

University of Groningen

## Distinguishing islet amyloid polypeptide fibril structures with infrared isotope-label spectroscopy

Saxena, Vishesh; Steendam, Ruben; Jansen, Thomas L. C.

*Published in:*  
The Journal of Chemical Physics

*DOI:*  
[10.1063/5.0082322](https://doi.org/10.1063/5.0082322)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2022

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Saxena, V., Steendam, R., & Jansen, T. L. C. (2022). Distinguishing islet amyloid polypeptide fibril structures with infrared isotope-label spectroscopy. *The Journal of Chemical Physics*, 156(5), [ 055101]. <https://doi.org/10.1063/5.0082322>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Distinguishing islet amyloid polypeptide fibril structures with infrared isotope-label spectroscopy

Cite as: J. Chem. Phys. **156**, 055101 (2022); <https://doi.org/10.1063/5.0082322>

Submitted: 14 December 2021 • Accepted: 13 January 2022 • Accepted Manuscript Online: 13 January 2022 • Published Online: 02 February 2022

 Vishesh Saxena,  Ruben Steendam and  Thomas L. C. Jansen

## COLLECTIONS

Paper published as part of the special topic on [Time-resolved Vibrational Spectroscopy](#)



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

### [Computational spectroscopy of complex systems](#)

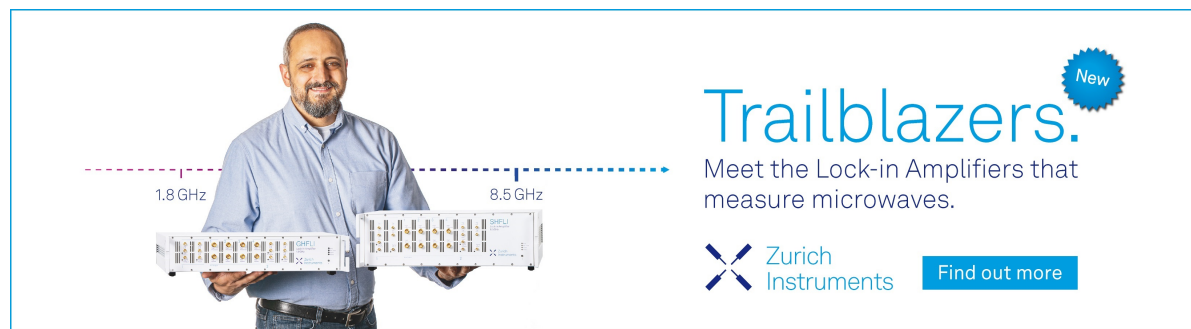
The Journal of Chemical Physics **155**, 170901 (2021); <https://doi.org/10.1063/5.0064092>


### [Probing functional conformation-state fluctuation dynamics in recognition binding between calmodulin and target peptide](#)

The Journal of Chemical Physics **156**, 055102 (2022); <https://doi.org/10.1063/5.0074277>


### [Wave packet dynamics in an harmonic potential disturbed by disorder: Entropy, uncertainty, and vibrational revivals](#)

The Journal of Chemical Physics **156**, 054303 (2022); <https://doi.org/10.1063/5.0079938>



**Trailblazers.** 

Meet the Lock-in Amplifiers that measure microwaves.

 Zurich Instruments [Find out more](#)

# Distinguishing islet amyloid polypeptide fibril structures with infrared isotope-label spectroscopy

Cite as: J. Chem. Phys. 156, 055101 (2022); doi: 10.1063/5.0082322

Submitted: 14 December 2021 • Accepted: 13 January 2022 •

Published Online: 2 February 2022



View Online



Export Citation



CrossMark

Vishesh Saxena,<sup>a)</sup>  Ruben Steendam,<sup>b)</sup>  and Thomas L. C. Jansen<sup>b)</sup> 

## AFFILIATIONS

University of Groningen, Zernike Institute for Advanced Materials, Nijenborgh 4, 9747 AG Groningen, The Netherlands

**Note:** This paper is part of the JCP Special Topic on Time-resolved Vibrational Spectroscopy.

<sup>a)</sup> **Current address:** Department of Physics, University of Hamburg, 20355 Hamburg, Germany.

<sup>b)</sup> **Author to whom correspondence should be addressed:** [t.l.c.jansen@rug.nl](mailto:t.l.c.jansen@rug.nl)

## ABSTRACT

Here, we performed spectral simulations of the amide-I vibrational spectra for three proposed fibril structures of the human islet amyloid polypeptide, which is involved in type II diabetes. We modeled both the overall absorption and two-dimensional infrared spectra for these structures. We further analyzed the isotope-labeled spectra, including the variation between structures. The analysis suggests that the infrared spectra of the cryo-electron microscopy structure provide the best match with experimental data. We further simulated isotope-labeled dilution spectroscopy investigating the correlation between the predicted spectral peak shift and the coupling between the amide units. While this correlation works in most cases, failures were observed when the isotope-labeled spectra were broad compared to the coupling or exhibited structure. These findings will be useful in the quest for potential toxic fibril formation intermediates.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0082322>

## I. INTRODUCTION

Amylin is a hormone produced in the pancreas.<sup>1</sup> It is also denoted as human islet amyloid polypeptide (hIAPP) and plays a role in type II diabetes.<sup>2–6</sup> In the normal function, it travels to the brain, where it suppresses the appetite.<sup>1</sup> When amylin becomes overproduced in a diabetic patient, it starts accumulating and forms plaques composed of fibers formed by aggregated hIAPP. These plaques are likely not the toxic species as demonstrated by toxicity experiments,<sup>7</sup> but rather nucleating oligomers<sup>8</sup> or intermediates<sup>9,10</sup> are responsible for the toxicity. The evidence of such protein folding intermediates has been reported with the help of two-dimensional infrared spectroscopy (2DIR).<sup>11</sup> Still, no such oligomers nor folding intermediates have been isolated or identified. The goal of this paper is to examine the infrared spectroscopy of several proposed structures for the already aggregated structures<sup>12,13</sup> as an understanding of these will be crucial for using infrared spectroscopy to study the potential oligomers.

Amyloids can appear in many other organs of the human body as either a symptom or a cause of (degenerative) diseases.<sup>14</sup> As of yet, at least 50 diseases are recognized that are characterized by the

deposition of amyloids.<sup>15,16</sup> These diseases include Parkinson's disease, Huntington's disease, rheumatoid arthritis, and spongiform encephalopathy.<sup>17–21</sup> A similar mechanism for the pathology, such as that of type II diabetes, may thus be possible in these diseases. Therefore, the understanding of the hIAPP system may provide valuable insights for a whole range of diseases.

It has been shown by various studies that hIAPP fibrils are polymorphic in nature.<sup>22–25</sup> From solid-state nuclear magnetic resonance (ss-NMR) studies, two possible structures for the  $A\beta_{1-40}$  fibrils, named 2fold<sup>26</sup> and 3fold,<sup>27–29</sup> were proposed. These structures were modeled with molecular dynamics and spectral modeling,<sup>12</sup> and the results were compared to experimental data. Since then, several primarily experimental two-dimensional infrared studies utilizing site-specific isotope-labeling<sup>7,11,30–33</sup> have been applied to learn more about the structure and formation of the fibrils.

Recently, cryogenic electron microscopy (cryo-EM)<sup>34</sup> has seen a rapid development allowing increasingly accurate structure determination with the method. In 2020, a new structure of hIAPP fibrils was determined with cryo-EM.<sup>13</sup> This structure has the protein databank entry 6Y1A. So far, the infrared spectroscopy of this structure has not been modeled. It is therefore still unclear what

consequences this has for the quest of identifying the missing intermediates using Fourier Transform Infrared (FTIR) spectroscopy and 2DIR spectroscopies.

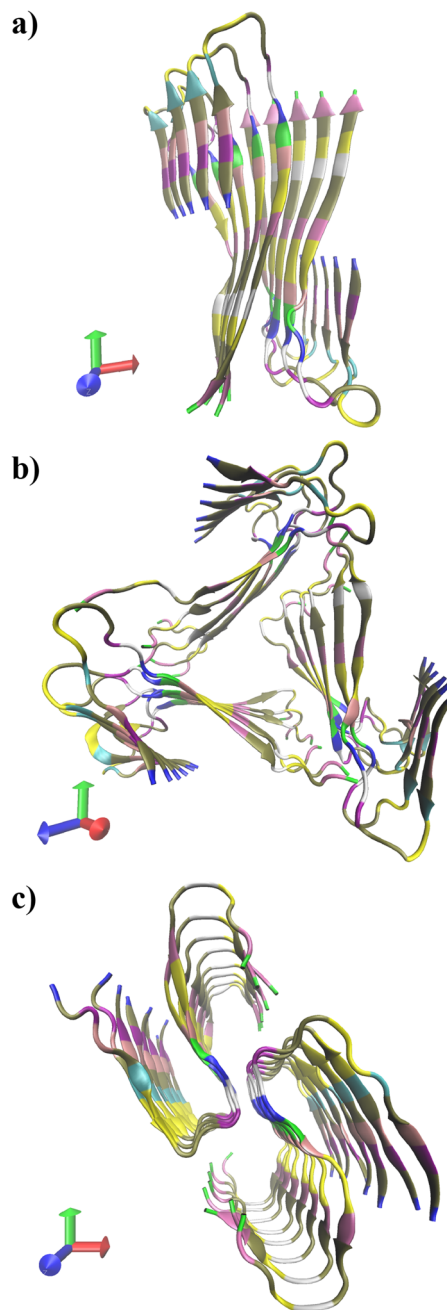
In this paper, we will present our computational examination of the infrared spectroscopy of the cryo-EM structure. For consistency, we repeat the examination of the 2fold and 3fold structures previously performed<sup>12</sup> but under identical conditions as the examination of the cryo-EM structure. A significant difference from the previous study is that we include only residues 13–37 as the coordinates as the remaining residues are not resolved in the 6Y1A cryo-EM structure. We follow a common computational spectroscopy simulation protocol<sup>35</sup> combining molecular dynamics,<sup>36</sup> vibrational frequency mapping,<sup>37</sup> and spectroscopic modeling using response functions.<sup>38</sup>

Isotope-labeling is a powerful tool to add site-specific resolution to infrared spectroscopy of proteins.<sup>39–45</sup> We follow an isotope-labeled protocol,<sup>30,46,47</sup> where spectra are determined for both systems where all copies of a specific residue are isotope-labeled and systems where only a dilute population of isotope-labels of that same residue is present. In both cases, the absorption of the isotope-labeled units is expected to be sufficiently shifted to red to effectively decouple the vibrations from all of the unlabeled residues. However, in the dilute case, the isotope-labeled units are also too distant from each other to be coupled, and ideally, this diluted isotope-labeled experiment should provide the spectrum of the base unit. When coupling with the other residues is present, a spectral shift arising from that coupling is expected.<sup>47</sup> In the simplest approximation, assuming a linear alignment of the residue, this shift is  $2J$ , where  $J$  is the coupling between neighboring residues in the linear arrangement. If the transition-dipole coupling (TDC) model is assumed, the shift will be  $2\zeta(3)J$ , where  $\zeta$  is the Riemann zeta function and  $2\zeta(3) \approx 2.404$ . In more general cases, the shift can be approximated by the signed sum of all couplings of one residue with all others,<sup>48</sup> also denoted as the coupling strength. The observed shift between fully site-specific isotope-labeled samples and dilute site-specific isotope-labeled samples is, thus, very sensitive to the aggregation pattern of the protein strands, which makes isotope-labeled Fourier Transform Infrared (FTIR) spectroscopy and 2DIR spectroscopy ideal for the examination of structures and dynamics within the aggregation processes of proteins.

The remainder of this paper is organized as follows: In Sec. II, we describe the applied methods for simulating the dynamics and spectroscopy of the three proposed fibril structures. In Sec. III, we first present the predicted overall spectra of the simulated structures followed by an analysis of the isotope-edited spectroscopy. Finally, in Sec. IV, we summarize the conclusions and provide an outlook.

## II. METHODS

The three structures for human islet amyloid polypeptide (hIAPP) fibrils were the 2fold and 3fold ss-NMR-based structures<sup>12</sup> as well as the recent cryo-EM structure (6Y1A).<sup>13</sup> To make the simulations as comparable as possible, we reduced all structures to the sequence—ANF LVH SSN NFG AIL SST NVG SNT Y—as reported in the cryo-EM structure.<sup>13</sup> Here, Ala13 is the first residue included and Tyr37 is the last. The 2fold, 3fold, and cryo-EM structures contain 10, 18, and 16 identical strands, respectively. The simulated structures are shown in Fig. 1. Each strand contains



**FIG. 1.** (a) 2fold, (b) 3fold, and (c) cryo-EM structures of islet amyloid polypeptide fibrils used in the present simulations. The colors indicate the different amino acid types according to the standard coloring scheme of VMD.<sup>49</sup>

five asparagine (N) units, which have a primary amide in the side chain.

The structural dynamics of the three structures were determined using molecular dynamics. The GROMACS-4.6.3<sup>36</sup> program was employed using the Optimized Potentials for Liquid Simulations–All Atoms (OPLS-AA)<sup>50</sup> force field for the peptide

and the extended simple point charge model (SPC/E) model for water.<sup>51</sup> The temperature was kept constant at 300 K using the V-rescale algorithm. The pressure was kept fixed at 1 bar using the Parrinello–Rahman algorithm.<sup>52</sup> The classical dynamics were calculated using 2 fs time steps with the Verlet algorithm. For long-range electrostatic forces, the particle mesh Ewald method<sup>53</sup> with a 1 nm cut-off radius was used. For the production run, snapshots of the structure were saved at 20 fs intervals with a total of 50 000 time steps providing one nanosecond of simulations time for each structure.

The Hamiltonians for the amide I vibrations were constructed using a frequency mapping procedure<sup>37</sup> using the AmideImaps program.<sup>54</sup> The total time-dependent Hamiltonian is

$$H(t) = \sum_i \hbar \omega_i(t) B_i^\dagger B_i + \sum_{j \neq i} J_{ij}(t) B_i^\dagger B_j - \frac{\Delta}{2} \sum_i B_i^\dagger B_i^\dagger B_i B_i - \sum_i \vec{\mu}_i(t) \cdot \vec{E}(t) (B_i^\dagger + B_i), \quad (1)$$

where  $\omega_i$  is the frequency of residue  $i$ ,  $B_i^\dagger$  and  $B_i$  are the Bosonic creation and annihilation operators for an amide I vibration on residue  $i$ ,  $J_{ij}$  represent the coupling between the amide I vibrations on two different residues, and  $\Delta$  is the anharmonicity, which was set to be  $16 \text{ cm}^{-1}$  for all amide I vibrations. The coupling with the external electric field,  $\vec{E}$ , is governed by the transition dipole  $\vec{\mu}_i$  of each amide I vibration. The frequency mapping developed in the Skinner group<sup>55</sup> was used including the mapping developed for side chains. A cut-off radius for the electrostatic interactions of 2 nm was employed, and the frequency shift from nearest neighbors along the backbone was included through a Ramachandran angle-based map.<sup>56,57</sup> The long-range couplings were modeled with the transition-dipole coupling (TDC) model,<sup>58</sup> while the nearest neighbor couplings along the backbone were described using a Ramachandran angle-based map.<sup>56,57</sup> This simulation protocol has been demonstrated to work well for proteins.<sup>59,60</sup>

Spectral simulations were performed using the Numerical Integration of the Schrödinger Equation (NISE).<sup>38,61,62</sup> In this algorithm, the response functions governing the spectra are calculated directly in the time-domain and the quantum dynamics of the amide I vibrations propagated for each time step are calculated by solving the time-dependent Schrödinger equation assuming the Hamiltonian to be time-independent for that brief time delay. This treatment has the advantage that it both accounts for non-adiabatic dynamics<sup>63</sup> as well as non-Gaussian fluctuations.<sup>64,65</sup> The spectra are finally obtained through a Fourier transform of the respective response functions. A more elaborate discussion of the procedure can be found in recent review papers.<sup>35,66,67</sup>

Isotope-labeled spectra were calculated in two different ways. The spectra, corresponding to the full isotope-labeling of a specific residue in all strands of the structures, were done by excluding all other residues in the Hamiltonian. This approximation requires that the isotope shift is large enough to separate the vibrational frequency of the specific residue sufficiently out of the main amide I band, such that the coupling with other residues can be neglected. This approximate procedure has the advantage over explicit isotope-labeling by shifting the frequency of the labeled sites that the result does not depend on the explicit label used ( $^{13}\text{C}$ ,  $^{18}\text{O}$ , or both). From the literature, it is known that  $^{13}\text{C}$  labeling results in a  $-41 \text{ cm}^{-1}$  systematic frequency shift<sup>42</sup> and  $^{13}\text{C}^{18}\text{O}$  labeling results in a  $-66 \text{ cm}^{-1}$

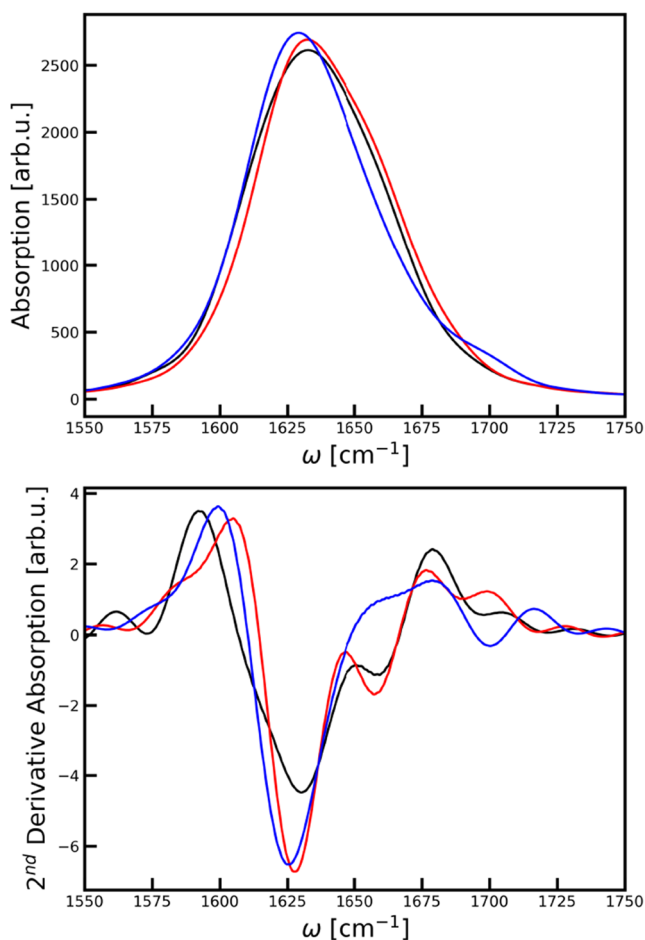
systematic frequency shift.<sup>30</sup> Additionally, the calculations are faster as the spectrum of the isotope-labeled residue is the only thing that needs to be calculated. The spectra corresponding to the isotope-labeling of a specific residue in a small fraction of the strands (dilute isotope-labeling) were obtained by repeating the previous procedure, but setting all couplings between the residues to zero, effectively treating them as if they were independent. In this way, the extreme dilute limit is simulated, as we average over the situation where the specific residue is isotope-labeled only in one of the strands at a time. With 18 strands as the largest number (in the 3fold structure), this, thus, corresponds to a maximum concentration of 5% of the labeled species.

### III. RESULTS AND DISCUSSION

We first calculated the FTIR spectra of the full structures. These spectra are shown in Fig. 2 together with the second derivatives. The spectra of the three structures are quite similar; however, the cryo-EM spectrum is narrower and has a high-frequency sideband peaking at  $1700 \text{ cm}^{-1}$ . The 2fold and 3fold spectra both exhibit a shoulder at  $1658 \text{ cm}^{-1}$  as highlighted by the second-derivative spectra. The main peaks are located at 1630, 1628, and  $1625 \text{ cm}^{-1}$  for the 2fold, 3fold, and cryo-EM structures, respectively. The similarity of the 2fold and 3fold structures can be understood from the fact that in these structures, the secondary structural elements are at the same position within the structures. The cryo-EM structure seems to exhibit less disorder resulting in a slightly narrower FTIR spectrum. The spectral shape for the cryo-EM structure is in slightly better qualitative agreement with spectra reported in the literature<sup>32</sup> even the spectral width for all three models is larger than the experimental observation. When comparing the overall spectra with previous experiments, it should be acknowledged that only the residues from Ala13 to Tyr37 were included in the current simulations and that the fibril structures are significantly shorter than those used within previous experiments. This smaller size likely leads to finite size effects resulting in more disorder at water exposed strands and a more limited effect of exchange narrowing.

We examined the effect of inter-strand couplings by calculating the spectra without coupling between the different strands (not-shown). Neglecting these couplings resulted in  $10\text{--}16 \text{ cm}^{-1}$  blue-shifts of the spectra as compared to the full spectra demonstrating that the spectroscopically relevant states are delocalized over multiple strands and that infrared spectroscopy is sensitive to the strand arrangement and not only the structure of the individual strands.

We proceeded to examine the overall 2DIR spectra for zero waiting time obtained using the perpendicular laser polarization. This polarization is known to enhance cross-peaks and structural information for  $\beta$ -sheet structures.<sup>68,69</sup> For the spectra presented in Fig. 3, it is clear that the 2fold and 3fold structures are very similar, as even minor differences are observable when examining the details. The spectrum of the cryo-EM structure is slightly narrower. From the FTIR or 2DIR spectra alone, it will be difficult to distinguish between the three structures from the other structures, as the spectra of the 2fold and 3fold structures are hardly distinguishable from each other. The spectral shape for the cryo-EM structure is in

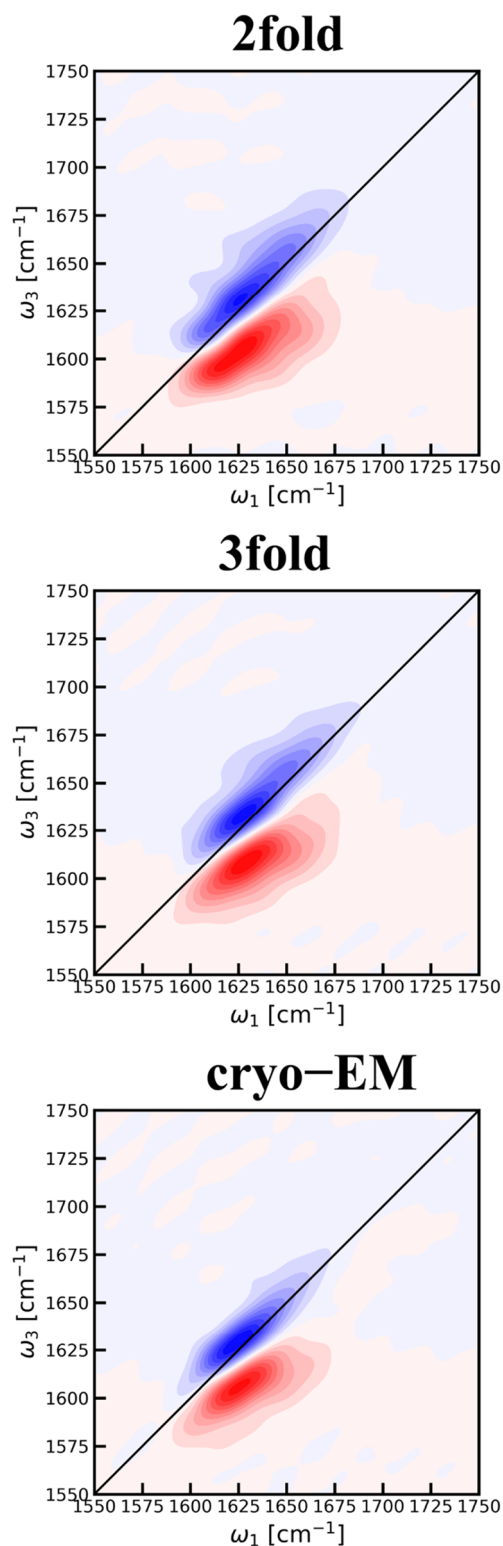


**FIG. 2.** Absorption spectra (top) and the corresponding second-derivative spectra (bottom) calculated for the three structures with 2fold in black, 3fold in red, and the cryo-EM structure in blue.

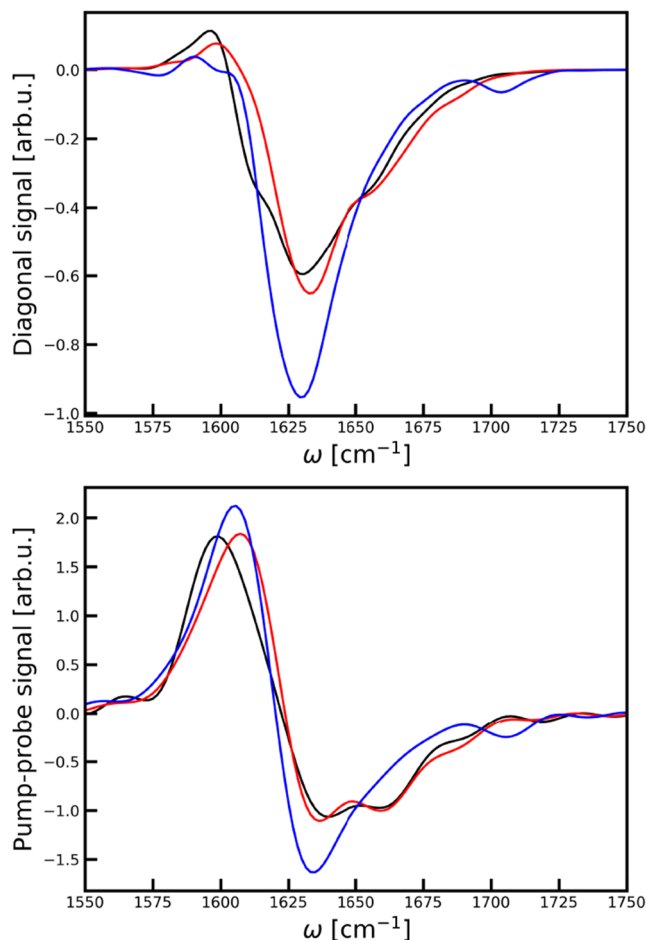
better qualitative agreement with spectra reported in the literature<sup>32</sup> than that of the 2fold and 3fold structures.

Diagonal cuts through the 2DIR spectra are presented in Fig. 4 together with the broadband pump-probe spectra calculated by integrating the 2DIR spectra over  $\omega_1$ . The similarity between the spectra of the NMR structures is very clear in both the diagonal cuts and the pump-probe spectrum. The pump-probe spectra further highlight the spectral component seen in the absorption spectra as shoulders at 1658 cm<sup>-1</sup> for the NMR structures and the peak at 1700 cm<sup>-1</sup> for the cryo-EM structure.

Isotope-labeled spectroscopy is sensitive to local structure and dynamics. For dilute isotope-labels, the absorption peak position is expected to be determined by the average site frequency but shifted, with this isotope shift depending on the choice of isotope. In Fig. 5, the average diagonal frequencies for all backbone residues are shown without applying a specific isotope shift. The lines are running averages over three consecutive residues meant as a guide to the eye for observing correlations. For many sites, the difference between the average frequencies for the different structures is



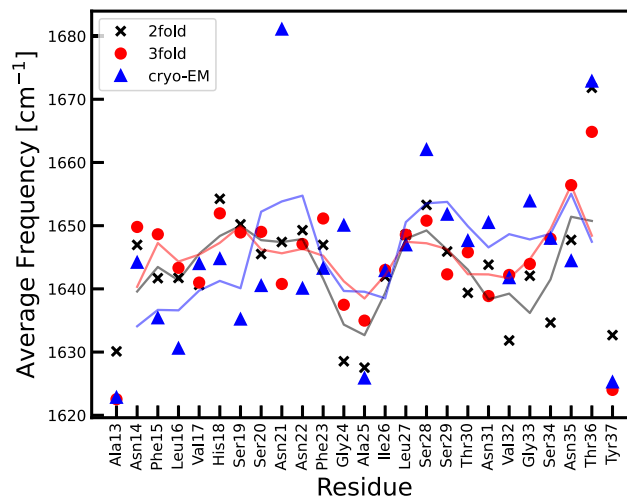
**FIG. 3.** Calculated 2DIR spectra for the three structures with perpendicular polarization at zero waiting time. The equidistant contour lines are drawn at each 10% of the maximal signal in each figure.



**FIG. 4.** Diagonal cuts through 2DIR spectra of Fig. 3 (top) and the calculated broadband pump-probe spectra for the three structures with perpendicular polarization at zero waiting time (bottom). 2fold is plotted in black, 3fold is plotted in red, and the cryo-EM structure is plotted in blue.

small, and measuring the peak position in an isotope-labeled experiment cannot be expected to distinguish between the structures. For example, the Ile26 and Leu27 (for Leu27, the symbol for 2fold is perfectly hidden behind that of 3fold in Fig. 5) residues will provide no insight in such an experiment. The Asn21 and Gly24 residues would reveal sufficient information to distinguish between the three structures.

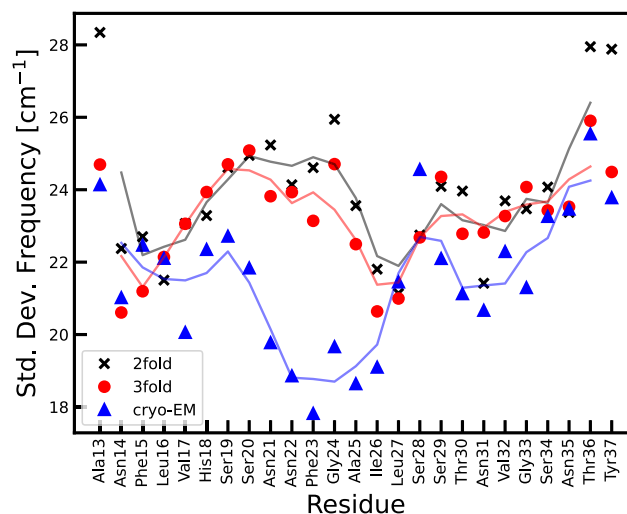
Assuming a Gaussian distribution of the frequencies for each amide I site, the standard deviation of the frequency fluctuations can be expected to correlate with the standard deviation extracted from a Gaussian fit to the experimental isotope-labeled spectrum of that site. The calculated standard deviation of the fluctuating site frequencies is presented in Fig. 6 for each site. The lines are running averages over three consecutive residues meant as a guide to the eye for observing correlations. The most evident observation is that the standard deviation is generally lower for the cryo-EM structure, especially in the segment starting around Val17 and ending at the Ile26. This is a direct reflection of the lower flexibility of the cryo-EM



**FIG. 5.** Average frequencies predicted for each amide I unit along the backbone. The lines are running averages over three consecutive residues.

structure as compared to the 2fold and 3fold structures. The lower variation in site frequencies for the cryo-EM structure can partially explain the narrower spectrum.

We define the (interchain) coupling strength as  $S_i = \sum_{jk} J_{ii}^{jk} (1 - \delta_{jk})$ ,<sup>70</sup> where the signed sum runs over all peptide chains and  $i$  labels a specific amide I site. The coupling,  $J_{ii}^{jk}$ , here denotes the coupling between identical sites numbered by  $i$  on different peptide chains labeled  $j$  and  $k$ . The coupling strength, thus, quantifies the coupling between vibrations corresponding to the same residue on different chains. This quantity is of particular interest when comparing spectra with full isotope-labeling of a specific residue



**FIG. 6.** The standard deviation of the frequency fluctuation predicted for each amide I unit along the backbone. The lines are running averages over three consecutive residues.

and spectra with dilute isotope-labeling, where, for example, just 10% of the strands are labeled resulting in spectra of uncoupled residues. The coupling strength is expected to correlate with the frequency shift observed between these full and dilute isotope-labeled experiments.<sup>12</sup> This correlation relies on the formation of delocalized exciton states involving vibrations with parallel transition dipoles in the full isotope-labeled case. The calculated coupling strengths are summarized in Table I and depicted in Fig. 7. The coupling strengths for the 2fold, 3fold, and cryo-EM structures were corrected with factors 1.25 (20/16), 1.2 (36/30), and 1.14 (32/28), respectively, to account for the lack of periodic boundary conditions in the molecular dynamics simulations. (These factors are given by  $\frac{2N_s}{2(N_s - N_F)}$ , where  $N_s$  is the number of strands in each structure and  $N_F$  is the fold symmetry numbers, which is three for the 3fold structure and two for the other structures.) The lines in Fig. 7 are running averages over three consecutive residues meant as a guide to the eye for observing correlations. In Table I, the coupling strength is also compared with the peak shift  $\Delta\omega_i$  extracted from simulated full and dilute isotope-labeled absorption spectra, where a negative number corresponds to a red-shift of the full isotope-labeled absorption peak. This comparison reveals that the coupling strength indeed is correlated well with the calculated peak shift; however, it is also clear that the correlation is not strong enough to allow a direct determination

of the coupling strength from the experimental data. This is because the simple relationship only applies when one neglects disorder in frequency fluctuations and transition-dipole orientations, which are visibly present in the fibrils. The specific case of Leu16, where a significant difference between the coupling strength and the frequency shift can be observed, will be discussed later.

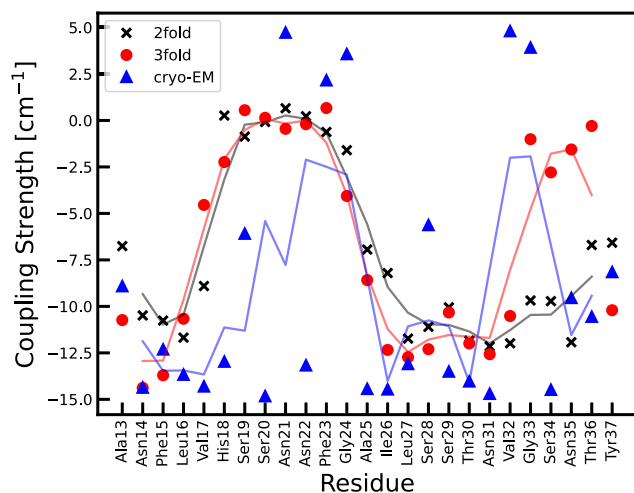
We found the absorption and 2DIR spectra of the 2fold and 3fold structures to be very similar. This can be understood due to their similarity in the secondary structure. When examining the coupling strength, it becomes apparent that these are very similar in the central His18 to Val32 region, but much more variation is observed in the C-terminal region. Isotope-labeled experiments in this region would, thus, offer the possibility of distinguishing these two structures. The cryo-EM structure exhibits large differences in coupling strengths from the NMR structures at many residue positions including, for example, His18, Ser20, Asn22, Val32, and Ser34.

For a more direct comparison with existing isotope-labeled experiment data,<sup>12</sup> the results were plotted for the sites for which experimental data are available in Figs. 8 and 9. This time the calculated frequencies were shifted to match the effect of the <sup>13</sup>C<sup>18</sup>O isotope-labeled shift<sup>12</sup> and the systematic error expected from the used frequency map<sup>60</sup> resulting in an overall red-shift of 50 cm<sup>-1</sup>.

**TABLE I.** Overview of coupling strengths  $S_i$  and isotope dilution peak shifts ( $\Delta\omega_i$ ) for all backbone sites in the three structures.

Residue	2fold		3fold		cryo-EM	
	$S_i$ (cm <sup>-1</sup> )	$\Delta\omega_i$ (cm <sup>-1</sup> )	$S_i$ (cm <sup>-1</sup> )	$\Delta\omega_i$ (cm <sup>-1</sup> )	$S_i$ (cm <sup>-1</sup> )	$\Delta\omega_i$ (cm <sup>-1</sup> )
Ala13	-6.8	-11.3	-10.7	-12.9	-8.9	-12.9
Asn14	-10.5	-15.2	-14.4	-15.7	-14.4	-15.4
Phe15	-10.8	-16.2	-13.7	-15.2	-12.3	-13.3
Leu16	-11.7	-13.2	-10.7	-13.7	-13.7	-25.1
Val17	-8.9	-8.8	-4.5	-5.9	-14.3	-16.4
His18	0.3	-1.7	-2.2	-1.3	-13.0	-16.3
Ser19	-0.9	0.1	0.6	-0.4	-6.1	-7.5
Ser20	-0.1	-0.4	0.1	0.8	-14.8	-15.9
Asn21	0.7	0.3	-0.5	-1.1	4.7	7.0
Asn22	0.2	0.2	-0.2	0.2	-13.2	-13.2
Phe23	-0.6	-1.3	0.7	0.9	2.1	8.1
Gly24	-1.6	-1.7	-4.0	-3.7	3.6	2.1
Ala25	-6.9	-9.9	-8.6	-10.3	-14.4	-17.3
Ile26	-8.2	-11.5	-12.3	-13.8	-14.5	-16.7
Leu27	-11.7	-13.6	-12.7	-14.2	-13.1	-11.8
Ser28	-11.1	-10.5	-12.3	-13.9	-5.6	-16.8
Ser29	-10.0	-11.1	-10.3	-14.0	-13.5	-14.6
Thr30	-11.8	-11.3	-12.0	-14.3	-14.0	-16.0
Asn31	-12.2	-14.3	-12.6	-17.6	-14.7	-16.2
Val32	-12.0	-13.5	-10.5	-13.5	4.8	4.8
Gly33	-9.7	-15.2	-1.0	0.9	3.9	5.1
Ser34	-9.7	-9.8	-2.8	-14.6	-14.5	-17.3
Asn35	-11.9	-14.9	-1.6	-1.4	-9.6	-13.2
Thr36	-6.7	-3.9	-0.3	1.6	-10.6	-15.0
Tyr37	-6.6	-13.5	-10.2	-13.3	-8.2	-12.0

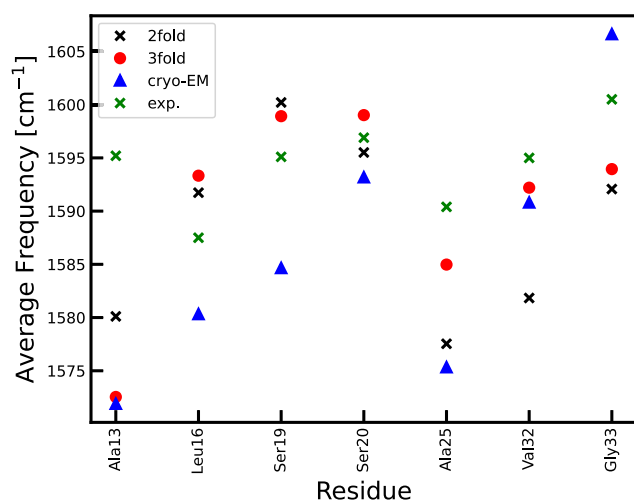




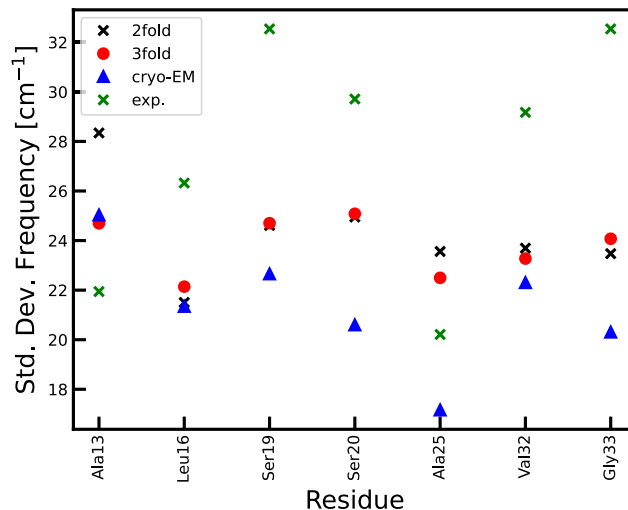
**FIG. 7.** Coupling strength  $S_i$  predicted for each amide I unit along the backbone. The lines are running averages over three consecutive residues.

For the spectral widths, the experimental standard deviation is estimated using an assumption of a Gaussian distribution.<sup>12</sup> The comparison between experiment and simulation does not provide a perfect match for any of the structures. However, general trends as the observation of lower average frequency and small spectral width for Ala25 are in best agreement with the cryo-EM model. Interestingly, the simulated frequency standard deviations are generally smaller than the experimental observations. This may suggest larger structural variations in the experimental system, which could also arise from an inhomogeneous distribution of structures.

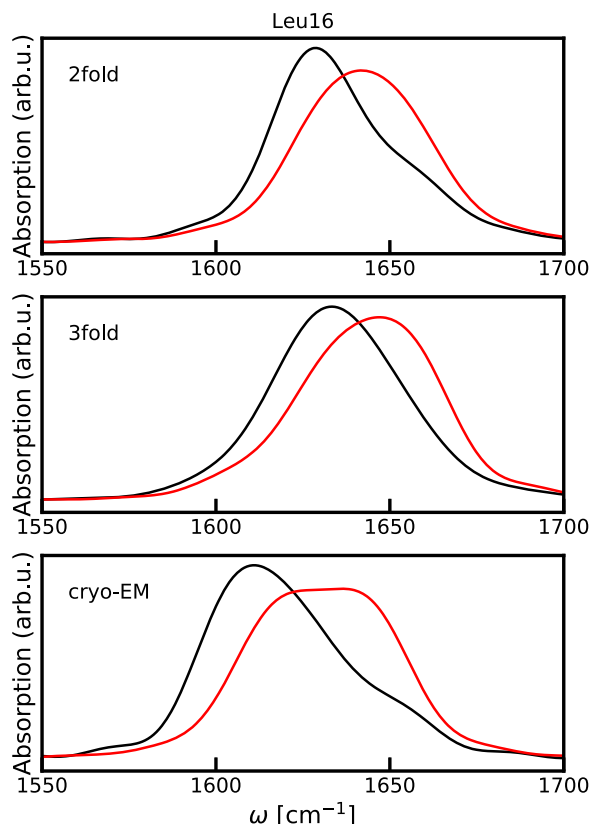
Below, the spectra for a few of the more noteworthy isotope-labels will be presented and discussed. For all the spectra, the



**FIG. 8.** Average frequencies predicted for selected amide I units along the backbone and compared with experimental peak positions from Ref. 12. The simulated frequencies are shifted by  $50 \text{ cm}^{-1}$  compared to Fig. 5 to account for the isotope-labeling and a systematic shift from the mapping.<sup>60</sup>



**FIG. 9.** The standard deviation of the frequency fluctuation predicted for selected amide I units along the backbone and compared with experimental peak widths from Ref. 12.

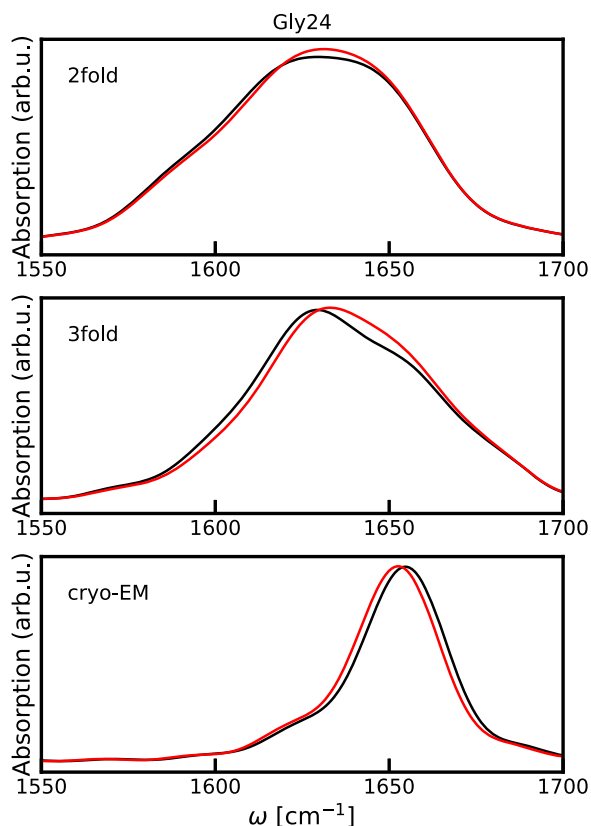


**FIG. 10.** Fully isotope-labeled (black) and diluted isotope-labeled (red) spectra for Leu16. The top panel contains 2fold spectra, the middle panel contains 3fold, and the bottom panel contains cryo-EM.

frequency used is the unshifted one, and for comparison to the actual experimental data, a shift corresponding to the specific label should be used along with a  $20\text{ cm}^{-1}$  blue-shift corresponding to the systematic error from the mapping method.<sup>60</sup>

In Fig. 10, the spectra for Leu16 (the pink residue closest to the N-terminus in Fig. 1) are presented. The shift upon isotope dilution is similar in the three cases. For the cryo-EM structure, a double peak structure is seen. This explains the exceptionally large shift reported in Table I. Essentially, the coupling in this case does not only result in a shift but also an increase in the relative intensity of the lower peak. The shift reported in Table I is between the extracted peak positions and, in this case, is not a good indicator of the actual coupling strength.

The isotope-labeled spectra for Gly24 are shown in Fig. 11. For Gly24, the difference between the average frequencies for the 2fold and 3fold structures is expected to be quite significant. However, the spectra are also quite broad and a two-peak structure is apparent. The spectrum for the cryo-EM structure on the other hand is distinguishable as the absorption is at a higher frequency and the peak is very narrow. The difference can be easily understood as Gly24 (central white unit in Fig. 1) is in the center of the cryo-EM structure, while it is located on the water exposed side of both the 2fold and 3fold structures.

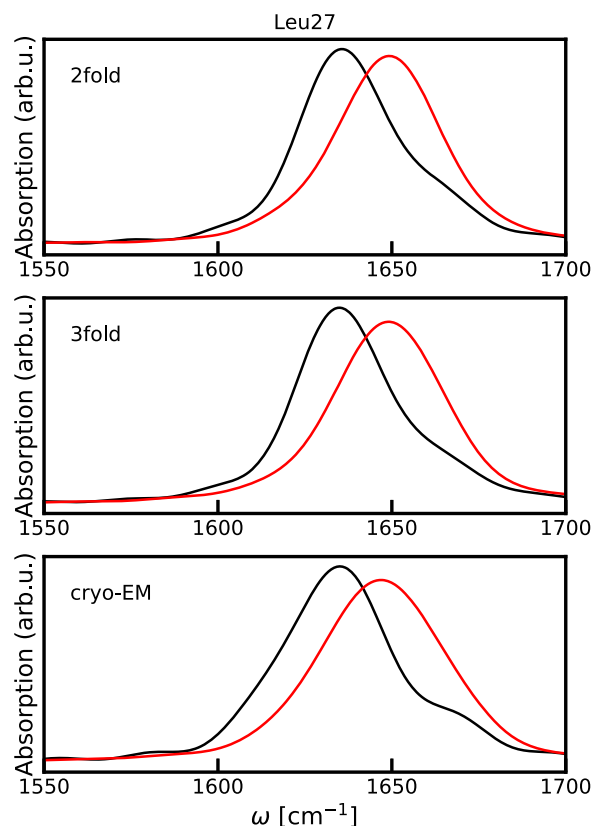


**FIG. 11.** Fully isotope-labeled (black) and diluted isotope-labeled (red) spectra for Gly24. The top panel contains 2fold spectra, the middle panel contains 3fold, and the bottom panel contains cryo-EM.

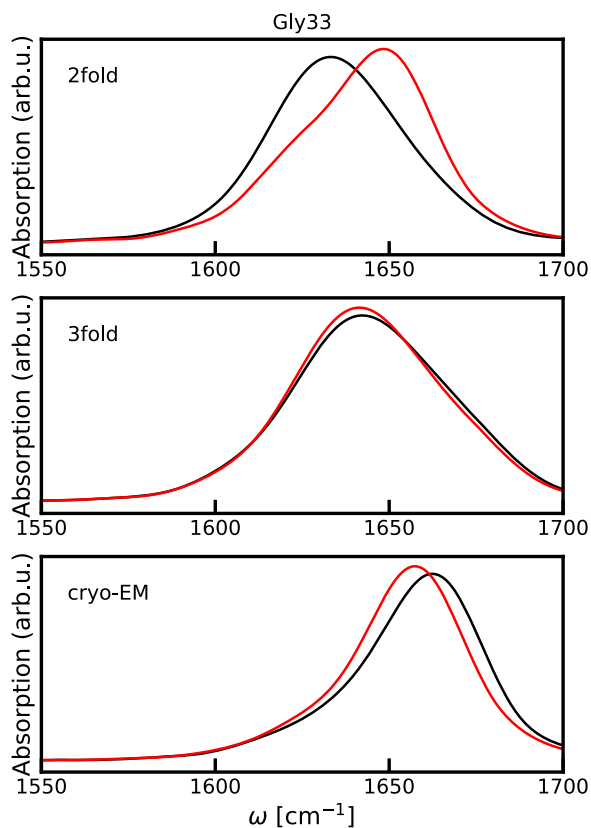
In Fig. 12, the spectra for Leu27 (the pink central residue in Fig. 1) are presented. For this residue, the isotope-labeled spectra for the three structures are very similar. This suggests that the local environment and couplings are essentially identical in these structures. Still, of course, this unit may be interesting for identifying intermediates as they could behave differently from the fully formed fibril.

In Fig. 13, the spectra for Gly33 (the white residue closest to the C-terminus in Fig. 1) are presented. Here, the 2fold structure exhibits a red-shift due to the coupling, while the 3fold structure is rather insensitive and the cryo-EM structure gives a blue-shift. This spectral variation suggests that this residue is well suited to distinguish between the three structures. The large variation can be understood as this residue is still in the fairly well protected  $\beta$ -sheet of the 2fold structure, while that part of the 3fold structure is more disordered, and in the cryo-EM structure, it is located in the outer turn region, explaining the positive value of the coupling strength.

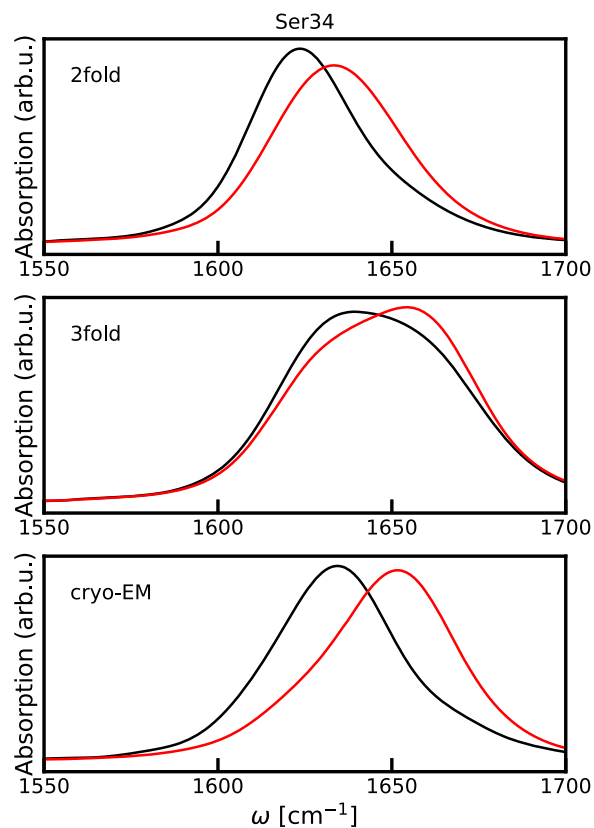
In Fig. 14, the spectra for Ser34 (the yellow residue closest to the C-terminus in Fig. 1) are presented. The difference between the 2fold and 3fold spectra is, similarly to Gly33, quite significant. The reason is probably the same as again this Ser34 residue is still part of the  $\beta$ -sheet in the 2fold structure, but not in the 3fold one. For the



**FIG. 12.** Fully isotope-labeled (black) and diluted isotope-labeled (red) spectra for Leu27. The top panel contains 2fold spectra, the middle panel contains 3fold, and the bottom panel contains cryo-EM.



**FIG. 13.** Fully isotope-labeled (black) and diluted isotope-labeled (red) spectra for Gly33. The top panel contains 2fold spectra, the middle panel contains 3fold, and the bottom panel contains cryo-EM.



**FIG. 14.** Fully isotope-labeled (black) and diluted isotope-labeled (red) spectra for Ser34. The top panel contains 2fold spectra, the middle panel contains 3fold, and the bottom panel contains cryo-EM.

cryo-EM spectra, the coupling induced shift is a red-shift as the unit is the first residue of the C-terminal  $\beta$ -sheet. The difference between the peak shift observed for the 3fold structure and the calculated coupling strength (see Table I) can be understood as the absorption peak is very broad and the coupling essentially induces a small deformation, which displaces the top of the peak.

In the discussion, the terminal groups have been left out. This is because one should be cautious about the quality of the modeling for these groups. The frequency mappings are not as well developed or tested<sup>56,71</sup> for these groups as for the internal groups. Furthermore, the simulated system was truncated near the termini and this may, of course, affect the local structure.

When comparing these results with experimental data, it should be kept in mind that the fibril structures may depend on the conditions under which they were produced and there could be a coexistence of different structures. In the simulations, we assume the proposed structure but do not account for the effects of the presence of buffer or salt concentrations. The mappings employed were tested against regular protein spectra<sup>60</sup> and gas-phase peptides<sup>71</sup> and for isotope-labels,<sup>30</sup> and it is well known that the details depend on the choice of force field and mappings. However, the overall trends are expected to be correctly predicted.

#### IV. CONCLUSION

In this paper, we compared the amide-I infrared spectroscopic properties of three proposed amyloid fibril structures. We found that the overall absorption spectra of the two NMR-based structures were more similar to each other than that of the cryo-EM-based structure. This similarity can be attributed to the local structure similarity imposed by the constraints obtained from the NMR experiments.<sup>29</sup> Still, some variation is seen in the two-dimensional spectra, and only when doing isotope-labeling, sensitivity to the tertiary structure is seen.

We found that in most cases, the coupling strength correlates well with the spectral shift observed between full and dilute isotope-labeling spectra. A few notable exceptions are observed, in particular, for Leu16, where a double peak feature resulting from local heterogeneity may result in misinterpretation of the isotope-labeled experiment for the cryo-EM structure. For Ser34, a similar issue is observed for one of the NMR structures, where deformation of the spectral peak results in an apparent shift that is much larger than the coupling strength. For broad and structured isotope-labeled peaks, caution should, thus, be taken when interpreting experimental isotope dilution spectra.

A comparison of the isotope-labeled spectral peak positions and widths with experiments suggests a slightly better match between the cryo-EM structure and the reported infrared data. However, the simulated peak widths for the isotope-labels are quite consistently narrower than the experimental numbers reported in Ref. 12. This could be an indication that more loose structures or a more heterogeneous distribution of structures exists under the given experimental conditions. We propose that isotope-labeled dilution experiments on the units His18, Ser20, Asn22, Val32, and Ser34 provide good discrimination between the three different structures examined here.

The data provided here may further be useful for discriminating between possible fibril structures and structures of folding intermediates or oligomers that have been proposed to be the toxic species involved in type II diabetes. Future steps could also predict and model such intermediate structures based on, for example, large-scale molecular dynamics simulations on dissolved monomers and oligomers.

## ACKNOWLEDGMENTS

The authors thank Martin T. Zanni for inspirational discussions and Lu Wang for kindly providing starting structures for 2fold and 3fold. We are grateful to Lauren E. Buchanan and Chris Middleton for helpful discussions. Christi A. Marocico is kindly acknowledged for data management support. We would like to thank the Center for Information Technology of the University of Groningen for their support and for providing access to the Peregrine high performance computing cluster.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES

- J. B. Buse, C. Weyer, and D. G. Maggs, *Clin. Diabetes* **20**, 137 (2002).
- L. Marzban, K. Park, and C. B. Verchere, *Exp. Gerontol.* **38**, 347 (2003).
- A. Kanatsuka, S. Kou, and H. Makino, *Diabetol. Int.* **9**, 143 (2018).
- S. E. Kahn, S. Andrikopoulos, and C. B. Verchere, *Diabetes* **48**, 241 (1999).
- A. Mukherjee, D. Morales-Scheihing, P. C. Butler, and C. Soto, *Trends Mol. Med.* **21**, 439 (2015).
- R. L. Hull, G. T. Westermark, P. Westermark, and S. E. Kahn, *J. Clin. Endocrinol. Metab.* **89**, 3629 (2004).
- A. Abedini, A. Plesner, P. Cao, Z. Ridgway, J. Zhang, L.-H. Tu, C. T. Middleton, B. Chao, D. J. Sartori, F. Meng, H. Wang, A. G. Wong, M. T. Zanni, C. B. Verchere, D. P. Raleigh, and A. M. Schmidt, *eLife* **5**, e12977 (2016).
- E. T. Powers and D. L. Powers, *Biochem. J.* **91**, 122 (2006).
- T. R. Serio, A. G. Cashikar, A. S. Kowal, G. J. Sawicki, J. J. Moslehi, L. Serpell, M. F. Arnsdorf, and S. L. Lindquist, *Science* **289**, 1317 (2000).
- J. Lee, E. K. Culyba, E. T. Powers, and J. W. Kelly, *Nat. Chem. Biol.* **7**, 602 (2011).
- A. L. Serrano, J. P. Lomont, L.-H. Tu, D. P. Raleigh, and M. T. Zanni, *J. Am. Chem. Soc.* **139**, 16748 (2017).
- L. Wang, C. T. Middleton, S. Singh, A. S. Reddy, A. M. Woys, D. B. Strasfeld, P. Marek, D. P. Raleigh, J. J. de Pablo, M. T. Zanni, and J. L. Skinner, *J. Am. Chem. Soc.* **133**, 16062 (2011).
- C. Röder, T. Kupreichyk, L. Gremer, L. U. Schäfer, K. R. Pothula, R. B. G. Ravelli, D. Willbold, W. Hoyer, and G. F. Schröder, *Nat. Struct. Mol. Biol.* **27**, 660 (2020).
- F. Chiti and C. M. Dobson, *Annu. Rev. Biochem.* **86**, 27 (2017).
- B. L. Kagan, R. Azimov, and R. Azimova, *J. Membr. Biol.* **202**, 1 (2004).
- M. D. Benson, J. N. Buxbaum, D. S. Eisenberg, G. Merlini, M. J. M. Saraiva, Y. Sekijima, J. D. Sipe, and P. Westermark, *Amyloid* **25**, 215 (2018).
- K. A. Conway, J. D. Harper, and P. T. Lansbury, *Biochemistry* **39**, 2552 (2000).
- E. Scherzinger, A. Sittler, K. Schweiger, V. Heiser, R. Lurz, R. Hasenbank, G. P. Bates, H. Lehrach, and E. E. Wanker, *Proc. Natl. Acad. Sci. U. S. A.* **96**, 4604 (1999).
- R. O'Hara, E. P. Murphy, A. S. Whitehead, O. FitzGerald, and B. Bresnihan, *Arthritis Res. Ther.* **2**, 142 (2000).
- P. Westermark, A. Andersson, and G. T. Westermark, *Physiol. Rev.* **91**, 795 (2011).
- R. N. Rambaran and L. C. Serpell, *Prion* **2**, 112 (2008).
- J. Zhao, X. Yu, G. Liang, and J. Zheng, *Biomacromolecules* **12**, 210 (2011).
- A. V. Kajava, U. Aebi, and A. C. Steven, *J. Mol. Biol.* **348**, 247 (2005).
- C. S. Goldsbury, G. J. S. Cooper, K. N. Goldie, S. A. Müller, E. L. Saafi, W. T. M. Gruijters, M. P. Misur, A. Engel, U. Aebi, and J. Kistler, *J. Struct. Biol.* **119**, 17 (1997).
- C. Goldsbury, K. Goldie, J. Pellaud, J. Seelig, P. Frey, S. A. Müller, J. Kistler, G. J. S. Cooper, and U. Aebi, *J. Struct. Biol.* **130**, 352 (2000).
- S. Luca, W.-M. Yau, R. Leapman, and R. Tycko, *Biochemistry* **46**, 13505 (2007).
- A. T. Petkova, W.-M. Yau, and R. Tycko, *Biochemistry* **45**, 498 (2005).
- A. K. Paravastu, R. D. Leapman, W.-M. Yau, and R. Tycko, *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18349 (2008).
- R. Tycko, *Annu. Rev. Phys. Chem.* **62**, 279 (2011).
- A. M. Woys, A. M. Almeida, L. Wang, C.-C. Chiu, M. McGovern, J. J. de Pablo, J. L. Skinner, S. H. Gellman, and M. T. Zanni, *J. Am. Chem. Soc.* **134**, 19118 (2012).
- P. Marek, A. M. Woys, K. Sutton, M. T. Zanni, and D. P. Raleigh, *Org. Lett.* **12**, 4848 (2010).
- L. E. Buchanan, E. B. Dunkelberger, H. Q. Tran, P.-N. Cheng, C.-C. Chiu, P. Cao, D. P. Raleigh, J. J. de Pablo, J. S. Nowick, and M. T. Zanni, *Proc. Natl. Acad. Sci. U. S. A.* **110**, 19285 (2013).
- L.-H. Tu, A. L. Serrano, M. T. Zanni, and D. P. Raleigh, *Biophys. J.* **106**, 1520 (2014).
- A. Doerr, *Nat. Methods* **14**, 34 (2017).
- T. L. C. Jansen, *J. Chem. Phys.* **155**, 170901 (2021).
- M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, *SoftwareX* **1-2**, 19 (2015).
- C. R. Baiz, B. Blasiak, J. Bredenbeck, M. Cho, J.-H. Choi, S. A. Corcelli, A. G. Dijkstra, C.-J. Feng, S. Garrett-Roe, N.-H. Ge, M. W. D. Hanson-Heine, J. D. Hirst, T. L. C. Jansen, K. Kwac, K. J. Kubarych, C. H. Londergan, H. Maekawa, M. Reppert, S. Saito, S. Roy, J. L. Skinner, G. Stock, J. E. Straub, M. C. Thielges, K. Tominaga, A. Tokmakoff, H. Torii, L. Wang, L. J. Webb, and M. T. Zanni, *Chem. Rev.* **120**, 7152 (2020).
- C. Liang and T. L. C. Jansen, *J. Chem. Theory Comput.* **8**, 1706 (2012).
- S. Woutersen and P. Hamm, *J. Chem. Phys.* **114**, 2727 (2001).
- R. Huang, V. Setnička, M. A. Etienne, J. Kim, J. Kubelka, R. P. Hammer, and T. A. Keiderling, *J. Am. Chem. Soc.* **129**, 13592 (2007).
- S. Roy, T. L. C. Jansen, and J. Knoester, *Phys. Chem. Chem. Phys.* **12**, 9347 (2010).
- A. W. Smith, J. Lessing, Z. Ganim, C. S. Peng, A. Tokmakoff, S. Roy, T. L. C. Jansen, and J. Knoester, *J. Phys. Chem. B* **114**, 10913 (2010).
- L. Wang and J. L. Skinner, *J. Phys. Chem. B* **116**, 9627 (2012).
- M. Reppert, A. R. Roy, and A. Tokmakoff, *J. Chem. Phys.* **142**, 125104 (2015).
- C. Falvo, W. Zhuang, Y. S. Kim, P. H. Axelsen, R. M. Hochstrasser, and S. Mukamel, *J. Phys. Chem. B* **116**, 3322 (2012).
- D. B. Strasfeld, Y. L. Ling, R. Gupta, D. P. Raleigh, and M. T. Zanni, *J. Phys. Chem. B* **113**, 15679 (2009).

- <sup>47</sup>S.-H. Shim, R. Gupta, Y. L. Ling, D. B. Strasfeld, D. P. Raleigh, and M. T. Zanni, *Proc. Natl. Acad. Sci. U. S. A.* **106**, 6614 (2009).
- <sup>48</sup>A. S. Bondarenko, I. Patmanidis, R. Alessandri, P. C. T. Souza, T. L. C. Jansen, A. H. de Vries, S. J. Marrink, and J. Knoester, *Chem. Sci.* **11**, 11514 (2020).
- <sup>49</sup>W. Humphrey, A. Dalke, and K. Schulten, *J. Mol. Graphics* **14**, 33 (1996).
- <sup>50</sup>W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.* **110**, 1657 (1988).
- <sup>51</sup>H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *J. Phys. Chem.* **91**, 6269 (1987).
- <sup>52</sup>M. Parrinello and A. Rahman, *Phys. Rev. Lett.* **45**, 1196 (1980).
- <sup>53</sup>U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, *J. Chem. Phys.* **103**, 8577 (1995).
- <sup>54</sup>See <https://github.com/ghlacour/amideimaps> for the amide-I mapping program.
- <sup>55</sup>L. Wang, C. T. Middleton, M. T. Zanni, and J. L. Skinner, *J. Phys. Chem. B* **115**, 3713 (2011).
- <sup>56</sup>T. L. C. Jansen, A. G. Dijkstra, T. M. Watson, J. D. Hirst, and J. Knoester, *J. Chem. Phys.* **125**, 044312 (2006).
- <sup>57</sup>T. L. C. Jansen, A. G. Dijkstra, T. M. Watson, J. D. Hirst, and J. Knoester, *J. Chem. Phys.* **136**, 209901 (2012).
- <sup>58</sup>H. Torii and M. Tasumi, *J. Raman Spectrosc.* **29**, 81 (1998).
- <sup>59</sup>A. S. Bondarenko and T. L. C. Jansen, *J. Phys. Chem.* **142**, 212437 (2015).
- <sup>60</sup>A. V. Cunha, A. S. Bondarenko, and T. L. C. Jansen, *J. Chem. Theory Comput.* **12**, 3982 (2016).
- <sup>61</sup>A. S. Sardjan, F. P. Westerman, J. P. Ogilvie, and T. L. C. Jansen, *J. Phys. Chem. B* **124**, 9420 (2020).
- <sup>62</sup>See [https://github.com/ghlacour/nise\\_2017](https://github.com/ghlacour/nise_2017) for the NISE2017 spectral simulation program.
- <sup>63</sup>T. L. C. Jansen and J. Knoester, *J. Phys. Chem. B* **110**, 22910 (2006).
- <sup>64</sup>K. Kwac, H. Lee, and M. Cho, *J. Chem. Phys.* **120**, 1477 (2004).
- <sup>65</sup>S. Roy, M. S. Pshenichnikov, and T. L. C. Jansen, *J. Phys. Chem. B* **115**, 5431 (2011).
- <sup>66</sup>T. L. C. Jansen and J. Knoester, *Acc. Chem. Res.* **42**, 1405 (2009).
- <sup>67</sup>T. L. C. Jansen, S. Saito, J. Jeon, and M. Cho, *J. Chem. Phys.* **150**, 100901 (2019).
- <sup>68</sup>C. M. Cheatum, A. Tokmakoff, and J. Knoester, *J. Chem. Phys.* **120**, 8201 (2004).
- <sup>69</sup>T. L. C. Jansen and J. Knoester, *Biophys. J.* **94**, 1818 (2008).
- <sup>70</sup>Y. El Khoury, G. Le Breton, A. V. Cunha, T. L. C. Jansen, L. J. G. W. van Wilderen, and J. Bredenbeck, *J. Chem. Phys.* **154**, 124201 (2021).
- <sup>71</sup>E. G. Buchanan, W. H. James, S. H. Choi, L. Guo, S. H. Gellman, C. W. Müller, and T. S. Zwier, *J. Chem. Phys.* **137**, 094301 (2012).