The Anomalous Infrared Amide I Intensity Distribution in ^{13}C Isotopically Labeled Peptide β-Sheets Comes from Extended, Multiple-Stranded Structures. An ab Initio Study

Jan Kubelka and Timothy A. Keiderling*

Contribution from the Department of Chemistry (M/C 111), University of Illinois at Chicago, 845 West Taylor Street, Room 4500 SES, Chicago, Illinois 60607-7061

Received January 30, 2001

Abstract: Ab initio based calculations of force fields and atomic polar tensors are used to simulate amide I infrared absorption spectra for a series of isotopically substituted $(Ac\text{-}A_{12}\text{-}NH\text{-}CH_3)_n$ peptides clustered in an antiparallel β-sheet conformation having a varying number of strands, $n = 2-5$. The results demonstrate that the anomalous intensity previously reported for the isotopically shifted amide I in ^{13}C labeled peptides is due to formation of multistranded β-sheet structures in this conformation. Computations show that the characteristic widely split amide I mode for β-sheet polypeptides as well as this anomalous intensity enhancement in isotopically substituted β-sheet peptides grows with increasing sheet size. For sheets of five strands, qualitative and near quantitative agreement with experimental amide I intensity patterns is obtained for both labeled and unlabeled peptides. The strongest transitions primarily represent in-phase coupled modes of the ^{13}C labeled, next nearest neighbor amides on the inner strands of the multistranded β-sheet. Long-range transition dipole coupling interactions do not promote the ^{13}C amide I intensity enhancement. Understanding of the IR intensity mechanisms with this level of detail for the isotopically labeled peptides permits design of site-specific probes of β-sheet folding and unfolding dynamics.

Introduction

The β-sheet conformation is the second most abundant secondary structure in globular proteins and is critical to the formation of many structural motifs that lead to specific protein functions. Despite this importance, β-sheet structure and formation have not until recently been studied as intensely as have those of the α-helix. The reasons for this have been manifold, but two fundamental impediments stand out.

First, well-defined peptide models for β-sheet formation that have a reasonable range of solubility have not been identified. The β-sheet structure, though classified as a secondary structure type, in fact involves tertiary or even quarternary interactions,[1,2] similar to those that lead to intermolecular aggregation. Controlling that interstand bonding to get a specific, unique structure has proven to be difficult, although several examples now exist of peptides whose properties indicate they form 2 and 3 strand hairpins or sheetlike segments, at least transiently, as part of dynamic equilibrium.[3−7]

Second, the structures of β-sheets found in proteins, as well as those made with peptide models, are highly variable. Beyond the central difference of parallel versus antiparallel orientation of the hydrogen-bonded strands, the sheets exhibit a wide distribution of strand lengths and various types and degrees of twists,[8] though predominately right-handed (which in turn makes the strands left-handed). Furthermore, all strands in proteins are distorted on their ends to form loop or turn structures.

In the past decade, since the highly aggregated β-sheet structures were identified as symptoms of amyloid and prion diseases,[9,10] the importance of understanding the dynamics and energetics of β-sheet formation has become more urgent. Thus development of techniques that have a high sensitivity to β-sheet formation and can sense structures in fluctuation between conformational states has important biomedical as well as biophysical applications.

Infrared (IR) spectroscopy is often used for detection of β-sheets[2,11,12] since their most intense amide I band (amide C=O stretch) component is sharply shifted down in frequency (to $\sim$1620−30 cm^{-1}) from typical helical and coil values, and a weak sideband appears at much higher values ($\sim$1680−90 cm^{-1}). While it has been common to attribute this lower frequency to increased H-bond strength, this work will show that it has another origin, in large part due to vibrational coupling of residues. Beyond this, the narrow, intense band seen in β-sheet polypeptides and some denatured proteins at 1610−1620 cm^{-1} is usually interpreted as being indicative of aggregated, antiparallel structure.[2]

* To whom correspondence should be addressed. Phone: (312) 996-2685. Fax: (312) 996-0431. E-mail: tak@uic.edu.

(1) Minor, D. L. J.; Kim, P. S. *Nature* **1994**, *371*, 264−267.
(2) Jackson, M.; Mantsch, H. H. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 95−120.
(3) Schenck, H. L.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4869−4870.
(4) Gellman, S. H. *Curr. Opin. Chem. Biol.* **1998**, *2*, 717−725.
(5) Sharman, G. J.; Searle, M. S. *J. Am. Chem. Soc.* **1998**, *120*, 5291−5300.
(6) Kortemme, T.; Ramirez-Alvarado, M.; Serrano, L. *Science* **1998**, *281*, 253−256.
(7) Ramirez-Alvarado, M.; Kortemme, T.; Blanco, F. J.; Serrano, L. *Bioorg. Med. Chem.* **1999**, *7*, 93−103.
(8) Salemme, F. R. *Prog. Biophys. Mol. Biol.* **1983**, *42*, 95−133.
(9) Harper, J. D.; Lansbury, P. T. *Annu. Rev. Biochem.* **1997**, *66*, 385−407.
(10) Harrison, P. M.; Bamborough, P.; Daggett, V.; Prusiner, S. B.; Cohen, F. E. *Curr. Opin. Struct. Biol.* **1997**, *7*, 53−59.
(11) Surewicz, W.; Mantsch, H. H.; Chapman, D. *Biochemistry* **1993**, *32*, 389−394.
(12) Haris, P. I.; Chapman, D. *Biopolymers* **1995**, *37*, 251−263.

10.1021/ja010270x CCC: $20.00
Published on Web 05/31/2001

Unfortunately, these frequency and intensity patterns correspond to global averages and do not yield any site resolved information. However, selective ^{13}C isotopic labeling[13−16] can add site-specific resolution to the vibrational spectra via both the frequency shift and intensity variation of the labeled vibrations. Strong, anomalous enhancement of the ^{13}C component of the amide I absorption has been observed for alternately doubly labeled antiparallel β-sheet peptides, while sequential or single labels give rise to weaker effects which none-the-less are easily detected and more intense than the statistical weight of labeled residues would suggest.[17,18] Isotopic intensity enhancement and frequency shifts have an important pragmatic effect: if well-understood, they can be used to probe site-specific structural aspects of a protein (or peptide) enriched with relatively few selected labels and yet provide a good experimental signal-to-noise ratio (S/N).

Brauner et al.[18] simulated the amide I spectra for their β-forming polypeptides, $K_2(LA)_6$, assuming formation of isotopically labeled pairs of β-strands and using an empirical coupled oscillator model that incorporated both through-bond and transition dipole coupling (TDC) between the C═O groups. With a single set of adjustable parameters representing the coupling constants, dipole moments, and unperturbed amide I frequencies, they obtained an excellent fit to the experimental data for several isotopomers. The anomalous enhancement in intensity for the labeled part was attributed by these authors[18] to a significant contribution of the unlabeled (^{12}C) amide groups coupled by the through-bond and TDC interactions to the ^{13}C labeled amide vibrations.

We previously demonstrated[16,19] that the amide IR absorption and VCD spectra for α-helical oligopeptides can be simulated with good accuracy using a transfer method of molecular property tensors obtained from an all atom, ab initio quantum mechanical calculation.[20] We here present simulations of IR absorption spectra for multistranded, large β-sheets (up to 65 amide groups) using this same approach. However, accurate β-sheet modeling is more complicated than that for an α-helix, first due to the lack of a reliable structural model. Second, if all important interactions are to be included, such modeling requires ab initio calculations on much larger peptide systems due to their multistranded nature. Simulations of large multistranded β-sheets presented here are based on fully quantum mechanical (density functional theory, DFT) property tensor calculations for "minimal" two- and three-stranded β-sheet peptides needed for the transfer, which, nevertheless, may be the largest peptide systems for which ab initio vibrational spectra have been computed.

Empirical treatments of β-sheet vibrational spectra, particularly by Krimm and co-workers,[21,22] suggested the importance of TDC between amide I modes. The quantum mechanical force field (FF) includes TDC, therefore it is implicitly present in our transferred FF for local interactions, i.e., within the range of the small fragment which is chosen as a source of parameters. However, the interactions between the parts of the large molecule beyond the size of this fragment are completely neglected. To assess the importance of such long-range couplings from a nonempirical basis, we test the effects of adding a TDC correction to the transferred FF[19] that is based on the same, small molecule ab initio calculation.

Methods

Experimental Details. Isotopically labeled samples of $K_2L^*A^*(LA)_5$ and $K_2LA^*LA^*(LA)_4$ prepared as C-terminal amides (an asterisk marks the residue that is ^{13}C labeled on the amide C═O) were kindly provided by Prof. Richard Mendelsohn. The unlabeled peptide $K_2(LA)_6$, C-terminal amide, was synthesized at the University of Illinois at Urbana-Champaign Protein facility. The latter was purified using HCl as an HPLC mobile phase to eliminate the trifluoroacetic acid (TFA) that forms an IR-interfering contaminant in most solid-phase synthesized peptides. The samples were deuterium preexchanged in an excess of deuterated ethanol overnight at room temperature, lyophilized, and redissolved in deuteriotrifluoroacetic acid (TFE-OD) at an approximate concentration of 10 mg/mL (∼7 mM) for spectral measurements. FTIR spectra were measured on a BioRad-Digilab FTS-60 in a semipermanent IR cell (Specac) consisting of CaF_2 windows and a Teflon spacer with path length of ∼100 μm. For each sample, the spectrum of the solvent (TFE-OD) was obtained first. Then the solvent was removed from the cell, which was subsequently refilled with the sample solution, to minimize effects of path length changes due to reassembling the cell. Absorbance spectra were calculated using the empty instrument spectrum as the reference, and solvent absorbance spectra were subtracted from the corresponding sample absorbance spectra. Additionally, the residual TFA signal for the $K_2L^*A^*(LA)_5$ and $K_2LA^*LA^*(LA)_4$ samples was eliminated by subtracting the spectrum of TFA in TFE-OD solution, prepared so that the absorbance of the interfering TFA peak matched that of the residual TFA peak in those samples. To check for concentration and solvent dependencies, the same set of experiments was carried out in deuterated methanol at a concentration of ∼1 mg/mL using a 200 μm path length.

Calculations. Two-, three-, four-, and five-stranded antiparallel β-sheet peptides, with a nominal strand sequence of Ac-A_{12}-NH-CH_3 (A = Ala), were constructed using coordinates from the crystal structure of the β-sheet form of poly-L-alanine.[23] Spectral simulations for the unlabeled and three isotopically labeled species Ac-A_{12}-NH-CH_3, Ac-AA*A*A_9-NH-CH_3, and Ac-AAA*AA*A_7-NH-CH_3 of each multistrand assembly (2, 3, 4, and 5 strands) were used as models for experimentally measured $K_2(LA)_6$, $K_2L^*A^*(LA)_5$, and $K_2LA^*LA^*(LA)_4$, respectively. Spectra for other isotopomers were also simulated as auxiliary tests of the results. Amide I IR absorption spectra were simulated by transfer of the ab initio force fields (FF) and atomic polar tensors (APT) in Cartesian coordinates[20] calculated at the DFT BPW91/6-31G** level. These high-level calculations were for smaller, two- and three-stranded antiparallel β-sheet peptides, based on Ac-A-A-NH-CH_3 strands with local (ϕ, ψ) geometries identical to those used to construct the target, large polypeptides. The two-stranded smaller peptide FF and APT tensor values were transferred onto the two-stranded long β-pair. The three-, four-, and five-stranded large β-sheet FF and APTs were transferred from the three-stranded smaller peptide. In these cases, the middle strand atomic tensors were assigned to atoms on the inner strands of the large sheet structure, and those from the outer strands were transferred to the outer strand residues. Quantum

(13) Tadesse, L.; Nazarbaghi, R.; Walters, L. *J. Am. Chem. Soc.* **1991**, *113*, 7036−7037.
(14) Fabian, H.; Chapman, D.; Mantsch, H. H. In *Infrared Spectroscopy of Biomolecules*; Mantsch, H. H., Chapman, D., Eds.; Wiley-Liss: Chichester, 1996; pp 341−352.
(15) Decatur, S. M.; Antonic, J. *J. Am. Chem. Soc.* **1999**, *121*, 11914−11915.
(16) Silva, R. A. G. D.; Kubelka, J.; Decatur, S. M.; Bour, P.; Keiderling, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8318−8323.
(17) Halverson, K.; Sucholeiki, I.; Ashburn, T. T.; Lansbury, R. T. *J. Am. Chem. Soc.* **1991**, *113*, 6701−6703.
(18) Brauner, J. W.; Dugan, C.; Mendelsohn, R. *J. Am. Chem. Soc.* **2000**, *122*, 677−683.
(19) Kubelka, J.; Silva, R. A. G. D.; Bour, P.; Decatur, S. M.; Keiderling, T. A. In *The Physical Chemistry of Chirality*, ACS Symp. Ser.; Hicks, J. M., Ed.; Oxford University Press: New York, 2001, in press.
(20) Bour, P.; Sopkova, J.; Bednarova, L.; Malon, P.; Keiderling, T. A. *J. Comput. Chem.* **1997**, *18*, 646−659.
(21) Krimm, S.; Abe, Y. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 2788−2792.
(22) Moore, W. H.; Krimm, S. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 4933−4935.
(23) Fraser, R. D. B.; McRae, T. P. *Conformation in Fibrous Proteins*; Academic Press: New York, 1973.

chemical calculations were performed using Gaussian 98.[24] Analytical harmonic FF and APT were computed for partially optimized geometries, where main-chain torsional angles were constrained to the experimental values.[23]

After transfer of the parameters[19,25] to the longer β-sheet models, calculations of projected vibrational frequencies, dipolar (D) strengths, and isotopic substitutions were performed using programs written in-house. To remove the spectral interference from the C−H deformation modes of the rigid (fixed rotamer) $-CH_3$ groups, deuterium was substituted for all methyl hydrogens. Even though all simulations are presented for N-protonated peptides, to test for deuteration effects several N−D substituted calculations were also carried out. Spectral band shapes for simulated spectra were obtained by assigning Lorentzian line-shapes to all transitions with a uniform width of 8 cm^{-1} and the area proportional to the dipole strength D.

Since the force constants are not known for the large molecules with respect to atom pairs beyond those encompassed in the small "fragment", these are normally set to zero. To test for any impact of these long-range interactions on the spectra, we approximated them by TDC, computed from ab initio APT values:[19]

$$F^{\mathrm{TDC}}_{\mathrm{A}\lambda,\mathrm{B}\xi} = \frac{\partial U^{\mathrm{TDC}}}{\partial q^{\mathrm{A}}_{\lambda}\,\partial q^{\mathrm{B}}_{\xi}} = \frac{1}{\epsilon R^3}\left[\sum_{i=x,y,z}\frac{\partial \mu^{\mathrm{A}}_i}{\partial q^{\mathrm{A}}_{\lambda}}\frac{\partial \mu^{\mathrm{B}}_i}{\partial q^{\mathrm{B}}_{\xi}} - 3\sum_{i,j=x,y,z}\left(\frac{\partial \mu^{\mathrm{A}}_i}{\partial q^{\mathrm{A}}_{\lambda}}e_i\right)\left(\frac{\partial \mu^{\mathrm{B}}_j}{\partial q^{\mathrm{B}}_{\xi}}e_j\right)\right] \quad (1)$$

for all pairs of atoms (A, B) whose force constants are absent in the transferred FF. Here q^{A}_{λ}, $\lambda = \{x,y,z\}$ are Cartesian displacements of the atom A, ϵ is a dielectric constant, which was chosen as 1 (maximal effect), R is the distance between atoms A and B, e_j are the components of a unit vector from atom A to B, and $\partial\mu_i/\partial q_{\lambda}$ are the components of the APT.

Results

Experimental Spectra. Measured amide I′ FTIR spectra for the $K_2(LA)_6$, $K_2L^*A^*(LA)_5$, and $K_2LA^*LA^*(LA)_4$ peptides in TFE-OD are shown in Figure 1 and are in full agreement with previously published results.[18] Identical IR spectra were also obtained in deuterated methanol at ∼10 times lower concentration. The unlabeled $K_2(LA)_6$ peptide gives rise to an amide I′ band shape reflecting that expected for a β-sheet, with an intense band at ∼1624 cm^{-1} and a weak one at ∼1689 cm^{-1}. In the sequentially ($K_2L^*A^*(LA)_5$) and alternately ($K_2LA^*LA^*(LA)_4$) labeled samples, the ^{12}C contribution shifts up in wavenumber to ∼1628 and ∼1634 cm^{-1} with the ^{13}C band appearing at ∼1606 and ∼1608 cm^{-1}, respectively. The relative intensity of the ^{13}C band is ∼26% of the total for the sequentially labeled species, while for the alternately labeled ones it is anomalously enhanced to ∼40%. Both are much more intense than would have been expected for only 2 of 14 residues being labeled. The weak band at ∼1550 cm^{-1} is a residual amide II absorption that suggests incomplete H/D exchange.

Unlabeled (Ac-A_{12}-NH-CH_3) Peptide Simulated Spectra. Calculated amide I IR spectra for an unlabeled Ac-A_{12}-NH-CH_3 β-strand peptide forming 2-, 3-, 4-, and 5-stranded sheets (referred to as **S2-5**, respectively) are shown in Figure 2. The **S2** simulation predicts an amide I band shape similar to the characteristic β-sheet amide I absorption pattern,[12,26] with the low-frequency main maximum and much weaker high-frequency maximum. However, these peaks are only ∼37 cm^{-1} apart as opposed to the ∼65 cm^{-1} seen experimentally, and the high-frequency peak is too intense compared to the main maximum. With an increasing number of strands, the simulated low-frequency peak shifts down in frequency, such that the splitting increases (up to ∼57 cm^{-1} in **S5**), and becomes more intense compared to the high-frequency peak, so that the overall band shape converges to the experimental pattern. In **S3**, **S4**, and **S5** a third maximum appears between the other two, but its relative intensity decreases with the sheet size.

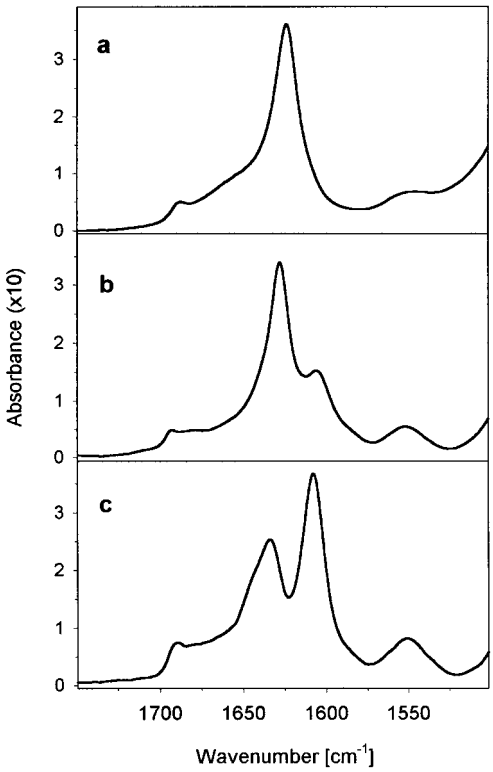

Figure 1. Experimental FTIR spectra of the β-sheet peptides: (a) unlabeled $K_2(LA)_6$, (b) sequentially doubly labeled $K_2L^*A^*(LA)_5$, and (c) alternately doubly labeled $K_2LA^*LA^*(LA)_4$.

The most intense amide I transitions arise from modes where each particular amide C═O vibrates out-of phase with those sequentially (thus in-phase alternately) on the same strand, and in-phase with the C═O from the amide hydrogen bonded to it on the other strands. This results in a strong net transition dipole oriented approximately perpendicular to the β-sheet axis and corresponds to the $\nu_{\perp}(\pi, 0)$ modes in Miyazawa's original notation for infinite, symmetric β-sheet models.[23,27] Even though Miyazawa's model is not rigorously applicable to our case of finite β-sheets, the outlined basic phase relationships can be clearly identified, and therefore this simple notation is used here for approximate characterization of the normal modes. This relative phasing of C═O stretches is illustrated in Figure 3 for selected modes of **S2** and in Figure 4 for **S5**. These schematic plots of the C═O stretching amplitudes are for the modes that most significantly contribute to the amide I band shape. In Figure 3 the high frequency (1711.4 cm^{-1}) mode is the strongest within

(24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.6; Gaussian, Inc.: Pittsburgh, PA, 1998.

(25) Bour, P.; Kubelka, J.; Keiderling, T. A. *Biopolymers* **2000**, *53*, 380−395.

(26) Jackson, M.; Haris, P. I.; Chapman, D. *Biochim. Biophys. Acta* **1989**, *998*, 75−79.

(27) Miyazawa, T. *J. Chem. Phys.* **1960**, *32*, 1647−1652.

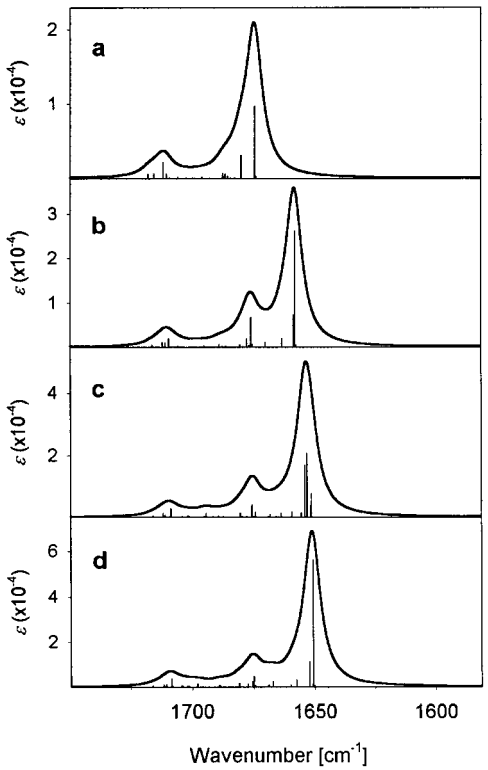

Figure 2. Simulated amide I IR for the unlabeled Ac-A_{12}-NH-CH_3 β-strand peptide in (a) 2-stranded (**S2**), (b) 3-stranded (**S3**), (c) 4-stranded (**S4**), and (d) 5-stranded (**S5**) antiparallel β-sheet conformation.

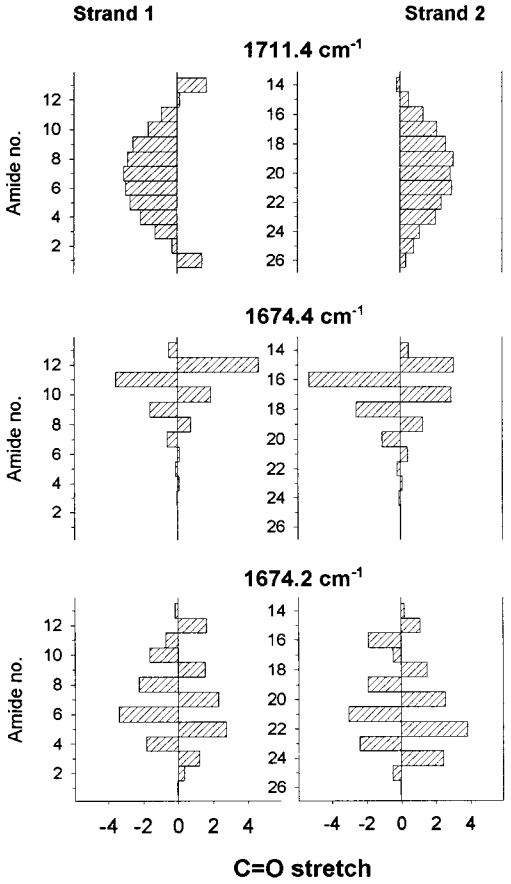

Figure 3. Plot of relative amide C═O bond stretching amplitudes in the 2-stranded antiparallel β-sheet (**S2**) for 1711.4 (top), 1674.4 (middle), and 1674.2 (bottom) cm^{-1} amide I normal modes. The 1711.4 cm^{-1} is the strongest mode near the high-frequency secondary maximum, the 1674.4 and 1674.2 cm^{-1} are nearly degenerate, most intense modes are at the main amide I band maximum. Positive values correspond to bond stretching, negative to bond contraction.

the secondary, high-frequency maximum, and the two, nearly degenerate, low-frequency modes (~1674 cm^{-1}) dominate the main amide I maximum. In Figure 4, the **S5** modes dominating the secondary maximum (1708.5 cm^{-1}) and the main maximum (1650.7 cm^{-1}) are shown, along with the strongest mode contributing to the third, central maximum (1674.8 cm^{-1}). The $\nu_\perp(0, \pi)$ character of the mode corresponding to the high-frequency (near 1710 cm^{-1}) maximum in both the **S2** and **S5** can be clearly seen.

Although it is not evident from Figures 3 and 4, due to their greatly simplified picture of the normal modes (effectively one-dimensional), the predicted transition dipoles of the high-frequency vibrational modes are weak but polarized approximately along the direction of the strands, as would be expected for $\nu_{II}(0, \pi)$. On the other hand, the two most intense normal modes, contributing to the main maximum, predicted ~1674 cm^{-1} in **S2** (Figure 3) and ~1650 cm^{-1} in **S5** (Figure 4), are clearly predominantly $\nu_\perp(\pi, 0)$ modes. In **S5**, the most intense normal modes near the main maximum are more localized on the inner strands of the sheet with the greatest amplitudes on the middle strand. The modes corresponding to the other two maxima (1674.2 and 1708.5 cm^{-1} in **S5**, Figure 4) show larger amplitudes for the outer strand amides. A similar C═O amplitude distribution can be seen in **S3** and **S4**. The modes around the central maximum computed at ~1675 cm^{-1} for **S3**, **S4**, and **S5** have complicated phase relationships, with contributions from both $\nu_\perp(\pi, \pi)$ and $\nu_\perp(\pi, 0)$ phases. Finally, a number of modes with little intensity, most of which are not seen in the simulated spectra plots, correspond, at least locally, to the $\nu(0, 0)$ combination. The central maximum, which is not experimentally observed in these model β-sheet peptides, is predicted to lose intensity for larger sheets. These central modes may not be apparent in multiple-stranded β-sheet polypeptides, but may be significant in proteins which have shorter, more distorted β-sheets.

The above qualitative description of the amide I normal mode patterns generally holds for the vibrations of the unlabeled groups in all the isotopically labeled peptides we have simulated. In particular, the main maximum (later referred to here as the main ^{12}C or unlabeled maximum) always arises due to $\nu_\perp(\pi, 0)$ type vibrations primarily on strands near the center of the sheet. Therefore, in the following we shall focus mainly on the isotopic labeling effects.

Sequentially Doubly Labeled (Ac-AA*A*A_9-NH-CH_3) Simulated Spectra. Simulated amide I IR spectra for the doubly labeled peptides, Ac-AA*A*A_9-NH-CH_3, in antiparallel sheets of 2, 3, 4 or 5 strands (denoted $\mathbf{S_{II}2-5}$, respectively) are shown in Figure 5. Compared to the unlabeled peptides**,** the main ^{12}C maximum is higher in frequency, and has a broadened shape. In $\mathbf{S_{II}2}$ the ^{13}C peak arises from two nearly degenerate normal modes, each corresponding to the out-of-phase vibrations ($\nu_\perp(\pi, 0)$) of neighboring labeled amide groups on each strand. In $\mathbf{S_{II}3}$ two ^{13}C maxima correspond to the out-of-phase vibrations of the labeled pair on the middle strand (lower frequency) and both labeled pairs on the outer-strand (higher frequency). In $\mathbf{S_{II}4}$ and $\mathbf{S_{II}5}$ the mode ordering has changed, as there are only seven (out of eight possible) and nine (out of ten), respectively, localized^{13}C modes below the main maximum. The remaining ^{13}C mode is "distributed", i.e., there are

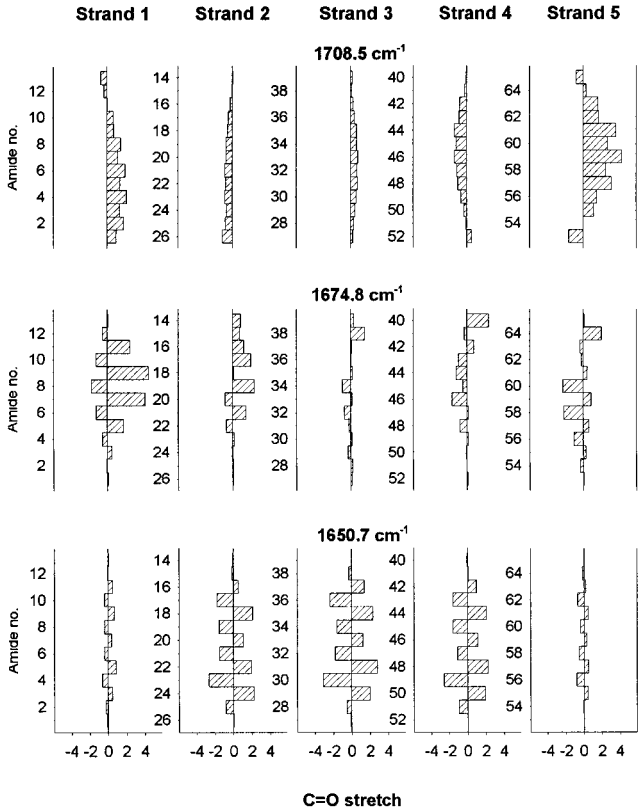

Figure 4. Plot of relative amide C═O bond stretching amplitudes in the 5-stranded antiparallel β-sheet (**S5**) for 1708.5 (top), 1674.8 (middle), and 1650.7 (bottom) cm^{-1} normal modes, which correspond to the strongest modes near the high-frequency maximum, the central maximum, and the low-frequency main amide I maximum, respectively. Positive values correspond to bond stretching, negative to bond contraction.

vibrations of one or more ^{13}C groups involved in a number of high-frequency ^{12}C transitions. The amide I band shapes for $\mathbf{S_{II}5}$ and $\mathbf{S_{II}4}$ are qualitatively very similar. The maximum of the ^{13}C component consists of the nearly degenerate, out-of-phase motions of the labeled neighboring groups on the inner strands with amides having the same orientation being in-phase.

The $\mathbf{S_{II}4}$ and $\mathbf{S_{II}5}$ simulations are closer to the experimentally observed band shapes than are those for $\mathbf{S_{II}2}$ or $\mathbf{S_{II}3}$. The relative intensity of the ^{13}C band is in both cases about 23% of the total and the splitting between the ^{12}C and ^{13}C maximum decreases from 35 cm^{-1} in $\mathbf{S_{II}4}$ to 33 cm^{-1} in $\mathbf{S_{II}5}$. The former is already very close to the experiment (26%), while the latter overestimates the experimental value of 23 cm^{-1}. The full dispersion of the amide I (highest ^{12}C to lowest ^{13}C) is $\sim$90 cm^{-1}, very close to the experimental spread from 1693 to 1606 cm^{-1}.

Alternately Doubly Labeled Ac-AA*AA*A_8-NH-CH_3 Simulated Spectra. The simulated amide I IR spectra for the alternate doubly labeled peptides, Ac-AA*AA*A_8-NH-CH_3, in β-sheets of 2, 3, 4, and 5 strands (labeled $\mathbf{S_{I-I}2-5}$, respectively) are shown in Figure 6. The enhancement of the ^{13}C band with respect to the population of the label is apparent in all cases. In addition, the broader, less intense character of the ^{12}C main band is also evident in the simulations. Both the main ^{12}C peak and the ^{13}C peak are shifted to higher frequency compared to sequentially labeled derivatives.

In $\mathbf{S_{I-I}2}$ the ^{13}C band is well-resolved from the unlabeled maximum and is broad due to frequency differences between the hydrogen-bonded (inward pointing) and non-hydrogen-bonded vibrations of the labeled pair. In $\mathbf{S_{I-I}3}$ the ^{13}C amide I

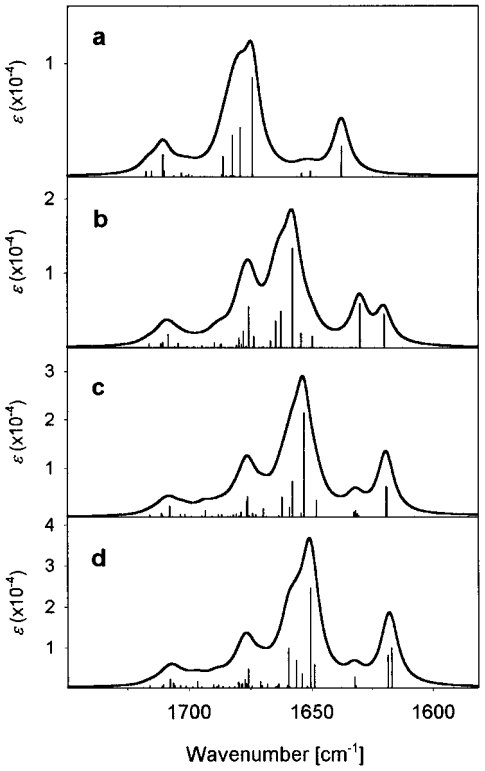

Figure 5. Simulated amide I IR for the sequentially doubly labeled Ac-AA*A*A_9-NH-CH_3 β-strand peptide in (a) 2-stranded ($\mathbf{S_{II}2}$), (b) 3-stranded ($\mathbf{S_{II}3}$), (c) 4-stranded ($\mathbf{S_{II}4}$), and (d) 5-stranded ($\mathbf{S_{II}5}$) antiparallel β-sheet conformation.

sideband is again broad consisting of two intense modes. The most intense ^{13}C component is also the most intense transition in the whole spectrum. In $\mathbf{S_{I-I}4}$ the amide I band shape becomes qualitatively different compared to that in $\mathbf{S_{I-I}3}$. The ^{13}C band results in a narrow peak with the highest intensity in the entire amide I band. Nearly all of this ^{13}C band intensity arises from a single mode with the largest contribution from the labeled groups on the two inner strands. The next strongest, highest frequency ^{13}C mode from the outer-strand labeled amides is so close in frequency to the ^{12}C modes that it contributes to the apparent intensity of the main ^{12}C band. Strong "anomalous" enhancement of the ^{13}C amide I sideband is clearly predicted from this ab initio simulation. The simulated amide I for $\mathbf{S_{I-I}5}$, qualitatively similar to that of $\mathbf{S_{I-I}4}$, has an even more intense ^{13}C band. Furthermore, its low-frequency ^{12}C transitions are stronger than those in the middle of the amide I contour, resulting in a more uniform band shape in closer agreement with experiment.

Schematic representations of the two most intense ^{12}C normal modes (1660.4 and 1652.3 cm^{-1}) and the two most intense ^{13}C normal modes (1648.2 and 1625.2 cm^{-1}) are shown in Figure 7. The 1648.2 cm^{-1} mode is within the band envelope of the main ^{12}C maximum, while the most intense amide I normal mode, 1625.2 cm^{-1}, is responsible for the anomalously strong intensity of the ^{13}C isotopic sideband. This 1625 cm^{-1} mode is dominated by in-phase vibrations of the ^{13}C labeled groups, which in this model are all pointing in the same direction. The labeled pair on the center strand has the largest amplitude. There are contributions from the unlabeled groups, but they are relatively small for the lowest energy modes. The phase pattern involving all groups, i.e., ^{12}C as well as ^{13}C labeled ones, corresponds to the intense $\nu_\perp(\pi, 0)$ mode. In this way the ^{12}C amide vibrations, even though much weaker in amplitude, have

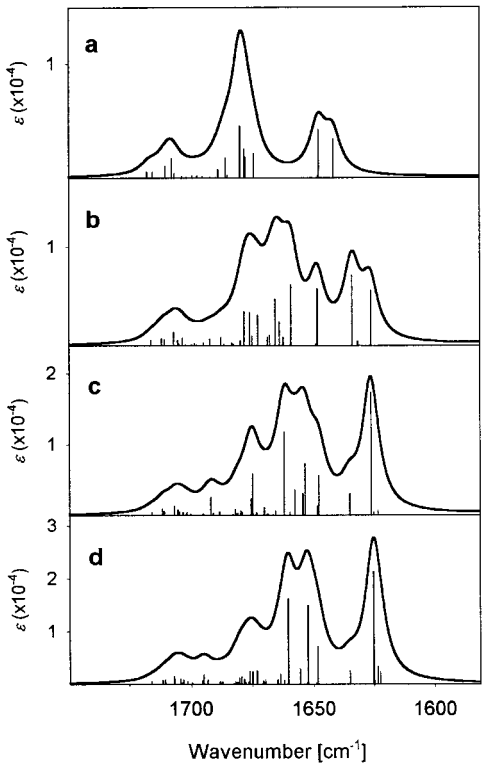

Figure 6. Simulated amide I IR for the alternately doubly labeled Ac-AA*AA*A_8-NH-CH_3 β-strand peptide in (a) 2-stranded (**S_{I-I}2**), (b) 3-stranded (**S_{I-I}3**), (c) 4-stranded (**S_{I-I}4**), and (d) 5-stranded (**S_{I-I}5**) antiparallel β-sheet conformation.

a relative phase that adds to the overall transition dipole moment. The 1648.2 cm^{-1} mode has a similar phase pattern but is dominated by the outer-strand amide vibrations. The two intense high-frequency modes have predominantly $\nu_{\perp}(\pi, 0)$ character and their largest contributions come from ^{12}C motions on the center strands of the sheet. Minor contributions of the ^{13}C groups can be seen for the higher frequency, 1660.4 cm^{-1} normal mode, often with a phase shift. The 1652 cm^{-1} mode has major intensity contributions from the strand ends rather than the centers of the three central strands.

The agreement between theory and experiment again improves as more β-strands are included in the β-sheet model. In this case, the dramatically enhanced intensity of the ^{13}C labeled peak is still underestimated: about 28% and 30% in **S_{I-I}4** and **S_{I-I}5**, respectively, compared to more than 40% in the experimental spectrum. Furthermore, the splitting between ^{12}C and ^{13}C maxima is slightly overestimated ($\sim$28 cm^{-1} in both **S_{I-I}4** and **S_{I-I}5** compared to an experimental value of 26 cm^{-1}). Also the ^{12}C intensity distribution is somewhat more complex than seen experimentally.

Transition Dipole Coupling. Comparison of the amide I band shape for the **S_{I-I}4** peptide calculated with the addition of a TDC correction (solid line) as compared to the normal transfer result (dashed line) is shown in Figure 8. The long-range TDC correction does not significantly change the amide I band shapes or the normal mode distributions. The TDC correction, in fact, leads to a slightly worse underestimation of the ^{13}C labeled band intensity than does the noncorrected FF. Therefore, the long-range TDC interactions do not seem to be significant and, specifically, do not improve the β-sheet amide I band shape predictions.

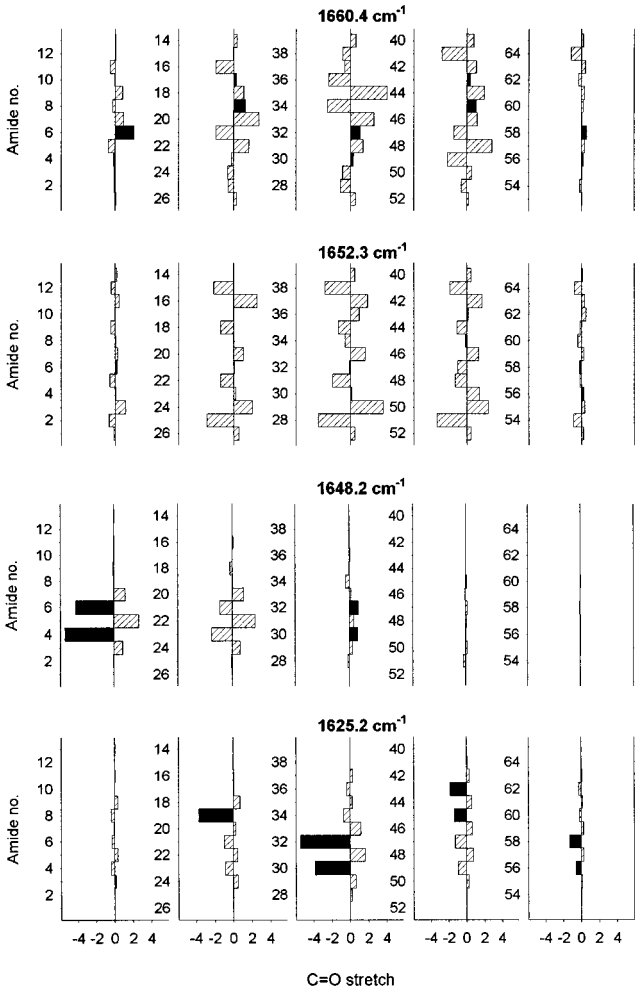

Figure 7. Plot of relative amide C═O bond stretching amplitudes in the alternately doubly ^{13}C labeled, 5-stranded antiparallel β-sheet (**S_{I-I}5**) for (from the top) 1660.4, 1652.3, 1648.2, and 1625.2 cm^{-1} amide I normal modes. Black bars represent the ^{13}C labeled residues. Positive values correspond to bond stretching, negative to bond contraction. The first two modes correspond to the two strongest ^{12}C transitions, the last two to the two strongest ^{13}C transitions. The 1648.2 cm^{-1} mode appears in the main ^{12}C band envelope, and 1625.2 cm^{-1} is the most intense amide I mode in **S_{I-I}5**.

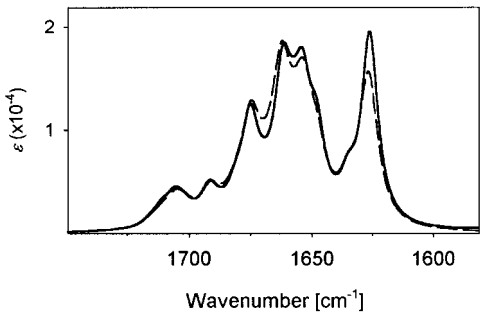

Figure 8. Comparison of simulated amide I IR with TDC correction (dashed line) and uncoorected (solid line) for the alternately doubly labeled Ac-AA*AA*A_8-NH-CH_3 peptide in the 4-stranded β-sheet (**S_{I-I}4**).

Discussion

Our simulated band shapes are in good agreement with experimental observations, considering that very little is known about the detailed structure of the model peptides studied experimentally. We note that peptides were N-protonated in our

simulations, while the experiments were performed in deuterated solvent, and the samples were preexchanged, and therefore (at least mostly) deuterated. Test simulations showed only minor differences between amide I band shapes for N-protonated and N-deuterated peptides. These included small frequency shifts of most components in the band and some small intensity variations. By no means would these alter the agreement reported above with the experimental results. Although the agreement of our simulations with experiment is not as precise as that of Brauner and co-workers,[18] our approach used only the results of ab initio calculations which had no adjustable empirical parameters beyond the choice of the method. Our agreement with experimental band shapes would certainly have been better had we scaled the calculated frequencies or the force field, or even used different line-widths for spectral components in our simulations. However, such an approach contains little additional physical insight. The limitations of our FF are clear, with the vibrational frequencies being too high compared to experiment, even though compared to our previous ab initio calculations on the helical peptides they are surprisingly close.[16,19,25] Unfortunately, overestimation of amide I frequencies seems to be a common pitfall of ab initio DFT calculations that can correlate to a short C═O bond length. This specific aspect can be modified by alteration of the basis set (Kubelka and Keiderling: to be submitted for publication). Aside from this absolute error, the relative frequencies within the amide I and the corresponding normal mode ordering can still give a reliable, useful picture for interpreting the IR spectra.

Polypeptide β-Sheets Consist of Extended, Multiple-Stranded Structures. This ab initio based computational study demonstrates that the observed "characteristic" β-sheet band shapes, as seen in oligo- and polypeptide β-sheet models, actually arise from extended, multistranded β-sheets. In the unlabeled peptides both the frequency splitting between the low- and high-frequency peaks and their relative intensities in the simulated amide I converge to experimental values with an increase in the number of strands used in the simulation. While the two-strand simulation has approximately the correct shape, the frequency splitting is too small and the high-frequency peak too intense. The correlation between the amide I component splitting and increasing size of the β-sheet is in agreement with the study of Chirgadze and Nevskaya,[28] who used a set of dipole-coupled oscillators having a β-sheet geometry to empirically simulate the amide I spectra. The necessity of a multiple-stranded β-sheet model to explain the experimental band shapes is most clearly demonstrated in the alternately doubly labeled peptide simulations. The anomalous intensity enhancement of the ^{13}C-labeled vibrations (whereby the ^{13}C mode dominates the amide I intensity) qualitatively only approaches that found experimentally for the four- and five-stranded β-sheet models in our computational tests and cannot be seen in the simple two-stranded model. In our hands, these peptides failed to dissolve at these concentrations in TFE unless sonicated for several hours, and then resulted in gel-like solution implying more extended structure had formed. The presence of a small, residual amide II absorption at $\sim$1550 cm^{-1} for these peptides (dissolved in TFE-OD) also suggests multiple strand formation, which could protect inner-strand amide groups from deuterium exchange.

There seems to be no reason to limit association for these $K_2(LA)_6$ peptides to only two strands. If strands were to align in an exact antiparallel fashion, the leucine side chains on neighboring strands would point alternately above and below the β-sheet plane, thus causing no packing hindrance. The same arrangement would occur if the strands were offset by two residues, so the two Lys residues at the end terminus were not paired with the neighboring strands, leading to improved solvation of charged residues. If two strands associate in this fashion, three or more could just as easily associate in the same fashion, resulting in multistranded structures of varying sizes. While Brauner et al.[18] report no solubility problems in methanol at somewhat lower concentrations, the spectra they obtained are virtually identical with ours, implying the same basic peptide structure occurs under both conditions. This observation is consistent with the fast convergence of the simulated amide I band shapes with respect to the number of associated β-strands. While significant changes in the simulated amide I band shapes occur between the two-, three-, and four-stranded β-sheets, addition of the fifth strand yields an amide I band shape qualitatively similar to the four-stranded as well as to the experimental band shapes. Therefore we do not expect significant qualitative simulated band shape changes to occur upon addition of more strands to the structure. Thus while our model assumes only a small number of strands to explain the observed spectral patterns, large aggregates could yield similar results.

Historically, IR, Raman,and CD spectra of the β-sheet form of poly-L-lysine (at high pH and $\sim$60 °C), which has an amide I IR pattern very similar to that of $K_2(LA)_6$, has been taken as a "standard" for β-sheets.[12,29,30] However, soluble globular proteins rarely absorb below 1630 cm^{-1} and have continuous absorption through the amide I region, normally attributed to their heterogeneous structures.[11] It has been suggested that the broadly "split" peptide β-sheet amide I IR band shape is due to aggregated, extended strands, stabilized by strong intermolecular hydrogen bonds.[2,26] This is also consistent with the same band shapes being characteristic for some thermally denatured, aggregated proteins.[31] A different spectral amide I pattern is expected for short strands, such as found in globular proteins as well as for model β-hairpins and 3-strand sheet models, now being synthesized by a number of researchers.[4,32−34] The length and structure effects on spectra can explain in part why β-sheets in proteins give rise to different amide I IR frequencies and band shapes than do β-sheet polypeptides. It further calls into question the use of assigned frequency components in the IR for structural analyses.[2,35,36] Specifically our spectral simulations predict that for distorted or less extensive β-sheets there will be a contribution to the amide I IR absorption between the extrema of the exciton splitting that directly overlaps frequencies characteristic of helices and random coils.

On the basis of our results it seems likely that the model peptides studied here adopt a conformation similar to that of the extended β-strands, represented by crystalline poly-L-alanine.[23] This is consistent with the observation that the larger the β-sheet becomes, the less it generally twists.[8] Our spectral

(28) Chirgadze, Y. N.; Nevskaya, N. A. *Biopolymers* **1976**, *15*, 607−625.
(29) Elliott, A.; Ambrose, E. J. *Nature* **1950**, *165*, 921−922.
(30) Fraser, R. D. B.; Suzuki, E. *J. Mol. Biol.* **1965**, *14*, 279−293.
(31) Van Stokkum, I. H. M.; Linsdell, H.; Hadden, J. M.; Haris, P. I.; Chapman, D.; Bloemendal, M. *Biochemistry* **1995**, *34*, 10508−10518.
(32) Nesloney, C. L.; Kelly, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 5836−5845.
(33) Griffiths-Jones, S. R.; Maynard, A. J.; Searle, M. S. *J. Mol. Biol.* **1999**, *292*, 1051−1069.
(34) Silva, R. A. G. D.; Sherman, S. A.; Perini, F.; Bedows, E.; Keiderling, T. A. *J. Am. Chem. Soc.* **2000**, *122*, 8623−8630.
(35) Pancoska, P.; Wang, L.; Keiderling, T. A. *Protein Sci.* **1993**, *2*, 411−419.
(36) Haris, P. I. In *Infrared Analysis of Peptides and Proteins: Principles and Applications*, ACS Symp. Ser.; Ram Singh, B., Ed.; American Chemical Society: Washington, DC, 2000, pp 54−95.

simulations of twisted β-sheet peptides, and sheets with shorter strands such as those found in proteins, show that their amide I band shape is sensitive to the degree of twist. In particular, for twisted sheets the most intense ^{12}C transition is significantly higher in frequency and the full exciton splitting less than for the extended models discussed here (Kubelka and Keiderling: to be submitted for publication). The ^{13}C band also shifts higher in frequency, but not so much. Therefore, to get the large splitting "characteristic" of the polypeptide β-sheets and correct relative frequencies in the ^{13}C-labeled spectra one needs multiple-strand ensembles of relatively regular, extended conformations. In such a 2-fold symmetric chain geometry, maximal exciton splitting of the amide I mode results as reflected in our simulations.

Isotopic Frequency Shifts. The low-frequency shift of the ^{13}C-labeled sideband in isotopically labeled β-sheet peptides from the unlabeled maximum ($\sim$20 cm^{-1}) is smaller than would have been originally expected ($\sim$40 cm^{-1}) based on masses of C isotopes. However, as our simulations point out, the true characteristic amide I frequency of the β-sheet peptide is not this low-frequency amide I maximum $\sim$1620$-$1630 cm^{-1}, but rather $\sim$1640$-$1650 cm^{-1} (approximate mean of exciton split components). The low frequency for β-sheets, which is often assumed to be due to increased hydrogen bond strength, is instead due to much stronger exciton splitting in the extended chain. This is further enhanced by interstrand coupling resulting in a split of the amide I by 50$-$70 cm^{-1}, while for example the α-helix amide I split is only $\sim$5 cm^{-1}. The ^{13}C labeled modes, as we have also shown in our simulations, are also distributed over a range of frequencies (e.g. 1621$-$1648 cm^{-1} in $\mathbf{S_{I-I}5}$) that is narrower than that for the ^{12}C, since there are only two labeled amides per strand. The ^{13}C modes still have the characteristic β-structure intensity distribution, with the main maximum at the low frequency and a weaker one at higher frequency. The positions of both the ^{12}C and the ^{13}C band maxima are then primarily dependent on the exciton splitting of these modes acting relatively independently, whose extent is dependent on the number and positions of the labeled residues.

Hydrogen bonding does play an important role in increasing the exciton splitting of the amide I for β-sheets, but it is one of introducing a coupling to the next chain rather than affecting the C═O bond strength (which would shift the centroid of the exciton band). As can be seen for the 2-stranded simulations, there is much less overall splitting predicted for both the unlabeled (Figure 2) and labeled (Figures 5 and 6) peptides than for the multistranded case where all the central residues are hydrogen bonded to another chain. This is consistent with our simulations of extended peptides with no hydrogen bonding which yield even less exciton splitting (Kubelka and Keiderling, to be submitted for publication). This coupling to another strand is just what the empirical parameters of Brauner et al.[18] achieved in their simulations. Thus there is a consistency between these two models even if interpretations of the resultant effects are different.

Isotopic Intensity Enhancement. The physical basis for the isotopic intensity enhancement effect that is derived from our simulations is different than that suggested by Brauner et al.,[18] who attributed the enhancement of the ^{13}C-labeled part to the participation of ^{12}C amide motions. While we observe a relatively small degree of mixing of the unlabeled amide vibrations into the internal ^{13}C modes, it would not account for the large intensity of the labeled modes. Despite this, the anomalous isotopic enhancement of intensity seen experimentally is apparent in our simulations for the multistranded sheets, which agree with experiment to an almost quantitative degree. The origin of this enhancement seems to come from a single, predominantly ^{13}C mode, whose largest amplitudes are on the center strand or strands of the multistranded structure. Interactions with the neighboring strands must cause the large amplitude to concentrate in relatively few amide groups in the sheet center. We have computationally investigated the influence of different positions of the alternate label pairs within the sheet. All tested label positions evidenced enhancement of the ^{13}C band; however, sheets constructed with irregularly placed pairs of labels, as would occur if the strands had no regular alignment, showed much weaker enhancement than did the regular cases.

Effect of Long-Range TDC. To address the problem of the long-range couplings, which are necessarily neglected in the transferred FF, we added the parameter-free TDC correction, which uses just the APT values obtained from the same ab initio calculation. The results imply that neglect of the longer range interactions did not cause substantial errors, which agrees with earlier empirical results.[28]

We cannot distinguish between the through-bond and TDC coupling contributions between locally coupled amide I modes, such as those within the repeating unit of the infinite antiparallel β-sheet,[22] since the quantum mechanical force field contains both, as well as higher order electrostatic couplings. However, including the TDC correction explicitly for the far off-diagonal part of the FF makes the predicted intensities worse (Figure 8). By contrast, this correction was shown to slightly improve the results of the IR absorption and VCD simulations of isotopically labeled α-helical peptides.[19] In the α-helical case, where the amide transition dipoles are nearly parallel and therefore their interaction is maximized, TDC probably represents better approximation for the long-range interactions. For the β-sheet, on the other hand, where transition dipoles of alternating amide orientations virtually cancel, this approximation is probably inadequate.

Conclusion

We presented simulations of amide I IR absorption spectra of multistranded oligopeptide β-sheets based on ab initio (i.e. using no empirical and/or adjustable parameters) quantum mechanical calculations and transfer of molecular property tensors. The results provide an alternative explanation of the ^{13}C isotopic enhancement effect in the model β-sheet peptides, that of arising from the aggregated nature of the extended oligopeptide β-strands. We demonstrate that the relative intensity of the ^{13}C-labeled part of the amide I increases with the number of associated β-strands. The most intense ^{13}C-labeled transitions are normally localized within labeled residues on the *central* strands of the multiple-stranded β-sheet. The ^{13}C results support the same conclusions for unlabeled peptides whose amide I band shape, namely the splitting between the high- and low-frequency maxima, converges to the experimentally observed "characteristic" β-sheet pattern with an increasing number of strands. They also confirm that the origin of the "low frequency" β-sheet amide I bands is exciton coupling and not primarily H-bonding strength. Therefore, our calculations provide a fundamental understanding of the hypothesis that the "standard" β-sheet IR spectrum as commonly found with polypeptide models arises from multistranded clusters of extended β-strands and explains why it poorly reflects protein β-sheet spectra. With the distinct spectral signature for such multistranded structure formation and the isotopic sensitivity as a new added diagnostic for β-sheet formation on a local, site-specific basis, time-dependent IR methods could provide a powerful tool for studying dynamics

of β-sheet formation. Finally, this study demonstrates the usefulness of ab initio based spectral simulations in structural interpretation of the experimental peptide vibrational spectra.

Acknowledgment. This work was supported by the Petroleum Research Fund, administered by the American Chemical Society (35443-AC). We thank Prof. Richard Mendelsohn, Rutgers University, for suggesting and discussing the problem as well as providing selected isotopically labeled samples for study.

JA010270X