Probing the spontaneous membrane insertion of a tail-anchored membrane protein by sum frequency generation spectroscopy

Khoi Nguyen[1**], Ronald Soong[1,2**], Sang-Choul Im[3], Lucy Waskell[3], Ayyalusamy Ramamoorthy[1,2*], Zhan Chen[1,2*]

[1]Department of Chemistry, [2]Department of Biophysics, [3]Department of Anesthesiology,

University of Michigan, Ann Arbor, MI 48109-1055.

List of abbreviations used in the main text is given below

Cyt-b_5: Cytochrome b_5

dDMPC: deuterated dimyristoylphosphatidylcholine

dDLPC: deuterated dilauroylphosphatidylcholine

dDPPC: deuterated dipalmitoylphosphatidylcholine

ER: endoplasmic reticulum

m-Cyt-b_5: Mutant Cytochrome b_5

ppp: p-polarized SFG signal, p-polarized input visible, and p-polarized input IR beam

SFG: Sum Frequency Generation

ssp: s-polarized SFG signal, s-polarized input visible, and p-polarized input IR beam

T_m: gel to liquid crystalline phase transition temperature

DKDVKYYTLEEIKKHNHSKSTWLIKHHKVYDLTKFLEEHPGGEEVLREQAGGDAT

ENFEDVGHSTDARELSKTFIIGELHPDDRSKLSKPMETLITTVDSNSSWWTNWVIPAI

SALIVALMY

Figure S1a. The amino acid sequence of a full-length wild-type rabbit cytochrome b_5. High-resolution structures of the soluble domain of the protein have been reported from solution NMR and X-ray crystallography studies, whereas the structure of the full-length protein is unknown as it has not been amenable for studies using high-resolution methods. However, the secondary structure of the transmembrane anchor has been determined by NMR.[1-2] Secondary structures such as α-helix (blue) and β-sheet (green) are indicated based on previous studies on the soluble domain of the protein. The different amino acid deletions are shown as follows: the 8 amino acid deletion mutant is shown in red, the 2 amino acids deletion mutant is highlighted in green and the 6 amino acid deletion mutant is highlighted in yellow.[1-2]

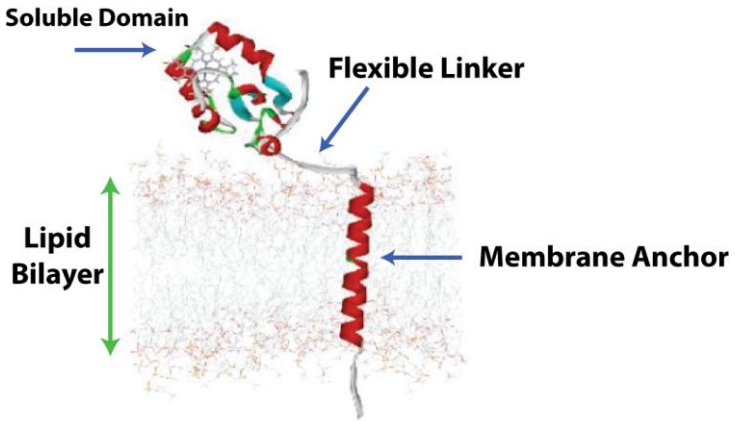


Figure S1b. A proposed model depicting the full-length structure of rabbit cytochrome b_5 derived from solid-state NMR studies. Various domains of the protein are labeled and their amino acid sequence is shown in Fig.S1a.

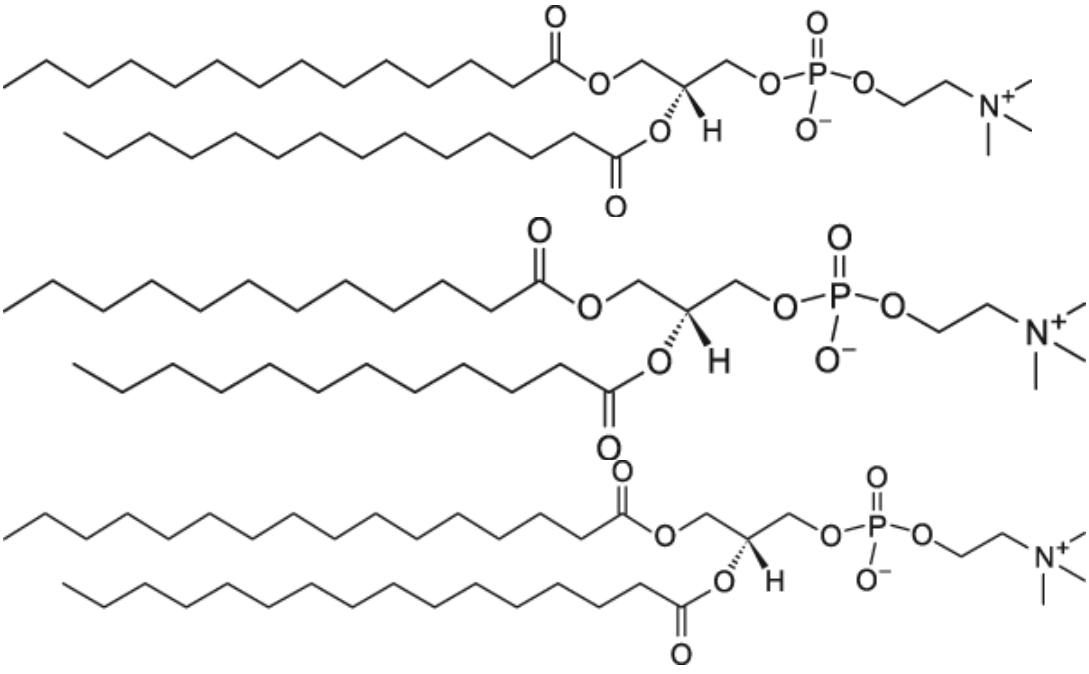


Figure S1c. Molecular strctures of DMPC, DLPC, and DPPC.

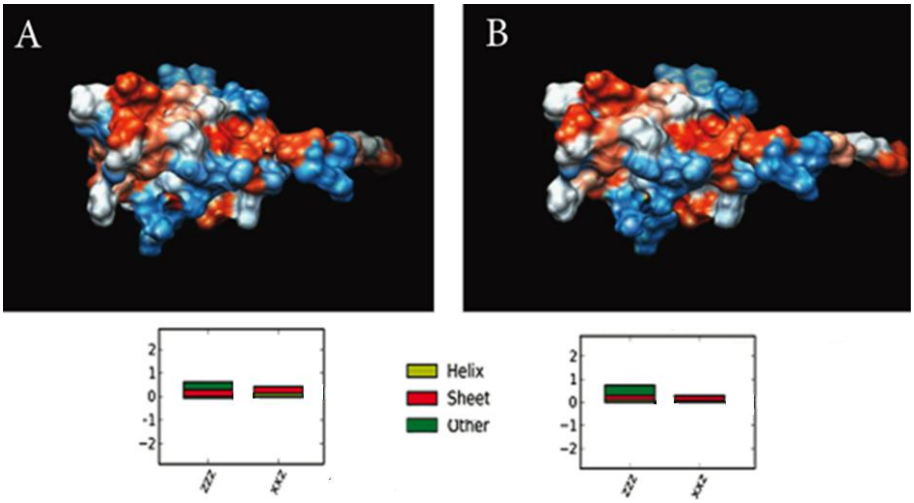


Figure S2. Macroscopic SFG $\chi^{(2)}$ quantities (in arbitrary units) of Cyt b5 water soluble domain calculated by NLOpredict.[3] The blue and red parts of the protein are the hydrophilic and hydrophobic regions of the protein, respectively. (A) When the anchoring tail is likely to adopt a transmembrane orientation in a lipid bilayer. (B) When the anchoring tail adopts a horizontal orientation on top of a lipid bilayer. For both cases, signal from the α-helical components is minimal.

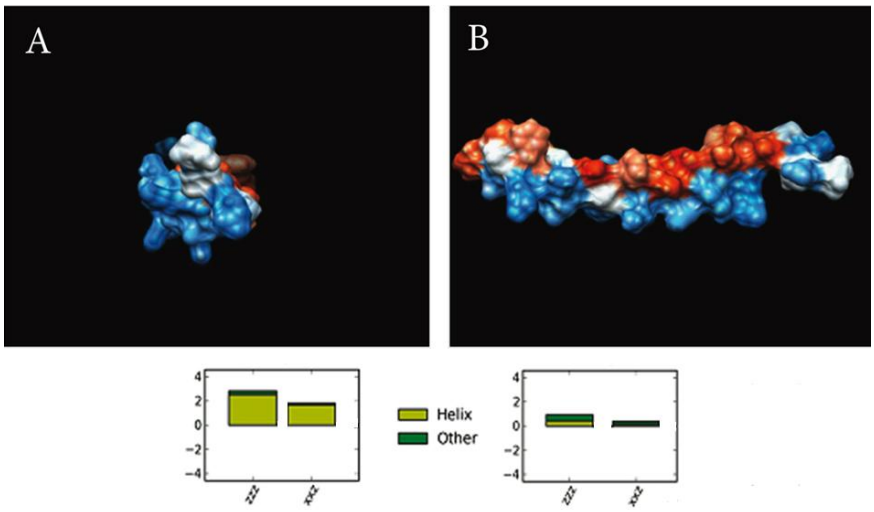


Figure S3. Macroscopic SFG $\chi^{(2)}$ quantities (in arbitrary units) of *Cyt b5*'s anchoring tail calculated by NLOpredict.[3] The blue and red parts of the protein are the hydrophilic and hydrophobic regions of the protein, respectively. (A) When the anchoring tail adopts a transmembrane orientation in a lipid bilayer. (B) When the anchoring tail adopts a horizontal orientation on top of a lipid bilayer.

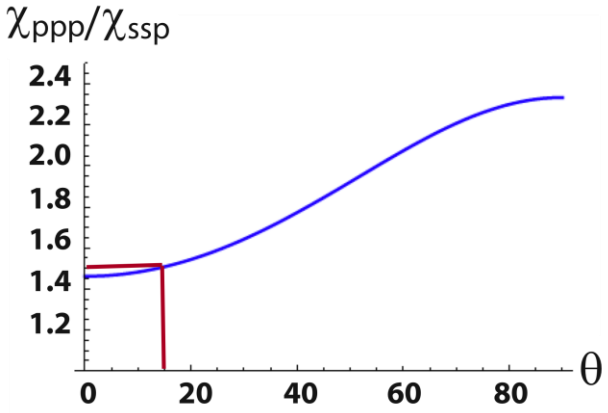
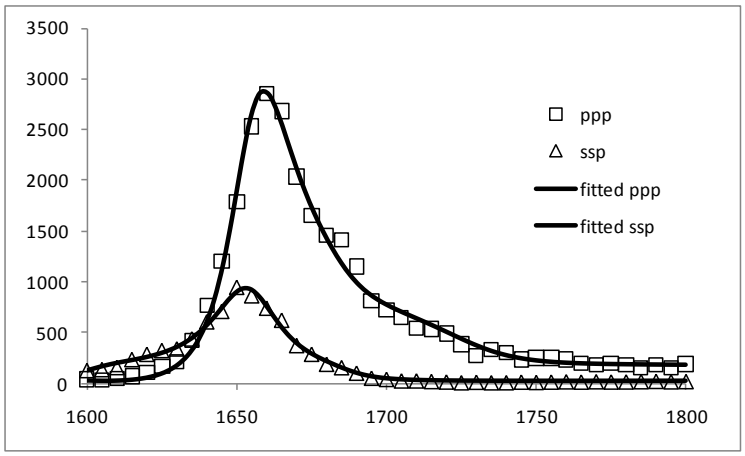
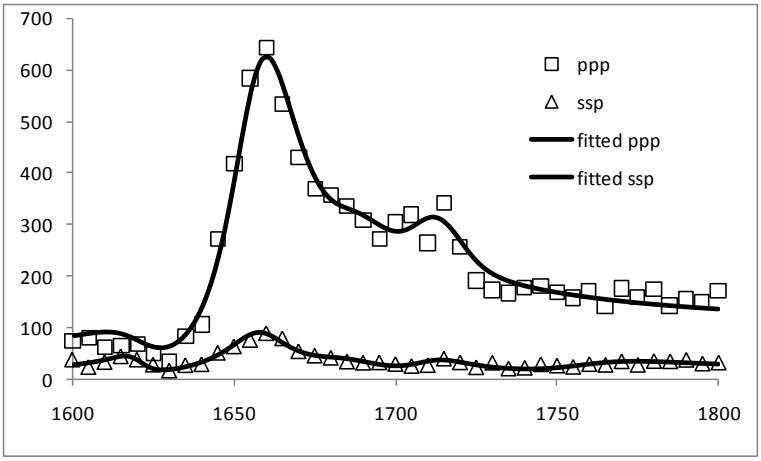


Figure S4. A 2D plot showing the χ_{ppp}/χ_{ssp} ratio of an α-helix as a function of the helical tilt angle.[4] The SFG ppp/ssp amide I signal strength ratio of an α-helix is related to the hyperpolarizability component ratio of the α-helix through the average orientation angle.[4] The SFG hyperpolarizability is a product of Raman polarizabilty and IR transition dipole moment. We deduced the Raman poarlizability and IR transition dipole moment using the bond additivity model.[4] The results were validated using Raman and IR experimental results reported in the literature. Therefore the SFG hyperpolarizability component ratio can be calculated. According to the measured SFG ppp/ssp signal strength ratio, the orientation of alpha-helices (assuming a delta orientation distribution) can be determined. More details can be found in ref (4).



	fitting papas for ppp				fitting papas for ssp					
		error				error				
yo	10.58175	0.44149		yo	-5.64104	0.26012				
f1	1655			f1	1655			alpha helix		
a1	754.2308	8.61305		a1	513.4858	4.91558				
w1	13.7			w1	14					
f2	1735			f2	1725			beta-sheet, side chains		
a2	-176.675	29.66876		a2	0					
w2	30			w2	30					
f3	1688			f3	1685			turns, beta-sheet		
a3	-18.8799	12.77632		a3	-297.664	12.83603				
w3	15			w3	20					
f4	1617			f4	1617			aggregated strands		
a4	-182.186	41.6274		a4	-113.232	19.01259				
w4	22			w4	25					

Figure S5a. Fitting results for the spectra shown in Figure 1A in the text. The parameters yo, fi, ai, and wi (i=1,2,3,4) are nonresonant background, peak center, signal strength, and peak width, respectively.



fittig paras ppp				fittig paras ssp						
		error				error				
yo	9.78652	0.1649		yo	3.72206	0.13577				
f1	1656			f1	1656			alpha helix		
a1	320.1908	3.88127		a1	-131.51	2.01359				
w1	13			w1	12					
f2	1720			f2	1730			beta-sheet, side chains		
a2	-31.7926	8.7425		a2	30.3564	2.54175				
w2	13			w2	13					
f3	1685			f3	1686			turns, beta-sheet		
a3	87.7461	10.99854		a3	94.21414	7.21773				
w3	20			w3	15					
f4	1617			f4	1617			aggregated strands		
a4	-190.869	8.9344		a4	-24.8212	2.64766				
w4	18			w4	7					

Figure S5b. Fitting results for the spectra shown in Figure 2A in the text. The parameters yo, fi, ai, and wi (i=1,2,3,4) are nonresonant background, peak center, signal strength, and peak width, respectively.

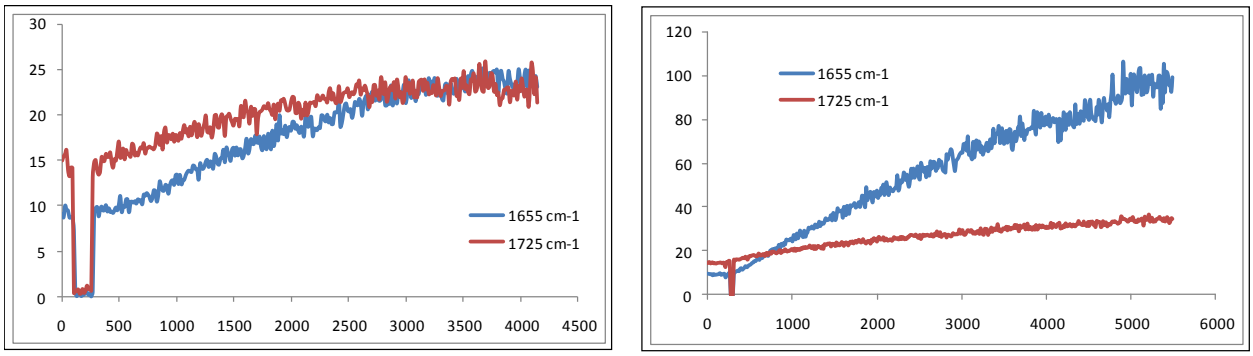


Figure S6. The time-dependent SFG amide I signals collected as a function of time after mutant Cyt b5 (left) and wild type b5 (right) solutions made contacts with DMPC bilayer. The 1725 cm^{-1} signal is from the lipid bilayer, while the 1655 cm^{-1} is due to the helical amide I mode. It can be seen that the amide I signals from the mutant Cyt b5 and the wild-type Cyt b5 reached equilibrium about 4000 and 5000 seconds after the initial contact. The zero signal around 250 seconds was collected when the IR beam was blocked, showing that the signal is SFG signal, not the scattering signal from the visible input light. The lipid bilayer SFG signals changed slightly, perhaps due to the asymmetry induced by the interactions between the lipid bilayer and Cyt b5 molecules. x-axis: Time in seconds. y-axis: SFG signal in arbitrary units.

Supplementary Methods

Materials

Cytochrome b_5 was expressed in *E.coli* C41 cells using the pLW01 plasmid with a purity of >90% as described in details elsewhere.[5] Deuterated dimyristoylphosphatidylcholine (dDMPC), deuterated dilauroylphosphatidylcholine (dDLPC) and deuterated dipalmitoylphosphatidylcholine *(*dDPPC*)* were purchased from Avanti Polar Lipids (Alabaster, AL) and used without any further purification. HEPES buffer (10 mM, pH= 7.3) was used in making the protein solution and stabilizing the pH of the reaction media. The CaF_2 prisms (purchased from Altos, Trabuco Canyon, CA) were cleaned in toluene, soap, methanol and then rinsed thoroughly with deionized water before being treated in a glow discharge plasma chamber for 4 minutes immediately before the deposition of lipid monolayers. The Langmuir-Blodgett and Langmuir-Schaefer (LB/LS) method was used to deposit the proximal and the distal leaflets of a single lipid bilayer onto the prisms.[6]

Sum Frequency Generation spectrometer setup

The SFG spectrometer used for this study is described elsewhere.[7] An experimental geometry similar to the total reflection geometry was used (input beam angles - IR: 62.9^o, visible: 67.9^o).[8] The angles of the visible and signal inside the substrate were close to the critical angle for total internal reflection at the interface of CaF_2 (with a layer of lipid bilayer) and water (or a dilute protein/peptide solution), which allowed for the collection of SFG vibrational spectra by the near total reflection geometry.

For the protein-bilayer interaction experiments, a specific volume of the protein aqueous stock solution (3 mM) was injected into a small reservoir of 2 mL to achieve the

desired concentration (0.3 mM). A magnetic microstirrer was used to ensure a homogeneous concentration distribution of cytochrome b_5 molecules and also to regulate the heat within the subphase below the bilayer. All experiments were carried out at room temperature (~25^{o}C), at which DMPC and DLPC bilayers are in the fluid phase. The temperature-controlled experiment was done using a hot plate and a thermal couple sensing the temperature of the reaction medium. The cytochrome b_5 solutions were left to interact with lipid bilayers for approximately one hour before SFG spectra in ssp and ppp polarization combinations were recorded. There was no SFG signal detected in the amide I region for α-helix from the lipid bilayer before the addition of cytochrome b_5.

Theory of SFG polarization and protein secondary structure

SFG is a 2nd order nonlinear optical spectroscopic technique that has proven powerful in dissecting interactions at both surface and interface. (For full review on SFG, please refer to references [9-15].) Under the electric-dipole approximation, the 2nd order nonlinear susceptibility vanishes for materials with inversion symmetry, leading to no SFG signal. Most bulk materials have an inversion center, resulting in SFG being surface or interface sensitive as the inversion symmetry must be broken at the surface or interface. For IR-visible SFG experiments conducted in this research, SFG signal is generated as the frequency sum of the two input beams at different frequencies: A frequency tunable infrared (IR) beam and a frequency fixed visible beam. The generated SFG signal is proportional to the square of the effective 2nd order nonlinear susceptibility tensor as described in Eqn (S1).

$$I_{SFG} \propto |\chi_{eff}^{(2)}|^2 \qquad [S1]$$

Where I_{SFG} is the intensity of the SFG signal beam and $\chi_{eff}^{(2)}$ is the effective 2nd order nonlinear susceptibility tensor. The effective 2nd order nonlinear susceptibility tensor components detected in the experiment can then be related to the 2nd order nonlinear susceptibility tensor components defined in the lab coordinate system.[7,16]

$$\begin{aligned}
\chi_{eff,ssp}^{(2)} &= L_{yy}(\omega)L_{yy}(\omega_1)L_{zz}(\omega_2)\sin\beta_2\chi_{yyz} \\
\chi_{eff,ppp}^{(2)} &= -L_{xx}(\omega)L_{xx}(\omega_1)L_{zz}(\omega_2)\cos\beta\cos\beta_1\sin\beta_2\chi_{xxz} \\
&\quad - L_{xx}(\omega)L_{zz}(\omega_1)L_{xx}(\omega_2)\cos\beta\sin\beta_1\cos\beta_2\chi_{xzx} \quad \text{[S2]} \\
&\quad + L_{zz}(\omega)L_{xx}(\omega_1)L_{xx}(\omega_2)\sin\beta\cos\beta_1\cos\beta_2\chi_{zxx} \\
&\quad + L_{zz}(\omega)L_{zz}(\omega_1)L_{zz}(\omega_2)\sin\beta\sin\beta_1\sin\beta_2\chi_{zzz}
\end{aligned}$$

where $\chi_{eff,ssp}^{(2)}$ and $\chi_{eff,ppp}^{(2)}$ are the effective 2nd order nonlinear susceptibility tensor components probed using the ssp (s polarized sum, s polarized visible, p polarized IR) and ppp polarization combinations of the input and output laser beams, respectively. χ_{yyz}, χ_{xxz}, χ_{xzx}, χ_{zxx}, and χ_{zzz} are different components of $\chi^{(2)}$ with the lab coordinates chosen such that z is along the interface normal and x in the incident plane. β, β_1, and β_2 are angles between the surface normal and the sum frequency beam, input visible beam, and the input IR beam, respectively. L_{ii} (i = x, y or z) are the Fresnel coefficients.

For the case of near total reflection geometry, which is used in the current SFG studies, we have:[17] $\chi_{eff,ppp}^{(2)} = L_{zz}(\omega)L_{zz}(\omega_1)L_{zz}(\omega_2)\sin\beta\sin\beta_1\sin\beta_2\chi_{zzz}$ [S3]

We can obtain the following equations according to equations [S2] and [S3]:

$$\chi_{eff,ssp}^{(2)} = F_{ssp}\chi_{yyz} \qquad \text{[S4]}$$

$$\chi_{eff,ppp}^{(2)} = F_{ppp}\chi_{zzz} \qquad \text{[S5]}$$

Here we call F_{ssp} and F_{ppp} the Fresnel coefficients for ssp and ppp SFG signals. For a particular vibrational mode *q*, the values for $\chi_{eff}^{(2)}$ components can be obtained from A_q/Γ_q by fitting the SFG spectra according to Eqn (S6) where $\chi_{nr}^{(2)}$ is the nonresonant background and A_q, w_q, and Γ_q are the strength, resonant frequency and damping coefficient (width) for the vibrational mode q. For the amide I mode of α-helical structures, we have shown in our previous research that the ratio between $\chi_{yyz}^{(2)}$ and $\chi_{zzz}^{(2)}$ can be used to determine the orientation angle θ.[6]

$$\chi_{eff}^{(2)} = \chi_{nr} + \sum \frac{A_q}{\omega_{IR} - \omega_q + i\Gamma_q} \qquad [S6]$$

In this research, SFG spectra were collected by tuning IR wavenumber from 1600 to 1800 cm^{-1}, in steps of 5 cm^{-1}. For data analysis, we assumed that the indices of refraction do not change in the amide I frequency range for the alpha helix, since this range only covers less than 20 cm^{-1}. We used the following values for the refractive indices: for IR beam: n1=1.385, n2=1.300. For the visible beam: n1=1.435, n2=1.334. For the visible and SFG beam, n'= 1.404. The n1 and n2 values are taken from the handbook[18] and the value of n' was calculated from ref. 19. We ignored the refractive index change at different temperatures; we believe such changes should be minor (One example to show that the changes of refractive indices on temperature are minor can be found in ref.20). We believe that the refractive indices which we used are reasonable, because our SFG data match the NMR data very well. Also, our previous study indicated that the results obtained from SFG measurements were correlated to the ATR-FTIR studies well.[6,21] The Fresnel coefficients for the IR, VIS and SFG beams depend on their

polarizations. The method used to calculate such values can be found in our previous publication,[4] with the above refractive indices. We did not calibrate the absolute intensity, thus we did not determine the $\chi^{(2)}$ values. Such calibration is not necessary if one uses the $\chi^{(2)}$ ratio as reported in the article.

Data analysis and further discussion

Here we are using a delta function in orientation analysis. We believe that the lipid bilayer and the α-helix interaction is more "specific" compared to that between proteins and polymer surfaces, and therefore the α-helical transmembrane anchor insert into the bilayer more or less with the same orientation. Previously we have deduced different orientation distribution functions for proteins on polymer surfaces[17] and peptides in lipid bilayers,[6,22] it was shown that the distribution of peptides in lipid bilayers are much narrower. Here our research is not focused on the orientation distribution determination. Instead, we compared the Cyt b5 membrane insertion helix orientation dependence on the linker length, lipid chain length, and temperature. Therefore, we used a delta-distribution in the orientation determination. We believe that this assumption is reasonable, because under this assumption, the deduced orientation angle using SFG is well correlated with the NMR result. Also, under this assumption, the ppp/ssp intensity ratio and relative SFG signal intensity of the wild and mutant Cyt b5 lead to a similar orientation angle.

It is difficult to give a quantitative number on the error for the final orientation angle deduced, but according to the SFG/NMR comparison, the comparison between the results obtained from SFG relative intensity (for wild and mutant Cyt b5) and the ppp/ssp intensity ratio for SFG spectra collected from the mutant Cyt b5 with different

polarization combinations in this article, comparisons between MD simulations[23] and SFG studies[6], and comparisons between our previous SFG and ATR-FTIR studies on various peptides[6,21], it is clear that the error should be less than ±5 degrees.

The objective of this research is not to apply SFG to determine atomistic structures of the protein molecules. Instead, here SFG was used to study orientation of the α-helical component of a protein in a lipid bilayer. It is difficult to determine detailed 3-D structure of proteins in liquid using vibrational spectroscopy alone. However, for orientation determination, vibrational spectroscopic techniques such as transmission FTIR and ATR-FTIR have been quite successful. FTIR and ATR-FTIR have been extensively used to determine orientation of α-helical peptides in lipid bilayers and excellent results have been obtained. It is difficult for FTIR or ATR-FTIR to measure orientation of each helix in a complex protein if there are more than one helix in the protein. Even so, they can measure the average orientation of all the α-helices in a complex protein to investigate the overall orientation of the protein. Different from FTIR or ATR-FTIR, under the electric dipole approximation, SFG can only detect signals of a sample with no inversion symmetry. For Cyt b5, the α-helices in the soluble domain more or less point to opposite directions, thus no SFG amide I signal can be observed from the α-helices in the soluble domain, as we deduced using *NLOpredit*. Therefore here the SFG amide I signal is dominated by the α-helical membrane anchor region. This is an ideal case for SFG studies.

References:

(1) Soong, R. et. al. Proton-evolved-local_field solid state NMR Studies of Cytochrome b_5 embedded in bicelles: revealing both structure and dynamical information. *J. Am. Chem. Soc.* **132**, 5779-5788 (2010)

(2) Durr. U. H. N. et. al. Solid-State NMR Reveals Structural and Dynamical Properties of a Membrane-Anchored Electron-Carrier Protein, Cytochrome b5. *J. Am. Chem. Soc.* **129**, 6670-6671 (2007)

(3) Moad A., et. al. NLOPredict: Visualization and data analysis software for nonlinear optics. *J. Comput. Chem.* **28**, 1996-2002 (2007)

(4) Nguyen, K., Le Clair, S. V., Ye, S., Chen, Z., Orientation Determination of Protein Helical Secondary Structure Using Linear and Nonlinear Vibrational Spectroscopy, *J. Phys. Chem. B*, **113**, 12358-1236 (2009)

(5) Clarke, T.A., Im, S.C., Bidwai A., Waskell L., The role of the length and sequence of the linker domain of cytochrome b5 in stimulating cytochrome P450 2B4 catalysis, *J. Bio. Chem.* **279**, 36809-36818 (2004)

(6) Nguyen, K., Le Clair, S., Ye, S., Chen, Z. Molecular Interaction between Magainin 2 and Model Membranes in Situ. *J. Phys. Chem. B*, **113**, 12358-12363 (2009)

(7) Wang, J., Chen, C., Buck, S. M. Zhan, C. Molecular Chemical Structure on Poly(methyl methacrylate) (PMMA) Surface Studied by Sum Frequency Generation (SFG) Vibrational Spectroscopy. *J. Phys. Chem. B*, **105**, 12118-12125, (2001)

(8) Wang, J. et. al. Detection of Amide Signal of Interfacial Proteins in situ Using SFG. *J. Am. Chem. Soc.*, **125**, 9914-9915 (2003)

(9) Shen, Y.R., Ostroverkhov, V. Sum-Frequency Vibrational Spectroscopy on Water Interfaces: Polar Orientation of Water Molecules at Interfaces. *Chem. Rev*. **106**, 1140-1154. (2006)

(10) Eisenthal, K. B. Liquid Interfaces Probed by Second-Harmonic and Sum-Frequency Spectroscopy. *Chem. Rev.* **96**, 1343-1360 (1996)

(11) Bain, C. D. Sum-frequency vibrational spectroscopy of the solid/liquid interface. *J. Chem. Soc. Faraday. Trans.* **91**,1281–1296 (1995)

(12) Richmond, G.L. Structure and bonding of molecules at aqueous surfaces. *Annu. Rev. Phys. Chem. 52*, 357–389 (2001)

(13) Shultz, M.J.; Schnitzer, C.; Simonelli, D.; Baldelli, S. Sum frequency generation spectroscopy of the aqueous interface: ionic and soluble molecular solutions. *Inter. Rev. Phys. Chem.* **19**, 123-153 (2000)

(14) Williams, C. T.; Beattie, D. A. Probing buried interfaces with non-linear optical spectroscopy. *Surf. Sci.* **500**, 545–576 (2002)

(15) Ye, S., Nguyen, K. T., Le Clair, S. V. Zhan, C. In Situ molecular level studies on membrane related peptides and proteins in real time using sum frequency generation vibrational spectroscopy. *J. Struct. Biol.* **168**, 61-77 (2009)

(16) Zhuang, X., Miranda, P. B., Kim, D., Shen, Y. R. Mapping molecular orientation and conformation at interfaces by surface nonlinear optics. *Phys. Re*v. *B* **59,** 12632-12640 (1999).

(17) Wang, J., Lee, S. –H., Chen, Z. Quantifying the ordering of adsorbed proteins in situ, *J. Phys. Chem. B*, **112**, 2281-2290 (2008).

(18) *Handbook of Chemistry and Physics*, 73rd Edition, Lide, D.R. Ed.; CRC Press: Boca Raton 1992

(19) Liu, J., Conboy, J. C. Structure of a gel phase lipid bilayer prepared by the Langmuir−Blodgett/Langmuir-Schaefer method characterized by sum-frequency vibrational spectroscopy. *Langmuir*, **21**, 9091–9097 (2005)

(20) Daimon, M., Masumura, A. Measurement of the refractive index of distilled water from the near-infrared region to the ultraviolet region. *Appl. Opt.*, **46**, 3811-3820 (2007)

(21) Ye, S., Nguyen, K., Boughton, A., Mello, C., Chen, Z. Orientation difference of chemically immobilized and physically adsorbed biological molecules on polymers detected at the solid/liquid interfaces in situ. *Langmuir*, **26**, 6471-6477 (2010)

(22) Chen, X., Wang, J., Boughton, A. P., Kristalyn, C. B., Chen, Z. Multiple Orientation of Melittin inside a Single Lipid Bilayer Determined by Combined Vibrational Spectroscopic Studies. *J. Am. Chem. Soc.*, **129**, 1420-1427 (2007)

(23) Murzyn, K., Pasenkiewicz-Gierula, M. Construction of a toroidal model for the magainin pore. *J. Mol. Model.* **9**, 217–224 (2003)