Determination of Absolute Orientation of Protein α-Helices at Interfaces Using Phase-Resolved Sum Frequency Generation Spectroscopy


DOI
10.1021/acs.jpclett.7b01059

Publication date
2017

Document Version
Final published version

Published in
The Journal of Physical Chemistry Letters

License
Article 25fa Dutch Copyright Act

Citation for published version (APA):
Determination of Absolute Orientation of Protein α-Helices at Interfaces Using Phase-Resolved Sum Frequency Generation Spectroscopy

Lars Schmäus,‡ Steven Roeters,‡ Helmut Lutz,‡ Sander Woutersen,‡ Mischa Bonn,*‡ and Tobias Weidner§†

1Molecular Spectroscopy Department, Max Planck Institute for Polymer Research, 55128 Mainz, Germany
2Van ’t Hoff Institute for Molecular Sciences, University of Amsterdam, 1098 EP Amsterdam, The Netherlands
3Department of Chemistry, Aarhus University, 8000 Aarhus C, Denmark

Supporting Information

ABSTRACT: Understanding the structure of proteins at surfaces is key in fields such as biomaterials research, biosensor design, membrane biophysics, and drug design. A particularly important factor is the orientation of proteins when bound to a particular surface. The orientation of the active site of enzymes or protein sensors and the availability of binding pockets within membrane proteins are important design parameters for engineers developing new sensors, surfaces, and drugs. Recently developed methods to probe protein orientation, including immunoassays and mass spectrometry, either lack structural resolution or require harsh experimental conditions. We here report a new method to track the absolute orientation of interfacial proteins using phase-resolved sum frequency generation spectroscopy in combination with molecular dynamics simulations and theoretical spectral calculations. As a model system we have determined the orientation of a helical lysine-leucine peptide at the air–water interface. The data show that the absolute orientation of the helix can be reliably determined even for orientations almost parallel to the surface.

A molecular understanding of the interaction of proteins or peptides with surfaces is essential for the rational design of biomaterials for use in implants and sensors but also for comprehending the interfacial biology at lipid membranes and mineralized tissue.1–3 An important question for biologists, bioengineers, and pharmacologists is the orientation of proteins when bound to surfaces: Is the enzyme active site within a protein chip pointing toward the solute for effective detection?4 Which parts of a membrane protein are accessible for interaction with therapeutic drugs?5 Which sites are in contact with minerals and control the nucleation of hard tissue?5 Despite the importance, the experimental observation of protein orientation is still challenging. The most established methods are based on enzyme-linked immunoassays,6,7 where the accessibility of antibody binding sites is observed. While this method provides important practical information for bioengineers, the detail of structural information obtained is limited. Time of flight secondary ion mass spectrometry (ToF-SIMS) has recently been developed to probe protein orientation using its extremely shallow probing depth to detect asymmetries in the amino acid distribution within ordered protein films to determine the molecular orientation.8 A drawback of ToF-SIMS is that the experiments require ultrahigh vacuum conditions, which makes sample preparation difficult and can alter or destabilize protein structures. Nuclear magnetic resonance9 and electron paramagnetic relaxation10 can determine the orientation of proteins but need a series of isotope- and spin-labeled proteins, which are time-consuming, often difficult to produce, and limit the size of proteins that can be studied.

Over the past decade sum frequency generation (SFG) spectroscopy has been developed into a reliable method to probe proteins of arbitrary size at interfaces in situ and in real time without the need for labels.11–13 In an SFG experiment, infrared (IR) and visible laser beams are overlapped in space and time at an interface to generate sum frequency photons by optical frequency mixing.14 When the IR light is in resonance with a vibrational mode of molecules at the interface, this leads to an enhancement of the SFG process and corresponding features in the resulting spectra. In analogy to IR and Raman spectroscopy, SFG in the amide I region (1600−1700 cm−1) allows the determination of the secondary and tertiary structure of proteins. SFG has been used to determine the structure of proteins on membranes, polymers, self-assembled monolayers, etc.

Received: May 2, 2017
Accepted: June 12, 2017
Published: June 12, 2017

DOI: 10.1021/acs.jpclett.7b01059
and inorganic surfaces. \textsuperscript{11,15–17} Reliable protocols have been developed to determine secondary structures and tilt angles of entire proteins or individual protein sites \textsuperscript{18} with respect to the surface.

However, published methods for determining the orientation of protein backbones are based on homodyne SFG intensity spectra \textsuperscript{19–21} and as such are unable to tell the absolute orientation of the protein, pointing “up” or “down”, relative to the surface. As discussed above, for applications the absolute orientation is often a key factor. The inability to infer absolute orientations from SFG intensity spectra originates from the fact that the intensity is proportional to the square of the SFG response of the protein ensemble. The SFG response is a complex variable, and in particular the sign of the imaginary part, or, equivalently, the phase of the SFG response, directly reflects the absolute orientation. Oppositely oriented moieties are precisely 180° out of phase. The phase of a signal can be determined by interfering the SFG signal with a reference signal of known phase.

Phase-resolved SFG has been used to determine the chirality, but not orientation, of interfacial proteins. \textsuperscript{22–25} It has been used to infer the absolute orientation of small molecules at interfaces, which requires, however, knowledge of the orientation of the molecular hyperpolarizability tensor (the outer product of the transition dipole moment and the transition polarizability) relative to the molecular axis; for methyl, carbonyl, and water molecules, \textsuperscript{26–29} this relation is well-known (see the Supporting Information for the orientation of the amide-I transition dipole moment and transition polarizability). However, because of the pluriform nature of proteins, the orientation of the molecular hyperpolarizability tensor of a given normal mode relative to the orientation of the protein is different for each protein, as it is determined by the positions and orientations of the local modes and the couplings between them. \textsuperscript{19} In other words, for protein systems, even knowing the phase of the signal does not directly provide absolute orientation information. This requires, in addition, knowledge of the orientation of the hyperpolarizability tensor in the protein molecular frame. Here, we show that for a model protein, combining heterodyne SFG with molecular hyperpolarizability calculations, we can determine the absolute orientation of proteins at an interface.

The experiments were performed with the well-established model peptide LK\textsubscript{14}, which is based only on leucine (L) and lysine (K) amino acids and an acetylated N-terminus (Ac-LKKLKKLKKLKL) and has been designed to adopt an α-helical secondary structure at the air–water interface (see Scheme 1). \textsuperscript{30,31}

\textbf{Scheme 1. Model Leucine (L) and Lysine (K) Peptides LK\textsubscript{14} Designed To Adopt α-Helical Secondary Structures at Interfaces} \textsuperscript{32}

\textsuperscript{31}LK\textsubscript{14} peptides are amphiphilic with hydrophobic leucines and hydrophilic lysines orienting to opposite sides of the helix.

To determine the molecular hyperpolarizability tensors of the amide-I normal modes of LK\textsubscript{14} at the air–water interface, we have first performed molecular dynamics simulations (MD) of the assembly of the peptides at the air–water interface. A snapshot of the MD simulation, taken at 100 ns, is shown in Figure 1a,b. The snapshot illustrates that the peptides assume a mostly α-helical secondary structure at the interface. The hydrophobic leucines point out of the water phase, and the hydrophilic lysines are in contact with the water. This general conformation is in agreement with previously published experimental and theoretical studies of LK peptides at water surfaces and other hydrophobic interfaces. \textsuperscript{32–37}

The simulation indicates the peptides are oriented largely parallel to the water surface, in good agreement with previous experimental and theoretical results. However, a more quantitative analysis of the orientation of the peptide backbone and the average amide-I transition dipole moment (TDM) orientation shown in Figure 1c shows that the peptides are not oriented entirely parallel to the water surface (see Experimental Methods and the Supporting Information for more details about the analysis). The amide-I TDM and the long peptide helix axis are tilted, with angles of 80° and 82°, respectively, with respect to the surface normal and the N-termine pointing towards the water phase.

Phase-resolved SFG spectra for a monolayer of LK\textsubscript{14} at the air–water interface are shown in Figure 2: panels a and b display the imaginary and real SFG spectra, respectively. The imaginary spectrum exhibits a pronounced peak near 1630 cm\textsuperscript{-1}. The real spectrum shows the related zero crossing close to 1630 cm\textsuperscript{-1}.

A convenient way to determine helical peptide orientation from SFG data is to analyze peak ratios obtained from fitted spectra for sps and spp polarization combinations. However, at the air–water interface, the spp combination consists of several hyperpolarizability tensor elements and is underdetermined due to the intrinsic uncertainty of the interfacial refractive index. We therefore chose to directly compare the experimental SFG data with the simulation, by calculating theoretical imaginary and real SFG spectra from the 100 ns MD snapshot shown in Figure 1. This approach also prevents any ambiguities that might arise from peak assignments to specific secondary structure elements. The results are included in Figure 2 (black lines). Clearly, the calculated spectra capture the spectral shape and polarity of both the imaginary and the real experimental spectra well. The sensitivity to the peptide absolute orientation, even at a “flat” angle of 82°, becomes apparent when calculating the related spectra for the opposite molecular orientation as a reference. We have rotated the simulation box with respect to the laboratory frame by 180° by rotating the snapshot about an x-axis within the peptide layer. The spectra of the inverted snapshot (red line) match neither the calculated spectra of the simulation nor the experimental spectra.

To verify the consistency of the measured phase-resolved spectra with the related and well-known homodyne detected spectra we calculated the absolute square of the imaginary and real spectra parts and compared the results with experimental homodyne spectra and homodyne spectra calculated from the simulation. The results are summarized in Figure 2c–e. The heterodyne and homodyne spectra as well as the calculations are in excellent agreement and exhibit a peak near 1645 cm\textsuperscript{-1}, which agrees well with literature values.

In summary, we have determined the absolute orientation of the model peptide LK\textsubscript{14} at the air–water interface using a combination of phase-resolved SFG data, MD simulations, and spectra calculations. We found that the widely studied model peptide LK\textsubscript{14} on average adopts a rather planar orientation at water surfaces and is slightly tilted (82°) with the N-terminus...
oriented towards the water phase (Scheme 2). Orientation measurements of water molecules, hydrocarbons, and protein components have been successfully used for years to answer questions in physical chemistry. Knowledge of the absolute orientation of proteins is important for protein-based biosensors, drug design, and the development of new biocompatible coatings. The procedure presented here allows direct assessment of absolute protein orientation by a comparison of simulated structures with SFG spectroscopic data.

**EXPERIMENTAL METHODS**

Details of the phase-specific SFG setup have been described before. Briefly, a 10 mJ laser pulse with a pulse duration of 40 fs is generated by a Ti:Sa amplified laser (Spitfire Ace, centered at 800 nm). A 2 mJ portion of the amplifier output is branched out to pump an optical paramagnetic amplifier (TOPAS, Light Conversion), which generates 2 μJ broadband pulses centered at 6.1 μm with a bandwidth of roughly 300 cm⁻¹. Another 1 mJ is guided through a Fabry–Perot Etalon (SLS Optics Ltd.) to narrow the broadband 800 nm pulses to a bandwidth of 11 cm⁻¹ with 3 μJ pulse energy. The IR and visible pulses are focused on a gold mirror with a 100 cm (visible) and a 5 cm (IR) focal length lens and overlapped in space and time to generate the local oscillator. Local oscillator, IR, and visible pulses are refocused to the sample using a concave mirror. The local oscillator is either delayed by a 1 mm thick fused silica plate for phase-specific heterodyne detected SFG spectra or

---

**Figure 1.** Snapshot of the MD simulation of LKα14 at the air–water interface at time point 100 ns: (a) side view and (b) top view of the simulation box. The LKα14 peptides predominantly assume α-helical secondary structures. While the leucine side chains are oriented away from the water surface, the lysine side chains are interacting with the water. (c) Orientation analysis of the amide-I bonds and the helix axes of the peptides.

**Figure 2.** Experimental and theoretical SFG spectra for LKα14 at the air–water interface. (a) Experimental (dots) imaginary spectrum along spectra calculated for the MD snapshot (black) and the snapshot rotated by 180° (red). (b) Calculated real SFG spectra. (c) Absolute squared spectrum calculated from the simulation spectra. (d) Absolute squared spectrum constructed from the experimental heterodyne spectra. (e) Homodyne detected SFG spectrum as reference.
The Journal of Physical Chemistry Letters

The spectra are calculated with an amide-I exciton model for the backbone amide groups, based on the formalism described in ref 19. Briefly, the Hamiltonian is constructed with (I) the local mode gas phase frequencies modulated by a hydrogen-bond induced red-shift on the diagonal; (II) for the nearest neighbors, a through-bond coupling model based on parametrized quantum-chemical calculations with the 6-31G+(d) basis set and B3LYP-functional that correlates the dihedral angles between the neighboring the local amide-I modes to the interaction strength; (III) for the non-nearest neighbors, the transition dipole coupling (TDC) model is used, which calculates the through-space coupling between amide groups based on their relative orientation and distance. When the Hamiltonian is diagonalized, the eigenvalues and eigenmodes of the normal modes are obtained, from which the IR, Raman, and SFG responses are calculated. The heterodyned spectra can then be obtained using

\[ \chi_{ijk,KE} = \text{Re} \left( \sum_v \frac{\chi_{ijk,v} \sqrt{\Gamma_{\text{Exc}} + \Gamma_v}}{\omega_0 + \omega - i(\Gamma_{\text{Exc}} + \Gamma_v)} \right) \]

\[ \chi_{ijk,lm} = \text{Im} \left( \sum_v \frac{\chi_{ijk,v} \sqrt{\Gamma_{\text{Exc}} + \Gamma_v}}{\omega_0 + \omega - i(\Gamma_{\text{Exc}} + \Gamma_v)} \right) \]

with \( \chi_{ijk,v} \), \( \Gamma_{\text{Exc}} \), and \( \omega_0 \) the hyperpolarizability, line width, and frequency of normal mode \( \nu_v \), respectively. We assume all \( \Gamma_v \) are 5 cm\(^{-1}\) and set \( \Gamma_{\text{Exc}} \) the width of the visible pulse, to 11 cm\(^{-1}\).

To achieve the best match between calculated and experimental data, we first determined a constant value for the nonresonant phase that was applied to all spectra. The only other fixed parameter used was an absolute intensity scaling factor for each imaginary, real, and absolute spectrum, i.e., the calculations are completely unbiased in terms of resonance position, relative intensity of the different modes, and phase.

Molecular Dynamics Simulations. Simulations were carried out using the GROMACS 4.6 simulation package. Peptide parameters are defined in the AMBER99SBildn force field; parameters for phosphate ions were adopted from ref 5, and the TIP3P water model was used. The initial state was packed using Packmol, with 23 peptides in the vacuum water interface of 8 × 8 × 7 nm of vacuum and 8 × 8 × 6.8 nm of water, containing 10 phosphate anions and sufficient chloride anions to neutralize the simulation box. Therefore the system can be described as 6.8 nm slabs of water separated by 7 nm of vacuum. Topology and coordinate files were generated with Tleap and subsequently converted to GROMACS-type files using Acype. Constraining all bonds with the LINCS algorithm, the simulations were run at a 2 fs time step with periodic boundary conditions in all three dimensions for 100 ns, using the particle mesh Ewald (PME) method for long-range Coulomb interactions. The distance for Lennard-Jones potential cutoff was set to 1 nm, and the temperature was maintained at 300 K using velocity rescaling with a stochastic term. The tilt of the peptide chain at the vacuum–water interface was calculated using a Python implementation of a method developed by Kahn, based on the \( \alpha \) carbon atoms along the peptide chain, disregarding the first and last two amino acids.

### ACKNOWLEDGMENTS

This work was supported by the Max Planck Graduate Center with the Johannes Gutenberg-Universität Mainz (MPGC), L.S. and T.W. thank the Deutsche Forschungsgemeinschaft (WE4478/4-1) for financial support. S.W. and S.J.R. thank the Netherlands Organization for Scientific Research (NWO) for financial support. We thank Dr. Ellen Backus for helpful discussions.

### REFERENCES


The Journal of Physical Chemistry Letters

Letter


