data* **1978**, *7(4)*, 1323–1444.

[249] Dubrin, J.; MacKay, C.; Pandow, M.; Wolfgang, R. *J. Inorg. Nuc. Chem.* **1964**, *26(12)*, 2113–2122.

[250] Xantheas, S. S.; Klaus, R. *Int. J. Quant. Chem.* **1994**, *49(4)*, 409–427.

[251] Congiu, E.; Minissale, M.; Baouche, S.; Chaabouni, H.; Moudens, A.; Cazaux, S.; Manicò, G.; Pirronello, V.; Dulieu, F. *Faraday Discuss.* **2014**, *168*, 151–166.

[252] Minissale, M.; Dulieu, F.; Cazaux, S.; Hocuk, S. *Astron. Astrophys.* **2016**, *585*, A24.

[253] Yamamoto, T.; Miura, H.; Shalabiea, O. M. *Mon. Not. R. Astron. Soc.* **2019**, *490(1)*, 709–717.

[254] Rosenblatt, M. *Ann. Math. Statist.* **1956**, *27(3)*, 832–837.

[255] Noble, J.; Congiu, E.; Dulieu, F.; Fraser, H. *Mon. Not. R. Astron. Soc.* **2012**, *421(1)*, 768–779.

[256] Hama, T.; Yokoyama, M.; Yabushita, A.; Kawasaki, M. *J. Chem. Phys.* **2010**, *133(10)*, 104504.

[257] Minissale, M.; Loison, J.-C.; Baouche, S.; Chaabouni, H.; Congiu, E.; Dulieu, F. *Astron. Astrophys.* **2015**, *577*, A2.

[258] Plattner, N.; Meuwly, M. *ChemPhysChem* **2008**, *9(9)*, 1271–1277.

[259] Lee, M. W.; Meuwly, M. *J. Phys. Chem. A* **2011**, *115(20)*, 5053–5061.

[260] Davidson, J. A.; Schiff, H. I.; Brown, T. J.; Howard, C. J. *J. Chem. Phys.* **1978**, *69(3)*, 1216–1217.

[261] Ogryzlo, E.; Reilly, J.; Thrush, B. *Chem. Phys. Lett.* **1973**, *23(1)*, 37–39.

[262] Raut, U; Baragiola, R. *Astrophys. J. Lett.* **2011**, *737(1)*, L14.

[263] Braunstein, M.; Duff, J. *J. Chem. Phys.* **2000**, *112(6)*, 2736–2745.

[264] Unke, O. T.; Meuwly, M. *J. Chem. Phys.* **2018**, *148(24)*, 241708.

[265] Meuwly, M. *Chem. Rev.* **2021**, *121(16)*, 10218–10239.

[266] Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell Jr, A. D. *J. Chem. Theory Comput.* **2012**, *8(9)*, 3257–3273.

[267] Adler, T. B.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2007**, *127(22)*, 221106.

[268] Peterson, K. A.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2008**, *128(8)*, 084102.

[269] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

[270] Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. *J. Phys. Chem. B* **2013**, *117(34)*, 9956–9972.

[271] Devereux, M.; Pezzella, M.; Raghunathan, S.; Meuwly, M. *J. Chem. Theory Comput.* **2020**, *16(12)*, 7267–7280.

[272] Arasa, C.; Hemert, M. C.; Dishoeck, E. F.; Kroes, G.-J. *J. Phys. Chem. A* **2013**, *117(32)*, 7064–7074.

[273] Ghesquière, P.; Mineva, T.; Talbi, D.; Theulé, P.; Noble, J.; Chiavassa, T. *Phys. Chem. Chem. Phys.* **2015**, *17(17)*, 11455–11468.

[274] Minissale, M.; Congiu, E.; Dulieu, F. *Astron. Astrophys.* **2016**, *585*, A146.

[275] Karssemeijer, L. J.; Cuppen, H. *Astron. Astrophys.* **2014**, *569*, A107.

[276] Karssemeijer, L.; Ioppolo, S.; Hemert, M.; Avoird, A.; Allodi, M.; Blake, G.; Cuppen, H. *Astrophys. J.* **2013**, *781(1)*, 1–15.

[277] Hagen, W.; Tielens, A.; Greenberg, J. *Chem. Phys.* **1981**, *56(3)*, 367–379.

[278] Jenniskens, P.; Blake, D. *Science* **1994**, *265(5173)*, 753–756.

[279] Bossa, J. B.; Isokoski, K.; Valois, M. S.; Linnartz, H. *Astron. Astrophys.* **2012**, *545*, A82.

[280] He, J.; Acharyya, K.; Vidali, G. *Astrophys. J.* **2016**, *825(2)*, 89.

[281] Kouchi, A.; Furuya, K.; Hama, T.; Chigai, T.; Kozasa, T.; Watanabe, N. *Astrophys. J. Lett.* **2020**, *891(1)*, L22.

[282] Oba, Y.; Miyauchi, N.; Hidaka, H.; Chigai, T.; Watanabe, N.; Kouchi, A. *Astrophys. J.* **2009**, *701*, 464–470.

[283] Keane, J.; Tielens, A.; Boogert, A.; Schutte, W.; Whittet, D. *Astron. Astrophys.* **2001**, *376(1)*, 254–270.

[284] Kouchi, A.; Tsuge, M.; Hama, T.; Oba, Y.; Okuzumi, S.; Sirono, S.-i.; Momose, M.; Nakatani, N.; Furuya, K.; Shimonishi, T.; et al. *Astrophys. J.* **2021**, *918(2)*, 45.

[285] Bossa, J.-B.; Isokoski, K.; Paardekooper, D. M.; Bonnin, M.; van der Linden, E. P.; Triemstra, T.; Cazaux, S.; Tielens, A. G. G. M.; Linnartz, H. *Astron. Astrophys.* **2014**, *561*, A136.

[286] Bossa, J.-B.; Maté, B.; Fransen, C.; Cazaux, S.; Pilling, S.; Rocha, W. R. M.; Ortigoso, J.; Linnartz, H. *Astrophys. J.* **2015**, *814(1)*, 47.

[287] Cazaux, S.; Bossa, J.-B.; Linnartz, H.; Tielens, A. G. G. M. *Astron. Astrophys.* **2015**, *573*, A16.

[288] Minissale, M.; Nguyen, T.; Dulieu, F. *Astron. Astrophys.* **2019**, *622*, A148.

[289] Tsuge, M.; Hidaka, H.; Kouchi, A.; Watanabe, N. *Astrophys. J.* **2020**, *900(2)*, 187.

[290] Watanabe, N.; Kouchi, A. *Astrophys. J.* **2002**, *567(1)*, 651.

[291] He, J.; Emtiaz, S.; Vidali, G. *Astrophys. J.* **2018**, *863(2)*, 156.

[292] Burnham, C. J.; Li, J. C.; Leslie, M. *J. Phys. Chem. B* **1997**, *101*, 6192–6195.

[293] Kumagai, N.; Kawamura, K.; Yokokawa, T. *Mol. Sim.* **1994**, *12(3-6)*, 177–186.

[294] Fredon, A.; Radchenko, A. K.; Cuppen, H. M. *Acc. Chem. Res.* **2021**, *54(4)*, 745–753.

[295] Stock, G. *Phys. Rev. Lett.* **2009**, *102(11)*, 118301.

[296] Talbi, D.; Chandler, G.; Rohl, A. *Chem. Phys.* **2006**, *320(2-3)*, 214–228.

[297] Liszt, H. T. *Astrophys. J.* **1978**, *224*, L73–L76.

[298] Ligterink, N. F. W.; Calcutt, H.; Coutens, A.; Kristensen, L.; Bourke, T.; Drozdovskaya, M. N.; Müller, H.; Wampfler, S.; Der Wiel, M.; Van Dishoeck, E. *Astron. Astrophys.* **2018**, *619*, A28.

[299] Codella, C.; Viti, S.; Lefloch, B.; Holdship, J.; Bachiller, R.; Bianchi, E.; Ceccarelli, C.; Favre, C.; Jiménez-Serra, I.; Podio, L. *Mon. Not. R. Astron. Soc.* **2018**, *474(4)*, 5694–5703.

[300] Ziurys, L.; McGonagle, D.; Minh, Y.; Irvine, W. *Astrophys. J.* **1991**, *373*, 535–542.

[301] McGonagle, D.; Ziurys, L.; Irvine, W. M.; Minh, Y. *Astrophys. J.* **1990**, *359*, 121–124.

[302] Ziurys, L.; Apponi, A.; Hollis, J.; Snyder, L. *Astrophys. J.* **1994**, *436*, L181–L184.

[303] Snyder, L. E.; Kuan, Y.-J.; Ziurys, L.; Hollis, J. *Astrophys. J.* **1993**, *403*, L17–L20.

[304] Barros, A.; Da Silveira, E.; Fulvio, D.; Boduch, P.; Rothard, H. *Mon. Not. R. Astron. Soc.* **2016**, *465(3)*, 3281–3290.

[305] Congiu, E.; Fedoseev, G.; Ioppolo, S.; Dulieu, F.; Chaabouni, H.; Baouche, S.; Lemaire, J. L.; Laffon, C.; Parent, P.; Lamberts, T. *Astrophys. J. Lett.* **2012**, *750(1)*, L12.

[306] Chou, S.-L.; Lo, J.-I.; Peng, Y.-C.; Lu, H.-C.; Cheng, B.-M.; Ogilvie, J. F. *Phys. Chem. Chem. Phys.* **2018**, *20*, 7730–7738.

[307] Arakawa, E. T.; Nielsen, A. H. *J. Mol. Spectrosc.* **1958**, *2(1-6)*, 413–427.

[308] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.;

Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; even, T. V.; Throssell, K.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian16 Revision C.01 Gaussian Inc. Wallingford CT: **2016**.

[309] Glendening, E. D.; Landis, C. R.; Weinhold, F. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*(*1*), 1–42.

[310] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.

[311] Vanommeslaeghe, K.; MacKerell, A. D. *J. Chem. Inf. Model.* **2012**, *52*(*12*), 3144–3154.

[312] Gaigeot, M.-P.; Martinez, M.; Vuilleumier, R. *Mol. Phys.* **2007**, *105*(*19-22*), 2857–2878.

[313] Futrelle, R. P.; McGinty, D. *Chem. Phys. Lett.* **1971**, *12*(*2*), 285–287.

[314] Mizutani, Y. Mizuno, M. *J. Chem. Phys.* **2022**, *157*(*24*), journal.

[315] Truhlar, D. G.; Muckerman, J. T. *Atom-Molecule Collision Theory*. Springers US: **1979**.

[316] T. Nagy, J. Y. R. Meuwly, M. *J. Chem. Theo. Comp.* **2014**, *10*, 1366–1375.

[317] Hama, T.; Yokoyama, M.; Yabushita, A.; Kawasaki, M. *The Journal of chemical physics* **2010**, *133*(*10*), journal.

[318] Yu, C.-C.; Chiang, K.-Y.; Okuno, M.; Seki, T.; Ohto, T.; Yu, X.; Korepanov, V.; Hamaguchi, H.-o.; Bonn, M.; Hunger, J. *Nat. Commun.* **2020**, *11*(*1*), 5977.

[319] Devlin, J. P. *J. Mol. Struct.* **1990**, *224*, 33–43.

[320] Minissale, M.; Nguyen, T.; Dulieu, F. *Astron. Astrophys.* **2019**, *622*, A148.

[321] Galvez, O.; Ortega, I. K.; Maté, B.; Moreno, M. A.; Martín-Llorente, B.; Herrero, V. J.; Escribano, R.; Gutiérrez, P. J. *Astron. Astrophys.* **2007**, *472(2)*, 691–698.

[322] Wakelam, V.; Loison, J.-C.; Mereau, R.; Ruaud, M. *Mol. Astrophys.* **2017**, *6*, 22–35.

[323] Ioppolo, S.; Fedoseev, G.; Minissale, M.; Congiu, E.; Dulieu, F.; Linnartz, H. *Phys. Chem. Chem. Phys.* **2014**, *16(18)*, 8270–8282.

[324] Acharyya, K. *Pub. Astron. Soc. Austr.* **2022**, *39*, e009.

[325] Käser, S.; Vazquez-Salazar, L. I.; Meuwly, M.; Töpfer, K. *Digital Discovery* **2023**, *2(1)*, 28–58.

[326] Upadhyay, M.; Töpfer, K.; Meuwly, M. *J. Phys. Chem. Lett.* **2023**, *15(1)*, 90–96.

[327] Fredon, A.; Lamberts, T.; Cuppen, H. *Astrophys. J.* **2017**, *849(2)*, 125.

[328] Bowman, J. M.; Qu, C.; Conte, R.; Nandi, A.; Houston, P. L.; Yu, Q. *J. Chem. Phys.* **2022**, *156(24)*, 240901.

[329] Groß, C.; Dillon, T. J.; Crowley, J. *Phys. Chem. Chem. Phys.* **2014**, *16(22)*, 10990–10998.

[330] Westermayr, J.; Gastegger, M.; Marquetand, P. *J. Phys. Chem. Lett.* **2020**, *11(10)*, 3828–3834.

[331] Sit, M. K.; Das, S.; Samanta, K. *J. Phys. Chem. A* **2023**, *127(10)*, 2376–2387.

[332] Antwi, E.; Ratliff, J. M.; Ashfold, M. N.; Karsili, T. N. *J. Phys. Chem. A* **2022**, *126(36)*, 6236–6243.

[333] Boittier, E. D.; Devereux, M.; Meuwly, M. *J. Chem. Theory Comput.* **2022**, *18(12)*, 7544–7554.