Biophysical Journal
Article

Site-Specific Peptide Probes Detect Buried Water in a Lipid Membrane

Jennifer C. Flanagan[1] and Carlos R. Baiz[1,*]
[1]Department of Chemistry, The University of Texas at Austin, Austin, Texas

ABSTRACT Transmembrane peptides contain polar residues in the interior of the membrane, which may alter the electrostatic environment and favor hydration in the otherwise nonpolar environment of the membrane core. Here, we demonstrate a general, nonperturbative strategy to probe hydration of the peptide backbone at specific depths within the bilayer using a combination of site-specific isotope labels, ultrafast two-dimensional infrared spectroscopy, and spectral modeling based on molecular dynamics simulations. Our results show that the amphiphilic pH-low insertion peptide supports a highly heterogeneous environment, with significant backbone hydration of nonpolar residues neighboring charged residues. For example, a leucine residue located as far as 1 nm into the hydrophobic bulk reports hydrogen-bonded populations as high as ~20%. These findings indicate that the polar nature of these residues may facilitate the transport of water molecules into the hydrophobic core of the membrane.

INTRODUCTION

Lipid membranes are essential to numerous biological processes, and their multiple functions are rooted in their diverse interfacial environments (1–4). On its own, a bilayer provides a nanometer-length gradient from a polar, aqueous environment at the interface to a nonpolar hydrophobic environment within the membrane center. However, the lipid-centric view of membranes may be too simplistic because in cell membranes, transmembrane proteins make up ~50% of the mass, which can significantly alter the bilayer properties (5–7). In general, residue polarity follows the hydrophilicity of its environment. However, the presence of hydrophilic residues in the alkyl region of a membrane is common. For example, in mitochondrial respiratory chain proteins, as much as ~24% of lipid-facing residues are polar or charged (4). Their purpose is not understood, but it has been hypothesized that these residues may be energetically favored within certain membrane potentials (8).

Polar residues may substantially alter the electrostatic potential and drive water toward the otherwise hydrophobic regions of the bilayer, but quantifying these effects remains challenging (5,9). Here, we report a general strategy to study local hydration environments using a combination of ultrafast nonlinear infrared (IR) spectroscopy and nonperturbative, site-specific spectroscopic probes within the backbone of a transmembrane helix.

In this study, we use pH-(low) insertion peptides (pHLIPs) as model membrane peptides. pHLIPs are a family of small amphiphilic peptides that have been extensively studied for their pH-responsive folding and insertion into lipid bilayers. Under acidic conditions, a pHLIP exists as a membrane-embedded α-helix, a property of interest for medicinal applications. For example, recent interest in pHLIP stems from its ability to selectively insert into tumor cells with potential applications for cancer diagnosis and drug delivery (10–16). The pHLIP sequence is composed of hydrophilic residues interspersed within the hydrophobic region of the lipid bilayer (Fig. 1), providing a polar environment and potentially reducing the free energy for water molecules to penetrate into the alkyl region of the membrane. The specific sequence of pHLIP is AAEQNPIYWARYADWLFTTPLLLLDLALLVDADEGT.

IR spectroscopy of amide-I vibrational modes is commonly used as a sensitive nonperturbative probe of peptide structure and hydration (17,18). A variety of strategies involving localized vibrational probes such as isotope labels or covalently modified residues have been valuable in extracting site-specific information (19–23). Here, we employ single-site $^{13}C{=}^{18}O$ isotope labels, which localize the IR signal to individual residue vibrations by shifting the labeled spectrum out of resonance with the main band, allowing its environment to be studied independently (24–31).

Submitted October 4, 2018, and accepted for publication March 6, 2019.
*Correspondence: cbaiz@cm.utexas.edu
Editor: Arne Gericke.
https://doi.org/10.1016/j.bpj.2019.03.002

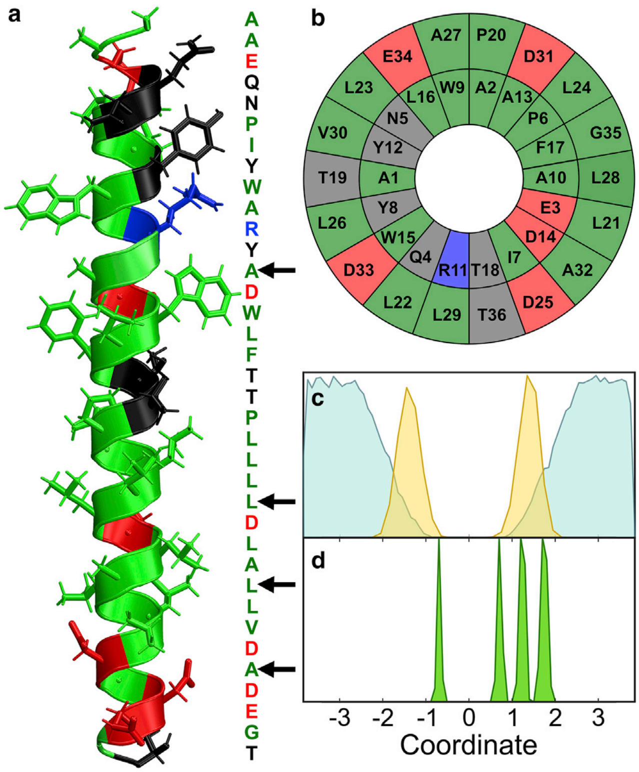

FIGURE 1 (*a*) pHLIP sequence, colored according to residue hydrophobicity. Green indicates hydrophobic, red indicates acidic, blue indicates basic, and black indicates polar residues. (*b*) A helix wheel representation. (*c*) Positions of simulated water molecules (*blue*) and lipid ester oxygen atoms (*gold*). (*d*) Positions of selected peptide residues in the bilayer (from *left* to *right*: A13, L24, L28, and A32). To see this figure in color, go online.

Isotope-edited IR spectroscopy has been fruitful for uncovering site-specific information of peptides in solution as well as in the membrane (26–29,31–38). For instance, Rothschild and co-workers (35) used isotope labeling to probe the local secondary structure and orientation of the membrane-anchoring region of an ion channel. However, the number of studies involving membrane peptides is small because of key challenges in resolving single-site isotope peaks. The low peptide/lipid ratio required to ensure monomers coupled with the relatively low concentration of the isotope-edited sites compared to unlabeled residues makes isotope peaks difficult to reliably interpret. Resolution-enhancing methods such as Fourier self-deconvolution can circumvent some of these issues, but this relies on strong signal/noise ratio to be effective (39). Two-dimensional (2D) IR (2D IR) spectroscopy has significant advantages for these vibrational modes that are difficult to measure. In brief, 2D IR is a pump-probe IR technique in which a pair of pump pulses excites a specific vibrational frequency, followed by a probe which detects vibrations over a broad frequency range. The signal scales with the fourth power of the dipole moment (40–43), compared to the second-order scaling of Fourier transform IR (FTIR) spectroscopy, resulting in narrower line shapes, particularly in the isotope region (44). This mitigates overlap in congested spectra resulting in peaks that are easier to isolate.

Spectral features in 2D IR spectroscopy can be interpreted similarly to FTIR spectroscopy (45). For example, an oscillator in a water-exposed heterogeneous environment will exhibit a broader peak than one in a homogeneous environment because of the more diverse environments and faster fluctuations. Most importantly, hydrogen-bonded oscillators will produce peaks red shifted by $\sim$15 cm^{-1} per hydrogen bond (H-bond) from its characteristic frequency (46).

To interpret 2D IR data from an atomistic point of view, computational models based on electrostatic maps have been developed for generating amide-I IR spectra from molecular dynamics (MD) simulations (47). Agreement between computed and experimental line shapes implies that MD simulations accurately sample the molecular ensembles observed in the experiment. Over the past decade, these methods have proven useful in interpreting 2D IR spectra of biomolecules, as well as providing a quantitative framework for selecting potential isotope-labeled sites within proteins (47–56). Here, we use 2D IR spectra of isotope-labeled pHLIP in combination with these computational tools to probe solvent penetration into a 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) bilayer surrounding a transmembrane peptide.

METHODS

Here, we provide a summary of the methods. Additional descriptions are provided in the Supporting Materials and Methods (Sections S1–S5).

Isotope site selection

Residues were selected through structure-based modeling of IR spectra as described below. First, computed main band amide-I spectra for pHLIP in a DMPC bilayer were used to confirm qualitative agreement with the experimental FTIR spectra of the unlabeled peptide. Similarly, isotope-labeled calculations were performed for all 36 residues to identify potential label sites according to the following criteria: 1) Peaks must be narrow and fully resolved from the main band, and 2) residues must be distributed throughout the backbone to sample different membrane depths. Using these two criteria, we selected four sites: A13 is located near the center of the membrane, L24 is primarily within the hydrophobic region but accessible to water, L28 lies at the lipid-water interface, and A32 is fully solvated according to coordinates obtained from simulations. All four isotope peaks are spectrally resolved from the main amide-I band (Fig. S1).

Isotope label synthesis

Isotope-enriched (^{13}C,^{18}O) alanine and leucine were synthesized in-house following the procedure of Marecek et al. (25) from FMOC-protected ^{13}C alanine and ^{13}C leucine (Sigma-Aldrich, St. Louis, MO). Labeled peptides were custom synthesized by Biosynthesis (Lewisville, TX). Peptides were lyophilized to remove residual trifluoroacetic acid before use.

Sample preparation

Unilamellar 100-nm DMPC vesicles were formed by reconstituting dried DMPC to 30 mM in 100 mM MOPS (3-(N-morpholino)propanesulfonic acid) buffer (pH 8.0) followed by extrusion. An aliquot of 2.4 mM pHLIP solution held at pH 8.0 was added to the vesicle suspension until the final peptide concentration was 0.6 mM for each sample. Samples were adjusted to their final pH of 4.0 (values reported here are uncorrected pH readings in D_2O) and incubated for 30 min at room temperature before the start of each measurement. Samples were held at pH 4 for two reasons: 1) Previous literature (15) and circular dichroism measurements indicate helical character at this value (see Supporting Materials and Methods, Section S2 c), and 2) deprotonated carboxylate side-chain absorption overlaps with the isotope labels. pHLIP contains six carboxylate side chains, including glutamate and aspartate (57).

2D IR experiments

2D IR spectra were measured at room temperature with a custom-built pulse-shaper-based spectrometer (58). Spectra were collected at a fixed pump-probe delay, or waiting time (t_2) of 500 fs and scanned coherence times (t_1) up to 3 ps. The frequency resolution along the probe axis is 2.7 cm^{-1}. Pump and probe polarization were perpendicular, and phase cycling was used to remove contributions due to pump scatter. 11.2 million laser shots were collected to produce each spectrum. Diagonal slices were taken 7 cm^{-1} from the detection axis (shown by the dashed line in Figs. 2 and 3) for comparison with analogous experimental and simulated spectra. All 2D spectra and diagonal slices were normalized to the maximal intensity of the main amide-I peak.

MD

MD simulations were carried out using Groningen Machine for Chemical Simulations. The initial box was constructed by inserting a single pHLIP into a bilayer consisting of 128 DMPC lipids (64 per leaflet) as generated using CHARMM-GUI (59–65). Simulations were carried out using the CHARMM36 force field and TIP3P water (66). The system was equilibrated for 2 ns, followed by a 10-ns NPT production run at 300 K and 1 atm. Proper equilibration and convergence was ensured by computing spectra at different points along the production trajectories.

Radial distribution functions (RDFs) and H-bond populations were calculated using built-in Groningen Machine for Chemical Simulations functions over the course of each simulation trajectory. Water RDFs were computed using the backbone carbonyl oxygen as a reference. The number of H-bonds in each simulation frame was computed using a 0.35-nm donor-acceptor distance cutoff and 30° (D-H-A) angle (67,68).

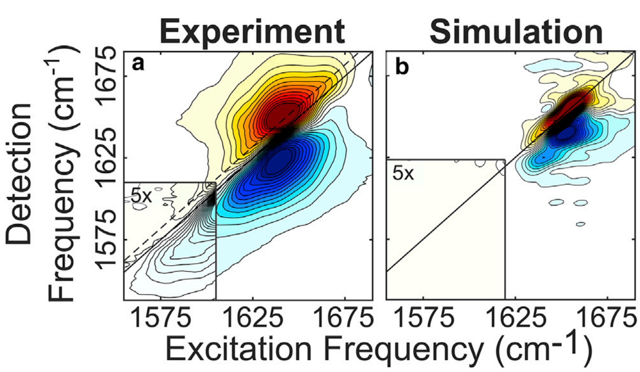

FIGURE 2 (*a*) Experimental (*left*) and (*b*) simulated (*right*) 2D IR spectra of unlabeled pHLIP in a DMPC bilayer (pH 4.0). The isotope-labeled region at frequencies below $[\omega_1, \omega_3] = 1600, 1610\ cm^{-1}$ is scaled by a factor of five to highlight the features in this region. The red (positive) peaks represent the ground-state bleach and stimulated emission, and the blue (negative) peaks are due to excited-state absorption. To see this figure in color, go online.

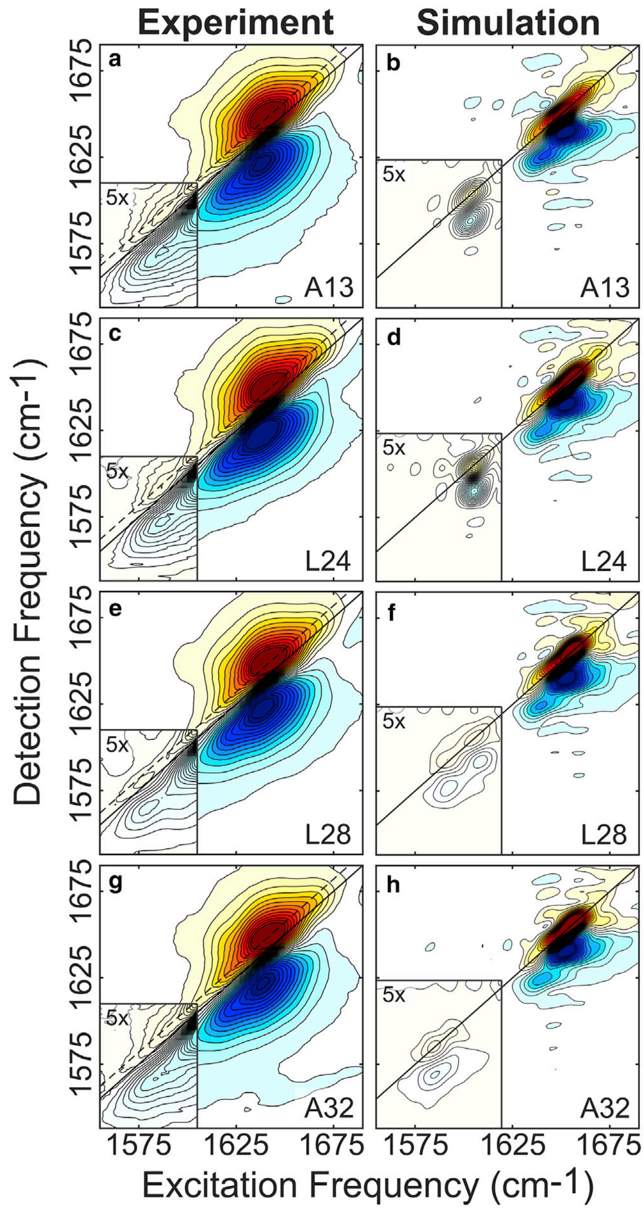

FIGURE 3 Experimental (*left*) and simulated (*right*) amide-I 2D IR spectra of all isotope-labeled pHLIPs in lipid vesicles. (*a* and *b*) A13-labeled peptide; (*c* and *d*) L24-labeled peptide; (*e* and *f*) L28-labeled peptide; (*g* and *h*) A32-labeled peptide. The insets indicate the label region. Because the isotope peaks are small compared to the main band, the insets are plotted at a lower contour level to highlight the features in this region. To see this figure in color, go online.

Spectral calculations

Spectra were calculated from the frequency map parameterized by Tokmakoff and co-workers (47,53,54), and the transition charge coupling map was constructed by Jansen and co-workers using the charges defined by the CHARMM36 force field (69,70). Frequency maps correlate electric fields at the C=O positions computed from MD to the frequency of each amide unit, which become the diagonal elements of the Hamiltonian. Off-diagonal elements, or couplings, are obtained using the nearest-neighbor map of Jansen and co-workers (48–52). Once a Hamiltonian trajectory is obtained, IR absorption and 2D IR spectra are computed through numerical integration of the Schrodinger equation (NISE) (49). Individual isotope spectra were calculated by red shifting the corresponding Hamiltonian site by 60 cm^{-1} to reflect the characteristic $^{13}C{=}^{18}O$ isotope shift and recomputing the spectrum (71,72).

RESULTS

Fig. 2 shows experimental (Fig. 2 *a*) and simulated (Fig. 2 *b*) 2D IR spectra of the main amide-I band of pHLIP in DMPC (pH 4) with the isotope region (*insets*). Both spectra are free of signals in the label region, aside from the sloping tail of the main band observed in the experimental spectra. The diagonal, where the excitation and detection frequency are the same, is qualitatively analogous to the signal obtained from an FTIR spectrum (26).

Fig. 3 shows the experimental and simulated spectra for the four isotope-labeled peptides studied here. The key observation is the appearance of peaks below 1610 cm^{-1} in both the experimental and simulated spectra corresponding to the labeled residues.

Because no distinctive off-diagonal features are observed, the relevant information can be extracted from diagonal slices of the 2D spectra. Further, the main band remains unchanged for the different isotope labels (see Supporting Materials and Methods, Section S4). We focus our analysis on the diagonal slices in the isotope-labeled region (1550–1610 cm^{-1}) as shown in Fig. 4 and interpret the experimental line shapes of each isotope-labeled site separately, comparing them to the simulations before discussing the overall environment of the peptide within the membrane.

Unlabeled pHLIP

The main amide-I band of the experimental spectrum is consistent with that of a primarily α-helical peptide, characterized by a single band centered at $\sim$1650 cm^{-1}. There are no resolved spectral features in the isotope region (*inset*), indicating an absence of deprotonated carboxylate groups. The baseline is more clearly shown in Fig. 4, which contains the diagonal slices in the isotope region (1550–1610 cm^{-1}). There are no resolved spectral features present in the unlabeled slice except a feature near 1600 cm^{-1} corresponding to the low frequency edge of the main band. There are also negative features on the low frequency side of several of these spectra. These arise from the interference between the fundamental (ground-state bleach) and overtone (excited-state absorption) bands of the 2D IR spectra. This feature has been observed in other proteins (26) and illustrates one challenge of interpreting 2D IR spectra without accurate structure-based models: this overlap and resultant negative peaks may alter or mask the underlying positive peaks.

The simulated 2D IR spectrum (Fig. 2 *b*) also appears as a single peak centered at 1650 cm^{-1}. As expected, the simulated spectrum of unlabeled pHLIP shows no features in the isotope region (Fig. 4). Main band line shapes, however, are notably narrower and overall blue-shifted in the simulation. We attribute the origin of this discrepancy to the membrane environment for which the computational model has not yet been parametrized, and thus its heterogeneity is not accurately captured. As described in Supporting Materials and Methods (Section S7), simulated pHLIP spectra in water agree semiquantitatively with experiments (Fig. S10). In the membrane, the low polarity of the lipid tail groups produces a "vacuum-like" electrostatic environment, which, combined with the rigid geometry of the peptide, causes the residue line shapes to be significantly narrower compared to the solution.

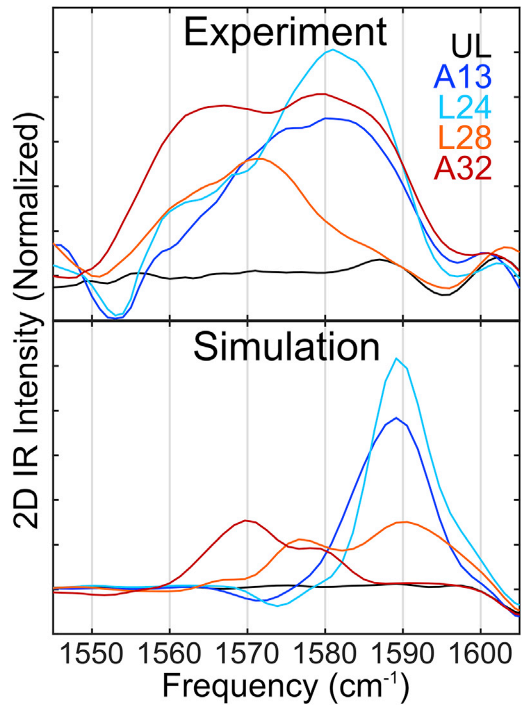

FIGURE 4 Diagonal slices of the experimental (*upper panel*) and simulated (*lower panel*) isotope region of all 2D IR spectra. These diagonals are also indicated as dashed lines in Fig. 3. To see this figure in color, go online.

Experimentally, main band spectra show no differences for the four isotope labels (Fig. 3), confirming that the peptide conformation is identical for the differently labeled peptides. Because isotopes do not perturb the conformation, these results confirm the consistency and reproducibility of the experiments, particularly because each sample was prepared starting from a different peptide. It is worth noting that there are no crosspeaks detected between the isotope peak and the main band; this is because of the weak coupling among residues in the vibrational Hamiltonian. Isotope peaks show qualitatively different features as a result of different environments. Below, we examine each isotope-labeled spectrum to assess the degree of hydration at each site by qualitative interpretation of the experimental diagonals combined with hydrogen bonding and analysis of water density near each site based on the MD simulations.

A13

The experimental A13 spectrum (Fig. 3 *a*) contains a single isotope peak centered at 1582 cm^{-1}. The diagonal slice (Fig. 4) shows that this is a narrow peak, indicative of a uniform solvation environment. This peak is blue-shifted with respect to other isotopes, which supports the interpretation that this C=O is primarily dehydrated, and the sharpness of the peak compared to the other label sites indicates conformational homogeneity in the A13 region of the peptide.

Line shapes are in good qualitative agreement between experimental and simulated spectra; in the simulation, a single, blue-shifted isotope peak is present at 1592 cm^{-1}, as expected for a residue in a nonpolar environment. Residue-specific coordinates extracted from the MD trajectory (Fig. 1) show that A13 is embedded near the center of the hydrophobic lipid tail region. The peak is further narrowed by the primarily hydrophobic neighboring residues within the helix, producing a nonpolar, homogeneous environment around this residue.

Further MD analysis involved counting the number of total H-bonds and donor-specific H-bonds per simulation frame using the backbone C=O at each labeled site as the H-bond acceptor. Donor-specific analysis included separately counting peptide-peptide and peptide-water H-bonds. Through this method, we determined that nearly all simulation frames included a single peptide-peptide H-bond. Any frames with two H-bonds contained one peptide-peptide and one peptide-water H-bond. Total H-bond counts for each site are given in Fig. 5 *a* and are further discussed in the next section.

With this designation, we assign this single A13 peak to a 1-HB population wherein this residue forms helix-stabilizing H-bonds between the carbonyl O and the helical backbone. This type of H-bond is more stable and long-lived than a water H-bond, which results in the peak narrowness observed here.

The intensity and narrow linewidth is consistent with a single H-bond between the carbonyl O and the helical backbone, which is more stable and long-lived than a water H-bond (26,46). Subsequent H-bond and hydration analysis from MD corroborates this interpretation, revealing a primarily singly hydrogen-bonded ensemble (Fig. 5 *a*). Compared to the peptide, H-bonds to water produce broader and less intense peaks (26,46).

Simulated peaks are narrower than the experimental peaks, as observed not only in the main band but also in the isotope signals. Possible contributing factors are discussed in Supporting Materials and Methods (Section S7), and the sharpness of the peak supports the conjecture that the peptide environment in the simulations may be artificially homogeneous because of the membrane electrostatics.

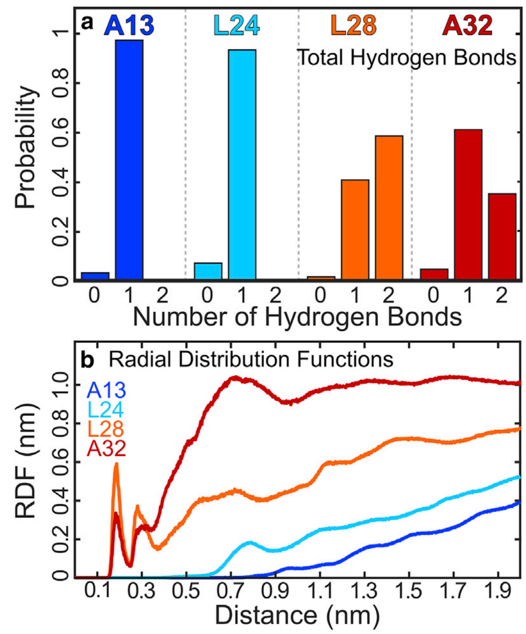

FIGURE 5 (*a*) Total H-bond populations. (*b*) Normalized radial distribution functions (RDFs) calculated for water, using the carbonyl oxygen atom as a reference. To see this figure in color, go online.

L24

The experimental L24 diagonal slice appears qualitatively similar to the A13 peak but with a shoulder to the red of the main isotope peak, which is suggestive of a second distinct population within the sample. We have assigned the blue-shifted peak to an H-bond with the peptide, making up roughly 80% of the total ensemble. The red-shifted peak, comprising the remaining 20% of the population, is assigned to an H-bond with water because the low intensity and approximate 15 cm^{-1} red shift is consistent with a solvent H-bond (46). The shoulder illustrates that water molecules penetrate into the L24 region to transiently H-bond with the backbone.

This shoulder is not observed in the simulated spectrum. There is partial qualitative agreement as the peak appears as a narrow stretch on the blue side of the window, but the shoulder is absent, suggesting the absence of water within the L24 region in the MD simulation. Likewise, H-bond counts reveal a predominantly single H-bonded ensemble with no H-bonds to water (Fig. 5 *a*). Similarly, the RDF indicates a lack of water penetration within the range of 0.35 nm from the carbonyl. This discrepancy may arise from three sources: 1) Water penetration may be slow, and by consequence, our trajectories may not be sufficiently long; 2) the force field underestimates the degree of hydration within the L24 region of the membrane; and 3) although unlikely, interactions between peptides may contribute to hydrogen bonding. Previous literature has not cited multimerization at the experimental pH or concentration (73,74), and control experiments have not detected any aggregation (Fig. S7), but peptide-peptide interactions cannot be entirely ruled out with the methods used here. The second suggestion is more plausible, as we have demonstrated in Supporting Materials and Methods (Section S7) that our system is converged within the 10 ns simulation time; although it is still conceivable that fluctuations slower than 10 ns are not captured, which indicates a need for more powerful, multiscale simulations, to more completely sample the structural ensembles within this model system.

Overall, the experimental line shapes suggest that this site is primarily enclosed within the hydrophobic bulk of the lipid membrane, as indicated by the sharp blue-shifted peak, and the shoulder on the red side corresponds to a small 2-HB population between the amide C=O and transient water.

L28

In contrast with the A13 and L24 sites, both experimental and simulated L28-labeled spectra appear broader and lower in intensity. Diagonal broadening supports the interpretation of multiple different H-bond populations within this region. The experimental diagonal band is composed of two peaks centered at 1573 and 1560 cm^{-1}, respectively, indicative of 1-HB and 2-HB populations. We approximate the single H-bond population to compose 60% of the total ensemble and the double H-bond population to compose 40%. The red shifts and lower intensity of these peaks suggest that L28 is partially water exposed. The low intensity of the red-shifted peak also indicates a more heterogeneous environment at this position.

The simulated spectrum shows similar broadening; however, it appears more blue-shifted than the experimental spectrum with the two peaks centered at 1590 and 1575 cm^{-1} with roughly equivalent populations to those observed in experiment. We attribute the blue shift to the same artificially homogeneous environment described earlier, which is further discussed in Supporting Materials and Methods (Section S7). According to the MD trajectory, the L28 site is characterized by two distinct H-bonding populations in which the peptide backbone is H-bonded to water in $\sim$70% of the snapshots. This agrees semiquantitatively with experiments, and we hypothesize that part of the discrepancy in the estimated populations between experiments and this analysis are due to the nonlinear scaling of 2D IR signals discussed earlier. Whereas the MD analysis presents the 2-HB population to be greater than the 1-HB, experimental 2D IR signals show the opposite trend. These observations can be attributed to the transient nature of a peptide-solvent H-bond compared to a peptide-peptide H-bond, explaining the origin of the larger peptide H-bond peak. The RDF (Fig. 5 *b*) further supports this interpretation because it shows significant water density within the vicinity of the backbone at the L28 position. The first and second solvation shell peaks at 0.18 and 0.28 nm are sharp and narrow, showing that waters molecules are locked into well-defined conformations around the peptide. The presence of confined water is expected because the MD trajectory reveals that L28 resides in an area of intermediate polarity at the interface between the hydrophobic and solvated regions of the membrane where water networks may be severely disrupted compared to the bulk (Fig. 1).

A32

The experimental A32 label peak is weak and broadened, appearing on the red side of the isotope region as two peaks of approximately equal intensity at 1580 and 1565 cm^{-1}. The bimodal character of this feature is reproduced in the simulated spectrum, albeit with an overall frequency shift toward lower frequencies, with the two peaks appearing at 1570 and 1580 cm^{-1}. Additionally, the populations do not match those represented in experiment, with the single H-bond peak representing only 30% of the ensemble. H-bond analysis of the MD trajectory reveals that A32 also contains two distinct environments corresponding to 1-HB and 2-HB populations (Fig. 5 *a*). Similar to L28, the two peaks can be assigned to a single peptide H-bond (1565 cm^{-1}) and a peptide H-bond plus a water H-bond (1580 cm^{-1}), respectively. In a more quantitative analysis, discrepancies arise between computed H-bond populations and the estimated populations based on the A32 2D IR spectra, wherein simulations underestimate the hydrogen-bond populations at this site. This system was equilibrated with structural restraints to maintain a transmembrane, helical conformation, and although these constraints were relaxed during equilibration and removed during the production MD run, it is possible that the simulation may undersample the peptide conformational landscape. In particular, slower fluctuations beyond the nanosecond timescale cannot be captured by this 10-ns simulation. This may result in a higher helical character, which could lead to the recorded over- and underestimation of peptide and solvent H-bonds, respectively.

DISCUSSION

In general, spectral predictions from MD simulations are in qualitative agreement with experimental 2D IR slices, which enables us to interpret the solvation environment of pHLIP in a DMPC bilayer. Consistently, spectral linewidths from simulation are narrower than their analogous experimental spectra in both the main amide-I and the amide-I isotope region. The source of the narrow line shapes can be attributed to two different factors: first, our trajectories are unlikely to sample the full conformational landscape of the peptide because structural rearrangements can occur on the microsecond timescales or beyond; and second, these computational models are predominantly parametrized for solvated peptides and may require further development to capture the heterogeneous environments produced by the bilayer. Supporting Materials and Methods (Section S7) show the effects of the membrane charges on the computed spectra. In brief, setting the lipid tail group charges to zero has minimal effect on the line shapes, which indicates that according to the model, the membrane has no effect on frequencies, essentially acting as a "vacuum" with respect to electrostatics. Thus, the line shapes may be artificially narrowed by this environment. As such, this leads to simulated spectra being more sensitive to single water molecules interacting with the backbone.

We have observed within our experiments that rigid residue-residue interactions that stabilize the helix backbone result in intense, blue-shifted line shapes compared to more solvated carbonyls. The A13 label indicates that effectively no water penetrates into the membrane and that this site is surrounded by both hydrophobic and bulky residues that likely also resist water penetration. L24, which is closely adjacent to the hydrophobic lipid tails and is neighbored by hydrophobic and bulky residues, reported a small degree of water penetration indicated by experimental hydrogen-bonded peaks. The L28 data depict two distinct states: one in which L28 is only H-bonding with other peptide residues and one in which it is water exposed. This site is an ideal candidate to investigate the role of the lipid structure, including aliphatic tail lengths, by performing similar analyses on membranes with distinct compositions. The A32 experimental data are consistent with a solvated residue, whereas the simulation underestimates the water content at this site.

Altogether, these site-specific peptide probes enabled us to view water penetration into the hydrophobic region of the bilayer, which may be otherwise thermodynamically unfavorable in peptide-free bilayers. Previous literature has observed negligible water presence beyond the ester carbonyl in pure lipid membranes (75–77) and has also suggested that embedded amino acids and peptides may facilitate water diffusion across lipid bilayers and into the alkyl bulk (6,78–82). Our results are consistent with this literature; however, additional methods beyond the scope of this study are required to rule out potential peptide multimers in the membrane.

CONCLUSIONS

Our results provide concrete experimental evidence of water permeation into the DMPC lipid bilayer using a series of site-specific probes of backbone hydration. Introducing single-residue $^{13}C{=}^{18}O$ isotope labels additionally allowed us to probe local environments at various embedding depths, and we captured increased hydration facilitated by a transmembrane peptide.

In summary, our conclusions are as follows: 1) We demonstrate the capabilities of isotope-edited 2D IR spectroscopy for examining site-specific hydration levels that are otherwise difficult to access with most common biophysical techniques; 2) residues in pHLIP that are fully inserted in the membrane core remain hydrated, and the polar residues may facilitate water penetration into the hydrophobic region of a DMPC bilayer; 3) MD simulations underrepresent solvation and water interactions; and 4) further improvements in electrostatic maps and MD should more accurately capture the effect of membrane electrostatics.

These studies showcase the advanced sensitivity of 2D IR over traditional FTIR experiments to successfully resolve low-concentration spectra above background signals. The diagonal slices of these 2D IR spectra give us an analog to FTIR spectra, which we can use to identify hydrogen-bond populations. Simulations serve to interpret experimental data from an atomistic perspective. Our evidence suggests that current models are unlikely to accurately reproduce hydrogen-bond ensembles within the bilayer, which may have important implications for modeling transmembrane proteins. Further improvements to the model will enable more quantitative comparisons between simulations and experiments.

SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2019.03.002.

AUTHOR CONTRIBUTIONS

C.R.B. conceived the study. C.R.B. and J.C.F. designed the experiments. J.C.F. carried out the experiments and simulations. J.C.F. analyzed the data. C.R.B. and J.C.F. wrote the manuscript.

ACKNOWLEDGMENTS

We thank Chi-Jui Feng (University of Chicago) for assistance in computing spectra from MD simulations, and Thomas la Cour Jansen (University of Groningen) for providing the NISE code.

This work has been supported by the Welch Foundation (F-1891) and the National Science Foundation (BIO-1815354). MD simulations were carried out at the Texas Advanced Computing Center.

SUPPORTING CITATIONS

References (83–85) appear in the Supporting Material.

REFERENCES

1. Henzler-Wildman, K., and D. Kern. 2007. Dynamic personalities of proteins. *Nature.* 450:964–972.
2. Oh, K. I., K. B. Smith-Dupont, …, F. Gai. 2015. Kinetics of peptide folding in lipid membranes. *Biopolymers.* 104:281–290.
3. Cymer, F., G. von Heijne, and S. H. White. 2015. Mechanisms of integral membrane protein insertion and folding. *J. Mol. Biol.* 427:999–1022.

4. Müller, D. J., N. Wu, and K. Palczewski. 2008. Vertebrate membrane proteins: structure, function, and insights from biophysical approaches. *Pharmacol. Rev.* 60:43–78.

5. Johansson, A. C., and E. Lindahl. 2009. Protein contents in biological membranes can explain abnormal solvation of charged and polar residues. *Proc. Natl. Acad. Sci. USA.* 106:15684–15689.

6. Heyden, M., J. A. Freites, …, D. J. Tobias. 2012. Assembly and stability of α-helical membrane proteins. *Soft Matter.* 8:7742–7752.

7. Siaw, H. M. H., G. Raghunath, and R. B. Dyer. 2018. Peripheral protein unfolding drives membrane bending. *Langmuir.* 34:8400–8407.

8. von Heijne, G. 1992. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* 225:487–494.

9. Wang, Z., J. M. Jumper, …, T. R. Sosnick. 2018. A membrane burial potential with H-bonds and applications to curved membranes and fast simulations. *Biophys. J.* 115:1872–1884.

10. Fendos, J., F. N. Barrera, and D. M. Engelman. 2013. Aspartate embedding depth affects pHLIP's insertion pK_a. *Biochemistry.* 52:4595–4604.

11. Andreev, O. A., D. M. Engelman, and Y. K. Reshetnyak. 2010. pH-sensitive membrane peptides (pHLIPs) as a novel class of delivery agents. *Mol. Membr. Biol.* 27:341–352.

12. Andreev, O. A., A. G. Karabadzhak, …, Y. K. Reshetnyak. 2010. pH (low) insertion peptide (pHLIP) inserts across a lipid bilayer as a helix and exits by a different path. *Proc. Natl. Acad. Sci. USA.* 107:4081–4086.

13. Vāvere, A. L., G. B. Biddlecombe, …, J. S. Lewis. 2009. A novel technology for the imaging of acidic prostate tumors by positron emission tomography. *Cancer Res.* 69:4510–4516.

14. Wiedman, G., W. C. Wimley, and K. Hristova. 2015. Testing the limits of rational design by engineering pH sensitivity into membrane-active peptides. *Biochim. Biophys. Acta.* 1848:951–957.

15. Hunt, J. F., P. Rath, …, D. M. Engelman. 1997. Spontaneous, pH-dependent membrane insertion of a transbilayer alpha-helix. *Biochemistry.* 36:15177–15192.

16. Barrera, F. N., D. Weerakkody, …, D. M. Engelman. 2011. Roles of carboxyl groups in the transmembrane insertion of peptides. *J. Mol. Biol.* 413:359–371.

17. Barth, A., and C. Zscherp. 2002. What vibrations tell us about proteins. *Q. Rev. Biophys.* 35:369–430.

18. Barth, A. 2007. Infrared spectroscopy of proteins. *Biochim. Biophys. Acta.* 1767:1073–1101.

19. van Wilderen, L. J., D. Kern-Michler, …, J. Bredenbeck. 2014. Vibrational dynamics and solvatochromism of the label SCN in various solvents and hemoglobin by time dependent IR and 2D-IR spectroscopy. *Phys. Chem. Chem. Phys.* 16:19643–19653.

20. Le Sueur, A. L., R. E. Horness, and M. C. Thielges. 2015. Applications of two-dimensional infrared spectroscopy. *Analyst.* 140:4336–4349.

21. Chalyavi, F., D. G. Hogle, and M. J. Tucker. 2017. Tyrosine as a non-perturbing site-specific vibrational reporter for protein dynamics. *J. Phys. Chem. B.* 121:6380–6389.

22. Ma, J., I. M. Pazos, …, F. Gai. 2015. Site-specific infrared probes of proteins. *Annu. Rev. Phys. Chem.* 66:357–377.

23. Serrano, A. L., M. M. Waegele, and F. Gai. 2012. Spectroscopic studies of protein folding: linear and nonlinear methods. *Protein Sci.* 21:157–170.

24. MacKenzie, K. R., J. H. Prestegard, and D. M. Engelman. 1997. A transmembrane helix dimer: structure and implications. *Science.* 276:131–133.

25. Marecek, J., B. Song, …, D. P. Raleigh. 2007. A simple and economical method for the production of 13C,18O-labeled Fmoc-amino acids with high levels of enrichment: applications to isotope-edited IR studies of proteins. *Org. Lett.* 9:4935–4937.

26. Baiz, C. R., and A. Tokmakoff. 2015. Structural disorder of folded proteins: isotope-edited 2D IR spectroscopy and Markov state modeling. *Biophys. J.* 108:1747–1757.

27. Shim, S. H., R. Gupta, …, M. T. Zanni. 2009. Two-dimensional IR spectroscopy and isotope labeling defines the pathway of amyloid formation with residue-specific resolution. *Proc. Natl. Acad. Sci. USA.* 106:6614–6619.

28. Smith, A. W., and A. Tokmakoff. 2007. Amide I two-dimensional infrared spectroscopy of β-hairpin peptides. *J. Chem. Phys.* 126:045109.

29. Ganim, Z., H. S. Chung, …, A. Tokmakoff. 2008. Amide I two-dimensional infrared spectroscopy of proteins. *Acc. Chem. Res.* 41:432–441.

30. Krejtschi, C., L. Wu, …, T. A. Keiderling. 2009. Site-specific folding dynamics of isotopically labeled peptides studied by time-resolved infrared-spectroscopy. *Biophys. J.* 96:73a.

31. Torres, J., A. Kukol, …, I. T. Arkin. 2001. Site-specific examination of secondary structure and orientation determination in membrane proteins: the peptidic (13)C=(18)O group as a novel infrared probe. *Biopolymers.* 59:396–401.

32. Sackett, K., and Y. Shai. 2005. The HIV fusion peptide adopts intermolecular parallel β-sheet structure in membranes when stabilized by the adjacent N-terminal heptad repeat: a 13C FTIR study. *J. Mol. Biol.* 350:790–805.

33. Manor, J., E. Arbely, …, I. T. Arkin. 2014. Use of isotope-edited FTIR to derive a backbone structure of a transmembrane protein. *J. Phys. Chem. Lett.* 5:2573–2579.

34. Ludlam, C. F., S. Sonar, …, K. J. Rothschild. 1995. Site-directed isotope labeling and ATR-FTIR difference spectroscopy of bacteriorhodopsin: the peptide carbonyl group of Tyr 185 is structurally active during the bR–>N transition. *Biochemistry.* 34:2–6.

35. Ludlam, C. F., I. T. Arkin, …, K. J. Rothschild. 1996. Fourier transform infrared spectroscopy and site-directed isotope labeling as a probe of local secondary structure in the transmembrane domain of phospholamban. *Biophys. J.* 70:1728–1736.

36. Tatulian, S. A. 2003. Attenuated total reflection Fourier transform infrared spectroscopy: a method of choice for studying membrane proteins and lipids. *Biochemistry.* 42:11898–11907.

37. Rodionova, N. A., S. A. Tatulian, …, L. K. Tamm. 1995. Characterization of two membrane-bound forms of OmpA. *Biochemistry.* 34:1921–1929.

38. Gordon, L. M., P. W. Mobley, …, A. J. Waring. 2004. Conformational mapping of the N-terminal peptide of HIV-1 gp41 in lipid detergent and aqueous environments using 13C-enhanced Fourier transform infrared spectroscopy. *Protein Sci.* 13:1012–1030.

39. Byler, D. M., and H. Susi. 1986. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers.* 25:469–487.

40. Ataka, K., T. Kottke, and J. Heberle. 2010. Thinner, smaller, faster: IR techniques to probe the functionality of biological and biomimetic systems. *Angew. Chem. Int.Engl.* 49:5416–5424.

41. Deflores, L. P., Z. Ganim, …, A. Tokmakoff. 2009. Amide I′-II' 2D IR spectroscopy provides enhanced protein secondary structural sensitivity. *J. Am. Chem. Soc.* 131:3385–3391.

42. Zanni, M. T. 2016. Two-dimensional infrared spectroscopy measures the structural dynamics of a self-assembled film only one molecule thick. *Proc. Natl. Acad. Sci. USA.* 113:4890–4891.

43. Grechko, M., and M. T. Zanni. 2012. Quantification of transition dipole strengths using 1D and 2D spectroscopy for the identification of molecular structures via exciton delocalization: application to α-helices. *J. Chem. Phys.* 137:184202.

44. Petti, M. K., J. P. Lomont, …, M. T. Zanni. 2018. Two-dimensional spectroscopy is being used to address core scientific questions in biology and materials science. *J. Phys. Chem. B.* 122:1771–1780.

45. Hamm, P., and M. T. Zanni. 2011. Concepts and Methods of 2D Infrared Spectroscopy. Cambridge University Press, Cambridge, UK.

46. Kim, Y. S., and R. M. Hochstrasser. 2009. Applications of 2D IR spectroscopy to peptides, proteins, and hydrogen-bond dynamics. *J. Phys. Chem. B.* 113:8231–8251.

47. Reppert, M., and A. Tokmakoff. 2016. Computational amide I 2D IR spectroscopy as a probe of protein structure and dynamics. *Annu. Rev. Phys. Chem.* 67:359–386.
48. Jansen, T., and J. Knoester. 2006. Nonadiabatic effects in the two-dimensional infrared spectra of peptides: application to alanine dipeptide. *J. Phys. Chem. B.* 110:22910–22916.
49. Jansen, T. L., and J. Knoester. 2009. Waiting time dynamics in two-dimensional infrared spectroscopy. *Acc. Chem. Res.* 42:1405–1411.
50. Jansen, T. L., B. M. Auer, …, J. L. Skinner. 2010. Two-dimensional infrared spectroscopy and ultrafast anisotropy decay of water. *J. Chem. Phys.* 132:224503.
51. Liang, C., and T. L. Jansen. 2012. An efficient N(3)-scaling propagation scheme for simulating two-dimensional infrared and visible spectra. *J. Chem. Theory Comput.* 8:1706–1713.
52. Liang, C., M. Louhivuori, …, J. Knoester. 2013. Vibrational spectra of a mechanosensitive channel. *J. Phys. Chem. Lett.* 4:448–452.
53. Reppert, M., and A. Tokmakoff. 2013. Electrostatic frequency shifts in amide I vibrational spectra: direct parameterization against experiment. *J. Chem. Phys.* 138:134116.
54. Reppert, M., A. R. Roy, and A. Tokmakoff. 2015. Isotope-enriched protein standards for computational amide I spectroscopy. *J. Chem. Phys.* 142:125104.
55. Ghosh, A., J. S. Ostrander, and M. T. Zanni. 2017. Watching proteins wiggle: mapping structures with two-dimensional infrared spectroscopy. *Chem. Rev.* 117:10726–10759.
56. Edington, S. C., J. C. Flanagan, and C. R. Baiz. 2016. An empirical IR frequency map for ester C═O stretching vibrations. *J. Phys. Chem. A.* 120:3888–3896.
57. Venyaminov, S. Yu, and N. N. Kalnin. 1990. Quantitative IR spectrophotometry of peptide compounds in water (H2O) solutions. I. Spectral parameters of amino acid residue absorption bands. *Biopolymers.* 30:1243–1257.
58. Edington, S. C., A. Gonzalez, …, C. R. Baiz. 2018. Coordination to lanthanide ions distorts binding site conformation in calmodulin. *Proc. Natl. Acad. Sci. USA.* 115:E3126–E3134.
59. Lee, J., X. Cheng, …, W. Im. 2016. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, openMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* 12:405–413.
60. Wu, E. L., X. Cheng, …, W. Im. 2014. CHARMM-GUI membrane builder toward realistic biological membrane simulations. *J. Comput. Chem.* 35:1997–2004.
61. Jo, S., J. B. Lim, …, W. Im. 2009. CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes. *Biophys. J.* 97:50–58.
62. Jo, S., T. Kim, and W. Im. 2007. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS One.* 2:e880.
63. Jo, S., T. Kim, …, W. Im. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29:1859–1865.
64. Brooks, B. R., C. L. Brooks, III, …, M. Karplus. 2009. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30:1545–1614.
65. Jo, S., X. Cheng, …, W. Im. 2014. CHARMM-GUI PDB manipulator for advanced modeling and simulations of proteins containing nonstandard residues. *Adv. Protein Chem. Struct. Biol.* 96:235–265.
66. Jorgensen, W. L., J. Chandrasekhar, …, M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
67. Abraham, M. J., T. Murtola, …, E. Lindahl. 2015. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 1–2:19–25.
68. van der Spoel, D., P. J. van Maaren, …, N. Tîmneanu. 2006. Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media. *J. Phys. Chem. B.* 110:4393–4398.
69. Huang, J., and A. D. MacKerell, Jr. 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* 34:2135–2145.
70. la Cour Jansen, T., A. G. Dijkstra, …, J. Knoester. 2006. Modeling the amide I bands of small peptides. *J. Chem. Phys.* 125:44312.
71. Husseini, F. S., D. Robinson, …, J. D. Hirst. 2017. Computing infrared spectra of proteins using the exciton model. *J. Comput. Chem.* 38:1362–1375.
72. Wang, L., C. T. Middleton, …, J. L. Skinner. 2011. 2DIR spectroscopy of human amylin fibrils reflects stable β-sheet structure. *J. Am. Chem. Soc.* 133:16062–16071.
73. Daumar, P., C. A. Wanger-Baumann, …, J. S. Lewis. 2012. Efficient ($^{18)}$F-labeling of large 37-amino-acid pHLIP peptide analogues and their biological evaluation. *Bioconjug. Chem.* 23:1557–1566.
74. Reshetnyak, Y. K., M. Segala, …, D. M. Engelman. 2007. A monomeric membrane peptide that lives in three worlds: in solution, attached to, and inserted across lipid bilayers. *Biophys. J.* 93:2363–2372.
75. Hansen, F. Y., G. H. Peters, …, A. Miskowiec. 2012. Diffusion of water and selected atoms in DMPC lipid bilayer membranes. *J. Chem. Phys.* 137:204910.
76. Das, J., E. Flenner, and I. Kosztin. 2013. Anomalous diffusion of water molecules in hydrated lipid bilayers. *J. Chem. Phys.* 139:065102.
77. von Hansen, Y., S. Gekle, and R. R. Netz. 2013. Anomalous anisotropic diffusion dynamics of hydration water at lipid membranes. *Phys. Rev. Lett.* 111:118103.
78. Kong, X., Z. Zhao, and J. Jiang. 2017. Dipeptides embedded in a lipid bilayer membrane as synthetic water channels. *Langmuir.* 33:11490–11495.
79. Kristensen, K., N. Ehrlich, …, T. L. Andresen. 2015. Single-vesicle detection and analysis of peptide-induced membrane permeabilization. *Langmuir.* 31:2472–2483.
80. Fiedler, S., J. Broecker, and S. Keller. 2010. Protein folding in membranes. *Cell. Mol. Life Sci.* 67:1779–1798.
81. Jusoh, S. A., and V. Helms. 2011. Helical integrity and microsolvation of transmembrane domains from Flaviviridae envelope glycoproteins. *Biochim. Biophys. Acta.* 1808:1040–1049.
82. Perkins, R., and V. Vaida. 2017. Phenylalanine increases membrane permeability. *J. Am. Chem. Soc.* 139:14388–14391.
83. Holzwarth, G., and P. Doty. 1965. The ultraviolet circular dichroism of polypeptides. *J. Am. Chem. Soc.* 87:218–228.
84. Venyaminov SYu, I. A. Baikalov, …, J. T. Yang. 1993. Circular dichroic analysis of denatured proteins: inclusion of denatured proteins in the reference set. *Anal. Biochem.* 214:17–24.
85. Lewis, R. N., and R. N. McElhaney. 2013. Membrane lipid phase transitions and phase organization studied by Fourier transform infrared spectroscopy. *Biochim. Biophys. Acta.* 1828:2347–2358.

Biophysical Journal, Volume *116*

Supplemental Information

Site-Specific Peptide Probes Detect Buried Water in a Lipid Membrane

Jennifer C. Flanagan and Carlos R. Baiz

SUPPORTING INFORMATION

Site-specific peptide probes detect buried water in a lipid membrane

Jennifer C. Flanagan[1] and Carlos R. Baiz[1,*]

[1]*Department of Chemistry, University of Texas at Austin*

*cbaiz@cm.utexas.edu

S1. Sample Preparation

a. Isotope site selection

Prior to selecting sites for isotope incorporation, representative FTIR spectra of helical pHLIP in DMPC bilayers were generated using the MD simulation and electrostatic map procedure described in Section S5 below. Isotope spectra were generated from the same trajectory applying a -60cm^{-1} shift to the residue of interest. Using the simulated spectra as a guide, we selected four sites displaying isotope peaks that were strong, narrow, and fully resolved from the main band. In addition, sites were selected such that they were located in different regions of the membrane system. A13 is located within the hydrophobic region of the lipid tails; L24 is primarily within the hydrophobic lipid tails but accessible to water; L28 lies at the lipid-water interface; and A32 is fully solvated, according to simulations. The calculated spectra for these four sites are given in **Figure S1**.

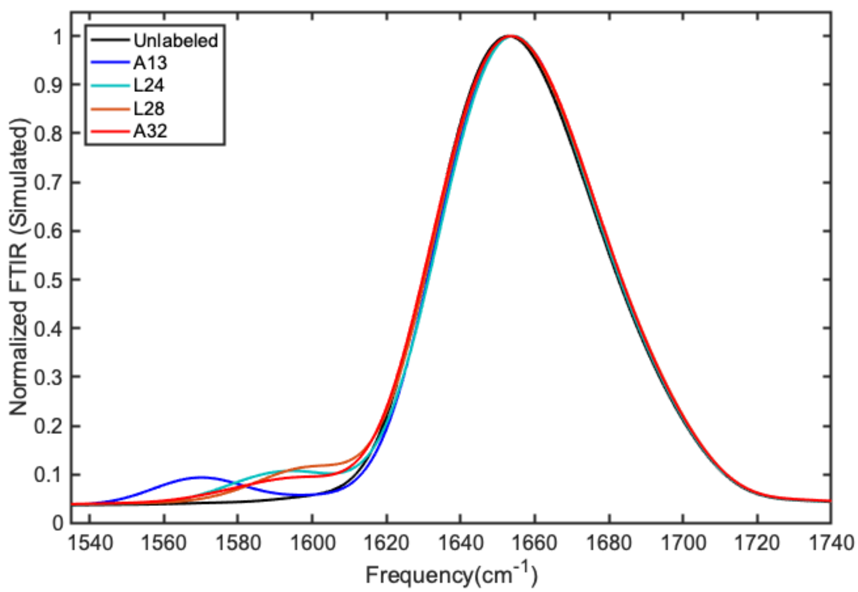

Figure S1. *Calculated IR absorption spectra of the selected isotope label sites.*

b. Isotope label synthesis

N-(9-Fluoroenylmethoxycarbonyl)-L-alanine-1-^{13}C (Fmoc-Ala-OH-1-^{13}C) and N-(9-Fluoroenylmethoxycarbonyl)-L-leucine-1-^{13}C (Fmoc-Ala-OH-1-^{13}C) were purchased from Sigma-Aldrich and used without further purification. $H_2{}^{18}O$ was purchased from Isoflex USA. Prior to use, a 50-mL round bottom flask equipped with a condenser column was assembled and heated to 115 °C for 15 minutes then cooled under nitrogen gas to ensure no water was present. One gram of the ^{13}C enriched amino acid was added to the round bottom flask and dissolved in 1mL ^{18}O water using 1,4-dioxane as a cosolvent in a 4:1 dioxane:water ratio. Acetyl chloride was slowly added to a final concentration of 1M in $H_2{}^{18}O$. The setup was purged with pure nitrogen gas for 15 minutes, then heated to 100 °C and refluxed for 48 hours. Products were recovered via lyophilization of remaining solvent and characterized by mass spectrometry. The lyophilized powder was dissolved in 20% acetonitrile in water and analyzed with flow injection analysis electrospray ionization mass spectrometry (FIA-ESI-MS) in both positive and negative ion modes to ensure ~85% enrichment. No further purification steps were required following MS analysis. The isotope-enriched FMOC-protected amino acids were sent to Biosynthesis, Inc. (Lewisville, TX) and used in solid-state peptide synthesis of singly labeled pHLIPs. pHLIP sequence is AAEQNPIYWARYADWLFTTPLLLLDLALLVDADEGT. Termini are NH_3 and COO^-, respectively.

c. IR sample preparation

Pure (3-(N-morpholino)propanesulfonic acid) (MOPS) was purchased from Sigma-Aldrich, and was dissolved in D_2O and lyophilized before use to exchange acidic hydrogens for deuterium. Deuterated-MOPS was dissolved in D_2O to a 100 mM concentration, and the pH was adjusted to 8.0 using DCl and NaOD. pHLIP was dissolved in DCl solution (1% v/v in D_2O) and lyophilized to remove remaining TFA counterions from synthesis. This was repeated three times until all TFA had been removed, as confirmed by the absence of the TFA carboxylate peak in the FTIR spectra. DMPC was purchased from Avanti Polar Lipids as 25 mg/mL solution in chloroform. Unilamellar DMPC vesicles were formed by drying DMPC under nitrogen overnight and reconstituting to 30 mM in 100mM MOPS buffer (pH 8.0). Samples were sonicated for 10 minutes then extruded with a 100 nm membrane. A 2 mM pHLIP solution was added to the DMPC solutions to a final peptide concentration of 0.6 mM pHLIP for IR data collection. Each sample was adjusted to a pH of 4 (uncorrected pH reading in D_2O) and incubated for 30 minutes before measurement. Samples and stock solutions were stored frozen at -20 °C prior to the IR measurements.

S2. Sample Characterization

a. Isotope synthesis – MS Results

Synthesis products were provided to the UT-Austin mass spectrometry (MS) facility for characterization. The raw results (**Figure S2** and **Figure S3**) from ESI MS in positive ion mode are below. Results showed the presence of $[M+Na]^+$ adducts corresponding to an unlabeled (^{13}C,2x^{16}O), a singly labeled (^{13}C,^{18}O,^{16}O), and a doubly labeled (^{13}C,2x^{18}O) product for both alanine and leucine. Percent enrichment was taken as the sum of the percent composition of the doubly labeled species plus half of the percent composition of the singly labeled species, with a target of >80% enrichment.

Alanine

Exact mass 316.3 g/mol; $[M+Na]^+$ adduct: 339.1179 m/z

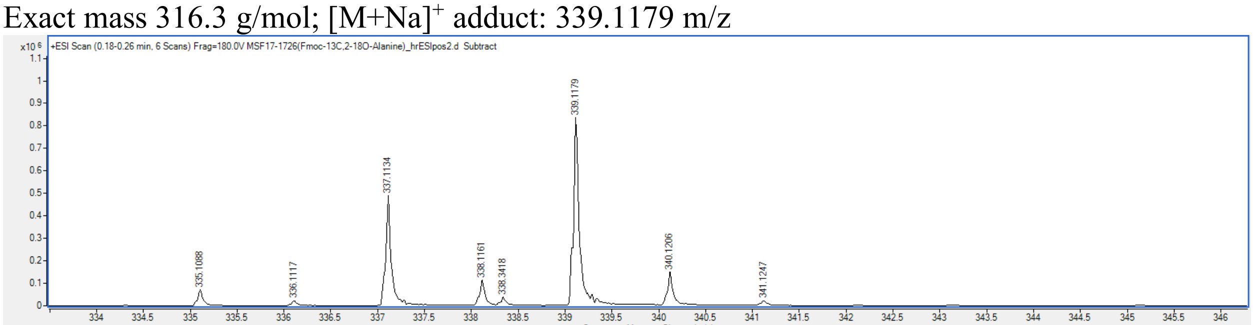

Figure S2. *Mass spectrum including $[M+Na]^+$ adducts obtained from +ESI scan of ^{13}C,^{18}O-Alanine synthesis product, courtesy of UT Austin Mass Spectrometry facility.*

m/z	Species	Abundance	% Composition
335.1084	[^{13}C,2x^{16}O]	21849.36	2.722
337.1130	[^{13}C,^{18}O,^{16}O]	193766.99	24.14
339.1176	[^{13}C,2x^{18}O]	587172.82	73.14

Calculated percent enrichment is 85.21%.

Leucine

Exact mass 358.17 g/mol; $[M+Na]^+$ adduct: 381.163 m/z

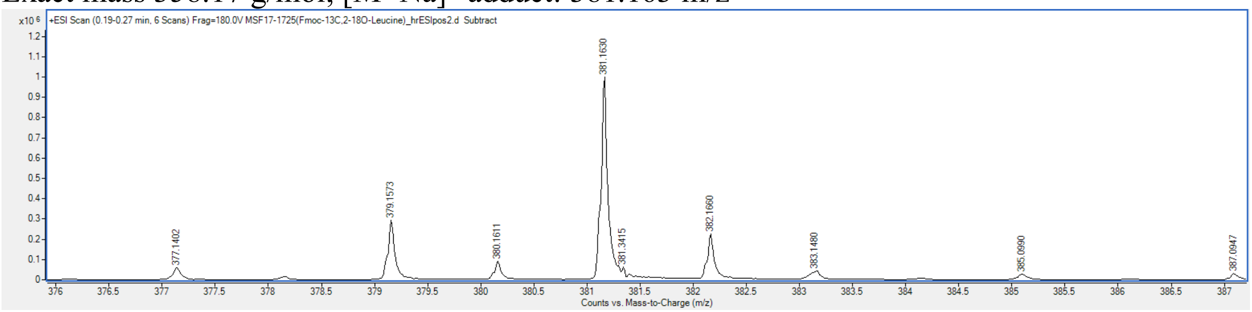

Figure S3. *Mass spectrum including $[M+Na]^+$ adducts obtained from +ESI scan of ^{13}C,^{18}O-Leucine synthesis product, courtesy of UT Austin Mass Spectrometry facility.*

m/z	Species	Abundance	% Composition
377.1402	[^{13}C,2x^{16}O]	58505.67	4.289
379.1563	[^{13}C,^{18}O,^{16}O]	297922.31	21.84
381.1630	[^{13}C,2x^{18}O]	1007579.8	73.87

Calculated percent enrichment is 84.79%.

b. Secondary structure determination – CD Results

Figure S4 shows CD results for pHLIP with DMPC vesicles in MOPS at pH 4 and pH 8. For CD, a 1:30 dilution of the FTIR samples was prepared in 100 mM MOPS buffer for a final peptide concentration of 0.02 mM. CD data was collected from 190 to 300 nm using a 1 mm quartz cuvette. The pH 4 result shows features characteristic of an alpha helix including the combination of a positive band around 193nm and two negative bands around 208nm and 222nm (1). In contrast, these features are absent from the sample held at pH 8, which also shows low ellipticity beyond 210nm. Both of these factors are characteristic of disordered proteins and peptides (2).

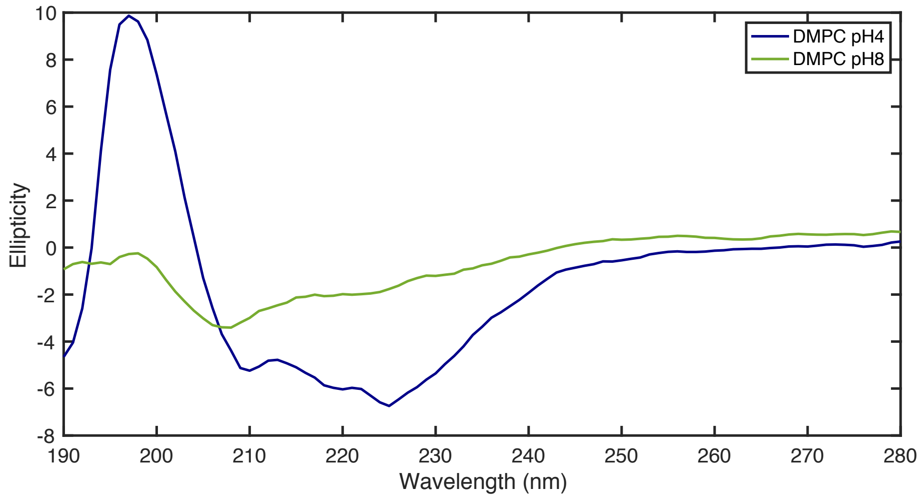

Figure S4. *Circular Dichroism results for pHLIP and DMPC vesicles in MOPS at pH 4 (blue) and pH 8 (green).*

S3. FTIR experiments.

FTIR spectra, given in **Figure S5**, were collected at 1 cm^{-1} resolution at room temperature using a Bruker Vertex 70 spectrometer. The samples were prepared as described in the previous section and were held between two CaF_2 windows with a 50 µm PTFE spacer. The spectra corresponding to the singly labeled peptides are not given as the isotope peaks were not resolved.

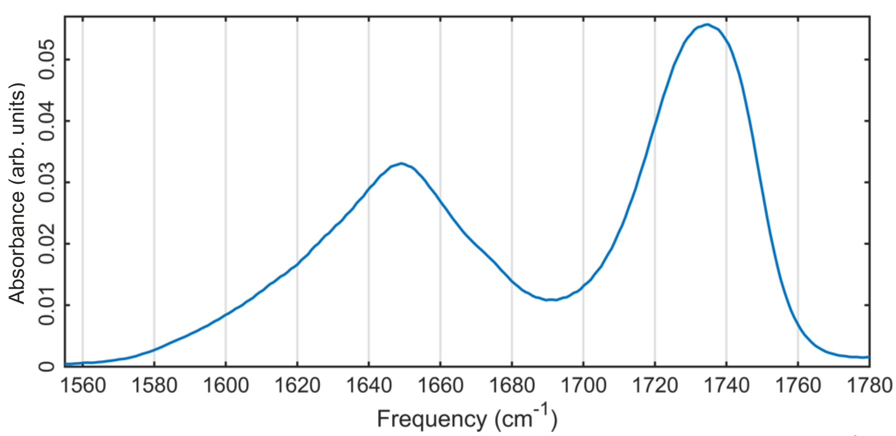

Figure S5. *Top: FTIR spectrum of pHLIP with DMPC in MOPS at pH 4. The band at 1720cm^{-1} corresponds to the ester C=O of the DMPC head group, and the band centered around 1650cm^{-1} is the amide I region.*

The key feature is the alpha helix peak centered at 1650 cm^{-1} and the lack of features between 1555 and 1610 cm^{-1}, which is the region we expect to see deprotonated carboxylate sidechain absorption. **Figure S5** features the FTIR spectrum after subtraction of a solvent (100mM MOPS in D2O) background.

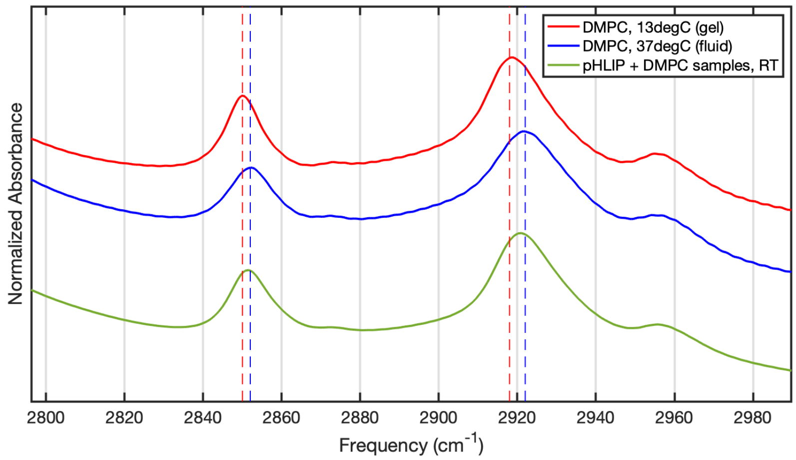

Figure S6. *FTIR spectrum of the asymmetric CH stretching region of gel-phase DMPC at 13°C (top, red), fluid-phase DMPC at 37°C (middle, blue), and DMPC with pHLIP inserted (bottom, green). Vertical dashed lines indicate the frequency$_{max}$ for the two control cases (red representing gel; blue representing fluid) for better visual comparison with experimental pHLIP samples.*

Using the CH asymmetric stretching features of the FTIR spectrum, we determine that these vesicles are in the fluid phase under our experimental conditions. This determination was made from the CH stretch of the FTIR spectrum shown in **Figure S6**, indicated by two peaks near 2850 and 2920 cm^{-1}, respectively, which exhibits a blue shift upon transition to the fluid phase (3). Control spectra of fluid DMPC at 37°C and gel DMPC at 10°C have been included on the same axes for comparative purposes.

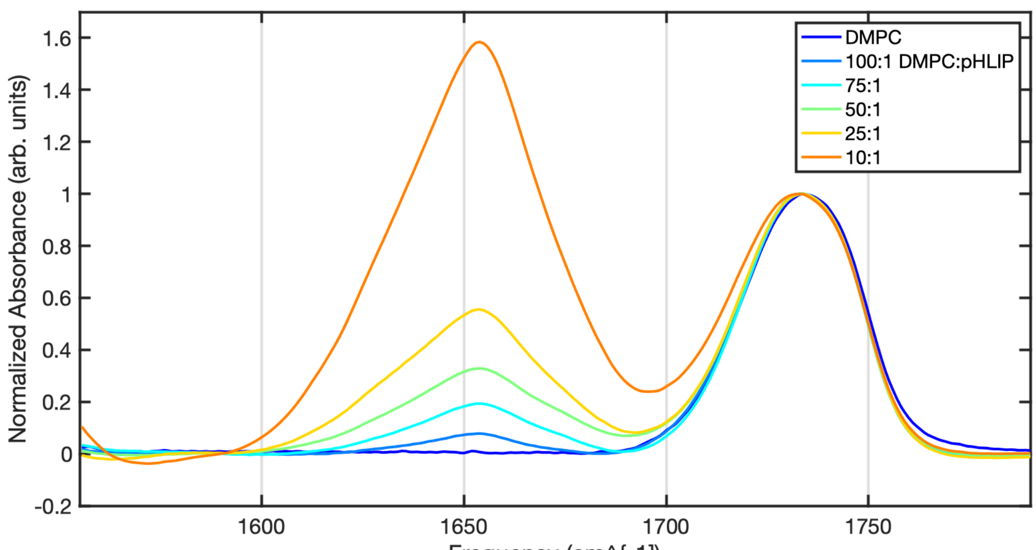

Figure S7*. Concentration-dependent FTIR of pHLIP in DMPC. Peaks are normalized to the maximum intensity of the ester C=O stretch (1730cm^{-1}).*

To rule out the possibility of aggregation, we performed concentration-dependent studies of pHLIP in DMPC, given in **Figure S7**. Protein aggregation is detectable by FTIR as a sharp peak around 1610 cm^{-1}. No aggregation is detected at the 50:1 DMPC:pHLIP ratio reported here, and there are no observed changes in the amide I lineshape upon changing pHLIP concentration. Thus, aggregates are not present these pHLIP samples, however, these spectra cannot rule out the possibility of peptide-peptide interactions.

S4. 2D IR experiments.

Two-dimensional IR experiments were performed in the time domain using an in-house spectrometer previously described by Edington et. al. (4). Spectra were collected using perpendicular polarization with a population time (T_2) of 500 fs and scanned coherence times (t_1) up to 3 ps in 7 fs steps. Four-frame phase cycling was used to remove scatter. Each spectrum was averaged for 11,200,000 laser shots (approx. 8 hours). Diagonal "slices" were taken +7 cm^{-1} from the diagonal axis for comparison with analogous 2D and MD data. All 2D IR spectra and diagonal slices are normalized to the maximum intensity of the amide-I peak. Diagonal slices, including the main amide I band, are shown in **Figure S8**, where we observe the similarity in lineshape between the unlabeled and each of the labeled samples. It should be noted that in **Figure 5**, a Gaussian fit correspondent to the main band has been subtracted from each spectrum to isolate the isotope region and remove nonzero baseline features.

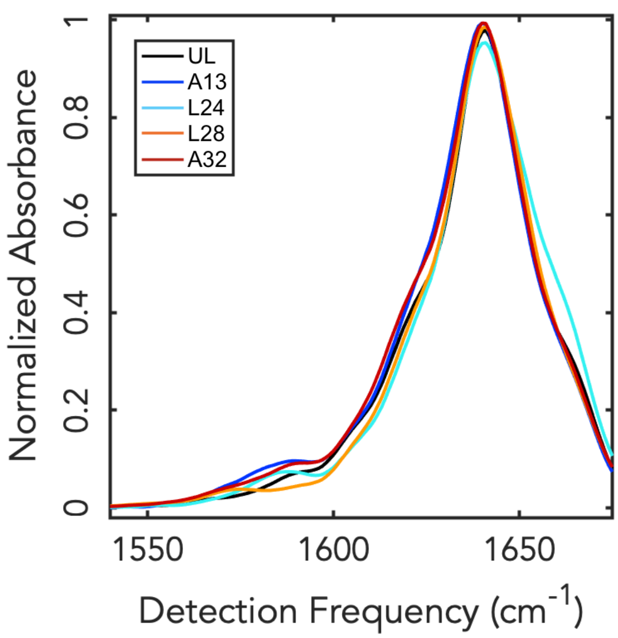

Figure S8. *Diagonal cuts of experimental 2D IR spectra of pHLIP with DMPC in MOPS, pH 4 for the unlabeled and for each isotope labeled peptide. The main band is centered around 1640cm-1.*

S5. Molecular Dynamics Simulations

Molecular dynamics (MD) simulations were carried out using GROMACS 5.1.2 (5). Topologies were generated using CHARMM-GUI as a 9 nm cubic box containing a DMPC bilayer of 64 DMPC molecules per leaflet with an inserted pHLIP as an idealized alpha helix, and solvated with TIP3P water molecules (6–12). Structures were equilibrated in six sequential steps. Trajectories were collected for 10 ns at 300K and 1 atm, storing snapshots every 20 fs for IR calculations and other analysis. Hydrogen bonding (H-bond) analysis was performed using the default hydrogen bond tool in GROMACS (gmx hbond) to calculate H-bond between the backbone C=O oxygen of each isotope labeled site with water as well as with other peptide atoms. RDF functions were computed using the RDF analysis tool in GROMACS for each labeled site.

a. Spectral Calculations

Spectral calculations used the electrostatic map published by Reppert and Tokmakoff using charges defined by the CHARMM27 force field (13). These frequency maps relate electrostatic potential, electric field, and gradient from classical MD simulations to generate matrix elements for an amide I Frenkel-type Hamiltonian (14). These Hamiltonians were calculated over a 1-ns segment of the trajectory, then diagonalized and propagated using the Trotter formula to compute the linear response function according to methods developed by Thomas la Cour Jansen (15–19). Solutions were then Fourier transformed to generate the frequency-domain spectrum. For 2D IR spectra, the waiting time was set to 0 fs and anharmonicity was defined to be -16 cm^{-1}. The same vibrational Hamiltonians were used to generate the third-order response function and subsequent frequency-domain 2D IR spectra for perpendicular polarization conditions.

S6. Additional MD data

a. Isotope-labeled spectra

Computed spectra of the selected isotope sites are given in **Figure S1** (See Section S1).

b. Main band diagonals

Figure S9 shows the diagonal slices of each of the main band diagonals taken at a +7 cm^{-1} offset from the main diagonal in the 2D plots. Solid lines correspond to experimental data, and dashed lines are the simulated spectra. The bands are conserved in the presence of isotope labels, indicating the structure of the peptide is not perturbed. Additionally, there is agreement between experiment and simulation in terms of line shape and peak frequency. Qualitatively, we see differences in linewidth between experiment and simulation. Possible causes for this disparity have been discussed in the main text (Discussion section) and is further detailed below (**Section S7**).

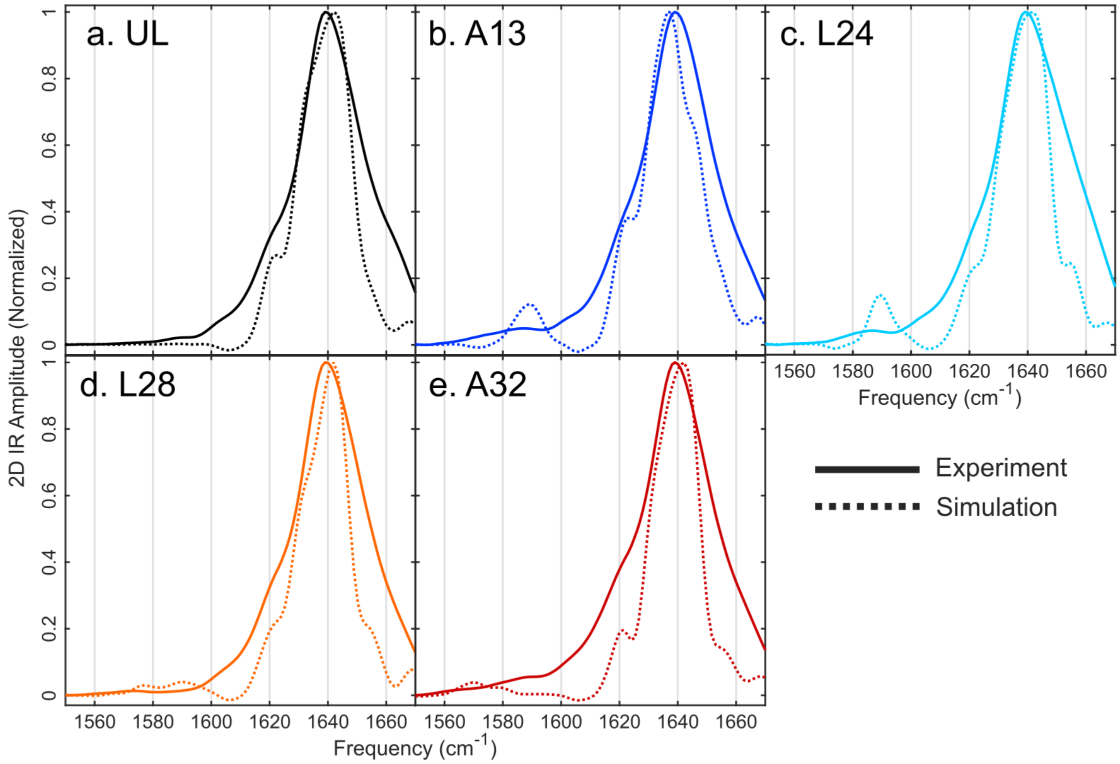

Figure S9. *Diagonal cuts of simulated 2D IR spectra of pHLIP (helix) with DMPC in water. Solid lines are experimental spectra and dashed lines are simulations. The main amide I band is centered around 1640 cm^{-1}, with the additional isotope peak showing up between 1560 and 1610 cm^{-1}.*

S7. Studies of Charge Effects on Spectral Predictions

To trace the origin of the narrowed peaks in the simulated data, we ran several additional spectral calculations on the original trajectory with altered electrostatics, modifying the topology to set the charges on certain molecule types to zero in the spectral calculations (note that coordinates are from the original trajectory). A 5 ns segment of each trajectory was analyzed. In total, we ran 5 different calculations:

1. pHLIP in a DMPC bilayer with standard CHARMM36 charges.
2. pHLIP in DMPC with the water charges set to 0.
3. pHLIP in DMPC with the lipid charges set to 0.
4. pHLIP in water with standard TIP3P charges.
5. pHLIP in water with the water charges set to 0.

Figure S10 summarizes the results of calculating an IR spectrum for each simulation, which demonstrates a lineshape dependence on the assigned charge of the water molecule. The results are further compared in **Figure S11**.

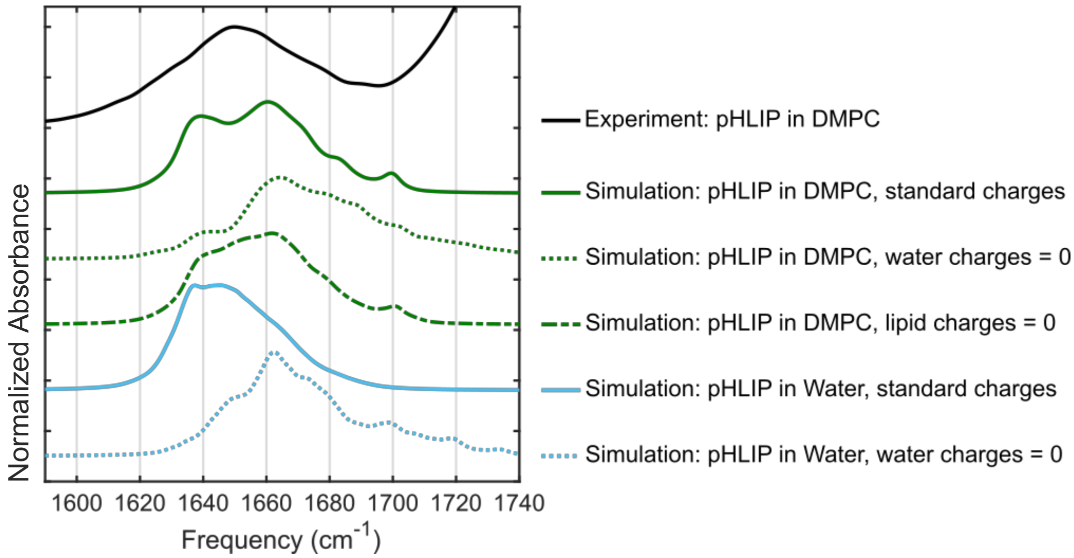

Figure S10. *Computed spectra from MD trajectories using altered electrostatics.*

We observe agreement between pHLIP/DMPC (standard charges) and experiment as well as between pHLIP/water (standard charges) and experiment. However, a key comparison is between pHLIP/DMPC (water = 0) and pHLIP/water (water = 0). While one of these simulations contains lipids and the other is effectively in vacuum, the spectra (**Figure S11**, left) have the same frequency and lineshape, indicating that the electrostatic environment is independent of the lipids.

Similarly, we see negligible changes in the spectrum in the presence of lipids between the calculated spectrum with standard lipid charges versus the spectrum calculated for zero lipid charges (**Figure S11**, right).

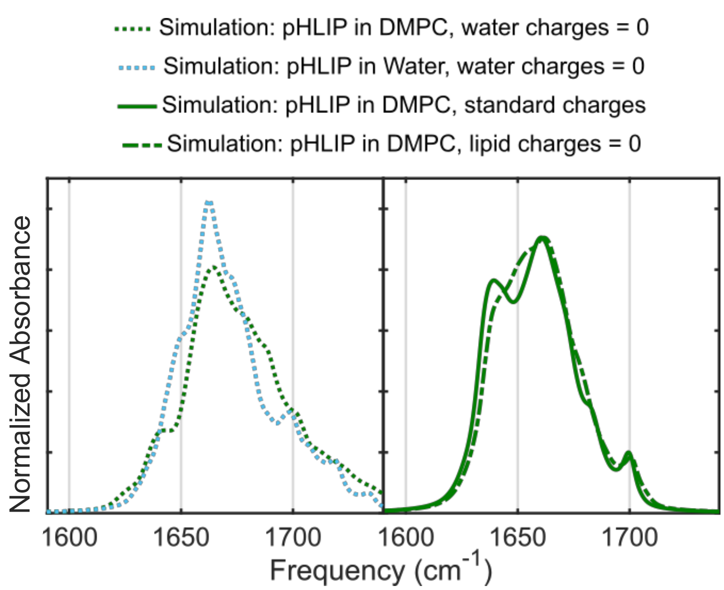

Figure S11. *Computed spectra with altered atomic charges.*

These observations support the hypothesis that the electrostatic maps used in these calculations do not accurately capture the lipid electrostatics, forcing a more homogeneous environment than in the actual system; thus, we attribute this as the source of the narrow peaks in the simulation.

a. Site frequencies: water electrostatics

To further support the claim that the residues within the lipid experience an artificially homogeneous environment, we analyzed the site-specific frequency for each residue throughout the simulation. The plots in **Figure S12** and **Figure S13** contain a histogram for each residue's frequency throughout the simulation, which is extracted from the diagonal elements of the Hamiltonian for each simulation frame.

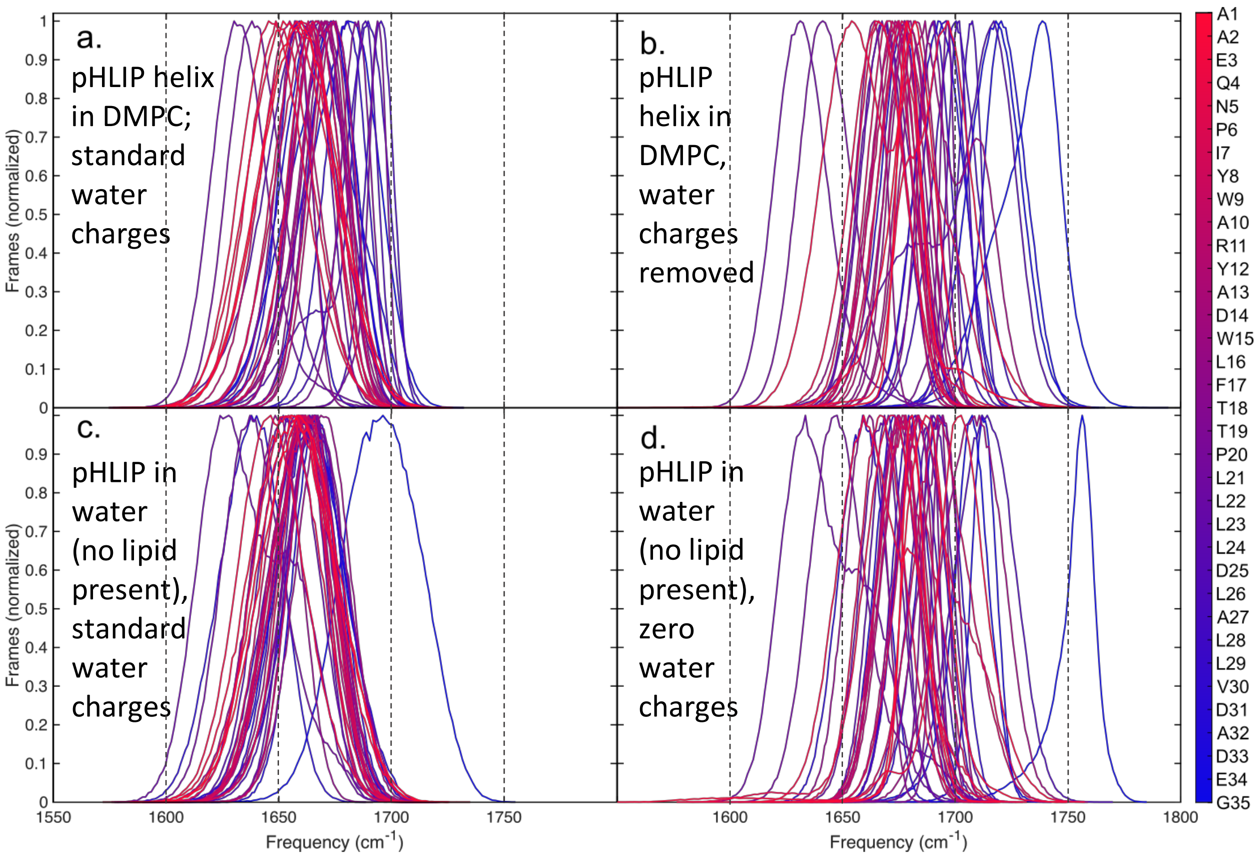

Figure S12*. Site energy histograms for varied electrostatics simulations. a) pHLIP in DMPC, standard charges; b) pHLIP in DMPC, water charges = 0; c) pHLIP helix in water without DMPC present, standard charges; d) pHLIP helix in water without DMPC present, water charges = 0.*

In the simulations where the water charges have been removed, we see a broader range of frequencies from residue to residue, but each individual histogram is narrower than its counterpart with standard water charges. This indicates that the local environment is more homogeneous without water electrostatics than with (including in the presence of lipid charges), which further supports the initial hypothesis.

b. Site frequencies: lipid electrostatics

Figure S13 highlights the difference in the effect of water on the site frequencies versus the effect of lipid electrostatics.

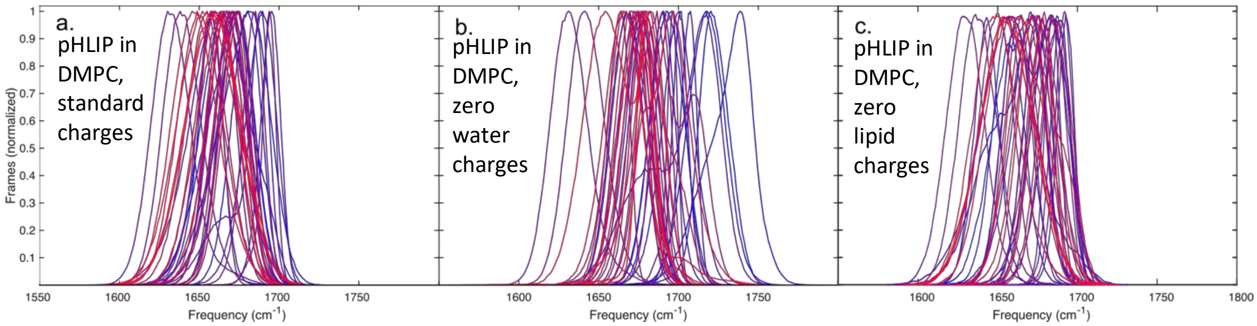

Figure S13. Site energy histograms for varied electrostatic simulations of pHLIP in DMPC bilayer. a) pHLIP in DMPC, standard charges; b) pHLIP in DMPC, water charges =0; c) pHLIP in DMPC, lipid charges = 0.

As expected from the spectral calculations (**Figure S12**), the site histograms show negligible discrepancy between the simulations with and without the lipid charges "turned on." Removing the water charges has a pronounced effect on the frequency distribution.

SUPPORTING REFERENCES

1. Holzwarth, G., and P. Doty. 1965. The Ultraviolet Circular Dichroism of Polypeptides. *J. Am. Chem. Soc*. 87:218-228.
2. Venyaminov, S.Y., I.A. Baikalov, Z.M. Shen, C.S.C. Wu, and J.T. Yang. 1993. Circular Dichroic Analysis of Denatured Proteins: Inclusion of Denatured Proteins in the Reference Set. *Anal. Biochem.* 214:17–24.
3. Lewis, R.N.A.H., and R.N. McElhaney. 2013. Membrane lipid phase transitions and phase organization studied by Fourier transform infrared spectroscopy. *BBA - Biomembr.* 1828:2347–2358.
4. Edington, S.C., A. Gonzalez, T.R. Middendorf, D.B. Halling, R.W. Aldrich, and C.R. Baiz. 2018. Coordination to lanthanide ions distorts binding site conformation in calmodulin. *Proc. Natl. Acad. Sci. USA*. 115:E3126–E3134.
5. Abraham, M.J., T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, and E. Lindahl. 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 1–2:19–25.
6. Jo, S., T. Kim, V.G. Iyer, and W. Im. 2008. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29:1859–1865.
7. Brooks, B.R., C.L. Brooks, A.D. Mackerell, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, and M. Karplus. 2009. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* 30:1545–1614.
8. Lee, J., X. Cheng, J.M. Swails, M.S. Yeom, P.K. Eastman, J.A. Lemkul, S. Wei, J. Buckner, J.C. Jeong, Y. Qi, S. Jo, V.S. Pande, D.A. Case, C.L. Brooks, A.D. MacKerell, J.B. Klauda, and W. Im. 2016. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* 12:405–413.
9. Jo, S., X. Cheng, S.M. Islam, L. Huang, H. Rui, A. Zhu, H.S. Lee, Y. Qi, W. Han, K. Vanommeslaeghe, A.D. MacKerell, B. Roux, and W. Im. 2014. CHARMM-GUI PDB Manipulator for Advanced Modeling and Simulations of Proteins Containing Nonstandard Residues. *Adv. Protein Chem. Struct. Biol.* 96:235–265.
10. Jo, S., J.B. Lim, J.B. Klauda, and W. Im. 2009. CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophys. J.* 97:50–58.
11. Klauda, J.B., R.M. Venable, J.A. Freites, J.W. O'Connor, D.J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A.D. MacKerell, and R.W. Pastor. 2010. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B.* 114:7830–7843.
12. Jorgensen, W.L., J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
13. Reppert, M., and A. Tokmakoff. 2013. Electrostatic frequency shifts in amide I vibrational spectra: Direct parameterization against experiment. *J. Chem. Phys.* 138:134116.
14. Reppert, M., and A. Tokmakoff. 2016. Computational Amide I 2D IR Spectroscopy as a Probe of Protein Structure and Dynamics. *Annu. Rev. Phys. Chem.* 67:359–386.
15. Liang, C., M. Louhivuori, S.J. Marrink, T. la Cour Jansen, and J. Knoester. 2013.

Vibrational Spectra of a Mechanosensitive Channel. *J. Phys. Chem. Lett.* 4:448–452.

16. la Cour Jansen, T., and J. Knoester. 2009. Waiting Time Dynamics in Two-Dimensional Infrared Spectroscopy. *Acc. Chem. Res.* 42:1405–1411.
17. Liang, C., and T. la Cour Jansen. 2012. An Efficient N 3 -Scaling Propagation Scheme for Simulating Two-Dimensional Infrared and Visible Spectra. *J. Chem. Theory Comput.* 8:1706–1713.
18. la Cour Jansen, T., and J. Knoester. 2006. Nonadiabatic Effects in the Two-Dimensional Infrared Spectra of Peptides: Application to Alanine Dipeptide. *J. Phys. Chem. B*, 110:22910-22916.
19. la Cour Jansen, T., B.M. Auer, M. Yang, and J.L. Skinner. 2010. Two-dimensional infrared spectroscopy and ultrafast anisotropy decay of water. *J. Chem. Phys.* 132: 224503.