

# Solvent, Structure, and Spectroscopy: the Role of Dynamics in Chemical Reactions and S-Nitrosylated Proteins

**Inauguraldissertation**

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Haydar Taylan Turan

2023

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel  
edoc.unibas.ch



This work is licensed under a Creative Commons  
Attribution-NonCommercial 4.0 International License.

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät  
auf Antrag von

Prof. Dr. Markus Meuwly

Prof. Dr. Anatole von Lilienfeld

Prof. Dr. Paolo Carloni

Basel, den 21. Juni 2022

Prof. Dr. Marcel Mayor

Dekan





---

## Abstract

Understanding the motions of atoms in a chemical system is core to chemistry. The dynamic motions of chemical events can occur at different time scales, ranging from a bond formation which occurs at the femtosecond timescale, or a displacement of a protein  $\beta$ -sheet which occurs over nanoseconds. Experimental characterization of such events at the atomistic level, especially in a condensed phase, requires refined temporal and spatial resolutions which still remains a challenge. Molecular Dynamic simulations can bring molecular-level insight into the dynamics and energetics of chemical processes for systems ranging from small molecules to proteins in the condensed phase.

The first part of the thesis emphasizes the chemical background of the conducted studies and the theoretical basis of the applied methods. In Chapter 3, the energetics and solvent distributions for the  $\text{NH}_3+\text{MeCl}$  and  $\text{Pyr}+\text{MeBr}$  reactions were investigated in explicit solvent (water, methanol, acetonitrile, benzene, cyclohexane) by means of reactive molecular dynamics simulations. In Chapter 4, The structural dynamics and vibrational spectroscopy of S-nitrosylation in the condensed phase are investigated for the methyl-capped cysteine model system and for myoglobin using conventional point charge and physically more realistic multipolar force fields for the -SNO group. In Chapter 5, the structural dynamics and hydration of the S-Nitrosylated KRAS protein with and without GDP binding are studied in the condensed phase. The detailed abstracts of each study are provided in their respective chapters. Finally, conclusions are drawn in Chapter 6.



---

## Acknowledgements

I would like to express my gratitude to Prof. Dr. Markus Meuwly who gave me an opportunity to do a Ph.D. in this great university and beautiful country. I would like to thank Prof. Dr. Anatole von Lilienfeld for accepting to be my second supervisor and Prof. Dr. Paolo Carloni who kindly accepted to act as my external examiner.

I also would like to thank my BSc. supervisor Prof. Dr. Nursel Acar, and my MSc. supervisor Prof. Dr. Viktorya Aviyente for guiding me towards academic research.

Thanks to all past and current members of the Meuwly group for all the good times we shared. I always enjoyed the chats we have at lunchtime and the beers we shared. I am grateful to Maryam, Marco, and Juan, I really enjoyed spending time with you. Many thanks to Eric and Juan for proofreading my thesis.

My warmest gratitude goes to my wonderful parents Aliye and Baki. You always stood behind me and I was very fortunate to have such a beautiful family.

Finally, I would like to thank Öyküm. I lost counting how many times we supported each other in these crazy 5 years. Knowing that you are there for me makes everything easier as I face life's challenges. Cheers to the next city we going to discover, the next cocktail we going to have, the next adventure we going to face.





# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Solvent Effects on Chemical Reactions and Proteins . . . . .	1
1.2	Post-translational Modifications . . . . .	3
1.3	PTM Detection by IR Spectroscopy . . . . .	5
<b>2</b>	<b>Theoretical Background</b>	<b>9</b>
2.1	Quantum Mechanical Methods . . . . .	9
2.1.1	Hartree-Fock Theory . . . . .	10
2.1.2	Møller Plesset Perturbation Theory . . . . .	11
2.1.3	Basis Sets . . . . .	13
2.2	Force Field . . . . .	15
2.2.1	Multipolar Force Field . . . . .	18
2.2.2	Multi-Surface Adiabatic Reactive Molecular Dynamics . . . . .	19
2.3	Molecular Dynamics Simulations . . . . .	21
2.3.1	Umbrella Sampling . . . . .	23
<b>3</b>	<b>Solvent Effects on the Menshutkin Reaction</b>	<b>25</b>
3.1	Abstract . . . . .	27
3.2	Introduction . . . . .	27
3.3	Computational Methods . . . . .	31
3.3.1	Molecular Dynamics Simulations . . . . .	31
3.3.2	Parametrization of the Force Field . . . . .	32
3.3.3	Umbrella Sampling . . . . .	34
3.3.4	Solvent Distribution . . . . .	34
3.4	Results and Discussion . . . . .	36
3.4.1	The Potential Energy Surfaces . . . . .	36
3.4.2	Umbrella Sampling Simulations and Potentials of Mean Force . . . . .	37
3.4.3	Solvent Distributions . . . . .	44
3.4.4	Solvent Energetics . . . . .	51

---

3.5	Conclusion . . . . .	57
<b>4</b>	<b>Spectroscopy, Dynamics and Hydration of S-Nitrosylated Myoglobin</b>	<b>61</b>
4.1	Abstract . . . . .	63
4.2	Introduction . . . . .	63
4.3	Computational Methods . . . . .	69
4.3.1	Molecular Dynamics . . . . .	69
4.3.2	Parametrisation of the Force Field . . . . .	70
4.3.3	Infrared Spectrum . . . . .	74
4.4	Results . . . . .	75
4.4.1	Dynamics and Spectroscopy of CysNO . . . . .	75
4.4.2	The Structural Dynamics and Spectroscopy of Mb-SNO . . . . .	79
4.4.3	Water Structure and Global Structural Changes . . . . .	89
4.5	Discussion and Conclusions . . . . .	99
<b>5</b>	<b>Spectroscopy, Dynamics and Hydration of S-Nitrosylated KRAS103</b>	
5.1	Abstract . . . . .	105
5.2	Introduction . . . . .	105
5.3	Computational Methods . . . . .	109
5.3.1	Molecular Dynamics . . . . .	109
5.3.2	Infrared Spectroscopy . . . . .	110
5.3.3	Dynamical Cross-Correlation Maps . . . . .	111
5.4	Results . . . . .	112
5.4.1	Water Structure . . . . .	112
5.4.2	Root Mean Square Fluctuation . . . . .	116
5.4.3	Structural Changes . . . . .	119
5.4.4	Infrared Spectrum . . . . .	127
5.5	Conclusion . . . . .	128
<b>6</b>	<b>Conclusion and Outlook</b>	<b>131</b>
	<b>Bibliography</b>	<b>135</b>
<b>A</b>	<b>MS-ARMD parameters for the Menshutkin reactions</b>	<b>165</b>
<b>B</b>	<b>Curriculum Vitae</b>	<b>169</b>





# Chapter 1

## Introduction

### 1.1 Solvent Effects on Chemical Reactions and Proteins

Most chemical reactions are carried out in the liquid phase, and, in the majority of these reactions, reactants and reagents are dissolved in a suitable medium to progress the reaction.<sup>1</sup> Although solvent-induced effects, such as modulation of reaction rates<sup>2</sup> and spectroscopic responses of solutes<sup>3</sup>, are frequently observed and discussed, identification of the underlying atomistic mechanisms of such effects is less studied. Solvents regulate chemical reactions in diverse manners. They can change the relative stabilization of the reactant, transition, and product state, exhibit entropic effects that can alter free energy barriers of reactions, or directly participate in the reaction steps and divert the reaction pathway.<sup>4</sup>

Organic solvents, broadly speaking, can be divided into two categories based on their polarity; polar and apolar. The relative static permittivity, or dielectric constant, of a solvent molecule, is usually proportional to its polarity.<sup>5</sup> Solvent molecules that are able to donate protons, usually via hydrogen bonding, are called protic polar solvents whereas solvents that do not have any hydrogen atoms bonded to an electronegative atom such as O, N, and F called aprotic polar solvents. The nature of the solvent determines the type of solute it can dissolve or the level of influence it can have on the progression of the chemical reaction. The stepwise  $S_N1$  mechanism reaction rate is known to significantly increase in protic

polar solvents whereas the concerted  $S_N2$  mechanism has higher rates in aprotic polar solvents.<sup>6</sup>

Solvent effects arise from a complex interplay between both enthalpic and entropic effects as a chemical reaction progresses.<sup>7</sup> The dynamics of the solute and solvent should be considered jointly to characterize these effects accurately. Any variation in solute induces changes in the solvent that in return influences the reactant dynamics.<sup>8</sup> Therefore, coupling between degrees of freedom of the solutes and solvent should be taken into account. Solvent molecules can collectively reorganize themselves around the solute throughout the different stages of a reaction. These collective motions of the solvent molecules can be required for the system to surmount the activation barrier, and observing the solvent distributions around the solute can provide valuable insights to quantify or qualify the underlying atomistic mechanisms. Solvent–solvent interactions also contribute to the free energy of the reaction. A solvent can be highly structured e.g. water with strong directional hydrogen bonds can form an intermolecular network, or weakly structured e.g. hydrocarbons with weak non-directional dispersion forces.<sup>9</sup> Changes in solvent distributions around the solute, i.e. reordering of non-bulk solvent molecules, are reflected directly in solvent–solvent energetics. In this context, solvent effects on the menshutkin reactions are investigated in Chapter 3.

Solute–solvent interactions are of vital importance in biochemical research as well. Tertiary structure stability of biomacromolecules such as proteins and enzymes depends on the interactions with water, both in the exterior and interior of the structure<sup>7</sup>, and water molecules evidently mediate protein folding.<sup>10</sup> Protein–ligand binding is commonly discussed in the terms of replacement of water molecules in the binding site by the ligand, although some water molecules might be retained in the binding site.<sup>10</sup> The stabilizing role of water to balance dipoles in proteins is still an ongoing discussion. However, accounting for electrostatic interactions around the protein is challenging due to inherent heterogeneous characteristics. The interior of a protein is usually non-polar whereas surrounding water molecules are highly polar. The charges on the protein residues not only interact with solvent molecules but also with each other. The charge–charge interactions on the protein are usually screened by the solvent. In general, the electrostatic energies of the proteins are significantly affected by charge–charge

and charge–solvent interactions.<sup>11</sup> In this context, the water structure around the specific regions of the myoglobin and KRAS proteins have been discussed in Chapter 4 and 5.

## 1.2 Post-translational Modifications

The central dogma of molecular biology was one of the major milestones in biochemical research. It was an explanation for how the genetic information flows within a biological system and is generally referred to as “DNA makes RNA, and RNA makes protein”.<sup>12</sup> It was first stated by Francis Crick as “the transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible.”<sup>13</sup> The information in this context means the precise determination of sequential information. It can be the sequence of bases in a nucleic acid or of amino acids in a protein.

In more detail, protein synthesis is a tightly controlled and complex process that occurs within cells. The first part of the central dogma (DNA makes RNA) is called transcription.<sup>14</sup> The process is initiated when the RNA polymerase enzyme binds to the region of a gene called the promoter, and breaks the hydrogen bonds between the complementary DNA nucleotides, unwinding the DNA. The enzyme adds RNA nucleotides that are complementary to one of the DNA strands and then the sugar-phosphate backbone is formed. Finally, hydrogen bonds in the DNA-RNA helix break and the messenger RNA (mRNA) strand is released.<sup>15</sup> If a cell is eukaryotic, additional processes such as capping, splicing and polyadenylation may occur.<sup>16</sup> After transcription is completed mRNA is exported to the cytoplasm where ribosome organelle reside.<sup>17</sup>

The second part of the central dogma (RNA makes protein) is called translation.<sup>18</sup> There are four RNA nucleobases but 20 amino acids, which means a nucleobase cannot code for a single amino acid residue. So, RNA bases are actually read off as groups of three during the translation. Each of these groups is called codons.<sup>19</sup> The genetic code consists of 64 different codons, and one type of amino acid can be encoded by more than one codon. The coding region of the mRNA ranges from the start codon (AUG) to one of the three stop codons (UAG, UGA, and

UAA) which neither encodes an amino acid residue.<sup>19</sup> The region between the 5' end of mRNA to the start codon (AUG) is called the untranslated region (UTR) and does not affect the amino acid sequence of the protein. However, the UTR is still a critical region of mRNA since it contains a ribosome binding site.<sup>20</sup>

The translation proceeds in three stages.<sup>21</sup> In the initiation stage, a small subunit of the ribosome binds to the UTR and proceeds to scan in the 5' to 3' direction until encounters the start codon. Then the large subunit of the ribosome is attached to the start codon, and transfer RNA (tRNA) binds to the P site on the ribosome. In the elongation stage, amino acids are brought to the ribosome by tRNA. Each codon on mRNA matched with a tRNA that contains complementary anti-codon and hence carries the corresponding amino acid.<sup>22</sup> In the termination stage, the ribosome releases the polypeptide chain, which will be folded into a protein, when the stop codon is encountered. Modifications that occur after this point are called post-translational modifications.

Post-translational modifications (PTM) are typically a covalent and enzymatic modification of proteins at one or more sites<sup>23</sup>, which can alter the physiochemical properties of the proteins and lead to structural, local hydration, and activity changes around the target site and binding partners.<sup>24</sup> They can impact the protein function both allosterically by the means of conformational changes induced at the functional site, or orthosterically via direct recognition by protein domains or through interference with binding.<sup>25</sup> The PTMs, which can be both reversible and irreversible, are precisely targeted and regulated by both spatial and temporal arrangements.

Protein PTM can be divided into two main groups. The first group classifies the covalent addition of a chemical group to the side chains of the amino acid residues. The attached chemical group is usually an electrophilic fragment of a cosubstrate and the modified side chain act as a nucleophile due electron-rich character.<sup>26</sup> 15 out of 20 common amino acid side chains can be subject to this group of PTMs whereas no side chain modifications of Leu, Ile, Val, Ala, or Phe are known.<sup>23</sup> The second group classifies the cleavage of peptide backbones by proteases or autocatalytic cleavage. Backbone modifications can govern both



the local and global conformation of the protein, to promote well-defined three-dimensional structures even in short peptides.<sup>27</sup> Backbone PTMs can range from small chemical modifications such as methylation of amide nitrogen<sup>28</sup> to significant alterations to the backbone<sup>29</sup>

The most common types of covalent side chain modifications are nitrosylation, phosphorylation, acylation, alkylation, glycosylation, and oxidation.<sup>26</sup> Protein kinase B is activated only after phosphorylation of Ser and Thr residues, and only then is able to regulate cell survival.<sup>30</sup> S-Palmitoylation, acylation type PTM, is involved in membrane anchoring of modified forms of Ras.<sup>31</sup> Further, mono- and polyubiquitylation modifications by ubiquitin, a protein with a weight of 8-kDa, is even known to unfold proteins.<sup>32</sup> The common chemical groups for alkylation are methyl, and C<sub>15</sub> or C<sub>20</sub> isoprenyl<sup>33</sup> which are introduced to the nearby environment to different levels depending on their size. Cysteine oxidation leads to protein aggregation via covalent cross-linking or hydrophobic interactions.<sup>34</sup> Lastly, S-nitrosylation, which is the PTM investigated in this thesis, is explained in detail in Chapters 4 and 5.

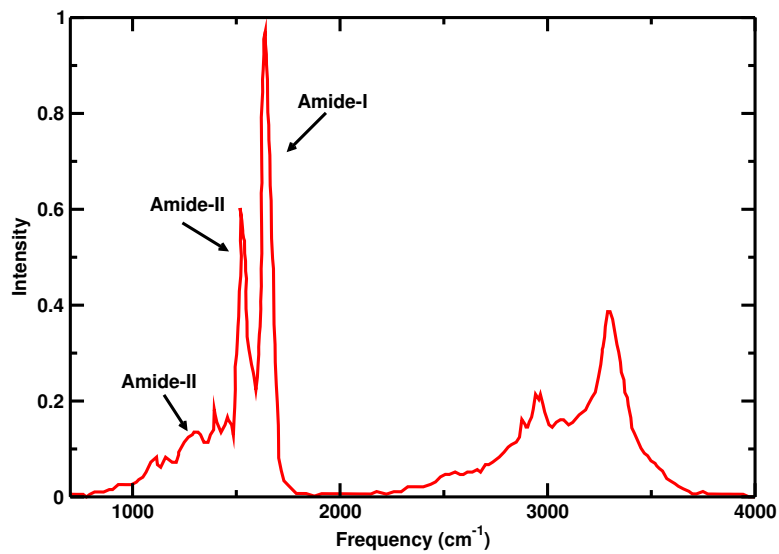
### 1.3 PTM Detection by IR Spectroscopy

Detection of post-translational modification is the first step to understanding its structure and its effect on function and regulation. Although Mass-spectroscopy is the main experimental method to detect the PTMs, a wide range of detection techniques has been proposed and applied. Detection techniques can be divided into two sub-groups. The direct detection techniques address the PTMs adducts directly whereas indirect methods usually break the covalent bond between the protein and modification, then capture the signals of the adduct part. Indirect methods such as Immunoprecipitation<sup>35</sup>, biotin-switch technique<sup>36</sup>, and Enzyme-Linked immunosorbent Assay (ELISA)<sup>37</sup> are robust methods but may have drawbacks and limitations in the terms of selectivity and reproducibility.<sup>38</sup> Indirect methods usually require multi-step sequential manipulation before final analysis hence decreasing chemoselectivity in each step can result in lower selectivity.<sup>39</sup> Direct methods such as erasable single-molecule blotting (eSiMBlot) or organophosphine probes are subject to similar selectivity problems as indirect

methods too, or in the case of phenylmercury can have limited applicability due to toxicity.<sup>39</sup>

As explained in the previous section, PTMs are the covalent modification of proteins. IR spectroscopy, which is one of the most direct methods to study bonds as it reports directly on the bond themselves, is a viable alternative to the aforementioned experimental methods to detect the PTMs on the proteins.<sup>40</sup> Inherently, the nature of a bond is sensitive to its surrounding local environment. The frequency, intensity, line shape, and number of vibrational absorptions of a given bond can be related to its local structure, environment, and dynamics.<sup>41</sup> Characterization of a modification by IR spectroscopy can also provide valuable information about the modification-induced effects on the local environment<sup>42</sup>, and it is used to investigate the molecular mechanism of protein reactions to protein folding and unfolding as well.<sup>43</sup>

Conformationally sensitive vibrations within the peptide backbone of the protein are the amide I ( $1600 - 1700 \text{ cm}^{-1}$ ), amide II ( $1510 - 1580 \text{ cm}^{-1}$ ), and amide III ( $1200 - 1350 \text{ cm}^{-1}$ ) bands which are common features in a protein IR spectrum, see Figure 1.1. The amide bands were used as site-specific structural and environmental reporters.<sup>44</sup> The amide I bands mainly arise from C=O stretching vibration and are directly related to the backbone conformations. The amide I has a high extinction coefficient, the distinct spectral signature of secondary structures. The amide II band results from N-H bending vibration and the C-N stretching vibration. Lastly, amide III is a complex band that arises from a mixture of several motions. As one of the drawbacks of IR spectroscopy, vibrational couplings and the spectral overlap between the amide bands and the signal of the modification can entangle the results, especially if the motion arises from the modification adducts have a low extension coefficient.



**Figure 1.1:** Infrared Spectrum of Myoglobin.<sup>45</sup> Peaks corresponds to amide I, amide II and amide II bands are labelled.



# Chapter 2

## Theoretical Background

### 2.1 Quantum Mechanical Methods

The Schrödinger equation is the fundamental equation of quantum mechanics that governs the wavefunction ( $\Psi$ ) of a quantum mechanical system. The wavefunctions can describe, and be directly correlated with the energetics and structure of a system. For many chemical systems, it is adequate to solve the time-independent Schrödinger equation:

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \quad (2.1)$$

where  $\hat{H}$  is the Hamiltonian operator for a system of nuclei and electrons and  $E$  is the energy eigenvalue. Although, the Schrödinger equation can be solved analytically for a single-electron system such as the hydrogen atom, for multi-electron systems the Hamiltonian is too complicated to solve exactly.

$$\hat{H} = \sum_{i=1}^n -\frac{\hbar^2}{2m_i} \nabla_i^2 + U(r) \quad (2.2)$$

where  $\nabla_i$  is the gradient operator and  $U(r)$  is the potential energy of the system. The main difficulty to in solving the Hamiltonian lies in the potential energy term which depends on the geometrical position of correlated electrons and nuclei. Consequently, approximations are needed for the multi-electron systems to solve the Schrödinger equation. The Born-Oppenheimer approximation<sup>46</sup> helps to separate the electron and nuclei motions. The approximation is based on the

fact that the nuclei moves more slowly compared to the electrons due to a three orders of magnitude heavier mass. Thus, the nuclei is approximated as being stationary. Therefore, the kinetic energy of the nuclei is neglected and repulsion among the nuclei is considered constant.

### 2.1.1 Hartree-Fock Theory

The Hartree-Fock (HF) theory is one of the pivotal methods to solve the Schrödinger equation for multi-electron systems. The HF approximation neglects the interaction between each individual electron, and instead considers the interaction of one electron to the average field created by the remaining electrons. Hence, it represents the  $N$  electron wave function as the product of the orthogonal wave function of the individual electrons which is known as the Hartree Product,

$$\Psi_{HP}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2) \cdots \chi_N(\mathbf{x}_N) \quad (2.3)$$

where  $\chi_i(\mathbf{x}_i)$  is the spin orbital of electron  $i$  which is a spatial orbital multiplied by a spin function. However, the Hartree Product (Eq. 2.3) does not satisfy the anti-symmetry principle. The wave function of identical fermions, electrons are fermionic particles with spin of 0.5, must be anti-symmetric with respect to an interchange between two fermions. The Slater Determinants (SDs) can be used to construct a wave function of  $N$  electrons which satisfy anti-symmetry principle,

$$\Psi_{SD} = |\chi_1\chi_2 \cdots \chi_N\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \quad (2.4)$$

A single Slater determinant is used as an approximation to the wave function. The variational principle, which HF is subject to, states that the best spin orbitals  $\chi_i$  for a given Slater determinant are the ones that minimize the electronic energy.

$$E_0 = \langle \Psi_{SD} | \hat{H} | \Psi_{SD} \rangle \quad (2.5)$$

where  $\hat{H}$  is the Hamiltonian operator. The electronic energy is minimized by using an iterative procedure called self-consistent field (SCF), and these results construct the HF equation,

$$\hat{F}_i \chi_i = \epsilon_i \chi_i \quad (2.6)$$

where  $\epsilon_i$  is the energy of the  $i$ -th spin orbital  $\chi_i$  and  $\hat{F}_i$  is the Fock operator. For a system that has  $2N$  electrons, the Fock operator is

$$\hat{F}_i = \hat{h}_i + \sum_j (\hat{J}_j - \hat{K}_j) \quad (2.7)$$

and  $\hat{J}_i$  and  $\hat{K}_i$  describe Coulomb electron-electron interaction and exchange operator while  $\hat{h}_i$  is the one-electron operator.<sup>47</sup>

### 2.1.2 Møller Plesset Perturbation Theory

The perturbation theory (PT) is one of the improvements over the simple HF method. The fundamental idea behind is to start from a simple model which can be solved exactly or approximately and progressively add small perturbations to this model. PT assumes that unperturbed reference systems are close to the perturbed system of interest. The PT partitions the Hamiltonian operator into the zeroth order, or called the reference part,  $\hat{H}_0$  and the perturbation part  $\hat{H}'$ .

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}' \quad (2.8)$$

where  $\lambda$  determines the strength of the perturbation. The energy and wave function continuously alter as  $\lambda$  values vary between 0 and 1. These terms can be expanded by using a Taylor expression, based on the parameter  $\lambda$ .

$$E = \lambda^0 E_0 + \lambda^1 E_1 + \lambda^2 E_2 + \dots + \lambda^n E_n + \dots \quad (2.9)$$

$$\Psi = \lambda^0 \Psi_0 + \lambda^1 \Psi_1 + \lambda^2 \Psi_2 + \dots + \lambda^n \Psi_n + \dots \quad (2.10)$$

where  $E_n$  and  $\Psi_n$  are the  $n$ -th order corrections to the energy and wave function, respectively. If  $\lambda = 0$ , the wave function is unperturbed, so  $\hat{H} = \hat{H}_0$ ; whereas If

$\lambda = 1$ , the  $n$ -th order energy or wave function becomes the sum of all terms up to  $n$ -th order. After the normalization of the perturbed wave function, the zero- to  $n$ -th order perturbation terms can be defined as;

$$\lambda^0 : \hat{H}_0 \Psi_0 = E_0 \Psi_0, \quad (2.11)$$

$$\lambda^1 : \hat{H}_0 \Psi_1 + \hat{H}' \Psi_0 = E_0 \Psi_1 + E_1 \Psi_0, \quad (2.12)$$

$$\lambda^2 : \hat{H}_0 \Psi_2 + \hat{H}' \Psi_1 = E_0 \Psi_2 + E_1 \Psi_1 + E_2 \Psi_0, \quad (2.13)$$

$$\lambda^n : \hat{H}_0 \Psi_n + \hat{H}' \Psi_{n-1} = \sum_{i=0}^n E_i \Psi_{n-i} \quad (2.14)$$

The Møller Plesset Perturbation Theory (MP) is a popular PT method that improves over the HF method by taking the electron correlation effect into account by using the Rayleigh–Schrödinger perturbation theory, to second (MP2), third (MP3), or fourth (MP4) order.<sup>48</sup> Although higher-order MP methods (e.g. MP5) exist, they are rarely used due to increased computational cost. The most popular MP method is MP2 due to the fact that it is the cheapest method that can yield a wave function based description of dispersive interactions.

MP method uses an HF calculation as a starting point for the perturbation expansion, thus the reference part (zeroth-order) part of MP2 is described by a HF density. To include the electron correlation effects, a transformation of the two-electron integrals from atomic orbitals (occupied;  $a, b$ ) to the basis of the canonical molecular orbitals (virtual;  $i, j$ ) is required.

$$(ia|jb) = \sum_{\mu\nu\lambda\sigma} c_{\mu i} c_{\nu a} c_{\lambda j} c_{\sigma b} (\mu\nu|\lambda\sigma) \quad (2.15)$$

where  $(ai|bj)$  integral is symmetrical, then MP2 energy can be calculated as,

$$E_{MP2} = - \sum_{ijab} \frac{2(ia|jb)(ia|jb) - (ia|jb)(ib|ja)}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \quad (2.16)$$

where  $\epsilon$  is the orbital energies.



### 2.1.3 Basis Sets

The aim of quantum mechanical methods is to derive information by solving the Schrödinger equation. However, as discussed in the previous sections, approximations are needed to solve it due to the complex nature of the equation. Basis sets, which are a set of known mathematical functions, are one of the fundamental approximations that are used to define an unknown function such as molecular orbital (MO) as expressed by a set of known functions.

An unknown molecular orbital ( $\psi_i$ ) can be expanded linearly from a set of  $N$  known basis functions  $\{\chi_\mu | \mu = 1, \dots, N\}$ ,

$$\psi_i = \sum_{\mu=1}^N c_{\mu i} \chi_\mu \quad (2.17)$$

where  $\chi_\mu$  is the basis set and  $c_\mu$  is the coefficient for a given  $\chi_\mu$ . This approach is known as the linear combination of atomic orbitals (LCAO).

Since only finite numbers of basis functions can be utilized, the size and the type of the basis sets are important parameters that should be taken into account to increase accuracy. Ideally, basis functions should allow wave function to be correctly constructed and have a low computational cost. If a basis function can successfully reproduce an unknown function, fewer of it would be necessary to achieve a given level of accuracy.<sup>48</sup>

The basis sets used in computational chemistry can be classified into two main types, Slater-type orbitals (STOs) and Gaussian-type orbitals (GTOs). The STOs have the functional form

$$\chi_{\zeta,n,l,m}(r, \theta, \varphi) = N Y_{l,m}(\theta, \varphi) r^{n-1} e^{-\zeta r} \quad (2.18)$$

where  $N$  is a normalization constant,  $r, \theta, \varphi$  are spherical coordinates and  $Y_{l,m}(\theta, \varphi)$  is spherical harmonic functions. Further,  $\zeta$  is the orbital exponent, and  $n, l$  and  $m$  are the quantum numbers for principal, angular momentum, and magnetic, respectively.  $l$  determines the shape of the orbital,  $l=0$  is for s-type,  $l=1$  is

for p-type, and  $l = 2$  is for d-type orbital etc. The  $m$  determines the spatial orientation of the orbital and takes values as  $-l \leq m \leq l$ . Another type of basis function is GTOs and they can be written in the terms of polar or coordinates as

$$\chi_{\zeta,n,l,m}(r, \theta, \varphi) = NY_{l,m}(\theta, \varphi)r^{2n-2-l}e^{-\zeta r} \quad (2.19)$$

$$\chi_{\zeta,l_x,l_y,l_z}(x, y, z) = Nx^{l_x}y^{l_y}z^{l_z}e^{-\zeta r^2} \quad (2.20)$$

where  $l_x, l_y, l_z$  are parameters. The sum of these parameters  $L = l_x + l_y + l_z$  is analogous to angular momentum which determines the orbital shape. The  $r^2$  dependence in the exponential terms causes problems for GTOs. Because GTOs do not have a cusp at the nuclear position, they have problems characterizing correct behavior in the proximity of the nucleus. Moreover, GTOs diminish rapidly compared to STOs, and it causes poor representation on the tail of the wave function. However, GTOs are much more computationally efficient due to fact that it is easier the work with Gaussian functions for the evaluation of two-electron integrals compared to STOs. Hence, a new approach called contracted basis functions is proposed, where the linear combinations of Gaussian functions are used to construct an approximate Slater-type function that allows atomic orbitals (AO) to have different signs in the different parts of the space. The number of contracted basis functions per occupied AO characterizes the degree of complexity. Minimal basis set for a basis function per AO, double- $\zeta$  basis set for two basis functions per AO, and three- $\zeta$  basis set for three basis functions per AO, etc.

The split-valence basis sets utilize the idea that it is more important to have flexibility in valence orbitals than core orbitals due to the fact that core orbitals have less contribution to the total energy of the system, and their shape is usually not affected by the bond formation. Pople basis, e.g. double- $\zeta$  6-31G, are among the most commonly used basis sets of this type. The polarization effect can be included in a basis set by using polarized basis functions ( $p$ ); these functions add  $d$  functions to polarize  $p$  basis functions and are denoted by (\*), e.g. 6-31G\*. Further, diffuse basis functions which have small exponents and decay slowly at large distances can be included for a better description of anions and van der Waals complexes, and are denoted by (+). Correlation consistent Dunning type basis sets that accounts for optimizing the basis sets using the electron correlation

methods are another type of commonly used basis sets. As an example, aug-cc-pVDZ, which is used in this thesis, is an augmented (aug, diffusion function), correlation-consistent (cc), polarized (p), valence (V), double- $\zeta$  (DZ) basis set.<sup>48</sup>

## 2.2 Force Field

Quantum mechanical methods are powerful tools to study chemical systems. However, they are only applicable to systems with a limited number of heavy atoms due to their high computational costs. So, many biological systems of interest which contain thousands of atoms are too large to study by using QM methods. Therefore, a simplified approach to study the structure, dynamics, and function of biological macromolecules such as proteins and enzymes was necessary.<sup>49</sup>

The electronic and nuclear motions of the molecules can be separated based on the aforementioned Born-Oppenheimer approximation. The force field (FF) and molecular mechanic (MM) methods neglect the electronic effects and calculate the total energy of the molecule as a function of nuclear coordinates. Nuclei are represented as a ball with mass and charge, and they are held together with “springs” that help to model different interactions, see Figure 2.1. The potential energy of the system,  $U(\vec{R})$ , is split into bonded and non-bonded terms,

$$U(\vec{R}) = U_{\text{bonded}} + U_{\text{non-bonded}} \quad (2.21)$$

The bonded term,  $U_{\text{bonded}}$ , includes the harmonic potentials representing bonds, angles, improper dihedrals, and periodic potential representing the dihedral angles.

$$U_{\text{bonded}} = \sum_{\text{bonds}} k_b (b - b_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 + \sum_{\text{dihedral}} k_\varphi (1 + \cos(n\varphi - \delta)) + \sum_{\text{impropers}} k_\omega (\omega - \omega_0)^2 + \quad (2.22)$$

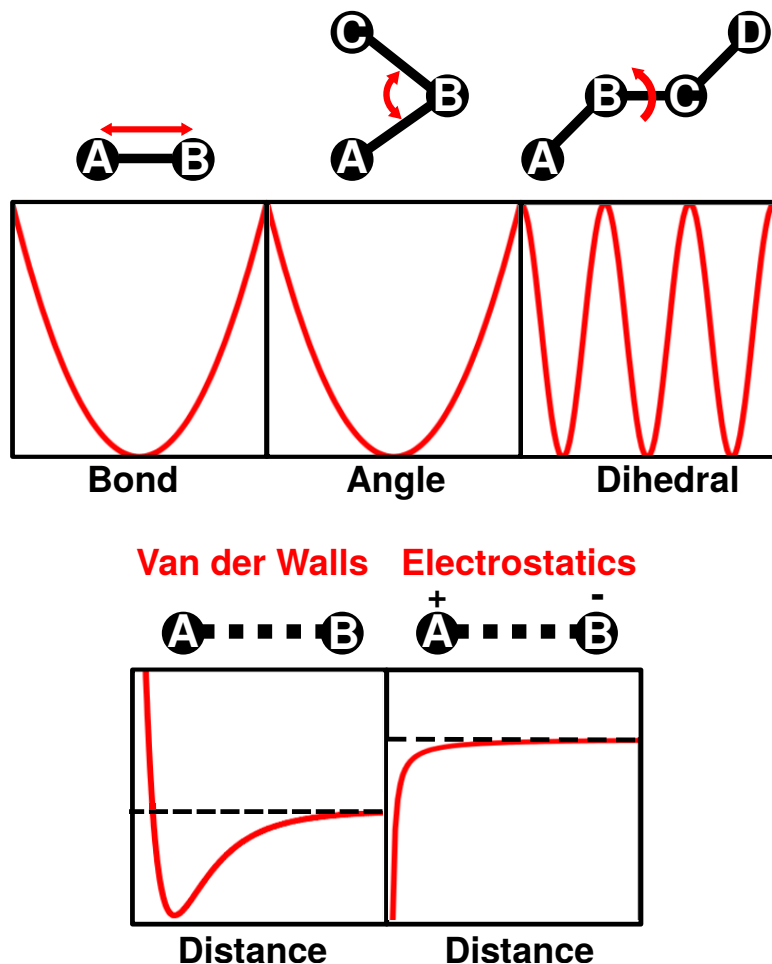
where  $k_b$ ,  $k_\theta$ , and  $k_\omega$  are the force constants and  $b_0$ ,  $\theta_0$ , and  $\omega_0$  are equilibrium values describing bonds, angles and improper dihedrals, respectively. For the di-

hedral angle,  $k_\varphi$  is the force constant,  $\varphi$  dihedral angle,  $n$  periodicity, and  $\delta$  the phase shift.

The non-bonded term,  $U_{\text{non-bonded}}$ , includes a Lennard-Jones (LJ) potential and electrostatic potential,

$$U_{\text{non-bonded}} = \sum_{i < j} \epsilon_{ij} \left[ \left( \frac{R_{\text{min},ij}}{r_{ij}} \right)^{12} - \left( \frac{R_{\text{min},ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (2.23)$$

The Lennard-Jones potential (LJ) is the first term in the Eq.2.23 which describes the van der Waals interaction between pairs of non-bonded atoms with  $i$ , and  $j$  indices. Where,  $r_{ij}$  is the distance between those atoms,  $\epsilon_{ij}$  is the depth of the Lennard-Jones potential, and  $R_{\text{min},ij}$  is the distance at which Lennard-Jones potential is zero. The LJ interactions have an attractive part with  $r^{-6}$  dependence, and a repulsive part with  $r^{-12}$  dependence. However,  $r^{-12}$  which is characteristic of the repulsive part is chosen for mathematical convenience rather than a physical motivation. The second part of the Eq.2.23 describes the electrostatic interactions,  $q_i$  and  $q_j$  are the partial charges of non-bonded atoms and  $\epsilon_0$  is the effective dielectric constant.<sup>49</sup>



**Figure 2.1:** Schematic representation of the force field terms with their corresponding potentials: bond stretching, bending, rotation, non-bonded interactions

The aforementioned terms in Eq. 2.22 and 2.23 are the standard terms of force fields. However, there are additional terms that have been incorporated into some force fields to increase their accuracy. One of the examples of such terms is the Urey-Bradley (UB) harmonic potential<sup>49</sup>  $U_{UB} = k_{UB}(R - R_0)^2$  at the distance ( $R$ ) between A and C atoms for bonded atoms A-B-C.  $k_{UB}$  is the force constant, and  $R_0$  is the equilibrium distance. Urey-Bradley is an important term to improve calculated vibrational spectra by treating symmetric and asymmetric bond stretching mode separation, as well as in-plane deformations.

The bond terms are defined as harmonic potential in the standard force fields thus they can not be broken or formed. Further, harmonic potential used to describe the bonds can not mimic the anharmonic nature of real bonds, and transitions to higher vibrational levels occur at the same frequency due to equal energy

spacing. Morse potential is used to overcome these problems, it can account for anharmonicity and non-zero transition probability for overtones and provide better characterization for the vibrational structure of the molecule;

$$U_{\text{Morse}} = D_e[1 - e^{-\beta(r-r_0)}]^2, \quad (2.24)$$

where  $D_e$  is the dissociation energy of the bond,  $\beta$  is the width of the potential and  $r_0$  is the equilibrium bond length.

### 2.2.1 Multipolar Force Field

Conventional force fields use atom-centered point charges (PCs), hence are unable to describe certain features of the electrostatic potential (ESP) of molecules. Realistic description of lone pairs,  $\sigma$ -holes, and  $\pi$ -electron density can not be achieved with PCs.<sup>50</sup> Therefore, multipoles (MTP) are incorporated into some FFs as a correction to atom-centered PCs. MTP are commonly obtain from *ab-initio* ESP. If  $\phi(q)$  is an electron density function of coordinate  $q$ ,  $\phi(q)$  values can be discretely obtained on a 3-dimensional (3D) grid with  $(r_k)$  coordinate points.

$$\phi(q_i) \approx \Phi(r_k) \quad (2.25)$$

$k = k^x, k^y, k^z$  is the coordinate where  $q_i$  is at the closest distance. Based on that, the ESP can be approximated at any grid point  $r$  using MTP (from monopoles to quadrupoles)<sup>51</sup> as follows,

$$\begin{aligned} \Phi(\mathbf{r}) &= \sum_i \sum_j Q_j^{(i)} f_j^{(i)}(\mathbf{r}) \\ &\approx \sum_i Q_{00}^{(i)} r^{-1} + Q_{10}^{(i)} r^{-2} \hat{r}_z + Q_{11c}^{(i)} r^{-2} \hat{r}_x + Q_{11s}^{(i)} r^{-2} \hat{r}_y \\ &\quad + Q_{20}^{(i)} r^{-3} (3\hat{r}_z^2 - 1)/2 + Q_{21c}^{(i)} r^{-3} \sqrt{3} \hat{r}_x \hat{r}_z \\ &\quad + Q_{21s}^{(i)} r^{-3} \sqrt{3} \hat{r}_y \hat{r}_z + Q_{22c}^{(i)} r^{-3} \sqrt{3} (\hat{r}_x^2 - \hat{r}_y^2)/2 \\ &\quad + Q_{22s}^{(i)} r^{-3} \sqrt{3} \hat{r}_x \hat{r}_y \end{aligned} \quad (2.26)$$

where  $i$  repeats for all atoms and  $j$  for all MTP coefficients.  $Q_{kl}$  is the  $l$ th MTP moment of rank  $k$  in spherical coordinates.  $f_j^{(i)}(\mathbf{r})$  are geometrical factors together with angular- and distance-dependent terms for the MTP moment  $Q_j^{(i)}$  at point  $\mathbf{r}$ .  $r = \|\mathbf{r}\|$  is the norm of vector  $\mathbf{r}$  and  $\hat{r}_a = \mathbf{r} \cdot \hat{a}/r$  is the norm of projection of vector  $\mathbf{r}$  on one of the three vectors  $x$ ,  $y$ , or  $Z$ .

Following target function then to be optimized to minimize the error between *ab-initio* ESP, and MTP estimated ESP,

$$\chi^2 = \min \sum_k (\Phi_{ab\text{-}initio}(r_k) - \Phi_{\text{MTP}}(r_k)) \quad (2.27)$$

## 2.2.2 Multi-Surface Adiabatic Reactive Molecular Dynamics

Conventional forces fields are formulated for a diabatic state where the atom connectivity does not change. This construct inherently hampers FFs ability to simulate chemical reactions in which bond breaking or forming takes place. *ab-initio* MD is one of the alternatives to simulate chemical reactions, the method computes the energy and forces from electronic structure calculations. However, evaluation of the force and energy at the *ab-initio* level is computationally expensive thus limiting to an extensive sampling of the configurational space. Several MM methods which allows smooth switching between diabatic states via use of switching function are developed to simulate chemical reactions. The reactive force field (Reax FF)<sup>52,53</sup>, multi-state VALBOND (MS-VALBOND)<sup>54</sup> and multi-surface Adiabatic Reactive Molecular Dynamics (MS-ARMD)<sup>55</sup> are examples of such force fields.

MS-ARMD is an energy conserving surface crossing algorithm based on the potential energy of the system of interests. The reactant and product states are described by parameterized FFs and called surfaces. The Morse potential is used in parameterized FFs to allow bond forming or breaking (see Section 2), and Generalized Lennard-Jones (GLJ) potential replaced 6-12 Lennard-Jones potentials to improve the description of van der Waals interaction between the atoms.

$$V_{GLJ}(r; n, m, \epsilon, r_{min}) = \frac{n\epsilon}{m-n} \left[ \left( \frac{r_{min}}{r} \right)^m - \frac{m}{n} \left( \frac{r_{min}}{r} \right)^2 \right] \quad (2.28)$$

where  $\epsilon$  is the depth of the GLJ potential and  $r_{min}$  is the separation at the minimum energy. The Morse and GLJ potentials are needed to accurately describe the potential energy along the reaction pathway.

The effective potential energy is defined as the linear combination of  $n$  diabatic PESs with coordinate-dependent weights.

$$V_{MS-ARMD}(\mathbf{x}) = \sum_{i=1}^n w_i(\mathbf{x})V_i(\mathbf{x}) \quad (2.29)$$

where  $w_i(\mathbf{x})$  calculated by renormalizing the raw weights  $w_{i,0}(\mathbf{x})$ ,

$$w_i(\mathbf{x}) = \frac{w_{i,0}(\mathbf{x})}{\sum_{j=1}^n w_{j,0}(\mathbf{x})} \quad (2.30)$$

where  $w_{i,0}(\mathbf{x})$  are obtained from exponential decay function of the energy difference between the minimum energy surface ( $V_{min}(\mathbf{x})$ ) and surface  $i$  ( $V_i(\mathbf{x})$ ) with a energy difference of  $\Delta V$ ,

$$w_{i,0}(\mathbf{x}) = \exp\left(-\frac{V_i(\mathbf{x}) - V_{min}(\mathbf{x})}{\Delta V}\right) \quad (2.31)$$

Combining the independently parameterized reactant and product state FFs leads to an unrealistic high-energies at the surface crossing region i.e. barrier regions. The products of GAussian and POLynomial functions (GAPOs) utilized to better characterize barrier region by adjusting it to match electronic structure calculations energies,

$$\Delta V_{GAPO,k}^{ij}(\mathbf{x}) = \exp\left(-\frac{(\Delta V_{ij}(\mathbf{x}) - V_{ij,k}^0)^2}{2\sigma_{ij,k}^2}\right) \times \sum_{l=0}^{m_{ij,k}} a_{ij,kl}(\Delta V_{ij}(\mathbf{x}) - V_{ij,k}^0)^l \quad (2.32)$$

where  $V_{ij,k}^0$  and  $\sigma_{ij,k}$  are the center and the standard deviation of the Gaussian function, respectively. Lastly, the global MS-ARMD potential energy surface is given by



$$V_{MS-ARMD}(\mathbf{x}) = \sum_{i=1}^n w_i(\mathbf{x}) V_i(\mathbf{x}) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n [w_i(\mathbf{x}) + w_j(\mathbf{x})] \sum_{k=1}^{n_{ij}} \Delta V_{GAPO,k}^{ij}(\mathbf{x}) \quad (2.33)$$

## 2.3 Molecular Dynamics Simulations

Molecular Dynamics Simulations (MD) is the primary tool to investigate the dynamics of chemical reactions and biomacromolecules. MD calculates the forces on each individual atom, based on the given positions of all atoms in a model system, by utilizing Newton's law of motion. If the position and velocity of each atom are known, then the following spectral position of each atom can be computed as a function of time. Hence, the consecutive configurations of motions in 3-dimensional space with respect to time, called trajectory, can be generated by integrating Newton's law of motion.<sup>56</sup> According to Newton's second law,

$$\mathbf{f}_i = m_i \mathbf{a}_i = m_i \frac{d^2 \mathbf{r}_i}{dt^2} \quad (2.34)$$

Where  $\mathbf{f}_i$  is the exerted force,  $m_i$ ,  $a_i$ ,  $r_i$  are the mass, acceleration, and position of atom  $i$  at time  $t$ . The exerted force on atom  $i$  is determined from the gradient of the potential energy  $V(\mathbf{r}^N)$  where  $r^N = (r_1, r_2, r_3, \dots, r_N)$  are the 3N spatial coordinates of atom  $i$ ,

$$\mathbf{f}_i = -\frac{\partial V}{\partial \mathbf{r}_i} \quad (2.35)$$

The differential Eq. 2.34 should be solved to generate the trajectory. There are multiple numerical algorithms used in MD to integrate equations of motion. The Velocity Verlet<sup>57</sup> and Leap-frog<sup>56</sup> algorithms used in these studies are among the most common.

In velocity verlet, the new positions and velocities of each atom at time  $t + \delta t$  can be obtained from the positions and velocities at time  $t$ ,

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2 \quad (2.36)$$

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t) + \frac{1}{2}[\mathbf{a}(t) + \mathbf{a}(t + \delta t)]\delta t \quad (2.37)$$

where  $\mathbf{r}(t)$ ,  $\mathbf{v}(t)$  and  $\mathbf{a}(t)$  are the positions, velocities and accelerations of atoms at time  $t$ . Then the coordinates and velocities can be propagated by solving Eq.2.36 and 2.37, and the trajectory can be recorded. The main advantage of velocity Verlet is that the algorithm allows calculating both atomic positions and velocities at the same time, unlike the leapfrog in which coordinates and positions are updated half time step apart.

In leap-frog algorithm, velocities at time  $t - \frac{1}{2}\delta t$  and accelerations at time  $t$  are used to calculate the velocities at  $t + \frac{1}{2}\delta t$ . Then the velocities at  $t + \frac{1}{2}\delta t$  utilized to get the positions at  $t + \delta t$  as follows,

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \mathbf{v}(t + \frac{1}{2}\delta t)\delta t \quad (2.38)$$

$$\mathbf{v}(t + \frac{1}{2}\delta t) = \mathbf{v}(t - \frac{1}{2}\delta t) + \mathbf{a}(t)\delta t \quad (2.39)$$

The size of  $\delta t$  regulates how much the system will move in a certain direction with a certain velocity. So, the correct selection of  $\delta t$  plays an important role in the both accuracy and feasibility of the trajectories. If  $\delta t$  is very small, the trajectory calculation is going to be less feasible due to the high computational cost. Although computational cost would be lower with a very large  $\delta t$ , It may lead to chemical instabilities e.g. two particles too close to each other could result in high repulsive energy. The  $\delta t$  should be smaller than the fastest periodic motion in the system, in most chemical/biochemical systems  $\delta t = 0.1$  fs is used to incorporate hydrogen bond stretch motion.

The initial conditions of MD simulations are drawn from statistical ensembles to propagate Newton's equations of motion and get consequential trajectories. Statistical ensembles consist of a large number of virtual copies of the system in different states and share common macroscopic properties. Three common ex-

amples of statistical ensembles, which are used in this thesis, are microcanonical ( $NVE$ ) with a constant number of particles  $N$ , volume  $V$  and total energy  $E$ , canonical ( $NVT$ ) with a constant number of particles  $N$ , volume  $V$  and temperature  $T$ , and lastly Isothermal–isobaric ensemble ( $NpT$ ) with a constant number of particles  $N$ , pressure  $p$  and temperature  $T$ .

### 2.3.1 Umbrella Sampling

The free energy difference between the states is one of the main driving forces of any chemical process. However, characterization of the free energy surfaces still remains a challenge with MD simulations due to the high energy barriers of some chemical and biological processes. Conventional MD methods usually infrequently sample these higher energy regions, which are represented as saddle points on the free energy surface, and their configurations may not appear along the reaction pathway. Further, there is still a gap between the time scale that can be achieved in MD simulations and the time scale required to observe a high energy process. Hence, enhanced sampling methods which accelerate the dynamics of such systems have been developed to overcome these aforementioned problems.

Umbrella sampling (US)<sup>58</sup> is an enhanced sampling method for determining free energy surfaces along with user-defined reaction coordinates. An external harmonic bias potential ( $U(r)$ ) is added to the system to provide sufficient sampling along the reaction coordinate.

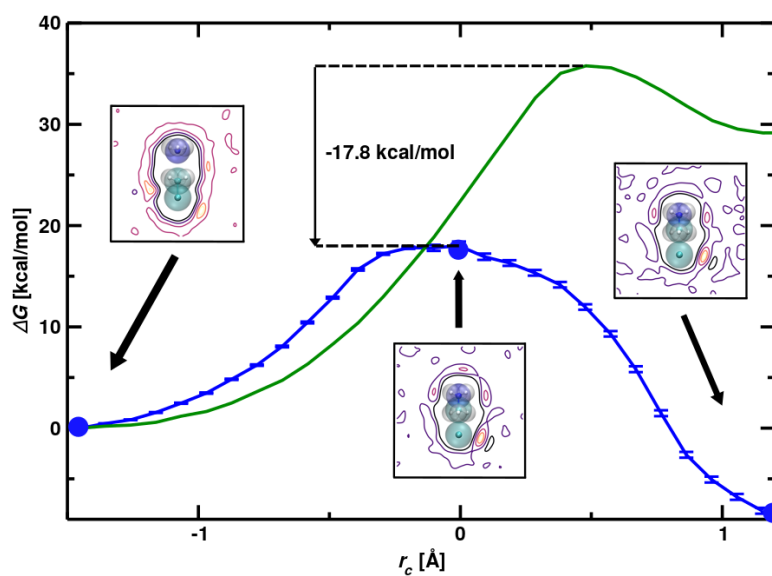
$$U(r) = k_u(r - r_0)^2 \tag{2.40}$$

where  $k_u$  is the force constant,  $r$  is the current value of the reaction coordinate  $r_0$  is the equilibrium value of the reaction coordinate for a given umbrella. Each independent simulation with a given reaction coordinate value is called a “window”. To generate the potential mean forces (PMF), multiple simulations with bias potential at the equidistantly spaced windows along the reaction coordinate are carried out. Histograms of the sampled reaction coordinate values are constructed for each window, and spacing between windows is chosen to confirm overlapping histograms. Following, statistics from all the windows were combined and analyzed with WHAM<sup>59,60</sup> to yield a one-dimensional PMF.



# Chapter 3

## Solvent Effects on the Menshutkin Reaction



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

*J. Phys. Chem. B. 2022, 126, 9, 1951-1961*

*doi: 10.1021/acs.jpcc.1c09710*

\*\*\*

*Dr. Sebastian Brickel contributed to this work as a second author.*



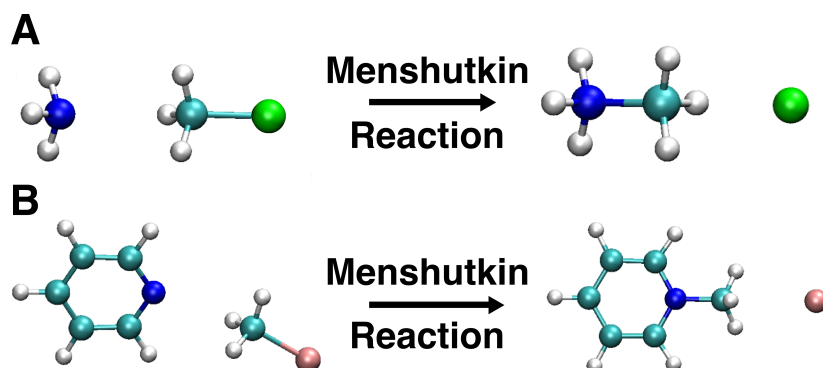
## 3.1 Abstract

The Menshutkin reaction is a methyl transfer reaction relevant in fields ranging from biochemistry to chemical synthesis. In the present work, the energetics and solvent distributions for  $\text{NH}_3+\text{MeCl}$  and  $\text{Pyr}+\text{MeBr}$  reactions were investigated explicit solvent (water, methanol, acetonitrile, benzene, cyclohexane) by means of reactive molecular dynamics simulations. For polar solvents (water, methanol, and acetonitrile) and benzene, strong to moderate catalytic effects for both reactions were found whereas apolar and bulky cyclohexane interacts weakly with the solute and does not show pronounced barrier reduction. The calculated barrier heights for the  $\text{Pyr}+\text{MeBr}$  reaction in acetonitrile and cyclohexane are 23.2 and 28.1 kcal/mol compared with experimentally measured<sup>61</sup> barriers of 22.5 and 27.6 kcal/mol, respectively. The solvent distributions change considerably between reactant and TS but comparatively little between TS and product conformations of the solute. As the system approaches the transition state, correlated solvent motions occur which destabilize the solvent-solvent interactions. This is required for the system to surmount the barrier. Finally, it is found that the average solvent-solvent interaction energies in the reactant, TS, and product state geometries are correlated with changes in the solvent structure around the solute.

## 3.2 Introduction

Solvation is essential in chemistry and directly influences properties such as reaction rates or spectroscopic responses of solutes.<sup>2,3,62-71</sup> Depending on the nature of the solvent (e.g. polar or apolar) the potential energy surface underlying the nuclear dynamics changes and affects mechanistic aspects of the dynamics.<sup>72-74</sup> As an example, for the Claisen rearrangement the reaction barrier in polar solvent decreases compared to the gas phase<sup>75-78</sup> whereas it increases for  $\text{S}_{\text{N}}2$  reactions

in going from the gas phase into solution.<sup>79–86</sup>



**Figure 3.1:** Schematic representation for A) the Menshutkin reaction of ammonia with chloromethane and B) the Menshutkin reaction of pyridine with bromomethane.

The Menshutkin reaction<sup>87,88</sup> is an important methyl transfer reaction and plays an essential role in many fields of chemistry ranging from biochemical processes (histone methylation<sup>89</sup>, methyltransferases<sup>90,91</sup>) to chemical synthesis.<sup>92</sup> Methyltransfer reactions in biological systems are catalyzed by methyltransferases for which the most prominent enzymatic cofactor is *S*-adenosyl-L-methionine (SAM).<sup>90</sup> SAM is derived from ATP by the enzyme methionine synthase<sup>93</sup> and serves as a regulator of a variety of processes including DNA, tRNA, and rRNA methylation as well as immune response.<sup>94</sup> DNA methylation is governed by DNA methyltransferase (DNMTs) enzymes such as DNMT1, DNMT3A, and DNMT3B.<sup>95</sup> More recent findings suggest that DNA and histone methylation cooperate to maintain the cellular epigenomic landscape.<sup>96</sup> Finally, histone methylation has been found to be undone by histone demethylases.<sup>97</sup>

Common examples of methylation reactions that have been studied extensively in the past are S<sub>N</sub>2 reactions<sup>82,98–106</sup> such as  $\text{Cl}^- + \text{MeBr} \rightarrow \text{MeCl} + \text{Br}^-$ <sup>85</sup> or  $\text{F}^- + \text{MeI} \rightarrow \text{MeF} + \text{I}^-$ <sup>105</sup> and the Menshutkin reaction.<sup>61,84,107–112</sup> The S<sub>N</sub>2 and Menshutkin reactions are of particular interest because of their importance in



chemical synthesis.<sup>113</sup> The main difference between a standard  $S_N2$  reaction and a Menshutkin reaction is that for an  $S_N2$  reaction the reactant state is charged whereas it is neutral for a Menshutkin reaction, e.g.  $\text{NH}_3 + \text{NH}_2\text{C}_6\text{H}_4\text{-Br} \rightarrow \text{NH}_2\text{-C}_6\text{H}_4\text{-NH}_3^+ + \text{Br}^-$  or  $\text{Pyr} + \text{MeCl} \rightarrow [\text{Pyr-Me}]^+ + \text{Cl}^-$ .<sup>61,111</sup> The Menshutkin reaction proceeds via a dipolar transition state (TS) to an ionic product and the TS is stabilized in polar solvents.<sup>61,109</sup>

The Menshutkin reaction is also a characteristic example for studying the effect of solvation on reactions<sup>112</sup> since it displays pronounced solvent effects on activation barriers.<sup>61,114,115</sup> Experimental studies have shown a decrease in the activation energy depending on the polarity of the solvent. The experimentally reported activation barrier for  $\text{Pyr} + \text{MeBr}$  in methanol is 22.5 kcal/mol which is 3.1 and 5.1 kcal/mol lower compared with apolar cyclohexane and di-n-butyl ether<sup>61</sup>. Further, an almost linear relationship between the natural log of the rate constants and polarity/polarizability of the solvents for the alkylation (Menshutkin) reaction between 1-bromodecane and 1,2-dimethylimidazol was observed.<sup>116</sup> DMSO was found to be the best solvent for accelerating the rate of heteroatom alkylations.

Although solvent effects are evidently relevant for the energetics of the Menshutkin reaction, surprisingly little is known about the molecular details of the reaction itself and the solvent ordering in particular. Previous calculations<sup>61</sup> for the  $\text{Pyr} + \text{CH}_3\text{Br}$  reaction were carried out with an implicit solvent model (Self consistent reaction field Self-Consistent Isodensity PCM (SCRF-SCIPCM)) which, however, does not provide molecular-level insight on the role of explicit solvent as the reaction develops from reactant to product. Gaining insight into the reaction barriers and how the solvent ordering changes with explicit solvent

present is one of the motivations for the present work. Only very recently, the effect of electric fields on the Pyr+CH<sub>3</sub>I reaction in explicit acetonitrile, chloroform and acetone was considered by analyzing individual snapshots using mixed quantum/molecular mechanics (QM/MM) calculations.<sup>117</sup> However, no activation free energies were determined in this study.

Solvent effects for the NH<sub>3</sub>+MeCl and Pyr+MeBr reactions were previously considered with quantum mechanical (QM) and semi-empirical methods. For NH<sub>3</sub>+MeCl electronic structure calculations at the HF/6-31++G\*<sup>118</sup> and B3PW91/6-31+G\*<sup>119</sup> levels with the implicit polarizable continuum model (PCM) model<sup>120</sup> as well as QM/MM simulations at the B3LYP/MM level were carried out.<sup>121</sup> Also, both reactions were investigated with multiscale reaction density functional theory (RxDFT) in acetonitrile.<sup>122</sup>

The present work addresses the question of how the explicit presence and the nature of the solvent affects the reaction barrier height for two methyl transfer reactions NH<sub>3</sub>+MeCl and Pyr+MeBr (see Figure 3.1). For a comprehensive sampling of the molecular dynamics along the progression coordinate the multistate adiabatic reactive molecular dynamics (MS-ARMD) framework was used which provides a computational means to investigate chemical reactions (bond breaking/bond formation) with an efficiency comparable to an empirical force field.<sup>55</sup> Both, the energetics of the reaction and the organization of the solvent for critical points along the progression coordinate is analyzed. First, the force fields and their parametrization is described. This is followed by free energy calculations and the analysis of the catalytic effect, the organization of the solvent around the transition state and the solvent energetics. Finally, the results are discussed in a

broader context.

## 3.3 Computational Methods

### 3.3.1 Molecular Dynamics Simulations

All molecular dynamics (MD) simulations were performed with CHARMM<sup>123</sup> with provision for forming and breaking bonds via MS-ARMD.<sup>55</sup> MD simulations were started following 500 steps of steepest descent and 500 steps of Adopted Basis Newton-Raphson minimization. Then, 500 ps of  $NpT$  dynamics were carried out using the leapfrog Verlet integrator<sup>57</sup> ( $\Delta t = 1$  fs) and a Hoover barostat<sup>124</sup> with a collision rate of  $5 \text{ ps}^{-1}$ . The  $NpT$  dynamics were followed by 2 ns ( $\Delta t = 1$  fs) of free dynamics in the  $NVT$  ensemble using SHAKE<sup>125</sup>. Periodic Boundary Conditions (PBC) together with the Particle Mesh Ewald (PME<sup>126</sup>) method were used for the long range electrostatic interaction. The cut-off for non-bonded, switching, and smoothing was 16 Å, 14 Å, and 12 Å, respectively.

The simulations in the different solvents were started by solvating the reactant states of the two reactions,  $\text{NH}_3 + \text{MeCl}$  and  $\text{Pyr} + \text{MeBr}$ , in pre-equilibrated, cubic solvent boxes. Due to the different shapes and sizes of the solvents, cubic box sizes were chosen accordingly. They were of length  $L = 30$  Å (water),  $L = 25$  Å (methanol),  $L = 28$  Å (acetonitrile),  $L = 27$  Å (benzene), and  $L = 30$  Å (cyclohexane), respectively. The parametrizations for the solvent molecules are those from the CHARMM General Force Field (CGenFF).<sup>127</sup>

### 3.3.2 Parametrization of the Force Field

All electronic structure calculations of reactant, transition, and product states were performed with Gaussian09<sup>128</sup> at the MP2/6-311++G(2d,2p) level of theory which has been found to be superior for reaction energetics compared with density functional theory treatments in a recent assessment for methyl transfer reactions.<sup>129</sup> For simulating Me-transfer between a donor and an acceptor a reactive force field to break and form chemical bonds is required. MS-ARMD is a reactive molecular dynamics implementation which combines individually weighted connectivities, describing different states by parametrized force fields. So called, GAPOs (GAussian  $\times$  POlynomials) for describing the adiabatic barrier are added to smoothly connect reactant and product side to form the global reactive Potential Energy Surface (PES). GAPOs are calculated from the energy difference  $\Delta V_{ij}(x)$  ( $= V_j(x) - V_i(x)$ ) between two states  $i$  and  $j$ .

For the Menshutkin reactions  $\text{NH}_3 + \text{MeCl}$  and  $\text{Pyr} + \text{MeBr}$  the parametrization of the reactant and product states started with optimization of the reactant and product geometries at the MP2/6-311++G(2d,2p) level of theory. Initial force fields and charges for the reactant and product molecules ( $\text{NH}_3$ , MeCl, methylammonium cation, pyridine, and methyl pyridine cation, respectively), were obtained from Swissparam<sup>130</sup> and from a natural bond order (NBO)<sup>131</sup> analysis of the wavefunctions.

For  $\text{NH}_3 + \text{MeCl}$  the reference reactant complex structures were generated from a 1 ns gas-phase MD simulation in the  $NVT$  ensemble using the Swissparam parametrization. In the MD simulations the distance between the ammonia and chloromethane was kept close to the equilibrium distance in the reactant state using a harmonic constraint. From this trajectory 1000 structures were stored for

which reference energies at the MP2/6-311++G(2d,2p) level were determined. For the Pyr+MeBr reaction a similar procedure was followed. The reference reactant complex structures were generated from a 1 ns gas-phase MD simulation in the *NVT* ensemble using the Swissparam parametrization. The distance between the pyridine and bromomethane was kept close to the equilibrium distance in the reactant state using a harmonic constraint. From this trajectory 1700 structures were stored for which reference energies at the MP2/6-311++G(2d,2p) level were determined. The different number of snapshots used in the parametrization (1000 vs. 1700) is due to the differing sizes and numbers of internal degrees of freedom of the two reaction systems.

In addition to the conformational ensemble of the reactant states, structures and energies along the internal reaction coordinate (IRC) path were used together with product state geometries which were generated from scanning along the C–Cl and C–Br distances, respectively. Because the focus of the present work is on the forward barrier, parametrizations were carried out for the reactant, TS, and ion-pair states but not for the ionic products ( $[\text{H}_3\text{NMe}]^+$  and  $[\text{PyrMe}]^+$ ). This would be required if the reaction is followed out and beyond the solvent-separated ion pairs which is not done in the following. Hence, on the product side the relaxed scan only involved C–Cl separations between 2.5 Å to 6 Å and C–Br distances ranging from 3 Å to 13 Å, respectively. Subsequently all force field parameters were fitted to the reference energies<sup>78,132–134</sup> by using a downhill simplex algorithm<sup>135</sup>. Finally, the reactant and product force fields were connected by fitting the GAPOs along the IRC using a genetic algorithm. The bond, angle, dihedral and van der Waals parameters of the force field are provided in Tables A.1 and A.2 in appendix A. Tables A.3 and A.4 summarize the NBO, the NBO-based refined and Mulliken (for comparison) charges for the two reactions. Further, the resulting fit of the Menshutkin reactions for  $\text{NH}_3+\text{MeCl}$  and  $\text{Pyr}+\text{MeBr}$  contains

three Gaussians and the GAPO parameters are provided in Tables A.5 and A.6 in appendix A.

### 3.3.3 Umbrella Sampling

Since direct sampling of the two reactions is not possible due to the high reaction barriers ( $> 20$  kcal/mol), umbrella sampling (US)<sup>136</sup> simulations were used to follow the reaction path. The reaction coordinate chosen here was the difference between the carbon-X bond of the reactant and the carbon-N bond in the product, i.e.  $r_c = d_{CX} - d_{CN}$  where X = Cl, Br for the NH<sub>3</sub>+MeCl and Pyr+MeBr reactions, respectively. Simulations were carried out for equidistant windows between  $r_c = -1.3$  and  $1.6$  Å with  $\Delta r_c = 0.1$  Å and with  $k_{umb} = 150$  kcal/mol/rad<sup>2</sup>.

US simulations for both systems were performed in a sequential manner, i.e. restarting a new US simulation from the structure and velocities of the previous simulation. Each window was simulated for 50 ps and window statistics was accumulated after equilibration for 5 ps. Statistics from all the windows were combined to yield the 1-dimensional potential of mean force (PMF) using the Weighted Histogram Analysis Method (WHAM)<sup>59,60</sup> with a tolerance of 0.001.

### 3.3.4 Solvent Distribution

For characterizing the solvent distributions in the reactant, TS, and product state structures of the solute, more extensive, separate 2 ns MD simulations for NH<sub>3</sub>+MeCl and Pyr+MeBr were carried out. The reactant, TS and product state structures were constrained at their MP2/6-311++G(2d,2p) optimized geometries. To analyze the solvent distributions the solute was reoriented and su-

perimposed to minimize the structural root mean squared deviation with respect to the first frame of the simulation. As the solute differs for the two reactions considered, separate schemes were followed; see Figure 3.1. For the  $\text{NH}_3+\text{MeCl}$  reaction first the methyl-carbon was translated to the origin of the box, then the ammonia-N was aligned along the  $x$ -axis and one of the three hydrogens bonded to the ammonia-N was placed into the  $xy$ -plane. For the  $\text{Pyr}+\text{MeBr}$  reaction the pyridine-N was translated to the origin of the box, the methyl-C was then aligned along the  $x$ -axis and one of the hydrogens bonded to the ammonia-N was placed in the  $xy$ -plane.

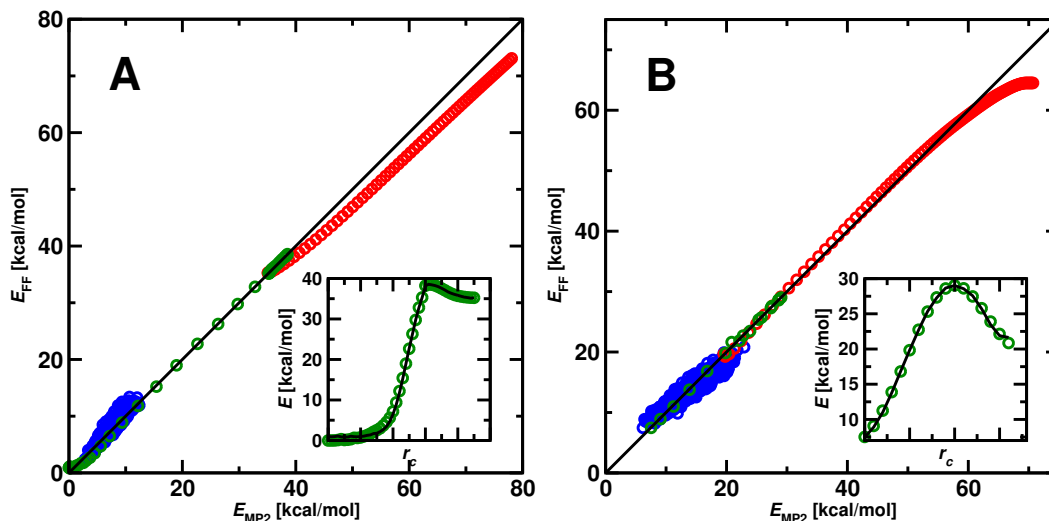
From each of the 2 ns simulations, 2000 snapshots were extracted and the water-oxygen ( $\text{O}_W$ ), methanol-oxygen ( $\text{O}_{\text{MeOH}}$ ) and acetonitrile-nitrogen ( $\text{N}_{\text{ACN}}$ ) coordinates were obtained whereas for benzene and cyclohexane the center of mass was determined and used for further analysis. The CoM was chosen for apolar solvents because selecting one specific atom as for the polar solvents is less meaningful as this would lead to skewing the results due to rotation of the solvent molecule. Based on these reference coordinates, all solvent molecules with the point of reference ( $\text{O}_W$ ,  $\text{O}_{\text{MeOH}}$ ,  $\text{N}_{\text{ACN}}$ , and the CoMs of benzene and hexane) within  $-1 \leq z \leq +1 \text{ \AA}$  were selected and further analyzed to give a 2-dimensional solvent distribution. The collected data points were then smoothed with a 2D bivariate normal kernel density estimation<sup>137,138</sup>. Solvent boxes were divided into 100 bins along the  $x$ - and  $y$ -axes to generate a  $100 \times 100$  grid. Then, the solvent density around each of the grid points was smoothed using a 2D Gaussian distribution with a width of  $1.5 \text{ \AA}$ . Isocontours were drawn at 90%, 75%, 50%, 25%, and 10% of the highest occupation for each solvent to allow direct comparison.

## 3.4 Results and Discussion

### 3.4.1 The Potential Energy Surfaces

The correlation between the fitted force field and the MP2 reference energies is reported in Figure 3.2. As the main interest of the present work is the forward process, i.e.  $\text{NH}_3 + \text{MeCl} \rightarrow [\text{H}_3\text{NMe}]^+ + \text{Cl}^-$  and  $\text{Pyr} + \text{MeBr} \rightarrow [\text{PyrMe}]^+ + \text{Br}^-$ , the parametrization was focused on the reactant state including the transition state region. For the 1000 and 1700 reference reactant state structures (see Methods) the RMSDs are 0.78 kcal/mol and 0.93 kcal/mol, respectively, for the two reactions. As the inset demonstrates, the minimum energy path (MEP) is well reproduced by these parametrizations, too. Such a quality is comparable with that found in previous studies. The reactive PES for the hydrogen abstraction reaction between  $\text{MgO}^+$  and ethane had a RMSD of 1.5 kcal/mol with respect to MP2/aug-cc-pVTZ for the reactant and 1.1 kcal/mol for the product state.<sup>139</sup> For the Diels-Alder reaction between 2,3-dibromobutadiene and maleic anhydride had a RMSD of 2.9 kcal/mol for the overall reaction.<sup>140</sup> For the product side (ion-pair) the RMSDs along the C–Cl and C–Br separations are 3.60 kcal/mol and 2.41 kcal/mol for the two reactions, respectively, and dominated by high energy structures for long C–Cl<sup>-</sup> separations.





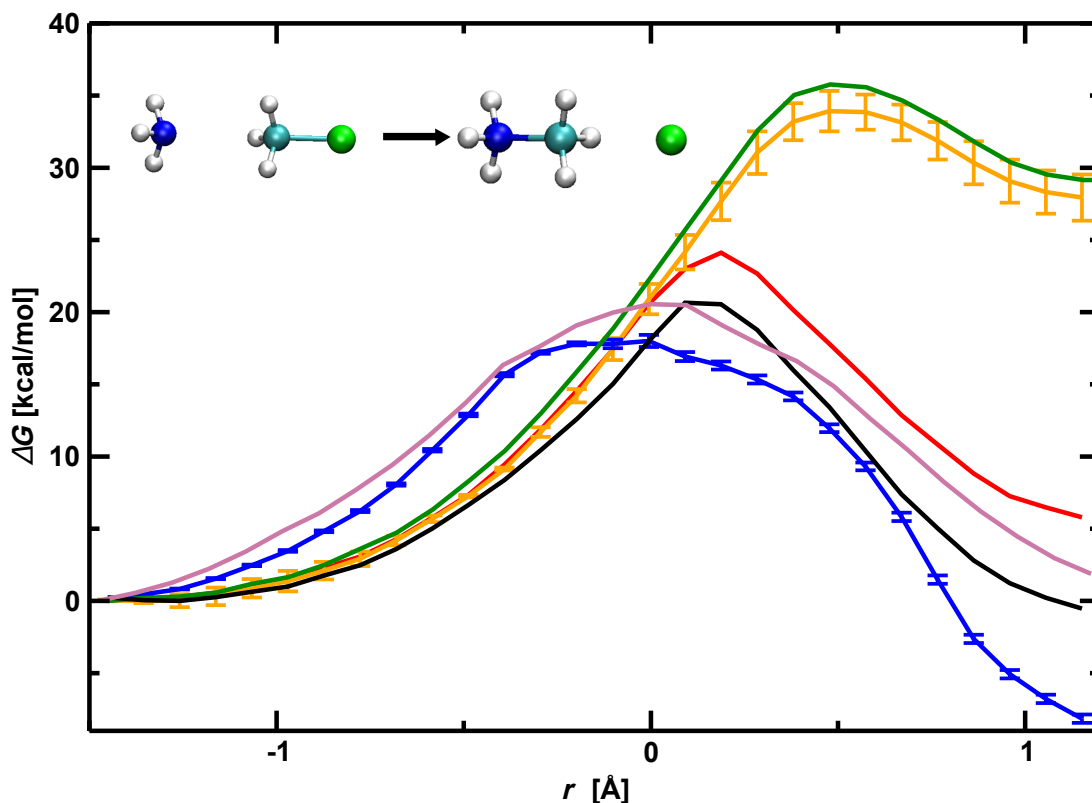
**Figure 3.2:** Panel A Main panel: Menshutkin reaction of ammonia with chloromethane to methylammonium chloride. Energy correlation between the fitted force field and the MP2 reference for 1000 reactant structures (blue, RMSD of 0.78 kcal/mol) and 70 product structures (red, RMSD of 3.60 kcal/mol). Panel B Main panel: Menshutkin reaction of pyridine with bromomethane to methyl pyridine and  $\text{Br}^-$ . Energy correlation between the fitted force field and the MP2 reference for 1700 reactant structures (blue, RMSD=0.93 kcal/mol) and 101 product structures (red, RMSD=2.41 kcal/mol). Insets: The IRC points (green circles) together with the energies from the reactive force field (black line) based on the GAPO-fits to connect the reactant and product force fields.

### 3.4.2 Umbrella Sampling Simulations and Potentials of Mean Force

*NH<sub>3</sub>+MeCl*: The 1-dimensional PMFs for the  $\text{NH}_3+\text{MeCl}$  reaction are reported in Figure 3.3 and Table 3.1 summarizes the activation free barriers. Umbrella sampling simulations based on the parametrized, reactive force field for the  $\text{NH}_3+\text{MeCl}$  reaction yields a free energy barrier  $\Delta G^\ddagger = 35.8$  kcal/mol in the gas-phase. This compares with an activation barrier of 38.5 kcal/mol from the IRC in the gas-phase at the MP2/6-311++G(2d,2p) level. Previous electronic structure calculations for the same reaction reported barrier heights of 36.2 kcal/mol and 32.7 kcal/mol at the HF/6-31++G<sup>\*118</sup> and B3PW91/6-31+G<sup>\*119</sup> levels, respectively.

	Gas	Water	Methanol	Acetonitrile	Benzene	Hexane
Sim.	35.8	$18.0 \pm 0.5$	20.5	20.6	24.1	$33.9 \pm 1.4$
Exp.		$23.5^{141}$	$20.8^{142}$			
Lit.		$16.1^{118}$				
		$21.7^{61}$				
	$36.2^{118}$	$25.6^{143}$		$16.8^{122}$		$32.6^{61}$
	$32.7^{119}$	$15.7 \pm 0.3^{121}$		$22.3^{122}$		
		$21.9 \pm 2.7^{121}$				

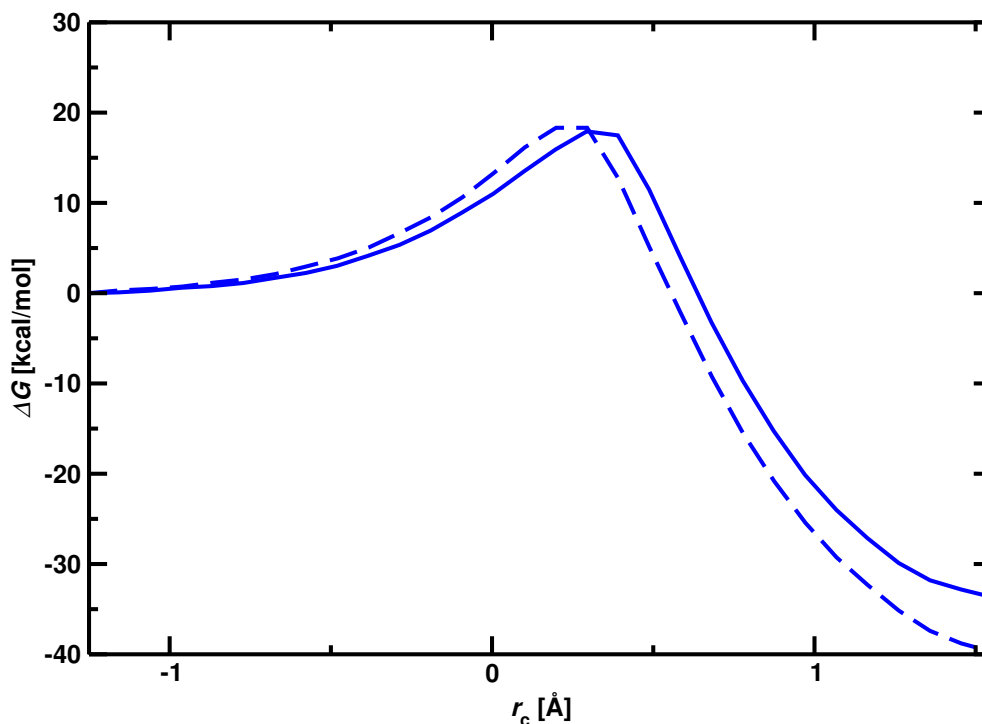
**Table 3.1:** Free energy barrier for the Menshutkin reaction of  $\text{NH}_3 + \text{MeCl}$  (in kcal/mol) calculated with MS-ARMD and US in different solvents (Sim.). The experimental (Exp.)  $\Delta G^\ddagger$  value is for the  $\text{NH}_3 + \text{MeI}$  reaction in water (0.05 M MeI)<sup>141</sup> and in methanol (0.1 M MeI).<sup>142</sup> Hexane = cyclohexane.



**Figure 3.3:** Potentials of mean force for the Menshutkin reaction for  $\text{NH}_3 + \text{MeCl}$  in water (blue), methanol (purple), benzene (red), acetonitrile (black), and cyclohexane (orange) from umbrella sampling simulations. For comparison, the PMF from US in the gas phase is also reported in green.

In water, methanol, acetonitrile, benzene, and cyclohexane the barrier heights from the 1-dimensional PMFs are  $18.0 \pm 0.5$ , 20.5, 20.6, 24.1, and  $33.9 \pm 1.4$  kcal/mol, respectively. Representative error bars were determined from boot-

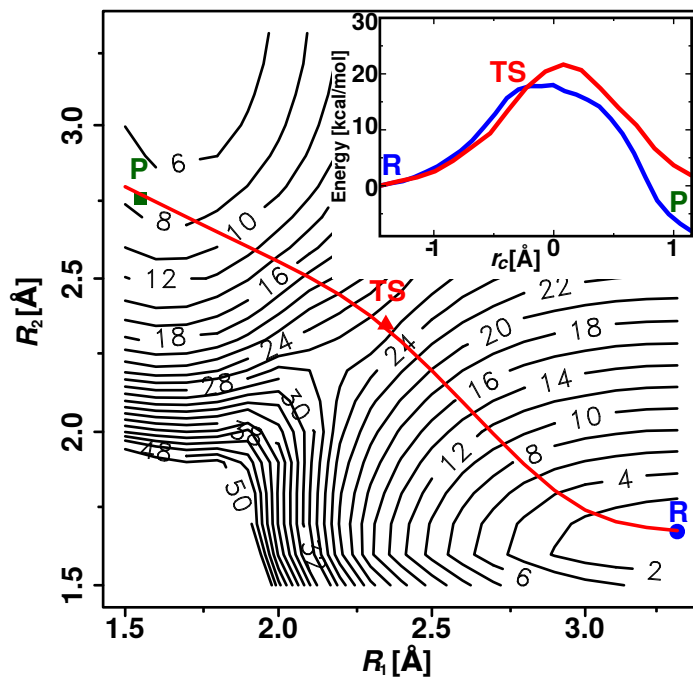
strapping for polar (water) and apolar (cyclohexane) solvent which indicate good convergence of the profiles. In addition, one simulation for Pyr+MeBr in water with 250 ps sampling per window was carried out which confirms the level of convergence; see Figure 3.4. For the simulations in water and methanol, the activation energies of  $18.0 \pm 0.5$  and  $20.6$  kcal/mol compare with the experimentally reported values of  $23.5$ <sup>141</sup> and  $20.8$  kcal/mol<sup>142</sup> for  $\text{NH}_3 + \text{MeI}$ , respectively. Only qualitative comparison between experiment and simulations is possible because a) results from experiments are only available for MeI and not MeBr, and b) the MeI concentration was 0.05 M in water and 0.1 M in methanol.<sup>141</sup> Previously, the minimum energy pathway for the  $\text{NH}_3 + \text{MeCl}$  reaction in explicit water was determined at the HF/6-31G\* level of theory, followed by single-point energy calculations at the MP4SDTQ/6-31+G\* level. The resulting activation barrier was  $25.6$  kcal/mol.<sup>143</sup> In a more recent study the barrier heights from a semi-empirical QM study were  $15.7 \pm 0.3$  and  $21.9 \pm 2.7$  kcal/mol, respectively.<sup>121</sup> The activation barrier was  $15.7 \pm 0.3$  kcal/mol when directly computed from the lower level PM3/MM calculations and increased to  $21.9 \pm 2.7$  kcal/mol by using a higher level, indirect B3LYP/MM methodology. In this method, configurational sampling and iterative pathway optimization were carried out with a lower-level PM3/MM Hamiltonian, and analyzed with WHAM to obtain the lower-level reaction free energy profile. This was followed by thermodynamic perturbation<sup>144</sup> to a higher level target B3LYP/MM Hamiltonian with 6-31G\* basis set to obtain a corrected free energy profile.<sup>121</sup>



**Figure 3.4:** Potentials of mean force for the Menshutkin reaction for Pyr + MeBr in water for 50 ps (solid) and 250 ps (dashed) simulations per umbrella window.

As expected, polar solvents show a stronger effect in reducing the reaction barrier height compared to apolar solvents; see Figure 3.3. All polar solvents and benzene find an earlier TS compared to vacuum for which it is at  $r_c = 0.5$  Å. For benzene, acetonitrile, methanol and water the TS progressively shifts towards the reactant with  $r_c$  ranging from  $r_c = 0.2$  Å to  $r_c = -0.1$  Å. Furthermore, the free energy profile widens considerably when going from the gas phase to methanol and water as a solvent and compared with acetonitrile and the apolar solvents. Water and methanol are the smallest and most polar solvents and are able to approach and interact with the solute more directly than the larger and apolar solvent molecules which leads to widening of the PMF. For benzene the packing is easier due to its planar structure compared to the chair conformation of cyclohexane. Further, benzene has stronger electrostatic interactions with the solute as a result of negatively charged carbon atoms. Thus, benzene displays a barrier

reduction closer to the polar water solvent than the apolar cyclohexane.



**Figure 3.5:** 2D PMF of  $\text{NH}_3+\text{MeCl}$  in water, scanned through reaction coordinates of  $R_1 = d_{\text{C-N}}$  and  $R_2 = d_{\text{C-Cl}}$ . Relative positions of reactant (blue circle), transition state (red triangle) and product (green square) shown on the plot. The red line denotes the minimum energy pathway from reactant to product (see text). Contours are drawn in increments of 2 kcal/mol. The inset shows the 1D PMF extracted from the 2D surface (red line) superimposed onto the 1D PMF (blue line) of  $\text{NH}_3+\text{MeCl}$  in water from Figure 3.3.

In order to assess how representative the 1-dimensional PMFs are, a 2-dimensional PMF surface was determined for  $\text{NH}_3+\text{MeCl}$  in water, see Figure 3.5. The two reaction coordinates  $R_1$  and  $R_2$  were the distance between the methyl-C and ammonia-N atom ( $R_1 = d_{\text{C-N}}$ ) and between the methyl-C and the Cl atoms ( $R_2 = d_{\text{C-Cl}}$ ), respectively. In the simulations,  $R_1$  and  $R_2$  were constrained by a harmonic potential with a force constant of 1000 kcal/mol and the N-C-Cl angle was constrained to remain linear. For a fixed  $R_1$  value the 2D PMF was scanned along the  $R_2$  coordinate and statistics accumulated from 50 ps sampling for each

value of the reaction coordinate. From this, the 2d PMF was constructed using 2D-WHAM<sup>59,60</sup> with a tolerance of 0.001.

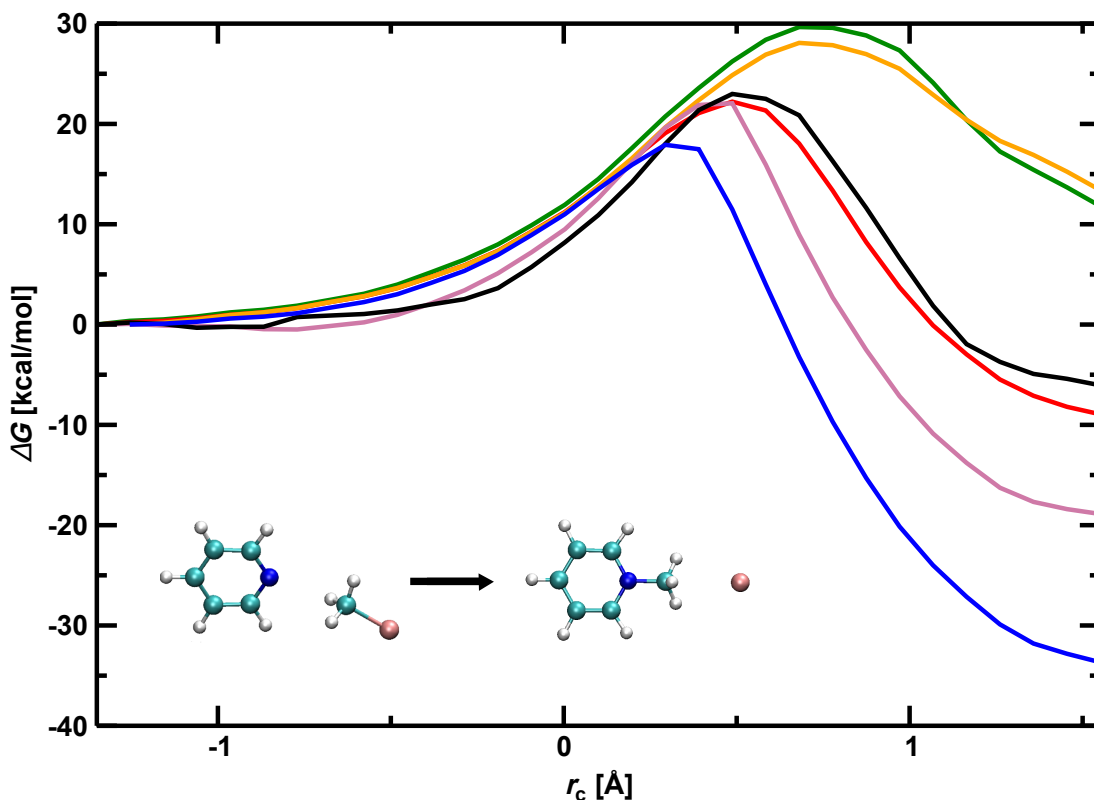
A direct comparison between the 1D PMF from the US (blue) and the minimum energy path on the 2D PMF (red) is shown in the inset of Figure 3.5. The minimum energy path on the 2D PMF was calculated by fixing  $R_2$  values on a grid with  $\Delta R_2 = 0.1 \text{ \AA}$  and minimizing along  $R_1$ . This provides a realistic path between reactant and TS. Between TS and the product the true MEP is slightly different which is, however, of lesser interest in the present work. The activation barrier calculated from the 2D PMF is 21.7 kcal/mol which is 3.7 kcal/mol higher than that of the 1D PMF (see Figure 3.5 inset) and closer to the experimentally determined activation barrier of 23.5 kcal/mol for the  $\text{NH}_3 + \text{MeI}$  reaction. The fact that the 1D and 2D PMF are quite close to one another supports the results of the 1D PMFs reported for the other solvents.<sup>141</sup> Part of the difference in the barrier heights between the 1D and 2D PMFs can be attributed to the constraints that were used in computing the 2D PMF. On the other hand, the 2D PMF yields additional insights into mechanistic aspects of the reaction. Specifically, the 2D PMF clarifies that the Menshutkin reaction is concerted from the perspective of the two reaction coordinates. As  $R_1$  contracts from the reactant (R) to the product (P),  $R_2$  elongates. Furthermore, the transition region is comparatively steep along the transition pathway but flat orthogonal to it. For separations  $R_1$  and  $R_2$  smaller than 2.2  $\text{\AA}$  a strong repulsion is found.

*Pyr+MeBr*: The 1-dimensional PMFs for the Pyr+MeBr Menshutkin reaction are reported in Figure 3.6 and Table 3.2 summarizes the activation free barriers. In the gas phase US yields an activation free energy of  $\Delta G^\ddagger = 29.7 \text{ kcal/mol}$  compared with the barrier height of 29.0 kcal/mol from the IRC at the MP2/6-

311++G(2d,2p) level of theory. In water, methanol, and acetonitrile the computed barriers are 17.9, 22.1, and 23.2 kcal/mol and in apolar solvents they are 22.2 and 28.1 kcal/mol for benzene and hexane, respectively. Experimentally reported barrier heights as determined from the measured rates for acetonitrile and hexane were 22.5 and 27.6 kcal/mol, respectively,<sup>61</sup> and the ordering and absolute values agree favourably with the present simulations.

	Gas	Water	Methanol	Acetonitrile	Benzene	Hexane
Sim.	29.7	17.9	22.1	23.2	22.2	28.1
Exp. <sup>61</sup>				22.5		27.6

**Table 3.2:** Free energy barrier for the Menshutkin reaction of Pyr+MeBr (in kcal/mol) calculated with MS-ARMD and US in different solvents (Sim.). The experimental (Exp.) values for  $\Delta G^\ddagger$  are from Ref.<sup>61</sup> for Pyr+MeBr. Hexane = cyclohexane.



**Figure 3.6:** Potentials of mean force for the Menshutkin reaction for Pyr+MeBr in water (blue), methanol (purple), acetonitrile (black), benzene (red), and cyclohexane (orange). For comparison, the result from US in the gas phase is reported in green.

For polar solvents the reaction barrier height decreases with increasing polarity of the solvent. Notably, for the reaction in benzene, the barrier height of 22.2 kcal/mol is close to that for methanol and lower than that for acetonitrile. The catalytic effect of benzene has been observed previously from computations for the Menshutkin reaction of  $\text{NH}_3 + \text{NH}_2\text{C}_6\text{H}_4\text{-Br} \rightarrow \text{NH}_2\text{-C}_6\text{H}_4\text{-NH}_3^+ + \text{Br}^-$  at the MP2/6-31+G(d) level<sup>111</sup>. In this case, the gas-phase activation barrier of 33.2 kcal/mol decreased to 22.8 kcal/mol in benzene.

The position of the TS in the different solvents is directly correlated with the catalytic effect. Cyclohexane, which displays a lower barrier to that in the gas phase has the TS at  $r_c \sim 0.7 \text{ \AA}$ , whereas solvents with catalytic effect shift the TS to progressively shorter values of  $r_c$ . Benzene, which reduces the barrier height slightly more strongly than methanol, has a TS at around  $r_c = 0.5 \text{ \AA}$  which is also the value for methanol. In water, which shows the strongest reduction in the barrier, the maximum of the PMF is shifted towards an even smaller value of  $r_c = 0.3 \text{ \AA}$ . This increase in barrier reduction and shift of the TS position to shorter separations follows the ordering of the dipole moments of the three polar solvents (water, methanol and acetonitrile) from strongest to weakest.<sup>145</sup> For cyclohexane the barrier height is nearly unaffected when compared to the gas phase whereas benzene achieves a barrier height reduction comparable to that of methanol. This is rationalized by the improved packing around the solute due to benzene's planar structure as opposed to apolar cyclohexane; see Figure 3.10.

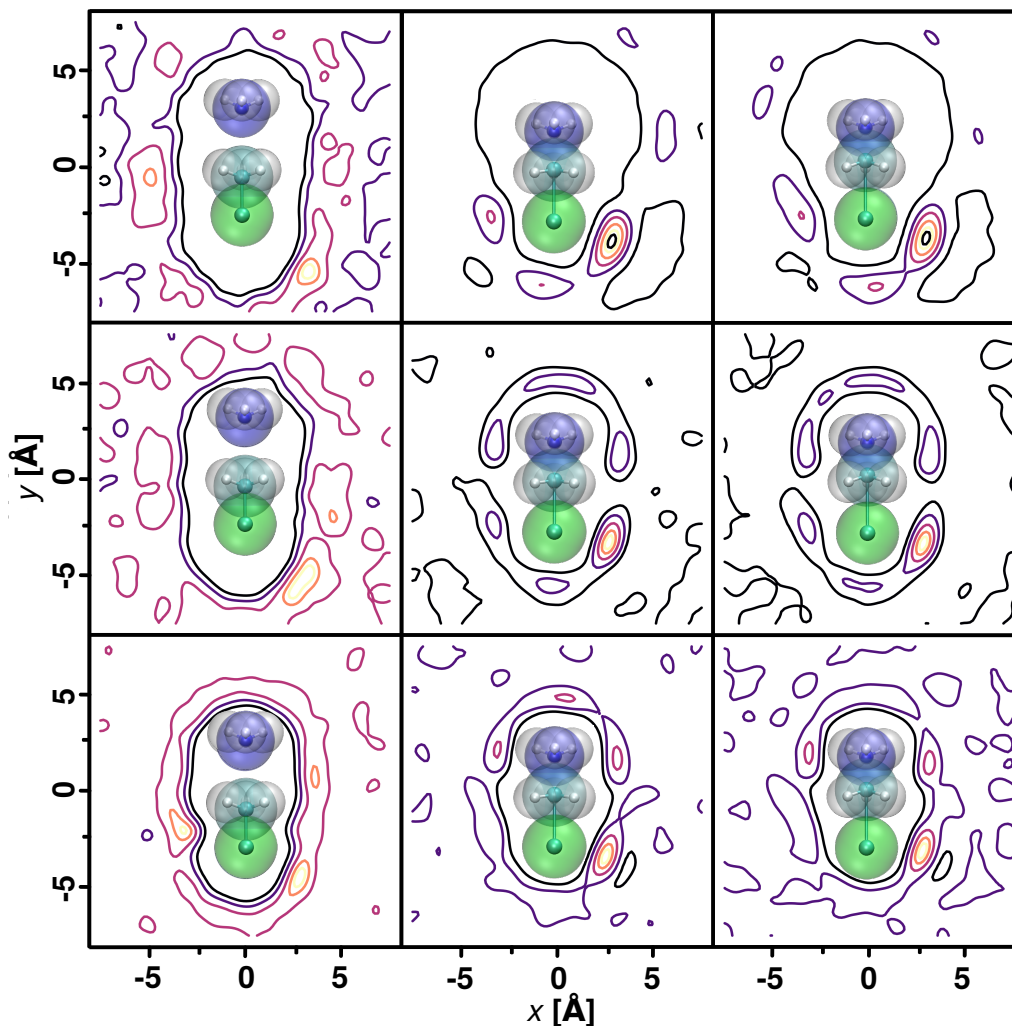
### 3.4.3 Solvent Distributions

The MD simulations also offer the opportunity to analyze the solvent distribution along the reaction at molecular resolution. For this, separate simulations were



carried out for the reactant, product, and transition states of both reactions in all solvents. The solute was constrained at the optimized geometries of the reactant, TS, and ion-pair structures from the MP2/6-311G++(2d,2p) calculations and the solvent was sampled from unbiased simulations.

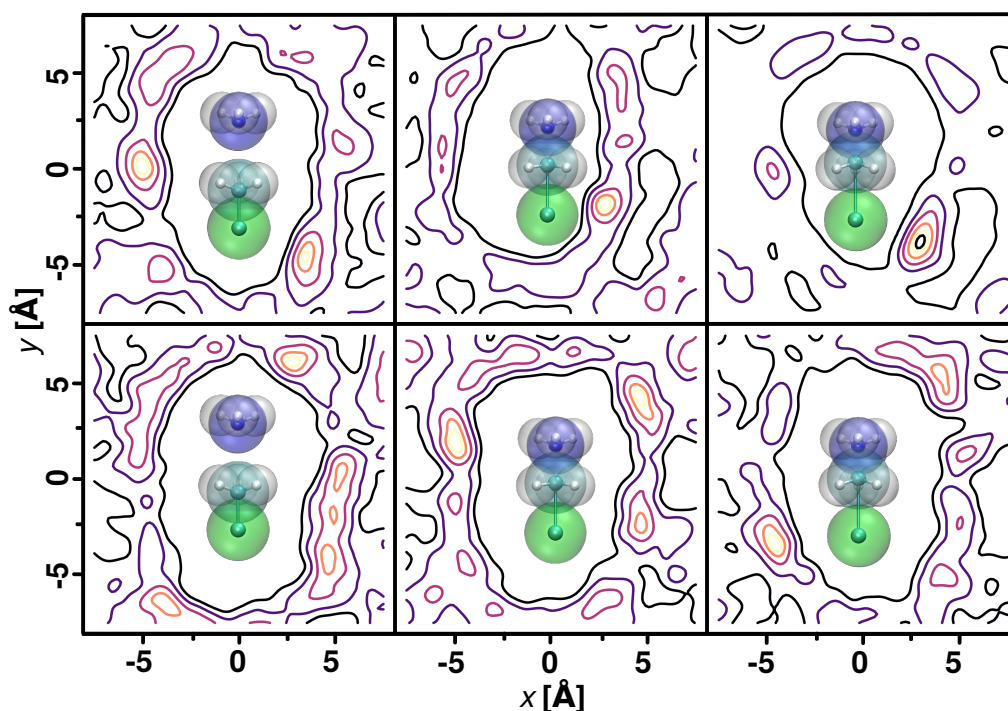
For the  $\text{NH}_3+\text{MeCl}$  reaction the solvent distribution around the solute in polar (water, methanol, acetonitrile) and apolar (benzene and hexane) solvent is presented in Figures 3.7 and 3.8, respectively. In general, for polar solvents the distribution around the reactant and the TS differs. This is most evident for acetonitrile and methanol (Figure 3.7 top and middle) but also for water (Figures 3.7 bottom). Solvent density maxima around the  $\text{Cl}^-$  anion in the TS were found for methanol, water, and acetonitrile. The asymmetry around  $\text{Cl}^-$  anion is partly due to the thin analysis slab ( $-1\text{\AA} < z < 1\text{\AA}$ ) which was necessary to avoid congestion of the solvent distribution for other parts of the molecule. The solvent density maxima are explained by the favourable interaction of the OH-hydrogen atoms (water and methanol) or the positively charged carbon atom (acetonitrile) with the  $\text{Cl}^-$  anion. For the TS in methanol and water there are three density maxima around the H-atoms of ammonia, which are potential H-bonding sites. For the reactants, the density maximum is around  $\text{Cl}^-$  in polar solvents. For all three structures, the innermost isocontour is closest to the solute for water as solvent as packing is tightest due to its small size.



**Figure 3.7:** 2-dimensional solvent distributions for  $\text{NH}_3+\text{MeCl}$  from 2 ns simulations in polar solvents: acetonitrile, methanol and water (from top to bottom) around the reactant, TS, and product state structures of the solute (from left to right) projected onto the  $xy$ -plane containing the chloride, carbon and nitrogen atoms. Units in  $\text{\AA}$ . In the simulations and the figure the solute is in its optimized structure for the reactant, TS, and product state, respectively, at the MP2/6-311++G(2d,2p) level of theory. The color code for atoms is H (white), C (cyan), N (blue) and Cl (green). Note the larger solvent rearrangement between reactant and TS compared with TS and product structures. Isocontours are reported at 90%, 75%, 50%, 25%, and 10% of the highest occupation for each solvent to allow direct comparison.

Solvent structures for the TS and the product state are comparable. This can, in part, be rationalized by the similarity of the TS and ion-pair structures of the solute and the differences between the reactant and TS structures: the N-C distances for the Cl-Me-NH<sub>3</sub> arrangement are 3.30  $\text{\AA}$ , 1.81  $\text{\AA}$ , and 1.55  $\text{\AA}$  for the

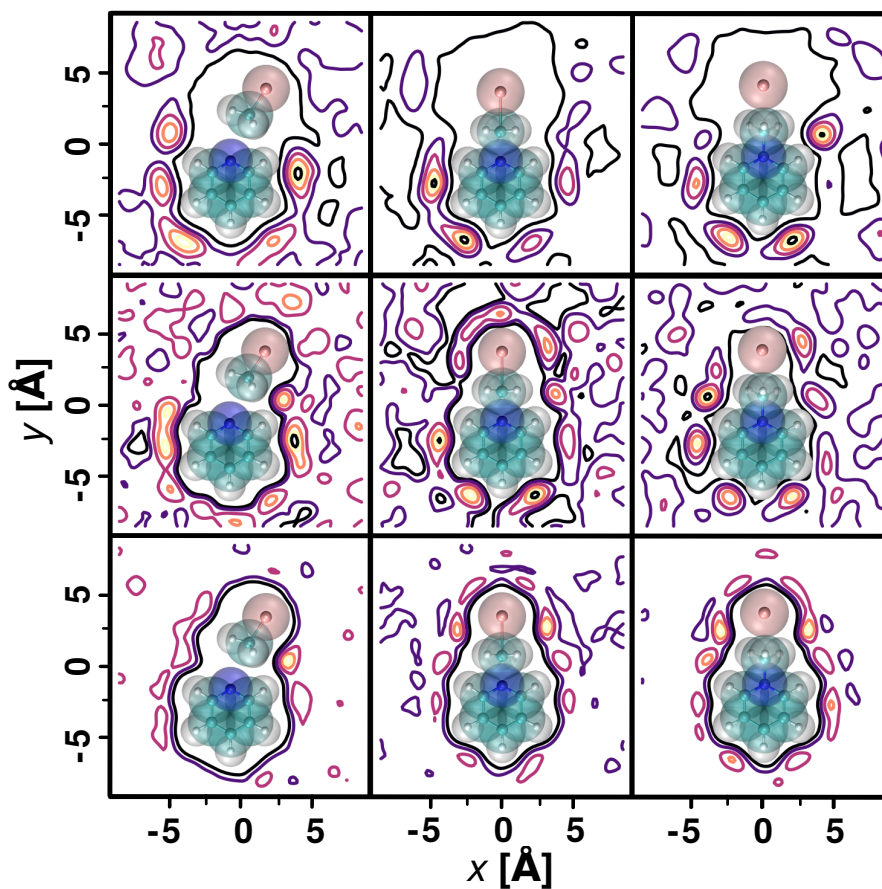
reactant, TS, and ion-pair in vacuum, respectively. Similarly, the C-Cl distances are 1.80 Å, 2.42 Å, and 2.76 Å for reactant, TS and ion-pair. This similarity between the TS and ion-pair structure is reminiscent of a manifestation of the Hammond postulate which states that if two states occur consecutively along a reaction and have nearly the same energy content, their interconversion will involve only a small reorganization of the molecular structures. For the solute structures this is evidently the case as the TS and the contact-ion-pair structures only differ little. Similarly, the solvent distributions between reactant and TS differ, in general, considerably more compared with the change in solvent structure between TS and the product.



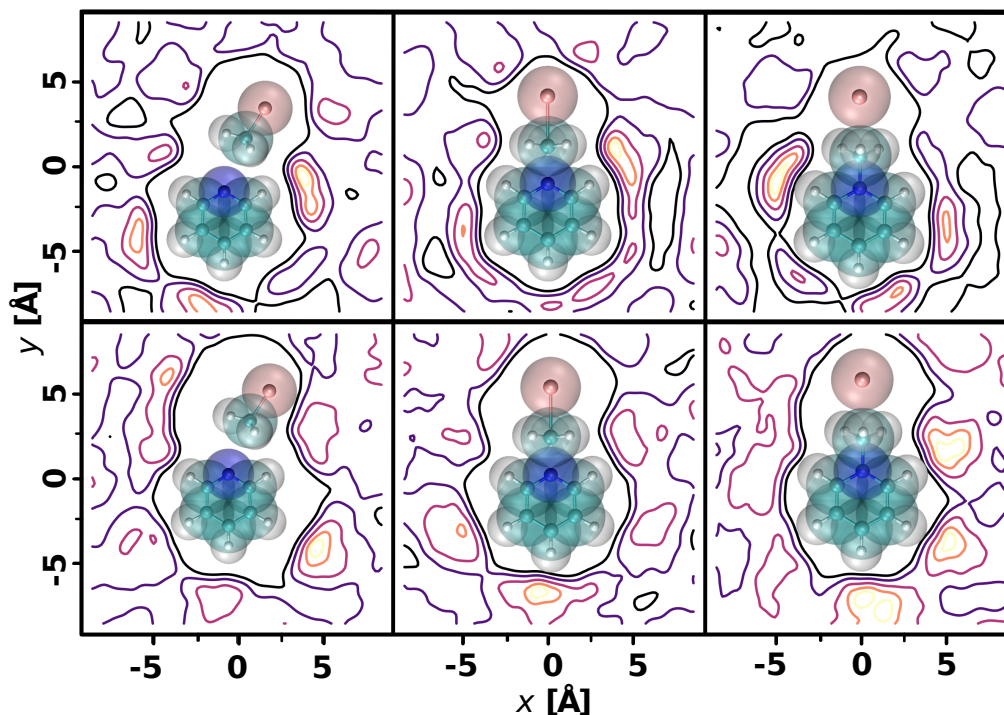
**Figure 3.8:** 2-dimensional solvent distributions for  $\text{NH}_3+\text{MeCl}$  in apolar solvents: benzene (top) and hexane (bottom) around the reactant, TS, and product state structures of the solute (from left to right) projected onto the  $xy$ -plane containing the chloride, carbon and nitrogen atoms. Units in Å. In the simulations and the figure the solute is in its optimized structure for the reactant, TS, and product state, respectively, at the MP2/6-311++G(2d,2p) level of theory. The color code for atoms is H (white), C (cyan), N (blue) and Cl (green). Isocontours are reported at 90%, 75%, 50%, 25%, and 10% of the highest occupation for each solvent to allow direct comparison.

For the apolar solvents, the distribution around the solute showed less pronounced maxima, see Figure 3.8, except for the product state in benzene. For cyclohexane the solvent distribution is considerably less structured than for the polar solvents which indicates the weak interaction between solute and solvent in this case and is the underlying reason for the limited catalytic effect of cyclohexane for the  $\text{NH}_3+\text{MeCl}$  reaction. Contrary to cyclohexane, the density maximum of solvent molecules around the  $\text{Cl}^-$  anion for the product in benzene, and to some degree for the TS, were present. The packing around the solute was more dense in benzene, especially at the TS.

For the  $\text{Pyr}+\text{MeBr}$  reaction the solvent distribution around the solute in polar and apolar solvents is presented in Figures 3.9 and 3.10, respectively. Similar to the  $\text{NH}_3+\text{MeCl}$  reaction, the solvent distribution around the reactant and the TS differed. For the reactant the  $\text{N}_{\text{Pyr}}-\text{C}-\text{Br}$  angle deviates from linearity whereas in the TS and for the contact-ion-pair it was linear. Due to the positively charged pyridine-H atoms multiple density maxima around the ring were observed for all states in polar solvents. In methanol, the solvent distribution was heterogeneous, with multiple density maxima around the solute and beyond the first solvation shell. Compared with acetonitrile, methanol is more tightly packed around the solute. Contrary to that, the solvent distribution for water was comparatively homogeneous. Although multiple density maxima for water near the solute are observed for the first solvation shell, the solvent distribution becomes more homogeneous beyond the first solvation shell for all three states of the solute. For the contact-ion-pair solvent structuring occurs for water and methanol, but not for acetonitrile.



**Figure 3.9:** 2-dimensional solvent distributions for Pyr+MeBr from 2 ns simulations in polar solvents: acetonitrile, methanol and water (from top to bottom) around the reactant, TS, and product state structures of the solute (from left to right) projected onto the  $xy$ -plane containing the chloride, carbon and nitrogen atoms. Units in Å. In the simulations and the figure the solute is in its optimized structure for the reactant, TS, and product state, respectively, at the MP2/6-311++G(2d,2p) level of theory. The color code for atoms is H (white), C (cyan), N (blue) and Br (pink). Isocontours are reported at 90%, 75%, 50%, 25%, and 10% of the highest occupation for each solvent to allow direct comparison.



**Figure 3.10:** 2-dimensional solvent distributions for Pyr+MeBr in apolar solvents: benzene (top) and hexane (bottom) around the reactant, TS, and product state structures of the solute (from left to right) projected onto the  $xy$ -plane containing the chloride, carbon and nitrogen atoms. Units in Å. In the simulations and the figure the solute is in its optimized structure for the reactant, TS, and product state, respectively, at the MP2/6-311++G(2d,2p) level of theory. The color code for atoms is H (white), C (cyan), N (blue) and Br (pink). Isocontours are reported at 90%, 75%, 50%, 25%, and 10% of the highest occupation for each solvent to allow direct comparison.

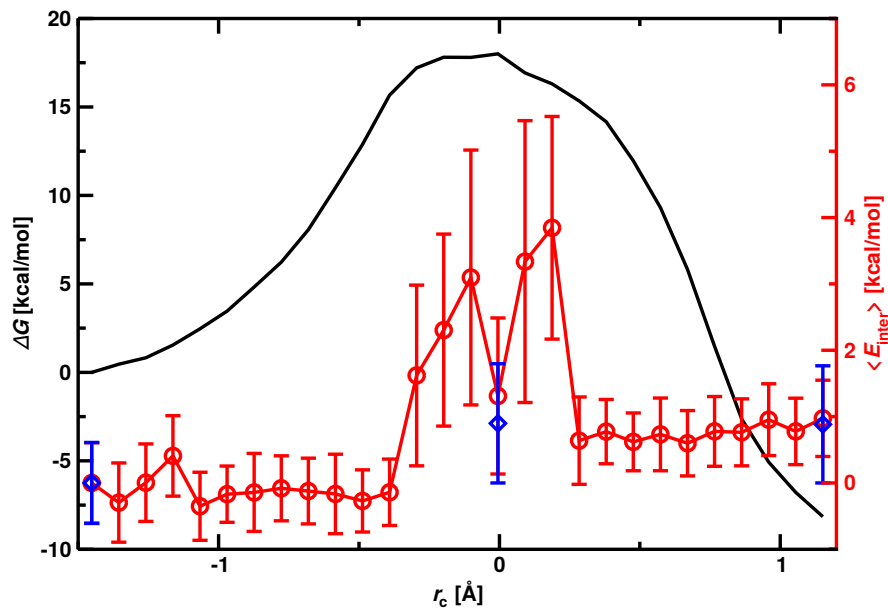
For apolar solvents, see Figure 3.10, the solvent distribution around the solute was less dense compared to the polar solvents. The maxima around the positively charged hydrogen atoms appear for benzene. However, the densities around the  $\text{Br}^-$  anion were less pronounced than for polar solvents. The packing around the  $\text{Br}^-$  anion was tighter for the TS and contact-ion-pair compared with the reactant state. Cyclohexane packs more densely around the solute for Pyr+MeBr compared to  $\text{NH}_3+\text{MeCl}$ . This can be explained by the larger size of the solute, which allows the bulkier, apolar solvent to accommodate better, than around a smaller solute. The trend of similar solvent distribution around TS and product

is also found for these two solvents.

### 3.4.4 Solvent Energetics

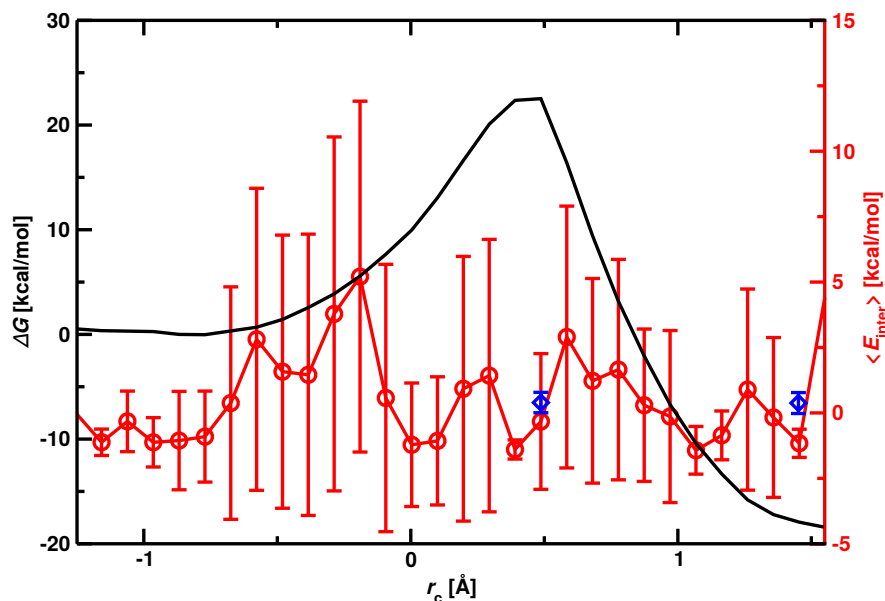
The simulations also provide quantitative information about the relative solvent-solvent interactions along the reaction pathway with respect to an arbitrary reference state which is the reactant in the following. Two different analyses were carried out for the energetics of solvent molecules as a function of the reaction coordinate. First, the energetics of the solvent molecules within the first solvation shell, i.e. molecules in direct contact with the solute, were investigated as a function of the reaction coordinate from the umbrella sampling simulations. Secondly, a similar analysis was carried out from 2 ns simulations with the solute frozen in the reactant, TS, and contact-ion-pair geometries at the MP2/6-311G++(2d,2p) level of theory for the first solvation shell.

For the analysis only solvent molecules within a cutoff distance of 5 Å of any atom of the solute were retained which corresponds approximately to the first solvation shell for each solvent. For each of the umbrellas from the US simulations the average solvent-solvent interaction energy per solvent molecule was determined together with the fluctuation around the mean for the  $\text{NH}_3+\text{MeCl}$  reaction in water (see Figure 3.11) and for the  $\text{Pyr}+\text{MeBr}$  reaction in methanol (see Figure 3.12).



**Figure 3.11:** Potentials of mean force for the Menshutkin reaction for  $\text{NH}_3+\text{MeCl}$  in water (black line) and the energy of solvent molecules within  $5 \text{ \AA}$  of the solute in kcal/mol/molecule (red line). The circles are the mean for a given umbrella together with the standard deviation from the mean as a bar. The blue diamonds are the average solvent interaction energies from the 2 ns simulations with the frozen solute (reactant, TS and contact-ion-pair) structures together with the standard deviation from the mean as a bar.



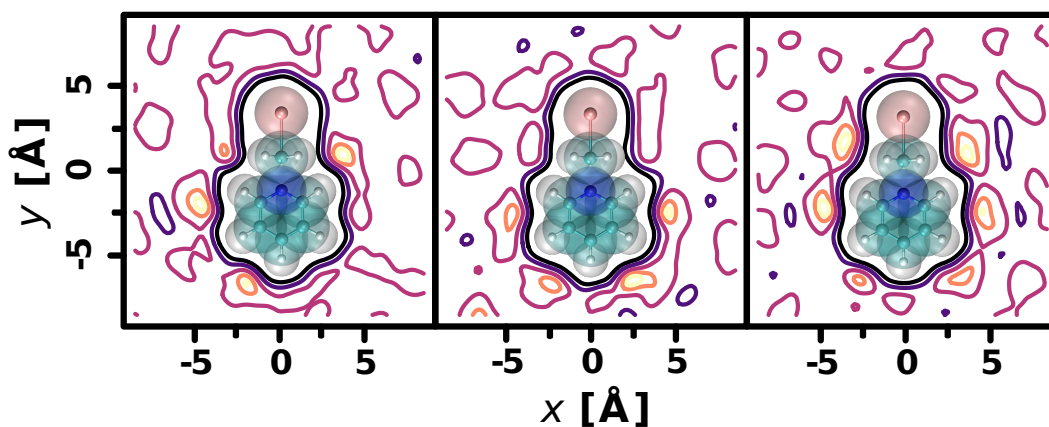


**Figure 3.12:** Potentials of mean force for Pyr+MeBr in methanol (black line) and the average solvent-solvent interaction for molecules  $5\text{\AA}$  around the solute in kcal/mol/molecule (red line). The circles are the mean for a given umbrella together with the standard deviation from the mean (fluctuation bar). The blue diamonds are average solvent-solvent interaction energies for molecules  $5\text{\AA}$  around the solute in kcal/mol/molecule from 2 ns *NVT* simulations in reactant, TS and product states, respectively, together with the standard deviation (fluctuation bar).

The average energy per solvent molecule with respect to the reactant within  $5\text{\AA}$  (red trace) superimposed on the PMF of  $\text{NH}_3+\text{MeCl}$  in water (black) is shown in Figure 3.11. Starting from the reactant structure at  $r_c = -1.2\text{\AA}$  the average interaction energy remains around  $\sim -0.5$  kcal/mol/molecule up to  $r_c \sim -0.4\text{\AA}$  after which it suddenly increases. Before the transition state the average interaction energy per water molecule is destabilized by  $\sim 2.5$  kcal/mol to accommodate the transition state of the solute. At the transition state,  $r_c = 0$ , the solvent is destabilized with respect to the reactant by about 1 kcal/mol per water molecule (but stabilized relative to the structures before the TS by about 1.5 kcal/mol). Beyond the TS the surrounding water molecules arrange again in a more unfavourable conformation before their average interaction energies return to levels comparable to that of the reactant of around  $\sim 0.8$  kcal/mol at  $r_c \sim 0.3\text{\AA}$ . This

demonstrates that the water solvent participates actively in the reaction progress.

For the Pyr+MeBr reaction in methanol the average energy per solvent molecule with respect to the reactant within 5 Å is shown in Figure 3.12. Starting from the reactant structure the average energy per solvent molecule remains at  $\sim -1$  kcal/mol/molecule up to  $r_c \sim -0.8$  Å. Upon approaching the transition state ( $r_c \sim -0.5$  Å) the average interaction energy between the solvent molecules increases (destabilization). Compared with the NH<sub>3</sub>+MeCl reaction in water the destabilization of the solvent starts earlier, though. Before reaching the TS at  $r_c \sim 0.4$  Å the average interaction energy decreases and reaches a value similar to that of the reactant albeit with larger variance. Between the TS and the product structure the interaction energy increases again and then drops back to levels of the reactant structure. Since the TS and product states are structurally similar to each other (see Figure 3.9 and its discussion), the effects on the average solvent energetics are less prominent in the later stage of the reaction.



**Figure 3.13:** 2-dimensional solvent distribution for Pyr+MeBr in methanol for  $r_c = [-0.3, -0.2, -0.1]$  projected onto the  $xy$ -plane, respectively. The color code for atoms is H (white), C (cyan), N (blue) and Br (pink). Isocontours are reported at 90%, 75%, 50%, 25%, and 10% of the highest occupation for each solvent to allow direct comparison.

The 2D-solvent distributions for  $r_c = [-0.3, -0.2, -0.1]$  Å are shown in Figure 3.13. It is found that the density maxima around MeBr decrease in amplitude as the reaction proceeds from reactant towards TS, especially for the region between the methyl-H and the Pyr-H. The protic nature of methanol promotes H-bonding which results in density maxima around the region between the methyl-H and the Pyr-H. As the isocontours are drawn at the same heights in all representations, the populations are directly comparable. It is found that as the system approaches the region with  $r_c = -0.2$  Å from either the reactant or the TS side, the ordering decreases appreciably thus, correlated motions of solvent molecules are required in this region of the reaction profile. This partly explains in a “time lapse” picture why reactions “take more time” (ps and longer) than the actual reactive step (which is rather fs). which is reflected in the energetic destabilization of the solvent at this position relative to the reactant or TS structure. Hence, changes in the solvent distribution are directly reflected in the average solvent-

solvent energetics.

Again, the more extended 2 ns *NVT* simulations with the solute frozen in its reactant, TS, and product state geometry were analyzed and the average interaction energies for solvent molecules within 5 Å of the solute were determined, see Tables 3.3 and 3.4, and the blue diamonds in Figure 3.12. These results compare favourably with the analysis of the US simulations and indicate that the sampling from the US simulations is representative.

Solvent	Reactant	TS	Product
Water	0 (0.06)	0.06 (0.06)	0.08 (0.06)
Water-45	0 (0.93)	0.07 (1.05)	0.09 (1.04)
Methanol	0 (0.65)	0.30 (0.41)	0.41 (0.32)
Acetonitrile	0 (0.28)	0.02 (0.28)	0.10 (0.27)
Benzene	0 (2.03)	0.67 (6.30)	0.49 (2.34)
Cyclohexane	0 (1.87)	0.21 (2.26)	0.24 (2.22)

**Table 3.3:** Solvent-solvent interaction for  $\text{NH}_3+\text{MeCl}$  for solvent molecules within 5 Å of any of the solute atoms. The average energy per solvent molecule of TS and product states with respect to the reactant, in kcal/mol. The standard deviation from the mean per solvent molecule given in parentheses. Water-45 labels the simulations in the 45 Å water box.

In addition, one simulation for the  $\text{NH}_3+\text{MeCl}$  reaction in a larger 45 Å cubic box (Water-45 in Table 3.3) was carried out to assess the influence of the box size on the solvent energetics. The solvent-solvent interaction energy per molecule for the entire 45 Å box was  $0.01 \pm 0.04$  kcal/mol and  $0.04 \pm 0.08$  kcal/mol compared with  $0.09 \pm 0.06$  kcal/mol and  $0.10 \pm 0.06$  kcal/mol in the smaller box for TS and product state relative to the reactant state, respectively. This is an insignificant difference and also confirms the results for the analysis of the first solvation shell in the 30 Å box, see Table 3.3. Finally, when analyzing the solvent-solvent energies in the first solvation shell from the simulation in the 45 Å box the average

interaction energies closely follow the results from simulations in the 30 Å box, see Table 3.3. However, a significant change is in the amplitude of the fluctuation around the average which increases by an order of magnitude. One reason for this may be the fact that water at the solute/bulk water interface is more affected in the larger box compared with the situation in the smaller box because bulk water is not really established in the smaller system. As the reported energies are all referenced to the reactant state but fluctuations are reported as absolute numbers, such effects show up more clearly in the fluctuations rather than in the average energies. These are effects worth to be explored but outside the scope of the present work.<sup>146</sup>

Solvent	Reactant	TS	Product
Water	0 (0.23)	0.61 (0.53)	0.62 (0.55)
Methanol	0 (0.13)	0.40 (0.39)	0.38 (0.40)
Acetonitrile	0 (1.10)	2.13 (1.16)	2.18 (1.20)
Benzene	0 (3.88)	0.65 (4.71)	0.70 (4.64)
Cyclohexane	0 (1.84)	0.19 (2.15)	0.25 (2.23)

**Table 3.4:** Solvent-solvent interaction for Pyr+MeBr for solvent molecules within 5 Å of any of the solute atoms. The average total energy of solvent boxes of TS and product per solvent molecule with respect to the reactant, in kcal/mol. The standard deviation from the mean per solvent molecule given in parentheses.

## 3.5 Conclusion

The energetics and solvent distributions for two Menshutkin reactions are quantitatively characterized and analyzed at molecular detail. Barrier height reductions in going from the gas phase to more polar solvents are consistent with what is known from experiments. The solvent distributions change appreciably between reactant, transition state, and product states. These changes in solvent structure are also reflected in the average solvent-solvent interactions. One notable feature of the solvent-solvent interactions is the fact that the fluctuation around the mean

increases considerably when going from small to larger solvent molecules.

Starting from the reactant structure the average solvent-solvent interaction within the first solvation shell ( $\sim 5 \text{ \AA}$  within the solute) remains small and constant up to about 50 % of the barrier height after which it increases rapidly (see Figures 3.11 and 3.12). Around the transition state the solvent-solvent strain relaxes, increases again and then returns to levels slightly higher than that of the reactant state. This points towards an intimate interplay between solute and solvent degrees of freedom along the reaction coordinate. It is also of interest to note that the fluctuations around the mean increase appreciably around the TS. Analysis of the solvent distributions also suggests that in approaching the TS collective motions of the solvent molecules are required. This may be one of the reasons why time scales for reactions can differ dramatically from the actual time to cross the barrier. It will be of interest to compare these findings with those from unbiased simulations. However, to obtain sufficient statistics a large number of reactive trajectories will be required, and ideally a system with a lower activation barrier is considered.

In summary, a computationally tractable and qualitatively correct description - as compared with the few experimental data available - of the Menshutkin reaction the gas phase and in various solvents has been presented. Analysis of the solvent degrees of freedom point towards tight coupling between solute and solvent dynamics with increased fluctuations in the solvent-solvent interactions around the transition state. The necessary solvent reorganization between reactant and TS structures of the solvent requires extensive sampling which is reflected in the time scale separation between the true rate of a reaction in solution and the time

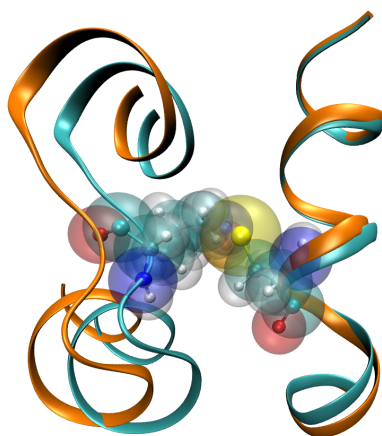
required to traverse the barrier.





# Chapter 4

## Spectroscopy, Dynamics and Hydration of S-Nitrosylated Myoglobin



---

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

*J. Phys. Chem. B. 2021, 125, 4262-4273*

*doi: 10.1021/acs.jpcc.0c10353*

\*\*\*



## 4.1 Abstract

S-nitrosylation, the covalent addition of NO to the thiol side chain of cysteine, is an important post-translational modification that can alter the function of various proteins. The structural dynamics and vibrational spectroscopy of S-nitrosylation in the condensed phase is investigated for the methyl-capped cysteine model system and for myoglobin. Using conventional point charge and physically more realistic multipolar force fields for the -SNO group it is found that the SN- and NO-stretch and the SNO-bend vibrations can be located and distinguished from the other protein modes for simulations of MbSNO at 50 K. The finding of stable cis- and trans-MbSNO is consistent with experiments on other proteins as is the observation of buried -SNO. For MbSNO the observed relocation of the EF loop in the simulations by  $\sim 3 \text{ \AA}$  is consistent with the available X-ray structure and the conformations adopted by the -SNO label are in good overall agreement with the X-ray structure. Despite the larger size of the -SNO group compared with -SH, MbSNO recruits more water molecules in the first two hydration shells due to stronger electrostatic interactions. Similarly, when comparing the hydration between the A- and H-helices they differ by up to 30 % between WT and MbSNO. This suggests that local hydration can also be significantly modulated through nitrosylation.

## 4.2 Introduction

Nitric oxide (NO) is a cell-signaling molecule relevant to function in the cardiovascular, nervous and immune systems.<sup>147</sup> Due to its high biological activity and diffusibility, NO plays an important role in many biological functions ranging from immune response, regulation of blood pressure to its function as a neurotransmitter.<sup>148,149</sup> Furthermore, its reversible binding to ferrous or ferric heme iron is well characterized.<sup>150</sup> One of the well known effects of NO on muscle tissue is the

activation of guanylate cyclase by the binding of NO to the Heme group, which leads to relaxation of the smooth muscle.<sup>151</sup> In addition, NO can also play an important role in post-translational modifications (PTM) of its target protein.<sup>152</sup>

S-Nitrosylation, i.e. the covalent addition of NO to the thiol side chain of cysteine, is an important PTM that mediates signal transduction. More than 3000 proteins, which are responsible for a wide range of cellular functions, have been identified as candidates for S-nitrosylation under physiological conditions.<sup>153</sup> S-nitrosylation is reversible, precisely targeted and regulated by temporal and spatial arrangements. Specificity of S-Nitrosylation can be governed by acid-base motifs i.e. electrostatic interactions that affect the  $pK_a$  of thiol, and the physiological concentration of NO.<sup>154</sup> Further, the relative hydrophobicity of the surrounding region of the thiol may provide 'hydrophobic motifs' that affect solvent and co-factor accessibility.<sup>155</sup> Moreover, S-nitrosylation is known to alter the function of a protein via allosteric regulation<sup>156-159</sup> and can inhibit or promote the formation of sulfide linkage within or between proteins.<sup>155</sup>

A number of studies have been proposed to characterize the underlying molecular mechanism of S-Nitrosylation. One of them is based on S-N decomposition that leads to thiyl radical formation, and thiol deprotonation as a first step.<sup>160</sup> Alternatively, a mechanism by which NO reacts directly with reduced thiol to generate a radical intermediate was proposed<sup>161</sup>, or  $\text{NO}_2$  reacts with thiol to produce thiyl radical intermediate, subsequently attacked by NO radical to form nitrosothiol.<sup>162</sup> Finally, co-localization of nitric oxide synthase (NOS) enzymes with target protein found to be a determinant of S-nitrosylation under physiological conditions.<sup>163</sup>

Beside the actual mechanism for SNO formation, the detection of the final product is equally important. A wide range of indirect detecting techniques have been proposed, considered, and applied. They include the biotin-switch technique<sup>36</sup>, His-tag switch<sup>164</sup> or fluorogenic probes<sup>165</sup> which can be utilized to infer S-nitrosylation. Indirect methods usually break the S-N bond and capture the signals of the sulfur or nitrogen parts, instead of addressing the SNO adducts directly. However, such techniques have limitations and drawbacks in terms of their selectivity and reproducibility.<sup>38</sup> Indirect methods need sequential manipulation before final analysis and can be hampered by decreasing chemoselectivity in each step.<sup>39</sup> Also, interference of other species such as  $\text{NO}_2^-$  can yield artifacts in quantitative measurements.<sup>166</sup>

Direct detection techniques such as gold nanoparticles (AuNP), phenylmercury, organophosphine or benzenesulfinate probes target the intact SNO moieties.<sup>39</sup> One limitation is incomplete free thiol blockage which can lead to false positive results with AuNP.<sup>167</sup> Also, depending on the compound used as a probe, the general applicability can be limited due to toxicity, as is the case with phenylmercury.<sup>39</sup> Further, it was reported that indirect sunlight can lead to reduction of biotin-HPDP to biotin-SH, and cause false-positives.<sup>168</sup>

An alternative direct technique to demonstrate that S-nitrosylation has occurred is infrared (IR) spectroscopy. IR spectroscopy is a widely used technique for the structural characterization of proteins.<sup>169,170</sup> Conformationally sensitive vibrations within the peptide backbone of the protein are the amide I (1600 to 1700  $\text{cm}^{-1}$ ), amide II (1510 to 1580  $\text{cm}^{-1}$ ), and amide III (1200 to 1350  $\text{cm}^{-1}$ ) bands which report on the peptide backbone vibrations. As such, the method provides a structural fingerprint by which target proteins can be identified. Similarly, the

SNO probe has the SN stretch, SNO bending and NO stretch modes which provide potentially useful signatures in the infrared to establish that S-nitrosylation has occurred. The main advantage of SNO is the chemical uniqueness of the probe. Only cysteine and methionine residues contain a sulfur atom and SN single and NO double bonds are absent in unmodified amino acids. On the other hand the frequency of the NO stretch mode at  $\sim 1525 \text{ cm}^{-1}$  overlaps with the amide II band which makes it potentially challenging for direct detection.<sup>171</sup> There are also experimental difficulties to measure the IR spectrum of proteins in water. Due to the overlap between the HOH bending vibration of water and the conformationally sensitive amide II region at  $\sim 1600 \text{ cm}^{-1}$ , the signals can be considerably improved by recording difference spectra for the  $^{14}\text{NO}$  and  $^{15}\text{NO}$  isotopes<sup>172</sup>, by subtracting the water background<sup>173,174</sup>, or by subtracting the spectrum of the WT system.

Several studies examined the spectral<sup>175–181</sup> and structural<sup>182–184</sup> properties of S-Nitrosothiols in various media. The vibrational spectra of S-nitrosoglutathione (GSNO) were measured in the solid state at room temperature<sup>171,176</sup> and reported NO-related frequencies from Raman spectroscopy at  $[1385 \text{ cm}^{-1}, 1526 \text{ cm}^{-1}]$ <sup>171</sup> and at  $[1467 \text{ cm}^{-1}, 1479 \text{ cm}^{-1}]$  from Raman and IR spectroscopy,<sup>176</sup> respectively. The assignments were based on difference spectra between glutathione (GSH) and GSNO. Depending on the RSNO compound considered, the frequencies from experiments in KBr pellets in the infrared can range from  $1479 \text{ cm}^{-1}$  to  $1514 \text{ cm}^{-1}$  and those from Raman spectroscopy cover  $1467 \text{ cm}^{-1}$  to  $1513 \text{ cm}^{-1}$ , slightly shifted to lower frequency (red shift).<sup>176</sup> GSNO in the solid state was studied as an analogue of S-nitrosylated Cysteine<sup>171</sup> which is also the reason why comparison in the present work is made with the band at  $1526 \text{ cm}^{-1}$ .<sup>171</sup> Also, the IR spectra for ethyl thionitrites were studied<sup>177</sup> in the gas phase for both, the cis- and trans-conformers. The NO stretch frequency was at  $1537 \text{ cm}^{-1}$  for the cis-

and at  $1559\text{ cm}^{-1}$  for the trans- conformer, respectively. Hence, the frequency of the trans-conformer is shifted to higher frequency (blue shift) with respect to the cis-orientation by  $22\text{ cm}^{-1}$ . For S-nitrosoglutathione (GSNO) in solution at room temperature only the cis-conformer was reported with a stretch frequency at  $1497\text{ cm}^{-1}$ .<sup>185</sup> Experiments on  $[\text{GSNO}+\text{H}]^+$  in the gas phase using IRMPD spectroscopy reported the NO stretch signal at  $1690\text{ cm}^{-1}$ .<sup>178</sup> Accompanying calculations at the DFT level of theory found a pronounced sensitivity of this band to the conformation for which the normal mode calculations were carried out. For 4 different conformers considered, the NO stretch covers a range from  $1640\text{ cm}^{-1}$  to  $1714\text{ cm}^{-1}$ .<sup>178</sup>

Isotopic substitution ( $^{14}\text{N}\rightarrow^{15}\text{N}$  and  $^{16}\text{O}\rightarrow^{18}\text{O}$ ) can also be used to assign vibrations and to determine their sensitivity to the local environment. Several studies have reported the vibrational spectroscopy of, mostly charged, isotopically-labeled S-nitroso analogues.<sup>176,178,186,187</sup> For experiments in the gas-phase, ions of S-nitrosocysteine (CysNO) were generated by electrospray ionization and probed via infrared multiple photon dissociation (IRMPD) spectroscopy.<sup>186</sup> The NO stretch for the  $[\text{CysNO}-\text{H}]^-$  anion was found at  $[1460,1488]\text{ cm}^{-1}$  which shifted to the red  $[1418,1454]\text{ cm}^{-1}$  upon  $^{15}\text{N}$  substitution, i.e. by  $\sim -40\text{ cm}^{-1}$ . Contrary to this, for the  $[\text{CysNO}-\text{H}]^+$  cation, a blue shift of  $2\text{ cm}^{-1}$  from ( $1716\text{ cm}^{-1}$  to  $1718\text{ cm}^{-1}$ ) was reported although the IR spectrum for the  $^{14}\text{N}$  isotope was not directly observed because the band was deemed covered by another vibration.<sup>186</sup> These results show that both, the absolute wavenumber and the isotope-induced shift for nitrosylated Cysteine depend on the charge of the species. From these results it is conceivable to expect a) the NO-stretch in neutral CysNO midway between  $1460\text{ cm}^{-1}$  and  $1716\text{ cm}^{-1}$  (consistent with reported frequencies of  $\sim 1525\text{ cm}^{-1}$  for neutral analogues in the solid state) and b) an isotope-induced red shift of an estimated  $\sim -20\text{ cm}^{-1}$  upon  $^{14}\text{N}\rightarrow^{15}\text{N}$  replacement. It has, however, to be

pointed out that no IR or Raman experiments are available for neutral CysNO in the gas phase or in solution.

Computational methods provide a viable alternative to investigate the relationship between structure, spectroscopy, and dynamics.<sup>188</sup> Earlier investigations of the thiol group included electronic structure studies of small model systems, such as CH<sub>3</sub>SNO<sup>182,183</sup> or empirical force fields based on harmonic potentials using point charge (PC) models.<sup>189,190</sup> In the present study, the spectral and structural properties of wild type (WT) blackfin tuna myoglobin (Mb) and its S-Nitrosylated analogue are studied by means of molecular dynamics simulations using Morse potentials and point charge and multipolar (MTP) force fields for the electrostatics. Two conformations of S-Nitrosylated myoglobin, cis-MbSNO and trans-MbSNO, with respect to the C<sub>β</sub>SNO angle are considered. The aim is to characterize the absorption features in the infrared and the structural effects induced by S-Nitrosylation of Cys10. Changes in the local structure are potentially important for modifications in the function of a protein. Hence, the global and local dynamics of the protein and its helices and loops, and changes in the solvent shells around the protein are investigated.

The present work is structured as follows. First, the force field parametrization and the atomistic simulations are described. This is followed by the discussion of calculated IR and power spectra of the CysNO model system. Then, the IR spectra of WT, cis-MbSNO and trans-MbSNO are presented and discussed. Finally, structural effects induced by S-Nitrosylation are discussed and conclusions are drawn.



## 4.3 Computational Methods

### 4.3.1 Molecular Dynamics

All molecular Dynamics (MD) simulations were performed using the CHARMM<sup>123</sup> software with the CHARMM36<sup>191</sup> force field. The equations of motion were propagated with a leapfrog integrator<sup>192</sup>, using a time step of  $\Delta t = 1$  fs and all bonds involving hydrogen atoms were constrained using SHAKE.<sup>193</sup> Non-bonded interactions were treated with a switch function<sup>194</sup> between 12 and 16 Å and electrostatic interactions were computed with the particle mesh Ewald method.<sup>195</sup>

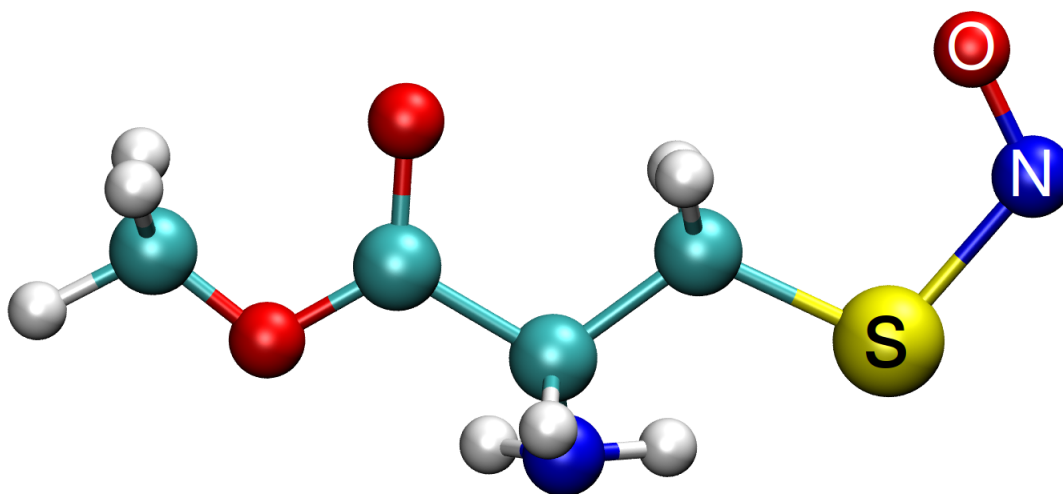
*Cys-NO*: For simulating nitrosylated cysteine (Cys-NO), the molecule was placed in the center of a cubic box (dimensions  $25 \times 25 \times 25$  Å<sup>3</sup>) of TIP3P<sup>196</sup> water and maintained there with a weak center-of-mass constraint with a force constant of 1 kcal/mol. Two sets of simulations were carried out: one using PC electrostatics and the other one with MTP interactions, see below. First, the systems were heated to 300 K and equilibrated at this temperature in the *NVT* ensemble for 500 ps. Production simulations of 10 ns were then performed in the *NVE* ensemble.

*WT and Nitrosylated Mb*: For the simulations involving wild type and S-nitrosylated Mb, eight different simulations were set up: wild-type myoglobin (PDB: 2NRL)<sup>197</sup> at 50 K and 300 K, cis- and trans-S-Nitrosylated myoglobin (S-Nitrosylation at Cys10) with a PC model for the -SNO moiety at 50 K and 300 K, and cis- and trans-S-Nitrosylated myoglobin with a multipolar charge model at 50 K and 300K. Because density for the terminal residue Gly147 was missing in the 2NRL structure only residues Ala2 to Ser146 were used to set up the WT and S-nitrosylated protein at position Cys10. With this setup, simulations for cis-

MbSNO and trans-MbSNO starting from the same initial structure except the dihedral angle  $\phi(C_\beta SNO)$  ( $0^\circ$  for cis-MbSNO and  $180^\circ$  for trans-MbSNO) were started.

All systems were solvated in a  $80 \times 80 \times 80 \text{ \AA}^3$  cubic box of TIP3P<sup>196</sup> water molecules with buffer regions of  $15 \text{ \AA}$  to the edges of the box. The protein was weakly constrained to the middle of the simulation box, minimized, heated to the desired temperature and equilibrated for 500 ps in the *NVT* ensemble. Production runs of 10 ns were then performed in the *NVE* ensemble.

### 4.3.2 Parametrisation of the Force Field



**Figure 4.1:** The structure of methyl-capped CysNO used for the parametrization of the force field. The CysNO ligand is shown as CPK. The color code for the atoms is H (white), C (cyan), O (red), N (blue), S (yellow).

The CHARMM36 force field was employed to describe the methyl-terminated Cys residue. The additional parameters required for Cys-NO were determined from electronic structure calculations. For this, the structure of nitrosylated,  $\text{CH}_3$ -terminated Cysteine (see Figure 4.1) was optimized at the MP2/aug-cc-pVDZ<sup>198,199</sup> level of theory. All electronic structure calculations were carried out

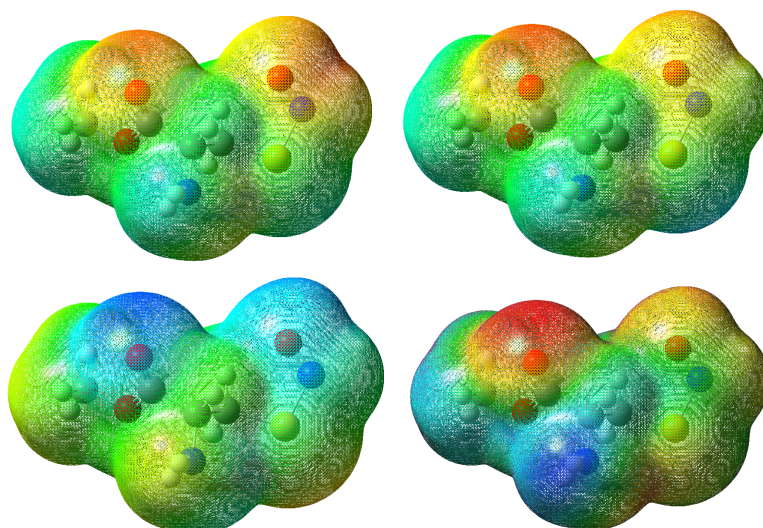
with Gaussian09.<sup>128</sup> The nature of the stationary point was verified by calculating vibrational frequencies. Then the SN bond was scanned along the bond separation  $r$  on a grid between 1.3 and 6 Å whereas the NO bond was scanned between 0.9 and 5.1 Å, both in increments of 0.1 Å. The coordinates of all atoms except for the NO moiety along the SN bond scan, and O for the NO bond scan were frozen.

Next, the reference MP2 energies were fit to Morse oscillator functions  $V(r) = D_e[1 - \exp(-\beta(r - r_e))]^2$  where  $D_e$  is the well-depth,  $r_e$  is the equilibrium bond distance, and  $\beta$  is the parameter that controls the steepness of the potential. Subsequently, the steepness parameters  $\beta$  are adjusted such as to reproduce experimental frequencies of the SN and NO stretch vibrations at 520 and 1526  $\text{cm}^{-1}$ , respectively.<sup>171</sup> The  $\beta$  values were 2.171 Å<sup>-1</sup> and 1.437 Å<sup>-1</sup> for the SN and NO stretch before fitted to the experimental vibration. For these parametrizations the energetically more stable cis-conformer was considered. The final Morse parameters after fitting  $\beta$  to the experimentally observed stretch frequencies are  $D_e = 202.7$  kcal/mol,  $\beta = 1.887$  Å<sup>-1</sup>,  $r_e = 1.207$  Å for the NO bond and  $D_e = 71.1$  kcal/mol,  $\beta = 1.987$  Å<sup>-1</sup>,  $r_e = 1.843$  Å for the SN bond.

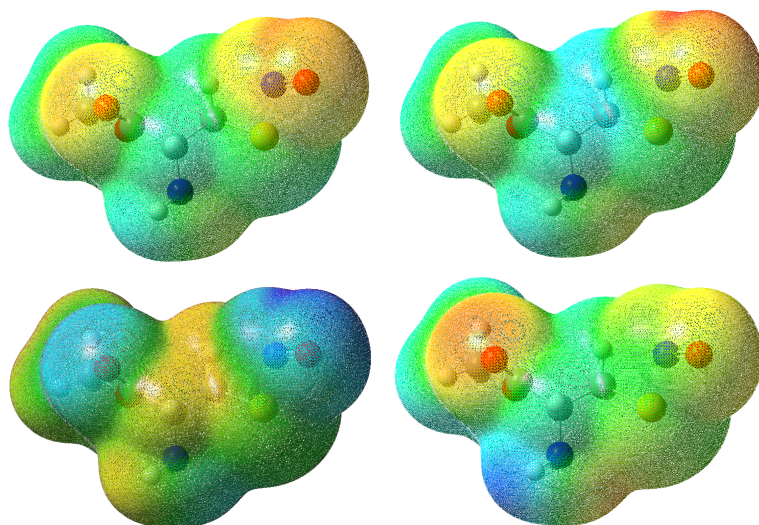
For the SNO bending, the angle was scanned between 0° and 180°. For the bending potential, the reference energies were fitted to  $V(\theta) = k_\theta(\theta - \theta_e)^2$  where  $k_\theta$  is the force-constant and  $\theta_e = 0$  is the equilibrium angle. Fitting the force constant to reproduce the experimentally determined frequency<sup>171</sup> of the SNO bending motion in GSNO at 886  $\text{cm}^{-1}$  yields a value of  $k_\theta = 132.5$  kcal/mol/rad<sup>2</sup> and  $\theta_e = 115.6^\circ$ .

For the electrostatic interactions two different models were developed. One of them is a standard point charge (PC) model and the second one uses multipoles

(MTPs) on the S, N, and O atoms.<sup>200,201</sup> For this, the charges of all atoms except for the S, N, and O atom (i.e. the -SNO label) were those of the CHARMM36 force field<sup>191</sup> which were kept unchanged. This was done in order to maintain the parametrization of the cysteine residue consistent with the remaining protein force field. On the other hand, the quality of the ESP fit will be affected by this strategy, see below. Two models were considered. The first is a PC model, fitted to the electrostatic potential (ESP) of the optimized structure of Cys-NO at the MP2/aug-cc-pVDZ level of theory, see Table 4.1. Next, these fitted PCs were frozen at the optimized values and the atomic dipole and quadrupole moments were fitted to improve the model.<sup>51</sup> All parameters are reported in Table 4.1.



**Figure 4.2:** The density cube generated from the cis-CysNO residue with PC model (Up-Left) and MTP model (Up-Right). The cube generated from the difference between PC and MTP model (Down-Left) and reference ESP cube at MP2/aug-cc-pVDZ level (Down-Right).



**Figure 4.3:** The density cube generated from the trans-CysNO residue with PC model (Up-Left) and MTP model (Up-Right). The cube generated from the difference between PC and MTP model (Down-Left) and reference ESP cube at MP2/aug-cc-pVDZ level (Down-Right).

The ESP maps for cis- and trans-CysNO with PCs and MTPs and the differences between the reference ESP (at the MP2 level) and those from the PC and MTP models are reported in Figures 4.2 and 4.3. The RMSE between reference ESP and the PC model is 8.36 kcal/mol and 7.44 kcal/mol for cis-CysNO and trans-CysNO, respectively. This rather larger difference originates from the fact that the CHARMM36 charges were retained and not allowed to adjust to the reference ESP in the fit. When including the MTPs on the -SNO label the RMSE decreased to 4.21 kcal/mol for cis-CysNO and 2.32 kcal/mol for trans-CysNO which is a significant improvement over a conventional PC model.

Atom	PC [ $e$ ]	Dipole [ $ea_0$ ]			Quadrupole [ $ea_0^2$ ]				
	$Q_{00}$	$Q_{10}$	$Q_{1C}$	$Q_{1S}$	$Q_{20}$	$Q_{21C}$	$Q_{21S}$	$Q_{22C}$	$Q_{22S}$
S	0.330	-0.373	0.079	0.000	0.032	0.006	0.000	-0.019	0.000
N	-0.203	0.000	0.359	0.128	0.370	0.000	0.000	-0.265	0.006
O	-0.127	0.104	-0.135	0.000	-0.040	0.040	0.000	-0.034	0.000

**Table 4.1:** The PC, dipole and quadrupole parameters generated at MP2/aug-cc-pVDZ level which been used in the MTP simulations

### 4.3.3 Infrared Spectrum

The infrared spectrum was obtained from the Fourier transform of the dipole moment autocorrelation function. For this, the molecular dipole moment was calculated from the MD trajectories and the partial charges. The autocorrelation function where  $i$  is the index number of a snapshot

$$C(t) = \langle \vec{\mu}(0) \cdot \vec{\mu}(t) \rangle \quad (4.1)$$

was accumulated over  $2^{15}$  time origins to cover 1/3 to 1/2 of the trajectory. From this, the absorption spectrum is determined according to

$$A(\omega) = \omega(1 - e^{-h\omega/(k_B T)}) \int C(t)e^{-i\omega t} dt \quad (4.2)$$

where  $T$  is the temperature in Kelvin,  $k_B$  is the Boltzmann constant and the integral is determined using a fast Fourier transform (FFT). IR spectra of WT, cis-MbSNO and trans-MbSNO have been generated for blocks of 100 ps simulation by correlating over  $2^{19}$  time origins. A total of 100 spectra were generated for each system (total simulation time of 10 ns) and averaged.

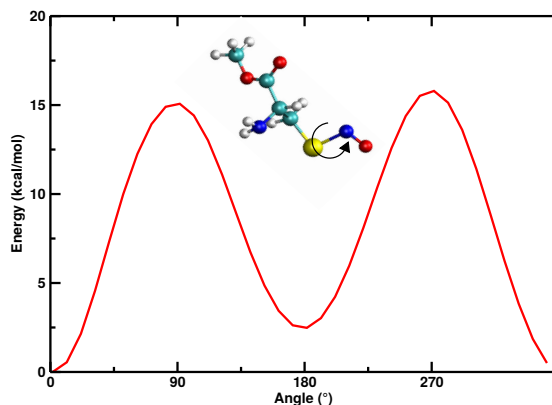
Additionally power spectra of the NO and SN bonds and the SNO angle have been calculated from the FFT of the bond length and angle time series autocorrelation functions. This provides assignments of the vibrational spectra and allows to detect couplings between modes.<sup>202</sup> These power spectra were not averaged and correlated over  $2^{15}$  time origins for the entire simulation time of 10 ns.

## 4.4 Results

First, the structural dynamics and IR spectroscopy of nitrosylated cysteine in solution was investigated. This provides a basis for characterization and interpretation the dynamics and spectroscopy of cysteine incorporated in a protein, such as myoglobin. Next, the spectroscopy and conformational dynamics of cis- and trans-MbSNO are discussed.

### 4.4.1 Dynamics and Spectroscopy of CysNO

Figure 4.1 shows the structure of cis-CysNO in aqueous solvent. Quantum chemical calculations at the MP2/aug-cc-pVDZ level show two minima with  $\phi_{\text{CSNO}}$  dihedral angles of 0 and 180°. The energy profile  $V(\phi)$  is reported in Figure 4.4 and shows a stability difference between the global minimum (cis-Cys-NO) and the trans-conformer ( $\phi = 180^\circ$ ) of 2.5 kcal/mol with the two states separated by a barrier of  $\sim 15$  kcal/mol at the MP2 level of theory. Hence, interconversion between the two states is expected to be slow and will be on the millisecond time scale according to transition state theory.

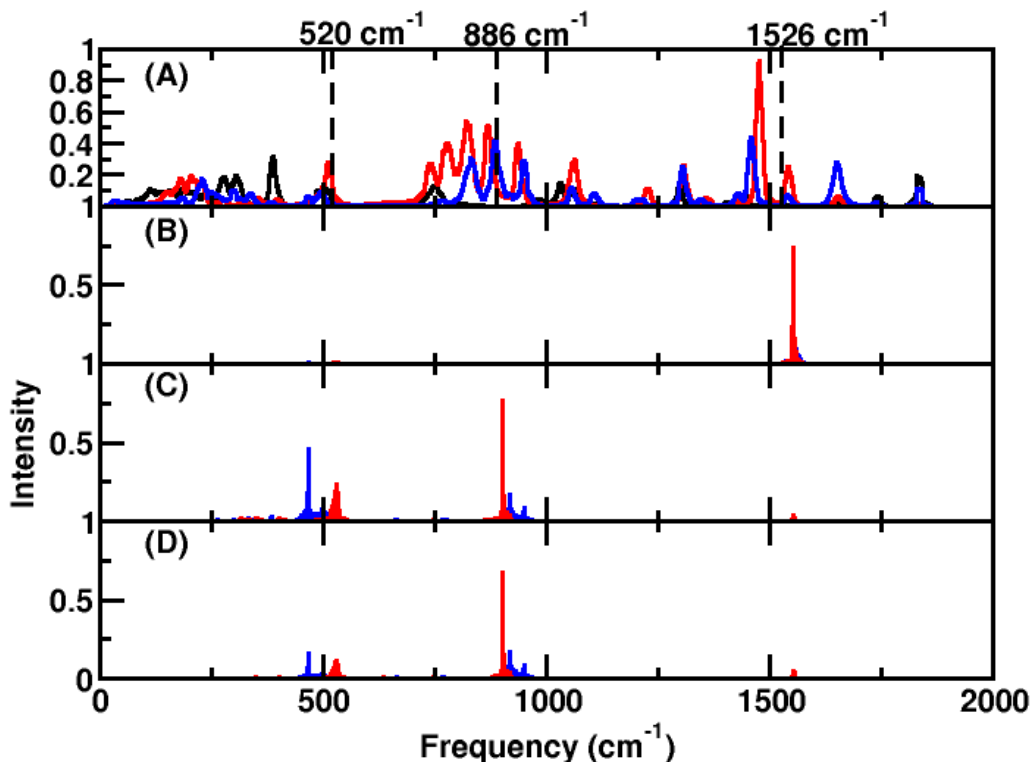


**Figure 4.4:** The energy profile of the CSNO torsion angle at the MP2/aug-cc-pVDZ level of theory.

The IR and power spectra for Cys and CysNO are shown in Figure 4.5 and compared with the experimental line positions of the vibrations from Raman spectroscopy of S-nitrosoglutathione (GSNO) in the solid state<sup>171</sup> which was the reference for the parametrization. The NO stretch peak is clearly visible in Figure 4.5A at  $1540\text{ cm}^{-1}$  for cis-CysNO and  $1538\text{ cm}^{-1}$  for trans-CysNO (red and blue traces, respectively) as the power spectrum in Figure 4.5B confirms. Contrary to that, the SN and SNO peaks are more difficult to locate in the power spectrum. Considering the power spectra in Figures 4.5C and D, the SN stretch and SNO bending modes are at  $514\text{ cm}^{-1}$  and  $871\text{ cm}^{-1}$  for cis-CysNO and  $495\text{ cm}^{-1}$  and  $881\text{ cm}^{-1}$  for trans-CysNO, respectively, which are close to the experimental values. The power spectra (Figures 4.5B to D) show that SN and SNO modes are strongly coupled whereas the coupling to the NO stretch is less pronounced.

Although the same force field was used for cis- and trans-CysNO, their spectroscopy for the SN and SNO modes differs. In the power spectra, the bands of trans-CysNO are shifted to lower frequency (red shift) by  $29\text{ cm}^{-1}$  for SN and blue shifted by  $14\text{ cm}^{-1}$  for SNO with respect to cis-CysNO, which should be suf-



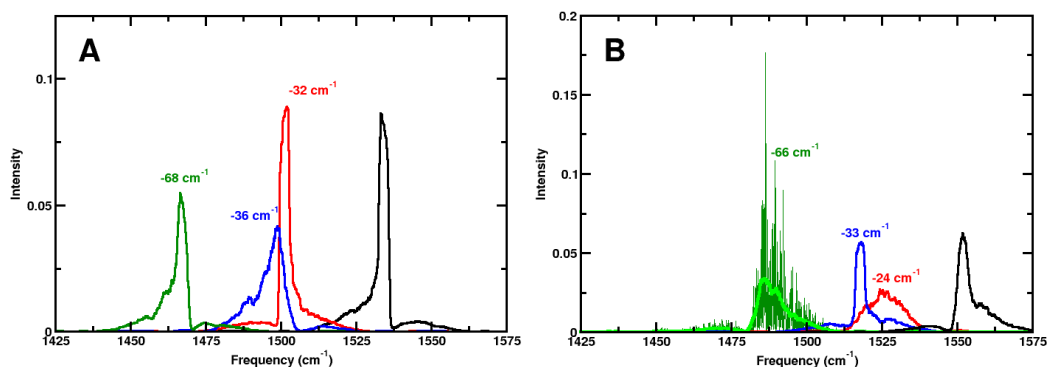


**Figure 4.5:** Infrared and power spectra for Cys (black) and Cys-NO (blue and red) in water. Panel A: calculated IR spectrum of Cys (Black), cis-Cys-NO (red) and trans-Cys-NO (blue) from the MD simulations at 50 K with the MTP model. Panels B to D: the NO, SNO and SN power spectra, respectively. The dashed lines at  $520\text{ cm}^{-1}$ ,  $886\text{ cm}^{-1}$  and  $1526\text{ cm}^{-1}$  are the experimental values for the SN, SNO and NO modes in GSNO in the solid state, respectively.<sup>171</sup> The dashed lines indicate the positions of the experimental values.

ficient to detect both isomers if they are present in solution. For the NO-stretch the splitting between cis- and trans-CysNO is  $3\text{ cm}^{-1}$  in the gas phase which compares with  $21\text{ cm}^{-1}$  for methyl thionitrite ( $\text{CH}_3\text{SNO}$ ) in an Argon matrix<sup>175</sup> whereas in solution, the present simulations only find an insignificant splitting. Consistent with experiment, the simulations also find that the trans-conformer absorbs at higher frequency. The different magnitude of the shift in the gas phase may be due to both, the different environment (argon in the experiments) and the somewhat different chemical environment of  $\text{CH}_3$  versus Cysteine. Further, the SN stretch is at  $376\text{ cm}^{-1}$  for cis- $\text{CH}_3\text{SNO}$  and  $371\text{ cm}^{-1}$  for trans- $\text{CH}_3\text{SNO}$  which amounts to a red shift by  $5\text{ cm}^{-1}$ .<sup>175</sup> This shift was more pronounced in

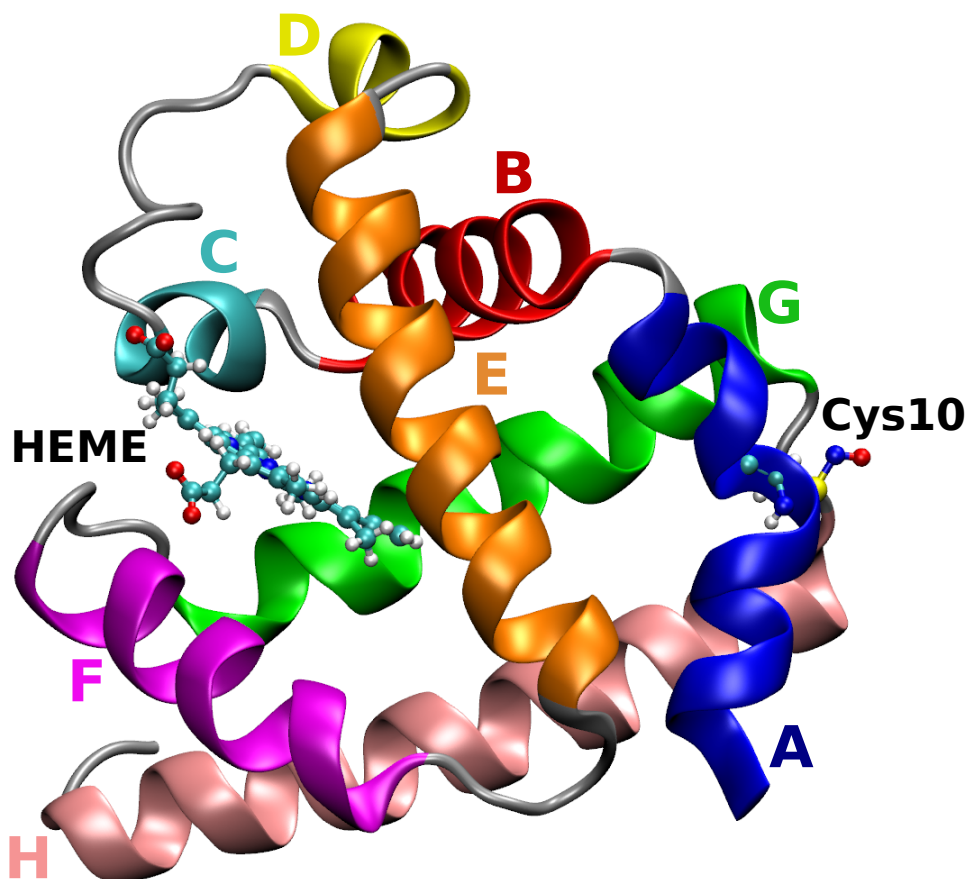
the present simulations with  $29\text{ cm}^{-1}$ .

Next, the frequency shifts following isotopic substitution ( $^{14}\text{N}\rightarrow^{15}\text{N}$  and  $^{16}\text{O}\rightarrow^{18}\text{O}$ ) are determined for cis- and trans-CysNO. The power spectra for the  $^{14}\text{N}^{16}\text{O}$ ,  $^{15}\text{N}^{16}\text{O}$ ,  $^{14}\text{N}^{18}\text{O}$  and  $^{15}\text{N}^{18}\text{O}$  for cis- and trans-CysNO in the gas phase are presented in Figures 4.6A and B. Identical simulations as those above for all three isotopic variants lead to red shifts of the  $^{15}\text{N}^{16}\text{O}$ -stretch by  $[-32, -24]\text{ cm}^{-1}$  for [cis,trans]-CysNO,  $[-36, -33]\text{ cm}^{-1}$  for the  $^{14}\text{N}^{18}\text{O}$ -stretch, and  $[-68, -66]\text{ cm}^{-1}$  for the  $^{15}\text{N}^{18}\text{O}$ -stretch relative to those of  $^{14}\text{N}^{16}\text{O}$ , respectively. These findings are consistent with expectations from experiments<sup>186</sup> on the ionic species which are estimated at  $\sim -20\text{ cm}^{-1}$  for  $^{15}\text{N}^{16}\text{O}$  compared with  $[-32, -24]\text{ cm}^{-1}$  from the present simulations.



**Figure 4.6:** Power spectra for all four isotopic variants of (panel A) cis- and (panel B) trans-CysNO from 1 ns simulations at 50 K in the gas phase. The isotope-induced red shifts for  $^{15}\text{N}^{16}\text{O}$  (red),  $^{14}\text{N}^{18}\text{O}$  (blue) and  $^{15}\text{N}^{18}\text{O}$  (green) compared to  $^{14}\text{N}^{16}\text{O}$  (black) are provided in respective colors. For all power spectra a running average is reported and panel B shows a comparison between the raw data and the running average for  $^{15}\text{N}^{18}\text{O}$ . The isotopic shifts are reported as the frequency differences between the maxima of the most pronounced peaks.

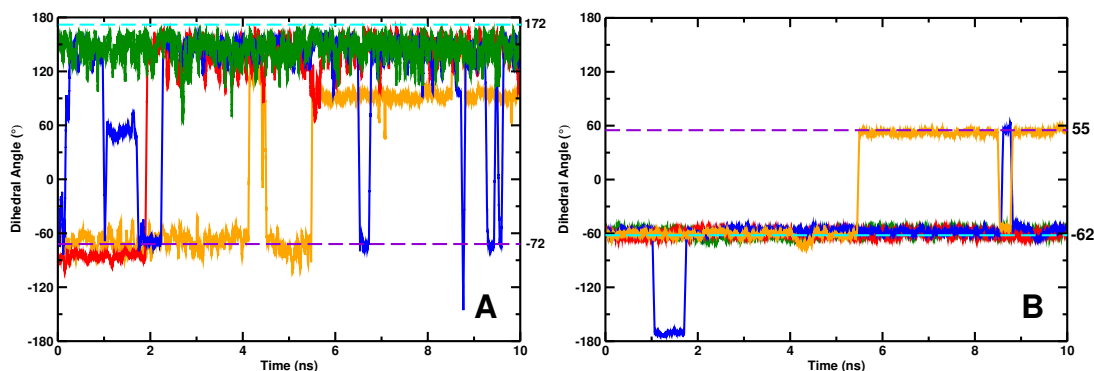
#### 4.4.2 The Structural Dynamics and Spectroscopy of Mb-SNO



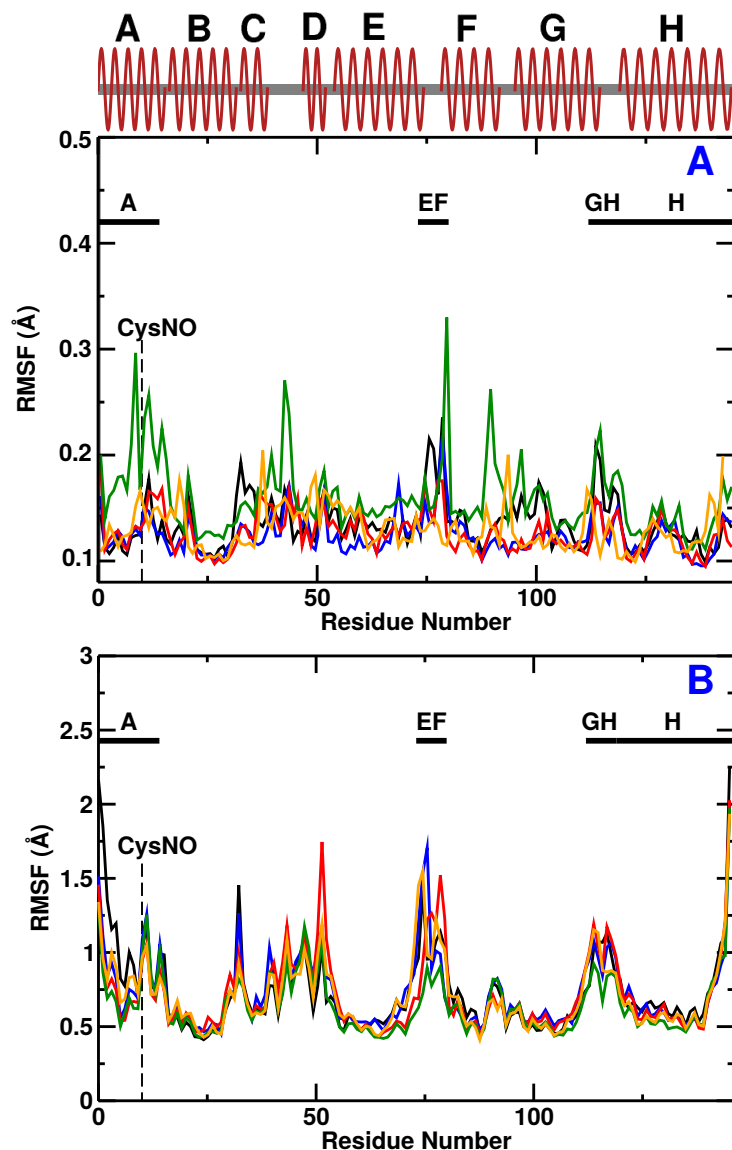
**Figure 4.7:** Structure of trans-MbSNO (Cartoon representation) with helices A to H in different color together with corresponding labels. The heme group and the S-Nitrosylated Cys10 residue are represented by CPK. Cys10 is part of the A-helix of Mb.

Next, the dynamics and spectroscopy of wild-type (WT) blackfin tuna myoglobin (PDB: 2NRL)<sup>197</sup> and its S-nitrosylated variant<sup>197</sup> at Cys10 (PDB: 2NRM) was considered, see Figure 4.7. For the S-nitrosylated cysteine in blackfin tuna myoglobin two different cis-conformers were reported.<sup>197</sup> The major and minor cis-conformers had  $\phi(\text{NC}_\alpha\text{C}_\beta\text{S})$  dihedral angles of  $-62^\circ$ ,  $55^\circ$  and  $\phi(\text{C}_\alpha\text{C}_\beta\text{SN})$  of  $172^\circ$ ,  $-72^\circ$  in the X-Ray structures, respectively. In the simulations at 300 K with both, PC and MTP models, the structure with  $\phi(\text{NC}_\alpha\text{C}_\beta\text{S}) \sim -62^\circ$  was present throughout the 10 ns simulation (see Figure 4.8B). With respect to the

dihedral  $\phi(C_\alpha C_\beta SN)$ , the simulation using MTPs sampled only the state with  $\sim 160^\circ$  which is close to the major cis- conformer value whereas the simulation with PCs sampled both states with  $\sim -85^\circ$  which is close to the minor cis- conformer value and  $\sim 160^\circ$ , see Figure 4.8A. The first, shorter lived state with  $\phi(NC_\alpha C_\beta S) \sim -85^\circ$  using PCs was sampled for the first 2 ns after which  $\phi(NC_\alpha C_\beta S)$  switched to  $\sim 160^\circ$  and remained there for the rest of the simulation (see Figure 4.8B). These results indicate that the simulations are able to account for the major cis- conformer with both charge models at 300 K. Given the different spectroscopic features for cis- and trans-CysNO, for MbSNO also both conformers were considered in the MD simulations. The dynamics and spectroscopy of the WT and S-nitrosylated variants were studied at 50 K and 300 K using PC and MTP charge models.



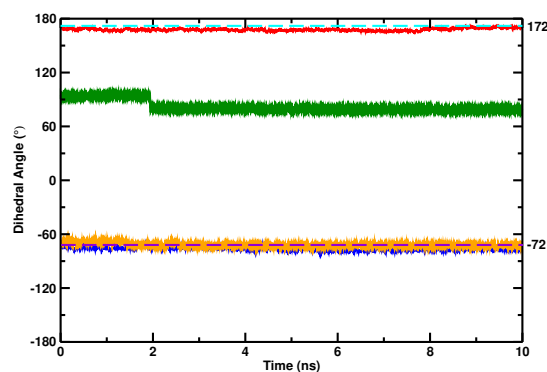
**Figure 4.8:** Panel A) Dihedral angle of  $C_\alpha C_\beta SN$  for Cys10 as a function of time in cis-MbSNO with PC (red), trans-MbSNO with PC (blue), cis-MbSNO with MTP (green) and trans-MbSNO with MTP (orange) at 300 K. Panel B) Dihedral angle  $NC_\alpha C_\beta S$  of Cys10 as a function of time in cis-MbSNO with PC (red), trans-MbSNO with PC (blue), cis-MbSNO with MTP (green) and trans-MbSNO with MTP (orange) at 300 K. Dashed lines represent the values from the major (cyan) and minor (violet) cis-MbSNO conformer in 2NRM.



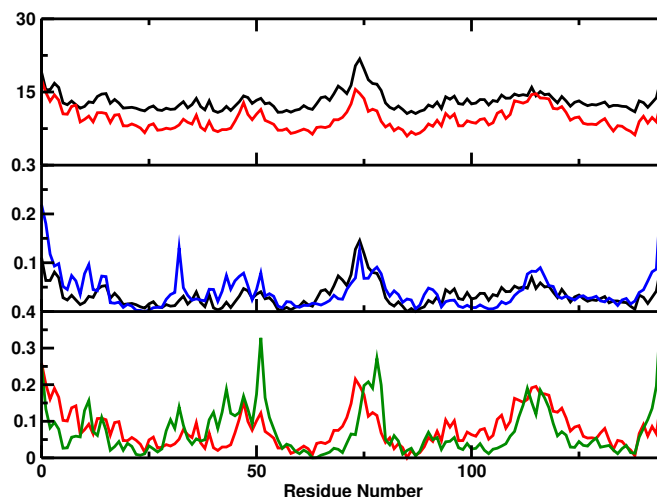
**Figure 4.9:** Root Mean Squared Fluctuation of each residue at 50 K (panel A) and 300 K (panel B) from simulations using the PC and MTP model. (Black line: WT, Red line: cis-MbSNO with PC, Blue line: trans-MbSNO with PC, Green line: cis-MbSNO with MTP and Orange line: trans-MbSNO with MTP). The location of CysNO (at position 10), and helix A (Ala2 to Glu15), helix H (Gly121 to Ser146), loop EF (Ala74 to Ile81) and loop GH (Glu113 to Gly120) are explicitly indicated. Note the different scales along the  $y$ -axis in panels A and B.

The root mean squared fluctuations of the  $C_{\alpha}$  atoms of every residue at 50 K and 300 K from 10 ns simulations with the PC and MTP models are reported in Figure 4.9 (top and bottom) for the WT (black), cis-MbSNO (PC, red), trans-MbSNO (PC, blue), cis-MbSNO (MTP, green), and trans-MbSNO (MTP, orange). At 50 K all RMSFs are small and S-nitrosylation of Cys10 decreases the

flexibility of cis- and trans-MbSNO with PC in the C-, E-, F-, and G-helix regions compared with WT. The RMSF for residues Lys31 to Glu35 (end of helix B and beginning of helix C) display reduced flexibility as a consequence of the chemical modification. The largest differences are found for the residues between Lys90 and Leu102 (end of helix F, FG loop, beginning of helix G). With MTPs on the -SNO label the RMSFs for cis-MbSNO are typically larger or equal compared with WT (PC) due to unfavorable conformation around the -SNO group whereas for trans-MbSNO they are comparable. Contrary to simulations at 50 K that sample the  $C_{\alpha}C_{\beta}SN$  dihedral only at one particular angle, the dynamics of cis-MbSNO samples states characterized by angles of  $94^{\circ}$  and  $80^{\circ}$  (see blue line in Figure 4.10). The conformational change of S-nitrosylated Cys10 from  $94^{\circ}$  to  $80^{\circ}$  at such low temperature indicates an unfavorable conformation that could lead to a higher flexibility of the protein.



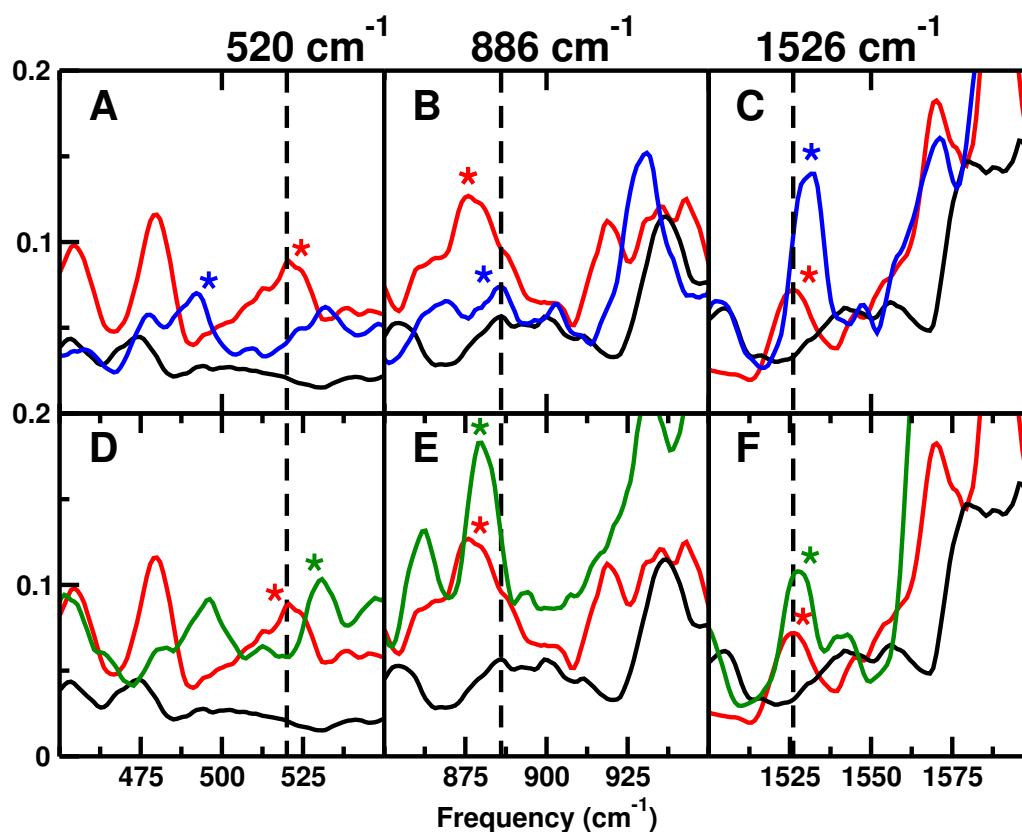
**Figure 4.10:** Dihedral angle of  $C_{\alpha}C_{\beta}SN$  for Cys10 as a function of time in cis-MbSNO with PC (red), trans-MbSNO with PC (blue), cis-MbSNO with MTP (green) and trans-MbSNO with MTP (orange) at 50 K. Dashed lines represent the values from the major (cyan) and minor (violet) cis-MbSNO conformer in 2NRM.



**Figure 4.11:** Superposition of the experimentally measured  $^{197}\text{C}_\alpha$  B-factors and the computed  $\text{C}_\alpha$  RMSFs from the present simulations. A) Experimental WT B-Factors (black) vs. Experimental MbSNO B-Factors (red) B) Scaled Experimental WT B-Factors (black) vs WT RMSF at 300 K (blue) C) MbSNO: Scaled experimental B-Factors (red) vs. RMSF from simulations for cis-MbSNO with PC at 300 K (green)

At 300 K, the differences in flexibility of each residue between WT, cis- and trans-MbSNO with both PC and MTP are much smaller. While the magnitude of the RMSF between 50 K and 300 K increases considerably, as expected, the differences between the cis-, trans-MbSNO and WT become more specific. NO attachment decreases the flexibility of residues Ala2 to Cys10 (helix A) for both, cis- and trans-MbSNO, compared with WT. This indicates a rigidification of helix A and suggests that nitrosylation at Cys10 impacts on the protein dynamics around the modification site. Moreover, increased flexibilities up to  $1.7 \text{ \AA}$  are found for residues Glu70 to Ile81 (EF loop). This is the region with the largest displacements in both, cis- and trans-MbSNO compared to WT. It is interesting to note that although attachment of NO to Cys10 may appear as a small perturbation, the  $\text{C}_\alpha$  B-factors from the experimental X-ray structures can differ throughout the protein, see Figure 4.11A. Comparing (scaled) experimental B-factors with computed RMSF for WT Mb shows that the two agree qualitatively, except for the region around residue 30 (Figure 4.11B). Similarly, there is qualitative agreement between experiment and simulations for MbSNO except

for a few residues around position 60. The results show the intrinsic flexibility of loops EF and GH which is correctly captured in the simulations. The lack of more quantitative agreement between experiment and computations can be due to the fact that two conformers exist for MbSNO. In such cases it was found that B-factors can decrease due to partial occupancies of the different substates.<sup>203</sup> Also, lattice disorder and crystal packing effects that occur in experiments are not included in the simulations<sup>204</sup> and may lead to differences between results from experiments and simulations.



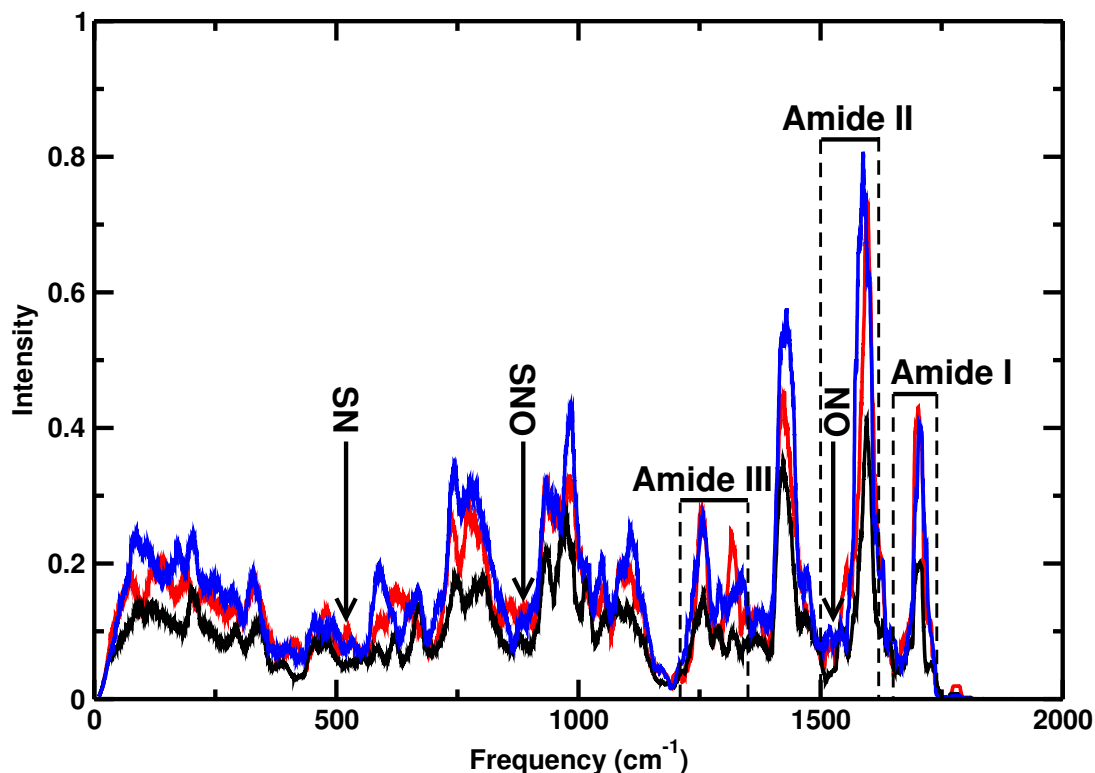
**Figure 4.12:** Comparison of the calculated IR spectra for WT, cis- and trans-MbSNO from 10 ns simulations at 50 K in the region of the SN stretch (panels A and D), SNO bend (panels B and E), and NO stretch (panels C and F). Top row from simulations with PC (black: WT, red: cis-MbSNO, blue: trans-MbSNO). Bottom row from simulations with PC and MTP (black: WT, red: cis-MbSNO with PC, green: cis-MbSNO with MTP). The dashed vertical lines indicate the experimental values of the corresponding vibrations in GSNO<sup>171</sup> and the colored stars label the spectral signatures to which the mode was assigned based on analysis of the power spectra. In general, the WT spectra (black) are featureless in the region where the -SNO label shows spectroscopic signatures.



The spectra related to the NO and SN stretches and the SNO bend are reported in Figures 4.12A to C from simulations for cis- and trans-SNO with PCs. A comparison for PC and MTP simulations is given in Figures 4.12D to F. For the NO stretch (Figure 4.12C) the peak for cis-MbSNO from simulations with PCs appears at  $1526\text{ cm}^{-1}$  compared with  $1532\text{ cm}^{-1}$  for trans-MbSNO. These absorption frequencies for the NO stretch compare with  $1526\text{ cm}^{-1}$  from experiments on GSNO in the solid state, i.e. identical to that for cis-MbSNO, and a blue shift of  $6\text{ cm}^{-1}$  for trans-MbSNO between simulations and experiment. It is also noted that the NO-stretch vibration in CysNO (at  $1538\text{ cm}^{-1}$  and  $1540\text{ cm}^{-1}$  for cis- and trans-CysNO with MTPs) differs from that for cis-MbSNO and trans-MbSNO by  $14$  and  $6\text{ cm}^{-1}$ , respectively, see Figure 4.5. Likewise, a blue-shift was observed for trans-CH<sub>3</sub>SNO relative to cis-CH<sub>3</sub>SNO for experiments in argon.<sup>175</sup> The intensity of the NO band for trans-MbSNO is higher compared with that of cis-MbSNO. Figure 4.12C also shows that the NO stretch is clearly set apart from the nearby the amide II band which is ranging from  $1500$  to  $1620\text{ cm}^{-1}$  in the simulations and from  $1500$  to  $1600$  from experiments.<sup>205</sup> This should make it possible to locate the NO stretch mode from experiments, in particular when reference spectra for the WT protein or for two different isotopes of the label (<sup>14</sup>NO and <sup>15</sup>NO) are subtracted.

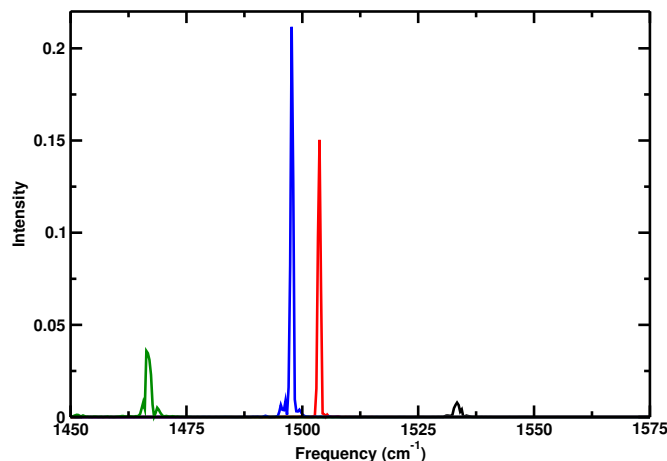
The SN stretch and SNO bend modes are more challenging to identify for both conformations. The spectrum in Figure 4.12B has the SNO bending mode at  $876\text{ cm}^{-1}$  for cis-MbSNO which shifts to  $886\text{ cm}^{-1}$  for the trans-conformer. Finally, the SN stretch appears at  $522\text{ cm}^{-1}$  for cis-MbSNO and at  $494\text{ cm}^{-1}$  for the trans-conformer, see Figure 4.12A. Hence, the orientation of the NO group (cis vs. trans) shifts these modes by  $6$  to  $28\text{ cm}^{-1}$ . Again, the magnitude of these shifts is consistent with the findings for CysNO in water.

In order to quantify the influence of the charge model used on the infrared signatures of the SN stretch, SNO bending, and the NO stretch, cis-MbSNO was considered. For this, a 10 ns MD simulation at 50 K was carried out with MTPs on the -SNO moiety. The corresponding spectra are shown in Figures 4.12D to F. The NO stretch from the simulations with MTP appears at  $1527\text{ cm}^{-1}$  which is shifted by  $1\text{ cm}^{-1}$  to the blue compared with the simulation with the PC model. The intensity of this peak increases considerably when using the more elaborate model for the electrostatics, see Figure 4.12F. For the SNO bending vibration the frequency maximum appears at  $880\text{ cm}^{-1}$  in the simulation with MTPs which is a shift of  $4\text{ cm}^{-1}$  to the blue compared with the PC simulation. Finally, the SN stretch is at  $531\text{ cm}^{-1}$ , shifted by  $9\text{ cm}^{-1}$  to the blue compared with the PC simulation. Hence, all three vibrations shift to the blue in simulations with the MTP model compared with PCs. Such shifts are typical for simulations with PCs and MTPs.<sup>206–210</sup> The increased intensity with MTP model made it easier to detect the vibrations on the total IR spectrum of the protein, especially for SNO and NO.

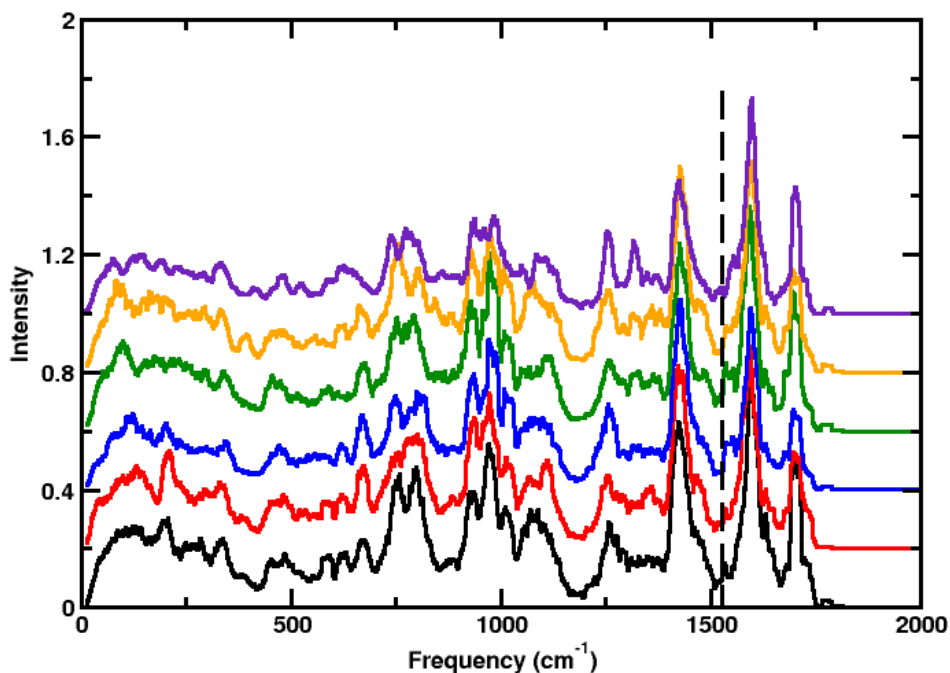


**Figure 4.13:** Infrared spectrum of WT (black), cis-MbSNO (red) and trans-MbSNO (blue) at 50 K from the total dipole moment of the protein.

In summary, it is noted that in the regions of the SN stretch, SNO bend and NO stretch modes the intensity of the spectrum for WT myoglobin (black trace in Figure 4.13) is characteristically low whereas those for the two nitrosylated variants show increased intensities (see arrows in Figure 4.13). The calculated IR spectra from simulations with PCs and the Fourier transform of the protein dipole autocorrelation function for WT, cis-MbSNO with PC and trans-MbSNO are reported in Figure 4.13. Spectral features of amide I, amide II and amide III absorption bands are present in all spectra with varying intensities. Also, the blue shifted frequencies in particular for the NO-stretch vibration for trans-versus cis-SNO is consistent with experiments on model compounds in the gas phase or in argon matrices.



**Figure 4.14:** Power spectra for all four isotopes of cis-MbSNO with PC from 1 ns simulations at 50 K in solution. The isotope-induced red shifts compared to  $^{14}\text{N}^{16}\text{O}$  (black) are  $-30\text{ cm}^{-1}$ ,  $-36\text{ cm}^{-1}$ , and  $-68\text{ cm}^{-1}$  for  $^{15}\text{N}^{16}\text{O}$  (red),  $^{14}\text{N}^{18}\text{O}$  (blue) and  $^{15}\text{N}^{18}\text{O}$  (green), respectively.



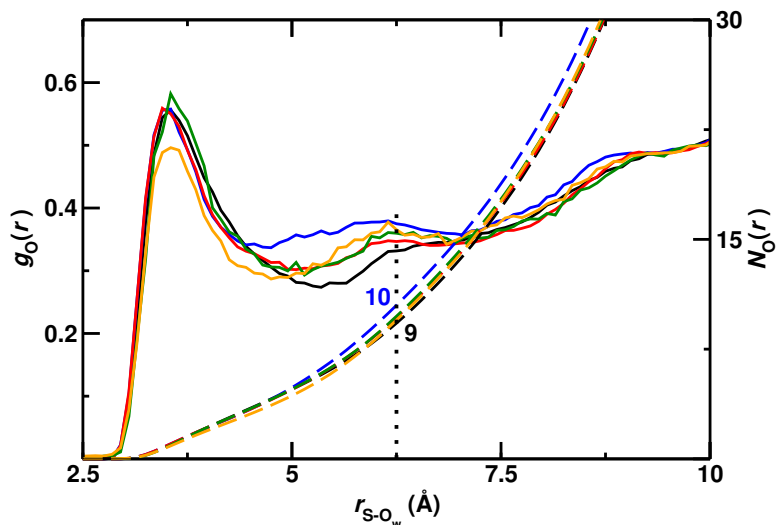
**Figure 4.15:** Infrared spectrum of cis-MbSNO with PC at 50 K computed from different 2 ns intervals, vertically displaced for clarity. Black line: 0 to 2 ns, red line: 2 to 4 ns, blue line: 4 to 6 ns, green line: 6 to 8 ns, orange line: 8 to 10 ns, indigo line: 0 to 10 ns. The vertical dashed line indicates the location of the NO-stretch vibration.

The frequency shifts for the NO-stretch following isotopic substitution ( $^{14}\text{N}\rightarrow^{15}\text{N}$  and  $^{16}\text{O}\rightarrow^{18}\text{O}$ ) were also determined for cis-MbSNO with PC, see Figure 4.14. The power spectra for the three isotopic substitutions yield red shifts of  $-30\text{ cm}^{-1}$  ( $^{15}\text{N}^{16}\text{O}$ ),  $-36\text{ cm}^{-1}$  ( $^{14}\text{N}^{18}\text{O}$ ), and  $-68\text{ cm}^{-1}$  ( $^{15}\text{N}^{18}\text{O}$ ), respectively. These shifts differ by  $-2\text{ cm}^{-1}$  at most from those found for CysNO. As a final point, the convergence of the spectra from the 10 ns simulations was assessed by considering cis-MbSNO with PC at 50 K. Figure 4.15 compares IR spectra from five successive 2 ns intervals and the IR spectrum from the entire 10 ns simulation. The important amide I, II and III bands (see Figure 4.13) as well as the NO stretch are present in all spectra but with varying intensities which provides a convergence measure of the IR spectra on the time scale of the present simulations.

### 4.4.3 Water Structure and Global Structural Changes

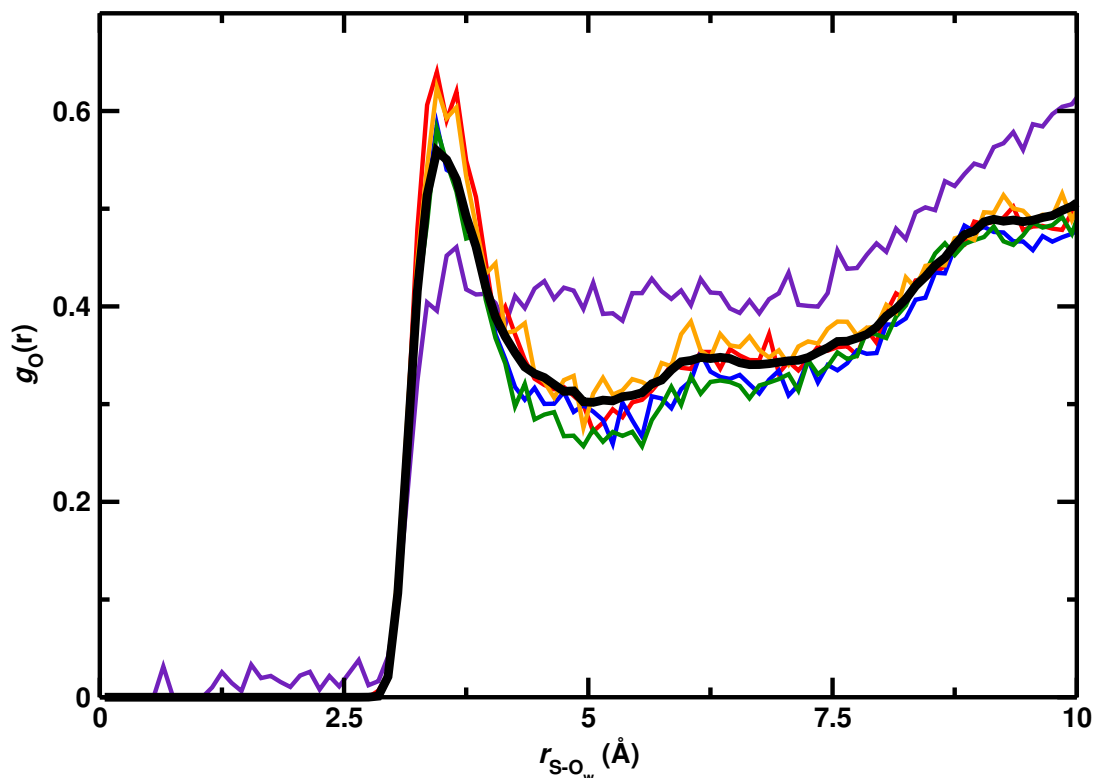
Next, it is of interest to consider the local water ordering around the modification site (Cys10) for WT and nitrosylated Mb. The radial distribution function  $g_{\text{S-OW}}(r)$  and corresponding coordination number  $N_{\text{S-OW}}(r)$  of water oxygen (OW) with respect to the sulfur atom of Cys10 in WT, cis-MbSNO, and trans-MbSNO are shown in Figure 4.16. At 300 K and with PC for the simulations the first solvation shell peak appears at  $3.5\text{ \AA}$  in all proteins. While the difference of the first solvation peak in terms of peak height and shape is marginal, the radial distribution functions for cis- and trans-MbSNO (red and blue) differ from that for WT (black) in the region between  $4\text{ \AA}$  and  $7.5\text{ \AA}$ , see Figure 4.16. Hence, with respect to the Cys10-sulfur atom both, cis-MbSNO and trans-MbSNO, are more highly hydrated in the range of  $r_{\text{S-OW}} \sim 5$  to  $7\text{ \AA}$  compared with WT. Since nitric oxide is more solvent exposed in trans-MbSNO (blue trace) compared to cis-MbSNO (red trace) the -SNO label is more highly hydrated in the trans-conformation. The occupation of the first solvation shell (up to  $r_{\text{S-OW}} \sim 6.25$

Å) for WT and nitrosylated Mb is found to differ by about 1 water molecule, which amounts to an increase of 10 %, see dashed line in Figure 4.16. Although the -SNO (modified Mb) group evidently occupies more space compared with -SH (WT Mb), attaching the -NO label recruits more water molecules. This is evidently an effect of the different electrostatics between water and the -SH and -SNO groups, respectively.



**Figure 4.16:** Radial distribution function of water oxygen and the corresponding coordination number of  $N_O(r)$  of water oxygen with respect to the sulfur atom obtained from *NVE* simulations at 300 K. Color code: WT (black), cis-MbSNO with PC (red), trans-MbSNO with PC (blue), cis-MbSNO with MTP (green), trans-MbSNO with MTP (orange). The vertical dotted line indicates the completion of the first solvent shell by which the number of hydration waters between WT and the modified proteins differs by  $\approx 1$  water molecule, i.e. a change of about 10 %, despite the larger volume the -SNO group occupies compared to -SH.

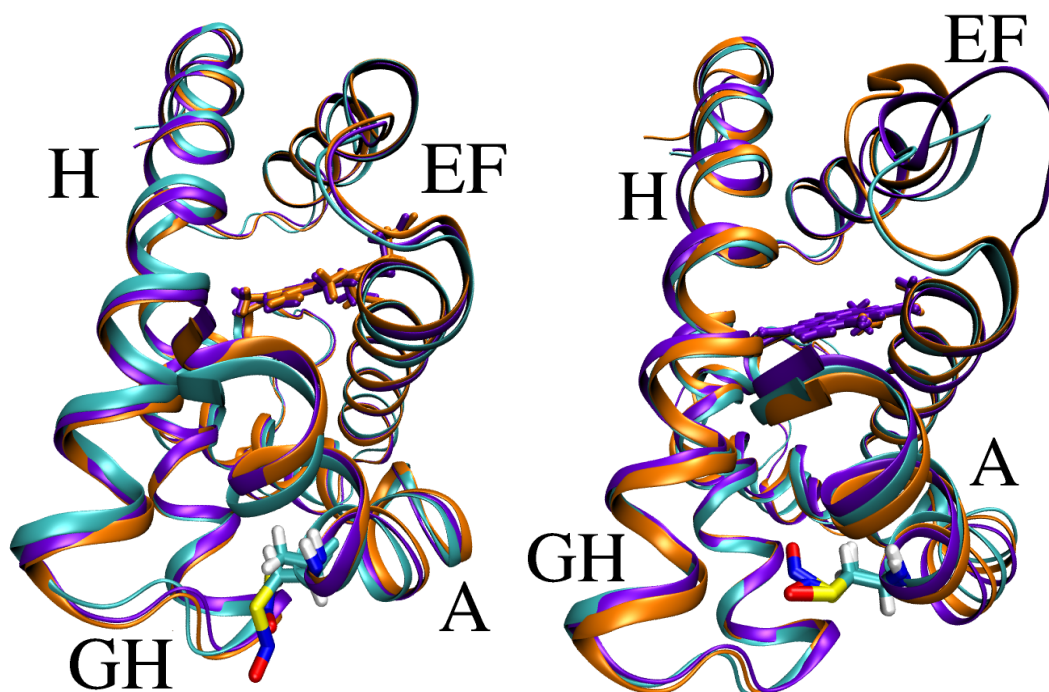
Similar to the IR spectra, the  $g_{S-O_w}(r)$  was computed for successive 2 ns intervals, see Figure 4.17. This shows that after 2 ns the  $g_{S-O_w}(r)$  fluctuates around a well defined average. Also, it is found that the first solvation shell around the -SNO label is established within the first 2 ns. This data also demonstrates that 10 ns simulations are sufficient for obtaining a meaningful radial distribution function for a particular conformational substate (here cis-MbSNO).



**Figure 4.17:** Radial distribution function of water oxygen with respect to the sulfur atom obtained from cis-MbSNO with PC in *NVE* ensemble at 300 K. The data collected from 2 ns time intervals. Indigo line: 0 to 2 ns, red line: 2 to 4 ns, blue line: 4 to 6 ns, green line: 6 to 8 ns, orange line: 8 to 10 ns, black line: 0 to 10 ns.

With the MTP model, trans-MbSNO displays a shallower first solvation shell compared to PC, while cis-MbSNO with MTP has comparable solvation to PC. The dissimilar impact of additional multipoles on the two conformers can be rationalized by the  $\phi(\text{NC}_\alpha\text{C}_\beta\text{S})$  angles sampled during the dynamics. For cis-MbSNO and both charge models this angle fluctuates around  $\sim -60^\circ$  throughout the 10 ns simulation (see red and green lines in Figure 4.8B). However, for trans-MbSNO (see Figure 4.8B) the  $\phi(\text{NC}_\alpha\text{C}_\beta\text{S})$  angle sampled differs between PC and MTP. With both models  $\phi(\text{NC}_\alpha\text{C}_\beta\text{S})$  samples structures with  $\sim -55^\circ$  - albeit only briefly for PCs - and  $\sim 52^\circ$ . For the simulation with PCs, in addition the orientation with  $\phi(\text{NC}_\alpha\text{C}_\beta\text{S}) = -170^\circ$  is sampled. Also, the residence time in the configuration with  $\phi(\text{NC}_\alpha\text{C}_\beta\text{S}) \sim -55^\circ$  differs between simulations with PC (9 ns) compared with MTP (5.75 ns) and the simulation with PCs spontaneously re-

turns to the conformation with  $\phi(\text{NC}_\alpha\text{C}_\beta\text{S}) = -62^\circ$ . Consequently, the different level of solvent exposure can be the underlying reason for the shallower solvation shells observed with MTP. However, longer simulations (and hence more transitions) are required for quantitative characterizations of such residence times.



**Figure 4.18:** The conformational changes in the protein structure induced by S-Nitrosylation, averaged over the last ns of free dynamic simulations at 50 (Left) and 300 K (Right) with PC model. WT (cyan), cis-MbSNO (orange) and trans-MbSNO (violet). S-Nitrosylated Cys10 and the heme unit are represented by licorice. A, H helices and EF, GH loops are labeled.

*Global Changes:* For characterizing more global structural changes, the last nanosecond of the 10 ns production run have been analyzed to investigate the structural changes in the protein at 50 K and 300 K, respectively. The cis-MbSNO and trans-MbSNO structures are superimposed onto the WT simulation and the structures averaged over the last ns of the free dynamic simulations at 50 K (left) and 300 K (right) in Figure 4.18 for WT (cyan), cis-MbSNO (orange), and trans-MbSNO (violet), respectively. The dihedral angles  $\phi(\text{C}_\beta\text{SNO})$  (for cis- versus trans-),  $\phi(\text{C}_\alpha\text{C}_\beta\text{SN})$  and  $\phi(\text{NC}_\alpha\text{C}_\beta\text{S})$  (for the position of the NO with respect



to nearby loops and  $\alpha$ -helices) determine the orientation of the -SNO label and reported in Figure 4.10 for 50 K and Figure 4.8 at 300 K. The time series show that all states sampled are also observed in the 2NRM X-ray structure.

	50K				300K				X-Ray
	cis-MbSNO		trans-MbSNO		cis-MbSNO		trans-MbSNO		2NRM
	PC	MTP	PC	MTP	PC	MTP	PC	MTP	
Helix A	0.57	0.60	0.51	0.63	1.45	1.48	1.37	1.33	0.681
Helix H	0.86	1.67	1.00	1.01	1.19	1.19	1.11	0.86	0.627
Loop GH	1.31	0.71	0.68	0.93	1.72	1.68	1.55	1.92	1.076
loop EF	1.17	1.07	1.16	1.35	2.63	2.02	3.23	2.54	3.843
Protein	0.83	0.78	0.81	0.80	1.27	1.14	1.28	1.17	0.989

**Table 4.2:** The average  $C_\alpha$  RMSD (in Å) for Helix A, Helix H, Loop GH, Loop EF and the entire protein for cis-MbSNO and trans-MbSNO with respect to the WT X-Ray structure for the last nanosecond of a 10 ns free dynamics simulation at 50 and 300 K with PC and MTP models. As a comparison, the differences between the 2NRL and 2NRM structure are also given.

Although the entire protein structure is affected by the modification to a certain degree, four regions revealed most prominent changes. They are helix A (residues Ala2 to Glu15), helix H (residues Gly121 to Ser146), loop GH (residues Glu113 to Gly120), and loop EF (residues Ala74 to Ile81). The  $C_\alpha$  RMSD of helix A, helix H, loop GH, loop EF and the entire protein for cis- and trans-MbSNO with respect to the WT X-ray structure during the last nanosecond of the 10 ns free dynamics simulations are summarized in Table 4.2.

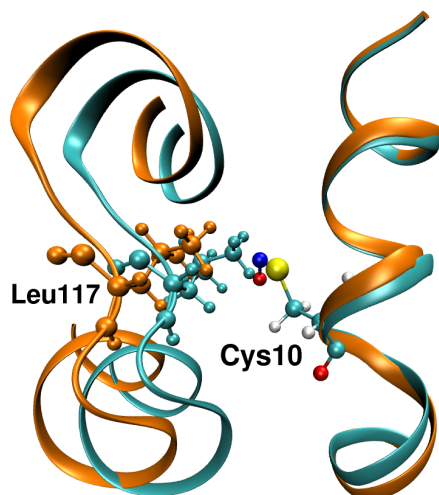
At 50 K, the  $C_\alpha$  RMSD for [cis-MbSNO, trans-MbSNO] compared to X-Ray WT are [0.83, 0.81] Å with PC and [0.78, 0.80] Å with MTP, respectively, which signals good preservation of the overall structure for the methods chosen. The  $C_\alpha$  RMSD calculated between the 2NRL and 2NRM structures is 0.99 Å which is compatible with the  $C_\alpha$  RMSD calculated from the simulations with both, PC and MTP electrostatic models for the NO-label. These results show that the

orientation of nitric oxide has a limited effect on the overall structure of the protein with only 0.02 Å difference between the two conformers with both charge models. Further, the same trend is observed in the local dynamics of the helices.  $C_\alpha$  RMSD values of the individual helices (A and H) deviate by 0.06 Å to 0.14 Å for cis-MbSNO compared to trans-MbSNO with PC, respectively. These small but consistent differences in the RMSD values between the two conformers can be rationalized by the orientation of nitric oxide with respect to  $\phi(C_\beta\text{SNO})$  which is  $0^\circ$  in cis-MbSNO and  $180^\circ$  in trans-MbSNO.

At 300 K, the  $C_\alpha$  RMSD of the cis-MbSNO and trans-MbSNO with respect to the WT X-Ray structure are 1.27 Å, 1.28 Å with PC and 1.14 Å, 1.17 Å with MTP, respectively. The local motion of the protein causes larger  $C_\alpha$  RMSD values of helix A in both, trans-MbSNO with PC and cis-MbSNO with PC. The values, with respect to X-Ray WT structure, were 1.45 Å and 1.37 Å respectively. The  $C_\alpha$  RMSD of helix H and loop GH were 1.19 and 1.72 Å in cis-MbSNO with PC, compared with 1.11 and 1.55 Å in trans-MbSNO with PC. Furthermore, the largest deviation from the X-Ray WT structure is observed for the EF loop, see Table 4.2. The relatively flexible loop moved closer to helix H in cis-MbSNO. Contrary to that, the movement of the loop was in the opposite direction for trans-MbSNO with PC. The  $C_\alpha$  RMSD of loop EF is 2.63 Å in cis-MbSNO with PC and 3.23 Å in trans-MbSNO with PC.

Depending on the orientation of the -SNO group (cis or trans), additional contacts with the protein can emerge. At 50 K, the  $\phi(C_\alpha C_\beta\text{SN})$  angles sampled for cis- and trans-MbSNO are shown in Figure 4.10. For cis-MbSNO the -SNO group is closer to the GH loop than for trans-MbSNO. Consequently, nitrosylation of Cys10 in its cis-conformer will lead to steric hindrance with the GH loop and

pushes it away from its position in WT Mb to avoid overlap between Leu117 and Cys10. Concomitantly, helix A (containing residue Cys10) is forced in the opposite direction of loop GH. Finally, helix H can push into the void created by the movement of helix A. This same steric hindrance was present to a lesser extent for trans-MbSNO and WT. The movement of the EF loop is also visible in Figure 4.18. The loop had a  $C_\alpha$  RMSD of 1.17 Å in cis-MbSNO and 1.16 Å in trans-MbSNO. The resulting crowding involving residues Cys10 and Leu117 without concomitant motion of loop GH and helix A upon nitrosylation is shown in Figure 4.19.



**Figure 4.19:** CPK representation of S-nitrosylated Cys10 and Leu117 (labelled) which demonstrates the crowding that would occur if there was no helix movement in cis-MbSNO. The WT and cis-MbSNO structures are shown by cyan and orange cartoon representations, respectively.

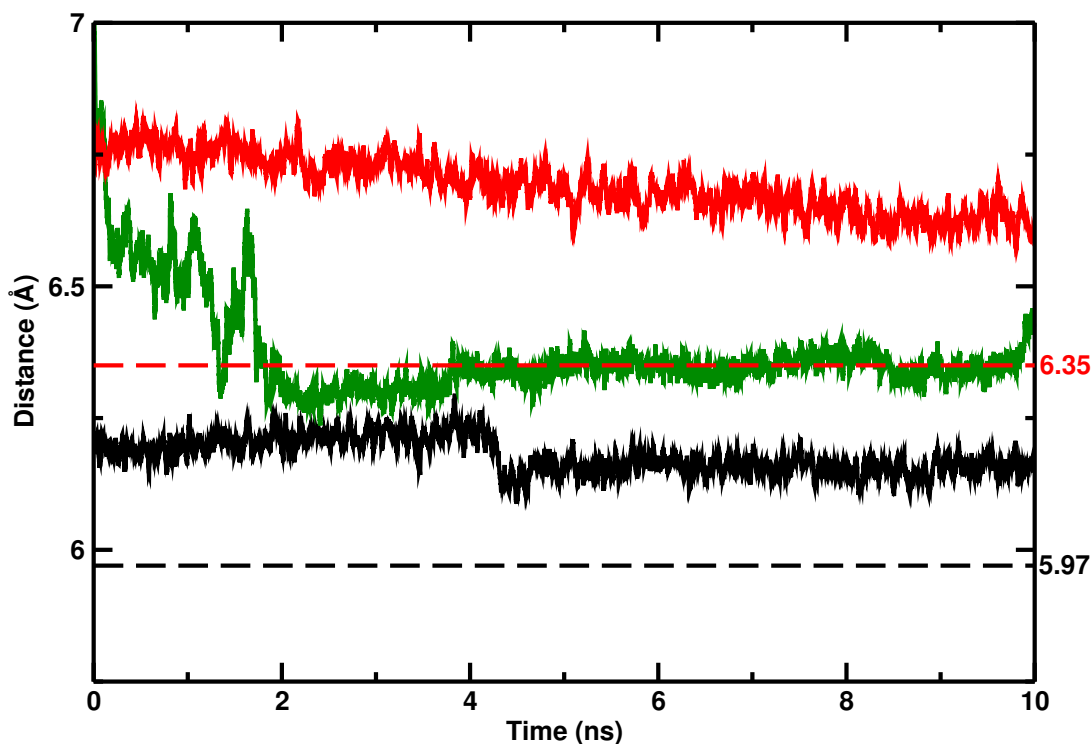
At 300 K, the influence of the NO-modification on the structure of Mb is more pronounced than for 50 K, specifically for loop EF. However, the similar  $C_\alpha$  RMSD for cis- and trans-MbSNO with respect to WT (see Table 4.2) shows that the orientation of the nitric oxide has only a limited effect on the overall structure of the protein. In the conformation with  $\phi(C_\alpha C_\beta SN) \sim -85^\circ$  (for cis-MbSNO with PC (red) and trans-MbSNO with MTP (orange), see Figure 4.8A) the nitric

oxide group resides midway between loop GH and helix A. The conformational transition to  $\phi(C_\alpha C_\beta SN) \sim 160^\circ$  moves the -NO group closer to helix H. For cis-MbSNO with MTPs (green) this is the only state sampled throughout the simulation and corresponds to the major component observed in the 2NRM X-ray structure. For cis-MbSNO with PCs (red) there is a spontaneous transition between the minor and the major conformer. This suggests that the major conformer is probably lower in energy but for a firm conclusion on this considerably more extended simulations are required.

For trans-MbSNO (blue and orange traces in Figure 4.8A) four metastable states for the orientation of  $\phi(C_\alpha C_\beta SN)$  were found in simulations with PCs. The orientation with  $\phi(C_\alpha C_\beta SN) \sim 155^\circ$  is prevalent (Figure 4.8A). In this conformation the NO modification faces towards helix A. After 1 ns, a transition to  $\phi(C_\alpha C_\beta SN) \sim 52^\circ$  occurs. Concomitantly, the  $\phi(NC_\alpha C_\beta S)$  dihedral changes from  $\sim -55$  to  $\sim -180^\circ$  (Figure 4.8B) which positions the nitric oxide towards the GH loop. A next transition leads to  $\phi(C_\alpha C_\beta SN) \sim -70^\circ$  with the NO midway between loop GH and helix A, as was found for cis-MbSNO. Finally, a transition to a state with  $\phi(NC_\alpha C_\beta S) \sim 55^\circ$ , and  $\phi(C_\alpha C_\beta SN) \sim 95^\circ$  occurred in which the NO faced towards the inner part of the protein. One should note that this was the only occasion on which SNO was buried into the protein. This conformation reduces the probability for the solvent accessibility of SNO.

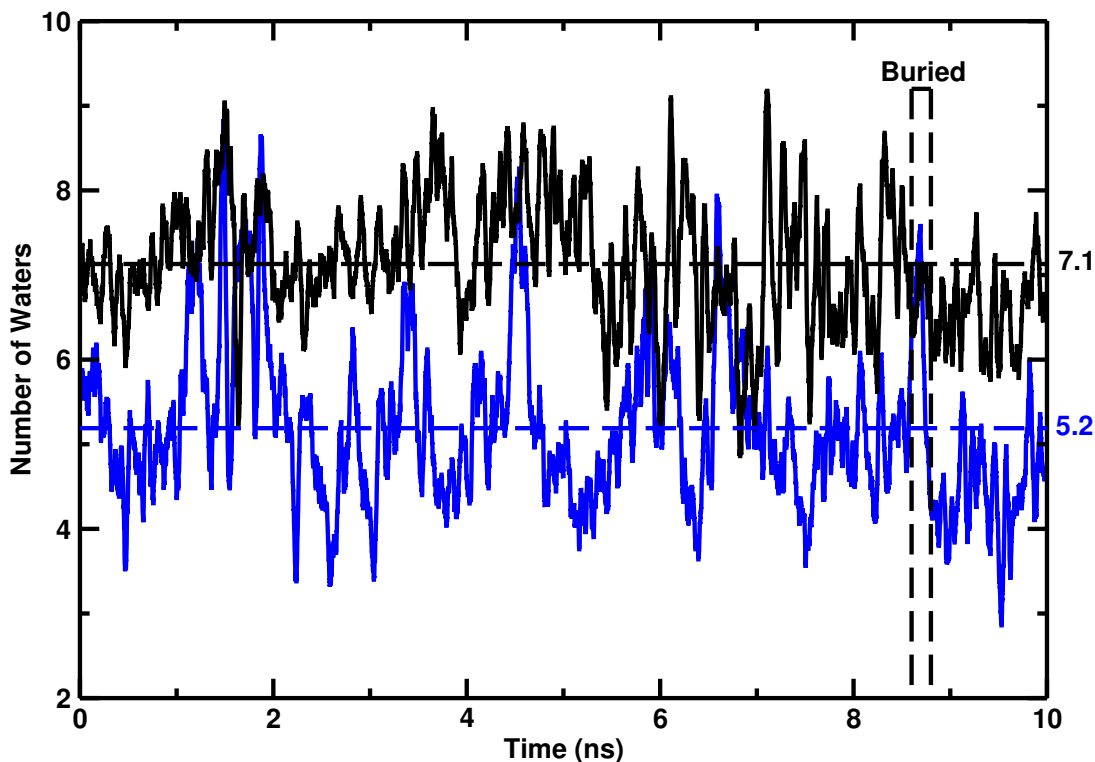
S-nitrosylation-induced structural changes have also been reported for the crystal structures of the WT (2NRL) and cis-MbSNO (2NRM). This is exacerbated by the high  $C_\alpha$  RMSD of loop EF as was observed in the present simulations (see Table 4.2). X-ray experiments revealed that the CysNO can induce crowding between Leu117 (loop GH) and Ala6 (helix A) if these parts of the protein struc-

ture did not move<sup>197</sup> which also been observed in the simulation of cis-MbSNO. The distance between the  $C_\alpha$  atoms of Cys10 and Leu117 ( $r_{10-117}$ ) in the X-Ray structure is 5.97 Å (for 2NRL, WT) and 6.35 Å (for the major conformer of 2NRM). This compares with an averaged  $r_{10-117}$  of 6.18 Å for WT (black trace in Figure 4.20) and 7.39 for cis-MbSNO with PC (red trace in Figure 4.20) in the simulations. The same simulation for cis-MbSNO with MTP leads to relaxation of the structure after  $\sim 2$  ns (green trace in Figure 4.20). Finally, restarting the cis-MbSNO simulation with PC after 10 ns but with MTPs on the -SNO group also leads to relaxation towards the value from the X-ray structure. So, with MTP model,  $r_{10-117}$  assumed the value found in X-Ray whereas with PC it does not. This finding shows that repositioning of loop GH and helix A occurs to accommodate residues Cys10 and Leu117 in cis-MbSNO. The effect is demonstrated in Figure 4.19. Although the averaged  $r_{10-117}$  decreases to 6.37 Å for cis-MbSNO with MTP, the steric overlap is prevented due to sampling the  $\phi(C_\alpha C_\beta SN)$  dihedral at different angles ( $85^\circ$  and  $94^\circ$ ) than cis-MbSNO with PC (red in Figure 4.10). Quantum chemical calculations indicate that the potential energy curve along the CCSN dihedral is flat between  $50^\circ$  and  $300^\circ$  with minima at  $100^\circ$  and  $250^\circ$  and a barrier between them of 2 kcal/mol at the MP2 level of theory.



**Figure 4.20:** Distance between the  $C_{\alpha}$  of Cys10 and Ile117 residues in WT (black), cis-MbSNO with PC (red) and cis-MbSNO with MTP (green) at 50 K. Dash lines represents the Cys10 - Ile117 distance in the crystal structure of WT and major conformer of cis-MbSNO in 2NRM

The structural changes induced by nitrosylation at Cys10 also influence local hydration between the helix A, helix H and loop GH, see Figure 4.21. The chemical modification leads to a decrease of hydration in this region by 30 % between the WT and trans-MbSNO. Given the prominent role that water molecules can play in protein folding,<sup>211</sup> and for function,<sup>212,213</sup> such a change in hydration may also be functionally relevant for a PTM such as nitrosylation. Also, the degree of hydration may affect the stability of the protein as has recently been demonstrated for insulin dimer. Mutation of residue PheB24 to Ala or Gly leads to water influx and destabilization of the dimer by a factor of 2 to 3.<sup>214-218</sup>



**Figure 4.21:** The number of water molecules between helices A and H and around loop GH. All water oxygen atoms within 5 Å of residues 5 to 12 (helix A) and residues 113 to 130 (loop GH, helix H) as a function of time are reported for WT (black) and trans-MbSNO (blue, simulation with PCs). The average occupation is 7.1 water molecules for WT compared with 5.2 for trans-MbSNO, i.e. a difference of 30 %.

## 4.5 Discussion and Conclusions

The present work reported on the structural, dynamical and spectroscopic implications of nitrosylation at cysteine. For this, CysNO as a model and nitrosylated Mb (MbSNO) were considered. For both systems it was found that cis- and trans-orientations can be spectroscopically distinguished. It is of interest to note that for nitrosylated Cysteine cis- and trans-orientations have been observed experimentally in human thioredoxin at position Cys69.<sup>219</sup> Also, nitrosylated Cys62 was completely buried and points towards the protein interior which was also found for a short time during the present simulations for the trans-conformer.

While for CysNO the spectroscopic signatures can be more readily differentiated from other vibrational modes this is more challenging for MbSNO due to the larger number of vibrations and overlap with other vibrational excitations. Nevertheless, the spectroscopic signatures can be clearly located with both, simulations using PC and MTP models, see Figure 4.12. From recording and subtracting the IR spectrum for WT Mb, or the difference spectrum between MbS<sup>14</sup>NO and MbS<sup>15</sup>NO (as for MbNO<sup>172</sup>) it should be possible to identify the IR signatures for nitrosylation in proteins under physiological conditions for which such experiments are possible.

For the properties that are of main interest to the present study (IR spectra, hydration and structural fluctuation around conformational substates) a sampling time of 10 ns is adequate, see Figures 4.15 and 4.17. Infrared spectra computed from different 2 ns intervals exhibit all the essential features, including the amide I to amide III and the NO-stretch bands. Similarly, the  $g(r)$  between water oxygen and the sulfur atom calculated from 2 ns intervals (see Figure 4.17) show that the first solvation shell around the sulfur probe is established after 2 ns. The results indicate that 10 ns simulations provide sufficient sampling for the analysis of the hydration around the SNO probe in a given conformational substate. On the other hand, sampling the transition between the cis- and trans- conformer of the -SNO label is associated with a barrier height of  $\sim 15$  kcal/mol and converging the interconversion dynamics requires micro- to milli-second simulations or using umbrella sampling simulations. Converging this property is, however, outside the scope of the present study. In fact, variable-temperature NMR spectroscopy on R-SNO compounds has reported that even at room temperature this rotation is hindered.<sup>176</sup>



Isotopic substitution ( $^{14}\text{N}\rightarrow^{15}\text{N}$  and  $^{16}\text{O}\rightarrow^{18}\text{O}$ ) leads to red shifts for CysNO and Mb-SNO between  $-24$  and  $-68\text{ cm}^{-1}$ , depending on the isotopes used. This is consistent with expectations from experiments on charged CysNO in the gas phase although no experimental data is available on the neutral systems for direct comparison.

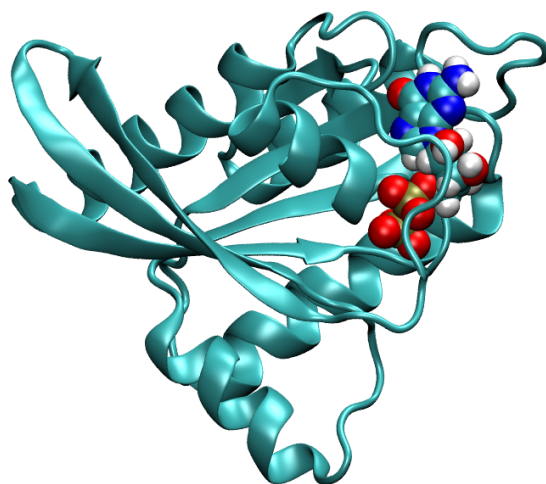
Although the -SNO group is larger compared with -SH, one consequence of nitrosylation is that local hydration increases around the modification site compared with the WT protein and hydration between the A- and H-helices changes as well. For the protein considered here, myoglobin, one additional water molecule was found to be recruited by -SNO compared with -SH for modification at Cys10 within the first solvation shell, which amounts to an increase of  $\sim 10\%$ . Direct observation of this change is likely to be very challenging, e.g. by infrared spectroscopy, as the expected intensity differences are insignificant given the small change in occupation. Although this number may appear small, given the functional role of individual, local water molecules that has been established in other proteins such as HIV-I protease<sup>220,221</sup> where a single catalytic water molecule was found in the active site or insulin<sup>215,216</sup> for which single water molecules attacking the dimerization interface can reduce the thermodynamic stability of the dimer by a factor of two. Thus, the difference in local hydration may also be functionally relevant in Mb-SNO or for solvent-exposed, nitrosylated cysteine residues. Furthermore, S-nitrosylation leads to discernible spectroscopic features in the infrared spectrum of MbSNO and to structural changes near the modification site.

In summary, S-nitrosylation in myoglobin was found to lead to detectable spectroscopic features in the infrared. Local hydration is enhanced as a consequence of the modification and the structure and dynamics at more distant sites in the

protein can be changed appreciably and are consistent with X-ray experiments.

# Chapter 5

## Spectroscopy, Dynamics and Hydration of S-Nitrosylated KRAS



---

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

*J. Phys. Chem. B, 2023, 127, 1526-1539*

*doi: 10.1021/acs.jpcc.2c07371*

\*\*\*



## 5.1 Abstract

S-nitrosylation is an important post-translational modification that can alter the functions of the proteins. The structural dynamics and hydration of S-Nitrosylation in the condensed phase are investigated for the oncogenic KRAS protein with and without GDP binding. For cis-KRASNO, relocation of Switch-I is observed whereas the modification rigidified Loop E, where nitrosylated Cys118 resides, for both cis- and trans-KRASNO. Correlated motions of  $C_{\alpha}$  atoms increased in KRASNO compared to WT, especially for trans-KRASNO. The -SNO group repels water molecules compared with -SH in the first solvation shell, they differ up to 43% between WT and S-nitrosylated variants. Likewise, hydration around the GDP decreased in cis-KRASNO compared to WT and trans-KRASNO by up to 5 water molecules due to the displacement of the Switch-I region. Findings suggest that local hydration can be modulated through nitrosylation. The NO-stretch vibration can be located and distinguished from the other protein modes for simulations of KRASNO at 50K with the  $^{14}\text{N}^{18}\text{O}$  isotope.

## 5.2 Introduction

Nitric oxide (NO) is a cell-signaling molecule that has diverse bio-regulatory functions in the immune, nervous and cardiovascular systems.<sup>147</sup> The production of NO by NO synthase enzymes is a complex and tightly controlled mechanism that orders the specificity of its signaling, limiting the toxicity to other cellular components despite being a free radical with high diffusibility and biological activity.<sup>149</sup> Furthermore, it mediates blood vessel relaxation and plays a role as a neurotransmitter in the central and peripheral nervous system<sup>222</sup>, and its reversible binding to ferrous and ferric heme iron is well characterized.<sup>150</sup> Additionally, NO can also play an important role in the post-translational modification (PTM) of its target

protein.<sup>152</sup>

S-Nitrosylation, the covalent addition of NO to the thiol side chain of cysteine, is an important post-translational modification that can moderate signal transduction. S-Nitrosylation is known to alter the biological functions of the proteins via allosteric regulation<sup>156–159</sup>. Further, it can inhibit or promote the formation of sulfide linkage within or between proteins.<sup>155</sup> More than 3000 proteins are shown to have detectable levels of S-nitrosylation under physiological conditions.<sup>153</sup> S-Nitrosylation is regulated by temporal and spatial arrangements, accurately targeted, and reversible. The specificity of the S-Nitrosylation may be governed by the characteristics, mainly hydrophobicity and  $pK_a$ , of the surrounding region of the target Cys and the physiological concentration.<sup>154</sup> The  $pK_a$  alteration caused by acid-base motifs, i.e. electrostatic interactions, and relative hydrophobic regions can affect the solvent and cofactor accessibility of the region.<sup>155</sup>

The KRAS protein is one of the three members of the RAS oncogene family along with the HRAS and NRAS that plays a role in human cancer.<sup>223</sup> Over 20% of human cancers contain mutated RAS genes which makes them the most frequent oncogenic drivers<sup>224</sup>, while KRAS accounts for 85% of all RAS mutations.<sup>225</sup> Pancreatic (88%), colorectal (50%) and lung cancers (35%) are types of human cancer with the highest rate of KRAS mutations.<sup>226</sup> The most common mutation spots found in RAS family are G12, G13, and Q61. However, KRAS is the only RAS isoform where the G12 mutations such as G12A, G12C, G12D, G12R, G12V are dominant alongside G13D and Q61H.<sup>227,228</sup>

KRAS protein switches between the guanosine diphosphate (GDP)-bound inactive state and the guanosine triphosphate (GTP)-bound active state. The trans-

formation from the stable GDP-bound state to the active GTP-bound state is stimulated by guanine nucleotide exchange factor (GEF) proteins. After the GEF displaces the GDP from the nucleotide-binding site, GTP, which has a higher intracellular concentration than GDP, replaces and binds to the nucleotide-binding site.<sup>229</sup> The transformation back to the inactive GDP-bound state is mediated by GTPase-activating proteins (GAP). The switch between active and inactive states is highly regulated and responsive to multiple signal inputs due to the fact that the switch is controlled by GEFs and GAPs.<sup>223</sup> Further, KRAS in an active GTP-bound state leads to the activation of intracellular transduction signaling pathways such as MAPK and AKT.

The G domain of KRAS encompasses six beta-strands and five alpha-helices, beta-strands form the core of protein while alpha-helices surround the core.<sup>230</sup> The G domain of KRAS is responsible for the biological functions whereas the carboxyl-terminal structure element called the hypervariable region (HVR) is responsible for membrane tethering of the KRAS protein that is required for signaling.<sup>231</sup> HVR is the main difference between the RAS isoforms. There are four main regions surrounding the nucleotide-binding pocket, the phosphate-binding loop (P-loop), switch-I, switch-II, and base binding loops. Switch-I and switch-II are important functional elements of the KRAS that forms a binding interface for effector proteins and RAS regulators. Multiple residue definitions are used for the switch regions but it is hard to define globally due to the high intrinsic flexibility of these regions. The switch-I region lays between helix A and sheet B whereas (residues ~30 to 40) switch-II region is between sheet C to helix B. The switch-II may include, partially or fully, helix B. The P-loop and switch-II are the hotspots for human cancer mutations.

KRAS protein has been studied extensively to be targeted directly/indirectly due to its high incidence and role in different cancer types. Until recently, it was identified as an undruggable oncoprotein due to its complex downstream signaling and lack of accessible binding pockets.<sup>232</sup> However, small molecules which can bind directly to specific KRAS conformers are discovered and changed the perception of “undruggable” KRAS.<sup>233</sup> Although these early-stage molecules are promising, improvements in the terms of potency and specificity for mutated forms of KRAS are needed to work adequately as therapeutics.<sup>234</sup> Further, bacterial Ras/Rap1-specific endopeptidase (RRSP) was studied as a candidate for direct KRAS targeting.<sup>235</sup> RRSP induced proteolytic cleavage to the Switch-I region of KRAS. The cleavage alters the structure around the Switch-I regions and inhibits the interaction between protein and GEFs, thus preventing the activation of the protein. Moreover, approaches developed to indirectly inhibit KRAS signaling, by targeting upstream activators and downstream effectors of the KRAS pathway, were considered.<sup>232</sup> However, no successful strategies have been developed yet.

The amount of S-nitrosylation *in situ* is a small fraction of the total protein mass for globular proteins such as myoglobin and hemoglobin. WT globular proteins are on a millimolar scale whereas S-nitrosylated variants are on a micromolar scale. KRAS is one of the rare examples where 100% S-nitrosylation occurs on Cys118 in the cell.<sup>236</sup> S-Nitrosylation on Cys118 of HRAS stabilizes the GTP-bound state by increasing the dissociation of guanine nucleotides<sup>237</sup>, and mutation of Cys118 inhibits NO-induced Ras activation.<sup>238</sup> It’s been hypothesized that S-Nitrosylation provides a way to diversity Ras-dependent oncogenic signaling alongside the RAS mutations.<sup>184</sup> Further, S-nitrosylation of wild-type Ras protein is essential for the initiation and maintenance of PDAC tumor growth.<sup>239</sup> So, an exhaustive understanding of KRAS structure, dynamics, and hydration



before and after the S-nitrosylation could provide beneficial insights into the effects induced by this modification and its allosteric regulation. In this context, the structural and spectral features of wild-type KRAS and its S-Nitrosylated analogues are studied by means of molecular dynamics simulations. Two conformations of S-Nitrosylated KRAS, *cis*-KRASNO and *trans*-KRASNO, with respect to the  $C_{\beta}$ SNO angle are considered. The present work is structured as follows, first, the atomistic simulations and computational details are described. This is followed by the discussion of the water structure around Cys118 and GDP. Then, structural effects induced by S-nitrosylation are discussed. Next, the IR and power spectra of WT, *cis*-KRASNO, and *trans*-KRASNO are presented and discussed. Finally, the conclusions are drawn.

## 5.3 Computational Methods

### 5.3.1 Molecular Dynamics

All molecular dynamics (MD) simulations were performed using the CHARMM<sup>123</sup> software with the CHARMM36<sup>191</sup> force field. The equations of motion were propagated with a leapfrog integrator<sup>192</sup>, using a time step of  $\Delta t = 1$  fs and all bonds involving hydrogen atoms were constrained using SHAKE.<sup>193</sup> Non-bonded interactions were treated with a switch function<sup>194</sup> between 12 and 16 Å and electrostatic interactions were computed with the particle mesh Ewald method.<sup>195</sup>

For the simulations involving wild-type and S-nitrosylated KRAS, twelve different simulations were set up: wild-type KRAS (PDB: 4OBE)<sup>230</sup> at 50 and 300 K, wild-type KRAS without GDP at 50 and 300 K, *cis*- and *trans*-S-nitrosylated KRASNO at 50 and 300 K, and *cis*- and *trans*-S-nitrosylated KRASNO without GDP at 50 and 300 K. Two chain with identical sequence were available in the

crystal structure. Chain A selected for the simulation setup. The cis-KRASNO and trans-KRASNO set up starting from the same initial structure except for the dihedral angle  $\phi(\text{C}_\beta\text{SNO})$  ( $0^\circ$  for cis- and  $180^\circ$  for trans-KRASNO).

All systems were solvated in a  $66 \times 66 \times 66 \text{ \AA}^3$  cubic box of TIP3P<sup>196</sup> water molecules. The protein was weakly constrained to the middle of the simulation box, minimized, heated to the desired temperature, and equilibrated for 500 ps in the  $NVT$  ensemble. Production runs of 10 ns were then performed in the  $NpT$  ensemble using the leapfrog Verlet integrator<sup>57</sup> ( $\Delta t = 1 \text{ fs}$ ) and a Hoover barostat<sup>124</sup> with a collision rate of  $5 \text{ ps}^{-1}$ . The force field employed for the -SNO moiety was parameterized as described in the previous study.<sup>240</sup>

### 5.3.2 Infrared Spectroscopy

The molecular dipole moment ( $\mu$ ) of the protein was calculated from the MD trajectories and the partial charges. Subsequently, the Fourier transform of the dipole moment autocorrelation function was computed to obtain the infrared spectrum. The autocorrelation function where  $i$  is the index number of a snapshot

$$C(t) = \langle \vec{\mu}(0) \cdot \vec{\mu}(t) \rangle \quad (5.1)$$

was accumulated over  $2^{16}$  time origins to cover  $1/3$  to  $1/2$  of the trajectory. From this, the absorption spectrum is determined according to

$$A(\omega) = \omega(1 - e^{-h\omega/(k_B T)}) \int C(t) e^{-i\omega t} dt \quad (5.2)$$

where  $T$  is the temperature in Kelvin,  $k_B$  is the Boltzmann constant, and the integral is determined using a fast Fourier transform (FFT). IR spectra of WT, cis-KRASNO and trans-KRASNO have been generated for blocks of 100 ps simu-

lation by correlating over  $2^{19}$  time origins. A total of 100 spectra were generated for each system (total simulation time of 10 ns) and averaged.

In addition to IR spectra, the power spectrum of the NO bond was calculated from the FFT of the bond length time-series autocorrelation function to provide assignments of the vibrational spectra and allows to detect couplings between modes.<sup>202</sup> These power spectra were not averaged and correlated over  $2^{16}$  time origins for the entire simulation time of 10 ns.

### 5.3.3 Dynamical Cross-Correlation Maps

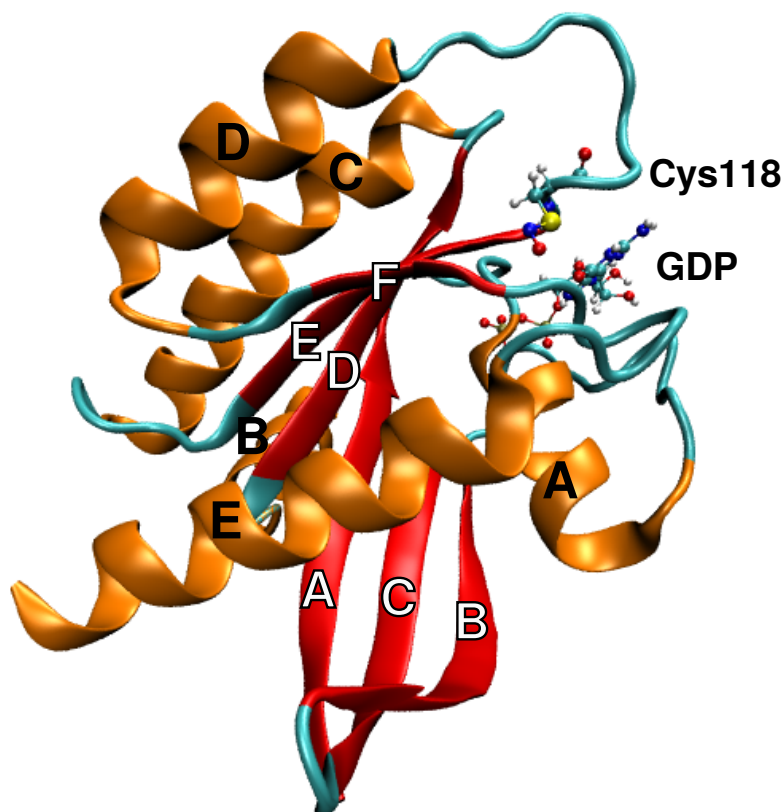
The dynamical cross-correlation maps (DCCM) and difference dynamical cross-correlation maps<sup>241,242</sup> ( $\Delta$ DCCM) were calculated to quantitatively characterize the effects of S-nitrosylation on the protein dynamics using the Bio3D package.<sup>243</sup> Dynamic cross-correlation maps matrices and coefficients

$$C_{ij} = \langle \Delta r_i \cdot \Delta r_j \rangle / (\langle \Delta r_i^2 \rangle \langle \Delta r_j^2 \rangle)^{1/2} \quad (5.3)$$

were determined from the position of  $C\alpha$  in amino acids  $i$  and  $j$  with positions  $r_i$  and  $r_j$ .  $\Delta r_i$  and  $\Delta r_j$  determine the displacement of the  $i$ th  $C\alpha$  from its average position throughout the trajectory. One should note that DCCM characterizes the correlated ( $C_{ij} > 0$ ) and anti-correlated ( $C_{ij} < 0$ ) motions in a protein whereas  $\Delta$ DCCM reports on the distinct differences between unmodified and modified protein.

## 5.4 Results

The present work reported on the solvation, structural, dynamical, and spectroscopic implication of S-nitrosylation at Cysteine for KRAS protein. The structure of trans-KRASNO with GDP is shown in Figure 5.1



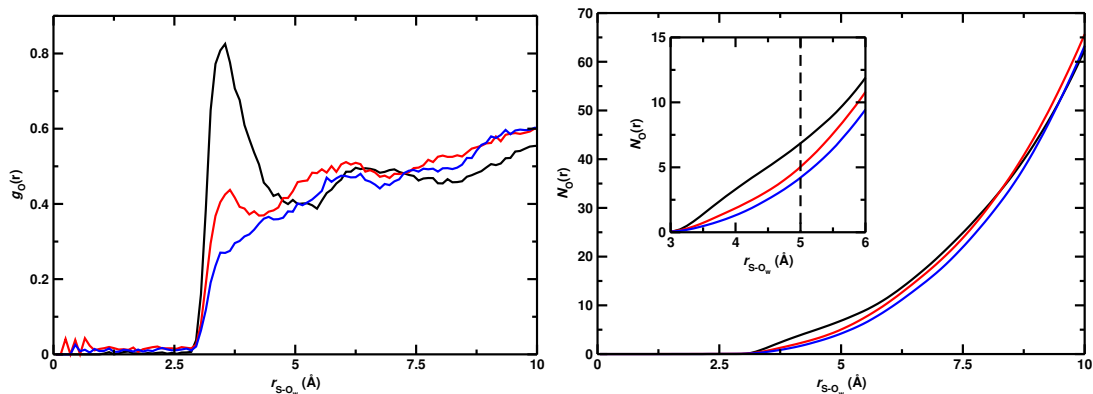
**Figure 5.1:** Structure of trans-KRASNO with GDP. The labeled  $\alpha$ -helices are shown in orange,  $\beta$ -sheets are in red and loops are in cyan. The S-Nitrosylated Cys118 and GDP are represented by CPK.

### 5.4.1 Water Structure

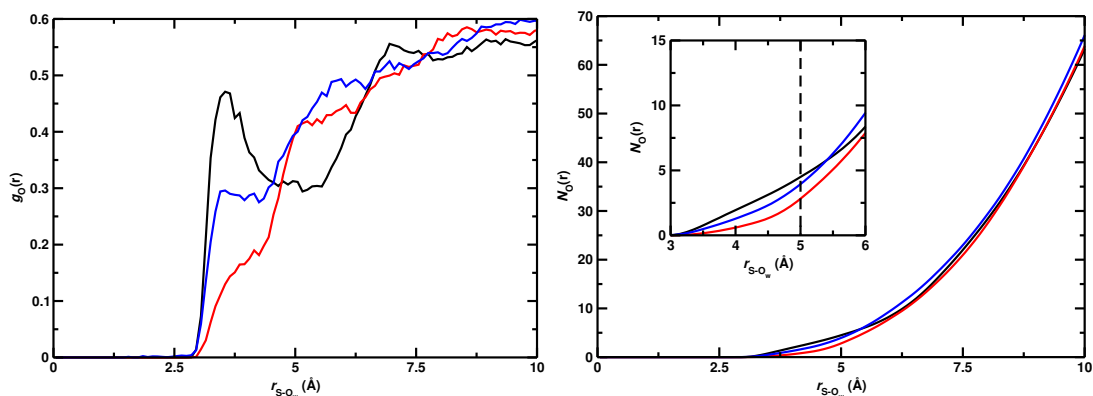
The local water ordering around the modification site (Cys118) for WT and S-nitrosylated KRAS is considered. The radial distribution function  $g_{S-OW}(r)$  and the corresponding number  $N_{S-OW}(r)$  of water oxygen (OW) with respect to the sulfur atom of Cys118 in WT, cis-KRASNO, and trans-KRASNO are shown for both with and without GDP simulations in Figures 5.2 and 5.3. For simulations

with GDP at 300 K, hydration around Cys118 substantially differs between WT and cis-, trans-KRASNO. Both nitrosylated variants were less hydrated with respect to the Cys118-sulfur atom in the range of  $r_{S-OW}$  3–5 Å compared with WT. The first solvation shell peak appears at 3.5 Å in WT and cis-KRASNO whereas no sharp first solvation peak was observed for the trans-KRASNO. The results show that -SNO moiety is less solvent exposed, thus less hydrated, in the trans-conformation compared to cis-. The occupation of the first solvation shell (up to  $r_{S-OW} \sim 5$ ) differs by 2 and 3 water molecules compared to WT for cis- and trans-KRASNO, respectively, which amounts to 43 and 28% decrease for the variants.

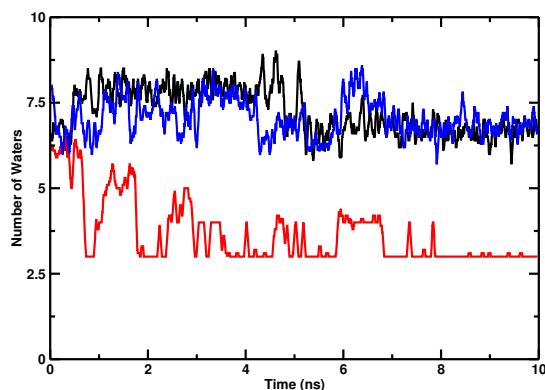
For simulations without GDP at 300 K, again, hydration around Cys118 substantially differs between WT and cis-, trans-KRASNO. The nitrosylated variants were less hydrated with respect to the Cys118-sulfur atom in the range of  $r_{S-OW}$  3–5 Å compared with WT. However, contrary to simulations with GDP, the trans-SNO were more solvent exposed and thus have higher hydration with respect to cis-SNO. The first solvation shell peak for WT and trans-KRASNO appears at 3.5 Å. The occupation of the first solvation shell (up to  $r_{S-OW} \sim 5$ ) differs by 2 and 1 water molecules compared to WT for cis- and trans-KRASNO, respectively, which amounts to 40 and 20% decrease for the variants.



**Figure 5.2:** Radial distribution function of water oxygen (Left) and the corresponding coordination number of  $N_{\text{O}}(r)$  of water oxygen (Right) with respect to the sulfur atom obtained from  $NpT$  simulations at 300 K with PC model. Color code: WT (black), cis-KRASNO (red), trans-KRASNO (blue).



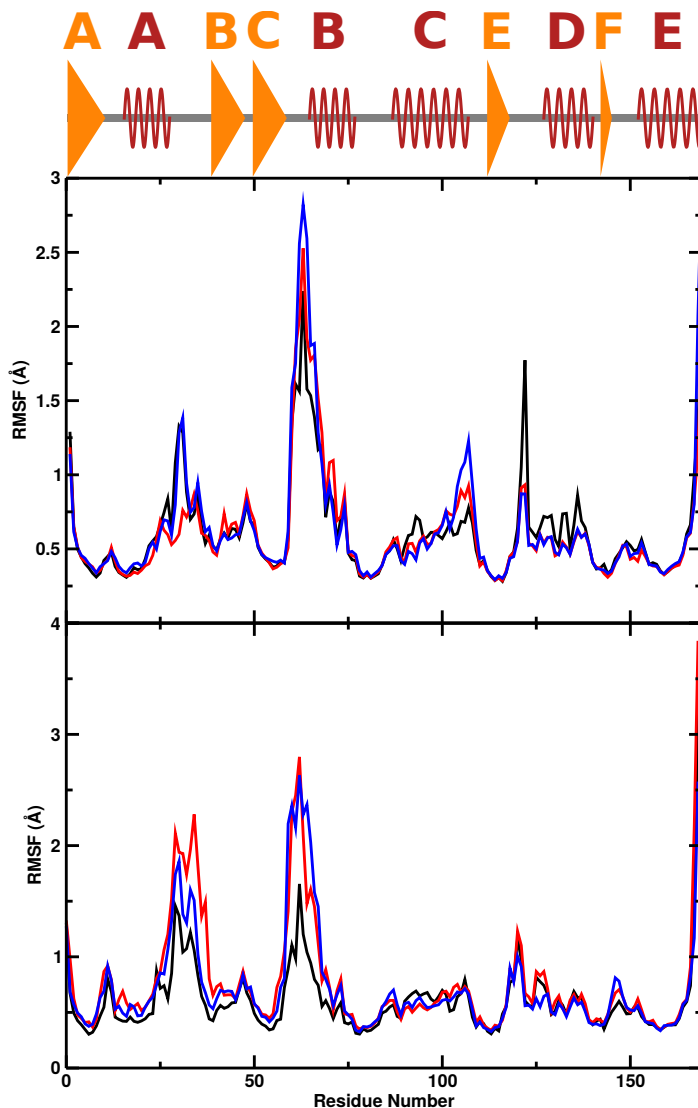
**Figure 5.3:** For the simulations without GDP, the radial distribution function of water oxygen (Left) and the corresponding coordination number of  $N_{\text{O}}(r)$  of water oxygen (Right) with respect to the sulfur atom obtained from  $NpT$  simulations at 300 K with PC model. Color code: WT (black), cis-KRASNO (red), trans-KRASNO (blue).



**Figure 5.4:** The number of waters around (5 Å) the phosphate atoms of GDP in WT (black), cis- (red), and trans-KRASNO (blue) at 300 K as a function of time.

The number of water molecules around ( $5 \text{ \AA}$ ) the phosphate atoms of GDP ( $P_{\text{GDP}}$ ) in WT, cis- and trans-KRASNO at 300 K as a function of time is shown in Figure 5.4. The hydration around  $P_{\text{GDP}}$  is drastically decreased in cis-KRASNO compared to WT and trans-KRASNO. The average number of water molecules was 7.2, 3.7, and 7.0 for WT, cis- and trans-KRASNO, respectively. The decrease in hydration can be rationalized with the increased displacement of the Switch-I region (Phe28 to Asp38) near the GDP in cis-KRASNO with respect to the WT X-Ray structure and WT, trans-KRASNO simulations, (see Table 5.1) which inhibits the solvent accessibility of  $P_{\text{GDP}}$ . Further, the importance of local hydration, even of single catalytic water, and its effects on the function and thermodynamical stability has been established in proteins such as HIV-I protease<sup>220,221</sup> or insulin.<sup>215,216</sup> Consequently, the difference in local hydration may be relevant for KRAS - GDP binding thus the functionality of the protein.

### 5.4.2 Root Mean Square Fluctuation



**Figure 5.5:** Root Mean Square Fluctuation of each residue at 300 K from simulations with (top) and without GDP (bottom) using the PC model for WT (black), cis-KRASNO (red), trans-KRASNO (blue). Orange triangles indicate the position of  $\beta$ -sheets and red helices indicate the position of  $\alpha$ -helices.

The root-mean-squared fluctuations of the  $C_{\alpha}$  atoms of every residue at 300 K from 10 ns simulations with and without GDP are reported in Figure 5.5 (top and bottom) for WT (black), cis-KRASNO (red), and trans-KRASNO (blue). For simulations with GDP, NO attachment decreased the flexibility of residues Phe28 to Asp33 (Loop B) for cis-KRASNO compared to WT and trans-KRASNO. The

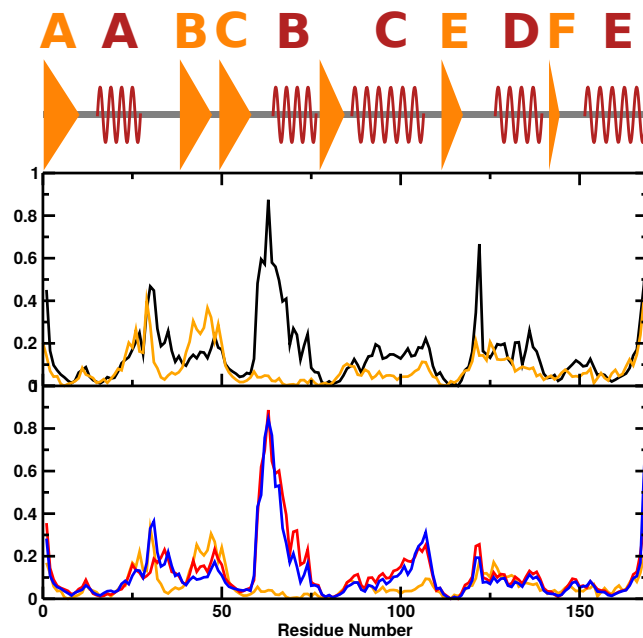


importance of this region is arising from the fact the residues are on the Switch-1 region (Phe28 to Asp38). Switch regions are acts as a binding interface for effector proteins and RAS regulators.<sup>227</sup> So, the rigidification of these regions would have an effect on the binding characteristics of KRAS protein. Moreover, the second switch region (Switch-II, Tyr58 to Tyr64) is also affected by the modification. The modification increased the flexibility of residues Tyr58 to Arg68 in both cis- and trans-KRASNO compared to WT. The RMSF values increased up to 2.8 Å and 2.5 Å for trans- and cis-KRASNO, respectively. Increased flexibility was also observed for the residues Val103 to Asp108 which are resides on the end of Helix C and the beginning of Loop D. Interestingly, the flexibility of Ser122 was significantly higher in WT compared to nitrosylated variants.

For the simulations without GDP, NO attachment induced increased flexibility of both Switch regions. The RMSF values increased for residues His27 to Ile45, the residues reside on Switch-1 (Loop B) and  $\beta$ -sheet B. However, unlike simulations with GDP, both cis- and trans-KRASNO had increased flexibility in this region compared to WT. The RMSF values increased up to 2.3 and 1.9 Å for cis- and trans-KRASNO, respectively. Further, the flexibility increase was also observed in the Switch-II for the residues between Tyr58 and Asp69. The RMSF values increased to 2.8 and 2.6 Å for cis- and trans-KRASNO, the values are on par with RMSF values from with GDP simulations. S-nitrosylation had a marginal effect on the flexibility/rigidification of the rest of the residues. Also, the flexibility of loop E (Cys118 to Asp126) where Cys118 resides was less affected by the NO attachment compared to Loop B and Loop C regions.

The calculated RMSF values at 300 K superimposed onto the measured  $C_{\alpha}$  B-factors and results are shown in Figure 5.6. Comparing the computed RMSF

values for WT with experimental B-factors shows a qualitative agreement between the two except for Switch-II and Helix B region which is range from Try58 to Gly77. The quantitative agreement between experiment and simulation remains consistent until His27 which is the end residue of Helix A. Similar to WT simulations, a qualitative agreement was observed between the experiment and simulations for cis- and trans-KRASNO except for Switch-II and Helix B region. The quantitative agreement again remains consistent until the end of helix A. Additional quantitative agreement from the beginning of Sheet E to the C-terminus of the protein is remarked for the nitrosylated variants. The lack of qualitative agreement for the Switch-II and Helix B region could be due to the fact that two chains are present in the crystal structure. The differences between experiments and simulations also may arise from effects such as crystal packing or lattice disorder which would not be present in simulations.<sup>204</sup>



**Figure 5.6:** Superposition of the experimentally measured  $C_{\alpha}$  B-factors and the computed  $C_{\alpha}$  RMSFs from the present simulations. Up) Experimental WT B-factors (orange) vs. WT (black) RMSF at 300 K. Bottom) WT B-factors vs cis- (red) and trans-KRASNO (blue) RMSF at 300 K. B-Factor and RMSF values are scaled and normalized. Orange triangles indicate the position of  $\beta$ -sheets and red helices indicate the position of  $\alpha$ -helices

### 5.4.3 Structural Changes

The last nanosecond of the 10 ns production run has been analyzed to characterize the structural changes at 50 and 300 K. Even though the structural changes induced by the modification were evident for the entire protein structure to a certain degree, eight regions revealed eminent changes. The regions are sheet D (Ser 65 to Glu76), sheet E (Val112 to Lys117), sheet F (Ile142 to Tyr144), helix C (Tyr87 to Ser106), helix D (Tyr127 to Ile139), helix E (Val152 to Lys169), loop B (Ser39 to Asp47) and Loop E (Cys118 to Asp126). The average  $C_\alpha$  RMSD (in Å) for sheet D, sheet E, sheet F, helix C, helix D, helix E, loop B, loop E, and the entire protein for WT, cis-KRASNO and trans-KRASNO with respect to the WT X-Ray structure for the last nanosecond of a 10 ns free dynamics simulation at 50 and 300 K with PC are summarized in Table 5.1 for simulations with GDP and in Table 5.2 for simulations without GDP.

For simulations with GDP at 50 K, the  $C_\alpha$  RMSD for WT, cis- and trans-KRASNO with respect to WT X-Ray structure are 0.56, 0.67, and 0.58 Å, respectively, which indicates good protein structure preservation. These results show that the addition of nitric oxide had a limited effect on the global structure of the protein at 50 K, but effects were observable at local levels. Especially, the RMSD of Sheet F doubled after the modification, the displacement increased from 0.17 Å in WT to 0.33 and 0.36 Å for cis- and trans-KRASNO, respectively. Moreover, loop B is rigidified compared to WT.  $C_\alpha$  RMSD decreased by 0.2 Å for both variants. The orientation of the NO also affected displacements at the local level. The displacement of helix E increased from 0.42 Å in WT to 0.70 Å in cis-KRASNO whereas it was unaffected in trans-KRASNO. Contrary, displacement increased from 0.37 Å to 0.70 Å in trans-KRASNO whereas marginal effect seen in cis-KRASNO. Since helix D and helix E are in the opposite directions with respect

to Cys118, this finding can be rationalized by the orientation of nitric oxide with respect to  $\phi(C_\beta SNO)$  which is  $0^\circ$  for cis- and  $180^\circ$  for trans-KRASNO. At 300 K, the  $C_\alpha$  RMSD for WT, cis- and trans-KRASNO with respect to WT X-Ray structure are 1.28, 1.75, and 1.53 Å. The effect of the modification to the global structure of the protein was more evident at 300 K, as expected. The RMSD increased by 0.47 and 0.25 Å for cis- and trans-KRASNO, respectively. The local motion of the protein had more evident effects on helix E, loop B, and loop E. The  $C_\alpha$  RMSD of increased to 1.45 Å and 1.87 Å in cis- and trans-KRASNO from 0.91 Å. The helix E represents the last part of the protein sequence from Val152 to the C terminus at Lys159. However, the most dramatic increase in  $C_\alpha$  RMSD was observed for loop B in cis-KRASNO. The displacement increased to 4.42 Å from 0.89 Å in WT. Since, loop B is also the Switch-I part of the protein, this increase in displacement after the S-nitrosylation can have a significant effect on the binding interface characteristics of KRAS to effector proteins and RAS regulators. Also, loop E, where Cys118 resides, showed lower displacement after the modification. The RMSD values are decreased in half for cis- and trans-KRASNO compared to WT.

	50K			300K		
	WT	cis	trans	WT	cis	trans
Sheet D	0.28	0.34	0.24	0.47	0.57	0.56
Sheet E	0.21	0.28	0.18	0.36	0.60	0.43
Sheet F	0.17	0.33	0.36	0.46	0.66	0.40
Helix C	0.37	0.55	0.50	0.99	1.15	1.07
Helix D	0.37	0.40	0.70	0.77	0.95	0.70
Helix E	0.42	0.70	0.43	0.91	1.45	1.87
Loop B	1.15	0.92	0.94	0.89	4.42	1.17
Loop E	0.37	0.49	0.40	1.53	0.85	0.72
Protein	0.56	0.67	0.58	1.28	1.75	1.53

**Table 5.1:** The average  $C_\alpha$  RMSD (in Å) for Sheet D, Sheet E, Sheet F, Helix C, Helix D, Helix E, Loop B, Loop E and the entire protein for WT, cis-KRASNO and trans-KRASNO with respect to the WT X-Ray structure for the last nanosecond of a 10 ns free dynamics simulation at 50 and 300 K with PC.

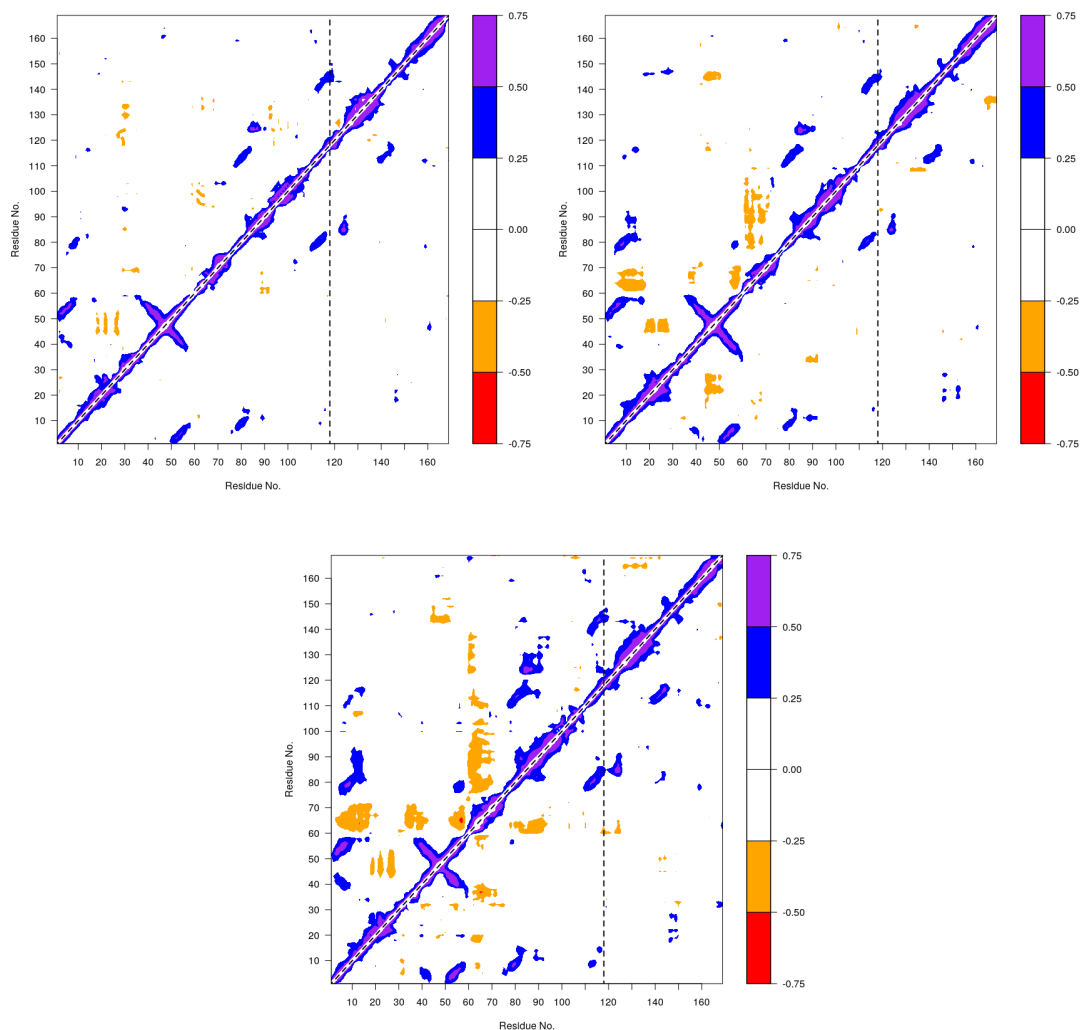
For simulations without GDP at 50 K, the  $C_\alpha$  RMSD for WT, cis- and trans-KRASNO with respect to WT X-Ray structure are 0.56, 0.67, and 0.58 Å, respectively. Again, the modification showed a limited effect on the global structure of the protein at 50 K. The  $C_\alpha$  RMSD was higher in trans-KRASNO compared to cis-KRASNO, except for loop E, on the local level. The highest displacement was seen on loop B (Switch-I), the  $C_\alpha$  RMSD was 1.17 Å in WT and it decreased to 0.99 for cis-KRASNO and 1.06 for trans-KRASNO. Contrary displacement trends for helix D and helix E in cis- and trans-KRASNO did not observe, as was the case for simulations with GDP. At 300 K, the  $C_\alpha$  RMSD for WT, cis- and trans-KRASNO with respect to WT X-Ray structure are 1.79, 1.45, and 1.31 Å, respectively. The  $C_\alpha$  RMSD increased by 0.48 and 0.14 Å for cis- and trans-KRASNO with compared to WT, respectively. The increase of displacement was on a par in magnitude compared to simulations with GDP. The helix E and loop E regions showed lesser displacement after the modification compared to WT. The loop E, where Cys118 resides, is again shown lower displacement in both cis- and trans-KRASNO. The RMSD values decreased to 1.12 and 1.08 Å in cis- and trans-KRASNO compared to 1.74 Å in WT. Contrary to simulations with GDP, higher displacement of loop B is observed both for cis- and trans-KRASNO compared to WT. The RMSD increased to 2.71 Å for cis and 2.27 Å for trans-KRASNO. Interestingly, very low displacement was observed for WT simulations with respect to WT crystal structure for helix D, the RMSD value was only 0.08 Å. Although absolute displacement for helix D was on par compared to simulations with GDP, a relatively substantial increase was observed for cis- and trans-KRASNO compared to WT simulations.

Overall, similar displacement trends were observed for simulations with GDP and without GDP. The cis-KRASNO had higher  $C_\alpha$  RMSD at 300 K compared to trans-KRASNO and WT with respect to WT crystal structure. The lower

displacement of Loop E, where Cys118 resides, demonstrates the effects of nitrosylation on the local motion of the protein. Primarily, the increased RMSD of the switch-I region after the modification in both with and without GDP simulations can alter the binding interface characteristics of the KRAS protein to effector proteins and regulators.

	50K			300K		
	WT	cis	trans	WT	cis	trans
Sheet D	0.43	0.29	0.42	0.53	0.48	0.49
Sheet E	0.40	0.32	0.40	0.75	0.49	0.48
Sheet F	0.26	0.25	0.43	0.57	0.49	0.43
Helix C	0.44	0.39	0.58	1.02	0.93	0.88
Helix D	0.61	0.42	0.52	0.08	1.01	0.81
Helix E	0.49	0.31	0.34	1.99	1.58	1.29
Loop B	1.17	0.99	1.06	1.27	2.71	2.27
Loop E	0.77	0.68	0.61	1.74	1.12	1.08
Protein	0.60	0.54	0.65	1.31	1.79	1.45

**Table 5.2:** For the simulations without GDP, the average  $C_\alpha$  RMSD (in Å) for Sheet D, Sheet E, Sheet F, Helix C, Helix D, Helix E, Loop B, Loop E and the entire protein for WT, cis-KRASNO and trans-KRASNO with respect to the WT X-Ray structure for the last nanosecond of a 10 ns free dynamics simulation at 50 and 300 K with PC.



**Figure 5.7:** The DCCM for WT (Upper Left), cis-KRASNO (Upper Right) and trans-KRASNO (Bottom) at 300 K. Upper matrices is for the simulations with GDP, and lower matrices for the simulations without GDP.

DCCMs characterize the correlated ( $C_{ij} > 0$ ) and anti-correlated ( $C_{ij} < 0$ ) motions within a protein. The DCCM for WT (Upper Left), cis-KRASNO (Upper Right), and trans-KRASNO (Bottom) at 300 K are shown in Figure 5.7, and upper matrices are for the simulations with GDP, and lower matrices for the simulations without GDP. There was three common feature present in the DCCM plots. 1) Positive correlation between sheet A and sheet C, sheet D and loop A, sheet D and sheet E, sheet D and loop E, sheet E and sheet F with  $0.25 \leq C_{ij} \leq 0.75$  is observed. Also, motions between residue Thr50 to Gly60 and Thr35

to Val45 were correlated with  $0.5 \leq C_{ij} \leq 0.75$ . 2) Fewer correlated motions are observed in proteins after removal of the GDP which emphasizes the dynamical effects induced by ligand binding to the KRAS protein. 3) DCCM plots showed that the anti-correlated movement increased after the S-nitrosylation.

To further analyze the effects of the modification, difference maps  $\Delta$ DCCM are calculated for cis-KRASNO and WT ( $\Delta$ DCCM<sub>cis</sub>), trans-KRASNO and WT ( $\Delta$ DCCM<sub>trans</sub>) and presented in Figure 5.8. WT DCCM was selected as a reference for  $\Delta$ DCCM to report on pronounced differences between modified and unmodified proteins. For the simulations without GDP, the  $\Delta$ DCCM results support the observation of less correlated motions for simulations, see Figure 5.8 right panel. For both  $\Delta$ DCCM<sub>cis</sub> and  $\Delta$ DCCM<sub>trans</sub>, no significant correlated motions between the different local parts of the protein have been observed. The residues mainly had correlated movements within their local part with adjacent residues.

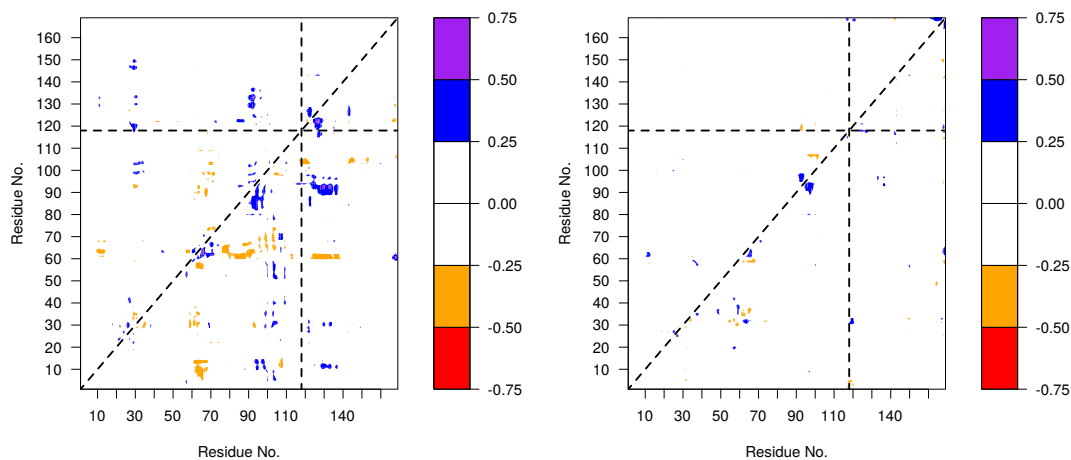
For the simulations with GDP, the local protein dynamics were correlated considerably more compared to simulations without GDP. For  $\Delta$ DCCM<sub>cis</sub>, helix C had a correlated movement with helix D with  $0.25 \leq C_{ij} \leq 0.75$ . Also, the correlated movement of the Switch-I region with multiple regions such as helix C, loop D, and loop F was observed. The results show that nitrosylation not only increased the  $C_\alpha$  RMSD of loop B but also its displacement started to correlate with other local parts of the protein. Especially, the correlated movement between the Switch-I region and helix C shows that residues both can be correlated locally or through space. For  $\Delta$ DCCM<sub>trans</sub>, correlation motion between helix C and helix D was observed with higher intensity. The helix C (Tyr87 to Ser106) and loop F (Glu 107 to Met111) had correlated and anti-correlated motions with residues from sheet A (Met 1 to Gly10) to sheet D (Phe78 to Ile84). Also, loop



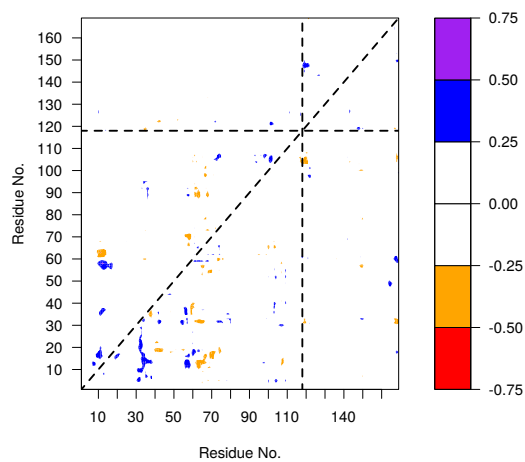
C had anti-correlated movement with residues from Gln70 to Phe90 and Leu120 to Pro140.

Further,  $\Delta$ DCCM are calculated between cis-KRASNO with GDP and without GDP, trans-KRASNO with GDP and without GDP, to quantify the effects of the removal of GDP, and the results are shown in Figure 5.9. In the terms of cis, loop A had correlated motions with helix A and sheet C whereas has anti-correlated motions with loop C. Since the diphosphate part of GDP resides near loop A and helix A, the motion of these regions with respect to other local parts of the protein are important to characterize. Similar trends with higher magnitude were also observed for trans-KRASNO. The prominent correlated motion was observed between sheet B and residues from Lys5 to Asp30 which reside on sheet A, loop A, and helix A. The sheet C had correlated motions with loop A and helix A whereas loop C showed anti-correlated motions with these regions.

Overall, DCCM maps show that S-nitrosylation not only has a significant effect on the individual displacement of the local structures but also affected the correlated motions of the local structures. Several correlated motion features were observed in all systems, but their broadness and intensity increased after the modification, especially for trans-KRASNO. Lastly, the removal of GDP decreased the correlated motions significantly which emphasizes the dynamical effects induced by ligand binding to the KRAS protein

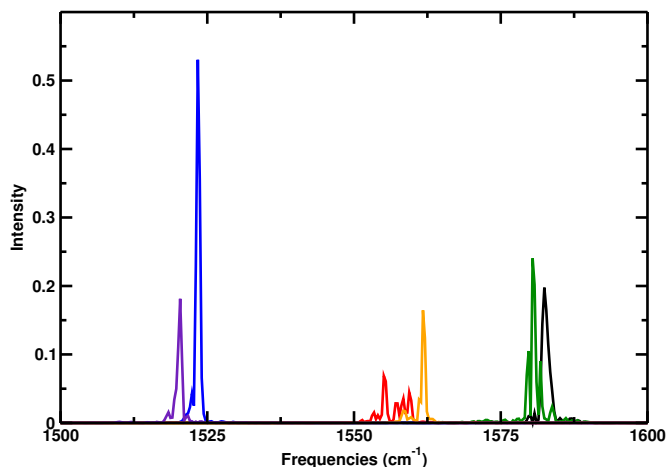


**Figure 5.8:** Left)  $\Delta$ DCCM for cis-KRASNO and WT (upper matrix) and trans-KRASNO and WT (Lower matrix) at 300 K. Right)  $\Delta$ DCCM for cis-KRASNO without GDP and WT without GDP (upper matrix) and trans-KRASNO without GDP and WT without GDP (Lower matrix) at 300 K. The vertical and horizontal lines indicate the position of Cys118.



**Figure 5.9:**  $\Delta$ DCCM for cis-KRASNO with and without GDP (upper matrix) and trans-KRASNO with and without GDP (Lower matrix) at 300 K. The vertical and horizontal lines indicate the position of Cys118.

### 5.4.4 Infrared Spectrum

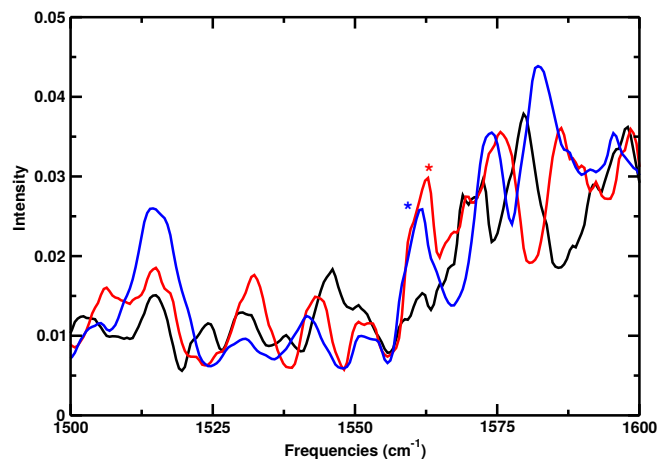


**Figure 5.10:** The power spectra of cis-KRASNO and trans-KRASNO at 50 K with isotopes. Color code as follows: cis-KRASNO with  $^{14}\text{N}^{16}\text{O}$  (black),  $^{14}\text{N}^{18}\text{O}$  (red) and  $^{15}\text{N}^{18}\text{O}$  (blue). trans-KRASNO with  $^{14}\text{N}^{16}\text{O}$  (green),  $^{14}\text{N}^{18}\text{O}$  (yellow) and  $^{15}\text{N}^{18}\text{O}$  (magenta)

The power spectra of cis-KRASNO with  $^{14}\text{N}^{16}\text{O}$  (black),  $^{14}\text{N}^{18}\text{O}$  (red) and  $^{15}\text{N}^{18}\text{O}$  (blue). trans-KRASNO with  $^{14}\text{N}^{16}\text{O}$  (green),  $^{14}\text{N}^{18}\text{O}$  (yellow) and  $^{15}\text{N}^{18}\text{O}$  (magenta) has shown in Figure 5.10. The identical simulations for all three isotopic variants lead to red-shifts of the  $^{14}\text{N}^{18}\text{O}$ -stretch by [-27, -20]  $\text{cm}^{-1}$  for [cis, trans]-KRASNO and  $^{15}\text{N}^{18}\text{O}$ -stretch by [-59, -61]  $\text{cm}^{-1}$  for [cis, trans]-KRASNO compared to the natural isotope  $^{14}\text{N}^{16}\text{O}$ . The NO peak is observed at  $1582\text{ cm}^{-1}$  for the cis-KRASNO with the natural isotope  $^{14}\text{N}^{16}\text{O}$ , and red-shifted to  $1555\text{ cm}^{-1}$  with  $^{14}\text{N}^{18}\text{O}$  and to  $1523\text{ cm}^{-1}$  with  $^{15}\text{N}^{18}\text{O}$ . Similarly, NO peak is observed at  $1581\text{ cm}^{-1}$  for the trans-KRASNO with the natural isotope  $^{14}\text{N}^{18}\text{O}$ , and red-shifted to  $1561\text{ cm}^{-1}$  with  $^{14}\text{N}^{18}\text{O}$  and to  $1520\text{ cm}^{-1}$  with  $^{15}\text{N}^{18}\text{O}$ . The red-shifts observed for the isotopic substitution are on par with previous examples.<sup>240</sup>

The NO peak couldn't be observed with the natural isotope due to congested spectra around the amide I region. Thus, the IR spectra related to the  $^{14}\text{N}^{18}\text{O}$  stretch are reported in Figure 5.11 for WT, cis- and trans-KRASNO with GDP

at 50 K. For the  $^{14}\text{N}^{18}\text{O}$  stretch, the peak for cis-KRASNO from the simulations appears at  $1561\text{ cm}^{-1}$  compared with  $1563\text{ cm}^{-1}$  for trans-KRASNO. Also, the peak intensity was higher for the trans variant with respect to cis. The identified peaks are in agreement with power spectra.



**Figure 5.11:** IR spectra of WT (black), cis-KRASNO (red), and trans-KRASNO with GDP at 50 K. The  $^{14}\text{N}^{18}\text{O}$  isotope is used for the cis- and trans-KRASNO. Spectrum range between  $1500$  and  $1600\text{ cm}^{-1}$  presented. The colored stars label the spectral signatures to which the mode was assigned based on analysis of the power spectra

## 5.5 Conclusion

The structural dynamics, hydration, and spectroscopic detection of S-nitrosylated KRAS protein are reported in the present study. The bulkier -SNO repelled water molecules compared with -SH in the first solvation shell. The number of water molecules around Cys118 with respect to WT is decreased by [2,3] for [cis-,trans]-KRASNO with GDP, and [2,1] for [cis-,trans]-KRASNO without GDP. Further, hydration around phosphate atoms of GDP is drastically decreased in cis-KRASNO due to the relocation of the Switch-I region near GDP. Since the functional role of an individual, local water molecules has been established in proteins, the changes in local hydration around solvent-exposed Cys118 residue

or GDP can be functionally relevant.

The displacement of loop E, where Cys118 resides, decreased after the modification with respect to WT whereas Switch-I substantially relocated in cis-KRASNO with GDP. Considering Switch-I and Switch-II regions are the binding interfaces for effector proteins and RAS regulators, the relocation of these regions may influence binding characteristics. Also, the high flexibility of these regions is captured in the simulations. For Switch-II, which is a hotspot for human cancer mutations, flexibility even increased after the modification.

DCCM maps showed that S-nitrosylation had a significant effect on the correlated motions of the local structures. Several correlated motion features were observed in all systems with varying intensities. Correlated motions decreased significantly after the removal of the GDP.

Distinguishing the spectroscopic signatures of -SNO was challenging for the KRASNO due to a large number of overlapping vibrations, and congested spectra around the amide I region. Isotopic substitutions ( $^{14}\text{N}\rightarrow^{15}\text{N}$  and  $^{16}\text{O}\rightarrow^{18}\text{O}$ ) leads to red-shifts for N-O stretch between  $-20$  to  $-61\text{ cm}^{-1}$  depending on the isotopes used. The red shifts are consistent with the previous studies. The spectroscopic feature of the N-O stretch is distinguished with the  $^{14}\text{N}^{18}\text{O}$  isotope.

In summary, local hydration is decreased as a result of the modification. The structure and dynamics of both local sites and binding interfaces of the protein are altered noticeably. The spectral feature of the N-O stretch mode were detectable in the infrared spectrum with isotopic substitution.



## Chapter 6

# Conclusion and Outlook

This thesis was focused on the research of solvent, structure, and spectroscopic properties of chemical reactions and S-Nitrosylated myoglobin and KRAS proteins. Although “Solvent” and “Structure” dynamics was the main theme throughout the thesis, the “Spectroscopy” refers to the detection of the S-Nitrosylation in proteins. The methods and tools used in this work to investigate the aforementioned properties can be applicable to many other chemical reactions and biomacromolecules.

In Chapter 3, solvent effects on two Menshutkin reactions are quantitatively characterized and analyzed in atomistic level using the MS-ARMD force field. Barrier height reductions going from vacuum to more polar solvents are consistent with the few available experimental studies. The solvent distributions around the solute change noticeably between reactant, transition, and product states. In particular, analysis suggests that collective motion of solvent molecules is needed when approaching the TS. Also, solvent – solvent energetics are directly affected by the changes in solvent structure. Analysis of the solvent degrees of freedom point toward tight coupling between solute and solvent dynamics with increased fluc-

tuations in the solvent-solvent interactions around the transition state. Since the study showed that solvent effects on methyl transfer reactions can be accurately characterized by the MS-ARMD, future research could focus on the biologically crucial phosphate transfer reactions. As an example, GDP to GTP phosphorylation can be an interesting reaction to investigate due to the importance of GTP in signal transduction, or KRAS activation. Also, phosphate ester hydrolysis is another potential reaction to study, that plays role in many biological processes such as energy production, replication of genetic material, and protein synthesis.<sup>2</sup>

In Chapter 4, structural, dynamical, and spectroscopic implications of nitrosylation at cysteine were reported for model system CysNO and S-Nitrosylated myoglobin. The spectroscopic features located for both systems using PC and MTP models, and cis- and trans- orientations were spectroscopically distinguished. Although it is challenging to differentiate spectroscopic signatures from other vibrational modes in MbSNO due to the large number of, and overlap with, other vibrations. It should be possible to identify the IR signatures for S-Nitrosylation on myoglobin under physiological conditions. The local hydration around the S-Nitrosylated Cys is increased after the modification, and -SNO recruits an additional water molecule compared to -SH. Also, the hydration between A- and H-helices is altered as well.

Chapter 5 reports on the structural dynamics, hydration, and spectroscopic detection of S-Nitrosylated KRAS protein with and without GDP binding. There were multiple structural and hydration alterations that can be important for the functionality of the protein. Substantial relocation of the Switch-I region is observed in cis-KRASNO with GDP, and flexibility of Switch-II is increased after the modification. Since Switch regions are binding interfaces for effector proteins



and RAS regulators, these alterations in local structure may affect the binding characteristics of the protein. DCCM maps showed that S-Nitrosylation had a significant effect on the correlated motions of the local structures. Correlated motions decreased significantly after the removal of the GDP. As a future study, it would be of interest to investigate GTP bound, active state, wild-type human KRAS protein and compare the effects of the S-Nitrosylation between inactive and active forms.

Two-dimensional solvent distributions also can be utilized for the local hydration around the modification sites of proteins. The dynamics of the protein can be incorporated with the solvent distribution plots, if the latter is generated as a function of time with given time intervals. The coupling between local structure, e.g. secondary structure, degrees of freedom, and solvent degrees of freedom can provide valuable insights into the understanding interplay between local hydration and structure.

Although the computational infrared spectroscopy is a useful tool to investigate a variety of biological processes, it can be challenged by the complexity of the biomacromolecules. This complexity can lead to spectral congestion and thus difficulty in analyzing and detecting the spectral bands and signatures. An overlap between the spectral signature of the modification and amide bands could result in the inability to detect the modification. The IR spectroscopy would be an even more useful tool for the detection of PTMs if a modification absorbs in the “transparent window” region of the protein spectra which is in the range between 1800 to 2500  $\text{cm}^{-1}$ .



# Bibliography

- [1] Buncel, E.; Stairs, R. A. *Solvent Effects in Chemistry*; John Wiley & Sons, 2015.
- [2] Kamerlin, S. C.; Haranczyk, M.; Warshel, A. Are Mixed Explicit/Implicit Solvation Models Reliable for Studying Phosphate Hydrolysis? A Comparative Study of Continuum, Explicit and Mixed Solvation Models. *ChemPhysChem* **2009**, *10*, 1125–1134.
- [3] Luzhkov, V.; Warshel, A. Microscopic Calculations of Solvent Effects on Absorption Spectra of Conjugated Molecules. *J. Am. Chem. Soc.* **1991**, *113*, 4491–4499.
- [4] Varghese, J. J.; Mushrif, S. H. Origins of Complex Solvent Effects on Chemical Reactivity and Computational Tools to Investigate Them: a Review. *React. Chem. Eng.* **2019**, *4*, 165–206.
- [5] Griffiths, T. R.; Pugh, D. C. Correlations Among Solvent Polarity Scales, Dielectric Constant and Dipole Moment, and a Means to Reliable Predictions of Polarity Scale Values from Cu. *Coord. Chem. Rev.* **1979**, *29*, 129–211.
- [6] Soderberg, T. *Organic Chemistry with a Biological Emphasis Volume I*; 2019.

- [7] Mabesoone, M. F.; Palmans, A. R.; Meijer, E. Solute–Solvent Interactions in Modern Physical Organic Chemistry: Supramolecular Polymers as a Muse. *J. Am. Chem. Soc.* **2020**, *142*, 19781–19798.
- [8] Berezhkovskii, A.; Dudko, S. Solvent Dynamics Influence on Chemical Reaction Dynamics. *J. Chem. Phys.* **1994**, *100*, 5949–5955.
- [9] Reichardt, C. Solvents and Solvent Effects: an Introduction. *Org. Process Res. Dev.* **2007**, *11*, 105–113.
- [10] Papoian, G. A.; Ulander, J.; Eastwood, M. P.; Luthey-Schulten, Z.; Wolynes, P. G. Water in Protein Structure Prediction. *PNAS* **2004**, *101*, 3352–3357.
- [11] Zhou, H.-X.; Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **2018**, *118*, 1691–1741.
- [12] Crick, F. Central Dogma of Molecular Biology. *Nature* **1970**, *227*, 561–563.
- [13] Crick, F. H. On Protein Synthesis. **1958**, *12*, 8.
- [14] Cramer, P.; Armache, K.-J.; Baumli, S.; Benkert, S.; Brueckner, F.; Buchen, C.; Damsma, G. E.; Dengl, S.; Geiger, S. R.; Jasiak, A. J., et al. Structure of Eukaryotic RNA Polymerases. *Annu. Rev. Biophys.* **2008**, *37*, 337–352.
- [15] Allison, L. A. *Fundamental Molecular Biology*; John Wiley & Sons, 2021.
- [16] Proudfoot, N. J.; Furger, A.; Dye, M. J. Integrating mRNA Processing with Transcription. *Cell* **2002**, *108*, 501–512.
- [17] Köhler, A.; Hurt, E. Exporting RNA from the Nucleus to the Cytoplasm. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 761–773.

- [18] Hershey, J. W.; Sonenberg, N.; Mathews, M. B. Principles of Translational Control: an Overview. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, a011528.
- [19] Clark, D. P.; Pazdernik, N. J. *Molecular Biology*; Elsevier, 2013.
- [20] Clancy, S.; Brown, W. Translation: DNA to mRNA to Protein. *Nature Education* **2008**,
- [21] Merrick, W. C.; Pavitt, G. D. Protein Synthesis Initiation in Eukaryotic Cells. *Cold Spring Harb. Perspect. Biol.* **2018**, *10*, a033092.
- [22] Marchingo, J. M.; Cantrell, D. A. Protein Synthesis, Degradation, and Energy Metabolism in T Cell Immunity. *Cell. Mol. Immunol.* **2022**, 1–13.
- [23] Walsh, C. *Posttranslational Modification of Proteins: Expanding Nature's Inventory*; Roberts and Company Publishers, 2006.
- [24] Scott, J. D.; Pawson, T. Cell Signaling in Space and Time: Where Proteins Come Together and When They're Apart. *Science* **2009**, *326*, 1220–1224.
- [25] Nussinov, R.; Tsai, C.-J.; Xin, F.; Radivojac, P. Allosteric Post-Translational Modification Codes. *Trends Biochem. Sci.* **2012**, *37*, 447–455.
- [26] Walsh, C. T.; Garneau-Tsodikova, S.; Gatto Jr, G. J. Protein Post-translational Modifications: the Chemistry of Proteome Diversifications. *Angew. Chem. Int. Ed.* **2005**, *44*, 7342–7372.
- [27] Walsh, C. T.; Malcolmson, S. J.; Young, T. S. Three Ring Post-translational Circuses: Insertion of Oxazoles, Thiazoles, and Pyridines into Protein-derived Frameworks. *ACS Chem. Biol* **2017**, *7*, 429–442.
- [28] Van Der Velden, N. S.; Kälin, N.; Helf, M. J.; Piel, J.; Freeman, M. F.; Künzler, M. Autocatalytic Backbone N-methylation in a Family of Ribosomal Peptide Natural Products. *Nat. Chem. Biol* **2017**, *13*, 833–835.

- [29] Burkhart, B. J.; Schwalen, C. J.; Mann, G.; Naismith, J. H.; Mitchell, D. A. YcaO-dependent Post-translational Amide Activation: Biosynthesis, Structure, and Function. *Chem. Rev.* **2017**, *117*, 5389–5456.
- [30] Case, N.; Thomas, J.; Sen, B.; Styner, M.; Xie, Z.; Galior, K.; Rubin, J. Mechanical Regulation of Glycogen Synthase Kinase 3 $\beta$  (GSK3 $\beta$ ) in Mesenchymal Stem Cells is Dependent on Akt Protein Serine 473 Phosphorylation via mTORC2 Protein. *J. Biol. Chem.* **2011**, *286*, 39450–39456.
- [31] Rocks, O.; Peyker, A.; Kahms, M.; Verveer, P. J.; Koerner, C.; Lumbierres, M.; Kuhlmann, J.; Waldmann, H.; Wittinghofer, A.; Bastiaens, P. I. An Acylation Cycle Regulates Localization and Activity of Palmitoylated Ras Isoforms. *Science* **2005**, *307*, 1746–1752.
- [32] Hagai, T.; Levy, Y. Ubiquitin not Only Serves as a Tag but Also Assists Degradation by Inducing Protein Unfolding. *PNAS* **2010**, *107*, 2001–2006.
- [33] Roskoski Jr, R. Protein Prenylation: a Pivotal Post-translational Process. *Biochem. Biophys. Res. Commun.* **2003**, *303*, 1–7.
- [34] Davies, M. J. Protein Oxidation and Peroxidation. *Biochem. J.* **2016**, *473*, 805–825.
- [35] Horita, H.; Law, A.; Middleton, K. Utilizing a Comprehensive Immunoprecipitation Enrichment System to Identify an Endogenous Post-translational Modification Profile for Target Proteins. *J. Vis. Exp.* **2018**, e56912.
- [36] Jaffrey, S. R.; Erdjument-Bromage, H.; Ferris, C. D.; Tempst, P.; Snyder, S. H. Protein S-nitrosylation: a Physiological Signal for Neuronal Nitric Oxide. *Nat. Cell Biol.* **2001**, *3*, 193–197.
- [37] Voller, A.; Bartlett, A.; Bidwell, D. Enzyme Immunoassays with Special Reference to ELISA Techniques. *J. Clin. Pathol.* **1978**, *31*, 507–520.

- [38] Wang, H.; Xian, M. Chemical Methods to Detect S-nitrosylation. *Curr. Opin. Chem. Biol.* **2011**, *15*, 32–37.
- [39] Alcock, L. J.; Perkins, M. V.; Chalker, J. M. Chemical Methods for Mapping Cysteine Oxidation. *Chem. Soc. Rev.* **2018**, *47*, 231–268.
- [40] Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta - Bioenerg.* **2007**, *1767*, 1073–1101.
- [41] Adhikary, R.; Zimmermann, J.; Romesberg, F. E. Transparent Window Vibrational Probes for the Characterization of Proteins with High Structural and Temporal Resolution. *Chem. Rev.* **2017**, *117*, 1927–1969.
- [42] Lorenz-Fonfria, V. A. Infrared Difference Spectroscopy of Proteins: from Bands to Bonds. *Chem. Rev.* **2020**, *120*, 3466–3576.
- [43] Winter, R.; Pühse, M.; Markgraf, J. *Protein Folding and Misfolding*; Springer, 2012; pp 117–146.
- [44] Reppert, M.; Tokmakoff, A. Computational Amide-I 2D IR Spectroscopy as a Probe of Protein Structure and Dynamics. *Annu. Rev. Phys. Chem.* **2016**, *67*, 359–386.
- [45] Miller, L.; Smith, G.; Carr, G. Synchrotron-Based Biological Microspectroscopy: From the Mid-infrared Through the Far-infrared Regimes. *J. Biol. Phys.* **2003**, *29*, 219–230.
- [46] Born, M.; Oppenheimer, R. Zur Quantentheorie der Molekeln. *Annalen der physik* **1927**, *389*, 457–484.
- [47] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Courier Corporation, 2012.
- [48] Jensen, F. *Introduction to Computational Chemistry*; John Wiley & Sons, 2017.

- [49] MacKerell Jr, A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S., et al. All-atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- [50] Stone, A. J. Distributed Multipole Analysis, or How to Describe a Molecular Charge Distribution. *Chem. Phys. Lett.* **1981**, *83*, 233–239.
- [51] Hedin, F.; El Hage, K.; Meuwly, M. A Toolkit to Fit Nonbonded Parameters from and for Condensed Phase simulations. *J. Chem. Inf. Model.* **2016**, *56*, 1479–1489.
- [52] Rappe, A.; Pietsch, M.; Wiser, D.; Hart, J.; Bormann-Rochotte, L.; Skiff, W. RFF, Conceptual Development of a Full Periodic Table Force Field for Studying Reaction Potential Surfaces. *Mol. Eng.* **1997**, *7*, 385–400.
- [53] Van Duin, A. C.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: a Reactive Force Field for Hydrocarbons. *J. Phys. Chem. A* **2001**, *105*, 9396–9409.
- [54] Schmid, M. H.; Das, A. K.; Landis, C. R.; Meuwly, M. Multi-State VAL-BOND for Atomistic Simulations of Hypervalent Molecules, Metal Complexes, and Reactions. *J. Chem. Theory Comput.* **2018**, *14*, 3565–3578.
- [55] Nagy, T.; Reyes, J. Y.; Meuwly, M. Multisurface Adiabatic Reactive Molecular Dynamics. *J. Chem. Theory. Comput.* **2014**, *10*, 1366–1375.
- [56] Leach, A. R.; Leach, A. R. *Molecular Modelling: Principles and Applications*; Pearson education, 2001.
- [57] Verlet, L. Computer Experiments on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **1967**, *159*, 98–103.



- [58] Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.
- [59] Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- [60] Souaille, M.; Roux, B. Extension to the Weighted Histogram Analysis Method: Combining Umbrella Sampling with Free Energy Calculations. *Comput. Phys. Commun.* **2001**, *135*, 40–57.
- [61] Castejon, H.; Wiberg, K. B. Solvent Effects on Methyl Transfer Reactions. 1. The Menshutkin Reaction. *J. Am. Chem. Soc.* **1999**, *121*, 2139–2146.
- [62] Warshel, A.; Weiss, R. M. An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes. *J. Am. Chem. Soc.* **1980**, *102*, 6218–6226.
- [63] Yadav, A.; Jackson, R. M.; Holbrook, J. J.; Warshel, A. Role of Solvent Reorganization Energies in the Catalytic Activity of Enzymes. *J. Am. Chem. Soc.* **1991**, *113*, 4800–4805.
- [64] Zhan, S.; Mårtensson, D.; Purg, M.; Kamerlin, S. C.; Ahlquist, M. S. Capturing the Role of Explicit Solvent in the Dimerization of Ru<sup>V</sup>(bda) Water Oxidation Catalysts. *Angew. Chemie - Int. Ed.* **2017**, *56*, 6962–6965.
- [65] Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. Transition State Stabilization and Substrate Strain in Enzyme Catalysis: ab initio QM/MM Modelling of the Chorismate Mutase Reaction. *Org. Biomol. Chem.* **2004**, *2*, 968–980.

- [66] Shaw, K. E.; Woods, C. J.; Mulholland, A. J. Compatibility of Quantum Chemical Methods and Empirical (MM) Water Models in Quantum Mechanics/Molecular Mechanics Liquid Water Simulations. *J. Phys. Chem. Lett.* **2010**, *1*, 219–223.
- [67] Carlsson, J.; Åqvist, J. Calculations of Solute and Solvent Entropies from Molecular Dynamics Simulations. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5385–5395.
- [68] Bingemann, D.; King, A. M.; Crim, F. F. Transient Electronic Absorption of Vibrationally Excited CH<sub>2</sub>I<sub>2</sub>: Watching Energy Flow in Solution. *J. Chem. Phys.* **2000**, *113*, 5018–5025.
- [69] Almlöf, M.; Carlsson, J.; Åqvist, J. Improving the Accuracy of the Linear Interaction Energy Method for Solvation Free Energies. *J. Chem. Theory Comput.* **2007**, *3*, 2162–2175.
- [70] Elles, C. G.; Cox, M. J.; Crim, F. F. Vibrational Relaxation of CH<sub>3</sub>I in the Gas Phase and in Solution. *J. Chem. Phys.* **2004**, *120*, 6973–6979.
- [71] Preston, T. J.; Shalowski, M. A.; Crim, F. F. Probing the Photoisomerization of CHBr<sub>3</sub> and CHI<sub>3</sub> in Solution with Transient Vibrational and Electronic Spectroscopy. *J. Phys. Chem. A* **2013**, *117*, 2899–2907.
- [72] Rivera, C. A.; Winter, N.; Harper, R. V.; Benjamin, I.; Bradforth, S. E. The Dynamical Role of Solvent on the ICN Photodissociation Reaction: Connecting Experimental Observables Directly with Molecular Dynamics Simulations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 8269–8283.
- [73] Orr-Ewing, A. J. Perspective: Bimolecular Chemical Reaction Dynamics in Liquids. *J. Chem. Phys.* **2014**, *140*, 090901.

- [74] Claeysens, F.; Ranaghan, K. E.; Lawan, N.; Macrae, S. J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. Analysis of Chorismate Mutase Catalysis by QM/MM Modelling of Enzyme-Catalysed and Uncatalysed Reactions. *Org. Biomol. Chem.* **2011**, *9*, 1578–1590.
- [75] Severance, D. L.; Jorgensen, W. L. Effects of Hydration on the Claisen Rearrangement of Allyl Vinyl Ether from Computer Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10966–10968.
- [76] Guest, J. M.; Craw, J. S.; Vincent, M. A.; Hillier, I. H. The Effect of Water on the Claisen Rearrangement of Allyl Vinyl Ether: Theoretical Methods Including Explicit Solvent and Electron Correlation. *Perkin Trans. 2* **1997**, 71–74.
- [77] Cramer, C. J.; Truhlar, D. G. What Causes Aqueous Acceleration of the Claisen Rearrangement? *J. Am. Chem. Soc.* **1992**, *114*, 8794–8799.
- [78] Brickel, S.; Meuwly, M. Molecular Determinants for Rate Acceleration in the Claisen Rearrangement Reaction. *J. Phys. Chem. B* **2019**, *123*, 448–456.
- [79] Hwang, J. K.; King, G.; Creighton, S.; Warshel, A. Simulation of Free Energy Relationships and Dynamics of  $S_N2$  Reactions in Aqueous Solution. *J. Am. Chem. Soc.* **1988**, *110*, 5297–5311.
- [80] Shaik, S. The Collage of  $S_N2$  Reactivity Patterns: A State Correlation Diagram Model. *Prog. Phys. Org. Chem.* **1985**, *15*, 197–337.
- [81] Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. Theoretical Examination of the  $S_N2$  Reaction Involving Chloride Ion and Methyl Chloride in the Gas Phase and Aqueous Solution. *J. Am. Chem. Soc.* **1985**, *107*, 154–163.

- [82] Merkel, A.; Zahradník, R.; Havlas, Z. Evaluation of the Rate Constant for the  $S_N2$  Reaction  $\text{CH}_3\text{F} + \text{H}^- \rightarrow \text{CH}_4 + \text{F}^-$  in the Gas Phase. *J. Am. Chem. Soc.* **1988**, *110*, 8355–8359.
- [83] Gao, J.; Xia, X. A Two-Dimensional Energy Surface for a type II  $S_N2$  Reaction in Aqueous Solution. *J. Am. Chem. Soc.* **1993**, *115*, 9667–9675.
- [84] Fradera, X.; Amat, L.; Torrent, M.; Mestres, J.; Constans, P.; Besald, E.; Marti, J.; Simon, S.; Lobato, M.; Oliva, J. M.; Luis, J. M.; Sol, M.; Carbb, R.; Duran, M. Analysis of the Changes on the Potential Energy Surface of Menshutkin Reactions Induced by External Perturbations. *J. Mol. Struct.* **1996**, *371*, 171–183.
- [85] Adamovic, I.; Gordon, M. S. Solvent Effects on the  $S_N2$  Reaction: Application of the Density Functional Theory-Based Effective Fragment Potential Method. *J. Phys. Chem. A* **2005**, *109*, 1629–1636.
- [86] Shaik, S.; Ioffe, A.; Reddy, A. C.; Pross, A. Is the Avoided Crossing State a Good Approximation for the Transition State of a Chemical Reaction? An Analysis of Menshutkin and Ionic  $S_N2$  Reactions. *J. Am. Chem. Soc.* **2006**, *116*, 262–273.
- [87] Menshutkin, N. Beiträge zur Kenntnis der Affinitätskoeffizienten der Alkylhaloide und der Organischen Amine. *Zeitschrift für Phys. Chemie* **1890**, *5*, 589–600.
- [88] Menshutkin, N. Über die Affinitätskoeffizienten der Alkylhaloide und der Amine. *Zeitschrift für Phys. Chemie* **1890**, *6*, 41–57.
- [89] Allfrey, V. G.; Faulkner, R.; Mirsky, A. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of RNA Synthesis. *Proc. Nat. Acad. Sci. USA* **1964**, *51*, 786.

- [90] Schubert, H. L.; Blumenthal, R. M.; Cheng, X. Many Paths to Methyltransfer: A Chronicle of Convergence. *Trends Biochem. Sci.* **2003**, *28*, 329–335.
- [91] Schmidt, T.; Schwede, T.; Meuwly, M. Computational Analysis of Methyl Transfer Reactions in Dengue Virus Methyltransferase. *J. Phys. Chem. B* **2014**, *118*, 5882–5890.
- [92] Arava, S.; Diesendruck, C. E. Strategies for the Synthesis of N-Arylammonium Salts. *Synthesis* **2017**, *49*, 3535–3545.
- [93] Ding, W.; Smulan, L. J.; Hou, N. S.; Taubert, S.; Watts, J. L.; Walker, A. K. S-Adenosylmethionine Levels Govern Innate Immunity Through Distinct Methylation-Dependent Pathways. *Cell Metab.* **2015**, *22*, 633–645.
- [94] Scavetta, R. D.; Thomas, C. B.; Walsh, M. A.; Szegedi, S.; Joachimiak, A.; Gumport, R. I.; Churchill, M. E. Structure of Rsr I Methyltransferase, a Member of the N 6-adenine  $\beta$  Class of DNA Methyltransferases. *Nucleic Acids Res.* **2000**, *28*, 3950–3961.
- [95] Tajima, S.; Suetake, I.; Takeshita, K.; Nakagawa, A.; Kimura, H. Domain Structure of the Dnmt1, Dnmt3a, and Dnmt3b DNA Methyltransferases. *Adv. Exp. Med. Biol.* **2016**, *945*, 63–86.
- [96] Li, Y.; Chen, X.; Lu, C. The Interplay Between DNA and Histone Methylation: Molecular Mechanisms and Disease Implications. *EMBO Rep.* **2021**, e51803.
- [97] Shi, Y.; Lan, F.; Matson, C.; Mulligan, P.; Whetstine, J. R.; Cole, P. A.; Casero, R. A.; Shi, Y. Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. *Cell* **2004**, *119*, 941–953.

- [98] Lieder, C. A.; Brauman, J. I. A Technique for Detection of Neutral Products in Gas-Phase, Ion-Molecule Reactions. *Int. J. Mass Spectrom. Ion Phys.* **1975**, *16*, 307–319.
- [99] Olmstead, W. N.; Brauman, J. I. Gas-Phase Nucleophilic Displacement Reactions. *J. Am. Chem. Soc.* **1977**, *99*, 4219–4228.
- [100] Hierl, P. M.; Ahrens, A. F.; Henschman, M.; Viggiano, A. A.; Paulson, J. F. Rate Constants and Product Distributions as Functions of Temperature for the Reaction of  $\text{OH}^-(\text{H}_2\text{O})_{0,1,2}$  with  $\text{CH}_3\text{CN}$ . *Int. J. Mass Spectrom. Ion Process.* **1987**, *81*, 101–122.
- [101] Barlow, S. E.; Van Doren, J. M.; Bierbaum, V. M. The Gas-Phase Displacement Reaction of Chloride Ion with Methyl Chloride as a Function of Kinetic Energy. *J. Am. Chem. Soc.* **1988**, *110*, 7240–7242.
- [102] Vande Linde, S. R.; Hase, W. L. A Direct Mechanism for  $\text{S}_{\text{N}}2$  Nucleophilic Substitution Enhanced by Mode Selective Vibrational Excitation. *J. Am. Chem. Soc.* **1989**, *111*, 2349–2351.
- [103] DePuy, C. H.; Gronert, S.; Mullin, A.; Bierbaum, V. M. Gas-Phase  $\text{S}_{\text{N}}2$  and  $\text{E}2$  Reactions of Alkyl Halides. *J. Am. Chem. Soc.* **1990**, *112*, 8650–8655.
- [104] Kato, S.; Davico, G. E.; Lee, H. S.; DePuy, C. H.; Bierbaum, V. M. Deuterium Kinetic Isotope Effects in Gas Phase  $\text{S}_{\text{N}}2$  Reactions. *Int. J. Mass Spectrom.* **2001**, *210-211*, 223–229.
- [105] Stei, M.; Carrascosa, E.; Kainz, M. A.; Kelkar, A. H.; Meyer, J.; Szabó, I.; Czakó, G.; Wester, R. Influence of the Leaving Group on the Dynamics of a Gas-Phase  $\text{S}_{\text{N}}2$  Reaction. *Nat. Chem.* **2016**, *8*, 151–156.

- [106] Carrascosa, E.; Meyer, J.; Zhang, J.; Stei, M.; Michaelsen, T.; Hase, W. L.; Yang, L.; Wester, R. Imaging Dynamic Fingerprints of Competing E2 and S<sub>N</sub>2 Reactions. *Nat. Commun.* **2017**, *8*, 1–7.
- [107] Viers, J. W.; Schug, J. C.; Stovall, M. D.; Seeman, J. I. MNDO Study of Reaction Pathways for S<sub>N</sub>2 Reactions. Menshutkin Reaction Potential Energy Surfaces. *J. Comp. Chem.* **1984**, *5*, 598–605.
- [108] Amovilli, C.; Mennucci, B.; Floris, F. M. MCSCF Study of the S<sub>N</sub>2 Menshutkin Reaction in Aqueous Solution within the Polarizable Continuum Model. *J. Phys. Chem. B* **1998**, *102*, 3023–3028.
- [109] Castejon, H.; Wiberg, K. B.; Sklenak, S.; Hinz, W. Solvent Effects on Methyl Transfer Reactions. 2. The Reaction of Amines with Trimethylsulfonium Salts. *J. Am. Chem. Soc.* **2001**, *123*, 6092–6097.
- [110] Sola, M.; Lledos, A.; Duran, M.; Bertran, J.; Abboud, J. L. M. Analysis of Solvent Effects on the Menshutkin Reaction. *J. Am. Chem. Soc.* **1991**, *113*, 2873–2879.
- [111] Jiang, L.; Orimoto, Y.; Aoki, Y. Substituent Effects on Menshutkin-Type Reactions in the Gas Phase and Solutions: Theoretical Approach from the Orbital Interaction View. *J. Chem. Theory Comput.* **2013**, *9*, 4035–4045.
- [112] Acevedo, O.; Jorgensen, W. L. Exploring Solvent Effects Upon the Menshutkin Reaction Using a Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 8425–8430.
- [113] Gronet, S.; DePuy, C. H.; Bierbaum, V. M. Deuterium Isotope Effects in Gas-Phase Reactions of Alkyl Halides: Distinguishing E2 and S<sub>N</sub>2 Pathways. *J. Am. Chem. Soc.* **1991**, *113*, 4009–4010.

- [114] Auriel, M.; de Hoffmann, E. Quantitative Study of Solvent Effects on the Menshutkin Reaction Between 1, 4-diazabicyclo [2.2.2] octane and 2-chloroethylbenzene, 2-bromoethylbenzene, and 2-iodoethylbenzene. Part 2. Mixed Solvents. *J. Chem. Soc., Perkin trans. 2* **1979**, 325–329.
- [115] Maran, U.; Pakkanen, T. A.; Karelson, M. Semiempirical Study of the Solvent Effect on the Menshutkin Reaction. *J. Chem. Soc., Perkin trans. 2* **1994**, 2445–2452.
- [116] Komarova, A. O.; Dick, G. R.; Luterbacher, J. S. Diformylxylose as a New Polar Aprotic Solvent Produced from Renewable Biomass. *Green Chem.* **2021**, *23*, 4790–4799.
- [117] Dutta Dubey, K.; Stuyver, T.; Kalita, S.; Shaik, S. Solvent Organization and Rate Regulation of a Menshutkin Reaction by Oriented External Electric Fields are Revealed by Combined MD and QM/MM Calculations. *J. Am. Chem. Soc.* **2020**, *142*, 9955–9965.
- [118] Poater, J.; Solà, M.; Duran, M.; Fradera, X. Effects of Solvation on the Pairing of Electrons in a Series of Simple Molecules and in the Menshutkin Reaction. *J. Phys. Chem. A* **2001**, *105*, 6249–6257.
- [119] Halls, M. D.; Schlegel, H. B. Chemistry Inside Carbon Nanotubes: the Menshutkin  $S_N2$  Reaction. *J. Phys. Chem. B* **2002**, *106*, 1921–1925.
- [120] Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum. A Direct Utilizaion of *ab initio* Molecular Potentials for the Prevision of Solvent Effects. *Chem. Phys.* **1981**, *55*, 117–129.
- [121] Pan, X.; Li, P.; Ho, J.; Pu, J.; Mei, Y.; Shao, Y. Accelerated Computation of Free Energy Profile at *ab initio* Quantum Mechanical/Molecular Mechanical Accuracy via a Semi-Empirical Reference Potential. II. Recalibrat-



- ing Semi-Empirical Parameters with Force Matching. *Phys. Chem. Chem. Phys.* **2019**, *21*, 20595–20605.
- [122] Tang, W.; Zhao, J.; Jiang, P.; Xu, X.; Zhao, S.; Tong, Z. Solvent Effects on the Symmetric and Asymmetric  $S_N2$  Reactions in the Acetonitrile Solution: A Reaction Density Functional Theory Study. *J. Phys. Chem. B* **2020**, *124*, 3114–3122.
- [123] Brooks, B. R. et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- [124] Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697.
- [125] VanGunsteren, W.; Berendsen, H. Algorithms for Macromolecular Dynamics and Constraint Dynamics. *Mol. Phys.* **1977**, *34*, 1311–1327.
- [126] Essmann, U.; Perera, L.; Darden, M. L. B.; Lee, H.; Pedersen, L. G.; Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- [127] Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I., et al. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31*, 671–690.
- [128] Frisch, M. J. et al. Gaussian09 Revision D.01. Gaussian Inc. Wallingford CT 2009.
- [129] Mahmood, A.; Longo, R. L. Structural and Relative Energy Assessments of DFT functionals and the MP2 Method to Describe the Gas Phase Methy-

- lation of Nitronates:  $[R^1R^2CNO_2]^- + CH_3I$ . *Phys. Chem. Chem. Phys.* **2016**, *18*, 17062–17070.
- [130] Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules. *J. Comput. Chem.* **2011**, *32*, 2959–2368.
- [131] Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural Population Analysis. *J. Chem. Phys.* **1985**, *83*, 735–746.
- [132] Brickel, S.; Das, A.; Unke, O.; Turan, H.; Meuwly, M. Reactive Molecular Dynamics for the  $[Cl-CH_3-Br]^-$  Reaction in the Gas Phase and in Solution: A Comparative Study Using Empirical and Neural Network Force Fields. *Electron. Struct.* **2019**, *1*, 024002.
- [133] Brickel, S.; Meuwly, M. OH-Stretching Overtone Induced Dynamics in  $HSO_3F$  from Reactive Molecular Dynamics Simulations. *J. Phys. Chem. A* **2017**, *121*, 5079–5087.
- [134] Reyes, J. Y.; Brickel, S.; Unke, O. T.; Nagy, T.; Meuwly, M.  $HSO_3Cl$ : a Prototype Molecule for Studying OH-stretching Overtone Induced Photodissociation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 6780–6788.
- [135] Nelder, J.; Mead, R. A Simplex Method for Function Minimization. *Chem. Phys.* **1965**, *7*, 308–313.
- [136] Kottalam, J.; Case, D. A. Dynamics of Ligand Escape from the Heme Pocket of Myoglobin. *J. Am. Chem. Soc.* **1988**, *110*, 7690–7697.
- [137] Rosenblatt, M. Remarks on Some Nonparametric Estimates of a Density Function. *Ann. Math. Stat.* **1956**, *27*, 832–837.
- [138] Parzen, E. On Estimation of a Probability Density Function and Mode. *Ann. Math. Stat.* **1962**, *33*, 1065–1076.

- [139] Sweeny, B. C.; Pan, H.; Kassem, A.; Sawyer, J. C.; Ard, S. G.; Shuman, N. S.; Viggiano, A. A.; Brickel, S.; Unke, O. T.; Upadhyay, M., et al. Thermal Activation of Methane by  $\text{MgO}^+$ : Temperature Dependent Kinetics, Reactive Molecular Dynamics Simulations and Statistical Modeling. *Phys. Chem. Chem. Phys.* **2020**, *22*, 8913–8923.
- [140] Rivero, U.; Turan, H. T.; Meuwly, M.; Willitsch, S. Reactive Atomistic Simulations of Diels-Alder Type Reactions: Conformational and Dynamic Effects in the Polar Cycloaddition of 2, 3-dibromobutadiene Radical Ions with Maleic Anhydride. *Mol. Phys.* **2021**, *119*, e1825852.
- [141] Okamoto, K.; Fukui, S.; Shingu, H. Kinetic Studies of Bimolecular Nucleophilic Substitution. VI. Rates of the Menschkin Reaction of Methyl Iodide with Methylamines and Ammonia in Aqueous Solutions. *Bull. Chem. Soc. of Jpn* **1967**, *40*, 1920–1925.
- [142] Okamoto, K.; Fukui, S.; Nitta, I.; Shingu, H. Kinetic Studies of Bimolecular Nucleophilic Substitution. VIII. The Effect of Hydroxylic Solvents on the Nucleophilicity of Aliphatic Amines in the Menschutkin Reaction. *Bull. Chem. Soc. of Jpn* **1967**, *40*, 2354–2357.
- [143] Gao, J. A Priori Computation of a Solvent-Enhanced  $\text{S}_{\text{N}}2$  Reaction Profile in Water: the Menschutkin Reaction. *J. Am. Chem. Soc.* **1991**, *113*, 7796–7797.
- [144] Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
- [145] Reichardt, C. *Solvents and Solvent Effects in Organic Chemistry*; Wiley, 2006.

- [146] El Hage, K.; Hedin, F.; Gupta, P. K.; Meuwly, M.; Karplus, M. Valid Molecular Dynamics Simulations of Human Hemoglobin Require a Surprisingly Large Box Size. *Elife* **2018**, *7*, e35560.
- [147] Sessa, W. C. The Nitric Oxide Synthase Family of Proteins. *J. Vasc. Res.* **1994**, *31*, 131–143.
- [148] Houk, K.; Hietbrink, B. N.; Bartberger, M. D.; McCarren, P. R.; Choi, B. Y.; Voyksner, R. D.; Stamler, J. S.; Toone, E. J. Nitroxyl Disulfides, Novel Intermediates in Transnitrosation Reactions. *J. Am. Chem. Soc.* **2003**, *125*, 6972–6976.
- [149] Kone, B. C.; Kuncewicz, T.; Zhang, W.; Yu, Z.-Y. Protein Interactions with Nitric Oxide Synthases: Controlling the Right Time, the Right Place, and the Right Amount of Nitric Oxide. *Am. J. Physiol. Renal. Physiol.* **2003**, *285*, F178–F190.
- [150] Cooper, C. E. Nitric Oxide and Iron Proteins. *Biochim. Biophys. Acta* **1999**, *1411*, 290–309.
- [151] Russwurm, M.; Koesling, D. NO Activation of Guanylyl Cyclase. *EMBO J.* **2004**, *23*, 4443–4450.
- [152] Astier, J.; Lindermayr, C. Nitric Oxide-Dependent Post-translational Modification in Plants: an Update. *Int. J. Mol. Sci.* **2012**, *13*, 15193–15208.
- [153] Hess, D. T.; Stamler, J. S. Regulation by S-nitrosylation of Protein Post-translational Modification. *J. Biol. Chem.* **2012**, *287*, 4411–4418.
- [154] Sun, J.; Steenbergen, C.; Murphy, E. S-nitrosylation: NO-related Redox Signaling to Protect Against Oxidative Stress. *Antioxid. Redox Signal.* **2006**, *8*, 1693–1705.

- [155] Hess, D. T.; Matsumoto, A.; Kim, S.-O.; Marshall, H. E.; Stamler, J. S. Protein S-nitrosylation: Purview and Parameters. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 150–166.
- [156] Yasukawa, T.; Tokunaga, E.; Ota, H.; Sugita, H.; Martyn, J. J.; Kaneki, M. S-Nitrosylation Dependent Inactivation of Akt/Protein Kinase B in Insulin Resistance. *J. Biol. Chem.* **2005**, *280*, 7511–7518.
- [157] Wu, M.; Katta, A.; Gadde, M. K.; Liu, H.; Kakarla, S. K.; Fannin, J.; Paturi, S.; Arvapalli, R. K.; Rice, K. M.; Wang, Y., et al. Aging-Associated Dysfunction of Akt/Protein Kinase B: S-Nitrosylation and Acetaminophen Intervention. *PLoS One* **2009**, *4*, e6430.
- [158] Chen, Y.-Y.; Chu, H.-M.; Pan, K.-T.; Teng, C.-H.; Wang, D.-L.; Wang, A. H.-J.; Khoo, K.-H.; Meng, T.-C. Cysteine S-nitrosylation Protects Protein-Tyrosine Phosphatase 1B Against Oxidation-Induced Permanent Inactivation. *J. Biol. Chem.* **2008**, *283*, 35265–35272.
- [159] Barrett, D. M.; Black, S. M.; Todor, H.; Schmidt-Ullrich, R. K.; Dawson, K. S.; Mikkelsen, R. B. Inhibition of Protein-Tyrosine Phosphatases by Mild Oxidative Stresses is Dependent on S-nitrosylation. *J. Biol. Chem.* **2005**, *280*, 14453–14461.
- [160] Zhou, C.; Liang, J.; Cheng, S.; Shi, T.; Houk, K.; Wei, D.-Q.; Zhao, Y.-L. Ab initio Molecular Metadynamics Simulation for S-nitrosylation by Nitric Oxide: S-Nitroxide as the Key Intermediate. *Mol. Simul.* **2017**, *43*, 1134–1141.
- [161] Gow, A. J.; Buerk, D. G.; Ischiropoulos, H. A Novel Reaction Mechanism for the Formation of S-Nitrosothiol in vivo. *J. Biol. Chem.* **1997**, *272*, 2841–2845.

- [162] Martinez-Ruiz, A.; Araújo, I. M.; Izquierdo-Alvarez, A.; Hernansanz-Agustin, P.; Lamas, S.; Serrador, J. M. Specificity in S-nitrosylation: a Short-Range Mechanism for NO Signaling? *Antioxid. Redox Signal.* **2013**, *19*, 1220–1235.
- [163] Foster, M. W.; McMahon, T. J.; Stamler, J. S. S-nitrosylation in Health and Disease. *Trends Mol. Med.* **2003**, *9*, 160–168.
- [164] Camerini, S.; Polci, M. L.; Restuccia, U.; Usuelli, V.; Malgaroli, A.; Bachi, A. A Novel Approach to Identify Proteins Modified by Nitric Oxide: the HIS-TAG Switch Method. *J. Proteome Res.* **2007**, *6*, 3224–3231.
- [165] Shao, S.; Chen, B.; Cheng, J.; Wang, C.; Zhang, Y.; Shao, L.; Hu, Y.; Han, Y.; Han, F.; Li, X. A Fluorogenic Probe for Imaging Protein S-Nitrosylation in Live Cells. *Biosens. Bioelectron.* **2017**, *94*, 162–168.
- [166] Paulsen, C. E.; Carroll, K. S. Cysteine-Mediated Redox Signaling: Chemistry, Biology, and Tools for Discovery. *Chem. Rev.* **2013**, *113*, 4633–4679.
- [167] Devarie-Baez, N. O.; Zhang, D.; Li, S.; Whorton, A. R.; Xian, M. Direct Methods for Detection of Protein S-Nitrosylation. *Methods* **2013**, *62*, 171–176.
- [168] Forrester, M. T.; Foster, M. W.; Benhar, M.; Stamler, J. S. Detection of Protein S-nitrosylation with the Biotin-Switch Technique. *Free Radic. Biol. Med.* **2009**, *46*, 119–126.
- [169] Koziol, K. L.; Johnson, P. J.; Stucki-Buchli, B.; Waldauer, S. A.; Hamm, P. Fast Infrared Spectroscopy of Protein Dynamics: Advancing Sensitivity and Selectivity. *Curr. Opin. Struc. Biol.* **2015**, *34*, 1–6.
- [170] Hamm, P.; Zanni, M. *Concepts and Methods of 2D Infrared Spectroscopy*; Cambridge University Press, Cambridge, 2011.

- [171] Walsh, G. M.; Leane, D.; Moran, N.; Keyes, T. E.; Forster, R. J.; Kenny, D.; O'Neill, S. S-Nitrosylation of Platelet  $\alpha\text{IIb}\beta\text{3}$  as Revealed by Raman Spectroscopy. *Biochem.* **2007**, *46*, 6429–6436.
- [172] Coyle, C. M.; Vogel, K. M.; Rush III, T. S.; and R. Williams, P. M. K.; Spiro, T. G.; Dou, Y.; Ikeda-Saito, M.; Olson, J. S.; Zgierski, M. Z. FeNO Structure in Distal Pocket Mutants of Myoglobin Based on Resonance Raman Spectroscopy. *Biochem.* **2003**, *42*, 4896—4903.
- [173] Manning, M. C. Use of Infrared Spectroscopy to Monitor Protein Structure and Stability. *Expert Rev. Proteomic* **2005**, *2*, 731–743.
- [174] Kramer, R. M.; Shende, V. R.; Motl, N.; Pace, C. N.; Scholtz, J. M. Toward a Molecular Understanding of Protein Solubility: Increased Negative Surface Charge Correlates with Increased Solubility. *Biophys. J* **2012**, *102*, 1907–1915.
- [175] Mueller, R.; Huber, J. R. Two Rotational Isomers of Methyl Thionitrite: Light-Induced, Reversible Isomerization in an Argon Matrix. *J. Phys. Chem.* **1984**, *88*, 1605–1608.
- [176] Arulsamy, N.; Bohle, D.; Butt, J.; Irvine, G.; Jordan, P.; Sagan, E. Interrelationships Between Conformational Dynamics and the Redox Chemistry of S-Nitrosothiols. *J. Am. Chem. Soc.* **1999**, *121*, 7115–7123.
- [177] Canneva, A.; Della Vedova, C. O.; Mitzel, N. W.; Erben, M. F. Conformational Properties of Ethyl- and 2, 2, 2-trifluoroethyl Thionitrites,  $(\text{CX}_3\text{CH}_2\text{SNO}, \text{X} = \text{H} \text{ and } \text{F})$ . *J. Phys. Chem. A* **2015**, *119*, 1524–1533.
- [178] Gregori, B.; Guidoni, L.; Chiavarino, B.; Scuderi, D.; Nicol, E.; Frison, G.; Fornarini, S.; Crestoni, M. E. Vibrational Signatures of S-

- Nitrosoglutathione as Gaseous, Protonated Species. *J. Phys. Chem. B* **2014**, *118*, 12371–12382.
- [179] Bartberger, M. D.; Houk, K.; Powell, S. C.; Mannion, J. D.; Lo, K. Y.; Stamler, J. S.; Toone, E. J. Theory, Spectroscopy, and Crystallographic Analysis of S-Nitrosothiols: Conformational Distribution Dictates Spectroscopic Behavior. *J. Am. Chem. Soc.* **2000**, *122*, 5889–5890.
- [180] Romeo, A. A.; Filosa, A.; Capobianco, J. A.; English, A. M. Metal Chelators Inhibit S-nitrosation of Cys $\beta$ 93 in Oxyhemoglobin. *J. Am. Chem. Soc.* **2001**, *123*, 1782–1783.
- [181] Hess, D. T.; Matsumoto, A.; Nudelman, R.; Stamler, J. S. S-nitrosylation: Spectrum and Specificity. *Nat. Cell Biol.* **2001**, *3*, E46–E48.
- [182] Khomyakov, D. G.; Timerghazin, Q. K. Toward Reliable Modeling of S-Nitrosothiol Chemistry: Structure and Properties of Methyl Thionitrite (CH<sub>3</sub>SNO), an S-Nitrosocysteine Model. *J. Chem. Phys.* **2017**, *147*, 044305.
- [183] Timerghazin, Q. K.; Peslherbe, G. H.; English, A. M. Structure and Stability of HSNO, the Simplest S-Nitrosothiol. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1532–1539.
- [184] Bignon, E.; Allegra, M. F.; Lucchetta, M.; Tiberti, M.; Papaleo, E. Computational Structural Biology of S-Nitrosylation of Cancer Targets. *Front. Oncol.* **2018**, *8*, 272.
- [185] Melvin, A. C.; Jones, W. M.; Lutzke, A.; Allison, C. L.; Reynolds, M. M. S-Nitrosoglutathione Exhibits Greater Stability than S-nitroso-N-acetylpenicillamine Under Common Laboratory Conditions: A Comparative Stability Study. *Nitric Oxide* **2019**, *92*, 18–25.



- [186] Lanucara, F.; Chiavarino, B.; Crestoni, M. E.; Scuderi, D.; Sinha, R. K.; Maître, P.; Fornarini, S. S-Nitrosation of Cysteine as Evidenced by IRMPD Spectroscopy. *Int. J. Mass Spectr.* **2012**, *330*, 160–167.
- [187] Coletti, C.; Re, N.; Scuderi, D.; Maître, P.; Chiavarino, B.; Fornarini, S.; Lanucara, F.; Sinha, R. K.; Crestoni, M. E. IRMPD Spectroscopy of Protonated S-Nitrosocaptopril, a Biologically Active, Synthetic Amino Acid. *Phys. Chem. Chem. Phys.* **2010**, *12*, 13455–13467.
- [188] El Hage, K.; Brickel, S.; Hermelin, S.; Gaulier, G.; Schmidt, C.; Bonacina, L.; van Keulen, S. C.; Bhattacharyya, S.; Chergui, M.; Hamm, P.; Rothlisberger, U.; Wolf, J.-P.; Meuwly, M. Implications of Short Time Scale Dynamics on Long Time Processes. *Struct. Dyn.* **2017**, *4*, 061507.
- [189] Han, S. Force field Parameters for S-Nitrosocysteine and Molecular Dynamics Simulations of S-Nitrosated Thioredoxin. *Biochem. Biophys. Res. Commun.* **2008**, *377*, 612–616.
- [190] Petrov, D.; Margreitter, C.; Grandits, M.; Oostenbrink, C.; Zagrovic, B. A Systematic Framework for Molecular Dynamics Simulations of Protein Post-Translational Modifications. *PLoS Comput. Biol.* **2013**, *9*, e1003154.
- [191] Huang, J.; MacKerell Jr, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*, 2135–2145.
- [192] Hairer, E.; Lubich, C.; Wanner, G. Geometric Numerical Integration Illustrated by the Störmer–Verlet Method. *Acta Numer.* **2003**, *12*, 399–450.
- [193] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

- [194] Steinbach, P. J.; Brooks, B. R. New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular Simulation. *J. Comput. Chem.* **1994**, *15*, 667–683.
- [195] Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An  $N \log(N)$  Method for EwaldSums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- [196] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- [197] Schreiter, E. R.; Rodríguez, M. M.; Weichsel, A.; Montfort, W. R.; Bonaventura, J. S-Nitrosylation-Induced Conformational Change in Blackfin Tuna Myoglobin. *J. Biol. Chem.* **2007**, *282*, 19773–19780.
- [198] Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 Energy Evaluation by Direct Methods. *Chem. Phys. Lett.* **1988**, *153*, 503–506.
- [199] Kendall, R. A.; Dunning Jr, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- [200] Kramer, C.; Gedeck, P.; Meuwly, M. Atomic Multipoles: Electrostatic Potential Fit, Local Reference Axis Systems, and Conformational Dependence. *J. Comput. Chem.* **2012**, *33*, 1673–1688.
- [201] Berau, T.; Kramer, C.; Monnard, F. W.; Nogueira, E. S.; Ward, T. R.; Meuwly, M. Scoring Multipole Electrostatics in Condensed-Phase Atomistic Simulations. *J. Phys. Chem. B* **2013**, *117*, 5460–5471.
- [202] Lammers, S.; Meuwly, M. Investigating the Relationship Between Infrared Spectra of Shared Protons in Different Chemical Environments: A Compar-

- ison of Protonated Diglyme and Protonated Water Dimer. *J. Phys. Chem. A* **2007**, *111*, 1638–1647.
- [203] Carugo, O. How Large B-factors Can be in Protein Crystal Structures. *BMC Bioinform.* **2018**, *19*, 61–69.
- [204] Sun, Z.; Liu, Q.; Qu, G.; Feng, Y.; Reetz, M. T. Utility of B-factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering thermostability. *Chem. Rev.* **2019**, *119*, 1626–1665.
- [205] van de Weert, M.; Haris, P. I.; Hennink, W. E.; Crommelin, D. J. Fourier Transform Infrared Spectrometric Analysis of Protein Conformation: Effect of Sampling Method and Stress Factors. *Anal. Biochem.* **2001**, *297*, 160–169.
- [206] Nutt, D.; Meuwly, M. Theoretical Investigation of Infrared Spectra and Pocket Dynamics of Photodissociated Carbonmonoxy Myoglobin. *Biophys. J.* **2003**, *85*, 3612–3623.
- [207] Devereux, M.; Plattner, N.; Meuwly, M. Application of Multipolar Charge Models and Molecular Dynamics Simulations to Study Stark Shifts in Inhomogeneous Electric Fields. *J. Phys. Chem. A* **2009**, *113*, 13199–13209.
- [208] Plattner, N.; Meuwly, M. Higher Order Multipole Moments for Molecular Dynamics Simulations. *J. Mol. Model.* **2009**, *15*, 687–694.
- [209] Plattner, N.; Meuwly, M. Atomistic Simulations of CO Vibrations in Ices Relevant to Astrochemistry. *ChemPhysChem* **2008**, *9*, 1271–1277.
- [210] Plattner, N.; Meuwly, M. The Role of Higher CO-Multipole Moments in Understanding the Dynamics of Photodissociated Carbonmonoxide in Myoglobin. *Biophys. J.* **2008**, *94*, 2505–2515.

- [211] Schiro, G.; Fichou, Y.; Gallat, F.-X.; Wood, K.; Gabel, F.; Moulin, M.; Haertlein, M.; Heyden, M.; Colletier, J.-P.; Orecchini, A.; Paciaroni, A.; Wuttke, J.; Tobias, D. J.; Weik, M. Translational Diffusion of Hydration Water Correlates with Functional Motions in Folded and Intrinsically Disordered Proteins. *Nat. Chem.* **2015**, *6*, 1–8.
- [212] Pocker, Y. Water in Enzyme Reactions: Biophysical Aspects of Hydration-Dehydration Processes. *Cell. Mol. Life Sci.* **2000**, *57*, 1008–1017.
- [213] Pal, S.; Zewail, A. Dynamics of Water in Biological Recognition. *Chem. Rev.* **2004**, *104*, 2099–2123.
- [214] Strazza, S.; Hunter, R.; Walker, E.; Darnall, D. W. The Thermodynamics of Bovine and Porcine Insulin and Proinsulin Association Determined by Concentration Difference spectroscopy. *Arch. Biochem. Biophys.* **1985**, *238*, 30–42.
- [215] Desmond, J. L.; Koner, D.; Meuwly, M. Probing the Differential Dynamics of the Monomeric and Dimeric Insulin from Amide-I IR Spectroscopy. *J. Phys. Chem. B* **2019**, *123*, 6588–6598.
- [216] Raghunathan, S.; El Hage, K.; Desmond, J. L.; Zhang, L.; Meuwly, M. The Role of Water in the Stability of Wild-type and Mutant Insulin Dimers. *J. Phys. Chem. B* **2018**, *122*, 7038–7048.
- [217] Zoete, V.; Meuwly, M.; Karplus, M. Study of the Insulin Dimerization: Binding Free Energy Calculations and Per-Residue Free Energy Decomposition. *Proteins* **2005**, *61*, 79–93.
- [218] Zoete, V.; Meuwly, M.; Karplus, M. A Comparison of the Dynamic Behavior of Monomeric and Dimeric Insulin Shows Structural Rearrangements in the Active Monomer. *J. Mol. Biol.* **2004**, *342*, 913–929.

- [219] Weichsel, A.; Brailey, J. L.; Montfort, W. R. Buried S-Nitrosocysteine Revealed in Crystal Structures of Human Thioredoxin. *Biochem.* **2007**, *46*, 1219–1227.
- [220] Baldwin, E. T.; Bhat, T. N.; Gulnik, S.; Liu, B.; Topol, I. A.; Kiso, Y.; Mimoto, T.; Mitsuya, H.; Erickson, J. W. Structure of HIV-1 protease with KNI-272, a Tight-Binding Transition-State Analog Containing Allophenyl-norstatine. *Structure* **1995**, *3*, 581–590.
- [221] Prashar, V.; Bihani, S.; Das, A.; Ferrer, J.-L.; Hosur, M. Catalytic Water Co-Existing with a Product Peptide in the Active Site of HIV-1 Protease Revealed by X-ray Structure Analysis. *PloS one* **2009**, *4*, e7860.
- [222] Dawson, T. M.; Dawson, V. L. REVIEW: Nitric Oxide: Actions and Pathological Roles. *Neuroscientist* **1995**, *1*, 7–18.
- [223] Simanshu, D. K.; Nissley, D. V.; McCormick, F. RAS Proteins and Their Regulators in Human Disease. *Cell* **2017**, *170*, 17–33.
- [224] Colicelli, J. Human RAS Superfamily Proteins and Related GTPases. *Sci. Signal.* **2004**, *2004*, re13–re13.
- [225] Reck, M.; Carbone, D.; Garassino, M.; Barlesi, F. Targeting KRAS in Non-Small-Cell Lung Cancer: Recent Progress and New Approaches. *Ann. Oncol.* **2021**, *32*, 1101–1110.
- [226] Prior, I. A.; Hood, F. E.; Hartley, J. L. The Frequency of Ras Mutations in Cancer. *Cancer Res.* **2020**, *80*, 2969–2974.
- [227] Pantsar, T. The Current Understanding of KRAS Protein Structure and Dynamics. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 189–198.
- [228] Mo, S. P.; Coulson, J. M.; Prior, I. A. RAS Variant Signalling. *Biochem. Soc. Trans.* **2018**, *46*, 1325.

- [229] Bos, J. L.; Rehmann, H.; Wittinghofer, A. GEFs and GAPs: Critical Elements in the Control of Small G Proteins. *Cell* **2007**, *129*, 865–877.
- [230] Hunter, J. C.; Gurbani, D.; Ficarro, S. B.; Carrasco, M. A.; Lim, S. M.; Choi, H. G.; Xie, T.; Marto, J. A.; Chen, Z.; Gray, N. S., et al. In Situ Selectivity Profiling and Crystal Structure of SML-8-73-1, an Active Site Inhibitor of Oncogenic K-Ras G12C. *PNAS* **2014**, *111*, 8895–8900.
- [231] Hancock, J. F.; Paterson, H.; Marshall, C. J. A Polybasic Domain or Palmitoylation is Required in Addition to the CAAX Motif to Localize p21ras to the Plasma Membrane. *Cell* **1990**, *63*, 133–139.
- [232] Merz, V.; Gaule, M.; Zecchetto, C.; Cavaliere, A.; Casalino, S.; Pesoni, C.; Contarelli, S.; Sabbadini, F.; Bertolini, M.; Mangiameli, D., et al. Targeting KRAS: the Elephant in the Room of Epithelial Cancers. *Front. Oncol.* **2021**, *11*, 361.
- [233] Ostrem, J. M.; Peters, U.; Sos, M. L.; Wells, J. A.; Shokat, K. M. K-Ras (G12C) Inhibitors Allosterically Control GTP Affinity and Effector Interactions. *Nature* **2013**, *503*, 548–551.
- [234] Shima, F.; Matsumoto, S.; Yoshikawa, Y.; Kawamura, T.; Isa, M.; Kataoka, T. Current Status of the Development of Ras Inhibitors. *J. Biochem.* **2015**, *158*, 91–99.
- [235] Biancucci, M.; Minasov, G.; Banerjee, A.; Herrera, A.; Woida, P. J.; Kiefer, M. B.; Bindu, L.; Abreu-Blanco, M.; Anderson, W. F.; Gaponenko, V., et al. The Bacterial Ras/Rap1 Site-Specific Endopeptidase RRSP Cleaves Ras Through an Atypical Mechanism to Disrupt Ras-ERK Signaling. *Sci. Signal.* **2018**, *11*, eaat8335.
- [236] Ntai, I.; Fornelli, L.; DeHart, C. J.; Hutton, J. E.; Doubleday, P. F.; LeDuc, R. D.; van Nispen, A. J.; Fellers, R. T.; Whiteley, G.; Boja, E. S.,

- et al. Precise Characterization of KRAS4b Proteoforms in Human Colorectal Cells and Tumors Reveals Mutation/Modification Cross-Talk. *PNAS* **2018**, *115*, 4140–4145.
- [237] Raines, K. W.; Cao, G.-L.; Lee, E. K.; Rosen, G. M.; Shapiro, P. Neuronal Nitric Oxide Synthase-Induced S-Nitrosylation of H-Ras Inhibits Calcium Ionophore-Mediated Extracellular-Signal-Regulated Kinase Activity. *Biochem. J.* **2006**, *397*, 329–336.
- [238] Raines, K. W.; Bonini, M. G.; Campbell, S. L. Nitric Oxide Cell Signaling: S-Nitrosation of Ras Superfamily GTPases. *Cardiovasc. Res.* **2007**, *75*, 229–239.
- [239] Lim, K.-H.; Ancrile, B. B.; Kashatus, D. F.; Counter, C. M. Tumour Maintenance is Mediated by eNOS. *Nature* **2008**, *452*, 646–649.
- [240] Turan, H. T.; Meuwly, M. Spectroscopy, Dynamics, and Hydration of S-nitrosylated Myoglobin. *J. Phys. Chem. B* **2021**, *125*, 4262–4273.
- [241] Ichiye, T.; Karplus, M. Collective Motions in Proteins: a Covariance Analysis of Atomic Fluctuations in Molecular Dynamics and Normal Mode Simulations. *Proteins* **1991**, *11*, 205–217.
- [242] Arnold, G. E.; Ornstein, R. L. Molecular Dynamics Study of Time-Correlated Protein Domain Motions and Molecular Flexibility: Cytochrome P450BM-3. *Biophys. J.* **1997**, *73*, 1147–1159.
- [243] Grant, B. J.; Rodrigues, A. P.; ElSawy, K. M.; McCammon, J. A.; Caves, L. S. Bio3d: an R Package For the Comparative Analysis of Protein Structures. *Bioinformatics* **2006**, *22*, 2695–2696.





# Appendix A

## MS-ARMD parameters for the Menshutkin reactions

	Reactant				Product			
Bond Harmonic	$K_b$	$d_{eq}$			$K_b$	$d_{eq}$		
N – H <sub>NH<sub>3</sub></sub>	437.150	1.016			490.522	1.039		
C – H <sub>Me</sub>	375.145	1.100			369.812	1.151		
Bond Morse	$D_e$	$d_{eq}$	$\beta$		$D_e$	$d_{eq}$	$\beta$	
C – Cl	145.720	1.785	0.582		X	X	X	
C – N	X	X	X		47.9544	1.520	1.682	
Angle	$K_\Theta$	$\Theta_{eq}$			$K_\Theta$	$\Theta_{eq}$		
H <sub>NH<sub>3</sub></sub> – N – H <sub>NH<sub>3</sub></sub>	40.594	107.500			44.274	108.806		
H <sub>Me</sub> – C – H <sub>Me</sub>	37.170	110.010			21.686	114.437		
H <sub>Me</sub> – C – Cl	45.335	112.040			X	X		
H <sub>NH<sub>3</sub></sub> – N – C	X	X			57.420	107.189		
H <sub>NH<sub>3</sub></sub> – C – H <sub>Me</sub>	X	X			25.311	105.421		
GVDW	$r$	$\epsilon$	$n$	$m$	$r$	$\epsilon$	$n$	$m$
N–C	1.209	0.115	6.232	12.323	X	X	X	X
C–Cl	X	X	X	X	2.500	0.300	5.518	12.196

**Table A.1:** The harmonic bond, Morse bond, angle and generalised van der Waals (GVDW) parameters for for NH<sub>3</sub>+MeCl.  $K_b$  in kcal/mol/Å<sup>2</sup>,  $d_{eq}$  in Å,  $D_e$  in kcal/mol,  $d_{eq}$  in Å,  $\beta$  in Å<sup>-1</sup>,  $K_\theta$  in kcal/mol/radian<sup>2</sup>,  $r$  in Å and  $\epsilon$  in kcal/mol.

## Appendix A. MS-ARMD parameters for the Menshutkin reactions

Bond Harmonic	Reactant				Product			
	$K_b$	$d_{eq}$			$K_b$	$d_{eq}$		
$C_{Pyr} - C_{Pyr}$	389.240	1.369			420.800	1.393		
$C_{Pyr} - N_{Pyr}$	446.700	1.319			464.880	1.345		
$C_{Pyr} - H_{Pyr}$	352.560	1.083			246.970	1.078		
$C_{Me} - H_{Me}$	293.370	1.083			226.040	1.074		
Bond Morse	$D_e$	$d_{eq}$	$\beta$		$D_e$	$d_{eq}$	$\beta$	
C - Br	87.908	1.972	0.957		X	X	X	
C - N	X	X	X		286.170	1.581	0.908	
Angle	$K_\Theta$	$\Theta_{eq}$			$K_\Theta$	$\Theta_{eq}$		
$C_{Pyr} - C_{Pyr} - H_{Pyr}$	29.657	122.100			27.082	120.800		
$C_{Pyr} - C_{Pyr} - N_{Pyr}$	39.676	126.390			37.792	121.650		
$N_{Pyr} - C_{Pyr} - H_{Pyr}$	43.001	118.010			52.771	115.520		
$C_{Pyr} - N_{Pyr} - C_{Pyr}$	81.961	112.800			96.761	119.040		
$H_{Me} - C_{Me} - H_{Me}$	30.810	104.220			32.760	115.190		
$H_{Me} - C_{Me} - Br$	40.252	102.410			X	X		
$N_{Pyr} - C_{Me} - H_{Me}$	X	X			48.194	104.190		
$C_{Pyr} - N_{Pyr} - H_{Me}$	X	X			40.800	121.640		
GVDW	$r$	$\epsilon$	$n$	$m$	$r$	$\epsilon$	$n$	$m$
N-C	1.296	0.751	8.096	16.710	X	X	X	X
C-Br	X	X	X	X	0.491	0.927	7.773	11.510

**Table A.2:** The harmonic bond, Morse bond, angle and generalised van der Waals (GVDW) parameters for for Pyr+MeBr.  $K_b$  in kcal/mol/ $\text{\AA}^2$ ,  $d_{eq}$  in  $\text{\AA}$ ,  $D_e$  in kcal/mol,  $d_{eq}$  in  $\text{\AA}$ ,  $\beta$  in  $\text{\AA}^{-1}$ ,  $K_\theta$  in kcal/mol/radian<sup>2</sup>,  $r$  in  $\text{\AA}$  and  $\epsilon$  in kcal/mol.

	Reactant			Product		
	Mulliken	NBO	FF	Mulliken	NBO	FF
C	-0.08	-0.42	-0.42	-0.15	-0.28	-0.25
C's H	0.13	0.35	0.35	0.18	0.41	0.28
N	-0.39	-1.04	-1.04	0.07	-0.69	-0.30
N's H	0.05	0.18	0.18	0.13	0.22	0.21
Cl	-0.09	-0.12	-0.12	-0.86	-0.94	-0.94

**Table A.3:** The Mulliken, NBO and fitted reactive force field charges (starting from the NBO charges) for the  $NH_3 + MeCl$  (in  $e$ ).

	Reactant			Product		
	Mulliken	NBO	FF	Mulliken	NBO	FF
N	-0.36	-0.53	-0.59	0.02	-0.38	-0.38
Pyr-C	0.13	0.12	0.16	-0.20	0.19	-0.03
Pyr-C	-0.16	-0.26	-0.15	-0.08	-0.24	-0.03
Pyr-C	0.00	-0.12	-0.15	-0.21	-0.06	-0.03
Pyr-H	0.11	0.17	0.15	0.14	0.22	0.22
Pyr-H	0.10	0.20	0.15	0.15	0.22	0.22
Pyr-H	0.10	0.20	0.15	0.13	0.22	0.22
C	-0.07	-0.51	-0.59	-0.12	-0.28	-0.28
C's H	0.09	0.19	0.16	0.12	0.22	0.22
Br	-0.18	-0.06	-0.15	-0.90	-0.96	-0.96

**Table A.4:** The Mulliken, NBO and fitted reactive force field charges (starting from the NBO charges) for the  $\text{NH}_3 + \text{MeCl}$  (in  $e$ ).

$k$	$V_{ij,k}^0$	$\sigma_{ij,k}$	$\alpha_{ij,k0}$
1	0.48578	22.15304	-17.47731
2	40.88660	12.45761	-1.34837
3	-36.67239	11.02057	-1.014567

**Table A.5:** GAPO parameters for  $\text{NH}_3 + \text{MeCl}$ :  $i$  labels the reactant,  $j$  labels the product,  $V_{ij,k}^0$  is the center of each of the  $k = 3$  Gaussian functions (in kcal/mol),  $\sigma_{ij,k}$  is the width of each Gaussian (in kcal/mol) and  $\alpha_{ij,k0}$  is the polynomial coefficient. In the present case polynomial order “0” was sufficient, hence  $\alpha_{ij,k0}$ .

$k$	$V_{ij,k}^0$	$\sigma_{ij,k}$	$\alpha_{ij,k0}$
1	-44.46610	12.24875	-0.747557
2	1.17231	28.18915	-22.93224
3	55.29621	23.35329	-3.89254

**Table A.6:** GAPO parameters for  $\text{Pyr} + \text{MeBr}$ :  $i$  labels the reactant,  $j$  labels the product,  $V_{ij,k}^0$  is the center of the Gaussian function (in kcal/mol), and  $\sigma_{ij,k}$  the width of the Gaussian (in kcal/mol).  $\alpha_{ij,k0}$  is the polynomial coefficient in kcal/mol. In the present case polynomial order “0” was sufficient, hence  $\alpha_{ij,k0}$ .



# Appendix B

## Curriculum Vitae

### Personal Information

**Name:** Haydar Taylan Turan

**Address:** Hebelstrasse 56, 4056 Basel, Switzerland

**E-mail:** haydartaylan.turan@unibas.ch

**Date of Birth:** 27.11.1990, Istanbul, Turkey

### Education

**Ph.D. in Chemistry**, University of Basel, 01/2018 – 08/2022

- **Thesis:** Solvent, Structure, and Spectroscopy: the Role of Dynamics in Chemical Reactions and S-Nitrosylated Proteins
- **Supervisor:** Prof. Markus Meuwly

**MSc. in Chemistry**, Bogazici University, 01/2015 – 06/2017

- **Thesis:** A computational approach to optical and charge carrier properties of organic semiconductors
- **Supervisor:** Prof. Viktorya Aviyente

**BSc. in Chemistry**, Ege University, 09/2009 – 06/2014

- **Thesis:** Investigation of Interactions between Serotonin and Caffeine Molecules by Molecular Modelling
- **Supervisor:** Prof. Nursel Acar

## Publications

Turan, H. T.; Eken, Y.; Marazzi, M.; Pastore, M.; Aviyente, V.; Monari A. “Assessing one-and two-photon optical properties of boron containing arenes”. *J. Phys. Chem. C*, **2016**, *120*, 17916-17926

Turan, H. T.; Yavuz, I.; Aviyente, V. “Understanding the impact of thiophene/furan substitution on intrinsic charge-carrier mobility”. *J. Phys. Chem. C*, **2017**, *121*, 25682-25690

Turan, H. T.; Kucur, O.; Kahraman. B.; Salman, S.; Aviyente, V. “Design of donor–acceptor copolymers for organic photovoltaic materials: a computational study”. *PCCP*, **2018**, *20*, 3581-3591

Brickel. S.,; Das, A. K.; Unke, O. T.; Turan, H. T.; Meuwly. M. ”Reactive molecular dynamics for the  $[\text{Cl-CH}_3\text{-Br}]^-$  reaction in the gas phase and in solution: a comparative study using empirical and neural network force fields”. *Electronic Structure*, **2019**, *1*, 024002

Kucur, O.; Turan, H. T.; Monari A.; Aviyente, V. “Computational Study of Photo-oxidative Degradation Mechanisms of Boron-Containing Oligothiophenes”. *J. Phys. Chem. A*, **2020**, *124*, 1390-1398

Rivero, U.; Turan, H. T.; Meuwly, M.; Willitsch, S.; “Reactive atomistic simulations of Diels-Alder-type reactions: conformational and dynamic effects in the polar cycloaddition of 2,3-dibromobutadiene radical ions with maleic anhydride”.

*Mol. Phys.*, **2021**, *119*, e1825852

Turan, H. T.; Meuwly, M. “Spectroscopy, Dynamics, and Hydration of S-Nitrosylated Myoglobin”. *J. Phys. Chem. B*, **2021**, *125*, 4262-4273

Turan, H. T.; Meuwly, M. “Solvent Effects on the Menshutkin Reaction”.

*J. Phys. Chem. B*, **2022**, *126*, 1951-1961

Turan, H. T.; Meuwly, M. “Local Hydration Control and Functional Implications Through S-Nitrosylation of Proteins: Kirsten Rat Sarcoma Virus (K-RAS) and Hemoglobin (Hb)”. *J. Phys. Chem. B*, **2023**, *127*, 1526-1539

Turan, H. T.; Boittier, E.; Meuwly, M. “Interaction at a distance: Xenon migration in Mb”. *J. Chem. Phys.*, **2023**, *158*, 125103