The dissertation of Yi Liu was reviewed and approved* by the following:

Paul S. Cremer
Professor of Chemistry
Professor of Biochemistry and Molecular Biology
Dissertation Adviser
Chair of Committee

Nicholas Winograd
Evan Pugh Professor of Chemistry

John Asbury
Associate Professor of Chemistry

Kurt Gibble
Professor of Physics

Barbara J. Garrison
Shapiro Professor of Chemistry
Head of the Department of Chemistry

*Signatures are on file in the Graduate School.
ABSTRACT

Vibrational Sum Frequency Spectroscopy of Osmolytes and Ions at Aqueous Interfaces.

(December 2014)

Yi Liu, B.S., Peking University, Beijing, China

Chair of Advisory Committee: Dr. Paul S. Cremer

Surfaces/interfaces are omnipresent in nature, ranging from physics and chemistry to biology and material sciences. To characterize the interfacial structures and understand the surface phenomena, many useful tools have been developed. Vibrational sum frequency generation (VSFG) spectroscopy has proven to be a powerful second-order non-linear optical technique for this purpose and is gaining more and more attention nowadays. With VSFG, two pulsed laser beams, one at infrared frequency and one at visible frequency, are incident on the surface and generate a new beam with a frequency equal to the sum of the IR and visible frequencies. When the IR frequency matches a surface vibrational mode frequency, this process would be resonantly enhanced. In this dissertation, this surface-sensitive technique was adopted to investigate several interfaces with special relevance to biology.

The first topic of interest is the unusual orientation of a strong protein stabilizer, trimethylamine N-oxide (TMAO), at two aqueous/hydrophobic interfaces (air/water interface and OTS/water interface). By interpretation of the relative phase of VSFG spectra coupled with a numerical algorithm, the maximum entropy method (MEM) analysis of the molecular orientation, it is found that the methyl groups of TMAO prefer to point into the aqueous medium, while the oxide moieties (N\(^+\)-O\(^-\)) orient towards the
This unusual orientation may be attributed to the more hydrophilic nature of methyl groups that is attached to a strong electron withdrawing atom such as a quaternary nitrogen. These results could help elucidate the stabilizing effect of TMAO on proteins the increased need to keep the methyl group hydrated would cause them to be excluded from protein interface and thereby lead to protein stabilization.

The other major issue focused in this dissertation is based on the ion specific interactions at a charged interface, which plays a decisive role in various physicochemical and biological processes. Binding affinity of different cations to monolayers of amphiphilic molecules (e.g. fatty acids, phospholipids) at the air/aqueous surfactant interfaces, may provide molecular level clues on various functions of cell membranes that are resembled by these amphiphilic molecules. Specifically, the binding events of several alkali cations to the hydrophilic carboxylate headgroups of long chain fatty acid, inferred from interfacial water structures, are thoroughly investigated by VSFG measurement. Results show that Li$^+$ binds strongest to the negatively charged carboxylate groups, followed by Na$^+$, then K$^+$ although the difference is slight. The ranking of the alkali metal cations’ binding abilities differs from the sequence predicted by the law of match water affinities (LMWA) and also varies with different headgroups in the model system, which may suggest the distinct solvation behaviors of these ions. Such findings should help to elucidate the molecular-level binding behavior to proteins in aqueous solutions.
# TABLE OF CONTENTS

List of Figures .................................................................................................................. vi
List of Abbreviations ......................................................................................................... ix
List of Symbols .................................................................................................................. xi
Acknowledgements .......................................................................................................... xiii

Chapter 1 Introduction ...................................................................................................... 1

Chapter 2 Sum Frequency Generation: Principles and Instrumentation ......................... 8

2.1 Mathematical and Physical Background of SFG ......................................................... 9
2.2 Retrieval of Phase Information from VSFG Intensity Spectra .................................. 26
    2.2.1 Phase-sensitive measurement ...................................................................... 30
    2.2.2 Maximum entropy method ......................................................................... 36
2.3 VSFG Instrumentation ............................................................................................. 40

Chapter 3 The Study of Trimethylamine N-Oxide Orientation at Hydrophobic Interfaces. 48

3.1 Background .................................................................................................................. 48
3.2 Experimental Section ............................................................................................... 52
3.3 Results ........................................................................................................................ 55
3.4 Discussion .................................................................................................................... 67

Chapter 4 Study of Ionic Binding of Some Alkali Metal Cations to the
    Carboxylate/Phosphonate Headgroups of Surfactant Monolayers ................................ 70

4.1 Background .................................................................................................................. 70
4.2 Experimental Section ............................................................................................... 73
4.3 Results and Discussion ............................................................................................ 75
4.4 Conclusion ................................................................................................................... 100

Chapter 5 Implications and Conclusions ......................................................................... 102

Reference ........................................................................................................................... 105

Appendix A The Matlab Scripts Used for VSFG Spectra Fitting ..................................... 114

Appendix B The Matlab Script Used for MEM Analysis ................................................ 125
LIST OF FIGURES

Figure 1-1. Infrared-visible sum frequency generation spectroscopy. The insert spectrum shown is the O-H stretch region (3000 ~ 3800 cm\(^{-1}\)) at neat water/air interface. ..........7

Figure 2-1. A planar surface and laboratory coordinate system adopted. The z axis is along the surface normal and the projections of the incoming and outgoing lights on the surface are travelling in the positive x direction. Here all three electromagnetic fields of the laser beams are oscillating within the plane of incidence (x-z plane) as a special case for the polarization combination. .........................................................18

Figure 2-2. Energy diagram of vibrational sum frequency generation (VSFG). Different line widths between the energy levels suggest the possibilities of the corresponding transition. \(|g|, |v|\) and \(|s|\) refer to the ground state, the excited vibrational state and any other state (a virtual state), respectively. The physical expression and explanation of the transitions have been given in equation (24). .........................................................25

Figure 2-3. (a) The schematic of an HD-VSFG optical configuration. (b) Raw spectra of the sample (blue) and z-cut quartz crystal (red). (c) Time domain interferogram obtained by the inverse Fourier transformation of raw data and the filter function (green, plotted to the right vertical axis). (d) Real (solid lines) and imaginary (dashed lines) parts of the heterodyne spectra by the Fourier transformation of the time-domain interferogram after filtration of the sample and reference. This figure is reproduced from Ref. 57. ..............................................................32

Figure 2-4. The experimental set up of the first VSFG spectrometer. The sum frequency signal is generated by temporal and spatial overlapping of input infrared and visible beams at the sample surface. ..................................................................................43

Figure 2-5. The experimental set up of the second VSFG spectrometer purchased from EKSPLA. ..................................................................................................................45

Figure 2-6. The Langmuir trough used as the sample holder for air/water interface experiments. ......................................................................................................................47

Figure 3-1. The experimental setup containing a Teflon flow cell with a quartz window on top and bottom. The top quartz window contains an OTS monolayer on the bottom face. The sum frequency generation spectrum is obtained by passing the visible and infrared beams through the top quartz window and selecting for the bottom face...............................51

Figure 3-2. (a) VSFG spectra for TMAO of various concentrations at the air/water interface. (b) The oscillator strength of the OH stretch of water (3700 cm\(^{-1}\) peak) and the symmetric CH\(_3\) stretches of TMAO (2750 cm\(^{-1}\) peak) vs TMAO concentration. Although the dangling OH bonds of water decrease with TMAO concentration in a linear fashion, the peak is still evident even at 5 M TMAO. ..........56

Figure 3-3. The fitting of VSFG data for 4 M TMAO at the air/water interface. The peak assigned to the symmetric stretches of the TMAO methyl group at ~2950 cm\(^{-1}\) is fit
with a negative sign in a) and positive sign in b). It is clear that the fit shown in a) is much better, indicating that the methyl groups of TMAO most likely point away from the hydrophobic interface with air.

Figure 3-4. MEM calculations of the imaginary portion of $\chi^{(2)}$ as a function of wavelength at the air/water interface in the presence and absence of 4 M TMAO. These calculations are in agreement with recent phase sensitive VSFS measurements at the air/water interface. In addition, the large negative amplitude of $\chi^{(2)}$ at ~2950 cm$^{-1}$ (symmetric stretch of the methyl group) indicates the methyl groups of TMAO point away from the air/hydrophobic interface.

Figure 3-5. VSFG spectra of the OTS/D$_2$O interface, taken in a quartz flow cell. The spectra shown in a) compare OTS in D$_2$O with OTS in 6 M TMAO. Since it is known that the methyl groups of OTS point down into solution, the constructive interference with the TMAO methyl group peak indicates this group is also pointing down. The two spectra overlaid in b) are OTS in D$_2$O and OTS in 6 M deuterated TMAO. The OTS monolayer appears to be relatively undisturbed by the presence of 6 M TMAO. Last, the spectra shown in c) indicate that no peak is observed when both OTS and TMAO are deuterated and a peak similar to that at air/water interface is observed for 6 M hydrogenated TMAO in perdeuterated OTS.

Figure 3-6. VSFG data inssp polarization for 4 M TMAO at the OTS/D$_2$O interface. The peak assigned to the symmetric stretches of the TMAO methyl group at ~2950 cm$^{-1}$ is fit with a negative sign in a) and positive sign in b). Once again, the fit corresponding to the methyl groups pointing away from the hydrophobic interface (shown in a)) is preferred.

Figure 3-7. MEM calculations of the imaginary portion of $\chi^{(2)}$ in the presence and absence of 6 M TMAO at the OTS/water interface. The largest difference is observed ~2950 cm$^{-1}$, where the peak due to the symmetric stretch of TMAO is observed. This peak from TMAO overlaps with the Fermi resonance of OTS (~2934 cm$^{-1}$). The two peaks have the same sign and lead to an increase in overall magnitude of Im[$\chi^{(2)}$].

Figure 4-1. VSFG spectra in ssp polarization combination of air/eicosanoic acid (C$_{19}$H$_{39}$COOH) monolayer/pure water interface (black trace) and air/eicosanoic acid monolayer/50 mM NaCl salt solution interface (red trace). Both spectra were taken at bulk pH 11 to make the acid headgroups fully deprotonated.

Figure 4-2. Cartoon depiction of the ionic binding events. The long chain fatty acid can form a good Langmuir monolayer over the aqueous subphase due to its amphiphilic property. Closely packed (condensed) phase for the surfactants was adopted and maintained during the whole time. The red and blue dots represent O and H atoms in water molecules and acid headgroups, while the yellow dots stand for monovalent ions (Li$^+$, Na$^+$ and K$^+$).

Figure 4-3. The comparison between the VSFG intensities of air/pure water interface and air/fatty acid monolayer/water interface. The green dots are the air/pure water
spectrum magnified by a factor of ten for clearer view about the relative low intensity...80

Figure 4-4. The VSFG spectra at the C-H stretching region for pure salt solutions (100 mM NaCl and 100 mM KCl are selected as examples). The nearly flat C-H intensities indicate no organic contamination existed in those solution. ....................82

Figure 4-5. VSFG experiments of ecosanoic acid (C_{20}) monolayer with different subphases (ultra-pure water and alkali metal chloride solutions). The salts solution subphases are all of the same concentration to ensure the identical ionic strength and available binding species, which help to elucidate the binding affinity to the charged headgroups. ........................................84

Figure 4-6. Gradual quenching effect of the signal of the O-H stretching region as the electrolyte concentration increases. 50 mM, 100mM, 200mM and 500 mM of KCl salt solutions were used as the subphases. All other experimental conditions were identical.................................................................86

Figure 4-7. VSFG spectra of ecosanoic acid (C_{20}) monolayer spread over ultra-pure deionized water (black squares) and with the presence of 100 mM alkali metal halide salts (LiCl, NaCl and KCl) under pH 4. At the acidic pH, the acid headgroups are fully protonated and the similar intensities in the O-H regions suggest no appreciable binding may occur to the charge neutral -COOH groups of the acid.......................89

Figure 4-8. Divalent ion effect. Very low concentrations of calcium ions were introduced and found to be extremely effective in quenching O-H signals........................................91

Figure 4-9. Control experiment of a solution containing 100 mM NaCl and 0.1 mM CaCl$_2$. Both inorganic salts are with highest purity available (99.999%). ......................93

Figure 4-10. VSFG experiments of an octadecylphosphonate monolayer with different subphases (ultra-pure water and alkali metal chloride solutions) at two concentrations. Both spectra reveal the unidentifiable binding strength of Na$^+$ vs K$^+$ ...95
LIST OF ABBREVIATIONS

IR           infrared
vis          visible
UV           ultra violet
SF           sum frequency
VSFG         vibrational sum frequency generation
PS-VSFG      phase-sensitive vibrational sum frequency generation
MEM          Maximum entropy method
BAM          Brewster angle microscopy
G            gas phase
LE           liquid expand phase
LC           liquid condensed phase
MMA          mean molecular area
TMAO         trimethylamine N-oxide
PA           palmitic acid
SA           stearic acid
EA           ecosanoic acid
DPPA         1,2-Dipalmitoyl-sn-glycero-3-phosphate sodium salt
DPPC         1,2-Dipalmitoyl-sn-glycero-3-phosphocholine
DPPE         1,2-Dipalmitoyl-sn-glycero-3-phosphoethanolamine
DPPG         1,2-Dipalmitoyl-sn-glycero-3-phosphoglycerol sodium salt
DPPS         1,2-Dipalmitoyl-sn-glycero-3-phospho-L-serine sodium salt
A.U.         arbitrary units
as           asymmetric stretch
ss  symmetric stretch
FR  Fermi resonance
SNR signal-to-noise ratio
fs  femtosecond
ps  picosecond
cm  centimeter
µm micrometer
nm  nanometer
cm\(^{-1}\) wavenumber/inverted centimeter
CW  continuous wave
CIP contact ion pair
SIP solvent shared/separated ion pair
FL  focal length
fwhm full width at half maximum
m.f. mole fraction
MD  molecular dynamics
LIST OF SYMBOLS

$\alpha$ linear polarizability

$\beta$ second-order hyperpolarizability

$\varepsilon$ absolute permittivity

$\varepsilon_0$ vacuum permittivity (electric constant)

$\varepsilon_r$ relative permittivity (dielectric constant)

$\theta_c$ critical angle

$\theta_i$ angle of incidence

$\theta_r$ angle of reflection

$\theta_t$ angle of refraction

$\lambda$ wavelength

$\mu$ dipole moment

$\nu$ frequency

$\phi$ phase angle

$\chi$ susceptibility

$\chi^{(1)}$ first-order (linear) susceptibility

$\chi^{(2)}$ second-order susceptibility

$\chi_{\text{eff}}$ effective susceptibility

$\omega$ angular frequency

$c$ speed of light in vacuum

e elementary charge

$E$ electric field

$E_i$ electric field of the incident wave field

$E_r$ electric field of the reflected wave field
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_t$</td>
<td>electric field of the transmitted wave field</td>
</tr>
<tr>
<td>$I$</td>
<td>intensity of light</td>
</tr>
<tr>
<td>$k$</td>
<td>angular wavenumber (the magnitude of the wave vector)</td>
</tr>
<tr>
<td>$k$</td>
<td>wave vector</td>
</tr>
<tr>
<td>$k_B$</td>
<td>Boltzmann constant</td>
</tr>
<tr>
<td>$k_i$</td>
<td>wave vector of the incident wave</td>
</tr>
<tr>
<td>$k_r$</td>
<td>wave vector of the reflected wave</td>
</tr>
<tr>
<td>$k_t$</td>
<td>wave vector of the transmitted wave</td>
</tr>
<tr>
<td>$K_a$</td>
<td>acid disassociation constant</td>
</tr>
<tr>
<td>$m$</td>
<td>moles of solute per kilogram of water (mol/kg)</td>
</tr>
<tr>
<td>$M$</td>
<td>moles per liter (mol/L)</td>
</tr>
<tr>
<td>$\Xi$</td>
<td>complex refractive index</td>
</tr>
<tr>
<td>$n_i$</td>
<td>refractive index of the incident medium</td>
</tr>
<tr>
<td>$n_t$</td>
<td>refractive index of the transmitted medium</td>
</tr>
<tr>
<td>$P$</td>
<td>polarization (dipole moment per unit volume)</td>
</tr>
<tr>
<td>$q_j$</td>
<td>$j_{th}$ vibrational normal mode</td>
</tr>
<tr>
<td>$t$</td>
<td>time</td>
</tr>
<tr>
<td>$T$</td>
<td>temperature</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Dr. Paul S. Cremer for his support, guidance and all the insights during my Ph. D. studies. It is he who leaded me into the SFG world and help me make every little step forward.

I also extend my appreciation to Dr. John Asbury, Dr. Nicholas Winograd and Dr. Kurt Gibble for their time and graciously serving on my committee.

I wish I can express my sincere gratitude to my former advisor, Dr. Wayne Goodman in Texas A&M University who instructed and inspired me at the onset of this unforgettable doctorate journey.

My acknowledgement is also devoted to Dr. Tinglu Yang for his great advice and countless help with the experimental data analysis and interpretation. Thanks also go to my former colleagues Dr. Sarah Flores and Dr. Jaibir Kherb for their guidance and willing to share their knowledge unconditionally, as well as my lab mates Da Huang, Kelvin Rembert, Halil Okur, Matthew Poyton, Xiao Cong, Yi-Ting Liao, Seung-yi Lee, William Hunn, Anne Sendecki, Alexis Baxter, Simou Sun, Saranya Pullanchery and so many more for their constant help, good advices and informative discussions through my studies.

Last of all, I am very grateful to my parents, to whom I owe the most of love. I would not be who I am today without their tremendous spiritual care, protection and support throughout these years.
Chapter 1

Introduction

The interface is described as a thin boundary that separates two different bulk media. It plays very important roles in many physical, chemical, and biological processes of the world due to its unique and interesting properties that are rather different from the bulk media.\textsuperscript{1,2} “God made the bulk; the surface was invented by the devil.”\textsuperscript{3} This famous sentence expressed by Nobel Laureate Wolfgang Pauli reveal the diabolical behavior of the interfaces. Considering the breaking of symmetry at the boundary, the interfacial molecules do not experience the same isotropic environment as the molecules do in the bulk. Therefore, the interface atoms encounter the asymmetry in forces at the interfacial region. One well-known example is the surface tension, caused by the unsymmetrical cohesive forces among liquid molecules of the surface layer.

Investigation of the interfacial phenomena, ranging from molecules adsorption to heterogenetic catalysis reactions, of course requires the comprehensive understanding of the chemical structures at this specific region. The key and the challenge in studying the complex physics and chemistry of interfaces is to just probe the topmost layers (a few angstroms) of the material surfaces. It is of no surprise that many pioneering research have been extensively done on this subject. The discipline “Surface Science”, established in the last century, introduces many modern analytical techniques and provided lots of useful information on single crystal surfaces, thin films, nano-structures, etc.\textsuperscript{2} Some of the representative techniques are atomic force microscopy (AFM),\textsuperscript{4,5} scanning tunneling microscopy (STM),\textsuperscript{5-7} transmission electron microscopy (TEM),\textsuperscript{8,9} low energy electron
diffraction (LEED),\textsuperscript{10,11} neutron reflection or neutron scattering,\textsuperscript{12} X-ray diffraction (XRD),\textsuperscript{13} X-Ray photoelectron spectroscopy (XPS),\textsuperscript{14} and so on. Thanks to these powerful experimental methods as well as theoretical molecular dynamics simulation studies, our understanding of surfaces has grown in number and in sophistication radically. Nonetheless, many of these surface science techniques are quite invasive and require samples in ultrahigh vacuum (UHV) conditions and hence limited in application or not suitable for some cases. For instance, few of these studies involve air/liquid and liquid/liquid interfaces because UHV conditions are hardly achievable. Therefore there is a huge need to develop novel and applicable methods to investigate those interfacial structures.

It is hence natural to turn to optical techniques that may help solve these problems. As a matter of fact, time-resolved total internal reflection fluorescence,\textsuperscript{15-18} attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR)\textsuperscript{19,20} and surface enhanced Raman scattering (SERS)\textsuperscript{21,22} are intensively used to draw information from the interfaces noninvasively. Still, these techniques have some drawbacks. For example, the surface sensitivity of ATR-FTIR is determined/limited by the penetration depth of the IR light (typically on the order of hundreds of nanometers to a few micrometers),\textsuperscript{23} which may go beyond the surface region. SERS usually requires a metal substrate to enhance the signal and become inapplicable to many interfaces.

There also exists another branch of optics called non-linear optics (NLO), describing the behavior of intense light in non-linear media (the media where the induced dielectric polarization responds non-linearly to the electric field of the light).\textsuperscript{24} The basic principles were theoretically proposed or predicted dating back to the early days of
quantum mechanics, according to the solutions to Maxwell’s equations in nonlinear dielectrics. However, it only become reality until the invention of strong light source such as pulsed lasers in 1950s. The first non-linear phenomenon people found experimentally is second harmonic generation (SHG) by P. A. Franken et al at University of Michigan in 1961. Surface SHG was then formulated by Bloembergen and Pershan at Harvard University in 1962. Its application, no longer limited to metal surfaces, vastly broadened the surface science studies. With the advancement of ultra-short laser sources, light conversion and detection techniques, another promising non-linear probe sum frequency generation (SFG) was also developed. The first SFG spectrum was based on a dye Coumarin monolayer on glass, reported by Yuen-Ron Shen and co-workers in Department of Physics, University of California at Berkeley in 1986. Since then, these powerful NLO techniques became an effective and versatile tool for all kinds of surfaces (including the solid/vacuum, solid/liquid, liquid/air and solid/solid interfaces), owing to their unique advantages:

a) Noninvasive to the samples surface (non-destructive)

b) *In-situ* probe

c) Applicable to all interfaces accessible by light

d) High and inherent surface specificity

e) Output signal is monochromatic and coherent that can be easily discerned from the input and background by spatial and spectral filtering

f) High resolution

The most attracting prospect of SHG and SFG is the inherent surface specificity. Due to the quantum selection rules, second-harmonic and sum frequency process can
only take place at regions where the inversion symmetry is broken, therefore it is inherently forbidden at bulk centrosymmetric phases, but very sensitive to study the molecular structures at the material surface/interfaces between two bulk phases (e.g. air/water interface in this study). This property will be discussed further in the next chapter.

In our group, we adopted the infrared-visible vibrational sum frequency generation spectroscopy to investigate many surface structures and phenomena at molecular level. In practice, one intense visible (typically at a fixed frequency $\omega_{\text{vis}}$) laser beam and another infrared (tunable in the mid-IR range, $\omega_{\text{IR}}$) laser beam are overlapped spatially and temporally at the interface to be studied. In the absence of inversion symmetry, one of the possible nonlinear processes involves the producing a new beam at the sum frequency:

$$\omega_{\text{SF}} = \omega_{\text{vis}} + \omega_{\text{IR}}.$$  

By measuring the SFG signal as a function of the infrared frequency, a vibrational spectrum of chemical moieties present at the interfacial region can be obtained (c. f. Figure 1-1).

In Chapter 2 I will briefly describe the fundamental physical principles of vibrational sum frequency generation as well as the instruments used in this work.

There is an ongoing argument about the molecular mechanism of stabilization/denaturation of proteins. Finding how the protein stabilizer or denaturant interact with protein functional groups is crucial in answering this question. In Chapter 3 we report a study on the absolute orientation of a powerful protein stabilizer, trimethylamine N-oxide (TMAO), at two aqueous/hydrophobic interfaces by means of
conventional VSFG spectroscopy coupled with a numerical algorithm, the maximum entropy method (MEM) calculation of the molecular orientation. Surprisingly, the findings reveal that the methyl groups of TMAO prefer to point into the aqueous phase, while the oxide moieties (N\(^+\)-O\(^-\)) orient towards the hydrophobic medium. This unusual orientation may be due to the more hydrophilic nature of methyl groups which are attached to a strongly electron withdrawing atom such as a quaternary nitrogen. These results may give us some hint on the stabilizing effect of TMAO on proteins as the orientation of the oxygen towards a hydrophobic protein surface may decrease hydrophobic solvation and thereby increase the propensity for hydrophobic collapse.

Understanding ion interactions with charged surfaces is of great interest to many researchers in the hope of solving important problems in biological fields. For example, the widely studied Langmuir films, formed by long-chain molecules (fatty acids, phospholipids with various headgroups etc.) spreading over an aqueous subphase, are often used to (partly) resemble cell membranes. Ion specific effects with these charged surfaces can be ordered in the so-called Hofmeister series\(^{29,30}\) and may account for many biological relevant phenomena such as protein solubility, enzyme activity, protein folding and many others.\(^{31-36}\) By analyzing the changes in the intensity of the VSFG spectra at the air/surfactant/liquid interface, we could obtain valuable information on the extent of interactions between the cations and negatively charged surfaces formed by surfactant monolayer (and vice versa). Specifically, the binding events of alkali cations to the hydrophilic carboxylate headgroups of long chain fatty acid, inferred from interfacial water structure, are thoroughly investigated. These studies are reported in Chapter 4 and 5 of this dissertation. We found the ranking of the alkali metal cations’ binding abilities
differs from the sequence predicted by the law of match water affinities (LMWA)\textsuperscript{37-39} and also varies with different headgroups in the model system, which may suggest the distinct solvation behaviors of these ions.
Figure 1-1. Infrared-visible sum frequency generation spectroscopy. The insert spectrum shown is the O-H stretch region (3000 ~ 3800 cm$^{-1}$) at neat water/air interface.
Chapter 2

Sum Frequency Generation: Principles and Instrumentation

In this chapter, a brief instruction on the mathematical and physical background of sum frequency generation will be given to address the physical picture of this unique nonlinear technique more clearly. This part is followed by a discussion on the difficulties of the interpretation of the SFG intensity spectra. To solve this problem, two novel methods, namely phase-sensitive (PS) measurement and maximum entropy method (MEM) are proposed and used in the experimental and data analysis process. The next section is our SFG instrument setup, coupled with some background information on the laser and non-linear optical processes associated with the light-conversion technique.
2.1 Mathematical and Physical Background of SFG

The following part of the physical origins and mathematical deduction of SFG follows the manuscript of Colin Bain and David J. Neiandt.

To understand the basic principles of SFG, we must first clarify the interaction of light (electromagnetic radiation) with matter. Generally speaking, when a molecule is placed in a weak electromagnetic field (e.g. a non-laser or weak laser light source in UV/vis or Raman spectroscopy), the induced dipole moment, \( \mu \), is directly proportional to the strength of that electric field, \( E \). In other words, the electron distribution in the molecule responds harmonically to the incident electric field. The dipole moment of the molecule can be written as

\[
\mu = \mu^0 + \alpha E
\]

where \( \mu^0 \) is the static dipole moment (or called permanent dipole moment) and \( \alpha \) is the (linear) polarizability of that molecule. Since the induced dipole moment is usually not in the same direction as the incident electric field, \( \alpha \) is a \((3 \times 3)\) tensor.

In condensed phases, we also would like to introduce another physical property to describe the interaction of the electric field and materials: the dipole moment per unit volume, or bulk polarization, \( P \):

\[
P = P^0 + \varepsilon_0 \chi^{(1)} E
\]

where \( P^0 \) represent static polarization. We can see now the macroscopic linear (or first-order) susceptibility \( \chi^{(1)} \) replaces the microscopic linear polarizability in Equation (1). Since very few materials (and none of the materials/systems studied in this thesis) have a
static polarization, we can set $P^{(0)}$ as 0 for simplicity from now on. The relationship between linear susceptibility and polarizability thus is easy to clarify:

$$\chi^{(1)} = \frac{N}{\varepsilon_0} \langle \alpha \rangle$$  \hspace{1cm} (3)

where $N$ represents the number of molecules per unit volume, i.e. molecular number density and the angular brackets $\langle \alpha \rangle$ denotes linear polarizability averaged over all the molecular orientations in the material.

Now we will move to the physical property of the incident light. The most relevant one is the complex refractive index, $\mathbb{m}$. It characterizes the linear properties of isotropic materials and follows

$$\mathbb{m}^2 = \chi^{(1)} + 1$$  \hspace{1cm} (4)

The real part of $\mathbb{m}$ is the factor by which the speed of light $v$ is reduced as it traverses the medium:

$$v = \frac{c}{\operatorname{Re}[\mathbb{m}]}$$  \hspace{1cm} (5)

while the imaginary part of $\mathbb{m}$ determines the absorption coefficient, which is coincidentally also denoted as $\alpha$ for historical reasons:

$$\alpha = \frac{4\pi}{\lambda} \operatorname{Im}[\mathbb{m}]$$  \hspace{1cm} (6)

How could we better understand this complex refractive index? Let’s consider the simple transmission infrared spectroscopy as an example. Infrared spectra usually plot the absorbance ($A$) or the transmittance ($T$) versus wavenumbers. Absorbance can be expressed as Equation (7) for a given sample thickness $l$, in accord with the Lambert-Beer Law:
\[ A = -\log_{10} \frac{I}{I_0} = \alpha l \]  

(7)

where \( I \) is the intensity of the beam after passing through the sample, \( I_0 \) is the intensity of the incident light, and \( \alpha \) is the absorption coefficient which is related to the extinction coefficient \( \kappa \) by

\[ \alpha = \frac{4\pi\kappa}{\lambda} \]  

(8)

Here \( \kappa \) is just the imaginary part of the complex refractive index \( n \) and represents the damping of an electromagnetic wave inside the material. We may now rewrite the complex refractive index as

\[ n = n + i\kappa = \sqrt{1 + \chi^{(1)}} \]  

(9)

Both \( n \) and \( k \) are functions of wavelength \( \lambda \).

For now we merely focus on a weak electric filed and its implication on matter. When referring to a strong electric field induced by an intense light (such as pulsed laser, usually \( 10^{10} - 10^{12} \) V/m), the electric field strength is so high that the electrons are no longer able to respond harmonically. Higher order terms must be included to describe the dipole moment:

\[ \mu = \mu^0 + \alpha E + \beta : EE + \gamma : EEE + ... \]  

(10)

where \( \beta \) and \( \gamma \) are known as the first and second hyperpolarizabilities and

\[ \beta : EE = \sum_{j,k} \beta_{ijk} E_j E_k \]  

(11)

The corresponding polarization of a material, \( P \), is expressed as

\[ P = P^{(1)} + P^{(2)} + P^{(3)} + ... \]

\[ = \varepsilon_0 \chi^{(1)} E + \varepsilon_0 \chi^{(2)} : EE + \varepsilon_0 \chi^{(3)} : EEE + ... \]  

(12)
where $\chi^{(2)}$ is a third-rank tensor, known as the second-order (non-linear) susceptibility, and $\chi^{(3)}$ is a fourth-rank tensor known as the third-order (non-linear) susceptibility. Optical processes that arise from $\chi^{(2)}$ and $\chi^{(3)}$ are known as second-order and third-order non-linear effects, respectively. Again, non-linear effects only become significant and observable when the applied electromagnetic field is comparable with the field experienced by the electrons in a molecule, which are only normally achievable with pulsed lasers.

The most striking aspect of these non-linear effects might be the generation of outgoing light with new frequencies. Let’s first consider a single input laser beam whose electric field is oscillating as time $t$ and can be described as a cosine wave function:

$$E(r, t) = E(r) \cos \omega t$$  \hspace{1cm} (13)

If we only care about the second-order effect, we have

$$P^{(2)} = \varepsilon_0 \chi^{(2)}:EE = \varepsilon_0 \chi^{(2)}:E(r)E(r)\cos^2 \omega t$$  \hspace{1cm} (14)

Since $\cos 2\varphi = 2\cos^2 \varphi - 1$, we have

$$P^{(2)} = \frac{1}{2} \varepsilon_0 \chi^{(2)}:E(r)E(r)(1 + \cos 2\omega t)$$  \hspace{1cm} (15)

As the second term in the bracket of this equation indicates, a new light at frequency $2\omega$ (twice the frequency of the incident light) will be generated. This process is called second harmonic generation (SHG).

If we let two input laser beams ($E_1$ and $E_2$) with different frequencies ($\omega_1$ and $\omega_2$) interact simultaneously at the sample media, the expression for second order polarization $P^{(2)}$ becomes:

$$P^{(2)} = \varepsilon_0 \chi^{(2)}:E_1E_2 = \varepsilon_0 \chi^{(2)}:E_1(r)E_2(r)\cos \omega_1 t \cos \omega_2 t$$  \hspace{1cm} (16)
According to trigonometric function, \(2\cos\alpha \cdot \cos\beta = \cos(\alpha + \beta) \cdot \cos(\alpha - \beta)\), so

\[
P^{(2)} = \frac{1}{2} \varepsilon_0 \chi^{(2)}(r) E_1(r) E_2(r) [\cos(\omega_1 + \omega_2) t + \cos(\omega_1 - \omega_2) t]
\] (17)

Now the two terms in the brackets represents two new oscillating fields: the first one with frequency \((\omega_1 + \omega_2)\) and the second with \((\omega_1 - \omega_2)\). This would give rise to the sum frequency generation (SFG) and difference frequency generation (DFG), the former of which together with SHG are widely used in surface chemistry study nowadays (In fact, we can treat SHG as a special case for SFG: two laser beams with the same frequency \(\omega\) overlapped would generate the new light of frequency \(2\omega\) in SF process). Though the simple classical electromagnetic approach adopted here is sufficient to demonstrate the physical origins of these second-order nonlinear processes, a rigorous and comprehensive derivation is only obtainable by quantum mechanical calculations, which can be found in the textbooks written by Y. R. Shen\textsuperscript{24} or Boyd.\textsuperscript{42}

From above discussion we know that SFG and DFG are two coherent processes, how could we separate them if we only want to detect one of them? The solution is rather simple. Because the direction of emission beam is well defined by the conservation of momentum parallel to the surface (known as the phase-matching condition), the output SF beam and DF beam with different frequencies/energies will be propagating in distinct directions.

For SFG conducted in the co-propagating, reflection geometry as illustrated in Fig. 1-1,

\[
k_{SF} \sin\theta_{SF} = k_{vis} \sin\theta_{vis} + k_{IR} \sin\theta_{IR}
\] (18)
where $\theta_{\text{vis}}$, $\theta_{\text{IR}}$ and $\theta_{\text{SF}}$ are the incident angles of the visible and IR laser beams, respectively and $\theta_{\text{SF}}$ is the angle of emission of the SF light. $k$ is the wave vector and its amplitude is determined by

$$k = \frac{2\pi}{\lambda} = \frac{\omega}{c}$$

Actually the most precise form for the wave vector in a dielectric medium is

$$k = \frac{2\pi n}{\lambda} = \frac{\omega n}{c}$$

Since the three beams are propagating in the same medium under current condition (only detecting the reflected SF beam rather than the refracted one because it is more accessible at most interfaces), we often omit $n$ and the equation has such a form:

$$\omega_{\text{SF}} \sin \theta_{\text{SF}} = \omega_{\text{vis}} \sin \theta_{\text{vis}} + \omega_{\text{IR}} \sin \theta_{\text{IR}}$$

(21)

Note if we scan IR through a spectrum (varying $\omega_{\text{IR}}$), $\theta_{\text{SF}}$ will gradually changes. Since the IR frequency range used in our study (2800 ~ 3800 cm$^{-1}$) is way smaller than the visible beam (532.1 nm, ~ 18793 cm$^{-1}$), the angle of the SF light $\theta_{\text{SF}}$ is mainly dependent on the incident visible beam with fixed frequency. Therefore when we only scan a short range of IR, the change of $\theta_{\text{SF}}$ is small enough that we do not need to adjust the SF light beam path to the detector for recording the spectrum.

Next we will elaborate the intensity and selection rule of SFG. The intensity of the emitted SF light depends on the absolute square of $P^{(2)}$, therefore proportional to $|\chi^{(2)}|^2$: 

$$I_{\text{SF}} \propto |P^{(2)}|^2 \propto |\chi^{(2)}|^2 \cdot I_{\text{vis}} \cdot I_{\text{IR}}$$

(22)
The second-order nonlinear optical susceptibility \( \chi^{(2)} \) is a macroscopic property of the sample. Like the linear (first-order) susceptibility \( \chi^{(1)} \), it is also the orientational average of all the molecular second order hyperpolarizabilities \( \beta \) in the material:

\[
\chi^{(2)} = \frac{N}{\varepsilon_0} \langle \beta \rangle
\]  

(23)

The expression for \( \beta \) can be derived from second-order perturbation theory as an infinite sum over the quantum states of the system. This would give a really sophisticated general solution, but can be simplified if neither \( \omega_{SF} \) nor \( \omega_{vis} \) is in resonance with an electronic transition of the material. Near a vibrational transition, \( \omega_0 \), of a molecule, the hyperpolarizability, \( \beta \), can be stated

\[
\beta_{lmn} = -\frac{1}{2\hbar} \sum_s \left\{ \langle g|\mu_l|s\rangle \langle s|\mu_m|v\rangle - \langle g|\mu_m|s\rangle \langle s|\mu_l|v\rangle \right\} \otimes \left\{ \frac{\langle v|\mu_n|g\rangle}{\omega_{IR} - \omega_0 + i\Gamma} \right\}
\]  

(24)

in which \( |g\rangle \) is the ground state, \( |v\rangle \) the excited vibrational state, \( |s\rangle \) any other state, \( \Gamma^{-1} \) the relaxation time of the excited vibrational state, and \( \mu = qr \) is the electric dipole operator.

Although the above expression for hyperpolarizability \( \beta \) still looks complex, it has a simple chemical interpretation. The first term in brackets can be recognized as the Raman transition dipole moment \( M_{lm} \). \( \langle v|\mu_n|g\rangle \) is also easy to be identified as the IR transition dipole moment \( A_n \) between the ground state \( |g\rangle \) and excited vibrational state \( |v\rangle \). Thus

\[
\beta_{lmn} \propto \frac{M_{lm}A_n}{\omega_{IR} - \omega_0 + i\Gamma}
\]  

(25)
From above statement we can see naturally that if a molecular vibration is sum frequency active, it has to be both infrared active and Raman active. The gross selection rule for IR requires the change of electric dipole moment of the molecule must not be zero if that vibrational normal mode is infrared active. Raman selection rule, on the other hand, asks for the change of polarizability during the vibration (i.e. a molecule must have anisotropic polarizability to be Raman active). These requirements could help determine whether a certain vibrational modes are SF active in many cases and lead to some simple criteria. For example, homonuclear diatomic molecules (H$_2$, O$_2$ etc.) are infrared inactive because their dipole moments remain zero during the stretching vibration. Thus these molecules are SF inactive. Moreover, according to the rule of mutual exclusion, in a centro-symmetric molecule, a vibrational mode may be either IR active or Raman active but not both when the atoms are displaced relative to each other. In the case of CO$_2$, we could find that those modes which are Raman active are IR inactive and vice versa. Therefore such molecules with a center of inversion symmetry are SF inactive, too.

If we view this SF inactive situation of centrosymmetric materials in a macroscopic aspect, we could find that its intrinsic symmetry prevents the SF process by purely electric dipole mechanisms. As we mentioned earlier, $\chi^{(2)}$ is a macroscopic property of the sample medium, representing the orientational average of the second order hyperpolarizability $\beta$ of all the molecules in the material. As a third-rank polar tensor, $\chi_{ijk}^{(2)}$ changes sign under the inversion operation (equivalent to the reversing of axis system):

$$\chi_{-i-j-k}^{(2)} = (-1)^3 \chi_{ijk}^{(2)} = -\chi_{ijk}^{(2)}$$

(26)
Centrosymmetric material is invariant under the inversion operation for two opposing directions since all directions are equivalent:

\[ \chi^{(2)}_{-i-j-k} = \chi^{(2)}_{ijk} \]  

(27)

To satisfy both (26) and (27), there’s only one solution:

\[ \chi^{(2)}_{ijk} = -\chi^{(2)}_{ijk} = 0 \]  

(28)

That is to say, materials or systems with inversion symmetry (such as the bulk gases, liquids or most solids) cannot give rise to SF signal due to the second-order susceptibility being zero under the electric dipole approximation. However, if we only consider the interfacial area of such systems, we can see at the interfaces (no matter solid/liquid, gas/liquid, liquid/liquid or other interfaces), the centrosymmetry is naturally broken along the surface normal and this second-order nonlinear optical phenomenon can occur. That is the intrinsic source of the surface specificity mentioned previously. Owing to this, SFG spectroscopy becomes very popular and is vastly applied in the study of surface/interface during recent decades.

Moreover, the planar surfaces itself are considered isotropic so we can treat the surface with a \( C_\infty \) symmetry along its surface normal as depicted in Figure 2-1. The surface Cartesian axis system \( (x, y, z) \) we adopted is following the convention: the surface normal is always the \( z \) axis, the surface forms \( x-y \) plane and the input and output beams are travelling in the \( x-z \) plane.
Figure 2-1. A planar surface and laboratory coordinate system adopted. The $z$ axis is along the surface normal and the projections of the incoming and outgoing lights on the surface are travelling in the positive $x$ direction. Here all three electromagnetic fields of the laser beams are oscillating within the plane of incidence ($x$-$z$ plane) as a special case for the polarization combination.
We continue to focus on the second-order term:

\[ P^{(2)}_i = \varepsilon_0 \chi^{(2)}_{ijk} E_j E_k \]  

(29)

It seems rather difficult to analyze since every direction may yield the second-order response. However, some of the 27 elements of the third rank tensor \( \chi^{(2)}_{ijk} \) must vanish due to symmetry constraint. As can be seen, with \( C_\infty \) symmetry about the surface normal, \( x \equiv -x \), \( y \equiv -y \) but \( z \neq -z \). If we reverse the \( x \) or \( y \) axis, the non-vanishing \( \chi^{(2)}_{ijk} \) on this surface should not change sign because no change has actually taken place. Nonetheless, the tensor rule still holds true: if the direction of any individual axis is reversed, the directionally dependent value of \( \chi^{(2)}_{ijk} \) must change sign.

For example, let’s consider \( \chi^{(2)}_{xxx} \). By reversing the \( x \) axis, it produces \( -x - x - x \).

Applying the tensor rule aforementioned, we have

\[ \chi^{(2)}_{-x-x-x-x} \equiv -\chi^{(2)}_{x-x-x} \equiv \chi^{(2)}_{xx-x} \equiv -\chi^{(2)}_{xxx} \]  

(30)

The sign of \( \chi^{(2)}_{xxx} \) is also reversed. This result breaks the symmetry constraint unless \( \chi^{(2)}_{xxx} \) equals zero (a non-contributing element).

If we treat \( \chi^{(2)}_{xxz} \) under similar operation, we may find

\[ \chi^{(2)}_{-x-x} \equiv -\chi^{(2)}_{x-x} \equiv \chi^{(2)}_{xxx} \]  

(31)

There is no overall change in the sign of \( \chi^{(2)}_{xxx} \) if we reverse the \( x \) axis, hence it could be a non-zero, contributing element.

For another example, \( \chi^{(2)}_{zzz} \) holds the exact the same form with reversal of the \( x \) or \( y \) axis, and it also contributes.
In summary, with these two restrictions, we may find only seven elements of $\chi_{ijk}^{(2)}$ out of the total 27 can be non-zero. They are

$\chi_{xzx}^{(2)}, \chi_{xxz}^{(2)}, \chi_{xyy}^{(2)}, \chi_{yxy}^{(2)}, \chi_{zzx}^{(2)}, \chi_{zxy}^{(2)}, \text{and } \chi_{zzz}^{(2)}.$

Moreover, the $x$ or $y$ axis are equivalent or interchangeable for the isotropic surface ($x$-$y$ plane). In other words, we can choose the propagation direction of the light either $x$ or $y$ without any actual change of the physical picture. As a consequence, there are only four independent non-vanishing $\chi_{ijk}^{(2)}$ element that can generate SF signals at the surface:

$\chi_{xzx}^{(2)} \equiv \chi_{yzy}^{(2)}$

$\chi_{xxz}^{(2)} \equiv \chi_{yyz}^{(2)}$

$\chi_{zxx}^{(2)} \equiv \chi_{zyy}^{(2)}$

$\chi_{zzz}^{(2)}$

These different tensor components of $\chi_{ijk}^{(2)}$ can be sampled by utilizing different laser polarizations combinations. As in the usual case, $p$ polarized incident light (electromagnetic wave) oscillates parallel to (or within) the plane of incidence (in this case, $x$-$z$ plane, the associated electric vector $E_p$ can be decomposed into two components, $E_x$ and $E_z$), while $s$ polarized light has an electric field perpendicular to that plane (hence $E_s$ is oscillating along the $y$ direction, $E_s = E_y$).

Note SFG is a three-wave mixing process and we have $s$ or $p$ control for each of the beams, so there should exist eight ($2^3 = 8$) different polarizations combinations. But again, due to the symmetry constraint mentioned above, some combinations do not
generate SFG signal. In fact of all the combinations, only four specific ones are commonly utilized, namely ssp, sps, pss and ppp modes, where the first letter refers to the output VSFG beam, the second letter for the input visible beam, and the last letter for the input infrared beam (these letters are arranged in the descend sequence of the beam frequencies by convention). Amongst these four combinations, ssp is the most widely used one. With the s-polarized visible and p-polarized IR beams, only \( \chi_{yyz}^{(2)} \) contributes and gives rise to s-polarized sum frequency emission. The relationship between the effective susceptibilities of each mode and the more intrinsic, actual non-linear susceptibility \( \chi_{ijk}^{(2)} \) can be calculated easily as follows:

\[
\begin{align*}
\chi_{ssp}^{(2)} &= L_{yy}(\omega_{SF})L_{yy}(\omega_{vis})L_{zz}(\omega_{IR}) \sin \theta_{IR} \chi_{yyz}^{(2)} \\
\chi_{sps}^{(2)} &= L_{yy}(\omega_{SF})L_{zz}(\omega_{vis})L_{yy}(\omega_{IR}) \sin \theta_{vis} \chi_{yyz}^{(2)} \\
\chi_{pss}^{(2)} &= L_{zz}(\omega_{SF})L_{yy}(\omega_{vis})L_{yy}(\omega_{IR}) \sin \theta_{SF} \chi_{yyz}^{(2)} \\
\chi_{ppp}^{(2)} &= L_{zz}(\omega_{SF})L_{xx}(\omega_{vis})L_{xx}(\omega_{IR}) \sin \theta_{SF} \cos \theta_{vis} \cos \theta_{IR} \chi_{xxx}^{(2)} \\
&\quad - L_{xx}(\omega_{SF})L_{xx}(\omega_{vis})L_{zz}(\omega_{IR}) \cos \theta_{SF} \cos \theta_{vis} \sin \theta_{IR} \chi_{xxz}^{(2)} \\
&\quad + L_{zz}(\omega_{SF})L_{zz}(\omega_{vis})L_{zz}(\omega_{IR}) \sin \theta_{SF} \sin \theta_{vis} \sin \theta_{IR} \chi_{zzz}^{(2)} \\
&\quad - L_{xx}(\omega_{SF})L_{zz}(\omega_{vis})L_{xx}(\omega_{IR}) \cos \theta_{SF} \sin \theta_{vis} \cos \theta_{IR} \chi_{xzx}^{(2)}
\end{align*}
\]

where \( L_{ii} \) corresponds to the non-linear Fresnel factor associated with each beam, and \( \theta_i \) denotes the angle of the associated beam versus surface normal. The non-linear Fresnel
factor acts as a local-field correction factor of each electric field when the beam propagates across the boundary that separates the interface and the adjacent isotropic medium. The Fresnel factors could be written as functions of the refractive indices of the beam in different media, the incident angle, the angle of reflection and the angle of refraction. A detailed derivation of all the tensorial Fresnel factors can be found in the work presented by Y. R. Shen and omitted here. Therefore, the Fresnel factors may affect VSFG intensity and even cause spectral line shapes distortion, which should be taken into consideration in spectral normalization and data processing. For the simplicity, most VSFG studies at the air/water interfaces where no prominent distortion caused by Fresnel factor occurs, only the effective susceptibilities instead of \( \chi_{ljk}^{(2)} \) are presented, indicating that the presented VSFG spectra are either normalized to the visible and IR intensities or to a standard such as z-cut quartz.

Another issue need to be addressed is that in SFG, if the underlying substrate is also SF active, it would generate another susceptibility, termed \( \chi_{NR}^{(2)} \). The subscript “NR” suggests its non-resonant nature that is IR frequency independent. For the susceptibility we discussed so far, it is related closely to the resonant behavior of the interfacial molecules (i.e. frequency dependent), and we can name this kind of susceptibility \( \chi_{R}^{(2)} \). Therefore,

\[
\chi^{(2)} = \chi_{NR}^{(2)} + \chi_{R}^{(2)} = \chi_{NR}^{(2)} + \sum_{\nu} \chi_{\nu}^{(2)}
\]  

(33)

for discrete vibrational modes.

In the end, we can finally get a simple expression for the intensity of the SFG beam. By combining equations (22) through (25) as well as (33), it is obvious that
\[ I_{SF} \propto |p^{(2)}|^2 \propto |\chi^{(2)} E_{vis} E_{IR}|^2 = |\chi^{(2)}|^2 \cdot I_{vis} \cdot I_{IR} \]

\[ = \left| \chi^{(2)}_{NR} + \sum_{\nu} \frac{A_{\nu}}{\omega_{IR} - \omega_{0,\nu} + i\Gamma_{\nu}} \right|^2 \cdot I_{vis} \cdot I_{IR} \]  

(34)

And the detailed form after geometry correction is

\[ I_{SF} = \frac{8\pi^2 \omega_{SF}^2 \sec^2 \theta_{SF}}{c^3 n(\omega_{SF}) n(\omega_{vis}) n(\omega_{IR})} \left| \chi^{(2)}_{NR} + \sum_{\nu} \frac{A_{\nu}}{\omega_{IR} - \omega_{0,\nu} + i\Gamma_{\nu}} \right|^2 \cdot I_{vis} \cdot I_{IR} \]  

(35)

where \( c \) is the speed of light in vacuum (2.99792458 \times 10^8 \text{ m/s}), \( \theta_{SF} \) is the angle between the output SF signal and surface normal (also can be noted as detection angle), \( n(\omega) \) is the refractive index of the bulk medium in which light travels at frequency \( \omega \), \( A_{\nu} \) is the amplitude of the VSFG transition moment of each mode \( \nu \) (\( A_{\nu} \) is also known as oscillator strength), \( \omega_{0,\nu} \) is the center frequency (usually in \( \text{cm}^{-1} \)) of the molecular vibrational transition, and \( \Gamma_{\nu} \) refers to the line width of that transition (\( \Gamma^{-1} \) is the relaxation time of the excited vibrational state). Note the existence of the term \( \left( \omega_{IR} - \omega_{0,\nu} \right) \) in the denominator of \( \chi^{(2)}_{\nu} \). As a consequence, we would expect to observe the Lorentzian peak shapes for collected SFG spectra, since \( |\chi^{(2)}|^2 \) is a Lorentzian function of the IR frequency.

For dielectric medium, such as pure water or salt solutions, \( \chi^{(2)}_{NR} \) is typically very small,\(^{40}\) hence the VSFG signal primarily originates from the resonant terms \( \sum_{\nu} \chi^{(2)}_{\nu} \). It basically means when the frequency of the incident infrared photons, \( \omega_{IR} \), is resonant with that of a vibrational mode \( \nu \) of the interfacial species, \( \left( \omega_{IR} - \omega_{0,\nu} \right) \) is approaching
zero and thus the value of $\chi^{(2)}_\phi$ would reach a local maximum and a VSFG intensity enhancement can be observed. This process is illustrated in Figure 2-2.
Figure 2-2. Energy diagram of vibrational sum frequency generation (VSFG). Different line widths between the energy levels suggest the possibilities of the corresponding transition. $|g\rangle$, $|v\rangle$, and $|s\rangle$ refer to the ground state, the excited vibrational state and any other state (a virtual state), respectively. The physical expression and explanation of the transitions have been given in equation (24).
2.2 Retrieval of Phase Information from VSFG Intensity Spectra

Since SFG is a coherent spectroscopic process, besides the magnitude, it should also contain the phase information of the output signal. As a matter of fact, by inverting the orientation of the interfacial molecules, we will change the sign of $\chi^{(2)}$ as well as the sign of $P^{(2)}$. The phase of the emitted light is hence changed by $\pi$. If we can retrieve the exact phase, we may then deduce the net orientation of such species that generate the SF signal. In other words, the phase of the emitted light can give us ideas whether the overall dipoles are aligned towards the interface or away from the interface (whether the molecules at the interfacial region are oriented “up” or “down” along the surface normal on average).

Let’s look at this process in another view. From the discussion in last section, we know that the second-order nonlinear susceptibility $\chi^{(2)}$ is a complex number and we have the resonant terms with such forms:

$$\text{Re}[\chi^{(2)}] = \sum_{v} \frac{A_v (\omega_{1R} - \omega_{0,v})}{(\omega_{1R} - \omega_{0,v})^2 + \Gamma_v^2}$$  \hspace{1cm} (36)

$$\text{Im}[\chi^{(2)}] = - \sum_{v} \frac{A_v \Gamma_v}{(\omega_{1R} - \omega_{0,v})^2 + \Gamma_v^2}$$  \hspace{1cm} (37)

The real part of $\chi^{(2)}_R$ has a dispersive shape as a function of frequency in the vicinity of a vibrational resonance, owing to the term $(\omega_{1R} - \omega_{0,v})$ in numerator and $(\omega_{1R} - \omega_{0,v})^2$ in denominator. On the other hand, the imaginary part of $\chi^{(2)}_R$ behave like a resonant peak with a maximum at $\omega_{0,v}$. The second-order nonlinear susceptibility $\chi^{(2)}_R$ intrinsically carries the absolute orientation, and this is information is directly correlated...
with the relative sign of the oscillator strength $A_\nu$ since it is determined by the IR transition dipole and the Raman transition polarizability of a molecule. Therefore, we can easily determine the sign of $A_\nu$ from the relative sign of $\text{Im}\chi^{(2)}_R$ and then deduce whether the net transition dipole of a molecule is the pointing “up” or “down” with respect to the interface normal. However, conventional VSFG can only yield an intensity spectrum due to its characteristic modulus square of the second-order response [see equation (35)]. By doing the square calculation, the inherent information of the sign of $A_\nu$ is lost. That is to say, it is challenging, or almost impossible to unambiguously obtain the net orientation of the transition moment solely with the intensity spectrum.

We can, however, fit the intensity spectrum by assigning relative sign for each vibrational mode using a Lorentzian model and derive the phase information. Unfortunately, fitting results are usually controversial between research groups because of the modulus square property in the intensity spectrum. Interference between the neighboring peaks can generate some intensity cross terms, and interference between the dispersed background in some instances and the vibrational resonances can lead to line shapes that are hard to interpret. Intrinsic difficulties for interpretation especially appear when we do not know in advance how many vibrational modes or peaks are present in the system being studied. Moreover, people usually adopt discrete resonances in fitting the intensity spectrum to equation (35). However, some spectral features such as the O-H stretching modes at the air/neat water interface are not discrete. The superposition of such vibrational resonances forms a broad continuum:

$$\chi^2 = \chi^{(2)}_{NR} + \int \frac{A_\nu \rho(\omega_q)}{\omega \omega_0(q) - \omega_0(q) + i \Gamma_q} d\omega_q$$

(38)
where $A_q$, $\omega_{\nu_q}$ and $\Gamma_q$ still denote the resonant amplitude (oscillator strength), central frequency and damping constant (line width) of the $q$th vibrational resonance, and $\rho(\omega_q)$ represents the density of modes at $\omega_q$.

Of course we can approximate the broad continuum with distinct resonances. But if we fit the experimental data with such distinct resonances without specifying a priori the resonant frequencies and the signs of such resonances, the fitting results usually are not unique.\textsuperscript{44} Therefore, different fitting methods and priori hypotheses existing in the VSFG research communities can often produce different results for the same VSFG intensity spectrum (sometimes the fitting results are even mutually exclusive from group to group), and accordingly lead to distinct interpretations of the structures of the same system.\textsuperscript{45-52} This situation calls for a direct measurement of the imaginary portion of the complex second-order nonlinear susceptibility ($\text{Im}[\chi^{(2)}]$) rather than $|\chi^{(2)}|^2$ because the imaginary portion accounts for absorption or emission spectroscopy and also contained the resonant information.\textsuperscript{44}

The need for the phase retrieval and the demand for an imaginary portion ($\text{Im}[\chi^{(2)}]$) spectrum actually are associated closely with each other since (the nonresonant part can be assumed as a constant and is omitted here for simplicity)

$$\chi_R^{(2)}(\omega_{1R}) = \text{Re}\left[\chi_R^{(2)}\right] + i \text{Im}\left[\chi_R^{(2)}\right] = |\chi^{(2)}|e^{i\phi(\omega_{1R})}$$  \hspace{1cm} (39)

where $|\chi^{(2)}|$ is the modulus (or absolute value) of the complex nonlinear susceptibility $\chi^{(2)}$. Once we get the phase information $\phi(\omega_{1R})$, it is quite easy to deduce the imaginary portion of $\chi^{(2)}$. 
Recently, people have developed two ways to gain the phase information. The first technique is the experimental measurement that can provide the complex $\chi^{(2)}$ spectrum by interference methods, first introduced by Yuen-Ron Shen and co-workers in UC Berkeley in 2005, using their scanning laser system.\textsuperscript{53-55} Traditionally we call this technique phase-sensitive vibrational sum frequency generation spectroscopy or PS-VSFG spectroscopy for short. Later, Alexander V. Benderskii group in Wayne State University\textsuperscript{56} and Tahei Tahara group in RIKEN, Japan\textsuperscript{57-59} developed this technique with their broad bandwidth (non-scanning) VSFG systems, respectively. They named this method as heterodyne-detected vibrational sum frequency generation (HD-VSFG) spectroscopy. Heterodyne detection is a method of detecting radiation/signal of interest by non-linear mixing with radiation of a reference (called local oscillator, LO), and the desired outcome is the difference frequency, which carries the information (amplitude, phase, and frequency modulation) of the original signal. No matter what the nomenclature people adopted, the basic principle still lies on interference of the sample’s SF signal with an additional SF response from the phase reference at the same frequency. I will briefly introduce the experimental set up and data processing method adopted in Tahei Tahara group\textsuperscript{57-59} and Heather Allen group.\textsuperscript{60-62}

Another phase retrieval method is based on a mathematical algorithm: maximum entropy method (MEM). MEM allows extraction of phase information of complex data from an intensity spectrum without any previous knowledge of the physical processes involved.\textsuperscript{63-66} It is widely used in geophysics, economics and spectroscopy such as coherent anti-Stokes Raman scattering (CARS).\textsuperscript{63-66} The application of this computational method on VSFG was proposed by Jung Y. Huang \textit{et al.} dating back to 1997\textsuperscript{67,68} and
Further discussed by Mischa Bonn in 2009. Various reports have shown that MEM is capable of reconstructing the absolute phase of the SF signal directly from VSFG intensity spectra at a wide variety of surfaces and interfaces. In part 2.2.2 we will focus on this subject and elaborate the procedure to extract the correct phase information.

2.2.1 Phase-sensitive measurement

The following experimental part is reported by Tahei Tahara and used here as an example of the HD-VSFG measurement.

The setup and optical configuration of a typical HD-VSFG is depicted in Figure 2-3(a) (adapted from Ref. 57). An optical parametric amplifier (OPA) and a difference frequency generator (DFG) are used to generate a broadband IR beam whose pulse covers the whole frequency region under investigation, free from scanning the IR center wavelength (In this example, water region is chosen). The visible beam (ω₁) and broadband IR beam (ω₂) are then spatially and temporally overlapped on a sample surface to get the sum frequency beam. The three laser beams reflected by the sample surface are refocused by a concave mirror onto a GaAs surface (acts as a local oscillator, LO) to generate another sum frequency beam. Note the first sum frequency pulse from the sample passes through a thick silica plate positioned in between the sample stage and the concave mirror, which delays the first sum frequency pulse relative to the input IR and visible pulses by a certain time T. The real function of this silica is a phase-modulation plate, in the purpose of controlling the relative phases of the three waves due to the dispersion of silica. As a consequence, this delay leads to the time and
phase difference between the first sum frequency pulse from the sample and the second one from the GaAs surface. The two sum frequency pulses are sent together into a polychromator and detected by charge coupled device (CCD). In the polychromator, the two sum frequency responses are stretched in time and interfere with each other to generate an interference pattern (fringe) in the frequency domain as shown in the raw spectrum in Figure 2-3(b).
Figure 2-3. (a) The schematic of an HD-VSFG optical configuration. (b) Raw spectra of the sample (blue) and z-cut quartz crystal (red). (c) Time domain interferogram obtained by the inverse Fourier transformation of raw data and the filter function (green, plotted to the right vertical axis). (d) Real (solid lines) and imaginary (dashed lines) parts of the heterodyne spectra by the Fourier transformation of the time-domain interferogram after filtration of the sample and reference. This figure is reproduced from Ref. 57.
The overall electric fields of the SF light in the time domain can be written as

\[ E_{\text{total}}(t) = E_{\text{sample}}(t - T) + E_{\text{LO}}(t) \]

where \( E_{\text{sample}} \) and \( E_{\text{LO}} \) represent the SF electric fields directly from the sample and the local oscillator (GaAs in this case), respectively. \( T \) is the time delay between them.

If we perform Fourier transformed into the frequency domain in the following way:

\[ E_{\text{total}}(\omega) = \int_{-\infty}^{+\infty} E_{\text{total}}(t) e^{i\omega t} dt = E_{\text{sample}}(\omega) e^{i\omega T} + E_{\text{LO}}(\omega) \]

It should be noted that \( E_{\text{sample}} \) and \( E_{\text{LO}} \) are both in the frequency domain now so we can use \( \tilde{E}_{\text{sample}} \) and \( \tilde{E}_{\text{LO}} \) to indicate this fact. The detected total signal [shown in the raw spectrum in Figure 2-3(b)] is:

\[ I = |\tilde{E}_{\text{total}}(\omega)|^2 = |\tilde{E}_{\text{sample}}|^2 + |\tilde{E}_{\text{LO}}|^2 + \tilde{E}_{\text{sample}} \tilde{E}_{\text{LO}}^* e^{i\omega T} + \tilde{E}_{\text{sample}}^* \tilde{E}_{\text{LO}} e^{-i\omega T} \]

Note there are some complex conjugates (the terms with asterisk) appearing in the right hand side of this equation. The two cross terms \( \tilde{E}_{\text{sample}} \tilde{E}_{\text{LO}}^* e^{i\omega T} \) and \( \tilde{E}_{\text{sample}}^* \tilde{E}_{\text{LO}} e^{-i\omega T} \) correspond to the interference fringe (fine) structures that contains the phase information. That is what people got for the raw \( |\tilde{E}_{\text{total}}(\omega)|^2 \) data. If we do the inverse Fourier transformation of the raw spectrum, we will re-get the time domain interferogram as Figure 2-3(c) shows. \( \tilde{E}_{\text{sample}} \tilde{E}_{\text{LO}}^* e^{i\omega T} \) would give the peaks at \( t = +T \) and \( \tilde{E}_{\text{sample}}^* \tilde{E}_{\text{LO}} e^{-i\omega T} \) is for \( t = -T \). The peak at origin \( (t = 0) \) is ascribed to the first and second terms \( |\tilde{E}_{\text{sample}}|^2 + |\tilde{E}_{\text{LO}}|^2 \). Another interesting character for the time domain interferogram is that the three groups of peaks are well separated. We can filter out one of the intensity cross terms like the one at \( t = +T \) with ease and preform Fourier
transformation back into the frequency domain. This time we will only obtain the third term of the equation, $\vec{E}_{\text{sample}} \vec{E}_{\text{LO}}^* e^{i\alpha T}$, depicted in Figure 2-3(d).

Now let’s center on the SF electric field $\vec{E}_{\text{sample}}$ and $\vec{E}_{\text{LO}}$. They can be described by the second-order nonlinear susceptibilities and

$$\vec{E}_{\text{sample}} = i a_{\text{sample}} \chi^{(2)}_{\text{sample}} \vec{E}_{\text{vis}} \vec{E}_{\text{IR}}$$

$$\vec{E}_{\text{LO}} = a_{\text{LO}} \chi^{(2)}_{\text{LO}} r_{\text{sample,vis}} \vec{E}_{\text{vis}} r_{\text{sample,IR}} \vec{E}_{\text{IR}}$$

where $a_{\text{sample}}$ and $a_{\text{LO}}$ are positive constants (real numbers) for each SF response and are functions of related Fresnel factors, $r_{\text{sample,vis}}$ and $r_{\text{sample,IR}}$ are the reflectivities of the sample for visible and IR beams, respectively. This reflectivity quantity is present because the SF response from LO (GaAs) is generated by the reflected visible and IR beams after hitting on sample surface. It should be pointed out that there is an imaginary unit in the expression for $\vec{E}_{\text{sample}}$ but disappear in $\vec{E}_{\text{LO}}$. That is because the bulk of the sample is not SF active while the bulk of GaAs is SF active.

So how can we get the $\chi^{(2)}_{\text{sample}}$ spectrum in the end? We may find a reference with known phase and do the normalization with the acquired sample intensity. To this purpose, a HD-VSFG spectrum was measured from the z-cut quartz crystal with the same GaAs. Similarly,

$$\vec{E}_{\text{quartz}} = a_{\text{quartz}} \chi^{(2)}_{\text{quartz}} \vec{E}_{\text{vis}} \vec{E}_{\text{IR}}$$

$$\vec{E}'_{\text{LO}} = a_{\text{LO}} \chi^{(2)}_{\text{LO}} r_{\text{quartz,vis}} \vec{E}_{\text{vis}} r_{\text{quartz,IR}} \vec{E}_{\text{IR}}$$

And $a_{\text{quartz}}$ is again real number and positive constant, $r_{\text{quartz,vis}}$ and $r_{\text{quartz,IR}}$ are the reflectivities of the quartz for the electric field induced by visible and IR beams. The bulk
of quartz is \( \chi^{(2)} \) active hence no \( i = \sqrt{-1} \) is included in the right hand side of the expression.\textsuperscript{24,25} Moreover, \( \chi^{(2)}_{\text{quartz}} \) can be treated as a real number because quartz does not have any electronic resonances in either IR, visible or sum frequency regions.\textsuperscript{24,25}

By dividing the sample interferogram [Figure 2-3(d), blue curves] by the quartz reference interferogram [Figure 2-3(d), red curves] (normalization process), we have

\[
\frac{\tilde{E}_{\text{sample}} \tilde{E}_{\text{LO}}^* e^{i\omega T}}{i \tilde{E}_{\text{quartz}} \tilde{E}_{\text{LO}}'^* e^{i\omega T}} = \frac{a_{\text{sample}} \chi^{(2)}_{\text{sample,vis} \text{sample,IR}}}{a_{\text{quartz}} \chi^{(2)}_{\text{quartz,vis} \text{quartz,IR}}}
\]

\[
\chi^{(2)}_{\text{sample}} = \frac{c_{\text{quartz}}}{c_{\text{sample}}} \frac{\tilde{E}_{\text{sample}} \tilde{E}_{\text{LO}} e^{i\omega T}}{\tilde{E}_{\text{quartz}} \tilde{E}_{\text{LO}}'^* e^{i\omega T}} \chi^{(2)}_{\text{quartz}}
\]

That is to say, we can finally retrieve the nonlinear susceptibility \( \chi^{(2)}_{\text{sample}} \) relative to the reference by dividing the filtered sample intensity by the filtered reference intensity coupled with experimental corrections (the constants \( c_{\text{sample}} \) and \( c_{\text{quartz}} \)) for reflectivity ratios and Fresnel factors as well as \( \chi^{(2)}_{\text{quartz}} \) (another real constant over the spectral range under the present electronically nonresonant condition). The imaginary part of the result \( \text{Im}[\chi^{(2)}] \) is directly associated with the vibrational resonance and thus most informative. It is a linear superposition of the resonances in \( \chi^{(2)} \):\textsuperscript{44}

\[
\text{Im}[\chi^{(2)}] = A_{\text{NR}} \sin \phi_{\text{NR}} - \sum_{v} \frac{A_v I_v}{(\omega_{\text{IR}} - \omega_{0,v})^2 + I_v^2}
\]

All the intensity cross terms that are generated by the square now get eliminated, and it is more straightforward and convenient to check each of the vibrational modes as well as the absolute orientation of the transition dipole of a molecule from the sign of \( A_v \).
2.2.2 Maximum entropy method

MEM have been carried out on different VSFG spectra with various nonresonant backgrounds and resonant peak widths and good agreements with phase-sensitive experimental measurements are obtained. The result deduced by maximum entropy corresponds to the most probable result in the space of all possible outcomes that satisfy the known information. It is rather fascinating since no model is involved in this process and therefore the requirement for experimental phase measurement can be avoided. A brief review of the procedure of phase-retrieval by MEM is as follows.

The entropy (information) for an intensity spectrum $I(\omega)$ in the frequency interval $[\omega_1, \omega_2]$ is expressed by definition as

$$h = \int_{\omega_1}^{\omega_2} \log I(\omega) d\omega$$

We can introduce a normalized frequency $v = \frac{\omega - \omega_1}{\omega_2 - \omega_1}$ with $\omega \in [\omega_1, \omega_2]$ to project the frequency interval to [0, 1]. Using variational calculus and the Lagrange multiplier method, we can get the solution that maximizes the spectral entropy under the constraint of the autocorrelation function calculated from the measured spectral points (which will be elaborated later). The resulting solution (also the expression for the spectral intensity) for $2M + 1$ spectral points is:

$$I(v) = \frac{|\beta|^2}{|1 + \sum_{k=1}^{M} a_k \exp(i2\pi kv)|^2}$$

The unknown coefficients $a_k$ and $|\beta|^2$ (total unknown number is $M+1$) can be determined from the Toeplitz matrix equation:
where $R(m)$ is the autocorrelation function, which is the Fourier transform of the power spectrum $I(v)$:

$$R(m) = \int_0^1 I(v) \exp(-i2\pi mv) \, dv$$

The solution of $\chi^{(2)}(v)$ with maximum entropy is given by

$$\chi^{(2)}(v) = \frac{|\beta| \exp(i\phi(v))}{1 + \sum_{k=1}^{M} a_k \exp(i2\pi kv)}$$

We can use another quantity $A_M(v)$ to replace the denominator for clarity:

$$A_M(v) = 1 + \sum_{k=1}^{M} a_k \exp(i2\pi kv) = |A_M(v)| \exp(-i\psi(v))$$

$$\chi^{(2)}(v) = \frac{|\beta| \exp(i\phi(v))}{|A_M(v)| \exp(-i\psi(v))}$$

Here $\phi(v)$ is the error phase introduced in the phase retrieval analysis and $\psi(v)$ is the one contained within $A_M(v)$, called MEM phase. It is the frequency-dependent quantity returned by the phase retrieval algorithm as the most likely phase associated with the SFG intensity spectrum.\(^{69}\)

The frequency-dependent absolute phase $\theta(v)$ associated with the surface vibrational resonances can now be expressed as

$$\theta(v) = \psi(v) + \phi(v)$$

Therefore, the imaginary part of the second-order nonlinear susceptibility originated from resonant part $\chi^{(2)}(v)$ is calculated as:
Im\left[\chi^{(2)}(v)\right] = \sqrt{I(v)} \cdot \sin \theta(v) = \sqrt{I(v)} \cdot \sin[\psi(v) + \phi(v)]

Nonetheless, the nonlinear susceptibility also contains a nonresonant (NR) contribution, \(\chi_{NR}^{(2)}\), which is a frequency independent quantity associated with the nonresonant SFG intensity. As a consequence, one also need to add a nonresonant phase, \(\phi_{NR}\), to get the correct overall absolute phase:

\[\text{Im}\left[\chi^{(2)}\right] = \sqrt{I(v)} \cdot \sin[\psi(v) + \phi(v) + \phi_{NR}]\]

As described earlier, because coefficients \(a_k\) and \(|\beta|^2\) can be obtained by the Toeplitz matrix, the MEM phase \(\psi(v)\) is what we know. The true phase (absolute phase) contains two additional contributions that cannot be deduced from the power spectrum \(I(v)\): the error phase \(\phi(v)\) and the nonresonant phase \(\phi_{NR}\). Therefore, the \(\text{Im}[\chi^{(2)}]\) spectrum can be obtained from the MEM phase but only after correcting for the error phase and the nonresonant phase. The error phase \(\phi(v)\) is most likely ascribed to the presence of other resonances (e.g. electronic resonances). We know this process is quite minimal in our VSFG study. The error phase is typically estimated with a slowly varying function over the frequency range, and has values close to zero (negligible). To perform the nonresonant phase correction, there are several criteria to follow: 67-69, 71

1. After correction, the constructed \(\text{Im}[\chi^{(2)}]\) should be in accordance with known molecular orientations. For example, at the air/pure water interface, there is a prominent peak centered on 3700 cm\(^{-1}\) which is attributed to the dangling O-H groups of the water molecules in the topmost layer. It is restricted to let the O-H vectors pointing upwards (towards the air along the surface normal), hence the dipole moment should be positive. When doing the
phase correction, we should use this pre-determined orientation as an internal reference to make the \( \text{Im}[\chi^{(2)}] \) spectrum positive at this region.

2. The imaginary part of the susceptibility should asymptotically approach a constant in the regions far away from resonance.

3. The resonances (especially for an isolated one) must appear as peaks in the \( \text{Im}[\chi^{(2)}] \) spectra with a frequency accuracy of a few inverse wavenumbers as that of the power spectrum (modulus).

4. For a sharp resonance peak, the imaginary part of the susceptibility is an even function relative to the resonant frequency. It should be entirely positive or entirely negative and follow a symmetric line shape.

5. \( \phi_{NR} \) lies in the interval \([0, 2\pi]\).

The biggest advantage of MEM analysis for the intensity spectrum is it can be applied to any SFG measurement and does not need a sophisticated PS-SFG setup. It does not require a reference sample either, but at the cost of the phase correction which should be performed with great care.

In a recent report by Heather Allen and Sylvie Roke, a direct comparison of PS-SFG water spectra with the result acquired by MEM calculation is done.\(^{43}\) Both methods indeed yield the same complex spectrum after corrections. This indeed support our idea that MEM does provide a valuable algorithm for the VSFG data analysis, helping prediction of some interfacial structures.

The Matlab script used for our MEM analysis can be found in Appendix B of this dissertation and we will show the application of this method in Chapter 3 in detail.
2.3 VSFG Instrumentation

Two VSFG systems were used in this work. The first system has been described in detail in our previous publication. Briefly, it consists of five major parts: a custom made active/passive mode-locked Nd:YAG laser (Continuum, model PY-61C, Santa Clara, CA), an optical parametric generation/optical parametric amplification (OPG/OPA) stage (LaserVision, Bellevue, WA), a sample stage, a signal detection/data acquisition unit and beam delivery optics between each components. The overall set up is illustrated in Figure 2-4.

Neodymium doped yttrium aluminum garnet (Nd:Y₃Al₅O₁₂ or Nd:YAG) crystal is the gain medium in the laser system, equipped with a negative feedback loop. The optical bench in this laser is composed of a oscillator pumped by single flashlamp, an amplifier pumped by two flashlamps in a close coupled configuration, and a Q-switch dye pump package DCP-03 for mode-locking. A fundamental 1064 nm laser with a very high pulse-to-pulse stability and beam uniformity is generated. The produced beam has a pulse width of 17 picoseconds, a repetition rate of 20 Hz and the highest output energy of 50 mJ/pulse. The environmental conditions (room temperature, relative humidity, cooling water flow rate, etc.) are strictly controlled to maintain proper optical head performance and avoid laser power fluctuation. The output beam profile can be observed by an IR detection card or shots on burn paper as circular spot with the ring diffraction pattern. Normally flashlamps are changed every 20 million shots and the Q-switch dye is replaced with freshly prepared solution in purified 1,2-dichloroethane (99+%, Acros Organics;
Geel, Belgium) when the operational threshold reaches ~1.39 kV or when the level of the dye solution is near half of the container.

The optical parametric generation/amplification (OPG/OPA) are then pumped with this 1064 nm fundamental beam with a minimum energy of 40 mJ/pulse (required for the downstream conversion). This OPG/OPA unit can convert the fundamental beam into two new lasers: one fixed frequency visible beam (532nm, green) and one frequency-tunable infrared beam (2000 cm\(^{-1}\) - 4000 cm\(^{-1}\), invisible to naked eye).

Firstly, the polarized 1064 nm laser beam is focused and magnified by a telescope and then divided into two streams with the original frequency by a beam splitter. One stream is sent directly towards the difference frequency generation (DFG) stage which will be mentioned later. The other stream is guided to a nonlinear crystal potassium titanyl arsenate (KTiOAsO\(_4\) or KTA) to double the frequency to 532 nm. The 532 nm beam, after sending the unconverted fundamental beam to dump, is once again spitted into two parts: one is sent out of the parametric system, being used as the input visible light in VSFG measurement, and one further for the OPG/OPA process. The nonlinear crystals used in this process are two potassium titanyl phosphate (KTiOPO\(_4\) or KTP) crystals. OPG, or more commonly known as spontaneous parametric down-conversion can happen with the help of these crystals, in which a pump photon is spitted into a pair of photons. People usually name the new photons as “signal” (with relative higher energy/shorter wavelength) and “idler” (with relative lower energy/longer wavelength). The frequencies of the signal and the idler pair are determined by the phase matching conditions (the laws of conservation of energy and momentum), which can be tuned by the angle between the incident pump laser and the optical axes of the crystal. The output
beams (signal and idler) in optical parametric generation are relatively weak in intensity, therefore they must be amplified. In our systems, this is done by pumping the crystals from both directions and recombine a portion of the 532 nm with the idler collinearly through KTP crystals for numerous turns. The OPA will weaken the pump 532 nm beam and amplify the idler beam. After this process, the remaining 532 nm beam and the signal are sent towards a beam dump by a dichroic mirror, only letting the idler travel to the next stage for the generation of desired mid-IR beam. This stage, coupled with two KTA crystals, can mix the path delayed 1064 nm beam (mentioned earlier) and the idler, where difference frequency generation (DFG) takes place. All four crystals (two KTP in OPG/OPA stage and two KTA in DFG stage) are mounted to motors that are connected to and controlled by LaserVision software. By gradually changing the positions of the motors, we can easily tune the angles of the nonlinear crystal, producing an idler of wavelength 1.35 to 1.85 μm and an infrared beam of 5 to 2.5 μm (2300-4000 cm$^{-1}$). The intensity of the IR beam is close to 600 μJ/pulse at 3200 cm$^{-1}$ under good beam alignment and extensive motor position optimization.
Figure 2-4. The experimental set up of the first VSFG spectrometer. The sum frequency signal is generated by temporal and spatial overlapping of input infrared and visible beams at the sample surface.
We adopted collinear and co-propagating geometry for the two input beams, with an incident angle of 45° for the green beam and 55° for IR. The two beams were focused by lenses and guided to achieve overlapping appropriately in time and space at the sampling area. The sum frequency response from the sample surface are generated at a certain directions in accordance with the law of conservation of momentum. After proper filtration of undesired optical signal, the sum frequency output was collected in the reflected direction through a series of steering optics and polarizers to a high-gain low-noise photomultiplier tube (model R647, Hamamatsu, Japan), then sent to a gated integrator/boxcar averager (Stanford Research Systems, Inc. Sunnyvale, CA) to improve the signal to noise ratio. The averaged digital data is acquired by a software based on LabVIEW.

All the data reported in this dissertation was obtained under ssp polarization combination (s polarized sum frequency, s polarized visible and p polarized infrared) and were subject to background subtraction and normalization against a non-resonant VSFG spectrum obtained from a right-handed z-cut quartz crystal to eliminate the spectral distortion caused by the uneven infrared beam energy distribution at different frequency in the spectral region of interest. Each VSFG spectrum was repeated at least three times and averaged to ensure reproducibility.

The second SFG spectrometer (current one being used in Pennsylvania State University) is purchased from EKSPLA (Vilnius, Lithuania). The setup is rather similar to the home-made first one and schematically shown in Figure 2-5.
Figure 2.5. The experimental set up of the second VSFG spectrometer purchased from EKSPLA.
This commercially available VSFG spectrometer is composed of a picosecond Nd:YAG laser (PL2251 series), a harmonic unit (model No. HM500), an optical parametric generation (OPG)/ optical parametric amplification (OPA)/difference frequency generation (DFG) system (model No. PG501/DFG1P), and a detection system coupled with monochromator and PMT. In this system, diode pumped master oscillator and one flashlamp pumped amplifier are used for the Nd:YAG laser. The output 1064 nm laser beam is of 20 ps pulse width and 50 Hz. This repetition rate can be adjusted if necessary. The visible beam (532 nm) is generated by frequency-doubling of the fundamental 1064nm pulses by potassium dideuterium phosphate (KD2PO4, KD*P) nonlinear crystals in the harmonic unit. OPG and OPA can produce a signal beam (680 to 1064 nm) and an idler beam (1065 to 2299 nm). The idler beam and the 1064 nm pump beam are used in a DFG stage to generate a tunable mid-IR light (2300 nm to 10000 nm, i.e. 1000 cm⁻¹ to 4300 cm⁻¹). For VSFG experiments carried out in this work, the input visible and IR energies are typically ~200 µJ/pulse and ~100 µJ/pulse, respectively. The incident angles of the visible and IR input beams are 60° and 55° to the surface normal. The VSFG signal from the surface is collected by a photomultiplier tube (model R585, Hamamatsu, Japan) attached to a monochromator (MS2001).

The same sample stage were used for both systems. For the air/water interface, a mini Langmuir trough (model 601M, NIMA, U. K.) can serve as a good sample holder of the aqueous solution. The image of the trough is shown in Figure 2-6. It is equipped with a pair of sliding barriers and a surface pressure sensor which allows measurement and control of the surface pressure by a paper Wilhelmy plate, to a minimum of 0.1 mN/m.
Figure 2-6. The Langmuir trough used as the sample holder for air/water interface experiments.
Chapter 3
The Study of Trimethylamine N-Oxide Orientation at Hydrophobic Interfaces

3.1 Background

Trimethylamine-N-oxide (TMAO) is one kind of protecting osmolytes used by nature that are of great importance in stabilizing intracellular proteins against a wide variety of adverse or extreme environmental conditions.\textsuperscript{75,76} In contrast to denaturing osmolytes (denaturants), protecting osmolytes thermodynamically favor the folded state of proteins. TMAO, for instance, has been shown to offset the effects of denaturants such as urea in the kidneys.\textsuperscript{77-79} However, the molecular level mechanism on how this molecule stabilizes proteins is much less explored and, hence, not entirely understood.\textsuperscript{80-83} Previous research related the stabilizing and denaturing effects of compounds to excluded volume and preferential hydration.\textsuperscript{84,85} These models are based on whether a cosolvent partitions to the protein-water interface and often involve surface tension studies. Specifically, denaturants typically decrease the surface tension of the air/water interface and reside at the protein/water interface, implying the direct interaction with hydrophobic proteins. On the other hand, protein stabilizers are generally believed to stay away from the protein/water interface. Thus, they should increase the surface tension of the air/water interface, which leads to a theory that protein stabilizers act through a more indirect fashion. Although many protein denaturants and stabilizers fit this general model well, there do exist some exceptions. For example, urea, a denaturant, slightly increases
the surface tension of the air/water interface,\textsuperscript{86} leading to a puzzling conclusion that it does not reside at the protein/water interface.

The calculated partition coefficients of TMAO, unlike many other protein stabilizing compounds, shows a mild accumulation at protein/water interface when this osmolyte is added to solution.\textsuperscript{87} This may imply that TMAO most likely resides at the protein/water interface and may directly interact with the protein.\textsuperscript{86} In fact, it has been shown that TMAO uniquely changes water structures at the protein/water interface, which could in turn change the solubility of the protein and stabilize the folded state.\textsuperscript{88} However, there is disagreement as to whether TMAO weakens or strengthens water structure at the protein/water interface.\textsuperscript{89} In addition, changes in water structures have not been universally observed for other protecting and denaturing osmolytes, making it difficult to explain a general mechanism based solely on changes in water structures.

Several recent studies have introduced a general mechanism for the protecting or denaturing ability of various osmolytes through preferential interactions.\textsuperscript{90-92} According to this model, denaturing osmolytes would possess favorable interactions with various polar groups in proteins, while stabilizing osmolytes would have unfavorable interactions with these same groups. However, due to the similar chemical makeup of protecting and denaturing osmolytes, these arguments often invoke changes in the exposed surface area of the protein. In addition, these studies typically focus on polar and charged groups in proteins, ignoring possible interactions between osmolytes and the hydrophobic groups. The fact that TMAO decreases the air/water surface tension implies that it builds up at hydrophobic interfaces and raises the possibility of important interactions with these surfaces. Since many studies have shown the importance of hydrophobic groups in
protein collapse and folding,\textsuperscript{93-96} we found strong impetus to carry out a study investigating TMAO at hydrophobic interfaces.

In this chapter, we probed TMAO at two hydrophobic interfaces, air/water and OTS/water, with the help of vibrational sum frequency generation spectroscopy. This interface-specific technique can provide not only a vibrational spectrum of molecules residing at the interface, but also the orientational information of the chemical moieties that give rise to the corresponding signal,\textsuperscript{45,46,97} as mentioned earlier. The molecular orientation of the methyl groups of TMAO was investigated at both hydrophobic interfaces, coupled with a numerical algorithm, the maximum entropy method (MEM) calculation. The results of both methods manifest that TMAO is oriented with its methyl groups pointing away from the hydrophobic surfaces (Figure 3-1). This observation may provide molecular-level clues into the stabilizing nature of TMAO since orienting the negatively charged oxygen atom toward to hydrophobic groups of proteins may increase the propensity for hydrophobic collapse.
Figure 3-1. The experimental setup containing a Teflon flow cell with a quartz window on top and bottom. The top quartz window contains an OTS monolayer on the bottom face. The sum frequency generation spectrum is obtained by passing the visible and infrared beams through the top quartz window and selecting for the bottom face.
3.2 Experimental Section

Preparing TMAO Solutions in D$_2$O

Trimethylamine N-oxide (dihydrate, 98% purity, Fisher Scientific, Waltham, MA) was dissolved in 99.9% D$_2$O (Cambridge Isotope Laboratories, Inc.). To exchange H$_2$O bound to the oxygen atom of TMAO for deuterium, the TMAO solutions were evaporated using a rotoevaporator, redissolved in D$_2$O and evaporated again. This process was repeated 4 to 5 times until no hydrogenated water was detected in the VSFG spectra between 3000 and 4000 cm$^{-1}$. The dried TMAO was redissolved in D$_2$O, sonicated for about 1 hour, and used for measurements.

Preparing and Characterizing OTS Monolayers on Quartz

IR grade quartz discs (1 inch diameter by 1/8 inch thick; Quartz Plus, Inc. Brookline, NH) were cleaned by the standard procedures. The quartz pieces immersed in a V:V=50:50 mixture of concentrated sulfuric acid and concentrated nitric acid for several hours. The quartz pieces were then rinsed thoroughly with deionized water and dried with nitrogen gas, followed by oven calcination for about half an hour.

These cleaned quartz discs were subsequently used as substrates for the preparation of the self-assembly of octadecyltrichlorosilane or OTS (Sigma-Aldrich, St. Louis, MO) monolayer. Prior to the monolayer formation, the cleaned quartz pieces were soaked in a 2 M NaOH solution for 15 minutes at room temperature, rinsed with copious amounts of deionized water, dried under nitrogen gas, and left in a drying oven for about 30 minutes. Upon cooling to room temperature, the quartz pieces were transferred to a 1
mM OTS solution in hexane and sit for 2 hours to form the OTS monolayer. The samples were then cleaned with ethanol, acetone and deionized water, ensured the removal of residual OTS. The samples were stored in deionized water and dried with nitrogen gas just prior to use.

The OTS monolayers on quartz were characterized by VSFG with air/quartz interface using three different polarization combinations (ssp, ppp and sps). The spectra agree well with previous literature results and indicate that the OTS monolayer is in an (nearly) all-trans alkyl chain configuration. This is in agreement with previous studies reporting the tilt angle is close to 10° from the surface normal.

**Experiments at the Quartz/Water Interface**

The homemade Teflon flow cell fabrication and setup for solid/water interface measurement has been described in detail elsewhere. A quartz disk coated with OTS was placed onto the cell face down so that the OTS was in contact with the solution inside it (Figure 3-1). The flow cell was assembled in H₂O or D₂O. Solutions of TMAO were subsequently flowed into the cell using a 60 mL syringe. The input beams were transmitted through the quartz substrate and focused at the OTS/solution interface.

**Vibrational Sum Frequency Generation Spectroscopy**

The VSFG system which was employed in these studies has been described in detail in the Chapter 2 (old system). The polarization combination used in all the TMAO experiments was ssp unless otherwise specified. The VSFG spectra reported were normalized to the nonresonant response from a piece of Z-cut crystalline quartz after
background subtraction. To help elucidate the differences in the oscillator strength $A_\nu$, the intensities were further fit to the following equation using the Matlab software:

$$I_{SF} \propto \left| \chi_{NR}^{(2)} + \sum_\nu \frac{A_\nu}{\omega_{1R} - \omega_{0,\nu} + iI_\nu} \right|^2$$

The Matlab script used for spectra fitting can be found in Appendix A of this dissertation.
3.3 Results

TMAO Solutions at the Air/Water Interface

The starting point for this discussion is a comparison of spectra of different concentrations of TMAO in aqueous solution at the air/water interface in a Langmuir trough. For all concentrations, prominent peaks in the C-H stretching region near 2950 cm\(^{-1}\) were observed, corresponding to the symmetric stretch frequency of the methyl groups on TMAO (Figure 3-2, left panel). This demonstrated that the osmolyte does in fact reside in a well oriented fashion at the air-water interface. It is intuitively correct to expect that the methyl groups (hydrophobic moieties) of organic compounds will replace the energetically unfavorable dangling O-H bonds from the topmost layer of water molecules at the interface. This is demonstrated by the disappearance of the stretch of dangling O-H bonds at ~3700 cm\(^{-1}\) and is observed with low concentrations of many surfactants.\(^{107,108}\) The 3700 cm\(^{-1}\) peak also attenuates in the present experiments, albeit much more slowly. Indeed, it is still observable when 5 M TMAO is present in solution.

Fitting the oscillator strength of both the 2950 and 3700 cm\(^{-1}\) peaks indicates a roughly linearly increase of the former (the C-H stretch) and a roughly linear decrease of the latter (the O-H stretch) between 2 and 5 M (see Figure 3-2, right panel). This relatively gradual change in the free O-H stretch signal together with the high concentration used appear to indicate that there is not substantial accumulation of TMAO at the air/water interface relative to bulk solution. And this also indicate TMAO molecules may adopt a different interfacial orientation than that of as a typical surfactant.
Figure 3-2. (a) VSFG spectra for TMAO of various concentrations at the air/water interface. (b) The oscillator strength of the OH stretch of water (3700 cm\(^{-1}\) peak) and the symmetric CH\(_3\) stretches of TMAO (2750 cm\(^{-1}\) peak) vs TMAO concentration. Although the dangling OH bonds of water decrease with TMAO concentration in a linear fashion, the peak is still evident even at 5 M TMAO.
Determining the absolute orientation of the methyl peaks in TMAO can give us ideas on how it arranges at the air/water interface. This can be demonstrated, for example, by fitting the 4 M TMAO data (Figure 3-3). Since the free OH stretch is still present at higher TMAO concentrations, we may assume that the water resonances and phases are similar to those at the neat air/water interface. To do the analysis, four peaks are employed to fit the O-H stretch modes from the water: ~3110 cm\(^{-1}\), 3220 cm\(^{-1}\), 3400 cm\(^{-1}\), and 3700 cm\(^{-1}\). In agreement with previous work at the air/water interface, the phases of resonances with the lowest and highest frequencies have positive signs, while the signs for the two intermediate resonances are negative.\(^{47}\) These assignments are consistent with phase sensitive measurements of the air/water interface which indicate a sign change from positive to negative above 3200 cm\(^{-1}\) and another change from negative back to positive near 3630 cm\(^{-1}\) for the imaginary portion of \(\chi^{(2)}\).\(^{55,109}\)

It can be seen that the fit to the experimental data are quite good when a negative sign is employed for the 2950 cm\(^{-1}\) C-H stretching resonance [Figure 3-3(a)], while the fit is poor when the sign is assumed to be positive [Figure 3-3(b)]. It should be noted that a very small additional resonance was employed in Figure 3-3 near 3020 cm\(^{-1}\), which corresponds to methyl asymmetric stretch in TMAO. The best fit to this resonance was found when its sign was positive. Again, the 3700 cm\(^{-1}\) peak for the dangling O-H stretch is well-known to face up into the air,\(^{45,46}\) which means that the methyl groups must point in the opposite direction (face down).
Figure 3-3. The fitting of VSFG data for 4 M TMAO at the air/water interface. The peak assigned to the symmetric stretches of the TMAO methyl group at ~2950 cm\(^{-1}\) is fit with a negative sign in a) and positive sign in b). It is clear that the fit shown in a) is much better, indicating that the methyl groups of TMAO most likely point away from the hydrophobic interface with air.
The sign of the oscillator strength of the 2950 cm\(^{-1}\) resonance (negative) indicate that their methyl group orientation is in fact opposite to most organic surfactants. As such, the methyl groups of TMAO should point away, instead of towards the hydrophobic interface.

This somewhat unusual finding is confirmed in the present work by employing the maximum entropy method (MEM)\(^{69-71}\) to calculate the sign of the imaginary portion of \(\chi^{(2)}\) (Figure 3-4). The maximum entropy method, whose basic principles are presented in detail in the last chapter, is a useful supplementary way for phase determination other than bare spectra fit. Its output also indicated that the methyl peaks from TMAO possessed the opposite sign from the dangling OH, thus should point down into the bulk solution.
Figure 3-4. MEM calculations of the imaginary portion of $\chi^{(2)}$ as a function of wavelength at the air/water interface in the presence and absence of 4 M TMAO. These calculations are in agreement with recent phase sensitive VSFS measurements at the air/water interface. In addition, the large negative amplitude of $\chi^{(2)}$ at $\sim$2950 cm$^{-1}$ (symmetric stretch of the methyl group) indicates the methyl groups of TMAO point away from the air/hydrophobic interface.
OTS Monolayers in TMAO Solution

We also wanted to test if the orientation of TMAO found at this air/water interface would be the general case in other aqueous/hydrophobic interfaces. Moreover, if we can find another interface with an easily known (or pre-known) phase reference, it would become much more convenient to assign the absolute orientation of TMAO. For these purposes, this osmolyte was investigated at the OTS/water interface (a model for oil/water interface).

The quartz with self-assembled OTS monolayers were used as the top window of a homemade Teflon flow cell, with the monolayer side facing towards D$_2$O aqueous medium. The data for such interface is shown as the in black spectrum of Figure 3-5a. We can clearly see the presence of two main peaks at 2876 cm$^{-1}$ and 2934 cm$^{-1}$, corresponding to the CH$_3$ symmetric stretch of the terminal methyl group as well as a Fermi resonance.$^{101}$ When 6 M TMAO solution dissolved in D$_2$O was flowed into the flow cell, the higher frequency peak increased in intensity by more than 10% and became a little blue-shifted (Figure 3-5a, red spectrum). On the other hand, the intensity as well as the location (2876 cm$^{-1}$) of CH$_3$ symmetric stretch peak remained unchanged.
Figure 3-5. VSFG spectra of the OTS/D$_2$O interface, taken in a quartz flow cell. The spectra shown in a) compare OTS in D$_2$O with OTS in 6 M TMAO. Since it is known that the methyl groups of OTS point down into solution, the constructive interference with the TMAO methyl group peak indicates this group is also pointing down. The two spectra overlaid in b) are OTS in D$_2$O and OTS in 6 M deuterated TMAO. The OTS monolayer appears to be relatively undisturbed by the presence of 6 M TMAO. Last, the spectra shown in c) indicate that no peak is observed when both OTS and TMAO are deuterated and a peak similar to that at air/water interface is observed for 6 M hydrogenated TMAO in perdeuterated OTS.
There were three possibilities accounting for the changes after the introduction of TMAO. First, the increased intensity of the higher frequency peak could be related to a reordering of the OTS monolayer, hence with a larger sum frequency active dipole. Second, the changes could reflect the presence of the osmolyte TMAO (the higher signal directly come from the methyl group of TMAO). Third, a combination of intensity from TMAO and a reordering of the OTS monolayer could both be responsible for the observed changes. To determine which is the best explanation, the OTS/D\textsubscript{2}O spectra were recorded in a 6 M deuterated TMAO solution and compared to the identical monolayer taken in pure D\textsubscript{2}O (no TMAO present) (see Figure 3-5b). As can be seen, the spectra in this case are essentially identical. This confirms the presence of the osmolyte TMAO does not disrupt or reorder the tightly order, well-packed OTS monolayer. Hence the first and the third possibility are ungrounded, making the second one the only reason for the spectra change in Figure 3-5a.

To check this explanation from another aspect, we performed an additional experiment with perdeuterated OTS and hydrogenated TMAO. No signal was obtained in the C-H stretch region when just the perdeuterated OTS monolayer was present at the quartz/OTS/water interface (Figure 3-4c, black spectrum). However, a weak peak near 2960 cm\textsuperscript{-1} was observed when 6 M TMAO solution flowed into the cell (Figure 3-5c, red spectrum). This is consistent with the small rise and blue-shift observed in Figure 3-5a if the two resonances (methyl Fermi resonance from OTS and methyl symmetric stretch from TMAO) have constructive interference (e.g. have the same phase). This can only be the case if the methyl groups on TMAO have the same orientation as the terminal methyl
groups on the OTS monolayer. As such, they must face away from the surface and point towards the aqueous medium as depicted schematically in Figure 3-1.

Another confirmation for the orientation of the methyl groups of TMAO at this interface was done by fitting the data in Figure 3-5a. Both the 2876 cm\(^{-1}\) and 2934 cm\(^{-1}\) resonances from the OTS monolayer possessed a negative sign in agreement with the fact that the terminal methyl groups face toward the aqueous solution. Figure 3-6a shows a fit to the data in which the additional resonance from TMAO at 2960 cm\(^{-1}\) is assumed to be negative (the same sign as the OTS peaks), while Figure 5b shows another fit result when this resonance has the opposite sign (positive). As can be seen, the fit with the negative sign (left panel) is superior. It is in good agreement with the idea that the methyl groups from TMAO face toward the aqueous medium rather than toward the OTS layer.

As a final check, MEM calculations were also performed which confirmed that the imaginary part of \(\chi^{(2)}\) becomes more negative near 2960 cm\(^{-1}\) after 6 M TMAO was introduced to the D\(_2\)O/OTS interface (Figure 3-7). Since the overall Im[\(\chi^{(2)}\)] spectrum is simply the algebraic sum of the Im[\(\chi^{(2)}\)] spectra of the resonant peaks, the two constituents (OTS and TMAO) must have the same sign and lead to an increase in overall magnitude of Im[\(\chi^{(2)}\)].
Figure 3-6. VSFG data in $ssp$ polarization for 4 M TMAO at the OTS/D$_2$O interface. The peak assigned to the symmetric stretches of the TMAO methyl group at $\sim$2950 cm$^{-1}$ is fit with a negative sign in a) and positive sign in b). Once again, the fit corresponding to the methyl groups pointing away from the hydrophobic interface (shown in a)) is preferred.
Figure 3-7. MEM calculations of the imaginary portion of $\chi^{(2)}$ in the presence and absence of 6 M TMAO at the OTS/water interface. The largest difference is observed ~2950 cm$^{-1}$, where the peak due to the symmetric stretch of TMAO is observed. This peak from TMAO overlaps with the Fermi resonance of OTS (~2934 cm$^{-1}$). The two peaks have the same sign and lead to an increase in overall magnitude of Im[$\chi^{(2)}$].
3.4 Discussion

TMAO is often considered to be the quintessential example of a stabilizing osmolyte. Therefore, the molecular level details of its interactions at hydrophobic interfaces may shed light onto the mechanism of protein stabilization. Vanderkooi and coworkers have classified TMAO as a hydrophobic solute.\textsuperscript{110,111} They note that the addition of this osmolyte to aqueous solutions increases the infrared adsorption of the O-H stretch band on the red side of the peak. Moreover, the population of water with less distorted hydrogen bond angles in its first hydration shell is increased, while the population with more distorted hydrogen bonds is decreased. Adding TMAO to water also leads to a positive change in the hydration heat capacity. These properties are classically associated with increasing the “ice-like” properties of bulk water and may play a role in TMAO’s propensity to partition to the air/water interface despite its relatively high solubility in aqueous solution. The current studies show that once TMAO resides at a hydrophobic/aqueous interface, it orients to have its methyl groups facing away from the hydrophobic phase and towards water.

The orientation of TMAO is perhaps somewhat surprising as one might have otherwise expected a hydrophobic interface to attract the methyl moieties of a solute molecule from an aqueous environment. However, TMAO is zwitterionic and the positively charged trimethylammonium moiety partially resembles the tetramethylammonium cation, which like most cations is excluded from the air/water surface.\textsuperscript{112,113} Moreover, the methyl groups on TMAO should not necessarily be considered hydrophobic.\textsuperscript{114} This is because these groups are attached to an electron...
withdrawing substituent. As such, the C-H moieties should be better able to engage in proton donation to surrounding water molecules than a methyl group attached, for example, to a methylene group.\textsuperscript{115} In fact, recent thermodynamic studies of TMAO have determined that TMAO’s methyl groups have hydrophilic rather hydrophobic properties.\textsuperscript{116} Therefore, it is energetically costly to dissociate water from them in order for direct interactions with air or hydrocarbon surfaces to take place.

Of course, it is also energetically costly for the negatively charged oxide moiety of TMAO to face toward the apolar surface. The question is which orientation is the least energetically costly. The current VSFG experiments clearly indicate that an orientation in which the methyl groups face downward into the aqueous solution is more favorable. As such, dehydration of the methyl groups may cost the molecule more potential hydrogen bonds, than loss of water at the oxygen. At any rate, requiring the oxide moiety to point toward the interface should be energetically unfavorable for exposed hydrophobic moieties from proteins and may provide a clue as to why the addition of TMAO to solution helps favor native folded structures.

**Implications for the Stabilization of Protein Structure**

Although proteins contain charged and polar groups, their hydrophobic content are sufficiently large to yield a dielectric constant that is generally considerably lower than the surrounding water. Most estimates of protein dielectric constants at 25°C are between 2 and 20,\textsuperscript{117} which is much less than the value of 78 for water. The dielectric constant of an OTS monolayer and air is approximately 2 and 1, respectively.\textsuperscript{118} Thus, in terms of hydrophobicity and bulk dielectric, both an OTS monolayer and air can resemble
proteins to at least some extent. It is therefore conceivable that TMAO would orient in a similar manner at the more hydrophobic portions of the protein/water interface.

Evidence that osmolyte orientation at hydrophobic interfaces is an important factor in protein stabilization/denaturation behavior comes from the fact that denaturants show a markedly interfacial behavior than TMAO. VSFG measurements of methylated urea compounds, such as tetramethylurea (TMU), at the air/water interface demonstrate that TMU is aligned with its methyl groups pointing towards the hydrophobic surface, not away (data not shown). Moreover, the denaturing efficacy of urea-like compounds scales directly with the hydrophobic content of these molecules.\textsuperscript{119} Thus, the ability of urea-like compounds to denature protein structure directly corresponds to both their hydrophobic content and orientational properties. TMU orientation is in agreement with thermodynamic studies which suggest that its methyl groups are relatively hydrophobic.\textsuperscript{120,121} This is expected, as they are attached to nitrogen atoms which are not quaternary and apparently don’t provide a sufficient electron withdrawing propensity to render the methyl groups hydrophilic. Moreover, molecules similar to TMAO, but possessing more hydrophobic cationic groups, such as triethylamine N-oxide, have less effect on stabilizing protein structure.\textsuperscript{119} These type of osmolytes simply behave like surfactants, which typically denature proteins.\textsuperscript{122,123} Indeed, recent studies have implicated hydrophobicity as being the dominant factor in many biological structuring and chemical processes ranging from micelle formation\textsuperscript{124-126} and enzyme catalysis\textsuperscript{127-129} to protein folding.\textsuperscript{130-132}
Chapter 4

Study of Ionic Binding of Some Alkali Metal Cations to the Carboxylate/Phosphonate Headgroups of Surfactant Monolayers

4.1 Background

Specific ionic effects appear ubiquitously and play an important role in many biological and chemical phenomena such as protein solubility and stability, enzyme activity, phase transition behavior of monolayers and macromolecules, surface tension of aqueous interfaces, etc. Out of all the ion-specific interactions, it is always of particular interest to understand the binding events between the alkali metal cation (M\(^+\)) and the carboxylate (-COO\(^-\)) groups in long-chain fatty acid monolayers because they can serve as a good model system for studying biological relevant interactions such as the selection of Na\(^+\) versus K\(^+\) in ion channels. Potassium and sodium ions are the predominant cationic components of intracellular and extracellular fluid, respectively. How does nature distinguish these ions with similar chemical properties when transporting them across biological membranes? To answer this intriguing question, decades of research on the model system, both theoretical and experimental, has been carried out. Kim Collins and others speculated that Hofmeister chemistry could maintain the ion gradient across cell membranes; moreover, the different binding strength was thought to be related to the law of matching water
affinities (LMWA). \(^{37-39}\) Pavel Jungwirth and co-workers combined molecular dynamic (MD) simulations, X-ray absorption and \textit{ab initio} calculations to determine relative interaction strengths between several cations and carboxylate groups in aqueous solution, and found sodium interacted more strongly than potassium. \(^{147-149}\) Nico F. A. van der Veg\textit{t et al} also performed MD simulation on this subject, and attributed the binding affinity with acetate ions to the preferable formation of weak solvent-shared ion pairs. \(^{150}\)

Recently some second-order non-linear optical techniques with surface specificity and molecular-level sensitivity such as second harmonic generation (SHG)\(^{25,151,152}\) and vibrational sum frequency generation (VSFG)\(^{45,153,154}\) spectroscopy are employed to probe the binding behaviors of different ions in subphases with a variety of charged surfaces or monolayers. For example, Julianne Gibbs-Davis and colleagues exploited second harmonic generation spectroscopy to study the influence of different alkali metal chlorides on the effective acid dissociation constant (\(pK_{a\text{eff}}\)) of the silanol groups at silica/water interface. \(^{155}\) Mischa Bonn group investigated the effect of sodium and calcium ions on the structures of different phospholipid monolayers by VSFG, finding that Ca\(^{2+}\) perturbed the lipid organization significantly while Na\(^{+}\) only showed a subtle effect. \(^{156}\) Marc Gurau \textit{et al}. in our lab got the same conclusion that divalent ions could greatly affect charged monolayers by measuring the VSFG intensity of eicosanoic acid at a low concentration of Mg\(^{2+}\) and Zn\(^{2+}\) salts. \(^{157}\) Xin Chen \textit{et al}. discussed the anion effects on interfacial water structures adjacent to charged protein, lipid and polymer monolayers at different pHs. \(^{158,159}\) Moreover, Sarah Flores \textit{et al}. explored the interactions of several cation chloride salts with highly negatively charged solid (TiO\(_2\) and SiO\(_2\))/aqueous interfaces and concluded that the different orderings of Hofmeister series may result from
disparities in the charge density, polarizability, and basicity of the two oxide surfaces.\textsuperscript{160} Heather Allen \textit{et al} did research on Na\textsuperscript{+} and K\textsuperscript{+} binding to palmitic acid monolayer at air/water interface by VSFG, and claimed K\textsuperscript{+} showed a higher binding affinity to palmitic acid headgroups than Na\textsuperscript{+} due to the formation of an ionic complex of K\textsuperscript{+} and -COO\textsuperscript{-}.\textsuperscript{161} However, this finding is not supported by other groups’ simulation work.\textsuperscript{147-150,162}

In this particular work, we utilized VSFG to systematically investigate the ionic binding events of biologically relevant cations Na\textsuperscript{+} and K\textsuperscript{+} as well as Li\textsuperscript{+} to some long-chain fatty acids, including palmitic acid (PA, C\textsubscript{15}H\textsubscript{31}COOH), stearic acid (SA, C\textsubscript{17}H\textsubscript{35}COOH) and eicosanoid acid (EA, C\textsubscript{19}H\textsubscript{39}COOH) in order to get a better understanding of the ion-selective interactions with carboxylate headgroups. By interpreting the VSFG intensity spectra, we found sodium ion, instead of potassium ion, is more effective in binding to the negatively charged carboxylate moieties of the fatty acids. Lithium binds the tightest, suggesting ion-pairing affinities can be affected by surface charge densities. There is no obvious difference between the three ions when the headgroups are protonated. This is in partial agreement with the law of matching water affinity but further explanations need to be addressed for the tightest binding lithium ion. In the meanwhile, we also studied the binding behaviors of these cations to phosphonate headgroups at air/aqueous interface and found different binding events existed: lithium will bind to the charged headgroups quite tightly while there is no distinguishable difference between sodium and potassium. Such findings should help to elucidate the molecular-level binding behavior to proteins in aqueous solutions.
4.2 Experimental Section

Preparation of high purity salts solutions

LiCl (99.995% trace metals basis, Alfa Aesar), NaCl (99.999% trace metals basis, Aldrich), KCl (99.999% trace metals basis, Fluka, Sigma-Aldrich) were used in this study. All glassware was soaked in concentrated H₂SO₄ followed by rinsing thoroughly with ultra-pure water before use. The salt solutions with different concentrations are obtained by dissolving proper amount of dry inorganic salts in fresh ultra-pure deionized water with a minimum resistivity of 18.2 MΩ·cm from NANOpure Ultrapure Water System (Barnstead, Dubuque, IA).

The desired pH value was attained by adding an appropriate amount of 0.1 mol/L NaOH or 0.1 mol/L HCl solution (also prepared with ultra-pure deionized water). To avoid contamination from heavy-metal ions, no other salt was added to keep the ionic strength constant.

Formation of Langmuir monolayers

Eicosanoic acid and palmitic acid were purchased from Sigma-Aldrich (powder, ≥99%). Stearic acid sodium salt was purchased from TCI, America (powder, ≥98%). N-Octadecylphosphonic acid [CH₃(CH₂)₁₇P(O)(OH)₂] was purchased from Alfa Aesar (powder, ≥97%). All products were used as received. A small Langmuir trough (model 601M, NIMA, U. K.) was used as a sample stage. To achieve a stable Langmuir monolayer over the subphase, several droplets of fatty acid solution in chloroform (~10 μL depending on the molecular weight of the surfactant with a concentration ~2 mg/mL)
were added to the top of a 35 mL subphase in the trough by a micro-syringe. The surface pressure was then adjusted to 20 mN/m by gently compressing two barriers, which was measured with a paper Wilhelmy plate and electronically controlled by KSV-NIMA software during the whole measurement. The adopted surface pressure was to ensure the surfactant exist in a liquid condensed phase. An equilibration time of ten minutes was allotted for complete chloroform evaporation and monolayer formation. Longer waiting time had no obvious evidence of spectral changes within experimental errors. The fatty acid monolayer was quite stable for many hours without major disruption or dissolution of surfactant. The same procedure was followed to make the phosphonate monolayer. Note the phosphonic acid cannot be dissolved thoroughly in pure chloroform so the solvent is a mixture of methanol and chloroform with V/V=1:1. This will require a longer stabilizing time for the formation of the monolayer which is not very stable and some appreciable dissolution are observed during the experiment. Each VSFG spectrum was taken on a freshly prepared monolayer on a subphase (ultra-pure water or salt solution) of appropriate pH.

**Vibrational Sum Frequency Generation Spectroscopy**

The VSFG system which was employed in these studies has been described in detail in the Chapter 2. The polarization combination used in all the experiments was $ssp$ unless otherwise noted.
4.3 Results and Discussion

Figure 4-1 shows the experimentally recorded ssp VSFG spectra of air/eicosanoic acid (C\textsubscript{19}H\textsubscript{39}COOH) monolayer/pure water interface (black trace) and air/eicosanoic acid monolayer/50 mM NaCl salt solution interface (red trace) at bulk pH 11. There are several prominent and characteristic features in these spectra. Deduced from the acid monolayer/pure water spectrum, it is obvious that the C-H stretching regions (2800 ~ 3000 cm\textsuperscript{-1}) have two major peaks centered at 2875 cm\textsuperscript{-1} and 2940 cm\textsuperscript{-1}, which have been thoroughly studied and assigned as the symmetric stretch of terminal methyl groups (CH\textsubscript{3,ss}) of the fatty acid for the lower frequency one and a methyl Fermi resonance for the higher frequency one (CH\textsubscript{3,FR}).\textsuperscript{101,164} Additional but relatively tiny feature is also found around 2840 cm\textsuperscript{-1}, corresponding to the CH\textsubscript{2} symmetric stretch (CH\textsubscript{2,ss}) due to Gauche defect.\textsuperscript{159} The overall spectra features (intensities and wavenumbers) indicate a high degree of ordering of the fatty acid molecules at the air/water interface, with their hydrophilic carboxylic acid headgroups facing aqueous medium and hydrophobic alkyl chain (tails) pointing up towards the air, just as the cartoon shown in the upper panel of Figure 4-2.
Figure 4-1. VSFG spectra in ssp polarization combination of air/eicosanoic acid (C\textsubscript{19}H\textsubscript{39}COOH) monolayer/pure water interface (black trace) and air/eicosanoic acid monolayer/50 mM NaCl salt solution interface (red trace). Both spectra were taken at bulk pH 11 to make the acid headgroups fully deprotonated.
Figure 4-2. Cartoon depiction of the ionic binding events. The long chain fatty acid can form a good Langmuir monolayer over the aqueous subphase due to its amphiphilic property. Closely packed (condensed) phase for the surfactants was adopted and maintained during the whole time. The red and blue dots represent O and H atoms in water molecules and acid headgroups, while the yellow dots stand for monovalent ions (Li⁺, Na⁺ and K⁺).
The peak assignment in the O-H stretching regions (3000 ~ 3800 cm\(^{-1}\)) is ambiguous since there exist different sources of the O-H oscillators as well as intermolecular and intramolecular coupling of the two O-H bonds in H-O-H. Besides, the width of O-H peaks are usually broad which would overlap with each other and the interferences between neighboring resonances and interferences between resonances and non-resonant background can give rise to intensity cross terms and spectral distortion, imposing more uncertainty to line shapes deconvolution.\(^{43,165,166}\) To answer this question, we must first know what the sources of the O-H oscillators are, i.e. whether the O-H moieties in the fatty acid headgroups can generate VSFG signal at this region. Various research groups have reported the surface pK\(_a\) of long-chain fatty acids are relatively larger than that in bulk aqueous solution, although a definite value are not reached (only a single pK\(_a\) was reported in each system, usually between 5.5 and 9).\(^{167-171}\) Under our experimental condition (pH 11), no matter how the real surface pK\(_a\) varies, the majority of the carboxylate headgroups are deprotonated (existing in the form of -COO\(^-\)). The lack of O-H oscillators in the headgroups hence cannot contribute to the O-H intensities in our measured spectra. In other words, the O-H bands can be primarily, if not solely, attributed to different interfacial water structures lying just below these charged headgroups. It was formerly believed that basically two broad bands exist in this O-H stretching region of the air/fatty acid monolayer/water interface, namely “ice-like” peak (broad, centered around 3200 cm\(^{-1}\)) and “liquid-like” peak (also broad, centered around 3450 cm\(^{-1}\)) because these two bands resemble the infrared O-H stretches of the ice and liquid water in bulk. The “ice-like” peak represents those symmetric tetrahedrally coordinated/hydrogen-bonded water species while the “liquid-like” one is less coordinated, or in a distorted tetrahedral
Recent isotopically diluted water study with heterodyne-detected VSFG (HD-VSFG) conducted by Tahara group suggested that at the highly charged interface, the two O-H bands found in VSFG spectra have the same structural origin (hydrogen-bonded environment essentially similar as in the bulk water but no tetrahedrally coordinated “ice-like” water), and they are possibly split by intramolecular and/or intermolecular anharmonic couplings of H-O-H vibrations.

Since the carboxylate headgroups of the fatty acid are deprotonated and bearing negative charges, they would introduce a local electric field within the interfacial layers. The water molecules nearby thus would be expected to be aligned by the surface electric field with their hydrogen atoms pointing “up” on average (transition dipoles of H-O-H pointing towards the surface) and become highly ordered. This can explain the huge intensity enhancement in the O-H region than that of neat water/air interface (compare to the pure air/water spectrum in Figure 4-3).
Figure 4-3. The comparison between the VSFG intensities of air/pure water interface and air/fatty acid monolayer/water interface. The green dots are the air/pure water spectrum magnified by a factor of ten for clearer view about the relative low intensity.
As the subphase was replaced with 50 mM NaCl salt solution, we could clearly see that the intensity of C-H peaks are nearly unchanged while the O-H intensities are greatly reduced. The nearly identical C-H peak positions and intensities suggest that the introduction of the Na\(^+\) and Cl\(^-\) ions does not disrupt the fatty acid monolayer arrangement since the C-H features are indicators of the overall ordering of the surfactants. The distinct disparity for the salt solution subphase is that both of the O-H bands are quenched to a certain degree. It is therefore reasonable to conclude that the existence of the cations would disrupt the interfacial water’s preferential orientation by binding themselves to the negatively charged headgroups. Through this “charge screening” event, the local electric filed strength is decreased (binding products are electric neutral), and its ability to align water molecules nearby is also reduced.\(^{179}\) Apparently, if a certain kind of cation binds stronger to the charged headgroups than other cations, more binding sites (-COO\(^-\)) will be occupied under the same ion concentration, and the surface charge neutralization effect at interfacial region would become more obvious, which will in turn quench the O-H intensity to a larger extent.

Within these ideas in mind, we continued to investigate three alkali metal salts with the same anion (LiCl, NaCl and KCl) at pH 11. To check and exclude the possible organic contamination in the high purity salt, we firstly took the VSFG spectra in the C-H stretching region (2800 ~ 3000 cm\(^{-1}\)) of pure inorganic salt solutions (no surfactant added) after filtration. All the C-H spectra of different types and concentrations of these salts solution are flat (c.f. Figure 4-4), indicating no organic contamination are present in our prepared solutions since the appearance of C-H peaks reveal the existence of organic species.
Figure 4-4. The VSFG spectra at the C-H stretching region for pure salt solutions (100 mM NaCl and 100 mM KCl are selected as examples). The nearly flat C-H intensities indicate no organic contamination existed in those solution.
Next we implemented the investigation with several concentrations of pure salt solutions covered by an eicosanoic acid ($C_{19}H_{39}COOH$) monolayer with exactly the same surface pressure (20 mN/m) and the results are shown in Figure 4-5 (results with 200 mM concentration are chosen as an instance). It also includes the spectrum of ultra-pure water subphase with the monolayer as a reference.
Figure 4-5. VSFG experiments of ecosanoic acid (C$_{20}$) monolayer with different subphases (ultrapure water and alkali metal chloride solutions). The salts solution subphases are all of the same concentration to ensure the identical ionic strength and available binding species, which help to elucidate the binding affinity to the charged headgroups.
According to the spectra, the VSFG intensities of the O-H stretching region are in the following order:

Ultra-pure water > KCl > NaCl > LiCl

So the intensity attenuation amount as well as the binding affinity of the cations to the negatively charged -COO\(^-\) are in the reverse order:

\[ K^+ < Na^+ < Li^+ \]

The next sets of experiments were conducted with different concentrations of the same electrolyte solution. Here we presented the results of KCl as an example. VSFG spectra were taken at four concentrations of KCl solution: 50 mM, 100 mM, 200 mM and 500 mM, all under pH 11, just as Figure 4-6 depicts. From the normalized intensities of these spectra, it is clear that together with the increase in ion concentration, a progressive reduction of the overall spectral intensity is found in the O-H stretching region. This is also in good accordance with our assumption: with higher ion concentrations, more and more negatively charged sites (-COO\(^-\)) are bonded by the metal cations to form a charge neutral product. This binding event will neutralize the surface charge and consequently induce the gradual intensity drop.
Figure 4-6. Gradual quenching effect of the signal of the O-H stretching region as the electrolyte concentration increases. 50 mM, 100mM, 200mM and 500 mM of KCl salt solutions were used as the subphases. All other experimental conditions were identical.
One may also question whether the alkyl chain length would play a role in the binding events. We repeated the same VSFG measurements with surfactants of different hydrophobic hydrocarbyl groups: palmitic acid (C_{16}) and stearic acid (C_{18}) Langmuir films. The order of O-H region quenching level was found to be the same as in the case of eicosanoic acid (C_{20}), i.e. K^+ < Na^+ < Li^+.

For now we have discussed the binding ability of the ions to the carboxylate headgroups when they are deprotonated (-COO^-). One may wonder what the binding trend would be if the headgroups are charge neutral ones (-COOH). Will the ions also show the specific binding with the protonated form? By varying the subphase pH, it is quite easy to switch between the deprotonated and protonated headgroups of the fatty acid surfactants. Firstly we examined the carboxylate anion and carbonyl stretch region at different pHs: pH 4, pH 7 and pH 11 to confirm that the adjustment of subphase (bulk) pH did affect the surface protonation/deprotonation degree. Note if the deprotonated anions are the predominant form of the monolayer’s headgroups, the symmetric stretch of carboxylate anions (-COO^-) would appear near 1410 cm\(^{-1}\) as a single strong symmetric peak.\textsuperscript{161} On the other hand, if most of the headgroups are protonated, there would be no such carboxylate stretch. Instead we would observe the carbonyl stretch (C=O bond) in -COOH. This vibrational mode is centered at 1710 cm\(^{-1}\).\textsuperscript{161} To get rid of C-H bending mode (also lies in this region), we adopted D$_2$O as the subphase. From the VSFG spectra of the carboxylate anion stretching region (1400 ~ 1600 cm\(^{-1}\)), we could see that there don’t exist any bands under pH 4. On the other hand, a carbonyl stretch peak is obvious. Hence it is safe to say at pH 4, nearly 100% of the carboxylic acid headgroups are protonated/charge neutral. Suppose the ions also follow a certain binding trend, we may
find distinct intensities in the O-H region as well: the presence of excess cations in bound complexes now act as the charge species, and the hydrogen-bonded water molecules in the interfacial layers reorganize, giving rise to the SF intensities. However, that is not so. The VSFG results of four subphases at pH 4 are presented in Figure 4-7. Within the statistical accuracy, no differences can be observed between the three cations as well as the pure water. The findings thus reveal that the three ions have no appreciable binding affinity with the charge neutral -COOH group.
Figure 4-7. VSFG spectra of ecosanoic acid (C$_{20}$) monolayer spread over ultra-pure deionized water (black squares) and with the presence of 100 mM alkali metal halide salts (LiCl, NaCl and KCl) under pH 4. At the acidic pH, the acid headgroups are fully protonated and the similar intensities in the O-H regions suggest no appreciable binding may occur to the charge neutral -COOH groups of the acid.
Divalent metal ions (M$^{2+}$) species has been proved to be able to bind quite strongly to the -COO$^-$ groups of eicosanoic acid even at a rather low concentration (below millimolar).\textsuperscript{151} We also examined this divalent metal ion effect by adding 100 mM and 0.1 mM (100\% and 0.1\% of the concentration of the contrast M$^+$ ion) CaCl$_2$ into the ultrapure water and recorded VSFG spectra respectively. As illustrated in Figure 4-8, the VSFG spectra of the fatty acid monolayers changed significantly in the presence of mere 0.1 mM Ca$^{2+}$ in the subphase. The intensities of the O-H stretching region of both spectra are attenuated almost to the bottom, which was never seen before with the monovalent ions even at 1 M concentration. The findings are consistent with the idea that the divalent metal cations (in this case, Ca$^{2+}$) can bind extremely tightly to the negatively charged carboxylate anions. Even at a rather low concentration they are still able to make the headgroups fully neutralized and cause the interfacial water molecules to be highly randomly distributed (with no preference for H-up or H-down configurations), resulting in low to almost none O-H intensities.
Figure 4-8. Divalent ion effect. Very low concentrations of calcium ions were introduced and found to be extremely effective in quenching O-H signals.
To get a clearer view of the sharp contrast to the monovalent ions, we specifically prepared a solution of 100 mM NaCl mixed with 0.1 mM CaCl₂. If, by any chance, the two cations possess similar binding affinity in this competing environment, it is natural to predict the VSFG spectrum of this mixture would be much alike the spectrum of 100 mM NaCl subphase in consideration of the fact that the concentrations of Na⁺ is 1000 times higher than Ca²⁺. Obviously this assumption is not true according to Figure 4-9. There is no surprise to us that the acquired data showed the reduction of intensities in the O-H region resembles the spectrum of the low concentration Ca²⁺, instead of the spectrum of the much more concentrated, 100 mM Na⁺. To sum up, Ca²⁺ does act as a strong binding ion to the carboxylate in respect to the monovalent K⁺/Na⁺/Li⁺, so it is capable of neutralizing surface charge at a much lower concentration.
Figure 4-9. Control experiment of a solution containing 100 mM NaCl and 0.1 mM CaCl₂. Both inorganic salts are with highest purity available (99.999%).
The ionic binding sequence could be varied if the carboxylate are substituted with other charged headgroups. We performed the measurements of phosphonate (-PO$_3^{2-}$) and found the results particularly intriguing. N-Octadecylphosphonic acid was used as the surfactant film and all experiments were conducted at pH 10.4 to get the negatively charged form. The ratio of -1 and -2 species are unknown since no clear surface pK$_a$ of phosphonate compounds are reported by far but pH 10.4 is enough to maintain the local electric field of negatively charged surface, as confirmed by pH dependent VSFG spectra (data not shown here). Although the hydrophilic part of the surfactants are changed, the observed large reduction of O-H stretching band can still be mainly attributed to the change of local density of aligned water molecules confined in between charged monolayer and counterions. The sequence of binding ability differs significantly from the one reported above: no distinguishment are present between the sodium and potassium ions. However, the attenuation effect for lithium is quite astonishing: the intensities of O-H stretching bands almost decrease to the bottom, as the divalent ion does in the carboxylate case. This different binding behavior will be discussed later in this text. Other headgroups such as phosphate and sulfate are to be investigated and will be presented in the future study.
Figure 4-10. VSFG experiments of an octadecylphosphonate monolayer with different subphases (ultra-pure water and alkali metal chloride solutions) at two concentrations. Both spectra reveal the unidentifiable binding strength of Na\(^+\) vs K\(^+\).
How should we rationalize the rank of binding abilities of these metal ions to the charged groups? The classic treatment is to check with the law of matching water affinity (LMWA).\textsuperscript{37-39} Kim Collins proposed this empirical law in order to better understand the underlying mechanism of ion-pairs formation/ionic binding events, focusing on the hydration energy of ions. It basically states that the relative affinity of ions in solution depends on the matching of their hydration enthalpies. In other words, contact ion pairs are much more easily formed between two ions with similar hydration enthalpies than those ions with relatively larger hydration enthalpies difference. In the present case, carboxylate moieties attached to small molecules, peptides, polymers and proteins would be expected to have higher affinity toward similarly hydrated cations on the basis of the LMWA. It would be expected that the Na\textsuperscript{+}-COO\textsuperscript{-} binding pair would be the tightest because Na\textsuperscript{+} has a hydration enthalpy of -416 kJ/mol, very close to the value of carboxylate group (-425 kJ/mol).\textsuperscript{180} More weakly hydrated monovalent cations like K\textsuperscript{+} (hydration enthalpy -334 kJ/mol)\textsuperscript{180} should pair with carboxylate to a lesser extent than Na\textsuperscript{+}.

However, there does exist an exception: lithium. According to its hydration enthalpy (-531 kJ/mol),\textsuperscript{180} it would form a less tight ion pair with carboxylate than sodium ion. Nevertheless, the spectroscopic results suggested a different trend with Li\textsuperscript{+} binding the tightest (K\textsuperscript{+} < Na\textsuperscript{+} < Li\textsuperscript{+}). Therefore we need to seek for extra explanation other than LMWA for its strongest binding ability with carboxylate. Li\textsuperscript{+} is the smallest monovalent metal cation and binds solvent water molecules more tightly than any other monovalent cations. As a consequence, it is relatively harder for lithium ion to get rid of the surrounding water molecules in its hydration shell and form a contact ion pair (CIP)
with the negatively charged carboxylate groups. Instead it may be more favorable to participate in forming a solvent-separated or solvent-shared ion pair (SIP). This deduction is partially supported by computational calculations. Berk Hess and Nico F. A. van der Vegt performed MD simulations with quantitatively accurate models acetate (mimicking carboxylate groups on protein surfaces). What they found is only very few CIPs are present in 1 m LiAce solution (0.02 per ion) while SIPs are predominantly observed. In these SIPs, water oxygens strongly coordinate/bind the positive lithium ion while the hydrogens in these surrounding water molecules are repelled from the hydration shell and form linkages with hydrogen-bonding acceptors, in this case, acetates, in the vicinity. In summary, the order $K^+ < Na^+ < Li^+$ of increasing binding affinity with carboxylate ions is primarily caused by a stronger preference for forming solvent-shared ion pairs.

This trend also holds true for other systems with similar carboxylate groups of the side chains of peptides at the air/water interface. For example, Jaibir Kherb in our group explored the Hofmeister series of cations with an elastin-like polypeptide (ELP) containing 16 aspartic acid residues. By measuring the phase transition temperature (lower critical solution temperature, LCST) and VSFG spectra of this biopolymer with monovalent cations, it is clear that lithium ions showed a relatively strong association to the carboxylate sites in respect to sodium and potassium ions.

It should be noted that Professor Allen’s group conducted similar research also by VSFG spectroscopy yet got a different result: it looks like there is a stronger preference of $K^+$ binding to the carboxylate groups of palmitic acid with respect to $Na^+$. This controversial finding naturally lead to a totally different explanation on the binding behaviors of sodium and potassium ions. However, improper use of the experimental
material may hurt the validity of this conclusion and discussions in turn. ACS grade salt, instead of ultra-pure (99.999% purity) salt solutions were used as the subphases in that work. Normally ACS grade salt of alkali metal chloride could contain 1% contamination of other metals, especially the divalent alkaline-earth metal ions such as Ba$^{2+}$, Ca$^{2+}$ and Mg$^{2+}$. Previous data in our group and this work have already shown the much stronger binding affinity of divalent ions to the carboxylate, which could drastically alter the interfacial water alignment even at micromolar concentration. Hence the effect of trace amount of divalent metal ions in ACS grade salts should not be neglected and the original results need to be reexamined. Later Huang and Allen et al. probed this problem and did see the significant influence of salt purity on the interactions between sodium and the carboxylate headgroups.\textsuperscript{182} Without knowing the exact contents in NaCl and KCl ACS grade salts, it is unsafe to use them in the ionic binding study and cannot draw any valid conclusion. In other words, only ultra-pure grade salts are allowed in the investigation of topics related to ion specific effect, otherwise contamination would ruin any discovery.

Recently, Tahara \textit{et al} implemented heterodyne-detected VSFG technique to examine the counter ion effect on interfacial water structures at positively charged cetyltrimethylammonium bromide (CTA$^+$Br$^-$)/electrolyte solution interface and negatively charged sodium dodecyl sulfate (Na$^+$DS$^-$)/electrolyte solution interface.\textsuperscript{183} For the CTA$^+$ aqueous interface, the intensity of O-H bands of the interfacial water molecules varied in the order of Hofmeister anion series (F$^-$ > Cl$^-$ > Br$^-$ > I$^-$), indicating direct interaction/contact ion pair formation of the large halide ions with trimethylammonium headgroups, which is in agreement with one of our group’s previous VSFG study.\textsuperscript{33} In contrast, for the DS$^-$ interface, the intensity depend on the cations only weakly. The
hydrogen-bond strength (the median frequency of the O-H stretch), on the other hand, correlate well with the cation effect. This means the effect of the cation on the water structure reaches beyond its first solvation shell, i.e. interacting through solvent separated ion-pairing. The findings are quite informative on the ion-pairing study: by changing the charged headgroups, the binding event may follow different mechanism (non-contact vs. contact adsorption). Similar to their model system of sulfate, the phosphonate headgroups in our experiments are also considered to be strongly hydrated (comparing to carboxylates) and therefore tend to the formation of the solvent shared/separated ion-pairs. This indirect interaction through changing of hydrogen bonding in the process eventually affects the attenuation amount in the O-H region, especially for the strongly hydrated lithium ions. Varying headgroups will affect the ratio of contact and solvent shared/separated ion pairs, and this may ultimately explain the Li$^+$ ion’s somewhat anomalous behavior in the cases tested. Although we may draw some empirical conclusion with the current results of VSFG intensity spectra, further spectroscopic studies and quantitative calculations by MD are required to probe this subject in more detail.
### 4.4 Conclusion

Ionic binding between the alkali metal cation and the fatty acid headgroups represents a vital ion specific interaction. To get a deeper insight into this binding phenomenon, we examined the interfacial water organization at air/surfactant monolayer/water interface by VSFG spectroscopy, with the variation of subphase of alkali metal chloride salts solutions. At high pH values, the -COOH headgroups of the fatty acid should be significantly deprotonated, leaving the monolayer negatively charged. The interfacial water is then preferentially oriented according to the direction of the electric field created by the charge of the surfactant headgroups. The introduction of different metal cations has a charge-screening effect, either through direct binding to the carboxylate headgroups or indirect interaction via hydrogen bonding, and induce the perturbation of the O-H oscillators in vicinity. By comparing and interpreting the intensities of the O-H stretching region and correlate the amount of intensity changes with binding affinity, the ionic interactions to carboxylate are found to increase in the sequence:

\[ \text{K}^+ < \text{Na}^+ < \text{Li}^+ \]

\(\text{Na}^+\) shows a stronger binding affinity to the fatty acid than \(\text{K}^+\) under basic pH. This trend is in good accordance with the findings by other research groups obtained with calculations and X-ray absorption spectroscopy.

At low pHs where carboxylic headgroups exist in the form of -COOH, all the spectra of the salts are essentially identical within experimental error. This means that fatty acids with protonated carboxylate sites lead to essentially no cation specific
differences in the salt concentration range employed. Our data also demonstrated that
divalent cations have a much greater impact on the interfacial hydrogen bonding network
than the monovalent cations, indicating a much larger binding ability at least for the case
of long-chain fatty acid monolayers.

Finally, we mentioned the influence of the headgroups of the surfactants. Different choices of the anionic headgroups may allow distinct binding mechanisms. The phosphonate headgroups, for instance, are strongly hydrated and therefore tend to avoid contact ion pairs but favor the formation of the solvent shared/separated ion-pairs, which suits the lithium hydration coordination best. Further studies are undergoing in purpose of determining the preference of contact ion pairs or solvent shared ion pairs in other systems such as the phosphate groups in phospholipids. The present work may help provide information about how cations interact with more complex biological systems.
Chapter 5

Implications and Conclusions

The work presented in this dissertation was motivated by the growing interest for a thorough understanding of the molecular organization (orientation and interaction) at the biological relevant interfaces. For this aim, vibrational sum frequency generation (VSFG), a surface specific nonlinear spectroscopic technique, is utilized to investigate a variety of interfacial phenomena. Owing to the complexity of interfacial systems, especially those involving macromolecules like polymers, peptides and proteins, it is essential to select proper and simple systems as starting points in research before elucidating more intricate processes.

TMAO is stereotyped as a protein stabilizer that is directly related to osmolytes’ stabilization/denaturation effect. The two aqueous/hydrophobic interfaces we use in this study can resemble proteins surfaces to at least some extent in terms of hydrophobicity.

The tremendous versatility of VSFG as surface analytical tools has made the orientation investigation possible. Besides, MEM calculations allow for direct interrogation of the average orientation of the transition dipole moment of interfacial species. What we found is counter-intuitive in the way that methyl groups of the organic compounds are normally thought to be hydrophobic and would prefer to be dehydrated due to hydrophobic interactions. The hydrophilic nature of these special methyl groups may be attributed to their attachment with an electron withdrawing group. Thus, the data presented herein indicate that the efficacy of the powerful protecting osmolyte, TMAO, may lie in its specific unfavorable interactions with hydrophobic groups on proteins.
Another focus lies on the fundamental understanding of ion pairing between ions and different kinds of counterions at the air/surfactant/aqueous interface. Surfactant monolayers may serve as a step toward better understanding the mechanism of cell membranes in many aspects. For examples, the interfacial water structures have always been important because of the involvement of water in many processes happened at the biological interfaces. Upon binding with the charged monolayer species, it necessarily change the arrangement of interfacial water molecules around the binding sites, which makes them a good indicator for the ion-specific interactions. In light of this, the work presented in this dissertation was directed towards the detailed quantitative information of ion-pairing affinities between carboxylate groups of fatty acids in a Langmuir film and several cations. The found sequence of the binding affinities for the monovalent ions, with lithium binding tightest, contradicts the one predicted by empirical law of matching water affinities. Moreover, once the negative charges on the surface are neutralized, much weaker salt interactions take place. On the other hand, divalent ions show much stronger binding affinities to the carboxylate as compared to less charged monovalent ions. By changing the charged headgroups, such as the of phosphonate group C-PO(OH)₂, we could clearly see that a different trend is followed that one can hardly differentiate sodium and potassium. In fact, currently we are still studying the interactions of phosphate groups C-OPO(OH)₂ in a lipid monolayer with these cations. The results, however, would be brought back to the one found with fatty acids, in sharp contrast to the phosphonate study. In summary, charge of the surfaces, charge densities and polarizabilities of the ions, chemical identity of the charged groups, and complexation ability of different metal ions could play a vital role in determining this significant
preference. With these ideas in mind, these opposite findings would suggest that the true nature of ionic binding in real biological systems, containing multiple ligands and various residues, is not as simple as ligand-specific ionic interactions.
Reference


Appendix A The Matlab Scripts Used for VSFG Spectra Fitting

Fit.m

pi=3.1415926;

data=
[
3798 0.240818896
3795 0.289624604
3792 0.284674714
3789 0.239369382
3786 0.298178082
3783 0.235603156
3780 0.301238034
3777 0.336163088
3774 0.274765075
3771 0.295794835
3768 0.290229303
3765 0.287686823
3762 0.29973856
3759 0.318571003
3756 0.278543793
3753 0.228209021
3750 0.23306272
3747 0.233423541
3744 0.199231091
3741 0.29612128
3738 0.269630131
3735 0.259312884
3732 0.325072325
3729 0.378637998
3726 0.3036391
3723 0.30452429
3720 0.302338197
3717 0.384845878
3714 0.328595245
3711 0.310191956
3708 0.325693375
3705 0.351925626
3702 0.372841242
3699 0.436909913
3696 0.50079855
3693 0.380827461
3690 0.343700572
3687 0.371925382
3684 0.45370163
<table>
<thead>
<tr>
<th>Value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>3475</td>
<td>2.238396313</td>
</tr>
<tr>
<td>3470</td>
<td>2.638341718</td>
</tr>
<tr>
<td>3465</td>
<td>2.474302928</td>
</tr>
<tr>
<td>3460</td>
<td>2.30728924</td>
</tr>
<tr>
<td>3455</td>
<td>2.66307899</td>
</tr>
<tr>
<td>3450</td>
<td>2.476959664</td>
</tr>
<tr>
<td>3445</td>
<td>2.693399626</td>
</tr>
<tr>
<td>3440</td>
<td>2.767810457</td>
</tr>
<tr>
<td>3435</td>
<td>2.855902474</td>
</tr>
<tr>
<td>3430</td>
<td>2.853543386</td>
</tr>
<tr>
<td>3425</td>
<td>3.141275019</td>
</tr>
<tr>
<td>3420</td>
<td>2.819618134</td>
</tr>
<tr>
<td>3415</td>
<td>2.933226717</td>
</tr>
<tr>
<td>3410</td>
<td>3.002687051</td>
</tr>
<tr>
<td>3405</td>
<td>3.18669098</td>
</tr>
<tr>
<td>3400</td>
<td>3.236410344</td>
</tr>
<tr>
<td>3395</td>
<td>3.082438336</td>
</tr>
<tr>
<td>3390</td>
<td>3.009195812</td>
</tr>
<tr>
<td>3385</td>
<td>2.973817429</td>
</tr>
<tr>
<td>3380</td>
<td>2.895873683</td>
</tr>
<tr>
<td>3375</td>
<td>2.878857487</td>
</tr>
<tr>
<td>3370</td>
<td>3.112071882</td>
</tr>
<tr>
<td>3365</td>
<td>2.910609056</td>
</tr>
<tr>
<td>3360</td>
<td>3.244247313</td>
</tr>
<tr>
<td>3355</td>
<td>3.095914489</td>
</tr>
<tr>
<td>3350</td>
<td>3.029554833</td>
</tr>
<tr>
<td>3345</td>
<td>3.048427519</td>
</tr>
<tr>
<td>3340</td>
<td>3.178802976</td>
</tr>
<tr>
<td>3335</td>
<td>2.883799936</td>
</tr>
<tr>
<td>3330</td>
<td>2.939752677</td>
</tr>
<tr>
<td>3325</td>
<td>3.189378271</td>
</tr>
<tr>
<td>3320</td>
<td>3.202809756</td>
</tr>
<tr>
<td>3315</td>
<td>3.050649106</td>
</tr>
<tr>
<td>3310</td>
<td>2.944645173</td>
</tr>
<tr>
<td>3305</td>
<td>3.0972291</td>
</tr>
<tr>
<td>3300</td>
<td>3.04628647</td>
</tr>
<tr>
<td>3295</td>
<td>3.069882537</td>
</tr>
<tr>
<td>3290</td>
<td>3.460492262</td>
</tr>
<tr>
<td>3285</td>
<td>3.137685859</td>
</tr>
<tr>
<td>3280</td>
<td>3.38032867</td>
</tr>
<tr>
<td>3275</td>
<td>3.108656918</td>
</tr>
<tr>
<td>3270</td>
<td>3.220438235</td>
</tr>
<tr>
<td>3265</td>
<td>3.290080937</td>
</tr>
<tr>
<td>3260</td>
<td>2.961776526</td>
</tr>
<tr>
<td>3255</td>
<td>3.018523868</td>
</tr>
<tr>
<td>3250</td>
<td>2.966749259</td>
</tr>
<tr>
<td>3245</td>
<td>3.082131963</td>
</tr>
<tr>
<td>3240</td>
<td>3.095695455</td>
</tr>
<tr>
<td>3235</td>
<td>3.40882107</td>
</tr>
<tr>
<td>3230</td>
<td>3.059691898</td>
</tr>
<tr>
<td>3225</td>
<td>3.10386702</td>
</tr>
<tr>
<td>3220</td>
<td>3.121395132</td>
</tr>
<tr>
<td>Value</td>
<td>Decimal Value</td>
</tr>
<tr>
<td>-------</td>
<td>---------------</td>
</tr>
<tr>
<td>3215</td>
<td>2.969411206</td>
</tr>
<tr>
<td>3210</td>
<td>3.015784737</td>
</tr>
<tr>
<td>3205</td>
<td>2.820181369</td>
</tr>
<tr>
<td>3200</td>
<td>2.821413328</td>
</tr>
<tr>
<td>3195</td>
<td>2.694434017</td>
</tr>
<tr>
<td>3190</td>
<td>3.137739431</td>
</tr>
<tr>
<td>3185</td>
<td>2.865847323</td>
</tr>
<tr>
<td>3180</td>
<td>2.718523799</td>
</tr>
<tr>
<td>3175</td>
<td>2.513132965</td>
</tr>
<tr>
<td>3170</td>
<td>2.613074944</td>
</tr>
<tr>
<td>3165</td>
<td>2.538996432</td>
</tr>
<tr>
<td>3160</td>
<td>2.158763318</td>
</tr>
<tr>
<td>3155</td>
<td>2.502015702</td>
</tr>
<tr>
<td>3150</td>
<td>2.363443368</td>
</tr>
<tr>
<td>3145</td>
<td>2.403587516</td>
</tr>
<tr>
<td>3140</td>
<td>2.738769801</td>
</tr>
<tr>
<td>3135</td>
<td>2.568618163</td>
</tr>
<tr>
<td>3130</td>
<td>2.421555284</td>
</tr>
<tr>
<td>3125</td>
<td>2.574552554</td>
</tr>
<tr>
<td>3120</td>
<td>2.416064697</td>
</tr>
<tr>
<td>3115</td>
<td>2.583130385</td>
</tr>
<tr>
<td>3110</td>
<td>2.493822344</td>
</tr>
<tr>
<td>3105</td>
<td>2.282601107</td>
</tr>
<tr>
<td>3100</td>
<td>2.053732555</td>
</tr>
<tr>
<td>3095</td>
<td>2.045692899</td>
</tr>
<tr>
<td>3090</td>
<td>2.209430907</td>
</tr>
<tr>
<td>3085</td>
<td>2.116262486</td>
</tr>
<tr>
<td>3080</td>
<td>1.745272542</td>
</tr>
<tr>
<td>3075</td>
<td>2.005543923</td>
</tr>
<tr>
<td>3070</td>
<td>1.76649443</td>
</tr>
<tr>
<td>3065</td>
<td>1.723536589</td>
</tr>
<tr>
<td>3060</td>
<td>1.686850163</td>
</tr>
<tr>
<td>3055</td>
<td>1.806762185</td>
</tr>
<tr>
<td>3050</td>
<td>1.521341612</td>
</tr>
<tr>
<td>3045</td>
<td>1.738870987</td>
</tr>
<tr>
<td>3040</td>
<td>1.806806207</td>
</tr>
<tr>
<td>3035</td>
<td>1.445862172</td>
</tr>
<tr>
<td>3030</td>
<td>1.376324858</td>
</tr>
<tr>
<td>3025</td>
<td>1.433141095</td>
</tr>
<tr>
<td>3020</td>
<td>1.49346069</td>
</tr>
<tr>
<td>3015</td>
<td>1.416724956</td>
</tr>
<tr>
<td>3010</td>
<td>1.405459713</td>
</tr>
<tr>
<td>3005</td>
<td>1.425812465</td>
</tr>
<tr>
<td>3000</td>
<td>1.147949857</td>
</tr>
<tr>
<td>2998</td>
<td>1.267903116</td>
</tr>
<tr>
<td>2995</td>
<td>1.381667532</td>
</tr>
<tr>
<td>2992</td>
<td>1.329871412</td>
</tr>
<tr>
<td>2989</td>
<td>1.311306531</td>
</tr>
<tr>
<td>2986</td>
<td>1.260371188</td>
</tr>
<tr>
<td>2983</td>
<td>1.300707573</td>
</tr>
<tr>
<td>2980</td>
<td>1.466547424</td>
</tr>
<tr>
<td>2977</td>
<td>1.871780411</td>
</tr>
<tr>
<td></td>
<td>Value</td>
</tr>
<tr>
<td>----</td>
<td>---------</td>
</tr>
<tr>
<td>2974</td>
<td>1.930298195</td>
</tr>
<tr>
<td>2971</td>
<td>2.204612894</td>
</tr>
<tr>
<td>2968</td>
<td>2.829265273</td>
</tr>
<tr>
<td>2965</td>
<td>3.507061034</td>
</tr>
<tr>
<td>2962</td>
<td>3.989318567</td>
</tr>
<tr>
<td>2959</td>
<td>5.991541413</td>
</tr>
<tr>
<td>2956</td>
<td>7.220460539</td>
</tr>
<tr>
<td>2953</td>
<td>6.801845156</td>
</tr>
<tr>
<td>2950</td>
<td>7.918493833</td>
</tr>
<tr>
<td>2947</td>
<td>8.623317771</td>
</tr>
<tr>
<td>2944</td>
<td>8.66264594</td>
</tr>
<tr>
<td>2941</td>
<td>8.705621664</td>
</tr>
<tr>
<td>2938</td>
<td>5.619437091</td>
</tr>
<tr>
<td>2935</td>
<td>4.090073437</td>
</tr>
<tr>
<td>2932</td>
<td>2.455213698</td>
</tr>
<tr>
<td>2929</td>
<td>2.171406285</td>
</tr>
<tr>
<td>2926</td>
<td>1.706736161</td>
</tr>
<tr>
<td>2923</td>
<td>1.368153806</td>
</tr>
<tr>
<td>2920</td>
<td>1.14363137</td>
</tr>
<tr>
<td>2917</td>
<td>0.997737736</td>
</tr>
<tr>
<td>2914</td>
<td>0.958203906</td>
</tr>
<tr>
<td>2911</td>
<td>0.858673252</td>
</tr>
<tr>
<td>2908</td>
<td>1.033983628</td>
</tr>
<tr>
<td>2905</td>
<td>1.311262603</td>
</tr>
<tr>
<td>2902</td>
<td>1.1497833</td>
</tr>
<tr>
<td>2899</td>
<td>1.185219083</td>
</tr>
<tr>
<td>2896</td>
<td>1.326586591</td>
</tr>
<tr>
<td>2893</td>
<td>1.324096893</td>
</tr>
<tr>
<td>2890</td>
<td>1.75884255</td>
</tr>
<tr>
<td>2887</td>
<td>2.005987053</td>
</tr>
<tr>
<td>2884</td>
<td>2.457786815</td>
</tr>
<tr>
<td>2881</td>
<td>4.288642629</td>
</tr>
<tr>
<td>2878</td>
<td>6.334886145</td>
</tr>
<tr>
<td>2875</td>
<td>5.926098018</td>
</tr>
<tr>
<td>2872</td>
<td>3.382564353</td>
</tr>
<tr>
<td>2869</td>
<td>1.614649301</td>
</tr>
<tr>
<td>2866</td>
<td>0.970550407</td>
</tr>
<tr>
<td>2863</td>
<td>0.625255901</td>
</tr>
<tr>
<td>2860</td>
<td>0.376532018</td>
</tr>
<tr>
<td>2857</td>
<td>0.402236638</td>
</tr>
<tr>
<td>2854</td>
<td>0.724658656</td>
</tr>
<tr>
<td>2851</td>
<td>0.879313033</td>
</tr>
<tr>
<td>2848</td>
<td>1.020170782</td>
</tr>
<tr>
<td>2845</td>
<td>0.92223248</td>
</tr>
<tr>
<td>2842</td>
<td>0.521675356</td>
</tr>
<tr>
<td>2839</td>
<td>0.45560846</td>
</tr>
<tr>
<td>2836</td>
<td>0.29175539</td>
</tr>
<tr>
<td>2833</td>
<td>0.224880171</td>
</tr>
<tr>
<td>2830</td>
<td>0.161767663</td>
</tr>
<tr>
<td>2827</td>
<td>0.081451719</td>
</tr>
<tr>
<td>2824</td>
<td>0.060089349</td>
</tr>
<tr>
<td>2821</td>
<td>0.061608973</td>
</tr>
</tbody>
</table>
x=data(:,1);
y=data(:,2);

% read parameter

% num of peaks

% for each peak, need parameters of
% A: Oscillator Strength
% wr: resonant frequency;
% Tau: width
% phi: phase angle
% sigma: distribution of phase angle

% y=SFG_signal_sum (x,p);

p=[0
0.1
-15
2850
44
0
-18
2880
15
0
-35
2940
15
0
140
3250
130
0
286
3450
125
0
-111

2818 0.034299287
2815 0.051198061
2812 0.042897359
2809 0.033908473
2806 0.031930712
2803 0.036102402
2800 0.041991864

];
num_peaks = (length(p) - 2) / 4; % num of peaks
n = 50; % set optimization cycles

Boundary = [
    0 1 % green light level
    -0.2 0.2 % non-resonant background
    -30 30
    2844 2860
    0 40
    0 1e-15
    -40 0
    2872 2888
    0 45
    0 1e-15
    -40 0
    2935 2950
    0 30
    0 1e-15
    0 444
    3200 3380
    55 380
    0 1e-15
    0 444
    3380 3499
    55 340
    0 1e-15
    -400 100
    3050 3150
    0 400
    0 1e-15
    0 33
    3690 3710
    0 30
    0 1e-15
];

p_LB = Boundary(:, 1);
p_UB = Boundary(:, 2);

% [p_LB, p_UB] = Xin_Set_Parameter_Boundary(p);
pall=zeros(length(p),3);
pall(:,1)=p_LB;
pall(:,2)=p;
pall(:,3)=p_UB;

%boundary=[50];
%p_LB=p-boundary;
%p_UB=p+boundary;

%Display the initial error
%y_predicted = SFG_signal_sum(p, x);
%[relative_residual,residual]=relative_residual(y_predicted,y);
%disp(strcat('Initial error = ',num2str(relative_residual)));
% optimization
options=optimset('Maxiter',500000,'TolFun',1.0*10^-9); %set iterative cycles
for index=1:n
    disp(strcat('********Starting optimization cycle #
','num2str(index),'********'));
    [p,residual_norm,residual]=lsqcurvefit('SFG_signal_sum',p,x,y,p_LB,p_UB,options);
    relative_residual = sqrt(residual_norm / sum (y.^2));
    disp(strcat('relative error = ',num2str(relative_residual)));
end

% print final error for individual points and sum_error
if n>=0
    peaks=zeros(length(y),10);
    for index =1:num_peaks
        temp=SFG_Lorentzian_Gaussian(p(index*4-1),p(index*4),x,p(index*4+1),p(index*4+2));
        peaks(:,index)=abs(temp).^2;
    end
    y_predicted = SFG_signal_sum(p, x);
    figure;
    plot(x,y,'ro',x,y_predicted,'b-',x,peaks(:,1),'r--',
         x,peaks(:,2),'g--',x,peaks(:,3),'c--',x,peaks(:,4),'r--',
         x,peaks(:,5),'g--',x,peaks(:,6)); % plot y against x
    % Print final variables
    disp(p);
end

% Print results
ys=zeros(10,length(y));
ys(1,:) = rot90(x);
ys(2,:) = rot90(y);
ys(3,:) = rot90(y_predicted);

for index = 1:num_peaks
    temp = SFG_Lorentzian_Gaussian(p(index*4-1), p(index*4), x, p(index*4+1), p(index*4+2));
    ys(index+3,:) = rot90(abs(temp).^2);
end

fid = fopen('data.txt', 'w');
fprintf(fid, '%6.6f %6.6f %6.6f %6.6f %6.6f %6.6f %6.6f %6.6f %6.6f
', ys);
fclose(fid);

relative_residual.m

function [relative_residual, residual] = relative_residual(y_predicted, y)
residual = y_predicted - y;
relative_residual = sqrt(sum(residual.^2)/sum(y.^2));

SFG_Lorentzian_Gaussian.m

% Susceptibility with a gaussian distributed phase angle centered at zero degree%
% OUTPUT: % return the susceptibility array (complex numberes) as function of frequencies for a resonent peak
% INPUT: % w frequencies, can be array % wr resonent frequency, Tau, width, A, oscillation strength. All scaler % sigma, distribution of phase angle, in arc unit; function ki = SFG_Lorentzian_Gaussian (A,wr,w,Tau,sigma)

if (sigma<=10) % wavenumbers
    ki = SFG_Lorentzian(A,wr,w,Tau);
    return;
ki=zeros(size(w));

num_per_sigma=30;  %number of point evaluated for each sigma
max_sigma=3;       %max sigma evaluated

norm=0;
for n= -num_per_sigma*max_sigma:num_per_sigma*max_sigma
    weight= exp(-1*n*n/(num_per_sigma*num_per_sigma)); % weight according to guassion distribution.
    ki=ki+SFG_Lorentzian(A,wr-n/num_per_sigma*sigma,w,Tau)*weight;
    norm=norm + weight;
end
ki=ki/norm;

SFG_signal_sum.m
% Sum of SFG signal %
% OUTPUT:
% return the signal array (real numberes) as function of frequencies
% INPUT:
% w frequencies, can be array
% n number of peaks
function y = SFG_signal_sum (parameters, frequency)

ki=zeros(size(frequency));
y=zeros(size(frequency));

num_peaks = (length(parameters)-2)/4;
for i = 1:num_peaks
    index = (i-1)*4 + 2 ;
    ki=ki+SFG_Lorentzian_Gaussian(parameters(index+1),parameters(index+2),frequency,parameters(index+3),parameters(index+4));
end

ki= ki+ parameters(2);  % non-resonent SFG signal
y = abs(ki).^2;
y= y+ parameters(1);  % Backgroud noise from green light scattering
Appendix B The Matlab Script Used for MEM Analysis

```matlab
function Diff = test(filename, M, DV)
% This function is to calc ....
% input : filename - the data file contents data set( WaveNumber,
%         SFGIntensity)
%        M, DV consistent, DV stands for 1/dv

if nargin == 1
    M = 40; DV = 400;
elseif nargin == 2
    DV = 400;
elseif nargin < 1
    filename = 'ySimuAllData.txt'; % 'y1.txt';
    M = 50; DV = 200;
end

fid = fopen(filename);
oriData = fscanf(fid, '%g %g', [2 inf]);
close(fid);

w = oriData(1,:)';
SFG = oriData(2,:)';

BetaSquare = 0;
[xo, xi, yi] = Smooth(w, SFG, DV);
[A, BetaSquare] = Coefficients(xi, yi, M);
AmV = Am(A, DV);

x = sqrt(BetaSquare) ./AmV;
MAmV = abs(AmV).*abs(AmV);
cSFG = BetaSquare./MAmV;

Im2Eq6 = sqrt(yi).*sin(-angle(AmV));
t = abs(x).*sin(angle(x));

Diff = sum(abs(yi - cSFG))/sum(abs(yi));
resut = [xo' yi' cSFG' t' Im2Eq6']; % only for export to EXCEL

plot(xo, yi, 'magenta', xo, cSFG, 'blue', xo, t, 'red',
     xo, Im2Eq6, 'cyan');
legend('SFG Intensity', 'Calc SFG', 'Im@Ben', 'Im@Eq6', 2);
disp('done!');
```
newfilename = sprintf('%s-M%dDV%d(Diff%4.2f).Result.xls', filename, round(M), round(DV), Diff);

[fid, msg] = fopen(newfilename, 'wt+');
if(fid < 0 )
    disp(msg);
else
    fprintf(fid,'Filename:%s
W=t%d
DV=t%d
Diff=t%4.2f
', filename, M, DV, Diff);
    fprintf(fid,'WaveLength t SFG t cSFG t Im@ben t Im2Eq6
');
    fprintf(fid,'%4.4f t%9.5f t%9.5f t%9.5f t%9.5f
', resut);
    fclose(fid);
end

% ========End of the main function test()=====}

function [a,beta] = Coefficients(x,y,M)
    rm= 0:1:2*M;
    for tmp= 0:2*M
        mm= tmp-M;
        rm(tmp+1) = trapz(x, y.*exp(1i*2*pi*mm*x));
    end
    r = zeros(M);
    for row=1:M
        for col = 1:M
            r(row,col)=rm(row-col+M+1);
        end
    end
    b= 0 - rm(M+2:2*M+1);
    a = r\b.,'
    beta = [r(1,:) rm(1)]*[1;a]; % same as [r(1,:)
    rm(1)].*[1;a].'
    beta = abs(beta);
end % ========End of the function Coefficients()=====}

function ret = Am(a,DV)
m = length(a);

K = (1:m)*2*pi;
%RA=abs(A);
K = K';
ret = 0:DV;

for tmp = 0:DV
    v = tmp/DV;
    ret(tmp+1) = 1+ sum(a.*exp(1i*v*K));
end

end % ===========End of the function Am ============

function [xo,xi,yi] = Smooth(x,y,DV)
% convn and interpl accordin
g the giving wavenumber and SFG
%   size of the averaging window in convn is 4
%  xo is reflect xi from 0-1 to original wavenumber

maxx = max(x);
minx = min(x);

sv=(x - minx)/(maxx-minx);

% ty = [ ones(5,1).*y(1) ; y ; ones(5,1).*y(length(y)) ]; % temp array for convn, duplicate first and last elements

span = 5; % Size of the averaging window
window = ones(span,1)/span;
sty=convn(ty,window,'same');

sy = sty(6: length(ty)-5); % sub array of sty, omit the first and last 5 elements

xi = 0:1/DV:1;

yi = interp1(sv,sy,xi, 'cubic');% can't use 'spline';
'linear' close to 'cubic'

xo = xi.* (maxx-minx) + minx;

end % =====end of function Smooth======
VITA

Yi Liu

Address: Department of Chemistry
Frear South Building, Room 233
Pennsylvania State University
University Park, PA, USA, 16802

E-mail: yzl155@psu.edu

Education:
2009...........................................B.S. in Chemistry, Peking University, China
2009 to 2010........Graduate Teaching Associate, Department of Chemistry, Texas A&M University
2010 to 2012........Graduate Research Associate, Department of Chemistry, Texas A&M University
2013 to present...Graduate Research Associate, Department of Chemistry, The Pennsylvania State University