



---

# Atomistic Simulations of Reactive Processes from Gas to Condensed Phase

---

INAUGURALDISSERTATION

zur  
Erlangung der Würde eines Doktors der Philosophie  
vorgelegt der  
Philosophisch-Naturwissenschaftlichen Fakultät  
der Universität Basel

von  
**Juvenal Yosa Reyes**  
aus Kolumbien

Basel, 2014

Genehmigt von der  
Philosophisch-Naturwissenschaftlichen Fakultät auf  
Antrag von:

**Prof. Dr. Markus Meuwly**

**Prof. Dr. Stefan Willitsch**

Basel, den 10. Dezember 2013

Prof. Dr. Jörg Schibler  
Dekan



# PREFACE

---

*“Somewhere, something incredible is waiting to be known.”*

- Carl Sagan -



*To Mariana and Jean Philip.*



## ACKNOWLEDGMENTS

---

I am very grateful to my supervisor Prof. Markus Meuwly for giving me the opportunity to study and work in one of the best interdisciplinary teams. This fruitful 4 years, gave me a new vision of the computational work in chemistry. I appreciated his support and advice in all the scientific fields. I am especially thankful to prof. Dr. Stefan Willitsch for reading and suggestions on the content of this thesis.

I am especially thankful to Tibor Nagy, Stefan Lutz, Tristan Bereau, Ana Gamiz and Prashant Gupta, for all the scientific support and for the very intensive and profitable scientific and "pseudoscientific" discussion. Sharing projects with Tibor sometimes is difficult, but at the end of the day you learn a lot from his experience, thanks for that Dr. Nagy.

Special thanks to Ana Gamiz and Prashant Gupta, Tristan Bereau, Tibor Nagy for reading parts of my thesis. Also I want to thank Myung Won Lee for his advices on bash scripting and FORTRAN programming. It was also a great pleasure to work with our group members: Juan Carlos Castro, Lixian Zhang, Christian Kramer, Manuela Utzinger, Tobias Schmidt, Jing Huang, Vijay Solomon, Franziska Hofmann, Nuria Plattner, Zhen-Hao, Akshaya Kumar, Livia Glanzmann, Pierre-Andre Cazade, Florent Hedin, Mike Devereux, Maksym Soloviov, Jaroslav Padevet, Jaroslaw Szymczak, Maurus Schmid and Ganna Berezovska.

Special thank goes to Carolina Clavijo, Carlos Manuel Estevez, Claudia Cifuentes, Gina Mendez, Orlando Acevedo and Monica Huertas, for the continuous motivation and support during all this years away from home. I would like to thanks to Prof. Mario Blanco for all the support, advice and specially for introducing me into the fascinating world of computational chemistry.

Also, and not least I would like to thank my parents Marina Reyes and Juvenal Yosa, my loving sister and brother, wherever they belong now, Sandra Yosa and Juan Manuel Yosa, my brother Juan Carlos Yosa, my wonderful kids Jean

Philip and Mariana, Carina Arenilla, Lorena Rubiano, Amparo Mahecha and Nelson Mahecha for all the help in the good and not so good moments, and also for being part of this journey.

Finally I would like to thank Prof. Leonardo Lareo who has unfortunately passed away, and whose enthusiasm for science has always been an inspiration to me.

---

## ABSTRACT

---

In this thesis, we use advance computational methods to explore molecular reaction dynamics in gas and condensed phase. The molecular reactions that we describe in here have been studied for the last 25 years, using computational methods, from stratospheric chemistry to biochemistry.

We will first describe the vibrationally induced dissociation of  $\text{H}_2\text{SO}_4$  to  $\text{H}_2\text{O}$  and  $\text{SO}_3$  in stratospheric conditions, which is one of the important reactive steps in the Earth's atmosphere. Although this reaction has great significance in atmospheric chemistry, the exact mechanism whereby the reaction happens is still unknown. It has been postulated that the photodissociation of  $\text{H}_2\text{SO}_4$  to form sulfur trioxide and water, can explain the anomalous enhancement of the polar stratospheric sulfate aerosol layer in the springtime. This photodissociation process was previously assumed to take place *via* absorption ultraviolet (UV) radiation, to produce a dissociative electronically excited state. However, the electronic absorption spectrum of  $\text{H}_2\text{SO}_4$  up to 140 nm could not be found experimentally. Alternative mechanisms, including vibrationally induced dissociation, were proposed. Using adiabatic reactive molecular dynamics (ARMD) simulations with validated force fields for the product and reactant channels, it is shown through explicit atomistic simulation that by exciting the  $\nu_9$  OH-stretching mode, photodissociation can occur on the picosecond time scale. With the potential energy surfaces used in the present work,  $\nu_9 = 4$  is sufficient for this process. Because ARMD simulations allow multiple and long-time simulations, both nonstatistical, impulsive H-transfer and statistical, intramolecular vibration energy redistribution (IVR) regimes of the decomposition were found.

Another exciting reaction that we studied, this time in the condense phase, is the [3, 3]-sigmatropic rearrangement of allyl vinyl ethers. This is an important reaction due to its special synthetic relevance, which allows the preparation of  $\gamma,\delta$ -unsaturated carbonyl compounds. The intramolecular cyclic character (six-membered ring) of the rearrangement is generally accepted. However, an

understanding of the precise nature and geometry of the transition state and the exact mechanism is still controversial. It is generally accepted, evidenced by ab initio calculations using implicit solvent, that the reaction proceed through chair-like intermediates. In this investigation, using high level ab initio calculations, and ab initio Molecular Dynamics, we found that the reaction proceeds *via* boat-like transition state. This finding allows to initiate new studies about the real characterization of the transition state and hence its reaction mechanism.

Finally, we will describe the initial mechanism whereby the activation of the diguanylate cyclase the PleD protein in *Caulobacter crescentus* is taking place. PleD is widely studied since it participate in the formation of the ubiquitous second messenger, cyclic-di-GMP, involved in bacterial biofilm formation and persistence. PleD protein has an import role during the cell cycle of "C. crescentus". The active form of diguanylate cyclase PleD localizes to the stalked pole of differentiating *C. crescentus* cells. PleD is responsible for turning off flagellum rotations and inhibiting motility before genome replication begins and also for regenerating motility after differentiation has completed. Here we describe the coupling between the Phe102 and Thr83, in the active form of the protein. This coupling is highly influenced by phosphorylation, and the mechanisms seem to be key to form active dimers.

---



# CONTENTS

---

<b>List of Figures</b>	<b>1</b>
<b>List of Tables</b>	<b>9</b>
<b>1 Background and Motivation</b>	<b>11</b>
1.1 Molecular Reaction Dynamics . . . . .	13
1.2 Enthalpies and Free Energies of Reaction . . . . .	14
1.3 Kinetics . . . . .	15
1.3.1 Arrhenius Equation and Activation Energy . . . . .	15
1.4 Collision Theory . . . . .	16
1.5 Saddle Points and Potential Energy Surface . . . . .	18
1.6 Transition State Theory . . . . .	21
1.7 Statistical Theories of Chemical Reactions . . . . .	23
1.7.1 Rice–Ramsperger–Kassel–Marcus RRKM Theory . . . . .	23
<b>2 Concepts and Methods</b>	<b>27</b>
2.1 Molecular Dynamics . . . . .	28
2.2 Force Fields . . . . .	34
2.2.1 Force Field Parameterization . . . . .	38
2.3 Nonadiabatic Dynamics . . . . .	44
2.3.1 Mean-Field (Ehrenfest) Method . . . . .	44
2.3.2 Surface Hopping . . . . .	46
2.4 Reactive Force Fields . . . . .	49
2.4.1 Adiabatic Reactive Molecular Dynamics . . . . .	50
2.4.2 Multi-Surface/ Adiabatic Reactive Molecular Dynamics	54
2.4.3 ReaxFF . . . . .	60
2.5 Rare Events Techniques . . . . .	65
2.5.1 Umbrella Sampling . . . . .	67
2.5.2 Metadynamics . . . . .	69
2.6 Quantum Mechanics/Molecular Mechanics . . . . .	73
<b>3 Reaction in Gas Phase</b>	<b>77</b>
3.1 Vibrationally induced dissociation of $\text{H}_2\text{SO}_4$ . . . . .	79
3.1.1 Multi-Surface ARMD . . . . .	98
3.1.2 abstract . . . . .	98
3.1.3 Introduction . . . . .	98
3.1.4 MS-ARMD and its Applicability . . . . .	101

---

3.1.5	Multi-Surface ARMD . . . . .	105
3.1.6	Photodissociation of H <sub>2</sub> SO <sub>4</sub> . . . . .	109
3.1.7	Discussion and Conclusions . . . . .	117
3.2	Competitive Reaction Pathways of H <sub>2</sub> SO <sub>4</sub> . . . . .	119
3.2.1	Introduction . . . . .	119
3.3	Computational Methods . . . . .	121
3.3.1	Force Field Parametrization . . . . .	121
3.3.2	Molecular Dynamics Simulations . . . . .	125
3.3.3	Final State Analysis . . . . .	126
3.4	Results . . . . .	127
3.4.1	Validation of the Force Field . . . . .	127
3.4.2	Reaction Dynamics . . . . .	128
3.4.3	Final State Analysis . . . . .	132
3.4.4	Kinetic analysis . . . . .	136
3.5	Conclusions . . . . .	142
<b>4</b>	<b>Reaction in Condensed Phase</b>	<b>145</b>
4.1	Claisen Rearrangement Mechanism . . . . .	147
4.1.1	Introduction . . . . .	147
4.1.2	Theoretical Methods . . . . .	152
4.1.3	Results . . . . .	154
4.2	PleD . . . . .	160
4.2.1	Introduction . . . . .	160
4.2.2	Theoretical Methods . . . . .	164
4.2.3	Results . . . . .	167
<b>5</b>	<b>General Conclusions and Outlook</b>	<b>179</b>
	<b>Bibliography</b>	<b>181</b>
<b>I</b>	<b>Appended Papers</b>	<b>199</b>

# LIST OF FIGURES

---

2.1	Molecular mechanics potential energy function . . . . .	36
2.2	Types of simplex moves . . . . .	42
2.3	Schematic representation of two potentials $V_1$ and $V_2$ showing bonding and non-bonding interactions between two atoms A and B along their displacement coordinate $r_{AB}$ . $\Delta$ corresponds to the energetical gap between $V_1$ and $V_2$ at their asymptotic limit. $H_{1\leftarrow 2}$ is the activation enthalpy for a transition from $V_2$ to $V_1$ defined as the energy difference of the potential minimum on $V_2$ and the crossing point of $V_1$ with $V_2$ . . . . .	51
2.4	Schematic representation of the five steps taking place during an ARMD simulation involving a transition between two potentials $V_1$ and $V_2$ along a fictitious trajectory. The dashed line describes the duration $m$ during which $V_2$ is slowly transferred into $V_1$ . . .	53
2.5	Simple model for estimating energy violation in ARMD simulations. The system (effective mass $m$ ) is approaching from the left on PES $V_1(x)$ (phase I). At $t = 0$ time it is at $x_0$ with velocity $v_0$ and kinetic energy $E_{\text{kin},0}$ . After crossing is detected at $x = 0$ the time is rewound by $t_s/2$ and the dynamics is re-simulated while $V_1(x)$ is being switched to $V_2(x)$ during $t_s$ (phase II). . .	56
2.6	The MS-ARMD switching method in one dimension ( $x$ ) for three surfaces ( $V_1$ : blue dotted, $V_2$ : black dashed, $V_3$ : magenta dot-dashed) defined by linear ( $V_1$ and $V_2$ ) and quadratic ( $V_3$ ) functions. The effective surface ( $V_{\text{MS-ARMD}}$ : red solid) always follows the lowest-energy surface ( $V_{\text{min}}$ : green solid), except for regions where several surfaces are close in energy (within a few times $\Delta V = 0.5$ a.u.). Here, the algorithm switches smoothly among them by varying their weights ( $w_1$ , $w_2$ , and $w_3$ ; see lower panel) as a function of $x$ . Within the grayish band above the minimum energy surface the mixing weights of surfaces drop by a factor of $e \approx 2.7$ . All quantities are in arbitrary units (a.u.). .	58

- 
- 2.7 The MS-ARMD switching method in two dimensions ( $x_1$  and  $x_2$ ) for three surfaces ( $V_1$ ,  $V_2$  and  $V_3$ ) defined by simple functions. The scales of axes are in arbitrary units (a.u.). The  $V_{\text{MS-ARMD}}$  surface is always the lowest-energy surface ( $V_{\text{min}}$ ), except for the region where more surfaces have similarly low energy (within a few times  $\Delta V = 0.2$  a.u.), here the algorithm switches smoothly among them as a function of the coordinates. . . . . 59
- 2.8 free energy surface in the space of pseudo-dihedral angle  $\phi$  and dihedral angle  $\psi$ . Energy is given in kcal/mol. . . . . 72
- 3.1 Algorithm of Adiabatic Reactive Molecular Dynamics (ARMD) simulation method is schematically shown in five steps for collinear atom transfer reaction (atom B is transferred) where the coordinates of donor (A) and acceptor (C) atoms are fixed. During the crossing the surfaces are switched in time and the Morse bond is replaced by van der Waals (vdW) interactions and vice versa. . . . . 102
- 3.2 Simple model for estimating energy violation in ARMD simulations. The system (effective mass  $m$ ) is approaching from the left on PES  $V_1(x)$  (phase I). At  $t = 0$  time it is at  $x_0$  with velocity  $v_0$  and kinetic energy  $E_{\text{kin},0}$ . After crossing is detected at  $x = 0$  the time is rewound by  $t_s/2$  and the dynamics is re-simulated while  $V_1(x)$  is being switched to  $V_2(x)$  during  $t_s$  (phase II). . . 103
- 3.3 The MS-ARMD switching method in one dimension ( $x$ ) for three surfaces ( $V_1$ : blue dotted,  $V_2$ : black dashed,  $V_3$ : magenta dot-dashed) defined by linear ( $V_1$  and  $V_2$ ) and quadratic ( $V_3$ ) functions. The effective surface ( $V_{\text{MS-ARMD}}$ : red solid) always follows the lowest-energy surface ( $V_{\text{min}}$ : green solid), except for regions where several surfaces are close in energy (within a few times  $\Delta V = 0.5$  a.u.). Here, the algorithm switches smoothly among them by varying their weights ( $w_1$ ,  $w_2$ , and  $w_3$ ; see lower panel) as a function of  $x$ . Within the grayish band above the minimum energy surface the mixing weights of surfaces drop by a factor of  $e \approx 2.7$ . All quantities are in arbitrary units (a.u.). . 106

- 3.4 The MS-ARMD switching method in two dimensions ( $x_1$  and  $x_2$ ) for three surfaces ( $V_1$ ,  $V_2$  and  $V_3$ ) defined by simple functions. The scales of axes are in arbitrary units (a.u.). The  $V_{\text{MS-ARMD}}$  surface is always the lowest-energy surface ( $V_{\text{min}}$ ), except for the region where more surfaces have similarly low energy (within a few times  $\Delta V = 0.2$  a.u.), here the algorithm switches smoothly among them as a function of the coordinates. . . . . 107
- 3.5 A) Reactions with the three lowest barriers for sulfuric acid considered in the present work: water elimination (pathway I), H-transfer between  $\text{sp}^3$  and  $\text{sp}^2$  oxygen atoms (pathway II), or inversion of the bond configuration around the sulfur center (pathway III). B) The inverted configuration of the oxygen atoms can be obtained via the low-energy pathway of elimination and subsequent addition (pathway -I) of water. Re-addition was not observed in our studies, due to the escape of water. . . . . 110
- 3.6 Potential energy along the IRC for water elimination (Figure A) and for intramolecular H-transfer (Figure B). Black circles are the *ab initio* reference data, the blue dotted curve corresponds to MS-ARMD surface ( $\Delta V = 5$  kcal/mol) without GAPO, and the red solid curve is the same with GAPOs. Structures of stationary states along the IRCs are also shown in ball-and-stick representation. . . . . 112
- 3.7 The network of states (vertices, circles) and possible reactions (lines) for  $\text{H}_2\text{SO}_4$  represented as a symmetric graph. States denoted by circles and labeled by two-digit index, see text. Allowed reactions at the investigated energies are intramolecular H-transfer and  $\text{H}_2\text{O}$ -elimination. The four states observed in a typical trajectory (see Figure 3.8) are highlighted and given as a ball-and-stick representation. Four thick-line circles correspond to states of  $\text{H}_2\text{O} + \text{SO}_3$ . . . . . 114
- 3.8 The total energy as a function of time for a simulation with several reactive steps. The significant fluctuations during the first 15 ps are caused by the initial preparation of the  $\nu_9 = 5$  state which contains a significant amount of energy in a single OH-stretch mode. The two-digit labels refer to the state of the system, see Figure 3.7, and the ball-and-stick models are corresponding representative structures. . . . . 116

- 
- 3.9 Distribution of total energy fluctuations based on simulations of the free dynamics of sulfuric acid after OH overtone excitation ( $\nu_9$ ) for 50 independent initial conditions sampled at 300 K. Standard deviation  $\sigma=0.0073$  kcal/mol was empirically estimated from the sample, and the corresponding Gaussian distribution is also plotted. Note, the expectation value ( $\mu$ ) of the fluctuation is 0 kcal/mol by construction. . . . . 117
- 3.10 Sulfuric acid with atom numbering used throughout the text. The librational motion of O4 and H5, related to the dihedral angle H5-O4-S1-O6, is indicated by light spheres. . . . . 123
- 3.11 Panel A shows the quality of the fitted PES relative to the *ab initio* reference energies. Points along the MEPs (red squares) and remaining reference points (black circles) are distinguished. Panel B shows the fully relaxed MS-ARMD PES for both, intramolecular H-transfer and water elimination, in the space defined by distances H5-O6 and H7-O6. The white lines indicate approximate MEPs for the two possible reactions, water elimination (left-hand side) and H-transfer (right-hand side). Selected isocontours (in kcal/mol) are labelled. . . . . 125
- 3.12 Calculated PES along dihedral angles H5-O4-S1-O6 ( $\phi$ ) and H7-O6-S1-O4 ( $\phi'$ ) at (A) the MP2/6-311G++(2d,2p) level (rigid scan) and (b) with the FF (rigid scan). The two minima are  $C_2$  and  $C_s$  which are connected through transition states  $C_{1a}$  (4.9 kcal/mol) and  $C_{1b}$  (5.4 kcal/mol). . . . . 129
- 3.13 Percentage of water elimination after excitation of  $\nu_9 = 5$  (A) and  $\nu_9 = 6$  (B). Black, red and green curves correspond to overall, direct water elimination, and water elimination preceded by intramolecular H-transfer, respectively. Panel C reports histograms for the distribution of reaction times for direct water elimination (black) and for water elimination preceded by intramolecular H-transfer (blue) for  $\nu_9 = 5$  (top) and  $\nu_9 = 6$  (bottom) on a semi-logarithmic scale. . . . . 130

- 3.14 Time series of selected internal coordinates for  $\text{H}_2\text{SO}_4$ , including the H5-O4-S1-O6 dihedral angle (top), the H5-O4-S1 angle (middle), and the H5-O4 (red curve) and S1-O6 (black curve) bonds (bottom). Left column reports the data for a rapid water elimination reaction (breakup after 1.26 ps) whereas the right column corresponds to the last 3 ps of a slow reaction (breakup after  $\approx 111$  ps), for two representative trajectories with excitation of  $v_9 = 6$ . Note the characteristic torsional dynamics in the left panel. . . . . 131
- 3.15 Energy distribution for the  $\text{H}_2\text{O}$  and  $\text{SO}_3$  fragments. Contribution of translational, rotational, vibrational and rovibrational energies are shown for the two processes; direct water elimination (black histograms) and water elimination preceded by intramolecular H-transfer (red histograms) for  $v_9 = 5$  (left) and 6 (right). Overall distributions are represented by the green curve. 133
- 3.16 Generation of the angular momenta in the final state. The blue arrows on  $\text{H}_2\text{SO}_4$  illustrate the H-transfer step which leads to water elimination. Due to the off-center breakup and the conservation of total angular momentum the two fragments,  $\text{SO}_3$  and  $\text{H}_2\text{O}$ , separate in a counterrotating manner with high fragment angular velocities, indicated by the curved arrows. . . 134
- 3.17 Distribution of the  $\text{H}_2\text{O}$  and  $\text{SO}_3$  rotational quantum numbers  $j$  for both, direct (black) and indirect (preceded intramolecular H-transfer) (red) water elimination for excitation of  $v_9 = 5$  and 6. The overall distribution is the green curve . . . . . 135
- 3.18 Product yield as a function of time for the decomposition of sulfuric acid in free dynamics simulations after exciting the OH stretching vibrations with  $v_9 = 4, 5$  and 6 quanta in a canonical ensemble of molecules at 300 K. A simple kinetic model (see text) was fitted to the data, which allows determination of characteristic times (see Table 3.1) for first order relaxation and decomposition processes ( $\tau_{\text{eq}}$ ) at the corresponding energies. Left panel reports data up to 1 ns with both axes on a logarithmic scale. The right panel (with yield on a linear scale) provides the data up to 40 ns which has only be determined for  $v_9 = 6$ . It is found that beyond  $\approx 4$  ns additional IVR time scales appear. . 138

---

3.19	After exciting OH-stretching vibrational mode with 4, 5 and 6 quanta in sulfuric acid, photolysis quantum yields ( $\Phi_i$ ) and fractional contributions ( $J_i/J_{\text{total}}$ ) to the total photolysis rate ( $J_{\text{total}}$ ) are shown as a function of the altitude in the stratosphere and lower mesosphere. . . . .	142
4.1	Two possible Transition State for Allyl vinyl ether and their derivatives . . . . .	148
4.2	More O’Ferrall-Jencks diagram for the aliphatic Claisen Rearrangement. Red curve, represents the synchronous concerted pathway reaction <i>via</i> an aromatic-like TS. Green dashed curve, represents the stepwise pathway reaction <i>via</i> bis-allyl-like TS. Gray dashed curve, represents the stepwise pathway reaction <i>via</i> a 1-4-diyl like TS. Black curve, represent the alternative TS. . .	149
4.3	Molecular system for Allyl vinyl ether (CPMD/Metadynamics simulation) . . . . .	153
4.4	Transition state for A) gas phase and B) water implicit solvent, shows a boat TS conformation. Calculations were done using QST-2 method at MP2/6-311G++(2d,2p) level. . . . .	154
4.5	Claisen-rearrangement intrinsic Reaction Path for Allyl vinyl ether, in gas phase and implicit solvent. . . . .	155
4.6	Schematic representation of the molecular systems used to investigate the effect of the substituents in the Claisen rearrangement reaction. Strong donor ( $\text{NH}_2$ ), acceptor (CN), and an alkyl group ( $\text{CH}_3$ ) at positions 2 and 6. . . . .	156
4.7	Claisen-Rearrangement free energy surface on the space on C9-C7 and O5-C12 distance. CPMD-Metadynamics simulations. .	157
4.8	Boat transition state from CPMD/Metadynamics simulations. .	158
4.9	Optimized geometries for reactant, product and transition state for parent and substituted allyl vinyl ethers at the MP2/6-311G++(2d,2p) level. . . . .	159
4.10	Interface <b>Act</b> and <b>Inact</b> in PleD protein. D1 is represented in green color while D2 is represented in red color. Each monomer has a phosphorylation site located in the Asp53 residue. . . .	163
4.11	Representation models for active and inactive PleD. D1 is represented in green color while D2 is represented in red color . . .	168
4.12	Dominant intermolecular Interaction for <b>Act</b> and <b>Inact</b> . . . .	172



- 
- 4.13 Dominant intermolecular Interaction, **Act** and **Inact**. A. Common salt bridge interactions at interdomain interface. B. Key salt bridge between Arg4 and Glu221 at the interdomain interface. C. Schematic representation of Phe102-Thr83 coupling. Transparent ball and stick representation corresponds to the active protein, while normal ball and stick represents the inactive protein. 173
- 4.14 Molecular dynamic simulation for active, A and inactive B. Black (D1) and green (D2) curves represent the time series for the pseudo-dihedral angle  $\phi$  in the monomer 1 and red and blue for monomer 2. For the rest of the coordinates, black corresponds to monomer 1 and red for monomer 2. . . . . 174
- 4.15 Molecular dynamic simulation for monomer ( time series) 5 different trajectories. . . . . 175
- 4.16 Biased Molecular dynamic simulation for active unphosphorylated dimer, Black (D1) and red (D2) curves represent the time series for every coordinate. . . . . 175
- 4.17 Distributions of the  $\chi$  angle for the Biased Molecular Dynamic simulation. Black distribution corresponds to monomer D1 and yellow to monomer D2. . . . . 176
- 4.18 free energy surface in the space of pseudo-dihedral angle  $\phi$  consisting of  $C\alpha$ -atoms of the  $\beta 4$ - $\alpha 4$ , residues Thr83-Ala84-Leu85-Asp86 D1 domain and dihedral angle  $\psi$  in Phe102, consisting of N,  $C\alpha$ ,  $C\beta$  and  $C\gamma$ . Figure shows the most likely mechanism in which, after removing the phosphate-moiety, the system goes from an active state to inactive one. Energy is given in kcal/mol. 177



# LIST OF TABLES

---

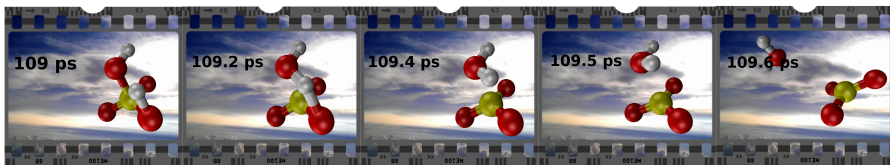
- 3.1 Number of IVR steps and corresponding relaxation ( $\tau_{\text{IVR},i}$ ) and decomposition ( $\tau_i$ ) times (in ps) for  $v_9 = 4, 5$ , and  $6$  inferred from the kinetic model by fitting to the trajectory data. For  $v_9 = 6'$ , no H-transfer was allowed during the simulation (see text) and only the IVR rate parameters were tuned, while other parameters were those from  $v_9 = 6$ . Relative fitting errors in the time scales  $\tau$  are around 5 % at most. The average total energy and its standard deviation are reported for all ensembles. 139
- 4.1 Activation Energies for the Claisen Rearrangements of Substituted Allyl Vinyl Ethers at the B3LYP/6-31G\*<sup>1</sup> and MP2/6-311G++(2d,2p) level. The energies are given in kcal/mol . . . 156



# 1

## BACKGROUND AND MOTIVATION

---



*“The fact that we live at the bottom of a deep gravity well, on the surface of a gas covered planet going around a nuclear fireball 90 million miles away and think this to be normal is obviously some indication of how skewed our perspective tends to be.”*

Douglas Adams, *The Salmon of Doubt: Hitchhiking the Galaxy One Last Time*

How memory works? How the information is stored in the brain? Why this information, in some cases, is lost? What are the chemical and physical processes that are behind some memory disorder? How creativity and imagination works? These are the questions that I was interested to know 10 years ago, unfortunately I did not find a concrete answer, but, instead many theories and hypothesis which associate the learning process to one particular protein, the N-methyl-D-aspartate NMDA. NMDA is a ion channel protein complex containing 3 to 5 subunits,<sup>2</sup> and in most configurations includes at least one copy of NR1 subunit, in addition to NR2 (A-D)<sup>3</sup> or NR3 (A-B).<sup>4</sup> Subunit NR1 appears to be the key for the formation of the channel. NR1 contains the glycine or glycine<sub>B</sub>-binding site and NR2 contains the glutamate-binding site. This molecular complex depends on glycine to be activated. Once it is activated, glutamate is attached to the NR2 subunit and the channel is opened to allow Ca<sup>2+</sup> cations flux into the cell. What happens after is a recruit of multiple types of synaptic proteins to new synapses with distinct mechanisms and time courses. How strong the synapses are, depends mainly on the stimulus in that particular channel of information. For example, if you want to remember a phone number, you should use it frequently, the more often you repeat the number, the stronger a particular information will be stored. But, even though there are more than 200000 publications, hundreds of research groups and millions of dollars spent, we do not have a clue about the exact mechanism, whereby memory and all the processes behind it, like creativity and imagination even consciousness works. Furthermore, how can we, at molecular level, manipulate it to treat neurodegenerative or neuropsychiatry disorders and increase memory processes.

We started to study one of the subunits of NMDA, the NR1, in order to find a pattern in which agonist and antagonist can be distinguish by their physicochemical or electronic properties.<sup>5</sup> If one wants to go into the atomistic details, one has to investigate the molecular dynamic behavior of the protein. Even when it is a large complex, with the computation power nowadays, one can explore some interesting mechanism (i.e the initial steps in the NMDA activation), using molecular dynamics. Furthermore, one can apply the theories in molecular reaction dynamics, in order to quantify and describe the kinetics of such molecular processes. That's how I joined the group of Professor Meuwly. Although, my research in brain proteins is not finish yet, new challenges in molecular reaction dynamics were achieved.

The time I spent here, gave the opportunity to learn and investigate reactive dynamic processes from the most simple molecules in gas phase, such as N<sub>2</sub>,

$\text{H}_2\text{SO}_4$  and  $\text{ClHSO}_3$  to organic molecules in solution such as derivatives of the allyl vinyl ether and macromolecules such as the diguanylate cyclase PleD protein.

As you will have realized, this thesis is not about NMDA, synaptic processes or learning processes, but about molecular reactive processes from gas to condensed phase. I will review in this first chapter about some concepts of reaction dynamics. In chapter 2, methods that were used in this thesis, are described. Chapter 3 will include reactions in gas phase (sulfuric acid photodissociation), and in chapter 4 I will describe the Claisen rearrangement for allyl vinyl ether and the effect of the substituents and the PleD protein activation, both in condensed phase.

## 1.1 Molecular Reaction Dynamics

---

Reaction dynamics is the study of the molecular level mechanism of elementary chemical and physical processes. It seeks to understand what actually takes place at that level when a change, chemical or physical, occurs. As an example, when a perturbation, such as phosphorylation in a protein induce dimerization, which is the case of response regulator proteins,<sup>6</sup> where the activation takes place once a phosphoryl chemical group is transfer to the protein, and thereafter, is activated to perform the biochemical function (i.e., in *Escherichia coli*, this perturbation leads to change the direction of flagellar rotation from counterclockwise to clockwise ).<sup>7</sup> Molecular reaction dynamics is the study of elementary processes and the means to probing them, understanding them and controlling them.<sup>8</sup> Molecular reaction dynamics is also applied to reactions in solution and to reactions in surfaces, exploring the elementary steps in catalysis. Also is not limited to neutral chemical species but also includes cations and anions. Biochemical reactions provide some examples of processes where electrostatic effects have a central role. Therefore, molecular reaction dynamics gave us valuable information about the microscopic mechanism in which atoms and molecules are involved, from the most simple interaction (i.e., gas particles interactions), to more complex processes at the cellular level.

## 1.2 Enthalpies and Free Energies of Reaction

The rate constant of a reaction for a chemical species A and B can be described in terms of a phenomenological second-order rate law:

$$A + B \rightleftharpoons C_1 + \cdots + C_n$$

$$-\frac{d[A]}{dt} = k[A][B] - k' \prod_{i=1}^n [C_i] \quad (1.1)$$

where  $k$  and  $k'$  are the rate coefficients,  $[A]$  and  $[B]$  are the concentration of the reactants  $A$  and  $B$ , and  $[C]$  is the concentration of the products  $C_i$  respectively. The equilibrium constant,  $K$  for the process is given by the quotient of the forward and reverse rate constants,<sup>9</sup> and the reaction quotient is defined by:

$$Q_k = \frac{\prod_{i=1}^n [C_i]}{[A][B]} \quad (1.2)$$

Usually the rate constant is measured under conditions where the second term in Equation 1.1 is negligible. In this case,  $k$  gives the total rate constant for the product formation. One can relate the temperature-dependent equilibrium constant with the standard-state Gibbs free energy of reaction,  $\Delta G_T^o(T)$  at temperature  $T$  as:

$$K = Q_K^o(T) e^{-\Delta G_T^o/RT} \quad (1.3)$$

where  $R$  is the gas constant,  $Q_K^o$  is the reaction quotient at the standard state, and  $\Delta G_T^o(T) = \Delta H_T^o(T) - T\Delta S_T^o$ , where  $\Delta H_T^o(T)$  and  $\Delta S_T^o$  are the standard-state enthalpy and entropy of the reaction, respectively. In a general view, the free energy change upon reaction can be written as:

$$\Delta G = RT \ln \frac{Q_K}{K}. \quad (1.4)$$

If the free energy change is 0, the reaction is at equilibrium. If  $\Delta G_T^o$  or  $\Delta G$  is negative, the reaction produces work and is called exergonic, on the other



hand, if  $\Delta G_T^0$  or  $\Delta G$  is positive, the reaction consume work and is called endergonic. The enthalpy of reaction is negative for an exothermic reaction (heat releasing) and positive for an endothermic reaction (heat absorption). Enthalpy of reaction can be obtained at a given temperature from the reactant and products enthalpies of formation. For an electronically adiabatic reaction, the enthalpy of reaction at 0 K can be calculated using electronic structure (ES) calculations, as the change in Born-Oppenheimer electronic energy (which includes nuclear repulsion) plus the change in zero point vibrational energy.<sup>10</sup> A reaction with a negative potential energy of reaction is called exoergic, and one with a positive energy of reaction is called endoergic.

## 1.3 Kinetics

### 1.3.1 Arrhenius Equation and Activation Energy

From a phenomenological point of view, one can see that the variation of the rate constant with temperature can be described by the Arrhenius equation<sup>11</sup>

$$k = Ae^{-E_a/RT} \quad (1.5)$$

where  $A$  is the preexponential or frequency factor, which may have a weak dependence on temperature and  $E_a$  is the activation energy. The activation energy can be approximately interpreted as the minimum energy (kinetic plus potential, relative to the lowest state reactants) that reactants must have to form the products, and the preexponential factor is a measure of the rate (collision frequency), at which collisions occur. Tolman,<sup>12,13</sup> has a more precise interpretation, in which the Arrhenius energy of activation is the average total energy (relative translational) plus internal of all reacting pairs of reactants minus the average total energy of all pairs of reactants, including nonreactive pairs.<sup>12,14</sup>

## 1.4 Collision Theory

For bimolecular reactions, simple collision theory gives us an approximation related with the temperature dependence and magnitude of rate constants. First of all, let us describe the different outcomes for collision of atom or molecule  $A$  in internal state  $i$  with molecule  $B$  in internal state  $j$ . Three possible outcomes can be: i) Elastic collision; in which neither the arrangement (composition and bonding pattern), nor internal state of the molecules, nor the relative translational energy changes; here the direction of the relative motion is the only significant change. ii) Inelastic collision; the two molecules retain their arrangement but the internal states  $i$  and  $j$  change. and iii) Reactive collision; in which the two molecules react to form one or more new molecules (i.e.  $C$  in internal state  $m$  and  $D$  in internal state  $n$ ). In the last case, where a number of  $A(i)$  are incident in a beam with relative velocity  $V_R$  upon scattering zone containing  $B(j)$ , one can define the state-selected rate constant  $k_{ij}$  and reaction cross section  $\sigma_{ij}$  as:

$$k_{ij}(V_R) = V_R \sigma_{ij}(V_R) \quad (1.6)$$

The average reaction cross section  $\sigma_r$  is obtained by averaging over all the reactants internal states:

$$\sigma_r = \sum_{i,j} w_j^A w_j^B \sigma_{ij}(V_R) \quad (1.7)$$

where  $w_j^A$  and  $w_j^B$  represent the Boltzmann weighting factors of the  $i$  and  $j$  reactants internal states, respectively.

The thermal rate constant for the process is given by averaging  $V_r \sigma_r$  over an equilibrium Maxwell-Boltzmann distribution of  $V_R$ ,<sup>15,16</sup> this can be written as:

$$k = \beta \left( \frac{8\beta}{\pi\mu} \right)^{1/2} \int_0^\infty dE_{rel} E_{rel} \sigma_r(E_{rel}) e^{-\beta E_{rel}} \quad (1.8)$$

where the relative translational energy  $E_{rel}$  can be written as:

$$E_{rel} = \mu V_R^2/2 \quad (1.9)$$

where  $\mu$  is the reduced mass of relative translational motion.

For state-selected thermal rate constants one can consider separately each of the internal states:

$$k_{ij} = \beta \left( \frac{8\beta}{\pi\mu} \right)^{1/2} \int_0^\infty E_{rel} \sigma_{ij}(E_{rel}) e^{-\beta E_{rel}} dE_{rel} \quad (1.10)$$

One can also define the reaction probability  $P_R$  as a function of the impact parameter  $b$ , which is defined as the distance of closest approach between the two molecules in absence of interparticle forces. The probability of reaction decreases to zero for large values of  $b$ . Actually, one can consider a value of  $b = b_{max}$  after which the reaction probability is negligible, and the reaction cross section will be given by:

$$\sigma_r = 2\pi \int_0^{b_{max}} P_R(b) b db \quad (1.11)$$

The simplest model is to consider the reactants like hard spheres that do not interact with each other if the intermolecular distance is larger than the arithmetic average  $d$  of their diameters, therefore  $P_R(b > d) = 0$ , but that react at all shorter distances ( $P_R(b \leq d) = 1$ ).<sup>14</sup> In this case the reaction cross section is  $\pi d^2$ , applying Equation 1.8 the reaction rate is:<sup>14</sup>

$$k(T) = \left( \frac{8}{\pi\mu\beta} \right)^{1/2} \pi d^2. \quad (1.12)$$

The thermal average value of the relative speed is:

$$\bar{V}_R = \left( \frac{8}{\pi\mu\beta} \right)^{1/2} \quad (1.13)$$

one can write Equation 1.12 as:

$$k = \bar{V}_R \pi d^2 \quad (1.14)$$

$k$  is then, usually the thermal average  $\bar{V}_R \sigma_r$  of  $V_R \sigma_r$ , but if  $\sigma_r$  is independent of the relative speed, then  $k$  becomes  $\bar{V}_R \sigma_r$ . Equation 1.12 does not account for the observed experimental behavior described by the Arrhenius equation, since it predicts a temperature dependence of  $T^{1/2}$  for the rate constant.<sup>8</sup> An improvement of this model is the reactive hard spheres model in which it assumes that the reaction occurs if  $\mu V_{LOC}^2/2$  exceeds a threshold energy  $E^0$ , where  $V_{LOC}$  is the relative velocity along the line of centers, i.e., in the direction connecting the centers of the two spheres. This velocity depends on the impact parameter at which the reaction is assumed to occur if:<sup>14</sup>

$$E^0 \leq E_{rel}(1 - b^2/d^2). \quad (1.15)$$

Then the reaction cross section can be written as:

$$\sigma_r = \pi b_{max}^2 = \pi d^2(1 - E^0/E_{rel}) \quad (1.16)$$

and the rate constant is:

$$k(T) = \pi d^2 \left( \frac{8}{\pi \mu \beta} \right)^{1/2} e^{-\beta E^0} \quad (1.17)$$

which is similar to the Arrhenius expression and predicts a variation with temperature of  $T^{1/2}$  for the preexponential factor.<sup>14</sup>

## 1.5 Saddle Points and Potential Energy Surface

The potential energy surface (PES) is related to the Born-Oppenheimer approximation, in quantum mechanics. In this approximation the total molecular wavefunction will be the product of an electronic wavefunction and a nuclear wavefunction. If one consider a system with  $M$  nuclei and  $N$  electrons, including electrostatic interaction, one can write the Hamiltonian of the system as:

$$H = - \sum_{\alpha=1}^M \frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 - \sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_i^2 + V(\mathbf{r}, \mathbf{R}), \quad (1.18)$$

where  $\mathbf{r}$  and  $\mathbf{R}$  are the electronic and nuclear coordinates, respectively. All electrostatic interactions (i.e. electron-electron, electron-nuclei and nuclei-nuclei interactions) are included in  $V(\mathbf{r}, \mathbf{R})$ . The mass of the nucleus  $\alpha$  is denoted by  $M_\alpha$  and  $m_e$  is the mass of the electron. One can rewrite the time-dependent Schrödinger equation expression in the time independent way as:

$$H\Psi_E(\mathbf{r}, \sigma, \mathbf{R}) = E\Psi_E(\mathbf{r}, \sigma, \mathbf{R}), \quad (1.19)$$

where  $\sigma$  corresponds to the electronic spin coordinates. In the Born-Oppenheimer approximation the wavefunction  $\psi_E$  is written as a product function:

$$\Psi_E \approx \Psi^{B.O.} = \psi_e(\mathbf{r}, \sigma, \mathbf{R})\phi(\mathbf{R}), \quad (1.20)$$

where the electronic wavefunction  $\psi_e$  is a solution of the electronic Schrödinger equation, then one can rewrite Equation 1.19 as:

$$\left\{ -\sum_{\alpha=1}^M \frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 + E_e(\mathbf{R}) \right\} \phi(\mathbf{R}) = E^{B.O.} \phi(\mathbf{R}) \quad (1.21)$$

where  $E^{B.O.}$  denotes the energy for the system within the Born-Oppenheimer approximation. The validity of this separation between electronic and nuclear motion is because the large ratio between electronic and nuclear masses. Equation 1.21 shows that the nuclei moves in an effective potential which is the electronic energy (also including nuclei-nuclei interaction) as a function of the internuclear distances in which  $E_e(\mathbf{R})$  is constant with respect to the translation and rotation of a fixed nuclear configuration.

One of the important things in classical molecular dynamics is to replace a quantum mechanical description of the nuclear motion by a classical one (see chapter 2), which makes the calculations in every step faster, and allowing one to compute large systems in a very fast computing time, compared with ES calculations.

### Saddle Points localizing transition state

Transition state structures can be computed looking in the first-order saddle points on the PES. That specific point on the PES indicates the present of a minimum in all dimension but only one.<sup>17</sup> One can compute this saddle point using ES calculations, packages like Gaussian,<sup>18</sup> GAMESS<sup>19</sup> and Jaguar,<sup>20</sup>

provide different methods to perform transition state search. However, locating them is often difficult and there is no method guaranteed to find the right transition state. Many methods of locating transition states also aim to find the minimum energy pathway (MEP) along the PES. Each method has its advantages and disadvantages depending on the particular reaction under investigation. Methods like linear and quadratic synchronous transit (LST, QST), search for the TS generating an estimate of the transition state by finding the highest point along the shortest line connecting two minima (LTS), and subsequently searching for a minimum along a line perpendicular to the previous one (QST). The path connecting minima and the found point may be searched for a saddle point.<sup>21</sup> Methods which guess the MEP that connects the two stationary points, using a discrete number of structures (images) which are placed along the guessed MEP, these images are moved according to the force acting on them, which are perpendicular to the path, the highest energy image gives a good estimate of the transition state. Examples of some algorithms that use this method are: Nudged elastic band method<sup>22</sup> and Climbing image nudged elastic band method.<sup>23</sup> Another similar method for searching saddle points is the String method.<sup>24</sup> This approximation also involves a series of the images along a guess of the MEP, but in this case the images are moved in two steps. Initially the images are moved according to the force acting on them perpendicular to the path. Using an interpolated path, the images are moved short distances along the MEP to make sure they are evenly spaced. Variations on the string method include the growing string method,<sup>25</sup> in which the guess of the pathway is generated as the program progresses. If there is no knowledge of the final structure one can use the dimer method,<sup>26</sup> in which a "dimer" is formed by two images very close to each other on the PES. The method works by moving the dimer uphill from the starting position whilst rotating the dimer to find the direction of lowest curvature. Finally a method called activation relaxation technique, in which following the direction of lowest negative curvature on the PES can reach the saddle point, relaxing in the perpendicular hyperplane between each "jump" (activation) in this direction.<sup>27, 28</sup>

## 1.6 Transition State Theory

For elemental chemical reactions, the transition state theory TST, can explain the reaction velocities if the reactants and activated transition state complexes are found in some grade of chemical equilibrium (quasi-equilibrium).<sup>29</sup> If one wants to understand, at least qualitatively the chemical reaction mechanism, you can use this theory, even when it was initially created for calculating the absolute reaction rate constants, which does not work well, mainly because a very well knowledge of the PES is required.<sup>30</sup> However, is useful in order to compute the standard enthalpy, entropy and Gibbs energy of activation, of course, if experimental rate constants from experiments are provided. The basic idea of the TST is that, initially, rates of reactions can be studied by examining activated complexes which are found close to the saddle point (TS), of a PES. Then it is important to consider that the activated complexes are in a "quasi-equilibrium" with the reactants molecules. And finally, the activated complexes can, eventually, convert into products. This allows one to compute the rate of the above conversion using kinetic theory. The rate constant for a reaction in gas phase, in solution or for an enzymatic reaction can be written as:

$$k(T) = \gamma(T)(k_B T/h)(C^0)^{1-n} e^{-\Delta G^{TS,0}(T)/RT} \quad (1.22)$$

where  $C^0$  is the standard state concentration,  $n$  is the order of reaction,  $R$  is the gas constant,  $T$  is the temperature, and  $\Delta G^{TS,0}$  is the standard-state quasithermodynamic free energy of activation.<sup>31</sup> The factor  $(k_B T/h)$  is a frequency factor for crossing the transition state, which for 300 K is about  $6 \text{ ps}^{-1}$ , this factor can be applied for reaction in solution or in gas phase.<sup>31</sup> The generalized transmission coefficient,  $\gamma(T)$ , relates the actual rate for the reaction to that obtained from simple transition state theory, which has  $\gamma(T) = 1$ . The choice of the reaction coordinate and the dividing surface are somewhat arbitrary and that the correct result for the rate constant  $k(T)$  is obtained if  $\Delta G^{TS,0}$  and  $\gamma(T)$  are calculated in a consistent manner.<sup>31,32</sup> When the number of degrees of freedom increases, and the molecular system is rather complex, many reaction paths can be found (i.e., proteic enzymes and ribozyme). Every reaction path has one or more saddle points, in that specific case, the reaction dynamics is expected to be more complex that those found in small-molecule

gas-phase reactions. For single molecules in gas phase, the potential surface often has a single valley leading to a single saddle point, and we sum over paths through that valley. In complex systems like enzymes we have to average over many reactant conformations and sum the rate over paths that proceed through many valleys, which can differ in conformational degrees of freedom, hydrogen bonding patterns, and so forth. However, this effect is automatically included in the transition state formalism because all reactive paths must pass through the dividing surface separating reactants from products.<sup>31</sup> The generalized transmission coefficient,  $\gamma(T)$ , can be expressed<sup>30</sup> as a product

$$\gamma(T) = \Gamma(T)k(T)g(T) \quad (1.23)$$

where  $\Gamma(T)$  emerges from dynamical recrossing. Thus,  $\Gamma(T) \leq 1$  because for any choice of transition state, some trajectories that cross in the direction of products originate as products or recross the dividing surface to return to the reactant region.<sup>31,33</sup> The second,  $k(T)$ , arises from the contribution of quantum mechanical tunneling; therefore, almost always,  $k(T) \geq 1$ . This correction is necessary because  $\Delta G^{TS,0}$  includes only one of the major quantum effects, namely, quantization of the bound vibrations; the increase in the rate due to the other major quantum effect, namely tunneling, is given by  $k(T)$ .<sup>30,31,34</sup>

The third contributing factor,  $g(T)$ , arises from deviations of the equilibrium distribution in phase space;  $g(T)$  can be either less than or greater than 1; i.e., there can be either an enhancement or a decrease of the reaction rate relative to that which would exist when all degrees of freedom have their equilibrium (Boltzmann) distribution.<sup>31,35</sup>

Transition state theory fails for some reactions at high temperature. The theory assumes the reaction system will pass over the lowest energy saddle point on the potential energy surface. While this description is consistent for reactions occurring at relatively low temperatures, at high temperatures, molecules populate higher energy vibrational modes; their motion becomes more complex and collisions may lead to transition states far away from the lowest energy saddle point.<sup>31</sup> Other generalized TST like microcanonical variational TST, canonical variational and improved canonical TST, can overcome the TST limitations and give in many cases good, results.<sup>13</sup>

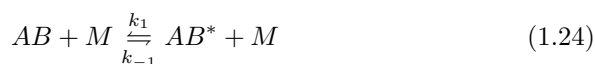


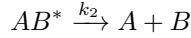
## 1.7 Statistical Theories of Chemical Reactions

For unimolecular reactions, the rate is determined from an approach that does not involve any explicit consideration of the reaction dynamics. The basic assumption in statistical theories is that, the initially prepared state, in an indirect unimolecular reaction  $A^*(E) \rightarrow$  products, prior to reaction has relaxed *via* intramolecular vibrational energy redistribution, frequently called IVR, such that any distribution of the energy  $E$  over the internal degrees of freedom occurs with the same probability. The above assumption is equivalent to the basic "equal a priori probabilities" postulated for statistical mechanics,<sup>10</sup> for microcanonical ensemble where every state within a narrow energy range is populated with the same probability. This uniform population of states describes the system regardless of where it is on the potential energy surface associated with the reaction.<sup>10</sup> An average decay rate  $k(E)$  is obtained, since one does not distinguish between various degenerate states.

### 1.7.1 Rice–Ramsperger–Kassel–Marcus RRKM Theory

Rice-Ramsperger-Kassel-Marcus theory or just RRKM theory,<sup>36</sup> is a chemical kinetics theory allowing estimation of unimolecular reaction rates from a few parameters of the potential energy surface. RRKM theory is an improved form of the Rice-Ramsperger-Kassel (RRK) theory<sup>37</sup> in which the rate of excited reactant molecule breaks down, is treated as a function of the energy it contains. It was mentioned in the TST section, that a TS is defined as the "dividing surface" between reactants and products, one assumption in the RRKM theory is that every trajectory, which passes this surface, will form the products without experiencing any recrossing. The second assumption, and the principal one, is the ergodicity assumption in which a fast and a complete randomization of the available freely distributable energy among all active mode after excitation. For the reaction:





the rate constant  $k_2(E)$ , will depend on the transition frequency ( $\nu^\ddagger$ ), assuming that the trajectory passes through the TS,  $\nu^\ddagger = kT/h$ , and the probability, which is the statistical weight ratio, to for the transition state geometry from all  $AB^*$ . Due that, in general, an enormous number of  $AB^*$  states can be found, is better to express these states *via* density of states ( $\rho(E)$ , number of states per energy intervals).<sup>38</sup> The final expression for  $k_2(E)$  in the RRKM theory is

$$k_2(E) = \frac{kT}{h} \frac{W^\ddagger(E - E_0)}{\rho(E)kT} = \frac{W^\ddagger(E - E_0)}{h\rho(E)} \quad (1.25)$$

where  $W$  is the statistical weight ratio, this can be determined from the sum of states,  $W(E)$  of  $AB^\ddagger$ . The unimolecular rate constant is obtained by integrating over all energy levels above  $E_0$ , weighted by the population distribution,  $g(E)$  as:

$$k_{uni} = \int_{E_0}^{\infty} \frac{k_2(E)k_1(E)/k_{-1}}{1 + K_2(E)/k_{-1}[M]} g(E) dE. \quad (1.26)$$

The ratio  $k_1(E)/k_{-1}$  is written as:

$$\frac{k_1(E)}{k_{-1}} = K_{eq} = \frac{Q_{AB^*}(E)}{Q_{AB}} \quad (1.27)$$

where  $Q$  is the partition function, and is defined as:

$$Q = \sum_{i=0}^{\infty} g_i e^{-E_i/kT} \quad (1.28)$$

applying Equation 1.28 to a "species" only in the range  $E, E + dE$

$$Q(E) = \left( \sum_E^{E+dE} g(E) \right) e^{-E/kT} = \rho(E) e^{-E/kT} \quad (1.29)$$

one can obtain:

$$\frac{k_1(E)}{k_{-1}} = \frac{\rho(E)}{Q_{AB}} e^{-E/kT}. \quad (1.30)$$

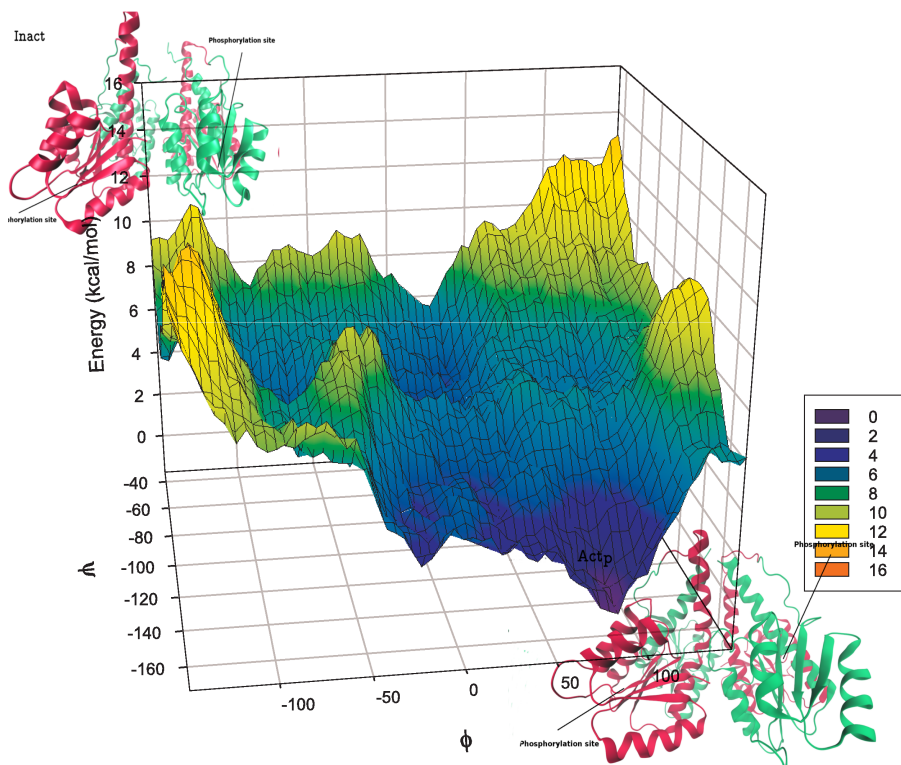
The density of states,  $\rho(E)$ , is then necessary to compute not only  $k_2$  but also for the ratio between the activation and deactivation rate constants.<sup>38</sup>

The power of RRKM theory lies in that it can be used for studying gas phase molecular reactions, such as those that occur in planetary atmospheres or in the interstellar medium.<sup>38</sup>



# 2

## CONCEPTS AND METHODS



*"The most exciting phrase to hear in science, the one that heralds the most discoveries, is not "Eureka!" (I found it!) but 'That's funny...'"*

Isaac Asimov

## 2.1 Molecular Dynamics

A molecular system can be described classically, using the classical Hamiltonian,  $H$  which is a function of both coordinates  $\mathbf{r}$  and momenta  $\mathbf{p}$ . The total energy of the molecular system is then the Hamiltonian  $H$ , which can be written like,

$$H = H(\mathbf{q}, \mathbf{p}) = \mathbf{K}(\mathbf{p}) + U(\mathbf{q}) = \sum_i \frac{p_i^2}{2m_i} + U(\mathbf{q}) \quad (2.1)$$

where  $K(\mathbf{p})$  is the kinetic energy,  $U(\mathbf{q})$  is the potential energy function, which is independent of time and velocity,  $p_i$  is the momentum of particle  $i$ , and  $m_i$  is the mass of particle  $i$ . A microscopic state of the system is therefore characterized by the set of values  $\{\mathbf{q}, \mathbf{p}\}$ , which corresponds to a point in the phase space and is defined by both coordinates  $\mathbf{r}$  and momenta  $\mathbf{p}$ .

In statistical mechanics, thermodynamic averages over a canonical ensemble, which is defined by the macroscopic variables  $(N, V, T)$ , where  $N$  is the number of particles,  $V$  is the volume of the system and  $T$  is the temperature, are obtained computing the probability of finding a system at every state of the phase space (probability distribution  $\rho(q, p)$ ). This probability distribution is described by the Boltzmann distribution function,

$$\rho(\mathbf{q}, \mathbf{p}) = \frac{e^{-H(\mathbf{q}, \mathbf{p})/k_B T}}{Z} \quad (2.2)$$

where  $Z$  is the canonical partition function which is an integral over all phase space of the Boltzmann factors  $e^{-H(\mathbf{q}, \mathbf{p})/k_B T}$ ,  $T$  is the temperature and  $k_B$  is the Boltzmann constant. Once this distribution function is known it can be used to calculate phase space averages of any dynamic variable  $A$  of interest (position, total energy, kinetic energy, fluctuations, etc). These thermodynamic averages are expressed like,

$$\langle A(\mathbf{q}, \mathbf{p}) \rangle_Z = \int_V d\mathbf{q} \int_{-\infty}^{\infty} d\mathbf{p} \rho(\mathbf{q}, \mathbf{p}) A(\mathbf{q}, \mathbf{p}) \quad (2.3)$$

where  $\rho(\mathbf{q}, \mathbf{p})$ , is the distribution probability. Nevertheless, the calculation of these thermodynamic averages is very difficult because it is necessary to know,

simultaneously, the Boltzmann probability for each state  $\{\mathbf{q}, \mathbf{p}\}$ , and this is a complex computational issue.

An alternative approach for computing averages of the entire system, is to follow the motion of a single point through phase space instead of averaging over the whole phase space all at once, which is extremely difficult. Specifically, in this way, the motion of a single molecular state through phase space is followed as a function of time, and the averages are calculated only over those states that were sampled during the process. The motion of a single state through phase space is obtained by integrating the equation of motion of the system. Starting from a state  $\{\mathbf{q}(0), \mathbf{p}(0)\}$ , the integration produces a trajectory that is the set of states  $\{\mathbf{q}(t), \mathbf{p}(t)\}$  describing the system at any successive time  $t$ . Dynamic averages of any dynamical variable  $A(\mathbf{q}, \mathbf{p})$  can be calculated along this trajectory as follows:

$$\langle A(\mathbf{q}, \mathbf{p}) \rangle_\tau = \frac{1}{\tau} \int_0^\tau A(\mathbf{q}(t), \mathbf{p}(t)) dt \quad (2.4)$$

where  $\tau$  is the time of the simulation. In contrast with the thermodynamic averages approach, the dynamic averaging is easier to perform, because the averaging, is made over a single state in phase space as a function of the time. The perfect case in the dynamic average, is that in which the state that is being dynamically followed, will eventually cover all the phase space. In order to achieve this, an infinitely long trajectory will be necessary (ergodic hypothesis). In this case, the thermodynamic ensemble average and the dynamic average become equivalent to each other,

$$\lim_{\tau \rightarrow \infty} \langle A(\mathbf{q}, \mathbf{p}) \rangle_\tau = \langle A(\mathbf{q}, \mathbf{p}) \rangle_Z \quad (2.5)$$

Here, when the trajectory becomes long enough, the state that generates the dynamics, will eventually cover all of the phase space, in such case, the two averages become identical. Two conditions are necessary to cover the ergodic hypothesis; i) the system has to be at equilibrium and, ii) there must not be any obstacle, such as a fragmented topology, that will prevent an infinitely long trajectory from covering all of phase space. A system that obeys these two conditions is said to be ergodic, and its hypothesis is the theoretical justification for using molecular dynamic simulations as a means for calculating thermodynamic averages of molecular systems.

**Newtonian Molecular Dynamics** The dynamic behavior of molecules is best described by the quantum mechanical equation of motion, (i.e., the time-dependent Schrödinger equation). However, because this equation is extremely difficult to solve for large molecular systems, a simpler classical mechanical description is often used to approximate the molecular motion of a system with many atoms. Newtonian molecular dynamics is widely used to simulate complex systems, using forces from both, empirical force fields (FFs) and first-principles (*ab initio*) calculations. Newton's second law of motion describes the force like,

$$F_i = m_i a_i = m_i \ddot{\mathbf{q}}_i \quad (2.6)$$

where  $F_i$  is the force acting on particle  $i$ ,  $m_i$  is the mass of particle  $i$ ,  $a_i$  is its acceleration, and  $\ddot{\mathbf{r}}_i$  is the second derivative of the particle position  $\mathbf{q}$  with respect to time. The force  $F_i$  is determined by the gradient of the potential energy function,  $U(\mathbf{q})$ , which is a function of all the atomic coordinates  $\mathbf{q}$ ,

$$F_i = -\nabla_i U(\mathbf{q}) \quad (2.7)$$

Equation 2.35 is a second-order differential equation. The formulation of Newton's equation of motion can be given in terms of the Hamiltonian  $H$ , in this view, the classical equation of motion is written as a pair of coupled first-order differential equations:

$$\dot{q}_k = \frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial p_k}; \quad \dot{p}_k = -\frac{\partial H(\mathbf{q}, \mathbf{p})}{\partial q_k} \quad (2.8)$$

The first first-order differential equation in Equation 2.8 becomes the standard definition of momentum, while the second turns into Equation 2.6. A set of two first-order differential equations is often easier to solve than a single second-order differential equation.

Characteristic properties can be derived from the Newton's equations of motion, these properties are:

*Conservation of energy.* if the potential function  $U$  and the Hamiltonian  $H$  do not depend explicitly on time or velocity (so that  $\partial H/\partial t = 0$ ), from Equation 2.8 the total derivative  $dH/dt$  is zero; i.e., the Hamiltonian is a constant of motion for Newton's equation, therefore total energy of the system will be conserved.<sup>39</sup>



*Conservation of linear and angular momentum.* If the potential function  $U$  depends only on particle separation (as is usual) and there is no external field applied, then Newton's equation of motion conserves the total linear momentum of the system,  $\mathbf{P}$ ,

$$\mathbf{P} = \sum_i \mathbf{p}_i \quad (2.9)$$

and the total angular momentum,  $\mathbf{L}$ ,

$$\mathbf{L} = \sum_i \mathbf{q}_i \times \mathbf{p}_i = \sum_i m_i \mathbf{q}_i \times \dot{\mathbf{q}}_i \quad (2.10)$$

*Time reversibility.* The third property of Newton's equation of motion is that it is reversible in time. Changing the signs of all velocities (or momenta) will cause the molecule to retrace its trajectory. If the equations of motion are solved correctly, then the numerical trajectory should also have this property. However, in practice this time reversibility can be reproduced by numerical trajectories only over very short periods of time because of the chaotic nature of large molecular systems.<sup>39</sup>

Numerical procedure for integrating the differential equation is necessary to solve the Newton's equation of motion. Finite-difference approach is a standard method to solve this kind of differential equations. In this approach, the molecular coordinates and velocities at a time  $t + \Delta t$  are obtained from the molecular coordinates and velocities at an earlier time  $t$ . Every step the equations are solved, the choice of time interval  $\Delta t$  depends on the properties of the molecular system, and  $\Delta t$  must be significantly smaller than the characteristic time of the molecular motion studied. The time step employed in MD simulations for advancing the positions and velocities of all the atoms, must generally be selected to be very small due to the need to represent the bond stretching motion. The smaller the values of  $\Delta t$ , the more accurate will be the numerical solution of Newton's equation of motions.

For example, one of the fastest bond stretching motions corresponds to the C-H bonds. A typical bond stretching frequency of a C-H bond is about  $3000 \text{ cm}^{-1}$ , the period of vibration is  $1.1 \times 10^{-14} \text{ s}$ , or 11 fs. Thus, in order to sample data points for C-H vibration, the time step has to be generally set to 1 fs. However, an small  $\Delta t$  required for MD simulations limits the total simulation time, in other words, while a total simulation time of 1 ps requires

only 1000 evaluations of the equation of motions, a total simulation time of 1 ns requires evaluation of  $10^6$  steps, whereas for a total simulation time of  $1 \mu\text{s}$  requires evaluation of  $10^9$  steps, and a total simulation time of 1 millisecond requires evaluation of  $10^{12}$  steps, which is out of the computing power; this is problematic for some simulations such as protein folding, which may take place on a time scale of seconds, even for a very large system where the degrees of freedom are enormous, like protein complexes.

With the finite-difference methods, the derivatives can be properly-behaved, by Taylor's theorem,

$$\mathbf{q}(t + \Delta t) = \mathbf{q}(t) + \dot{\mathbf{q}}(t)\Delta t + \frac{1}{2}\ddot{\mathbf{q}}(t)\Delta t^2 \dots \quad (2.11)$$

Alternatively, this can be written as

$$\mathbf{q}(t + \Delta t) = \mathbf{q}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2}\mathbf{a}(t)\Delta t^2 + \dots \quad (2.12)$$

where  $\mathbf{v}(t)$  is the velocity vector and  $\mathbf{a}(t)$  is the acceleration. Because the integration proceeds in a stepwise fashion, and recalling Equation 2.6, it is convenient to rewrite the above expansion in a discrete form. Using  $\mathbf{q}_n$  to indicate the position at step  $n$  (at time  $t$ ) and  $\mathbf{q}_{n+1}$  to indicate the position at the next step,  $n + 1$  (at time  $t + \Delta t$ ), Equation 2.12 can be written as

$$\mathbf{q}_{n+1} = \mathbf{q}_n + \mathbf{v}_n\Delta t + \frac{1}{2}\left(\frac{\mathbf{F}_n}{\mathbf{m}}\right)\Delta t^2 + O(\Delta t^3) \quad (2.13)$$

where  $O(\Delta t^n)$  are the terms of order  $\Delta t^n$  or smaller. With this information the velocity  $\mathbf{v}_{n+1}$  at time  $n + 1$  can be crudely estimated, for example, as

$$\mathbf{v}_{n+1} = (\mathbf{q}_{n+1} - \mathbf{q}_n)/2 \quad (2.14)$$

Together, Eqs. 2.12 and 2.14 form an integration algorithm. Given the position  $\mathbf{q}_n$ , the velocity  $\mathbf{v}_n$ , and the force  $\mathbf{F}_n$  at step  $n$ , these equations allow to estimate the position  $\mathbf{q}_{n+1}$  and velocity  $\mathbf{v}_{n+1}$  at step  $n + 1$ . The formulation results in a low quality integration algorithm.<sup>39</sup> More accurate algorithms have been developed using the same kind of reasoning, like those that are summarized below.

**Verlet integrator** The Verlet integrator<sup>40</sup> is based on two Taylor expansions, a forward expansion ( $t + \Delta t$ ) and a backward expansion ( $t - \Delta t$ ),

$$\mathbf{q}_{n+1} = \mathbf{q}_n + \mathbf{v}_n \Delta t + \frac{1}{2} \left( \frac{\mathbf{F}_n}{\mathbf{m}} \right) \Delta t^2 + O(\Delta t^3) \quad (2.15)$$

$$\mathbf{q}_{n-1} = \mathbf{q}_n - \mathbf{v}_n \Delta t + \frac{1}{2} \left( \frac{\mathbf{F}_n}{\mathbf{m}} \right) \Delta t^2 - O(\Delta t^3) \quad (2.16)$$

The sum of the two expansions yields an algorithm for propagating the position,

$$\mathbf{q}_{n+1} = 2\mathbf{q}_n - \mathbf{q}_{n-1} + \frac{\mathbf{F}_n}{\mathbf{m}} \Delta t^2 + O(\Delta t^4) \quad (2.17)$$

The algorithm that describe the above equations follows two steps:

i) Use the current position  $\mathbf{q}_n$  to calculate the current force  $\mathbf{F}_n$  and, ii) use the current and previous positions  $\mathbf{q}_n$  and  $\mathbf{q}_{n-1}$  with the current force  $\mathbf{F}_n$  to compute the position in the next step,  $\mathbf{q}_{n+1}$ , Equation 2.17.

The above procedure is repeated for each atom in the molecule. Velocities can be propagated by subtracting Equation 2.15 from Equation. 2.16,

$$\mathbf{v}_n = \frac{\mathbf{q}_{n+1} - \mathbf{q}_{n-1}}{2\Delta t} + O(\Delta t^2) \quad (2.18)$$

**Leap-Frog Integrator** Because Verlet algorithm experience serious problems computing the velocities, some modifications to the Verlet scheme have been proposed to address this point. One of those are leap-frog algorithm<sup>41</sup> and velocity Verlet algorithm.<sup>42</sup>In the leap-frog algorithm, velocities are evaluated at the midpoint of the position and vice versa,

$$\mathbf{q}_{n+1} = \mathbf{q}_n + \mathbf{v}_{n+1/2} \Delta t \quad (2.19)$$

$$\mathbf{v}_{n+1/2} = \mathbf{v}_{n-1/2} + \frac{\mathbf{F}_n}{\mathbf{m}} \Delta t \quad (2.20)$$

Where  $\mathbf{v}_{n\pm 1/2}$  is the velocity at the mid-step time  $[t \pm (1/2)\Delta t]$ , if one eliminates the velocities from Equation 2.19 and Eq 2.20, algebraically the equation will be equivalent to the Verlet algorithm. The verlet algorithm uses the current

position  $\mathbf{q}_n$  to compute the current force  $\mathbf{F}_n$ , then, the current  $\mathbf{F}_n$  and the previous mid-step velocity  $\mathbf{v}_{n-1/2}$  are used to compute the next mid-step velocity  $\mathbf{v}_{n+1/2}$ . Finally, the current position  $\mathbf{q}_n$  and the next mid-step  $\mathbf{v}_{n+1/2}$ , from the previous procedure, to compute the position in the next step  $\mathbf{q}_{n+1}$ . The current velocity  $\mathbf{V}_n$  can be calculated as

$$\mathbf{v} = (\mathbf{v}_{n+1/2} + \mathbf{v}_{n-1/2})/2 \quad (2.21)$$

**Velocity Verlet Integrator** As it was mentioned before, other algorithm, that can handle the large errors caused by the Verlet scheme, is the Velocity Verlet.<sup>42</sup> This algorithm is a variant of the basic Verlet integrator. The velocity Verlet algorithm stores positions, velocities, and accelerations all at the same time  $t$  and minimizes roundoff errors. The equation can be written as,

$$\mathbf{q}_{n+1} = \mathbf{q}_n + \mathbf{v}_n \Delta t + \frac{1}{2} \left( \frac{\mathbf{F}_n}{\mathbf{m}} \right) \Delta t^2 \quad (2.22)$$

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \frac{1}{2} \left[ \frac{\mathbf{F}_n}{\mathbf{m}} + \frac{\mathbf{F}_{n+1}}{\mathbf{m}} \right] \Delta t \quad (2.23)$$

Current force  $\mathbf{F}_n$  is evaluated, after that, the position  $\mathbf{q}_{n+1}$  at time  $t + \Delta t$  is calculated from Equation 2.22. Then, the velocity at mid-step  $\mathbf{v}_{n+1/2}$  is calculated using the next expression:

$$\mathbf{v}_{n+1/2} = \mathbf{v}_n + \frac{1}{2} \left( \frac{\mathbf{F}_n}{\mathbf{m}} \right) \Delta t \quad (2.24)$$

At this point, the kinetic energy at time  $t + \Delta t$  is available, then a new step start again.

## 2.2 Force Fields

Molecular mechanics (MM) is based on a mathematical model of a molecule as a collection of spheres, which corresponds to atoms held together by springs (corresponding to bonds). The energy of the molecular system changes as a function of the geometry, which makes the mathematical model, conceptually

very intuitive. The MM model clearly ignores electrons. The principle behinds MM, is to express the energy of a molecule as a function of its resistance toward bond stretching, bond bending, and atom-atom repulsion or attraction, hence, this energy equation is used to find the bond lengths, angles, an dihedral corresponding to the minimum-energy geometry, or more precisely, to the various possible potential energy surface (PES) minima.

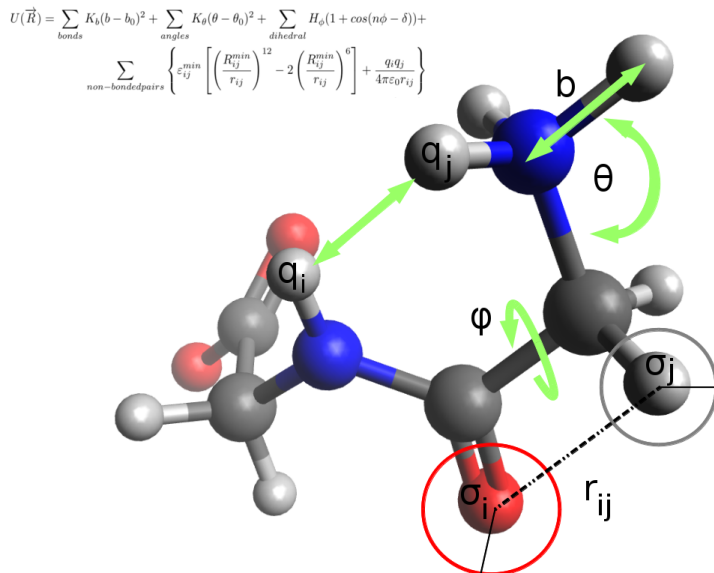
The mathematical expression for the energy constitutes a FF, and MM methods are some times called FF methods. The method makes no reference to electrons, therefore, cannot handle electronic properties like charge distribution, or nucleophilic and electrophilic behavior for the molecules, except for some specialized algorithm.

MM started as an attempt to obtain quantitative information about chemical reactions, at time when the possibility of doing quantitative quantum mechanical calculations, on anything much bigger than the hydrogen molecule, was technically impossible. The principles of MM, as a potentially general method for studying the variation of energy of molecular systems with their geometry was formulated by D. H. Andrews, which proposed extending spectroscopic FF ideas to doing molecular mechanics.<sup>43</sup> In 1940 F. H. Westheimer performed the only molecular mechanics calculation done by hand to determine the transition state of a tetrasubstituted biphenyl.<sup>44</sup> In 1965 K. B. Wiberg publishes first general molecular mechanics type program with ability to find energy minimum. In 1968 Lifson and Warshel, published a Consistent Force Field (CFF),<sup>45</sup> and 1976 N. L. Allinger publishes the first (MM1) in a series of highly popular FFs.<sup>46</sup> Further developments were done which allows us to simulated large and complex systems. Indeed the Nobel Prize in Chemistry 2013 was awarded jointly to Martin Karplus, Michael Levitt and Arieh Warshel for the development of multiscale models for complex chemical systems.

The potential energy of a molecule can be written as

$$E = \sum_{\text{bonds}} E_{\text{stretch}} + \sum_{\text{angles}} E_{\text{bend}} + \sum_{\text{dihedrals}} E_{\text{torsion}} + \sum_{\text{pairs}} E_{\text{nonbond}} \quad (2.25)$$

Where each term are contributions from bond stretching, angle bending, torsional motion (rotation) around single bonds, and interactions between atoms or groups which are nonbonded (not directly bonded together). The sum are over all the bonds, all the angles defined by three atoms A-B-C, all dihedral



**Figure 2.1:** Molecular mechanics potential energy function

angles defined by for atoms A-B-C-D, and all pairs of significant nonbonded interactions. The mathematical form of these terms and the parameters in them constitute a particular forcefield, Figure 2.1.

Here, I mostly focus principally in the CHARMM (Chemistry at HARvard Macromolecular Mechanics) FF,<sup>47</sup> CHARMM is a versatile and widely used molecular simulation program with broad application to many-particle systems. It has been developed with a primary focus on the study of molecules of biological interest, including peptides, proteins, prosthetic groups, small molecule ligands, nucleic acids, lipids, and carbohydrates.

The general form of the potential energy function most commonly used in CHARMM for macromolecular simulations is based on fixed point charges and can be written as

$$\begin{aligned}
U(\vec{R}) = & \sum_{bonds} K_b(b - b_0)^2 + \sum_{angles} K_\theta(\theta - \theta_0)^2 + \sum_{Urey-Bradley} K_{UB}(S - S_0)^2 \\
& \sum_{dihedral} H_\phi(1 + \cos(n\phi - \delta)) + \sum_{impropers} K_\omega(\omega - \omega_0)^2 + \\
& \sum_{non-bondedpairs} \left\{ \varepsilon_{ij}^{min} \left[ \left( \frac{R_{ij}^{min}}{r_{ij}} \right)^{12} - 2 \left( \frac{R_{ij}^{min}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\} + \\
& \sum_{residues} U_{CMAP}(\phi\psi)
\end{aligned} \tag{2.26}$$

The potential energy,  $(U(\vec{R}))$ , is a sum over individual terms representing the internal and nonbonded contributions as a function of the atomic coordinates. Internal terms include bond ( $b$ ), valence angle ( $\theta$ ), Urey–Bradley (UB, $S$ ), dihedral angle ( $\phi$ ), improper angle ( $\omega$ ), and backbone torsional correction (CMAP,  $\phi$ ,  $\psi$ ) contributions, Equation 2.26 The parameters  $K_b$ ,  $K_\phi$ ,  $K_{UB}$ ,  $K_\theta$ , and  $K_\omega$  are the respective force constants and the variables with the subscript 0 are the respective equilibrium values. All the internal terms are taken to be harmonic, except the dihedral angle term, which is a sinusoidal expression, in which  $n$  is the multiplicity or periodicity of the dihedral angle and  $\delta$  is the phase shift. For the protein main chain, a numerical correction term, called CMAP is introduced.<sup>47</sup> For three bonded atoms A-B-C, the Urey-Bradley term is a quadratic function of the distance,  $S$ , between atoms A and C. The improper dihedral angle term is used at branchpoints; that is, for atoms A, B, and D bonded to a central atom, C, the term is a quadratic function of the (pseudo) dihedral angle defined by A-B-C-D.

Nonbonded terms include Coulombic interactions between the point charges ( $q_i$  and  $q_j$ ) and the Lennard-Jones (LJ) 6-12 term are described in Equation 2.26. In the LJ potential the  $r^{12}$  term, which is the repulsive term, describes Pauli repulsion at short ranges due to overlapping electron orbitals and the  $r^6$  term, which is the attractive long-range term, describes attraction at long ranges (van der Waals force, or dispersion force). Nonbonded interactions are calculated between all atom pairs, except for covalently bonded atom pairs (1,2 interactions) and atom pairs separated by two covalent bonds (1,3 interactions). The relative dielectric constant,  $\varepsilon$ , is set to one in calculations with explicit solvent, corresponding to the permittivity of vacuum,  $\varepsilon_0$ . In the LJ term, the well depth is represented by  $\varepsilon_{min}^{ij}$ , where  $i$  and  $j$  are the indices of the

interacting atoms,  $r_{ij}$  is the interatomic distance, and  $R_{min}^{ij}$  is the distance at which the LJ term has its minimum. Typically,  $\varepsilon_{min}^{ij}$  and  $R_{min}^{ij}$  are obtained for individual atom types and then combined to yield  $\varepsilon_{min}^{ij}$  and  $R_{min}^{ij}$  and for the interacting atoms via a standard combination rule. In the current CHARMM FFs, the  $\varepsilon_{min}^{ij}$  values are obtained *via* the geometric mean ( $\varepsilon_{min}^{ij} = \sqrt{\varepsilon_{min}^i \varepsilon_{min}^j}$ ) and  $R_{min}^{ij}$  *via* the arithmetic mean,  $R_{min}^{ij} = (R_{min}^i + R_{min}^j)/2$

### 2.2.1 Force Field Parameterization

The potential energy in Equation 2.26 is described by phenomenological functions, expressed in terms of coordinates, coordinate deviations and coefficients, which can be optimized to reproduce experimental data and electronic structure (ES) calculations. The potential energy function is itself empirical, and is expressed as a function of the position of each atom. The parameters in the potential energy terms can be obtained from experimental structural data (X-ray crystallography or NMR for bond lengths and angles), spectroscopy (IR and Raman, for the force constants) and electronic structure calculations for charges and LJ potential parameters. Some FFs have been parameterized using data from ES calculations only, this allows the determination of FF parameters for molecules where little or no data exists. One example of this, is the Merck Molecular FF (MMFF).<sup>48</sup> MMFF achieve the MM3-like accuracy for small molecules, mainly those that interact with proteins, like cofactors and coenzymes for example. MMFF has been derived from high-quality computations, with structures optimized at high level ES calculations. Furthermore, MMFF has been parameterized for a wide variety of chemical systems of interest, from organic to medicinal chemistry. Whereas, experimental data relevant to one set of parameters may be available in the gas phase, while that for another may be available for a solid, parameters from electronic structure calculations can all be found from the same phase (usually the gas phase). The most difficult parameters to determine, whether from experimental or ES data, are the non-bonded parameters. Van der Waals parameters are often found from experimental crystal structures. They can also be calculated from high level (electronic correlation) ES calculations. However, parameters derived from gas phase cannot reproduce, in many cases, condense properties, mainly because gas phase ES calculations can neglect interactions involving three or more atoms. These can have a large contribution to the energy of a system, so neglecting these leads to large errors in the calculated values of some quantities.



FFs like the Optimized Parameters for Liquid Simulation (OPLS) are adjusted to reproduce experimental values of thermodynamic quantities such as density and heats of vaporization, using Monte Carlo simulations on pure liquids.

Since more small molecules with biological and chemical relevance are known nowadays, more efficient algorithms are required in order to parameterize the FFs. A variety of plugins and packages, that perform the complete parameterization protocol of small organic molecules in a relative short time currently exist. One example of these programs is the plugin Paratool, which is found in the Visual Molecular Dynamics (VMD) program. In Paratool the Hessian is transformed into internal coordinates to get the force constants of bond, angle, proper and improper dihedral angles, charges are obtained from the Mulliken charges and LJ parameters are obtained by analogy with similar chemical species. Of course to transform the Hessian matrix is necessary to provide *ab initio* calculations to the plugin (electronic optimized structure, frequency calculations and Hessian transform files), which is the most time-consuming part. As it was mentioned above, in order to develop a FF, empirical, semi-empirical or pure *ab initio* calculation, has been used to find the best combination of parameters that describe the PES of a chemical system. To determine the best combination of parameters, least-squares minimization and combination rules to optimize the set of the FF parameters has been used to parameterize many chemical species in CHARMM, OPLS, Amber and consistent valence forcefield (CVFF).

Because there are not empirical information for all the chemical compounds, unfortunately, one alternative approach, assuming that the true PES can be approximated using *ab initio* calculations, is that in which the FF can be directly fitted to a calculated QM PES by numerically matching the gradients (or energy).<sup>49</sup> This approach has been used to create so-called quantum mechanical FFs (QMFF) for simple di and triatomics, as well as hydrocarbons.<sup>50-52</sup> On the other hand, some approach used, instead of ES energies calculation as a reference, the vibrational frequencies to be fitted having as a reference both vibrational frequencies and eigenvalue projections derived from ES calculation.<sup>53</sup> Fitting algorithms are generally based on optimizing a "merit function" that corresponds to minimizing a weighted sum of square deviations between the classical data and reference data points from the ES PES, the goal, then, is to find the best algorithm that finds the best set of parameters that represent the PES obtained by ES calculations. This is not an easy task and can take a lot of time before to start a simulation (eg. MD simulation). Here we describe an algorithm that does not use second derivatives to minimize the sum of the square deviations between both ES calculation and MM PES, and it is

described in the next section.

### 2.2.1.1 Downhill Simplex algorithm

Given a set of  $N$  data points  $(x_i, y_i)$  with errors  $\sigma_i$  and a function  $f(x; \mathbf{c})$  which depends on the vector of coefficients  $\mathbf{c} = (c_1, c_2, \dots, c_n)$ , least-square fitting is a problem of finding a choice of coefficients  $\mathbf{c}_0$  which minimizes the distance between the function and the datapoints. In this context the distance is defined as:

$$\chi^2 = \sum_{i=1}^N [y_i - f(x_i; \mathbf{c})]^2 / \sigma_i^2 \quad (2.27)$$

In the FF parameterization, from Equation 2.27  $y_i$  corresponds to ES reference energies,  $f(x_i; \mathbf{c})$  corresponds to the energies from MM calculations,  $r_i$  corresponds to the coordinates of the system and  $\mathbf{c}$  are the set of parameters (force constants, equilibrium bonds, equilibrium angles, etc.). If  $f(x_i; \mathbf{c})$  depends linearly on the  $c_i$ 's the problem can be solved using linear algebra, but if  $f(x_i; \mathbf{c})$  is nonlinear in some or all of the  $c_i$ 's then other methods have to be used. In the most general case  $\chi^2$  has an arbitrary dependence on the  $c_i$ 's, and the problem of finding  $\mathbf{c}_0$  is thus equivalent to finding a global minimum in  $n$ -dimensional space. One can use the downhill simplex method to find a local minimum of  $\chi^2$  in  $n$ -dimensional space. This method is very robust and it requires relatively few evaluations of the function. On the other hand, in order to find the local minimum, is necessary to define an initial simplex, which are related with the initial parameters of the FF, if the FF does not start with a good guess of the parameter set, the minimum that the algorithm finds, will not be necessarily a global minimum. To mend the problem of finding the global minimum the initial simplex can be chosen near to the global minimum, or in other words one has to be able to guess the approximate location of the global minimum of  $\chi^2$  in  $n$ -dimensional space. It means that if a good set of parameters are provided, the Downhill simplex method will return  $\mathbf{c}_0$  with a given accuracy. But let's describe in some detail the Downhill Simplex algorithm:

**The Simplex algorithm** From a geometrical point of view, a simplex is a polyhedron containing  $N + 1$  points in a  $N$  dimensional space, for example, a 2-simplex is a triangle, a 3-simplex is a tetrahedron, and a 4-simplex is a 5-cell

and so on. A single point may be considered a 0-simplex, and a line segment may be considered a 1-simplex. A simplex may be defined as the smallest convex set containing the given vertices. When the points are equidistant, the simplex is said to be regular. The simplex method compares the values of the objective function at the  $N + 1$  vertices and then moves this polyhedron gradually towards the optimum point during the iterative process. For each iteration of the algorithm, it attempts to replace simplex vertices that yield high function values with vertices whose new function values are lower. The aim of this algorithm is to move the simplex, by replacing vertices, into the neighborhood of a minimizer. The motion of the simplex is given by three operations: reflection, expansion and contraction.

**Reflection:** in the reflection motion, the worst vertex, corresponding to the worst value of the objective function among all vertices, is reflected in the opposite face to obtain a new value. Since the direction of this motion is always away from the worst result, the reflection moves towards a favorable direction. When the objective function does not have steep valleys, many reflections are performed until lead to the function optimum point, Figure 2.2. In a 2-dimension simplex, like in the figure 2.2, if the vertex A is the highest and vertex C the lowest, one can expect the point D obtained by the reflecting the point A in the opposite face to have the smallest value. If this is the case, to construct a new simplex by rejecting the point A from the original simplex and including the point D. This process can be written as

$$D = (1 + \alpha)H - \alpha A \quad (2.28)$$

Where  $\alpha > 0$  is a *refletion coefficient*, A is the vertex corresponding to the maximum function value and H is the centroid of all the points except A.

**Expansion:** When the reflection process gives a new minimum, one can generally expect to decrease the function value further by moving along the same direction expanding the simplex. In the 2-dimension simplex of Figure 2.2, has its minimum in the vertex D. One can expect the point E obtained by expanding the simplex in the same direction gives a better value. If this is the case, a new simplex is made by rejecting the point D from the previous simplex and including the new point E. This process can be written as

$$E = \beta D + (1 - \beta)H \quad (2.29)$$

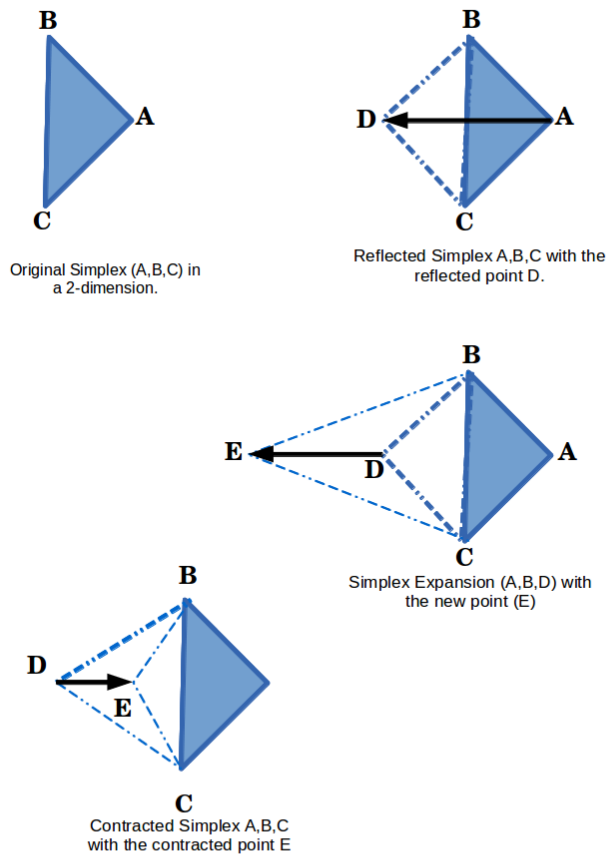


Figure 2.2: Types of simplex moves

where  $\beta > 1$  is the expansion coefficient, D is the vertex created by the reflection and H is the centroid of all the points except D.

**Contraction :** when the reflection process gives a new point that is worse than all the other vertices except the worst, the simplex is contracted by exchanging the new point with the worst point. A new simplex, then, is constructed by rejecting the point D from the previous simplex and including the new point E. Figure 2.2 is showing the process in which the vertex E is created by shrinking the reflected simplex. This process can be written as

$$E = \beta D + (1 - \beta)H \quad (2.30)$$

where  $\beta \in [0, 1]$  is a contraction coefficient, D is the vertex created by reflection and H is the centroid of all the points except D.

The original Nelder and Mead downhill method,<sup>54</sup> has been modified to consider not only the minimization of a function, but also the enforcement of feasibility. During the last years, several methods have been proposed for handling constraints. The methods based on penalty functions are based on the concept that the objective function  $f(x)$  is increased according to the intensity of the constraint violation. This is done to ensure that the vertices do not deviate too much from the constraints. The penalty function can be written as

$$P(x, \rho) = f(x) + \rho * c(x) \quad (2.31)$$

where  $\rho$  is the penalty parameter and  $c(x)$  is the sum of all constraint violation. A deficiency of this methodology is that, the penalty parameter can be problem-dependent and might draw the simplex to premature converge. A way to overcome this problem is to adapt the penalty parameter to the present maximum and minimum function values.

$$\rho = f_{max}(x) - f_{min}(x) \quad (2.32)$$

Each function  $c_i(x)$  can be transformed into a fuzzy, introducing a tolerance parameter  $k$  in the  $i^{th}$  constraint definition. If  $c_i(x) \geq k$  then  $x$  is completely unfeasible. For  $0 < C_i(x) < k$  the distinction between feasible and unfeasible values is fuzzy.

The termination criteria in any multidimensional algorithm is crucial. Because the simplex algorithm does not use derivatives, there is not termination criterion which is based on the gradient of the objective function. Therefore, the algorithm will stop when it cannot find solution with improvements higher than some tolerance.

## 2.3 Nonadiabatic Dynamics

### 2.3.1 Mean-Field (Ehrenfest) Method

If one consider that nonadiabaticity involve changes in the adiabatic state population, designed by  $|a_i^2|$ , changing the nuclear configuration, the distortion of the electron cloud will influence the nuclear trajectory. Although in some situations this is negligible, for many chemical systems it is important to incorporated electronic-nuclear feed-back.<sup>55,56</sup>

If the atomic nuclei are propagated by solving Newton's equations:<sup>56</sup>

$$M_K \ddot{\mathbf{R}}_K = \mathbf{F}_K(\mathbf{R}), \quad (2.33)$$

where

$$\mathbf{F}_K(\mathbf{R}) = -\nabla_K E_i(\mathbf{R}), \quad (2.34)$$

and replacing the adiabatic potential energy surface APES  $E_i$  by the energy expectation value, then one can obtain:<sup>56</sup>

$$E_{eff} = \langle \Phi | \mathcal{H}_{el} | \Phi \rangle = \sum_i |a_i|^2 E_i, \quad (2.35)$$

where,

$$\Phi(\mathbf{r}; t) = \sum_j a_j(t) \phi_j(\mathbf{r}, \mathbf{R}) e^{-\frac{i}{\hbar} \int E_j(\mathbf{R}) dt} \quad (2.36)$$

In that way, the atoms evolve on an effective potential, representing an average over the adiabatic states weighted by their population  $|a_i|^2$ .<sup>55</sup>

The expression for the nuclear forces either from the gradient of the equation 2.35 or using the Hellmann-Feymann theorem:<sup>56</sup>

$$\mathbf{F}_K = -\langle \Phi | \nabla_K \mathcal{H}_{\text{el}} | \Phi \rangle, \quad (2.37)$$

the relation can be written as:

$$\begin{aligned} \langle \Phi | \nabla_K \mathcal{H}_{\text{el}} | \Phi \rangle &= \nabla_K E_i \delta_{ij} \\ &= \langle \nabla_K \phi_i | \mathcal{H}_{\text{el}} | \phi_j \rangle + \langle \phi_i | \nabla_K \mathcal{H}_{\text{el}} | \phi_j \rangle + \langle \phi_i | \mathcal{H}_{\text{el}} | \nabla_K \phi_j \rangle \\ &= \langle \phi_i | \nabla_K \mathcal{H}_{\text{el}} | \phi_j \rangle + (E_j - E_i) \mathbf{d}_{ji}, \end{aligned} \quad (2.38)$$

where the nonadiabatic coupling vectors,  $\mathbf{d}_{ji}$ , is defined as

$$\mathbf{d}_{ji} = \langle \phi_j | \nabla_K | \phi_i \rangle, \quad (2.39)$$

and assuming that the solution of the time-dependent (electronic) Schrödinger equation as

$$\mathcal{H}_{\text{el}}(\mathbf{r}, \mathbf{R}) \phi_i(\mathbf{r}, \mathbf{R}) = E_i(\mathbf{R}) \phi_i(\mathbf{r}, \mathbf{R}), \quad (2.40)$$

and together with the hermiticity of  $\mathcal{H}_{\text{el}}$

$$\langle \phi_i | \mathcal{H}_{\text{el}} | \nabla_K \phi_j \rangle = \langle \nabla_K \phi_j | \mathcal{H}_{\text{el}} | \phi_i \rangle^* = \langle \nabla_K \phi_j | E_j \phi_j \rangle^* = E_i \mathbf{d}_{ij}^* = -E_i \mathbf{d}_{ij}. \quad (2.41)$$

Note that

$$\mathbf{d}_{ij}^* = -\mathbf{d}_{ij}, \quad (2.42)$$

because

$$\begin{aligned}\nabla_K \langle \phi_i | \phi_j \rangle &= \nabla_K \delta_{ij} = 0 \\ &= \langle \lambda_K \phi_i | \phi_j \rangle + \langle \phi_i | \nabla_K \phi_j \rangle = \mathbf{d}_{ij}^* + \mathbf{d}_{ij}.\end{aligned}\quad (2.43)$$

Equating the rhss in equation 2.38 one obtains after rearranging:<sup>56</sup>

$$\langle \phi_i | \nabla_K \mathcal{H}_{\text{el}} | \phi_j \rangle = \nabla_K E_i \delta_{ij} - (E_j - E_i) \mathbf{d}_{ij}.\quad (2.44)$$

The nuclear forces are thus given by

$$\mathbf{F}_K = - \sum_i |a_i|^2 \nabla_K E_i + \sum_{i,j} a_i^* a_j (E_j - E_i) \mathbf{d}_{ij}.\quad (2.45)$$

Last equation illustrate the two contributions to the nuclear forces: the first term is simply the population-weighted average force over the adiabatic states, while the second term, takes into account nonadiabatic changes of the adiabatic state occupation. Is important to emphasize that the nonadiabatic contributions to the nuclear forces are in the direction of the nonadiabatic coupling vectors  $\mathbf{d}_{ij}$ .<sup>56</sup>

Even when the method has some limitations, it has been applied with great success to a number of chemical problems, including energy transfer at metal surfaces.<sup>56</sup>

### 2.3.2 Surface Hopping

It was mention above that after exiting a well localized nonadiabatic coupling region, it is unphysical for nuclear motion to be governed by a mixture of adiabatic states. Rather it, would be desible that in asymptotic regions the system evolves on a pure adiabatic PES. This idea is fundamental to the surface hopping approach. Instead of calculating the “best” (i.e., state-averge) path like in the Ehrenfest method, the surface hopping technique involve an ensemble of trajectories. At any moment in time, the system is propagated on some pure adiabatic state  $i$ , which is selected according to its state population  $|a_i|^2$ . Changing adiabatic state occupations can thus result in nonadiabatic transition between different adiabatic PESs. The ensemble averaged number of



trajectories evolving on adiabatic state  $i$  at any time is equal to its occupation number  $|a_i|^2$ .<sup>56</sup>

In the original formulation of the surface hopping method by Tully and Preston,<sup>57</sup> switches between adiabatic states were allowed only at certain locations, defined prior the simulation. Tully later generalized the method in such a way that nonadiabatic transitions can occur at any point in configuration space.<sup>58</sup> At the same time, an algorithm, known as the fewest switches criterion, was proposed which minimizes the number of surfaces hops per trajectory, whilst guaranteeing the correct ensemble averaged state population at all times.<sup>56</sup> The latter is important because excessive surface switching effectively results in weighted averaging over the adiabatic states much like in the case of the Ehrenfest method.<sup>56</sup>

The derivation of the fewest switches criterion is as follows: out of a total of  $N$  trajectories,  $N_i$  will be in state  $i$  a time  $t$ ,

$$N_i(t) = \rho_{ii}(t)N. \quad (2.46)$$

The density matrix notation is represented by,

$$\rho_{ij}(t) = a_i^*(t)a_j(t). \quad (2.47)$$

At a later time  $t' = t + \Delta t$  the new occupation numbers are

$$N_i(t') = \rho_{ii}(t')N. \quad (2.48)$$

If one assume that  $N_i(t')$ ,  $N_i(t)$  or  $\Delta N = N_i(t) - N_i(t') > 0$ , then the minimum number of transitions required to go from  $N_i(t)$  to  $N_i(t')$  is  $\Delta N$  hops from state  $i$  to any other state and zero hops from any other state to state  $i$ .<sup>56</sup> The probability  $P_i(t, \Delta t)$  for a transition out of state  $i$  to any other state during the time interval  $[t, t + \Delta t]$  is given by

$$P_i(t, \Delta t) = \frac{\Delta N}{N} = \frac{\rho_{ii}(t) - \rho_{ii}(t')}{\rho_{ii}} \approx -\frac{\dot{\rho}_{ii}\Delta t}{\rho_{ii}}, \quad (2.49)$$

where

$$\dot{\rho}_{ii} \approx \frac{\rho_{ii}(t') - \rho_{ii}(t)}{\Delta t}. \quad (2.50)$$

The above expression can be rewritten as

$$\dot{\rho}_{ii} = \frac{d}{dt}(a_i^* a_i) = \dot{a}_i^* a_i + a_i^* \dot{a}_i = (a_i^* \dot{a}_i) + a_i^* \dot{a}_i = 2\mathcal{R}(a_i^* \dot{a}_i), \quad (2.51)$$

if  $\dot{a}_i$  is define as

$$\dot{a}_i = - \sum_j a_j C_{ij} e^{-\frac{i}{\hbar} \int (E_j - E_i) dt}, \quad (2.52)$$

where

$$C_{ij} \equiv \langle \phi_i | \frac{\partial}{\partial t} | \phi_j \rangle, \quad (2.53)$$

Then inserting eq 2.51 into 2.52

$$\dot{\rho}_{ii} = -2\mathcal{R} \left( \sum_j \rho_{ij} C_{ij} e^{-\frac{i}{\hbar} \int (E_j - E_i) dt} \right). \quad (2.54)$$

Substituting expression 2.54 into 2.49 the probability  $P_i$  can be rewritten as follows

$$P_i(t, \Delta t) = \frac{2\mathcal{R} \left( \sum_j \rho_{ij} C_{ij} e^{-\frac{i}{\hbar} \int (E_j - E_i) dt} \right) \Delta t}{\rho_{ii}}. \quad (2.55)$$

Since the probability,  $P_i$ , for a switch from state  $i$  to any other state must be the sum over all states of the probabilities,  $P_{ij}$ , for a transition from state  $i$  to a specific state  $j$ ,<sup>56</sup>

$$P_i(t, \Delta t) = \sum_j P_{ij}(t, \Delta t), \quad (2.56)$$

it follows from eq. 2.55 that

$$P_{ij}(t, \Delta t) = \frac{2\mathcal{R} \left( \rho_{ij} C_{ij} e^{-\frac{i}{\hbar} \int (E_j - E_i) dt} \right) \Delta t}{\rho_{ii}}. \quad (2.57)$$

A transition from state  $i$  to state  $k$  is invoked if  $P_i^k < \zeta < P_i^{(k+1)}$ , where  $\zeta (0 \leq \zeta \leq 1)$  is a uniform random number and  $P_i^k$  is the sum of the transition probabilities for the first  $k$  states,<sup>56</sup>

$$P_i^{(k)} = \sum_j^k P_{ij}. \quad (2.58)$$

In order to conserve total energy after surface hop takes place, the atomic velocities has to be rescaled. The usual procedure is to adjust only the velocity components in the direction of the nonadiabatic coupling vector  $d_{ik}(\mathbf{R})$ . One can follow this procedure qualitatively if one assume that the nonadiabatic contribution to the Ehrenfest forces also are in the direction of the nonadiabatic coupling vector  $d_{ik}(\mathbf{R})$ . Such discontinuities in nuclear velocities must be regarded as a flaw of the surface hopping approach. In mos physical scenarios, however, nonadiabatic surface switches take place only at relatively small potential energy separations, so that necessary adjustment to the nuclear velocities is small.<sup>56</sup>

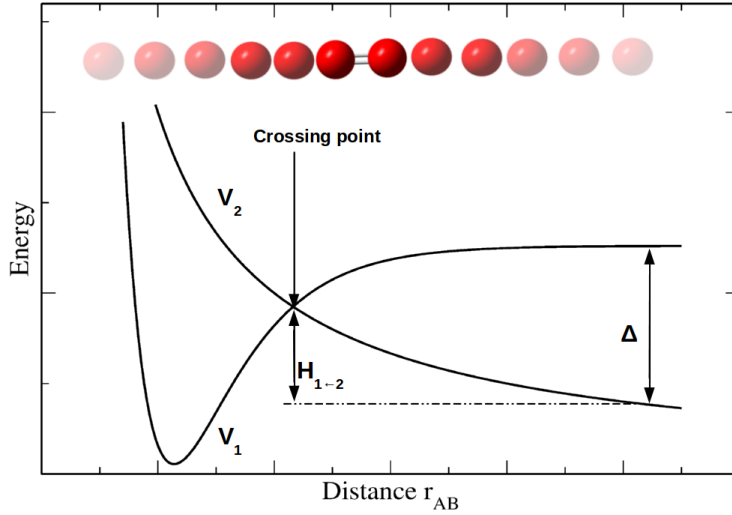
## 2.4 Reactive Force Fields

The most exciting part of the chemistry, according to my point of view, is the reaction mechanism and the dynamics of such mechanism of chemical reactions. Molecules like to be in a minimum of energy depending the conditions of temperature and pressure of the environment where they belong. One perturbation over the system or directly in the molecule or in the ensemble of molecules, may produce a chemical reaction. There are many experimental and computational techniques to, somehow, describe the mechanism, the dynamics and the kinetics of some chemical reactions. Unfortunately for many interesting reactions experiments cannot be perform, and also some molecular systems are or to big or the reaction is too slow (*nstos*) to perform *ab initio* molecular dynamics (AMD). On the other hand, the regular FFs can describe the PES

of a molecular entity (gas phase) or many molecular entities which interact each other (condense phase), in a minimum of energy, which means that a regular FF cannot simulate chemical reaction. A FF that can describe a global PES of reaction is called reactive force fields RFFs. The RRFs overcome the problem of regular FF allowing bond breaking and bond formation, and the parameterization should be addressed to this purpose. The first RFF was created by the 2013 Nobel Laureate Arieh Warshel, the empirical valence bond (EVB)<sup>59</sup> which is one of the most powerful tool for examining chemical reactivity in the condensed phase, according to Kamerlin and Warshel.<sup>60</sup> Another RRF created in the past decade is ReaxFF, which is a bond order based RFF, developed by Adri van Duin, William A. Goddard, III and co-workers at the California Institute of Technology. ReaxFF aims to be as general as possible and has been parameterized and tested for hydrocarbon reactions, transition-metal-catalyzed nanotube formation, and high-energy materials.<sup>61</sup> Water is so important in organic and biochemical reactions, indeed, interesting redox biochemical reactions like photosynthesis and oxidative phosphorylation, involve bond breaking and bond formation of water molecules. Exchange of protons with water to make that, an acid or basic aminoacid, change the total charge and the result be a functional protein (eg. catalytic activity, ligand-protein interaction etc.) and other many examples where water play a principal role in the reaction, make water an attractive molecule to develop a RRF which can exchange protons with some other water and with the molecules around them. One attempt to achieve the above aim is the reactive force field for water (RWFF) developed by Detlef W. M. Hofmann, Liudmila N. Kuleshova and Bruno D'Aguzzo.<sup>62</sup> RWFF reproduces the experimental data of neutron scattering accurately, and allows the simulation of bond formation/breaking of water and acids. Finally, but not less important than the other RFF mentioned above, is the Adiabatic Reaction Molecular Dynamics (ARMD) and multi-state ARMD, which I will describe in more detail below. Also a description of ReaxFF RFF will be written in detail.

### 2.4.1 Adiabatic Reactive Molecular Dynamics

Classical FF, like for example CHARMM FF, are defined by the equation 2.26, which means that bonds, angles, proper and improper dihedral angles and non-bonded interactions describe a molecule as the sum of different functional forms. Now let's, for example, describe a simple reaction in which atoms A



**Figure 2.3:** Schematic representation of two potentials  $V_1$  and  $V_2$  showing bonding and non-bonding interactions between two atoms A and B along their displacement coordinate  $r_{AB}$ .  $\Delta$  corresponds to the energetical gap between  $V_1$  and  $V_2$  at their asymptotic limit.  $H_{1 \leftarrow 2}$  is the activation enthalpy for a transition from  $V_2$  to  $V_1$  defined as the energy difference of the potential minimum on  $V_2$  and the crossing point of  $V_1$  with  $V_2$ .

and B are bonded with a covalent bond (A-B), the bond, eventually, is broken forming a van der Waals complex (A - - B). If one assume that the bond is described by a Morse potential, then the equation 2.26 will be like

$$V_{r_{AB}} = D_e(1 - e^{-\beta(r_{AB}-r_e)})^2 \quad (2.59)$$

where  $r_{AB}$  is the distance between the atoms A and B,  $r_e$  is the equilibrium bond distance between atom A and B,  $D_e$  is the well depth (defined relative to the dissociated atoms), and  $\beta$  is a constant parameter that controls the "width" of the potential, Figure 2.3. On the other hand, the van der Waals complex can be defined by solely non-bonded interactions of the form

$$V_{AB} = \varepsilon_{AB}^{min} \left[ \left( \frac{R_{AB}^{min}}{r_{AB}} \right)^{12} - 2 \left( \frac{R_{AB}^{min}}{r_{AB}} \right)^6 \right] + \frac{q_A q_B}{4\pi\epsilon_0 r_{AB}} \quad (2.60)$$

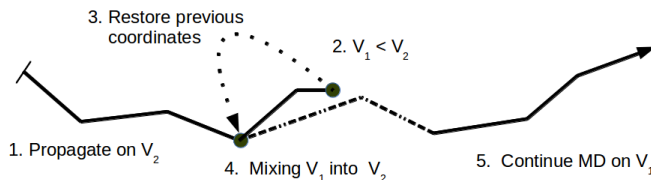
where the terms and parameters are the same as for Equation 2.26. Now

Equation 2.59 represents our PES of the bonded state of the atoms A and B and Equation 2.60 the PES for the the van der Waals complex, respectively. A problem arise because that the global PES must be a combination of the two different FFs, and the algorithm which describes the reaction, somehow, must be able to switch between one surface to the other (Figure 2.3). In other words, once the reaction is observed, the bond which is described for the morse potential, has to be turned off and the van der Waals complex, defined by the potential in the Equation 2.60, has to be turned on. ARMD overcome the problem of mixing both surfaces as follows.<sup>63,64</sup> First let's call the bonded potential  $V_1$  and the non-bonded one  $V_2$ . Figure 2.4 shows a schematic representation of the ARMD algorithm.  $V_1$  represents the potential on which the MD simulation gets initiated in case  $r_{AB}$  is fluctuated in the bond equilibrium distance ( $r_e$ ) and  $V_1$  is expected to be energetically favored over the unbound state  $V_2$  (step 1 in Figure 2.4). In each MD step of the ongoing simulation  $V_1$  and  $V_2$  are compared in terms of their potential energy difference  $\Delta$  found in the asymptotic limit of each potential (Figure 2.3). When A and B loose their bond nature (A starts to separate from B)  $V_1$  starts to increase and  $V_2$  is becoming more stable. When the distance is larger than its equilibrium  $r_e$ ,  $V_1$  and  $V_2$  cross each other. In this case  $V_2$  becomes lower in energy compared to  $V_1$  ( $\Delta < 0$ ) and the ARMD module stops the simulation (step 2 in Figure 2.4) to rewind coordinates and velocities of the system for a user defined amount of timesteps  $m/2$  (step 3 Figure 2.4). Subsequently, the simulation is continued from this point (step 4 Figure 2.4) for  $m$  timesteps where  $V_1$  is replaced by an intermediate potential  $V_{1\leftarrow 2}$  of the form:

$$V_{1\leftarrow 2} = V_1 \frac{\tanh[a(t - t_0)] + 1}{2} + V_2 \left( 1 - \frac{\tanh[a(t - t_0)] + 1}{2} \right) \quad (2.61)$$

where  $t$  is the current time,  $t_0$  is the time at which the crossing occurred and  $a$  is a constant which is related to the chosen mixing time  $m$ . The functional form of  $V_{1\leftarrow 2}$  describes a smooth transition from  $V_2$  towards  $V_1$  within  $m$  timesteps which accounts for an adiabatic process. After the convergence of  $V_{1\leftarrow 2}$  into  $V_2$  is completed, the MD simulation is continued (step 5 Figure 2.4).

ARMD introduces an additional variable  $\Delta_{1\leftarrow 2}$  which modifies  $\Delta$  between  $V_1$  and  $V_2$  to a meaningful magnitude to circumvent the lack of a reference energy.  $\Delta_{1\leftarrow 2}$  represents the potential shift between the two potentials which is, in every ARMD simulation, directly subtracted from  $V_2$ . Appropriate values for  $\Delta_{1\leftarrow 2}$  can be fitted to the transition enthalpy  $H_{1\leftarrow 2}$  between  $V_1$  and  $V_2$  or



**Figure 2.4:** Schematic representation of the five steps taking place during an ARMD simulation involving a transition between two potentials  $V_1$  and  $V_2$  along a fictitious trajectory. The dashed line describes the duration  $m$  during which  $V_2$  is slowly transferred into  $V_1$ .

free energies of activation  $\Delta G^\ddagger$  which are related to experimentally accessible transition rates  $k$  by the transition-state-theory.<sup>30</sup>

As it was mentioned above, the CHARMM FF describes dispersion interactions between two nonbonded atoms  $i$  and  $j$  with a  $LJ$  12-6 potential as specified in the second term of Equation 2.26. In particular for CHARMM,  $\sigma_{ij}$  is defined as the sum of the individual van der Waals radii  $\sigma_i$  and  $\sigma_j$ , and  $\varepsilon_{ij}$  describes the root of the product of the atom related well depths  $\varepsilon_i$  and  $\varepsilon_j$  in Equation 2.26.  $\sigma_{ij}$ , which corresponds to the equilibrium distance of the dispersion interaction, is usually fitted to unnaturally large values. The reason for this overestimation is to prevent clashes between different tertiary structure elements in a protein or at the interface of the protein. However, this form of the potential and its parameters are generally too repulsive for the description of an energetically favorable bond formation. To address this problem, ARMD implements the possibility to replace the potential form of the  $LJ$  interaction between bond forming atom pairs, like  $i$  and  $j$  by a general potential of the form:

$$V_{LJ_{ij}} = \varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r} \right)^x - 2 \left( \frac{\sigma_{ij}}{r} \right)^y \right] \quad (2.62)$$

This potential form allows to replace the  $LJ12 - 6$  function coming from the CHARMM FF with a user defined  $LJx - y$  function, where  $x$  and  $y$  have to be defined uniquely for all interactions describing a dissociable bond in a single ARMD simulation.  $\varepsilon_{ij}$  and  $\sigma_{ij}$  can be assigned individually for each of these atom pairs and correspond to the depth of the potential well and the sum of the van der Waals radii between  $i$  and  $j$  on the dissociated potential. For certain systems, too large values of  $\sigma_{ij}$  can even have an impact on an atom  $k$  which is constantly bound (on  $V_1$  and  $V_2$ ) to atom  $j$  and therefore is connected to atom  $i$  over two bonds on  $V_1$ . To obtain similar benefits for the van der Waals interaction between  $i$  and  $k$  on  $V_2$  as for the interaction between  $i$  and  $j$  through the introduction of the modified potential in the equation 2.62, the standard CHARMM  $LJ$  potential acting between the former atoms on  $V_2$  is similarly replaced by a potential that can be written as

$$V_{LJ_{ij}} = \sqrt{\varepsilon_k \varepsilon_{ij}} \left[ \left( \frac{\sigma_k + \sigma_{ij}/2}{r} \right)^{12} - 2 \left( \frac{\sigma_k + \sigma_{ij}/2}{r} \right)^6 \right] \quad (2.63)$$

where  $\varepsilon_C$  and  $\sigma_C$  are the general CHARMM force field  $LJ$  parameters for atom  $C$  and  $\varepsilon_{AB}$  and  $\sigma_{AB}$  are described by the same parameters used in Equation 2.62. This representation of the van der Waals interaction basically corresponds to the standard CHARMM  $LJ$  potential where  $\varepsilon_i = \varepsilon_C$ ,  $\varepsilon_j = \varepsilon_{AB}$ ,  $\sigma_i = \sigma_C$  and  $\sigma_j = \sigma_{AB}/2$ . In this way, the van der Waals interaction between atoms  $A$  and  $C$  experiences a similar but weaker degree of scaling as the interaction between  $A$  and  $B$  which is replaced by Equation 2.62. It assures that a smooth transition of the van der Waals interactions between the non-bonding 1-2 and 1-3 interactions is still maintained in case  $\sigma_{AB}$  is getting significantly reduced over the sum of  $\sigma_i$  and  $\sigma_j$  of atoms  $A$  and  $B$ .

### 2.4.2 Multi-Surface/ Adiabatic Reactive Molecular Dynamics

As during surface crossing the ARMD potential energy is explicitly time-dependent, the total energy of the system cannot be conserved in a strict sense (see below). For large systems (e.g. proteins in solution) total energy was found to be conserved to within  $\approx 1$  kcal/mol which is sufficient for most applications.<sup>64,65</sup> However, for small systems in the gas phase this is not necessarily true. In cases where only the reaction time is of interest (i.e. the time to reach the transition state), exact energy conservation is not necessarily



required. This was the case for vibrationally induced photodissociation of  $\text{H}_2\text{SO}_4$ .<sup>66</sup> If, however, several crossings between the states involved can take place or the course of the dynamics after the reaction is of interest - e.g. for a final state analysis - energy conservation becomes crucial. In order to explore the range of applicability of the ARMD method, an approximate expression for the energy violation is derived and analyzed in the following. One can consider a curve crossing between two one-dimensional PESs  $V_1(x)$  and  $V_2(x)$ , as in Figure 2.4. In ARMD the dynamics in the crossing region evolves on a mixed PES:

$$V_{\text{ARMD}}(x, t) = (1 - f(t)) V_1(x) + f(t) V_2(x). \quad (2.64)$$

When the crossing between the two PESs is taking place (see Figure 2.5), the two 1D PESs cross at  $x = 0$ . The following analysis starts from a linear approximation of the PESs which is valid sufficiently close to the crossing point ( $V_1(x) = \alpha x$  and  $V_2(x) = \beta x$ , where  $\alpha > \beta$ ) For the switching function  $f(t)$ , a function which allows analytical work and is defined on the interval  $(0, t_s)$  is  $f(t) = t/t_s$ .

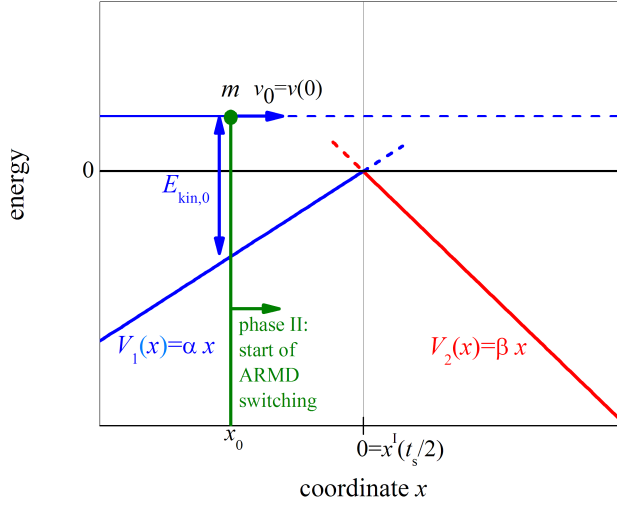
During the first phase (phase I, corresponding quantities are denoted with superscript I), we consider a system with effective mass  $m$  on  $V_1(x)$  (thus  $F^{\text{I}} = -\partial_x V^{\text{I}}(x, t) = -\alpha$ ) with initial position  $x_0 < 0$  and velocity  $v_0 > 0$  (i.e. regular MD, see Figure 2.5). At  $t = 0$   $x_0$  is chosen such that the system will reach the crossing point ( $x = 0$ ) exactly at half of the user-defined switching time ( $t_s/2$ ).

$$x^{\text{I}}(t_s/2) = 0 = x_0 + \int_0^{t_s/2} dt' \left( v_0 + \int_0^{t'} dt'' \frac{F^{\text{I}}}{m} \right) = x_0 + \frac{v_0}{2} t_s - \frac{\alpha}{8m} t_s^2 \quad (2.65)$$

According to ARMD, when crossing is detected, the time is rewound by  $t_s/2$  and the dynamics is restarted on the mixed PES (phase II). The ARMD potential and its partial derivatives are:

$$V^{\text{II}}(x, t) = \alpha x + \frac{\beta - \alpha}{t_s} x t, \quad (2.66)$$

$$\partial_t V^{\text{II}}(x, t) = \frac{\beta - \alpha}{t_s} x, \quad (2.67)$$



**Figure 2.5:** Simple model for estimating energy violation in ARMD simulations. The system (effective mass  $m$ ) is approaching from the left on PES  $V_1(x)$  (phase I). At  $t = 0$  time it is at  $x_0$  with velocity  $v_0$  and kinetic energy  $E_{\text{kin},0}$ . After crossing is detected at  $x = 0$  the time is rewound by  $t_s/2$  and the dynamics is re-simulated while  $V_1(x)$  is being switched to  $V_2(x)$  during  $t_s$  (phase II).

$$F^{\text{II}}(x, t) = -\partial_x V^{\text{II}}(x, t) = -\alpha + \frac{\alpha - \beta}{t_s} t. \quad (2.68)$$

The time evolution of the coordinates of the system during phase II follows from:

$$x^{\text{II}}(t) = x_0 + \int_0^t dt' \left( v_0 + \int_0^{t'} dt'' \frac{F^{\text{II}}(x(t''), t'')}{m} \right) = x_0 + \quad (2.69)$$

$$v_0 t - \frac{\alpha}{2m} t^2 + \frac{\alpha - \beta}{6m t_s} t^3.$$

The total energy change  $\Delta E^{\text{II}}$  during the crossing period is obtained by integration of  $dE^{\text{II}}/dt = \partial_t V^{\text{II}}(x, t)$ , defined in Equation 3.4, and subsequent transformation from Equation 3.2:

$$\Delta E^{\text{II}} = \int_0^{t_s} dt \partial_t V^{\text{II}}(x(t), t) = \frac{\beta - \alpha}{t_s} \int_0^{t_s} dt x(t) = \quad (2.70)$$

$$(\beta - \alpha) \left( x_0 + \frac{v_0}{2} t_s - \frac{3\alpha + \beta}{24m} t_s^2 \right) = \frac{\beta(\alpha - \beta)t_s^2}{24m}.$$

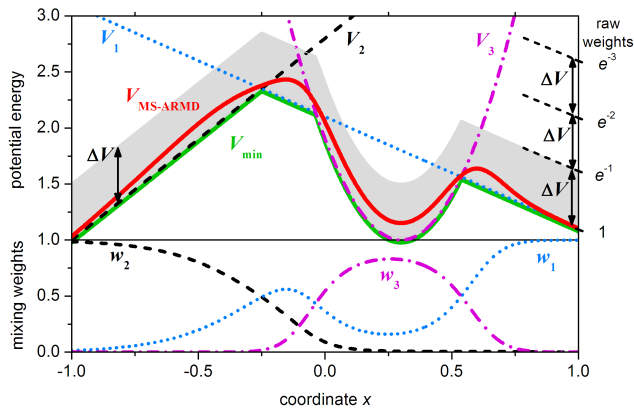
Hence, exact or nearly exact energy conservation,  $\Delta E^{\text{II}} \approx 0$ , can be achieved **a)** if the gradients of the two PESs are the same ( $\alpha \approx \beta$ ), **b)** if the system has a large effective mass. This is often the case in simulations of biomolecular systems, where either the moieties involved in the reaction are heavy or the reaction is coupled with rearrangement of solvation shell, which is composed of several solvent molecules, **c)** for the special case where the second surface has exactly zero or a negligible slope ( $\beta \approx 0$ ) in the crossing region then the positive and negative energy violations will cancel each other, **d)** if the switching time goes to zero. However, for  $t_s \rightarrow 0$  a smooth connection between the PESs is not possible and numerical integrators with finite step size will cause energy violation. As the switching time is the only variable parameter during ARMD simulations, it is useful to consider a criterion for its magnitude to guarantee energy conservation. For this purpose, one may require that energy violation compared to the initial kinetic energy ( $E_{\text{kin},0} = mv_0^2/2$ ) be small, i.e.  $|\Delta E^{\text{II}}| \ll E_{\text{kin},0}$ , which implies the switching time has to fulfill  $t_s \ll \sqrt{12}mv_0/\sqrt{|\beta(\alpha - \beta)|}$ .

#### 2.4.2.1 Multi-Surface ARMD

If formally exact energy conservation is required, switching between the PESs should be time-independent. In the following, we present a multi-surface (MS) variant of ARMD in which the weights ( $w_i(\mathbf{x})$ ) of the PESs ( $V_i(\mathbf{x})$ ) are also functions of all Cartesian coordinates  $\mathbf{x}$  and thereby the total energy can be conserved during crossing. According to MS-ARMD, the effective surface is always the lowest-energy surface ( $V_{\min}(\mathbf{x}) = \min_i V_i(\mathbf{x})$ ), except for those geometries where other surfaces get close to it in energy (within several  $\Delta V$ ), in which case the algorithm switches smoothly among them by changing their weights.  $\Delta V$  is a user defined parameter which regulates how the weights of the PESs involved in a crossing reduce with increasing energy separation from the lowest-energy surface at a given point  $\mathbf{x}$ . It plays a similar role as the

switching time  $t_s$  in ARMD. The optimum value of  $\Delta V$  is determined during parametrization of the global PES, which is usually done by fitting it to *ab initio* data. The effective potential energy is written as a linear combination of  $n$  PESs with weights  $w_i(\mathbf{x})$ .

$$V_{\text{MS-ARMD}}(\mathbf{x}) = \sum_{i=1}^n w_i(\mathbf{x}) V_i(\mathbf{x}) \quad \text{where} \quad w_i(\mathbf{x}) = \frac{w_{i,0}(\mathbf{x})}{\sum_{i=1}^n w_{i,0}(\mathbf{x})} \quad (2.71)$$

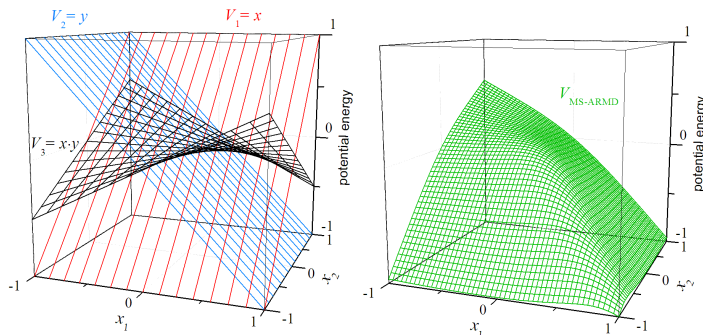


**Figure 2.6:** The MS-ARMD switching method in one dimension ( $x$ ) for three surfaces ( $V_1$ : blue dotted,  $V_2$ : black dashed,  $V_3$ : magenta dot-dashed) defined by linear ( $V_1$  and  $V_2$ ) and quadratic ( $V_3$ ) functions. The effective surface ( $V_{\text{MS-ARMD}}$ : red solid) always follows the lowest-energy surface ( $V_{\text{min}}$ : green solid), except for regions where several surfaces are close in energy (within a few times  $\Delta V = 0.5$  a.u.). Here, the algorithm switches smoothly among them by varying their weights ( $w_1$ ,  $w_2$ , and  $w_3$ ; see lower panel) as a function of  $x$ . Within the grayish band above the minimum energy surface the mixing weights of surfaces drop by a factor or  $e \approx 2.7$ . All quantities are in arbitrary units (a.u.).

The  $w_i(\mathbf{x})$  are obtained by renormalizing the raw weights  $w_{i,0}(\mathbf{x})$ , calculated by using a simple exponential decay function with a characteristic length of  $\Delta V$ .

$$w_{i,0}(\mathbf{x}) = \exp\left(-\frac{V_i(\mathbf{x}) - V_{\text{min}}(\mathbf{x})}{\Delta V}\right) \quad (2.72)$$

Using then the normalized weights, the global smoothed PES can be computed according to Equation 2.71. The capabilities of MS-ARMD are illustrated



**Figure 2.7:** The MS-ARMD switching method in two dimensions ( $x_1$  and  $x_2$ ) for three surfaces ( $V_1$ ,  $V_2$  and  $V_3$ ) defined by simple functions. The scales of axes are in arbitrary units (a.u.). The  $V_{\text{MS-ARMD}}$  surface is always the lowest-energy surface ( $V_{\text{min}}$ ), except for the region where more surfaces have similarly low energy (within a few times  $\Delta V = 0.2$  a.u.), here the algorithm switches smoothly among them as a function of the coordinates.

for surfaces in one and two dimensions in Figures 2.6 and 2.7, respectively. Smooth global surfaces are obtained in all cases. The MS-ARMD surface always corresponds to the lowest PES except for regions where other surfaces are close in energy where they mix. Furthermore, the algorithm also leads to a smooth MS-ARMD surface in regions where more than two surfaces are close in energy, as it is shown in both figures.

Equation 3.9 implies that the weights of surfaces compared to the weight of the momentary lowest-energy surface drop exponentially with increasing energy, that is by a factor of  $e \approx 2.7$  at every  $\Delta V$ . Therefore only those surfaces will have significant weights, whose energy is within a few times  $m$  of  $\Delta V$  (i.e.  $V_i(\mathbf{x}) - V_{\text{min}}(\mathbf{x}) < m\Delta V$ ) from the lowest energy surface  $V_{\text{min}}(\mathbf{x})$ . The term  $V_{\text{min}}(\mathbf{x})$  is necessary to avoid possible overflow and underflow in numerical implementation of the algorithm when the raw weights of the lowest energy surface is calculated. This shifting term will cancel from the expressions of weights and the MS-ARMD potential energy (Equation 2.71). Opposed to the raw weights, the final weights are analytical functions (infinitely differentiable) of the surface energies, therefore the resulting MS-ARMD surface ( $V_{\text{MS-ARMD}}(\mathbf{x})$ ) is a smooth function of the coordinates ( $\mathbf{x}$ ).

The smooth MS-ARMD PES can be used both in Molecular Dynamics (MD) and Monte Carlo (MC) simulations. For MD simulations, the first derivative of

the PES with respect to the coordinates  $\mathbf{x}$  is required and includes derivatives of the weighting functions as well:

$$\nabla V_{\text{MS-ARMD}}(\mathbf{x}) = \sum_{i=1}^n w_i(\mathbf{x}) \cdot \nabla V_i(\mathbf{x}) + \nabla w_i(\mathbf{x}) \cdot V_i(\mathbf{x}) \quad (2.73)$$

The new algorithm can be applied to the simultaneous switching of any number of PESs. Thus, it is possible to follow reactions that involve different pathways, such as intramolecular proton transfer and water-elimination in  $\text{H}_2\text{SO}_4$  which is discussed below.

---

### 2.4.3 ReaxFF

---

ReaxFF was developed to bridge the gap between EC and FF based computational chemical methods. Where EC methods are, applicable to all chemical systems, regardless of connectivity, their computational cost makes them inapplicable for large systems. Their computational expense also makes EC methods primarily applicable for single point or local energy minimization; high-temperature molecular dynamics (MD) simulations are generally too time-consuming. The method is based on bond-order-dependent potential, in which the Van der Waals and Coulomb forces are including from the beginning and dissociation and reaction curves are derived fro EC.<sup>67</sup>

---

#### 2.4.3.1 The Force Field in ReaxFF

---

The FF in ReaxFF divides the system energy up into various partial energy terms as follows:

$$E_{\text{system}} = E_{\text{bond}} + E_{\text{over}} + E_{\text{under}} + E_{\text{val}} + E_{\text{pen}} + E_{\text{tors}} + E_{\text{conj}} + E_{\text{vdWaals}} + E_{\text{Coulomb}}, \quad (2.74)$$

all this terms will be explained in the next subsection.

**Bond order and energy** The main assumption in ReaxxFF is that the bond order  $BO'_{ij}$  between a pair of atoms can be obtained from the interatomic distance  $r_{ij}$ , following the next expression:

$$BO'_{ij} = \exp \left[ p_{\text{bo},1} \cdot \left( \frac{r_{ij}}{r_0} \right)^{p_{\text{bo},2}} \right] + \exp \left[ p_{\text{bo},3} \cdot \left( \frac{r_{ij}^\pi}{r_0} \right)^{p_{\text{bo},4}} \right] + \exp \left[ p_{\text{bo},5} \cdot \left( \frac{r_{ij}^{\pi\pi}}{r_0} \right)^{p_{\text{bo},6}} \right]. \quad (2.75)$$

The equation 2.75 consist of three exponential terms; i) the sigma bond ( $p_{\text{bo},1}$  and  $p_{\text{bo},2}$ ) which is unity below  $\sim 1.5$  Å, but become negligible above  $\sim 2.5$  Å. ii) the first po bond ( $p_{\text{bo},3}$  and  $p_{\text{bo},4}$ ) which is unity below  $\sim 1.2$  Å and negligible above  $\sim 1.75$  Å, and iii) the second pi bond ( $p_{\text{bo},5}$  and  $p_{\text{bo},6}$ ), which is unity below  $\sim 1.0$  Å and negligible above  $\sim 1.4$  Å

This leads to a carbon-carbon bond with a maximum bond order of 3. For carbon-hydrogen and hydrogen-hydrogen bonds, only the sigma-bond contribution is considered, leading to a maximum bond order of 1.<sup>67</sup> The bond order need to be corrected in case of overcoordination and for residual 1~3 bond orders in valence angles using the scheme described in the angle term part. This description is applied to all bonds in the molecule, while the overcoordination correction is only to bonds that contain carbon atoms.<sup>67</sup> The overcoordination correction can be summarized in the next set of equations:

$$BO_{ij} = BO'_{ij} \cdot f_1(\Delta'_i, \Delta'_j) \cdot f_4(\Delta'_i, BO'_{ij}) \cdot f_5(\Delta'_j, BO'_{ij}) \quad (2.76)$$

$$f_1(\Delta_i, \Delta_j) = \frac{1}{2} \cdot$$

$$\left( \frac{\text{Val}_i + f_2(\Delta'_i, \Delta'_j)}{\text{Val}_i + f_2(\Delta'_i, \Delta'_j) + f_3(\Delta'_i, \Delta'_j)} + \frac{\text{Val}_j + f_2(\Delta'_i, \Delta'_j)}{\text{Val}_j + f_2(\Delta'_i, \Delta'_j) + f_3(\Delta'_i, \Delta'_j)} \right) \quad (2.77)$$

$$f_2(\Delta'_i, \Delta'_j) = \exp(-\lambda_1 \cdot \Delta'_i) + \exp(-\lambda_1 \cdot \Delta'_j) \quad (2.78)$$

$$f_3(\Delta'_i, \Delta'_j) = \frac{1}{\lambda_2} \cdot \ln \left\{ \frac{1}{2} \cdot [\exp(-\lambda_2 \cdot \Delta'_i) + \exp(-\lambda_2 \cdot \Delta'_j)] \right\} \quad (2.79)$$

$$f_4(\Delta'_i, BO'_{ij}) = \frac{1}{1 + \exp(-\lambda_3 \cdot (\lambda_4 \cdot BO'_{ij} \cdot BO'_{ij} - \Delta'_i) + \lambda_5)} \quad (2.80)$$

$$f_5(\Delta'_j, BO'_{ij}) = \frac{1}{1 + \exp(-\lambda_3 \cdot (\lambda_4 \cdot BO'_{ij} \cdot BO'_{ij} - \Delta'_j) + \lambda_5)}, \quad (2.81)$$

where  $\text{Val}_i$  is the valency of atom  $i$ .  $\Delta'_i$  is the degree of deviation of the sum of the uncorrected bond orders around an atomic center from its valency  $\text{Val}_i$ , described by the next expression:

$$\Delta'_i = \sum_{j=1}^{nbond} \text{BO}'_{ij} - \text{Val}_i. \quad (2.82)$$

In order to compute the energy of the bond the next potential is used:

$$E_{\text{bond}} = -D_e \cdot \text{BO}_{ij} \cdot \exp[p_{\text{be},1}(1 - \text{BO}_{ij}^{p_{\text{be},1}})] \quad (2.83)$$

**Angles** For angles following the same rules for bonds, it is important that the energy contribution from valence angles terms goes to zero. In order to reach this, the expression that represent the valence angle in ReaxFF is summarized in the next set of equations:

$$E_{\text{val}} = f_7(\text{BO}_{ij}) \cdot f_7(\text{BO}_{ik}) \cdot f_8(\Delta_j) \cdot \{k_a - k_a \exp[-k_b(\Theta_0 - \Theta_{ijk})^2]\} \quad (2.84)$$

$$f_7(\text{BO}_{ij}) = 1 - \exp(-\lambda_{11} \cdot \text{BO}_{ij}^{\lambda_{12}}) \quad (2.85)$$

$$f_8(\Delta_j) = \frac{2 + \exp(-\lambda_{13} \cdot \Delta_j)}{1 + \exp(-\lambda_{13} \cdot \Delta_j) + \exp(-p_{v,1} \cdot \Delta_j)} \cdot \left[ \lambda_{14} - (\lambda_{14} - 1) \cdot \frac{2 + \exp(\lambda_{15} \cdot \Delta_j)}{1 + \exp(\lambda_{15} \cdot \Delta_j) + \exp(-p_{v,2} \cdot \Delta_j)} \right] \quad (2.86)$$

$$\text{SBO} = \Delta_j - 2 \cdot \left\{ 1 - \exp \left[ -5 \cdot \left( \frac{1}{2} \Delta_j \right)^{\lambda_{16}} \right] \right\} + \sum_{n=1}^{\text{neighbors}(j)} \text{BO}_{jn,\pi} \quad (2.87)$$

$$\Delta_{j,2} = \Delta_j \text{ if } \Delta_j < 0$$

$$\Delta_{j,2} = 0 \text{ if } \Delta_j \geq 0$$

$$\text{SBO2} = 0 \text{ if } \text{SBO} \leq 0$$

$$\text{SBO2} = \text{SBO}^{\lambda_{17}} \text{ if } 0 < \text{SBO} < 1 \quad (2.88)$$

$$\text{SBO2} = 2 - (2 - \text{SBO})^{\lambda_{17}} \text{ if } 1 < \text{SBO} < 2$$

$$\text{SBO2} = 2 \text{ if } \text{SBO} > 2$$

$$\Theta_0 = \pi - \Theta_{0,0} \cdot \{1 - \exp[-\lambda_{18} \cdot (2 - \text{SBO2})]\}.$$

ReaxFF use bond-order-dependent form to calculate the energy associate with deviation in valence angle  $\Theta_{ijk}$  from the equilibrium value  $\Theta_0$ . The  $f_7$  term



in equation 2.4.3.1, makes the valence angle energy contribution disappear smoothly during bond dissociation.<sup>67</sup> The equation 2.4.3.1 has implicit the effect of over/undercoordination in central atom  $j$  on the valence angle energy, The equilibrium angle  $\Theta_0$  for  $\Theta_{ijk}$  depends on the sum of  $\pi$ -bond orders (SBO) around the central atom  $j$ . In that vein, the equilibrium angle changes from around  $109.47^\circ$  for  $sp^3$  hybridization ( $\pi$ -bond = 0) to  $120^\circ$  for  $sp^2$  ( $\pi$ -bond = 2) based on the geometry of the central atom  $j$  and its neighbors. In addition the effects of  $\pi$ -bonds on the central atom  $j$ , needs to take into account the effects of over and undercoordination in central atom  $j$  ( $\Delta_j$ ) on the equilibrium valency angle, including the influence of a lone electron pair. The functional form in the equation 2.4.3.1 was created in ReaxFF to avoid singularities when  $SBO = 0$  and  $SBO = 2$ .<sup>67</sup>

**Torsion angles** As with the angle potential, there is a dependence of the torsion angle energy which accounts properly for  $BO \rightarrow 0$  and for  $BO$  greater than 1. As for angles and bonds a set of equation represent the torsional angle potential to follow the above rules, these equation are:

$$E_{\text{tors}} = f_{10}(BO_{ij}, BO_{jk}, BO_{kl}) \cdot \sin \Theta_{jjk} \cdot \sin \Theta_{jkl} \cdot \left[ \frac{1}{2} V_2 \cdot \exp\{p_l(BO_{jk} - 3 + f_{11}(\Delta_j, \Delta_k))^2\} \cdot (1 - \cos 2\omega_{ijkl}) + \frac{1}{2} V_3 \cdot (1 + \cos 3\omega_{ijkl}) \right] \quad (2.89)$$

$$f_{10}(BO_{ij}, BO_{jk}, BO_{kl}, ) = [1 - \exp(-\lambda_{23} \cdot BO_{ij})] \cdot [1 - \exp(-\lambda_{23} \cdot BO_{jk})] \cdot [1 - \exp(-\lambda_{23} \cdot BO_{kl})] \quad (2.90)$$

$$f_{11}(\Delta_j, \Delta_k) = \frac{2 + \exp[-\lambda_{24} \cdot (\Delta_j + \Delta_k)]}{1 + \exp[-\lambda_{24} \cdot (\Delta_j + \Delta_k)] + \exp[\lambda_{25} \cdot (\Delta_j + \Delta_k)]}. \quad (2.91)$$

The  $V_2$ -cosine term in the equation 2.91 depends on the bond order of the central bond  $BO_{jk}$ . For torsion angles with a central double bond ( $BO_{jk} = 2$ ), the  $V_2$ -term has a maximum.<sup>67</sup> If  $BO_{jk}$  deviates from 2 the magnitude of the  $V_2$ -term rapidly decrease its value. The valence-angle-dependent term  $\sin(\Theta_{jkl})$  in equation 2.91, ensures that the torsion energy contribution disappears when either of the two valence angles ( $\Theta_{ijk}$  or  $\Theta_{jkl}$ ) approaches to  $\pi$ .<sup>67</sup>

**Conjugated systems** ReaxFF also describes conjugated systems contribution in energy, when those are present in the molecular system. A maximum

contribution of conjugation energy is obtained, when successive bonds have bond-order values of 1.5 as in benzene and other aromatics. The expression which describes the energetic contribution of the conjugated systems are:

$$E_{\text{conj}} = f_{12}(\text{BO}_{ij}, \text{BO}_{jk}, \text{BO}_{kl}) \cdot \lambda_{26} \cdot [1 + (\cos^2 \omega_{ijkl} - 1) \cdot \sin \Theta_{ijk} \cdot \sin \Theta_{jkl}] \quad (2.92)$$

$$f_{12}(\text{BO}_{ij}, \text{BO}_{jk}, \text{BO}_{kl}) = \exp \left[ -\lambda_{27} \cdot \left( \text{BO}_{ij} - 1 \frac{1}{2} \right)^2 \right] \cdot \exp \left[ -\lambda_{27} \cdot \left( \text{BO}_{ij} - 1 \frac{1}{2} \right)^2 \right] \cdot \exp \left[ -\lambda_{27} \cdot \left( \text{BO}_{kl} - 1 \frac{1}{2} \right)^2 \right] \quad (2.93)$$

**Nonbonded and van der Waals interactions** In addition to valence interactions which depend on overlap, there are repulsive interactions at short interatomic distances due to Pauli principle orthogonalization and attraction energies at long distances due to dispersion.<sup>67</sup>

These interactions, comprised of van der Waals and Coulomb forces, are included for all atom pairs, thus avoiding awkward alterations in the energy description during bond dissociation. In this respect, ReaxFF is similar in spirit to the central valence force fields that were used earlier in vibrational spectroscopy. To account for the van der Waals interactions ReaxFF uses a distance-corrected Morse-potential.<sup>67</sup> By including a shielded interaction, excessively high repulsions between bonded atoms and atoms sharing a valence angle are avoided.<sup>67</sup> The potential for van der Waals interaction has the next expression:

$$E_{vdWaal} = D_{ij} \cdot \left\{ \exp \left[ \alpha_{ij} \cdot \left( 1 - \frac{f_{13}(r_{ij})}{r_{vdW}} \right) \right] - 2 \cdot \exp \left[ \frac{1}{2} \cdot \alpha \cdot \left( 1 - \frac{f_{13}(r_{ij})}{r_{vdW}} \right) \right] \right\} \quad (2.94)$$

$$f_{13}(r_{ij}) = \left[ f_{ij}^{\lambda_{29}} + \left( \frac{1}{\lambda_w} \right)^{\lambda_{28}} \right]^{1/\lambda_{28}} \quad (2.95)$$

**Coulomb interactions** As other FF Coulomb interactions are taken into account between all atom pairs. In order to adjust the orbital overlap between atoms at close distances, ReaxFF apply a shielded Coulomb potential:

$$E_{Coulomb} = C \cdot \frac{q_i \cdot q_j}{[r_{ij}^3 + (1/\gamma_{ij})^3]^{1/3}} \quad (2.96)$$

Atomic charges are obtained with Electron Equilibration Method (EEM) approach. The  $\gamma_{ij}$  parameter can be optimized to reproduce the QEq orbital overlap correction.<sup>67</sup>

#### 2.4.3.2 Applicability

ReaxFF aims to be as general as possible and has been parameterized and tested for hydrocarbon reactions, transition-metal-catalyzed nanotube formation, and high-energy materials. Recently, ReaxFF is developed to study oxygen interactions with realistic silica surfaces. This version of ReaxFF is based on highly accurate and benchmarking density functional studies. Highly accurate density functional results are achieved by employing Minnesota functionals.<sup>68-74</sup>

## 2.5 Rare Events Techniques

We are living in a great computation age, where the ideas and scientific concepts from the last century, have been applied and converted in very efficient software packages, which makes the atomistic simulations of complex systems relatively easy to perform, which allows one to investigate interesting phenomena in several fields like solid state physics, biophysics and chemistry. However, in several important cases, simulations are still not competitive, when is compared with more empirical methods. This mainly because using atomistic models for long or big systems, is computationally expensive, even with current resources. The time scales that can be currently simulated are in the range of  $\mu\text{s}$  for classical MD and in the range of hundreds of picoseconds for *ab initio* MD (ABMD). Those are small time intervals for most of the exciting natural phenomena, which are in the orders of magnitude larger. Those events that are sampled in the phase space very rarely are consider as **rare events**. Some examples of rare events are chemical reactions, structural phase transitions, protein folding, protein-protein interactions, protein complex rearrangements and molecular recognition. Probably in a near future, and because the efficiency

of computation is increasing every day more, is feasible that will be possible to observe, by direct simulation, the reactive trajectories of more complex system. For the moment, moreover, observing one or few reactive trajectories might not be sufficient for converging statistical averages and computing observables that can be compared with experiments. Now, for massive systems like protein complexes, where millions of atoms are interacting, one can renounce the all-atom description and use instead a coarse-grained model. This would require a detailed knowledge of the systems that is often unavailable.

All the above mentioned drawbacks can be overcome using a methodology called *accelerating rare events* allowing us to simulate some awesome and interesting phenomena with the available computer resources. These methodologies can be classified in four categories:

- 1). Methods in which the phase space is explored simultaneously at different values of temperature, such as parallel tempering<sup>75</sup> and replica exchange,<sup>76</sup> or as a function of the potential energy, such as multicanonical MD<sup>77</sup> and Wang-Landau.<sup>78</sup> These methods exploit more or less explicitly the potential energy as a generalized collective variables (CV). In several cases, ordered and disordered states may correspond to the same value of potential energy, or be present in the thermal ensemble at the same temperature. This may lead to hysteresis and convergence problems.<sup>79</sup>
- 2). Methods which explore the transition mechanism and constructing reactive trajectories,<sup>80</sup> examples of these methods are: transition path sampling,<sup>81–83</sup> transition interface sampling,<sup>84</sup> nudged elastic band,<sup>22</sup> finite-temperature string method,<sup>24,85</sup> milestoning<sup>86</sup> and forward flux method.<sup>87</sup> This kind of method requires knowledge of the initial and final state of the process that has to be simulated. For example, in order to observe the structural changes during the activation of a protein, like phosphorylation-dependent protein,<sup>7,88–91</sup> which is in the range of hundreds of nanoseconds, one can apply the method if the experimental structure for the inactive (non-phosphorylated) and active (phosphorylated) are available, or any kind of putative or hypothetical protein are available.
- 3). Methods that construct the probability distribution or enhance the sampling as a function of one or few predefined CVs. In chemical reactions, for example, where bonds are forming and breaking, a natural CV can be a distance between atoms that are covalently bonded. Examples of these methods include steered MD,<sup>92</sup> free energy perturbation,<sup>93</sup> conformational flooding,<sup>94</sup> weighted

histograms techniques,<sup>95,96</sup> Jarzynski's identity-based methods,<sup>97,98</sup> thermodynamic integration,<sup>99,100</sup> adaptive force bias,<sup>101,102</sup> umbrella sampling<sup>103</sup> and adiabatic molecular dynamics.<sup>104</sup> These approaches require that the CVs describe very well the molecular system, which make the choice of the CVs a critical point in this kind of simulation. In other words if the CVs does not describe well, for example, the reaction coordinate or a set of reaction coordinates, the calculation will be describing something that is not real, or simply, the calculation, does not converge.

4). Methods exploring the PES and localizing the saddle points that correspond to the transition state such as eigenvalues following,<sup>105</sup> the dimer method,<sup>26</sup> hyperdynamics,<sup>106</sup> and multiple-time scale accelerated MD.<sup>107</sup> These approaches are extremely powerful for exploring the PES of low dimensionality, but their reliability degrades with the complexity of the system. Indeed, for very large or complex systems the number of possible transition states surrounding a minimum becomes rapidly too large for a deterministic search. The problem for this methods is that in solvated systems the concept of saddle point on the PES becomes fuzzy, and these approaches cannot easily be applied.

Here I will introduce two of these techniques; umbrella sampling<sup>103</sup> and metadynamics,<sup>108</sup> which were used for some of the simulations all along my PhD studies.

### 2.5.1 Umbrella Sampling

In order to understand reactivity, one needs to know the free energy surface of the process. Computing free energies requires sampling of the elemental PES to generate ensembles. In equilibrium, the free energy difference  $\Delta G_{A \rightarrow B}$  between the reactant state  $A$  and the product state  $B$ , is determined by their populations  $p$ :

$$\Delta G_{A \rightarrow B} = -k_B T \ln \frac{p_B}{p_A}, \quad (2.97)$$

where  $k_B$  is the Boltzmann constant, and  $T$  the temperature. However, for chemical reactions, the barriers separating the state  $A$  and  $B$  are high, and transitions are considered as rare events. Therefore, it is not likely that both states  $A$  and  $B$  are sampled sufficiently in a normal MD simulation to provide a reasonable estimate for  $\Delta G_{A \rightarrow B}$ .

To bias the sampling toward a region of interest that would not otherwise be significantly populated, a restraining potential  $U_r(q)$  is added to the potential energy of the system.  $U_r$  is often referred to as an umbrella potential.<sup>103</sup> For concreteness, we assume the harmonic form of the potential

$$U_r(q; \lambda) = k_h[q - q_0(\lambda)]^2 \quad (2.98)$$

where  $k_h$  is a force constant,  $q_0(\lambda)$  is a target value of  $q$ , and  $\lambda$  a coupling parameter. However, umbrella potentials are by no means limited to a harmonic form. The reaction coordinate  $q$  could be a dihedral angle, a distance between two selected atoms, or a more complicated, collective degree of freedom.  $q_0(\lambda)$  is constructed so that as the coupling coordinate  $\lambda$  varies,  $q_0(\lambda)$  traverses the region or regions of interest.

The free energy difference between two stable conformations can be obtained by a thermodynamic integration approach.<sup>109</sup> Let  $q_A$  and  $q_B$  represent the centers of the two corresponding energy wells. The free energy derivative is seen to be

$$\frac{\partial F}{\partial \lambda}(\lambda) = -2k_h(\langle q \rangle_\lambda - q_0(\lambda)) \frac{dq_0}{d\lambda}(\lambda) \quad (2.99)$$

which can be obtained from a simulation with the restraining potential  $U_r(q; \lambda)$ . Equation 2.99 is a integrating between  $q_A$  and  $q_B$  gives the free energy difference between the two wells, but with restraints present at each endpoint. Additional steps are needed in which the restraints are removed at the endpoints. The corresponding free energies can be obtained from the thermodynamic perturbation formula.<sup>39</sup>

This approach is easily generalized to nonharmonic restraint terms and to cases where several restraint terms are used. Thus, if restraints are applied to several dihedral angles or several interatomic distances, each will contribute a term of the Equation 2.99 to the free energy derivative.

Application of Equation 2.97 corrects the free energies of the endpoints but not those of the intermediate conformations. Therefore, the above approach yields a free energy profile between  $q_A$  and  $q_B$  that is altered by the restraint(s). In particular, the barrier height is not that of the natural, unrestrained system. It is possible to correct the probability distributions  $P_r$  observed all along the pathway (with restraints) to obtain those of the unrestrained system.<sup>39</sup> From the relation  $P_{(q)Z_{ur}} = P_r(q)Z_r \exp(U_r/kT)$ , this can be written as

$$P(q) = P_r(q)e^{U_r(q)/kT} / \langle e^{-U_r/kT} \rangle_r \quad (2.100)$$

$$= P_r(q)e^{U_r(q)/kt} \langle e^{-U_r/kT} \rangle_{ur} \quad (2.101)$$

where the subscripts  $r$  and  $ur$  refer to the restrained and unrestrained systems, respectively. The resulting  $P(q)$  is expected to be accurate only if there is a large overlap between probable conformations of the unrestrained system and conformations where the restraint energy is small. Thus in practice,  $q_0(\lambda)$  must be close to a stable energy minimum of the unrestrained system, and  $P(q)$  will be accurate only close to  $q_0(\lambda)$ . To obtain  $P(q)$  over a broader range, a series of umbrella potentials are required, covering a range of  $q_0$  values. Let  $U'_r$  be a second umbrella potential, corresponding to a  $q_0(\lambda')$  slightly displaced relative to  $q_0(\lambda)$ . One can show that<sup>110</sup>

$$P(q) = P_{r'}(q)e^{U_{r'}(q)/kT} / (\langle e^{U_r/kT} \rangle_r \langle e^{(U_{r'}-U_r)/kT} \rangle_{r'}) \quad (2.102)$$

This formula is expected to be accurate close to  $q_0(\lambda')$ . Continuing in this manner, one can obtain accurate formulas for  $P(q)$  over a broad range of  $q$ , provided the regions sampled with the successive umbrellas overlap. For many problems, the ideal umbrella potential would be one that completely flattens the free energy profile along  $q$ , i.e.,  $U_r(q) = -W(q)$ . Such a potential cannot be determined in advance. However, iterative approaches exist that are known as adaptive umbrella sampling [41]. Such approaches are especially important for large-scale sampling of many very different conformations, such as folded and unfolded conformations of a protein or peptide.

### 2.5.2 Metadynamics

Metadynamics<sup>108</sup> can be seen as a method, that can bring together many features from the previously described techniques to compute free energies and accelerate rare events. The algorithm is based on the work of Kevrekidis,<sup>111</sup> and requires the preliminary identification of a CV set which are able to describe the process of interest. One of the most interesting features of the metadynamics algorithm, is the possibility to treat several CVs simultaneously. In contrast with other methods, the algorithm can be proficiently used for

both, reconstructing the free energy and accelerating rare events. Enhanced by a history dependent potential, constructed as a sum of Gaussians, which are centered along the trajectory followed by the CVs. Similar approach is also found in the Wang and Landau algorithm,<sup>78</sup> adaptive force bias<sup>101</sup> and self-healing umbrella sampling.<sup>112</sup>

Just because I like the Italian analogies, to explain this method, that is used from one of the algorithm developers, I will use his "qualitative example":<sup>113</sup> Now, Imagine a walker who, during the night, falls into an empty swimming pool. The swimming pool wall is too steep for the walker to climb, and the complete darkness makes it difficult for him to localize the shallowest point (lowest saddle). In these conditions the walker will move more easily downhill, and it is rather unlikely that he will find by chance the lowest saddle. His walk in these conditions resembles that performed by microscopic systems in normal MD or Monte Carlo: a random walk with a bias in the direction of lower free energy, with a very small probability to explore transition regions (climb out of the pool). In metadynamics, the walker has access to a large source of sand that he can deposit in his current position. The sand will slowly fill the pool. Thus, even if at the beginning he visits more often the bottom region of the pool, little by little he fills the pool with sand (Gaussians), and he almost deterministically starts exploring regions that are higher and higher. Sooner or later, the walker is destined to fill sufficiently the pool to be able to climb out of it. And most probably he will climb out from the shallowest point of the pool. The novel idea that differentiates metadynamics from pre-existing methods is that if the walker is able to keep memory of all the positions in which he has deposited sand (Gaussians), he will be able to reconstruct a negative image of the underlying pool (the free energy).<sup>113</sup>

In this respect, a system which is in a free energy minimum, and is biased with metadynamics, can escape from that well, following the minimum energy path *via* a saddle point. This characteristic can bring the possibility to explore new reaction pathways, and of course accelerate the observation of rare events. It sounds like the ultimate method for accelerating rare events, but, remember this algorithm depends on a good selection of the CVs that describe the event, if few or not information about which CV describes a determined process, could be not possible to use this technique.

### The algorithm

A molecular system can be described by a set of coordinates  $x$  and a potential  $V(x)$  evolving in a MD, if one is interested in exploring the properties of a



system as a function of a finite number of CV  $S_\alpha(x)$ ,  $\alpha = 1, d$  (where  $d$  is a small number). The CV can be any explicit function of  $x$  such as an angle, a distance, a coordination number, etc. The equilibrium behavior of these variables is defined by a probability distribution of the form

$$P(s) = \frac{e^{-(1/T)F(s)}}{\int ds e^{-(1/T)F(s)}}, \quad (2.103)$$

Where  $s$  denotes the  $d$  dimensional method  $(s_1, \dots, s_d)$  and the free energy  $F(s)$  is given by

$$F(s) = -T \ln \left[ \int dx e^{(-1/T(V(x)))} \delta(s - S(x)) \right]. \quad (2.104)$$

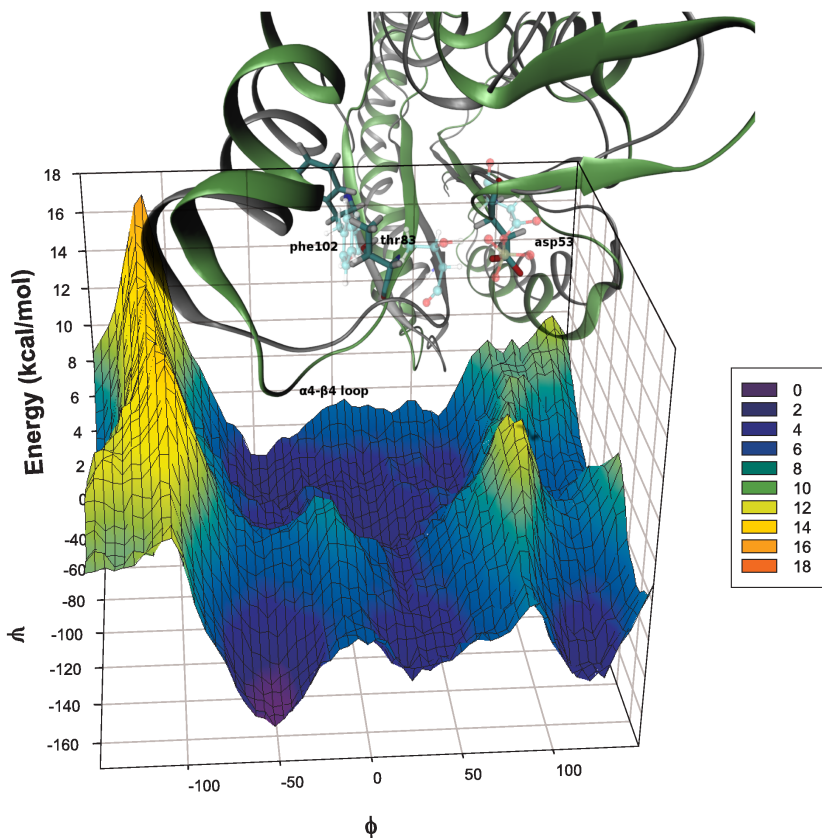
In metadynamics, the free energy is reconstructed recursively, starting from the bottom of the well by a history dependent random walk that explores a large portion of the configuration space. In the simplest MD implementation of the algorithm the mathematical equivalent of the sand deposited by the walker, in the example of the introduction, is a small repulsive Gaussian potential added every  $\tau_G$  MD steps. This manner of biasing the dynamics is usually referred to as "direct metadynamics". The external (metadynamics) potential acting on the system at time  $t$  is given by

$$V_G(S(x), t) = w \sum_{\substack{t'=\tau_G, 2\tau_G, \dots \\ t' < t}} e^{\left( -\frac{(S(x) - s(t'))^2}{2\delta s^2} \right)}, \quad (2.105)$$

where  $s(t) = S(x(t))$  is the value taken by the CV at time  $t$ ,  $w$  is the gaussian height,  $\delta s$  is the Gaussian width and  $\tau_G$  is the frequency at which Gaussians are added.

$w$ ,  $\delta s$  and  $\tau_G$  influence the accuracy and efficiency of the free energy reconstruction. If the Gaussians are large, the free energy surface will be explored a fast space, but the reconstructed profile will be affected by large errors. Instead, If the Gaussians are small or are placed infrequently the reconstructions will be accurate, but it will take a longer time.

Figure 2.8 shows an example of metadynamics in action. The simulation was done with NAMD<sup>114</sup> using the Collective Variable package. The molecular system that was used is the diguanylate cyclase PleD protein from *Caulobacter*



**Figure 2.8:** free energy surface in the space of pseudo-dihedral angle  $\phi$  and dihedral angle  $\psi$ . Energy is given in kcal/mol.

*crescentus* which belongs to response regulator family, and its activity is controlled by phosphorylation (See paper 4). Here we applied metadynamics over two CVs; dihedral angles  $\phi$  and  $\psi$ , which are essential for PleD dimerization (see paper 4). Briefly, the protein is very stable after the phosphorylation, and the two CV were chosen can explain why this stability happens(see paper 4). Figure 2.8 shows the free energy surface was reconstructed using a Gaussian height  $w$  of 0.01 kcal/mol, a GAussian width  $\delta s$  of  $5.0^\circ$  and frequency  $\tau_G$  of 100 steps. Details about this particular 3D plot can be found in (see paper 4).

## 2.6 Quantum Mechanics/Molecular Mechanics

From the great work of the Nobel Laureates Warshel and Levitt,<sup>115</sup> the hybrid quantum mechanics/molecular mechanics (QM/MM) combined methods, help to understand some molecular systems, chemical reaction mechanisms and other relevant processes where electronic influence of the system is important and cannot be treated only with MM methods (remember MM methods does not include electronic effects). In the combined QM/MM method, the system to be studied is primarily partitioned into a quantum mechanical (QM) part and a molecular mechanical (MM) part.<sup>116</sup> The QM part corresponds to what is to be studied electronically. This part may be a molecule (such as a solute molecule in a solution) or several molecules, a fragment (or part) of a large molecule. Atoms in this part are explicitly expressed as electrons and nuclei. The motion of the electrons are described quantum mechanically. When a combined QM/MM method is used to study a system involving charge transfer, electron excitation or chemical reactions, the corresponding region is always treated quantum mechanically. That is, the region is always included in the QM part. The MM part is the environment to the QM part. The part is non-reactive and is treated by using classical MM FFs. nonreactive also means that there is no charge transfer or other chemical exchange between the QM and MM parts.

In the combined QM/MM approach, the conventional methods to treat the boundary part can be introduced directly. One can assume that the system is divided into a QM and an MM part, without any specific boundary. If one has this in mind and according to this model, the effective Hamiltonian of the system can be written as:

$$\hat{H}^{eff} = \hat{H}_{QM}^0 + \hat{H}_{MM} + \hat{H}_{QM/MM} \quad (2.106)$$

and the corresponding Schrödinger equation

$$\hat{H}^{eff}|\Psi\rangle = E_{tot}|\Psi\rangle, \quad (2.107)$$

where  $E_{tot}$  is the total energy of the system. The term  $\hat{H}_{QM}^0$  in Equation 2.106 is the Hamiltonian of the isolated QM part and can be expressed as:

$$\begin{aligned} \hat{H}_{QM}^0 = & -\frac{1}{2} \sum_i \nabla_i^2 - \sum_i \sum_{\alpha} \alpha \frac{Z_{\alpha}}{r_{i\alpha}} + \frac{1}{2} \sum_i \sum_{j \neq i} \frac{1}{r_{ij}} \\ & + \sum_{\alpha} \sum_{\beta > \alpha} \frac{Z_{\alpha} Z_{\beta}}{R_{\alpha\beta}}, \end{aligned} \quad (2.108)$$

where  $i, j$  and  $\alpha, \beta$  are the electrons and the nuclei, respectively.  $r$  is the electron-electron or electron-nucleus distance.  $R_{\alpha\beta}$  is the distance between the nuclei  $\alpha$  and  $\beta$ .  $Z_{\alpha}$  is the charge of the nucleus  $\alpha$ .

The term  $\hat{H}_{MM}$  corresponds to the Hamiltonian of the isolated MM part. It is treated by regular MM force fields and corresponds to the MM energy of the isolated MM part,

$$\hat{H}_{MM} = E_{MM} \quad (2.109)$$

The last term in Equation 2.106 ( $\hat{H}_{QM/MM}$ ), is the interaction between the QM atoms and the MM atoms, which mainly refers to the coupling between the QM and MM parts. Because the QM/MM Hamiltonian involves the coordinates of the QM electrons, it is involved in the QM calculation. In that sense, the QM calculation can be written according to the following Schrödinger equation:

$$(\hat{H}_{QM}^0 + \hat{H}_{QM/MM})|\Psi_{QM}\rangle = (E_{QM} + E_{QM/MM})|\Psi_{QM}\rangle, \quad (2.110)$$

where  $|\Psi_{QM}\rangle$  is the state of the QM electrons. The total energy  $E_{total}$ , then can be written as

$$E_{tot} = E_{QM} + E_{MM} + E_{QM/MM} \quad (2.111)$$

How the theoretical coupling between the QM and MM parts is established depends very much on how the atoms on the MM parts are represented. In the MM FFs, the intermolecular interactions are most often described by the electrostatic and van der Waals interactions. The MM atoms are normally represented by point charges and Lennard-Jones parameters (usually centered on atoms) in the calculation of intermolecular interactions. The Hamiltonian can be expressed in the following form:

$$\begin{aligned}
\hat{H}_{QM/MM} = & - \sum_i \sum_m \frac{Q_m}{r_{im}} + \sum_\alpha \sum_m \frac{Z_\alpha Q_m}{R_{\alpha m}} \\
& + \sum_\alpha \sum_m 4\varepsilon_{\alpha m} \left[ \left( \frac{\sigma_{\alpha m}}{R_{\alpha m}} \right)^{12} - \left( \frac{\sigma_{\alpha m}}{R_{\alpha m}} \right)^6 \right]
\end{aligned}
\tag{2.112}$$

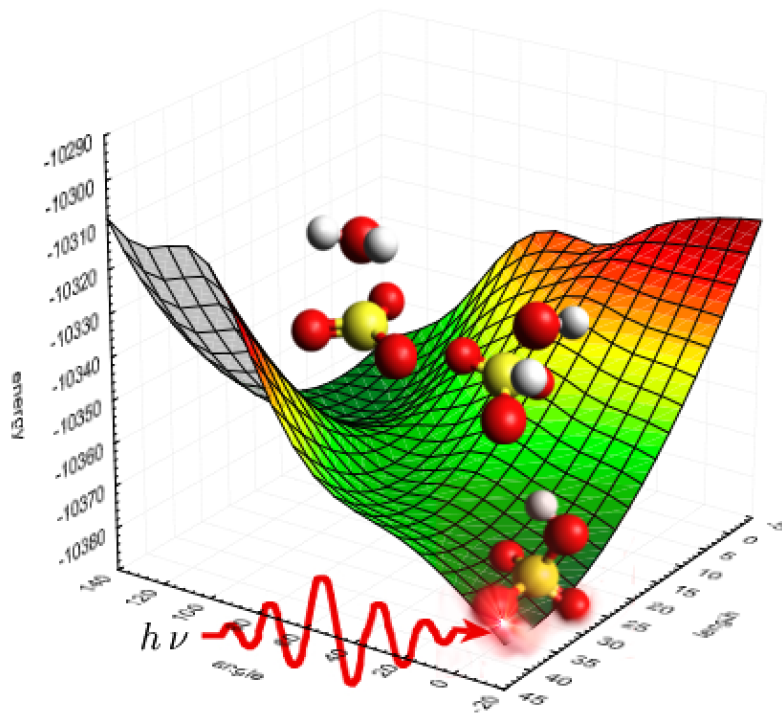
where  $i, \alpha$  denote the QM electrons and nuclei, respectively.  $Q_m$  is the point charge on the MM atom  $m$ . The first and second term on the right hand side, are the electrostatic interaction of the MM charges with QM electrons and nuclei, respectively. The last term on the right hand side of the equation, is the van der Waals interaction between the QM atoms and the MM atoms. It involves the corresponding exchange-repulsive and dispersion interactions between the two parts. Other coupling including; the polarization from the MM part and fluctuating charge model can be also added to the coupling part of the Equation 2.112. Many software packages nowadays provide this features in which the user can define the QM and MM part accordingly with their own choice. For example in CHARMM, exist at least 4 different schemes; 1) QM/MM using QChem, 2) QM/MM with SCC-DFTB, 3) QM/MM with Gaussian09 and 4) QM/MM with GAMESS.



# 3

## REACTION IN GAS PHASE

---



*“If you wish to make an apple pie from scratch, you must first invent the universe.”*

—Carl Sagan, *Cosmos*.

## Contents

---

1.1	Molecular Reaction Dynamics . . . . .	13
1.2	Enthalpies and Free Energies of Reaction . . . . .	14
1.3	Kinetics . . . . .	15
1.3.1	Arrhenius Equation and Activation Energy . . . . .	15
1.4	Collision Theory . . . . .	16
1.5	Saddle Points and Potential Energy Surface . . . . .	18
1.6	Transition State Theory . . . . .	21
1.7	Statistical Theories of Chemical Reactions . . . . .	23
1.7.1	Rice–Ramsperger–Kassel–Marcus RRKM Theory . . . . .	23

---



This chapter has 3 sections; section 1 describe the initial work on the vibrational induced dissociation of sulfuric acid using ARMD. Because the initial ARMD algorithm does not conserves energy, in section 2, a new algorithm namely MS-ARMD will be described, together with the test case, which is the same reaction we are describing in this chapter (sulfuric acid photodissociation). With the new MS-ARMD it is possible to take in to account various reaction channels, this new features allows one to simulate the most probable reactions of sulfuric acid (intramolecular H-transfer and water elimination). In section 3 we will describe the reaction of sulfuric acid, considering all the reactive channels, furthermore, with this new MS-ARMD algorithm the total energy of the system is conserved, which allows one to perform a final state analysis.

### **3.1** Vibrationally induced dissociation of H<sub>2</sub>SO<sub>4</sub>

---

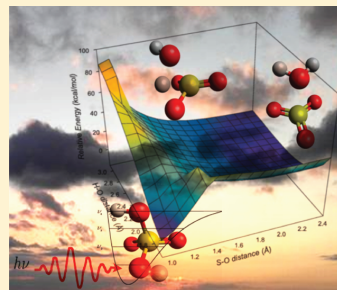
Vibrationally Induced Dissociation of Sulfuric Acid (H<sub>2</sub>SO<sub>4</sub>)

Juvenal Yosa and Markus Meuwly\*

Department of Chemistry, University of Basel, Klingelbergstrasse 80, 4056 Basel, Switzerland

5 Supporting Information

**ABSTRACT:** One of the important reactive steps in Earth's atmosphere is the decomposition of H<sub>2</sub>SO<sub>4</sub> to H<sub>2</sub>O and SO<sub>3</sub>. However, because the UV spectrum of H<sub>2</sub>SO<sub>4</sub> was not found up to 140 nm, alternative mechanisms, including vibrationally induced dissociation, were proposed. Using adiabatic reactive molecular dynamics (ARMD) simulations with validated force fields for the product and educt channels, it is shown through explicit atomistic simulation that by exciting the  $\nu_9$  (OH-stretching-) mode, photodissociation can occur on the picosecond time scale. With the potential energy surfaces used in the present work,  $\nu_9 = 4$  is sufficient for this process. From a statistically significant number of trajectories (several thousands), vibrationally induced dissociation times are found to follow Gamma-distributions with most likely reaction times between 40 and 200 ps by depositing energies ranging from 40 to 60 kcal/mol, corresponding to 4 and 6 vibrational quanta in the OH stretching vibration. Because ARMD simulations allow multiple and long-time simulations, both nonstatistical, impulsive H-transfer and statistical, IVR-regimes of the decomposition reaction can be discussed in detail at an atomistic level.



## I. INTRODUCTION

In the troposphere, sulfur is emitted mainly as sulfur dioxide (SO<sub>2</sub>), which is oxidized in the atmosphere to form SO<sub>3</sub> and subsequently hydrolyzed to sulfuric acid (H<sub>2</sub>SO<sub>4</sub>).<sup>1,2</sup> This cycling between SO<sub>2</sub> and H<sub>2</sub>SO<sub>4</sub> has been studied mainly because it plays an important role in the troposphere aerosol layer formation, which is related to global atmospheric chemistry,<sup>3</sup> including cloud condensation<sup>4</sup> altering the chemical composition and influencing the course of aqueous reactions in clouds. Recent experiments have shown that formation rates and H<sub>2</sub>SO<sub>4</sub> concentrations are comparable to those found in the atmosphere, supporting the idea that sulfuric acid is involved in the nucleation of atmospheric aerosols.<sup>5</sup> The mechanisms and pathways involved in the dissociation of H<sub>2</sub>SO<sub>4</sub> to SO<sub>3</sub> and H<sub>2</sub>O are still debated. Initially, it was assumed that hydrolysis occurs via absorption of ultraviolet (UV) radiation to produce a dissociative electronically excited state.<sup>6</sup> However, the electronic absorption spectrum of H<sub>2</sub>SO<sub>4</sub> up to 140 nm could not be found experimentally.<sup>7</sup> In addition, UV photons are absorbed at higher altitudes.<sup>8</sup> Consequently, only few photons penetrate into the atmosphere with wavelengths shorter than 179 nm,<sup>9</sup> and UV photodissociation of H<sub>2</sub>SO<sub>4</sub> is unlikely to occur at lower altitudes.<sup>10</sup>

Following this, it was proposed that some OH-containing species in the atmosphere, including HONO, HONO<sub>2</sub>, or HO<sub>2</sub>NO<sub>2</sub>, absorb visible radiation and induce vibrational overtone transitions to states with several quanta in the OH stretching vibration.<sup>11</sup> Related calculations suggest that dissociation of H<sub>2</sub>SO<sub>4</sub> is possible mainly from highly excited OH-stretching vibrational overtones,<sup>10</sup> which has been also proposed to occur for HNO<sub>x</sub> systems.<sup>12</sup> Experimentally, the asymmetric OH stretching vibration  $\nu_9$  with  $\nu_9 = 4$  and  $\nu_9 = 5$  has been characterized by using cavity ring-down spectroscopy, but vibrational photodissociation

of H<sub>2</sub>SO<sub>4</sub> has as yet not been observed directly.<sup>13</sup> The vibrational overtones were found at 13490 and 16494 cm<sup>-1</sup>, which correspond to 38.57 and 47.16 kcal/mol, respectively. The barrier for sulfuric acid hydrolysis is predicted to be between 32 and 40 kcal/mol ( $\approx 12600$  cm<sup>-1</sup>), which opens the possibility for vibrationally induced photodissociation.

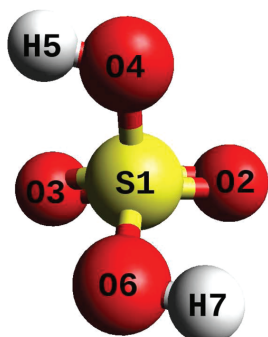
Because experimentally the vibrationally induced decomposition reaction H<sub>2</sub>SO<sub>4</sub>  $\rightarrow$  SO<sub>3</sub>+H<sub>2</sub>O was not directly observed, computational studies were used to address this problem. A pioneering atomistic simulation at the PM3 level investigated the dynamics of vibrationally excited sulfuric acid by depositing 21 511 cm<sup>-1</sup> (61.5 kcal/mol corresponding to  $\nu_9 = 6$  quanta<sup>15,16</sup>) of vibrational energy, and the subsequent dynamics was followed for  $\sim 100$  trajectories.<sup>15</sup> In 5% of the cases, photodissociation was found to take place with an average reaction time of 9 ps. However, the number of trajectories that could be run and analyzed was too small to arrive at a conclusive picture due to the high computational effort involved in MD/PM3 simulations.<sup>15</sup> Subsequently, Rice–Ramsperger–Kassel–Marcus theory (RRKM) was used to calculate rate constants ( $k_{RRKM}$ ) assuming a statistical energy distribution in the excited state.<sup>17</sup> Together with an estimate of the vibrational de-excitation rate,  $k_{collision}$ , the behavior of H<sub>2</sub>SO<sub>4</sub> after excitation at different altitudes was studied. This work showed that the photodissociation quantum yield increases with altitude and that, at high altitude (above 35 km), the low pressure slows down collisional vibrational relaxation.

In order to treat both possible limits—statistical (RRKM-like) and nonstatistical (impulsive) reactions<sup>18</sup>—on an equal footing,

Received: August 24, 2011

Revised: November 2, 2011

Published: November 04, 2011



**Figure 1.** Minimum energy structure ( $C_2$  symmetry) for  $H_2SO_4$  together with atom labeling.

the current work uses adiabatic reactive molecular dynamics (ARMD)<sup>19–22</sup> to investigate vibrationally induced  $H_2SO_4$  decomposition from a statistically significant number of molecular dynamics (MD) trajectories. ARMD is a surface-crossing algorithm that allows one to follow bond-breaking and bond-forming processes in MD simulations. Because the method was developed for empirical force fields, extensive sampling of phase space is possible. To complement the study, simulations based on the semiempirical self-consistent charges-density functional tight binding (SCC-DFTB) level are also carried out and analyzed.

## II. THEORETICAL AND EXPERIMENTAL METHODS

**A. MD Simulations.** All MD simulations were carried out with CHARMM<sup>23</sup> with provisions for bond-breaking and bond-formation through ARMD.<sup>19</sup> In addition, several hundred “on the fly” MD simulations with SCC-DFTB (see below) were run for up to 1 ns. Starting from a geometry optimized structure of  $H_2SO_4$  (see Figure 1) with 1000 steps of steepest descent minimization, the system was heated to 300 K. The equations of motion were propagated with a leapfrog Verlet algorithm with a time step of  $\Delta t = 0.1$  fs during 40 ps and equilibrated for 40 ps, followed by 50 ps of free dynamics simulations in the NVE ensemble. Vibrational excitation was induced by preparing a nonequilibrium state, which corresponds to depositing a specific amount of energy into the relevant degrees of freedom. Different schemes to prepare nonequilibrium initial conditions exist.<sup>24</sup> One attractive possibility is to scale the instantaneous velocity vector along the normal mode direction corresponding to the mode of interest. Alternatively, both positions and velocities can be scaled. The first approach—previously and successfully used to, e.g., study proton transfer in small H-bonded complexes<sup>25</sup>—assumes that the entire excitation is converted into kinetic energy, whereas the second method uses modifications in both the kinetic and potential energies to prepare a nonequilibrium state. One possible drawback of modifying both positions and momenta is the fact that, after excitation, a short equilibration period (a few picoseconds) is required to avoid artifacts due to close proximity of atoms.<sup>24</sup> On the other hand, such an approach is probably more general. In the present work, the nonequilibrium state was prepared according to the first scheme, i.e., scaling the instantaneous velocity vector along the normal mode  $\nu_9$  with energies of 13490 and 16494  $cm^{-1}$  (38.57 and 47.16 kcal/mol) from experiment<sup>13</sup> and 21511  $cm^{-1}$  (61.5 kcal/mol) estimated from

computation.<sup>15</sup> After this nonequilibrium preparation, 5000 independent ARMD trajectories in the NVE ensemble were run for 0.1–1 ns depending on the level of excitation.

For the simulations in a water droplet,  $H_2SO_4$  was solvated within a sphere of radius 16 Å of TIP3P water.<sup>26</sup> Spherical boundary conditions were applied to avoid dissociation of the droplet. The restraining potential is a quartic polynomial  $k\delta^2(\delta^2 - V_P)$  where  $\delta = r - r_{off}$ ,  $k = 0.2$  kcal/(mol Å<sup>4</sup>),  $r$  is the distance of the restraint atom from the center of the simulation sphere, and  $V_P$  is an offset value taken to be 2.25 Å<sup>2</sup>. The cutoff distance  $r_{off} = 16$  Å corresponds to the size of the droplet. Again, the O6–H7 bond is vibrationally excited in the same fashion as for the gas phase simulations. In order to maintain the  $H_2SO_4$  molecule near the center of the sphere, a weak harmonic constraint was applied with a force constant of 0.5 kcal/mol/Å<sup>2</sup>. Upon vibrational excitation the boundary of the water sphere was set to  $r_{off} = 25$  Å to allow expansion of the water droplet.

**B. Force Fields for Educt and Product States.** In order to study vibrationally induced dissociation of  $H_2SO_4$ , a suitably parametrized force field is required. For this, electronic structure calculations were carried out with Gaussian 03 at the MP2/6-311G++(2d,2p) level of theory in the gas phase.<sup>27</sup> Initially, the structures of  $H_2SO_4$  and the product  $SO_3 + H_2O$  were optimized. Equilibrium parameters and force constants for bonds and angles were obtained from optimized structures and five single point electronic structure calculations along each of the internal coordinates. Gradients required for the force constants were evaluated from the five-point rule. Mulliken charges were determined on the MP2-optimized structure (see Figure 1 and Table S-I, Supporting Information). To allow bond breaking (distances O4–S1 and O6–H7 in  $H_2SO_4$ ) and bond formation (O4–H7 for water), Morse potentials were employed:

$$V(r) = D_e(1 - e^{-\beta(r - r_e)})^2 \quad (1)$$

Here,  $r$  is the distance between the atoms involved in the bond,  $D_e$  is the dissociation energy,  $\beta$  controls the steepness and width of the potential, and  $r_e$  is the equilibrium distance. The parameters  $D_e$ ,  $\beta$ , and  $r_e$  were obtained by fitting Morse potentials to the energies which yields the parameters given in Table S-II (Supporting Information). Equilibrium distances and dissociation energies were obtained from relaxed energy scans along the respective coordinate  $r$  from 20 steps between 1.0 and 3.0 Å for O4–S1 and 16 steps between 0.5 and 2.0 Å for the O6–H7 distance. All remaining bonds are represented by harmonic potentials. The parametrization for the water molecule in the product state ( $SO_3 + H_2O$ ) followed similar procedures, and the parameters are given in Table S-III (Supporting Information). However, it should be noted that the simulations are not sensitive to the particular parametrization in the product channel, as dissociation is irreversible.

Force field parameters for dihedral angles O2–S1–O4–H5 and H5–O4–S1–O6 were obtained by fitting to energies from relaxed scans starting from the stable  $C_2$  conformer and using increments of 10° each. Subsequently, the standard CHARMM force field terms for dihedral angles were fitted using I-NoLLS<sup>28</sup> and the CHARMM/I-NoLLS front end.<sup>29</sup> For van der Waals interactions, Lennard-Jones potentials with standard CHARMM parameters were initially used. However, as will be discussed further below, ARMD simulations require revised van der Waals parameters for atoms involved in the reaction.

With a validated force field, the only free parameter in ARMD is  $\Delta$ , which describes the energy difference between the product

**Table 1. Optimized Geometry of Educt ( $\text{H}_2\text{SO}_4$ ) and Product  $\text{SO}_3 + \text{H}_2\text{O}$  Calculated with SCC-DFTB, MP2(full)/6-31G(d,p), and Force Field**

Educt ( $\text{H}_2\text{SO}_4$ )			
bonds [Å]	MP2	CHARMM	SCC-DFTB
H7–O6	0.97	0.97	0.97
O6–S1	1.60	1.60	1.65
H5–O4	0.97	0.97	0.97
S1–O4	1.60	1.60	1.65
S1–O2	1.43	1.43	1.45
S1–O3	1.43	1.43	1.45
angles [°]			
S1–O4–H5	107.5	110.1	116.3
S1–O6–H7	107.5	110.1	116.3
O3–S1–O2	124.7	124.4	123.3
O3–S1–O4	108.7	106.1	109.8
O2–S1–O4	105.2	108.0	105.1
O6–S1–O4	101.7	101.5	101.8
O6–S1–O3	105.3	108.0	105.1
O6–S1–O2	108.7	106.1	109.8
Product ( $\text{SO}_3 + \text{H}_2\text{O}$ )			
bond [Å]	MP2	CHARMM	SCC-DFTB
S1–O3	1.43	1.43	1.46
S1–O2	1.43	1.43	1.46
S1–O4	1.43	1.43	1.45
O6–H7	0.96	0.96	0.97
O6–H5	0.96	0.96	0.97
angle [°]			
O3–S1–O2	119.9	119.3	119.5
O2–S1–O4	119.9	119.3	120.0
O4–S1–O3	119.9	119.3	120.0
H7–O6–H5	104.1	104.1	108.0

and the educt channels.<sup>19</sup> This energy difference encompasses two contributions: (i) the asymptotic energy difference between the states involved and (ii) a contribution from the fact that the zero of energies for the force fields for the states involved do not coincide. MP2 calculations show that the asymptotic energy difference is 18.7 kcal/mol, which yields  $\Delta = 95.0$  kcal/mol. With this value, the asymptotic energy difference between the force fields for  $\text{H}_2\text{SO}_4$  and  $\text{SO}_3 + \text{H}_2\text{O}$  is 18.0 kcal/mol, which is sufficiently accurate for the following.

To compare the structure, activation energy, and reaction path, the transition state (TS) structure was calculated using quadratic synchronous transit-guided (QST2) calculations,<sup>30</sup> and the reaction path was followed from intrinsic reaction coordinate (IRC) calculations.<sup>31</sup> Complementary to the ab initio calculations, potential energy profiles and optimized structures were also determined from semiempirical SCC-DFTB calculations.<sup>32</sup> Such an approach also allows one to carry out SCC-DFTB/MD simulations and to compare reaction probabilities with those obtained from ARMD.

### III. RESULTS

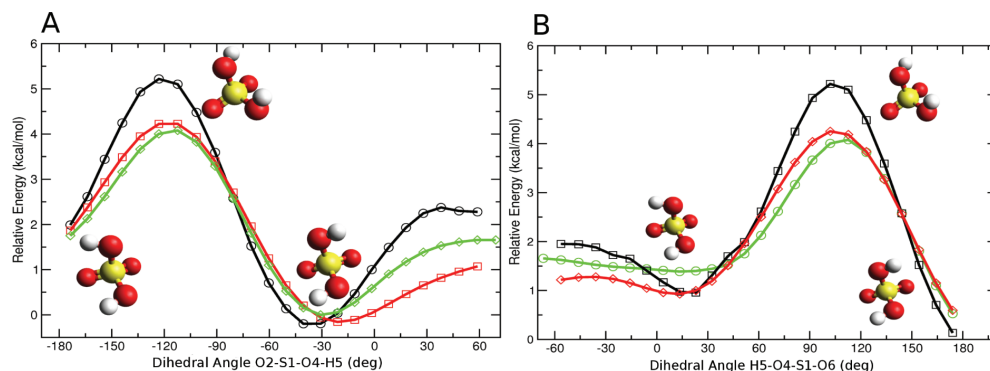
**A. Validation of the Force Field.** The energy-minimized structure using the force field leads to a  $C_2$ -symmetric structure

(see Figure 1).<sup>33</sup> A comparison of important bond lengths, angles and dihedral angles between educt and product structures, optimized from MP2, SCC-DFTB, and force field calculations, is summarized in Table 1. The root-mean-square deviation (rmsd) between the force field or SCC-DFTB and MP2 are 0.08 and 0.69 Å for educt and 0.13 and 0.50 Å for products, respectively. This suggests that the force field is parametrized in a meaningful way, and it reproduces MP2 geometries better than SCC-DFTB on average. For bonds, the difference between force field and MP2 calculations is 0.03 Å, whereas for angles it is 3° (see Table 1).

In the dynamics simulations discussed further below, it was found that the parametrization of the dihedral angles O2–S1–O4–H5 and H5–O4–S1–O6 requires particular attention. Previous computational work on rotamers for sulfuric acid revealed the existence of two minimum energy geometries ( $C_2$  and  $C_s$ ) and two saddle points separating them.<sup>34,35</sup> In order to capture these effects, a more detailed parametrization strategy was pursued (see Theoretical and Experimental Methods). Figure 2 compares potential energy profiles for relaxed scans along the two dihedral angles for the force field, MP2, and SCC-DFTB calculations. The force field calculations find the higher saddle point 4 kcal/mol above the most stable structure ( $C_2$  symmetry), whereas the second saddle point is 1.8 kcal/mol less stable. This is in good agreement with previous studies<sup>34,35</sup> and with the present MP2 calculations. It is also found that the force field is able to better capture the potential energy surface (PES) for the  $\text{H}_2\text{SO}_4$  rotamers than SCC-DFTB (Figure 2).

Additional validations of the force field for  $\text{H}_2\text{SO}_4$  were carried out for the OH-stretch overtones. Treating the OH-mode as a localized mode with the OH-reduced mass, the Morse potential yields frequencies of 3701  $\text{cm}^{-1}$  for the  $\nu_9$  fundamental and 14583, 17427, 20103, and 22683  $\text{cm}^{-1}$  for 4, 5, 6, and 7 quanta, respectively. The fundamental overestimates the experimental value of 3609  $\text{cm}^{-1}$  by  $\sim 100$   $\text{cm}^{-1}$ , and differences for overtones are correspondingly larger. Additional information can be gained from explicit MD simulations. To this end, 30 independent simulations, each 250 ps in length, were carried out at 300 K. The center of the power spectra of the OH-mode for the fundamental and overtones with 4, 5, 6, and 7 quanta peak at 3703, 14713, 17367, 20075, and 22706  $\text{cm}^{-1}$ . These values are similar to those calculated from the Morse potential, which indicates that the OH-mode is indeed quite localized. Additional improvements to the force field are certainly possible to make the parametrization spectroscopically accurate, which, however, was not the purpose of the present work. It should be noted that previous work found an excitation energy of 19289.5  $\text{cm}^{-1}$  (55.15 kcal/mol) for  $\nu_9 = 6$  at the CCSD(T) level,<sup>36</sup> which compares with 57.39 kcal/mol from the present work. As we are primarily concerned with the vibrationally induced dynamics following the deposition of a specific amount of energy, and to avoid confusion, we will refer to “excitation of  $\nu_9 = 6$ ” upon delivering 61.5 kcal/mol (original PM3 energy for  $\nu_9 = 6$ ) of energy. However, it should be noted that for the force field used in the present work, 61.5 kcal/mol corresponds to an excitation level closer to  $\nu_9 = 7$ .

As a final experimental observable the  $\text{H}_2\text{SO}_4$  dipole moment was calculated from the force field and from SCC-DFTB for an ensemble of 10000 snapshots from an equilibrium simulation. The values are  $\mu = 2.69 \pm 0.54$  D and  $3.71 \pm 0.56$  D, compared with 3.41 D from MP2/6-311G++(2d,2p) calculations and 2.30 D from experiment.<sup>37</sup> Thus, the force field parametrization best reproduces the dipole moment. However, again, the dipole moment is only of limited relevance for the present work which



**Figure 2.** PES along dihedral angles H5–O4–S1–O6 (A) and O2–S1–O4–H5 (B) from MP2 (green), the force field (red), and SCC-DFTB (black).

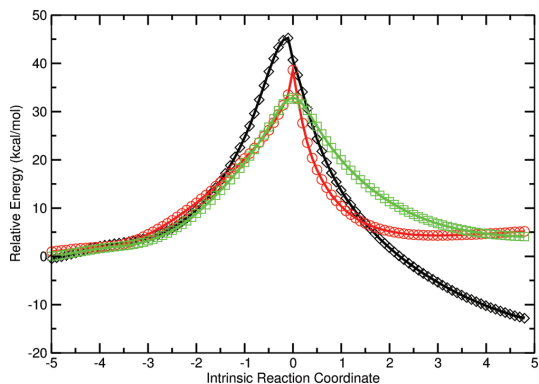
is concerned with the investigation of vibrationally induced photodissociation.

**B. Parametrization of van der Waals Interactions.** In conventional force fields, van der Waals interactions are represented by Lennard-Jones potentials

$$V(r) = \varepsilon[(\sigma/r)^{12} - 2(\sigma/r)^6] \quad (2)$$

Here, the  $r^{12}$  term represents short-range repulsion due to overlap of electron orbitals, and the  $r^6$  term describes dispersion interactions at long-range. In ARMD, the covalent bond between two atoms is replaced by nonbonded interactions upon breaking, which is determined from an energy criterion involving the entire system.<sup>19</sup> Because the instantaneous “bound” configuration can be quite different from a typical “dissociated” configuration, this can lead to large energy changes upon dissociation. If the standard van der Waals parameters are used upon breaking the chemical bond, the two atoms are likely to partially overlap. For atom separations deep within the bonded region, the notion of nonbonded interactions (as done in force fields) is not meaningful. Thus, for dissociating bonds in ARMD the van der Waals parameters need to be adapted to avoid artificial speedup of the separating atoms/moieties.

To determine meaningful van der Waals- $\sigma$  values for atoms involved in forming and breaking bonds (the O6–H7 and S1–O4 bonds), ARMD simulation were carried out. Figure S1 (Supporting Information) shows the O6–H7 (A) and S1–O4 (B) separation as a function of time after dissociation for different van der Waals ranges. Rapid increase of internuclear distances between atoms indicates a strong and instantaneous repulsion upon switching the force fields—based on ARMD<sup>19</sup>—from the bound ( $\text{H}_2\text{SO}_4$ ) to the dissociated ( $\text{SO}_3 + \text{H}_2\text{O}$ ) state. This rapid increase is, however, an artifact of using separately parametrized force fields for the bound and dissociated states, respectively. To avoid such artifacts, the van der Waals ranges  $\sigma$  of the atoms involved in the O6–H7 and S1–O4 interaction are gradually reduced. This leads to the following changes in the van der Waals ranges (see Table S-IV (Supporting Information)) for which no artifact in the dynamics upon breaking the bond was observed: from  $R_{\text{min}}^{\text{H}}/2 = 0.23 \text{ \AA}$ ,  $R_{\text{min}}^{\text{O}}/2 = 1.77 \text{ \AA}$ , and  $R_{\text{min}}^{\text{S}}/2 = 2.00 \text{ \AA}$  for the H, O, and S atoms to  $R_{\text{min,new}}^{\text{H}}/2 = 0.22 \text{ \AA}$ ,  $R_{\text{min,new}}^{\text{O}}/2 = 1.22 \text{ \AA}$ , and  $R_{\text{min,new}}^{\text{S}}/2 = 1.30 \text{ \AA}$ . This corresponds to a reduction of the  $\sigma$  values for the O–S and O–H pairs by 44 and 26%, respectively. Interestingly these  $\sigma$  values ( $\sigma(\text{S–O}) = 2.52$  and  $\sigma(\text{O–H}) = 1.44$ )



**Figure 3.** Reaction path for  $\text{H}_2\text{SO}_4$  hydrolysis. Structure numbers correspond to those structures that follow the reaction path in IRC calculation at the MP2 level. Here each structure from MP2 was taken, and single point energy was evaluated using force field and SCC-DFTB. Black represents calculations using SCC-DFTB, red is for force field, and green is for MP2. The barriers from MP2, force field, and SCC-DFTB are 32.8, 40.0, and 45 kcal/mol, respectively.

are close to the O–S distance in  $\text{H}_2\text{O–SO}_3$  of 2.42 Å in the product state and O–H distance (1.23 Å) in the TS, which supports the empirical adjustment of the van der Waals radii described above.

**C. Geometry of the TS.** Ab initio calculations for the TS between  $\text{H}_2\text{SO}_4$  and  $\text{SO}_3 + \text{H}_2\text{O}$  yielded structures comparable to those reported earlier.<sup>2,14</sup> Here, the barrier is 32.9 kcal/mol (including zero point energy corrections) compared to 32.2 kcal/mol from previous calculations.<sup>14</sup> Figure 3 reports the energy barrier from IRC calculations in the gas phase together with structures of the product, educt, and TS geometries.

Structures along the reaction path were used to evaluate the force field energy and that from SCC-DFTB. This allows one to compare the activation energy and to further validate the force field (see Figure 3). It is found that the overall shape of the PES connecting educt and product through the TS from the force field is in quite good agreement with the MP2 calculations. The relative energies of products and educts agree to within 2 kcal/mol, whereas the forward barrier is overestimated by 7.2 kcal/mol.



**Table 2.** Comparison for TS Structure for H<sub>2</sub>SO<sub>4</sub> from the Different Methods

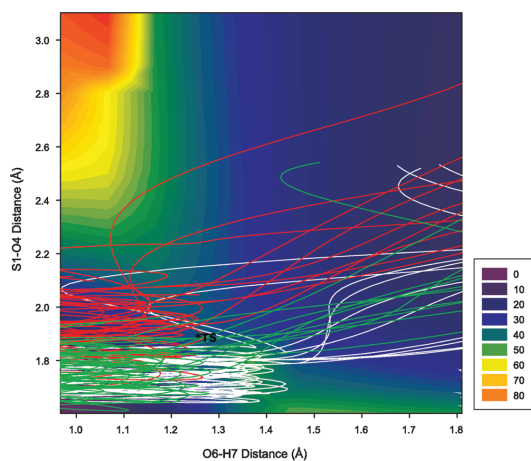
coordinate	literature <sup>14</sup>	MP2	force field <sup>a</sup>	SCC-DFTB
bonds [Å]				
O4–H5	0.97	0.97	1.02 ± 0.03	1.05
S1–O4	1.92	1.89	1.87 ± 0.01	2.01
O4–H7	1.24	1.25	1.33 ± 0.04	1.09
O6–H7	1.24	1.25	1.18 ± 0.03	1.28
S1–O6	1.53	1.52	1.49 ± 0.02	1.59
angle [deg]				
S1–O4–H5	110.8	106.6	110.9 ± 7.6	105.0
S1–O4–H5	66.9	67.7	64.4 ± 8.2	63.2
O4–S1–O6	80.4	81.6	87.2 ± 6.2	74.9

<sup>a</sup> For the force field calculations averages over 50 snapshots are reported.

Thus, all rates reported in the following are likely to be too slow. Toward the product channel the profiles start to deviate from each other. However, this is of less concern, as we are primarily interested in the decomposition reaction. SCC-DFTB also follows the MP2 energy profile quite closely but leads to a higher TS energy (by 12.2 kcal/mol) than MP2. On the product side, SCC-DFTB has a very different PES profile compared to MP2. However, this again does not affect the conclusions for the dissociation reaction. The structure of the TS obtained by ARMD is also near those obtained by MP2 calculations and reported previously (see Table 2).<sup>14</sup>

The PES for hydrolysis of H<sub>2</sub>SO<sub>4</sub> as a function of the O6–H7 and S1–O4 distances at the MP2/6-311G++(2d,2p) level is shown in Figure 4. It is of interest to consider projections of reactive trajectories onto this surface in order to better understand how the reactive seam is crossed. As will be described later, excitation of  $\nu_9 = 5$  leads to a broad distribution of reaction times ranging from a few to hundreds of picoseconds. From all the reactive trajectories calculated with ARMD, 10 with short, intermediate, and long reaction times were randomly chosen. They correspond to reaction times of  $\tau \approx 5$ , 250, and 600 ps and are also shown in Figure 4. Trajectories with  $\tau \approx 5$  ps (white) show pronounced stretching movement along the vibrationally excited (O6–H7) bond and rapidly cross the energy barrier without transferring appreciable amounts of energy into the S1–O4 bond. Due to the high vibrational excitation energy, rapid dissociation occurs, and the trajectory does not cross the reactive seam at the TS but away from it. The system explores energy regions uphill before dissociating into SO<sub>3</sub> + H<sub>2</sub>O. Trajectories with  $\tau \approx 250$  and 600 ps (green and red), show energy transfer into the (S1–O4) stretching mode but to different extents. The crossing of the reactive seam is also very rapid, but the distribution of crossing geometries is much wider compared to trajectories with early reaction times.

**D. Reaction Dynamics.** With a validated force field (correct product/educt geometries, qualitatively correct shape, and TS energy of PES), atomistic simulations can now be run to collect a statistically significant number of independent trajectories and to analyze the geometrical, energetic, and temporal aspects of vibrationally induced dissociation in H<sub>2</sub>SO<sub>4</sub>. In order to follow the temporal evolution of the system following vibrational excitation, ARMD simulations were carried out (see Theoretical and Experimental Methods). The parametrization presented so far allows two product states to be populated: dissociation into SO<sub>3</sub> + H<sub>2</sub>O and SO<sub>3</sub> + OH + H. However, the dissociation energy of water into OH + H in the gas phase is 117 kcal/mol,<sup>38,39</sup> which has to

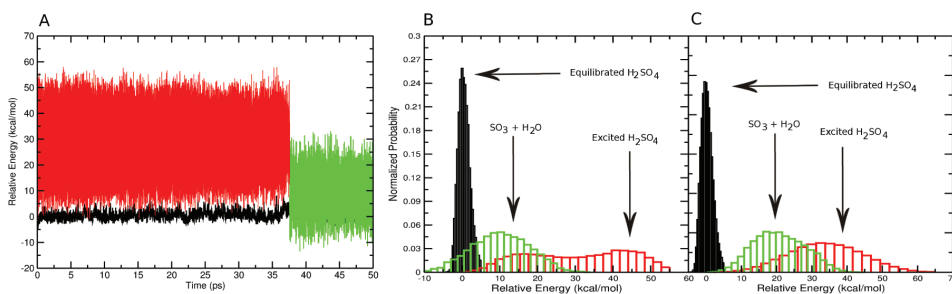


**Figure 4.** Calculated PES for hydrolysis of H<sub>2</sub>SO<sub>4</sub> in the space defined by distances O6–H7 and S1–O4 at the MP2/6-311G++(2d,2p) level. Color lines represent different reactive trajectories for  $\nu_9 = 5$ : 10 trajectories with  $\tau \approx 5$  ps (white), 10 trajectories with  $\tau \approx 250$  ps (green), and 10 trajectories with  $\tau \approx 600$  ps (red).

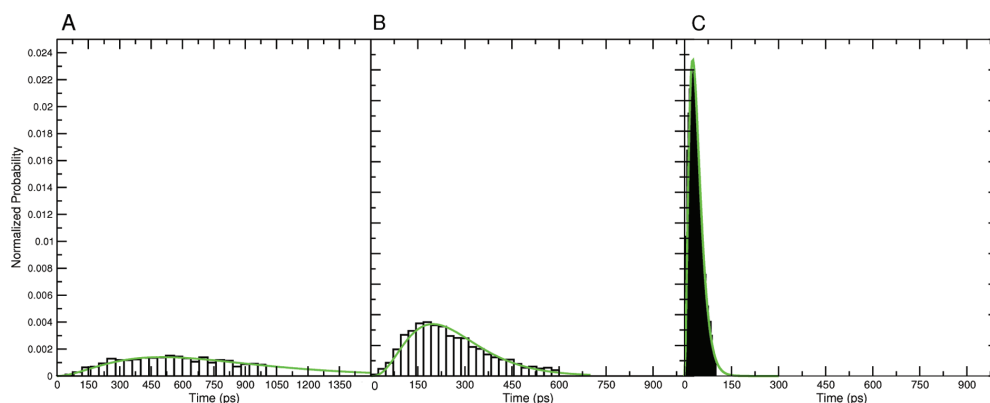
be included in order to correctly account for the asymptotic energetics. Thus, the ARMD simulations were run with one educt state (H<sub>2</sub>SO<sub>4</sub>) and two product states (SO<sub>3</sub> + H<sub>2</sub>O and SO<sub>3</sub> + OH + H) with correctly accounting for the asymptotic energetics.

As an illustration of a reactive trajectory, the potential energy as a function of time for a single ARMD trajectory (with excitation of  $\nu_9 = 6$ ) is reported in Figure 5A, which shows equilibrated H<sub>2</sub>SO<sub>4</sub> (black) and excited (nonequilibrium) H<sub>2</sub>SO<sub>4</sub> (red) dissociating after  $\sim 37$  ps. At this point, the potential energy changes considerably, indicating a transition from H<sub>2</sub>SO<sub>4</sub> to SO<sub>3</sub> + H<sub>2</sub>O. Figure 5B shows the distribution of the potential energy during the same simulation (histogram of the data in Figure 5A). Black corresponds to equilibrated H<sub>2</sub>SO<sub>4</sub> fluctuating around 0 kcal/mol, and red corresponds to the vibrationally excited state, where a bivariate distribution shows two distinct populations. One fluctuates around 10 kcal/mol and corresponds to the product state (SO<sub>3</sub> + H<sub>2</sub>O), whereas the second one fluctuates around 44 kcal/mol and characterizes the vibrationally excited educt state (H<sub>2</sub>SO<sub>4</sub>). Figure 5C shows the same quantities for a representative SCC-DFTB simulation. Again, the equilibrium and product state distributions (black) is narrow, whereas the nonequilibrium and product state distributions are broad.

For ARMD the total energy for vibrationally excited H<sub>2</sub>SO<sub>4</sub> fluctuates between 0 and 60 kcal/mol (Figure 5B) and overlaps with the product channel (SO<sub>3</sub> + H<sub>2</sub>O) between 0 and 30 kcal/mol. Contrary to that, the overlap between the two distributions in the SCC-DFTB/MD simulations is shifted to higher energies (Figure 5C), despite the same level of excitation. For these two randomly selected trajectories, crossing the reactive seam occurs at higher energy in SCC-DFTB/MD simulations compared to the ARMD simulations. This also agrees with the energy profiles from Figure 3. On the basis of these two observations, it is expected that ARMD compared to SCC-DFTB/MD gives more rapid (because of the lower barrier) and more reliable (because of better capturing the reaction energetics compared to the MP2 calculations) time scales for average reaction times.



**Figure 5.** (A) Potential energy as a function of time for ARMD. (B) Histogram of the data in panel A. (C) Energy distribution from an SCC-DFTB simulation. Black traces correspond to equilibrated H<sub>2</sub>SO<sub>4</sub>, and red corresponds to H<sub>2</sub>SO<sub>4</sub> excited with 61.5 kcal/mol. Bivariate distribution shows two populations corresponding to product (SO<sub>3</sub> + H<sub>2</sub>O) and educt (H<sub>2</sub>SO<sub>4</sub>).



**Figure 6.** Normalized distribution of dissociation times  $p(\tau)$  from ARMD simulations with excitation of  $\nu_9 = 4$  (A),  $\nu_9 = 5$  (B), and  $\nu_9 = 6$  (C). The green curves represent fits of the data to a  $\Gamma$ -distribution (see text). Note the different  $t$  axes in the three panels. Average dissociation times are 753, 201, and 40 ps for  $\nu_9 = 4$ ,  $\nu_9 = 5$  and  $\nu_9 = 6$ , respectively.

Initially, 5000 ARMD simulations were run for 100 ps. For excitation of  $\nu_9 = 4$ , a total of 96 (2%) dissociation events (H<sub>2</sub>SO<sub>4</sub> → SO<sub>3</sub> + H<sub>2</sub>O) out of the 5000 trajectories was observed. This fraction increased to 683 (14%) for  $\nu_9 = 5$  and 3015 (60%) for  $\nu_9 = 6$ . Therefore, the dissociation probability is a strong function of the excitation energy, as expected. Because for  $\nu_9 = 4$  and  $\nu_9 = 5$  only a small fraction of trajectories lead to dissociation, the trajectories were continued to  $t_{\text{max}} = 200$  ps. This increases the number of dissociative trajectories to 6% (324 events) and 34% (1707 events) for  $\nu_9 = 4$  and  $\nu_9 = 5$ , respectively. To further improve the quality of  $p(\tau)$ , which is the distribution of dissociation times, the simulation time was increased to 1 ns for  $\nu_9 = 4$  and to 600 ps for  $\nu_9 = 5$ , which finally leads to 56% and 87% of dissociation events for  $\nu_9 = 4$  and  $\nu_9 = 5$ , respectively.

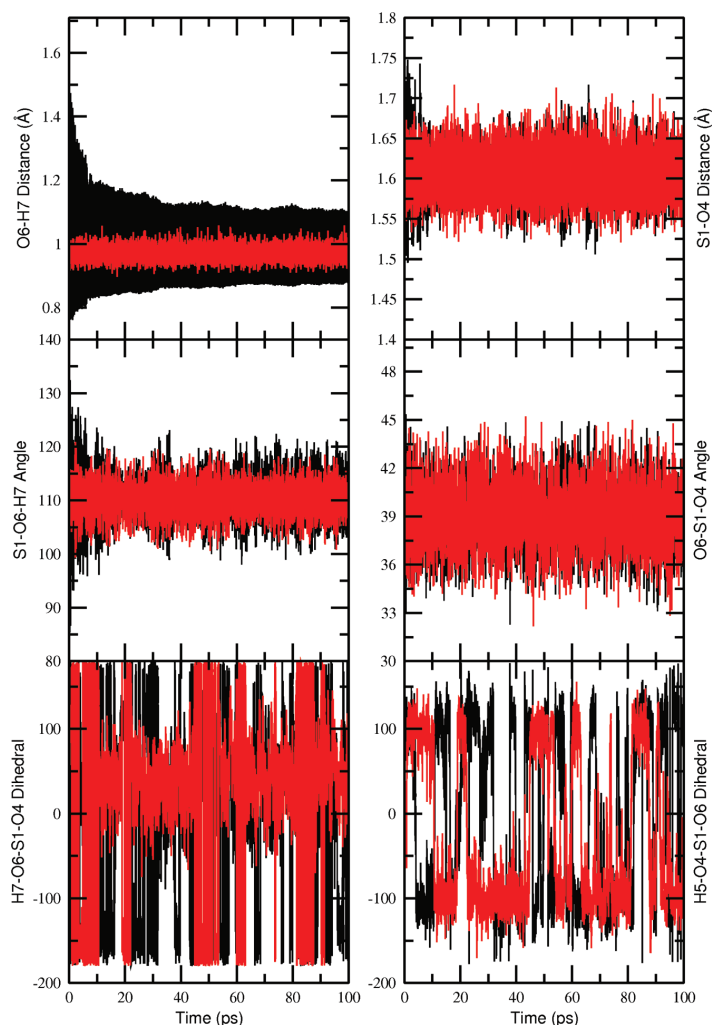
The distributions of dissociation times  $p(\tau)$  are reported in Figures 6A to C for all levels of excitation considered. The distributions  $p(\tau)$  can be conveniently characterized by  $\Gamma$ -distributions, which represent a probabilistic model for waiting times:

$$f(x; k, \theta) = x^{k-1} (e^{-x/\theta}) / (\theta^k \Gamma(k)) \quad (3)$$

Here,  $k$  is a shape parameter and  $\theta$  is a scale parameter, both of which are positive, and  $\Gamma(k)$  is the  $\Gamma$  function. The mean of the

$\Gamma$  distribution is  $k\theta$ , from which average dissociation times of  $\tau_d = 201$  ps (for  $\nu_9 = 5$ ) and  $\tau_d = 40$  ps (for  $\nu_9 = 6$ ) are obtained. Thus, by exciting the OH stretching mode with 61.5 kcal/mol, photodissociation is faster by almost 1 order of magnitude compared to excitation with 47.2 kcal/mol.

An alternative to ARMD are “on the fly” MD simulations with interactions calculated at an affordable (quantum or semiempirical) computational level. Such a model is SCC-DFTB, which for the present case allows one to run several hundred extended SCC-DFTB/MD simulations. The potential energy distribution for a single trajectory (with  $\nu_9 = 6$ ) is reported in Figure 5C. However, such trajectories are computationally much more demanding than ARMD. Therefore, only 280 trajectories, each 1 ns in time, were run with 61.5 kcal/mol of excess energy. A total of 101 dissociation events (36%) was observed. This compares with a dissociation probability of 60% from 100 ps simulations based on the parametrized force field. The distribution of reaction times,  $p(\tau)$ , is reported in Figure S2 (Supporting Information) and demonstrates that the average reaction time is considerably longer with SCC-DFTB/MD than for ARMD simulations with the same excitation energy. This reflects the higher activation energy for SCC-DFTB (46.5 kcal/mol) compared to the force field (40.0 kcal/mol; see Figure 3). On the basis of the finding



**Figure 7.** Time evolution of different coordinates for vibrationally excited  $\nu_9 = 6$   $\text{H}_2\text{SO}_4$  along the normal mode involving the (O6–H7) bond in a water droplet. The equilibrated system is shown as the red trace, whereas the black trace corresponds to vibrationally excited  $\text{H}_2\text{SO}_4$ . Note the very rapid initial decay along the vibrationally excited O6–H7 bond, which prevents the system from undergoing vibrationally assisted photodissociation.

that SCC-DFTB overestimates the energy barrier for the reaction, this result is not surprising. Furthermore, compared to MP2 and the force field parametrization, semiempirical DFTB is not sufficiently accurate to capture other important characteristics of the PES, including torsional barriers. On the other hand, SCC-DFTB also finds both physically relevant processes for the decomposition: statistical and nonstatistical vibrationally induced photodissociation.

**E. Dissociation of Sulfuric Acid in a Water Droplet.** In order to demonstrate the possibilities ARMD can offer to study reactive processes, and to compare with previous efforts on the binary  $\text{H}_2\text{SO}_4$ –water complexes, we also briefly consider dissociation of sulfuric acid in a water droplet. However, it should be noted that this process is probably not relevant in the atmosphere because previous work has shown that only three water molecules are

sufficient to induce deprotonation, i.e., formation of  $\text{HSO}_4^-$ .<sup>40</sup> Therefore, it is expected that in solution the ionic species  $\text{HSO}_4^-$  is more prevalent than neutral  $\text{H}_2\text{SO}_4$ . Hydrolysis of sulfuric acid has previously been investigated for hydrated  $\text{H}_2\text{SO}_4$ .<sup>2,14</sup> With one water molecule it was proposed that due to the lower barrier (25 kcal/mol) for the dissociation of  $\text{H}_2\text{SO}_4 \cdot \text{H}_2\text{O}$  compared with the reaction in gas phase, the process could occur by exciting  $\nu_9 \geq 3$ .<sup>10,14</sup> However, dynamics studies of the same system in 100 different trajectories reported no dissociation, but the decomposition of the cluster into the monomers  $\text{H}_2\text{SO}_4$  and  $\text{H}_2\text{O}$  was observed.<sup>15</sup> Therefore, photodissociation of  $\text{H}_2\text{SO}_4$  in water is likely to be suppressed due to collisions with surrounding molecules.<sup>17</sup>

This is in fact what the present simulations find as well. None of the 100 trajectories exhibited vibrationally induced dissociation.



Contrary to excitation in the gas phase, internal energy redistribution within  $\text{H}_2\text{SO}_4$  is unlikely, as collisions with surrounding water molecules occur. This energy flow, which takes place on a picosecond time scale, is irreversible and considerably reduces the probability for vibrationally induced photodissociation of solvated  $\text{H}_2\text{SO}_4$ . Figure 7 reports different internal coordinates of equilibrated (red) and vibrationally excited (black)  $\text{H}_2\text{SO}_4$  as a function of simulation time for a single trajectory. The results demonstrate that the vibrationally excited O6–H7 bond relaxes on the picosecond time scale without transferring the energy to other internal modes. Rather, the vibrational energy is lost through collisions with surrounding water molecules. Thus, the present simulations support the finding that vibrational energy flow to surrounding degrees of freedom is sufficiently rapid to suppress vibrational dissociation of  $\text{H}_2\text{SO}_4$ .

#### IV. DISCUSSION

In the present work, dissociation of  $\text{H}_2\text{SO}_4$  into  $\text{SO}_3 + \text{H}_2\text{O}$  was investigated by ARMD from a statistically significant number of events following vibrational excitation of the O–H stretching vibration. It was found that starting from  $\nu_9 = 4$  the reaction occurs on the picosecond to nanosecond time scale. Photodissociation on the picosecond time scale is observed upon excitation with  $\sim 50$  kcal/mol, which corresponds to  $\nu_9 \geq 5$ . This supports previous conclusions based on ab initio calculations<sup>13</sup> and PM3/MD simulations.<sup>15</sup>

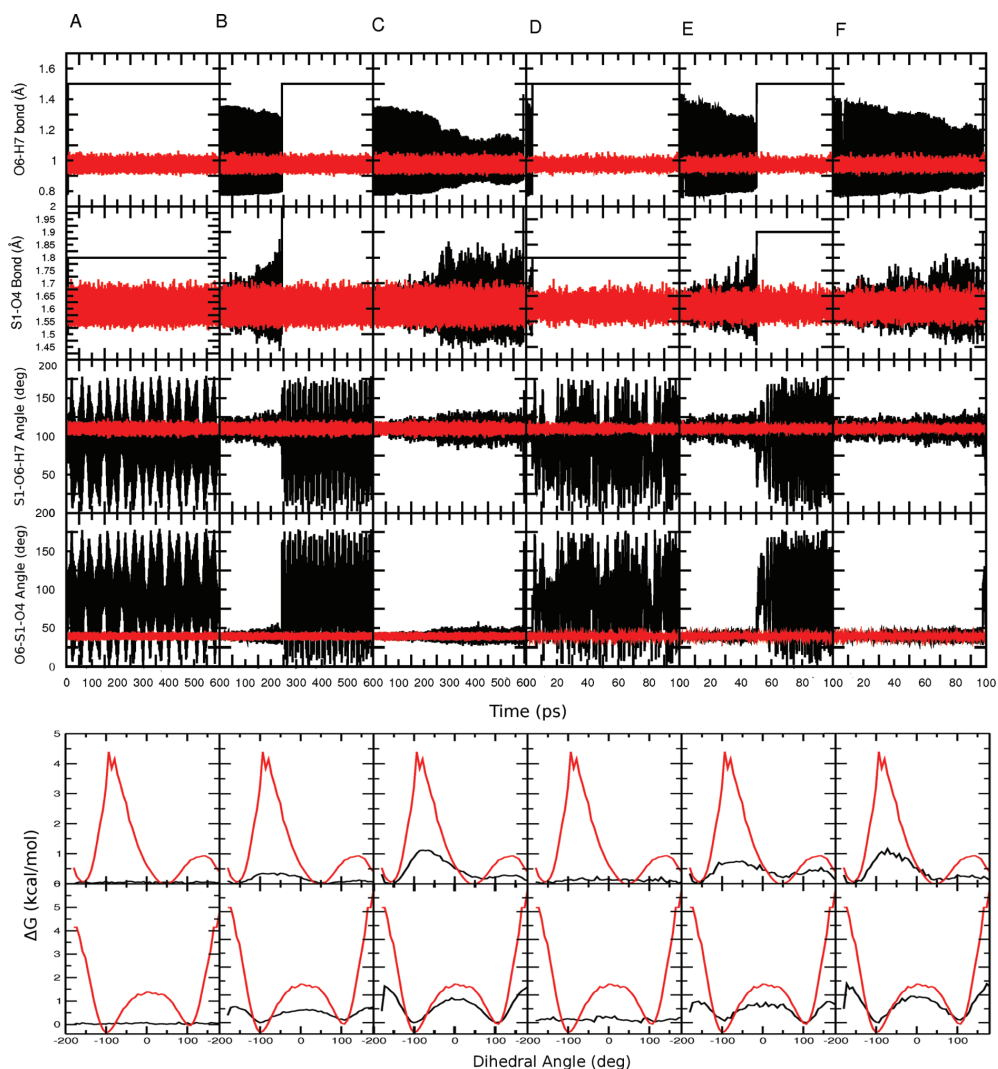
So far, semiempirical PM3, force field-based, and SCC-DFTB methods have been used to study the energetics and dynamics of sulfuric acid dissociation following vibrational excitation. Therefore, a comparison of the three methods is quite instructive. For  $\text{H}_2\text{SO}_4$  geometries between MP2/6-311++G(2d,2p) and the parametrized force field differ by 0.002 Å for all bonds and 2.1° for all angles. This compares with 0.024 Å and 2.7° in comparing SCC-DFTB with the MP2 calculations and deviations of up to 0.1 Å and between 1.1 and 4.6° for PM3 relative to MP2/TZP for  $\text{HNO}_3$  and *cis*- and *trans*-HONO<sup>41</sup> and less than 2% for bonds and 6% for angles for  $\text{H}_2\text{SO}_4$ .<sup>17</sup> For harmonic frequencies, reported average differences<sup>41</sup> between MP2/TZP and PM3 for  $\text{HNO}_3$  and *cis*- and *trans*-HONO are 130  $\text{cm}^{-1}$ , compared to 78  $\text{cm}^{-1}$  for the parametrized force field and 61  $\text{cm}^{-1}$  for SCC-DFTB. Finally, the activation energy for H-transfer from MP2/6-311++G(2d,2p) calculations is 32.8 kcal/mol compared with 42.8 kcal/mol from PM3,<sup>17</sup> 40.0 kcal/mol from the parametrized force field, and 46.5 from SCC-DFTB. These comparisons suggest that the parametrized force field performs best for structures and energetic barriers, whereas SCC-DFTB is on average slightly better for frequencies. The barrier for H-transfer from PM3 is somewhat higher than that from the force field. However, direct comparison for all, geometric, spectroscopic, and energetic data is somewhat difficult because for the performance of PM3 only explicit data for  $\text{HNO}_3$  and *cis*-/*trans*-HONO were available. The most relevant quantity for the present work is the barrier for H-transfer, which is best captured by the parametrized force field.

Reaction times with  $\nu_9 \geq 5$  range from 6 to several hundred picoseconds, respectively. To better understand the *mechanism* leading to dissociation, individual trajectories were considered in more detail. This was done for very rapid, average and long reaction times for the two excitation energies. Trajectories with reaction times of  $\tau = 6, 244,$  and 589 ps for excitation of  $\nu_9 = 5$ , and  $\tau = 7, 50,$  and 98 ps for  $\nu_9 = 6$ , respectively, were analyzed. The analysis focused in particular on the time scale and degree of

energy redistribution after vibrational excitation of the OH-bond. Figure 8 reports the time series for different internal coordinates. Three types of processes can be distinguished: (i) impulsive hydrogen hopping for which no internal vibrational relaxation (IVR) is found (see Figure 8, columns A and D); (ii) partial IVR (Figure 8, columns B and E); and (iii) almost complete IVR where the vibrational energy leaks into other degrees of freedom before returning into the dissociative coordinate after many vibrational periods (Figure 8, columns C and F). For process (iii), the amplitude of the OH stretching vibration decays to levels comparable to the equilibrium simulations before dissociation occurs without, however, reaching it. In view of previously reported RRKM rates, this is relevant because direct comparison of  $k_{\text{RRKM}}$  with rates from the present work would require complete redistribution of the energy following vibrational excitation. We would, therefore, expect that even the longest reaction times from the present work are still faster than  $k_{\text{RRKM}}$ , which is what is found (see below).

For  $\nu_9 = 5$  energy redistribution can occur on time scales of several hundred picoseconds before the actual reaction takes place. The initial energy is predominantly redistributed into angular and torsional degrees of freedom from where it flows back into the dissociative coordinate and the reaction takes place. This differs from simulations with 61.5 kcal/mol of available excess energy where only partial energy redistribution is observed. Under such circumstances, impulsive H-hopping is the dominant process. The same observation was made in previous simulations where partial energy redistribution was observed, and H-hopping was the proposed mechanism for dissociation in  $\text{H}_2\text{SO}_4$ .<sup>15</sup> For events more characteristic of the most probable reaction time ( $\tau > 40$  ps), partial to almost complete IVR is observed. Considering the torsional degrees of freedom, isomerization from  $C_2$  to  $C_s$  was observed in all trajectories. Figure 8 (bottom) shows a free energy barrier of 4.3 kcal/mol, which is in good agreement with previous studies of rotamers for  $\text{H}_2\text{SO}_4$ .<sup>35</sup> The isomerization between  $C_2$  and  $C_s$  is more pronounced when IVR occurs following vibrational excitation. On the contrary, when no IVR is seen and the kinetics of the reaction is faster, isomerization dynamics is virtually absent (see black curves in columns A and D, Figure 8 bottom).

Two different regimes of the reaction can therefore be distinguished: (i) Impulsive, rapid, and nonstatistical H-transfer where one of the hydroxyl-H atoms is transferred to the other OH resulting in  $\text{SO}_3$  and  $\text{H}_2\text{O}$  formation. This model was the predominant channel found in earlier work (4 trajectories out of 98 in total based on PM3/MD simulations).<sup>15</sup> (ii) Photodissociation following partial or complete intramolecular vibrational energy redistribution (IVR), whereby the vibrational energy is redistributed among all remaining degrees of freedom of  $\text{H}_2\text{SO}_4$ . It should be emphasized that with the present approach the distribution of reaction times can be converged arbitrarily well because a sufficient number of trajectories can be run. Previous work required one to resort to separate computations for the two processes and did not allow one to treat them on an equal footing.<sup>15,17</sup> The RRKM rates from the statistical treatment are  $1.7 \times 10^{-7}$ ,  $8.9 \times 10^{-9}$ , and  $1.1 \times 10^{-9}$  s for  $\nu_9 = 4, 5,$  and 6 quanta. This compares with average reaction times of  $\tau_d = 7 \times 10^{-10}$ ,  $2 \times 10^{-10}$ , and  $4 \times 10^{-11}$  from the present work. However, the average reaction times  $\tau_d$  do not correspond to the statistical limit. Rather, complete energy redistribution requires probably at least 1 order of magnitude more time (see Figure 8), which brings the RRKM and explicitly simulated decomposition

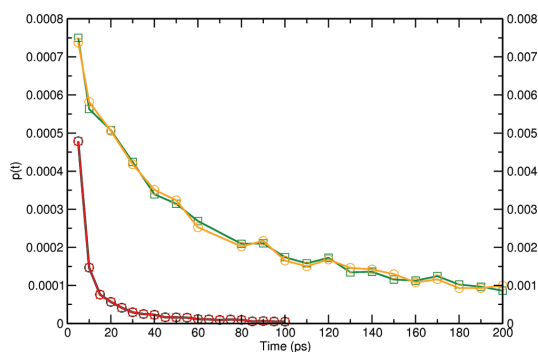


**Figure 8.** Coordinates of sulfuric acid in the gas phase as a function of time. Black traces correspond to the vibrationally excited state, and red corresponds to the equilibrated system (i.e., before excitation). Columns A–C are for excitation of  $\nu_9 = 5$ ; columns D–F are for  $\nu_9 = 6$ . Columns show reaction time in the next order: (A)  $\tau = 6$  ps, (B)  $\tau = 244$  ps, (C)  $\tau = 589$  ps, (D)  $\tau = 7$  ps, (E)  $\tau = 50$  ps, and (F)  $\tau = 98$  ps. Trajectories with rapid, average, and long reaction times are shown for both cases (for details see text). From top to bottom, the amplitudes of the O6–H7 bond, S1–O4 bond, S1–O6–H7 angle, and O6–S1–O4 angle are shown. Rows 5 and 6 report Gibbs free energy as a function of the dihedral angles (H5–O4–S1–O6 and H7–O6–S1–O4).

times within about 1 order of magnitude. A possible reason for the remaining difference is in the way the different degrees of freedom couple and exchange energy with each other, which is included in the atomistic simulations but not in the RRKM theory.

In order to characterize whether structure or dynamics drives the process, the following situations were considered in more detail. 100 trajectories with excitation of  $\nu_9 = 5$  were randomly selected with reaction times longer than 500 ps. In such trajectories, IVR is essentially complete. Likewise, 100 trajectories from simulations with 61.5 kcal/mol of excess energy were randomly

chosen provided that they showed dissociation within 60 ps of vibrational excitation. In such trajectories IVR is incomplete. For every trajectory, coordinates and velocities of all atoms were recorded 5 ps prior to photodissociation. Next, (i) every structure was started with all 100 velocities, and (ii) every velocity file was used with all 100 structures and the reaction time was recorded. This gives a total of 10 000 combinations, and every one of them was propagated for 200 ps for  $\nu_9 = 5$  and for 100 ps for  $\nu_9 = 6$ . Figure 9 shows that in both cases the reaction probability is not skewed toward either structure or dynamics.



**Figure 9.** Dissociation probability as a function of time. The green and red curves report reaction times when individual coordinates are combined with all velocities. The orange and black curves correspond to combining individual velocities with all coordinates. Red and black curves refer to  $\nu_9 = 6$ , in the same way orange and green curves refer to  $\nu_9 = 5$ . For details see text.

The decay times for  $\nu_9 = 6$  are about an order of magnitude more rapid compared to those for  $\nu_9 = 5$  (see Figure 9). This again suggests that for  $\nu_9 = 5$  dissociation is predominantly driven by IVR, whereas for  $\nu_9 = 6$  H-hopping is the more likely process. However, no preference of a “dynamically” over a “structurally” driven process could be found.

## V. CONCLUSION

In conclusion, extensive ARMD simulations with tailored force fields support the hypothesis that in the gas phase (and at higher altitudes in the atmosphere), vibrational excitation of  $\nu_9 > 4$  leads to photodissociation of  $\text{H}_2\text{SO}_4$  into  $\text{SO}_3 + \text{H}_2\text{O}$ . Typical average reaction times for  $\nu_9 \geq 4$  are on the picosecond time scale or faster. Given that the parametrized force field overestimates the reaction barrier by  $\sim 7$  kcal/mol, all reaction times reported here are probably 2 orders of magnitude too slow. However, this does not affect the general conclusion that dissociation through vibrational excitation of  $\nu_9 = 4$  or higher OH-overtone in  $\text{H}_2\text{SO}_4$  is a process that occurs on the picosecond time scale. Mechanistically, anything between impulsive H-transfer (sub-picoseconds) to almost complete IVR with subsequent dissociation (picoseconds to nanoseconds) are observed. The time scales found from the present simulations agree qualitatively with those from separate previous investigations and are, to our knowledge, the first that cover the entire range of dynamics following overtone excitation of  $\text{H}_2\text{SO}_4$ . Such types of simulations hold much promise to provide atomistically detailed pictures of reactive processes, not only for large biomolecular systems but also for activated processes in the gas phase and in clusters.<sup>22</sup>

## ■ ASSOCIATED CONTENT

**Supporting Information.** Details on the force field parameters and distribution times from the SCC-DFTB/MD simulations. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ ACKNOWLEDGMENT

The authors gratefully acknowledge financial support from the Swiss National Science Foundation through Grants 200021-117810

(to M.M.) and 200021-124936 (NCCR Molecular Ultrafast Science and Technology (MUST)). We thank Dr. S. Lutz for stimulating discussions.

## ■ REFERENCES

- (1) Clement, F.; Ford, I. J. *Atmos. Environ. A: Gen* **1999**, *33*, 475–487.
- (2) Larson, L. J.; Kuno, M.; Tao, F. M. *J. Chem. Phys.* **2000**, *112*, 8830–8838.
- (3) Hofmann, D. J.; Rosen, J. M. *Nature* **1982**, *297*, 120–124.
- (4) Raes, F.; Dingenen, R. V.; Vignati, E.; Wilson, J.; Pataud, J.-P.; Seinfeld, J. H.; Adams, P. *Atmos. Environ. A: Gen* **2000**, *34*, 4215–4240.
- (5) Duplissy, J.; et al. *Atmos. Environ. A: Gen* **2010**, *10*, 1635–1647.
- (6) Rinsland, C.; Gunson, M.; Ko, M.; Weisenstein, D.; Zander, R.; Abrams, M.; Goldman, A.; Sze, N.; Yue, G. *Geophys. Res. Lett.* **1995**, *22*, 1109–1112.
- (7) Hintze, P. E.; Kjaergaard, H. G.; Vaida, V.; Burkholder, J. B. *J. Chem. Phys.* **2003**, *107*, 1112–1118.
- (8) Vaida, V. *J. Phys. Chem. A* **2009**, *113*, 5–18.
- (9) Seinfeld, J.; Pandis, S. *Atmospheric Chemistry and Physics From Air Pollution to Climate Change*; John Wiley and Sons: New York, 1998.
- (10) Vaida, V.; Kjaergaard, H. G.; Hintze, P. E.; Donaldson, D. J. *Science* **2003**, *299*, 1566–1568.
- (11) Donaldson, D.; Frost, G.; Rosenlof, K.; Tuck, A. F.; Vaida, V. *Geophys. Res. Lett.* **1997**, *24*, 2651–2654.
- (12) Donaldson, D. J.; Tuck, A. F.; Vaida, V. *Chem. Rev.* **2003**, *103*, 4717–4729.
- (13) Feierabend, K. J.; Havey, D. K.; Brown, S. S.; Vaida, V. *Chem. Phys. Lett.* **2006**, *420*, 438–442.
- (14) Morokuma, K.; Muguruma, C. *J. Am. Chem. Soc.* **1994**, *116*, 10316–10317.
- (15) Miller, Y.; Gerber, R. B. *J. Am. Chem. Soc.* **2006**, *128*, 9594–9595.
- (16) Gerber, R. B. University of California, Irvine, School of Physical Sciences; Private communication, 2011.
- (17) Miller, Y.; Gerber, R. B.; Vaida, V. *Geophys. Res. Lett.* **2007**, *34*, 1–5.
- (18) Carpenter, B. K. *Annu. Rev. Phys. Chem.* **2005**, *56*, 57–89.
- (19) Danielsson, J.; Meuwly, M. *J. Chem. Theory. Comput.* **2008**, *4*, 1083–1093.
- (20) Nutt, D.; Meuwly, M. *Biophys. J.* **2006**, *90*, 1191–1201.
- (21) Mishra, S.; Meuwly, M. *Biophys. J.* **2006**, *96*, 2105–2118.
- (22) Lutz, S.; Meuwly, M. *ChemPhysChem* **2011**, DOI: 10.1002/cphc.201100575.
- (23) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (24) Nguyen, P. N.; Stock, G. *J. Chem. Phys.* **2003**, *119*, 11350–11358.
- (25) Meuwly, M.; Müller, A.; Leutwyler, S. *Phys. Chem. Chem. Phys.* **2003**, *5*, 2663–2672.
- (26) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (27) Frisch, M. J. et al. *Gaussian 03*, revision C.01; Gaussian, Inc.: Wallingford CT, 2004.
- (28) Law, M.; Hutson, J. *Comput. Phys. Commun.* **1997**, *102*, 252–268.
- (29) Devereux, M.; Meuwly, M. *J. Chem. Inf. Model.* **2010**, *50*, 349–357.
- (30) Halgren, T.; Lipscomb, W. *Chem. Phys. Lett.* **1977**, *49*, 225–232.
- (31) Gonzales, C.; Schlegel, H. *J. Chem. Phys.* **1989**, *90*, 2154–2161.
- (32) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.
- (33) Demaison, J.; Herman, M.; Lievin, J.; Rudolph, H. D. *J. Phys. Chem. A* **2007**, *111*, 2602–2609.
- (34) Lohr, L. *J. Mol. Struct. (THEOCHEM)* **1982**, *4*, 221–227.
- (35) Havey, D. K.; Feierabend, K. J.; Vaida, V. *J. Mol. Struct. (THEOCHEM)* **2004**, *680*, 243–247.
- (36) Sabin, J.; Brandas, E.; Goodsite, M. E.; Johnson, M. *Advances in Quantum Chemistry. Applications of Theoretical Methods to Atmospheric Science*; Elsevier: New York, 2008.
- (37) Sedo, G.; Schltz, J.; Leopold, R. *J. Mol. Spectrosc.* **2008**, *251*, 4–8.

- (38) Ruscic, B.; Feller, D.; Dixon, D.; Peterson, K.; Harding, L.; Asher, R.; A., W. *J. Phys. Chem. A* **2001**, *105*, 1–4.
- (39) Fanourgakis, S.; Xantheas, S. *J. Chem. Phys.* **2008**, *128*, 07450601–07460611.
- (40) Arstila, H.; Laasonen, K.; Laaksonen, A. *J. Chem. Phys.* **1998**, *108*, 1031–1039.
- (41) Miller, Y.; Chaban, G.; Finlayson-Pitts, B.; Gerber, R. *J. Phys. Chem. A* **2006**, *110*, 5342–5354.

**Supporting Information for: Reactive Molecular Dynamics study of sulfuric acid (H<sub>2</sub>SO<sub>4</sub>) Dissociation.**

Juvenal Yosa<sup>a</sup> and Markus Meuwly<sup>a,\*</sup>

*<sup>a</sup>Department of Chemistry, University of Basel,  
Klingelbergstrasse 80, 4056 Basel, Switzerland*

(Dated: November 29, 2011)

In the following, force field parameters and additional figures for validation of the simulations are provided.

TABLE I: Bond, angles, torsional and nonbonded parameters for H<sub>2</sub>SO<sub>4</sub>

Bonds	$k$ [kcal/mol/Å <sup>2</sup> ]	$r_{\text{eq}}$ [Å]	
S1-O2	936.140	1.430	
S1-O3	936.140	1.430	
S1-O4	356.340	1.629	
Angles	$k_{\theta}$ [kcal/mol/rad <sup>2</sup> ]	$\theta$ [°]	
O3-S1-O2	127.462	114.037	
O4-S1-O2	111.110	99.776	
O4-S1-O3	111.110	99.776	
O4-S1-O6	103.490	96.771	
H5-O4-O1	94.304	115.043	
O6-S1-O3	111.110	99.776	
O6-S1-O2	111.110	99.776	
H7-O6-S1	94.304	115.04	
Dihedral angles	$k_{\chi}$ [kcal/mol]	n	$\delta$ [°]
O3-S1-O4-H5	0.600	2	180.00
O3-S1-O6-H7	0.600	2	180.00
O2-S1-O4-H5	0.600	2	180.00
O2-S1-O6-O7	0.600	2	180.00
O4-S1-O6-O7	0.430	3	0.00
H5-O4-S1-O6	0.430	3	0.00
Atom Type [VdW]	charge [ $e^{-}$ ]	$\epsilon$ [kcal/mol]	$R_{\text{min}}/2$
S1	1.603	-0.470	2.000
O2	-0.631	-0.152	1.700
O3	-0.631	-0.152	1.700
O4	-0.465	-0.152	1.770
H5	0.295	-0.046	0.225
O6	-0.465	-0.152	1.770
H7	0.295	-0.046	0.225

TABLE II: Parameters for the Morse potentials (see Eq. (1) in the manuscript) for the various bonds in the ARMD simulations, fit to MP2 calculations.  $D_e$  is the dissociation energy,  $\beta$  controls the width of the potential and  $r_e$  is the equilibrium distance.

Bond	$D_e$ [kcal/mol]	$\beta$ [ $\text{\AA}^{-1}$ ]	$r_e$ [ $\text{\AA}$ ]
S1-O4	85.112	2.165	1.629
O6-H7	127.524	2.392	0.975
O4-H7	117.024	0.937	0.948

TABLE III: Bond, angles, torsional and non bonded parameters for the product state  $\text{SO}_3 + \text{H}_2\text{O}$

Bonds	$k$ [kcal/mol/ $\text{\AA}^2$ ]	$r_{\text{eq}}$ [ $\text{\AA}$ ]	
S1-O2	936.140	1.430	
S1-O3	936.140	1.430	
S1-O4	936.140	1.430	
Angles	$k_\theta$ [kcal/mol/rad <sup>2</sup> ]	$\theta$ [ $^\circ$ ]	
O3-S1-O2	127.464	120.000	
O4-S1-O2	127.464	120.000	
O4-S1-O3	127.464	120.000	
O4-S1-O6	55.000	104.520	
Atom Type [VdW]	charge [ $e^-$ ]	$\epsilon$ [kcal/mol]	$R_{\text{min}}/2$
S1	1.431	-0.470	2.000
O2	-0.477	-0.152	1.700
O3	-0.477	-0.152	1.700
O4	-0.477	-0.152	1.700
H5	0.234	-0.046	0.225
O6	-0.468	-0.152	1.770
H7	0.234	-0.046	0.225



TABLE IV: Original and modified Lennard-Jones ranges. For details see the main text.

Atom	$R_{\min}/2$ [ $\text{\AA}$ ]	modified $R_{\min}/2$ [ $\text{\AA}$ ]
S1	2.00	1.30
O4	1.77	1.22
O6	1.77	1.22
H7	0.23	0.22

TABLE V: Experimental and calculated (Force Field and SCC-DFTB) frequencies. the main text. <sup>a</sup> Hintze et al. (*J. Phys. Chem. A.* 2003, 107, 1112-1118)

Assignment	$\nu$ [ $\text{cm}^{-1}$ ] Literature <sup>a</sup>	$\nu$ [ $\text{cm}^{-1}$ ] Force Field	$\nu$ [ $\text{cm}^{-1}$ ] SCC-DFTB
SO <sub>2</sub> bend	550.5	428.8	501.7
SO <sub>2</sub> rock		525.7	517.4
S(OH) <sub>2</sub> sym str	834.1	803.2	718.5
S(OH) <sub>2</sub> asym str	891.4	861.7	744.1
SOH sym bend		978.2	978.0
SOH asym bend	1157.1	1027.6	992.9
SO <sub>2</sub> sym str	1220.1	1333.3	1165.6
SO <sub>2</sub> asym str	1464.7	1603.2	1396.7
OH str	3609.2	3700.8	3793.4

### Figure Captions:

Figure S1 Evaluation of different  $\sigma$  values for one reactive coordinate. The plot shows distances (O6-H7 and S1-O4) - between atoms involve in bond-breaking for  $\text{H}_2\text{SO}_4$  hydrolysis - as a function of the simulation time. The reference value, in black, correspond to the default  $\sigma$  value for VdW parameters in CHARMM. A. O6-H7 bond and B. S1-O4.

Figure S2 Distribution of dissociation times for  $v_9 = 6$  using SCC-DFTB.

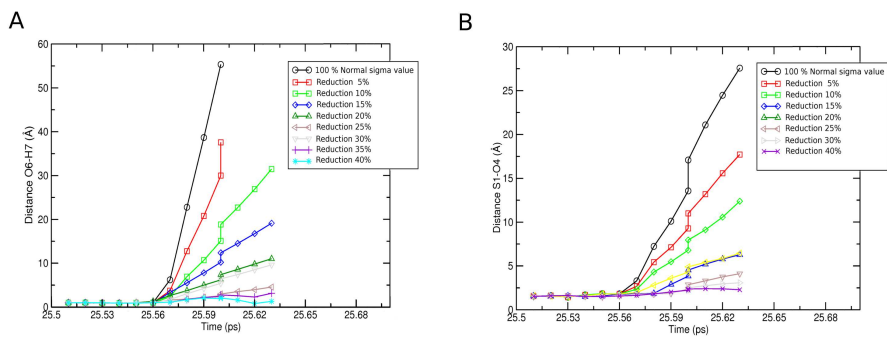


FIG. 1: Yosa and Meuwly

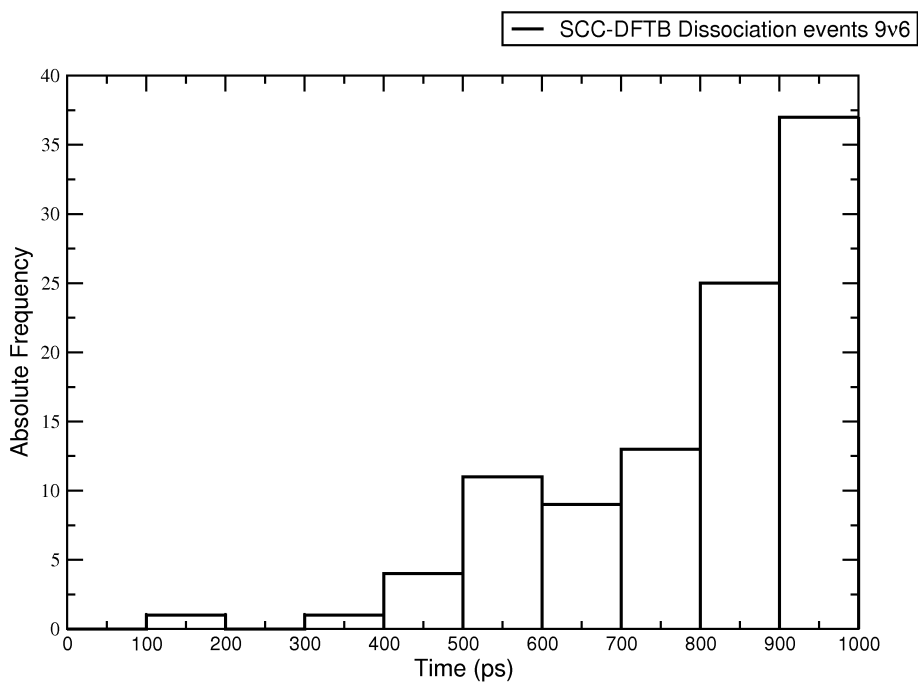


FIG. 2: Yosa and Meuwly

**3.1.1** Multi-Surface ARMD

---

**3.1.2** abstract

---

Adiabatic reactive molecular dynamics (ARMD) is a surface-crossing algorithm for modelling chemical reactions in classical molecular dynamics simulations using empirical force fields. As the ARMD Hamiltonian is time-dependent during crossing, it allows only approximate energy conservation. In the present work the range of applicability of conventional ARMD is explored and a new, multi-surface ARMD (MS-ARMD) method is presented, implemented in CHARMM and applied to the vibrationally induced photodissociation of sulfuric acid ( $\text{H}_2\text{SO}_4$ ) in the gas phase. For this, an accurate, global potential energy surface (PES) involving 12  $\text{H}_2\text{SO}_4$  and 4  $\text{H}_2\text{O}+\text{SO}_3$  force fields fitted to MP2/6-311G++(2d,2p) reference energies is employed. The MS-ARMD simulations conserve total energy and feature both, intramolecular H-transfer reactions and water elimination. Using mild assumptions on the PESs in the crossing region it is found that conventional ARMD can approximately conserve total energy for limiting cases. In one of them, the reduced mass of the system is large which often occurs for simulations of solvated, biomolecular systems. On the other hand, MS-ARMD is a general approach for modeling chemical reactions including gas-phase, homogeneous, heterogeneous and enzymatic catalytic reactions while conserving total energy in atomistic simulations.

**3.1.3** Introduction

---

Following the dynamics and energetics of chemical reactions is of fundamental importance in all branches of chemistry and biology. Experimentally, reaction mechanisms often need to be inferred from phenomenological kinetic models because simultaneous determination of the physical trajectories of the participating particles and their energetics is typically not possible. Hence, computational methods play an essential role in elucidating, at atomic resolution, the possible reaction pathways followed for a particular chemical reaction. Analysis of the atoms' trajectories ultimately provides deeper insight about the sequence of

events that connect reactants (educts) and products and are therefore the most elementary level at which reaction mechanisms can be analyzed and understood.

The most rigorous computational treatment of such processes uses quantum methods for the electronic structure and nuclear dynamics. However, this is usually impractical due to the computational requirements. Alternatively, the nuclear motion can be treated quantum mechanically while using a parametrized potential energy surface (PES). This has been successfully done for a number of topical systems.<sup>117–122</sup> As a thermal rate constant is an ensemble average over a large number of initial conditions, a statistically significant number of trajectories needs to be run and analyzed. Therefore, approaches which treat the nuclear dynamics at the classical level have become an attractive and often meaningful alternative. Such simulations can, under favourable circumstances, be run by resorting to quantum mechanical (QM)<sup>123–125</sup> or mixed quantum/molecular mechanics (QM/MM) treatments.<sup>126–128</sup> Alternatively, if empirical and fully-dimensional force fields of sufficient accuracy can be parametrized, they offer a viable means to exhaustively sample phase space from which meaningful rate parameters can be determined.<sup>64, 129–133</sup>

Over the past few years a number of approaches have been developed which use empirical force fields for following chemical reactions in the gas phase and in solution. One of them is the empirical valence bond (EVB) technique which was particularly relevant to (proton transfer) reactions in solution.<sup>134</sup> The generalization of EVB to multi-state EVB has played an important role for investigating proton transfer in solution.<sup>135</sup> The EVB Hamiltonian usually consists of two or more diagonal terms which are force field expressions for all states of interest. The off-diagonal terms are coupling matrix elements which depend parametrically on one (or several) internal coordinates of the system.<sup>136</sup> This introduces a dependence of the reaction on one (or several) coordinates which is not always desirable. Alternatively, a chemical reaction can be followed along time as the progression coordinate, which is the situation encountered in experiments.<sup>137</sup> This is the purpose of adiabatic reactive molecular dynamics (ARMD) which was originally developed for reactions in the condensed phase.<sup>64, 130, 138</sup> More recently, ARMD has also been applied to gas-phase systems such as the vibrationally induced photodissociation of sulfuric acid ( $\text{H}_2\text{SO}_4$ ). Here, the excitation of a higher overtone ( $\nu_9 \leq 4$ ) of a local OH stretch vibration can lead to photodissociation into water and

sulfur-trioxide ( $\text{H}_2\text{O} + \text{SO}_3$ ) on the pico- to nanosecond time scale.<sup>66,139</sup> However, the ARMD-trajectories were not suitable for final state analysis of the reaction products because they were based on an explicitly time-dependent Hamiltonian which does not conserve total energy during crossing.

Because significant energy violation was observed during crossing in gas-phase reactions, ARMD was modified to multi-surface ARMD (MS-ARMD) which leads to energy conservation in reactive molecular dynamics simulations.<sup>140</sup> It has already been successfully applied to the charge transfer reaction  $\text{N}_2^+ + \text{N}_2 \rightarrow \text{N}_2 + \text{N}_2^+$  and explain recent experimental findings based on an accurate global adiabatic potential energy surface (PES).<sup>140</sup> The reaction at the investigated conditions involves the formation of a vibrationally highly excited  $[\text{N}-\text{N}\cdots\text{N}-\text{N}]^+$  complex with lifetimes on the 1 to 100 picosecond time scale during which several bond rearrangements take place. The study required 24 force fields (FFs) to construct a permutationally invariant, global PES capable of describing all atom arrangements within the complex which can evolve into the various final states.<sup>140</sup> That work served as a validation study and the general underlying concepts are discussed in the present work together with an application to the decomposition reaction of  $\text{H}_2\text{SO}_4$  involving competing pathways.

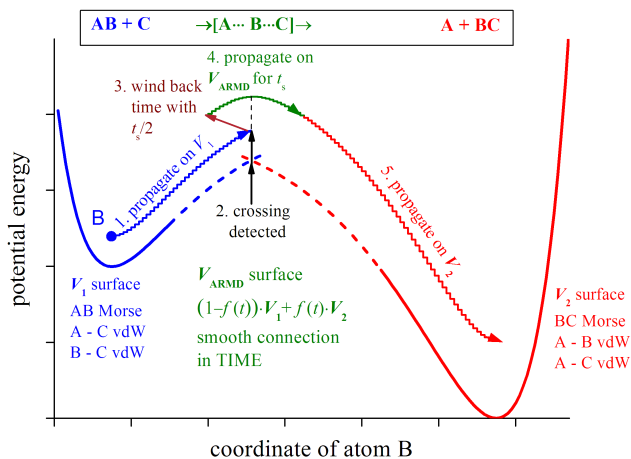
The present work is structured as follows. In chapter II, the original ARMD implementation is explored for its applicability and generalized to Multi-Surface ARMD (MS-ARMD) in chapter III. Next, the methodology is applied to the vibrationally induced photodissociation of  $\text{H}_2\text{SO}_4$  which is a suitable test system for several reasons. First, it offers the possibility of investigating competing pathways, including intramolecular H-transfer and water elimination. Both processes have been found in earlier MD simulations at the semiempirical level.<sup>139</sup> Second, the process is expected to occur on the picosecond to nanosecond time scale which allows us to exhaustively sample phase space prior to excitation and to follow the subsequent dissociation dynamics. Finally, vibrationally induced photodissociation is a physically interesting process with potential atmospheric importance as it might be a source of  $\text{SO}_3$  in the stratosphere.<sup>141-144</sup> Hence, the system considered here provides the necessary balance of complexity and tractability required for a rigorous validation of MS-ARMD. Finally, comparisons with other existing methods are made and conclusions are drawn.

**3.1.4** MS-ARMD and its Applicability

Adiabatic reactive molecular dynamics (ARMD) is a surface-crossing method that allows one to follow bond-breaking and forming processes using empirical force fields.<sup>64,65</sup> In ARMD at least two parametrized potential energy surfaces (PESs),  $V_1$  for the educt and  $V_2$  for the product states, are considered. The adiabatic dynamics of the nuclei takes place on the lowest PES while the energy of the higher state(s) is also determined. Whenever the energy of the current state  $I$  equals that of a higher PES, state  $II$ , the simulation is restarted a few time steps prior to the detected crossing. Following this, the method interpolates on-the-fly between the two PESs.<sup>64</sup> During a finite time interval  $t_s$ , called switching time (typically a few fs), which is largely determined by the process in question, the PESs are mixed in different proportions by multiplying them with a suitable time-dependent switching function  $f(t)$  (e.g. a tanh function). At the beginning of the mixing the system is fully in state  $I$  ( $f(t) = 0$ ), while at the end it is fully in state  $II$  ( $f(t) = 1$ ). One of the particular advantages of ARMD is that the reaction develops with time as the progression coordinate, which is a “natural” choice. Thus, no definition of a necessarily approximate reaction coordinate - which even might change as the reaction progresses - is required. The algorithm of ARMD is schematically shown for a collinear atom transfer reaction in Figure 3.1.

As during surface crossing the ARMD potential energy is explicitly time-dependent, the total energy of the system can not be conserved in a strict sense (see below). For large systems (e.g. proteins in solution) total energy was found to be conserved to within  $\approx 1$  kcal/mol which is sufficient for most applications.<sup>64</sup> However, for small systems in the gas phase this is not necessarily true. In cases where only the reaction time is of interest (i.e. the time to reach the transition state), exact energy conservation is not necessarily required. This was the case for vibrationally induced photodissociation of H<sub>2</sub>SO<sub>4</sub>.<sup>66</sup> If, however, several crossings between the states involved can take place or the course of the dynamics after the reaction is of interest - e.g. for a final state analysis - energy conservation becomes crucial.

To explore the range of applicability of the ARMD method, an approximate expression for the energy violation is derived and analyzed in the following.



**Figure 3.1:** Algorithm of Adiabatic Reactive Molecular Dynamics (ARMD) simulation method is schematically shown in five steps for collinear atom transfer reaction (atom B is transferred) where the coordinates of donor (A) and acceptor (C) atoms are fixed. During the crossing the surfaces are switched in time and the Morse bond is replaced by van der Waals (vdW) interactions and vice versa.

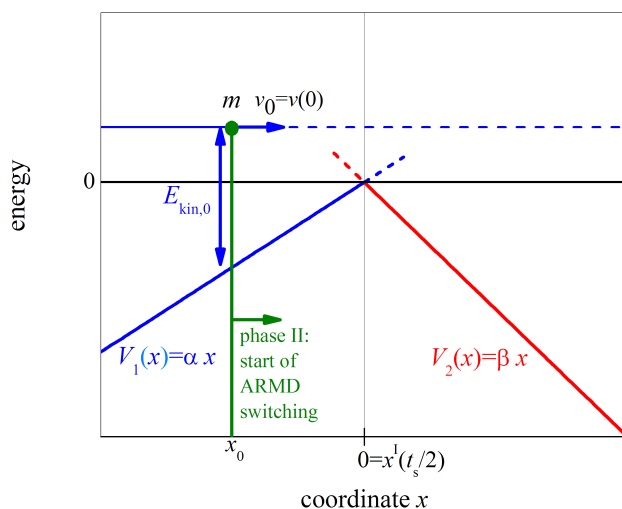
For this, we consider a curve crossing between two one-dimensional PESs  $V_1(x)$  and  $V_2(x)$ . In ARMD the dynamics in the crossing region evolves on a mixed PES:

$$V_{\text{ARMD}}(x, t) = (1 - f(t)) V_1(x) + f(t) V_2(x). \quad (3.1)$$

For the situation in Figure 3.2, the two 1D PESs cross at  $x = 0$ . The following analysis starts from a linear approximation of the PESs which is valid sufficiently close to the crossing point ( $V_1(x) = \alpha x$  and  $V_2(x) = \beta x$ , where  $\alpha > \beta$ ). For the switching function  $f(t)$ , a function which allows analytical work and is defined on the interval  $(0, t_s)$  is  $f(t) = t/t_s$ .

During the first phase (phase I, corresponding quantities are denoted with superscript I), we consider a system with effective mass  $m$  on  $V_1(x)$  (thus  $F^{\text{I}} = -\partial_x V^{\text{I}}(x, t) = -\alpha$ ) with initial position  $x_0 < 0$  and velocity  $v_0 > 0$  (i.e. regular MD, see Figure 3.2). At  $t = 0$   $x_0$  is chosen such that the system will reach the crossing point ( $x = 0$ ) exactly at half of the user-defined switching time ( $t_s/2$ ).





**Figure 3.2:** Simple model for estimating energy violation in ARMD simulations. The system (effective mass  $m$ ) is approaching from the left on PES  $V_1(x)$  (phase I). At  $t = 0$  time it is at  $x_0$  with velocity  $v_0$  and kinetic energy  $E_{\text{kin},0}$ . After crossing is detected at  $x = 0$  the time is rewound by  $t_s/2$  and the dynamics is re-simulated while  $V_1(x)$  is being switched to  $V_2(x)$  during  $t_s$  (phase II).

$$x^{\text{I}}(t_s/2) = 0 = x_0 + \int_0^{t_s/2} dt' \left( v_0 + \int_0^{t'} dt'' \frac{F^{\text{I}}}{m} \right) = x_0 + \frac{v_0}{2} t_s - \frac{\alpha}{8m} t_s^2 \quad (3.2)$$

According to ARMD, when crossing is detected, the time is rewound by  $t_s/2$  and the dynamics is restarted on the mixed PES (phase II). The ARMD potential and its partial derivatives are:

$$V^{\text{II}}(x, t) = \alpha x + \frac{\beta - \alpha}{t_s} x t, \quad (3.3)$$

$$\partial_t V^{\text{II}}(x, t) = \frac{\beta - \alpha}{t_s} x, \quad (3.4)$$

$$F^{\text{II}}(x, t) = -\partial_x V^{\text{II}}(x, t) = -\alpha + \frac{\alpha - \beta}{t_s} t. \quad (3.5)$$

The time evolution of the coordinates of the system during phase II follows from:

$$x^{\text{II}}(t) = x_0 + \int_0^t dt' \left( v_0 + \int_0^{t'} dt'' \frac{F^{\text{II}}(x(t''), t'')}{m} \right) = x_0 + v_0 t - \frac{\alpha}{2m} t^2 + \frac{\alpha - \beta}{6mt_s} t^3. \quad (3.6)$$

The total energy change  $\Delta E^{\text{II}}$  during the crossing period is obtained by integration of  $dE^{\text{II}}/dt = \partial_t V^{\text{II}}(x, t)$ , defined in Equation 3.4, and subsequent transformation from Equation 3.2:

$$\begin{aligned} \Delta E^{\text{II}} &= \int_0^{t_s} dt \partial_t V^{\text{II}}(x(t), t) = \frac{\beta - \alpha}{t_s} \int_0^{t_s} dt x(t) = \\ &(\beta - \alpha) \left( x_0 + \frac{v_0}{2} t_s - \frac{3\alpha + \beta}{24m} t_s^2 \right) = \frac{\beta(\alpha - \beta)t_s^2}{24m} \end{aligned} \quad (3.7)$$

Hence, exact or nearly exact energy conservation,  $\Delta E^{\text{II}} \approx 0$ , can be achieved **a)** if the gradients of the two PESs are the same ( $\alpha \approx \beta$ ), **b)** if the system has a large effective mass. This is often the case in simulations of biomolecular systems, where either the moieties involved in the reaction are heavy or the reaction is coupled with rearrangement of solvation shell, which is composed of several solvent molecules, **c)** for the special case where the second surface has exactly zero or a negligible slope ( $\beta \approx 0$ ) in the crossing region then the positive and negative energy violations will cancel each other, **d)** if the switching time goes to zero. However, for  $t_s \rightarrow 0$  a smooth connection between the PESs is not possible and numerical integrators with finite step size will cause energy violation. As the switching time is the only variable parameter during ARMD simulations, it is useful to consider a criterion for its magnitude to guarantee energy conservation. For this purpose, one may require that energy violation compared to the initial kinetic energy ( $E_{\text{kin},0} = mv_0^2/2$ ) be small, i.e.  $|\Delta E^{\text{II}}| \ll E_{\text{kin},0}$ , which implies the switching time has to fulfill  $t_s \ll \sqrt{12}mv_0/\sqrt{|\beta(\alpha - \beta)|}$ .

---

**3.1.5** Multi-Surface ARMD
 

---

**3.1.5.1** Constructing global PESs with MS-ARMD
 

---

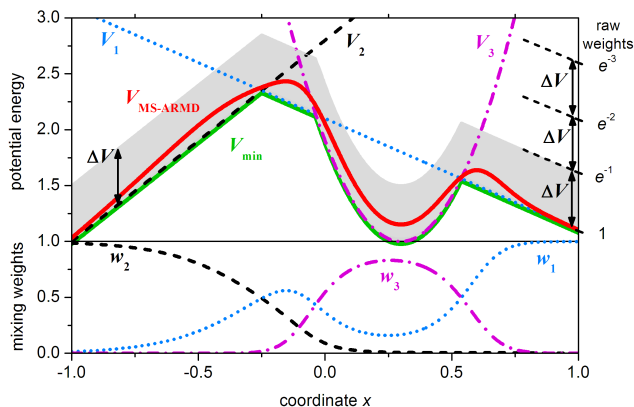
If formally exact energy conservation is required, switching between the PESs should be time-independent. In the following, we present a multi-surface (MS) variant of ARMD in which the weights ( $w_i(\mathbf{x})$ ) of the PESs ( $V_i(\mathbf{x})$ ) are also functions of all Cartesian coordinates  $\mathbf{x}$  and thereby the total energy can be conserved during crossing. According to MS-ARMD, the effective surface is always the lowest-energy surface ( $V_{\min}(\mathbf{x}) = \min_i V_i(\mathbf{x})$ ), except for those geometries where other surfaces get close to it in energy (within several  $\Delta V$ ), in which case the algorithm switches smoothly among them by changing their weights.  $\Delta V$  is a user defined parameter which regulates how the weights of the PESs involved in a crossing reduce with increasing energy separation from the lowest-energy surface at a given point  $\mathbf{x}$ . It plays a similar role as the switching time  $t_s$  in ARMD. The optimum value of  $\Delta V$  is determined during parametrization of the global PES, which is usually done by fitting it to *ab initio* data. The effective potential energy is written as a linear combination of  $n$  PESs with weights  $w_i(\mathbf{x})$ .

$$V_{\text{MS-ARMD}}(\mathbf{x}) = \sum_{i=1}^n w_i(\mathbf{x}) V_i(\mathbf{x}) \quad \text{where} \quad w_i(\mathbf{x}) = \frac{w_{i,0}(\mathbf{x})}{\sum_{i=1}^n w_{i,0}(\mathbf{x})} \quad (3.8)$$

The  $w_i(\mathbf{x})$  are obtained by renormalizing the raw weights  $w_{i,0}(\mathbf{x})$ , calculated by using a simple exponential decay function with a characteristic length of  $\Delta V$ .

$$w_{i,0}(\mathbf{x}) = \exp\left(-\frac{V_i(\mathbf{x}) - V_{\min}(\mathbf{x})}{\Delta V}\right) \quad (3.9)$$

Armed with the normalized weights, the global smoothed PES can be computed according to Equation 3.8. The capabilities of MS-ARMD are illustrated for surfaces in one and two dimensions in Figures 3.3 and 3.4, respectively. Smooth global surfaces are obtained in all cases. The MS-ARMD surface always corresponds to the lowest PES except for regions where other surfaces are close in energy where they mix. Furthermore, the algorithm also leads to a smooth MS-ARMD surface in regions where more than two surfaces are close

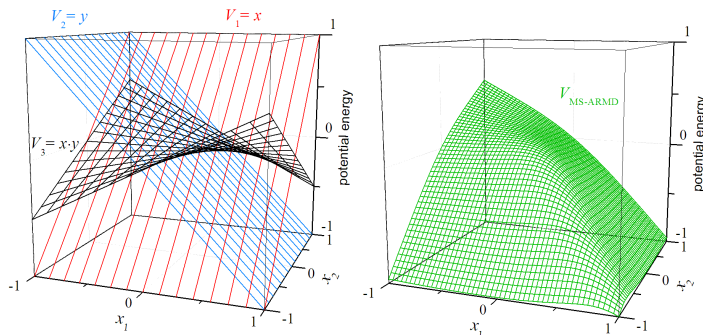


**Figure 3.3:** The MS-ARMD switching method in one dimension ( $x$ ) for three surfaces ( $V_1$ : blue dotted,  $V_2$ : black dashed,  $V_3$ : magenta dot-dashed) defined by linear ( $V_1$  and  $V_2$ ) and quadratic ( $V_3$ ) functions. The effective surface ( $V_{\text{MS-ARMD}}$ : red solid) always follows the lowest-energy surface ( $V_{\text{min}}$ : green solid), except for regions where several surfaces are close in energy (within a few times  $\Delta V = 0.5$  a.u.). Here, the algorithm switches smoothly among them by varying their weights ( $w_1$ ,  $w_2$ , and  $w_3$ ; see lower panel) as a function of  $x$ . Within the grayish band above the minimum energy surface the mixing weights of surfaces drop by a factor of  $e \approx 2.7$ . All quantities are in arbitrary units (a.u.).

in energy, as it is shown in both figures.

Equation 3.9 implies that the weights of surfaces compared to the weight of the momentary lowest-energy surface drop exponentially with increasing energy, that is by a factor of  $e \approx 2.7$  at every  $\Delta V$ . Therefore only those surfaces will have significant weights, whose energy is within a few times  $m$  of  $\Delta V$  (i.e.  $V_i(\mathbf{x}) - V_{\text{min}}(\mathbf{x}) < m\Delta V$ ) from the lowest energy surface  $V_{\text{min}}(\mathbf{x})$ . The term  $V_{\text{min}}(\mathbf{x})$  is necessary to avoid possible overflow and underflow in numerical implementation of the algorithm when the raw weights of the lowest energy surfaces are calculated. This shifting term will cancel from the expressions of weights and the MS-ARMD potential energy (Equation 3.8). Opposed to the raw weights, the final weights are analytical functions (infinitely differentiable) of the surface energies, therefore the resulting MS-ARMD surface ( $V_{\text{MS-ARMD}}(\mathbf{x})$ ) is a smooth function of the coordinates ( $\mathbf{x}$ ).

The smooth MS-ARMD PES can be used both in Molecular Dynamics (MD) and Monte Carlo (MC) simulations. For MD simulations, the first derivative of the PES with respect to the coordinates  $\mathbf{x}$  is required and includes derivatives



**Figure 3.4:** The MS-ARMD switching method in two dimensions ( $x_1$  and  $x_2$ ) for three surfaces ( $V_1$ ,  $V_2$  and  $V_3$ ) defined by simple functions. The scales of axes are in arbitrary units (a.u.). The  $V_{\text{MS-ARMD}}$  surface is always the lowest-energy surface ( $V_{\text{min}}$ ), except for the region where more surfaces have similarly low energy (within a few times  $\Delta V = 0.2$  a.u.), here the algorithm switches smoothly among them as a function of the coordinates.

of the weighting functions as well:

$$\nabla V_{\text{MS-ARMD}}(\mathbf{x}) = \sum_{i=1}^n w_i(\mathbf{x}) \cdot \nabla V_i(\mathbf{x}) + \nabla w_i(\mathbf{x}) \cdot V_i(\mathbf{x}) \quad (3.10)$$

The new algorithm can be applied to the simultaneous switching of any number of PESs. Thus, it is possible to follow reactions that involve different pathways, such as intramolecular proton transfer and water-elimination in H<sub>2</sub>SO<sub>4</sub> which is discussed below.

### 3.1.5.2 Extending flexibility of the MS-ARMD PES

MS-ARMD as presented in the previous section has been implemented into the latest CHARMM version. It allows to add/remove or reparameterize the following standard force field terms: harmonic bond, Morse bond, harmonic angle and Urey-Bradley potential, Fourier series for proper dihedral angle potentials, improper dihedral angle potential, Coulomb potential with point charges, and Lennard-Jones potential also with a separate parameter set for 1-4 neighbors. As the energy of each force field is measured from its own global

minimum, an additive constant can be defined for each force field to level all PESs relative to each other in order to reproduce reaction energies.

These standard potential terms were devised to describe low-energy vibrations and conformational motions of semi-rigid molecules. Even with Morse potentials, simultaneously and accurately describing regions close to the equilibrium *and* the crossing region (i.e. the barrier or asymptotic region of the potential) is not possible. In order to increase the flexibility of the MS-ARMD force field, we introduce a generalized Lennard-Jones potential (equivalent to the Mie-potential<sup>145, 146</sup>) to better describe the van der Waals interaction between atoms around the reaction center, which results in a better description of the reactant and product complexes and the transition state region, too.

$$V_{\text{GLJ}}(r ; n, m, \varepsilon, r_{\min}) = \frac{n\varepsilon}{m-n} \left[ \left( \frac{r_{\min}}{r} \right)^m - \frac{m}{n} \left( \frac{r_{\min}}{r} \right)^n \right] \quad (3.11)$$

The quantities  $\varepsilon$  ( $> 0$ ) and  $r_{\min}$  are the well-depth and the corresponding separation at this energy minimum, respectively - in analogy with the meaning of Lennard-Jones parameters. When a bond between atoms A and B is broken, the bonded (stretch) interactions are replaced by nonbonded interactions (van der Waals and electrostatics). Because in typical force fields the sum of the van der Waals radii of two atoms is significantly larger than the equilibrium bond length between them, unphysical, strong repulsions can occur upon switching. This artifact can also be avoided by using a MIE potential.

Force fields separately optimized for reactant and product states sometimes lead to an unrealistic, high-energy crossing point. In order to best capture the barrier region determined from electronic structure calculations, products of Gaussian and polynomial functions (GAPO) of the energy difference  $\Delta V_{ij}(\mathbf{x}) = V_j(\mathbf{x}) - V_i(\mathbf{x})$  are used. Multiple GAPOs ( $k = 1, \dots, n_{ij}$ ) can be used between any two surfaces ( $i$  and  $j$ ):

$$\Delta V_{\text{GAPO},k}^{ij}(\mathbf{x}) = \exp \left( - \frac{\left( \Delta V_{ij}(\mathbf{x}) - V_{ij,k}^0 \right)^2}{2\sigma_{ij,k}^2} \right) \cdot \sum_{l=0}^{m_{ij,k}} a_{ij,kl} \left( \Delta V_{ij}(\mathbf{x}) - V_{ij,k}^0 \right)^l \quad (3.12)$$

Here,  $V_{ij,k}^0$  and  $\sigma_{ij,k}$  denote the center and the standard deviation of the Gaussian function, respectively. Whenever the energy difference between the

two PESs deviates from  $V_{ij,k}^0$  more than a few times of  $\sigma_{ij,k}$ , the corresponding GAPO function will be negligible. Therefore, if  $V_{ij,k}^0$  and  $\sigma_{ij,k}$  are small, the correction functions will be non-negligible only close to the crossing region and rapidly decaying away from it. The global PES is then a weighted sum of PESs and the sum of correction functions weighted with the sum of the weights of the two corresponding surfaces:

$$V_{\text{MS-ARMD}}(\mathbf{x}) = \sum_{i=1}^n w_i(\mathbf{x})V_i(\mathbf{x}) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n [w_i(\mathbf{x}) + w_j(\mathbf{x})] \sum_{k=1}^{n_{ij}} \Delta V_{\text{GAPO},k}^{ij}(\mathbf{x}) \quad (3.13)$$

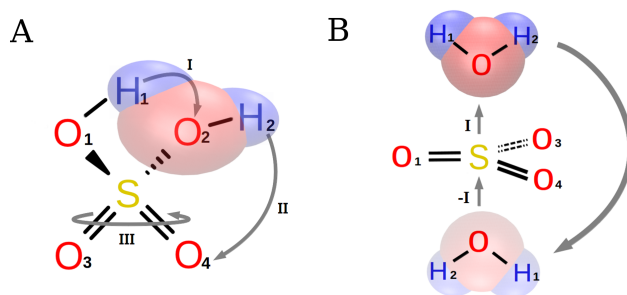
In actual MD simulations, where the gradient of the PES with respect to the coordinates is required, the term from Equation 3.10 needs to be supplemented by the derivative of the second term in Equation 3.13.

### 3.1.6 Photodissociation of H<sub>2</sub>SO<sub>4</sub>

#### 3.1.6.1 Force fields for H<sub>2</sub>SO<sub>4</sub> and H<sub>2</sub>O+SO<sub>3</sub>

To demonstrate the capabilities of the MS-ARMD method we apply it to the vibrationally induced photodissociation of sulfuric acid (H<sub>2</sub>SO<sub>4</sub>).<sup>147</sup> Previous semiempirical on-the-fly MD simulations have already shown that after significant vibrational excitation, H<sub>2</sub>SO<sub>4</sub> can undergo two different chemical transformations including intramolecular H-transfer and water elimination (H<sub>2</sub>SO<sub>4</sub> → H<sub>2</sub>O+SO<sub>3</sub>).<sup>139</sup> In order to follow both processes within MS-ARMD, the global PES needs to describe all states involved (see below). As the current work focuses on the theory and implementation of MS-ARMD, the details of fitting and thoroughly exploring the PES and the dynamics on it will be discussed elsewhere.<sup>148</sup>

For the singlet ground state PES, electronic structure calculations were carried out at the MP2/6-311G++(2d,2p) level with the Gaussian03 suite of programs.<sup>18</sup> The energy at the equilibrium geometry of H<sub>2</sub>SO<sub>4</sub> was chosen as the global zero reference for leveling all *ab initio* energies. Representative



**Figure 3.5:** A) Reactions with the three lowest barriers for sulfuric acid considered in the present work: water elimination (pathway I), H-transfer between  $sp^3$  and  $sp^2$  oxygen atoms (pathway II), or inversion of the bond configuration around the sulfur center (pathway III). B) The inverted configuration of the oxygen atoms can be obtained via the low-energy pathway of elimination and subsequent addition (pathway -I) of water. Re-addition was not observed in our studies, due to the escape of water.

geometries for *ab initio* reference points were collected from MD simulations using previously employed FFs for H<sub>2</sub>O, SO<sub>3</sub> and H<sub>2</sub>SO<sub>4</sub>.<sup>66</sup> Starting from an optimized geometry, the system was heated to 300 K, and equilibrated for 40 ps, which was followed by 10 ns of free microcanonical dynamics. Along the three trajectories, approximately 1000, 1000 and 3000 geometries were selected for H<sub>2</sub>O, SO<sub>3</sub> and H<sub>2</sub>SO<sub>4</sub>, respectively. Furthermore, for H<sub>2</sub>SO<sub>4</sub> a rigid 2D scan of the H<sub>1</sub>-O<sub>1</sub>-S=O<sub>3</sub> and H<sub>2</sub>-O<sub>2</sub>-S=O<sub>4</sub> proper dihedral angles (for labels see Figure 3.5) was performed resulting in a further 1296 (= 36<sup>2</sup>) reference data points which were added to the previous 3000.

The force fields for the three molecules were constructed using Morse potentials for the bonds and harmonic potentials with a Urey-Bradley potential for the angles. For SO<sub>3</sub> the same improper dihedral angle potential was employed symmetrically (i.e.  $\phi[\text{SO}_1\text{O}_3\text{O}_4]$ ,  $\phi[\text{SO}_3\text{O}_4\text{O}_1]$ , and  $\phi[\text{SO}_4\text{O}_1\text{O}_3]$ ) whereas for H<sub>2</sub>SO<sub>4</sub> proper dihedral angle potentials were used. Coulomb interactions at the level of atomic point charges and generalized Lennard-Jones potentials (Equation 3.11) were used between 1-4 and 1-5 neighboring atoms of H<sub>2</sub>SO<sub>4</sub>. The initial guess for the point charges was obtained by fitting the molecular electrostatic potential on a grid using the CHELPG scheme.<sup>149</sup> Initial values for all other parameters were taken or calculated (i.e. for generalized Lennard-Jones) from standard CHARMM parameters. A standalone Fortran code based on the simplex downhill algorithm<sup>54</sup> was used for optimizing all force field and surface-leveling parameters to match the *ab initio* reference



energies. The final RMSDs were 0.02, 0.47 and 1.23 kcal/mol for H<sub>2</sub>O, SO<sub>3</sub> and H<sub>2</sub>SO<sub>4</sub>, respectively.

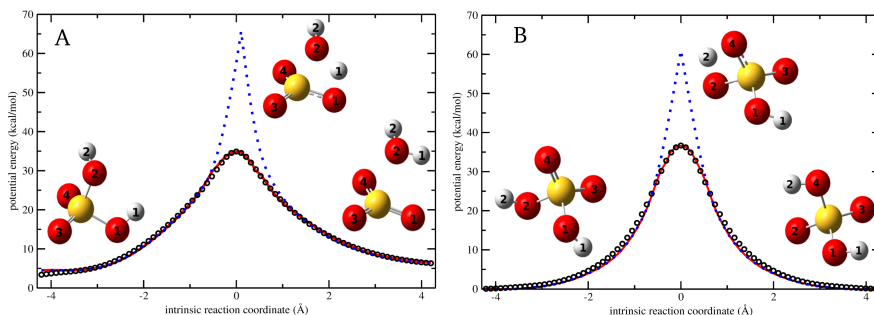
For the H<sub>2</sub>O ··· SO<sub>3</sub> van der Waals complex, separate MD simulations were carried out at 300 K and 2000 geometries were sampled for electronic structure calculations. The nonbonded terms, including parameters for generalized Lennard-Jones potentials and point charges, were optimized while the bonded parameters, refined above, remained unchanged. The final RMSD of this fit was 0.55 kcal/mol. Such accuracy in fitting several thousand reference points has also been found in previous work.<sup>140</sup>

#### 3.1.6.2 Minimum energy paths

For quantitative simulations of the reactions of interest, the global PES also has to describe the barrier regions adequately. One possibility is to determine the reaction paths from the transition state leading to either reactant or product. This was done by first locating the transition states from quadratic synchronous transit (QST) calculations,<sup>150</sup> followed by an intrinsic reaction coordinate (IRC) calculation at the MP2/6-311G++(2d,2p) level of theory. The two minimum energy paths for water elimination and intramolecular H-transfer are shown in Figure 3.6. The reference *ab initio* calculations yield a barrier height of 32.0 kcal/mol for water elimination, and 36.6 kcal/mol for intramolecular H-transfer, which also supports that the latter process is competitive and should be included for realistic simulations. In previous computational studies the energy barrier for water elimination has been found to be in the range of 32-40 kcal/mol depending on the level of theory used.<sup>66, 142, 151</sup> The system can be involved in another non-dissociative chemical transformation, namely the inversion of the bond configuration around the sulfur center, see Figure 3.5. A corresponding transition state calculation finds a barrier height of 90.4 kcal/mol which makes this process highly improbable compared to the other two transformations. Hence, it is not investigated any further.

Figure 3.6 compares the energy along the IRCs from the electronic structure calculations with those from the MS-ARMD force field with  $\Delta V = 4$  kcal/mol. Evidently, the global, separately parametrized PES for H<sub>2</sub>SO<sub>4</sub> and H<sub>2</sub>O ··· SO<sub>3</sub>

leads to barrier heights 20-30 kcal/mol higher than the reference data (dotted line). To better capture the barrier region, GAPO functions were used. For this, the parameters in the GAPO functions, the switching parameter  $\Delta V$  and the overall shift of the FFs were optimized. Using 3 GAPOs with 1<sup>st</sup>-order polynomials for the H<sub>2</sub>O elimination, and 2 GAPOs with 2<sup>nd</sup>-order polynomials for the H-transfer reactions, the energy profile along the reference MEP can be reproduced very closely (RMSD = 0.55 kcal/mol), as is shown in Figure 3.6 (red continuous line).



**Figure 3.6:** Potential energy along the IRC for water elimination (Figure A) and for intramolecular H-transfer (Figure B). Black circles are the *ab initio* reference data, the blue dotted curve corresponds to MS-ARMD surface ( $\Delta V = 5$  kcal/mol) without GAPO, and the red solid curve is the same with GAPOs. Structures of stationary states along the IRCs are also shown in ball-and-stick representation.

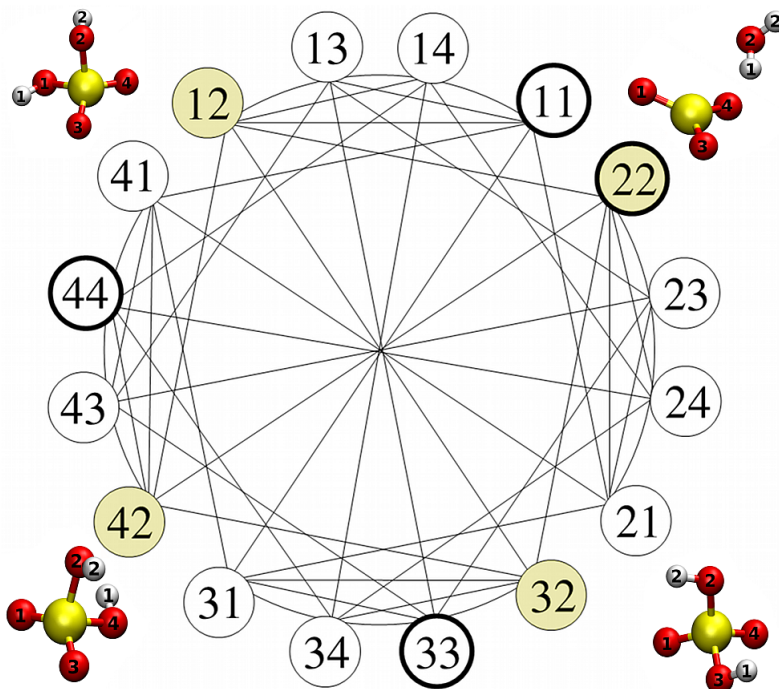
### 3.1.6.3 Construction of the global surface

In force fields, atom indices are used to define the connectivity of atoms. Intramolecular H-transfer in sulfuric acid of a given atom indexing can give 4 additional states which differ only in the oxygen-atom indices to which the transferred H-atom is bound. Water elimination from sulfuric acid of a given atom indexing can lead to H<sub>2</sub>O+SO<sub>3</sub> states with 2 different indexing of atoms. All possible H<sub>2</sub>SO<sub>4</sub> states, and the H<sub>2</sub>O+SO<sub>3</sub> states are separately chemically equivalent to each other, therefore they have to have exactly the same FF parametrization, the only difference being that the indices of some oxygen atoms involved in the H-transfer are swapped and the signs of H-O-S=O dihedral angles potential phases are reversed. The global PES has to describe all these equivalent states with all possible reaction pathways

between them. For this, copies of surfaces and sets of GAPO functions are used.

To enumerate all possible states the oxygen atoms are labelled from 1 to 4, i.e. O1 to O4. Altogether 16 states ( $4^2$ ) can then be distinguished by a two-digit index (e.g. 12, or 34) which denotes to which O-atoms the two H-atoms are bound. Two identical digits (e.g. 11, or 22) represent a state in which the H-atoms are bound to the same O-atom, i.e. an  $\text{H}_2\text{O}$  molecule corresponding to the dissociated state  $\text{H}_2\text{O} + \text{SO}_3$ . In the present gas-phase, single-molecule, free-dynamics studies the water molecule escapes after elimination from  $\text{H}_2\text{SO}_4$  and there is no possibility for recombination with the  $\text{SO}_3$  molecule. In situations where the reaction products can re-react, the global PES would need to support the possibility for addition of  $\text{H}_2\text{O}$  also to the other side of the  $\text{SO}_3$  molecule (“top” and “bottom” side, see Figure 3.5B). The resulting sulfuric acid molecule is a mirror image of the other with respect to the configuration of the oxygen atoms. It would require 12 more copies of the sulfuric acid PESs which are identical except for the phases of the H-O-S=O dihedral angle potentials, which all would reverse their sign. To describe these mirror image configurations without adding further surfaces, we chose the phase shifts of all dihedral angle potentials within sulfuric acid to be either  $0^\circ$  or  $180^\circ$  ( $\equiv -180^\circ$ ). Therefore our final surface can describe the states and reactions of all possible configurations of sulfuric acid.

The network of states and possible reactions between them can be represented as a graph (Figure 3.7). Each state (vertex of the graph) can be involved in six different reactions (degree of the vertex), during which one of their digits changes to one of the three other possibilities (e.g.  $11 \rightarrow 21, 31, 41, 12, 13, 14$ ). In a reaction (edge of the graph) only one of the digits can change, therefore each state can be reached from any other state in two reaction steps (diameter of a graph). Hence, the reaction network corresponds to a symmetric graph with 16 equivalent (!) vertices of degree 6, with a diameter of 2, as shown in Figure 3.7. The 16-state reaction network has altogether 48 edges corresponding to 48 transition states, 24 of which involve H-transfer and 24 of which correspond to  $\text{H}_2\text{O}$ -elimination, respectively. In summary, the global surface, which is invariant to the permutation of chemically equivalent atoms, was generated from 12  $\text{H}_2\text{SO}_4$  surfaces and 4  $\text{H}_2\text{O}+\text{SO}_3$  surfaces and 48 ( $24 + 24$ ) sets of GAPO functions.



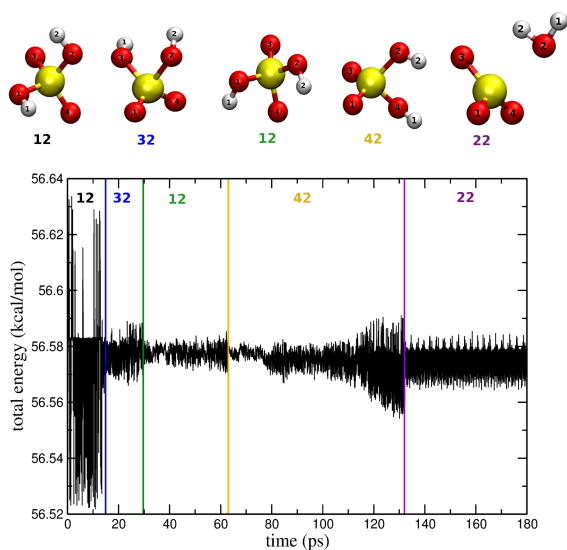
**Figure 3.7:** The network of states (vertices, circles) and possible reactions (lines) for  $\text{H}_2\text{SO}_4$  represented as a symmetric graph. States denoted by circles and labeled by two-digit index, see text. Allowed reactions at the investigated energies are intramolecular H-transfer and  $\text{H}_2\text{O}$ -elimination. The four states observed in a typical trajectory (see Figure 3.8) are highlighted and given as a ball-and-stick representation. Four thick-line circles correspond to states of  $\text{H}_2\text{O} + \text{SO}_3$ .

**3.1.6.4** Vibrational excitation and reaction dynamics

The dynamics of sulfuric acid after OH vibrational overtone excitation was investigated in order to validate the MS-ARMD implementation in CHARMM. Starting from the optimized H<sub>2</sub>SO<sub>4</sub> geometry the molecule was heated to 300 K in 20 ps, equilibrated for 30 ps, and evolved freely for 20 ps, which was followed by vibrational excitation (see below). The equations of motion were propagated using the leapfrog Verlet algorithm with a time step of  $\Delta t = 0.1$  fs. Vibrational excitation was invoked by scaling the instantaneous velocities along the  $\nu_9$  normal mode which corresponds approximately to the OH-local mode even in highly excited states. For the present simulations, 5 quanta of vibrational energy, corresponding to 47.15 kcal/mol, were deposited into this mode<sup>66,152</sup> which is significantly larger than the thermal energy of H<sub>2</sub>SO<sub>4</sub> of  $\approx 9$  kcal/mol at 300 K. After excitation, propagation continued for 500 ps or until water elimination took place. Overall, 50 independent trajectories were run.

Figure 3.8 reports the total energy along a part of one typical such trajectory. Three intramolecular H-transfer processes were observed at times 15 ps, 30 ps, and 63 ps, which were eventually followed by water elimination at 132 ps. Figure 3.8 also shows the transition times and the states visited.

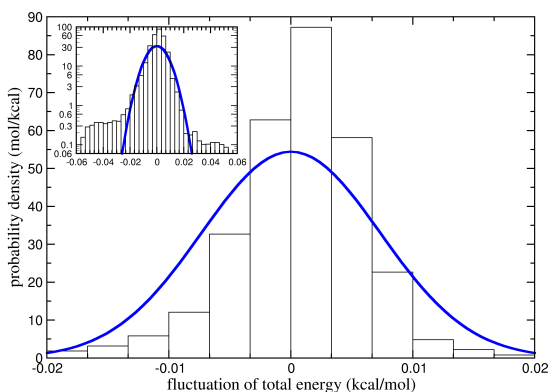
The energy fluctuations of up to  $\pm 0.06$  kcal/mol at the beginning of the excited trajectory are due to the localization of the large excess energy in a single OH stretch mode. The rapid motion of the H atoms can not be described accurately with the stepsize, especially at the short bond length turning point. This could be improved by using an adaptive integrator which was done in our previous work.<sup>140</sup> But even with a Verlet-type integrator, as is commonly used in biomolecular simulations, the degree of energy conservation is sufficient for large-scale applications. Once the excitation energy has been redistributed, particularly into the hindered O=S-O-H rotations, the fluctuations decrease which occurs after 15 ps for the particular trajectory shown in Figure 3.8. After the hindered rotations gain sufficient energy to surpass their barriers the optimal conformations for H-transfer and water elimination can be accessed. Before each crossing an increase of the energy fluctuations can be observed, which originates from a re-localization of energy into one of the OH-stretching modes. Throughout the course of several H-transfer reactions and the final



**Figure 3.8:** The total energy as a function of time for a simulation with several reactive steps. The significant fluctuations during the first 15 ps are caused by the initial preparation of the  $\nu_9 = 5$  state which contains a significant amount of energy in a single OH-stretch mode. The two-digit labels refer to the state of the system, see Figure 3.7, and the ball-and-stick models are corresponding representative structures.

water elimination step no drift in the total energy occurs which validates the present implementation in CHARMM.

For further statistical characterization a histogram of the root mean square deviation ( $E_i - \langle E_i \rangle$ ,  $i = 1$  to 50) for all 50 trajectories was constructed and analyzed. This probability distribution ( $\rho(E - \langle E \rangle)$ ) is shown in Figure 3.9. By construction,  $\rho$  has an expectation value of  $\mu = 0$  and the standard deviation is  $\sigma = 0.0073$  kcal/mol. The corresponding Gaussian distribution  $\rho(x) = \exp[-(x - \mu)^2 / 2\sigma^2] / \sqrt{2\pi}\sigma$  is also shown in Figure 3.9. In summary, we conclude that stable, energy conserving simulations can be carried out with MS-ARMD on the nanosecond time scale even for such a highly excited system.



**Figure 3.9:** Distribution of total energy fluctuations based on simulations of the free dynamics of sulfuric acid after OH overtone excitation ( $\nu_9$ ) for 50 independent initial conditions sampled at 300 K. Standard deviation  $\sigma=0.0073$  kcal/mol was empirically estimated from the sample, and the corresponding Gaussian distribution is also plotted. Note, the expectation value ( $\mu$ ) of the fluctuation is 0 kcal/mol by construction.

### 3.1.7 Discussion and Conclusions

In the present manuscript we have presented a multi-surface, energy conserving generalization of the ARMD algorithm. It allows to construct global, reactive PESs by smoothly connecting parametrized FFs involving several states. The MS-ARMD method was implemented in the latest version of CHARMM in a general way.

Other reactive molecular dynamics methods based on empirical force fields have been recently and thoroughly reviewed.<sup>133</sup> A general and widely used procedure is the Empirical Valence Bond Model (EVB) which requires the solution of an  $n \times n$  eigenvalue problem if  $n$  states are involved to obtain the effective potential energy surface. This can be done analytically and efficiently up to four surfaces. If more states are close in energy, the relevant quantities need to be evaluated numerically which is computationally more expensive. On the contrary, for MS-ARMD the time for evaluating energies and forces scales linearly with the number of low-lying surfaces. By construction, the

MS-ARMD surface always passes through the crossing point, which is not necessarily the case for EVB surfaces. Furthermore, with the help of Gaussian and polynomial product functions, the MS-ARMD PES is very flexible to reproduce reference data in the barrier region. similar to the Chang-Miller<sup>153</sup> and the distributed Gaussian<sup>154</sup> extension of EVB, to adjust the frequencies of stationary states or the shape of PES around any selected geometry locally.

To validate the implementation, H-transfer and water elimination in  $\text{H}_2\text{SO}_4$  induced by high vibrational excitation of the OH stretching mode was investigated. The MS-ARMD algorithm was used for constructing a permutationally invariant PES for sulfuric acid based on 12  $\text{H}_2\text{SO}_4$  and 4  $\text{H}_2\text{O}+\text{SO}_3$  force fields. The trajectories establish that total energy is conserved on the nanosecond time scale during which several reactive processes take place. The present implementation allows modeling of chemical reactions including gas-phase reactions of small molecules, homogeneous and heterogeneous catalytic reactions, and enzyme catalysis. Hence, the functionality is comparable to that of QM/MD albeit at approximately the speed of a conventional force field simulation. The added effort is evidently the parametrization of the force fields involved.



## 3.2 Competitive Reaction Pathways of H<sub>2</sub>SO<sub>4</sub>

### 3.2.1 Introduction

The chemistry of sulfur-containing compounds is of great importance in atmospheric sciences. Sulfur is mainly emitted in the form of dimethyl sulfide (DMS)<sup>155–157</sup> and sulfur dioxide (SO<sub>2</sub>).<sup>158,159</sup> One of the main intermediates formed from the oxidation of DMS is SO<sub>2</sub><sup>160,161</sup> which is subsequently oxidized to sulfur trioxide (SO<sub>3</sub>) by reaction with the hydroxyl radical (HO·) and with O<sub>2</sub>.<sup>162–164</sup> Subsequently, SO<sub>3</sub> is hydrated to sulfuric acid (H<sub>2</sub>SO<sub>4</sub>) and sulfates (SO<sub>4</sub><sup>2-</sup>). These reactions have been extensively studied, mainly because of their important role in tropospheric aerosol layer formation, which is related to global atmospheric chemistry<sup>165</sup> including cloud condensation,<sup>166</sup> altering the chemical composition and influencing the course of aqueous reactions in clouds. On the other hand, the anomalous enhancement of SO<sub>2</sub> in the polar stratospheric aerosol layer during springtime has been postulated to originate from photodissociation of H<sub>2</sub>SO<sub>4</sub> to form water and sulfur trioxide which rapidly decays to SO<sub>2</sub>.<sup>167</sup> This photodissociation process was previously assumed to take place *via* absorption of ultraviolet (UV) radiation, to produce a dissociative excited state. However, the electronic absorption spectrum of H<sub>2</sub>SO<sub>4</sub> up to 140 nm could not be found experimentally<sup>168,169</sup> or in computations.<sup>170–172</sup> The lowest energy electronic transition in H<sub>2</sub>SO<sub>4</sub>, using coupled cluster methods and correlation consistent basis sets, was calculated to be at 139 nm.<sup>173</sup> Because UV photons are absorbed at higher altitudes only a small fraction of photons with wavelengths shorter than 179 nm will penetrate into the lower mesosphere and upper stratosphere,<sup>174</sup> making UV photodissociation of H<sub>2</sub>SO<sub>4</sub> unlikely to occur at lower altitudes.<sup>168</sup>

Following this, it was proposed that some OH-containing species in the atmosphere absorb visible radiation which can potentially induce transitions to vibrational overtones with several quanta in the OH stretch mode.<sup>143,175</sup> Related calculations suggest that dissociation of H<sub>2</sub>SO<sub>4</sub> is possible mainly from highly excited OH-stretching vibrational overtones, which could provide the necessary energy for photodissociation.<sup>144</sup> Hence, the vibrational spectroscopy

of  $\text{H}_2\text{SO}_4$  in the near-infrared and visible regions has been intensely studied, both experimentally<sup>168,176–179</sup> and computationally.<sup>171,172,179</sup> Even though experimentally the vibrationally induced decomposition reaction  $\text{H}_2\text{SO}_4 \rightarrow \text{SO}_3 + \text{H}_2\text{O}$  has not been observed, it is possible that exciting higher OH-stretching vibrational overtones in  $\text{H}_2\text{SO}_4$  by visible light provides sufficient energy for photodissociation. This process is likely because the energy barrier for the dissociation of  $\text{H}_2\text{SO}_4$  is in the range of 32–40 kcal/mol, according to calculations.<sup>66,142,151</sup> On the other hand, vibrational mode-specific reactivity has already been observed, for example for the reaction of methane on a nickel surface.<sup>180</sup>

Previously, the dynamics of the photodissociation reaction following excitation with  $v_9 = 4$  to 6 quanta in the OH-stretching vibration has been investigated by classical trajectory simulations together with semiempirical PM3 electronic structure calculations,<sup>139</sup> or parametrized force fields (FFs).<sup>66</sup> It was found that the photodissociation reaction occurs typically on a picosecond timescale - depending on the level of excitation - which is significantly faster than the collision rate in the atmosphere.<sup>181</sup> The molecular dynamics (MD) simulations also showed that photoexcitation can lead to a) intramolecular H-transfer or b) nonstatistical, impulsive  $\text{H}_2\text{O}$ -elimination and statistical, intramolecular vibrational-energy redistribution (IVR).<sup>66,139</sup>

Because vibrationally induced photodissociation has as yet not been observed experimentally for  $\text{H}_2\text{SO}_4$ , suitable observables are required to detect such a decay mechanism. One possibility is to determine final state distributions of the individual energy components for the fragments  $\text{H}_2\text{O}$  and  $\text{SO}_3$  after photodissociation. Such information is now available from suitable experiments.<sup>182–184</sup> Also a competition between the two reaction channels, which have similar activation energies, may affect the decay rate. In order to understand in detail the dynamics of vibrationally induced decomposition of  $\text{H}_2\text{SO}_4$ , reactive molecular dynamics simulations are performed using multisurface adiabatic reaction dynamics (MS-ARMD).<sup>185</sup> This method, based on adiabatic reactive MD,<sup>186,187</sup> is capable of treating multiple reaction pathways and conserve total energy along a reactive trajectory. As it is based on force fields to describe the total energy of the system, a statistically significant number of trajectories can be run and analyzed. In the present work we investigate the dynamics on a new fully dimensional reactive potential energy surface capable of following intramolecular H-transfer and water elimination in  $\text{H}_2\text{SO}_4$ . This provides the

information to analyze the reaction products' energetics including translational, rotational and vibrational contributions which will eventually be relevant for comparison with experiments.

## 3.3 Computational Methods

### 3.3.1 Force Field Parametrization

The present work employs a global, parametrized and reactive potential energy surface (PES) to follow intramolecular H-transfer and water elimination ( $\text{H}_2\text{SO}_4 \rightarrow \text{H}_2\text{O} + \text{SO}_3$ ) in  $\text{H}_2\text{SO}_4$  upon vibrational excitation of the OH-stretching mode. For this, MS-ARMD was employed.<sup>185</sup> The global PES is based on reference electronic structure calculations at the MP2/6-311G++(2d,2p) level. All *ab initio* calculations were performed with Gaussian09.<sup>18</sup> The energy at the equilibrium geometry of  $\text{H}_2\text{SO}_4$  was chosen as the global zero and serves as a common origin for all *ab initio* energies. Using previously parametrized FFs for  $\text{H}_2\text{O}$ ,  $\text{SO}_3$ , the  $\text{H}_2\text{O} \cdots \text{SO}_3$  van der Waals complex, and  $\text{H}_2\text{SO}_4$ , representative geometries were collected from MD simulations which cover the regions of the multidimensional PES relevant to the equilibrium dynamics.<sup>66</sup> Starting from the optimized geometry, each system was heated to 300 K, and subsequently equilibrated for 40 ps, followed by 10 ns of free dynamics at constant total energy. Along the three trajectories, approximately 1000 geometries were selected for  $\text{H}_2\text{O}$ ,  $\text{SO}_3$  and  $\text{H}_2\text{SO}_4$ , respectively. A similar procedure was followed for the  $\text{H}_2\text{O} \cdots \text{SO}_3$  van der Waals complex (see below). Additionally, a rigid 2D scan along the dihedral angles H6-O7-S1-O4 and H5-O4-S1-O6 (Figure 3.10) in  $\text{H}_2\text{SO}_4$  was performed resulting in an extra 1296 reference data points (648 independent calculations) which were added to the previous 3000 structures.

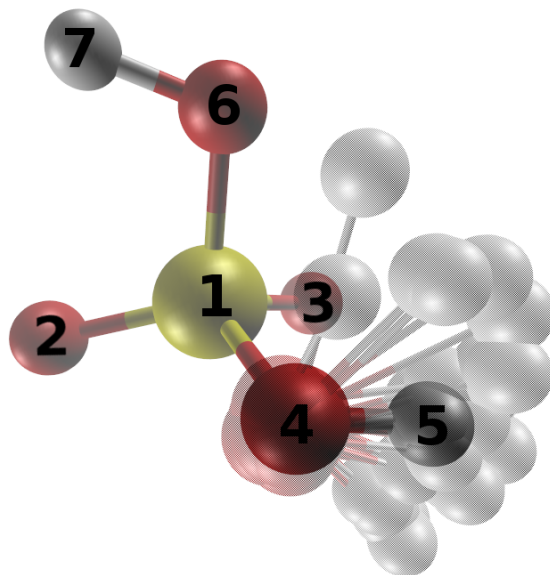
The functional form of the force field which describes the  $\text{H}_2\text{SO}_4$  and  $\text{H}_2\text{O} + \text{SO}_3$

states, including the  $\text{H}_2\text{O} \cdots \text{SO}_3$  van der Waals complex, is

$$\begin{aligned}
 V(\mathbf{x}) = & \sum_{\text{bonds}} D_e(1 - e^{-\beta(r-r_0)})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 \\
 + & \sum_{\text{Urey-Bradley}} K_{\text{UB}}(S - S_0)^2 + \sum_{\text{dihedral}} H_\phi(1 + \cos(n\phi - \delta)) \\
 & + \sum_{\text{impropers}} K_\omega(\omega - \omega_0)^2 \quad (3.14) \\
 + & \sum_{ij} \left\{ \frac{n\varepsilon_{ij}}{m-n} \left[ \left( \frac{r_{\min,ij}}{r_{ij}} \right)^m - \frac{m}{n} \left( \frac{r_{\min,ij}}{r_{ij}} \right)^n \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\}
 \end{aligned}$$

where  $r_0$  is the equilibrium bond distance,  $D_e$  is the dissociation energy, and  $\beta$  controls the steepness and width of the Morse potential. In all FFs the bonds were treated with Morse potentials. For every angle in the three molecules  $\text{H}_2\text{SO}_4$ ,  $\text{H}_2\text{O}$  and  $\text{SO}_3$  (i.e., H5-O4-S1, O2=S1=O3, etc., see Figure 3.10) two potentials were used; one of them is the usual harmonic potential involving three atoms, whereas the second is a Urey-Bradley term which is a quadratic interaction potential of the distance between atoms separated by two bonds (1,3 interaction). All dihedral angles in  $\text{H}_2\text{SO}_4$  are described by proper dihedral terms. For  $\text{SO}_3$  additional improper dihedral potentials were employed symmetrically to maintain its planarity. Generalized-exponent Lennard-Jones potentials (or Mie potentials)<sup>145</sup> were used for the  $\text{H}_2\text{SO}_4$  molecule between 1-4 (e.g. H5 and O2 see Figure 3.10) and 1-5 (between H5 and H7 see Figure 3.10) neighboring atoms. Point charges were obtained by fitting the molecular electrostatic potential (ESP) using the CHELPG scheme within Gaussian09.<sup>149</sup> Initial values for all other parameters were taken or calculated (i.e. for generalized Lennard-Jones) from standard CHARMM parameters<sup>188</sup> and a standalone Fortran code based on the downhill simplex algorithm<sup>54</sup> was used for optimizing the parameters to match the *ab initio* reference energies.

With the refined FF parameters for  $\text{H}_2\text{O}$  and  $\text{SO}_3$ , the generalized Lennard-Jones (GLJ) parameters and atomic charges (last term in Eq. 3.15) were fitted in order to describe the intermolecular interactions for the  $\text{H}_2\text{O} \cdots \text{SO}_3$  van der Waals complex. Initial guesses for the GLJ parameters are obtained from the Lennard-Jones parameters between pairs of atoms  $i$  and  $j$  using the Lorentz-Berthelot combination rules in which  $\varepsilon_{ij}$  is the geometric mean of  $\varepsilon_i$  and  $\varepsilon_j$  and  $r_{\min,ij}$  is the arithmetic mean of  $r_{\min,i}$  and  $r_{\min,j}$ . For the GLJ potential,  $\varepsilon_{ij}$  ( $> 0$ ) and  $r_{\min,ij}$  are also the well-depth and the corresponding separation at this energy minimum, respectively. As a reference for the fitting,



**Figure 3.10:** Sulfuric acid with atom numbering used throughout the text. The librational motion of O4 and H5, related to the dihedral angle H5-O4-S1-O6, is indicated by light spheres.

electronic structure calculations were carried out on approximately 1000 geometries collected from an MD simulation at 300 K for the van der Waals complex (see above). Two different parameter sets describe the reactant and product.

The FF in Eq. 3.15 is also suitable to follow H-transfer in sulfuric acid. Because there are 12 different ways how the two H-atoms can be arranged, the final PES needs to be permutationally invariant. In total, there exist 4 different final states which give rise to 16 FFs that are required to describe all possible reaction pathways of interest in the present work. The 16 states are connected through 48 transition states.<sup>185</sup>

In a next step, the FFs for the states involved in the reactions ( $\text{H}_2\text{SO}_4$ , and  $\text{H}_2\text{O}+\text{SO}_3$ ) are combined to a global PES for reactive molecular dynamics simulations. For this, geometrically dependent weights  $w_i(\mathbf{x})$  are used to generate the global PES from the individual FFs  $V_i(\mathbf{x})$ .<sup>185</sup> Increased flexibility for the MS-ARMD PES is required in the barrier regions where the FFs of

the states - i.e.  $\text{H}_2\text{SO}_4$  and  $\text{H}_2\text{O}+\text{SO}_3$  - cross. This is accomplished by using products of Gaussian and polynomial functions (GAPOs) of the energy difference ( $\Delta V_{ij}$ ) between the PESs involved<sup>185</sup>

$$\Delta V_{\text{GAPO},k}^{ij}(\mathbf{x}) = \exp\left(-\frac{(\Delta V_{ij}(\mathbf{x}) - V_{ij,k}^0)^2}{2\sigma_{ij,k}^2}\right) \times \sum_{l=0}^{m_{ij,k}} a_{ij,kl} (\Delta V_{ij}(\mathbf{x}) - V_{ij,k}^0)^l. \quad (3.15)$$

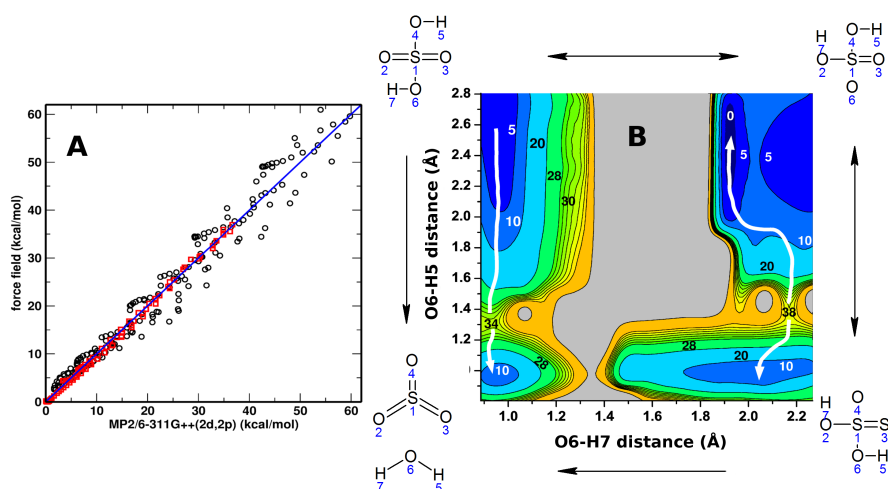
where  $V_{ij,k}^0$  and  $\sigma_{ij,k}$  denote the center and the standard deviation of the Gaussian functions, respectively,  $\Delta V_{ij}(\mathbf{x}) = V_i(\mathbf{x}) - V_j(\mathbf{x})$ , and  $a_{ij,kl}$  are fitting parameters. The global PES is then a weighted sum of PESs (first term in Eq. 3.16) combined with the weighted GAPOs (second term in Eq. 3.16) which allows to better capture the barrier region

$$V_{\text{MS-ARMD}}(\mathbf{x}) = \sum_{i=1}^n w_i(\mathbf{x}) V_i(\mathbf{x}) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n [w_i(\mathbf{x}) + w_j(\mathbf{x})] \sum_{k=1}^{n_{ij}} \Delta V_{\text{GAPO},k}^{ij}(\mathbf{x}) \quad (3.16)$$

and the sum runs up to  $n = 16$ . Parameters of the GAPO functions were optimized to reproduce the barrier height for both processes of interest, water elimination and intramolecular H-transfer. The reference points in the fit were augmented by points along both minimum energy paths (MEPs) determined from quadratic synchronous transit (QST2) calculations<sup>150</sup> followed by an intrinsic reaction coordinate (IRC) calculation at the MP2/6-311G++(2d,2p) level of theory. The MS-ARMD switching parameter  $\Delta V$  and the overall shifts of the FFs (see ref.<sup>185</sup> for definitions) were optimized with the downhill simplex algorithm using  $k = 3$  GAPOs with  $m_{ij,k} = 1$  (first order polynomial) for  $\text{H}_2\text{O}$  elimination, and  $k = 2$  GAPOs with  $m_{ij,k} = 2$  for the H-transfer reaction.<sup>185</sup>

The energy profile along the reference minimum energy path can be reproduced very closely (RMSD = 0.55 kcal/mol), as is shown in Figure 3.11A (red squares) whereas the remaining electronic structure data are somewhat further away from a 1:1 correlation, but still close to it. Figure 3.11B illustrates the two minimum energy pathways for water elimination and intramolecular H-transfer.

The reference *ab initio* calculations yield a barrier height of 32.0 kcal/mol for water elimination, and 36.6 kcal/mol for intramolecular H-transfer, which is also closely captured by the fitted PES. The energetics establishes that for a meaningful treatment of the reactive dynamics both pathways need to be included. The present barrier for water elimination compares with previous estimates ranging from 32 to 40 kcal/mol depending on the level of theory used.<sup>66, 142, 151</sup>



**Figure 3.11:** Panel A shows the quality of the fitted PES relative to the *ab initio* reference energies. Points along the MEPs (red squares) and remaining reference points (black circles) are distinguished. Panel B shows the fully relaxed MS-ARMD PES for both, intramolecular H-transfer and water elimination, in the space defined by distances H5-O6 and H7-O6. The white lines indicate approximate MEPs for the two possible reactions, water elimination (left-hand side) and H-transfer (right-hand side). Selected isocontours (in kcal/mol) are labelled.

### 3.3.2 Molecular Dynamics Simulations

All MD simulations were carried out with CHARMM<sup>189</sup> with provisions for bond-breaking and bond-formation through MS-ARMD.<sup>185</sup> Starting from a geometry optimized structure of H<sub>2</sub>SO<sub>4</sub>, the system was heated to 300 K. The equations of motion were solved with the leapfrog Verlet algorithm with

a time step of  $\Delta t = 0.1$  fs during 40 ps and equilibrated for 40 ps, followed by 50 ps of free dynamics simulations (i.e. constant energy). The small time step is required to appropriately follow the rapid dynamics of the H-atoms.

Vibrational excitation of the OH stretching mode was modelled by scaling the instantaneous velocities along the  $v_9$  normal mode which is predominantly an OH-stretching mode and highly localized. The OH stretch potential can be realistically described by a Morse potential, for which the exact energy levels are known. The excitation energies include  $13490\text{ cm}^{-1}$  (38.6 kcal/mol,  $v_9 = 4$ ),  $16494\text{ cm}^{-1}$  (47.2 kcal/mol,  $v_9 = 5$ ), and  $19328\text{ cm}^{-1}$  (55.3 kcal/mol,  $v_9 = 6$ ). Such an approach has also been successfully used for studying proton transfer in H-bonded complexes.<sup>190</sup> Alternatively, positions and velocities can be scaled whereby both the kinetic and potential energies are modified.<sup>152</sup> One possible drawback of modifying both positions and momenta is the fact that, after excitation, a short equilibration period (a few picoseconds) is required to avoid artifacts due to close proximity of atoms.

### 3.3.3 Final State Analysis

For the final state analysis, the total energy of the system was decomposed into translational, rotational and vibrational components of each fragment. The translational energy of a reaction product is  $E_{\text{trans}} = \frac{1}{2} M v_{\text{CM}}^2$  where  $M$  and  $v_{\text{CM}}$  denote the mass and the velocity of the center of mass for the given fragment ( $\text{H}_2\text{O}$  and  $\text{SO}_3$ ). Following classical mechanics, the angular momentum vector of a reaction product (with  $N$  atoms) is  $\mathbf{L} = \sum_{i=1}^N m_i (\mathbf{r}_i - \mathbf{r}_{\text{CM}}) \times (\mathbf{v}_i - \mathbf{v}_{\text{CM}})$ . Here,  $\mathbf{r}_{\text{CM}}$  and  $\mathbf{v}_{\text{CM}}$  denote the position and the velocity vectors of the center of mass for the given fragment. This yields the rotational energy  $E_{\text{rot}} = \frac{1}{2} \mathbf{L}^T \Theta^{-1} \mathbf{L}$  where  $\Theta$  is the moment of inertia tensor of the fragment  $\Theta = \sum_{i=1}^N m_i (\mathbf{E} r_i^2 - \mathbf{r}_i \mathbf{r}_i^T)$  and  $\mathbf{E}$  is the unit tensor. From  $E_{\text{trans}}$  and  $E_{\text{rot}}$  the vibrational energy of each product is determined through  $E_{\text{vib}} = E_{\text{kin}} - E_{\text{trans}} - E_{\text{rot}} + V$  where  $V$  denotes the intramolecular potential energy of the given reaction product. Once the two fragments have moved sufficiently far from the interaction region (here taken as 20 Å between the  $\text{SO}_3$ -sulfur and the  $\text{H}_2\text{O}$ -oxygen atoms), the translational and



rovibrational energy of each fragment will be constant. As the vibrational and rotational modes in each fragment keep on exchanging energy if the angular momentum is nonzero, these quantities are averaged over periods much longer (i.e. 5 ps) than the characteristic time for vibrations within the molecules.

The angular momentum quantum number  $j$  can be calculated from the classical-quantum correspondence relationship  $L^2 = j(j+1)\hbar^2$  in which  $\hbar$  is Planck's constant. The orbital angular momentum ( $\mathbf{L}_{\text{orbital}}$ ) can be determined from the individual angular momenta of the fragments and from the conservation of total angular momentum. Considering the photodissociation reaction as an inverse scattering process in the framework of classical mechanics, it is useful to introduce the perpendicular distance ( $b_{\text{scatter}}$ ) between the asymptotic paths of the fragments in analogy with the impact parameter of bimolecular reactions:  $b_{\text{scatter}} = L_{\text{orbital}}/\sqrt{2\mu E_{\text{trans,tot}}}$ . Here,  $\mu$  and  $E_{\text{trans,tot}}$  are the reduced mass and the sum of translational energies of the fragments, respectively, and  $L_{\text{orbital}} = |\mathbf{L}_{\text{orbital}}|$ . This scattering parameter provides information about the relative displacements of the fragmentation partners at the moment of breakup.

## 3.4 Results

### 3.4.1 Validation of the Force Field

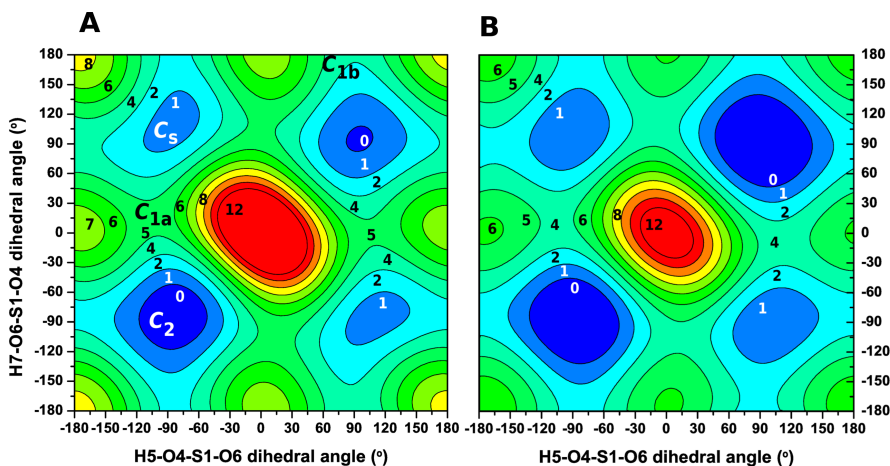
The energy-minimized structures at the MP2/6-311G++(2d,2p) level and using the force field has  $C_2$ -symmetry (see Figure 3.10).<sup>191</sup> A comparison of equilibrium bond lengths, angles and dihedral angles between the two treatments is provided in Table S-I (supporting information). The positional root-mean-square deviations (RMSDs) are 0.04 Å for H<sub>2</sub>SO<sub>4</sub> and 0.001, 0.001 and 0.13 Å for SO<sub>3</sub>, H<sub>2</sub>O and the vdW complex, respectively. For bond lengths and valence angles the average difference between the MP2 reference and the FF structures is less than 0.03 Å and around 3°, respectively. This, together with the results from Figure 3.11A, suggests that the force field is parametrized in a meaningful way.

Previous computational work on rotamers for sulfuric acid revealed the existence of two minimum energy geometries with point group symmetries  $C_2$  and  $C_s$ , connected by two saddle points with  $C_1$  symmetry, 4.9 kcal/mol and 5.4 kcal/mol higher in energy than the  $C_2$  structure. The lower of the two saddle points is called  $C_{1a}$  whereas the higher one is  $C_{1b}$ . The energies of the two saddle points at the current level of theory are around 1 kcal/mol higher than those computed by Lohr and Havey, who found 3.34 and 4.16 kcal/mol (at the HF/STO-3G level) and 1.45 and 4.06 kcal/mol (at the B3LYP/6-311G(2d,2p) level) in the two studies, respectively.<sup>176,192</sup>

A comparison between the reference and fitted FF energies for the torsional degrees of freedom is presented in Figure 3.12. The two saddle points on the FF PES are 3.7 kcal/mol and 4.3 kcal/mol above the global  $C_2$  minimum which compares with 4.9 kcal/mol and 5.4 kcal/mol in the MP2 reference calculations. In addition, a normal mode analysis was performed in order to compare the experimentally observed vibrational frequencies with those from the MP2 and FF calculations. All MP2 frequencies were scaled with a factor of 0.97 (MP2/6-31+G(d,p)<sup>193</sup> and MP2-fc/6-311G(d,p)<sup>194</sup>). As the FF is based on MP2 calculations, the vibrational frequencies obtained from the normal mode analysis using the FF were also scaled with the same scaling factor. Comparisons of individual frequencies are reported in Table S-II (supporting information). The average differences between experimental,<sup>195-197</sup> MP2 and FF harmonic frequencies for  $H_2SO_4$ ,  $SO_3$ ,  $H_2O$  and  $H_2O \cdots SO_3$  are [72, 20, 22, 41]  $cm^{-1}$  (MP2 vs. FF), [53, 52, 87, 65]  $cm^{-1}$  (expt. vs. MP2), and [88, 68, 65, 69]  $cm^{-1}$  (expt. vs. FF).

### 3.4.2 Reaction Dynamics

With a suitably formulated and accurately parametrized FF which can reproduce optimized geometries, torsional and reaction barrier heights, and vibrational frequencies (see Tables S-I and S-II (supporting information) and Figures 3.11 and 3.12) close to those computed from reference electronic structure calculations at the MP2 level, reactive molecular dynamics simulations were carried out. In order to arrive at a statistically significant number of reactive events, 7000 independent trajectories were run using

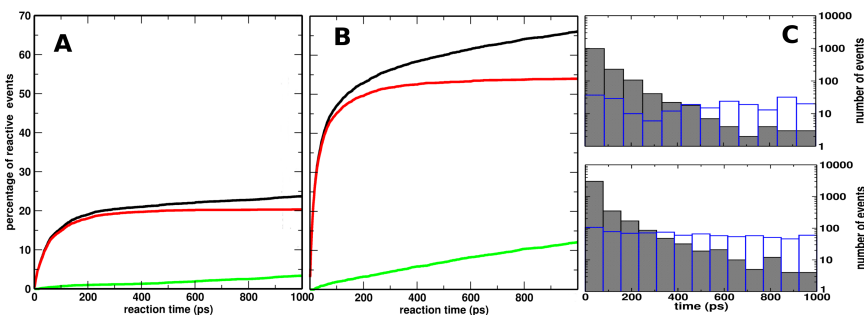


**Figure 3.12:** Calculated PES along dihedral angles H5-O4-S1-O6 ( $\phi$ ) and H7-O6-S1-O4 ( $\phi'$ ) at (A) the MP2/6-311G++(2d,2p) level (rigid scan) and (b) with the FF (rigid scan). The two minima are  $C_2$  and  $C_s$  which are connected through transition states  $C_{1a}$  (4.9 kcal/mol) and  $C_{1b}$  (5.4 kcal/mol).

MS-ARMD for  $v_9 = 4, 5,$  and  $6$  quanta. The global surface can describe the two energetically accessible channels: water elimination and intramolecular H-transfer. Because the barrier heights for both processes are accurately described by the MS-ARMD PES the competition between the two reactive channels can also be studied.

From 7000 independent trajectories with a maximum simulation time of 1 ns, 4048, 5418 and 5599 trajectories for excitation of  $v_9 = 4, 5$  and  $6$ , respectively, showed either water elimination or intramolecular H-transfer. Of those, 57% ( $v_9 = 4$ ), 54% ( $v_9 = 5$ ) and 14% ( $v_9 = 6$ ), experienced only intramolecular H-transfer without subsequent water elimination. For water elimination, two different mechanisms were found: i) only water elimination occurs and ii) intramolecular H-transfer precedes water elimination. For the first mechanism, 35, 1495 and 3779 trajectories were found whereas for the second pathway 0, 236 and 853 events were observed for  $v_9 = 4, 5$  and  $6$ , respectively. These results suggest that excitation of  $v_9 = 4$  is probably not sufficient to obtain significant amounts of  $\text{SO}_3$  on the nanosecond time scale whereas excitation of  $v_9 \geq 5$  results in rapid decay.

Panels A and B of Figure 3.13 reports the percentage of reactive events as a function of reaction time for water elimination with  $v_9 = 5$  and 6, respectively. For excitation with  $v_9 = 5$  a maximum lifetime for direct water elimination of around 300 ps is found which reduces to approximately 100 ps for excitation with  $v_9 = 6$ . Contrary to that, for the reaction in which the system experiences one or more intramolecular H-transfers prior to water elimination, the percentage of decays increases linearly for both, excitation of  $v_9 = 5$  and 6 (green curves in Figure 3.13).

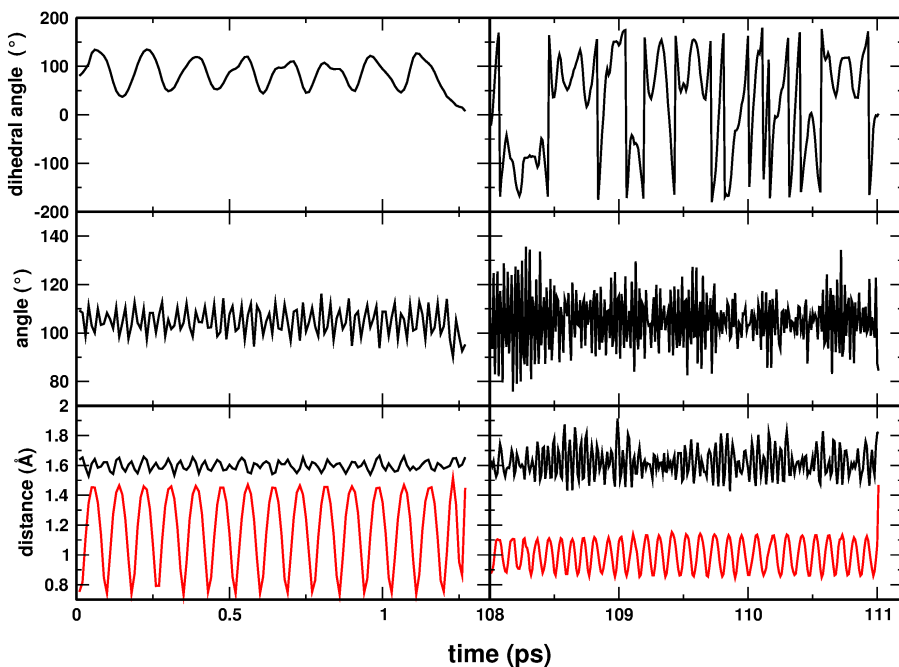


**Figure 3.13:** Percentage of water elimination after excitation of  $v_9 = 5$  (A) and  $v_9 = 6$  (B). Black, red and green curves correspond to overall, direct water elimination, and water elimination preceded by intramolecular H-transfer, respectively. Panel C reports histograms for the distribution of reaction times for direct water elimination (black) and for water elimination preceded by intramolecular H-transfer (blue) for  $v_9 = 5$  (top) and  $v_9 = 6$  (bottom) on a semi-logarithmic scale.

Figure 3.13 C shows the lifetime distributions for the two possible mechanisms on a semi-logarithmic scale. The rate for direct water elimination (grey) for  $v_9 = 5$  (top) is slower than that for  $v_9 = 6$  (bottom) while for the process in which water elimination is preceded by intramolecular H-transfer (blue), the rates are constant in time. However, the absolute number of events differs by a factor of  $\approx 5$  between the two excitation levels. For excitation of  $v_9 = 6$  the trajectories in which  $\text{H}_2\text{SO}_4$  did not decay within 1 ns the dynamics was continued out to 40 ns. By that time 6844 (98 %) out of the 7000 trajectories decayed to  $\text{H}_2\text{O}$  and  $\text{SO}_3$  of which 3995 (57 %) went through direct water elimination and the remaining 2849 (40 %) was preceded by intramolecular H-transfer. 150 events (2 %) experienced only intramolecular H-transfer and 6 events (0.08 %) did not exhibit any reaction at all.

For excitation of  $v_9 = 6$ , 90 % of events are found to occur in less than 10 ns

(Figure S-I black curve). For direct water elimination, a rapid increment is observed until 400 ps after which very few additional new events are observed. Conversely, water elimination preceded by intramolecular H-transfer continues to increase until  $\approx 15$  ns after vibrational excitation. The total conversion to products is  $\approx 98$  % after 40 ns.



**Figure 3.14:** Time series of selected internal coordinates for  $\text{H}_2\text{SO}_4$ , including the H5-O4-S1-O6 dihedral angle (top), the H5-O4-S1 angle (middle), and the H5-O4 (red curve) and S1-O6 (black curve) bonds (bottom). Left column reports the data for a rapid water elimination reaction (breakup after 1.26 ps) whereas the right column corresponds to the last 3 ps of a slow reaction (breakup after  $\approx 111$  ps), for two representative trajectories with excitation of  $\nu_9 = 6$ . Note the characteristic torsional dynamics in the left panel.

Figure 3.14 reports time series of selected internal coordinates along two trajectories with very different reaction times for water elimination. The left panels correspond to a rapid reaction occurring in less than 2 ps whereas the right hand panels are for a situation in which the vibrational excitation redistributes on the 100 ps time scale before water elimination. For the rapid process, the H5-O4-S1-O6 dihedral angle at time 0 starts close to the  $C_2$  minimum ( $90^\circ$ ). Initially, the motion of this dihedral angle is similar to a

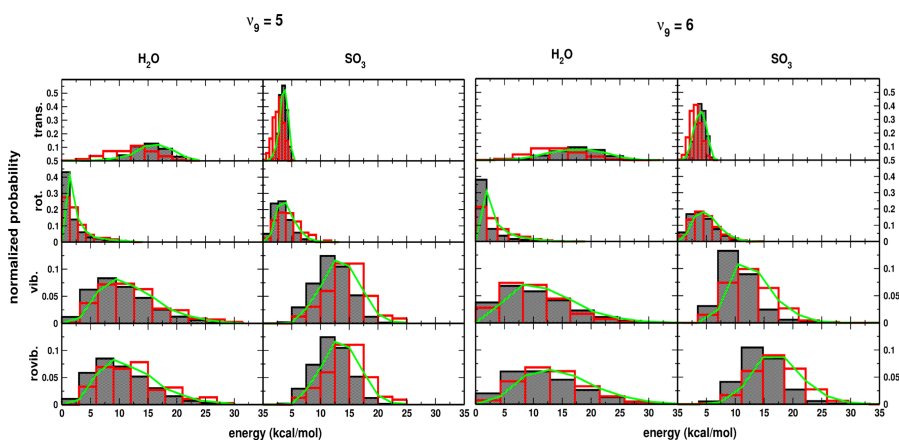
librational mode, fluctuating around  $80^\circ$ . When the reaction occurs at around 1.26 ps, the H5-O4-S1-O6 dihedral angle reaches  $0^\circ$ , which is the lowest saddle point in the transition from the  $C_2$  to the  $C_s$  minimum (see Figure 3.10 and 3.12). On the other hand, the H5-O4 bond length fluctuates around an average value of about 1.1 Å and reaches a maximum of 1.45 Å when the reaction takes place at 1.26 ps. At this point, the distance between atoms H5 and O6 is  $\approx 1.4$  Å which is close to the transition state geometry, see Figure 3.10.<sup>66,142,151</sup> Figure 3.14 suggests that for trajectories with such a rapid reaction, these two coordinates (dihedral angle and OH bond length) play an important role in the reaction path. However, a more detailed reaction coordinate analysis is required to better understand this aspect.<sup>198</sup>

For the trajectory which shows water elimination after  $\approx 111$  ps, see Figure 3.14B, several intramolecular H-transfers occur before  $t = 108$  ps. In addition, large torsional and angular excursions are observed which leads to visiting both possible minimum energy structures, see Figure 3.12. Such motions occur on the picosecond time scale, as can be seen in Figure 3.14B. Concomitantly, the H5-O4-S1 angle and S1-O6 bond also pick up a considerable amount of energy which assists the eventual dissociation along coordinates H5-O4 and S1-O6.

### 3.4.3 Final State Analysis

Final state analysis is a meaningful way to characterize the outcome of a gas phase reaction. Because experimentally the direct investigation of the water elimination reaction is challenging, a computational exploration of the likely outcomes of such an experiment is potentially helpful. Contrary to previous efforts, the present simulations are capable of making concrete and quantitative statements about this aspect for several reasons. Firstly, MS-ARMD conserves total energy which is essential if a meaningful final state analysis is to be performed. Secondly, the current reactive PES is much improved over the previous force field in that it realistically captures the important torsional degrees of freedom. Thirdly, the present treatment allows both, intramolecular H-transfer and water elimination which is essential because the barrier height for both processes is similar.

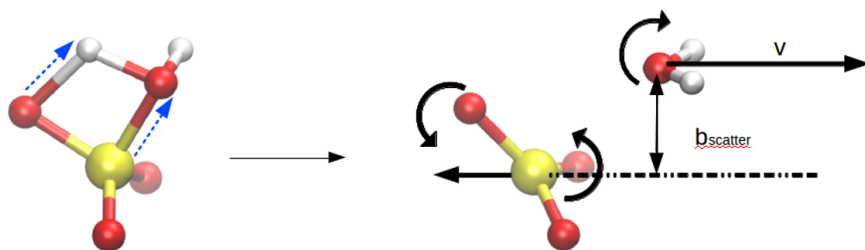
In order to investigate the energy distribution after vibrational excitation and subsequent water elimination, the translational, rotational and vibrational energies in the reaction products were determined following established methodologies.<sup>199</sup> The analysis was carried out for all trajectories that experienced water elimination within 1 ns. The final state analysis starts when the sulfur-oxygen(water) distance exceeds 20 Å. Trajectories were then continued for an additional 5 ps during which 500 snapshots were collected and analyzed. Translational, rotational, vibrational, rovibrational energies and the rotational quantum number were computed for every snapshot and rotational and vibrational energies were averaged over all 500 snapshots from which probability distributions of the averaged values can be determined.



**Figure 3.15:** Energy distribution for the H<sub>2</sub>O and SO<sub>3</sub> fragments. Contribution of translational, rotational, vibrational and rovibrational energies are shown for the two processes; direct water elimination (black histograms) and water elimination preceded by intramolecular H-transfer (red histograms) for  $v_9 = 5$  (left) and 6 (right). Overall distributions are represented by the green curve.

Figures 3.15 shows distributions for all energy components of the products for excitation of  $v_9 = 5$  and 6, respectively. Three distributions, each normalized individually, are reported. One of them (solid line) is the overall distribution of the particular energy component which would be the expected experimental signal. The two histograms are distributions for direct water elimination (without prior intramolecular H-transfer) and energy distributions for events in which intramolecular H-transfer precedes water elimination, respectively. It is found that the two partial energy distributions peak at different maximal energies in general. This is most prominent for the translational and vibrational

energies in  $\text{SO}_3$ . For  $v_9 = 6$ , a maximum peak at about 20 kcal/mol is observed for the  $\text{H}_2\text{O}$ -translational energy for the direct reaction, whereas for  $\text{SO}_3$  this quantity peaks at 5 kcal/mol (Figure 3.15 grey trace). Contrary to that, reactions with prior H-transfer lead to a maximum peak in the translational energy distribution at 12 kcal/mol for  $\text{H}_2\text{O}$  and  $\approx 3$  kcal/mol in the case of  $\text{SO}_3$ . The translational energies of the products are correlated due to conservation of linear momentum, which was initially zero by preparation. On the other hand the rotational energies are similarly distributed for trajectories showing water elimination with or without previous H-transfer. The peaks for the two differ only by about 1 kcal/mol. The angular momenta of the products, and therefore their rotational energies, are not correlated, contrary to the linear momentum. This is due to the presence of orbital angular momentum.

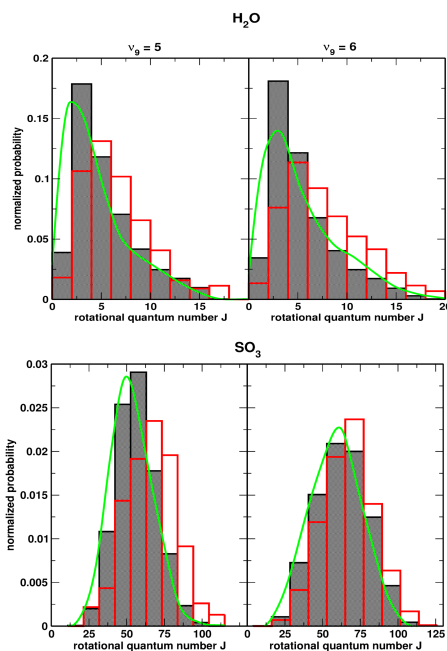


**Figure 3.16:** Generation of the angular momenta in the final state. The blue arrows on  $\text{H}_2\text{SO}_4$  illustrate the H-transfer step which leads to water elimination. Due to the off-center breakup and the conservation of total angular momentum the two fragments,  $\text{SO}_3$  and  $\text{H}_2\text{O}$ , separate in a counterrotating manner with high fragment angular velocities, indicated by the curved arrows.

The angular momentum distributions can also be discussed within the framework of a scattering process. The geometrical situation is schematically summarized in Figure 3.16 where perpendicular distances of the asymptotic paths of the fragments are characterized by a Gaussian distribution centered around  $b_{\text{scatter}} \sim 0.7\text{\AA}$ . While the distributions for the direct and indirect mechanisms were distinctly different, they proved to be largely independent of the level of excitation:  $b_{\text{scatter}}^{\text{direct}} = \mu \pm \sigma = 0.72 \pm 0.18\text{\AA}$  ( $v_9 = 4$ ),  $0.67 \pm 0.19\text{\AA}$  ( $v_9 = 5$ ) and  $0.66 \pm 0.19\text{\AA}$  ( $v_9 = 6$ ) and  $b_{\text{scatter}}^{\text{indirect}} = \mu \pm \sigma = 0.89 \pm 0.31\text{\AA}$  ( $v_9 = 5$ ) and  $0.85 \pm 0.27\text{\AA}$  ( $v_9 = 6$ ). I.e. the impact parameter for direct water elimination is typically 25 % smaller than that for the indirect process. The magnitudes of  $b_{\text{scatter}}$  and the translational energies of the fragments explain the large orbital angular momentum ( $L_{\text{orbital}} \approx 60$ ). The high orbital



angular momentum is compensated for by the rotational excitation of  $\text{SO}_3$  (see Figure 3.17 panel B) to conserve the total angular momentum. This suggests that when the system slides down on the product side of the barrier the O-atom which loses its hydrogen atom will experience a significant force which is transmitted as a torque to the leaving  $\text{SO}_3$  fragment. The relative magnitudes of the rotational excitation of the fragments ( $j_{\text{max}}^{(\text{H}_2\text{O})} \sim 5$  vs.  $j_{\text{max}}^{(\text{SO}_3)} \sim 50$ ) can also be rationalized by noting that the rotational constants of  $\text{H}_2\text{O}$  and  $\text{SO}_3$  differ by almost 2 orders of magnitude (9 to 27  $\text{cm}^{-1}$  compared to 0.17 and 0.34  $\text{cm}^{-1}$ , respectively).<sup>200,201</sup> Noting that the two fragments receive similar amounts of rotational energy, the ratios of the maximum  $j$  follows from  $B^{(\text{H}_2\text{O})}/B^{(\text{SO}_3)} \sim 100 \sim (j_{\text{max}}^{(\text{SO}_3)}/j_{\text{max}}^{(\text{H}_2\text{O})})^2$  from which one obtains  $j_{\text{max}}^{(\text{SO}_3)}/j_{\text{max}}^{(\text{H}_2\text{O})} \sim 10$  which is what is found in Figure 3.17.



**Figure 3.17:** Distribution of the  $\text{H}_2\text{O}$  and  $\text{SO}_3$  rotational quantum numbers  $j$  for both, direct (black) and indirect (preceded intramolecular H-transfer) (red) water elimination for excitation of  $v_9 = 5$  and 6. The overall distribution is the green curve

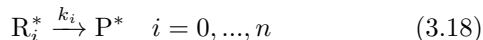
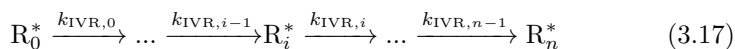
Overall, the analysis suggests that when water elimination is preceded by H-transfer the excess energy goes into vibrational and rotational degrees of freedom. In other words, H-transfer during the nonequilibrium dynamics

promotes energy redistribution within the complex. On the other hand, for direct water elimination the initial energy disposed in the system preferably goes into translational energy.

#### 3.4.4 Kinetic analysis

Photoexcitation generates a non-equilibrium distribution of sulfuric acid molecules from the initial canonical ensemble at 300 K. Due to coupling between vibrational modes the ensemble relaxes to a microcanonical distribution at every energy if ergodicity for the undecayed species is assumed. However, because vibrational excitation to  $v_9 \geq 4$  provides energy above the dissociation limit the available phase space can not be further explored by  $\text{H}_2\text{SO}_4$  which limits the degree to which this system is ergodic. As the time scales for direct water elimination and IVR are similar, statistical rate theories which assume microcanonical or canonical ensembles, can not be employed.

In the following a simple kinetic model is set up to extract first-order rate parameters from the trajectories in order to characterize the long-time behavior of the system. The relaxation of the initial excited population of reactant molecules ( $\text{R}_0^* = \text{H}_2\text{SO}_4^*(v_9 = 4,5,6)$ ) during the free dynamics is modeled as a process of  $n$  consecutive steps, including states  $\text{R}_i^*$  ( $i = 1, \dots, n$ ) which are relaxed to various extents. The final population ( $\text{R}_n^*$ ) is considered to be fully relaxed which is a mixture of microcanonical ensembles with weights reflecting the initial canonical ensemble before excitation. Stepwise IVR to yield  $\text{R}_n^*$  (i.e. complete redistribution of the available energy) is related to the fact that under the present conditions (absence of a collision partner M), the deactivation process  $\text{R}_n^* + \text{M} \rightarrow \text{R}_n + \text{M}$  can not occur. In other words, full relaxation is only possible if the time scales for IVR are shorter than the average collision time under atmospheric conditions. From each  $\text{R}_i^*$  the formation of excited (translationally, rotationally, vibrationally) products ( $\text{P}^* = (\text{H}_2\text{O} + \text{SO}_3)^*$ ) is possible which is assumed to follow a first-order kinetics. Note that no distinction is made between products obtained from different reactant ensembles. The proposed kinetic model is the following:



In the following it is assumed that a single rate coefficient,  $k_{\text{IVR},i}$  and  $k_i$  can characterize the average behavior of the various energy states in a given relaxation or decomposition step, respectively. This is a reasonable assumption as the average energy and its spread within the ensemble ( $\bar{E}_{300\text{K}} = 8.16$  kcal/mol and  $\sigma_E = 2.27$  kcal/mol) are small compared to its average energy after excitation ( $\approx 47$  kcal/mol or more). The decomposition rate from the fully relaxed ensemble can be considered as an approximation to the microcanonical rate coefficient at the particular average energy. The corresponding system of kinetic differential equations for the time-evolution of the probabilities for the various states is:

$$\frac{d[\text{R}_0^*]}{dt} = - (k_{\text{IVR},0} + k_0) [\text{R}_0^*] \quad (3.19)$$

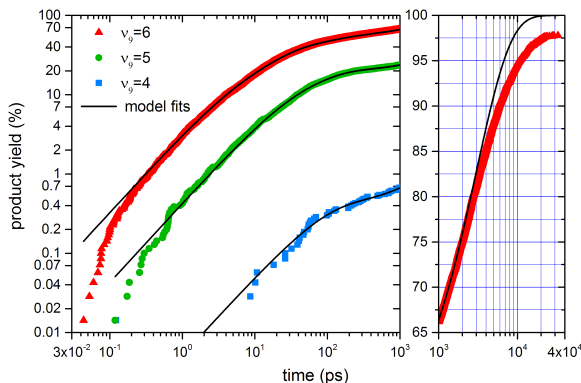
$$\frac{d[\text{R}_i^*]}{dt} = - (k_{\text{IVR},i} + k_i) [\text{R}_i^*] + k_{\text{IVR},i-1} [\text{R}_{i-1}^*] \quad i = 1, \dots, n-1 \quad (3.20)$$

$$\frac{d[\text{R}_n^*]}{dt} = - k_n [\text{R}_n^*] + k_{\text{IVR},n-1} [\text{R}_{n-1}^*] \quad (3.21)$$

$$\frac{d[\text{P}^*]}{dt} = \sum_{i=0}^n k_i [\text{R}_i^*] \quad (3.22)$$

Initially, the entire population (100%) is in  $\text{R}_0^*$  and the probability of all other  $\text{R}_i^*$  is zero:  $[\text{R}_i^*](t=0) = 0$  for  $i > 0$  and  $\text{P}(t=0) = 0$ . The product yield ( $[\text{P}^*](t)/[\text{R}_0^*](t=0)$ ) as a function of time can be expressed in closed analytical form for a system of linear differential equations. Increasing the number of IVR steps (i.e. larger  $n$ ) between the initial ( $\text{R}_0^*$ ) and the fully relaxed ensembles ( $\text{R}_n^*$ ) allows to more accurately describe the relaxation process. The solution of the kinetic differential equations were fitted to the results of the trajectory calculations for the first nanosecond with increasing  $n$  values, corresponding to increasingly detailed relaxation models. A good fit is obtained if the  $v_9 = 4$  excitation relaxes with a single step ( $n = 1$ ) whereas for excitations  $v_9 = 5$  and 6 two and three-step relaxation models were required ( $n = 2$  and  $n = 3$ ), respectively. The need for additional relaxation steps to describe the observed kinetics for higher excitations is related to the fact that more energy has to be

redistributed from a single mode, indicating that their  $R_0^*$  state is further from the fully equilibrated state. The first-order rate coefficients allow the definition of characteristic times  $\tau = k^{-1}$  for the processes. The product yields and the fitted curves of the corresponding relaxation models are shown in Figure 3.18, and the fitted characteristic times are reported in Table 3.1.



**Figure 3.18:** Product yield as a function of time for the decomposition of sulfuric acid in free dynamics simulations after exciting the OH stretching vibrations with  $v_9 = 4, 5$  and  $6$  quanta in a canonical ensemble of molecules at  $300$  K. A simple kinetic model (see text) was fitted to the data, which allows determination of characteristic times (see Table 3.1) for first order relaxation and decomposition processes ( $\tau_{\text{eq}}$ ) at the corresponding energies. Left panel reports data up to  $1$  ns with both axes on a logarithmic scale. The right panel (with yield on a linear scale) provides the data up to  $40$  ns which has only been determined for  $v_9 = 6$ . It is found that beyond  $\approx 4$  ns additional IVR time scales appear.

The general observation is that decomposition steps ( $\tau_n$ ) speed up with higher excitation and slow down with the progress of relaxation. This is in accord with the fact that a potential barrier has to be surmounted for the reaction to take place and the excitation of the OH stretching mode plays an important role in this. Decomposition from a non-relaxed ensemble ( $\tau_{n-1}$ ) is 1 to 2 orders of magnitude faster than decomposition from the relaxed ensembles ( $\tau_n$ ):  $19.5$  ns  $<$   $360$  ns ( $v_9 = 4$ ),  $[233$  ps,  $680$  ps]  $<$   $18.6$  ns ( $v_9 = 5$ ), and  $[30.6$  ps,  $64.0$  ps,  $403$  ps]  $<$   $3$  ns ( $v_9 = 6$ ). In contrast to that, relaxation times for the IVR steps,  $\tau_{\text{IVR}}$ , for all excitations are similar ( $\tau_{\text{IVR},n-1} = 80.7$  ps,  $92.9$  ps,  $148$  ps for  $v_9 = 4, 5$  and  $6$  and  $\tau_{\text{IVR},n-2} = 23$  ps,  $28.5$  ps for  $v_9 = 5$  and  $6$ , respectively). This implies that the relaxation kinetics is less energy dependent the closer the ensemble is to its fully relaxed state.

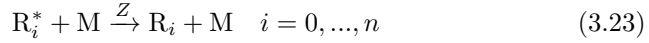
$v_9$	4	5	6	6'
$E$ (kcal/mol)	46.8( $\pm$ 2.3)	55.4( $\pm$ 2.3)	63.5( $\pm$ 2.3)	
$n$	1	2	3	
$\tau_{\text{IVR},n-3}$	-	-	3.9	5.2
$\tau_{\text{IVR},n-2}$	-	23	28.5	33.7
$\tau_{\text{IVR},n-1}$	80.7	92.9	148	145
$\tau_{n-3}$	-	-	30.6	
$\tau_{n-2}$	-	233	64.0	
$\tau_{n-1}$	19500	680	403	
$\tau_n$	358000	18600	3000	

**Table 3.1:** Number of IVR steps and corresponding relaxation ( $\tau_{\text{IVR},i}$ ) and decomposition ( $\tau_i$ ) times (in ps) for  $v_9 = 4, 5$ , and 6 inferred from the kinetic model by fitting to the trajectory data. For  $v_9 = 6'$ , no H-transfer was allowed during the simulation (see text) and only the IVR rate parameters were tuned, while other parameters were those from  $v_9 = 6$ . Relative fitting errors in the time scales  $\tau$  are around 5 % at most. The average total energy and its standard deviation are reported for all ensembles.

Phenomenologically, this can be understood by considering the molecule as a system of coupled oscillators with well-defined coupling constants between them. The coupling parameters determine the relaxation time – i.e. the rate of energy exchange between various modes. With increasing energy content of individual vibrational modes higher order couplings become more important. Consequently, the relaxation timescales for the different excitations become similar as the ensembles get closer to their equilibrium.

For excitations  $v_9 = 4$  and 5 the relaxation is significantly faster than decomposition from the initial non-equilibrium ensemble, which results in low conversion ( $\approx 0.7\%$  and  $24\%$ ) within 1 ns. For excitation of  $v_9 = 6$  vibrational relaxation and direct decomposition have more similar rates, which leads to 66% conversion within 1 ns. Almost complete (i.e. 99% conversion) is expected within 1.6  $\mu$ s, 80 ns and 11.6 ns for excitations of  $v_9 = 4, 5$  and 6, respectively. As a comparison, the 40 ns trajectories with excitation of  $v_9 = 6$  show conversion of 95.5 % within 11.6 ns after photoexcitation which compares with 99 % from the kinetic model (see right panel of Figure 3.18). Although the two approaches are in reasonable agreement with each other, the difference suggests that there is an additional process at longer times which can in fact be modelled if the number of relaxation steps is increased. However, whether or not such a process is relevant (and observable) depends on the environment. Under upper stratospheric-lower mesospheric conditions the deactivation of

excited  $[\text{H}_2\text{SO}_4^*]$  molecules ( $\text{R}_i^* \rightarrow \text{R}_i$ ) due to collisions with other molecules (“third bodies” (M)) with a second-order rate coefficient equaling the collision frequency ( $Z$ ) need to be included.



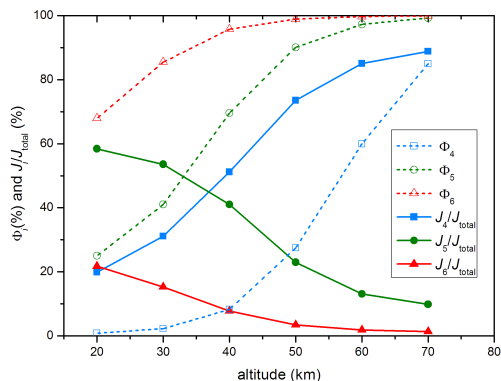
Assuming typical polar (75°S) conditions in the stratopause (at a height of 50 km) with a pressure  $p \sim 70$  Pa, a temperature  $T \sim 280$  K and a sum of the collision radii  $r_{\text{O}_2/\text{N}_2} + r_{\text{H}_2\text{SO}_4} \sim 5\text{\AA}$  within the strong collision assumption (i.e. one collision is sufficient to deactivate the excited molecule), quenching occurs on the 130 ns timescale (inverse of the collision frequency,  $\tau_{\text{quench}} = (Z[\text{M}])^{-1}$ ), with quantum yields of approximately 30%, 90% and 99% for the 4, 5 and 6 excitation, respectively. This shows that the lowest ( $v_9 = 4$ ) excitation cannot lead to full conversion in the ensemble due to quenching. Previous computational studies suggested integrated absorption cross sections of  $2.0 \cdot 10^{-21}$ ,  $1.9 \cdot 10^{-22}$ ,  $2.6 \cdot 10^{-23}$   $\text{cm} \cdot \text{molecule}^{-1}$  for excitations with 4, 5 and 6 quanta, respectively.<sup>144</sup> They assumed unit photolysis quantum yields and concluded that the contribution of the highest excitation is negligible. Hence, first-order photolysis rate coefficients were considered only for excitations with 4 and 5 quanta. According to their calculations, the total photolysis rate coefficient varies only a little as a function of altitude due to slight changes in actinic flux with height.

From the lower mesosphere (70 km) down to lower altitudes in the stratosphere (down to 20 km) the pressure and thereby the quenching rate increases by 1 to 3 orders of magnitude, which suggests that excitations with  $v_9 = 5$  can dominate over  $v_9 = 4$  and excitation into  $v_9 = 6$  can also play a significant role in determining the photolysis rate at lower altitudes. In order to estimate the altitude-variation of the contribution of these excitations to the photolysis rates we carried out more rigorous calculations at various altitudes.

Considering pressures and temperatures typical in the relevant altitude range (see Figures 3.1 and 3.2 in reference<sup>202</sup>) at a latitude of 75°S (Antarctic), the photolysis quantum yields were calculated from the asymptotic ( $t \rightarrow \infty$ ) solution of the kinetic model (Eq. 3.18) including quenching (Eq. 3.23). At this latitude the temperature varies non-monotonically between 200 – 280 K within the altitude range of 20 – 70 km, which results in a reduction by 0.55 – 2.72 kcal/mol in the average thermal energies of sulfuric acid molecules

compared to our simulations starting from thermalized sulfuric acid molecules at 300 K ( $8.16 \pm 2.27$  kcal/mol). This reduction is negligible compared to the average total vibrational energy of sulfuric acid molecules after photoexcitation (46.8, 55.4 and 63.4 kcal/mol, see Table 3.1). Therefore, the decomposition and relaxation lifetimes ( $\tau_i$ ) determined from fitting are good approximations for the relevant altitude range. The actinic flux in the investigated altitude and wavelength range (absorption lines at 741, 606 and 517 nm for  $v_9 = 4, 5, 6$ ) varies only little with wavelength assuming clear sky conditions. Beside these assumptions, the photolysis quantum yields ( $\Phi_i$ , where  $i = v_9 = 4, 5, 6$ ) and fractional contributions ( $J_i/J_{\text{total}}$ ) to the total photolysis rate are calculated for the altitude range of 20 – 70 km (see Fig. 3.19). At these altitudes the characteristic quenching time will be comparable to the decomposition lifetime, which causes significantly reduced quantum yields for the lower excitations. Below 40 km photolysis via excitation of the OH stretch with 5 quanta becomes dominant and at an altitude of 20 km excitations with 4 and 6 quanta will have a 20 % contribution to the photolysis rates (see Fig. 3.19 at 20 km  $J_4/J_{\text{total}} = J_6/J_{\text{total}} = 20$  % and  $J_5/J_{\text{total}} = 60$ %). Most importantly, the reduced quantum yields with excitation of  $v_9 = 4$  and 5 in the stratosphere (below 50 km) cause significantly reduced photolysis rates compared to the case when quantum yields of 1 are assumed for all excitations. Due to quenching at a height of 70 km we predict a reduction of 14 % in total vibrationally induced photolysis rate coefficient. And this rate drops further due to even faster collisional quenching by a factor of 2.6 and 24 when descending to altitudes of 50 km and 20 km, respectively. Our calculations an extension and specification to the results of Vaida et al.<sup>144</sup> and show that a multi-level modeling is required to understand the significance of vibrationally induced decomposition pathways of sulfuric acid in the middle atmosphere.

In order to investigate the role of H-transfer in the IVR process, 1000 additional 300 ps trajectories were run for  $v_9 = 6$  in which the H-transfer channel was closed. Hence, only one FF for  $\text{H}_2\text{SO}_4$  and two for  $\text{H}_2\text{O}+\text{SO}_3$  were used in the simulations. The same three-step kinetic model was fitted, but only the IVR timescales were allowed to vary while the water elimination parameters  $\tau_n$  remained unchanged. An accurate fit was achieved with similar, small parameter variations (see column  $v_9 = 6'$  in Table 3.1). The removal of H-transfer increased the timescales of first two IVR steps significantly, by 34% and 18%, which suggests that H-transfer plays a significant role in accelerating IVR.



**Figure 3.19:** After exciting OH-stretching vibrational mode with 4, 5 and 6 quanta in sulfuric acid, photolysis quantum yields ( $\Phi_i$ ) and fractional contributions ( $J_i/J_{total}$ ) to the total photolysis rate ( $J_{total}$ ) are shown as a function of the altitude in the stratosphere and lower mesosphere.

### 3.5 Conclusions

In the present work it is demonstrated that rigorous reactive molecular dynamics simulations for small molecules in the gas phase can be used to characterize the final state distribution of photofragmentation processes. For the specific case of vibrationally induced photodissociation of  $H_2SO_4$  to form  $H_2O$  and  $SO_3$  the final state distributions will be useful in experimentally identifying such a reaction mechanism. Related work on Cl-CN following a more simplistic approach than the one pursued in the present work showed that a direct link between experiment and atomistic simulations can be established.<sup>203,204</sup> Contrary to previous simulations for water elimination in  $H_2SO_4$  which were more qualitative in nature,<sup>66,139</sup> the present work employs a refined and physically more meaningful PES which allows to follow both, H-transfer and water elimination within the same framework. This is necessary because the two processes have similar barrier heights and can compete with each other. The results suggest that direct water elimination without preceding H-transfer follows exponentially decaying rate laws whereas the rate of the



process proceeding through H-transfer is largely time-independent on the time scale of the present simulations (1 ns).

It is also interesting to compare the present results with earlier and more qualitative work.<sup>66, 139</sup> Although a statistically significant number of reactive trajectories could be analyzed from the previous ARMD simulations, the computational model only allowed to follow one reaction channel, namely water elimination. This led to quite different reaction time distributions compared to the present work which exhibits a more diverse and realistic reaction dynamics. Furthermore, the important torsional degrees of freedom are much more realistically represented in the current work but were only treated qualitatively in the earlier parametrization. The barriers for isomerization were much higher and therefore it was more difficult to transfer energy from the vibrationally excited OH stretching vibration into the torsions. Finally, the reaction barrier in the MS-ARMD PES is close to the best *ab initio* calculations but overestimated by several kcal/mol on the earlier PES. All these differences lead to modifications in the reaction time distributions. Nevertheless, the overall conclusion from the previous investigation, namely that vibrationally induced photodissociation is a relevant and probable decay mechanism in H<sub>2</sub>SO<sub>4</sub> under atmospheric conditions, remains unchanged.

The kinetic model used to analyze the rates following excitation suggests that a one-step relaxation process is sufficient to explain the data when exciting  $v_9 = 4$  whereas an increasing number of relaxation steps is required for excitation of  $v_9 \geq 5$ . This is consistent a) with the fact that no water elimination with previous H-transfer has been observed (i.e. only direct water elimination) in the simulations with excitation of  $v_9 = 4$  and b) with the fact that the barrier for water elimination (32.0 kcal/mol) is lower than that for H-transfer (36.6 kcal/mol). However, a mechanistic interpretation of the various time scales in terms of elementary processes is not straightforward/evident. This is related to the fact that all elementary processes (H-transfer directly after excitation; subsequent H-transfers; water elimination) are characterized by distributions of reaction times and depending on the degree of relaxation, subsequent processes sensitively depend on the history of the trajectory. Including the possibility for third-body collisions and using realistic physical conditions for the relevant altitudes suggests that excitation of  $v_9 = 5$  can become dominant for vibrationally induced photolysis below 40 km. For excitation by 4 quanta the quantum yield drops quickly with decreasing altitude, thereby excitation by 6

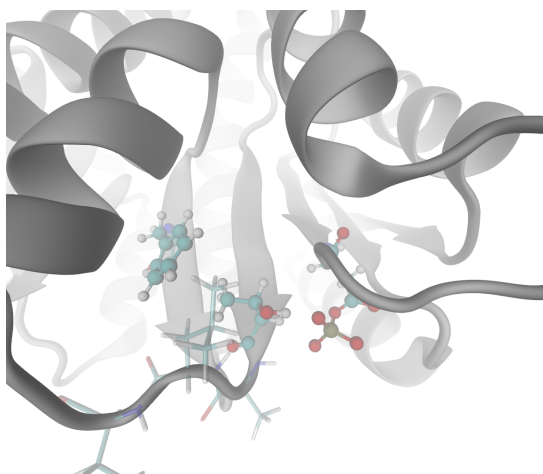
quanta can also become relevant for the photolysis rate in the lower stratosphere.

The present framework is a generic computational strategy to investigate non-RRKM decomposition reactions following vibrational excitation of small molecules in the gas phase. A challenging extension of the present work will consist of applying it to the reaction dynamics of excited molecules in solution. Several interesting experimental studies which provided fragmentation patterns following the excitation of small molecules in solution have recently appeared.<sup>131,205</sup> As the present and previous investigations have shown, simulations can provide a deeper understanding of energy migration pathways in such complex systems.<sup>131,206</sup>

# 4

## REACTION IN CONDENSED PHASE

---



*“Science, you don’t know, looks like magic.”*

—Christopher Moore.

**Contents**

---

---

<b>2.1</b>	<b>Molecular Dynamics . . . . .</b>	<b>28</b>
<b>2.2</b>	<b>Force Fields . . . . .</b>	<b>34</b>
2.2.1	Force Field Parameterization . . . . .	38
<b>2.3</b>	<b>Nonadiabatic Dynamics . . . . .</b>	<b>44</b>
2.3.1	Mean-Field (Ehrenfest) Method . . . . .	44
2.3.2	Surface Hopping . . . . .	46
<b>2.4</b>	<b>Reactive Force Fields . . . . .</b>	<b>49</b>
2.4.1	Adiabatic Reactive Molecular Dynamics . . . . .	50
2.4.2	Multi-Surface/ Adiabatic Reactive Molecular Dynamics	54
2.4.3	ReaxFF . . . . .	60
<b>2.5</b>	<b>Rare Events Techniques . . . . .</b>	<b>65</b>
2.5.1	Umbrella Sampling . . . . .	67
2.5.2	Metadynamics . . . . .	69
<b>2.6</b>	<b>Quantum Mechanics/Molecular Mechanics . . . . .</b>	<b>73</b>

---

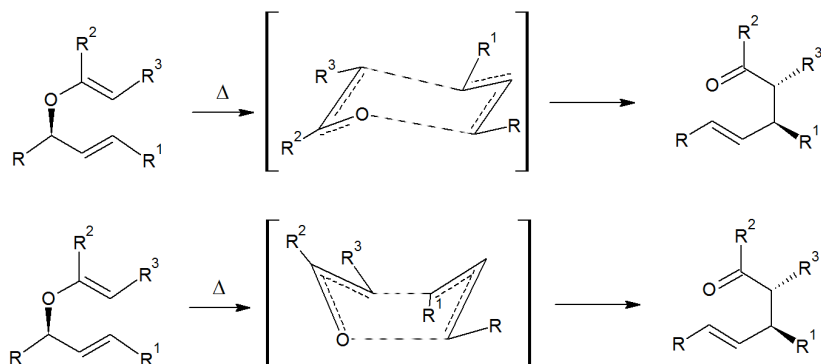
## 4.1 Claisen Rearrangement Mechanism

### 4.1.1 Introduction

#### 4.1.1.1 Claisen Rearrangement

The [3,3]-sigmatropic rearrangement of allyl vinyl ethers, which allows the preparation of  $\gamma,\delta$ -unsaturated carbonyl compounds, is an important reaction due to its special synthetic relevance, which make the [3,3]-sigmatropic rearrangement one of the widely used in organic synthesis. Many different version of the [3,3]-sigmatropic rearrangement, after the Claisen reaction was proposed, has been developed. For example Carroll Rearrangement, which the final product is the  $\gamma,\delta$ -unsaturated ketones, starting from a thermal rearrangement of allylic  $\beta$ -ketoesters, followed by decarboxylation. This reaction require temperatures from 130-220 °C.<sup>207</sup> Eschenmoser Rearrangement, exchange amide acetals with allylic alcohol, here the rearrangement occurs on the N,O-ketene acetals to yield  $\gamma,\delta$ -unsaturated amides. This reaction has a great applicability in alkaloid synthesis.<sup>208-210</sup> Furthermore, other important [3,3]-sigmatropic rearrangement such as Jhonson Rearrangement, Ireland-Claisen Rearrangement, reformatsky-Claisen Rearrangement, Thio-claisen Rearrangement, Aza-Claisen Rearrangement, Chelate Claisen Rearrangement, Retro-Claisen Rearrangement and Metallo-Claisen rearrangement are summarized in the excellent review written by Martin.<sup>211</sup>

**Chemical Kinetics** Synchronic evolution for these reactions, through aromatic transition states was accepted, consisting of a combination of  $\sigma$  and  $\pi$  overlap of 2p atomic orbitals of the carbon atoms for both allyl fragments.<sup>212</sup> It was concluded that, out of the two feasible geometries for the transition state, the reaction proceed through chair-like intermediates, instead of boat-like intermediates, Figure 4.1. Both transition states corresponds to supra-supra processes and, therefore, allowed by the Woodward-Hoffmann rules and the frontier orbital theory.<sup>213</sup>

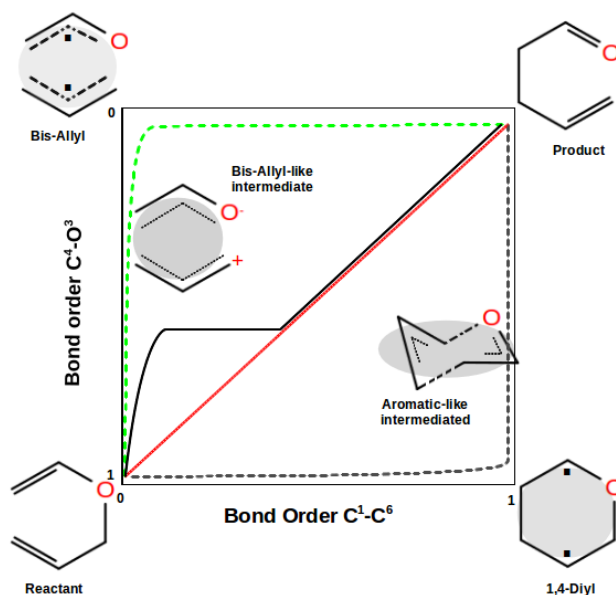


**Figure 4.1:** Two possible Transition State for Allyl vinyl ether and their derivatives

The intramolecular cyclic character (six member ring) of the rearrangement is generally accepted. However, understand the precise nature and geometry of the transition state is controversial. Theoretical calculations, which predict the structures for the transition states involved in the Claisen rearrangement, have been reported.<sup>214–216</sup> In general, a concerted rearrangement through a chair-like transition state is the common pathway. However, some other calculations, focuses in the stereoselectivity of the Ireland-Claisen rearrangement of cyclohexenyl silyl enol ethers, predicts both the chair or boat preferences in the transition state, Figure 4.1. This feature depends of the substituents on the cyclohexenyl ring.<sup>215</sup> Experimental based on kinetic isotope effects, have been performed in order to determine the geometry of the transition state of aliphatic and aromatic Claisen rearrangement.<sup>217–220</sup> However, no general agreement about the geometry of the transition state can be related from theoretical predictions.

In general, 3 different mechanisms can be drawn from QM calculation and experiments: i) synchronus concerted, in which the reaction take place *via* an aromatic-like TS. ii) Stepwise mechanism through a bis-allyl-like TS, in this pathway the C-O bond breaking take place in the first step and iii) 1-4-diyll-like TS in which C-C bond formation take place in the first step, Figure 4.2. An interesting experiment using real-time vibrational spectroscopy, suggest a new mechanism apart from the three above mentioned. This new mechanism, was proposed based on the sub-5 fs laser pulse experiments.<sup>221</sup> Three steps are proposed: First, the C-O bond is weakened to generate a bis-allyl-like intermediate. Consequently, formation of a weak C-C bond results in the

generation of an aromatic-like intermediated. Finally, C-O bond breaks and C-C bond formation occurs simultaneously to generate allyl acetaldehyde, Figure 4.2.



**Figure 4.2:** More O'Ferrall-Jencks diagram for the aliphatic Claisen Rearrangement. Red curve, represents the synchronous concerted pathway reaction *via* an aromatic-like TS. Green dashed curve, represents the stepwise pathway reaction *via* bis-allyl-like TS. Gray dashed curve, represents the stepwise pathway reaction *via* a 1-4-diyl like TS. Black curve, represent the alternative TS.

**Factors Affecting the Reaction Rate** The most frequently reported Claisen rearrangement-thermal isomerization of allyl vinyl ethers, is a process that requires high temperatures and proceeds quite slowly at atmospheric pressure. To transform this reaction into a synthetically useful procedure, numerous attempts to find milder experimental conditions have been reported. The introduction of different substituents in the carbon skeleton of the substrate as well as variation of the catalyst are worth mentioning.

**Influence of the substituents** Claisen rearrangement depends on the inductive or mesomeric effects of electron-withdrawing, or electron-donating substituents,

located at different positions of the carbon skeleton. For example, the effect of the cyano group at different positions of allyl vinyl ethers.<sup>222</sup> The electronic effect product of the substituent, was modeled using Hückel molecular orbital (HMO) theory. The difference of  $\pi$  energy of HMO ( $\Delta E\pi$ ), between the ground state and the transition state, with the value of  $\Delta E\pi$  for the unsubstituted analogue compound predict the magnitude, of the effect of the substituent, in the activation enthalpy of the reaction. In this model, the electron-withdrawing and -donating substituents, are represented as a carbocations and carbanions, respectively. From the delocalized model of the transition state, some qualitative predictions about the effects of the substituents, for Claisen rearrangement reaction, could be made. The main inconvenience of this model, is that the cyano group, is not only an electron-withdrawing group, but also a radical-stabilizing group, so that the acceleration resulting from the presence of a cyano group at C-2 and C-4 may not be a consequence of its electron-withdrawing character. So, what is the real effect of substituents at position 4 or 2? Studies on this particular topic were done, changing the substituent from Cyano group to  $\text{CF}_3$  group.<sup>223</sup> A  $\text{CF}_3$  group is an electron-withdrawing substituent with an inductive character but with no mesomeric one.<sup>211</sup> The allyl vinyl ethers, having a  $\text{CF}_3$  group at C-2, is 73 times faster with respect to the unsubstituted substrate. On the other hand, a  $\text{CF}_3$  group at C-4 has no influence in the reaction rate. These results suggests that, the electron-withdrawing character of the substituents at C-4, is not responsible for the increase in the rate, but rather to stabilize radicals, which is reflected in the stabilization of the transition state. Similarly for  $\text{CF}_3$  group at C-1 does not modify the reaction rate, whereas the effect of fluorine atom, at the same position, will depend on the influence of an alkyl substituent R at C-2.<sup>224</sup> The prediction of such effects, using several substituents in the Claisen rearrangement rate have been proposed.<sup>225</sup> Those model suggest that, the structure of the transition state adopts the features of the substrate or product depending of the exothermic properties of the reaction. In addition, it will have associative or dissociative character, according to the way that the substituents can stabilize such character, Figure 4.2. Also theoretical calculations on the effect of cyano, amino, and trifluoromethyl substituents on the rate, which are in agreement with the experimental studies, have been reported.<sup>1</sup> The effect of alkoxy groups has been largely studied by Curran.<sup>226,227</sup> An electron-donating substituent (alkoxy group) at C-6 sharply accelerates the Claisen rearrangement.<sup>226</sup>

In general, donor and acceptor substituents, at position 1,2 and 4, increase the rate in comparison with just an hydrogen atom in the same position. At position



5 and 6 the effect are complementary. A rate acceleration is observed with donor groups at position 6 and acceptor groups at position 5, in contrast, the reaction is decelerated when the donor and acceptor groups are interchanged.<sup>211</sup>

**Influence of the Solvent** Water, as a solvent, influence the Claisen rearrangement reaction when is used with aliphatic substrates. Measurements of the first-order rate constant of the allyl vinyl ether, in solvents, decrease in the order: water > trifluoroacetic acid > metanol > ethanol > cyclohexane, which reinforce the polar character of the transition state. A similar accelerating effect has been observed with other substrates.<sup>228</sup> Theoretical calculations, have corroborated the role of polar solvents.<sup>229</sup> Explicit water molecules make hydrogen bonds with oxygen atoms of the solute, which are stronger in the transition state than in the substrate. In addition, the presence of two water molecules promote more dissociative and polarized transition state.<sup>211</sup>

The polar nature of the reaction, was reinforced by the fact, that the rate of the rearrangement of allyl *p*-tolyl ether gradually increased as the polarity of the solvent was higher.<sup>230</sup> The contribution of the polar effects, to the increase in the reaction rate, was first observed in ortho Claisen rearrangement of allyl *p*-X-phenyl ethers.<sup>231</sup>

**Enzymatic Claisen Rearrangement** The conversion of chorismate to prephenate, is catalyzed by chorismate mutase, an essential enzyme in the Shikimate pathway. The pathway, is responsible for the biosynthesis of aromatic amino acids in bacteria, fungi and plants.<sup>232</sup> The absence of this pathway in mammals, has made it a desirable target for designing novel antibiotics, fungicides and herbicides. Chorismate mutase, is inherently interesting, since it is a rare example of an enzyme that catalyzes a pericyclic reaction, formally a Claisen rearrangement. This rearrangement also proceeds independent of the enzyme, providing a directly comparison between the catalyzed and uncatalyzed reaction. Despite the extensive experimental and theoretical work, the enzymatic reaction mechanism and the molecular basis of the catalytic behavior of chorismate mutase remains controversial. Electrostatic stabilization of the transition state by the enzyme, and restriction of chorismate to the reactive, but higher energy pseudoaxial conformer, have been proposed as a contributing factors for enzymatic rate enhancement.<sup>233</sup>

---

**4.1.2** Theoretical Methods

---

**4.1.2.1** Claisen rearrangement of the Allyl vinyl ether

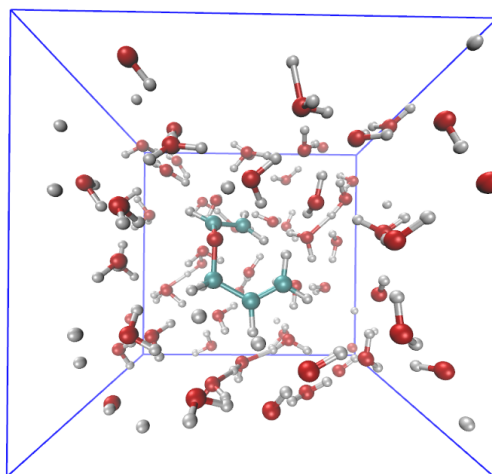
---

*Ab initio* molecular calculations employing Moller-Plesset perturbation theory at second-order (MP2)<sup>234</sup> level of theory with the 6-311G++(2d,2p) basis set, were carried out for all geometry optimizations. Initially, optimized geometries were performed for the reactant and product of the Claisen rearrangement reaction for allyl vinyl ether compound. All theoretical calculations were performed using Gaussian09 package.(reference) Localization of the transition state was done using quadratic synchronous transit (QST) calculations,<sup>150</sup> followed by an intrinsic reaction coordinate (IRC) calculation at the same level of theory. This procedure was done for the reaction in gas phase and in solution. For the compound in solution, implicit solvent (H<sub>2</sub>O) effects, were taken into account using the PCM formalism as implemented in Gaussian09.<sup>18</sup>

All calculated equilibrium structures (local minima and saddle points) of the allyl vinyl ether and the transition states connecting them were characterized by harmonic vibrational frequency calculations at the same level of theory, which also furnished the zero-point vibrational energies (ZPVE) and other thermochemical data.

In order to investigate the role of water molecules in the Claisen rearrangement reaction, a Carr Parrinello Molecular Dynamics (CPMD) together with well tempered metadynamics was performed.<sup>123</sup> Optimization was achieved by CPMD with annealing of the atomic velocities, using a time step of 4 a.u. and a fictitious electron mass of 450 a.u. to solve the equations of motion. Geometries were considered optimized when the largest component of the nuclear gradient was smaller than  $1 \times 10^{-6}$  a.u. Valence orbitals were expanded with a plane wave basis set up to 70 Ry. Perdew-Burke-Ernzerhof functional PBE were used to describe the core electron-valence shell electron interactions.<sup>235</sup> Initially a wavefunction optimization was performed, later, a MD at 370 K was carried out with temperature rescaling for the atoms and not thermostat for the electrons. Velocities and coordinates from the previous step were used for production step. The complete system consists of a unit cell of 12.459 Å , with 64 water molecules in total, in which periodic boundary conditions were applied in all three directions, Figure 6. Well tempered metadynamics was performed with

plumed plugin.<sup>236</sup> A height of  $4.75 \times 10^{-4}$  au, a bias factor of 50 steps and the sigma parameter for the collective variables of 0.35 au were used. The collective variables (CVs), which describe the reaction, were the C7-C9 and O5-C12 distances, these distances belongs to the bond breaking and bond formation during the Claisen rearrangement reaction.



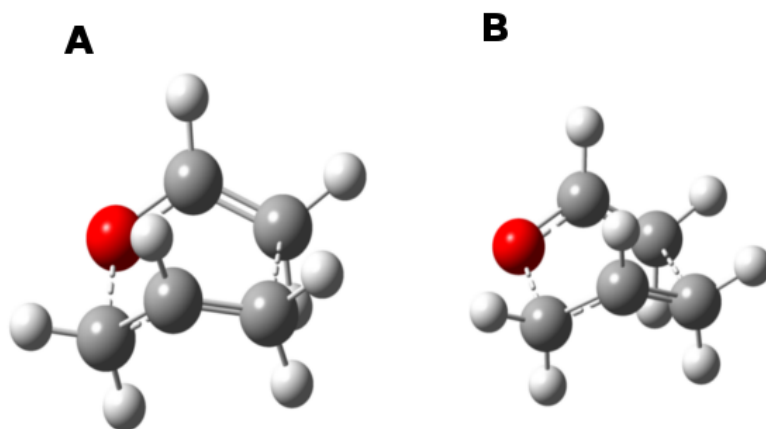
**Figure 4.3:** Molecular system for Allyl vinyl ether (CPMD/Metadynamics simulation)

#### 4.1.2.2 Effect of the substituents

Substituents at C-2 and C-6 (Figure 4.6) were prepared, having in total 7 different molecular systems, including the substitution only with Hydrogen atoms (Allyl vinyl ether, compound 1). This substitution corresponds to a strong donor ( $\text{NH}_2$ ), acceptor (CN), and an alkyl group ( $\text{CH}_3$ ). *Ab initio* calculations were carried out with Gaussian 09,<sup>18</sup> using Moller-Plesset perturbation theory at second-order (MP2)<sup>234</sup> with the 6-311G++(2d,2p) basis set. The reactants, products and transition structures (Figure 4.9) were fully optimized and characterized by frequency analyses. Activation energies are reported in Table 4.1.

**4.1.3** Results

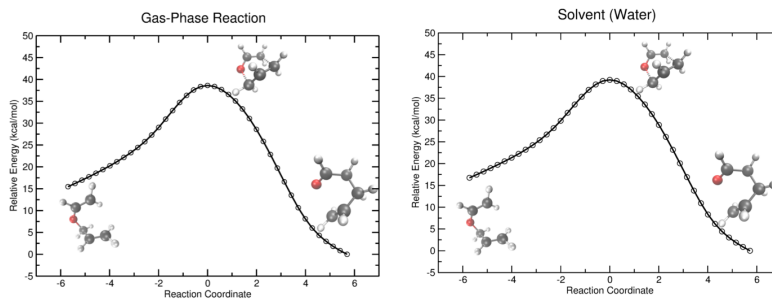
The optimized geometries for the reactants and products showed a nonplanar geometry with the anti conformation about the C2-C3 bond, as was reported in previous studies (Figure 4.9).<sup>237</sup> The dihedral angle about the C2-O bond is found to be close to  $0^\circ$ , in order to maintain conjugation of the oxygen with the vinyl fragment. In contrast, the staggered geometries about the C4-O bond yield preferred dihedral angles of  $180^\circ$ , as was reported by Jorgensen thorough analysis of the conformations of allyl vinyl ether.<sup>237</sup> Transition states, for all the cases, are found to be in in a boat conformation, contrary to those reported in previous studies using Density Functional Theory (DFT), in which the lowest energy saddle point was assumed to be in a chair TS conformation, for the same set of substituents.<sup>237</sup>



**Figure 4.4:** Transition state for A) gas phase and B) water implicit solvent, shows a boat TS conformation. Calculations were done using QST-2 method at MP2/6-311G++(2d,2p) level.

Representative transition state is showed in Figure 4.4 and 4.9, for both, gas phase and implicit solvent.

The barrier height energy, for the Claisen rearrangement in allyl vinyl ether, using implicit solvent, is 28.98 kcal/mol, exceeding by 1 kcal/mol the energy of

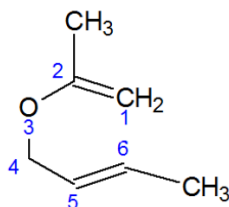


**Figure 4.5:** Claisen-rearrangement intrinsic Reaction Path for Allyl vinyl ether, in gas phase and implicit solvent.

the barrier height compared with previous calculations using DFT.<sup>1</sup> This may indicate that, the barrier height energy for the two possible TS conformation (boat and chair), have very close energies. The minimum energy path for the Claisen rearrangement is shown in Figure 4.5, for the reaction in gas phase and implicit solvent. In both cases, barrier height shows similar energetic values (28.98 kcal for implicit solvent and 28.96 in gas phase), and similar TS conformations, Figure 4.5.

Is well known that water as a solvent, does not affects the barrier height in the Claisen rearrangement reaction. The effect of water molecules is rather to stabilize the TS (see Introduction section). In order to investigate if explicit water molecules can affect the TS conformation, CPMD was performed together with metadynamics simulations (see methods section). The  $\Delta G^\ddagger$  is 28.75 kcal/mol as is shown in Figure 4.7 for CPMD simulation, compared with 25.79 kcal/mol obtained with ES calculation at 298.150 K, using MP2/6-311G++(2d,2p) level.

Figure 4.8 shows the TS found in metadynamics simulation, a boat-like was found as in ES calculations. Interestingly one water molecule is stabilizing the TS, Figure 4.8. The surface does not show any intermediates structures, and maybe because the resolution of the metadynamics. Highest resolution (decreasing the sigma parameter) probably will allows one to better understand the intermediates (i.e., bis-allyl-like) and the electronic mechanism involved in the reaction, such inductive or mesomeric effects.



1	R2 = H	R6 = H
2	R2 = CN	R6 = H
3	R2 = CH3	R6 = H
4	R2 = NH2	R6 = H
5	R2 = H	R6 = CN
6	R2 = H	R6 = CH3
7	R2 = H	R6 = NH2

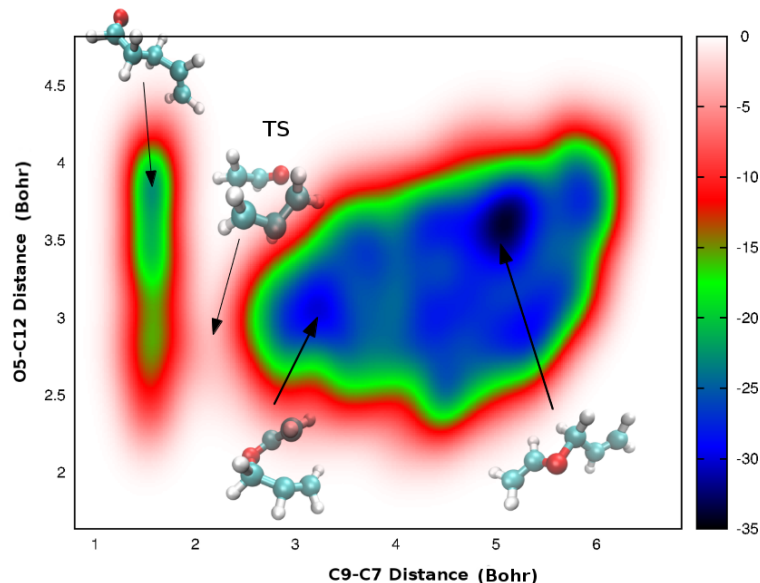
**Figure 4.6:** Schematic representation of the molecular systems used to investigate the effect of the substituents in the Claisen rearrangement reaction. Strong donor (NH<sub>2</sub>), acceptor (CN), and an alkyl group (CH<sub>3</sub>) at positions 2 and 6.

#### 4.1.3.1 Effect of the substituents on Activation Energies

Table 4.1 lists calculated activation energies from this work, and compared with those obtained by previous calculations, using B3LYP/6-31G\*.<sup>1</sup> With the exception of the Allyl vinyl ether, activation energies for the substituents are

**Table 4.1:** Activation Energies for the Claisen Rearrangements of Substituted Allyl Vinyl Ethers at the B3LYP/6-31G\*<sup>1</sup> and MP2/6-311G++(2d,2p) level. The energies are given in kcal/mol

Subst C2	Subst C6	$\Delta E$ MP2	$\Delta E$ B3LYP
R2 = H	R6 = H	29.0	27.9
R2 = CN	R6 = H	18.0	24.8
R2 = CH3	R6 = H	22.1	28.8
R2 = NH2	R6 = H	22.0	21.7
R2 = H	R6 = CN	22.6	31.7
R2 = H	R6 = CH3	23.0	30.1
R2 = H	R6 = NH2	21.4	27.2

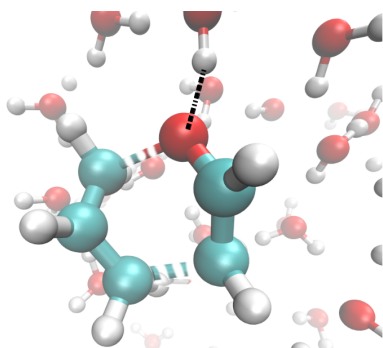


**Figure 4.7:** Claisen-Rearrangement free energy surface on the space on C9-C7 and O5-C12 distance. CPMD-Metadynamics simulations.

overestimates, in some cases with a value of around 9 kcal/mol (e.i., substitution at C-6 for CN). Transition states for all the substituents (see Figure 4.9), shows a boat transition state as for Allyl vinyl ether compound. In this work we found some energetic and structural differences with respect to previous computational works.<sup>1</sup> In order to find the correct method for calculating the activation energy, a further calculation has to be done.

#### 4.1.3.2 Further Procedure

One of the aim of this work is to study the influence of a potent electron-donor groups like NH<sub>2</sub>, electron-withdrawing group like CN, and an alkyl group and the influence on the reaction mechanism at atomistic level. For this, a better description of the TS has to be done, computing the two possible TS (chair-like and boat-like) using accurate level of theory such Coupled cluster (CC), can give insight of the real, or approximated activation energies. Thermochemistry will be necessary to compute the rate constants, and compare with the experimental ones. Calculation of the enthalpies and free energies of reaction can be obtained



**Figure 4.8:** Boat transition state from CPMD/Metadynamics simulations.

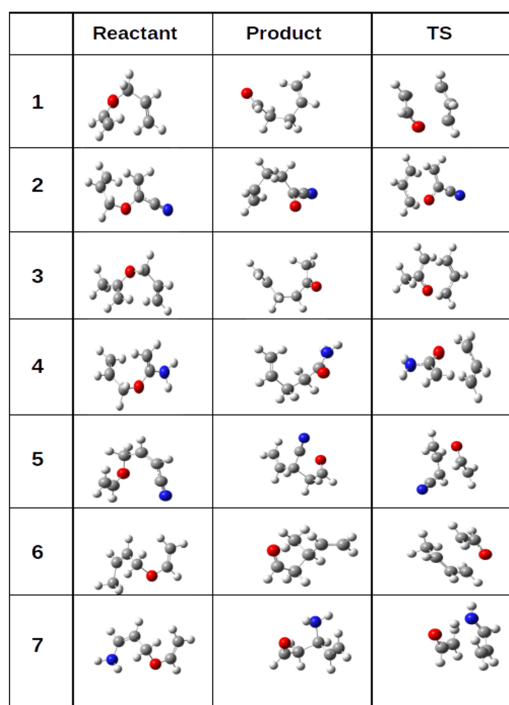
to compute the rate of the reactions, using Arrhenius equation (see introduction). Furthermore, in order to understand the real effect of the substituents in the energetics of the reaction Marcus theory formalism<sup>238</sup> as proposed by Murdoch,<sup>239</sup> to separate intrinsic and thermodynamic contributions to the activation energy, this expression can be written as:

$$\Delta E^\ddagger = \Delta E_0^\ddagger + \frac{1}{2} \Delta E_{rxn} + \frac{(\Delta E_{rxn})^2}{16 \Delta E_0^\ddagger} \quad (4.1)$$

where  $\Delta E^\ddagger$  is the activation barrier of the reaction,  $\Delta E_{rxn}$  is the energy of the reaction, and the intrinsic barrier,  $\Delta E_0^\ddagger$ , is that of a hypothetical thermoneutral process. This permits an estimate of how much of the change in activation energy caused by the substituent is due to an alteration of the heat of reaction, based on the assumption that the extent of transition state lowering by a change in the reaction thermodynamics is a function of the position along the reaction coordinate. The remaining energy change is assumed to be due to an intrinsic effect specific to the method of calculation.<sup>1, 238</sup>

The second aim of this work is to provide an atomistic description of the reaction mechanism. As a dynamical process the mechanism will be evaluated using CPMD as for Allyl vinyl ether together with metadynamics technique, for all the substituents, also relative new technique, to sampling rare events namely selective integrated tempering method (SIT),<sup>240</sup> will be used in order to understand and bring new insight about the real mechanism of Claisen rearrangement reactions.





**Figure 4.9:** Optimized geometries for reactant, product and transition state for parent and substituted allyl vinyl ethers at the MP2/6-311G++(2d,2p) level.

## 4.2 PleD Mechanism mediated by phosphorylation

### 4.2.1 Introduction

Sensing environmental signals and rapid metabolic response plays a crucial role in survival of bacteria. A typical cascade to integrate external stimuli is the so-called "two-component" signal transduction system.<sup>241</sup> In general they consist of a histidine kinase, which is a transmembrane protein, that senses a specific environmental stimulus and transfers a phosphoryl group to a conserved aspartate of a response regulator (RR) protein, who is in turn the second components of the system, which mediates the cellular response.<sup>88</sup> The diguanylate cyclase PleD protein from *Caulobacter crescentus* belongs to RR family, its activity is controlled by phosphorylation where two cognate kinases are involved; DivJ and PleC. PleD is required for polar differentiation in the bacterial cell cycle. Bacterial cells without functional PleD are hypermobile and fail to accomplish the swarmer-to-stalked cell transition.<sup>242–245</sup> PleD contains an intrinsic nucleotide cyclase activity, which converts two molecules of GTP into 3',5'-cyclic diguanylic acid (c-di-GMP).<sup>243</sup> c-di-GMP is a molecule of great interest which regulates surface-adhesion properties and motility in bacteria.<sup>246,247</sup> In addition, c-di-GMP are involved in bacterial biofilm formation and persistence.<sup>248</sup> Because c-di-GMP is found exclusively in bacteria, makes it a potential target for medicinal applications. PleD is a homodimer protein, where each monomer contains two receivers domains; Rec1 ( $\approx$  aa 4 to 122) and Rec2 ( $\approx$  aa 154 to 269), Rec1 serves as phosphoryl acceptor (Asp-53) whereas, Rec2 was proposed to acts as an adaptor for dimerization.<sup>249,250</sup> Also PleD contain GGDEF domain ( $\approx$  160-residue), or catalytic domain, which is a conserved amino acid sequence motif with dyguanylate cyclase activity (DGC). The signaling molecule c-di-GMP is synthesized from two GTP molecules in a symmetric condensation reaction, where each monomer accommodate one GTP in the catalytic domain. Thus, the productive encounter of two GTP-loaded active sites of PleD is required. Indeed, it has been demonstrated by cross-linking experiments that exclusively dimeric PleD is catalytically active.<sup>251</sup> Once two monomers have been associated, the catalytic step solely depends on the correct positioning of the flexibly attached DGC domains. As a con-

sequence, the overall production rate of c-di-GMP strongly depends on the dimerization rate. PleD activity is modulated by its phosphorylation state and dimerization is significantly enhanced upon phosphorylation (or pseudo-phosphorylation by  $\text{BeF}_3^-$ ).<sup>250,251</sup> Allosteric regulation of PleD includes not only product inhibition<sup>249,252</sup> but also phosphorylation-mediated dimerization control.<sup>250,251</sup> The PleD X-ray structure reveals that the dimerization interface are the stem domains D1 and D2.<sup>249,250</sup> The active structure seems to have a phosphorylated aspartic acid in the 53 position (pseudo-phosphorylation by  $\text{BeF}_3^-$  in the case of crystal structures founded in the protein data bank). The pseudo-phosphorylated aspartate form 4 H-bonds with the binding pocket (acidic pocket) which in turn, contributes to the coordination of the adjacent  $\text{Mg}^{2+}$  ion. Hereinafter, and, according with the crystal structure examination, a restructuring of the  $\beta 4\alpha 4$  loop, with Thr83 moving more than 3 Å relative the position in the nonactive state, this motion causes an H-bond formation with the  $\text{BeF}_3^-$ . The space left by Trh83 is occupied by Phe102 which change the side conformation from *gauche+* to *trans* to become buried, which causes that, part of  $\beta 5$  main chain moves up to 2 Å.<sup>249,250</sup> All this structural changes result in a shift/rotation of domain D2 with respect to D1 ( $14^\circ/6$  Å), leading to a more optimal tight-packed dimer interface in activated PleD.<sup>250</sup>

The phosphorylation from one side of the protein, resulting in a chemical reaction in an other different side, could be explain by allostery mechanism. Allostery is a complicated process that still is not fully understood. Many models supported by experiments and atomistic simulations, have been proposed to the date. Those that explain allostery in the classical way (MWC and KNF models),<sup>253</sup> we will called along the text, “the old view”, following the same guidelines used in many publications that addresses the issue. The Monod, Wyman and Chageux model (MWC) describe allostery as a binding event which one side of a protein affecting the activity at the another side due to conformational changes. The MWC model describe the transition as a concerted action between two co-existing states (R and T). The Koshland, Nemethy and Filmer model (KNF), describe the allosteric process as a sequence of events cause by a conformational change in the binding site; "induced fit model", in the old view two components are mandatory; i) there are two distinct R and T conformations, in which, when the ligand is not attached to the protein, they exist in a ratio governed by an equilibrium constant; and ii) allostery involve a change in shape. In summary the all view can then, be described by three characteristics: i) there are only two states T and R; ii) The conformational change in binding site, altered by the substrate; and iii) the signal after the

perturbation is transmitted via single pathway. In contrast more evidence indicates that the native state is a conformational ensemble, and allosteric perturbation involves a shift of pre-existing conformations, even without, there is not a change in the binding site shape (for review see ref.<sup>254–259</sup>). However, the new view definition of allostery change from the old view, in the sense that it emphasizes rather that only two conformational states, proteins exist in ensembles, which can be define allostery in thermodynamic terms permits dividing allosteric proteins into three types of cases. i) governed only by entropy; ii) governed by enthalpy and entropy; and iii) governed by enthalpy. In type I there is not or a subtle backbone structural change; in type II minor conformational changes are coupled with entropic effects; and in type III there is a relatively large domain or local conformational changes. The existence of multiple conformational and dynamics states, implies multiple pathways through which the strain energy is released from the allosteric site following a perturbation event like phosphorylation.<sup>254–259</sup>

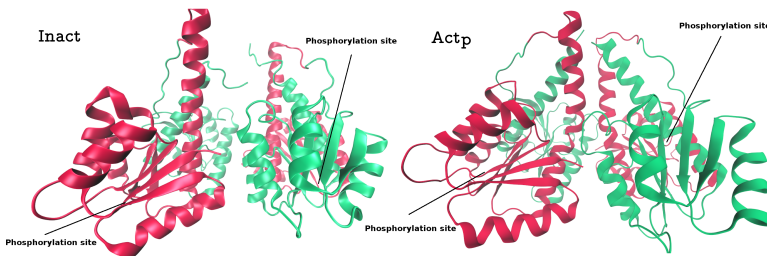
But how phosphorylation can affect allosterically PleD?. At the date, the knowledge about phosphorylation in PleD tells that is a key step for dimerization in active state, indeed, non-phosphorylated protein was found to be in the nonactive state.<sup>251</sup> What happens dynamically after phosphorylation is a subject of debate and research. Phosphorylation is a ubiquitous mechanism of post-translational modification, the phosphate group carries a  $-2$  charge at physiologically relevant pH, resulting in a large electrostatic perturbation, which modulates the energy landscapes, governing the catalytic activity. In many cases, phosphorylation results in switch-like changes in proteins.<sup>89,260</sup>

The cascade proposed by Wassmann et al.<sup>250</sup> envisages the following key characteristics: Phosphorylation at the phospho-acceptor residue Asp53 is communicated to the  $\alpha 4$ - $\beta 5$ - $\alpha 5$  face leading to a reorientation of domain D1 relative to D2 ( $14^\circ/6 \text{ \AA}$ ). Eventually, the quaternary structure is altered such that the dimerization interface is optimized and the interaction between two monomers increases. This suggest that the chain of events, is based on the comparison of the inactive and  $\text{BeF}_3^-$ -activated X-ray conformation and yields useful information on structural changes that occur upon phosphorylation. However, there is lacks information on the signal propagation, from the phospho-acceptor to the domain and dimerization interfaces, and the dynamics of the activation process in general.

Sequence alignment of domains D1 and D2 showed 25% to 37% identity with the members of the response regulator superfamily<sup>249,261</sup> annotated D1 as receiver domain (REC-1) whereas D2 is termed as an adapter domain (REC-2).

Structural changes following phosphorylation have been studied extensively on the single-domain bacterial chemotaxis protein Y (CheY) and they were found to be characteristic to REC domains in general.<sup>6,7,256,262–264</sup> Changes upon phosphorylation are rather localized and involve a major repositioning of the  $\beta 4 - \alpha 4$  loop and the reorientation of the sidechain of a conserved Tyr/Phe residue from solvent exposed to a buried position. Moreover, a highly conserved Thr/Ser moves closer to the phosphorylation site to form a hydrogen-bond with the phospho-moiety ( $\text{PO}_3^{2-}$ ). This mechanism traditionally is the so-called "Y-T" coupling.<sup>7,264,265</sup> The Thr/Ser due to hydrogen-bond formation leaves empty space for the Tyr/Phe to adopt the buried rotameric state. However, the steric interaction between these two residues was shown to be minimal by experimental<sup>266</sup> and computational<sup>256,262,263</sup> works. Free energy calculations found that the role of Thr/Ser is rather energetic where Thr/Ser stabilizes the active conformation of the  $\beta 4 - \alpha 4$  loop introducing a population shift. Previous atomistic simulations strongly suggest a coupling of the Tyr/Phe rotation and the activation state of the  $\beta 4 - \alpha 4$  loop in that Tyr/Phe rotation facilitates the activation of the loop in CheY.<sup>256,262,263</sup> Many NMR experiments on the regulators Spo0F, CheY and NtrC showed that the apo-forms of the molecules can be divided into subpopulations having different conformations, supporting the population shift model for the mechanism of this RR proteins.<sup>90,266,267</sup>

However, many information support the population shift model for the allosteric process in RRs proteins, but is it a common evolutionary mechanism for this kind of proteins?, is the mechanism changing depending of the structural complexity level of the molecules? The present computational study focuses on the characterization of processes associated with the relaxation from the active to the nonactive state in the diguanylate cyclase PleD protein, i.e. dynamic events that occur after removal the phosphate group.



**Figure 4.10:** Interface **Act** and **Inact** in PleD protein. D1 is represented in green color while D2 is represented in red color. Each monomer has a phosphorylation site located in the Asp53 residue.

---

**4.2.2** Theoretical Methods

---

**4.2.2.1** Setup of the "PleD" monomers and dimers

---

The activation of PleD relies on oligomerization to bring two active sites in close proximity for GTP condensation reaction, while the inhibition is usually mediated by the enzymatic product, c-di-GMP, *via* immobilization of catalytic domains in a nonproductive orientations.<sup>250</sup> Experiments have proved that non-activated PleD partly dimerizes in concentration-dependent manner at high protein concentration, but the most relevant dimerization was found to be enhanced by the addition of the mimic phosphoryl-moiety  $\text{Be}_3^-$ , and cations like  $\text{Mg}^{2+}$  or  $\text{Mn}^{2+}$ .<sup>250</sup> Indeed, domains D1 and D2 experience mayor structural changes after the pseudo-phosphorylation compared with the non phosphorylated-inactive state,<sup>250</sup> suggesting that there are initial path for activation that probably start with the reorientation of the domains D1 and D2. In that sense, absence of the stabilizing  $\text{BeF}_3$  moiety in the active conformation will possibly lead to a change from active to inactive in a dynamic transition events. As we are interested in the D1/D2 transitions, the catalytic part of the protein was not having into account. Residues 1 to 285, for each monomer, which correspond to domains D1 and D2 were kept, whereas all other residues were deleted,  $\text{Mg}^{+2}$  ions were conserved in the active and inactive form (PDB codes 1W25<sup>249</sup> and 2V0N<sup>250</sup>). Additional ligand molecules were removed, Figure 4.10. Simulation of domains D1/D2 should be sufficient to investigate the dimerization process including the observed rearrangement of domains D1 and D2 of PleD.

Phosphorylation perturbation lies in the structural changes in the PleD protein,<sup>250</sup> furthermore, removal of such perturbation allows the protein to relax from the active to inactive state, for this reason two new system were prepared; i) for the phosphorylated active state, the phosphate moiety in the Asp53 was removed, in order to differentiate with respect to the active phosphorylated system, we refer as **Act<sub>p</sub>** to the molecular sytem that is in the active state and phosphorylated. For the active non-phosphorylated state we called **Act**. ii) For the non-phosphorylated inactive system, a phosphate was add to the Asp53, as before for the inactive non-phosphorylated system we called **Inact** and for the inactive but phosphoorylated we called **Inact<sub>p</sub>**.

All titratable sidechains were generated in their standard protonation state for pH 7, the HSD model was applied to all Histidine residues and the overall charge of  $-3$  in each monomer was neutralized. Positions of added hydrogen atoms were relaxed using 5000 steps of Conjugate Gradient in NAMD.

#### 4.2.2.2 Molecular Dynamics Simulations

All MD simulations were carried out with NAMD<sup>114</sup> using the CHARMM22 force field.<sup>268</sup> MD simulations were carried out in explicit TIP3 water solvent. The NAMD solvation package was used with a protein-box distance of 10 Å resulting in system sizes of 71,000 and 52,400 atoms approximately for each system. Periodic boundary conditions were used in all simulations and images were updated every 1000 steps. A 12 Å cutoff was used for shifted electrostatic and switched van der Waals interactions, hydrogen atoms were constrained with SHAKE<sup>269</sup> and a timestep of 1 fs was used. The Particle Mesh Ewald (PME) was applied to correct for long range electrostatic interactions.  $K^+$  ions were added to neutralize the system. Starting from the solvated X-ray structure, the system was minimized for 1,000, and then equilibrated for 2ns. Afterwards, equilibration in the NPT ensemble were carried out at  $T = 300$  K and constant pressure of 1 atm for 55 and 65 ns of simulation for **Act<sub>p</sub>** and **Inact** respectively, 40 and 60 ns for **Act** and **Inact<sub>p</sub>** respectively.

#### 4.2.2.3 Biased Molecular Dynamics Simulations

Biased Molecular Dynamics BMD, was done as follows: coordinates and velocities from **Act** were used for initial conditions. an harmonic constraint was applied over the pseudo-dihedral angle  $\phi$ , starting from  $0^\circ$  until  $-110^\circ$ , with a force of  $7 \text{ kcal/mol}/\text{\AA}^2$ . The pseudo-dihedral angle was gradually decreasing every  $20^\circ$  and NPT ensemble were carried out at  $T = 300$  K and constant pressure of 1 atm for 5 ns each, until reach the target angle ( $-110^\circ$ ). The constraint was applied to every monomer in D1/D2 dimer. After reach the target angle, the boundary values of  $-110^\circ$  for  $\phi$  were kept up to 190 ns

**4.2.2.4** Metadynamics

---

In order to observe the barrier energy that is involved in the transition from active to inactive PleD, metadynamic simulations were carried out with NAMD. Pseudo dihedral angle  $\phi$  and  $\psi$  angle (the corresponding dihedral angle in Phe102 consisting of N, C $_{\alpha}$ , C $_{\beta}$  and C $_{\gamma}$  atoms) were chosen as a collective variables. Upper and lower boundaries of  $-140^{\circ}$  to  $140^{\circ}$  for the pseudo dihedral angle  $\phi$  and  $-165^{\circ}$  to  $65^{\circ}$  for  $\psi$  angle were used. The height of the Gaussian hills was 0.01 kcal/mol and the width of  $5.0^{\circ}$ .

**4.2.2.5** Data analysis

---

The  $\beta 4$ - $\alpha 4$  loop conformation which is related to the activation state of response regulators, can be quantified by the pseudodihedral angle,  $\phi$ , consisting of C $_{\alpha}$  atoms of the  $\beta 4$ - $\alpha 4$ , residues Thr83-Ala84-Leu85-Asp86 D1 domain. An angle of  $110^{\circ}$  which corresponds to the fully active conformation whereas a range of  $-100^{\circ}$  to  $-90^{\circ}$  is found in the inactive state. The angle  $\chi$  formed by the N of peptide bond of the  $\beta$  carbon and  $\gamma$  carbon for Phe102 was measured to determine the gauche(+) conformer ( $90^{\circ}$ ) in the inactive state and *trans* conformer ( $145^{\circ}$ ) in the active state. Interaction between the phosphate in the Asp53 and Thr83 residues, has been seen in active state, the distance between oxygen-hydroxyl of Thr83 and C phospho-ester in Asp53 also has been monitored, for evidence of H-bond formation. there are two key salt-bridges between domain D1 and D2 (asp257-arg88 and glu260-arg91), which are exchange when the transition from inactive to active occurs. The distance between the Carbonyl carbon of the two negative residues (Asp-257 and Glu-260) and the carbon of the amino group of the two positive residues (Arg-88 and Arg-91) were used for monitoring the salt bridges in the D1-D2 interface for each monomer.



---

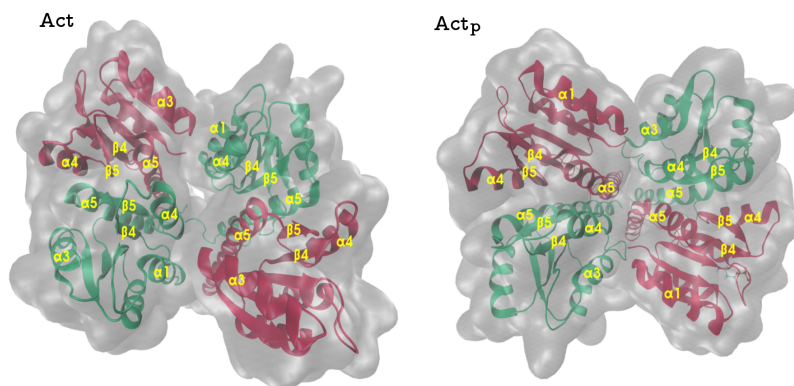
**4.2.3** Results

---

**4.2.3.1** **Act<sub>p</sub>** and **Inact** dynamics

---

Molecular dynamics simulations were carried out for **Act<sub>p</sub>** and **Inact**, during 55 ns. Several structural features, which keep the dimer in the active or inactive form, were found to be common compared with the crystal structure.<sup>249,250</sup> In the case of the **Act<sub>p</sub>** system, not further structural changes, related with the crystal, were found. In the acidic pocket, Mg<sup>2+</sup> cation is coordinated by phosphate in Asp53, also the carbonyl oxygen of the Asp53, Asp10, the C carbonyl of the backbone in Met55 and two water molecules, that were found to be very stable during the all 55 ns. In contrast, and because the absence of the phosphate moiety, for the **Inact** system, the Mg<sup>2+</sup> cation is coordinate by Asp53, Asp9, Asp53, and 3 water molecules (Figure 4.10). Asp9 and an additional water molecule coordinates the Mg<sup>2+</sup> cation in the **Inact** state, instead of Met55 in the **Act<sub>p</sub>** state. One of the main features, which is observed in other RRs, is the apparently coupling between, the Thr83 and Phe102. this coupling is mediated by the structural change of  $\beta$ 4- $\alpha$ 4 loop. In the crystal structures the Thr83 moves more than 3 Å, relative to the inactive state position, forming an H-bond with the BeF<sub>3</sub><sup>-</sup> moiety (phosphate group in this study), Figure 4.12A. The vacated space, is occupied by Phe102, which change its side chain conformation from gauche+ to *trans*, to become buried. This causes that the  $\beta$ 5 (residues 102-104) moves up to 2 Å, Figure 4.13C.<sup>250</sup> As it was observed in the crystal, in the dynamic simulation the interaction of Thr83 and phosphate-moiety, through H-bond is very stable during the all simulation, Figure 4.13C and 4.14. Phe102, which change its side chain, in the inactive with respect to the active crystal structures, is also stable during the simulation, Figure 4.14. The  $\chi$  angle fluctuates around 140° during the 55 ns of simulation, which is the approximated angle in the active crystal structure. Occasionally, the  $\chi$  angle change close to the inactive value in around 33, 45, 45.7 47 and 110 ns, but only for few ps, then returns to its predominant structure in the active form, Figure 4.14. For the **Inact** dynamics, the Thr83 was found to be apart from Asp53 with a fluctuate distance value around 3.7Å, which is how is found in the crystal structure. The  $\chi$  angle for Phe102, in the dynamics also conserves the features seen in the crystal structure, the angle fluctuates around a value of 90°.



**Figure 4.11:** Representation models for active and inactive PleD. D1 is represented in green color while D2 is represented in red color

Domains D1 and D2 are assigned to the class of CheY-like receiver domains for which repositioning of the  $\beta 4 - \alpha 4$  loop, accompanied by a side chain rotation of the conserved residues Thr(T)/Ser and Tyr(Y)/Phe (Thr83 and Phe102 in PleD), is characteristic. This rotation is often termed "Y-T" coupling in CheY protein. The activation state of the  $\beta 4 - \alpha 4$  loop can be characterized by the pseudodihedral angle,  $\phi$ , of the loop C $\alpha$ -atoms,<sup>6,262</sup> residues 83 to 86 (Thr-Ala-Leu-Asp) in domain D1 according to the alignment by Hecht et al.<sup>261</sup> Unphosphorylated CheY, including meta-active CheY where Tyr/Phe occupies the buried position in absence of phosphorylation, shown that the range of  $\phi$  angle between  $10^\circ$  to  $30^\circ$  corresponds to the inactive conformation, whereas values of  $\approx 110^\circ$  characterize BeF<sub>3</sub>-activated CheY.<sup>6,262</sup> The  $\beta 4 - \alpha 4$  loop, Figure 4.13C, is also very stable in its active and nonactive state, Figure 4.14. The pseudo-dihedral angle ( $\phi$ ), which measures the change of the  $\beta 4 - \alpha 4$  (see methods), fluctuates around  $110^\circ$  and  $-110^\circ$  for **Act<sub>p</sub>** and **Inact** respectively.

The  $\alpha 4 - \beta 5 - \alpha 5$  faces from D1 and D2 since to be a conserved domain in RRs, Figure 4.10 in which intramolecular and intermolecular inhibitory and activating interactions with themselves and effector domains, are performed. This molecular surface makes a enormous differences between the nonactive and active conformations.<sup>270</sup> In many RRs, overlapping surfaces of the  $\alpha 4 - \beta 5 - \alpha 5$  face are used for interactions with different targets, in the nonactive and active states. In PleD Domain D1 interacts with the adaptor domain D2 across the  $\alpha 4 - \beta 5 - \alpha 5$  faces of both domains. Activation results in a domain rotation/shift of  $14^\circ / 6 \text{ \AA}$  of D2 relative to D1 and a massive repacking of the domain interface, Figure 4.10.<sup>250</sup> In the molecular dynamics simulations for **Act<sub>p</sub>** and **Inact**,

the  $\alpha 4$ - $\beta 5$ - $\alpha 5$  faces, does not experience large changes from those found in the crystal structure, Figure 4.11. Surprisingly, the **Inact** dimer does not fall apart each other, suggesting that the nonactive dimers are stables. Arg91, Val110, and Ile251, contribute to the interface in the **Inact** state, while, Ar88, Ile92, Leu95, Val241, Gln250 and Ala263 contribute only in the activated state, Figure 4.12B-D. The three interdomain salt bridges (Asp108-Arg237, Arg115-Asp250, Asp101-Arg264) are common interactions for the **Act<sub>p</sub>** and **Inact** state, and as the intermolecular interactions above mentioned, this salt bridges are very stables during the 55 ns of simulation, Figure 4.13A. Another interdomain salt bridge interaction, between Arg4 from monomer 1 and Glu221 of monomer 2, showed to be stable during the molecular dynamics simulation only in the **Act<sub>p</sub>** state, Figure 4.13B

#### 4.2.3.2 **Act** and **Inact<sub>p</sub>** dynamics

After 5 ns of molecular dynamic simulation, a phosphate group was added to the **Inact** state and for the **Act<sub>p</sub>** state, the phosphate group was removed, creating two new molecular systems (see methods); **Act** and **Inact<sub>p</sub>**, in order to observe the changes after adding or removing the perturbation in the systems. The 4 system were simulated as is described in the methods section.

Not further changes were observed when the phosphate was added to the **Inact** state, the principal interactions described above, were not disrupt during the simulation. Even the main event described in the crystal structure,<sup>250</sup> which are the H-bond between Thr83 and the phosphate-moiety of Asp53, is not seen in the dynamic simulation, Figure 4.14. The Phe102 preserves the structural features (gauche+). As in the **Inact** state the  $\beta 4$ - $\alpha 4$  loop in the **Inact<sub>p</sub>** remains in its inactive form, Figure 4.14. Opposite to the **Inact<sub>p</sub>** state, once the phosphate-moiety is removed from the **Act<sub>p</sub>** state, an spontaneous relaxation of the  $\beta 4$ - $\alpha 4$  loop, is observed in both monomers after 1.5 ns of simulation. This relaxation is about  $100^\circ$ , and does not fully goes to the non nonactive state, instead it leads to an intermediate, but stable state, Figure 4.14. This relaxation of the  $\beta 4$ - $\alpha 4$  loop, causes that the Thr83 move apart from the Asp53 approximately 4 Å in each monomer, Figure ??, this motion also promotes steric restrictions with the Phe102, which is in the *trans* conformation, to populated those conformation, close but no fully in the nonactive state, the  $\chi$  angles changes is about  $40^\circ$ , Figure 4.13C and 4.14. Although the Phe102

start to populated the nonactive state, this conformational changes only last a few picoseconds, after return to de active conformation, Figure 4.14.

In order to evaluate if the relaxation of the  $\beta 4$ - $\alpha 4$  loop, is really an spontaneous event, 5 different trajectories where simulated under exactly the same condition as before (see methods), but this time only one of the monomers was used. The simulation were carried out with the **Act<sub>p</sub>** and **Act** during 10 ns. The results are summarized in Figure 4.15. Three of the five trajectories showed spontaneous relaxation after 2 ns, those where the relaxation was observed, showed a full rupture of the H-bond between Thr83 and Asp53, whereas for those where the  $\beta 4$ - $\alpha 4$  loop did not relax, the H-bond is continuously breaking and forming in the time series, Figure 4.15. Phe102 shows the same behavior as it was observed with the dimers in the previous simulations, continuously change its conformations in those trajectories where the relaxation was observed. As we expected, not further changes were observed for the monomers which contain the phospho-moiety, suggesting the great stability that the phosphorylation perturbation gives to the active state.

#### 4.2.3.3 Biased Molecular Dynamics Simulations

Figure 4.16 shows the results for the BMD. Interestingly when the pseudo-dihedral angle  $\phi$  reach an angle of around  $-110^\circ$  (28 ns) , the  $\chi$  angle changes dramatically in both monomers D1/D2. Only this main feature was observed, the remaining coordinates are maintained with not big variations. Figure 4.17 shows the distribution of the  $\chi$  angle in the BMD. Two maximum in the case of D1 monomer (black distribution), are found, corresponding to  $145^\circ$  which is characteristic to the active state, and  $115^\circ$  which is close to the inactive state ( $90^\circ$ ). In the case of monomer D2 (yellow distribution) three maximum peak are found at  $145^\circ$ ,  $110^\circ$  and  $90^\circ$ . The distributions Figure 4.17 and the time series Figure 4.16 shows that once the pseudo-dihedral angle  $\phi$  reach the coordinates characteristic for the inactive state, the  $\chi$  angle gradually start to change to become buried.

**4.2.3.4** Metadynamics

In order to investigate the energetics for the relevant changes of the  $\beta 4\text{-}\alpha 4$  loop and the conformational change in the Phe102, from *gauche +* to *trans*, a metadynamic simulation was conducted using this two representatives coordinates as a collective variables. Figure 4.18 shows the free energy surface in the space of the pseudo-dihedral angle  $\phi$  ( $\beta 4\text{-}\alpha 4$  loop) and the  $\psi$  angle for the conformational change in the Phe102. The collective variable  $\psi$  is the dihedral angle in which the Phe102 can rotate to go from active to nonactive conformer, a value of about  $-60^\circ$  is related with the nonactive state ( $\chi$  angle of  $140^\circ$ ) and a value of about  $-160^\circ$  y related to the active state ( $\chi$  angle of  $90^\circ$ ). The free energy landscape for the **Act<sub>p</sub>** state, describes several minima, the one which is related with the activate state is found at  $110^\circ$  and  $-160^\circ$  for  $\phi$  and  $\psi$  respectively. A barrier energy of about 7 kcal/mol between the **Act<sub>p</sub>** and **Act** was found to describe the minimum energy path that can follow the active state after removal the phosphate-moiety group, Figure 4.18. This results are in agreement with the molecular dynamic simulations, this barrier makes the **Act<sub>p</sub>** state to be stable without sampling the states of the **Act**, because the high free energy barrier. Amazingly the free energy landscape for the **Act** shows that once the phosphate-moiety is removed, the free energy barrier height, shift to a value of 2 kcal/mol, which explains why the faster relaxation in the **Act** state. Also a transformation of the landscape is seen when the free energy surface for the two states (active and nonactive) are compared. Small barrier height is found also (after the removal of the phopho-moiety) to change the conformation from *gauche+* to *trans*, this conformational change is more likely once the relaxation of the  $\beta 4\text{-}\alpha 4$  is reached. figures 4.16-4.18.

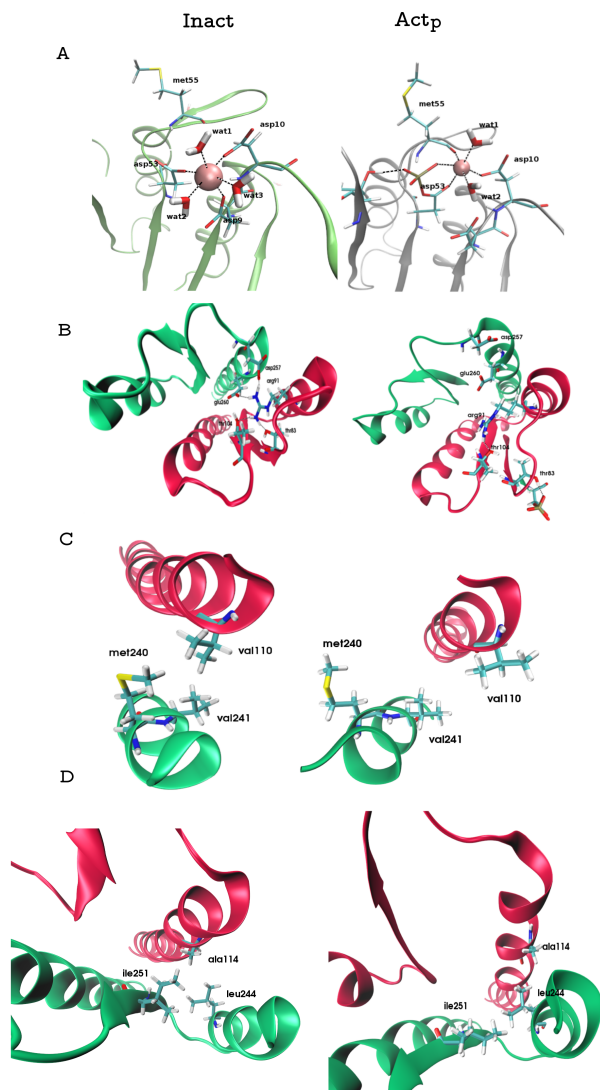
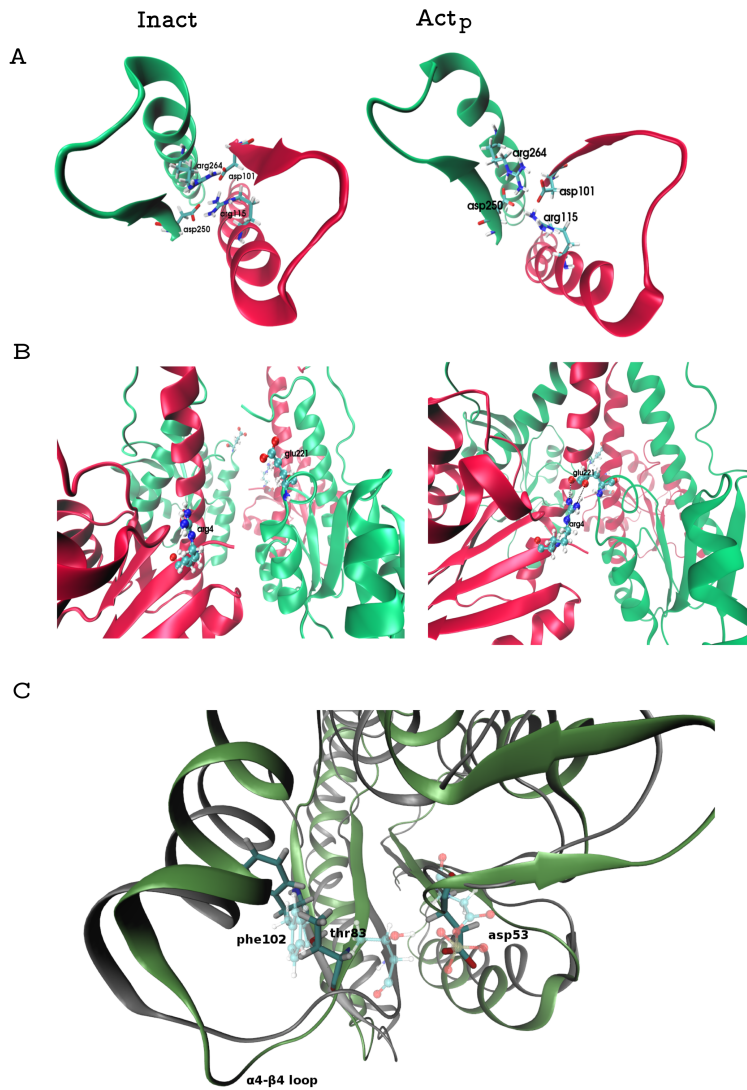
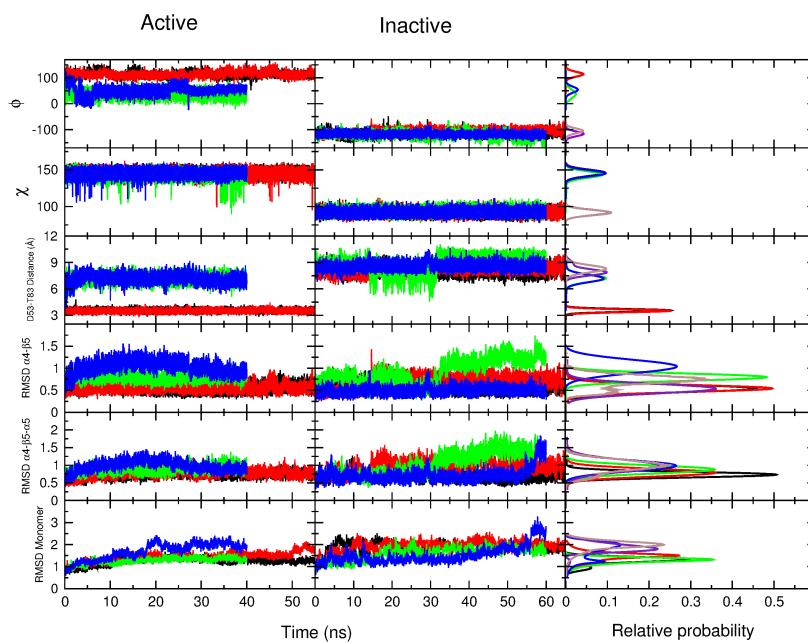


Figure 4.12: Dominant intermolecular Interaction for Act and Inact

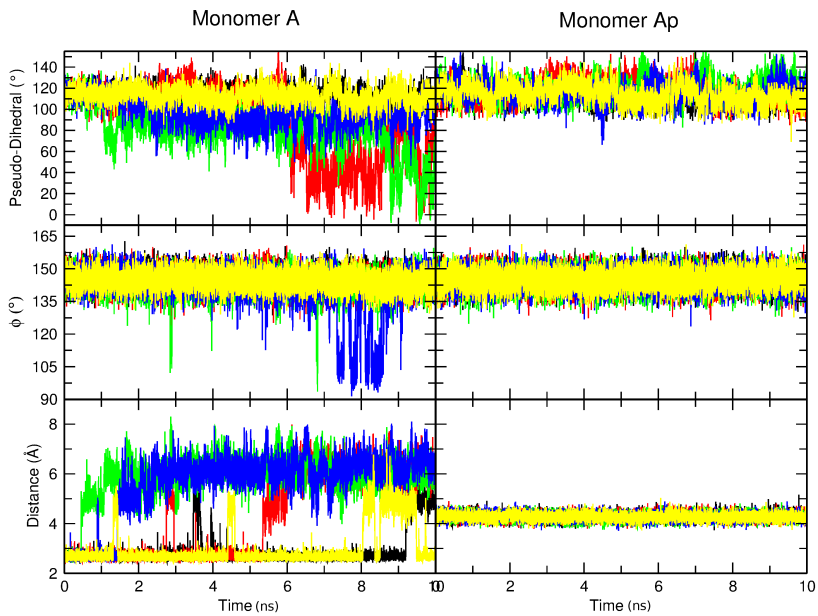


**Figure 4.13:** Dominant intermolecular interaction, **Act** and **Inact**. A. Common salt bridge interactions at interdomain interface. B. Key salt bridge between Arg4 and Glu221 at the interdomain interface. C. Schematic representation of Phe102-Thr83 coupling. Transparent ball and stick representation corresponds to the active protein, while normal ball and stick represents the inactive protein.

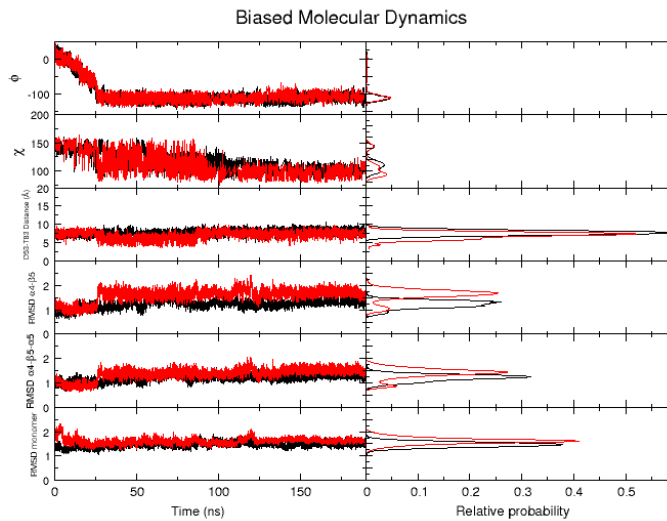


**Figure 4.14:** Molecular dynamic simulation for active, A and inactive B. Black (D1) and green (D2) curves represents the time series for the pseudo-dihedral angle  $\phi$  in the monomer 1 and red and blue for monomer 2. For the rest of the coordinates, black corresponds to monomer 1 and red for monomer 2.

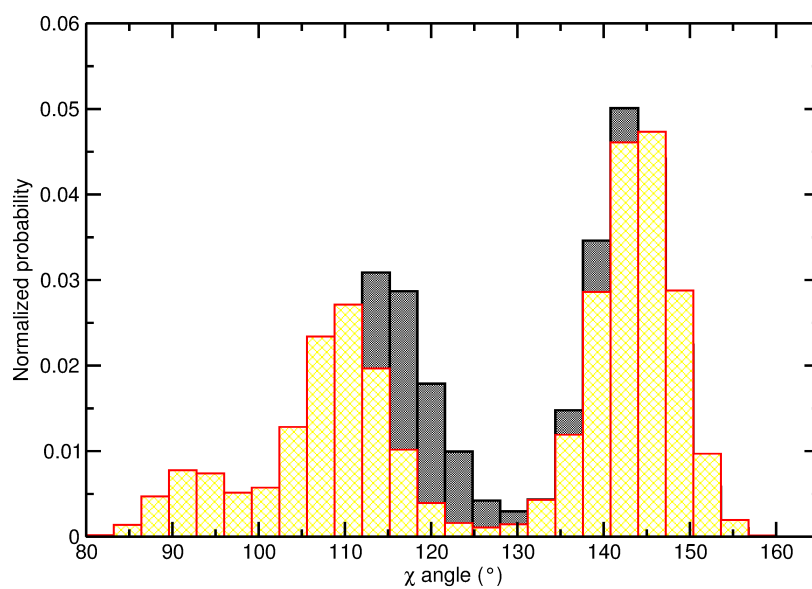




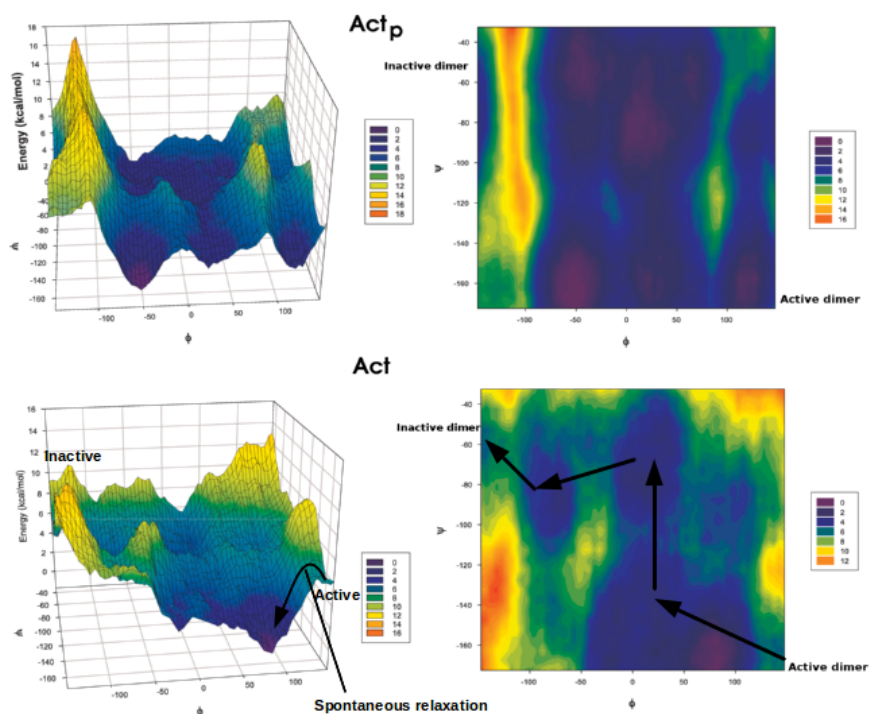
**Figure 4.15:** Molecular dynamic simulation for monomer (time series) 5 different trajectories.



**Figure 4.16:** Biased Molecular dynamic simulation for active unphosphorylated dimer, Black (D1) and red (D2) curves represents the time series for every coordinate.



**Figure 4.17:** Distributions of the  $\chi$  angle for the Biased Molecular Dynamic simulation. Black distribution corresponds to monomer D1 and yellow to monomer D2.



**Figure 4.18:** free energy surface in the space of pseudo-dihedral angle  $\phi$  consisting of  $C\alpha$ -atoms of the  $\beta 4$ - $\alpha 4$ , residues Thr83-Ala84-Leu85-Asp86 D1 domain and dihedral angle  $\psi$  in Phe102, consisting of N,  $C\alpha$ ,  $C\beta$  and  $C\gamma$ . Figure shows the most likely mechanism in which, after removing the phosphate-moiety, the system goes from an active state to inactive one. Energy is given in kcal/mol.



# 5

## GENERAL CONCLUSIONS AND OUTLOOK

---

The amazing potential of experiments and computations working together, bring us the opportunity, nowadays, to understand chemical phenoma at an atomistic level that were not possible in the past Here using atomistic simulations we investigated some reactive processes in exciting molecular systems.

We can now describe the vibrationally induced dissociation of sulfuric acid, which support the hypothesis that in the stratosphere, vibrational excitation of OH-stretching to  $\nu_9 = 4$  and beyond can leads to dissociation into  $\text{SO}_3 + \text{H}_2\text{O}$ , on the ps time scale. We develop high accurate FF for reactant and product and with the novel MS-ARMD method to describe the competitive reaction channels, the intramolecular H-transfer and water elimination. Mechanistically, everything between impulsive water elimination and (sub-picoseconds) to almost complete IVR with subsequent dissociation (picoseconds) are observed. The time scales found from the present simulations agree qualitatively with those from previous independent studies and are, to our knowledge, the first that cover the entire range of dynamics following overtone excitation of  $\text{H}_2\text{SO}_4$ . Such types of simulations hold much promise to provide atomistically detailed insight into reactive processes.

Even though we can explain the most probable mechanism by which, after photoexitation under atmospheric conditions, the sulfuric acid is dissociated, the experimental confirmation of the OH-stretching overtone-induced photodissociation of  $\text{H}_2\text{SO}_4$  has remained elusive. There are inherent difficulties in monitoring such an experiment, as  $\text{H}_2\text{SO}_4$  has an extremely low vapor pressure and exists only in chemical equilibrium with  $\text{H}_2\text{O}$ , in the gas phase.

Analogous reaction with the chlorosulfonic acid was recommended by Kjaergaard,<sup>177</sup> because the detection of product is free from the aforementioned difficulties. The tools we develop and the skills we gain, allows us to investigate the system and compare our results with the experimental findings.

In relation to the [3, 3]-sigmatropic rearrangement of Allyl vinyl ethers, even when it is a preliminary study, many question came up from it. We report that the most preferable conformation of the TS for the Allyl vinyl ether and its derivatives, is a boat conformation. Here we used *ab initio* calculation at MP2 level, parallel with quantum mechanical molecular dynamics. In both approach a boat conformation for TS was found, contradicting results that has been established for more than 20 years. Further investigations using high-level *ab initio* calculations has to be done, to be consistent with this hypothesis that we are proposed here. For final conclusions higher level calculations has to be done for the determinations of the free energies of the two possible TS conformations.

Finally, the work with the PleD protein open a new way to elucidate the exact mechanism of the activation in the response regulator proteins. We found that the coupling between Phe102 and Thr83 is very strong and it is mediated by phosphorylation. This mechanism was primarily proposed to occur in the CheY proten, which is part of the response regulator protein member, with this new findings, one can suggest an evolutionary mechanism conserved over the whole respond regulator family. Interrupting this coupling can provide a basis for the development of generic antibiotics. Further investigations are require on how to affect the coupling to achieve this ambitions goal.

Additional investigations has to be done on other members of the response regulator family to prove that this is a real conserved evolutionary mechanism in prokaryote cells.

# BIBLIOGRAPHY

---

- [1] V Aviyente and N Houk. . *J. Phys. Chem. A*, 105:383, 2001.
- [2] T Ishii, K Moriyoshi, H Sugihara, K Sakurada, H Kadotani, M Yokoi, C Akazawa, R Shigemoto, N Mizuno, and M Masu. . *J. biol. chem.*, 268(4):2836–43, 1993.
- [3] H Meguro, H Mori, K Araki, E Kushiya, T Kutsuwada, M Yamazaki, T Kumanishi, M Arakawa, K Sakimura, and M Mishina. . *Nature*, 357(6373):70–74, 1992.
- [4] Stuart G Cull-Candy and Daniel N Leszkiewicz. . *Science*, 2004(255):re16, 2004.
- [5] J Yosa, M Blanco, O Acevedo, and L R Lareo. . *Eur. J. med. chem.*, 44(7):2960–2966, 2009.
- [6] Collin M Dyer and Frederick W Dahlquist. . *Journal Of Bacteriology*, 188(21):7354–7363, 2006.
- [7] S Y Lee, H S Cho, J G Pelton, D Yan, R K Henderson, D S King, L Huang, S Kustu, E A Berry, and D E Wemmer. . *Nature*, 8(1):52–56, 2001.
- [8] R. D. Levine. , 1997.
- [9] Robert K. Boyd. Macroscopic and microscopic restrictions on chemical kinetics. *Chem. Rev.*, 77(1):93–119, 1977.
- [10] Niels Engholm Henriksen and Flemming Yssing Hansen. *Theories of molecular reaction dynamics*. Oxford Graduate Texts, 2008.
- [11] S Arrhenius. . *Zeitschrift für Physikalische Chemie*, 4:226, 1889.
- [12] Richard Chace Tolman. *The Principles of Statistical Mechanics*. Courier Dover Publications, 1938.
- [13] D G Truhlar and B C Garrett. . *Acc. Chem. Res.*, 13:440, 1980.
- [14] A Fernandez-ramos, J Mille, S J Klippenstein, and D Truhlar. . *Chem. Rev.*, 106:4518, 2006.
- [15] M A Eliason and J O Hirschfelder. . *J. Chem. Phys.*, 83:171, 1959.
- [16] A Kuppermann and E F Greene. . *J. Chem. Educ.*, 45:361, 1968.
- [17] Frank Jensen. *Introduction to Computational Chemistry*, volume 10. Wiley, 2006.
- [18] M J Frisch, G W Trucks, H B Schlegel, G E Scuseria, M A Robb, J R Cheeseman, J A Jr. Montgomery, T Vreven, K N Kudin, J C Burant, J M Millam, S S Iyengar, J Tomasi, V Barone, B Mennucci, M Cossi, G Scalmani, N Rega, G A Petersson, H Nakatsuji, M Hada, M Ehara, K Toyota, R Fukuda, J Hasegawa, M Ishida, T Nakajima, Y Honda, O Kitao, H Nakai, M Klene, X Li, J E Knox, H P Hratchian, J B Cross, C Adamo, J Jaramillo, R Gomperts, R E Stratmann, O Yazyev, A J Austin, R Cammi, C Pomelli, J W Ochterski, P Y Ayala, K Morokuma, G A Voth, P Salvador, J J Dannenberg, V G Zakrzewski, S Dapprich, A D Daniels, M C Strain, O Farkas, D K Malick, A D Rabuck, K Raghavachari, J B Foresman, J V Ortiz, Q Cui, A G Baboul, S Clifford, J Cioslowski, B B Stefanov, G Liu, A Liashenko, P Piskorz, I Komaromi, R L Martin, D J Fox,

- T Keith, M A Al-Laham, C Y Peng, A., M Nanayakkara, P M W Challacombe, B Gill Johnson, W Chen, M W Wong, C Gonzalez, and J A Pople. *Gaussian 03, Revision C.01*. Gaussian, Inc., Wallingford CT, U.S.A., 2004.
- [19] M W Schmidt, K K Baldridge, J A Boatz, S T Elbert, M S Gordon, J J Jensen, S Koseki, N Matsunaga, K A Nguyen, S Su, T L Windus, M Dupuis, and J A Montgomery. General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.*, 14:1347, 1993.
- [20] Art D. Bochevarov, Edward Harder, Thomas F. Hughes, Jeremy R. Greenwood, Dale A. Braden, Dean M. Philipp, David Rinaldo, Mathew D. Halls, Jing Zhang, and Richard A. Friesner. Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *International Journal of Quantum Chemistry*, 113(18):2110–2142, 2013.
- [21] Chunyang Peng and H Bernhard Schlegel. Combining Synchronous Transit and Quasi-Newton Methods to Find Transition States. *Isr. J. Chem.*, 33:449–454, 1993.
- [22] G Henkelman and H Jonsson. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *Journal Of Chemical Physics*, 113(22):9978–9985, 2000.
- [23] G Henkelman, B P Uberuaga, and H Jonsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *Journal Of Chemical Physics*, 113(22):9901–9904, 2000.
- [24] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden. Finite temperature string method for the study of rare events. *The journal of physical chemistry. B*, 109(14):6688–6693, 2005.
- [25] Baron Peters, Andreas Heyden, Alexis T Bell, and Arup Chakraborty. A growing string method for determining transition states: comparison to the nudged elastic band and string methods. *The Journal of chemical physics*, 120(17):7877–7886, 2004.
- [26] G Henkelman and H Jonsson. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *Journal Of Chemical Physics*, 111(15):7010–7022, 1999.
- [27] Eduardo Machado-Charry, Laurent Karim Béland, Damien Caliste, Luigi Genovese, Thierry Deutsch, Normand Mousseau, and Pascal Pochet. Optimized energy landscape exploration using the ab initio based activation-relaxation technique. *The Journal of chemical physics*, 135(3):034102, 2011.
- [28] N Mousseau, P Derreumaux, G T Barkema, and R Malek. Sampling activated mechanisms in proteins with the activation-relaxation technique. *Journal of molecular graphics & modelling*, 19(1):78–86, 2001.
- [29] A Wilkinson, A McNaught, M Nic, J Jirat, and B Kosata. In *IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book")*. Blackwell Scientific Publications, 1997.
- [30] Donald G Truhlar, Bruce C Garrett, and Stephen J Klippenstein. Current Status of Transition-State Theory. *The Journal of Physical Chemistry*, 100(31):12771–12800, 1996.



- [31] Mireia Garcia-Viloca, Jiali Gao, Martin Karplus, and Donald G Truhlar. How Enzymes Work: Analysis by Modern Rate Theory and Computer Simulations. *Science*, 303:186–195, 2004.
- [32] C L Brooks III, M Karplus, and B M Pettitt. *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*. John Wiley & Sons, New York, 1988.
- [33] C. Bruce and Donald G. Truhlar. Generalized transition state theory. Classical mechanical theory and applications to collinear reactions of hydrogen molecules. *J. Phys. Chem.*, 83(8):1059 – 1079, 1979.
- [34] H Eyring and M Polanyi. Concerning simple gas reactions. *Zeitschrift fur Physikalische Chemie-Abteilung B-Chemie der Elementarprozesse Aufbau der Materie*, 12(4):279–311, 1931.
- [35] B Shizgal and M Karplus. . *J. Chem. Phys.*, 52:4262, 1970.
- [36] R. A. Marcus. Unimolecular dissociations and free radical recombination reactions. *J. Chem. Phys.*, 20:359, 1952.
- [37] L. S. Kassel. Studies in Homogeneous Gas Reactions. I. *J. Phys. Chem.*, 32(2):225–242, 1928.
- [38] H. Carstensen and A. Dean. The Kinetics of Pressure-Dependent Reactions” in Comprehensive Chemical Kinetics. *Comprehensive Chemical Kinetics*, 89:503, 2007.
- [39] Oren M. Becker, Alexander D. MacKerell, Benoît Roux, and Masakatsu Watanabe. *Computational Biochemistry and Biophysics*. Marcel Dekker AG, 2001.
- [40] Loup Verlet. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review*, 159(1):98–103, 1967.
- [41] R Hockney. The potential calculation and some applications(Potential calculation from given source distribution, including direct and iterative methods, error analysis. *csacom*, 1970.
- [42] William C Swope, Hans C Andersen, Peter H Berens, and Kent R Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, 76(1):637, 1982.
- [43] C. Kettering, D. Andrews, and L. Shutts. A Representation of the Dynamic Properties of Molecules by Mechanical Models, 1930.
- [44] Paul V. R. Schleyer Edward M. Engler , Joseph D. Andose. Critical evaluation of molecular mechanics. *Journal of the American Chemical Society*, 95(24):8005—8025, 1973.
- [45] F. Lifson and A. Warshel. Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and nAlkane Molecules . *J. Chem. Phys.*, 49:5116, 1968.
- [46] N.L. Allinger. Calculation of Molecular Structure and Energy by Force-Field Methods. *Advances in Physical Organic Chemistry*, 13:1–82, 1976.

- [47] B R Brooks, C L Brooks, A D Mackerell, L Nilsson, R J Petrella, B Roux, Y Won, G Archontis, C Bartels, S Boresch, A Caffisch, L Caves, Q Cui, A R Dinner, M Feig, S Fischer, J Gao, M Hodoscek, W Im, K Kuczera, T Lazaridis, J Ma, V Ovchinnikov, E Paci, R W Pastor, C B Post, J Z Pu, M Schaefer, B Tidor, R M Venable, H L Woodcock, X Wu, W Yang, D M York, and M Karplus. CHARMM: the biomolecular simulation program. *Journal of computational chemistry*, 30(10):1545–1614, 2009.
- [48] Thomas A Halgren. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comp. Chem.*, 17:490–519, 1998.
- [49] J R Maple, U Dinur, and A T Hagler. Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, 85(15):5350–5354, 1988.
- [50] M J J Hwang, T P Stockfisch, and A T Hagler. Derivation of Class-Ii Force-Fields .2. Derivation and Characterization of a Class-Ii Force-Field, Cff93, for the Alkyl Functional-Group and Alkane Molecules. *Journal of the American Chemical Society*, 116(6):2515–2525, 1994.
- [51] G D SMITH, R L JAFFE, and D Y YOON. A FORCE-FIELD FOR SIMULATIONS OF 1,2-DIMETHOXYETHANE AND POLY(OXYETHYLENE) BASED UPON AB-INITIO ELECTRONIC-STRUCTURE CALCULATIONS ON MODEL MOLECULES. *JOURNAL OF PHYSICAL CHEMISTRY*, 97(49):12752–12759, 1993.
- [52] Yi-Ping Liu, Kyungsun Kim, B J Berne, Richard A Friesner, and Steven W Rick. Constructing ab initio force fields for molecular dynamics simulations. *J. Chem. Phys. FIELD Full Journal Title:Journal of Chemical Physics*, 108(12):4739–4755, 1998.
- [53] J.C. Smith, I.B. Costescu, A.C. Vaiana, and Z. Cournia. AFMM: A molecular mechanics force field vibrational parametrization program, 2005.
- [54] J.A. Nelder and R Mead. A simplex method for function minimization. *Chem. Phys.*, 7(4):308 – 313, 1965.
- [55] T. Donald. *Modern Methods for Multidimensional Dynamics Computations in Chemistry*. World Scientific, 1998.
- [56] J. Grotendorst, D. Marx, and A. Muramatsu. *Quantum simulations of complex many-body systems: from theory to algorithms*. NIC Series Volume, Netherlands, 2002.
- [57] J. R. Krenos, R. K. Preston, R. Wolfgang, and J. C. Tully. Molecular beam and trajectory studies of reactions of h+ with h2. *J. Chem. Phys.*, 60(4):1634–1659, 1974.
- [58] John C. Tully. Molecular dynamics with electronic transitions. *J. Chem. Phys.*, 93(2):1061–1071, 1990.
- [59] G Hong, E Rosta, and A Warshel. Using the constrained DFT approach in generating diabatic surf aces and off diagonal empirical valence bond terms for modeling reactions in con densed phases. *J. Phys. Chem. B*, 110:19570–19574, 2006.
- [60] Shina C L Kamerlin and Arieh Warshel. The empirical valence bond model: theory and applications. *Wiley Interdisciplinary Reviews Computational Molecular Science*, 1(1):30–45, 2011.

- [61] Adri Van Duin. ReaxFF reactive force field : A new link from QM to MM. Technical report, 2005.
- [62] D W M Hofmann, L Kuleshova, and B D'Aguanno. A new reactive potential for the molecular dynamics simulation of liquid water. *Chemical Physics Letters*, 448:138–143, 2007.
- [63] D R Nutt, M Meuwly, and M Karplus. . *in preparation*, 2004.
- [64] J Danielsson and M Meuwly. Atomistic simulation of adiabatic reactive processes based on multi-state potential energy surfaces. *J. Chem. Theo. Comp.*, 4:1083–1093, 2008.
- [65] D Nutt and M Meuwly. Studying reactive processes with classical dynamics: Rebinding dynamics in {MbNO}. *Biophys. J.*, 90:1191–1201, 2006.
- [66] J. Yosa and M. Meuwly. Vibrationally induced dissociation of sulfuric acid (h2so4). *J. Phys. Chem. A*, 115(50):14350 – 14360, 2011.
- [67] A C T van Duin, S Dasgupta, F Lorant, and W A Goddard III. ReaxFF: A Reactive Force field for hydrocarbons. *J. Phys. Chem. A*, 105:9396–9409, 2001.
- [68] Kevin D. Nielson, Adri C. T. van Duin, Jonas Oxgaard, Wei-Qiao Deng, and William A. Goddard. Development of the reaxff reactive force field for describing transition metal catalyzed reactions, with application to the initial stages of the catalytic formation of carbon nanotubes. *The Journal of Physical Chemistry A*, 109(3):493–499, 2005.
- [69] M. J. Buehler, A. van Duin, and W. Goddard. Multiparadigm modeling of dynamical crack propagation in silicon using a reactive force field. *Phys. Rev. Lett.*, 96:095505, Mar 2006.
- [70] A. Strachan, E. Kober, A. van Duin, J. Oxgaard, and W. Goddard. Thermal decomposition of rdx from reactive molecular dynamics. *J. Chem. Phys.*, 122(5):54502 – 54515, 2005.
- [71] A. Strachan, A. van Duin, D. Chakraborty, S. Dasgupta, and W. Goddard. Shock waves in high energy materials the initial chemical events in nitramine rdx. *Phys. Rev. Lett.*, 91(9):098301, 2003.
- [72] Markus J. Buehler, Harvey Tang, Adri C. T. van Duin, and William A. Goddard. Threshold crack speed controls dynamical fracture of silicon single crystals. *Phys. Rev. Lett.*, 99:165502, Oct 2007.
- [73] J. Ojwang, R. van Santen, G. Kramer, A. van Duin, and W. Goddard. Modeling the sorption dynamics of nah using a reactive force field. *J. Chem. Phys.*, 128(16):16714, 2008.
- [74] A. Kulkarni, D. Truhlar, S. Goverapet Srinivasan, A. van Duin, P. Norman, and T. Schwartzentruber. Oxygen interactions with silica surfaces: Coupled cluster and density functional investigation and the development of a new reaxff potential. *The Journal of Physical Chemistry C*, 117(1):258–269, 2013.
- [75] W. Wenzel and H. Merlitz. Comparison of stochastic optimization methods for receptor–ligand docking, 2002.

- [76] Y Sugita and Y Okamoto. Replica-exchange molecular dynamics method for protein folding. *Section Title: General Biochemistry*, 314(1,2):141–151, 1999.
- [77] Junichi Higo, Nobuyuki Nakajima, Akinori Kidera, and Haruki Nakamura. Flexible docking of a ligand peptide to a receptor protein by multicanonical molecular dynamics simulation, 1997.
- [78] F Wang and D P Landau. Efficient, multiple-range random walk algorithm to calculate the density of states. *Physical review letters*, 86(10):2050–2053, 2001.
- [79] Simon Trebst, Matthias Troyer, and Ulrich H E Hansmann. Optimized parallel tempering simulations of proteins. *The Journal of chemical physics*, 124(17):174903, 2006.
- [80] R. Elber and M. Karplus. A method for determining reaction paths in large molecules: Application to myoglobin, 1987.
- [81] Christoph Dellago, Peter G. Bolhuis, Felix S. Csajka, and David Chandler. Transition path sampling and the calculation of rate constants. *The Journal of Chemical Physics*, 108(5):1964, 1998.
- [82] Christoph Dellago, P Bolhuis, and PL Geissler. Transition path sampling. *Advances in chemical . . .*, 123(October):1–78, 2002.
- [83] Baron Peters and Bernhardt L Trout. Obtaining reaction coordinates by likelihood maximization. *The Journal of chemical physics*, 125(5):054108, 2006.
- [84] Daniele Moroni, Peter G Bolhuis, and Titus S van Erp. Rate constants for diffusive processes by partial path sampling. *The Journal of chemical physics*, 120(9):4055–4065, 2004.
- [85] Luca Maragliano, Alexander Fischer, Eric Vanden-Eijnden, and Giovanni Ciccotti. String method in collective variables: minimum free energy paths and isocommittor surfaces. *The Journal of chemical physics*, 125(2):24106, 2006.
- [86] Anton K Faradjian and Ron Elber. Computing time scales from reaction coordinates by milestoning. *The Journal of chemical physics*, 120(23):10880–10889, 2004.
- [87] Rosalind J Allen, Patrick B Warren, and Pieter Rein Ten Wolde. Sampling rare switching events in biochemical networks. *Physical review letters*, 94(1):018104, 2005.
- [88] Thorsten Mascher, John D Helmann, and Gottfried Unden. Stimulus Perception in Bacterial Signal-Transducing Histidine Kinases. *Microbiology and Molecular Biology Reviews*, 70(4):910–938, 2006.
- [89] Arjun Narayanan and Matthew P Jacobson. Computational studies of protein regulation by post-translational phosphorylation. *Current Opinion in Structural Biology*, 19(2):156–163, 2009.
- [90] V A Feher and J Cavanagh. Millisecond-timescale motions contribute to the function of the bacterial response regulator protein Spo0F. *Nature*, 400(6741):289–293, 1999.
- [91] Newton A Hecht GB. Identification of a novel response regulator required for the swarmer-to-stalked-cell transition in *Caulobacter crescentus*. *J. Bacteriol.*, 177(21):6223–6229, 1995.

- [92] Justin R Gullingsrud, Rosemary Braun, and Klaus Schulten. Reconstructing Potentials of Mean Force through Time Series Analysis of Steered Molecular Dynamics Simulations. *Journal of Computational Physics*, 151(1):190–211, 1999.
- [93] P A Bash, U C Singh, F K Brown, R Langridge, and P A Kollman. Calculation of the relative change in binding free energy of a protein-inhibitor complex. *Science (New York, N. Y.)*, 235(4788):574–576, 1987.
- [94] Helmut Grubmüller. Predicting slow structural transitions in macromolecular systems: Conformational flooding, 1995.
- [95] Alan Ferrenberg and Robert Swendsen. New Monte Carlo Technique for Studying Phase Transitions, 1989.
- [96] Shankar Kumar, John M. Rosenbergl, Djamal Bouzida, Robert H. Swendsen, and Peter A. Kollman. Multidimensional Free-Energy Calculations Using the Weighted Histogram Analysis Method. *Journal of computational chemistry*, 16(11):1339–1350, 1995.
- [97] C. Jarzynski. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach, 1997.
- [98] Gavin E Crooks. Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems. *Journal of Statistical Physics*, 90(5):1481–1487, 1998.
- [99] Raymond Kapral, James T. Hynes, Giovanni Ciccotti, and E.A. Carter. Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Lett.*, 156(5):472–477, 1989.
- [100] Michiel Sprik and Giovanni Ciccotti. Free energy from constrained molecular dynamics. *Journal of Chemical Physics*, 109(18):7737, 1998.
- [101] Eric Darve and Andrew Pohorille. Calculating free energies using average force. *The Journal of Chemical Physics*, 115(20):9169, 2001.
- [102] David Rodriguez Gomez, Eric Darve, and Andrew Pohorille. Assessing the efficiency of free energy calculation methods. *The Journal of Chemical Physics*, 120(8):3563–3578, 2004.
- [103] G. N. Patey and J. P. Valleau. A Monte Carlo method for obtaining the interionic potential of mean force in ionic solution. *The Journal of Chemical Physics*, 63(6):2334, 1975.
- [104] Lula Rosso, Peter Minary, Zhongwei Zhu, and Mark E. Tuckerman. On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles. *The Journal of Chemical Physics*, 116(11):4389, 2002.
- [105] R Fletcher and M J D Powell. A rapidly convergent descent method for minimization. *The Computer Journal*, 6(2):163–168, 1963.
- [106] Arthur Voter. Hyperdynamics: Accelerated Molecular Dynamics of Infrequent Events, 1997.

- [107] Radu A Miron and Kristen A Fichthorn. Multiple-time scale accelerated molecular dynamics: addressing the small-barrier problem. *Physical review letters*, 93(12):128301, 2004.
- [108] Alessandro Laio and Michele Parrinello. Escaping free-energy minima. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):12562–12566, 2002.
- [109] Jan Hermans, R.H. Yun, and Amil G. Anderson. Precision of free energies calculated by molecular dynamics simulations of peptides in solution. *J Comput Chem*, 13(4):429–442, 1992.
- [110] C Haydock, J C Sharp, and F G Prendergast. Tryptophan-47 rotational isomerization in variant-3 scorpion neurotoxin. A combination thermodynamic perturbation and umbrella sampling study. *Biophysical journal*, 57(6):1269–1279, 1990.
- [111] Ioannis G. Kevrekidis, C. William Gear, and Gerhard Hummer. Equation-free: The computer-aided analysis of complex multiscale systems. *AIChE Journal*, 50(7):1346–1355, 2004.
- [112] Simone Marsili, Alessandro Barducci, Riccardo Chelli, Piero Procacci, and Vincenzo Schettino. Self-healing umbrella sampling: a non-equilibrium approach for quantitative free energy calculations. *The journal of physical chemistry. B*, 110(29):14011–14013, 2006.
- [113] Alessandro Laio and Francesco L Gervasio. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science, 2008.
- [114] James C Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D Skeel, Laxmikant Kalé, and Klaus Schulten. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–802, 2005.
- [115] A Warshel and M Levitt. Theoretical studies of enzymic reactions: {D}ielecric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.*, 103:227–249, 1976.
- [116] M J Field, P A Bash, and M Karplus. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.*, 11:700–733, 1990.
- [117] J K Hwang and A Warshel. How important are quantum mechanical nuclear motions in enzyme catalysis? *J. Am. Chem. Soc.*, 118(47):11745–11751, November 1996.
- [118] K Hinsen and B Roux. Potential of mean force and reaction rates for proton transfer in acetylacetone. *J. Chem. Phys.*, 106(9):3567–3577, March 1997.
- [119] M Pavese, S Chawla, D Lu, J Lobaugh, and G A Voth. Quantum effects and the excess proton in water. *J. Chem. Phys.*, 107:7428, 1997.
- [120] William H Miller, Yi Zhao, Michele Ceotto, and Sandy Yang. Quantum instanton approximation for thermal rate constants of chemical reactions. *J. Chem. Phys.*, 119:1329–1342, 2003.

- [121] Kim F Wong, Jason L Sonnenberg, Francesco Paesani, Takeshi Yamamoto, Jiff Vanicek, Wei Zhang, H Bernhard Schlegel, David A Case, Thomas E Cheatham III, William H Miller, and Gregory A Voth. Proton Transfer Studied Using a Combined Ab Initio Reactive Potential Energy Surface with Quantum Path Integral Methodology. *J. Chem. Theo. Comp.*, 6(9):2566–2580, 2010.
- [122] J Huang, M Buchowiecki, T Nagy, J Vanicek, and M Meuwly. Kinetic isotope effect in malonaldehyde from path integral Monte Carlo simulations. *PCCP*, in print(in print):in print, 2013.
- [123] R Car and M Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett. FIELD Full Journal Title:Physical Review Letters*, 55(22):2471–2474, 1985.
- [124] Herbert, JM and Head-Gordon, M. Accelerated, energy-conserving Born-Oppenheimer molecular dynamics via Fock matrix extrapolation. *PCCP*, 7(18):3269–3275, 2005.
- [125] Anders M N Niklasson, C J Tymczak, and Matt Challacombe. Time-reversible Born-Oppenheimer molecular dynamics. *Phys. Rev. Lett.*, 97(12), September 2006.
- [126] A Warshel and M Karplus. . *Chem. Phys. Lett.*, 32:11, 1975.
- [127] M J Field, P A Bash, and M Karplus. . *J. Comp. Chem.*, 11:700, 1990.
- [128] P A Bash, M J Field, and M Karplus. Free energy perturbation method for chemical reactions in the condensed phase: a dynamic approach based on a combined quantum and molecular mechanics potential. *J. Am. Chem. Soc.*, 109:8092–8094, 1987.
- [129] A Warshel and R M Weiss. An empirical valence bond approach for comparing reactions ins olutions and in enzymes. *J. Am. Chem. Soc.*, 102(20):6218–6226, 1980.
- [130] S Mishra and M Meuwly. Atomistic Simulation of NO Dioxygenation in Group<sup>VI</sup> Truncated Hemoglobin. *J. Am. Chem. Soc.*, 132:2968, 2010.
- [131] C. A. Rivera, N. Winter, R. V. Harper, I. Benjamin, and S. E. Bradforth. The dynamical role of solvent on the ICN photodissociation reaction: connecting experimental observables directly with molecular dynamics simulations. *Phys Chem Chem Phys.*, 13(18):8269 – 8283, 2011.
- [132] J Y Reyes and M Meuwly. Vibrationally Induced Dissociation of Sulfuric Acid (H<sub>2</sub>SO<sub>4</sub>). *J. Phys. Chem. B*, 115:14350, 2011.
- [133] Karim Farah, Florian Mueller-Plathe, and Michael C Boehm. Classical Reactive Molecular Dynamics Implementations: State of the Art. *Chem. Phys. Chem.*, 13(5):1127–1151, 2012.
- [134] S Braun-Sand, M Strajbl, and A Warshel. Studies of proton translocations in biological systems: Simulating proton transport in carbonic anhydrase by EVB-based models. *Biophys. J.*, 87:2221–2239, 2004.
- [135] U W Schmitt and G A Voth. Multistate empirical valence bond model for proton transport in water. *J. Phys. Chem. B*, 102(29):5547–5551, 1998.
- [136] G Y Hong, E Rosta, and A Warshel. Using the constrained DFT approach in generating diabatic surfaces and off diagonal empirical valence bond terms for modeling reactions in condensed phases. *J. Phys. Chem. B*, 110:19570, 2006.

- [137] Prakash Dayal, Sabine A Weyand, Joanne McNeish, and Nicholas J Mosey. Temporal quantum mechanics/molecular mechanics: Extending the time scales accessible in molecular dynamics simulations of reactions. *Chem. Phys. Lett.*, 516(4-6):263–267, November 2011.
- [138] K Nienhaus, S Lutz, M Meuwly, and G U Nienhaus. Reaction-Pathway Selection in the Structural Dynamics of a Heme Protein. *Chem. Eur. J.*, page in print, 2013.
- [139] Y. Miller and R. B. Gerber. Dynamics of vibrational overtone excitations of  $\text{h}_2\text{so}_4$ ,  $\text{h}_2\text{so}_4\text{-h}_2\text{o}$ : Hydrogen-hopping and photodissociation processes. *J. Am. Chem. Soc.*, 128(30):9594 – 9595, 2006.
- [140] Xin Tong, Tibor Nagy, Juvenal Yosa Reyes, Matthias Germann, Markus Meuwly, and Stefan Willitsch. State-selected ion-molecule reactions with Coulomb-crystallized molecular ions in traps. *Chem. Phys. Lett.*, 547:1–8, September 2012.
- [141] C F Clement and I J Ford. Gas-to-particle conversion in the atmosphere: I. Evidence from empirical atmospheric aerosols. *Atmos Environ A-Gen*, 33:475, 1999.
- [142] L. J. Larson, M. Kuno, and F. M. Tao. Hydrolysis of sulfur trioxide to form sulfuric acid in small water clusters. *J. Chem. Phys.*, 112(20):8830 – 8838, 2000.
- [143] D.J. Donaldson, G. J. Frost, K. H. Rosenlof, and V. Tuck, A. F. Vaida. Atmospheric radical production by excitation of vibrational overtones via absorption of visible light. *Geophys. Res. Lett.*, 24(21):2651 – 2654, 1997.
- [144] V. Vaida, H. G. Kjaergaard, P. E. Hintze, and D. J. Donaldson. Photolysis of sulfuric acid vapor by visible solar radiation. *Science*, 299(5612):1566 – 1568, 2003.
- [145] Gustav Mie. Zur kinetischen theorie der einatomigen korper. *Ann. Phys.*, 316(8):657–697, 1903.
- [146] Christian Kramer, Peter Gedeck, and Markus Meuwly. Multipole-based Force Fields from  $\text{ab initio}$  Interaction Energies and the Need for Jointly Refitting all Intermolecular Parameters. *J. Chem. Theo. Comp.*, in print, 2013.
- [147] V Vaida, H G Kjaergaard, P E Hintze, and D J Donaldson. . *Science*, 299:1566, 2003.
- [148] QQ and M Meuwly. in preparation. in preparation.
- [149] C. Breneman and K. Wiberg. Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis. *J. Chem. Comp.*, 11(3):361–373, 1990.
- [150] T.A. Halgren and W.N. Lipscomb. The synchronous-transit method for determining reaction pathways and locating molecular transition states. *Chem. Phys. Lett.*, 49(2):225 – 232, 1977.
- [151] K. Morokuma and C. Muguruma.  $\text{ab initio}$  molecular-orbital study of the mechanism of the gas-phase reaction  $\text{so}_3+\text{h}_2\text{o}$  - importance of the 2nd water molecule. *J. Am. Chem. Soc.*, 116(22):10316 – 10317, 1994.
- [152] P. N Nguyen and G. Stock. Nonequilibrium molecular-dynamics study of the vibrational energy relaxation of peptides in water. *J. Chem. Phys.*, 119(21):11350–11358, 2003.



- [153] Y.-T. Chang and W H Miller. An empirical valence bond model for constructing global potential energy surfaces for chemical reactions of polyatomic molecular systems. *J. Phys. Chem.*, 94:5884–5888, 1990.
- [154] H Bernhard Schlegel and Jason L Sonnenberg. Empirical Valence-Bond Models for Reactive Potential Energy Surfaces Using Distributed Gaussians. *J. Chem. Theo. Comp.*, 2(4):905–911, 2006.
- [155] R. Charlson, J. Lovelock, M. Andreae, and S. Warren. *Nature*, 326(6114):655 – 661, 1987.
- [156] R. Kiene. *Nature*, 402(6760):363 – 368, 1999.
- [157] R. Simo and C. Pedros-Alio. *Nature*, 402(6760):396 – 399, 1999.
- [158] B. Huebert. *Nature*, 400(6746):713 – 714, 1999.
- [159] K. Capaldo, J. Corbett, P. Kasibhatla, P. Fischbeck, and S. Pandis. *Nature*, 400(6746):743 – 746, 1999.
- [160] H. Niki, P. Maker, C. Savage, and L. P Breitenbach. *Int. J. Chem. Kinet.*, 15(6746):743 – 746, 1986.
- [161] D. Grosjean. Photooxidation of methyl sulfide, ethyl sulfide, and methanethiol. *Environ. Sci.*, 18(188):460 – 468, 1984.
- [162] R. Steudel. *Angew. Chem. Int. Ed.*, 34(178):1313 – 1315, 1995.
- [163] W-K. Li and M. McKee. Theoretical study of oh and h<sub>2</sub>o addition to so<sub>2</sub>. *J. Phys. Chem. A*, 101(50):9778 – 9782, 1995.
- [164] A. Kukui, V. Bossoutrot, G. Laverdet, and G. Le Bras. Theoretical study of oh and h<sub>2</sub>o addition to so<sub>2</sub>. *J. Phys. Chem. A*, 104(5):9935 – 946, 2000.
- [165] D. J. Hofmann and J. M. Rosen. Stratospheric condensation nuclei variations may relate to solar-activity. *Nature*, 297(5862):120–124, 1982.
- [166] Frank Raes, Rita Van Dingenen, Elisabetta Vignati, Julian Wilson, Jean-Philippe Pataud, John H. Seinfeld, and Peter Adams. Formation and cycling of aerosols in the global troposphere. *Atmos. Environ. A-Gen.*, 34(25):4215 – 4240, 2000.
- [167] J. Rosen and D. Hofmann. Unusual behavior in the condensation nuclei concentration at 30 km. *J. Geophys. Res.*, 88(6):3725 – 3731, 1983.
- [168] J. Burkholder, M. Mills, and S. McKeen. Upper limit for the uv absorption cross sections of h<sub>2</sub>so<sub>4</sub>. *Geophys. Res. Lett.*, 27(16):2493 – 2496, 2000.
- [169] P. E. Hintze, H. G. Kjaergaard, V. Vaida, and J. B. Burkholder. Vibrational and electronic spectroscopy of sulfuric acid vapor. *J. Chem. Phys.*, 107(8):1112 – 1118, 2003.
- [170] S. J. Wrenn, L. J. Butler, G. A. Rowland, C. J. Knox, and L. F. Phillips. The necessity for multiphoton processes in the 193-nm photochemistry of sulphuric acid aerosols. *J. Chem. Phys.*, 129(3):1112 – 1118, 1999.

- [171] H. G. Kjaergaard, J. R. Lane, A. L. Garden, D. P. Schofield, T. W. Robinson, and M. Mills. The necessity for multiphoton processes in the 193-nm photochemistry of sulphuric acid aerosols. *Adv. Quantum Chem.*, 55(137):1112 – 1118, 2008.
- [172] T. W. Robinson, D. P. Schofield, and H. G. Kjaergaard. High level ab initio studies of the excited states of sulfuric acid and sulfur trioxide. *J. Chem. Phys.*, 118(7226):7226 – 7232, 2003.
- [173] J. R. Lane and H. G. Kjaergaard. Calculated electronic transitions in sulfuric acid and implications for its photodissociation in the atmosphere. *J. Phys. Chem. A*, 112(22):4958 – 4964, 2008.
- [174] J. Langmer, H. Rodhe, P. J. Crutzen, and P. Zimmermann. Anthropogenic influence on the distribution of tropospheric sulphate aerosol. *Nature*, 359(6397):712 – 716, 1992.
- [175] V. Vaida and D. J. Donaldson. Red-light initiated atmospheric reactions of vibrationally excited molecules. *Phys Chem Chem Phys.*, 16:827–836, 2014.
- [176] D. K. Havey, K. J. Feierabend, and V. Vaida. Ab initio study of h<sub>2</sub>so<sub>4</sub> rotamers. *J. Mol. Struct-Teochem*, 680(5):243 – 247, 2004.
- [177] J. R. Lane, H. G. Kjaergaard, K. L. Plath, and V. Vaida. Overtone spectroscopy of sulfonic acid derivatives. *J. Phys. Chem. A*, 111(25):5434 – 5440, 2007.
- [178] P. E. Hintze, K. J. Feierabend, D. K. Havey, and V. Vaida. High-resolution spectroscopy of h<sub>2</sub>so<sub>4</sub>, hdso<sub>4</sub>, and d<sub>2</sub>so<sub>4</sub> vapor in the region 1200-10,000 cm<sup>(-1)</sup>. *Spec. Acta A-Mol. Biomol. Spec.*, 61(4):559 – 566, 2005.
- [179] K. J. Feierabend, D. K. Havey, S. S. Brown, and V. Vaida. Experimental absolute intensities of the 4v(9) and 5v(9) o-h stretching overtones of h<sub>2</sub>so<sub>4</sub>. *Chem. Phys. Lett.*, 420(6):438 – 442, 2006.
- [180] R. D. Beck, P. Maroni, D. C. Papageorgopoulos, T. T. Dang, M. P. Schmid, and T. R. Rizzo. Vibrational mode-specific reaction of methane on a nickel surface. *Science*, 302(5642):98–100, 2003.
- [181] Y. Miller, R. B. Gerber, and V. Vaida. Photodissociation yield for vibrationally excited states of sulfuric acid under atmospheric conditions. *Geophys. Res. Lett.*, 34(16):1 – 5, 2007.
- [182] DA Blank, WZ Sun, AG Suits, YT Lee, SW North, and GE Hall. Primary and secondary processes in the 193 nm photodissociation of vinyl chloride. *J. Chem. Phys.*, 108(13):5414–5425, 1998.
- [183] F. Fleming Crim. Chemical dynamics of vibrationally excited molecules: Controlling reactions in gases and on surfaces. *Proc. Nat. Acad. Sci. USA*, 105(35):12654–12661, 2008.
- [184] Michael P. Grubb, Michelle L. Warter, Kurt M. Johnson, and Simon W. North. Ion Imaging Study of NO<sub>3</sub> Radical Photodissociation Dynamics: Characterization of Multiple Reaction Pathways. *J. Phys. Chem. A*, 115(15):3218–3226, 2011.
- [185] T. Nagy, J. Yosa, and M. Meuwly. Multi-state adiabatic reactive molecular dynamics. *J. Chem. Theory. Comput.*, page accepted, 2014.

- [186] D. R. Nutt and M. Meuwly. Studying reactive processes with classical dynamics: Rebinding dynamics in mbno. *Biophys. J.*, 90:1191–1201, 2006.
- [187] J. Danielsson and M. Meuwly. Atomistic simulation of adiabatic reactive processes based on multi-state potential energy surfaces. *J. Chem. Theory. Comput.*, 4:1083, 2008.
- [188] AD MacKerell, D Bashford, M Bellott, RL Dunbrack, JD Evanseck, MJ Field, S Fischer, J Gao, H Guo, S Ha, D Joseph-McCarthy, L Kuchnir, K Kuczera, FTK Lau, C Mattos, S Michnick, T Ngo, DT Nguyen, B Prodhom, WE Reiher, B Roux, M Schlenkrich, JC Smith, R Stote, J Straub, M Watanabe, J Wiorcikewicz-Kuczera, D Yin, and M Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102(18):3586–3616, 1998.
- [189] B. R. Brooks, C. L. Brooks, III, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caffisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. CHARMM: The Biomolecular Simulation Program. *J. Chem. Comp.*, 30(10, SI):1545–1614, 2009.
- [190] M. Meuwly, A. Müller, and S. Leutwyler. Energetics, dynamics and infrared spectra of the dna base-pair analogue 2-pyridone 2-hydroxypyridine. *Phys Chem Chem Phys.*, 5:2663–2672, 2003.
- [191] J. Demaison, M. Herman, J. Lievin, and H. D. Rudolph. Equilibrium structure of sulfuric acid. *J. Phys. Chem. A*, 111(13):2602 – 2609, 2007.
- [192] L.L. Lohr. An abinitio study of the structure and torsional modes of the h2so4 molecule. *J. Mol. Struct-Teochem*, 4(3):221 – 227, 1982.
- [193] Yan Zhao, Benjamin J. Lynch, and Donald G. Truhlar. Doubly Hybrid Meta DFT: New Multi-Coefficient Correlation and Density Functional Methods for Thermochemistry and Thermochemical Kinetics. *J. Phys. Chem. A*, 108(21):4786–4791, 2004.
- [194] Anthony P. Scott and Leo Radom. *The Journal of Physical Chemistry*, 100(41):16502–16513, 1996.
- [195] T. L. Tso and E. Lee. Formation of sulfuric acid and sulfur trioxide/water complex from photooxidation of hydrogen sulfide in solid oxygen at 15 k. *J. Phys. Chem.*, 88(13):2776–2781, 1984.
- [196] L. Schriver, D. Carrere, A. Schriver, and K. Jaeger. Matrix-isolation photolysis of so2, o3 and h2o : evidence for the h2o:so3 complex. *Chem. Phys. Lett.*, 181(6):505–511, 1991.
- [197] A. Givan, A Larsen, A. Loewenschuss, and C. Nielsen. Matrix isolation mid- and far-infrared spectra of sulfuric acid and deuterated sulfuric acid vapors. *J. Mol. Str.*, 509(1–3):35–47, 1999.
- [198] A Ma and AR Dinner. Automatic method for identifying reaction coordinates in complex systems. *J. Phys. Chem. B*, 109(14):6769–6779, 2005.

- [199] R Schinke. Rotational Distributions in Direct Molecular Photodissociation. *Annu. Rev. Phys. Chem.*, 39:39–68, 1988.
- [200] G. Herzberg. *Infrared and Raman Spectra*. Van Nostrand, Princeton, 1962.
- [201] V. Meyer, D. H. Sutter, and H. Dreizler. The centrifugally induced pure rotational spectrum and the structure of sulfur trioxide. A microwave Fourier transform study of a nonpolar molecule. *Z. Naturforsch. A.*, 46(8):710–714, 1991.
- [202] G. P. Brasseur and S. Solomon. *Aeronomy of the Middle Atmosphere*. Springer, P.O.Box 17,3300 AA Dordrecht, The Netherlands., 2005.
- [203] Stephan Lutz and Markus Meuwly. Photodissociation Dynamics of ClCN at Different Wavelengths. *ChemPhysChem*, 13(1):305–313, 2012.
- [204] S. A. Barts and J. B. Halpern. Photodissociation of ClCN between 190-nm and 213-nm. *J. Phys. Chem.*, 93(21):7346 – 7351, 1989.
- [205] Andrew C. Crowther, Stacey L. Carrier, Thomas J. Preston, and F. Fleming Crim. Time-Resolved Studies of the Reactions of CN Radical Complexes with Alkanes, Alcohols, and Chloroalkanes. *J. Phys. Chem. A*, 113(16):3758–3764, 2009.
- [206] Myung Won Lee and Markus Meuwly. On the role of nonbonded interactions in vibrational energy relaxation of cyanide in water. *J. Phys. Chem. A*, 115(20):5053–5061, 2011.
- [207] Walter Kimel and Arthur C Cope. The Rearrangement of Allyl-Type Esters of  $\beta$ -Keto Acids. *Journal of the American Chemical Society*, 65(10):1992–1998, 1943.
- [208] Stephen R Wilson and Martyn F Price. The ester enolate Carroll rearrangement. *The Journal of Organic Chemistry*, 49(4):722–725, 1984.
- [209] A E Wick, Dorothee Felix, Katharina Steen, and A Eschenmoser. CLAISEN'sche Umlagerungen bei Allyl- und Benzylalkoholen mit Hilfe von Acetalen des N, N-Dimethylacetamids. Vorläufige Mitteilung. *Helvetica Chimica Acta*, 47(8):2425–2429, 1964.
- [210] Dorothee Felix, Katharina Gschwend-Steen, A E Wick, and A Eschenmoser. CLAISEN'sche Umlagerungen bei Allyl- und Benzylalkoholen mit 1-Dimethylamino-1-methoxy-äthen. *Helvetica Chimica Acta*, 52(4):1030–1042, 1969.
- [211] Ana M Martin. Claisen Rearrangement over the Past Nine Decades. *Chemical Reviews*, 104(6):2939–3002, 2004.
- [212] M G Evans. The activation energies of reactions involving conjugated systems. *Trans. Faraday Soc.*, 35:2939–3002, 1939.
- [213] Kenichi Fukui. Recognition of stereochemical paths by orbital interaction. *Acc. Chem. Res.*, 4:57–64, 1971.
- [214] Robert L Vance, Nelson G Rondan, K N Houk, Frank. Jensen, Weston Thatcher. Borden, Andrew. Komornicki, and Erich. Wimmer. Transition structures for the Claisen rearrangement. *Journal of the American Chemical Society*, 110(7):2314–2315, 1988.

- [215] M Khaledy, S Kalani, S Khuong, N Houk, V Aviyente, N Soldermann, R Neier, and J Velker. . *J. Org. Chem.*, 68:572, 2003.
- [216] S Dewar and C Jie. . *J. Am. Chem. Soc.*, 111:511, 1989.
- [217] P Meyer, J DelMonte, and A Singleton. . *J. Am. Chem. Soc.*, 121:10865, 1999.
- [218] D McMichael and L Korver. . *J. Am. Chem. Soc.*, 101:2746, 1979.
- [219] J Gajewski and D Conrad. . *J. Am. Chem. Soc.*, 101:2747, 1979.
- [220] W Schuler and W Murphy. . *J. Am. Chem. Soc.*, 72:3155, 1950.
- [221] Izumi Iwakura, Yu Kaneko, Shigehiko Hayashi, Atsushi Yabushita, and Takayoshi Kobayashi. The reaction mechanism of Claisen rearrangement obtained by transition state spectroscopy and single direct-dynamics trajectory. *Molecules (Basel, Switzerland)*, 18(2):1995–2004, 2013.
- [222] J Burrows and K Carpenter. . *J. Am. Chem. Soc.*, 103:6983, 1981.
- [223] J Gajewski, R Gee, and J Jurayj. . *J. Org. Chem.*, 55:1813, 1990.
- [224] J Cooper, C Olivares, and G Sandford. . *J. Org. Chem.*, 66:4887, 2001.
- [225] J Gajewski and J Emrani. . *J. Am. Chem. Soc.*, 106:5733, 1984.
- [226] P Curran and Y Suh. . *J. Am. Chem. Soc.*, 106:5002, 1984.
- [227] M R Coates, B Rogers, J Hobbs, R Peck, and P Curran. . *J. Am. Chem. Soc.*, 109:1160, 1987.
- [228] A Grieco, B Brandes, McCann S., and D Clark. . *J. Org. Chem.*, 54:5849, 1989.
- [229] N Hu, M Kobrak, C Xu, and H Hammes-Schiffer. . *J. Phys. Chem. A*, 104:8058, 2000.
- [230] F Kincaid and S Tarbell. . *J. Am. Chem. Soc.*, 61:3085, 1939.
- [231] N White, D Gwynn, R Schlitt, C Girard, and W Fife. . *J. Am. Chem. Soc.*, 80:3271, 1958.
- [232] Andrew R Knaggs. The biosynthesis of shikimate metabolites. *Natural product reports*, 20(1):119–136, 2003.
- [233] Helmut Goerisch. On the mechanism of the chorismate mutase reaction. *Biochemistry*, 17(18):3700–3705, 1978.
- [234] M. S. Plesset and Chr. Moller. Note on an Approximation Treatment for Many Electron Systems, 1934.
- [235] Jp Perdew, K Burke, and M Ernzerhof. Generalized Gradient Approximation Made Simple. *Physical review letters*, 77(18):3865–3868, 1996.
- [236] Massimiliano Bonomi, Davide Branduardi, Giovanni Bussi, Carlo Camilloni, Davide Provasi, Paolo Raiteri, Davide Donadio, Fabrizio Marinelli, Fabio Pietrucci, Ricardo A Broglia, and Michele Parrinello. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications*, 180(10):1961–1972, 2009.

- [237] Daniel L Severance and William L Jorgensen. Effects of hydration on the Claisen rearrangement of allyl vinyl ether from computer simulations. *Journal of the American Chemical Society*, 114(27):10966–10968, 1992.
- [238] R a Marcus. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. II. Applications to Data on the Rates of Isotopic Exchange Reactions. *The Journal of Chemical Physics*, 26(4):867, 1957.
- [239] Morgan Y Chen and Joseph R Murdoch. The application of Marcus-like equations to processes which have no corresponding identity reactions. Separation of thermodynamic and intrinsic contributions to barriers to internal rotation and conformational rearrangements. *Journal of the American Chemical Society*, 106(17):4735–4743, 1984.
- [240] Yi Qin Gao. An integrate-over-temperature approach for enhanced sampling. *The Journal of chemical physics*, 128(6):064105, 2008.
- [241] I Hwang and J Sheen. Two-component circuitry in Arabidopsis cytokinin signal transduction. *Nature*, 413(6854):383–389, 2001.
- [242] G J Burton, G B Hecht, and A Newton. Roles of the histidine protein kinase pleC in *Caulobacter crescentus* motility and chemotaxis. *Journal Of Bacteriology*, 179(18):5849–5853, 1997.
- [243] Ralf Paul, Stefan Weiser, Nicholas C Amiot, Carmen Chan, Tilman Schirmer, Bernd Giese, and Urs Jenal. Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes & Development*, 18(6):715–727, 2004.
- [244] Phillip Aldridge, Ralf Paul, Patrick Goymer, Paul Rainey, and Urs Jenal. Role of the GGDEF regulator PleD in polar development of *Caulobacter crescentus*. *Molecular Microbiology*, 47(6):1695–1708, 2003.
- [245] P Aldridge and U Jenal. Cell cycle-dependent degradation of a flagellar motor component requires a novel-type response regulator. *Molecular Microbiology*, 32(2):379–391, 1999.
- [246] Urs Jenal. Cyclic di-guanosine-monophosphate comes of age: a novel secondary messenger involved in modulating cell surface structures in bacteria? *Current Opinion in Microbiology*, 7(2):185–191, 2004.
- [247] Urs Jenal and Jacob Malone. Mechanisms of cyclic-di-GMP signaling in bacteria. *Annual Review of Genetics*, 40(July):385–407, 2006.
- [248] Rita Tamayo, Jason T Pratt, and Andrew Camilli. Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. *Annual Review of Microbiology*, 61(26):131–48, 2007.
- [249] Carmen Chan, Ralf Paul, Dietrich Samoray, Nicolas C Amiot, Bernd Giese, Urs Jenal, and Tilman Schirmer. Structural basis of activity and allosteric control of diguanylate cyclase. *Proceedings of the National Academy of Sciences of the United States of America*, 103(49):9470–9475, 2006.

- [250] Paul Wassmann, Carmen Chan, Ralf Paul, Andreas Beck, Heiko Heerklotz, Urs Jenal, and Tilman Schirmer. Structure of BeF<sub>3</sub>-modified response regulator PleD: implications for diguanylate cyclase activation, catalysis, and feedback inhibition. *Structure London England* 1993, 15(8):915–927, 2007.
- [251] Ralf Paul, Sören Abel, Paul Wassmann, Andreas Beck, Heiko Heerklotz, and Urs Jenal. Activation of the diguanylate cyclase PleD by phosphorylation-mediated dimerization. *The Journal of Biological Chemistry*, 282(40):29170–7, 2007.
- [252] Beat Christen, Matthias Christen, Ralf Paul, Franziska Schmid, Marc Folcher, Paul Jenoe, Markus Meuwly, and Urs Jenal. Allosteric control of cyclic di-GMP signaling. *The Journal of Biological Chemistry*, 281(42):32015–32024, 2006.
- [253] J Monod, J Wyman, and J P Changeux. ON THE NATURE OF ALLOSTERIC TRANSITIONS: A PLAUSIBLE MODEL. *Journal of Molecular Biology*, 12(December):88–118, 1965.
- [254] Matthew J Whitley and Andrew L Lee. Frameworks for understanding long-range intra-protein communication. *Current protein peptide science*, 10(2):116–127, 2009.
- [255] K Gunasekaran, Buyong Ma, and Ruth Nussinov. Is allostery an intrinsic property of all dynamic proteins? *Proteins*, 57(3):433–443, 2004.
- [256] Liang Ma and Qiang Cui. Activation mechanism of a signaling protein at atomic resolution from advanced computations. *Journal of the American Chemical Society*, 129(33):10261–10268, 2007.
- [257] Chung-Jung Tsai, Antonio Del Sol, and Ruth Nussinov. Allostery: absence of a change in shape does not imply that allostery is not at play. *Journal of Molecular Biology*, 378(1):1–11, 2008.
- [258] Chung-Jung Tsai, Antonio Del Sol, and Ruth Nussinov. Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. *Molecular Biosystems*, 5(3):207–216, 2009.
- [259] Sven Brüschweiler, Paul Schanda, Karin Kloiber, Bernhard Brutscher, Georg Kontaxis, Robert Konrat, and Martin Tollinger. Direct observation of the dynamic process underlying allosteric signal transmission. *Journal of the American Chemical Society*, 131(8):3063–3068, 2009.
- [260] Joachim Lätzer, Tongye Shen, and Peter G Wolynes. Conformational switching upon phosphorylation: a predictive framework based on energy landscape principles. *Biochemistry*, 47(7):2110–2122, 2008.
- [261] Jeryl L Appleby and Robert B Bourret. Identification of a novel response regulator required for the swarmer-to-stalked-cell transition in *Caulobacter crescentus*. *Journal Of Bacteriology*, 180(21):6223–6229, 1995.
- [262] Mark S Formanek, Liang Ma, and Qiang Cui. Reconciling the "old" and "new" views of protein allostery: a molecular simulation study of chemotaxis Y protein (CheY). *Proteins*, 63(4):846–867, 2006.
- [263] Michael H Knaggs, Freddie R Salsbury, Marshall Hall Edgell, and Jacquelyn S Fetrow. Insights into Correlated Motions and Long-Range Interactions in CheY Derived from Molecular Dynamics Simulations. *Biophysical Journal*, 92(6):2062–2079, 2007.

- [264] S Y Lee, H S Cho, J G Pelton, D Yan, E A Berry, and D E Wemmer. Crystal structure of activated CheY. Comparison with other activated receiver domains. *The Journal of Biological Chemistry*, 276(19):16425–16431, 2001.
- [265] Volz K. Structural Conservation in the CheY Superfamily. *Biochemistry*, 32:11741, 1993.
- [266] M Simonovic and K Volz. A Distinct Meta-active Conformation in the 1.1-Å Resolution Structure of Wild-type ApoCheY. *J. Biol. Chem.*, 276:28637, 2001.
- [267] B F Volkman, D Lipson, D E Wemmer, and D Kern. Two-state allosteric behavior in a single-domain signaling protein. *Science*, 291(5512):2429–2433, 2001.
- [268] Jr. A. D. MacKerell, D Bashford, M Bellott, Jr. R. L. Dunbrack, J D Evan Seck, M J Field, S Fischer, J Gao, H Guo, S Ha, D Joseph-McCarthy, L Kuchnir, K Kuczera, F T K Lau, C Mattos, S Michnick, T Ngo, D T Nguyen, B Prodhom, I I I W. E. Reiher, B Roux, M Schlenkrich, J C Smith, R Stote, J Straub, M Watanabe, J Wiorcikiewicz-Kuczera, D Yin, and M Karplus. All atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102:3586–3616, 1998.
- [269] J Ryckaert, G Ciccotti, and H Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23(3):327–341, 1977.
- [270] Robert B Bourret. Receiver domain structure and function in response regulator proteins. *Current Opinion in Microbiology*, 13(2):142–149, 2010.



# Part I

APPENDED

PAPERS



## Paper I

*In this paper we review the force-field-based approaches to investigate - through computer simulations - reactive processes in chemical and biological systems. This review was written together with Dr. Pierre-Andre Cazade, Dr. Jing Huang, Dr. Jaroslaw J. Szymczak and Prof. Markus Mewly.*

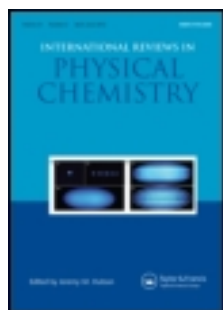


This article was downloaded by: [ETH Zurich]

On: 22 August 2012, At: 01:55

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## International Reviews in Physical Chemistry

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/trpc20>

### Atomistic simulations of reactive processes in the gas- and condensed-phase

Pierre-André Cazade<sup>a</sup>, Jing Huang<sup>a</sup>, Juvenal Yosa<sup>a</sup>, Jaroslaw J. Szymczak<sup>a</sup> & Markus Meuwly<sup>a</sup>

<sup>a</sup> Department of Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056 Basel, Switzerland

Version of record first published: 06 Jul 2012

To cite this article: Pierre-André Cazade, Jing Huang, Juvenal Yosa, Jaroslaw J. Szymczak & Markus Meuwly (2012): Atomistic simulations of reactive processes in the gas- and condensed-phase, *International Reviews in Physical Chemistry*, 31:2, 235-264

To link to this article: <http://dx.doi.org/10.1080/0144235X.2012.694694>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## Atomistic simulations of reactive processes in the gas- and condensed-phase

Pierre-André Cazade, Jing Huang, Juvenal Yosa, Jaroslaw J. Szymczak and Markus Meuwly\*

*Department of Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056 Basel, Switzerland*

*(Received 5 February 2012; final version received 14 May 2012)*

This review focuses on force-field-based approaches to investigate – through computer simulations – reactive processes in chemical and biological systems. Both, reactions in the gas-phase and in condensed-phase environments are discussed and opportunities and the potential for further developments are pointed out. Where available, results are compared with alternative methods and the advantages and drawbacks of the methods are compared. Particular applications include vibrationally and electronically induced (photo)dissociation of small molecules, proton transfer in the gas- and condensed phase and ligand un- and re-binding in proteins.

**Keywords:** Molecular Dynamics; Reactive MD; Force Field; Proton Transfer; Ligand Binding

	Contents	PAGE
<b>1.</b>	<b>Introduction</b>	236
<b>2.</b>	<b>Gas phase reactions</b>	238
2.1.	PT reactions with force fields	238
2.2.	Gas phase dissociation reactions	241
2.2.1.	Vibrationally induced dissociation of sulphuric acid (H <sub>2</sub> SO <sub>4</sub> )	241
2.2.2.	Final state distribution following photodissociation	245
<b>3.</b>	<b>Condensed phase reactions</b>	246
3.1.	PT in proteins	246
3.2.	Ligand binding in proteins	247
3.2.1.	Denitrication in truncated hemoglobin	253
3.2.2.	Fe–NO substitution reaction	254
<b>4.</b>	<b>Advantages and disadvantages of various methods</b>	256
<b>5.</b>	<b>Conclusion and outlook</b>	258
	<b>Acknowledgements</b>	258

---

\*Corresponding author. Email: m.meuwly@unibas.ch

## References

### 1. Introduction

Calculating the rate at which chemical reactions occur from first principles is a formidable task. The problem involves several steps, each of which poses its own challenges. Most fundamentally, one has to represent the intermolecular forces which govern the interactions between the atoms involved. Next, the dynamical problem of how the atoms move from educt to product on the potential energy surface (PES) has to be solved. Finally, a rate constant can be computed from integrating over the initial state distribution.

Fundamentally, the total interaction energy of an aggregate of atoms can be computed in two different ways: either by (approximately) solving the exact electronic Schrödinger equation or by a more or less empirical energy expression. Accurately solving the electronic Schrödinger equation has made tremendous progress since the 1960s and by now it is possible to converge total electronic energies to better than fractions of a kcal/mol for small- to medium-sized molecules (up to a few heavy atoms). On the other hand, repeated evaluation of energies as necessary for rate calculations becomes computationally prohibitive. Consequently, alternative means have been developed in parallel. They include:

- Additional simplifications to the rigorous electronic Schrödinger equation which lead to semiempirical methods, including parameterisations such as Parameterised Model 3 (PM3) or Austin Model 1 (AM1) and their variants. A more recent semiempirical theory is self-consistent charge-density functional tight binding (SCC-DFTB).
- Fully dimensional PESs which can be pre-calculated from electronic structure methods and then represented through a parameterised function. This is generally only applicable to systems with few degrees of freedom because fitting high-dimensional functions also becomes prohibitive with increasing dimensionality of the problem.
- A fully empirical energy expression.

This latter class is considered in more detail in the following.

Empirical force fields emerged in the late 1960s to study primarily conformational properties of alkanes [1,2]. They were subsequently used to refine protein structures, the first of which was myoglobin [3]. This development paved the way for the first molecular dynamics (MD) simulation of a protein on the picosecond time scale [4]. While very suitable for ground state conformational dynamics of a molecular system, force fields do not explicitly contain electronic degrees of freedom which is – strictly speaking – necessary to follow photochemical or reactive processes. Therefore, an additional layer needs to be introduced to investigate these phenomena. For reactive processes, one very promising step forward was to combine a force field description for the larger part of a system and to treat the reactive part with a quantum chemical method. Such a mixed quantum mechanics/molecular mechanics (QM/MM) ansatz has been used to study chemical reactions in larger systems. In QM/MM the system is divided into a reactive region, containing fewer atoms and described quantum mechanically (QM), and the rest of the system, containing the rest of the atoms and treated with an empirical force field (MM) [5–7]. The computationally demanding QM part is the bottleneck of the method that

dictates the accessible time scale of the simulations. Except for cases where semi-empirical methods are used for the QM-part, the nuclear dynamics and conformational sampling are typically not included in such studies. Improvements in the quality and speed of QM/MM methods, including SCC-DFTB and ‘neglect of differential overlap’ (NDDO), have been presented in recent reviews [8,9].

An alternative to study chemical reactions is to use empirical force fields with provisions to form and break bonds [10]. One approach, based on valence bond theory, is the ‘atoms in molecules’ (AIM) and the ‘diatomics in molecules’ (DIM) method, derived from it [11,12]. Originally developed to explain the electronic structure of polyatomic molecules from information about all possible diatomic and monoatomic fragments, it later matured into a technique which built globally valid PESs for molecular systems including the description of bond breaking and bond formation [13]. Related to this is empirical valence bond (EVB) [14], which has been extensively used to investigate catalytic reactions in biological systems with particular emphasis on including environmental effects. An alternative to EVB is the adiabatic reactive MD (ARMD) that has been recently proposed [15–17]. In ARMD the reactant and product are treated with individual force fields and transitions between the states occur on the basis of an energy criterion. The algorithm employs an energy criterion to decide when a crossing should occur. After a crossing is detected, the two states are mixed over a short period of time and the simulation continues on the second surface [16,17]. Such an approach allows to treat efficiently bond breaking and bond formation reactions in the framework of classical MD simulations (for some more technical details, see Section 2.2). Another promising method is ReaxFF, which is a reactive force field based on bond-order potentials [18]. Applications of ReaxFF were more focused on material science and high-energy processes including explosive materials and combustion [19,20]. Alternatively, empirical potentials consisting of Coulomb- and dispersion interactions and the Born–Meyer–repulsion, fitted to Hartree–Fock calculations on a  $\text{SiO}_4^{4-}$  cluster, allowed to estimate structural and energetic differences between four alternative crystal structures [21]. A somewhat refined parameterisation allowed to directly study the pressure-induced phase transition between various polymorphs of silica [22]. During the transitions, bonds between atoms need to be formed and broken and hence the force field is also ‘reactive’. A more recent extension is temporal QM/MM that combines conventional FFs with semiempirical and quantum chemical methods [23]. For the nonreactive portions of a trajectory, a conventional FF is employed whereby the conformations sampled are scored with a coordinate-dependent metric. Whenever the metric suggests a reaction to take place, a semiempirical calculation is used to decide whether or not the actual structure is likely to move towards the product. If this probability is high, the simulation is continued with quantum chemical methods and potential energies are mixed over a finite time window, similar to ARMD.

Compared to most other methods, ARMD does not employ geometrical descriptors to decide whether the system samples the product or educt state. This makes ARMD quite unique because possible reaction (or progression) coordinates are not determined *a priori* but rather can be extracted *a posteriori* from the reactive trajectories themselves. Suitable algorithms to do this have been presented in the literature [24]. All classical dynamics-based methods are limited in that they are not rigorously applicable in cases where the quantum nature (e.g. tunnelling, non-adiabaticity, coherence) of the reaction dominates. This is most obviously the case in small systems for which other and potentially more



suitable approaches are available [25–27]. Whenever a barrier-crossing is more appropriately described by a tunnelling frequency, ARMD is not the method of choice. In cases where a reaction rate is the more appropriate picture, ARMD provides a lower bound for the rate as it does not include transitions away from the crossing seam.

## 2. Gas phase reactions

The use of force fields to investigate reaction dynamics for gas phase reactions is by now an established area of research. Since the pioneering work on the  $\text{H} + \text{H}_2$  exchange reaction [28,29], which used a London–Eyring–Polanyi–Sato PES [30,31], many gas phase reactive processes have been studied using more or less empirical potential energy functions. Subsequently, approximations were made as to the dimensionality of the problem while retaining as much accuracy in the remaining degrees of freedom [32] or the problem was addressed in full dimensionality while sacrificing accuracy in treating details of the nuclear dynamics [33,34]. One of the largest calculations carried out so far concerns approximate diffusion Monte Carlo calculations on a parameterised, fully dimensional PES for proton transfer (PT) in malonaldehyde (MA) [35].

### 2.1. PT reactions with force fields

PT is fundamental for the function, structure and dynamics of chemically and biologically relevant systems. Direct investigation of PT dynamics is difficult experimentally. The most direct evidence for PT to occur are splitting of spectral features [36]. However, this is only available from high-resolution spectroscopy of small molecules in the gas phase [37–40]. Recently, high-resolution vibrational spectra have been recorded for protonated clusters. In 2003 Asmis *et al.* reported the vibrational spectrum of the shared proton in the protonated dimer [41]. Johnson *et al.* used vibrational predissociation spectroscopy to characterise the spectral signature of protonated molecular clusters [42,43]. Duncan *et al.* [44] reported photodissociation infrared spectra of water clusters  $\text{H}^+(\text{H}_2\text{O})_n$  ( $n = 2-5$ ) with and without Ar tagging. Time-resolved vibrational spectroscopy [45,46] and photoelectron spectroscopy [47] are also used to study PT in bulk water. The analysis of these experimental spectra is usually accompanied by computations. The vibrational transitions associated with the intermolecular PT or ‘low-barrier’ hydrogen bonds are very sensitive to the chemical environment, and the infrared signatures exhibited are diffuse [48,49]. Such highly dynamical systems involving anharmonic vibrations makes computational simulations, especially MD simulations valuable for analysing spectra of the studied system [50–53].

During the past few years, a force field-based method – Molecular Mechanics with Proton Transfer (MMPT) – has been developed and successfully applied to a variety of chemical and biological systems [51,54–60]. MMPT has been applied to several model PT systems. Examples include intermolecular PT in protonated water dimer [51], protonated ammonia dimer [55] and 2-pyridone-2-hydroxypyridine (2PY2HP) dimer [55], as well as intramolecular PT in protonated diglyme [51], MA [57] and acetylacetone [59].

In MMPT the interaction within a general  $\text{D-H}\cdots\text{A}$  motif is described by a parameterised three-dimensional PES  $V(R, r, \theta)$  where  $R$  is the distance between donor (D) and acceptor (A),  $r$  is the distance between the D and H atom, and  $\theta$  is the angle between  $\vec{R}$

and  $\vec{r}$  (Figure 1). To facilitate the parameterisation of MMPT potentials, the internal coordinate  $r$  is replaced by a dimensionless coordinate

$$\rho = (r - r_{\min}) / (R - 2r_{\min}) \quad (1)$$

where  $r_{\min} = 0.8 \text{ \AA}$  is in principle arbitrary but should be sufficiently small to cover the shortest possible D–H separations to avoid  $\rho$  becoming negative. Such 3D PESs  $V(R, \rho, \theta)$  are employed together with a standard force field which is CHARMM in this case [61]. To allow the breaking of the D–H and the formation of the H–A bond, bonded and non-bonded interactions on both the donor and acceptor side are explicitly evaluated, and are switched on and off depending on the position of the transferring atom (D–H  $\cdots$  A or D  $\cdots$  H–A).

MMPT PESs are available for generic proton- or hydrogen-transfer motifs including symmetric single minimum ( $\text{H}_2\text{O}-\text{H}^+ \cdots \text{OH}_2$ , Figure 1), symmetric double minimum ( $\text{H}_3\text{N}-\text{H}^+ \cdots \text{NH}_3$ ) and asymmetric double minimum ( $\text{H}_2\text{O}-\text{H}^+ \cdots \text{NH}_3$ ), calculated at the MP2/6-311++G(d, p) level of theory. The parameterised MMPT PESs  $V(R, r, \theta)$  have been fit to different functional forms. The radial dependence is always modelled by a double-Morse potential whereas for the  $\theta$ -dependence harmonic oscillators or Legendre polynomials are available. The detailed expressions and parameters are given in the literature [55,60]. These ‘generic’ or zeroth-order PESs can then be subsequently ‘morphed’ to adapt their overall shapes to topologically similar, but energetically different PESs depending on their chemical environment. In practice, morphing is achieved by modifying the MMPT parameters in a systematic way [62]. This can be a simple scaling [63] or a more general coordinate transformation depending on whether the purpose of this study and the experimental data justify such a more elaborate approach. For example, the MMPT PES of acetylacetone (AcAc, methyl-substituted MA) can be generated from that of MA [57] via the following PES morphing transformations [59]:

$$V^{\text{AcAc}}(R, \rho, \theta) = \lambda V^{\text{MA}}(R - R_0, \rho, \theta) \quad (2)$$

The morphing parameters  $\lambda = \Delta E^{\text{AcAc}} / \Delta E^{\text{MA}} = 0.796$  and  $R_0 = R_{\text{opt}}^{\text{MA}} - R_{\text{opt}}^{\text{AcAc}} = 0.03 \text{ \AA}$  are determined by comparing the PT energy barrier ( $\Delta E$ ) and the equilibrium donor–acceptor distance ( $R_{\text{opt}}$ ) of AcAc and MA computed by electronic structure calculations at the same level (MP2/aug-cc-PVTZ). PES morphing can also be performed to reproduce the X-ray structure [53] or even involves explicit fitting to dynamically averaged data [56,64].

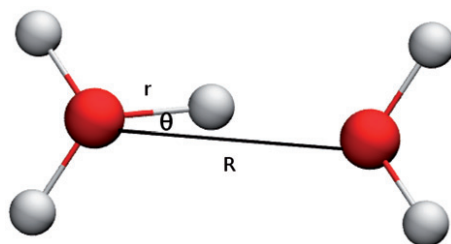


Figure 1. (Colour online) Protonated water dimer  $\text{H}_2\text{O}-\text{H}^+ \cdots \text{OH}_2$  as a prototype system for MMPT.  $R$ ,  $r$  and  $\theta$  are the distance between the donor and acceptor atom, the distance between the donor and transferring proton atom and the angle between them, respectively.

Constructed along the lines described above, MMPT is effectively a mixed QM/MM method where the QM part is precalculated and not evaluated on-the-fly. It has the additional advantage that no link atoms are required to connect the quantum and the classical part [65].

It is important to validate the PESs by comparing with experimental observables or other computational methods. For the protonated water dimer ( $\text{H}_5\text{O}_2^+$ ), MMPT/MD simulations find the asymmetric O–H $\cdots$ O stretching vibration at  $830\text{ cm}^{-1}$  which compares favourably with  $861\text{ cm}^{-1}$  computed from MD simulations by Bowman and co-workers on a 15-dimensional PES of  $\text{H}_5\text{O}_2^+$  at 100 K [66,67]. In  $\text{H}_5\text{O}_2^+$  this band strongly depends on temperature ( $724\text{ cm}^{-1}$  at 5 K compared with  $830\text{ cm}^{-1}$  at 100 K) whereas in protonated diglyme the same vibration is at  $870\text{ cm}^{-1}$  with no temperature dependence [51].

One advantage of MMPT is its ability to directly investigate PT. With atomistic MD simulations, the time series of donor–acceptor and donor–proton distances can be studied in detail [55,57,59]. Over-the-barrier rates for PT reactions can be calculated from the slope of Hazard plots [55,68]. For example, the intramolecular PT rate at 300 K in MA in vacuum and water is computed to be 2.4 and  $10\text{ ns}^{-1}$ , respectively. The fact that solvent enhances the rate by a factor of 5 implies a strong coupling between the transferring hydrogen atom and the solvent water [57]. For AcAc, the proton hopping rate in gas-phase is  $\approx 0.25\text{ ps}^{-1}$  [59], which is two orders of magnitude faster than that of MA. Other detailed information of the PT reaction, for example the fundamental role of the N–N vibration as the gating mode for PT in  $\text{NH}_4^+ \cdots \text{NH}_3$ , can also be captured [55]. It should be noted that with classical MD simulations only the over-the-barrier hopping rate can be determined, which should be considered as the lower limits for the PT rates. Including quantum tunnelling effects will increase the corresponding rates.

In situations where tunnelling plays a role, PT should rather be analysed in terms of a splitting than in terms of a rate constant. To estimate the tunnelling splitting, a suitable Hamiltonian for the quantum dynamics calculation is required. One suitable Hamiltonian is the harmonic bath averaged (HBA) Hamiltonian [32,57,69]:

$$H^{\text{HBA}}(p_s, s) = \frac{1}{2} \frac{p_s^2}{1 + \Delta_s} + V_0(s) + \sum_{k=1}^{3N-7} \frac{\omega_k(s)}{2} \left( 1 + \frac{\Delta_k(s)}{1 + \Delta_s} \right) \quad (3)$$

where  $s = \frac{\sqrt{m_H}(r_1 - r_2)}{2}$  is the hydrogen mass-weighted reaction path coordinate for PT and  $p_s$  is the corresponding momentum. It is important to emphasise that, although only 2 degrees of freedom are explicitly treated in the dynamics, such a Hamiltonian takes into account the effect of all degrees of freedom of the system. Based on this Hamiltonian, the experimentally observed H\*-tunnelling splitting of  $21.6\text{ cm}^{-1}$  is found if the mass of the transferring hydrogen atom H\* is chosen appropriately. This, however, determines the mass of the transferring deuterium atom D\* for which the tunnelling splitting has also been measured. The calculated D\*-tunnelling splitting is  $2.8\text{ cm}^{-1}$ , which is in very good agreement with the experiment ( $2.9\text{ cm}^{-1}$ ). The fact that the transferring particle is associated with a different mass than its physical one reflects the kinetic coupling between its degrees of freedom and the environment [57]. Furthermore, the IR active H\* and D\* fundamental vibrations can be obtained which are found at  $1573\text{ cm}^{-1}$  and  $1136\text{ cm}^{-1}$ , similar to recent computational results using alternative methodologies [35,70,71].

## 2.2. Gas phase dissociation reactions

PT reactions described above are particularly attractive because the time scales involved are short. On the other hand, quantum effects related to tunnelling and zero-point motion can become important. A different type of reaction concerns processes whereby chemical bonds between atoms are broken and new molecules are formed. Such reactions can involve one or several electronic states.

### 2.2.1. *Vibrationally induced dissociation of sulphuric acid ( $H_2SO_4$ )*

The notion that molecules in different internal states react more or less efficiently is a fascinating topic which is best studied under controlled conditions. An example for a vibrational mode-specific reaction involves  $CH_4$  on a Ni(100) surface. It was found that the reactivity of methane changes depending on whether two quanta are vibrationally excited along the same CH stretch or whether one quantum is deposited in two separate CH stretching modes [72]. This insight precludes the applicability of statistical models to correctly describe the underlying reaction mechanism. Another example is the  $Cl + CHD_3 \rightarrow HCl + CD_3$  reaction for which it was found that vibrational excitation of the CH-stretch in  $CHD_3$  was no more effective in driving the reaction than translational energy [73]. Computationally, this process was recently investigated in much detail based on fully dimensional PESs and quasiclassical MD simulations [74]. From  $2 \times 10^6$  trajectories, the cross sections for the reaction were determined as a function of the collision energy. The simulations found that compared to the vibrational ground state, vibrational excitation of stretch and bend vibrations indeed do enhance the reaction rate. However, at the same total energy  $E_{tot} = E_{coll} + E_{vib}$ , translational energy is more effective than the vibrational energy to drive the reaction for low  $E_{coll}$ .

Another process for which the reactivity of a molecule depends on its internal state is the vibrationally induced photodissociation of  $H_2SO_4$  into water and  $SO_3$ . In the atmosphere, sulphur is emitted mainly as sulphur dioxide ( $SO_2$ ), which is oxidised to  $SO_3$  and subsequently hydrolysed to  $H_2SO_4$  [75,76]. The cycling between  $SO_2$  and  $H_2SO_4$  has been studied mainly because it plays an important role in the troposphere aerosol layer formation, which is related to global atmospheric chemistry [77,78]. At high altitudes,  $SO_2$  molecules could be generated from  $H_2SO_4$  photolysis according to



with subsequent rapid photolysis of  $SO_3$  to produce  $SO_2$  [79,80]. However, there is no experimental observation of the electronic absorption spectrum of  $H_2SO_4$  up to 140 nm [81]. Moreover, UV photons are absorbed at higher altitudes [82], therefore only few photons penetrate into the atmosphere with wavelengths shorter than 179 nm [83].

An alternative process was proposed whereby the dissociation of  $H_2SO_4$  proceeds mainly from highly excited OH-stretching vibrational overtones [84,85]. Experimentally, using cavity ring-down spectroscopy, the asymmetric OH stretching vibration  $\nu_9$  of  $H_2SO_4$  with  $\nu_9 = 4$  and  $\nu_9 = 5$  has been characterised [86]. They were found at 13,490 and 16,494  $cm^{-1}$ , respectively. The barrier energy for  $H_2SO_4$  hydrolysis is predicted to be between 32 and 40 kcal/mol [76,87]. The amount of energy absorbed by several quanta in the OH stretching vibration  $\nu_9$  would then be sufficient for  $H_2SO_4$  hydrolysis.

The reaction dynamics after excitation of  $\text{H}_2\text{SO}_4$  is still a matter of debate. Initially, the vibrational dynamics following IR-excitation was investigated using classical trajectory simulations at the PM3 level [88]. For 98 trajectories, propagated for 400 ps each, it was found that in 5% of the photodissociation cases takes place with an average reaction time of 9 ps. To characterise the statistical limit of the reaction, Rice-Ramsperger–Kassel–Marcus (RRKM) theory was used to calculate reaction constants ( $k_{\text{RRKM}}$ ) and yields  $k_{\text{RRKM}} = 1.7 \times 10^7$ ,  $8.9 \times 10^9$  and  $1.1 \times 10^9$  s for  $\nu_9 = 4, 5$  and 6 quanta, respectively [89].

To atomistically analyse the reaction dynamics of vibrationally induced  $\text{H}_2\text{SO}_4$  decomposition (Figure 2) from a statistically significant number of trajectories, adiabatic reactive molecular dynamics (ARMD) was used [16,17,90]. Because ARMD was developed for empirical force fields, extensive sampling of phase space is possible and several thousand trajectories can be run and analysed [91]. ARMD involves two or multiple PESs defined by individual sets of force-field parameters corresponding to reactant and product states. For macromolecular systems the number of energy terms by which the PESs differ is much smaller compared to the total number of energy terms. Thus, by providing only a controllable number of additional parameters compared to a standard MD simulation, it is possible to describe the difference between the states of the system which limits computational overhead [17]. The dynamics of the system is initiated and propagated in its initial state and the energy difference  $\Delta E$  between all states is monitored at each time step. When  $\Delta E$  changes sign, a crossing is detected. The trajectory is restored in the configuration at time  $T_{\text{mix}}/2$  earlier and restarted on an appropriately mixed surface for a time  $T_{\text{mix}}$  such that at the end of the mixing time the system smoothly approaches the target state [17]. The choice of mixing time, which is often a few tens of fs, needs to be assessed, but so far has had only a small effect on the resulting trajectories [16,17]. During crossing, the energies and forces from the newly created or broken bonded terms are added or subtracted and the corresponding nonbonded interactions are updated.

Different schemes to prepare nonequilibrium initial conditions exist [92]. For example, both positions and velocities can be scaled, which leads to modifications in the kinetic and

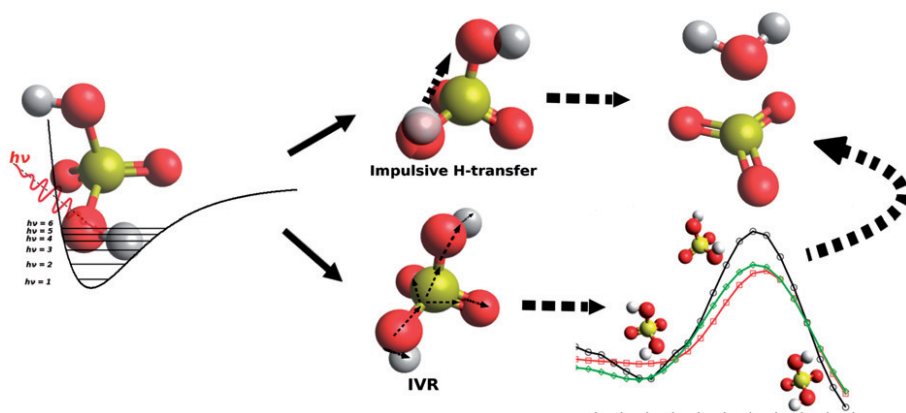


Figure 2. (Colour online) Photodissociation of  $\text{H}_2\text{SO}_4$ : after IR-excitation of one OH stretching vibration  $\nu_9$  the two limiting cases (a) impulsive H-transfer and (b) partial or complete IVR followed by dissociation are illustrated. One of the important degrees of freedom in IVR are torsions which are discussed in more detail in Figure 4.

potential energy to prepare a nonequilibrium state. One possible drawback of such a procedure is the fact that after excitation a short equilibration period (a few ps) is required to avoid artefacts due to close proximity of atoms [92]. Alternatively, the instantaneous velocity vector along the normal mode direction corresponding to the mode of interest can be scaled [50]. For excitation of the asymmetric OH-stretching mode ( $\nu_9$ ), the velocity vectors along the O–H bond were scaled to energies equivalent to  $\nu_9=4, 5$  and 6 quanta, respectively [91]. The OH-stretch is to a good approximation a local mode.

From the large number of reactive events (several thousand) the distributions of dissociation times  $p(\tau)$  for the excitation of  $\nu_9=4-6$  is determined (Figure 3). The distributions  $p(\tau)$  can be conveniently characterised by Gamma-distributions where the mean values for different degrees of excitation are  $\tau_d=201$  ps (for  $\nu_9=5$ ) and  $\tau_d=40$  ps (for  $\nu_9=6$ ) [91]. Both,  $p(\tau)$  and time series for the internal coordinates support the notion of a gradual transition from fully impulsive (very short reaction times) to almost complete internal vibrational redistribution (IVR) for the decomposition reaction. For the impulsive process one of the hydroxyl H is transferred to the other OH, resulting in  $\text{SO}_3$  and  $\text{H}_2\text{O}$  formation. This was observed by Miller and Gerber, where the process was called H-hopping [88]. Trajectories that follow this process are located at the beginning of the gamma distribution, few picoseconds after photoexcitation (Figure 3). For IVR it is necessary that the initially deposited vibrational energy rapidly redistributes over the remaining degrees of freedom before a sufficiently large amount returns into the dissociative coordinates (Figure 3) [89,91].

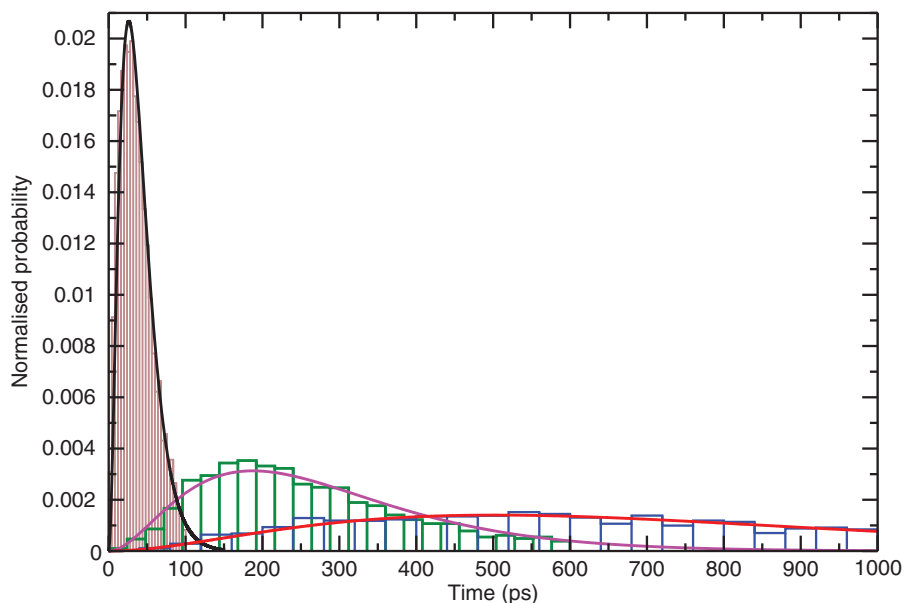


Figure 3. (Colour online) Normalised distribution of dissociation events  $p(\tau)$  as a function of time  $\tau$  for 5000 different trajectories with excitations of  $\nu_9=4, \nu_9=5$  and  $\nu_9=6$ , respectively. The blue distribution (red curve) correspond to  $\nu_9=4$ , green (magenta) to  $\nu_9=5$  and brown (black) to  $\nu_9=6$ . Red, magenta and black curves were determined by fitting to gamma distributions. Average dissociation times are 753, 201 and 40 ps for  $\nu_9=4, \nu_9=5$  and  $\nu_9=6$ , respectively.



Figure 4 reports probability distribution functions  $p(\phi)$  for the O6–S1–O4–H5 torsional angle  $\phi$  from reactive trajectories at  $\nu_9 = 6$  with different reaction times (6 (A), 244 (B) and 588 (C) ps). For each reactive trajectory (Figure 4 (red)) only one equilibrium trajectory (Figure 4 (black)) was used to compare the values for the torsional angle  $\phi$ . For each case the amount of data to build the distribution from equilibrium and nonequilibrium simulations was the same (e.g. if the reaction time was 6 ps as in Figure 4(A) only 6 ps from the equilibrium trajectory were used). Probability distributions for the equilibrium state (before excitation) are in black whereas those corresponding to the activated state are in red. The two equilibrium positions are at  $\phi = \pm 40^\circ$ . An impulsive reaction is analysed in Figure 4(A). For this case the equilibrium and nonequilibrium distribution closely match each other and the ratio  $p_{\text{activated}}/p_{\text{equil}}$  is basically featureless – also because of the short simulation time (the reaction occurs after 6 ps). In particular, the transition state  $\phi = 0$  is sampled equally by both trajectories. However, for longer reaction times (Figure 4B and C) the transition state is much more readily sampled from the nonequilibrium trajectories. This indicates that this particular coordinate is considerably activated through direct or indirect coupling to the vibrationally excited O–H stretching vibration. Also, a larger angular range is explored as is suggested from  $p_{\text{activated}}/p_{\text{equil}}$  beyond  $\pm 40^\circ$ .

Approximate IVR rates can also be extracted from ARMD simulations. They are ( $\tau_d = 7 \times 10^{10}$ ,  $2 \times 10^{10}$  and  $4 \times 10^{11}$ ) which is about two orders of magnitude more rapid than those from the RRKM theory (see above). However, in view of complete IVR the simulations carried out so far have probably not reached the statistical limit and further work is required. One can expect that complete IVR takes about an order of magnitude

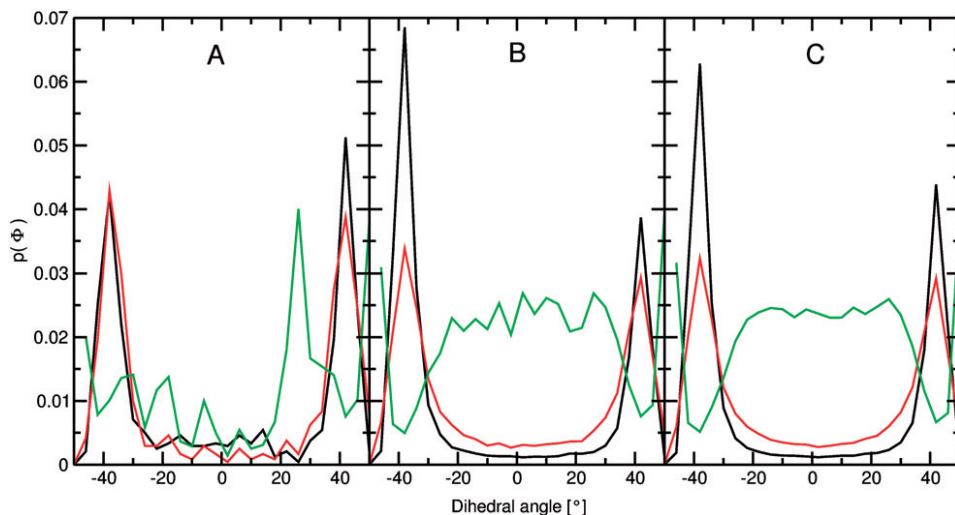


Figure 4. (Colour online) Probability distribution for one O–S–O–H torsional angle  $\phi$  (see Figure 2) for equilibrated (black) and vibrationally activated (red)  $\text{H}_2\text{SO}_4$ . The angle  $\phi$  describes the transition between the  $\text{C}_2$  and the  $\text{C}_s$  structure. Panels A, B and C report  $\phi$ -distributions for trajectories exhibiting different reaction times (6, 244 and 588 ps, respectively). The green curve is the ratio  $p_{\text{act}}(\phi)/p_{\text{eq}}(\phi)$ . For long reaction times the transition state region separating the two metastable states ( $\text{C}_2$  and  $\text{C}_s$ ) is more readily sampled.

longer which brings the RRKM and explicitly simulated decomposition times within about 1 order of magnitude.

Overall, atomistic simulations support the suggestion that vibrational excitation of  $\nu_9 > 4$  allows the photodissociation of  $\text{H}_2\text{SO}_4$  into  $\text{SO}_3 + \text{H}_2\text{O}$  in the gas phase [88,91]. The reaction times for  $\nu_9=4$  to  $\nu_9=6$  are on the picosecond time scale. Two limiting mechanisms were found to be operative: impulsive H-transfer and almost complete IVR with subsequent dissociation. All intermediate scenarios are also found as is evident from the smooth reaction probability functions. It is expected that trajectory calculations along these lines can considerably improve our understanding of energy redistribution on longer time scales.

### 2.2.2. Final state distribution following photodissociation

Photodissociation reactions, i.e. dissociation reactions induced through the absorption of photons, have played an important role in elucidating intermolecular interactions, ground- and excited state dynamics and energy flow in small molecules [93,94]. In the gas phase such reactions are typically considered to occur in a two-step fashion. Starting from the electronic ground-state  $X$  of molecule ABC, the absorption of a photon leads to an activated complex  $\text{ABC}^*$  which decays upon energy transfer between electronic and rovibrational degrees of freedom to  $\text{A} + \text{BC}$ . In its final state, BC will again assume a wide range of internal states whose population depends on the available energy and the details of the intermolecular dynamics and energy redistribution mechanisms.

A particularly well-studied example concerns the photodissociation of ClCN [94–99]. Classical and quantum simulations were carried out and had been found to be in very good agreement [99]. Compared to the experiment, the computed rotational state distribution  $P(j)$ , where  $j$  is the angular momentum of the photodissociated CN molecule, were found to be in qualitative agreement [96,97]. However, contrary to the experiments, simulations did not find a characteristic wavelength-dependence of both, the width of  $P(j)$  and the location  $j_{\text{max}}$ .

Extending ARMD to electronic transitions allowed to more realistically model this process [100]. Using the original PESs [96,101], several thousand trajectories were statistically analysed and found good agreement between computed and experimentally observed  $P(j)$  was found. Based on this agreement, further adjustments of the interaction potentials could now be attempted. Because photodissociation is an intrinsically rapid process and ARMD simulations can be carried out at MD-speed, the process could also be investigated in a water-droplet. It was found that  $P(j)$  is strongly quenched due to the presence of the solvent. Furthermore, the wavelengths at which photodissociation was possible differed between the reaction gas- and solution-phase. Also, the proportion of reactive trajectories differed between the two environments. Corresponding work has also been carried out for ICN in solution and at surfaces using MD-based approaches. These simulations focused on the dissociation dynamics at fixed wavelength and included nonadiabatic effects in the excited states [102–104]. These examples illustrate that MD-based techniques with refined interaction potentials hold much promise to investigate the adiabatic and non-adiabatic dynamics in realistic gas- and solution-phase processes.



### 3. Condensed phase reactions

Compared to gas-phase processes, reactions in the condensed phase have an environment (solvent, surrounding protein) as an active participant. Amongst others, the environment can electrostatically shield or assist reaction partners or modify the dynamics through its viscosity and intermolecular interactions. As most biological and the majority of chemically interesting processes occur in the presence of an actual environment, taking its role into account is of great interest. Methods based on electronic structure calculations are difficult to extend to the relevant spatial scales. It is for these types of processes that methods such as EVB [14] or ARMD[17] have been developed.

#### 3.1. PT in proteins

Recent advances in molecular biology and structural biology have highlighted the importance of PT in biological systems [105]. It serves as a rapid means to transport charge in systems including bacteriorhodopsin [106,107] and cytochrome c oxidase [108,109] PT also participates in numerous enzymatic catalysis – often as an elementary step [110] including liver alcohol dehydrogenase (LADH) [111–114] and dihydrofolate reductase [115–117]. LADH catalyses the oxidation of alcohols to aldehydes with the reduction of the coenzyme nicotinamide adenine dinucleotide (NAD<sup>+</sup> to NADH). A key step has been suggested to be the transport of one proton from the alcohol to solvent through a hydrogen-bonded network that consists of the hydroxyl group of Ser<sup>48</sup>, NAD<sup>+</sup> and His<sup>51</sup> (Figure 5). This three-step PT has been studied with QM/MM simulations where the QM and MM regions are treated with SCC-DFTB [118,119] and the CHARMM22 force field [61], respectively [112]. Classical MD simulations with the MMPT force fields are currently being carried out to investigate such PT processes. For this, the enzyme is solvated in a pre-equilibrated 90.0 Å × 68.3 Å × 65.2 Å water box and periodic boundary conditions were applied. As a preliminary result, the one-dimensional potential of mean force (PMF) for the first PT step was computed using umbrella sampling [120,121], shown in Figure 5. A free energy barrier of 16 kcal/mol is found, and the structures and energies of reactant and product for this PT step compare favourably with the results by Cui *et al.* [112]. Extensive atomistic simulations concerning the remaining two PT steps, as well as the refinement of MMPT parameters and development of algorithms allowing continuous PT, are currently in progress.

The first atomically resolved PT reaction in a biological system was suggested to be found in ferredoxin I of *Azotobacter vinelandii* [122]. However, preliminary MD simulations [122] only found rare and quite long-distance acceptor-hydrogen distances, which make direct transfer of the proton from the protein to the buried [3Fe–4S] cluster unlikely. In detail, the reaction suggested was Asp<sup>15</sup>–COOH + [3Fe–4S]<sup>0</sup> → Asp<sup>15</sup>–COO<sup>−</sup> + [3Fe–4S]H<sup>+</sup>. Subsequent MD simulations suggested that water molecules can have extended lifetimes (sub-ns) between the Asp15 residue and the buried [3Fe–4S]<sup>0</sup> cluster [123,124]. Consequently, the alternative water-assisted process Asp15–COOH + H<sub>2</sub>O + [3Fe–4S]<sup>0</sup> → Asp15–COO<sup>−</sup> + H<sub>2</sub>O + [3Fe–4S]<sup>0</sup>H<sup>+</sup> was studied [58]. MMPT simulations and electronic structure calculations yield a barrier of 11.7 kcal/mol for the forward reaction which is in good agreement with the experimental value (13.3 kcal/mol) [125]. These simulations also suggest that no stable intermediate hydronium ion (H<sub>3</sub>O<sup>+</sup>) is expected. Contrary to that, the originally proposed, water-

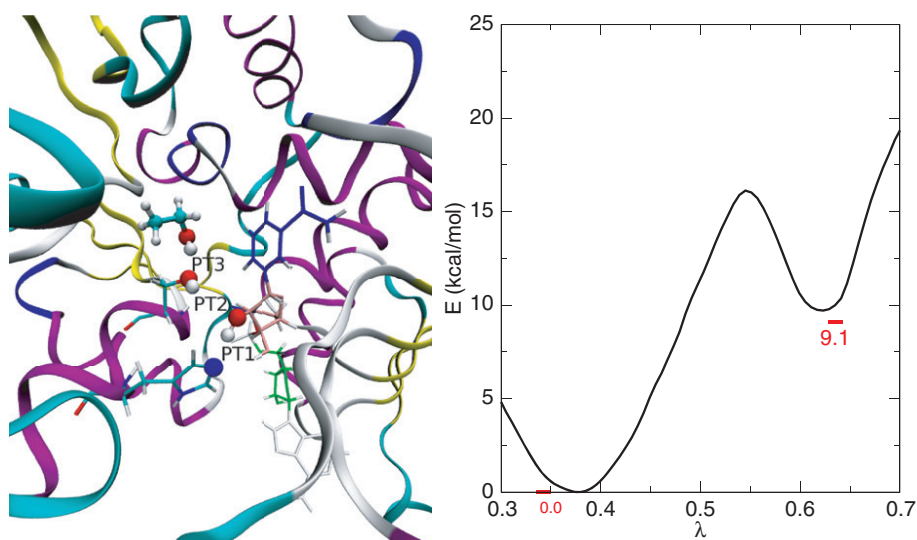


Figure 5. Left panel: the active site of LADH. The atoms directly involved in PT processes together with the substrate  $C_2H_5OH$  are shown in ball-and-stick (red: oxygen; blue: nitrogen; white: hydrogen; green: carbon); the cofactor  $NAD^+$  and protein residue 48 (Ser) and 51 (His) are shown in lines; the rest of the protein is shown in ribbon. Right panel: the potential of mean force for PT1 computed by MMPT/MD simulations along  $\lambda = r_{OH}/R_{ON}$ . Related PT states ( $\lambda$  and energies) computed with QM/MM methods by Cui *et al.* [112] are shown in red.

unassisted process is found to have a barrier of  $\Delta E \approx 35$  kcal/mol [58] and therefore can be ruled out.

### 3.2. Ligand binding in proteins

One of the paradigm reactive processes in proteins is ligand binding and unbinding in globins – especially hemo (Hb) – and myoglobin (Mb). Small ligands including CO, NO or  $O_2$  can diffuse into the protein and form more or less strongly bound ligand–Fe bonds with the heme-iron. Using flash photolysis, the kinetics of ligand rebinding to the heme-group in Mb was studied in detail across a wide range of temperatures [126–135], viscosities and in different regions of the electromagnetic spectrum. For CO, geminate rebinding occurs on the sub-microsecond time scale whereas for NO and  $O_2$  this process involves only small barriers and happens on the picosecond time-scale. Experimental studies of the rebinding of CO after photodissociation at different temperatures have shown that the superficially simple rebinding reaction is surprisingly complex [126–135]. The room temperature rebinding of CO is approximately exponential (with a time constant of 100 ns), while at low temperatures the rebinding is non-exponential in time and varies in a non-Arrhenius fashion with temperature [126,129,130,132,136]. This behaviour has been attributed to a distribution of barrier heights for ligand rebinding in the effectively frozen protein [126,130,137,138]. This assumes that the relaxation of the protein at low temperatures is slow relative to the rebinding reaction so that the system is inhomogeneous with each protein molecule having a different barrier for rebinding.

Nitric oxide, on the other hand, rebinds extremely rapidly and non-exponentially at all temperatures [133]. Assuming a power law dependence for the rebinding time yields a time constant of around 33 ps while a fit to a double exponential gave time scales of 28 and 280 ps [133]. A recent time-resolved infrared spectroscopy study [139] has reproduced this non-exponential behaviour but finds that the rebinding is even quicker, with time constants of 5.3 and 133 ps. Femtosecond mid-infrared experiments [140] have found that rebinding occurs mainly from a metastable state  $B_1$  with time-dependent rates, in agreement with earlier simulations [15]. In addition to  $B_1$ , two other populations ( $B_0$  and  $B_2$ ) were identified.

Given these widely different time scales, different computational approaches need to be considered. For CO, which will not be discussed any further, various approximations ranging from dimensional reduction [137,138] to application of MD-quenching techniques [141] based on Marcus theory and separating the actual dynamics from the energetics have been pursued in the past [142,143]. Contrary to that, the picosecond time-scale reaction of NO with heme-iron is ideally suited to be investigated with all-atom, fully dimensional MD simulations. It is for this type of processes for which ARMD was originally developed. The primary incentive is to follow chemical reactions along the time-coordinate without imposing a geometrical reaction coordinate.

Interactions of myoglobin and its mutants with NO have been studied experimentally and theoretically for almost four decades now [15,16,126,144–158]. Experiment [147,159] has shown that the rebinding of NO to Mb features non-exponential kinetics at all temperatures. Rebinding times vary from 5.3 to 34 ps and from 133 to 280 ps for the fast and slow components, respectively [139,147,160,161], with the most recent study pointing to shorter recombination times. However, the experiment itself is not able to provide an atomistically resolved mechanistic picture of the processes involved.

Several models were put forward to explain the observed non-exponential rebinding times. First, they have been attributed to distributed barriers [162] related to internal relaxation of the heme unit, second, a related multiple binding site model [163] within Mb from which a fraction of the photodissociated ligands can rebind has been proposed, and finally the relaxation model which features a time-dependent barrier that modulates ligand rebinding [133]. ARMD simulations were carried out whereby two states of the system were included: the hexacoordinate, bound MbNO conformation and the pentacoordinate, unbound Mb–NO state. The analysis of several thousand all-atom simulations in explicit solvent demonstrated that rebinding of NO indeed is nonexponential and occurs on the picosecond time scale [16]. Apart from the independently validated force fields for the two states, the only free parameter in such simulations is the asymptotic energy separation  $\Delta$  between the two energy manifolds [16,164]. The free parameter is usually estimated from electronic structure calculations and a conformational-independent value is assumed which, however, is an approximation. In these simulations [16], rebinding times of 3.8 and 18.0 ps compared with 28 and 280 ps [133], 5.3 and 133 ps [139] from the experiment were found. Given the substantial differences between the reported experimental data [133,139,165], the computed results can be considered to qualitatively agree with the experimental data. More importantly, the simulations allow to analyse individual families of trajectories. It is found that the early processes after photodissociation are governed by

heme relaxation, resulting from the displacement  $d$  of the iron out of the heme plane. This is reflected in the iron-out-of-plane probability distribution  $p(d)$  for specific rebinding times. For rapid rebinding,  $p(d)$  has a broad maximum between 0.2 and 0.3 Å whereas for slower rebinding the maximum in  $p(d)$  shifts to 0.33 Å which agrees quite well with results from X-ray crystallography [166]. Furthermore, on the nanosecond time scale NO molecules are found in both the distal heme pocket and the Xe4 pocket which suggests that rebinding can occur from different sites within the protein, and thus supports the multiple site model for longer time scales.

Recent time-resolved resonance Raman experiments suggested that the photodissociated ligand can rebind to the domed, Fe-out-of-plane conformation of the heme unit [156]. This experiment found that the heme-iron motion and NO rebinding kinetics are not synchronised. Upon NO-photolysis, the transition from an in-plane to a domed-Fe conformation occurs on the subpicosecond time scale whereas the opposite process (domed to in-plane) takes  $\approx 30$  ps. On the other hand, in the infrared the rapid rebinding component occurs within 5 ps which suggests that the returning ligand can rebind to a domed-iron conformation. This has already been suggested by Ionascu *et al.* [159]. To shed additional light on the structural dynamics underlying this process, an extended model for the heme-NO interaction is currently being developed. Apart from the previously considered relative orientation of the NO ligand to the heme-iron [150,152,153,158,167,168], the iron-out-of-plane (doming) coordinate will be explicitly considered. Experimentally, the existence of two nitrosyl linkage isomers (FeNO and FeON) was confirmed by infrared difference spectroscopy [150,167,168].

PESs for the in-plane  $d=0$  and one of the out-of-plane  $d=-0.2$  Å positions of the iron atom are reported in Figure 7. The PESs are characterised by two minima including the known Fe-NO bent minimum (Fe-N-O angle  $\phi \approx 146^\circ$ ), and a more shallow Fe-ON ( $\phi = 23^\circ$ ) minimum  $\approx 22$  kcal/mol higher than the first minimum for the Fe-in-plane ( $d=0$ ) conformation. Unrestrained *ab initio* optimisations lead to structures with  $\phi = 24^\circ$  (approximately 18 kcal/mol higher than the global Fe-NO minimum) which have  $\phi = 141^\circ$ . The barrier for the Fe-ON  $\rightarrow$  Fe-NO transition is  $\approx 3$  kcal/mol. For  $d = -0.2$  Å (Fe-out-of-plane) the primary minimum is less stable by 10 kcal/mol than for the  $d=0$  PES. The Fe-ON minimum determined on the  $d = -0.2$  Å surface located  $\approx 21.5$  kcal/mol above the global Fe-NO minimum, is, however, shifted toward smaller values of  $\theta$  ( $\approx 120^\circ$ ), with a geometry resembling more the metastable Fe- $\eta$ NO described previously [153]. The barrier between these minima is less than 0.5 kcal/mol. Initially, the effect of the different interaction potentials on the ligand motion in its bound state was briefly considered. Starting from geometries close to the FeON geometry, 100 trajectories were run and their relaxation into the global FeNO minimum was followed. The investigated system contained a total number of 2532 protein atoms, nitric oxide [16,153], and 178 water molecules represented by a modified TIP3P potential [169]. Classical MD simulations were performed using CHARMM [170] with the CHARMM22 force field [171] supplemented with parameters described earlier [153]. Trajectories were propagated for 10 ps with a time step of 1 fs. All simulations were performed at 300 K. A few selected trajectories are projected onto the corresponding PES and reported in Figure 7. Trajectories on the  $d=0$  PES sample the FeON minimum for up to 4 ps before isomerisation to FeNO occurs. This contrasts with the  $d=-0.2$  PES where the vanishing barrier leads to almost immediate isomerisation. It is also instructive to consider the crossing geometries for the

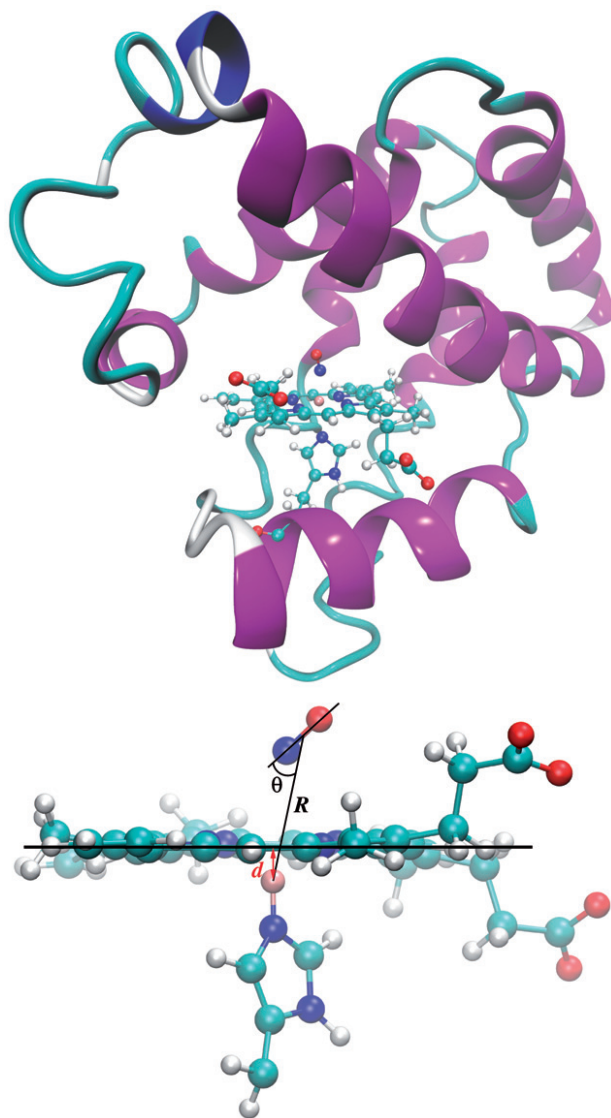


Figure 6. The nitrosyl myoglobin complex (top) with displayed heme-NO system. Black lines on the bottom panel present schematic picture of Jacobi coordinates  $\text{Fe-CoB}_{\text{NO}}$  ( $R$ ), and  $\text{Fe-CoB}_{\text{NO-N}}$  ( $\theta$ ) used to construct Fe-NO binding potential. Heme doming parameter  $d$  is shown in red.

two PESs. Both surfaces feature very wide crossing seams, which are shown in Figure 8. Although crossing occurs in both cases at almost the same distance of NO from the iron atom (see Figure 8 top), for the domed heme  $\theta$  angles are shifted significantly towards larger values as shown in the middle panel of Figure 8. These characteristics along with much narrower distributions of the transition times (see bottom of Figure 8) are consistent



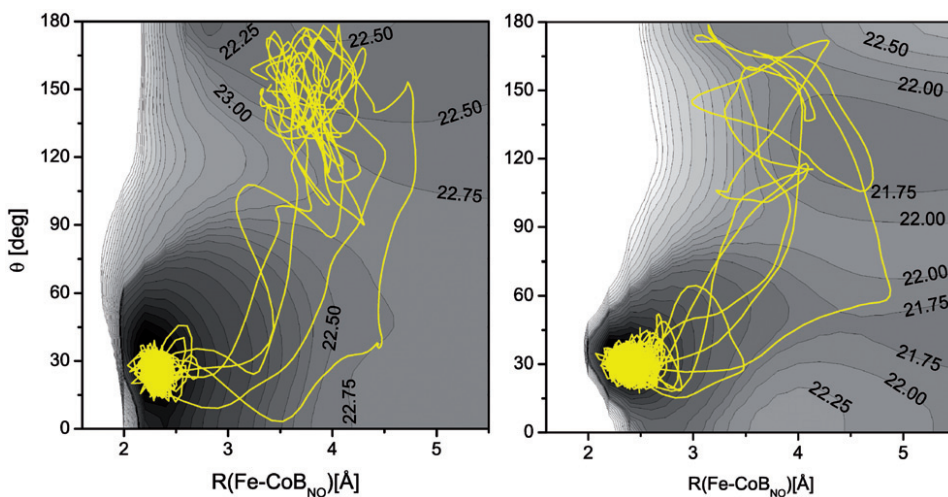


Figure 7. (Colour online) Projection of selected trajectories (yellow) onto two-dimensional representations of the PESs determined for the conformation corresponding to the in-plane (left) and domed (right) heme. Energies are reported in kcal/mol,  $R$  and  $\theta$  in Å and degrees, respectively.

with the picture obtained from the projection of trajectories shown in the right panel of Figure 7 showing minor sampling of FeON geometries followed by almost immediate transition to the FeNO minimum. Simulations in the full 3D PES and including the unbound state to follow the rebinding dynamics should provide further insight into the complex dynamics following NO photodissociation from Mb.

The NO rebinding dynamics has been investigated with a range of techniques. The information about time scales of NO recombination is available from the early room temperature-induced absorption measurements in the visible light regime by Petrich *et al.* [133]. They reported double exponential character of the NO rebinding, with rebinding times of 28 and 280 ps for the short and the long component, respectively. Later studies [139,151,156,159] also found two time scales with quite different rebinding time scales. For example, the picosecond absorption and the time-resolved resonance Raman measurements [151,156] report short  $\sim 10$  and long  $\sim 100$  ps rebinding times, while the corresponding values obtained from mid-IR [139] and transient absorption [159] report 5.3 and 13.8 ps for a short and 133 and 200 ps for the long component, respectively. Time-resolved resonance Raman measurements provided structural information about process that has not been reported yet by the X-ray diffraction measurements. For example, the picosecond absorption and the time-resolved resonance Raman measurements by Martin and co-workers [151,156] provide information about the in-plane movement of ferrous heme-iron following NO rebinding after fs-photodissociation. Femtosecond transient absorption [159] experiments investigate the influence of protein environment (temperature, viscosity) on the NO-rebinding barriers whereas mid-IR absorption spectroscopy [139] probes the influence of conformational substates of the protein on the rebinding dynamics by monitoring changes in the intensity of the NO intensity stretching mode

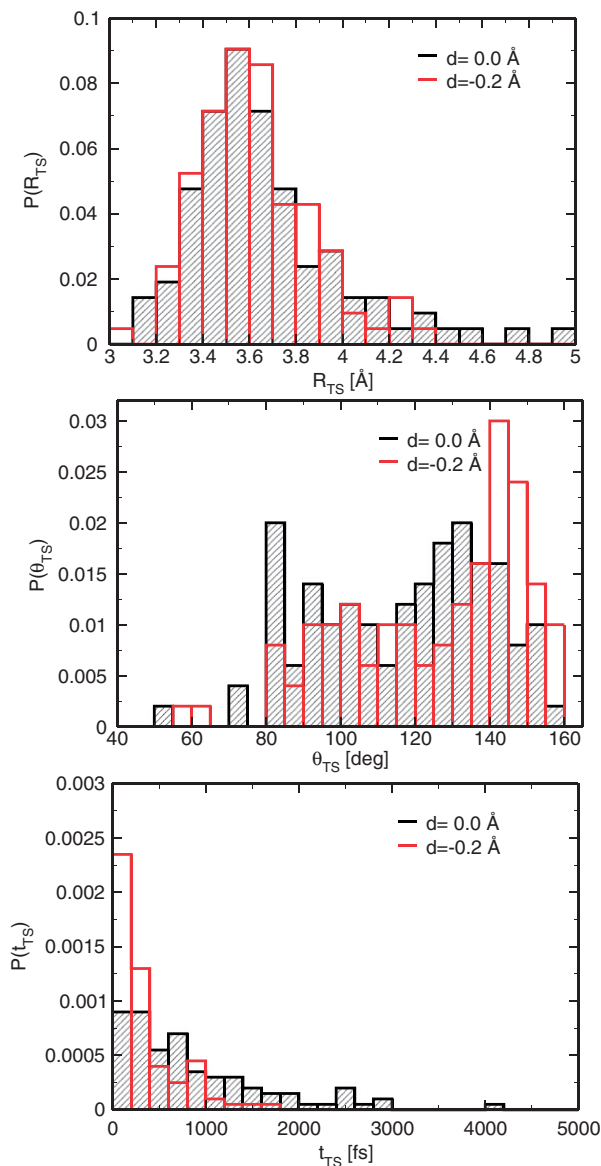


Figure 8. (Colour online) Probability distribution function of the Fe-CoBNO ( $R_{TS}$ ) distance (top), Fe-CoBNO-N ( $\theta_{TS}$ ) angle (middle), and time ( $t_{TS}$ ) (bottom) determined for the crossing geometries. Values for in-plane  $d = 0$  and out-of-plane  $d = -0.2 \text{ \AA}$  positions of the iron atom are shown in black and red, respectively.

upon photolysis. Although all these different experimental techniques find two time scales for the rebinding kinetics, no uniform picture as to the atomistic interpretation underlying the process has been obtained so far. This underlines the need for reliable computational investigations of the underlying processes.

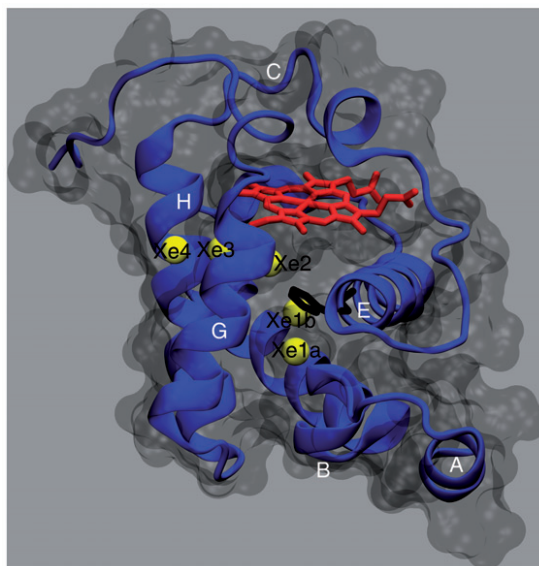
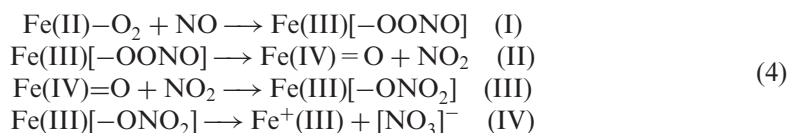


Figure 9. X-ray structure of truncated hemoglobin (PDB code 1s56 [179]). The backbone is represented in blue and the helices are labeled. The heme and Phe62 residues are shown in red and black sticks, respectively. The five xenon pockets found experimentally in the X-ray data are represented by yellow spheres to indicate the main docking sites. The solvent accessible surface is represented as a transparent black volume around the protein.

### 3.2.1. Denitrification in truncated hemoglobin

Truncated hemoglobin N (trHbN) (Figure 9) – a member of the truncated hemoglobin (trHb) family – is involved in efficient NO scavenging [172,173] whereby toxic NO is converted to harmless nitrate ( $\text{NO}_3^-$ ). Due to its physiological significance the entire family has received much recent interest, both experimentally and computationally [174–186]. The conversion of  $\text{O}_2$  and NO to  $\text{NO}_3^-$  has been studied with a range of experimental techniques in globins from different organisms. On the basis of stopped flow kinetic measurements, it was proposed [187] that the reaction pathway for NO oxidation in Mb involves the peroxyxynitrite ( $\text{Fe(III)-OONO}$ ) intermediate which undergoes O–O homolysis to form free nitrite radicals ( $\text{NO}_2$ ). In this model  $\text{NO}_2$  then binds to the oxo-ferryl ( $\text{Fe(IV)=O}$ ) group to form the nitrate complex ( $\text{Fe(III)-ONO}_2$ ), which ultimately leads to free nitrate. Alternatively, rapid-scan ultraviolet-visible (UV-vis) spectroscopy [187–191] and electron paramagnetic resonance (EPR) spectroscopy [192]. provided evidence for a  $\text{Fe(III)-OONO}$  (high-spin) intermediate in both oxy-Mb and oxy-Hb under alkaline conditions ( $\text{pH}=9.5$ ). At neutral pH, the reaction proceeds without any observed intermediates [188], suggesting a rapid rearrangement of the  $\text{Fe(III)-OONO}$  intermediate to nitrate. These different processes are summarised in the following reaction equations:





Stopped-flow spectroscopy yielded a second-order rate constant of  $7.5 \times 10^8 \text{ M}^{-1}\text{s}^{-1}$  [193]. Compared with the values obtained for vertebrate Mbs and Hbs, the reaction is 10–20 times faster in trHbN. Such rates make the NO dioxygenation reaction ( $\text{Fe(II)}-\text{NO} + \text{O}_2 \rightarrow \text{Fe}^+ \text{ (III)} + [\text{NO}_3]^-$ ) ideal to be studied by atomistic simulations. Irrespective of the reaction mechanism (see below), it is commonly agreed that in trHbN the diffusion-controlled migration of the ligands (NO and O<sub>2</sub>) to the heme active site is the slowest process in NO detoxification and, hence, it is the rate determining step [194,195].

Experimentally, no intermediates have been found so far. This and the fact that the studies in Mb and Hb provide conflicting and inconclusive views [187,188,196,197] suggested that computational studies should be of value to clarify mechanistic aspects of the reaction. To allow bond breaking and ARMD formation, simulations were employed in order to account for fluctuations in the nuclear motion [15–17]. The necessary force fields for the different states in reactions I–IV (see above) were derived from electronic structure calculations for a model system containing the porphyrin ring, Fe atom and an imidazole ring representing the proximal histidine and the O<sub>2</sub> and NO ligands [198]. ARMD simulations allow to follow specific pathways. Earlier suggestions involved pathway A including all four steps I–IV described above [187,190,199,200], whereas pathway B involves a rearrangement reaction from Fe(III)[–OONO] to the Fe(III)[–ONO<sub>2</sub>] complex [188,201]. ARMD simulations for steps I, III, and IV, along pathway A yield rate constants on the picosecond time scale [198]. However, step II was not found to be reactive. As the simulations were carried out for several nanoseconds, step II must be considerably slower than this time scale. The free-energy profile along the O1–O2 distance was calculated from umbrella sampling simulations at 300 K. It exhibits NO<sub>2</sub> dissociation barriers of 12–15 kcal/mol which precludes second-order rate constants of  $7.5 \times 10^8 \text{ M}^{-1}\text{s}^{-1}$  based on transition state theory.

On the other hand, ARMD simulations for the rearrangement reaction (pathway B) yield rate constants in the picosecond range [198], which readily explains the rapid overall NO dioxygenation that is associated with a second-order rate constant of  $7.5 \times 10^8 \text{ M}^{-1}\text{s}^{-1}$  in trHbN [193]. Moreover, Herold *et al.* were unable to observe any dissociated NO<sub>2</sub> which should emerge along pathway A. All these observations suggest that pathway B may be preferred [188,191]. Such a model has also been supported recently through experiments on a hexacoordinate oxy-globin model Fe(Por)(O<sub>2</sub>)(NH<sub>3</sub>) which was reacted with NO at  $\approx 80 \text{ K}$  [202]. FTIR and visible spectra were recorded upon warming the sample. Spectroscopic signatures characteristic of the –O<sub>2</sub> complex disappeared whereas new bands corresponding to an  $\eta^1$ -coordinated nitrate (–ONO<sub>2</sub>) compound emerged. The barrier for the overall process was estimated to be  $\approx 7 \text{ kcal/mol}$  which favourably compares with the computations (upper limit of 8.6 kcal/mol) [198].

### 3.2.2. Fe–NO substitution reaction

One last point that deserves discussion is the fact that the NO-bound protein is more stable than the O<sub>2</sub>-bound one. In group II truncated hemoglobin (trHbO) of *Mycobacterium Tuberculosis*, the rate constant for NO binding to ligand-free trHbO ( $0.18 \mu\text{M}^{-1}\text{s}^{-1}$  (80%) and  $0.95 \mu\text{M}^{-1}\text{s}^{-1}$  (20%)) is slightly larger than the corresponding rate constant for O<sub>2</sub> binding ( $0.11 \mu\text{M}^{-1}\text{s}^{-1}$  (80%) and  $0.85 \mu\text{M}^{-1}\text{s}^{-1}$  (20%)) [181,203]. Given this and the typically higher affinity of Fe(II) towards NO than to O<sub>2</sub>, it is likely that the

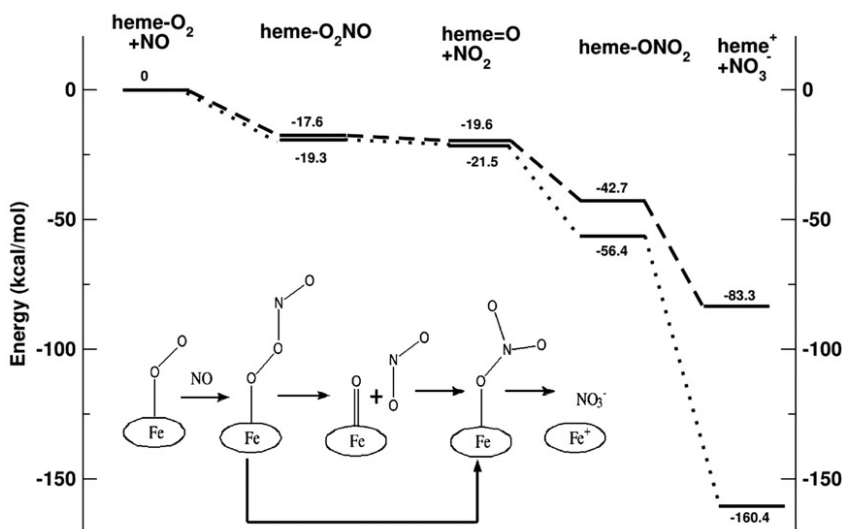
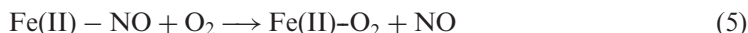


Figure 10. Schematic representation of the full reaction NO detoxification and the associated energies. The inset represents the succession of the reactions with the bond breaking and/or formation. The thick arrow indicates which steps have been studied so far.

reaction sequence discussed above (see Equation (4)) is preceded by a ligand exchange step (Figure 10):



In other words, starting from the ligand-free protein, the likely sequence of elementary steps includes NO and O<sub>2</sub> diffusion in the protein, NO-binding, replacement of NO by O<sub>2</sub> and finally denitrication as described above. Nitric oxide migration has been analysed for O<sub>2</sub>-bound trHbN and a network of ligand migration pathways has been found [90]. From several thousand transitions between the metastable binding sites [90,179], a ligand migration network was built and analysed with a transition network analysis which captures the essential ligand dynamics [204]. NO migration to the reactive site is about 10 times slower than the reaction time found in ARMD which supports the notion that ligand migration is the rate limiting step. More recently, we also started to investigate the network for O<sub>2</sub>-migration in NO-bound trHbN [205]. By comparing O<sub>2</sub> and NO as free ligands, it is found that both ligands, have their own favourable docking sites.

The ligand-exchange reaction – either SN1 (bond breaking and formation are two independent steps) or SN2 (bonds are created and broken in a concerted step) – can also be studied by means of ARMD simulations. New features in ARMD now allow to consider multiple PESs corresponding to different chemical connectivity [91]. Thus, the Fe+NO+O<sub>2</sub> (S1), Fe-NO+O<sub>2</sub> (S2) and Fe-O<sub>2</sub>+NO (S3) states can be involved simultaneously. From propagation on S2, the system can proceed through two different mechanisms without choosing *a priori* any path (S1 for SN1 or S3 for SN2), as illustrated in Figure 11. Initial ARMD simulations suggest that the substitution reaction is an SN1 mechanism as the unbinding of NO always takes place before O<sub>2</sub> eventual binding

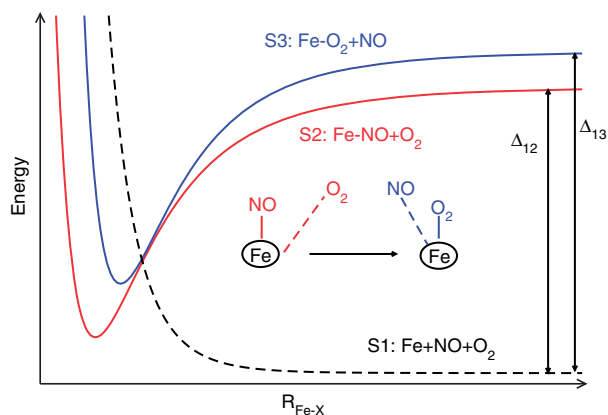


Figure 11. (Colour online) Representation of the three PESs corresponding to the different states involved in the substitution mechanism which is schematically represented by the reaction equation.

after a few tens of picoseconds of dynamics on the unbound PES (S1). These results are also supported by quantum calculations where no concerted substitution has been observed.

#### 4. Advantages and disadvantages of various methods

In the following, advantages and disadvantages of the different approaches are briefly summarised. If entropic effects are expected to be important, use of rigorous QM(/MM) methods is precluded due to their computational demands to sample a representative number of conformations. This will be the case for most, if not all condensed-phase reactions but is not necessarily true for gas-phase – in particular PT – reactions. If one chooses to use electronic structure-based schemes, (reparameterised) semiempirical methods including AM1, PM3 or SCC-DFTB are attractive alternatives. Parameterised PESs, such as MMPT, are not only computationally efficient but also require adjusted PESs to include environmental effects. In any case, reparameterization of these methods is mandatory either to improve their performance for a given problem and/or to include environmental effects.

Fully dimensional, fitted PESs based on accurate *ab initio* calculations are only a viable approach for small systems for several reasons. First, the computation of the energies to be fitted has to be carried out on a grid. Assuming 10 points per internal degree of freedom (dof),  $10^n$  points are needed for  $n$  internal dofs. In other words, for a nonlinear molecule with 6 atoms,  $10^{12}$  single point calculations are in principle needed although the actual number of evaluations may be reduced, e.g. due to symmetry. Fitting such a PES to a global, parameterised form is the next obstacle which is not easily overcome. Finally, the utility of this parameterised PES is limited to gas phase investigations as its shape depends on the chemical environment. One possibility would then be to morph the entire PES given either reference experimental (solution) data or additional computations with an environment present.

In general, semi-empirical methods are faster than the Hartree–Fock or post-Hartree–Fock methods. The fact that such methods are parameterised can be seen as a strength or a weakness, but it is clear that this provides a large degree of flexibility. However, the reparameterisation step can be tedious and is usually non-unique. On the other hand, such methods allow to treat proteins (up to 15,000 atoms), for example, with the MOZYME algorithm, which uses localised molecular orbitals (LMO) [206]. As another example, phosphoryl transfer reactions have been studied with *ab initio* (MP2), DFT and several semi-empirical methods (AM1, PM3, MNDO, MNDO/d and SCC-DFTB). All methods give large errors in the proton affinity and although the phosphoryl transfer mechanism is still controversial, the re-parameterisation of SCC-DFTB for phosphor-containing systems gives improved structures and energetics compared to previous studies [207]. The disadvantage of semi-empirical methods is that the parameterisation is seldom transferable and only properties of molecules similar to those used in the parameterisation can be expected to be reproduced reliably.

Force field-based approaches now provide several ways to treat efficiently bond breaking and bond formation reactions in the framework of classical MD simulations. Their main advantage is that they allow long time scale simulations on large systems including a realistic chemical environment (solvent) or multiple (thousands) simulations for fast reactions from which a detailed picture of the inter- and intramolecular dynamics can be obtained. Moreover, the full system (reactive/non-reactive) is treated at a consistent level of accuracy. As an example, EVB treats the dynamics on a well-defined potential surface and is relatively easy to implement. Nevertheless, the parameterisation can again be tedious. This is largely because the off-diagonal terms in EVB, i.e. the coupling between the states, needs to be determined ‘empirically’ [208,209]. Usually a functional expression for the coupling term needs to be assumed *a priori* and the corresponding parameters are determined by fitting to *ab initio* calculations. These parameters are, however, usually non-transferable and also implicitly define a reaction coordinate through the functional dependence of the off-diagonal element.

This is one of the advantages of ARMD where possible reaction coordinates are not determined *a priori* but rather can be extracted *a posteriori* from the reactive trajectories themselves. This advantage, due to the absence of geometrical descriptors, is also a limitation of the technique as it makes fitting the force field for the reactive states more difficult. In particular, careful optimisation of the van der Waals ranges is necessary because they lose their meaning in the reactive region [91].

In conclusion, QM or QM/MM methods are most accurate but suffer from the computational burden they present for large-scale and long-time simulations. Often, only one or a few trajectories can be run [88], which leaves the question open of how representative the results are. On the other hand, semi-empirical and entirely force-field-based methods can be applied to a broad range of reactive processes even including a realistic environment, but they all require a level of refinement for the intermolecular interactions. This step is non-trivial and future efforts should seriously address this point. Once suitable computational tools are available for this step, such methods will find more widespread applications and acceptance in chemistry, physics and biology. Furthermore, the accuracy of force fields is still limited and may not be uniform throughout configuration space but steady progress is made in improving their performance [210,211].

## 5. Conclusion and outlook

The current review summarises developments to follow reactive processes at the atomic scale using parameterised methods. The distinguishing feature of ARMD over alternative methods (such as EVB or ReaxFF) is the fact that instead of geometrical reaction coordinates, time as the 'natural' coordinate is used as the progression variable. The majority of applications so far uses classical MD simulations. However, if the process in question requires nuclear dynamics to be followed at a quantum level, this is also possible. Future extensions of ARMD include multi-surface dynamics, which is currently being developed [91], and non-classical transitions which can be effectively included by surface hopping whereby a transition between two PESs is modelled as a stochastic process [25,26].

Multi-surface dynamics in multi-dimensional systems at the classical dynamics level has recently been used for characterising the vibrationally induced photodissociation in  $\text{H}_2\text{SO}_4$ . Because the parameterisation allowed two product states to be populated ( $\text{SO}_3 + \text{H}_2\text{O}$  and  $\text{SO}_3 + \text{OH} + \text{H}$ ), the dynamics involves three states. Including the correct asymptotic energetics of each of the product channels allowed to realistically populate the possible final states [91]. Provided that suitable force fields are available for each of the states involved, extensions to more PESs is straightforward.

A particularly interesting application of reactive MD simulations of any flavour are processes in solution. Recent examples include photodissociation dynamics of ICN [103,104] and ClCN [100] in explicit solvent. For solvated ICN, a combined experimental and computational study provided detailed insight into laser-induced reaction dynamics in the condensed phase [104]. We anticipate that approaches as those discussed here will be of considerable importance to understand condensed phase reactions on multiple time and spatial scales at an atomistic level and provide insights that are complementary and sometimes inaccessible to direct experimentation.

## Acknowledgements

We gratefully acknowledge continued financial support from the Schweizerischer Nationalfonds through grant 200021-117810 and the NCCR MUST. Generous allocation of computing time at the CSCS in Manno, Switzerland, is acknowledged.

## References

- [1] J.B. Hendrickson, *J. Am. Chem. Soc.* **83**, 4537 (1961).
- [2] S. Lifson and A. Warshel, *J. Chem. Phys.* **49**, 5116 (1968).
- [3] M. Levitt and S. Lifson, *J. Mol. Biol.* **46**, 269 (1969).
- [4] J.A. McCammon, B.R. Gelin, and M. Karplus, *Nature* **267** (5612), 585 (1977).
- [5] A. Warshel and M. Levitt, *J. Mol. Biol.* **103**, 227 (1976).
- [6] U.C. Singh and P.A. Kollman, *J. Comput. Chem.* **7**, 718 (1986).
- [7] M.J. Field, P.A. Bash, and M. Karplus, *J. Comput. Chem.* **11**, 700 (1990).
- [8] K.W. Sattelmeyer, J. Tirado-Rives, and W.L. Jorgensen, *J. Phys. Chem. A* **110** (50), 13551 (2006).
- [9] T.J. Giese and D.M. York, *Theor. Chem. Acc.* **131** (3), 1145 (2012).

- [10] K. Farah, F. Müller-Plathe, and M.C. Böhm, *Chem. Phys. Chem.* **13** (5), 1127 (2012).
- [11] W. Moffitt, *Proc. Roy. Soc.* **210** (1101), 245 (1951).
- [12] F. Ellison, *J. Am. Chem. Soc.* **85** (22), 3540 (1963).
- [13] P. Kuntz, *Chem. Phys. Lett.* **16** (3), 581 (1972).
- [14] A. Warshel and R.M. Weiss, *J. Am. Chem. Soc.* **102**, 6218 (1980).
- [15] M. Meuwly, O.M. Becker, R. Stote, and M. Karplus, *Biophys. Chem.* **98**, 183 (2002).
- [16] D.R. Nutt and M. Meuwly, *Biophys. J.* **90**, 1191 (2006).
- [17] J. Danielsson and M. Meuwly, *J. Chem. Theo. Comp.* **4**, 1083 (2008).
- [18] A.C.T. van Duin, S. Dasgupta, F. Lorant, and W.A. Goddard, *J. Phys. Chem. A* **105**, 9396 (2001).
- [19] A. Strachan, A. van Duin, D. Chakraborty, S. Dasgupta, and W. Goddard, *Phys. Rev. Lett.* **91** (9), 098301 (2003).
- [20] S. Agrawalla and A.C.T. van Duin, *J. Phys. Chem. A* **115** (6), 960 (2011).
- [21] S. Tsuneyuki, M. Tsukada, H. Aoki, and Y. Matsui, *Phys. Rev. Lett.* **61**, 869 (1988).
- [22] S. Tsuneyuki, Y. Matsui, H. Aoki, and M. Tsukada, *Nature* **339** (6221), 209 (1989).
- [23] P. Dayal, S.A. Weyand, J. McNeish, and N.J. Mosey, *Chem. Phys. Lett.* **516** (4–6), 263 (2011).
- [24] A. Ma and A. Dinner, *J. Phys. Chem. B* **109** (14), 6769 (2005).
- [25] J.C. Tully and R. Preston, *J. Chem. Phys.* **55**, 562 (1971).
- [26] J.C. Tully, *J. Chem. Phys.* **93**, 1061 (1990).
- [27] W.H. Miller and T.F. George, *J. Chem. Phys.* **56**, 5637 (1972).
- [28] M. Karplus, R.D. Sharma, and R.N. Porter, *J. Chem. Phys.* **40** (7), 2033 (1964).
- [29] M. Karplus and R.D. Sharma, *J. Chem. Phys.* **43** (9), 3259 (1965).
- [30] H. Eyring and M. Polanyi, *Zeitschr. Phys. Chem.-Abteilung B-Chem. Element. Aufbau Mater.* **12** (4), 279 (1931).
- [31] S. Sato, *J. Chem. Phys.* **23** (12), 2465 (1955).
- [32] W.H. Miller, N.C. Handy, and J.E. Adams, *J. Chem. Phys.* **72**, 99 (1979).
- [33] H.S. Mei, M.E. Tuckerman, D.E. Sagnella, and M.L. Klein, *J. Phys. Chem. B* **102**, 10446 (1998).
- [34] M. Meuwly and M. Karplus, *J. Chem. Phys.* **116**, 2572 (2002).
- [35] Y. Wang, B.J. Braams, J.M. Bowman, S. Carter, and D.P. Tew, *J. Chem. Phys.* **128**, 224314 (2008).
- [36] D.R. Borst, J.R. Roscioli, D.W. Pratt, G.M. Florio, T.S. Zwier, A. Mueller, and S. Leutwyler, *Chem. Phys.* **112**, 3717 (2002).
- [37] S.L. Baughcum, R.W. Duerst, W.F. Rowe, Z. Smith, and E.B. Wilson, *J. Am. Chem. Soc.* **103**, 6296 (1981).
- [38] D.W. Firth, K. Beyer, M.A. Dvorak, S.W. Reeve, A. Grushow, and K.R. Leopold, *J. Chem. Phys.* **94**, 1812 (1991).
- [39] K. Tanaka, M. Toshimitsu, K. Harada, and T. Tanaka, *J. Chem. Phys.* **120**, 3604 (2004).
- [40] A.M. Daly, P.R. Bunker, and S.G. Kukolich, *J. Chem. Phys.* **132**, 201101 (2010).
- [41] K.R. Asmis, N.L. Pivonka, G. Santambrogio, M. Brümmer, C. Kaposta, D.M. Neumark, and L. Wöste, *Science* **299** (5611), 1375 (2003).
- [42] W.H. Robertson, E.G. Diken, E.A. Price, J.W. Shin, and M.A. Johnson, *Science* **299**, 1367 (2003).
- [43] J.M. Headrick, E.G. Diken, R.S. Walters, N.I. Hammer, R.A. Christie, J. Cui, E.M. Myshakin, M.A. Duncan, M.A. Johnson, and K.D. Jordan, *Science* **308**, 1765 (2005).
- [44] G.E. Douberly, R.S. Walters, J. Cui, K.D. Jordan, and M.A. Duncan, *J. Phys. Chem. A* **114**, 4570 (2010).
- [45] S. Woutersen and H.J. Bakker, *Phys. Rev. Lett.* **96**, 138305 (2006).
- [46] W. Amir, G. Gallot, F. Hache, S. Bratos, J.C. Leicknam, and R. Vuilleumier, *J. Chem. Phys.* **126**, 034511 (2007).



- [47] B. Winter, M. Faubel, I.V. Hertel, C. Pettenkofer, S.E. Bradforth, B. Jagoda-Cwiklik, L. Cwiklik, and P. Jungwirth, *J. Am. Chem. Soc.* **128**, 3864 (2006).
- [48] O. Vendrell, F. Gatti, and H.D. Meyer, *Angew. Chem. Intern. Ed.* **46**, 6918 (2007).
- [49] J.R. Roscioli, L.R. McCunn, and M.A. Johnson, *Science* **316**, 249 (2007).
- [50] M. Meuwly, A. Müller, and S. Leutwyler, *PCCP* **5**, 2663 (2003).
- [51] S. Lammers and M. Meuwly, *J. Phys. Chem. A* **111**, 1638 (2007).
- [52] U. Gellrich, J. Huang, W. Seiche, M. Keller, M. Meuwly, and B. Breit, *J. Am. Chem. Soc.* **133**, 964 (2011).
- [53] J. Huang, D. Häussinger, U. Gellrich, W. Seiche, B. Breit, and M. Meuwly, in preparation (2012).
- [54] S. Lammers and M. Meuwly, *Aust. J. Chem.* **57**, 1223 (2004).
- [55] S. Lammers, S. Lutz, and M. Meuwly, *J. Comp. Chem.* **29**, 1048 (2008).
- [56] J. Huang and M. Meuwly, *J. Chem. Theo. Comp.* **6**, 467 (2010).
- [57] Y. Yang and M. Meuwly, *J. Chem. Phys.* **133**, 064503 (2010).
- [58] S. Lutz, I. Tubert-Brohman, Y. Yang, and M. Meuwly, *J. Biol. Chem.* **286** (27), 23679 (2011).
- [59] J. Huang, Y. Yang, and M. Meuwly, submitted (2012).
- [60] J. Huang, Ph.D. thesis, University of Basel, 2011.
- [61] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorcikiewicz-Kuczera, D. Yin, and M. Karplus, *J. Phys. Chem. B* **102**, 3586 (1998).
- [62] M. Meuwly and J.M. Hutson, *J. Chem. Phys.* **110** (17), 8338 (1999).
- [63] J.M. Bowman and B. Gazdy, *J. Chem. Phys.* **95**, 816 (1991).
- [64] J. Huang and M. Meuwly, *Chem. Phys.* **13** (5), 1127 (2012).
- [65] P. Amara and M.J. Field, *Theor. Chem. Acc.* **109** (1), 43 (2003).
- [66] M. Kaledin, A.L. Kaledin, and J.M. Bowman, *J. Phys. Chem. A* **110**, 2933 (2006).
- [67] X.C. Huang, B.J. Braams, and J.M. Bowman, *J. Chem. Phys.* **122**, 044308 (2005).
- [68] E. Helfand, *J. Chem. Phys.* **69**, 1010 (1978).
- [69] Y. Yang and O. Kühn, *Mol. Phys.* **106**, 2445 (2008).
- [70] M. Schroeder, F. Gatti, and H.D. Meyer, *J. Chem. Phys.* **134** (23), 234307 (2011).
- [71] D.P. Tew, N.C. Handy, and S. Carter, *J. Chem. Phys.* **125**, 084313 (2006).
- [72] R.D. Beck, P. Maroni, D.C. Papageorgopoulos, T.T. Dang, M.P. Schmid, and T.R. Rizzo, *Science* **302** (5642), 98 (2003).
- [73] S. Yan, Y.T. Wu, B. Zhang, X.F. Yue, and K. Liu, *Science* **316** (5832), 1723 (2007).
- [74] G. Czako and J.M. Bowman, *Science* **334** (6054), 343 (2011).
- [75] C.F. Clement and I.J. Ford, *Atmos. Environ.* **33**, 1352 (1999).
- [76] L.J. Larson, M. Kuno, and F.M. Tao, *J. Chem. Phys.* **112**, 8830 (2000).
- [77] D.J. Hofmann and J.M. Rosen, *Nature* **297**, 120 (1982).
- [78] F. Raes, R.V. Dingenen, E. Vignati, J. Wilson, J.P. Pataud, J.H. Seinfeld, and P. Adams, *Atmos. Environ.* **34**, 4215 (2000).
- [79] M.J. Mills, O.B. Toon, and S. Solomon, *Geophys. Res. Lett.* **26**, 1133 (1999).
- [80] C.P. Rinsland, M.R. Gunson, M.K.W. Ko, D.W. Weisensten, R. Zander, M.C. Abrams, A. Goldman, N.D. Sze, and G.K. Yue, *Geophys. Res. Lett.* **22**, 1109 (1995).
- [81] P.E. Hintze, H.G. Kjaergaard, V. Vaida, and J.B. Burkholder, *J. Phys. Chem. A* **107**, 1112 (2003).
- [82] V. Vaida, *J. Phys. Chem. A* **113**, 5 (2009).
- [83] J.H. Seinfeld and S.N. Pandis, *Atmospheric Chemistry and Physics From Air Pollution to Climate Change* (John Wiley and Sons, New York, 1998).
- [84] D.J. Donaldson, A.F. Tuck, and V. Vaida, *Chem. Rev.* **103**, 4717 (2003).
- [85] V. Vaida, H.G. Kjaergaard, P.E. Hintze, and D.J. Donaldson, *Science* **299**, 1566 (2003).

- [86] K.J. Feierabend, D.K. Havey, S.S. Brown, and V. Vaida, *Chem. Phys. Lett.* **420**, 438 (2006).
- [87] K. Morokuma and C. Muguruma, *J. Am. Chem. Soc.* **116**, 10316 (1994).
- [88] Y. Miller and R.B. Gerber, *J. Am. Chem. Soc.* **128**, 9594 (2006).
- [89] Y. Miller, R.B. Gerber, and V. Vaida, *Geophys. Res. Lett.* **34**, L16820 (2007).
- [90] S. Mishra and M. Meuwly, *Biophys. J.* **96**, 2105 (2009).
- [91] J. Yosa and M. Meuwly, *J. Phys. Chem. A* **115**, 14350 (2011).
- [92] P.H. Nguyen and G. Stock, *J. Chem. Phys.* **119** (21), 11350 (2003).
- [93] R. Schinke, *Annu. Rev. Phys. Chem.* **39**, 39 (1988).
- [94] B. Schinke, *Photodissociation Dynamics: Spectroscopy and Fragmentation of Small Polyatomic Molecules* (Cambridge University Press, Cambridge, 1995), Vol. 2.
- [95] J.B. Halpern and W.M. Jackson, *J. Phys. Chem.* **86** (6), 973 (1982).
- [96] B.A. Waite and B.I. Dunlap, *J. Chem. Phys.* **84** (3), 1391 (1986).
- [97] S.A. Barts and J.B. Halpern, *J. Phys. Chem.* **93** (21), 7346 (1989).
- [98] R. Schinke and V. Engel, *Faraday Discuss.* **82**, 111 (1986).
- [99] R. Schinke, *J. Chem. Phys.* **92** (4), 2397 (1990).
- [100] S. Lutz and M. Meuwly, *Chem. Phys. Chem.* **13**, 305 (2012).
- [101] B.I. Dunlap, J.W.D. Connolly, and J.R. Sabin, *J. Chem. Phys.* **71** (8), 3396 (1979).
- [102] N. Winter and I. Benjamin, *J. Chem. Phys.* **121** (5), 2253 (2004).
- [103] M.L. Johnson and I. Benjamin, *J. Phys. Chem. A* **113** (26), 7403 (2009).
- [104] C.A. Rivera, N. Winter, R.V. Harper, I. Benjamin, and S.E. Bradforth, *PCCP* **13** (18), 8269 (2011).
- [105] C.A. Wraight, *Biochimica et Biophysica Acta* **1757**, 886 (2006).
- [106] U. Haupts, J. Tittor, and D. Oesterhelt, *Ann. Rev. Biophys. Biomole. Struct.* **28**, 367 (1999).
- [107] J.K. Lanyi, *Ann. Rev. Physiol.* **66**, 665 (2004).
- [108] M. Wikstrom, *Nature* **266**, 271 (1997).
- [109] M. Wikstrom, *Biochimica et Biophysica Acta* **1655**, 241 (2004).
- [110] Z.D. Nagel and J.P. Klinman, *Chem. Rev.* **106**, 3095 (2006).
- [111] S.R. Billeter, S.P. Webb, P.K. Agarwal, T. Jordanov, and S. Hammes-Schiffer, *J. Am. Chem. Soc.* **123**, 11262 (2001).
- [112] Q. Cui, M. Elstner, and M. Karplus, *J. Phys. Chem. B* **106**, 2721 (2002).
- [113] L. Masgrau, A. Roujeinikova, L.O. Johannissen, P. Hothi, J. Basran, K.E. Ranaghan, A.J. Mulholland, M.J. Sutcliffe, N.S. Scrutton, and D. Leys, *Science* **312**, 237 (2006).
- [114] J.N. Bandaria, C.M. Cheatum, and A. Kohen, *J. Am. Chem. Soc.* **131**, 10151 (2009).
- [115] P.K. Agarwal, S.R. Billeter, P.T.R. Rajagopalan, S.J. Benkovic, and S. Hammes-Schiffer, *Proc. Natl. Acad. Sci.* **99**, 2794 (2002).
- [116] M. Garcia-Viloca, J. Gao, M. Karplus, and D.G. Truhlar, *Science* **303**, 186 (2004).
- [117] S. Hammes-Schiffer and J.B. Watney, *Phil. Trans. R. Soc. B* **361**, 1365 (2006).
- [118] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, *Phys. Rev. B* **58**, 7260 (1998).
- [119] Q. Cui, M. Elstner, E. Kaxiras, T. Frauenheim, and M. Karplus, *J. Phys. Chem. B* **105**, 569 (2001).
- [120] G. Torrie and J. Valleau, *J. Comp. Phys.* **23**, 187 (1977).
- [121] J. Kottalam and D.A. Case, *J. Am. Chem. Soc.* **110**, 7690 (1988).
- [122] R. Camba, Y.S. Jung, L.M. Hunsicker-Wang, B.K. Burgess, C.D.S. CD, J. Hirst, and F.A. Armstrong, *Biochem.* **42**, 10589 (2003).
- [123] M. Meuwly and M. Karplus, *Faraday Discuss. Chem. Soc.* **124**, 297 (2003).
- [124] M. Meuwly and M. Karplus, *Biophys. J.* **86**, 1987 (2004).
- [125] J. Hirst, J.L.C. Duff, G.N.L. Jameson, M.A. Kemper, B.K. Burgess, and F.A. Armstrong, *J. Am. Chem. Soc.* **120** (28), 7085 (1998).
- [126] R.H. Austin, K.W. Beeson, L. Eisenstein, H. Frauenfelder, and I.C. Gunsalus, *Biochem.* **14**, 5355 (1975).



- [127] P.A. Cornelius, R.M. Hochstrasser, and W.A. Steele, *J. Mol. Biol.* **163**, 119 (1983).
- [128] J.L. Martin, A. Migus, C. Poyart, Y. Lecarpentier, R. Astier, and A. Antonetti, *Proc. Natl. Acad. Sci.* **80**, 173 (1983).
- [129] A. Ansari, J. Berendzen, S.F. Brown, H. Frauenfelder, I.E.T. Iben, T.B. Sauke, E. Shyamsunder, and R.D. Young, *Proc. Natl. Acad. Sci.* **82**, 5000 (1985).
- [130] A. Ansari, J. Berendzen, D. Braunstein, B.R. Cowen, H. Frauenfelder, M.K. Hong, I.E.T. Iben, J.B. Johnson, P. Ormos, T.B. Sauke, R. Scholl, A. Schulte, P.J. Steinbach, J. Vittitow, and R.D. Young, *Biophys. Chem.* **26**, 337 (1987).
- [131] P.A. Anfinrud, C. Han, and R.M. Hochstrasser, *Proc. Natl. Acad. Sci.* **86**, 8387 (1989).
- [132] P.J. Steinbach, A. Ansari, J. Berendzen, D. Braunstein, K. Chu, B.R. Cowen, D. Ehrenstein, H. Frauenfelder, J.B. Johnson, D.C. Lamb, S. Luck, J.R. Mourant, G.U. Nienhaus, P. Ormos, R. Philipp, A.H. Xie, and R.D. Young, *Biochem.* **30**, 3988 (1991).
- [133] J.W. Petrich, J.C. Lambry, K. Kuczera, M. Karplus, C. Poyart, and J.L. Martin, *Biochem.* **30**, 3975 (1991).
- [134] Q.H. Gibson, R. Regan, R. Elber, J.S. Olson, and T.E. Carver, *J. Biol. Chem.* **267**, 22022 (1992).
- [135] M. Karplus, *J. Phys. Chem. B* **104**, 11 (2000).
- [136] H. Frauenfelder, N. Alberding, A. Ansari, D. Braunstein, B. Cowen, M.K. Hong, I. Iben, B.J. Johnson, S. Luck, M. Marden, J. Mourant, P. Ormos, L. Reinisch, R. Scholl, A. Schulte, E. Shyamsunder, I.B. Sorensen, P.J. Steinbach, A.H. Xie, R.D. Young, and K.T. Yue, *J. Phys. Chem.* **94**, 1024 (1990).
- [137] N. Agmon and J.J. Hopfield, *J. Chem. Phys.* **79**, 2042 (1983).
- [138] V. Srajer, L. Reinisch, and P.M. Champion, *J. Am. Chem. Soc.* **110**, 6656 (1988).
- [139] S. Kim, G. Jin, and M. Lim, *J. Phys. Chem. B* **108**, 20366 (2004).
- [140] S. Kim and M. Lim, *J. Am. Chem. Soc.* **127**, 8908 (2005).
- [141] D. Zheng, V. Makarov, and P.G. Wolynes, *J. Am. Chem. Soc.* **118**, 2818 (1996).
- [142] P. Banushkina and M. Meuwly, *J. Phys. Chem. B* **109**, 16911 (2005).
- [143] P. Banushkina and M. Meuwly, *J. Chem. Phys.* **127**, 135101 (2007).
- [144] E.R. Henry, M. Levitt, and W.A. Eaton, *Proc. Natl. Acad. Sci.* **82**, 2034 (1985).
- [145] R. Elber and M. Karplus, *Science* **235**, 318 (1987).
- [146] J.A. McCammon and S.C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, Cambridge, 1987).
- [147] J.W. Petrich, J.C. Lambry, K. Kuczera, M. Karplus, C. Poyart, and J.L. Martin, *Biochem.* **30**, 3975 (1991).
- [148] T.G. Traylor and V.S. Sharma, *Biochem.* **31**, 2847 (1992).
- [149] E.A. Brucker, J.S. Olson, M. Ikeda-Saito, and J.G.N. Phillips, *Prot. Struct. Funct. Genom.* **30**, 352 (1998).
- [150] G.B. Richter-Addo, P. Legzdins, and J. Burstyn, *Chem. Rev.* **102**, 857 (2002).
- [151] M. Negrerie, S.G. Kruglik, J.C. Lambry, M.H. Vos, J.L. Martin, and S. Franzen, *J. Biol. Chem.* **281**, 10389 (2006).
- [152] D. Nutt, M. Karplus, and M. Meuwly, *J. Phys. Chem. B* **109**, 2118 (2005).
- [153] D. Nutt and M. Meuwly, *Chem. Phys. Chem.* **8**, 527 (2007).
- [154] K. Nienhaus, P. Palladino, and G.U. Nienhaus, *Biochem.* **47**, 935 (2008).
- [155] M. Radoul, M. Sundararajan, A. Potapov, C. Riplinger, F. Neese, and D. Goldfarb, *PCCP* **12**, 7276 (2010).
- [156] S.G. Kruglik, B.K. Yoo, S. Franzen, M.H. Vos, J.L. Martin, and M. Negrerie, *Proc. Natl. Acad. Sci.* **107**, 13678 (2010).
- [157] A. Arcovito, C. Ardiccioni, M. Cianci, P. D'Angelo, B. Vallone, and S. Della Longa, *J. Phys. Chem. B* **114**, 13223 (2010).
- [158] N. Xu, J. Yi, and G.B. Richter-Addo, *Inorg. Chem.* **49**, 6253 (2010).

- [159] D. Ionascu, F. Gruia, X. Ye, A. Yu, F. Rosca, C.B.A. Demidov, J.S. Olson, and P.M. Champion, *J. Am. Chem. Soc.* **127**, 16921 (2005).
- [160] Y. Kholodenko, E.A. Gooding, Y. Dou, M. Ikeda-Saito, and R.M. Hochstrasser, *Biochem.* **38**, 5918 (1999).
- [161] X. Ye, A. Demidov, and P.M. Champion, *J. Am. Chem. Soc.* **124**, 5914 (2002).
- [162] R.H. Austin, K.W. Beeson, L. Eisenstein, H. Frauenfelder, and I.C. Gunsalus, *Biochemistry* **14**, 5355 (1975).
- [163] H. Li, R. Elber, and J.E. Straub, *J. Biol. Chem.* **268**, 17908 (1993).
- [164] J. Danielsson and M. Meuwly, *J. Chem. Theo. Comp.* **4**, 1083 (2008).
- [165] S. Kim and M. Lim, *J. Am. Chem. Soc.* **127**, 5786 (2005).
- [166] H. Hartmann, S. Zinser, P. Komninos, R.T. Schneider, G.U. Nienhaus, and F. Parak, *Proc. Natl. Acad. Sci.* **93**, 7013 (1996).
- [167] L. Cheng, I. Novozhilova, C. Kim, A. Kovalevsky, K.A. Bagley, P. Coppens, and G.B. Richter-Addo, *J. Am. Chem. Soc.* **122**, 7142 (2000).
- [168] J. Lee, A.Y. Kovalevsky, I.V. Novozhilova, K.A. Bagley, P. Coppens, and G.B. Richter-Addo, *J. Am. Chem. Soc.* **126**, 7180 (2004).
- [169] W.L. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, and M. Klein, *J. Chem. Phys.* **79**, 926 (1983).
- [170] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, *J. Comp. Chem.* **4**, 187 (1983).
- [171] A.D. MacKerell Jr, D. Bashford, M. Bellott, R.L. Dunbrack Jr, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, and L. Kuchnir, *J. Phys. Chem. B* **102**, 3586 (1998).
- [172] A. Couture, S.R. Yeh, B.A. Wittenberg, J.B. Wittenberg, Y. Ouellet, D.L. Rousseau, and M. Guertin, *Proc. Natl. Acad. Sci.* **96**, 11223 (1999).
- [173] R. Poole, *Globins and other Nitric Oxide-reactive Proteins*, Vol. 436 (Academic Press, New York, 2008).
- [174] S.R. Yeh, M. Couture, Y. Ouellet, M. Guertin, and D.L. Rousseau, *J. Biol. Chem.* **275**, 1679 (2000).
- [175] M. Milani, A. Pesce, Y. Ouellet, P. Ascenzi, M. Guertin, and M. Bolognesi, *EMBO Rep.* **20**, 3902 (2001).
- [176] M. Mukai, P.Y. Savard, H. Ouellet, M. Guertin, and S.R. Yeh, *Biochem.* **41**, 3897 (2002).
- [177] M. Milani, P.Y. Savard, H. Ouellet, P. Ascenzi, M. Guertin, and M. Bolognesi, *Proc. Natl. Acad. Sci.* **100**, 5766 (2003).
- [178] M. Milani, Y. Ouellet, H. Ouellet, M. Guertin, A. Boffi, G. Antonini, A. Bocedi, M. Mattu, M. Bolognesi, and P. Ascenzi, *Biochem.* **43**, 5213 (2004).
- [179] M. Milani, A. Pesce, Y. Ouellet, S. Dewilde, J. Friedman, P. Ascenzi, M. Guertin, and M. Bolognesi, *J. Biol. Chem.* **279**, 21520 (2004).
- [180] M. Mukai, Y. Ouellet, H. Ouellet, M. Guertin, and S.R. Yeh, *Biochem.* **43**, 2764 (2004).
- [181] M. Milani, A. Pesce, M. Nardini, H. Ouellet, Y. Ouellet, S. Dewilde, A. Bocedi, P. Ascenzi, M. Guertin, L. Moens, J.M. Friedman, J.B. Wittenberg, M. Bolognesi, M.E. Johnson, and T. Head-Gordon, *J. Inorg. Biochem.* **99**, 97 (2005).
- [182] Y. Ouellet, M. Milani, M. Couture, M. Bolognesi, and M. Guertin, *Biochem.* **45**, 8770 (2006).
- [183] A. Bidon-Chanal, M.A. Martí, A. Crespo, M. Milani, M. Orozco, M. Bolognesi, F.J. Luque, and D.A. Estrin, *Proteins* **64**, 457 (2006).
- [184] A. Bidon-Chanal, M.A. Martí, D.A. Estrin, and F.J. Luque, *J. Am. Chem. Soc.* **129**, 6782 (2007).
- [185] A. Pesce, M. Nardini, M. Milani, and M. Bolognesi, *IUBMB Life* **59**, 535 (2007).
- [186] A. Pesce, M. Milani, M. Nardini, and M. Bolognesi, *Methods Enzymol.* **436**, 303 (2008).
- [187] J.L. Bourassa, E.L. Ives, A.L. Marqueling, R. Shim anovich, and J.T. Groves, *J. Am. Chem. Soc.* **123**, 5142 (2001).

- [188] S. Herold, M. Exner, and T. Nauser, *Biochem.* **40**, 3385 (2001).
- [189] P.R. Gardner, A.M. Gardner, W.T. Brashear, T. Suzuki, A.N. Hvitved, K.D. Setchell, and J.S. Olson, *J. Inorg. Biochem.* **100** (4), 542 (2006).
- [190] J. Su and J.T. Groves, *J. Am. Chem. Soc.* **131** (36), 12979 (2009).
- [191] S. Herold, *FEBS Letters* **443**, 81 (1999).
- [192] J.S. Olson, E.W. Foley, C. Rogge, A.L. Tsai, M.P. Doyle, and D.D. Lemon, *Free Radical Biol. Med.* **36** (6), 685 (2004).
- [193] H. Ouellet, Y. Ouellet, C. Richard, M. Labarre, B. Wittenberg, J. Wittenberg, and M. Guertin, *Proc. Natl. Acad. Sci.* **99**, 5902 (2002).
- [194] R.F. Eich, T. Li, D.D. Lemon, D.H. Doherty, S.R. Curry, J.F. Aitken, A.J. Mathews, K.A. Johnson, R.D. Smith, J. George, N. Phillips, and J.S. Olson, *Biochem.* **35**, 6976 (1996).
- [195] M. Angelo, A. Hausladen, D. Singel, and J.S. Stamler, *Methods in Enzymology* **436**, 131 (2008).
- [196] S. Goldstein, G. Merenyi, and A. Samuni, *J. Am. Chem. Soc.* **126**, 15694 (2004).
- [197] E.T. Yukl, S. de Vries, and P. Moenne-Loccoz, *J. Am. Chem. Soc.* **131**, 7234 (2009).
- [198] S. Mishra and M. Meuwly, *J. Am. Chem. Soc.* **132**, 2968 (2010).
- [199] L.M. Blomberg, M.R.A. Blomberg, and P.E.M. Siegbahn, *J. Biol. Inorg. Chem.* **9**, 923 (2004).
- [200] A. Crespo, M.A. Martí, S.G. Kalko, A. Morreale, M. Orozco, J.L. Gelpi, F.J. Luque, and D.A. Estrin, *J. Am. Chem. Soc.* **127**, 4433 (2005).
- [201] P.R. Gardner, *J. Inorg. Biochem.* **99** (1), 247 (2005).
- [202] T.S. Kurtikyan and P.C. Ford, *Chem. Commun.* **46**, 8570 (2010).
- [203] H. Ouellet, L. Juszczak, D. Dantsker, U. Samuni, Y.H. Ouellet, P.Y. Savard, J.B. Wittenberg, B.A. Wittenberg, J.M. Friedman, and M. Guertin, *Biochem.* **42**, 5764 (2003).
- [204] S. Mishra and M. Meuwly, *Biophys. J.* **99** (12), 3969 (2010).
- [205] P.A. Cazade and M. Meuwly, submitted (2012).
- [206] J. Stewart, *J. Mol. Struct.* **401**, 195 (1987).
- [207] Y. Yang, H. Yu, D. York, M. Elstner, and Q. Cui, *J. Chem. Theory. Comput.* **4**, 2067 (2008).
- [208] Y.T. Chang and W.H. Miller, *J. Phys. Chem.* **94**, 5884 (1990).
- [209] Y. Kim, J. Corchado, J. Villa, J. Xing, and D. Truhlar, *J. Chem. Phys.* **112** (6), 2718 (2000).
- [210] J.W. Ponder, C. Wu, P. Ren, V.S. Pande, J.D. Chodera, M.J. Schnieders, I. Haque, D.L. Mobley, D.S. Lambrecht, R.A. DiStasio, M. Head-Gordon, and G.N.I. Clark, *J. Phys. Chem. B* **114** (8), 2549 (2010).
- [211] C. Kramer, P. Gedeck, and M. Meuwly, *J. Comp. Chem.* in print, DOI: 10.1002/jcc.22996 (2012).



## Paper II

*In this paper we performed computational simulations to explain finding in the pioneering experiments of Prof. Stefan Willitsch. I was involved in the generation of the global potential energy surface for the charge-transfer reaction  $N_2^+ + N_2 \longrightarrow N_2 + N_2^+$ .*





## FRONTIERS ARTICLE

## State-selected ion–molecule reactions with Coulomb-crystallized molecular ions in traps

Xin Tong, Tibor Nagy, Juvenal Yosa Reyes, Matthias Germann, Markus Meuwly\*, Stefan Willitsch\*

Department of Chemistry, University of Basel, Klingelbergstrasse 80, 4056 Basel, Switzerland

## ARTICLE INFO

## Article history:

Available online 27 June 2012

## ABSTRACT

State-selected Coulomb-crystallized molecular ions were employed for the first time in ion–molecule reaction studies using the prototypical charge-transfer process  $N_2^+ + N_2 \rightarrow N_2 + N_2^+$  as an example. By preparing the reactant ions in a well-defined rovibrational state and localizing them in space by sympathetic cooling to milliKelvin temperatures in an ion trap, state- and energy-controlled reaction experiments with sensitivities on the level of single ions were performed. The experimental results were interpreted with quasi-classical trajectory simulations on a six-dimensional potential-energy surface which provided detailed insight into translation-to-rotation energy transfer occurring during charge transfer between  $N_2$  and  $N_2^+$ .

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

Achieving full control over the internal quantum states as well as the collision energy of the reaction partners to probe the detailed state-to-state dynamics and energy dependence of chemical processes has been one of the long-standing aspirations in gas-phase chemical studies. In the past, a range of experimental methods has been developed to address these objectives, e.g., internal cooling of the reactants in supersonic expansions, experiments with crossed molecular beams at variable relative velocities and laser preparation of the reactants in specific quantum states (see, e.g., Refs. [1–3] and literature cited therein). Very recently, new techniques have been established which allow the preparation of molecules at extremely low translational temperatures <1 K and at the same time enable an unprecedented degree of control over their kinetic energy [4–6]. In conjunction with the simultaneous preparation of the molecules in well-defined rotational–vibrational and even hyperfine states, ‘cold molecules’ methods are now starting to pave the way for studies of collisional processes and chemical reactions in new physical regimes and at levels of detail and sensitivity which have not been possible before [7–11].

Whereas most efforts have so far concentrated on neutral molecules, techniques for cooling and controlling molecular ions have recently made impressive progress as well. From an experimental perspective, the sensitivity of the ion motion to weak stray electric fields in the apparatus renders the precise control of their kinetic energy difficult. This problem can be overcome by trapping the ions and cooling them sympathetically to milliKelvin temperatures

by the interaction with co-trapped laser-cooled atomic ions [12]. Under these conditions, the ions localize in space to form ordered structures usually referred to as ‘Coulomb crystals’ in which it is possible to observe, manipulate and address single ions [13,6].

Moreover, the preparation of molecular ions in well-defined quantum states is experimentally challenging and indeed only a handful of reaction studies with rotationally state-selected ions have been reported thus far (see, e.g., [14–18]). Only very recently it has become possible to prepare Coulomb-crystallized molecular ions in well-defined internal quantum states [19–22], so that a simultaneous control over the kinetic energies, positions and internal states of the ions can now be achieved.

In the present article, we report the first study of a chemical reaction with rovibrationally state-selected Coulomb-crystallized molecular ions. We investigate the prototypical symmetric charge-transfer (CT) reaction  $N_2^+ + N_2 \rightarrow N_2 + N_2^+$  between state-prepared sympathetically-cooled  $N_2^+$  ions and internally cold  $N_2$  molecules from a supersonic expansion at a well-defined collision energy.

The mechanism and kinetics of this reaction have been studied previously, see, e.g., Refs. [23–28] and references therein. At low energies, the reaction was found to proceed via a long-lived  $N_4^+$  reaction complex which forms at the collision (Langevin) rate [24,25,28]. Using isotopically labeled nitrogen ions, Frost et al. showed that near-thermal collisions between  $N_2^+$  ions and  $N_2$  molecules in their vibrational ground states lead to a symmetric CT with a rate constant of  $k = 4.2 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$  which amounts to one half of the collision rate constant [25]. This observation was rationalized in terms of a symmetric sharing of the charge between the two  $N_2$  moieties in the  $N_4^+$  reaction complex so that the probability for CT amounts to 50% upon its breakup. Vibration-to-vibration (V–V) and vibration-to-translation (V–T) energy transfer

\* Corresponding authors.

E-mail addresses: [m.meuwly@unibas.ch](mailto:m.meuwly@unibas.ch) (M. Meuwly), [stefan.willitsch@unibas.ch](mailto:stefan.willitsch@unibas.ch) (S. Willitsch).

occurring during CT have been investigated in this system both experimentally and theoretically, see Refs. [29,30,25–27] and references therein. Kato et al. observed multi-quantum vibrational-energy transfer at near-thermal collision energies which was rationalized in terms of energy redistribution within the strongly bound  $N_4^+$  reaction complex [26]. Also, evidence for vibration-to-rotation and translation (V–R, T) energy transfer was found.

To our knowledge, the present work is the first to achieve rotational state selection of the reactant ions and rotational resolution in the analysis of the product-ion state populations allowing to gain insight into the dynamics of translation-to-rotation (T–R) energy transfer occurring during CT. The Coulomb-crystal technology employed in the present study enabled us to monitor reactive collisions in small ensembles of 20–30 state-selected, spatially localized  $N_2^+$  ions with single-particle sensitivity. The experimental results were interpreted with quasi-classical trajectory calculations on a full-dimensional potential-energy surface (PES) which provided insight into the charge- and energy-transfer dynamics underlying this fundamental reaction.

## 2. Experimental methods

The experimental setup for generating quantum-state selected Coulomb-crystallized  $N_2^+$  ions (see Figure 1) has been described in detail in previous publications [19,20]. Briefly,  $N_2^+$  ions in the rovibronic ground state  $X^+ \ ^2\Sigma_g^+, v^+ = 0, N^+ = 0, F_1$  were prepared by a  $[2 + 1']$  resonance-enhanced threshold-photoionization sequence via the  $a' \ ^1\Sigma_g^+, v' = 0, J' = 2$  intermediate level of neutral  $N_2$  as shown in Figure 2a. Here,  $v^+$  and  $v'$  ( $N^+$  and  $J'$ ) stand for the vibrational (rotational) quantum numbers of the cationic ground and neutral intermediate states, respectively, and  $F_1$  denotes the spin-rotational component. The  $N_2^+$  ions were generated inside a linear quadrupole ion trap [13,6] from a collimated, pulsed supersonic molecular beam of pure  $N_2$  gas. The molecular beam was positioned at a distance of  $\approx 400 \mu\text{m}$  from the Coulomb crystal in order to prevent collisions between ions and neutrals during the loading phase. The rotational temperature of the  $N_2$  molecules in the beam was determined to be  $T_{\text{rot}} \approx 10 \text{ K}$  by resonance-enhanced multiphoton-ionization (REMPI) spectroscopy [20] corresponding to populations of 50%, 25%, 22%, and 3% in the  $J = 0, 1, 2, 3$  rotational states. The beam velocity was estimated to be  $\approx 787 \text{ m s}^{-1}$  from flow-dynamics models [31].

Immediately after their generation, typically 25 state-selected  $N_2^+$  ions were sympathetically cooled to translational temperatures

of  $\approx 10 \text{ mK}$  by the interaction with laser-cooled  $\text{Ca}^+$  ions to form bi-component Coulomb crystals [13,6].  $\text{Ca}^+$  ions were produced in the center of the trap by non-resonant photoionization of Ca atoms emanating from a Ca oven. The  $\text{Ca}^+$  ions were laser cooled on the  $4s \ ^2S_{1/2} \rightarrow 4p \ ^2P_{1/2}$  transition using diode-laser radiation at 397 nm. Another diode-laser beam at 866 nm was used to re-pump population on the  $3d \ ^2D_{3/2} \rightarrow 4p \ ^2P_{1/2}$  transition to close the laser cooling cycle. The resulting  $\text{Ca}^+/\text{N}_2^+$  bi-component crystals were imaged by collecting the spatially resolved laser-cooling fluorescence of the  $\text{Ca}^+$  ions using a microscope coupled to camera. In the images, the positions of the non-fluorescing molecular ions are visible as a dark region in the center of the crystals, see Figure 3a.

Reactive collisions between state-selected  $N_2^+$  ions and neutral  $N_2$  molecules were initiated by overlapping the molecular beam with the bi-component Coulomb crystal immediately after ion loading and sympathetic cooling. The collision energy  $E_{\text{col}} \approx 0.045 \text{ eV}$  in the present experiments was entirely dominated by the kinetic energy of the  $N_2$  molecules in the beam. Neglecting the small initial rotational excitation of neutral  $N_2$ , the maximum kinetic energy of the product ions cannot exceed the total kinetic energy of the reactants which is much smaller than the trap depth ( $>2 \text{ eV}$ ). Consequently, the product ions remained trapped and were sympathetically re-cooled into the Coulomb crystal. Because the product ions were chemically identical to the reactant ions, their presence manifested itself in a time-dependent increase of the population in rotationally excited states in the ensemble of sympathetically-cooled  $N_2^+$  ions. Vibrational excitation of the products was precluded on energetic grounds. Rate constants were determined by measuring the  $N_2^+$  spin-rotational state populations as a function of the effective time of reaction with  $N_2$  molecules from the beam and fitting the results to a kinetic model as detailed in Section 4.

The rotational-state populations of the  $N_2^+$  product ions were probed by optical pumping to vibrationally excited levels to promote CT reactions with Ar atoms which are energetically forbidden in the vibrational ground state (laser-induced charge-transfer (LICT) spectroscopy [19,20]). LICT was initiated by exciting transitions to selected rotational levels of the  $X^+ \ ^2\Sigma_g^+, v^+ = 0 \rightarrow A^+ \ ^2\Pi_u, v = 4$  state with a pulsed dye laser operating at 613 nm, see Figure 2b. Subsequent fluorescent decay lead to the population of vibrationally excited states  $v^+ \geq 1$  in the electronic ground state with a probability of 93%. The removal of vibrationally excited  $N_2^+$  ions by CT with Ar gas leaked into the chamber was directly

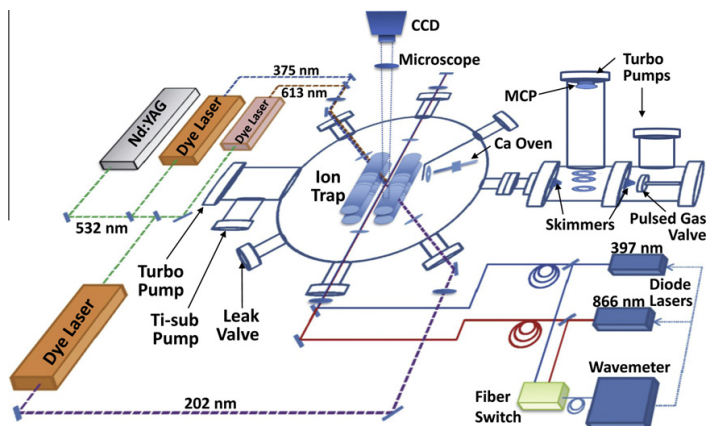
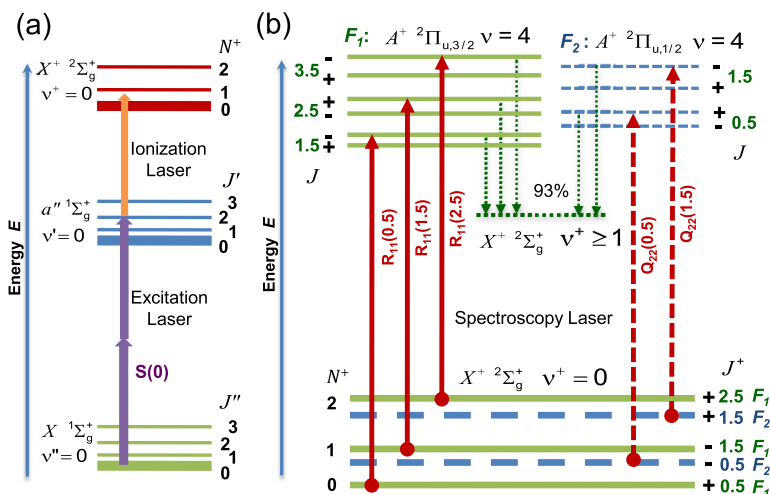


Figure 1. Schematic of the experimental setup.





**Figure 2.** (a) Resonance-enhanced threshold-photoionization scheme used to generate  $N_2^+$  ions in the  $X^+ 2^2\Sigma_g^+$ ,  $v^+=0$ ,  $N^+=0$ ,  $F_1$  spin-rovibrational ground state. (b) Laser-induced charge-transfer (LICT) scheme for probing the populations in the spin-rotational levels  $N^+$ ,  $F_{1,2}$  of the vibronic ground state of  $N_2^+$  produced by CT reactions. See text for details.

observed by the reduction of the core of non-fluorescing ions in the images, see Figure 3b. The number of  $N_2^+$  ions lost by LICT represented a direct measure of the population in the initially excited spin-rotational state. This number was determined by a comparison of the experimental fluorescence images with simulated images generated from molecular-dynamics simulations [19,20], see Figure 3a.

### 3. Theoretical methods

#### 3.1. Potential energy surfaces

The ground state six-dimensional potential-energy surface (PES) of the  $N_2-N_2^+$  system was computed at the UCCSD level [32,33] with Dunning's correlation consistent polarized valence basis set (cc-pVTZ) [34] using Gaussian03 [35]. Single-point calculations were performed for 5565 non-equivalent geometries with energies up to 1.9 eV relative to the minimum of the  $N_4^+$  complex on a non-equidistant rectangular grid including the N–N separations  $r_1$  and  $r_2$  in  $N_2$  and  $N_2^+$ , respectively, the center of mass distance  $R$  between the two diatomic molecules, and the angles  $\theta_1$ ,  $\theta_2$  and  $\phi$  (see the inset in Figure 6). The calculated well-depth ( $D_e$ ) of the complex is 1.254 eV, in good agreement with previous even higher-level theoretical estimates (1.26 eV at the (RCCSD (T)/VQZ (spdfg)) level of theory [36]). For comparison, the experimentally determined dissociation energy of  $N_4^+$  is  $D_0 = 1.06$  eV [37].

The global PES was represented by three surfaces, referred to as one 'bound' ( $N_4^+$ ) and two 'unbound' ones ( $N_2-N_2^+$  and  $N_2^+-N_2$ ). They were distinguished by the distance  $R$  and the localization of the majority of the charge. For  $R \leq 7.09a_0$  (3.75 Å) the system is considered to be bound whereas for  $R \geq 7.56a_0$  (4.0 Å) it is unbound. The intermediate region (7.09–7.56 $a_0$ ) represents the transition region during the dynamics, where the surfaces are connected by a smooth switching-function [38].

To allow (a) bond formation between any two atoms of  $N_2$  and  $N_2^+$  and (b) dissociation of the complex into either the charge-preserving or charge-transferring state at least 8 force fields (FFs) are needed for the bound state and 4 FFs are necessary for each of the unbound states. These FFs are related to each other through per-

mutation of their parameters. Within each set of FFs always the lowest energy surface is followed, except for when they are close in energy (within  $\approx 0.02$  eV) in which case they are smoothly joined by an energy difference-based switching function [38].

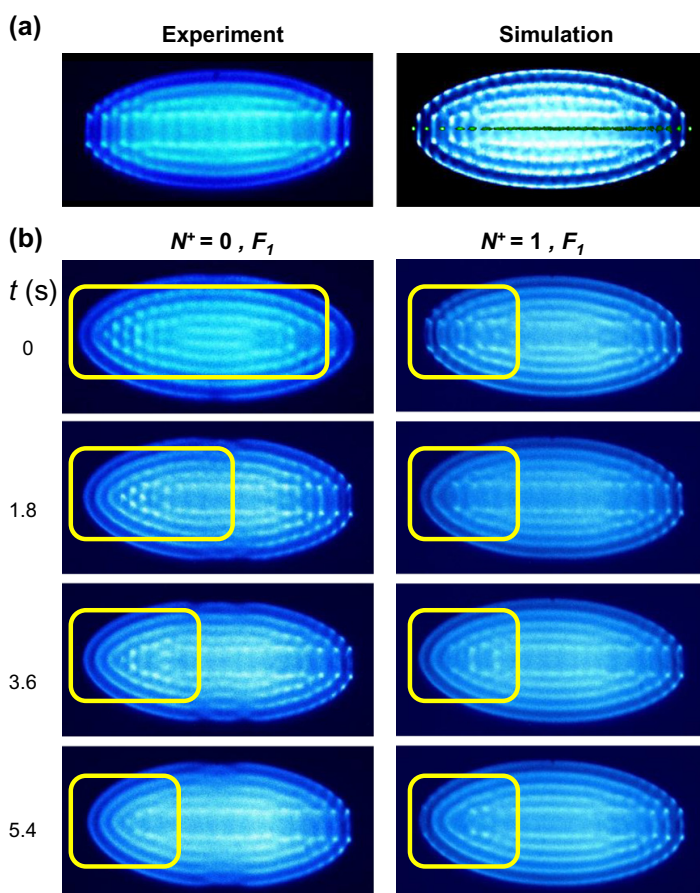
Morse and Lennard-Jones potentials were used for the bond stretches and the intermolecular interactions, respectively, for the unbound states. Morse parameters for the isolated molecules were determined from PES scans at the UCCSD level of theory. For the bound state, Morse potentials were used for all bonds of the complex and Lennard-Jones and electrostatic potentials between the edge atoms.

The force field parameters were fitted to the 5565 ab initio points with a simplex algorithm [39]. The ab initio energies for the unbound region were reproduced with a root-mean-square-deviation (RMSD) of  $\approx 0.013$  eV. It was considerably more difficult to obtain a good fit for the bound state due to strong angular dependence of the potential. The final fit had an RMSD of  $\approx 0.06$  eV by using two minimal symmetrized sets of FFs (twice 8, altogether 16 FFs). Figure 6 shows an example of the quality of the fit. The well depth of the bound state was exactly reproduced by the global surface. In the following, the bound state FFs are numbered 1–16, whereas those for the unbound states are labeled 17–20 and 21–24, respectively.

#### 3.2. Molecular-dynamics (MD) simulations

For the quasi-classical trajectory calculations a code with provisions for adiabatic reactive MD (ARMD) [40] was used. The Hamiltonian equations of motion were solved in Cartesian coordinates using the adaptive timestep Modified Extended Backward Differentiation Formulas method (MEBDFSO) [41].

The dynamics was initiated in the unbound state (FFs 21–24). Between center-of-mass separations of  $R = 7.56a_0$  and  $R = 7.09a_0$  the momentarily active unbound surface was smoothly switched to the bound FFs (numbers 1–16). The complex was considered to be formed for  $R \leq 7.09a_0$  and the dynamics was continued in the bound state. When a separation of  $R = 7.56a_0$  was reached again, it was determined which unbound FF was lowest in energy at the corresponding geometry and the dynamics was followed on this FF by smoothly switching from the bound state. Following the



**Figure 3.** (a) Experimental false-color fluorescence image and molecular-dynamics simulation of a  $\text{Ca}^+/\text{N}_2^+$  bicomponent Coulomb crystal before reaction. The spatial distribution of the non-fluorescing molecular ions has been made visible in green in the simulated image. (b) LICT experiments probing the populations in the  $\text{N}_2^+$   $N^+ = 0, F_1$  and  $N^+ = 1, F_1$  spin-rotational states. The panels show fluorescence images obtained after LICT as a function of the effective time of reaction  $t$  with neutral  $\text{N}_2$  molecules from the molecular beam. The LICT efficiency (corresponding to the number of non-fluorescing ions removed from the crystal in the boxed areas) decreases in  $N^+ = 0$  and increases in  $N^+ = 1$  with increasing reaction time, indicating the generation of rotationally excited  $\text{N}_2^+$  ions from the CT reaction  $\text{N}_2^+ + \text{N}_2$ . See text for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

dynamics further, the system either returned to the bound state (recrossing) or it decayed into  $\text{N}_2 + \text{N}_2^+$  (charge preserving, FFs 21–24) or  $\text{N}_2^+ + \text{N}_2$  (charge transferring, FFs 17–20). The dynamics was followed until either the lifetime of the bound state exceeded 100 ps or the initial separation of the two fragments was reached again. Once the complex survives a few vibrational periods, energy is expected to have randomized completely. Our results show that if the complex survives for at least 0.4 ps the correlation between the reactant and product states is negligible.

### 3.3. Analysis of final states

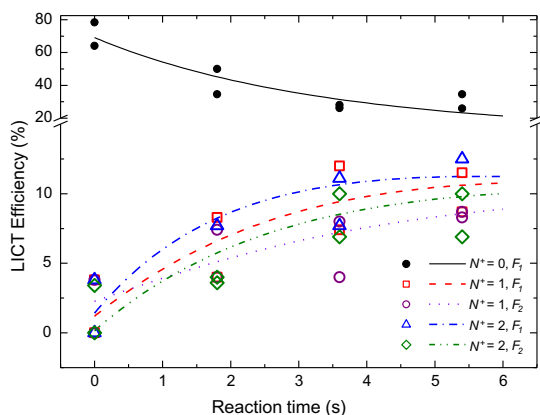
As a consequence of using classical dynamics, products are formed with rovibrational energies and angular momenta which correspond to fractional quantum numbers. Also, products having less than zero-point vibrational energy (ZPE) can be formed. This ZPE leakage is a shortcoming of quasi-classical simulations and various methods were proposed for either avoiding or correcting it [42–44]. In the present study, only those trajectories were analyzed further for which the total ZPE of the two product molecules was conserved within  $\pm 10\%$ .

## 4. Results and discussion

### 4.1. Product rotational-state distributions

Figure 3b shows LICT measurements probing the populations in the  $N^+ = 0, F_1$  and  $N^+ = 1, F_1$  spin-rotational states as a function of the effective time of reaction with neutral  $\text{N}_2$  molecules from the molecular beam. For the  $N^+ = 0$  state, a marked decrease of the LICT efficiency was observed with increasing reaction time, indicating the removal of the initially prepared  $\text{N}_2^+, N^+ = 0$  ions by CT with  $\text{N}_2$ . Conversely, an increase in the LICT efficiency out of the  $N^+ = 1$  state was observed, indicating the generation of ions in rotationally excited states as a consequence of the reactive collisions.

LICT measurements on the population of the  $N^+ = 0, 1, 2 F_{1,2}$  levels as a function of the reaction time are shown in Figure 4. Two independent LICT measurements were performed for each spin-rotational level at four different effective reaction times. As a general trend, a decrease of the population in  $N^+ = 0$  in favor of an increase of the population in  $N^+ \geq 1$  was observed. This result indicates that product ions over the whole range of rotation-



**Figure 4.** LICT efficiencies reflecting the populations in the lowest five spin-rotational levels of  $N_2^+$  as a function of the reaction time with neutral  $N_2$  molecules. Over the course of the reaction, the initial ensemble of state-selected  $N_2^+$  ions is replaced by rotationally excited product ions. The lines are a fit of the data to the kinetic model shown in Figure 5a.

ally excited states probed in the present study were generated as a consequence of CT between  $N_2^+ + N_2$ . Because the neutral products leave the trap, no conclusions on their rotational-state distribution can be drawn.

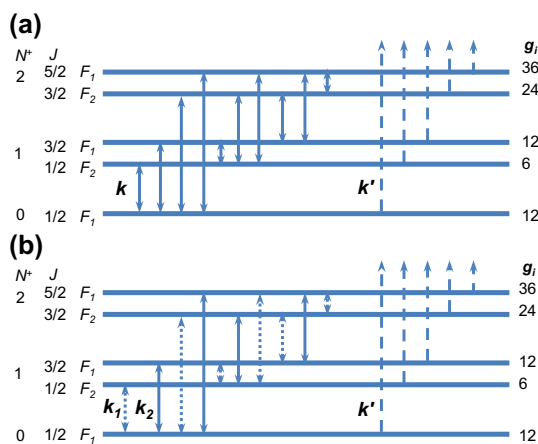
The observed redistribution of population in the ensemble of Coulomb-crystallized  $N_2^+$  ions is the result of a sequence of CT reactions with neutral molecules from the beam. The initially state-selected  $N_2^+$  ions undergo CT collisions with rotationally cold  $N_2$  molecules. The resulting  $N_2^+$  product ions are sympathetically cooled into the Coulomb crystal whereupon they in turn can undergo collisions. In this way, the ensemble of originally state-selected  $N_2^+$  ions in the Coulomb crystal is lost and replaced with rotationally excited product ions.

Close inspection of the data in Figure 4 reveals two important details. First, already in the beginning of the experiment when most of the reactive collisions occur with ions in  $N^+ = 0$ , the generation of ions in all rotational levels probed in the present study was observed. This observation suggests that the production of product ions is feasible over a broad range of rotational states by reactions with  $N_2^+$  in  $N^+ = 0$ . The relevant rates appear to be of a similar magnitude, at least over the range of product states probed in the present study. Second, for a specific rotational level  $N^+$ , there appears to be a slight preference for the generation of the  $F_1$  spin-rotational component in comparison to the  $F_2$  component.

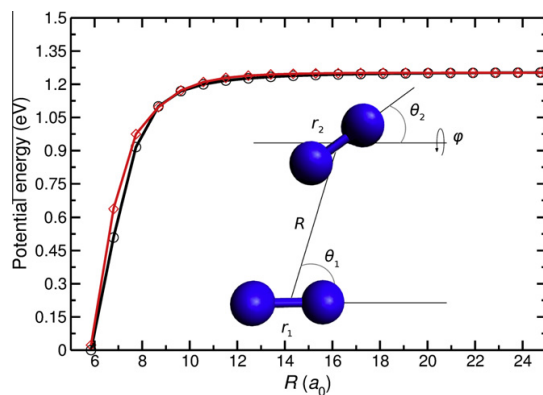
#### 4.2. Kinetics

The data were analyzed in terms of kinetic models taking into account CT collisions of the initially prepared  $N_2^+$  ( $N^+ = 0$ ) reactant ions as well as consecutive reactions of the product  $N_2^+$  ions, see Figure 5. Because the number density of neutral  $N_2$  in the molecular beam is much larger than the number density of the  $N_2^+$  ions in the trap, pseudo-first-order kinetics was assumed in both models.

Based on the observation that all product rotational states probed in the present study ( $N^+ \leq 2$ ) seem to be produced with similar rates, it was assumed in a first model that all processes of the form  $N_2^+(N^+) + N_2(J) \rightarrow N_2(\bar{J}) + N_2^+(\bar{N}^+)$  (bars indicate the rotational quantum numbers after the decay of the reaction complex) occur with the same  $k$ , see Figure 5a. The complete loss of population from the manifold of states with  $N^+ \leq 2$  was taken into account by an effective loss rate constant  $k'$ . In the absence of de-



**Figure 5.** Schematic of kinetic models used to fit the redistribution of rotational populations in the ensemble of sympathetically-cooled  $N_2^+$  ions as a consequence of CT collisions. Model assuming (a) an uniform rate constant  $k$  and (b) different rate constants  $k_1$  and  $k_2$  for processes connecting unlike and like spin-rotational levels  $F_{1,2}$  in the reactant and product ions, respectively. See text for details.  $g_i$  denotes the degeneracies of the spin-rotational levels including nuclear-spin statistical weights.



**Figure 6.** Comparison of the UCCSD/cc-pVTZ (black) and fitted energies (red) for linear structures with  $r_1 = r_2 = 2.08a_0$  as a function of  $R$ . The inset shows the relevant coordinates scanned in the ab initio calculations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tailed information on the neutral product state distribution,  $k$  and  $k'$  represent effective rate constants averaged over all neutral product states. The spin-rotational level populations  $n_i$  were obtained from the set of rate equations

$$\frac{dn_i}{dt} = k \sum_{j \neq i} (g_j n_j - g_i n_i) - k' n_i, \quad (1)$$

where  $g_{ij}$  represent degeneracy factors (see Figure 5). The initial state populations  $n_i(t=0)$  were also treated as a fit parameter, accounting for the uncertainty in determining the starting time of the measurement following the re-alignment of the molecular beam after ion loading (see Section 2). The fit of this model to the experimental data yielded the pseudo-first-order rate constants  $k = 0.0023(4) \text{ s}^{-1}$  and  $k' = 0.04(3) \text{ s}^{-1}$ . Because the density of neutral molecules in the molecular beam is not precisely known, it was not possible to obtain second-order rate constants from these

results. Previous rate measurements with thermal samples of  $N_2$  gas leaked into the chamber [20] were consistent with the result obtained by Frost et al. [25] that the total CT rate amounts to one half of the Langevin collision rate  $k_L = 8.3 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$ .

The time-dependent populations of the  $N^+ = 0, 1, 2, F_{1,2}$  states computed with the kinetic model using the fitted parameters are shown in Figure 4. The agreement between calculated and experimental level populations is satisfactory given the uncertainty of the measured level populations. The agreement vindicates the assumption that the state-specific rate constants do not (or only weakly) depend on the rotational state of the product ion within the uncertainty limits of the present measurement and the range of states studied. The kinetic model also reproduces the preferential generation of ions in the  $F_1$  compared to the  $F_2$  spin-rotational components of the same  $N^+$  state suggesting that this effect is caused by the higher statistical weight of the  $F_1$  components.

To check the validity of this model, its main simplification, i.e., the restriction to a ‘universal’ state-to-state rate constant  $k$ , was relaxed. In a second kinetic model (see Figure 5b), different rate constants  $k_1$  and  $k_2$  for reactions connecting states with different and like spin-rotational labels  $F_{1,2}$ , respectively, were assumed to account for the preferential production of product ions in the  $F_1$  levels. The fit yielded rate coefficients  $k_1 = 0.0023(9) \text{ s}^{-1}$ ,  $k_2 = 0.0022(7) \text{ s}^{-1}$  and  $k' = 0.04(3) \text{ s}^{-1}$ . The results obtained for  $k_1$  and  $k_2$  agree with each other and with the value of  $k$  within the uncertainty limits, supporting the conclusion that the increased production rates of ions in the  $F_1$  levels is indeed solely caused by the higher statistical weights associated with these channels.

#### 4.3. Trajectory calculations

Quasi-classical trajectory calculations were performed to interpret the experimental findings and gain further insight into the energy-transfer mechanisms involved in the reaction. Initial states for the trajectory calculations correspond to semi-classically quantized states of rotating Morse oscillators [45]. In the experiment the molecules are exclusively in their ground vibrational state ( $\nu = 0$ ). The initial rotational angular momentum  $N^+$  of the  $N_2^+$  ions was set to zero as prepared in the experiment whereas that of  $N_2$  was randomly sampled according to the experimentally measured distribution (see Section 2). Initial coordinates and momenta were generated by randomly sampling the phase-space distribution of the rotating Morse-oscillators and the spatial orientation of each molecule was taken from a uniform distribution within  $4\pi$  steradian. The magnitude of the angular velocity for a given vibrational phase was set to the angular momentum calculated from  $L = (J(J+1))^{1/2}\hbar$  and its direction was randomly sampled from a uniform distribution in the plane perpendicular to the axis of the molecule.

Impact parameters  $b$  were calculated between 0 and  $15a_0$  in steps of  $1a_0$ , and in steps of  $0.2a_0$  between 12 and  $13a_0$  where the opacity function drops to zero (see below). The initial separation of the center of mass of the two molecules was  $20 \text{ \AA}$  ( $\approx 38a_0$ ), at which distance the intermolecular interaction energy was less than  $1.4 \times 10^{-4} \text{ eV}$ . According to the Langevin–Gioumousis–Stevenson (LGS) model of ion–molecule reactions the rate constant does not depend on the relative velocity of the partners [46]. Therefore, the relative velocity of the colliding molecules was set to the calculated beam flow rate (787 m/s) without dispersion.

To validate the simulations, the probability of complex formation and the probability of CT were determined for a range of impact parameters  $b$  based on 500–1000 trajectories at each value (see Figure 7). According to our quasi-classical model, the probability for CT was found to be 50% (within the statistical uncertainty) up to impact parameters  $b = 9a_0$ . The CT probability starts drop-

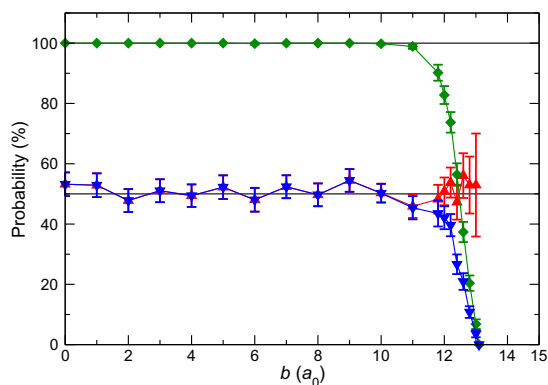
ping slowly between 10 and  $12a_0$ , and above  $12a_0$  it decays steeply to zero around  $13a_0$ . The probability of complex formation is 100% up to  $9a_0$  and drops proportionally to the CT probability. Once the complex is formed, the probability for CT is  $\approx 50\%$  regardless of the impact parameter. This finding is expected based on symmetry arguments and considering that rapid charge-redistribution takes place before the complex decays. It is also in agreement with previous experimental results [25].

From the integrated opacity function an integral cross section for CT of  $\sigma_{\text{tot}} = (243 \pm 19)a_0^2$  was obtained. Multiplication with the relative velocity ( $787 \text{ m s}^{-1}$ ) yields  $k = (5.36 \pm 0.42) \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$  for the second order rate coefficient for CT. This value compares favorably with previous experimental results of  $4.24 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$ ,  $6.6 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$  and  $5.0 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$  obtained for the CT reaction between  $^{15}N_2^+(\nu = 0)$  and  $^{14}N_2(\nu = 0)$  [25,47,48].

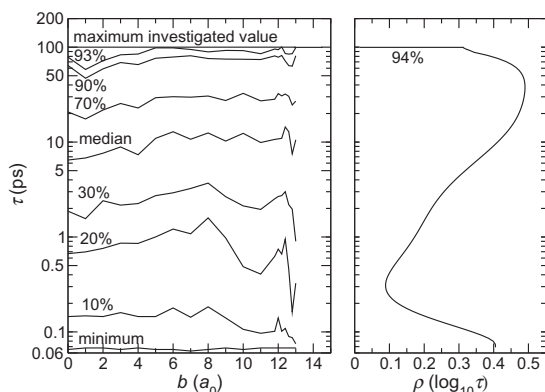
The lifetime of the complex was determined as the time difference between the last and the first crossing of the surface separation radius  $7.09a_0$  (see Section 3.1). As Figure 8 suggests, the median lifetime of  $N_4^+$  is  $\approx 10 \text{ ps}$  and does not or only slightly depends on  $b$ . Lifetimes up to 100 ps are found for 94% of the trajectories which qualitatively agree with previous estimates [24,49].

#### 4.4. Translation-to-rotation energy transfer

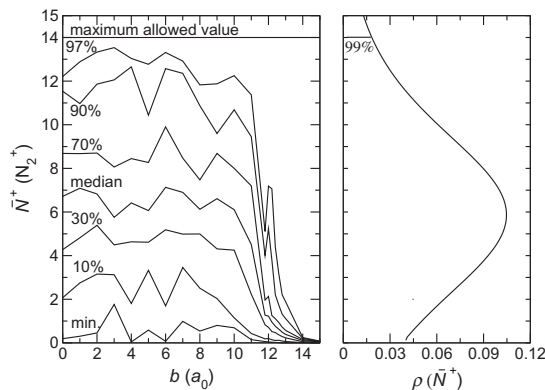
The salient quantity which can be extracted from the trajectories to aid in the interpretation of the experimental results is the distribution of rotational angular momentum  $\rho(\bar{N}^+)$  of the  $N_2^+$  products after complex formation and decay. Figure 9 reports  $\rho(\bar{N}^+)$  and suggests that for the majority (97%) of cases  $\bar{N}^+ \leq 13$ . On energetic grounds, a total of 0.045 eV is available in collisional energy for the reactants, and the initial rotational quantum number of  $N_2$  can be as high as 3 (corresponding to a rotational energy of  $\approx 0.003 \text{ eV}$ ), which implies  $\bar{N}^+ \leq 14$ . The  $N_2^+$  products are predominantly formed in excited rotational states, which supports the experimental findings of depletion of the  $N_2^+$  ground state population caused by reactive collisions with the molecules in the beam. Note that the theoretical product state distribution cannot be compared directly with the experimental findings reported in Figure 4, because the  $N_2^+$  ions probed in the experiment result from a sequence of CT reactions, whereas the simulations only reflect single collision events.



**Figure 7.** Opacity functions for complex formation (green diamonds) and charge transfer (blue triangles), and probability curve for charge transfer (red triangles) if the complex is formed. Error bars correspond to a  $2\sigma$  standard deviation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



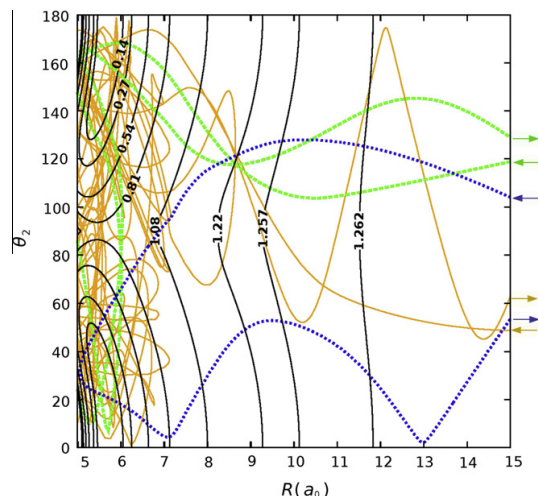
**Figure 8.** Percentiles representing the distribution of lifetimes  $\tau$  for the  $N_4^+$  complex as a function of the impact parameter  $b$  (left panel) and integrated probability density  $\rho$  (right panel). The median lifetime is around 10 ps, the complex decays within 100 ps in  $\approx 94\%$  of all observed cases.



**Figure 9.** Percentiles representing the rotational-state distributions for the  $N_2^+$  products as a function of the impact parameter  $b$  (left panel) and integrated probability density  $\rho(N_2^+)$  after breakup of the reaction complex (right panel). The products are predominantly formed in rotationally excited states.

However, the trajectory calculations clearly reproduce the rotational excitation of the products observed in the experiment and reveal the mechanisms of the underlying T–R energy transfer occurring during CT. The reactant molecules approach each other in a random orientation and are accelerated towards the deep potential well of the linear reaction complex. The torque exerted while forcing the reactants towards a linear configuration excites a counter-rotation of the  $N_2$  moieties which results in highly excited bending and torsional vibrations of the complex. Moreover, anharmonic couplings lead to a practically complete redistribution of the available energy over all vibrational degrees of freedom during the lifetime of the complex. Upon its breakup, large-amplitude bending and torsional vibrations are converted into product rotations. Representative trajectories, projected on the PES, are shown in Figure 10 to illustrate this effect.

Close inspection of the trajectories reveals the likely presence of additional processes resulting in the rotational excitation of the products, including rotation–vibration coupling and recrossing which all may contribute to the final state distribution. Under the present experimental conditions, the orbital angular momen-



**Figure 10.** Projection of the PES onto the  $R$  and  $\theta_z$  coordinate (see Figure 6) with  $r_1 = r_2 = 2.08a_0$ ,  $\theta_1 = 0$ , and  $\phi = 0$ . Contour lines are drawn between 0 and 1.262 eV. Three typical trajectories leading to rotational excitation of the products are reported. The arrows indicate the direction of motion (incoming and outgoing). The complex lifetimes amount to 87 fs (blue), 341 fs (green) and 6604 fs (orange). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tum of complex-forming collisions is computed to be  $L \lesssim 140h$  which is available for conversion into rotational motion of the reaction complex. Coriolis forces may lead to a coupling of the complex rotation to its internal motion, providing another mechanism for the transfer of angular momentum to the fragments. Furthermore, the trajectories shown in Figure 10 suggest that a wide range of scenarios is possible. They include direct mechanisms, illustrated by the blue trajectory, and can range to bound states with almost full redistribution (randomization, i.e. IVR, also recrossing) of the internal energy as shown for the orange trace. This is reminiscent of the situation recently encountered in the vibrationally induced photodissociation of sulfuric acid where all regimes from prompt reaction to complete IVR were found, depending on the amount of internal energy made available to the molecule [50].

## 5. Summary and conclusions

In the present study we have demonstrated for the first time ion–molecule reaction studies with state-selected Coulomb-crystallized molecular ions using the symmetric CT reaction  $N_2^+ + N_2$  as an example. By rotational state selection of the reactant ions and their localization in space by sympathetic cooling in an ion trap, their internal and translational motions were completely controlled. By simultaneously cooling the neutral co-reactants to the lowest rotational states and collimating their kinetic-energy distribution in a supersonic molecular beam, ion–molecule reactions were performed with high degree of control over both the collision energy and internal states of the reaction partners. The analysis of the product-ion rotational state distribution yielded information on T–R energy transfer occurring during CT. The experimental results were analyzed and interpreted by making contact with quasi-classical trajectory simulations on a full-dimensional PES. The simulations reproduced the experimental findings and yielded important insights at an atomistic level into the mechanisms underlying T–R energy transfer.



## References

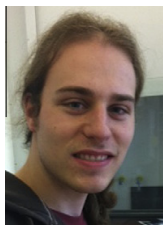
- [1] K. Liu, *J. Chem. Phys.* 125 (2006) 132307.
- [2] A. Teslja, J.J. Valentini, *J. Chem. Phys.* 125 (2006) 132304.
- [3] X. Yang, *Phys. Chem. Chem. Phys.* 13 (2011) 8112.
- [4] L.D. Carr, D. DeMille, R.V. Krems, J. Ye, *New J. Phys.* 11 (2009) 055049.
- [5] M. Schnell, G. Meijer, *Angew. Chem. Int. Ed.* 48 (2009) 6010.
- [6] S. Willitsch, *Int. Rev. Phys. Chem.* 31 (2012), <http://dx.doi.org/10.1080/0144235X.2012.667221>.
- [7] J.J. Gilijamse, S. Hoekstra, S.Y.T. van de Meerakker, G.C. Groenenboom, G. Meijer, *Science* 313 (2006) 1617.
- [8] S. Willitsch, M.T. Bell, A.D. Gingell, S.R. Procter, T.P. Softley, *Phys. Rev. Lett.* 100 (2008) 043203.
- [9] S. Ospelkaus et al., *Science* 327 (2010) 853.
- [10] M.H.G. de Miranda et al., *Nat. Phys.* 7 (2011) 502.
- [11] F.H.J. Hall, M. Aymar, N. Bouloufa-Maafa, O. Dulieu, S. Willitsch, *Phys. Rev. Lett.* 107 (2011) 243202.
- [12] K. Mølhave, M. Drewsen, *Phys. Rev. A* 62 (2000) 011401.
- [13] S. Willitsch, M.T. Bell, A.D. Gingell, T.P. Softley, *Phys. Chem. Chem. Phys.* 10 (2008) 7200.
- [14] S.R. Mackenzie, T.P. Softley, *J. Chem. Phys.* 101 (1994) 10609.
- [15] Th. Glenwinkel-Meyer, D. Gerlich, *Isr. J. Chem.* 37 (1997) 343.
- [16] R.J. Green, J. Qian, H.-T. Kim, S.L. Anderson, *J. Chem. Phys.* 113 (2000) 3002.
- [17] R.A. Dressler et al., *J. Chem. Phys.* 125 (2006) 132306.
- [18] L. Paetow, F. Unger, W. Beichel, G. Frenking, K.-M. Weitzel, *J. Chem. Phys.* 132 (2010) 174305.
- [19] X. Tong, A.H. Winney, S. Willitsch, *Phys. Rev. Lett.* 105 (2010) 143001.
- [20] X. Tong, D. Wild, S. Willitsch, *Phys. Rev. A* 83 (2011) 023415.
- [21] P.F. Staunum, K. Højbjerg, P.S. Skyt, A.K. Hansen, M. Drewsen, *Nat. Phys.* 6 (2010) 271.
- [22] T. Schneider, B. Roth, H. Duncker, I. Ernstring, S. Schiller, *Nat. Phys.* 6 (2010) 275.
- [23] B. Friedrich, S.L. Howard, A.L. Rockwood, W.E. Trafton, D. Wen-Hu, J.H. Futrell, *Int. J. Mass Spectrom.* 59 (1984) 203.
- [24] P.A.M. van Koppen, M.F. Jarrold, M.T. Bowers, L.M. Bass, K.R. Jennings, *J. Chem. Phys.* 81 (1984) 288.
- [25] M.J. Frost, S. Kato, V.M. Bierbaum, S.R. Leone, *J. Chem. Phys.* 100 (1994) 6359.
- [26] S. Kato, V.M. Bierbaum, S.R. Leone, *J. Chem. Phys.* 102 (1998) 6659.
- [27] K. Sohlberg, *Chem. Phys.* 246 (1999) 307.
- [28] M.J. Frost, C.R.J. Sharpe, *Phys. Chem. Chem. Phys.* 3 (2001) 4536.
- [29] K.B. McAfee, C.R. Szmamda, R.S. Hozack, *J. Phys. B: At. Mol. Phys.* 14 (1981) L243.
- [30] K. Sohlberg, J. Futrell, K. Szalewicz, *J. Chem. Phys.* 94 (1991) 6500.
- [31] D.R. Miller, in: G. Socles (Ed.), *Atomic and Molecular Beam Methods*, vol. 1, Oxford University Press, New York, 1988, p. 14.
- [32] G. Purvis, R. Bartlett, *J. Chem. Phys.* 76 (1982) 1910.
- [33] A. Scheiner, G. Scuseria, J. Rice, T. Lee, H. Schaefer, *J. Chem. Phys.* 87 (1987) 5361.
- [34] T.H. Dunning, *J. Chem. Phys.* 90 (1989) 1007.
- [35] M.J. Frisch, et al., *Gaussian 03, Revision C.01*, Gaussian, Inc., Wallingford CT, USA, 2004.
- [36] C. Léonard, P. Rosmus, S. Carter, N.C. Handy, *J. Phys. Chem. A* 103 (1999) 1846.
- [37] K.-M. Weitzel, J. Mähner, *Int. J. Mass Spectrom.* 214 (2002) 175.
- [38] B.R. Johnson, *J. Chem. Phys.* 29 (1985) 1204.
- [39] J.A. Nelder, R. Mead, *Chem. Phys.* 7 (1965) 308.
- [40] J. Danielsson, M. Meuwly, *J. Chem. Theory Comput.* 4 (2008) 1083.
- [41] T.J. Abdulla, J.R. Cash, M.T. Diamantakis, *Comput. Math. Appl.* 42 (2001) 121.
- [42] J.M. Bowman, B. Gazdy, Q.Y. Sun, *J. Chem. Phys.* 91 (1989) 2859.
- [43] W.H. Miller, W.L. Hase, C.L. Darling, *J. Chem. Phys.* 91 (1989) 2863.
- [44] G.H. Peslherbe, W.L. Hase, *J. Chem. Phys.* 100 (1994) 1179.
- [45] R.N. Porter, L.M. Raff, W.H. Miller, *J. Chem. Phys.* 63 (1975) 2214.
- [46] G. Gioumousis, D.P. Stevenson, *J. Chem. Phys.* 29 (1958) 294.
- [47] N.G. Adams, D. Smith, *Astrophys. J. Lett.* 247 (1981) L123.
- [48] T.B. McMahon, P.G. Miasek, J.L. Beauchamp, *Int. J. Mass Spectrom.* 21 (1976) 63.
- [49] L.F. Phillips, *J. Chem. Phys.* 92 (1990) 6523.
- [50] J. Yosa, M. Meuwly, *J. Phys. Chem. A* 115 (2011) 14350.



**Tibor Nagy** studied Chemistry and Physics at Eötvös Lóránd University in Budapest (ELTE, Hungary) and Chemistry at Leeds University (UK), and he received his PhD in Chemistry from ELTE in 2009. Between 2009–2011, he was a postdoctoral research fellow at Galway University (Ireland) and ELTE. His research interests lie in the kinetics and dynamics of elementary chemical reactions and optimization and reduction of large chemical mechanisms. Currently he is working on atomistic simulations at the University of Basel (Switzerland).



**Juvenal Yosa Reyes** received his B.Sc. in Chemistry from Distrital University (Bogota, Colombia) and a M.Sc. in Biology and Bioinformatics from Javeriana University (Bogota, Colombia) in 2005. Currently he is working on his Ph.D. at the Department of Chemistry at the University of Basel (Switzerland). His research interests include molecular dynamics, catalysis and atomistic simulations of biological processes.



**Matthias Germann** graduated from the University of Zürich (Switzerland) in 2008. After a short research stay at the University of Oxford (UK) in 2009, he joined the Department of Chemistry at the University of Basel (Switzerland) in 2010 where he is currently working on his Ph.D. His research interests focus on the applications of sympathetically-cooled molecular ions in precision spectroscopy and quantum technology.



**Markus Meuwly** studied Physics and obtained a PhD in Physical Chemistry from the University of Basel (1997). After postdoctoral work with Prof. J.M. Hutson (Durham, UK) and Prof. M. Karplus (Strasbourg, Harvard) as a Swiss National Science Foundation Fellow he joined the Department of Chemistry of the University of Basel as a Swiss National Science Foundation Professor. Since 2006 he is Professor for Physical and Computational Chemistry. He also holds an Adjunct Research Professorship from Brown University (since 2009). His research interests include reactive processes in complex systems and the development of force fields for spectroscopic applications.



**Stefan Willitsch** graduated from the Eidgenössische Technische Hochschule (ETH) Zürich (Switzerland) and received his PhD from the Laboratory of Physical Chemistry of ETH in 2004. From 2004–2007, he was a Junior Research Fellow at the University of Oxford (UK). In 2007, he was appointed lecturer in physical chemistry at University College London (UK) and since 2008 he is an assistant professor at the Department of Chemistry at the University of Basel (Switzerland). His research interests lie in chemical and spectroscopic applications of cold molecules and ions.



**Xin Tong** graduated from the University of Science and Technology of China in 2000 and obtained his PhD from the University of York (UK) in 2005 with a thesis on the high-resolution photoelectron spectroscopy of cluster molecules. From 2005 to 2008, he worked as a postdoc at the Photon Science Institute of the University of Manchester and since 2008 at the Department of Chemistry at the University of Basel (Switzerland). His research interests involve quantum-state controlled sympathetically-cooled molecular ions and their reaction dynamics.



# Juvenal Yosa

## Curriculum Vitae

*"Science without passion is just a wasting time"*

### Education

- 2009–2013 **PhD. Physical Chemistry**, *University of Basel*, Basel-Switzerland.  
2002–2005 **MSc. Biology**, *Javeriana University*, Bogota-Colombia.  
1995–2000 **Bs. degree in Chemistry**, *Distrital University*, Bogota-Colombia.

### PhD Thesis

- Title *Atomistic Simulations of Reactive Processes from Gas to Condensed Phase.*  
Supervisor Prof. Dr. Markus Meuwly  
Description This thesis explore different computational methods to model chemical reactions, from *ab initio* to reactive force fields, using molecular dynamics scheme.

### MSc Thesis

- Title *Molecular Orbital differentiation of Agonist and Antagonist Activity in the Glycine<sub>B</sub>-iGluR-NMDA receptor.*  
Supervisor Dr. Leonardo Lareo  
Description A simple and consistent signature of agonists versus antagonist action, the energy of the lowest unoccupied molecular orbital (LUMO), in the iGluR-NMDA, allows to better understand the receptor activation.

### Publications

- **Molecular orbital differentiation of agonist and antagonist activity in the Glycine<sub>B</sub>-iGluR-NMDA receptor.** Juvenal Yosa Reyes, Orlando Acevedo, Mario Blanco and Leonardo Lareo. *J. Med. Chem.* 2009, 44 (7), 2960.

*Schonastrasse 22 – Basel (Switzerland)*

☎ +41-789565410 • ☎ +41-612673815 • ✉ [juvenal.yosa@unibas.ch](mailto:juvenal.yosa@unibas.ch)

🌐 <http://juvenalyosa.wix.com/juvenal>

- **Semiempirical methods for the rapid evaluation of frontier orbitals in the classification of agonists and antagonists of the NR1 subunit of the iGluR-NMDA receptors.** Juvenal Yosa Reyes, Diana Clavijo Buritica, Carlos Estevez Breton. *Universitas Scientiarum*. 2011, 16 (1), 5.
- **Vibrationally Induced Dissociation of Sulfuric Acid (H<sub>2</sub>SO<sub>4</sub>).** Juvenal Yosa Reyes and Markus Meuwly. *J. Phys. Chem. A* 2012, 115 (50), 14350.
- **Influence of Mg<sup>2+</sup> ions on the interaction between 3, 5-dicaffeoylquinic acid and HTLV-I integrase.** Angela Pena, Juvenal Yosa Reyes, Yesid Cuesta, Orlando Acevedo, Leonardo Lareo and Felipe Garcia. *Universitas Scientiarum* 2012, 17 (1), 5.
- **Atomistic simulations of reactive processes in the gas- and condensed-phase.** Pierre-Andre Cazade, Jing Huang, Juvenal Yosa Reyes, Jaroslaw Szymczak and Markus Meuwly. *Int. Rev. Phys. Chem.* 2012, 31 (2), 235.
- **State-selected ion-molecule reactions with Coulomb-crystallized molecular ions in traps.** Xin Tong, Tibor Nagy, Juvenal Yosa Reyes, Matthias Germann, Markus Meuwly and Stefan Willitsch. *Chem. Phys. Lett.* 2012, 547 (21), 1.
- **Multi-Surface adiabatic reactive molecular dynamics.** Tibor Nagy, Juvenal Yosa Reyes and Markus Meuwly *J. Chem. Theory. Comput.* 2014, DOI: 10.1021/ct400953f
- **Competitive Reaction Dynamics between intra- and intermolecular channels in vibrationally induced photodissociation in H<sub>2</sub>SO<sub>4</sub>.** Juvenal Yosa Reyes, Tibor Nagy and Markus Meuwly *J. Phys. Chem. A* submitted.

Schonastrasse 22 – Basel (Switzerland)

☎ +41-789565410 • ☎ +41-612673815 • ✉ juvenal.yosa@unibas.ch

🌐 <http://juvenalyosa.wix.com/juvenal>