Sum Frequency Generation and Molecular Dynamics Studies of Biomolecular Orientation at Model Interfaces

by

Andrew P. Boughton

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2011

Doctoral Committee:

Professor Zhan Chen, Chair
Professor Ayyalusamy Ramamoorthy
Associate Professor John J.G. Tesmer
Associate Professor Ioan Andricioaei, University of California, Irvine

Acknowledgements

The research herein would not have been possible without the help of many talented colleagues. First, I am grateful to Prof. Chen and Prof. Andricioaei for their mentoring and sound advice. On a daily basis, I am also grateful to a variety of colleagues from both labs who provided the encouragement and technical know-how needed to explore complex scientific questions. I am indebted to Pu Tian, Catherine Musselman, Jeffrey Wereszcynski, Joanne Chen, Maria Mills, Jeremiah Nummela, Jie Wang, Matthew Clarke, Xiaoyun Chen, Arthur McClelland, Ann Vázquez, Shuji Ye, Pei Yang, Bei Ding, Josh Jasensky, and Lauren Soblosky.

Additional thanks goes to my committee members: Professors Crippen, Ramamoorthy, and Tesmer; their feedback and assistance has been invaluable. I am similarly beholden to my many teachers for their inspiration and encouragement, particularly Prof. Steimle.

Financial support and technical resources were provided by a number of agencies and programs, including NIH, NSF, and GAANN. The simulations described in this work would not have been possible without computing support from the UM Department of Chemistry, UM Center for Advanced Computing, and UC Irvine Greenplanet cluster and molecular modeling facility.

Lastly, I am grateful to all my family and friends for their support and assistance at all stages, with particular thanks to Amanda Cleveland.

Table of Contents

List of figures

List of Tables

Chapter 1: Introduction

1.1 Motivation and Background

Although many biological phenomena have been studied in solution, few experimental techniques possess the sensitivity and specificity to probe molecules at interfaces, which are of biological importance. Biomolecules at interfaces control many important phenomena- examples include blood coagulation on implant surfaces, membrane protein functions, marine biofouling, biosensing, and antimicrobial potency and selectivity.[1-6]

Proteins exist at interfaces quite naturally as well. For example, the lipid membranes that separate cells from their surroundings contain a wide range of proteins. These molecules are vital for transport of nutrients, survival of the cell, and the ability to sense and adapt to changes in the surrounding environment. Yet despite the availability of high-quality crystal structures for many peripheral and integral membrane proteins, comparatively little information is available about the orientation of membrane proteins and peptides when in their native biological environment. Changes in solvent or surroundings can have a significant effect on protein orientation, conformation, or dynamics, and it is necessary to investigate the structures of membrane associated proteins *in situ*.

Surface sensitive *in situ* probes of biomolecules at interfaces are needed to address these and other questions, yet such techniques are often limited- either due to a requirement for high vacuum, interfering signals from the surrounding environment, or a

need for chemically added labels which may alter the native behavior of the biomolecule in question.

To this end, Sum Frequency Generation (SFG) Vibrational Spectroscopy has recently been applied to a range of biological systems.[7-41] Valuable information about molecular orientation and conformation can be obtained, yet full characterization of complicated biological systems is hindered by the limited number of measurements available. The work described herein is aimed at improving interpretation of complex spectra through several computational strategies: molecular dynamics (MD) simulations, free energy sampling, and a novel method for determining overall protein orientation by calculating the combined nonlinear optical response for the molecule.

1.2 Sum Frequency Generation (SFG): Theory, Experiment, and Data Analysis

1.2.1 SFG: A Nonlinear Optical Process

SFG is a second-order nonlinear optical process. One practical consequence of this fact is that output signals will be observed only in media that lack inversion symmetry (under the electric dipole approximation[42]). Bulk materials such as air, liquids, amorphous solids, and many crystalline solids possess inversion symmetry. At many surfaces and interfaces, inversion symmetry is broken; thus SFG is highly sensitive to surface-adsorbed molecules and interfacial molecular ordering.[43, 44]

In the typical SFG experiment, a fixed-frequency visible beam is overlapped in space and time with a frequency tunable infrared laser, producing an output beam whose frequency is the sum of the two input beams (Figure 1).

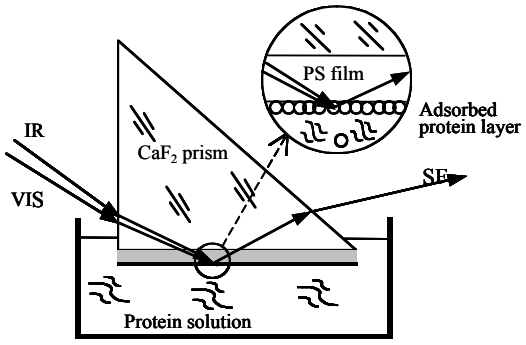


Figure 1-1: The "near" total internal reflection geometry used in my experiments.

This process is greatly enhanced when one of the beams is resonant with the vibrational transitions of the molecule, resulting in a surface-sensitive vibrational spectroscopic technique. By using polarized beams, molecular orientation can also be studied. Data is presented as the output (sum frequency) signal intensity plotted against the frequency of the tunable infrared beam, and SFG vibrational peak centers appear similar to those found in Fourier Transform Infrared (FTIR) and Raman spectra.

However, there are important differences. Vibrational modes are observed in FTIR or Raman spectra when either the dipole or polarizability transition moments change, respectively. The response of the molecule in SFG depends on a property known as the hyperpolarizability (β), which is the *product* of IR dipole and Raman polarizability transition moments, such that $\beta_{ijk} = \left(\frac{d\alpha^*}{dQ}\right)_{i,j} * \left(\frac{d\mu}{dQ}\right)_k$. If either the change in polarizability (α) or the change in dipole moment (μ) of a vibrational mode is equal to zero, then no SFG signal will be observed: the vibrational peaks observed in SFG are only those that are present in *both* infrared ***and*** Raman spectroscopy. This can be advantageous in the study of adsorbed proteins, as the elimination of extraneous peaks simplifies the resulting spectrum. Further, the selection rule of SFG is such that the interfacial layer is probed

selectively, and with good sensitivity. Both experiments and theoretical simulations indicate that SFG is submonolayer surface sensitive. By contrast, the surface sensitivity in Attenuated Total Reflectance Fourier Transform IR spectroscopy (ATR-FTIR, a commonly used linear spectroscopy), is determined by the penetration depth of the evanescent wave (on the order of hundreds of nanometers or even microns). In practice, ATR-FTIR probes many layers of molecules, beyond those at the surface or interface- as a result, large water bending signals must be subtracted when studying backbone signals from proteins. This background subtraction is not necessary for SFG.

To date, much work has been done in SFG to study the CH stretching modes in various chemical groups (such as methyl, methylene, and phenyl groups). More recent studies have extended to the highly informative Amide I band, which contains signals from backbone carbonyl groups in the protein.[19] The Amide I peak center frequencies and peak widths are sensitive to the secondary structure and conformational changes of the protein, as in FTIR[45-47]. SFG is complementary to other vibrational spectroscopies, with a range of new advantages for the study of complex biological systems.

1.2.2 SFG Experimental Procedures

The typical SFG experiment employs a fixed-frequency visible beam and an infrared laser (either tunable-frequency or broadband). Representative SFG instruments are described in refs 48-53, and in recent years commercial instruments have become increasingly available from several manufacturers. The input laser beams are spatially and temporally overlapped to produce the third "output" frequency, which is collected by photomultiplier tube, CCD, or other detection scheme. Since the beam sizes are quite small (~500 μm), sample requirements are modest. An example SFG system is shown in

Figure 1-2. This system has four components: 1) a pico-second Nd:YAG laser, 2) a harmonic unit with two KD*P crystals, 3) an optical parametric generation (OPG)/optical parametric amplification (OPA) and difference frequency generation (DFG) system based on LBO and $AgGaS_2$ crystals, and 4) a detection system. The visible beam (532 nm) is generated by frequency-doubling the fundamental output pulses of 20 ps pulsewidth from the Nd:YAG laser. The tunable IR beam is generated from the OPG/OPA and DFG system, and can be tuned from 1000 to 4300 cm^{-1}. By replacing the $AgGaS_2$ crystal with a GaSe crystal, this range may be extended down to 650 cm^{-1}. In the system used by the Chen group, the output signal is collected by a photomultiplier tube and processed with a gated integrator. Alternatively, CCD detectors are becoming increasingly available. By monitoring the power of the visible and IR beams using photodiodes, SFG spectra can be normalized by the power of the input laser beams.

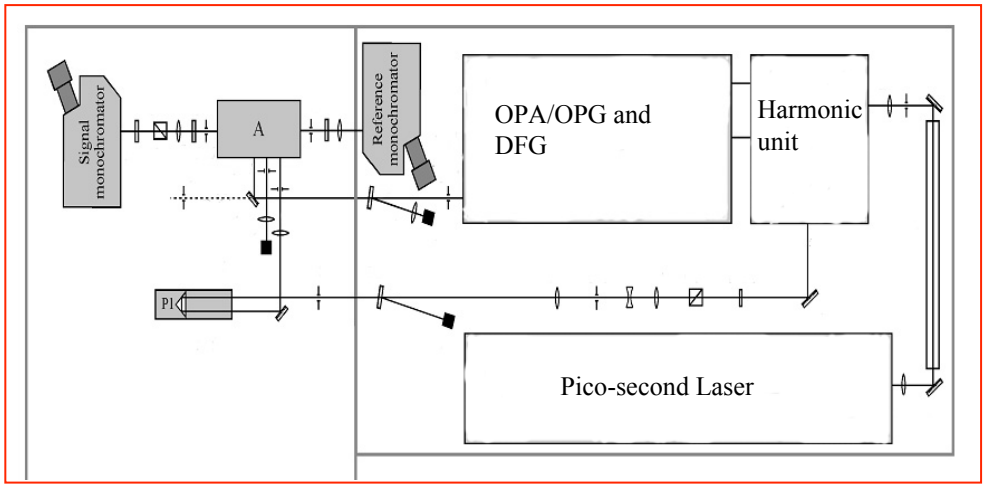


Figure 1-2: Schematic of the SFG spectrometer used in this research

Weak protein signals may be enhanced via changes in experimental geometry. As shown in section 1.2.3, the output signal is affected by the Fresnel factors (L_{ijk}), which depend on the incident angles of the beams and the indices of refraction of the media the

beams pass through. Many existing SFG experiments employ a straightforward reflection geometry, which has the advantage of being relatively easy to align. By changing to a "near" total internal reflection right-angle prism geometry instead, the angle-dependent Fresnel factors are increased substantially, making it easier to observe weak signals experimentally.[19] This provides advantages for fitting and quantitative analysis, as well as improved detection limits.

One important advantage of SFG in the study of biomolecules is that unlike vacuum techniques, SFG is capable of measuring protein orientation and conformation in a biologically relevant, aqueous environment. The surface sensitivity of SFG is such that as little as a few micrograms of the protein of interest are enough for reasonable signals to be observed, making it feasible to study protein behavior even at low concentrations.

1.2.3 Spectral Fitting and Orientation Determination

SFG signal intensity is proportional to the square of the effective second order nonlinear optical susceptibility: $I \propto \left|\chi_{eff}^{(2)}\right|^2$. As discussed herein, the use of polarized laser beams provides a means to relate $\chi_{eff}^{(2)}$ to the orientation of vibrational modes, and hence functional groups or molecules, at the interface.

The effective susceptibility component measured in a given polarization can be obtained by fitting spectra to a Lorentzian lineshape[54]:

$$\chi_{eff}^{(2)} = \chi_{nr} + \sum_q \frac{A_q}{\omega_2 - \omega_q + i\Gamma_q} \qquad \textit{Equation 1-1}$$

where ω_2 and ω_q represent the frequencies of the infrared and q^{th} peak center, respectively, Γ_q is the damping coefficient, χ_{nr} is a constant nonresonant background signal, and A_q is the signal strength. When different polarization combinations of

input/signal beams (e.g., ssp or ppp) are used, different components of $\chi_{eff}^{(2)}$ (such as $\chi_{eff,ssp}^{(2)}$ or $\chi_{eff,ppp}^{(2)}$) can be determined. The experimentally measured parameters can be related to the actual second order nonlinear optical susceptibility components, which are defined in the lab coordinate system (where z is usually defined as the surface normal and x-y plane is the surface/interface. The x-z plane contains the input and signal beams). For the near total reflection geometry:

$$\chi_{eff,ssp}^{(2)} = L_{xxz}\chi_{xxz}^{(2)} \qquad \textit{Equation 1-2}$$

$$\chi_{eff,ppp}^{(2)} = L_{zzz}\chi_{zzz}^{(2)} \qquad \textit{Equation 1-3}$$

where L_{xxz} and L_{zzz} are Fresnel factors that take into account the amount of light that is reflected or transmitted at the interface; such factors depend on polarization[55]. These relations hold for the ssp and ppp polarization combinations of the sum, visible, and infrared beams, respectively. The $\chi_{zzz}^{(2)}/\chi_{xxz}^{(2)}$ ratio is the principal measurement used for orientation studies, and it may be measured from the SFG ssp and ppp spectra.

The $\chi_{zzz}^{(2)}/\chi_{xxz}^{(2)}$ ratio is a property of many molecules. To link it to the orientation of an individual molecule, the hyperpolarizability tensor for a single molecule must first be calculated based on knowledge of the functional group or molecule. To date, parameters and equations have been derived for a wide variety of functional groups[49, 56-62] and protein secondary structures.[7, 8, 14, 20, 63-67] In a broad sense, the molecular and macroscopic properties are linked by relating the response of a single molecule to the response measured in the lab coordinate system, and performing averaging for the ensemble of molecules as needed. The orientation of a surface functional group or molecule can be characterized by either one tilt angle θ (e.g. in case of a single α-helix)

or tilt and twist angles θ and ψ, respectively, (e.g. in the case of a β-sheet or a complicated protein).

1.2.4 Limitations

SFG measurements offer the ability to probe changes in molecular orientation and conformation with high surface sensitivity, and without the need for rigorous background subtraction. However, due to the relations that govern observed signal intensities, it is not always directly possible to distinguish changes in orientation from interfacial number density: a dense layer of molecules lying parallel to the surface might yield no signal.[68]

Also, SFG can only provide a limited number of experimental measurements. Hence data interpretation often requires assuming that all molecules adopt the same orientation (δ distribution) or a known singular orientation distribution (such as a Gaussian distribution), which may not be realistic for biological systems. Furthermore, although the Amide I band is somewhat sensitive to changes in protein secondary structure, small changes in protein conformation may not be detectable within the limited resolution of our spectrometer. A more detailed picture of protein structure may be obtained using a complementary technique such as molecular dynamics.

1.3 Molecular Dynamics

1.3.1 Theory and Background

Classical molecular dynamics simulations rely on a simple physical model in which the interactions between atoms are depicted using a model of charged balls (atoms) connected by flexible springs (bonds). By solving Newton's equations of motion at every step, and considering the interactions (bonded and non-bonded alike) between all atoms

in the system, it is possible to model the movement of all atoms in the system over time and extract information about molecular structure in atomistic detail.

To develop the popular CHARMM[69] force field for biomolecular simulation, parameters for atom partial charges were determined based on quantum mechanical calculations for model compounds in the presence of TIP3P water molecules.[70] Thus, the parameters extracted are very specifically tuned to reproduce the behavior of proteins in an aqueous environment, using explicit solvent. The following interactions are considered:

$$U_{CHARMM} = U_{bonded} + U_{non-bonded} \qquad \textit{Equation 1-4}$$

where U_{bonded} represents bond stretches and rotation, as well as additional terms for in-plane deformations, dihedral angles, and additional interactions found necessary to reasonably reproduce the behavior of biological molecules. Non-bonded interactions are modeled with electrostatic and Lennard-Jones potential terms. To reduce the number of terms that must be calculated in a system with many pairs of atoms, a switching function is typically employed so that (weak) interactions between pairs of atoms more than ~12.0 Å apart can be ignored.

The CHARMM potential energy function is as follows:

$$\begin{aligned} U(\vec{R}) = & \sum_{bonds} K_b(b - b_o)^2 + \sum_{UB} K_{UB}(S - S_o)^2 \\ & + \sum_{angle} K_\theta(\theta - \theta_o)^2 \\ & + \sum_{dihedrals} K_\chi(1 + cos\,(n\chi - \delta)) \\ & + \sum_{impropers} K_{imp}(\phi - \phi_o)^2 + \sum_{nonbond} \epsilon\left[\left(\frac{R_{min,ij}}{r}\right)^{12} - \left(\frac{R_{min,ij}}{r}\right)^{6}\right] + \frac{q_i q_j}{\epsilon_1 r_{ij}} \end{aligned}$$

Equation 1-5

where K_b, K_{UB}, K_θ, K_χ, and K_{imp} are the bond, Urey-Bradley, angle, dihedral angle, and improper dihedral angle force constants, respectively; b, S, θ, χ, and ϕ are the bond length, Urey-Bradley 1,3-distance, bond angle, dihedral angle, and improper torsion angle, respectively. The subscript zero represents the equilibrium values for the individual terms[70]. It should be noted that the above force field provides only the potential energy of the system. Entropic contributions to the total adsorbed free energy must be determined using more advanced sampling methods[71-77], as discussed in Chapter 2.

1.3.2 Information Obtainable (and Limitations)

Molecular Dynamics (MD) simulations are a powerful tool for studying the motion of atoms and molecules in space and time. In principle, they provide the ability to study both structural and dynamic properties of the molecule at an atomistic level of detail, with femtosecond time resolution. From there, information about stable states and free energy can be extracted.

The primary limitation of molecular dynamics, then, is that these interactions must be calculated at femtosecond timescales (limited by the frequency of the fastest bond vibrations). The calculation of non-bonded interactions between all atoms in the system is particularly time consuming. Although a variety of methods may be used to speed these calculations[69, 78, 79], the timescales accessible are ultimately limited by the amount of computing power available. To date, most reported simulations involving protein adsorption have been on the order of nanoseconds to hundreds of nanoseconds. Even with specially optimized hardware and algorithms, the longest atomistic simulations of any biological system represent < 1 ms of total run time; most simulations are much shorter.[80-83] As a result, it is difficult, if not impossible, to calculate final protein structure

or free energy profiles using brute-force simulations. Such information can be extracted from shorter simulations by taking advantage of a variety of cleverly designed sampling methods, and thus comparisons can be made to experimental results that represent the equilibrium state attained over much longer timescales.

1.4 Prior Work

1.4.1 SFG Studies of Proteins: The CH and NH Spectral Region

Partially due to limitations in the tunable frequency range of older laser instruments, many early SFG studies of proteins focused on CH stretching signals: methyl groups, phenyl rings, and other functional groups of interest in protein amino acid side chains. Although the sheer number of similar amino acids would hamper any serious attempt to study the orientation and conformation of the entire protein at the interface based on CH signals, the requirement that a medium lack inversion symmetry allows signals to be interpreted in terms of side chain ordering at the interface. Early studies showed that protein CH signals could be detected readily on a range of biomedically important polymers[84], and these signals were interpreted as demonstrating the segregation of hydrophobic and hydrophilic residues at the polymer/water boundary. As a later test, Somorjai and coworkers used a simple peptide designed to form a facially amphiphilic helix at interfaces, and found that strong CH signals could indeed be obtained.[26, 27] These results confirmed that the hydrophobic residues are strongly ordered at a hydrophobic/hydrophilic interface (e.g. polymer/solution). Such ordering of hydrophobic groups has also been inferred for a variety of amino acids at an oil/water interface.[85] By also studying the NH stretch, it has been demonstrated that specific amino acids remain buried after the conversion of fibrinogen to fibrin.[35] Water molecules can also order in

the vicinity of a protein, providing additional clues about protein adsorption and the burial of hydrophobic regions.[86, 87]

Side chain ordering at an interface can be influenced by changing the contacting media of the protein.[16, 26, 88] Different spectra were collected from bovine serum albumin (BSA) at the polymer/water interface as opposed to the polymer/air interface, due to different ordering of hydrophobic groups. Replacing the water with a hydrophobic solvent (such as benzene) also leads to spectral changes. Stronger ordering was observed at hydrophobic/hydrophilic interfaces, whereas only weak signals were observed from hydrophilic/hydrophilic interfaces (such as SiO_2/water). It is important to note that SFG is very sensitive to molecular orientation, and as a result, weaker signals can sometimes be observed even when a large amount of adsorbed protein is present (if those molecules are less well-ordered).[24]

As protein concentration increases, protein-protein interactions become increasingly important. Protein CH signals were found to change at a protein solution/air interface as a function of concentration, indicating that interactions between nearby proteins can be very important in determining the ordering of hydrophobic side chains (as well as the entire structure).[24, 89] Various surfactants were likewise shown to affect protein aggregation.[23]

Investigating protein structures at interfaces using SFG CH signals provides direct, *in situ* evidence that interfacial protein side chain ordering can change when the environment is altered. For simpler biomolecules such as amino acids and peptides with well-defined structures, more detailed structural information may be inferred from such studies, and this is valuable for the study of structural changes related to biocompatibility

and biofouling.[26-28, 90] However, the very large number of similar hydrophobic groups in many larger proteins makes it difficult to describe structure in detail using CH signals alone. Isotope labeling of specific amino acids is one way to gain additional information from interfacial proteins.[91-93]

1.4.2 Studying Amide I Protein Signals with SFG

1.4.2.1 Advantages of the Amide I Band

The Amide I band of proteins is very useful for vibrational studies of proteins. This band can be loosely described as a stretching mode of backbone carbonyls[94]. The exact peak center frequency depends on the secondary structure and conformation of the protein. These frequencies are roughly the same as those observed via FTIR, but SFG Amide I spectra do not suffer the interference from strong water bending signals that usually obscures the Amide I region in ATR-FTIR. As newer and more broadly tunable SFG instruments become increasingly available, it is expected that this informative spectral region will be increasingly studied.

The unique peak center frequencies of various secondary structural elements in a protein (Table 1-1) can be combined with the orientation information obtainable by SFG to create a more complete picture of interfacial orientation. Although vibrational spectroscopy cannot directly inform us *which* segments of the protein have adopted a given helical or β-sheet conformation, the increasing availability of high-quality protein structures (via the Protein Data Bank, PDB) facilitates determination of overall molecular orientation based on a known structure. If regions of the protein undergo bending or deformation, changes in the spectrum will be observed relative to this. As discussed above, CH signals are often difficult to interpret in terms of detailed molecular

orientation. In contrast, the Amide I band presents a picture of backbone molecular orientation that is well-suited to the task.

Frequency (cm^{-1})	Polarization	Assignment
1635	ssp, ppp, spp	β-sheet (B_2 mode)
1652	ssp, ppp	α-helix (A, E_1 modes) in D_2O
1655	ssp, ppp	α-helix (A, E_1 modes) in H_2O
1645	ssp, ppp	Coiled-coils (A, E1 modes)
1652-1665	Ssp	Turns/random coils
1660-1670	ssp, ppp	3_{10} helix
1685	ssp, ppp, spp	β-sheet (B_1 mode)/ interstrand
1730	ssp, spp	β-sheet (B_3 mode)/ side chains

Table 1-1: Correlations between protein secondary structures and peak center frequency for Amide I vibrations.

Amide I signals have been examined by several research groups to date. The first SFG Amide I signals were collected from the polymer/protein interface using the near total reflection geometry, and it was shown that SFG is sensitive enough to distinguish between proteins with different secondary structures.[19] By using antimicrobial peptides such as α-helical MSI-594 and the β-sheet peptide tachyplesin I, it has been shown that SFG Amide I signals originating from proteins can be detected from both α-helices and β-sheets distinguishably.[11, 95] Using the thin-film optical model for the interpretation of spectra, it was proven that the signals obtained are from proteins adhered to the surface, rather than only part of a layer.[96]

1.4.2.2 Qualitative Studies: Orientation and Ordering

As was discussed above for the CH and NH spectral region, the strongest signals are obtained for greater molecular ordering or certain orientations, though for the Amide I region, the signals obtained are a guide to overall ordering/orientation of the protein backbone. Thus strong signals were observed for long and well-aligned poly-γ-benzyl-L-glutamate (PBLG) helices, [97] yet only very weak signals were obtained from albumin

despite a large number of α-helical segments in the latter.[19] As with other chemical groups, this molecular ordering or orientation can be affected by the composition of the surface in question, as was demonstrated for time-dependent SFG Amide I signal changes of fibrinogen on a variety of polymer surfaces.[68]

Methods for quantitative orientation determination of α-helices have been presented[14], and applied to a variety of systems on polymers[20, 98, 99] and in lipid bilayers[8, 64, 100, 101]. These results demonstrate that the measured orientation may be affected by factors such as concentration, aggregation, and the length and specific structure of the peptide being studied.

Information can also be extracted by considering the phase of the signals interfering with non-resonant background from a thin metal layer such as gold[37, 38], but this is not always appropriate for biological samples, and the metal films are susceptible to burning under intense laser power. Hence the use of polarized Amide I signal ratios is more broadly applicable.

1.4.2.3 Combining Spectroscopies Yields More Measurements

SFG provides a number of advantages for the study of biomolecules at interfaces: it is an inherently surface sensitive technique with many of the advantages and applications of FTIR, yet is also capable of providing more measurements to characterize the orientation of biological systems. Even a single-helix peptide in a well-defined environment cannot always be described by assuming only a single orientation. For larger proteins with multiple similar structural elements, obtaining more measurements from additional vibrational spectroscopies is critical.[102]

Different vibrational spectroscopic techniques measure different structural parameters. For example, if the tilt angle of a structural element (e.g., an α-helix) is θ versus the surface normal, ATR-FTIR can determine the angle through measuring $\langle\cos^2\theta\rangle$. For ATR-FTIR, this is the only parameter obtained. Therefore, any orientation analysis requires assuming that all chemically identical groups at the interface adopt the same orientation (a "δ distribution"). In contrast, SFG can measure both $\langle\cos\theta\rangle$ and $\langle\cos^3\theta\rangle$. Thus, SFG can be used to deduce two different structural parameters. In a Gaussian distribution, this would correspond to the average orientation and distribution width. More complex distributions require more measurements, which can be obtained by combining the results from multiple vibrational spectroscopic techniques.

By combining ATR-FTIR with SFG, the orientation of the helical peptide melittin in a bilayer was measured at the threshold concentration for antimicrobial activity, and it was found that two distinct orientations of the peptide were present simultaneously.[8] Using ATR-FTIR, only one parameter can be measured: namely, the intensity ratio of the s- and p-polarized ATR-FTIR absorbances. Thus, these results alone can yield only the average orientation. If every melittin molecule adopted the same orientation, ATR-FTIR would predict a tilt angle of 54^o relative to the surface normal. However, the unique $\langle\cos\theta\rangle$ and $\langle\cos^3\theta\rangle$ parameters measured by SFG were not consistent with either a δ-distribution or a single Gaussian distribution.

Therefore, the orientation distribution must be more complicated. By assuming that two orientations of melittin existed simultaneously, combined SFG and ATR-FTIR results demonstrated that approximately one fourth of the melittin molecules laid down at an angle of 100^o, and the remainder stood up in the bilayer at an angle of 13^o.

1.4.2.4 Orientation of Large Proteins Based on α-helical Domains

As discussed above, the orientation of larger proteins may be determined by calculating the combined hyperpolarizability of all the helical segments in the molecule. In a study of the multi-helix membrane signaling protein, Gβγ, the orientation of α-helical portions of the protein and the static crystal structure were used to determine overall molecular orientation of the protein bound in a membrane, and it was demonstrated that lipid composition can modulate that orientation.[7] Gβγ is anchored to the bilayer via a geranylgeranyl group with high affinity, and much weaker signal intensities were observed in the absence of this group. SFG experiments suggest that the geranylgeranyl group also anchors the protein in a specific orientation.

1.4.2.5 β-sheets Yield Chiral Signals

Because β-sheets may be quasi-centrosymmetric, there was originally some question as to whether signals could be detected for this structure using SFG. In fact, it has been shown that β-sheets possess D_2 symmetry, and that weak signals can be enhancd using a series of intermediate polarizations.[95] Recently, new relations have been derived that allow SFG to be used for more quantitative studies of β-sheet orientation.[63]

1.4.3 Molecular Dynamics: Previous Work

Applications of molecular dynamics to proteins have been reported as early as 1977, when the dynamics of a 58-residue peptide were studied in the presence of four water molecules; this simulation spanned some 8.8ps.[103] Thanks to advances in computer processing power, parallelization, and more efficient algorithms, much larger systems can now be routinely studied on nanosecond or longer timescales: an increase of several orders of magnitude, even when the protein is fully solvated.[80, 81, 104-106] Although fully

atomistic simulations have gained in popularity, coarse grained and implicit solvent/membrane models remain a source of ongoing interest as well. These approaches may ignore fine structural details, but in return, the reduced configurational space allows for greatly increased sampling and longer timescales.[79, 107-114]

1.4.3.1 Model Interfaces

As simulations of larger systems become more feasible, it has become possible to study interfaces, and thus protein adsorption. A variety of methods may be used. In many simulations, hydrophobic interfaces have been represented as idealized paraffin-like plates[115], using a liquid hexane-water boundary[116], or in atomistic detail using rigid graphite or SiO_2 surfaces.[28, 117, 118] In some cases, these simulations use few water molecules, or neglect them completely in favor of computationally inexpensive implicit solvent or implicit membrane models that may neglect important interactions in the system.[119, 120] Whereas periodic boundary conditions have been commonly used for simulations elsewhere, they typically assume infinite periodicity in three dimensions, which may lead to artifacts when applied to two-dimensional systems.[121, 122]

For a more detailed representation of the surface, Latour and coworkers have employed self-assembled monolayers in implicit or explicit solvent for the study of adsorption, and demonstrated that commonly used implicit solvent models tend to overestimate the adsorption energies of peptides.[123] A disadvantage of their model system is that it incorporates a layer of fixed-position bulk water molecules, which may constrain re-arrangement of the solvent around the protein.[121] For studies of short peptides in lipid bilayers, coarse grained simulations have also been employed, in which the number of atoms in the system is reduced by representing entire functional groups or residues as beads.[124] A variety of studies have also been performed in explicit and implicit membrane

models, though the long timescales required for equilibration and dynamics are a deterrent to methods that rely on detailed sampling of the system.

1.4.3.2 Information Obtainable

MD simulations are widely used to probe information about local structure and dynamics. For example, they have been used to study which residues of the helical peptide Magainin 2 are involved in membrane insertion[104], revealing important interactions with the lipid headgroups and hydrophobic tails. Final molecular orientation [104, 105, 117, 118, 123, 125], side chain ordering[28], water ordering[126, 127] and molecular deformability[110, 124, 128-130] have also been studied at interfacial environments.

For quantitative studies, a variety of procedures exist for the calculation of energy[131-133] or free energy change upon adsorption. Entropic contributions must be considered via one of a variety of sampling methods, many of which rely on a constraining potential to increase sampling at disfavored positions[71-77, 134, 135]. Two commonly used reaction coordinates for the study of adsorption are the distance of the protein from the surface[118, 129, 130] and the orientation of the molecule[98, 105, 136, 137]. In the latter case, the structure of the protein is typically held fixed to avoid artificial deformation. Free energy sampling will be discussed in more detail in Chapter 2.

1.5 Presented Research

In the last decade, SFG has been applied to biological systems with increasing frequency. The goal of the research presented herein is to enhance and extend our ability to interpret complex protein spectra through a variety of computational methods.

In Chapter 2, the orientation of the α-helical peptide magainin 2 was explored at a model hydrophobic/hydrophilic interface that was chosen to be amenable to detailed

simulations even with limited computing power. In the past, SFG studies on similar model systems have focused on side chain CH stretching signals, and concluded that the peptides laid down completely at high concentrations. In this work, newly measured Amide I region spectra reveal that the actual orientational behavior is more complicated. Furthermore, these ratios are concentration dependent, even for concentrations several orders of magnitude below those previously studied. Using a simplified polymer surface representation with explicit solvent, it was shown that these results would not be correlated to the most stable orientation for a single peptide. A slightly different orientation may be possible if peptides form dimers on the surface. Such peptide-peptide interactions have widely been ignored in previous simulations, but larger oligomers are worthy of future study.

In Chapter 3, I turn from model helical peptides to larger proteins with multiple helical segments. Determining the orientation of complex proteins requires calculating the net nonlinear optical response of all helical segments in the protein, but to date, such calculations have been extremely tedious and error-prone. These methods were revised and extended to provide greater accuracy, implemented as a new program that facilitates the analysis of any arbitrary protein structure. This was applied to study the formation and lipid bilayer orientation of the GRK2-G$\beta_1\gamma_2$ complex, which is involved in regulating cell signaling. The results suggest that subunits must reorient in order to form the complex, and demonstrate that protein orientation at membranes cannot always be inferred from the crystal structure alone.

Although α-helical protein secondary structures were among the first to be studied, recent advances have made it possible to study β-sheet peptides as well. In

chapter 4, I report molecular dynamics simulation studies on the 17-residue β-sheet peptide Tachyplesin I. No single most stable orientation was found, and in fact, this peptide was found to exhibit significant conformational flexibility- despite the fact that orientation analysis methods typically assume both structural rigidity and a single uniform orientation of molecules (referred to as a δ distribution). Based on these simulation results, previous experimental results were revisited, and a new orientation analysis relation was derived that allows more information to be extracted from the combination of ATR-FTIR and SFG measurements. When using the new relations, it was found that no single orientation provided a match for all experimental results. In this case, simulations were able to inform the interpretation of experimental results by examining the validity of commonly used assumptions.

Chapter 5 summarizes the findings in all chapters, and outlines future directions of interest.

1.6 References

(1) Zasloff, M. *Nature* **2002**, *415*, 389-395.
(2) Jiang, Y.; Lee, A.; Chen, J.; Ruta, V.; Cadene, M.; Chait, B. T.; MacKinnon, R. *Nature* **2003**, *423*, 33-41.
(3) Horbett, T., A.; Brash, J., L. *Proteins at interfaces II: Fundamentals and applications*; American Chemical Society: Washington, D.C., 1995; pp 561.
(4) Yebra, D. M.; Kiil, S.; Dam-Johansen, K. *Progress in Organic Coatings* **2004**, *50*, 75-104.
(5) Gray, J., J. *Curr. Op. Struct. Biol.* **2004**, *14*, 110-115.
(6) Nakanishi, K.; Sakiyama, T.; Imamura, K. *The Society for Biotechnology, Japan Biosci. Bioeng.* **2001**, *91*, 233-244.
(7) Chen, X.; Boughton, A., P.; Tesmer, J., J. G.; Chen, Z. *J. Am. Chem. Soc.* **2007**, *129*, 12658-12659.
(8) Chen, X.; Wang, J.; Boughton, A., P.; Kristalyn, C., B.; Chen, Z. *J. Am. Chem. Soc.* **2007**, *129*, 1420-1427.
(9) Chen, X.; Clarke, M., L.; Wang, J.; Chen, Z. *Int. J. Mod. Phys. B* **2005**, *19*, 691-713.
(10) Chen, X.; Wang, J.; Paszti, Z.; Wang, F.; Schrauben, J. N.; Tarabara, V. V.; Schmaier, A. H.; Chen, Z. *Anal Bioanal Chem* **2007**, *388*, 65-72.
(11) Chen, X.; Wang, J.; Sniadecki, J., J.; Even, M., A.; Chen, Z. *Langmuir* **2005**, *21*, 2662-2664.
(12) Even, M. A.; Wang, J.; Chen, Z. *Langmuir* **2008**, *24*, 5795-5801.
(13) Le Clair, S. V.; Nguyen, K.; Chen, Z. *J. Adhesion* **2009**, *85*, 484-511.
(14) Nguyen, K., Tan; Le Clair, S. V.; Ye, S.; Chen, Z. *J. Phys. Chem B* **2009**, *113*, 12169–12180.
(15) Pászti, Z.; Wang, J.; Clarke, M., L.; Chen, Z. *J. Phys. Chem. B* **2004**, *108*, 7779-7787.
(16) Wang, J.; Buck, S., M.; Chen, Z. *J. Phys. Chem. B* **2002**, *106*, 11666-11672.
(17) Wang, J.; Chen, X.; Clarke, M., L.; Chen, Z. *J. Phys. Chem. B* **2006**, *110*, 5017-5024.
(18) Wang, J.; Clarke, M., L.; Chen, X.; Even, M., A.; Johnson, W., C.; Chen, Z. *Surf. Sci.* **2005**, *587*, 1-11.
(19) Wang, J.; Even, M., A.; Chen, X.; Schmaier, A., H.; Waite, J., Herbert; Chen, Z. *J. Am. Chem. Soc.* **2003**, *125*, 9914-9915.
(20) Wang, J.; Lee, S.-H.; Chen, Z. *J. Phys. Chem. B* **2008**, *112*, 2281-2290.
(21) Ye, S.; Nguyen, K. T.; Le Clair, S. V.; Chen, Z. *J. Struct. Biol.* **2009**, *168*, 61-77.
(22) Chen, X.; Sagle, L., B.; Cremer, P., S. *J. Am. Chem. Soc.* **2007**, *129*, 15104-15105.
(23) Kim, J.; Koffas, T., S.; Lawrence, C., C.; Somorjai, G., A. *Langmuir* **2004**, *20*, 4640-4646.
(24) Kim, J.; Somorjai, G., A. *J. Am. Chem. Soc.* **2003**, *125*, 3150-3158.
(25) Koffas, T., S.; Kim, J.; Lawrence, C., C.; Somorjai, G., A. *Langmuir* **2003**, *19*, 3563-3566.
(26) Mermut, O.; Phillips, D., C.; York, R., L.; McCrea, K., R.; Ward, R., S.; Somorjai, G., A. *J. Am. Chem. Soc.* **2006**, *128*, 3598-3607.

(27) Phillips, D., C.; York, R., L.; Mermut, O.; McCrea, K., R.; Ward, R., S.; Somorjai, G., A. *J. Phys. Chem. C* **2007**, *111*, 255-261.
(28) York, R., L.; Browne, W., K.; Geissler, P., L.; Somorjai, G., A. *Israel J. Chem.* **2007**, *47*, 51-58.
(29) Jung, S., Y.; Lim, S., M.; Albertorio, F.; Kim, G.; Gurau, M., C.; Yang, R., D.; Holden, M., A.; Cremer, P., S. *J. Am. Chem. Soc.* **2003**, *125*, 12782-12786.
(30) Kim, G.; Gurau, M.; Kim, J.; Cremer, P., S. *Langmuir* **2002**, *18*, 2807-2811.
(31) Kim, G.; Gurau, M., C.; Lim, S., -M; Cremer, P., S. *J. Phys. Chem. B* **2003**, *107*, 1403-1409.
(32) Dreesen, L.; Humbert, C.; Sartenaer, Y.; Caudano, Y.; Volcke, C.; Mani, A. A.; Peremans, A.; Thiry, P. A.; Hanique, S.; Frere, J. M. *Langmuir* **2004**, *20*, 7201-7207.
(33) Dreesen, L.; Sartenaer, Y.; Humbert, C.; Mani, A. A.; Lemaire, J. J.; Methivier, C.; Pradier, C. M.; Thiry, P. A.; Peremans, A. *Thin Solid Films* **2004**, *464*, 373-378.
(34) Dreesen, L.; Sartenaer, Y.; Humbert, C.; Mani, A. A.; Methivier, C.; Pradier, C. M.; Thiry, P. A.; Peremans, A. *ChemPhysChem* **2004**, *5*, 1719-1725.
(35) Evans-Nguyen, K., M.; Fuierer, R., R.; Fitchett, B., D.; Tolles, L., R.; Conboy, J., C.; Schoenfisch, M., H. *Langmuir* **2006**, *22*, 5115-5121.
(36) Doyle, A. W.; Fick, J.; Himmelhaus, M.; Eck, W.; Graziani, I.; Prudovsky, I.; Grunze, M.; Maciag, T.; Neivandt, D. J. *Langmuir* **2004**, *20*, 8961-8965.
(37) Baugh, L.; Weidner, T.; Baio, J. E.; Nguyen, P. C.; Gamble, L. J.; Stayton, P. S.; Castner, D. G. *Langmuir* **2010**,
(38) Weidner, T.; Apte, J. S.; Gamble, L. J.; Castner, D. G. *Langmuir* **2010**, *26*, 3433-3440.
(39) Weidner, T.; Breen, N. F.; Drobny, G. P.; Castner, D. G. *J Phys Chem B* **2009**, *113*, 15423-15426.
(40) Weidner, T.; Samuel, N. T.; McCrea, K.; Gamble, L. J.; Ward, R. S.; Castner, D. G. *Biointerphases* **2010**, *5*, 9-16.
(41) Fu, L.; Ma, G.; Yan, E. C. *J. Am. Chem. Soc.* **2010**, *132*, 5405-5412.
(42) Shen, Y., R. *Principles of Nonlinear Optics*; Wiley-Interscience: New York, 1984; pp 563.
(43) Lambert, A., G.; Davies, P., B.; Neivandt, D., J. *App. Spec. Rev.* **2005**, *40*, 103.
(44) Guyot-Sionnest, P.; Hunt, J. H.; Shen, Y. R. *Phys. Rev. Lett.* **1987**, *59*, 1597-1600.
(45) Tamm, L., K.; Tatulian, S., A. *Quart. Rev. Biophys.* **1997**, *30*, 365-429.
(46) Tatulian, S., A. *Biochemistry* **2003**, *42*, 11898-11907.
(47) Vigano, C.; Manciu, L.; Buyse, F.; Goormaghtigh, E.; Ruysschaert, J.-M. *Biopolymers* **2000**, *55*, 373-380.
(48) Ong, T., Hui; Davies, P., B.; Bain, C., D. *Langmuir Langmuir* **1993**, *9*, 1836-1845.
(49) Wang, J.; Chen, C.; Buck, S., M.; Chen, Z. *J. Phys. Chem. B* **2001**, *105*, 12118-12125.
(50) Esenturk, O.; Walker, R. A. *J. Chem. Phys.* **2006**, *125*, 174701.
(51) Holman, J.; Davies, P. B.; Nishida, T.; Ye, S.; Neivandt, D. J. *J Phys Chem B* **2005**, *109*, 18723-18732.
(52) Beattie, D., A.; Haydock, S.; Bain, C., D. *Vib. Spectrosc.* **2000**, *24*, 109-123.
(53) Ma, G.; Liu, J.; Fu, L.; Yan, E. C. Y. *Appl. Spectrosc.* **2009**, *63*, 528-537.

(54) Bain, C., D.; Davies, P., B.; Ong, T., H.; Ward, R., N.; Brown, M., A. *Langmuir* **1991**, *7*, 1563-1566.
(55) Hecht, E. *Optics (4th Edition)*; Addison Wesley: Boston, MA, 2001;
(56) Hirose, C.; Akamatsu, N.; Domen, K. *The Journal of Chemical Physics* **1992**, *96*, 997.
(57) Hirose, C.; Yamamoto, H.; Akamatsu, N.; Domen, K. *The Journal of Physical Chemistry* **1993**, *97*, 10064-10069.
(58) Hirose, C.; Akamatsu, N.; Domen, K. *App. Spec.* **1992**, *46*, 1051-1072.
(59) Chen, X.; Tang, H.; Even, M., A.; Wang, J.; Tew, G., N.; Chen, Z. *J. Am. Chem. Soc.* **2006**, *128*, 2711-2714.
(60) Yang, C., S. -C; Wilson, P., T.; Richter, L., J. *Macromolecules* **2004**, *37*, 7742-7746.
(61) Moad, A., J.; Simpson, G., J. *Journal of Physical Chemistry B* **2004**, *108*, 3548-3562.
(62) McClelland, A. *Applied Physics* **2009**, *PhD*,
(63) Nguyen, K. T.; King, J. T.; Chen, Z. *J Phys Chem B* **2010**, *114*, 8291-8300.
(64) Ye, S.; Nguyen, K. T.; Chen, Z. *J Phys Chem B* **2010**, *114*, 3334-3340.
(65) Moad, A. J.; Moad, C. W.; Perry, J. M.; Wampler, R. D.; Goeken, G. S.; Begue, N. J.; Shen, T.; Heiland, R.; Simpson, G. J. *J. Comput. Chem.* **2007**, *28*, 1996-2002.
(66) Moad, A., J.; Simpson, G., J. *J. Phys. Chem A* **2005**, *109*, 1316-1323.
(67) Perry, J., M.; Moad, A., J.; Begue, N., J.; Wampler, R., D.; Simpson, G., J. *J. Phys. Chem B* **2005**, *109*, 20009-20026.
(68) Clarke, M. L.; Wang, J.; Chen, Z. *J Phys Chem B* **2005**, *109*, 22027-22035.
(69) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187-217.
(70) MacKerell, A., D. Jr; Bashford, D.; Bellott, M.; Dunbrack, R., L. Jr; Evanseck, J., D.; Field, M., J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F., T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D., T.; Prodhom, B.; Reiher, W., E. III; Roux, B.; Schlenkrich, M.; Smith, J., C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586-3616.
(71) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1988**, *61*, 2635-2638.
(72) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195-1198.
(73) Kumar, S.; Bouzida, D.; Swendsen, R., H.; Kollman, P., A.; Rosenberg, J., M. *J. Comp. Chem.* **1992**, *13*, 1011-1021.
(74) Grossfield, A. "An implementation of WHAM: the Weighted Histogram Analysis Method". http://membrane.urmc.rochester.edu/Software/WHAM/WHAM.html **April 14, 2008**
(75) Roux, B. *Comp. Phys. Commun.* **1995**, *91*, 275-282.
(76) Bartels, C.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 865-880.
(77) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187-199.
(78) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H., J. C. *J. Comput. Phys.* **1977**, *23*, 327-341.
(79) Chen, J.; Brooks, C. L. r.; Khandogin, J. *Curr. Opin. Struct. Biol.* **2008**, *18*, 140-148.

(80) Piana, S.; Sarkar, K.; Lindorff-Larsen, K.; Guo, M.; Gruebele, M.; Shaw, D. E. *J. Mol. Biol.* **2011**, *405*, 43-48.
(81) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. *Science* **2010**, *330*, 341-346.
(82) Lee, H.; Larson, R. G. *J Phys Chem B* **2008**, *112*, 7778-7784.
(83) Vellore, N. A.; Yancey, J. A.; Collier, G.; Latour, R. A.; Stuart, S. J. *Langmuir* **2010**, *26*, 7396-7404.
(84) Chen, Z.; Ward, R.; Tian, Y.; Malizia, F.; Gracias, D. H.; Shen, Y. R.; Somorjai, G. A. *J Biomed Mater Res* **2002**, *62*, 254-264.
(85) Watry, M. R.; Richmond, G. L. *J. Phys. Chem. B* **2002**, *106*, 12517-12523.
(86) Kim, J.; Cremer, P., S. *ChemPhysChem* **2001**, *2*, 543-546.
(87) Kim, J.; Kim, G.; Cremer, P., S. *Langmuir* **2001**, *17*, 7255-7260.
(88) Wang, J.; Buck, S., M.; Even, M., A.; Chen, Z. *J. Am. Chem. Soc.* **2002**, *124*, 13302-13305.
(89) Wang, J.; Buck, S., M.; Chen, Z. *Analyst* **2003**, *128*, 773-778.
(90) York, R., L.; Mermut, O.; Phillips, D., C.; McCrea, K., R.; Ward, R., S.; Somorjai, G., A. *J. Phys. Chem. C* **2007**, *111*, 8866-8871.
(91) Wang, J.; Clarke, M., L.; Zhang, Y.; Chen, X.; Chen, Z. *Langmuir* **2003**, *19*, 7862-7866.
(92) Arkin, I. *Curr. Op. Chem. Biol.* **2006**, *10*, 394-401.
(93) Marsh, D. *J. Mol. Biol.* **2004**, *338*, 353-367.
(94) Moore, W., H.; Krimm, S. *Proceedings of the National Academy of Sciences Proc. Natl. Acad. Sci. USA* **1975**, *72*, 4933-4935.
(95) Wang, J.; Chen, X.; Clarke, M., L.; Chen, Z. *Proc. Natl. Acad. Sci.* **2005**, *102*, 4978-4983.
(96) Wang, J.; Paszti, Z.; Even, M., A.; Chen, Z. *J. Phys. Chem. B* **2004**, *108*, 3625-3632.
(97) Knoesen, A.; Pakalnis, S.; Wang, M.; Wise, W., D.; Lee, N.; Frank, C., W. *Journal of IEEE J Sel. Top. Quant.* **2004**, *10*, 1154-1163.
(98) Boughton, A. P.; Andricioaei, I.; Chen, Z. *Langmuir* **2010**, *26*, 16031-16036.
(99) Ye, S.; Nguyen, K. T.; Boughton, A. P.; Mello, C. M.; Chen, Z. *Langmuir* **2009**, *26*, 6471-6477.
(100) Nguyen, K. T.; Le Clair, S. V.; Ye, S.; Chen, Z. *J. Phys. Chem. B* **2009**, *113*, 12358-12363.
(101) Nguyen, K. T.; Soong, R.; Lm, S. C.; Waskell, L.; Ramamoorthy, A.; Chen, Z. *J. Am. Chem. Soc.* **2010**, *132*, 15112-15115.
(102) Wang, J.; Paszti, Z.; Clarke, M., L.; Chen, X.; Chen, Z. *J. Phys. Chem. B* **2007**, *111*, 6088-6095.
(103) McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature* **1977**, *267*, 585-590.
(104) Kandasamy, S. K.; Larson, R. G. *Chem. Phys. Lipids* **2004**, *132*, 113-132.
(105) Sun, Y.; Welsh, W., J.; Latour, R., A. *Langmuir* **2005**, *21*, 5616-5626.
(106) Grossfield, A.; Feller, S. E.; Pitman, M. C. *Proteins: Struct., Funct., Bioinf.* **2007**, *67*, 31-40.
(107) Liu, S., M.; Haynes, C., A. *J. Colloid Interface Sci.* **2004**, *275*, 458-469.
(108) Euston, S. *Current Opinion in Colloid & Interface Science* **2004**, *9*, 321-327.

(109) Clancy, T., C.; Jang, J., Hwan; Dhinojwala, A.; Mattice, W., L. *J. Phys. Chem. B* **2001**, *105*, 11493-11497.
(110) Lee, H.; Larson, R., G. *Molecules* **2009**, *14*, 423-438.
(111) Levy, Y.; Onuchic, J., N. *PNAS* **2004**, *101*, 3325-3326.
(112) Feig, M.; Brooks, C. L. r. *Curr. Opin. Struct. Biol.* **2004**, *14*, 217-224.
(113) Grossfield, A. *Current Topics in Membranes* **2008**, *Computational Modeling of Membrane Bilayers*, 131-157.
(114) Edholm, O.; Jähnig, F. *Biophysical Chemistry* **1988**, *30*, 279-292.
(115) Huang, X.; Margulis, C., J.; Berne, B., J. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 11953-11958.
(116) Chipot, C.; Maigret, B.; Pohorille, A. *Proteins: Structure, Function, and Bioinformatics* **1999**, *36*, 383-399.
(117) Raffaini, G.; Ganazzoli, F. *J. Mat. Sci: Mat. in Med.* **2007**, *18*, 309-316.
(118) Ganazzoli, F.; Raffaini, G. *Phys Chem Chem Phys* **2005**, *7*, 3651-3663.
(119) Sun, Y.; Dominy, B., N.; Latour, R., A. *J. Comput. Chem.* **2007**, *28*, 1883-1892.
(120) Im, W.; Brooks, C., L. III. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6771-6776.
(121) Collier, G.; Vellore, N. A.; Latour, R. A.; Stuart, S. J. *Biointerphases* **2009**, *4*, 57-64.
(122) Hünenberger, P., H.; McCammon, J., A. *Biophys. Chem.* **1999**, *78*, 69-88.
(123) Raut, V., P.; Agashe, M., A.; Stuart, S., J.; Latour, R., A. *Langmuir* **2005**, *21*, 1629-1639.
(124) Lee, H.; Larson, R., G. *Journal of Physical Chemistry B J. Phys. Chem. B* **2006**, *110*, 18204-18211.
(125) Raffaini, G.; Ganazzoli, F. *Phys Chem Chem Phys* **2006**, *8*, 2765-2772.
(126) Lee, C.-Y.; McCammon, J., Andrew; Rossky, P., J. *J. Chem. Phys.* **1984**, *80*, 4448-4455.
(127) Brown, E. C.; Mucha, M.; Jungwirth, P.; Tobias, D. J. *J Phys Chem B* **2005**, *109*, 7934-7940.
(128) Lee, H.; Baker, J., R.; Larson, R., G. *Journal of Physical Chemistry B J. Phys. Chem. B* **2006**, *110*, 4014-4019.
(129) Kelly, C. V.; Leroueil, P. R.; Nett, E. K.; Wereszczynski, J. M.; Baker, J. R. J.; Orr, B. G.; Banaszak Holl, M. M.; Andricioaei, I. *J Phys Chem B* **2008**, *112*, 9337-9345.
(130) Kelly, C. V.; Leroueil, P. R.; Orr, B. G.; Banaszak Holl, M. M.; Andricioaei, I. *J Phys Chem B* **2008**, *112*, 9346-9353.
(131) Lee, S., J.; Park, K. *J. Vac. Sci. Technol. A* **1994**, *12*, 2949-2955.
(132) Lu, D., R.; Lee, S., J.; Park, K. *Journal of Biomaterials Science, Polymer Edition* **1992**, *3*, 127-147(21).
(133) Cormack, A. N.; Lewis, R. J.; Goldstein, A. H. *J. Phys. Chem. B* **2004**, *108*, 20408-20418.
(134) Knotts, T. A. t.; Rathore, N.; de Pablo, J. J. *Proteins* **2005**, *61*, 385-397.
(135) Boczko, E., M.; Brooks, C., L. III. *J. Phys. Chem.* **1993**, *97*, 4509-4513.
(136) Lee, J.; Im, W. *Chem. Phys. Lett.* **2007**, *441*, 132-135.
(137) Lee, J.; Im, W. *Phys. Rev. Lett.* **2008**, *100*, 018103.

Chapter 2: Surface Orientation of Magainin 2: Molecular Dynamics Simulation and Sum Frequency Generation Vibrational Spectroscopic Studies

2.1 Introduction

A variety of biosensing, antibiofouling, and biomaterial applications depend on understanding the early events that occur when peptides adsorb onto a polymer surface layer. To date, adsorption has been studied by techniques such as surface plasmon resonance (SPR), ellipsometry, atomic force microscopy (AFM), X-ray photoelectron spectroscopy (XPS), and others, but the common denominator in these techniques is that they are not able to directly probe interfacial molecular orientation. Vibrational spectroscopic techniques are a promising way to measure orientation, which may provide direct information about the interactions involved in adsorption. In previous studies, the complexity of the system that could be studied was limited by the small number of measurements obtainable and/or unique sample requirements of techniques such as ATR-FTIR or surface enhanced raman scattering (SERS).

SFG provides a way to directly probe proteins *in situ* in biologically relevant environments, and information about molecular orientation can be extracted by taking advantage of recently developed data analysis methodologies for helical peptides. This may provide a direct means to probe how peptides interact with the surface. Yet even seemingly simple interfaces can be the home to unexpectedly complicated behavior, as revealed by polarized spectra focused on the Amide I vibrational modes.

In order to explore this behavior and test possible explanations, molecular dynamics simulations were performed as a supplement to experimental measurements. The origins of the adsorption behavior were explored in terms of free energy sampling, surface area burial, and the contribution of rotational entropy. Increases in computing power and progress in simulation algorithms have made it possible to study interfacial systems, yet these simulations are still limited to short (sub-microsecond) timescales. Thus I believe that using experimentally suggested configurations as a starting point for molecular dynamics will ultimately enable a more detailed picture of adsorption behavior than could be determined from either MD or SFG alone.

In this chapter, I report studies on a simplified planar hydrophobic/hydrophilic interface consisting of α-helical magainin 2 peptides adsorbed onto polystyrene (PS) in an aqueous environment. These studies were designed to lend insights into molecular orientation, and thus advance the ability to interpret complex SFG spectra. Here magainin 2 serves as a model α-helix, and PS as a model polymer surface. This reduction in complexity compared to more complicated membrane systems allows for molecular dynamics simulations and free energy sampling that would not be feasible in an atomistically detailed lipid bilayer. As shown in Figure 2-1, magainin 2 (PDB ID 2mag, sequence GIGKFLHSAKKFGKAFVGEIMNS) is a 23-residue, facially amphiphilic peptide, and is therefore expected to be helical at hydrophobic/hydrophilic interfaces.

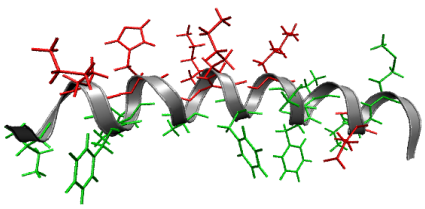


Figure 2-1: Magainin 2 (PDB ID 2mag) is an alpha helix in which hydrophobic residues (green) and charged residues (red) are segregated to different faces of the peptide.

2.2 Theory

2.2.1 SFG Orientation Analysis for α-helical Peptides

As described in section 1.2.3, polarized SFG spectra can be used to determine the orientation of a vibrational mode. In the case of α-helical peptides, Amide I signals originate from the well-ordered peptide backbone, and so the orientation can be related to the overall orientation of the molecule.

Relations allowing determination of molecular orientation for α-helices have been derived previously[1-3], and are presented below. There are two Amide I vibrational modes (denoted A and E1) whose frequencies cannot be resolved by the resolution of our spectrometer. The total measured response in a given polarization combination is considered as the sum of the two individual modes.

For the A mode:

$$\chi^{(2)}_{xxz,A} = \frac{1}{2} * N_s[(1+r)\langle cos\,\theta\rangle - (1-r)\langle cos^3\,\theta\rangle]\beta_{ccc} \qquad \textit{Equation 2-1}$$

$$\chi^{(2)}_{zzz,A} = N_s[r\langle cos\,\theta\rangle + (1-r)\langle cos^3\,\theta\rangle]\beta_{ccc} \qquad \textit{Equation 2-2}$$

For the E1 mode:

$$\chi^{(2)}_{xxz,E} = -N_s[\langle cos\,\theta\rangle - \langle cos^3\,\theta\rangle]\beta_{aca} \qquad \textit{Equation 2-3}$$

$$\chi^{(2)}_{zzz,E} = 2N_s[\langle cos\,\theta\rangle - \langle cos^3\,\theta\rangle]\beta_{aca} \qquad \textit{Equation 2-4}$$

In the above relations, N_s denotes the surface coverage of peptide molecules, and β_{ijk} denotes the molecular hyperpolarizability tensor component probed in the selected polarization combination. The requisite values of β_{ijk} have been determined for linear α-helices[3], and calculations for more complicated proteins structures will be presented in Chapter 3:.

The experimental measured quantity $\chi^{(2)}_{eff}$ is thus related to the calculated quantities as follows:

$$\chi^{(2)}_{eff,ssp} = L_{xxz}\chi^{(2)}_{xxz,tot} = L_{xxz}\left(\chi^{(2)}_{xxz,A} + \chi^{(2)}_{xxz,E}\right) \quad \textit{Equation 2-5}$$

$$\chi^{(2)}_{eff,ppp} = L_{zzz}\chi^{(2)}_{zzz,tot} = L_{xxz}\left(\chi^{(2)}_{zzz,A} + \chi^{(2)}_{zzz,E}\right) \quad \textit{Equation 2-6}$$

where L_{ijk} values represent Fresnel factors to correct for the fraction of light transmitted or reflected at the interface. In my studies, the ratio L_{zzz}/L_{xxz} falls in the range 0.84-0.92, depending on small differences in experimental geometry.

When the ratio $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ is taken, the dependence of signal intensity on surface coverage cancels out, making the experimentally fitted ppp/ssp signal strength ratio a useful parameter for orientation analysis. The variation of this curve is shown in Figure 2-2.

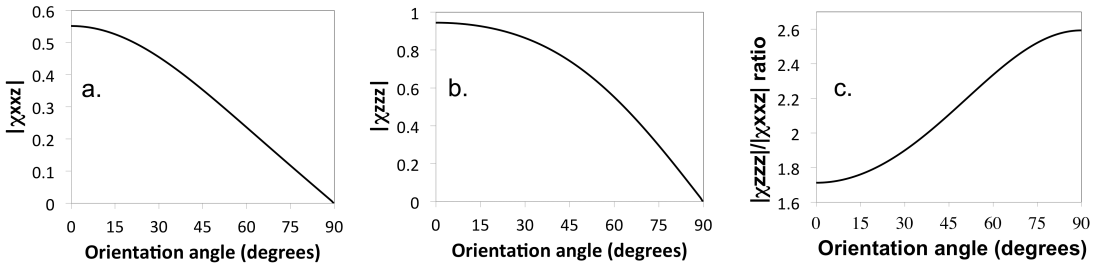


Figure 2-2: The magnitude of a) $\chi^{(2)}_{xxz}$ and b) $\chi^{(2)}_{zzz}$ shows variation in fitted signal strength as a function of helix tilt angle. c) The fitted ppp/ssp signal strength ratio is a useful parameter for determining molecular orientation.

2.2.2 Free Energy Sampling: The Weighted Histogram Analysis Method

In Section 2.3.2.3, the problem of determining the overall most stable state is discussed. A one-point potential energy function evaluation using CHARMM or similar packages does not incorporate entropic contributions, so that calculation of the overall free energy change requires detailed sampling. However, normal sampling tends to

overemphasize energetically favored regions of phase space (local minima), and hence direct sampling in short simulations cannot be used to obtain an accurate estimate of the entropy. To encourage sampling of regions that are energetically disfavored, the technique of umbrella sampling[4] employs harmonic restraining potentials of the form $U = k_{umb}(x - x_0)^2$ in a series of windows along the reaction coordinate, thus greatly enhancing the range of states sampled during the course of simulations, and commensurately improving estimates of the energy profile. The size of the force constants, and number of sampling windows/initial positions, are chosen somewhat arbitrarily as needed to ensure good sampling and overlap between windows.

However, this creates a complication, as the variety of windows and force constants chosen makes it difficult to combine all simulation data for a single uniform and reliable estimate of the free energy profile. To overcome this problem, a means of representing the probability distribution of states using histograms was devised.[5, 6] The Weighted Histogram Analysis Method (WHAM) equations are optimized iteratively to produce a set of coupling parameters that allow free energy information to be derived from the overlap between windows sampling adjacent regions along the reaction coordinate of interest.[7-10] This provides an efficient way to ensure adequate sampling across the entire reaction coordinate, and has become a widely used tool for the estimation of free energy changes in protein folding, subunit binding, molecular deformation, and other processes.

2.3 Materials and Methods

2.3.1 SFG Measurements

A PS thin film was prepared by spin-coating at 2500 rpm from a 2% wt solution of PS (or deuterated PS) in toluene on an IR-grade CaF_2 prism. The prism was pre-cleaned via soaking overnight in solutions of toluene and Contrex (Decon Labs, King of Prussia, PA) to remove contaminants, followed by 20 min in methanol and plasma treatment to remove any remaining hydrocarbons. No SFG signals were observed in the C-H stretching or Amide I frequency regions from the cleaned CaF_2 substrates prior to deposition of the polymer film.

Prior to collection of peptide signals, the polymer surface was contacted with Millipore water for 15 min and allowed to equilibrate while the laser shutters were blocked. Then SFG spectra were collected from the polymer surface in both the Amide I and C-H stretching frequency regions. For deuterated PS samples, additional SFG spectra were also collected as needed in the C-D stretching region. A "near" total reflection geometry was used.[11]

Stock solutions of magainin 2 (Genscript Inc) were prepared by dissolving 1 mg of peptide in 1 mL of Millipore water. Sufficient amounts of this stock solution were added to the water already in contact with the polymer surface to produce final peptide concentrations of approximately 200, 400, and 800 nM. In each case, the polymer-peptide system was allowed to equilibrate for 20 min, after which SFG spectra were collected in the ssp and ppp polarization combinations across multiple spectral ranges. The resulting spectra remained stable over time, suggesting that sufficient time had elapsed for peptides to diffuse and organize at the polymer/peptide solution interface.

2.3.2 Molecular Dynamics Simulations

2.3.2.1 The Model Polystyrene Surface

Polymer surface simulations are limited in part by system size, specifically by the number of atoms required to adequately represent the solvent, surface, and peptide. Thus, although some work has been done on atomistic-detail simulations of PS[12-14], the relatively long equilibration times (10-20 ns) in the presence of explicit solvent required are non-trivial, which complicates setup of any studies involving peptide adsorption.[15] One common strategy employed to overcome this limitation is to make use of a simplified surface model (such as graphite[16] or SiO_2[17]). However, these models are unable to directly capture variations in surface topography that may have important consequences for hydration or protein orientation. Likewise, the use of uniform implicit solvent models, though common, may fail to capture interfacial phenomena or small changes in peptide orientation due to dewetting behavior.[18]

As a compromise, I employed a simplified representation of the polymer surface that allowed for smaller system sizes than the use of full polymer chains, thus facilitating the use of explicit solvent and incorporation of certain aspects of the polymer surface (such as local reorganization of surface groups or the ability to vary surface roughness) not captured by more abstract, rigid model surfaces. This simplified, but still atomistic-level representation employs an array of individual ethylbenzene monomers to represent units of the PS backbone, with unit spacings adjusted to reproduce some basic experimentally measured properties of the system (such as a bulk density of 1.05 g/mL and the known preferred orientation of surface aromatic rings in water).[19]

To form the surface, an array of individual ethylbenzene molecules was arranged in a grid. Adjacent points in this grid were offset by half the length of an individual

monomer, with residue spacing adjusted to reproduce the bulk density. To avoid steric clashes, the monomers in the initial model did not lie flat, but were instead rotated so that the plane of the monomer was inclined at an angle of 65° with respect to the surface normal. The spacing between individual monomer units was 6.5 Å lengthwise, 5.5 Å in width, and 6.0 Å in height (for additional layers of surface residues). Several layers of the surface were replicated to form a bulk region, since steric effects from the presence of the bulk were necessary to preserve surface group orientation. This bulk region served a secondary function by preventing unrealistic penetration of explicit water molecules below the surface. A single alpha carbon in each monomer of the surface and bulk was constrained with a harmonic potential of 1 kcal/mol/Å^2 to mimic the presence of a backbone, and to prevent the monomers from moving unreasonably relative to each other. For a sphere of 36 Å radius, approximately 115 ethylbenzene units were required to represent the surface, and an additional 380 residues were used for the polymer bulk.

As discussed in section 1.4.3.1, periodic boundary conditions are not ideally applicable to the two dimensional interface of interest in this study. Instead, a spherical model system was chosen that was more amenable to simulations on limited computational hardware. Explicit solvent was added using approximately 4300 molecules of TIP3P water above the surface[20], creating a sphere of radius ~36 Å suitable for fully solvating the peptides studied. No ions were included in this system.

The system was then modeled as a droplet in a vacuum[21, 22] using a spherical quartic boundary potential to constrain the system inside the 36 Å radius.[23] This system consists of half water molecules and half surface/bulk residues (modeled by the ethylbenzene monomers), as shown in Figure 2-3.

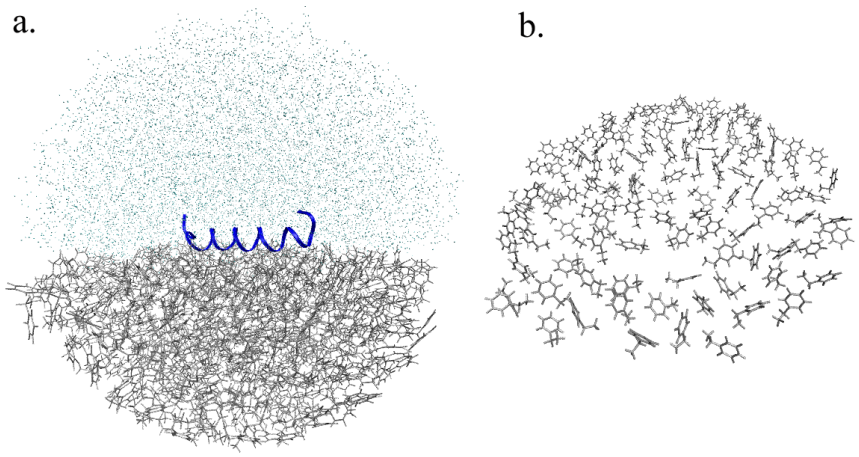


Figure 2-3: The spherical model for solvent and surface used for molecular dynamics simulations. a) Ethylbenzene residues are shown as grey lines, and the solvent molecules are represented as single points (light blue). The radius of the model system used was approximately 36 Å. The helical peptide is in the center, represented by the continuous coiled line (dark blue). b) The surface as an array of monomers (top view).

For this droplet model, fewer solvent molecules were required than would be the case for periodic boundary conditions, due to simple geometric considerations. In total, ~21,000 atoms were used to represent the solvent, surface, and peptide. A comparable simulation using periodic boundary conditions would have required at minimum 30% more atoms for a surface of comparable area, or 70% more atoms for a surface of comparable width. Other factors being equal, the length of a simulation is dictated by the need to calculate nonbonded interactions between all pairs of atoms in the system[24], and a 30% reduction in the number of atoms will lead to very significant reductions in the amount of computing time required to obtain simulations of useful length. For this reason, the spherical system is more amenable to use with limited computational resources, and enables more detailed free energy sampling simulations than would otherwise be possible.

2.3.2.2 Equilibration and Dynamics

The model solvent/surface system was built and equilibrated using version c34b2 of the CHARMM molecular dynamics package[25] with the CHARMM22 all-atom force field for proteins.[26]

Heating and early equilibration of the surface system were performed for 400 ps using the leapfrog Verlet algorithm, rescaling velocities every 1 ps to a final temperature of 298 K. The resulting initial coordinates for the solvent and surface were used in all peptide adsorption simulations described.

Following heating of the solvent/surface system, the peptide was placed at the interface, and overlapping waters were deleted. Peptide coordinates were obtained from the Protein Databank (PDB), and the exact position of the peptide was varied depending on the simulation conditions chosen. Peptide coordinates were held fixed as an additional 500 ps of equilibration was then performed for the solvent and surface in the presence of peptide.

For production runs, dynamics were performed using a Nosé-Hoover thermostat at 298 K with the velocity Verlet algorithm; the exact length of the simulation varied (as described below). To avoid discretization errors when evaluating the potential energy of the system, the time between steps of dynamics must be shorter than the fastest bond vibrational frequency, typically that of the C-H bonds. The SHAKE algorithm was used to constrain the length of all bonds involving hydrogen atoms, and a 2 fs timestep was used for dynamics.

2.3.2.3 Free Energy Calculations

The orientational preference of magainin 2 was studied by using umbrella sampling[4] and the WHAM[8] to calculate the free energy as a function of angular

orientation relative to the surface normal. The peptide was placed at a defined orientation relative to the surface, and the angle between the helical axis backbone vector and the surface normal was restrained using a biasing harmonic potential. The surface normal was defined using two fixed-position and non-interacting dummy atoms. The backbone vector was defined in terms of two centers of mass for groups of backbone atoms. In order to prevent this constraint from causing artificial deformation of the structure, the backbone dihedral angles of helical residues 4-19 were held constrained using a strong harmonic potential.

Simulations were divided into 5° increments for sampling purposes (with an additional window at tilt=72°). Each window was sampled for 3 ns of dynamics, with force constants ranging from 400 kcal/mol/radian^2 to 1,500 kcal/mol/radian^2 as needed to ensure good sampling and convergence in the chosen window. The highest force constants were required for orientation angles of 70-80°. For other angles, force constants of 400-600 kcal/mol/radian^2 were sufficient). Data was then unbiased using WHAM[8-10] as implemented by Grossfield[7] in order to create a plot of free energy as a function of helix orientation angle. For these simulations, an angle of 0° corresponds to the helix standing up, with the N-terminal end of the peptide closest to the surface.

2.4 Results and Discussion

2.4.1 Simulations of Single Peptides

A series of exploratory molecular dynamics simulations were performed in which magainin 2 was placed at a variety of orientations relative to the surface (perpendicular, parallel, or tilted, with hydrophobic residues facing towards or away from the surface in the latter two cases). In all cases, the peptide adopted the same final orientation within 10

ns, parallel to the surface with hydrophobic residues facing away from the solvent to maximize burial of hydrophobic surface area (Figure 2-4).

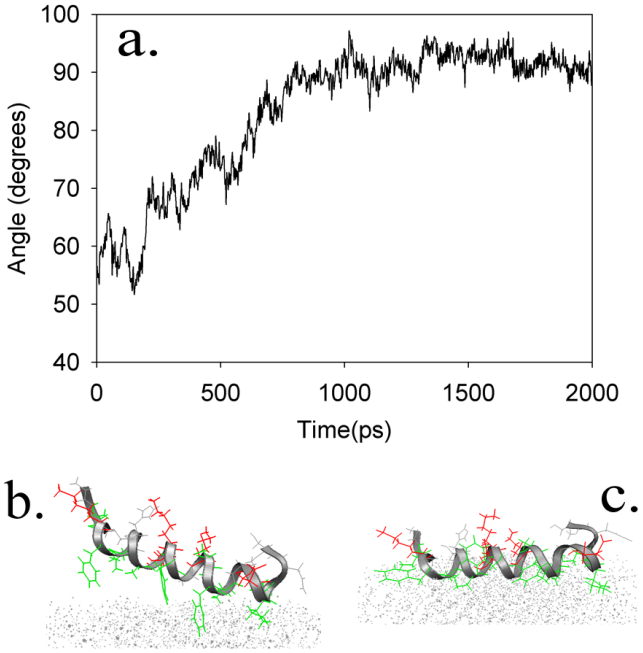


Figure 2-4: a) Change in the orientation angle from an initial orientation (tilted) to a final orientation (lying down) during a 2ns simulation, b) The peptide as it begins to lie down on the surface. (For clarity in viewing the protein, the surface is represented as an array of points). c) A characteristic final orientation of the magainin 2 peptide after 2 ns of simulation.

During the length of the simulation, the helix backbone radius of gyration and end-to-end distance remained consistent, suggesting that secondary structure remained intact. If unfolding does occur, it is likely to occur on substantially longer timescales.[27] The helical structure enforces excellent segregation of hydrophobic and charged residues, and therefore I believe that unfolding at the hydrophobic/hydrophilic interface will be unlikely even over longer timescales.

Despite the apparent orientation preference revealed by unconstrained simulations, the limited timescales accessible to molecular dynamics simulations

prevented a clear conclusion as to whether this was the overall most stable state. Also, experimental studies on other peptides have demonstrated that orientation is more realistically modeled using some non-uniform distribution function which may be better understood given an estimate of the free energy difference between various orientations.

To that end, free energy sampling was performed as described in Section 2.3.2.4. Constraints were varied in order to provide sufficient sampling across all windows, and a free energy profile was calculated (Figure 2-5).

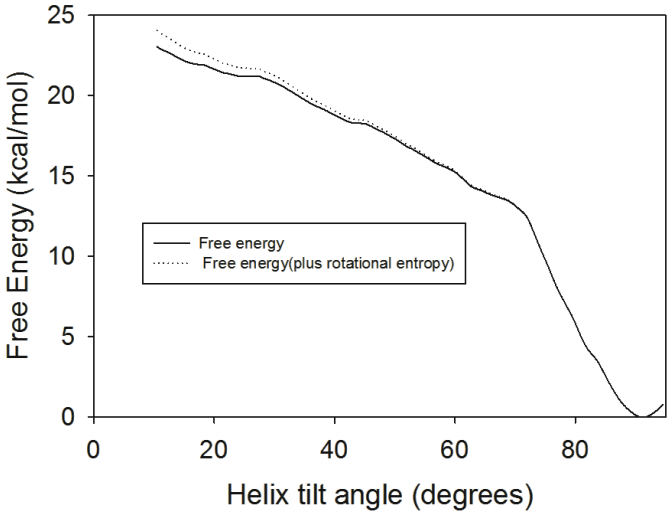


Figure 2-5: Free energy profile for magainin 2 as a function of helix angle relative to the surface normal. An angle of 0^o corresponds to the peptide oriented perpendicular to the surface. The solid line represents the calculated free energy as a function of helix angle; the dotted line is the free energy curve with the addition of the rotational entropy term derived by Lee and Im.[28,29] (see text). The actual free energy curve lies somewhere between these two.

The drop in free energy at high tilt angles is partially due to the extra overall rotational entropy in the azimuthal direction; this arises from changes in the volume elements sampled due to variations in solid angle. Lee and Im have derived[28, 29] an analytical formula for the contribution of this rotational entropy to free energy: in the

case of free rotation in the azimuthal direction, this is equal to $-KTln|\sin\theta|$. In my simulations, the actual variation in azimuthal angle is small but nonzero, and hence so is the included contribution of rotational entropy. The curves shown in Figure 2-5 represent both the original free energy profile (which inherently includes a limited amount of azimuthal angle variation) and the free energy curve calculated by adding the entropy due to completely random azimuthal rotation. The actual curve should thus lie in between these two. Unlike the case of the transmembrane peptide studied by Lee and Im, in this case the contribution of rotational entropy has a relatively minor effect on the preferred orientation. This is due to the different peptide and model system in question. The calculated free energy profile revealed that the preference for a horizontal peptide orientation dominates at room temperature, and that the slope (and magnitude) of the free energy change correlates reasonably well with estimates of the free energy change due to burial of hydrophobic surface area.[30-33] Indeed, this horizontal orientation was only adopted when hydrophobic residues were oriented towards the surface, further suggesting that peptide orientation is driven by burial of hydrophobic residues for this facially amphiphilic structure. (Figure 2-6)

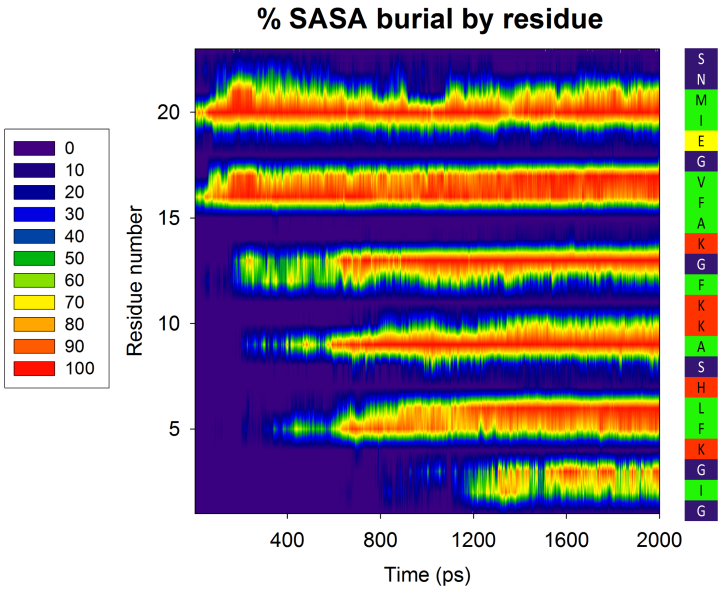


Figure 2-6: Burial of hydrophobic surface area drives the final orientation of the peptide. Left and center: colors indicate the percent change in solvent accessible surface area (SASA) of each residue during a 2 ns simulation, with an initial peptide orientation of 60^o. Sidebar, right: the peptide sequence, colored by residue type (according to the scheme in Figure 2-1)

As an additional check on the hypothesis that hydrophobic surface area burial drives peptide orientation, the results in Figures 2-4, 2-5, and 2-6 can be compared to a simple model of hydrophobic surface area burial in which groups of hydrophobic residues on every turn of the helix are considered as single points spaced 6 Å apart on a line (Figure 2-7). By considering a given cluster of hydrophobic residues as "buried" when water can no longer be accommodated between the peptide and surface (a peptide-surface distance < 2 Å), then hydrophobic groups placed at 4, 10, 16, 22, 28, and 34 Å would yield expected orientations of 60, 78, 83, 85, 86, and 87 degrees respectively. These values are in reasonable agreement with points where sharp changes in orientation or free energy occur. Minor deviations can be attributed to the fact that the hydrophobic

residues are bulky aromatic rings rather than single points that do not extend out away from the backbone.

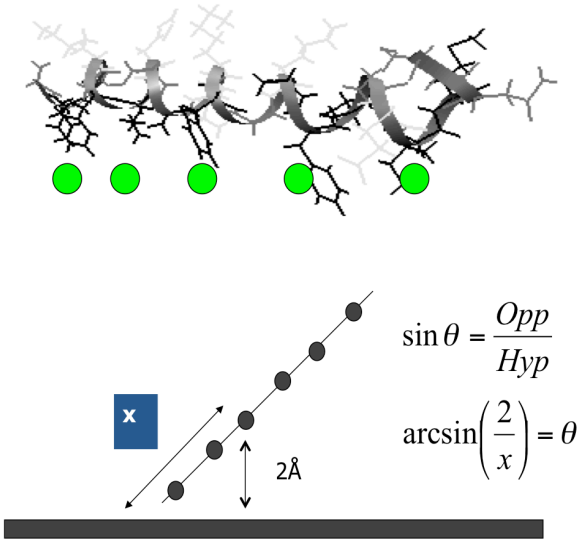


Figure 2-7: A simple scheme to predict the most stable orientations based on hydrophobic surface area burial alone. The angles predicted are in reasonable agreement with those from simulation results.

Thus from my simulation results, it is reasonable to assume that peptide-surface interactions are dictated purely by the burial of hydrophobic surface area, and that the peptide would display a strong, singular tendency to lie down parallel to the plane of the surface.

2.4.2 Experimental Results

In order to determine whether my above simulations captured important aspects of adsorption, the interface was probed experimentally using SFG spectroscopy. Spectra were collected from the PS/magainin 2 solution interface for a range of magainin concentrations (Figure 2-8). At the lowest magainin concentration range (200 nM), a

clear reduction in interfacial water signal was observed, indicating that peptides adsorbed to the interface. For a fairly smooth polymer surface, simulation results suggest that water molecules may orient preferentially due to hydrophobic considerations.[34] It is also possible that this water signal is induced by trace amounts of chemical impurities in the polymer (such as excess initiator) that impart a slight negative charge to the polymer surface.[35] Regardless, the adsorption of magainin 2 peptides to the surface results in a net disordering of water molecules and a drop in observed water signals. The drop in water O-H stretching signal intensity serves as an indication that peptides are adsorbing to the surface, despite the lack of observable peptide Amide I or C-H stretching signals at the lowest peptide concentration. At this concentration, any SFG signals generated from the interfacial peptides are below the detection limit of our spectrometer. As the magainin 2 concentration was increased to 400 nM, weak SFG Amide I signals were observable in the ppp polarization combination, but no Amide I signals were detected from the ssp polarization combination. The SFG signal intensity in the ppp polarization was roughly 3-4 times stronger than the detection limit. Although a single ppp measurement is insufficient to characterize molecular orientation, the lack of detectable ssp signal indicates that it must be at least 3-4 times weaker than the ppp signals. Based on previously published methods for SFG orientation analysis for interfacial α-helical structures[3], these ratios would only be observed if the peptides were oriented at a large angle relative to the surface normal (and thus, mostly lying down) (Figure 2-2). In the 400 nM concentration range, weak SFG C-H stretching signals from magainin 2 side chains were also observed, which originate from hydrophobic residues becoming ordered at the PS/solution interface.[17, 36]

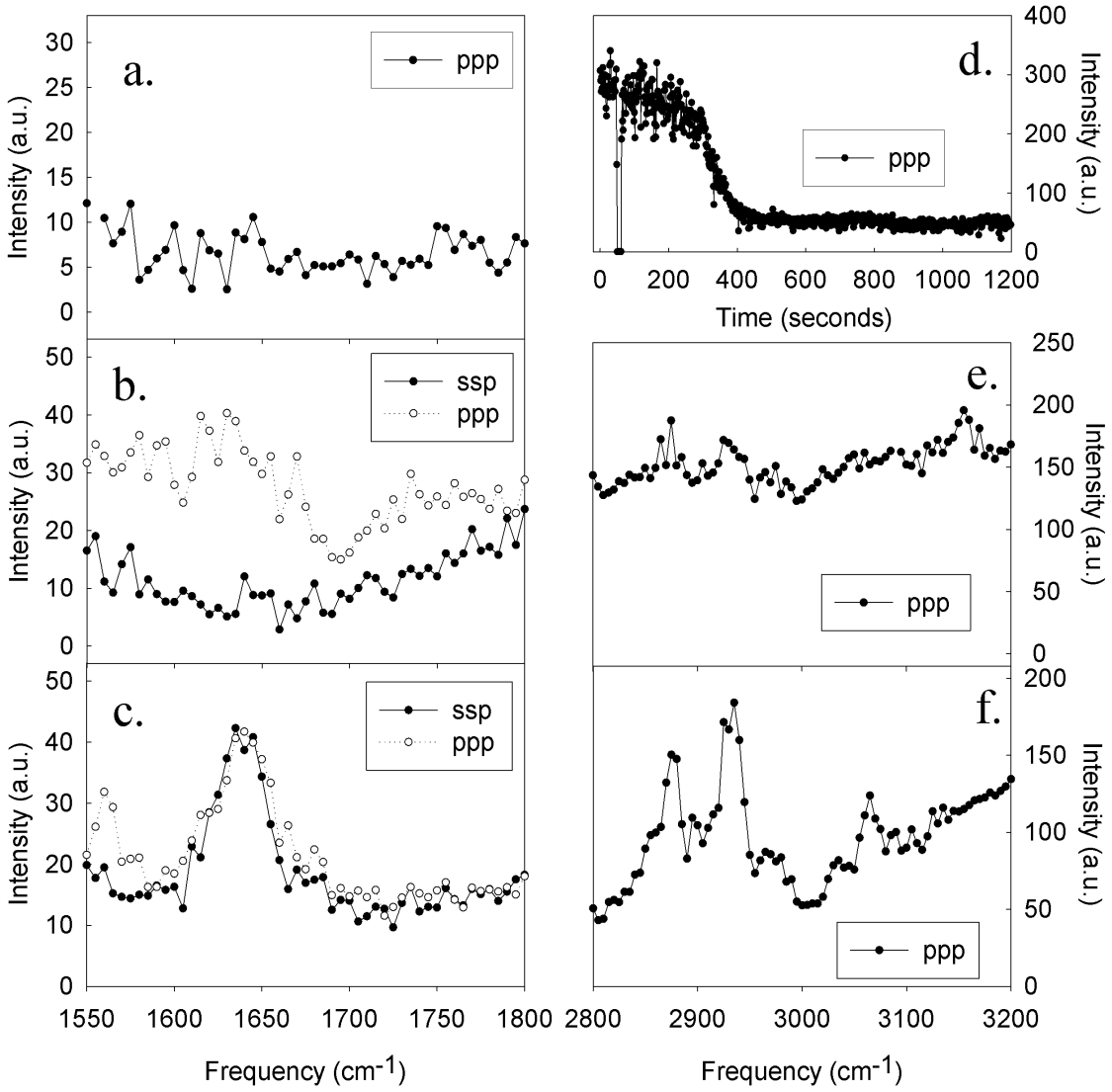


Figure 2-8: Polarized SFG spectra collected in the Amide I region from magainin 2 on the PS surface. Shown for concentrations of a) 200 nM, b) 400 nM, and c) 800 nM. d) Time-dependent ppp SFG signal at 3070 cm^{-1} collected from the deuterated PS/water interfaces; peptide was added at t=100 s. The laser was blocked at t=70 s to demonstrate that signal was present.

Polarized SFG spectra collected in the C-H/O-H stretching frequency region from magainin 2 at concentrations of e) 400 nM, and f) 800 nM.

As peptide concentration was increased to 800 nM, these C-H stretching signals became stronger. Even at much higher concentrations, these CH signals alone have previously been interpreted as evidence of either increased peptide number density or increased side chain ordering towards the surface. In a previous study, this was cited as an indication that the peptide was lying down.[37] However, the SFG instrument in our lab is capable of studying a broader spectral range, and thus taking advantage of new information in the Amide I spectral region.

Polarized Amide I measurements reveal a different behavior than that seen from sidechain signals: unlike the intensity ratios at lower concentrations, ppp and ssp signal intensities at the 800 nM concentration are very similar, and can be fitted by a single peak centered at ~1645 cm^{-1} (with stronger signals than were observed at lower concentrations). Due to limitations in spectral resolution, secondary structure could not be definitively assigned, but the peak center is consistent with the reported frequencies for coiled-coil structures of tightly associated helices.[38, 39]

The change in intensity ratios as compared to lower concentrations can be interpreted as evidence of a change in peptide orientation or structure at the interface– showing that the interfacial behavior of magainin 2 molecules changes as a function of peptide concentration. Furthermore, the fitted Amide I ppp/ssp intensity ratio at 800 nM (corrected for variations in Fresnel coefficient) was approximately 0.94, which is far below the range of possible ratios that would correspond to any single delta or Gaussian distribution of orientations (Figure 2-2c). It has previously been shown that such ratios can however be produced if multiple orientations are adopted.[2] The results from my molecular dynamics simulations suggest that single isolated peptides would express a

strong and singular tendency to lie down. Combined with the concentration dependent SFG results, I therefore believe that the experimentally observed behavior is due to peptide-peptide interactions.

2.4.3 Dimer Simulation

When the peptide concentration increases, peptides may interact with each other in addition to interacting with the surface. These dimers or oligomers may mediate the interfacial peptide orientation, and thus single-peptide simulations may not be adequately represent the real experimentally observed system.

Predicting the most stable structure of multi-peptide aggregates in the presence of a surface is a non-trivial task, but as a first approximation, I employed the known structure of an existing covalently bound magainin 2 dimer (PDB ID: 1dum).[40] This structure was modified to remove the covalent bond, and the resulting noncovalently interacting dimer was placed in contact with the surface. Based on my previous simulations and the observation that hydrophobic interactions were responsible for peptide adsorption, the dimer was oriented so as to maximize these hydrophobic interactions with the surface, and I monitored the evolution of the structure over the course of 7 ns of simulation. It was found that this dimer remained intact during this time, but flattened out (Figure 2-9).

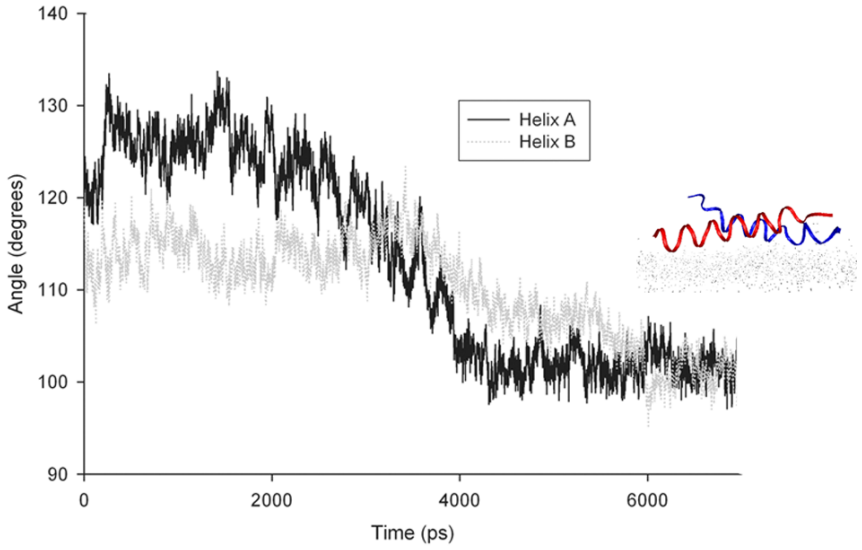


Figure 2-9: Plots of the orientation angle relative to the surface normal for each helix in the dimer. A sample view of the magainin 2 dimer is shown after 7 ns of simulation time.

The final orientations of the individual peptides in this system were observed to be approximately 100^{o} in each case. This slight tilt allowed peptides to interact with each other as well as with the surface, and the dimer remains intact during the entire course of the simulation. However, this simple dimer is in itself not consistent with the SFG Amide I intensity ratios observed. Thus, it is likely that the actual peptide oligomer at the surface adopts a more complex structure, or one that cannot be captured on short timescales from known initial conditions. It is interesting to note that the concentration at which oligomers are observed to form in these experiments is well below the minimum inhibitory concentration (MIC) at which magainin 2 is effective at breaking down lipid bilayers, and is in line with the concentration in which peptides are observed to insert into the bilayer[41, 42]; it is believed that peptide oligomers play a role in these phenomena.[40, 42] Given my experimental results, it is likely that at least some of the peptides present at the interface are present in the form of an oligomer. In creating a molecular-level picture of

peptides at interfaces, the role of peptide self-association must therefore be studied in the future.

2.5 Conclusions

Few experimental techniques are able to provide direct information on molecular orientation or proteins and peptides at solid/liquid interfaces, due to limitations in signal sensitivity, surface specificity, or *in situ* capability. In this chapter, a combination of MD and SFG was used to demonstrate that burial of hydrophobic surface residues strongly drives the facially amphiphilic AMP magainin 2 to lie down parallel to the surface at a hydrophobic/hydrophilic interface, and that orientation may be modulated somewhat by the formation of dimers or larger oligomers.

When these results are compared to concentration-dependent SFG spectra, results at low peptide concentrations can be correlated to MD simulation results on a single helical peptide at the interface. When the peptide concentration increases, a single uniform orientation of α-helices is insufficient to explain the observed SFG spectral intensity ratios. Thus, inter-peptide interactions must be significant, and preliminary exploratory simulations are presented that demonstrate that dimers can remain associated at the interface. In the past, simulation work by a wide variety of groups has focused on single-peptide simulations. I believe that greater attention to inter-peptide interactions will reveal a new layer of complexity that is necessary to describe how even relatively simple peptides interact with the polymer surface. Insights gained from studies of simple model systems may then be applicable to more complex biological systems in the future.

2.6 Acknowledgements

I am grateful to Jeremiah Nummela and Khoi Nguyen for numerous helpful discussions, as well as Dr. Jeff Wereszczynski and Dr. Shuji Ye for technical assistance. Computing resources were provided by the Center for Advanced Computing at the University of Michigan.

2.7 References

(1) Wang, J.; Lee, S.-H.; Chen, Z. *J. Phys. Chem. B* **2008**, *112*, 2281-2290.
(2) Chen, X.; Wang, J.; Boughton, A., P.; Kristalyn, C., B.; Chen, Z. *J. Am. Chem. Soc.* **2007**, *129*, 1420-1427.
(3) Nguyen, K., Tan; Le Clair, S. V.; Ye, S.; Chen, Z. *J. Phys. Chem B* **2009**, *113*, 12169–12180.
(4) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187-199.
(5) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1988**, *61*, 2635-2638.
(6) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195-1198.
(7) Grossfield, A. "An implementation of WHAM: the Weighted Histogram Analysis Method". http://membrane.urmc.rochester.edu/Software/WHAM/WHAM.html (accessed **April 14, 2008**)
(8) Kumar, S.; Bouzida, D.; Swendsen, R., H.; Kollman, P., A.; Rosenberg, J., M. *J. Comp. Chem.* **1992**, *13*, 1011-1021.
(9) Roux, B. *Comp. Phys. Commun.* **1995**, *91*, 275-282.
(10) Boczko, E., M.; Brooks, C., L. III. *J. Phys. Chem.* **1993**, *97*, 4509-4513.
(11) Wang, J.; Even, M., A.; Chen, X.; Schmaier, A., H.; Waite, J., Herbert; Chen, Z. *J. Am. Chem. Soc.* **2003**, *125*, 9914-9915.
(12) Ayyagari, C.; Bedrov, D.; Smith, G., D. *Macromolecules* **2000**, *33*, 6194-6199.
(13) Khare, R.; Paulaitis, M. E.; Lustig, S. R. *Macromolecules* **1993**, *26*, 7203-7209.
(14) Clancy, T., C.; Jang, J., Hwan; Dhinojwala, A.; Mattice, W., L. *J. Phys. Chem. B* **2001**, *105*, 11493-11497.
(15) Raffaini, G.; Ganazzoli, F. *Phys Chem Chem Phys* **2006**, *8*, 2765-2772.
(16) Ganazzoli, F.; Raffaini, G. *Phys Chem Chem Phys* **2005**, *7*, 3651-3663.
(17) York, R., L.; Browne, W., K.; Geissler, P., L.; Somorjai, G., A. *Israel J. Chem.* **2007**, *47*, 51-58.
(18) Sun, Y.; Welsh, W., J.; Latour, R., A. *Langmuir* **2005**, *21*, 5616-5626.
(19) Yang, C., S. -C; Wilson, P., T.; Richter, L., J. *Macromolecules* **2004**, *37*, 7742-7746.
(20) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926-935.
(21) Belch, A., C.; Berkowitz, M. *Chem. Phys. Lett.* **1985**, *113*, 278-282.
(22) Wang, L.; Hermans, J. *Molec. Sim.* **1996**, *17*, 67-74.
(23) Edholm, O.; Jähnig, F. *Biophysical Chemistry* **1988**, *30*, 279-292.
(24) Gruber, C. C.; Pleiss, J. *J. Comput. Chem.* **2011**, *32*, 600-606.

(25) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187-217.
(26) MacKerell, A., D. Jr; Bashford, D.; Bellott, M.; Dunbrack, R., L. Jr; Evanseck, J., D.; Field, M., J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F., T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D., T.; Prodhom, B.; Reiher, W., E. III; Roux, B.; Schlenkrich, M.; Smith, J., C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586-3616.
(27) Kuczera, K.; Jas, G. S.; Elber, R. *J. Phys. Chem. A* **2009**, *113*, 7461-7473.
(28) Lee, J.; Im, W. *Chem. Phys. Lett.* **2007**, *441*, 132-135.
(29) Lee, J.; Im, W. *Phys. Rev. Lett.* **2008**, *100*, 018103.
(30) Vallone, B.; Miele, A., E.; Vecchini, P.; Chiancone, E.; Brunori, M. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6103-6107.
(31) Lee, B.; Richards, F., M. *J. Mol. Biol.* **1971**, *55*, 379-400.
(32) Lum, K.; Chandler, D.; Weeks, J., D. *J. Phys. Chem. B.* **1999**, *103*, 4570-4577.
(33) Wei, Y.; Latour, R. A. *Langmuir* **2009**, *25*, 5637-5646.
(34) Lee, C.-Y.; McCammon, J., Andrew; Rossky, P., J. *J. Chem. Phys.* **1984**, *80*, 4448-4455.
(35) Stone-Masui, J.; Watillon, A. *J. Colloid Interface Sci.* **1975**, *52*, 479-503.
(36) Wang, J.; Chen, X.; Clarke, M., L.; Chen, Z. *J. Phys. Chem. B* **2006**, *110*, 5017-5024.
(37) Phillips, D., C.; York, R., L.; Mermut, O.; McCrea, K., R.; Ward, R., S.; Somorjai, G., A. *J. Phys. Chem. C* **2007**, *111*, 255-261.
(38) Reisdorf, W. C. J.; Krimm, S. *Biochemistry* **1996**, *35*, 1383-1386.
(39) Tamm, L., K.; Tatulian, S., A. *Quart. Rev. Biophys.* **1997**, *30*, 365-429.
(40) Wakamatsu, K.; Takeda, A.; Tachi, T.; Matsuzaki, K. *Biopolymers* **2002**, *64*, 314-327.
(41) Nguyen, K. T.; Le Clair, S. V.; Ye, S.; Chen, Z. *J. Phys. Chem. B* **2009**, *113*, 12358-12363.
(42) Mukai, Y.; Matsushita, Y.; Niidome, T.; Hatekeyama, T.; Aoyagi, H. *J. Pept. Sci.* **2002**, *8*, 570-577.

Chapter 3: $G\beta_1\gamma_2$ and GRK2 Binding and Orientation Change Observed on a Model Membrane In Situ

3.1 Introduction

It has been roughly a decade since SFG was first applied to the study of proteins, and less than that since the first application to the more informative Amide I range. As a result, previous quantitative studies of protein orientation- including those reported in the previous chapter- have largely focused on simple peptides, such as single α-helices[1-9] or β-sheets.[10-12] Many important membrane-active molecules (such as antimicrobial peptides) fit that description, but these studies only scratch the surface of what SFG is capable of.

It has been estimated that half of all known proteins are either membrane-bound, or interact with the membrane at some stage.[13] These proteins carry out a huge number of biological functions: they mediate the transport of ions[14, 15] or small molecules[16], or serve to recognize, respond to, and regulate signaling in cells. Yet despite a wealth of information obtained from mutation experiments, NMR, and X-ray crystallography experiments, the most detailed structural studies require removal of the protein from its native environment. As a result, although it is possible to describe the three-dimensional conformation of the protein in detail, the most commonly used structural techniques are generally unable to directly reveal which part of that structure is in contact with the bilayer. SFG provides a means of performing *in situ* measurements, but limitations in available measurements and data analysis methodologies have hampered previous efforts.

This study is focused on a multi-subunit protein complex responsible for regulating cell signaling. G protein coupled receptors (GPCRs) are integral membrane proteins that are involved in a wide variety of biological processes in eukaryotic cells.[17] In response to extracellular cues such as hormones, odorants, and light, they activate heterotrimeric GTP-binding proteins found on the inner surface of the cell membrane, which in turn directly interact with membrane-associated effectors such as adenylyl cyclase or cGMP phosphodiesterase, which control the level of second messengers.[18]

The native lipid environment plays many roles in organizing the subunits and facilitating catalysis.[19, 20] Understanding how the protein subunits are oriented at and with the plasma membrane will enhance our understanding of how the membrane facilitates higher affinity interactions between signaling molecules, how these molecules are optimally aligned through these interactions for catalysis, and how higher order signaling scaffolds are assembled at the membrane. The interactions of heterotrimeric G proteins, in particular their βγ subunits (Gβγ), with the lipid bilayer facilitate GPCR-catalyzed GTP exchange on the Gα subunit.[21, 22] Gβγ subunits also bind the enzyme GRK2, which is recruited to the membrane by G proteins to phosphorylate and help inactivate GPCRs (Figure 3-1). Because of its apparent ability to simultaneously interact with heterotrimeric G proteins $G\alpha_q$ and Gβγ, as well as GPCRs and the membrane, GRK2 may be involved in the assembly and organization of signaling complexes at GPCRs.[23] Although the crystal structure of GRK2 in complex with both $G\alpha_q$ and Gβγ is known[24], its orientation while engaged at the cell membrane is difficult to infer from the crystal structure alone, because many of the known membrane binding determinants of the complex are disordered or involved in crystal contacts. *In situ* measurements are needed.

In addition to its biological significance, the GRK2-Gβγ complex is an attractive candidate for SFG study due to its well-conserved structure and well-defined interactions with the membrane. These factors support the key assumption that the complex will adopt a narrow and singular orientation distribution at the lipid/buffer interface.

In this chapter, I report efforts to characterize the formation and orientation of the larger GRK2-Gβγ complex in a more biologically relevant lipid bilayer. In particular, I sought to determine (1) whether the orientation of the GRK2-Gβγ complex at the model cell membrane is consistent with predictions made based on an existing crystal structure, and (2) whether or not Gβγ reorients in order to accommodate binding of GRK2. In the process of studying this problem, I have also developed new approaches and analysis tools that greatly enhance the size and complexity of the proteins that can be studied with SFG.

3.2 Materials and Methods

3.2.1 Protein Preparation

Samples of bovine GRK2 and geranylgeranylated Gβγ were expressed in baculovirus-infected cell culture, purified as described elsewhere, and frozen in liquid nitrogen.[25] These samples were obtained separately, and also mixed to create the pre-formed complex. In order to determine whether or not the spectra at the interface originated from the bound GRK2-Gβγ complex, a Gβγ-binding deficient mutant GRK2-R587Q was similarly prepared from virus obtained from J. Benovic (Thomas Jefferson University).[26] Using a flow cytometry protein interaction assay[27], it was confirmed that GRK2-R587Q does not bind Gβγ under conditions that wild-type GRK2 binds with a K_D of 50 nM (data not shown). A complex containing both GRK2 and Gβγ was isolated

using S200 size exclusion columns from protein isolated from cells that were co-infected with baculoviruses encoding the $G\beta_1$, $G\gamma_2$ and GRK2 proteins. On the day of experiments, protein aliquots were thawed and centrifuged through Nanosep MF 0.2 μm filters (Pall Corporation) to remove aggregated protein, and the concentration was calculated by A_{280}. Samples were diluted directly into a solution containing 20 mM HEPES (pH 8.0), 50 mM NaCl, and 5 mM DTT. This buffer mixture was also used as the liquid subphase for the lipid bilayer.

3.2.2 Bilayer Preparation

The membrane environment was modeled using Planar supported lipid bilayers (PSLBs)[28], deposited using the Langmuir-Blodgett / Langmuir-Schaefer deposition method as described elsewhere.[29] To mimic the composition of mammalian cell membranes where this protein-protein complex is found, a 9:1 mixture of POPC/POPG lipids was used. Stock solutions of lipids were purchased from Avanti Polar Lipids Inc., dissolved in chloroform, and mixed as needed to produce the desired lipid mixture composition. Following bilayer creation, the bilayer subphase was flushed three times with fresh buffer to remove excess lipids prior to addition of the protein. All samples were prepared on clean right-angle CaF_2 prism substrates (Altos Photonics, Bozeman MT) and SFG spectra were collected from the proteins associated with the lipid bilayer in a "near" total internal reflection geometry.[30]

3.2.3 SFG Experiments

Membrane-bound proteins were studied by injecting a protein stock solution into the bilayer subphase and monitoring time-dependent SFG spectra until the protein signal intensity at 1652 cm^{-1} was stable (~1 hr). Spectral intensities were the same under

continuous collection as when the input laser beams were blocked for 30 min between scans, and no change was observed over the timescale of these experiments. This indicates that the sample did not sustain damage due to laser irradiation. In all cases, Gβγ, GRK2, or the GRK2-Gβγ complex were added to the aqueous buffer for a final concentration of 336 nM. The following samples were examined: 1) GRK2 or 2) Gβγ alone, 3) a preformed GRK2-Gβγ complex, 4) a GRK2-Gβγ complex formed by sequential addition of equimolar Gβγ and GRK2 and 5) Gβγ after the addition of equimolar GRK2-R587Q, a GRK2 mutant deficient in binding Gβγ.[26] It has previously been shown that GRK2 binding to the membrane is greatly enhanced by the presence of Gβγ.[22] Therefore, sequential addition experiments in which GRK2 was added first were not performed.

3.3 SFG Data Analysis

3.3.1 Nonlinear Optical Response for Large Proteins

As discussed in section 2.2.1, SFG spectral intensities depend on the orientation and number of molecules at the interface. This sensitivity to molecular orientation originates from the fact that SFG uses polarized sum, visible, and infrared beams. Different polarizations measure different components of the surface susceptibility tensor $\chi_{eff}^{(2)}$, which can be related to different components of the molecular hyperpolarizability tensor, β. In this work, I demonstrate a method for calculating the required values of β for an entire protein, based on a known crystal structure.

The orientation of the molecule may be characterized by relating the response of a single molecule to the response measured in the lab coordinate system. The orientation of a surface functional group or molecule can be characterized by either one tilt angle θ (e.g.

in case of a single α-helix) or tilt and twist angles θ and ψ, respectively, (e.g. in the case of a β-sheet or a complicated protein). The rotations corresponding to these angles are shown in Figure 3-1. In this study, data analysis is focused on the α-helical Amide I peak observed at ~1652 cm^{-1}, which dominates the spectra for Gβγ alone[31] as well as for the GRK2-Gβγ complex. The signals from β-sheet regions were too weak to be analyzed.

The GRK2-Gβγ complex contains 31 helical segments[32], and consequently, analyzing the molecular orientation relies on determining the combined response for all α–helices in the protein.[1, 31, 33-35] The net α–helical hyperpolarizability, $\beta_{protein}$, can be determined because the length and relative orientations of all helical segments are known from the known crystal structure (PDB entry 1omw). The α-helical segments were assigned using the DSSP algorithm as implemented in UCSF Chimera.[36] For the purposes of this analysis, two angles were used to describe the orientation of the helical subunits: the tilt angle (relative to the +z axis) and the azimuthal angle (derived from the rotation matrix for a vector along the helix backbone, $vec = [x \quad y \quad z]^T$ and calculated according to the formula $\phi = \sin^{-1}\left(\frac{y}{|vec| * \sin\theta}\right)$). The third Euler angle, ψ (representing twist) was ignored due to the cylindrical symmetry of each single α-helix.

The hyperpolarizability tensor elements for each individual α-helical segment of known residue length were calculated according to the bond additivity model described previously.[3, 37] Each of the 27 elements of the combined hyperpolarizability tensor for the entire protein was then calculated as the sum of the response for each individual helix, using a rotation of the axis system to place the z-aligned helical segment into the protein coordinate frame according to the angles calculated above.

$$\beta_{ijk,helix} = \left(\frac{d\alpha^*_{helix}}{dQ}\right)_{i,j} * \left(\frac{d\mu_{helix}}{dQ}\right)_k \qquad \textit{Equation 3-1}$$

$$\beta_{ijk,protein} = \sum_{n=1}^{\#helices} \frac{1}{2\pi} \int_0^{2\pi} \left(R * \alpha_{helix,n} * R^T\right)_{i,j} \left(R * \mu_{helix,n}\right)_k d\psi \qquad \textit{Equation 3-2}$$

In the equation above, R represents the Euler rotation matrix in the zyz convention:

$$R = R_z(\phi)R_y(\theta)R_z(\psi)$$

$$= \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} * \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} cos\,(\phi)cos\,(\theta)cos\,(\psi) - \sin(\phi)\sin(\psi) & cos(\phi)\,cos(\theta)\,sin(\psi) + \sin(\phi)\cos(\psi) & -\cos(\phi)\sin(\theta) \\ -\sin(\phi)\,cos(\theta)\,cos(\psi) - cos(\phi)\,sin(\psi) & -\sin(\phi)\cos(\theta)\sin(\psi) + \cos(\phi)\cos(\psi) & \sin(\phi)\sin(\theta) \\ \sin(\theta)\cos(\psi) & \sin(\theta)\sin(\psi) & \cos(\theta) \end{bmatrix}$$

Equation 3-3

The tilt angle θ and the twist angle ψ for a protein are defined according to Equation 3-3. This method was also used to calculate $\beta_{protein}$ for Gβγ, which contains four helical segments.

As discussed above, measured SFG observable (e.g., $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio) is a function of the protein orientation angles and the molecular hyperpolarizability. Quantitatively, the SFG surface susceptibility tensor can be calculated by rotating a single molecule into the lab coordinate frame and assuming that all molecules were randomly distributed in the plane of the surface (azimuthal angle ϕ). Expressed in a single compacted equation suitable for algorithmic solution:

$$\chi^{(2)}_{ijk} = \frac{1}{2\pi} \int_0^{2\pi} (R * R^T)_{i,j} * R_k \beta_{ijk,protein} d\phi \qquad \textit{Equation 3-4}$$

As discussed in Chapter 2, there are two SFG-active Amide I vibrational modes to be considered. The peak centers for these two modes are very close and cannot be

separated within the resolution of our SFG spectrometer. The total molecular hyperpolarizability was calculated as the sum of the A and E1 modes.

3.3.2 Software for Data Analysis.

The required calculations become prohibitively complicated as the size of the protein to be studied increases: a previous study of the four-helix protein Gβγ required manual setup of more than 216 integral equations, and the protein structure was analyzed by hand to determine the relative positions of all helical segments. For the study of large protein complexes to become routine, automated methods are necessary. In 2007, the Simpson group released the NLOPredict software package[35] to address this need, using parameters derived from quantum mechanical calculations. The Chen group has published an alternate method and series of parameters using a bond additivity model based on parameters from tetragonal crystals of aspartame.[38, 39] These results agree with the dipole moment and Amide I polarizability tensor components derived from experimental IR and Raman studies on highly ordered protein samples.[40-42]

Building on existing methodology for single helices[3, 11, 43], I have developed and implemented the methods of Section 3.2.1 to study the net response from proteins with many α-helices. My program allows for rapid analysis of any arbitrary protein structure from the Protein Databank (PDB), and was designed to easily export data and combine results from multiple measurements. This is essential for the study of anything larger than the simplest proteins. As an additional improvement over previous calculations[31], variations in helix length are explicitly treated in this calculation for greater accuracy. This program outputs SFG signal strength ratios ($\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$) for orientation analysis, and also makes it possible to compare observed signal intensities across different protein

samples or subunits. As discussed Section 3.4.3, this provides the potential for additional constraints that can be used to more uniquely assign molecular orientation.

The software performs four tasks. Each script is modular, so that these tasks can be performed at any time after the previous step. Repeating the analysis does not require repeating each step.

1) The protein is oriented into an initial reference position (Section 3.3.4). Although this step is not required, it can be helpful for interpreting results.
2) Relative to the reference orientation, the length and relative orientation of each helical segment is determined. The resulting information is saved to a csv file that describes the protein structure.
3) SFG experimental observables are calculated, and the results are saved to a text-based file format for future analysis. My program employs the Matplotlib library to produce high quality plots, and the panels of Figure 3-4 were prepared directly from images generated in this step. Good computational performance in this step was ensured by pre-calculating the symbolic algebraic forms of Equations 3-2 and 3-4. By removing the slow numerical integration step, significant performance gains were seen versus a reference implementation in a commercial mathematical software package.
4) Calculated SFG observables were compared to experimental measurements, using a newly designed set of plots aimed at summarizing results from multiple measurements.

3.3.3 Graphical Display of Results

Given the complexity of the system studied, and the sensitivity of the results to small experimental errors, multiple solutions must be considered. A set of heat map style plots was developed to display all orientations for which calculated values fall within ±10% of the measured Gβγ $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio (equivalent to about ±20% of the measured intensity ratio), and ±25% of the observed signal strength $\chi^{(2)}_{xxz}$ change between the measurements that compare Gβγ and GRK2-Gβγ. These plots display a score assigned based on how close the calculated value at the n[th] datapoint was to each experimental (target) value:

$$Score_n = 1 - \left|\frac{ratio_n - ratio_{exp}}{tol * ratio_{exp}}\right| \qquad \textit{Equation 3-5}$$

All orientations for which *any* criterion was not met were assigned a score of 0 by default. Thus this scoring function simply rates how close a given calculated parameter is to the experimental measurement, within the tolerance ("tol") given above. The final quality of the match at that point was determined as the product of i criteria, $Score_{tot} = \prod_i Score_i$. A score of 100% indicates that the orientation in question yields an exact match for all experimental measurements. Colors are then assigned based on score, so that related results can be put into context within experimental error estimates.

3.3.4 Defining the Reference Orientation

The orientation of the protein is described in terms of two angles (tilt and twist). To make these angles easier to interpret- and to put the results in a physically relevant context- all calculations herein are described in terms of a reference orientation. For this position, I chose the expected position of the GRK2-Gβγ complex when interacting with

the membrane. In the absence of any previous direct measurements, this was defined from a set of residues that are believed to be in close proximity to the bilayer. These include the β1-β2 loop of the PH domain and the C-terminal region of Gγ, which is geranylgeranylated.[26, 32, 44, 45]

A linear least squares fitting procedure was applied to define a plane through the backbone Cα atoms of the following residues: 30-31, 209, 528-547, and 569-576 of GRK2; residues 46 and 48 of Gβ; residues 62 and 66-68 of Gγ. Then, PDB entry 1omw was rotated so that a plane through these residues would coincide with the membrane plane. The chosen residues, and the final resulting orientation, are shown in Figure 3-1.

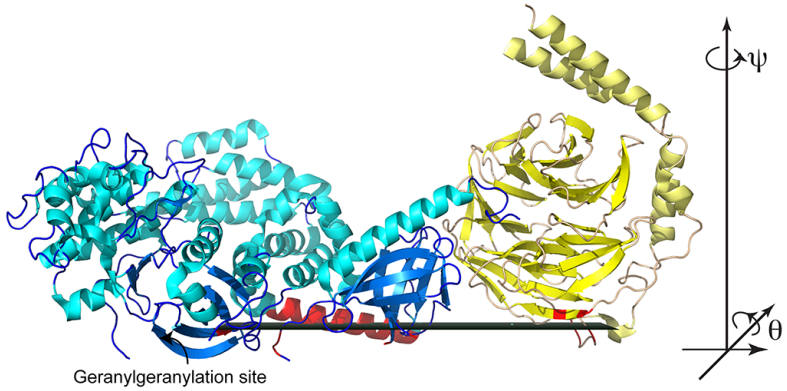


Figure 3-1:The GRK2-Gβγ complex in the reference orientation, based on a set of potential membrane-interacting residues (shown in red). The fitted plane, which is expected to be parallel to the plane of the lipid bilayer, is shown in green. GRK2 is colored blue with cyan helices. Gβγ is colored yellow, with helices shown in pale yellow. The C-terminal residue of the G protein γ subunit is geranylgeranylated, and thus must be in close proximity to the bilayer.

3.3.5 Graphical Depiction of Physically Allowed Orientations

Although Euler angles provide an efficient way to perform rotations, the resulting combinations of tilt and twist angles specified are not always straightforward to visualize. Furthermore, not all mathematically allowed solutions are in fact physically reasonable

given restraints such as the position of the hydrophobic geranylgeranyl group that anchors the protein to the bilayer. For the case of GRK2-Gβγ, unfavorable orientations typically featured a large number of atoms colliding with the membrane. Hence, each possible orientation of the protein was scored on a binary "allowed/disallowed" scale: physically allowed positions were defined as those where less than 10% of the backbone α carbons collided with the membrane (Figure 3-2). The resulting plot can be added to the matches for experimental measurements (Section 3.3.3) to better indicate which combinations of tilt and twist angle are physically meaningful. Because it is known that the geranylgeranyl group of Gβγ plays a key role in anchoring the protein to the membrane, this criterion was implemented by accepting all positions where less than 10% of backbone atoms were below the C-terminal residue of the Gγ chain. (the geranylgeranyl attachment point)

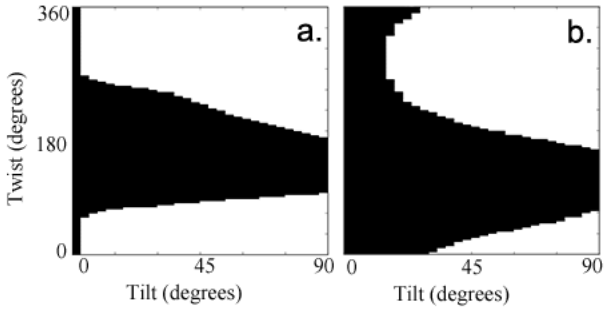


Figure 3-2: Orientation angles corresponding to physically reasonable orientations of (a) the GRK2-Gβγ complex, or (b) Gβγ alone. Allowed positions are shown in black, and defined as those where less than 10% of protein atoms collide with the plane of the bilayer.

3.4 Results and Discussion

3.4.1 SFG results: Formation of the Complex

As seen in Figure 3-3a, only weak signals were observed in SFG Amide I spectra collected from GRK2 alone at the lipid bilayer, presumably due to the relatively weak interactions between the lipid bilayer and GRK2. In contrast, much stronger spectra were observed for the Gβγ subunit alone in the lipid bilayer (Figure 3-3b), which is consistent with the fact that Gβγ is geranylgeranylated and thus has a much higher lipid affinity than GRK2.[22] The Gβγ spectra are dominated by a peak at ~1652 cm^{-1}, contributed by the α-helical regions of the protein.

To demonstrate the formation of the complex *in situ*, SFG spectra were collected from the pre-formed GRK2-Gβγ complex at the membrane interface (Figure 3-3c), and compared these to the results from serial addition of GRK2 to Gβγ (Figure 3-3d). The spectral intensities (Figure 3-3b) and fitted $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratios (as calculated in the "Methods" section) were similar, but markedly different from those observed for Gβγ alone. As expected, serial addition of the GRK2-R587Q mutant to Gβγ did not alter the Gβγ spectra (Figure 3-3e). These results clearly demonstrate that SFG can be used to study the formation of a specific complex *in situ*.

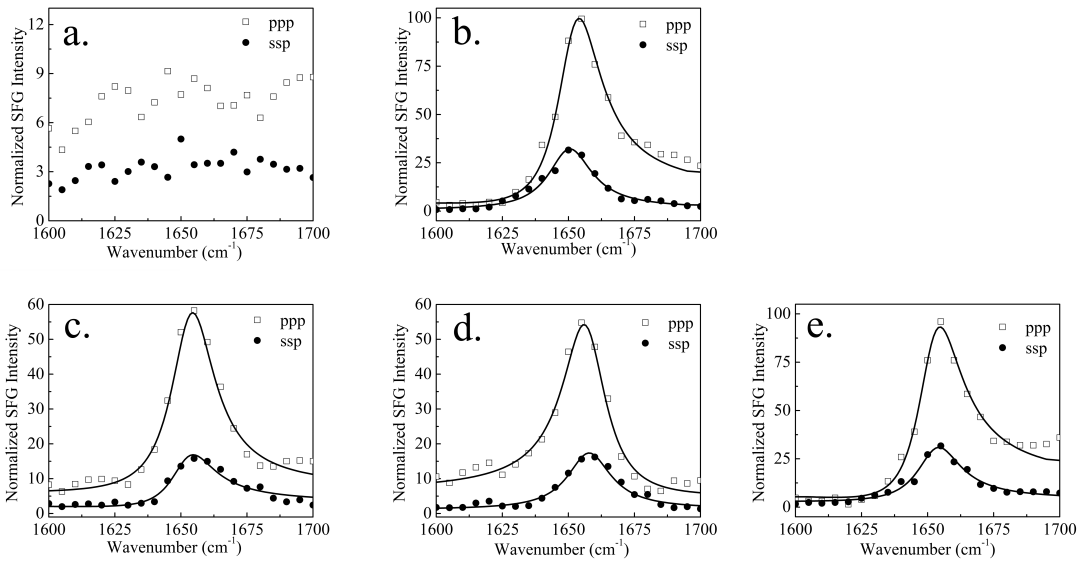


Figure 3-3: SFG measurements of GRK2, Gβγ, and their complexes on a POPC:POPG (9:1 weight ratio) lipid bilayer. a) Only weak Amide I signals were observed from GRK2 injected alone, consistent with the fact that GRK2 itself does not bind to lipids with high affinity. b) Stronger signals were observed from Gβγ, consistent with the fact that the protein is lipid modified. c) Signals and ppp/ssp ratios from the preformed complex and d) from two subunits added sequentially are similar, indicating that the complex can be formed in situ. In both cases, the fitted zzz/xxz ratio is ~2.20. e) Signals from Gβγ alone were unchanged upon binding of nonbinding mutant GRK2-R587Q, supporting the claim that signals in panels c and d originate from complex formation at the interface.

Fitting of the experimental spectral data shown in Figure 3-3 yields the following results: the Gβγ $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio is 2.01, the fitted signal strength in the ssp polarization drops by roughly a factor of 1.39 when GRK2 is added to form the complex, and upon formation of GRK2-Gβγ, the $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio is 2.20.

3.4.2 Orientation of the GRK2-Gβγ Complex

A marked drop in signal intensity was observed in the transition from Gβγ alone to the complex with GRK2 (Figure 3-3b, d). As discussed in section 3.3.1, SFG signal intensities are a function of molecular orientation, the net molecular hyperpolarizability

$\beta_{protein}$, *and* the number of molecules on the surface.[46] Given the expectation that the molar concentration of proteins at the membrane will be driven primarily by the presence of Gβγ (i.e. geranylgeranylated Gβγ will not dissociate from the lipid bilayer upon addition of GRK2), one can assume that the drop in intensity is due to a change in the net molecular hyperpolarizability upon addition of GRK2, reorientation of the protein segments, or both. Because the observed SFG spectra are dominated almost entirely by single peak centers corresponding to α-helices, it is reasonable for orientation analysis to focus on the nonlinear optical response ($\beta_{protein}$) for the α-helical regions of the protein. (Figure 3-1) No signals were observed from the β-propeller due to the symmetry selection rule of SFG.[31]

The crystal structure of Gβγ in complex with GRK2 (PDB entry 1omw) was used as the starting point for all data analysis.[32] Based on the fact that the overall conformation of the protein was preserved in two unique crystal structures[24] and due to the fact that GRK2 is stabilized by Gβγ,[45] once can assume that a similar overall conformation of this complex would be preserved in a lipid environment. The reference orientation depicted in Figure 3-1, which is denoted as the tilt (θ)=0˚, twist (ψ)=0˚ position of the complex. For consistency, this orientation was also used as the reference position for the Gβγ subunit coordinates alone (as extracted from the GRK2-Gβγ complex).

For the purposes of orientation analysis, the SFG macroscopic observables $\chi^{(2)}_{zzz}$, $\chi^{(2)}_{xxz}$, and $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ were calculated for all unique combinations of the tilt and twist angles of Gβγ and the GRK2-Gβγ complex (Figure 3-4).

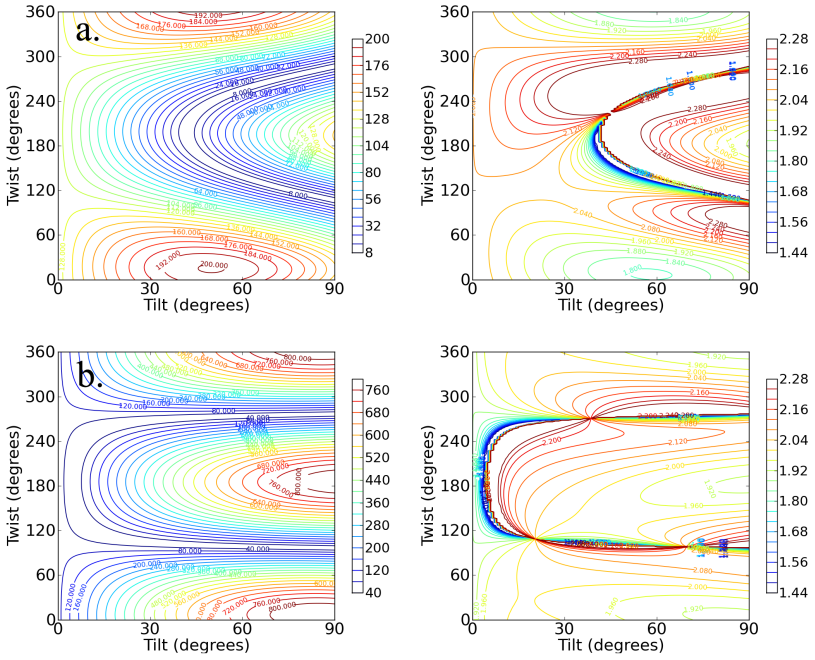


Figure 3-4: Contour plots showing the calculated molecular response in the ssp polarization (left), and the predicted ratio of fitted signal strengths $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ (right) for a) Gβγ alone, and b) the GRK2-Gβγ complex.

Because it is unreasonable to assume that the protein would be upside-down with the geranylgeranyl group far from the bilayer, I have limited the plots to focus on the range of interest (tilt angles 0-90°). It is apparent from the resulting contour plots that many possible combinations of tilt and twist angle can yield computed ratios that are close to the experimentally measured $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio. However, only a subset of these positions are physically reasonable. Many would pull the geranylgeranyl group of Gβγ unreasonably far from the membrane, or force the protein into an orientation that would require a collision with the membrane in order for the geranylgeranyl group at the C-terminus of Gγ to insert into the bilayer. Interestingly, the results in Figure 3-6b show that the experimentally measured $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio for the GRK2-Gβγ complex does not

match the calculated value at the reference position (θ =0˚, ψ=0˚). The nearest high-scoring match is at the position θ =10˚, ψ=180º. There is, however, a swath of closely related high scoring orientations with small tilt angles (10-15˚) and twists ranging from 120-240˚. There is also a narrower stretch of matches in the range from tilt 30-60˚ and twist of 90˚, but in these orientations the PH domain is not positioned such that it can make its expected contacts with the membrane. My results thus suggest that the most likely orientations of the GRK2-Gβγ complex would tilt the bulk of the kinase domain of GRK2 away from the membrane surface by ~10 to 15 degrees. A recent crystal structure of activated GRK6 (a related enzyme) revealed the structure and location of the GPCR receptor docking site of the kinase[47], which should be conserved in GRK2. The residues that constitute this site were disordered in the GRK2-Gβγ structure, and were not considered when defining the reference position. SFG results are consistent with a tilt that positions the receptor docking site of GRK2 to interact more efficiently with the cytosolic surface of the GPCR, in a similar orientation to that proposed for activated GRK6.

Due to the complexity of the molecules studied and the limited number of measurements available, it is not possible to narrow the range of orientations of GRK2-Gβγ to a single unambiguous position. However, the above results do demonstrate that *in situ* measurements can provide more information about protein orientation at a phospholipid bilayer than can be inferred from the crystal structure alone.

3.4.3 Orientation of Gβγ Alone and in the Complex

Next, I consider whether or not Gβγ must reorient in order to bind GRK2. The $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio is a single measurement, and more information is needed to narrow the range of possible orientations. Experimentally, signal intensity drops markedly upon

formation of the complex (Figure 3-3). This observation is most likely if both Gβγ and the GRK2-Gβγ complex adopt relatively small tilt angles (Figure 3-4). This may seem surprising since the GRK2-Gβγ complex is larger, but the net SFG response can be thought of as a vector quantity, and signal intensity depends strongly on the relative orientation of the α-helices whose signal dominates the observed spectrum. In the Gβγ subunit, the helical segments are relatively well aligned, and thus SFG signals can be quite strong at low tilt angles. For the GRK2-Gβγ complex, there are more but less well aligned helical segments, with a different net orientation that may result in weaker SFG signals for certain orientations. Because the molar amount of membrane associated GRK2-Gβγ should be similar to Gβγ alone, the drop in signal intensity upon formation of the GRK2-Gβγ complex provides an additional constraint that can help to determine the orientation of a single subunit, provided that the orientation of the other subunit is known or can be reasonably assumed. In total, two ppp/ssp ratios and one intensity change measurement are available as measurements. The best matches for these criteria are shown in Figure 3-5.

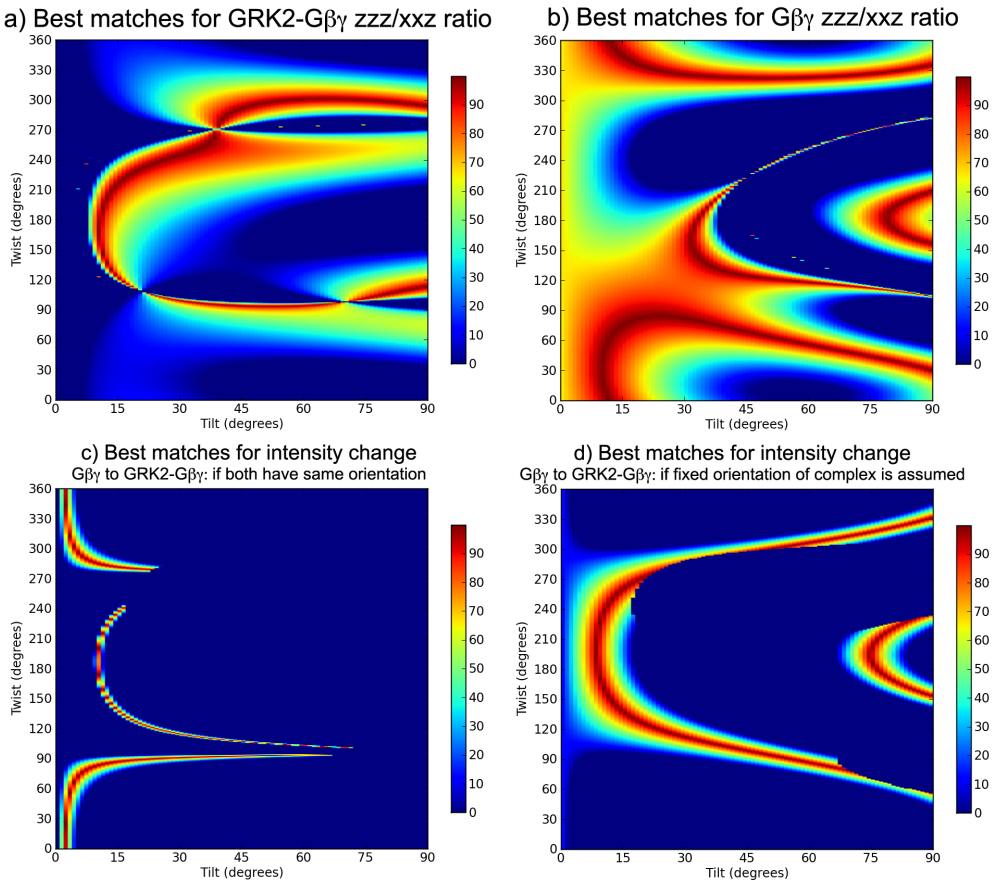


Figure 3-5: Best matches for each separate experimental measurement: the zzz/xxz ratios for (a) the GRK2-Gβγ complex and (b) Gβγ alone. Orientations that satisfy the observed drop in signal intensity upon formation of the complex, (c) if Gβγ does not reorient to bind GRK2, or (d) if it does, provided that the GRK2-Gβγ complex is assumed to adopt a single fixed orientation of θ= 10°, ψ=180° relative to the reference position. Colors indicate the quality of the match (100% = exact). The combination of panels a-c results in Figure 3-6a. The combination of panels b and d results in Figure 3-6b.

We might combine the experimental measurements with the assumption that Gβγ alone has the same orientation as it does in complex with GRK2. Thus, in addition to satisfying two $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratios (one each for Gβγ alone and the GRK2-Gβγ complex), the orientation must additionally provide a good match for the signal strength change $\frac{\chi^{(2)}_{G\beta\gamma,xxz}(\theta,\psi)}{\chi^{(2)}_{complex,xxz}(\theta,\psi)}$. Alternatively, one might assume that Gβγ alone can adopt an orientation

that is distinct from that in the GRK2-Gβγ complex, but that the overall orientation of the entire complex is a fixed, known position (such as the reference orientation shown in Figure 3-1). The measured $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio for Gβγ can then be combined with the signal strength change $\frac{\chi^{(2)}_{G\beta\gamma,xxz}(\theta,\psi)}{\chi^{(2)}_{complex,xxz}(\theta_{fixed},\psi_{fixed})}$, yielding the possible orientations of the Gβγ subunit in the absence of GRK2.

If instead one assumes that Gβγ does not reorient upon formation of the complex, two $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratios plus the intensity change can be applied. As shown in Figure 3-6a, the range of possible orientations is quite narrow, and even the best match has a score of only ~60%. The low scores in this analysis indicate less than optimal matches for all experimental measurements, supporting the hypothesis that Gβγ does not have the same orientation alone as it does when in complex with GRK2. Solutions that are physically plausible (as well as mathematically allowed) are further highlighted using a semi-transparent overlay (Figure 3-2) to indicate the positions where the geranylgeranyl anchoring group of Gβγ is close to the bilayer interface when in complex with GRK2. In this analysis, it appears that the overlay does not eliminate any of the possibilities, but neither does it improve the poor scores of the matches for experimental measurements.

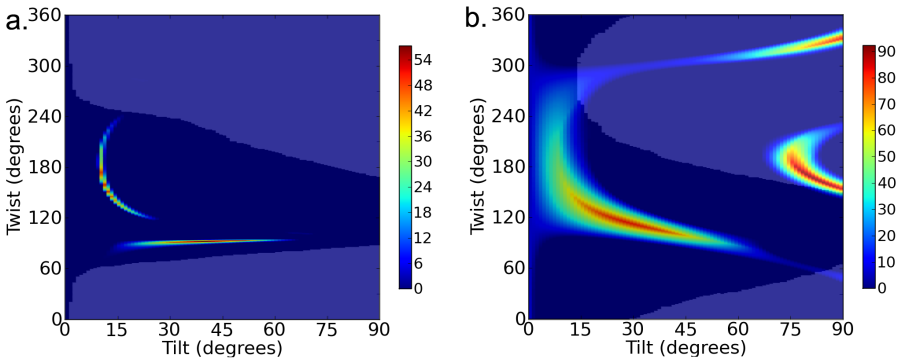


Figure 3-6: Best matches for all experimental measurements combined. (color scale indicates overall match quality for all criteria). (a) If Gβγ does not reorient, adding a constraint for intensity change yields possible orientations of both Gβγ and the GRK2-Gβγ complex. Using all three available measurements narrows the range of possibilities. (b) Orientation of Gβγ if the orientation of the complex is known or assumed. Use of two measurements helps to narrow down the best matches. For both panels, matches may be further limited down by adding an additional requirement that the positions of the protein be physically as well as mathematically allowed (dark overlay).

Another means of incorporating the drop in signal intensity as the GRK2-Gβγ complex is formed is to assume that the orientation of the GRK2-Gβγ complex is known. Then, the combined measurements can be used to characterize the orientation of Gβγ. As discussed above, the reference orientation (Figure 3-1) was based on a set of presumed membrane-interacting regions, but the measured $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio is not consistent with this orientation (Figure 3-5a). Based on the likely orientations of GRK2-Gβγ, I revised the membrane orientation to include a small tilt for the complex (θ=10°, ψ=180°), so that the second measurement becomes $\frac{\chi^{(2)}_{xxz,G\beta\gamma}(\theta,\psi)}{\chi^{(2)}_{xxz,complex}(10°,180°)}$. This new orientation of the complex is reasonable for the purposes of this analysis because other similarly high scoring orientations that match the experimentally measured $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio would generally yield similar values of $\chi^{(2)}_{xxz,complex}$, and therefore would not qualitatively alter the final

results (see supporting information for further discussion). The resulting matches produce high scores (>90%) and physically reasonable positions for Gβγ in terms of the resulting proximity of its geranylgeranyl group with respect to the membrane (Figure 3-6b). In this case, considering only physically reasonable positions helps to limit the range of matches in this analysis. The highest scoring positions are centered around θ=25°, ψ=120°, and do not overlap the new assumed position of the GRK2-Gβγ complex. Thus, regardless of how the intensity comparison is incorporated, this data indicates that Gβγ will likely alter its orientation upon engaging effectors such as GRK2. This is reasonable because both proteins are expected to form direct interactions with the cell membrane.

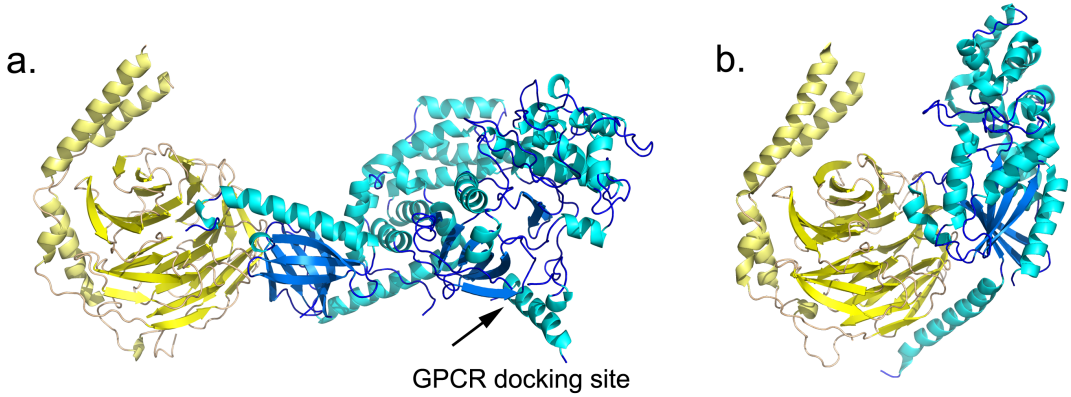


Figure 3-7: Predicted orientations of Gβγ complexes at phospholipid bilayers. a) The GRK2-Gβγ complex in a likely membrane orientation, with the membrane plane running along the bottom of the panel. The receptor docking site of GRK2 was homology modeled based on the structure of GRK6 (PDB entry 3NYN). The small tilt angle suggested by SFG measurements prevents this newly crystallized region from colliding with the lipid bilayer. b) The Gαβγ heterotrimer modeled in the same orientation, using Gβγ for alignment. The Gα subunit is shown in blue, and in this orientation it would maintain reasonable contacts with the lipid bilayer.

3.4.4 The Effect of Choosing Alternate Assumed Initial Positions of the Complex to Determine Gβγ Orientation

In section 3.4.2, I describe a method of orientation analysis that uses two experimental measurements to reveal the allowed orientations of the Gβγ subunit, provided that the initial orientation of the complex is known. As described above, the reference membrane was assigned via linear least-squares fitting of a plane passing through known (or expected) membrane-interacting residues. However, the calculated $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio for the GRK2-Gβγ complex did not match the experimentally measured value at this position, indicating that the actual orientation is somewhat different. If one assumes that the known membrane orientation is dictated by the geranylgeranyl anchoring group, then a very good match for the experimentally measured ppp/ssp ratio can be achieved by assuming that the complex adopts a small tilt (such as at the position $\theta = 10^o$, $\psi = 180^o$).

However, other physically allowed orientations of the complex can also produce the expected $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio. At these alternate positions of the complex, the calculated value of $\chi^{(2)}_{xxz}$ may be different from the value assumed in this work. This would in turn affect the measurement of the signal intensity change upon complex formation, by altering the value of $\chi^{(2)}_{xxz,complex}$ (determined by fitting the spectra) in the denominator of $\frac{\chi^{(2)}_{xxz,G\beta\gamma}(\theta,\psi)}{\chi^{(2)}_{xxz,complex}(\theta_{fixed},\psi_{fixed})}$.

Figure 3-8a shows the values of $\chi^{(2)}_{xxz}$ (fitted intensity, in a.u.) for the GRK2-Gβγ complex at all of the positions that match within ±10% of the experimentally measured $\chi^{(2)}_{zzz}/\chi^{(2)}_{xxz}$ ratio for the complex (one measurement), subject to the additional constraint that the possible matches are physically reasonable orientations (as determined above).

No scoring is applied. Insofar as the drop in signal upon formation of the GRK2-Gβγ complex depends on the numerical value $\chi^{(2)}_{xxz}$ rather than a position, then the search space for use in the intensity/fitted signal strength change measurement only requires using different values for $\chi^{(2)}_{xxz,complex}$ in the denominator. This is a much simpler way to perform the intensity comparison than exhaustively considering every possible position of the GRK2-Gβγ complex. Because many of the allowed orientations of the complex would generate similar SFG signal, the ratio of signals from the two samples does not have to be calculated for all 32,400 points on the 360^o by 90^o plot representing angles of interest. Rather, the comparison of intensities need only be performed for the relatively small number of different values of $\chi^{(2)}_{xxz,complex}$. This greatly simplifies the data analysis, and makes it possible to show whether Gβγ reorients upon complex formation- even if the assumed orientation of the complex is slightly different from the actual orientation.

In order to determine which contour lines representing similar intensity are of greatest interest, the values depicted in Figure 3-8a were histogrammed (Figure 3-8b).

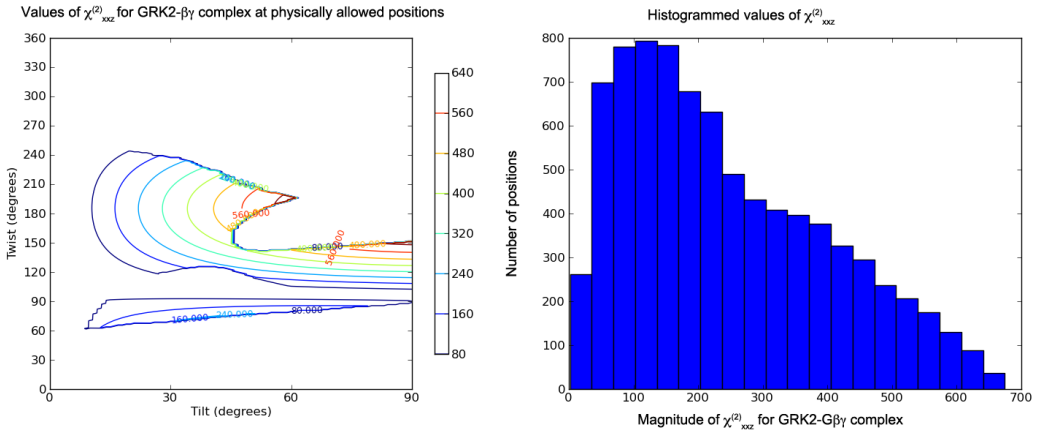


Figure 3-8: a) Possible values of $\chi^{(2)}_{xxz}$ for the complex, based on positions that match experimental measurements and physically reasonable positions. b) A histogram of the values of $\boldsymbol{\chi^{(2)}_{xxz}}$ displayed in (a). A range of values is possible.

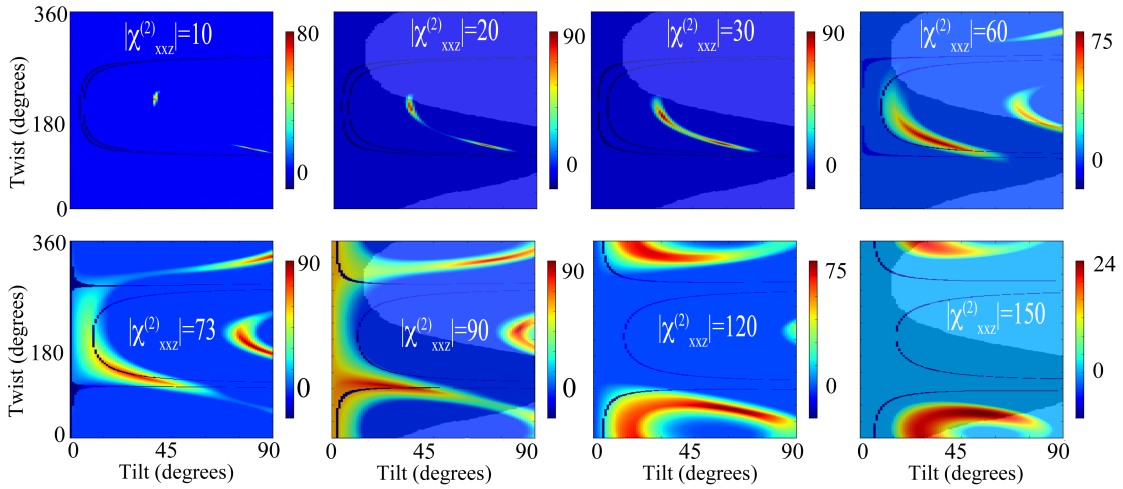


Figure 3-9: Allowed positions for Gβγ are depicted, based on various assumed positions (or signal strengths) for the complex. In order to better reflect that each value of $\chi^{(2)}_{xxz}$ can correspond to more than one position, the contour line corresponding to the chosen value of $\chi^{(2)}_{xxz}$ is drawn. These lines are shaded in dark blue by the simple expedient of assigning a value of -10 by default, or -20 if the contour line intersects any match for all constraints). Showing the contour line corresponding to the initial position(s) makes it easier to assess whether the allowed matches represent reorientation of Gβγ (these lines have been thickened to enhance visibility). The overlay representing physically allowed positions of Gβγ is also added. The $\chi^{(2)}_{xxz,complex}$=73 panel represents the orientation of the complex assumed in this work (θ=10°, ψ=180°).

In Figure 3-9, the effect of a change in the assumed position of the complex is explored. Each panel shows the results from the second method of orientation analysis, in which allowed positions of Gβγ are indicated by considering which positions match experimental measurements when one particular value of $\chi_{xxz}^{(2)}$ is assumed for the overall GRK2-Gβγ complex.

It is apparent that for < 30 or $\chi_{xxz}^{(2)}$ > 120, all possible matches become either low scoring or non-existent. Thus, the most likely positions of the complex are those that lie along a contour line within that range (such as the revised membrane orientation shown in Figure 3-7). The experimental measurements indicate that the signals from Gβγ are stronger than the signals obtained for GRK2-Gβγ, and this becomes harder to achieve as the value of $\chi_{xxz,complex}^{(2)}$ increases. Across this subset of images shown, it may be seen that the most likely positions of Gβγ do shift somewhat as the assumed position of the complex is varied. In all cases, there is practically no overlap between the original assumed position of the complex and the final allowed positions of Gβγ. At best, the points where Gβγ does not reorient upon binding GRK2 have scores no greater than 60-80%, and those positions would result in less optimal contacts with the lipid bilayer. Thus I conclude that even if the actual position of the complex is slightly different than that assumed in section 3.4.3, the conclusion that Gβγ likely reorients upon binding GRK2 would be qualitatively unchanged.

3.5 Conclusions

In this chapter, I demonstrated, for the first time, that SFG can be used to study the formation of a multi-subunit protein-protein complex *in situ*. In the process of

performing this research, a new software package was developed to facilitate SFG data analysis for proteins with many α-helical segments. To date, very few quantitative SFG studies have been attempted for large and multi-subunit proteins, and I expect this program to be of value as the field expands.

Based on experimental measurements, I found that Gβγ likely reorients slightly to facilitate binding of GRK2. The GRK2-Gβγ complex orientation slightly differs from its position that might be expected based on examination of the crystal structure alone, but is consistent with the optimal membrane orientation of the receptor docking site on the GRK2 homolog GRK6 (Figure 3-7). It should be noted that the GRK6 crystal structure was only solved quite recently[47], and that this orientation was predicted independently using SFG. The proposed orientation of Gβγ in the GRK2-Gβγ complex is also compatible with its expected orientation in complex with the heterotrimeric Gα subunit. The Gαβγ heterotrimer (PDB entry 1gp2) can be aligned with Gβγ from the GRK2-Gβγ complex in the specified position (θ=10°, ψ=180°) without collisions between the Gα subunit and the membrane (Figure 3-7b). As for the docking site of GRK2, this may allow the C-terminal helix of Gα, which is typically disordered in crystal structures, to dock productively with the cytoplasmic domain of activated receptors.

We have shown that it is possible to assess protein reorientation upon complex formation using a small set of readily obtainable SFG measurements. Calculated ratios also show good agreement with experimental values and physically reasonable positions. In the future, additional measurements such as SFG absolute intensities, polarized infrared or higher order spectroscopies, and isotope labeling can be introduced to yield a clearer and more specific picture of protein orientation for each individual subunit[48].

3.6 Acknowledgements

Spectra were collected and fit by Dr. Pei Yang. Dr. Valerie Tesmer prepared and filtered the protein samples. This work was supported by the National Institute of Health grants GM081655 (to Z.C.) and HL071818 and HL086865 (to J.T.). The authors are grateful to Dr. R. J. Lefkowitz (Duke University) for cell pellets containing GRK2 and Gβγ. I would also like to acknowledge B. Bucholz for conducting the flow cytometry protein interaction assays to measure binding between GRK2 (or GRK2-R587Q) and Gβγ.

3.7 References

(1) Chen, X.; Wang, J.; Boughton, A., P.; Kristalyn, C., B.; Chen, Z. *J. Am. Chem. Soc.* **2007**, *129*, 1420-1427.
(2) Nguyen, K. T.; Le Clair, S. V.; Ye, S.; Chen, Z. *J. Phys. Chem. B* **2009**, *113*, 12358-12363.
(3) Nguyen, K., Tan; Le Clair, S. V.; Ye, S.; Chen, Z. *J. Phys. Chem B* **2009**, *113*, 12169–12180.
(4) Ye, S.; Nguyen, K. T.; Boughton, A. P.; Mello, C. M.; Chen, Z. *Langmuir* **2009**, *26*, 6471-6477.
(5) Ye, S.; Nguyen, K. T.; Le Clair, S. V.; Chen, Z. *J. Struct. Biol.* **2009**, *168*, 61-77.
(6) Boughton, A. P.; Andricioaei, I.; Chen, Z. *Langmuir* **2010**, *26*, 16031-16036.
(7) Nguyen, K. T.; Soong, R.; Lm, S. C.; Waskell, L.; Ramamoorthy, A.; Chen, Z. *J. Am. Chem. Soc.* **2010**, *132*, 15112-15115.
(8) Ye, S.; Nguyen, K. T.; Chen, Z. *J Phys Chem B* **2010**, *114*, 3334-3340.
(9) Knoesen, A.; Pakalnis, S.; Wang, M.; Wise, W., D.; Lee, N.; Frank, C., W. *Journal of IEEE J Sel. Top. Quant.* **2004**, *10*, 1154-1163.
(10) Weidner, T.; Apte, J. S.; Gamble, L. J.; Castner, D. G. *Langmuir* **2010**, *26*, 3433-3440.
(11) Wang, J.; Lee, S.-H.; Chen, Z. *J. Phys. Chem. B* **2008**, *112*, 2281-2290.
(12) Fu, L.; Ma, G.; Yan, E. C. *J. Am. Chem. Soc.* **2010**, *132*, 5405-5412.
(13) Karp, G. *Cell and Molecular Biology: Concepts and Experiments*; Wiley: 2009;
(14) Gandhi, C. S.; Isacoff, E. Y. *The Journal of general physiology* **2002**, *120*, 455.
(15) Jiang, Y.; Ruta, V.; Chen, J.; Lee, A.; MacKinnon, R. *Nature* **2003**, *423*, 42-48.
(16) Olson, A. L.; Pessin, J. E. *Annu Rev Nutr* **1996**, *16*, 235-256.
(17) Neves, S. R.; Ram, P. T.; Iyengar, R. *Science* **2002**, *296*, 1636-1639.
(18) Sprang, S. R.; Chen, Z.; Du, X. *Adv. Protein Chem.* **2007**, *74*, 1-65.
(19) Tesmer, J. J. G.; Sunahara, R. K.; Gilman, A. G.; Sprang, S. R. *Science* **1997**, *278*, 1907-1916.

(20) Waldo, G. L.; Ricks, T. K.; Hicks, S. N.; Cheever, M. L.; Kawano, T.; Tsuboi, K.; Wang, X.; Montell, C.; Kozasa, T.; Sondek, J.; Harden, T. K. *Science* **2010**,
(21) Scheerer, P.; Park, J. H.; Hildebrand, P. W.; Kim, Y. J.; Krauss, N.; Choe, H. W.; Hofmann, K. P.; Ernst, O. P. *Nature* **2008**, *455*, 497-502.
(22) Pitcher, J. A.; Inglese, J.; Higgins, J. B.; Arriza, J. L.; Casey, P. J.; Kim, C.; Benovic, J. L.; Kwatra, M. M.; Caron, M. G.; Lefkowitz, R. J. *Science* **1992**, *257*, 1264.
(23) Dupre, D. J.; Robitaille, M.; Rebois, R. V.; Hebert, T. E. *Annu Rev Pharmacol Toxicol* **2009**, *49*, 31-56.
(24) Tesmer, V. M.; Kawano, T.; Shankaranarayanan, A.; Kozasa, T.; Tesmer, J. J. G. *Science* **2005**, *310*, 1686-1690.
(25) Lodowski, D. T.; Barnhill, J. F.; Pitcher, J. A.; Capel, W. D.; Lefkowitz, R. J.; Tesmer, J. J. G. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 936-939.
(26) Carman, C. V.; Barak, L. S.; Chen, C.; Liu-Chen, L. Y.; Onorato, J. J.; Kennedy, S. P.; Caron, M. G.; Benovic, J. L. *J. Biol. Chem.* **2000**, *275*, 10443.
(27) Shankaranarayanan, A.; Thal, D. M.; Tesmer, V. M.; Roman, D. L.; Neubig, R. R.; Kozasa, T.; Tesmer, J. J. *J. Biol. Chem.* **2008**, *283*, 34923-34934.
(28) Tamm, L., K.; McConnell, H., M. *Biophys. J.* **1985**, *47*, 105-113.
(29) Chen, X.; Chen, Z. *Biochim. Biophys. Acta* **2006**, *1758*, 1257-1273.
(30) Wang, J.; Even, M., A.; Chen, X.; Schmaier, A., H.; Waite, J., Herbert; Chen, Z. *J. Am. Chem. Soc.* **2003**, *125*, 9914-9915.
(31) Chen, X.; Boughton, A., P.; Tesmer, J., J. G.; Chen, Z. *J. Am. Chem. Soc.* **2007**, *129*, 12658-12659.
(32) Lodowski, D. T.; Pitcher, J. A.; Capel, W. D.; Lefkowitz, R. J.; Tesmer, J. J. G. *Science* **2003**, *300*, 1256-1262.
(33) Wang, J.; Paszti, Z.; Even, M., A.; Chen, Z. *J. Phys. Chem. B* **2004**, *108*, 3625-3632.
(34) Perry, J., M.; Moad, A., J.; Begue, N., J.; Wampler, R., D.; Simpson, G., J. *J. Phys. Chem B* **2005**, *109*, 20009-20026.
(35) Moad, A. J.; Moad, C. W.; Perry, J. M.; Wampler, R. D.; Goeken, G. S.; Begue, N. J.; Shen, T.; Heiland, R.; Simpson, G. J. *J. Comput. Chem.* **2007**, *28*, 1996-2002.
(36) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605-1612.
(37) Denbigh, K. G. *Trans. Faraday Soc.* **1940**, *36*, 936-948.
(38) Tsuboi, M.; Ikeda, T.; Ueda, T. *J. Raman Spectrosc.* **1991**, *22*, 619-626.
(39) Tsuboi, M.; Thomas, G., J. *Applied Spectroscopy Reviews* **1997**, *32*, 263.
(40) Lee, S.-H.; Krimm, S. *Chem. Phys.* **1998**, *230*, 277-295.
(41) Lee, S.-H.; Krimm, S. *J. Raman Spectrosc.* **1998**, *29*, 73-80.
(42) Lee, S.-H.; Krimm, S. *Biopolymers* **1998**, *46*, 283-317.
(43) Lee, S. H.; Wang, J.; Krimm, S.; Chen, Z. *J Phys Chem A* **2006**, *110*, 7035-7044.
(44) Pitcher, J. A.; Freedman, N. J.; Lefkowitz, R. J. *Annu. Rev. Biochem.* **1998**, *67*, 653-692.
(45) Lodowski, D. T.; Barnhill, J. F.; Pyskadlo, R. M.; Ghirlando, R.; Sterne-Marr, R.; Tesmer, J. J. G. *Biochemistry* **2005**, *44*, 6958-6970.
(46) Shen, Y., R. *Principles of Nonlinear Optics*; Wiley-Interscience: New York, 1984.

(47) Boguth, C. A.; Singh, P.; Huang, C. C.; Tesmer, J. J. *EMBO J*. **2010**, *29*, 3249-3259.
(48) Wang, J.; Paszti, Z.; Clarke, M., L.; Chen, X.; Chen, Z. *J. Phys. Chem. B* **2007**, *111*, 6088-6095.

Chapter 4: Orientation and Secondary Structure Change in Tachyplesin I: Molecular Dynamics and Sum Frequency Generation Spectroscopy Studies

4.1 Introduction:

In Chapter 2, I presented simulation studies focused on α-helical protein structures. These represent the first protein secondary structure motif for which SFG orientation measurements became possible, due to the fact that the well-ordered structure leads to strongly detectable signals, and because only a single angle (tilt) needs to be characterized. This is feasible even with a small set of measurements.

Recently, it has been shown that SFG is similarly capable of studying the orientation of β-sheet secondary structures, which consist of two hydrogen bonded β-strands in an antiparallel orientation.[1-9] Although this structure is highly symmetrical, it is not centrosymmetric, and weak Amide I signals are detectable. Since signal intensities are known to be proportional to the number of molecules at the interface, a recent study on β-sheet orientation increased the surface coverage of β-sheets by using the small peptide tachyplesin I, which consists of 17 residues (including six in the β-sheet conformation).

In order to better understand β-sheet orientation, a number of challenges remain that must be addressed. First, the need to characterize two orientation angles forces us to assume that all molecules adopt the same orientation (δ distribution). In fact, the structure of tachyplesin I is not as inherently facially amphiphilic[10] as magainin 2 (chapter 2), and

so it is unclear whether a δ distribution is actually realistic at a model hydrophobic/hydrophilic boundary such as the polystyrene/water interface. Second, despite the relatively high surface coverage of β-sheets, the actual region of tachyplesin I that adopts that secondary structure is fairly short: three residues per β-strand.[11] As a result, it is quite possible that the structure could deform significantly when adsorbed. Third and lastly, although a combination of ATR-FTIR and SFG measurements was used in the previous study, the orientation relations used for ATR-FTIR were not capable of extracting the full range of information in the spectrum. This is a consequence of the fact that FTIR is only capable of obtaining one measurement, yet the orientation of β-sheets must be described using two angles (tilt and twist). Previous orientation analysis methodologies have therefore required assuming that one orientation angle can be averaged out[12, 13] or held fixed.[5, 14, 15]

In this chapter, I report molecular dynamics simulation studies that are aimed at addressing the above concerns for a more complete understanding of β-sheet orientation at interfaces. The combination of SFG, new orientation relations for ATR-FTIR, and molecular dynamics simulations makes it possible to revisit previous studies in light of the above questions, and will facilitate future studies of β-sheet structures at interfaces.

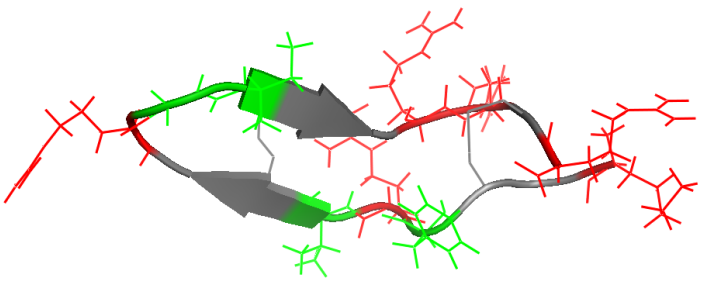


Figure 4-1: Tachyplesin I is a 17-residue β-sheet peptide with both hydrophobic (green) and charged (red) residues.

4.2 Methods

4.2.1 SFG and FTIR Spectroscopy

Synthetic C-terminal amidated tachyplesin I (NH_2-K-W-C-F-R-V-C-Y-R-G-I-C-Y-R-R-C-R-$CONH_2$) was obtained from GenScript Inc. (Piscataway, NJ) at >95% purity.

The experimental setup has been described in a previous publication. [5] The model polystyrene thin film/peptide interface was similar to that described for magainin 2 in chapter 2. SFG spectra were collected by contacting the PS film with a ~550nM solution of tachyplesin I in the ssp and spp polarization combinations. To confirm that the signals originate from β-sheet secondary structures (and not random coils), spectra were also collected after the reducing agent dithiothreitol (DTT) was added to the solution in order to cleave the disulfide bonds (residues 3-16 and 7-12) that are known to stabilize the structure.[16]

ATR-FTIR spectra were collected on a Nicolet 550 spectrometer (Thermo Fisher Scientific, MA, USA) with 1 cm^{-1} resolution. The polystyrene film was deposited by solution casting from a 0.1 wt% solution onto a clean germanium substrate, and brought into contact with D_2O. 50 μL of a 0.5mg/mL solution were injected into the trough, for a final peptide concentration of ~550nM (signals were not detectable at lower concentrations). The chamber was purged with nitrogen, and after the sample had been allowed to equilibrate for one hour, spectra were collected in the s and p polarizations.

4.2.2 SFG Data Analysis

It has only recently become possible to analyze the orientation of β-sheets, following the derivation of the required molecular hyperpolarizability and surface susceptibility tensor components.[5, 6] Briefly, the β-sheet structure possesses D_2 symmetry,

with three SFG-active vibrational modes: the B_1, B_2, and B_3 modes. This high symmetry can complicate orientation analysis: signals from β-sheets may be very weak (though detectable), and unlike α-helices, two angles (tilt and twist) must be characterized in order to develop a clear picture of molecule orientation. The most relevant experimentally observed quantities are:

$$\chi^{(2)}_{eff,ssp} = L_{xxz}\chi^{(2)}_{xxz}$$

$$\chi^{(2)}_{eff,spp} = L_{yxz}\chi^{(2)}_{yxz} + L_{yzx}\chi^{(2)}_{yzx} \qquad \text{Equation 4-1}$$

The orientation relations required for the B_1, B_2, and B_3 modes are of identical form for the ssp and ppp polarizations, but depend on different components β_{ijk} of the molecular hyperpolarizability tensor: β_{abc}, β_{abc}, and β_{bca}, respectively.

$$\begin{aligned}\chi^{(2)}_{xxz} &= \quad N_s\left\langle \cos^3\theta\cos\psi\sin\psi - \cos\theta\cos\psi\sin\psi\right\rangle\beta_{ijk}\\ \chi^{(2)}_{zzz} &= -2N_s\left\langle \cos^3\theta\cos\psi\sin\psi - \cos\theta\cos\psi\sin\psi\right\rangle\beta_{ijk}\\ \chi^{(2)}_{yxz} &= 0\end{aligned} \qquad \textit{Equation 4-2}$$

The relations for the chiral components probed in the spp polarization have different forms depending on the vibrational mode, as indicated. The experimental observables and fitted signal ratio $\frac{\chi^{(2)}_{ssp}}{\chi^{(2)}_{spp}}$ are shown in *Figure 4-2*.

$$\begin{aligned}\chi^{(2)}_{yzx,B1} &= \frac{1}{2}N_s\left\langle \sin^2\theta\sin^2\psi - \sin^2\theta\cos^2\psi\right\rangle\beta_{abc}\\ \chi^{(2)}_{yzx,B2} &= -\frac{1}{2}N_s\left\langle \cos^2\theta - \sin^2\theta\cos^2\psi\right\rangle\beta_{acb}\\ \chi^{(2)}_{yzx,B3} &= -\frac{1}{2}N_s\left\langle \cos^2\theta - \sin^2\theta\sin^2\psi\right\rangle\beta_{bca}\end{aligned} \qquad \textit{Equation 4-3}$$

At first glance, this would appear to enable characterization of rich structural detail by providing many experimentally measured parameters from spectral fitting results (three vibrational modes each across the ssp, ppp, and chiral spp polarizations). In

fact, calculated values of the hyperpolarizability tensor components β_{abc} and β_{bca} (for the B_1 and B_3 modes, respectively) are roughly one third the value of β_{acb}, and thus the B_2 mode will be roughly nine-fold stronger in spectral intensity. This observation is supported by previous SFG results in which spectra were dominated by the B_2 peak at 1635 cm^{-1}.

Within the subset of nine potential observables (three polarizations and three vibrational modes), the orientation relations for $\chi^{(2)}_{xxz}$ and $\chi^{(2)}_{zzz}$ are identical within a constant factor of two. Thus, the primary SFG result is the single intensity ratio $\frac{\chi^{(2)}_{xxz}}{\chi^{(2)}_{yzx}}$.

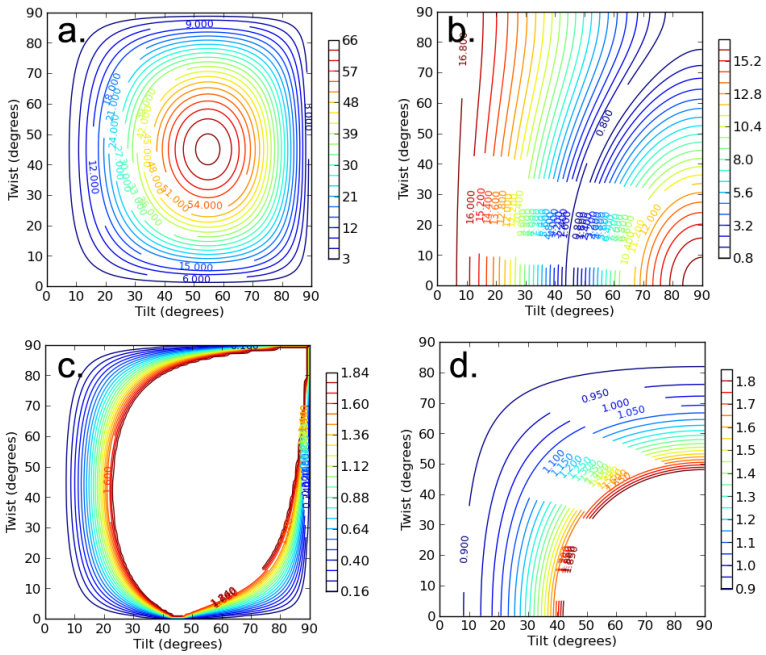


Figure 4-2: Calculated contour plots showing the relative magnitudes of a) $\chi^{(2)}_{ssp}$, b) $\chi^{(2)}_{spp}$, and c) the fitted ssp/spp ratio (these ratios incorporate the fresnel factors directly). d) Contour plot showing the value of the FTIR dichroic ratio.

4.2.3 Molecular Dynamics Simulations

The solvent/surface model system described in chapter 2 was used for these simulations. The structure of the tachyplesin I peptide was obtained from the Protein Databank (PDB ID 1ma2). Structure 9 of the 31-conformer NMR ensemble was extracted and used as the starting structure for all simulations. This starting structure was chosen due to the roughly parallel orientation of the two strands, particularly in the highly flexible disordered regions of the structure. This conformation therefore provides the best chance for all portions of the peptide to interact with the surface in a variety of initial simulation conditions. For a hydrophobic polymer surface, one would expect that hydrophobic interactions would dominate the adsorption behavior.[17-19] However, the structure of tachyplesin I does not possess the same strong facial amphilicity[10] as was present in magainin 2 (studied previously). Therefore, a variety of simulations were performed with the peptides placed in different starting orientations.

As described in chapter 2, the peptide was placed into the model solvent/surface system above the surface, so that the closest point of the peptide was ~2.5 Å from the surface. The peptide was oriented along the x-axis of the coordinate system, then rotated into the final coordinates as specified. Overlapping water molecules were removed, and peptide coordinates were held fixed as an additional 500 ps of equilibration was then performed for the solvent and surface in the presence of peptide.

For production runs, dynamics were performed using a Nosé-Hoover thermostat at 298 K with the velocity Verlet algorithm. The SHAKE algorithm was used to constrain all bonds involving hydrogen atoms, allowing the use of a 2 fs timestep. All simulation parameters were taken from the CHARMM22 parameter set.

Experiments have shown that the presence of the two disulfide bonds in tachyplesin I is key to retention of the β-sheet structure[16]. As a result, initial conditions shown to produce adsorption of the peptide were performed again with the disulfide bonds removed.

The most relevant exploratory simulations are summarized in Table 4-1, where the angles specified represent rotation of the entire PDB coordinate file. To facilitate comparison to results from SFG spectroscopy, these initial positions were also calculated in terms of the Euler angles required to rotate a β-sheet from the yz plane to the specified position. In Table 4-1, not all tilt angles are 90 (lying down perfectly). This is because the six β-sheet residues are not perfectly aligned with the overall peptide plane defined by all 17 total residues. The Euler angles are defined based on two vectors located parallel and perpendicular to the strand axis, respectively, and may be calculated from the rotation matrix used for SFG:

$$\theta = cos^{-1}\left(\frac{-z}{|length\ vec|}\right) \qquad \textit{Equation 4-4}$$

$$\psi = sin^{-1}\left(\frac{z}{|widthvec| * sin\ \theta}\right) \qquad \textit{Equation 4-5}$$

Identifier	Length (ns)	PDB rotation		Euler Angles	
		Tilt	Twist	Tilt	Twist
tach2*	12 (18*)	90°	270°	100°	65°
tach3*	4.8 (18*)	90°	0°	110°	-32°
tach3x315*	12	90°	315°	119°	21°
Tach3x180*	12	90°	180°	76°	28°
Tach3x235*	6 (18*)	90°	235°	109°	42°
Tach4	2	90°	90°	77°	-75°

Table 4-1: Initial conditions used for simulations, where PDB rotation indicates the angles used to rotate the entire peptide molecule (based on all backbone atoms). Angles are also listed for the β-sheet portion of the peptide, in a convention consistent with SFG analysis. Simulation identifiers marked with an asterisk were also performed a second time with disulfide bonds removed.

For final data analysis, description of the resulting orientation is hampered by the fact that a significant portion of the peptide may be disordered. (At best, only 6 of the 17 residues comprise the β-sheet region) Thus, in all results reported, the orientation of the molecule was defined solely from the β-sheet portion of the molecule (residues 6-8 and 11-14) using the latter Euler angle convention described above.

4.3 Molecular Dynamics Results

Including simulations in which the disulfide bonds were removed, a total of 11 molecular dynamics simulations spanning a total of ~120ns were evaluated here. Of these, two initial conditions (tach4 and tach3x235) showed no or poor interaction with the surface, and were therefore excluded from the analysis.

Name	With 3 residues			With 2 residues		
	Tilt	Twist	Interstrand Angle	Tilt	Twist	Interstrand Angle
Tach2	107°	52°	12°	106°	28°	34°
Tach3	94°	60°	39°	92°	48°	12°
Tach3x180	88°	-8°	28°	102°	-5°	28°
Tach3x315	85°	12°	9°	83°	13°	23°

Table 4-2: Final orientation of the β-sheet section of the peptide when disulfide bonds are retained. Angles are from the average of the last 50 frames in a rotation convention consistent with SFG data analysis. In order to account for deformation of the sheet structure, results for the entire sheet (residues 6-8 and 11-13) were compared to angles for a shorter segment (residues 6-7 and 12-13).

In general, all simulations showed the peptide lying roughly parallel to the plane of the surface (tilt of 90°), but a variety of twist angles were observed; the two most common local minima in time-dependent plots were twist angles centered around 0 and 45°. The changing angle between the β-strands also indicates that structural deformation of the β-sheet region may affect the orientation of the peptide. Since the orientation is defined based on the position of the backbone atoms, it is important to note that it will only have physical meaning if the peptide retains the β-sheet conformation at the end of the simulations.

In order to evaluate loss of secondary structure, the following metrics were used:

a. Root-mean-squared displacement (RMSD) was evaluated as a metric for deformation of the β-sheet portion of the peptide (residues 6 to 8 and 11 to 13), as well as for the backbone atoms and the overall peptide.
b. In order to separate strands in the extended conformation from fully hydrogen-bonded β-sheets, the number of hydrogen bonds in the β-sheet and backbone was

quantified over time. Hydrogen bonds were quantified using the default CHARMM parameters.

c. Plots of secondary structure over time were prepared using the STRIDE algorithm within VMD.[20, 21]

These results are summarized in the following several sections.

4.3.1 Adsorption and Orientation

In order to evaluate the feasibility of multiple peptide orientations, a variety of molecular dynamics simulations were performed in which the peptide was placed at different initial orientations. The results (Table 4-2) are striking for the variety of final orientations and structures observed. The simulations reported in Table 4-1 focus on the subset of ten trajectories that showed some interaction with the surface: five in which the disulfide bonds were kept intact, and five additional windows in which the disulfide bonds were removed. It is interesting to note that in one case (tach3x235), the peptide only adsorbed when disulfide bonds were broken. This suggests that reorganization of the backbone may help to accommodate rearrangements of side chains and alleviate unfavorable contacts with the surface. For example, in the simulation window tach2, the loss and recovery of secondary structure correlates to initial burial of the aliphatic chain of an arginine at the hairpin turn (residue 9). The adsorbed peptide eventually forms more (and more favorable) hydrophobic contacts, and recovery of the β-sheet structure in tach2 coincides with release of arginine from the surface. In the analogous simulation without disulfide bonds, however, no such disruption of secondary structure was seen on the timescales studied. Reorganization of the bulky tyrosine side chains at residues 8 and 13

was also seen to influence the total disruption of secondary structure in the window tach3x180.

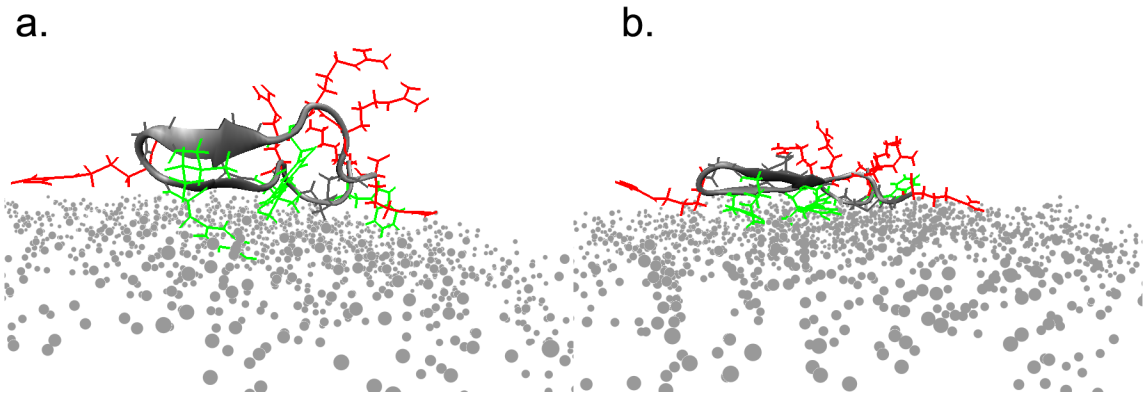
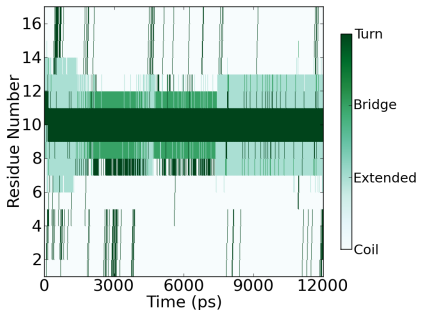
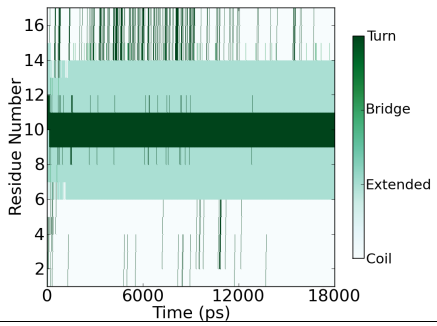


Figure 4-3: Two candidate final orientations of the peptide show qualitatively similar trends in surface area burial. Hydrophobic side chains are shown in green, and charged side chains are drawn in red. The peptide adopts a twist angle of a) ~50° (simulation tach2) or b) ~10° (simulation tach3x315).

4.3.2 Structural Deformation

A variety of structural deformation metrics were applied in an attempt to better understand the behavior of the peptide at the interface. These were presented in the results, with the most direct visual comparison presented in Table 4-3.

Name	With disulfide bonds	Without disulfide bonds
Tach2		

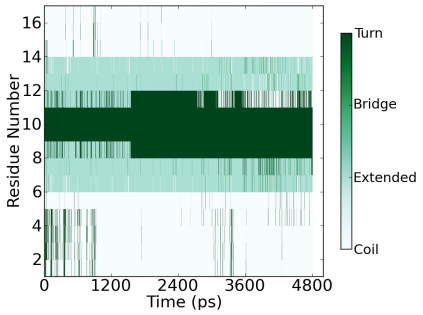
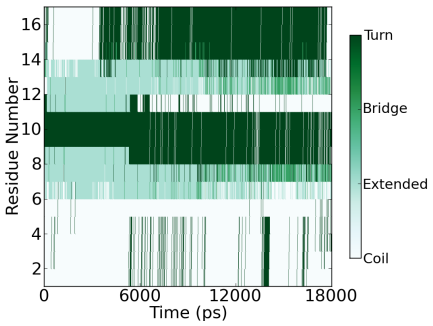
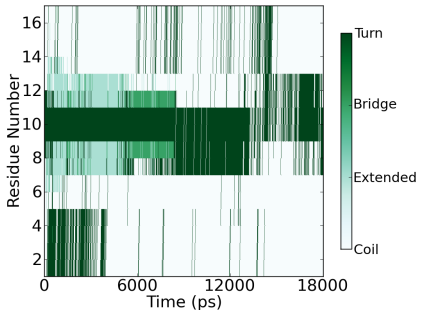
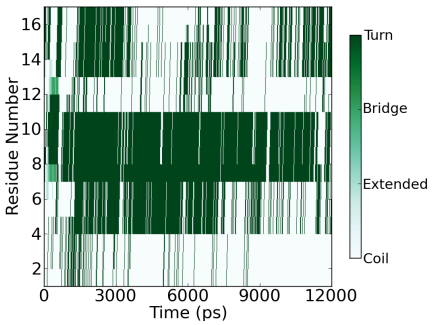
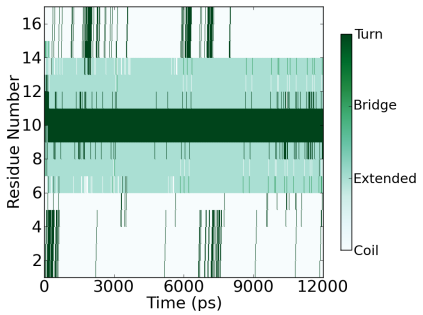
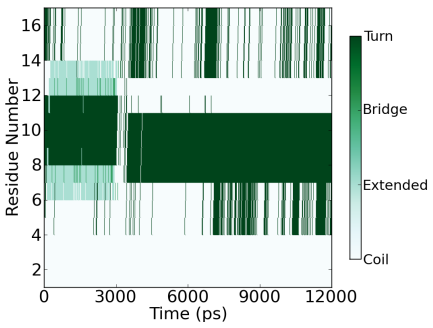
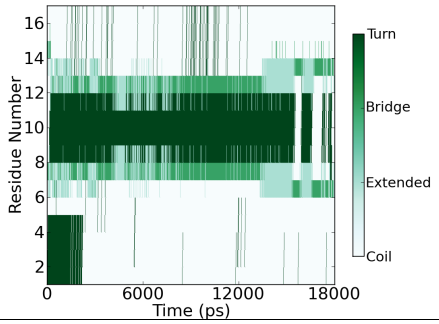
Tach3		
Tach3x180		
Tach3x315		
Tach3x235	Did not adsorb to surface	

Table 4-3: Secondary structure assigned to each residue of the peptide over time. Left: with disulfide bonds present. Right: with disulfide bonds removed.

Strikingly, I found that after 12-18ns of simulation, the very short β-sheet region of the peptide can be deformed even when disulfide bonds are present (as in tach3x180)- though as discussed above, this may be an artifact of side-chain reorganizations on short

timescales. When the disulfide bonds are removed, the β-sheet structure was only retained in window tach2. In runs tach3 and tach3x235, only a single residue remained transiently in the bridged conformation.

By contrast, when disulfide bonds were retained, secondary structure was generally retained to a much higher degree. Only in one of the trajectories was secondary structure completely lost during ongoing reorganizations of side chain residues as the peptide reoriented from a non-optimal initial position (tach3x180). Partial retention or even recovery of the secondary structure was observed in two of the four trajectories for which the peptide was found to adsorb to the surface (tach2 and tach3). The results in Table 4-3 are in reasonable agreement with the hydrogen bond and RMSD quantification in Table 4-4. It is interesting to note that neither metric alone is sufficient to unequivocally indicate peptide folding, however. This is likely due to the fact that the β-sheet portion of tachyplesin I is very short, and so small local deformations are possible without loss of secondary structure. This highlights the importance of backbone dynamics in interpreting molecular orientation, as most orientation analysis methods require a known and fixed structure. Even when the β-sheet secondary structure is preserved (as for tach3x315), my simulations show that the strands are not arranged in a completely ideal and extended conformation, with inter-strand angles of 10° or higher.

Name	# hydrogen bonds		RMSD (Å)	
	βsheet	Backbone	βsheet	Backbone
Tach2	2.6	3.5	0.6	2.5
No disulfide bonds	4.8	5.4	0.5	4.6
Tach3	1.5	1.6	1.1	3.0
No disulfide bonds	1.5	2.1	1.1	5.2
Tach3x180	2.8	4.0	1.9	2.1
No disulfide bonds	0.6	0.8	3.7	5.6
Tach3x315	4.1	4.3	0.75	3.2
No disulfide bonds	1.3	0.3	2.8	4.1
Tach3x235 (No disulfides, only)	1.8	2.7	1.1	3.5

Table 4-4: Average number of hydrogen bonds over the last 1 ns of each simulation trajectory.

4.3.3 The Stability of the Final Adsorbed State

The highly deformable peptide backbone of tachyplesin I prevents the calculation of a free energy profile along the lines of the methods used in chapter 2: it appears that such deformation is important to the initial adsorption process, but addition of an artificial constraint might lead to other less meaningful deformation of the peptide rather than restraining peptide orientation. Based on the assumption that hydrophobic surface area burial drives a change in free energy upon adsorption[17-19], solvent accessible surface area was calculated for all windows in the simulation based on the surface area of the residue + surface, minus the surface area of the side chain in solvent.[22]

Name	Change in solvent accessible surface area (SASA)- units of Å^2			
	Overall	Hydrophobic	Hydrophilic	Charged
Tach2	887	639	466	277
No disulfides	1659	850	1110	894
Tach3	1273	800	765	719
No disulfides	1582	862	852	651
Tach3x180	1295	614	696	632
No disulfides	1411	931	772	769
Tach3x315	1458	792	786	593
No disulfides	1147	589	690	670
Tach3x235 (No disulfides, only)	1287	304	872	726

Table 4-5: Burial of solvent accessible surface area by residue type. Individual tallies do not add up to 100% due to limitations in the algorithm used.

In all but one simulation window (tach3x315), the removal of disulfide bonds leads to greater surface area burial (*Table 4-5*). Yet since simulations with high hydrophobic surface area are accompanied by higher burial of other residues as well, it is unclear how much of a driving force this will provide, and calculations of the total potential energy of the system were inconclusive. By contrast, the strong facial segregation of residues in magainin 2 (Chapter 2:) created a clear driving force for adsorption via burial of hydrophobic surface area. In tachyplesin I, residues are poorly segregated, and side chains are free to reorganize. This leads to a variety of candidate orientations with tradeoffs between favorable and unfavorable peptide-surface interactions. Other factors, such as structural deformation and entropic considerations, may therefore play a significant role in determining the overall free energy of adsorption.[23] In these

simulations, the potential energy and/or peptide RMSD is still changing after 18ns, suggesting that final conclusions about the most stable peptide orientations may be timescale limited. However, the lack of any single unambiguously preferred state from the variety of initial orientations examined points to the possibility of a more diverse free energy landscape in which a range of twist values would be accommodated.

4.4 Experimental Results and Discussion

4.4.1 SFG and ATR-FTIR Spectra

SFG spectra were collected for ~550 nM concentrations of tachyplesin I at the polystyrene surface. Spectra were collected in the ssp and spp polarizations.[5, 6] The ssp spectrum shows contributions from β-sheet structures and disordered regions, with the dominant contribution from the B_2 mode at 1635 cm^{-1}.(Figure 4-4).

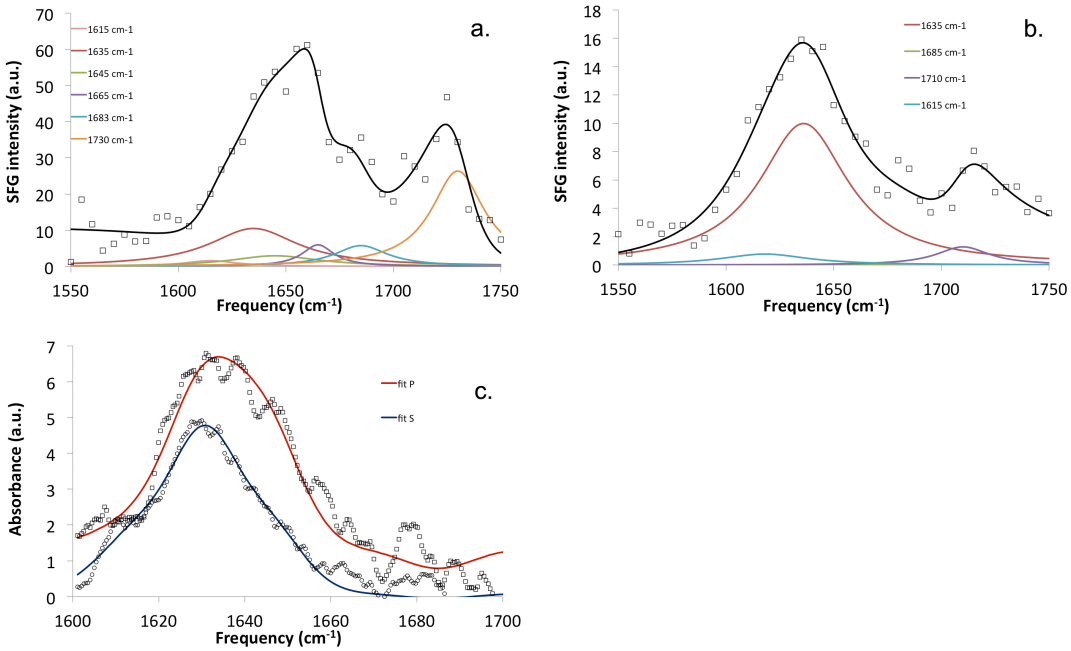


Figure 4-4: Amide I region spectra collected from ~550 nM solutions of tachyplesin I on a polystyrene surface in the a) ssp and b) spp polarization combinations for SFG. The component peaks and overall fitting results are shown as solid lines. c) ATR-FTIR spectra in the s and p polarizations.

Peak assignments[6, 14, 24] are summarized in Table 1-1. It is known that the secondary structure is lost when the disulfide bonds are cleaved, and these peaks indeed disappeared when DTT was added. Other signals could still be observed from random coil structures.[6] Signals disappeared completely in the spp polarization, which is uniquely sensitive to chiral signals (such as those from β-sheets). This observation generally agrees with my simulation results showing rapid loss of secondary structure when the disulfide bonds are removed.

As discussed above, SFG data analysis and orientation determination was focused on the B_2 mode peak at 1635 cm^{-1}, since other signals were too weak to ensure reliable ratios for analysis. Spectra were fit as described in chapter 1. For total internal reflection geometry, and taking $n' = \frac{n_1+n_2}{2}$ as the corrected index of refraction at the interface[25, 26], the normalized Fresnel factors L_{xxz}, L_{yxz}, and L_{yzx} are 10.1, 3.5, and 1.0, respectively.[27]

4.4.2 Combined Spectroscopies for Orientation Determination

Orientation analysis may be performed by considering the ratio of the experimental observables $\chi^{(2)}_{eff}$ in the polarizations of interest. As discussed in section 4.2.2, only one vibrational mode (B_2) produces sufficiently intense SFG amide I mode signals for orientation analysis, so that the SFG measured quantity for the B_2 mode peak centered at 1635 cm^{-1} is:

$$\frac{\chi^{(2)}_{ssp}}{\chi^{(2)}_{spp}} = \frac{L_{yyz}}{L_{yxz}} \frac{N_s\langle cos^3\theta cos\psi sin\psi - cos\theta cos\psi sin\psi\rangle\beta_{acb}}{-\frac{1}{2}N_s\langle cos^2\theta - sin^2\theta cos^2\psi\rangle\beta_{acb}} \qquad \textit{Equation 4-6}$$

In order to characterize both of the required orientation angles (tilt and twist), a second measurement was obtained from Polarized Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) Spectroscopy (Figure 4-4c).

Orientation relations for β-sheet structures have been published previously for ATR-FTIR[14, 28], and were employed in a previous study. Here, molecular orientation was considered only in terms of the tilt angle θ:

$$R^{ATR}(amide\ \ I) = \frac{E_x^2}{E_y^2} + \frac{2\langle\cos^2\theta\rangle E_z^2}{3 - \langle\cos^2\theta\rangle E_y^2} \qquad \textit{Equation 4-7}$$

In this equation, the twist angle is implicitly considered and assumed to be fixed. Although an equation that is dependent on a single variable is appropriate for analysis of ATR-FTIR spectra (for which only one measurement is available), two angles are in fact necessary to describe molecular orientation. The combination of SFG with ATR-FTIR provides the measurements necessary. Therefore, I have re-derived these orientation expressions in a zyz rotation convention consistent with the relations for SFG, where the ratio of signals from s and p polarized IR beams is given by:

$$R^{ATR} = \frac{E_x^2}{E_y^2} + \frac{\langle M_z^2\rangle}{\langle M_y^2\rangle}\frac{E_z^2}{E_y^2} \qquad \textit{Equation 4-8}$$

The angled brackets indicate averaging over the azimuthal angle ϕ, indicating that the molecules are randomly oriented in the plane of the surface, and M_x, M_y, and M_z refer to the respective components of the dipole moment in the lab frame coordinate system with axes x, y, and z. If one assumes that the entire dipole moment of a peptide unit lies along the c axis of its coordinate system, and is projected by an angle Θ into the peptide coordinate frame, then the expected dipole moment components for the β-sheet in the lab frame required by equation 4-8 can be derived. Here, the molecular coordinate system (axes a, b, c) for the peptide is defined so that the plane of the sheet corresponds to the bc plane.

$$\mu_{lab} = R_z(\phi)R_y(\theta)R_z(\psi)R_y(\Theta)\begin{bmatrix}0\\0\\|\mu|\end{bmatrix} \qquad \textit{Equation 4-9}$$

where the rotation matrices about the intrinsic molecular y and z axes for an arbitrary rotation angle α are

$$R_y(\alpha) = \begin{bmatrix}cos\,\alpha & 0 & -sin\,\alpha\\0 & 1 & 0\\sin\,\alpha & 0 & cos\,\alpha\end{bmatrix} \qquad R_z(\alpha) = \begin{bmatrix}cos\,\alpha & sin\,\alpha & 0\\-sin\,\alpha & cos\,\alpha & 0\\0 & 0 & 1\end{bmatrix}$$

Equation 4-10

Provided that the individual transition moments are oriented perpendicular to the strand axis, then the angle Θ=90° for the Amide I mode, resulting in a new expression for the dichroic ratio that takes into account both tilt and twist angles:

$$R^{ATR}(amide\ I) = \frac{E_x^2}{E_y^2} + \frac{2\langle cos^2\psi\rangle\langle sin^2\theta\rangle}{1+\langle sin^2\psi\rangle\langle sin^2\theta\rangle-\langle sin^2\theta\rangle}\cdot\frac{E_z^2}{E_y^2} \qquad \textit{Equation 4-11}$$

This equation allows more information to be extracted from the combination of SFG and ATR-FTIR measurements, by removing inherent assumptions about one angle. Expressions for the electric field amplitudes E_x^2, E_y^2, and E_z^2 have been derived previously,[24] and depend on the indices of refraction for the substrate, interfacial layer (in this case, the polystyrene film), and solvent. These indices are taken to be 4.0, 1.55, and 1.328, respectively.[5, 24, 29-31] For the polymer interface, this leads to values for E_x^2, E_y^2, and E_z^2 of 1.97, 2.25, and 1.36, respectively. The E_z^2 component is quite sensitive to small changes in the refractive index, so again, care must be taken to consider the effect of these errors on the final ratios obtained.

In order to relate the experimental observables to calculated molecular properties for both measurements, one may consider all orientations that lie within +/- 5% of both experimentally calculated quantities, satisfying the constraints:

$$\begin{cases} \dfrac{\chi^{(2)}_{ssp}}{\chi^{(2)}_{spp}} = \dfrac{L_{yyz}}{L_{yzx}} \dfrac{N_s\langle cos^3\theta cos\psi sin\psi - cos\theta cos\psi sin\psi\rangle\beta_{acb}}{-\frac{1}{2}N_s\langle cos^2\theta - sin^2\theta cos^2\psi\rangle\beta_{acb}} = 1.04 \\ R^{ATR}(amide\ I) = \dfrac{E_x^2}{E_y^2} + \dfrac{2\langle cos^2\psi\rangle\langle sin^2\theta\rangle}{1 + \langle sin^2\psi\rangle\langle sin^2\theta\rangle - \langle sin^2\theta\rangle} \cdot \dfrac{E_z^2}{E_y^2} = 0.95 \end{cases}$$

Equation 4-12

If one assumes that all molecules of peptide on the surface adopt exactly the same orientation, then the angles θ and ψ can be solved for directly from the above as a system of nonlinear equations. However, this may result in multiple possible solutions, and the relationship between results is not always clear- particularly when large error bars must be quoted due to uncertainty in the Fresnel coefficients and interfacial electric field amplitudes. Hence I have chosen to apply the graphical presentation developed in chapter 3 to display the positions that best match experimental results when equation 4-11 is used (Figure 4-5). All results within 5% error bars of all criteria are shown, but due to the large potential variation in Fresnel coefficients, no scoring is applied for match quality.

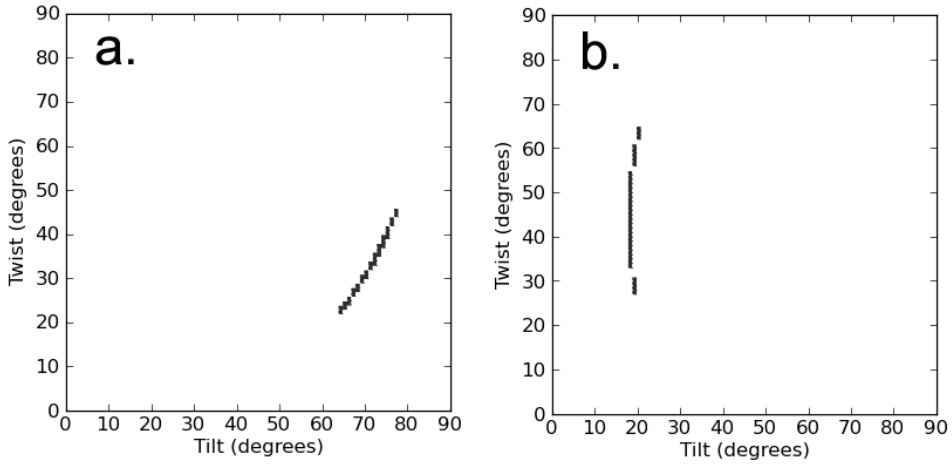


Figure 4-5: Best matches for all experimental measurements: a) Using equation 4-7, and b) using the new FTIR equation 4-11. The calculated values must satisfy three constraints: a match for the SFG ssp/spp ratio, a match for the ATR-FTIR dichroic ratio, and B_2 signal greater than B_1 signal. (In line with the experimental observation)

Upon incorporating the new equation for ATR-FTIR data analysis, only orientations that are mathematically allowed (Figure 4-5b) are at very low tilt angles (15-20°), at which the molecular axis is perpendicular to the plane of the surface. In order to better illustrate the effect of possible larger uncertainties in the experiments and data analysis, the scoring function and graphical presentation developed in chapter 3 was applied to illustrate all matches within ±40% of the target criteria (Figure 4-6). Even with these very large error bars, the peptide could adopt a tilt angle of no more than 60°, and then only by adopting very high twist angles that would remove one strand from being in contact with the surface completely.

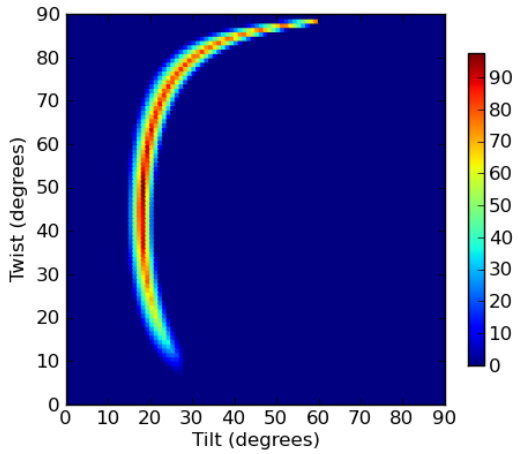


Figure 4-6: Possible orientations of Tachyplesin I are shown for all matches within ±40% of the target criteria, with the scoring functions developed in Chapter 3.

Low tilt angles would provide extremely poor contact between the peptide and the surface, and could only be adopted if either the charged residues of the disordered tail or the charged arginine at the β-turn (residue 9) were buried on the surface while hydrophobic contacts are minimal. In my simulations, such positions are explored. Relevant initial conditions either quickly collapsed down to lie flat on the surface, failed to adsorb (tach3x235), or (in the case of tach3x180) were accompanied by significant distortion of the peptide backbone as dramatic side chain reorganizations took place. High twist angles were not seen as a candidate stable orientation in my simulations at all. The physical likelihood of the best match positions is also not improved by assuming that all peptide molecules adopt roughly the same singular orientation within a Gaussian distribution of tilt and/or twist angles. It has previously been shown that even a small fraction of peptide molecules adopting a different orientation (or set of orientations) can lead to dramatic and sharp changes in the experimentally observed signal intensity ratios[32], and my simulation results point to the possibility of multiple distinct twist angles that allow similarly favorable peptide-surface contacts to form.

4.5 Conclusions

New experimental techniques provide a powerful means to probe the orientation of peptides at interfaces, but often require key assumptions. In this chapter, I explore several of those assumptions with regards to the adsorption behavior of the small β-sheet peptide tachyplesin I. The orientation and secondary structure change was characterized using a variety of metrics. General agreement was found between simulations and experiments with regards to the role of disulfide bonds in preserving the secondary structure, but simulations revealed a flexible interfacial structure that belies the assumption that all molecules would adopt a δ distribution. In some simulation windows, the overall secondary structure of tachyplesin I was partially disrupted. Thus despite the observation of signals from β-sheet regions of the peptide, care must be taken in applying orientation analysis methodologies to such short sequences. Due to limitations in simulation timescale, it is possible that this conformational flexibility is an artifact of the short simulations, but the poor facial segregation of residues in the tachyplesin I structure appears to result in a variety of meta-stable orientations or conformations.

These findings from molecular dynamics were used to inform the interpretation of results from SFG, and a new equation was derived that incorporated the twist angle to take advantage of the combination of ATR-FTIR and SFG measurements. Experimental results show that the physically most likely orientations revealed by molecular dynamics do not agree with the assumption that all peptides adopt a single δ distribution in tilt and twist angles. It is likely that all tachyplesin I molecules adopt a small range of tilt angles that place the strand axis roughly parallel to the surface, with different twist angles

possible. The exact angles could not be predicted from these simulations due to limitations in accessible simulation timescales.

4.6 Acknowledgements:

SFG and ATR-FTIR spectral data were collected by Khoi Nguyen and John King. Thanks to Dr. Pei Yang for assistance in calculating the Fresnel factors. Computational resources for simulations were provided by the UC Irvine School of Physical Sciences computing cluster, Greenplanet.

4.7 References

(1) Pauling, L.; Corey, R., B. *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 251-256.
(2) Pauling, L.; Corey, R., B. *Proceedings of the National Academy of Sciences Proc. Natl. Acad. Sci. USA* **1951**, *37*, 729-740.
(3) Pauling, L.; Corey, R., B. *Proceedings of the National Academy of Sciences Proc. Natl. Acad. Sci. USA* **1953**, *39*, 247-252.
(4) Chen, X.; Wang, J.; Sniadecki, J., J.; Even, M., A.; Chen, Z. *Langmuir* **2005**, *21*, 2662-2664.
(5) Nguyen, K. T.; King, J. T.; Chen, Z. *J Phys Chem B* **2010**, *114*, 8291-8300.
(6) Wang, J.; Chen, X.; Clarke, M., L.; Chen, Z. *Proc. Natl. Acad. Sci.* **2005**, *102*, 4978-4983.
(7) Fu, L.; Ma, G.; Yan, E. C. *J. Am. Chem. Soc.* **2010**, *132*, 5405-5412.
(8) Baugh, L.; Weidner, T.; Baio, J. E.; Nguyen, P. C.; Gamble, L. J.; Stayton, P. S.; Castner, D. G. *Langmuir* **2010**,
(9) Weidner, T.; Apte, J. S.; Gamble, L. J.; Castner, D. G. *Langmuir* **2010**, *26*, 3433-3440.
(10) Oishi, O.; Yamashita, S.; Nishimoto, E.; Lee, S.; Sugihara, G.; Ohno, M. *Biochemistry* **1997**, *36*, 4352-4359.
(11) Laederach, A.; Andreotti, A. H.; Fulton, D. B. *Biochemistry* **2002**, *41*, 12359-12368.
(12) Matsuzaki, K.; Shioyama, T.; Okamura, E.; Umemura, J.; Takenaka, T.; Takaishi, Y.; Fujita, T.; Miyajima, K. *Biochim. Biophys. Acta* **1991**, *1070*, 419-428.
(13) Rodionova, N. A.; Tatulian, S. A.; Surrey, T.; Jaehnig, F.; Tamm, L. K. *Biochemistry* **1995**, *34*, 1921-1929.
(14) Marsh, D. *Biophys. J.* **1997**, *72*, 2710-2718.
(15) Páli, T.; Marsh, D. *Biophys. J.* **2001**, *80*, 2789-2797.
(16) Matsuzaki, K.; Nakayama, M.; Fukui, M.; Otaka, A.; Funakoshi, S.; Fujii, N.; Bessho, K.; Miyajima, K. *Biochemistry* **1993**, *32*, 11704-11710.
(17) Wei, Y.; Latour, R. A. *Langmuir* **2009**, *25*, 5637-5646.
(18) Vallone, B.; Miele, A., E.; Vecchini, P.; Chiancone, E.; Brunori, M. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6103-6107.

(19) Chipot, C.; Maigret, B.; Pohorille, A. *Proteins: Structure, Function, and Bioinformatics* **1999**, *36*, 383-399.
(20) Humphrey, W.; Dalke, A.; Schulten, K. *J. Molec. Graphics* **1996**, *14*, 33-38.
(21) Frishman, D.; Argos, P. *Proteins* **1995**, *23*, 566-579.
(22) Lee, B.; Richards, F., M. *J. Mol. Biol.* **1971**, *55*, 379-400.
(23) Yang, A. S.; Honig, B. *J. Mol. Biol.* **1995**, *252*, 366-376.
(24) Tamm, L., K.; Tatulian, S., A. *Quart. Rev. Biophys.* **1997**, *30*, 365-429.
(25) Lambert, A., G.; Davies, P., B.; Neivandt, D., J. *App. Spec. Rev.* **2005**, *40*, 103.
(26) Zhuang, X.; Miranda, P., B.; Kim, D.; Shen, Y., R. *Phys. Rev. B* **1999**, *59*, 12632-12640.
(27) Hecht, E. *Optics (4th Edition)*; Addison Wesley: Boston, MA, 2001
(28) Zbinden, R. *Infrared Spectroscopy of High Polymers*; Academic Press: New York, 1964
(29) Hu, X.; Shin, K.; Rafailovich, M.; Sokolov, J.; Stein, R.; Chan, Y.; Wlwu, K. W.; Wu, W. L.; Kolb, R. *High Performance Polymers* **2000**, *12*, 621-629.
(30) Malitson, I., H. *Applied Optics Appl.Opt.* **1963**, *2*, 1103-1107.
(31) *CRC Handbook of Chemistry and Physics, 91st Edition*; CRC Press: 2010;
(32) Chen, X.; Wang, J.; Boughton, A., P.; Kristalyn, C., B.; Chen, Z. *J. Am. Chem. Soc.* **2007**, *129*, 1420-1427.

Chapter 5: Conclusions and Future Directions

It has long been known that the structure of proteins determines function, and the advent of ever-more-capable surface-sensitive techniques has only served to fuel interest in the role that interfaces play in biological phenomena. New vibrational spectroscopies provide a means to explore these phenomena *in situ* in biologically relevant environments, but the major limitation has been the complexity of interpreting the data.

In Chapter 2, I describe the use of molecular dynamics to explore unexpectedly complex behavior in a seemingly very simple model system. Both simulations and methods of interpreting SFG experimental results have focused on single peptides as model systems due to the difficulty of simulating or measuring more complex behavior. As additional complementary techniques become available, this assumption can and should be revisited. For these studies, a simple pre-determined dimer structure was used to test the hypothesis that aggregates could exist at interfaces. There has recently been rapid progress in methods for predicting oligomer structures, and in the near future it should be possible to design simulations around candidate oligomers.

Continuing advances in computing power also provide the potential for longer timescale studies, though here the simulation model system I present is not without its limitations. In particular, modeling the system as a droplet in a vacuum means that there is more than one hydrophobic/hydrophilic boundary present, and in long simulations, the peptide has a natural tendency to diffuse randomly so that it will eventually migrate to the very edge of the droplet at the water/vacuum interface. The simulations reported were

carefully monitored to exclude this behavior, and the surface area was chosen to be sufficiently large so that this behavior did not tend to occur on the timescales of interest. For longer simulations, or simulations with ions present, periodic boundary conditions may offer advantages.

In Chapter 3, my focus turns from small peptides with complex orientation distributions to a larger protein-protein complex. Here, I demonstrated that by extending published data methodologies for single helices, it is possible to study the formation of a multi-subunit lipid-bound protein complex, and to examine whether the subunits change orientation upon binding. Determining the exact orientation of all subunits in detail will depend on obtaining more measurements from other techniques, and efforts are underway to extend this orientation analysis program to include Four Wave Mixing (FWM) and polarized ATR-FTIR results. Within the realm of SFG measurements, the program that I have developed should also facilitate the use of absolute intensity measurements (rather than the more commonly used ppp/ssp ratio). Absolute intensity measurements depend on knowing the absolute hyperpolarizability tensor components rather than the relative values derived from experimental measurements, but such a determination would be both feasible and fruitful. In principle, SFG can also be used to examine the orientation of βγ in the Gαβγ heterotrimer relative to its position in the GRK2-Gβγ complex. This would provide an overlapping set of measurements that would help to characterize the orientation of all subunits, and such work is in progress.

The work in Chapters 2 and 3 focuses on helical secondary structures, but many large membrane proteins also contain sizable regions organized into extended β-sheets that would contribute to the overall signal. However, β-sheets are more complicated to

study for a variety of reasons, including the deformability of the secondary structure away from an extended conformation, which precludes the use of common adsorption free energy estimate procedures that rely on assuming a rigid structure. If these factors are understood, than more accurate measurements of β-sheet orientation can be obtained. For a protein with multiple types of secondary structure present, the benefits would be twofold. First, if the orientations of helical and β-sheet regions are both measured, then more measurements will be available to characterize complicated molecules. Secondly, SFG spectra cannot be used to directly assess changes in secondary structure due to the fact that polarized signals depend on both orientation *and* the number of groups at the surface. By measuring the orientation of both structures, however, then the dependence of signals on the number of groups could be extracted: this would in principle enable the use of SFG to characterize secondary structure change of some proteins at the interface *in situ*, addressing of the largest assumptions required for interpreting the orientation of complex and dynamic biomolecules. The studies reported in the preceding chapters represent a first step towards this goal, and will help *in situ* spectroscopic measurements to reach their full potential.