Quantification of transition dipole strengths using 1D and 2D spectroscopy for the identification of molecular structures via exciton delocalization: Application to α -helices

Maksim Grechko and Martin T. Zanni^{a)}

Department of Chemistry, University of Wisconsin, 1101 University Avenue, Madison, Wisconsin 53706, USA

(Received 6 August 2012; accepted 9 October 2012; published online 13 November 2012)

Vibrational and electronic transition dipole strengths are often good probes of molecular structures, especially in excitonically coupled systems of chromophores. One cannot determine transition dipole strengths using linear spectroscopy unless the concentration is known, which in many cases it is not. In this paper, we report a simple method for measuring transition dipole moments from linear absorption and 2D IR spectra that does not require knowledge of concentrations. Our method is tested on several model compounds and applied to the amide I' band of a polypeptide in its random coil and α -helical conformation as modulated by the solution temperature. It is often difficult to confidently assign polypeptide and protein secondary structures to random coil or α -helix by linear spectroscopy alone, because they absorb in the same frequency range. We find that the transition dipole strength of the random coil state is $0.12 \pm 0.013 \text{ D}^2$, which is similar to a single peptide unit, indicating that the vibrational mode of random coil is localized on a single peptide unit. In an α -helix, the lower bound of transition dipole strength is $0.26 \pm 0.03 \text{ D}^2$. When taking into account the angle of the amide I' transition dipole vector with respect to the helix axis, our measurements indicate that the amide I' vibrational mode is delocalized across a minimum of 3.5 residues in an α -helix. Thus, one can confidently assign secondary structure based on exciton delocalization through its effect on the transition dipole strength. Our method will be especially useful for kinetically evolving systems, systems with overlapping molecular conformations, and other situations in which concentrations are difficult to determine. © 2012 American Institute of Physics. [<http://dx.doi.org/10.1063/1.4764861>]

I. INTRODUCTION

Infrared spectroscopy is a well-established tool for studying protein and polypeptide secondary structures. Due to the vibrational coupling between amide groups, α -helices, β -sheets, random coils, and other backbone structures have characteristic absorption frequencies in the so-called amide I band (the backbone carbonyl stretch, amide I' for peptides in D_2O).¹ This relationship between secondary structure and amide I absorption frequency has been utilized for decades and is still especially important in membrane, aggregate, and kinetic protein studies.² However, the amide I frequency itself does not always provide definitive assignments. For instance, random coils absorb at about 1645 cm^{-1} , which overlaps with α -helices that absorb between 1635 cm^{-1} and 1655 cm^{-1} , depending on whether they are soluble or membrane bound.^{3–8} In principle, one can fit the amide I absorption band, but it is difficult to be confident in fits because structural disorder causes symmetry forbidden transitions to appear and because the absorption bands themselves have complex lineshapes, among other problematic issues.^{9–16}

In this paper, we report a method for measuring transition dipole strengths from 1D and 2D IR spectroscopy and illustrate the utility of using transition dipole strengths to identify secondary structure. Transition dipole strengths

can oftentimes be more sensitive to secondary structure than the couplings themselves. Vibrational couplings not only alter the amide I frequency, but they also redistribute the oscillator strength.^{16–24} The redistribution of oscillator strength occurs even for couplings that are insufficient to produce appreciable frequency shifts. As an example, consider two carbonyl stretches on two molecules separated by a distance d and at an angle of $\theta = 25^\circ$, like shown in Fig. 1(a). In our example, the local mode frequencies of these two molecules differ by 30 cm^{-1} , such as might be caused by differences in environment, hydrogen bonding, or isotopic labeling. The absorption spectrum for these oscillators is simulated by diagonalizing a 2×2 Hamiltonian to generate normal modes that are linear combinations of the original local modes. If d is large enough so that there is negligible coupling, the absorption spectrum (Fig. 1(b), solid line) is unchanged from that of the oscillators themselves because the vibrational modes of the two carbonyl groups are independent of one another, resulting in two absorption bands of equal intensity. If instead the molecules are close enough to be coupled, then the vibrational modes become delocalized to form normal modes, the frequencies shift apart, and the intensities change. For our two oscillators, a coupling strength of 7 cm^{-1} produces the absorption spectrum shown in Fig. 1(b) (dashed line) in which the two absorption bands are now separated by 33 cm^{-1} and differ in intensity by 70%. Thus, one can detect coupled modes and thereby identify the molecular structures by either the frequency

^{a)} Author to whom correspondence should be addressed. Electronic mail: zanni@chem.wisc.edu.

FIG. 1. Illustration of the effects of coupling on transition dipole strengths, frequencies and intensities in 1D and 2D IR spectra. (a) Two carbonyl groups at angle 25° separated by distance d . (b) Simulated absorbance spectrum of the carbonyls shown in (a) in the uncoupled (solid line) and coupled (dashed line) regime. (c) Simulated absorbance spectrum and (d) diagonal slice of simulated 2D IR spectrum of two near degenerate parallel ($\theta = 0$) oscillators. In (c) and (d) uncoupled and coupled regimes are shown by solid and dashed lines, respectively.

separation or the intensity caused by delocalization of the vibrational modes. An interesting consequence regarding the intensity is that whenever an absorption band in an infrared spectrum has a transition dipole moment that is larger than any of the individual modes themselves, a structure must exist that is coupling the modes. Thus, transition dipoles are good indicators of peptide secondary structure, for example, since many individual amide I bonds are coupled together in beta-sheets and α -helices, creating strong transition dipoles.^{22,25–27}

Of course, one can also detect coupling by the frequency separation, but frequency shifts are not always easily measurable. In the example above, the coupling caused each peak to shift by only 1.5 cm^{-1} , which may be difficult to experimentally detect considering that the linewidths of most condensed phase vibrational modes are $>10 \text{ cm}^{-1}$. In contrast, the 70% difference in intensities would be quite straightforward to measure. Thus, identifying molecular structures via couplings is often better accomplished through intensities rather than frequencies. A good example of this phenomenon is in experiments on nitrile labeled DNA in which the coupling between two nitrile groups was insufficient to create a measurable shift in frequencies, but still produced a $\sim 60\%$ change in relative intensities.²⁸

One might think that nothing more than absorption (1D) spectroscopy is needed to measure the transition dipole strengths in the above 2-chromophore example. After all, FTIR spectroscopy has been used for decades to measure transition dipole strengths. But absorption spectroscopy has intrinsic as well as practical limitations that prevent the measurement of the transition dipole strength in many situations, especially for coupled chromophores like above. The example above was for two chromophores whose individual absorption spectra were well resolved from one another. Instead, consider two chromophores with $\theta = 0$ that are not individually resolved because they have a frequency separation that is smaller than the linewidth (Fig. 1(c)). When they are uncoupled (solid line) they have equal intensities and when they are coupled, they have unequal normal mode intensities due to delocalization of the vibrations like shown in Fig. 1(c) (dashed line) in which the spectrum was calculated with 10 cm^{-1} of coupling. Besides the frequency shift the two linear spectra are quite similar even though the nature of the vibrational modes is very different. For degenerate local modes, the frequency shift might even be negligible. This example illustrates that the linear spectrum

is relatively insensitive to couplings, and thus molecular structure, when the coupling is comparable to or smaller than the linewidths, which is often the case for the amide I' band of peptides and many other condensed phase vibrational modes. Frequency shifts still occur, but because the integrated area of an absorption spectrum is independent of the vibrational delocalization caused by coupling,¹⁶ congested bands are largely insensitive to changes in transition dipole intensities. Thus, the absorption spectrum of a highly coupled system with strong transition dipole strengths like an α -helix has the same integrated intensity as an uncoupled random coil.³

Non-linear spectroscopies are more sensitive to transition dipole strengths than are linear spectroscopies. The integrated area of absorption spectroscopies is independent of coupling because it scales as the transition dipole squared, e.g., $|\mu|^2$ whereas 2D IR spectra scale as $|\mu|^4$. As a result, the integrated area of a 2D IR spectrum is not conserved when the coupling changes. Consider the two degenerate modes discussed above ($\theta = 0$). Diagonal slices through simulated 2D IR spectra are shown in Fig. 1(d) for the uncoupled (solid) and coupled (dashed) chromophores simulated in Fig. 1(c). Notice that the intensity of the 2D IR peaks is much more intense in the coupled than in the uncoupled case. If the two oscillators were degenerate, then the integrated area of the coupled 2D IR spectrum would be twice that of the uncoupled case (because the entire oscillator strength of both chromophores would contribute to the symmetric stretch). The enhancement is even larger for extended systems of coupled oscillators. Thus, non-linear spectroscopy is more sensitive than linear spectroscopy to secondary structures that have delocalized vibrational excitons.

In this paper, we report a procedure to measure transition dipole strengths from the combination of 1D and 2D spectroscopy. Neither spectroscopy alone is sufficient to measure transition dipole strengths. To measure the transition dipole strength from an absorption spectrum, one needs to know the path length of the cell and the concentration of the analyte. The path length can be accurately set with a properly calibrated sample cell, but the concentration may not be known, such as for inhomogeneous samples, systems in equilibrium between multiple conformations, or kinetically evolving structures. It is difficult to extracting transition dipole strengths from 2D spectroscopy alone, because it would require knowing beam overlaps and pump intensities in addition to the concentration requirements. We take advantage of

the fact that the amplitudes of both 1D and 2D spectra scale linearly with concentration, C , but non-linearly with transition dipole strengths. That is, 1D spectra scale as $C|\mu|^2$ while 2D spectra scale as $C|\mu|^4$. This fact is well documented in the literature and has been used on occasion to approximate transition dipole strengths.²⁹ In this paper, we develop a robust and simple method for extracting the transition strength that can be applied to most situations, including kinetically evolving samples. We demonstrate our procedure using 1D and 2D infrared spectroscopy of a model compound and then apply it to a polypeptide whose secondary structure we control with temperature. We discuss how the measurement of the transition dipole strength is a measure of the vibrational excitonic delocalization in a polypeptide and how this delocalization can be used to identify secondary structures. In this paper we focus on infrared spectroscopy, but the same principles apply to 1D and 2D electronic/visible spectroscopy as well, such as for light harvesting proteins consisting of coupled pigment chromophores or J-aggregates formed from coupled dyes.

II. MATHEMATICAL FORMALISM

The optical density, $OD(\omega)$, of a sample in a linear absorption measurement is given by

$$OD(\omega) \propto |\mu|^2 \Phi^{(1)}(\omega) CL, \quad (1)$$

where μ is the transition dipole moment, L is the width of the sample cell, C is the concentration, and $\Phi^{(1)}(\omega)$ is the linear vibrational lineshape. For the 2D IR signal, we consider pump-probe beam geometry, in which case the signal strength is given by

$$S(\omega_3) = 2I_{prb}(\omega_3) + C G I_{prb}(\omega_3) \int d\omega_1 [R_{NR}^{(3)}(\omega_3, 0, \omega_1) + R_R^{(3)}(-\omega_3, 0, \omega_1)] I_{pmp}(\omega_1) e^{-i\omega_1 t_1}, \quad (2)$$

in which we have neglected the homodyne contribution since it is much weaker, as is standard practice in 2D IR spectroscopy. Here $R_R^{(3)}$ and $R_{NR}^{(3)}$ are the sum of rephasing and non-rephasing third order response functions, respectively; ω_3 (ω_1) and I_{prb} (I_{pmp}) are the frequency and spectral intensity of the probe (pump) beam; t_1 is the time delay between the two pump pulses. The factor G accounts for all constants as well as experimental parameters that affect the 2D IR signal, such as the diameter of the pump and probe beams in the sample and their overlap. Dependence of G on wavelength is neglected in this work based on sufficiently narrow frequency distribution of the signals. As will be shown later this term vanishes and so its explicit form is not important.

In our measurements we take the logarithm of the signal given by Eq. (2) and perform a Taylor expansion based on the smallness of the third order signal compared to the intensity of the probe beam, to obtain

$$\log(S(\omega_3)) = \log(2I_{prb}(\omega_3)) + \frac{1}{2} C G \int d\omega_1 [R_{NR}^{(3)}(\omega_3, 0, \omega_1) + R_R^{(3)}(-\omega_3, 0, \omega_1)] I_{pmp}(\omega_1) e^{-i\omega_1 t_1}. \quad (3)$$

The first term on the right side of Eq. (3) is subtracted during the experiment with a chopper or phase cycling.³⁰ Af-

ter Fourier transforming over the time delay t_1 the collected signal, which we term the 3rd order optical density, ΔOD , is given by

$$\Delta OD(\omega_3, \omega_1) \propto C G [R_{NR}^{(3)}(\omega_3, 0, \omega_1) + R_R^{(3)}(-\omega_3, 0, \omega_1)] I_{pmp}(\omega_1). \quad (4)$$

Using the Condon approximation we write ΔOD in the form analogous to the expression for linear absorption:

$$\Delta OD(\omega_3, \omega_1) \propto C G |\mu|^4 \Phi^{(3)}(\omega_3, \omega_1) I_{pmp}(\omega_1), \quad (5)$$

where $\Phi^{(3)}(\omega_3, \omega_1)$ is a 2D lineshape that is determined not only by the vibrational dynamics of the molecule but also by the 2D IR pulse sequence. In the experiments presented here, we use a pulse sequence that creates an absorptive 2D lineshape and set the waiting time to zero so that the vibrational lifetime does not decrease the peak intensity. We also consider only diagonal peaks, not cross peaks, and so $\omega_1 = \omega_3 = \omega$.

Combining Eqs. (1) and (5), we obtain an expression for the ratio of the 3rd order optical density to 1D absorbance:

$$\frac{\Delta OD(\omega, \omega)}{OD(\omega)} \propto |\mu|^2 \frac{\Phi^{(3)}(\omega, \omega)}{\Phi^{(1)}(\omega)} \frac{G I_{pmp}(\omega)}{L}. \quad (6)$$

Notice that this ratio does not depend on C and scales as $|\mu|^2$. However, it also depends on the experimental parameters in G . In order to eliminate these variables, which are difficult to precisely characterize *a priori*, we use a calibrant molecule with a known transition dipole strength. The ratio of the 2D to 1D signals of the calibrant follows the same expression as Eq. (6), which we use to normalize the expression of our unknown molecule by taking the following ratio:

$$\begin{aligned} \frac{\Delta OD(\omega, \omega)}{OD(\omega)} &\div \frac{\widetilde{\Delta OD}(\omega^{\max}, \omega^{\max})}{\widetilde{OD}(\omega^{\max})} \\ &= \frac{|\mu|^2}{|\tilde{\mu}|^2} \frac{\Phi^{(3)}(\omega, \omega)}{\Phi^{(1)}(\omega)} \frac{\tilde{\Phi}^{(1)}(\omega^{\max})}{\tilde{\Phi}^{(3)}(\omega^{\max}, \omega^{\max})} \frac{I_{pmp}(\omega)}{I_{pmp}(\omega^{\max})}, \end{aligned} \quad (7)$$

where tilde designates values for the calibration molecule and $\omega^{\max}$ is the frequency at the maximum of the spectrum of the calibration molecule. Notice that normalization eliminates all of the experimental parameters except the intensity of the pump beam, $I_{pmp}(\omega)$. $I_{pmp}(\omega)$ depends upon the laser pulse center frequency and bandwidth. We define $d(\omega)$ to be

$$d(\omega) \equiv \frac{\Delta OD(\omega, \omega)}{OD(\omega)} \frac{\widetilde{OD}(\omega^{\max})}{\widetilde{\Delta OD}(\omega^{\max}, \omega^{\max})} \frac{I_{pmp}(\omega^{\max})}{I_{pmp}(\omega)} |\tilde{\mu}|^2. \quad (8)$$

Using Eq. (7) we obtain

$$d(\omega) = |\mu|^2 \frac{\Phi^{(3)}(\omega, \omega)}{\Phi^{(1)}(\omega)} \frac{\tilde{\Phi}^{(1)}(\omega^{\max})}{\tilde{\Phi}^{(3)}(\omega^{\max}, \omega^{\max})}, \quad (9)$$

in which $d(\omega)$ is a “spectrum” of the transition dipole strength of the unknown molecule.

III. EXPERIMENTAL SETUP AND SAMPLE PREPARATION

The detailed description of our experimental setup can be found elsewhere.^{30,31} In brief, a 60 fs (FWHM) light pulse

at about $6\text{ }\mu\text{m}$ is generated by Ti:sapphire laser system combined with optical parametric amplifier. A small fraction the pulse is split by a CaF_2 window and serves as both probe and local oscillator pulses. The rest is directed into the pulse shaper. The pulse shaper is used to create two pump pulses and to control the time delay t_1 between them. The pump-probe delay is $t_2 = 0$. The pump and probe beams are focused and overlapped in a sample. Because of the pump-probe geometry of our setup, both rephasing and non-rephasing third order signals are emitted along the direction of the probe beam. The sum of emitted signal and probe pulses is directed into a spectrometer (Acton SpectraPro 2150i) and is detected by a mercury cadmium telluride array.

In this work, we used 1,3-cyclohexanedione (Sigma-Aldrich, 97%), *N,N*-dimethylformamide (DMF) (Sigma-Aldrich, $\geq 99.9\%$), L-serine (Sigma-Aldrich), and *N*-methylacetamide (NMA) (Sigma-Aldrich, 99+%). 1,3-cyclohexanedione was dissolved in chloroform-*d* (Sigma-Aldrich, 99.8 atom%D) at concentrations of 107 mM and 27 mM. DMF was dissolved in carbon tetrachloride (Sigma-Aldrich, 99.9%) at concentration of 11 mM. L-serine was first isotope exchanged in D_2O at room temperature overnight, then lyophilized and dissolved in D_2O (Sigma-Aldrich, 99.9 atom%D) to the concentration of 40 mM. NMA was dissolved in D_2O to 119 mM and isotope exchanged overnight.

The 25-residue peptide $(\text{AAAAK})_4\text{AAAAAY}$ (AKA) was synthesized with CEM Liberty1 Single Channel Automated Peptide Synthesizer by standard Fmoc-techniques and purified by reverse phase HPLC (Jasco, Vydac C18 column). The purity of the peptide was confirmed by MALDI mass spectrometry. The peptide was isotope exchanged in D_2O at room temperature for 24 hours, lyophilized and then dissolved in D_2O to the concentration of 6.7 mM (measured by absorbance at 280 nm).

Samples were placed into CaF_2 sample cells with $56 \pm 1.8\text{ }\mu\text{m}$ Teflon spacers. The absorbance of 1,3-cyclohexanedione, NMA, and AKA was calculated as negative common logarithm of the ratio of voltages at the detector array measured back-to-back for the sample and corresponding solvent placed in a similar sample cell with $56 \pm 1.8\text{ }\mu\text{m}$ Teflon spacer. The details of the derivation of absorbance using 2D IR spectrometer are given in the supplementary material.³² The use of 2D IR spectrometer for

derivation of absorbance spectrum, though resulting in larger error bar, allows application of the technique to inhomogeneous and rapidly evolving samples. The absorbance of DMF and L-serine was measured using Thermo Scientific Nicolet iS10 FTIR spectrometer. Cooling and heating of samples was done by FTS Multi-Cool chiller and the temperature was measured by Exttech IR thermometer.

IV. RESULTS

A. Assessing the accuracy of the approach using a model compound

To ascertain the reliability of our method for determining transition dipole strengths, we applied the approach to 1,3-cyclohexanedione and NMA. We used these molecules to assess the robustness of our approach to experimental conditions such as the sample concentration, intensity of the pump beam, alignment of the laser beams, and solvent background absorption.

Figure 2(a) presents a typical 2D IR spectrum of 1,3-cyclohexanedione. 1,3-cyclohexanedione in chloroform has antisymmetric and symmetric stretch absorbances at 1711 cm^{-1} and 1735 cm^{-1} , respectively, which results in two sets of diagonal peaks in the 2D IR spectrum.³³ A slice along the diagonal of 2D spectrum is plotted in Fig. 2(b). The 1D absorption spectrum, measured using the probe beam in the 2D IR spectrometer, is shown in Fig. 2(c). The ratio of the antisymmetric and symmetric peaks is about 10:1 in the 2D IR spectrum while it is about 2.9:1 in the linear spectrum. This ratio of intensities is largely dictated by the angle between the carbonyl bonds, which is 60° .

To calculate $d(\omega)$ according to Eq. (8), a calibrant molecule is needed with a known or measurable transition dipole. We utilized DMF in carbon tetrachloride, because carbon tetrachloride has no background absorbance and DMF absorbs at 1685 cm^{-1} , which is reasonably close to the 1,3-cyclohexanedione absorbances. To our knowledge, the dipole strength of DMF in carbon tetrachloride has not been published previously, so we determined it ourselves to be $0.138 \pm 0.006\text{ D}^2$ using an FTIR spectrometer and a sample with a known concentration and pathlength. Its 2D IR and FTIR spectra are presented in Figs. S8(a) and S8(b), respectively.³²

FIG. 2. Spectra of 1,3-cyclohexanedione. (a) Absorptive 2D IR spectrum. (b) Diagonal slice of the absorptive 2D IR spectrum measured at 107 mM (black lines) and 27 mM (red lines), full (solid lines) and 22% (dashed lines) intensity of the pump beam. Blue line presents the result of the measurement at 107 mM and full intensity but with intentionally misaligned pump beam. (c) Absorbance spectrum at 107 mM (black line) and 27 mM (red line). (d) Calculated $d(\omega)$ (line assignment is the same as in (b)).

With the calibrant in place, $d(\omega)$ for 1,3-cyclohexanedione was calculated using Eq. (8) (Fig. 2(d)), under a range of experimental conditions that are expected to vary in the laboratory. To explore the effects of optical density, we measured data at 107 mM and 27 mM (black and red lines, respectively). To characterize the effects of pump pulse energy, we varied the pump pulse from full to 22% intensity (solid and dashed lines, respectively). Spectra were also measured for a purposely misaligned pump beam that reduced the 2D IR signal strength by 2.6 (blue). These experimental conditions result in very different intensities of both nonlinear and linear spectra. Nonetheless, $d(\omega)$ is remarkably consistent. $d(\omega)$ gives a ratio of transition dipole moments for the antisymmetric and symmetric stretch states to be ≈ 3.2 : 1, which agrees well with the linear spectrum, as it should for this well-resolved monomeric molecule.

To test the approach for peptides, we calculated $d(\omega)$ for NMA (in D_2O), which is a commonly used model for the peptide bond. Rather than use DMF as a calibrant molecule, we used L-serine (in D_2O) because the frequency of its carboxyl stretch vibration (1621 cm^{-1}) is more similar to the amide I' frequency of NMA (1622 cm^{-1}). Its dipole strength was determined to be $|\mu|^2 = 0.20 \pm 0.012\text{ D}^2$ using FTIR spectroscopy. 2D IR and FTIR spectrum of L-serine is shown in Figs. S9(a) and S9(b), respectively.³² The 2D IR spectrum of the amide I' band in NMA and its diagonal slice are shown in Figs. 3(a) and 3(b), respectively. The absorbance spectrum obtained using the 2D IR spectrometer is plotted in Fig. 3(c). The resulting $d(\omega)$ is shown in Fig. 3(d). The error bars were calculated from stochastic error generated by several

FIG. 3. Spectra of NMA. (a) Absorptive 2D IR spectrum. (b) Diagonal slice of the absorptive 2D IR spectrum. (c) Absorbance spectrum. (d) Calculated $d(\omega)$.

measurements (this error inherently includes dispersion of the Teflon spacer thickness), error bars on the optical density ($\pm 5\%$), and transition dipole strength of the calibrant. Notice that the maximum of $d(\omega)$ is located 6 cm^{-1} higher than the absorption maximum, which indicates that non-Condon effects are present. The value of $d(\omega)$ is $0.13 \pm 0.014\text{ D}^2$ at the maximum and $0.10 \pm 0.013\text{ D}^2$ at the frequency of maximum absorbance. This result is consistent with dipole strength of $0.12 \pm 0.007\text{ D}^2$ that we measured independently for NMA using FTIR spectroscopy and agrees with previous reports.^{34,35}

B. Application to a helical peptide

Having verified our method on two compounds, we apply it to assess the α -helical content of a 25 residue α -helical polypeptide. The helical content of the AKA peptide varies with temperature, as was demonstrated by the study of its capped analog.³ Using the same procedure as above, the 1D and 2D IR spectra for AKA were measured at 60°C , 23°C , and 6°C . 2D IR spectra are shown in Fig. 4. Diagonal slices

FIG. 4. Absorptive 2D IR spectra of AKA peptide measured at (a) 60°C , (b) 23°C , and (c) 6°C .

FIG. 5. Spectra of AKA peptide at different temperatures. (a) Diagonal slice of the absorptive 2D IR spectrum at 60 °C (red line), 23 °C (blue line), and 6 °C (black line). (b) Absorbance spectrum. (c) Calculated $d(\omega)$. Line assignment in (b) and (c) is the same as in (a).

through the 2D spectra averaged over 7 measurements are shown in Fig. 5(a) alongside linear spectra in Fig. 5(b). At 60 °C, the maximum absorbance is at 1646 cm⁻¹, which is consistent with the majority of the peptide adopting a random coil conformation. At 6 °C, the peak appears at 1635 cm⁻¹, which is typical of soluble α -helices and consistent with previous work on this polypeptide.^{3,36} Of course, for any temperature there exists a weighted distribution of α -helical and random coil conformations. The large overlap between the two makes it difficult to fit the spectra to determine the relative contributions of each, which highlights the difficulty of monitoring secondary structure content with FTIR spectroscopy alone, especially for an unknown sample.

As discussed in the Introduction, the integrated area of an absorption spectrum should be independent of molecular geometry in an excitonic system. To test this statement for the AKA peptide, we calculated the total transition dipole strength $|\mu^{tot}|^2$ by integrating the absorption spectrum according to

$$|\mu^{tot}|^2 = 9.18 \times 10^{-3} \int \frac{\varepsilon(\omega)}{\omega} d\omega, \quad (10)$$

where $\varepsilon(\omega)$ is the molar extinction coefficient (in M⁻¹ cm⁻¹) obtained from absorbance spectra, known concentration of the peptide and path length. To determine the average dipole strength of a single peptide unit we divide $|\mu^{tot}|^2$ by 25, the number of carbonyl groups, which gives 0.11 ± 0.017 D² at 60 °C, 0.11 ± 0.014 D² at 23 °C, and 0.11 ± 0.013 D² at 6 °C. We draw three conclusions from these values: (1) the amide I' transition dipoles of the local modes do not themselves

depend on temperature, (2) the amide I' absorption band is well-described by an excitonic Hamiltonian, since the integrated area is conserved with temperature, and (3) one cannot conclusively distinguish between a random coil and an α -helix by the absorption intensity alone.

The integrated absorption spectra provide the transition dipole strength for a local amide I stretch of a single oscillator, but of course the normal mode transition dipole strengths of an α -helix and random coil should be very different. To quantitatively determine the transition dipole strength, we calculate $d(\omega)$ following the procedure outlined in supplementary material, which provides a step-by-step example.³² We use L-serine as the calibrant. Plotted in Fig. 5(c) is $d(\omega)$, averaged over 7 measurements for each temperature. At the random coil absorption maximum, $d(1646 \text{ cm}^{-1}) = 0.12 \pm 0.013$ D² at 60 °C. This value is similar to the result for NMA, indicating that the peptide units are acting as independent absorbers when in the random coil conformation. In contrast, at 6 °C at the frequency of the α -helix $d(1635 \text{ cm}^{-1}) = 0.21 \pm 0.019$ D², which is 1.8 fold larger than NMA. Thus, the measurement of $d > 0.12$ D² (of NMA), straightforwardly reveals the formation of a vibrational excitonic state, which is a signature for the formation of an ordered secondary structure.

V. DISCUSSION

The above results on the model compounds and helical peptide indicate that our method for extracting the transition dipole strengths is robust. Before discussing the scientific relevance of the transition dipole strengths for the AKA α -helix, we discuss some affects that must be considered when utilizing this approach.

A. Technical issues that must be accounted for when applying the approach

There are several technical issues that must be considered when applying this approach. One consideration is that the fundamental and sequence band transitions that create the diagonal pairs of peaks in 2D IR spectra will affect $d(\omega)$ if they interfere with one another. 2D IR spectra consist of negative/positive pairs of peaks. The negative peaks appear along the diagonal while the positive peaks are shifted along the probe axis by the anharmonic shift of the vibrational mode. (In some texts the diagonal peaks are plotted positive and the anharmonically shifted peaks are negative. Regardless, they are always 180° out-of-phase.) When the anharmonic shift is smaller than the homogeneous line width, the anharmonically shifted peak overlaps with the fundamental. When this occurs, the fundamental peak is less intense than dictated solely by the transition dipole strength and concentration. Different vibrational modes can have different anharmonic shifts and/or homogeneous line widths and thus the relative intensities of diagonal peaks in 2D IR spectra can be affected as well. For the amide I' mode of peptides, the anharmonic shift of $\approx 16 \text{ cm}^{-1}$ ³⁷ is comparable with the homogeneous line width of $5 - 10 \text{ cm}^{-1}$.^{37,38} One can determine both of these parameters by fitting the 2D IR spectrum, and then scale the 2D IR peak intensities accordingly. If the peaks are not scaled, then

$d(\omega)$ will be diminished. The approach we have used here is to use a calibrant molecule that has about the same anharmonic shift and line width as the molecule that we are studying. As a result, the ratio of Eq. (8) already accounts for the scaling.

It should be noted that this approach does not account for the change of amide I' anharmonicity or linewidth upon coupling of peptide units. The delocalization of amide I' vibration tends to decrease diagonal anharmonicity,³⁹ thus increasing the interference of positive and negative peaks in 2D IR spectrum and decreasing measured $d(\omega)$, while the homogeneous linewidth is also typically narrower, thereby decreasing the interference. One could scale the intensity of the coupled system to account for this effect, but since the two factors cancel to some extent, we have not done so here.

Cross peaks can also interfere with the diagonal peaks, which primarily occurs when the coupled modes are separated by less than the vibrational linewidth. Cross peaks are typically much smaller than diagonal peaks, but their contribution should not be forgotten in the analysis.

Another consideration when interpreting the data are the lineshapes, since $d(\omega)$ is frequency dependent. In the limit that the vibrational dynamics lead to perfectly homogeneous broadening, then the linear spectrum has a Lorentzian lineshape, which we define as $A(\omega)$. The corresponding absorptive 2D IR spectrum is Lorentzian along both coordinates so that its spectrum along the diagonal ($\omega_1 = \omega_3 = \omega$) is $A(\omega_1)A(\omega_3) = A^2(\omega)$. As a result, $d(\omega)$ will be Lorentzian. In the opposite extreme, when the inhomogeneous linewidth is much larger than the homogeneous linewidth, then both the linear and 2D IR spectra are Gaussian, assuming that the cumulant expansion is valid. In this situation, $d(\omega)$ is constant across the peak. When the frequency fluctuation correlation functions (FFCF) of the sample and calibrant are similar, the frequency dependant factor in Eq. (9) will be equal to 1 at the maximum absorbance of the sample. Here we assume that FFCFs of L-serine, NMA, random coil, and α -helical peptide are similar, which is confirmed by similarity of the lineshapes, consistency of $d(\omega)$ and $|\mu|^2$ of NMA, and previous theoretical results.^{40–42} In this case, $d(\omega)$ at the maximum of the peptide provides the value of its transition dipole strength.

Perhaps the most important consideration when evaluating $d(\omega)$ is identifying overlapping transitions. Consider a mixture of molecular structures, such as an equilibrium between an α -helix and a random coil polypeptide. The 1D and 2D spectra are given by the simple sum of the α -helix and random coil spectra, weighted by their prevalence in the solution. However, $d(\omega)$ spectrum cannot be obtained in such a simple way. For a sample composed of more than one molecular species, $d(\omega)$ has the following form:

$$d(\omega) = \frac{\sum_i c_i |\mu_i|^4 \Phi_i^{(3)}(\omega, \omega)}{\sum_i c_i |\mu_i|^2 \Phi_i^{(1)}(\omega)} \frac{\tilde{\Phi}^{(1)}(\omega^{\max})}{\tilde{\Phi}^{(3)}(\omega^{\max}, \omega^{\max})}, \quad (11)$$

where c is the concentration and the sum runs over all species which are labeled by i . From this equation it is clear that when two species are both present, then the value of $d(\omega)$ will depend on concentrations c_i at frequencies in which they overlap. If a strong absorber and a weak absorber are present in

equal concentrations, the measured $d(\omega)$ will be in between the two individual values. This has implications for the interpretation of $d(\omega)$ in the experiments on an α -helix and random coil, which we discuss below.

Another issue is the effects of non-specific solvent background absorbance. FTIR spectra typically have a baseline absorbance that varies with the solvent. D₂O is the most common solvent in peptide infrared studies because the bending mode of H₂O overlaps with the amide I band. Even though the bend of D₂O is 400 cm⁻¹ lower than the amide I', it still has a broad non-specific absorption in the region of the amide I' band. This non-specific absorption creates an offset in the linear spectrum. The offset is negligible in the 2D IR spectrum, because the non-specific absorbers have very weak transition dipoles even though the solvent concentration is very high. Nonetheless, the solvent transition dipoles should decrease the calculated $d(\omega)$ according to Eq. (11) since there are multiple overlapping transition dipoles. To avoid this problem, the experimental procedure contains a step in which the background is subtracted from the linear spectrum.³² Subtraction enables the use of Eq. (9), and provides accurate transition dipole strength as proven by the experiments above.

B. Insight into the AKA structural distribution

As noted above, the measured transition dipole strength of the AKA peptide at 60 °C is about the same as NMA and at 6 °C it is about 1.8 times that of NMA. Thus, upon formation of an α -helix, the vibrational exciton extends over neighboring residues in the peptide. However, the 1.8-fold increase is a lower limit because the α -helical and random coil spectra are overlapped. As discussed in Sec. V A, $d(\omega)$ provides the weighted average for overlapping transitions. Thus, the true transition dipole strength of the helix is larger.

To determine the transition dipole strength of the pure α -helix, we subtract off the random coil spectrum from the 6 °C and 23 °C spectra in a manner similar to subtracting off the solvent background. To do so, we assume that 100% of AKA at 60 °C is in random coil conformation, which is confirmed by the melting curve and previously published experiments using CD spectroscopy for the capped AKA.³ We then normalize the 60 °C absorption spectrum to the 6 °C and 23 °C spectra at frequencies above 1680 cm⁻¹. Since this region only contains random coil features (confirmed by comparison to capped AKA polypeptides at 1.6 °C³), the difference spectra is that of the pure helix (Fig. 6(b)). We note that the difference spectra are independent of temperature, which indicates that the helix-to-coil transition is two-state, as expected from the isosbestic point (Fig. 5(b)). The scaling factors for the difference spectra are 0.7 (23 °C) and 0.5 (6 °C), which is equal to the random coil content in the sample. These fractions were then used to perform a similar subtraction for the 2D IR spectra (Fig. 6(a)). Using the difference spectra we calculate $d_{hlx}(\omega)$ for AKA α -helix according to Eq. (8) (Fig. 6(c)). At the frequency of the helix $d_{hlx}(1635 \text{ cm}^{-1}) = 0.26 \pm 0.03 D^2$ at both 23 °C and 6 °C, which is (once again) consistent with a two-state

FIG. 6. Derived spectra for AKA α -helix. (a) Diagonal slice of the absorptive 2D IR spectrum at 23 °C (blue line) and 6 °C (black line). (b) Absorbance spectrum. (c) Calculated $d(\omega)$. Line assignment in (b) and (c) is the same as in (a).

equilibrium. Thus, the dipole strength of the A mode of the helix is actually ≈ 2.2 times larger than that of NMA. In fact, this value is probably larger, because the α -helix absorption spectrum consists of both the A and E modes, which are not resolved, and so lowers the measured value (Eq. (11)).

Our interpretation of the data rests on the assumption that the transition dipole moments are additive. Transition dipoles can exhibit non-linear character with respect to the vibrational coordinates. Such effects have been observed before in HOD,⁴³ silane,⁴⁴ and in coupled DNA bases.⁴⁵ Some vibrational modes in peptides, such as amide II and amide III modes, may also be non-linear.⁴⁶ If the extent of non-linearity is known, perhaps by calculations, then one could still backout the number of oscillators participating in the normal mode. If the non-linearity is not known, then the measured value would be an upper bound to the number of oscillators involved in the normal mode. Either way, an increase in the transition dipole strength assuredly indicates the formation of a structure with coupled oscillators. We believe that nonlinearity of the amide I' transition dipole moment is negligible. Assuming that amide I' transition dipole of individual peptide unit is oriented at 38° relative to the helix axis⁴⁷ we find that the exciton delocalization is greater than $n \approx 2.2 \cos^{-2}(38^\circ) = 3.5$ peptide units. This lower limit is about twice smaller than previously reported delocalization for some small peptides containing α -helical motif.³⁷

VI. CONCLUSIONS

The characteristic frequencies of molecules are not always unique, which makes their assignment in absorption

spectra ambiguous. This statement is especially true for the secondary structures of peptides as measured by the amide I' vibrational mode, which often have unresolved and congested IR spectra. The approach reported here provides a method for identification of secondary structures via excitonic delocalization. That is, if a transition dipole is measured that is larger than that of a single amide mode, then a well-ordered structure must exist that delocalizes that vibration. When correlated with frequency information, transition dipole strength provides a much more rigorous assessment of typical secondary structures. Neither absorption nor 2D IR spectroscopy alone has this capability. It is the ratio of the two that provides the transition dipole strengths independent of concentration. Our procedure is very simple to implement, since it only requires the measurement of a model compound prior to and/or after performing an experiment. We intend to use it to identify transient secondary structures in evolving protein kinetics experiments. In this application here, to a model soluble α -helix, the FTIR peak intensity increases by 40% upon helix formation. Our measurements reveal that this occurs because the transition dipole is at least 100% larger in the folded state.

ACKNOWLEDGMENTS

Support for this research was provided by the National Institutes of Health (Grant DK79895 to M.T.Z.).

- ¹S. Krimm and J. Bandekar, *Adv. Protein Chem.* **38**, 181 (1986).
- ²A. Barth, *Biochim. Biophys. Acta* **1767**, 1073 (2007).
- ³D. Gangani, R. A. Silva, J. Y. Nguyen, and S. M. Decatur, *Biochemistry* **41**, 15296 (2002).
- ⁴D. K. Graff, B. Pastrana-Rios, S. Y. Venyaminov, and F. G. Prendergast, *J. Am. Chem. Soc.* **119**, 11282 (1997).
- ⁵J. Manor, E. S. Feldblum, M. T. Zanni, and I. T. Arkin, *J. Phys. Chem. Lett.* **3**, 939 (2012).
- ⁶G. Martinez and G. Millhauser, *J. Struct. Biol.* **114**, 23 (1995).
- ⁷R. A. G. D. Silva, J. Kubelka, P. Bour, S. M. Decatur, and T. A. Keiderling, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8318 (2000).
- ⁸S. D. Zakharov, M. Lindeberg, Y. Griko, Z. Salamon, G. Tollin, F. G. Prendergast, and W. A. Cramer, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4282 (1998).
- ⁹D. R. Turner and J. Kubelka, *J. Phys. Chem. B* **111**, 1834 (2007).
- ¹⁰J.-H. Choi, S. Ham, and M. Cho, *J. Chem. Phys.* **117**, 6821 (2002).
- ¹¹H. S. Chung and A. Tokmakoff, *J. Phys. Chem. B* **110**, 2888 (2006).
- ¹²S. Hahn, S.-S. Kim, C. Lee, and M. Cho, *J. Chem. Phys.* **123**, 084905 (2005).
- ¹³A. S. Reddy, L. Wang, Y.-S. Lin, Y. Ling, M. Chopra, M. T. Zanni, J. L. Skinner, and J. J. De Pablo, *Biophys. J.* **98**, 443 (2010).
- ¹⁴S. Roy, J. Lessing, G. Meisl, Z. Ganim, A. Tokmakoff, J. Knoester, and T. L. C. Jansen, *J. Chem. Phys.* **135**, 234507 (2011).
- ¹⁵J. Wang and R. M. Hochstrasser, *Chem. Phys.* **297**, 195 (2004).
- ¹⁶P. Hamm and M. Zanni, *Concepts and Methods of 2D Infrared Spectroscopy* (Cambridge University Press, New York, 2011).
- ¹⁷P. W. Higgs, *Proc. R. Soc. London, Ser. A* **220**, 472 (1953).
- ¹⁸S. Krimm, *J. Mol. Biol.* **4**, 528 (1962).
- ¹⁹S.-H. Lee and S. Krimm, *Chem. Phys.* **230**, 277 (1998).
- ²⁰T. Miyazawa, *J. Chem. Phys.* **32**, 1647 (1960).
- ²¹T. Miyazawa and E. R. Blout, *J. Am. Chem. Soc.* **83**, 712 (1961).
- ²²D. Abramavicius, W. Zhuang, and S. Mukamel, *J. Phys. Chem. B* **108**, 18034 (2004).
- ²³T. C. Cheam and S. Krimm, *J. Chem. Phys.* **82**, 1631 (1985).
- ²⁴H. Torii and M. Tasumi, *J. Chem. Phys.* **96**, 3379 (1992).
- ²⁵N. Demirdoven, C. M. Cheatum, H. S. Chung, M. Khalil, J. Knoester, and A. Tokmakoff, *J. Am. Chem. Soc.* **126**, 7981 (2004).
- ²⁶H. Maekawa, C. Toniolo, Q. B. Broxterman, and N.-H. Ge, *J. Phys. Chem. B* **111**, 3222 (2007).

- ²⁷D. B. Strasfeld, Y. L. Ling, R. Gupta, D. P. Raleigh, and M. T. Zanni, *J. Phys. Chem. B* **113**, 15679 (2009).
- ²⁸A. T. Krummel and M. T. Zanni, *J. Phys. Chem. B* **112**, 1336 (2008).
- ²⁹D. E. Moilanen, D. Wong, D. E. Rosenfeld, E. E. Fenn, and M. D. Fayer, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 375 (2009).
- ³⁰S.-H. Shim and M. T. Zanni, *Phys. Chem. Chem. Phys.* **11**, 748 (2009).
- ³¹C. T. Middleton, A. M. Woys, S. S. Mukherjee, and M. T. Zanni, *Methods* **52**, 12 (2010).
- ³²See supplementary material at <http://dx.doi.org/10.1063/1.4764861> for the details of the derivation of $d(\omega)$, 2D IR and absorbance spectra of *N,N*-dimethylformamide and L-serine.
- ³³M. T. Zanni, N.-H. Ge, Y. S. Kim, and R. M. Hochstrasser, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11265 (2001).
- ³⁴L. Ackels, P. Stawski, K. E. Amunson, and J. Kubelka, *Vib. Spectrosc.* **50**, 2 (2009).
- ³⁵J. Kubelka and T. A. Keiderling, *J. Phys. Chem. A* **105**, 10922 (2001).
- ³⁶C. Fang, J. Wang, Y. S. Kim, A. K. Charnley, W. Barber-Armstrong, A. B. Smith, S. M. Decatur, and R. M. Hochstrasser, *J. Phys. Chem. B* **108**, 10415 (2004).
- ³⁷P. Hamm, M. Lim, and R. M. Hochstrasser, *J. Phys. Chem. B* **102**, 6123 (1998).
- ³⁸A. M. Woys, Y.-S. Lin, A. S. Reddy, W. Xiong, J. J. de Pablo, J. L. Skinner, and M. T. Zanni, *J. Am. Chem. Soc.* **132**, 2832 (2010).
- ³⁹J. Zhao and J. Wang, *J. Chem. Phys.* **136**, 214112 (2012).
- ⁴⁰S. Ham, S. Hahn, C. Lee, T.-K. Kim, K. Kwak, and M. Cho, *J. Phys. Chem. B* **108**, 9333 (2004).
- ⁴¹K. Kwac and M. Cho, *J. Chem. Phys.* **119**, 2247 (2003).
- ⁴²K. Kwac and M. Cho, *J. Chem. Phys.* **119**, 2256 (2003).
- ⁴³J. R. Fair, O. Votava, and D. J. Nesbitt, *J. Chem. Phys.* **108**, 72 (1998).
- ⁴⁴J.-L. Liao and Q.-S. Zhu, *Chem. Phys. Lett.* **274**, 328 (1997).
- ⁴⁵A. T. Krummel and M. T. Zanni, *J. Phys. Chem. B* **110**, 13991 (2006).
- ⁴⁶A. Moran and S. Mukamel, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 506 (2004).
- ⁴⁷D. Marsh, M. Müller, and F.-J. Schmitt, *Biophys. J.* **78**, 2499 (2000).