Simulation of Two-Dimensional Infrared Spectroscopy of Peptides Using Localized Normal Modes

Magnus W. D. Hanson-Heine, Fouad S. Husseini, Jonathan D. Hirst, and Nicholas A. Besley*

School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom

ABSTRACT: Nonlinear two-dimensional infrared spectroscopy (2DIR) is most commonly simulated within the framework of the exciton method. The key parameters for these calculations include the frequency of the oscillators within their molecular environments and coupling constants that describe the strength of coupling between the oscillators. It is shown that these quantities can be obtained directly from harmonic frequency calculations by exploiting a procedure that localizes the normal modes. This approach is demonstrated using the amide I modes of polypeptides. For linear and cyclic diimides and hexapeptide Z-Aib-L-Leu-(Aib)2-Gly-Aib-O\textsubscript{t}Bu, the computed parameters are compared with those from existing schemes, and the resulting 2DIR spectra are consistent with experimental observations. The incorporation of conformational averaging of structures from molecular dynamics simulations is discussed, and a hybrid scheme wherein the Hamiltonian matrix from the quantum chemical local-mode approach is combined with fluctuations from empirical schemes is shown to be consistent with experiment. The work demonstrates that localized vibrational modes can provide a foundation for the calculation of 2DIR spectra that does not rely on extensive parametrization and can be applied to a wide range of systems. For systems that are too large for quantum chemical harmonic frequency calculations, the local-mode approach provides a convenient platform for the development of site frequency and coupling maps.

INTRODUCTION

Conventional infrared (IR) spectroscopy is a long-established technique for probing the structure of molecules. For larger molecules, IR spectra often become congested with many overlapping bands that can make spectra hard to interpret, with the consequence that much of the information contained within the spectra is lost. Two-dimensional infrared spectroscopy (2DIR) is the IR analogue to two-dimensional nuclear magnetic resonance spectroscopy, and reviews of this technique are available elsewhere.1−5 In 2DIR spectroscopy, the signal is generated from three ultrashort IR pulses and probes multiple quantum transitions, providing information on the coupling between vibrational modes, vibrational anharmonicities, and line broadening. Another feature of 2DIR spectroscopy is the polarization conditions of the IR pulses, which can be used to suppress or enhance different spectral features. Through spreading the transitions over a second frequency domain, 2DIR reveals more information by exposing cross peaks that correlate with the coupling between different vibrational modes.6,7,12,13 In recent years, 2DIR has emerged as a powerful technique for studying chemical and biological systems.6,11−19 One key advantage of 2DIR, compared with two-dimensional nuclear magnetic resonance, is its time resolution that can probe structural changes on the picosecond time scale. The amide I band of polypeptides and proteins is predominantly associated with the carbonyl stretch of the amide group and shows distinctive spectral characteristics depending on the nature of the secondary structure. The amide I band is particularly amenable to 2DIR spectroscopy, and 2DIR has been applied to study polypeptides and proteins,6−11,13−19 for example, monitoring the unfolding of a β-turn through the evolution of the cross peaks as the β-turn opened and inter-residue hydrogen bonds weakened.13 Complementary to the advancement of experimental 2DIR measurements, computational methods for the simulation of nonlinear spectroscopies have been developed through the inclusion of doubly excited states within the exciton method.6,8 The elements of the one-exciton Hamiltonian are given by

$$\langle \tilde{i} | H_{\tilde{i} \tilde{i}} \rangle = \varepsilon_{\tilde{i}}$$

(1)

$$\langle \tilde{i} | H_{\tilde{j} \tilde{j}} \rangle = \tilde{\beta}_{\tilde{j}}$$

(2)

with the elements of the two-exciton Hamiltonian defined as

$$\langle i_1 \tilde{i} | H_{i_1 \tilde{i}} | i_2 \tilde{i} \rangle = 2\varepsilon_{i_1} - \Delta$$

(3)

$$\langle i_1 \tilde{j} | H_{i_1 \tilde{j}} | i_2 \tilde{j} \rangle = \varepsilon_{i_1} + \varepsilon_{i_2}$$

(4)

$$\langle i_1 \tilde{j} | H_{i_1 \tilde{k}} | i_2 \tilde{k} \rangle = \tilde{\beta}_{i_1 \tilde{k}}$$

(5)

$$\langle i_1 \tilde{j} \tilde{i} | H_{i_1 \tilde{i}} | i_2 \tilde{i} \tilde{j} \rangle = \sqrt{2} \beta_{ij}$$

(6)
where \( \epsilon_i \) and \( \beta_i \) are the site frequencies (in the absence of coupling) and coupling constants between the singly excited states, respectively. \( \Delta \) is the anharmonicity value and corresponds to the difference in the frequency for the fundamental (0 \( \rightarrow \) 1) transition compared with its overtone (1 \( \rightarrow \) 2). For amide I vibrations, a value of 16 cm\(^{-1}\) is usually taken for \( \Delta \). The coupling between singly and doubly excited states is neglected, allowing the resulting Hamiltonian matrix to be separated into zero, one, and two quantum subspaces. The two-quantum transition dipole moments are expressed in terms of the local transition dipole moments according to

\[
\langle i_j | \hat{W}^{kk}_{ll} | i_j \rangle = 0
\]

where \( \hat{W}^{kk}_{ll} \) are the site frequencies and coupling constants between the singly excited states, respectively. \( \Delta \) is the anharmonicity value and corresponds to the difference in the frequency for the fundamental (0 \( \rightarrow \) 1) transition compared with its overtone (1 \( \rightarrow \) 2). For amide I vibrations, a value of 16 cm\(^{-1}\) is usually taken for \( \Delta \). The coupling between singly and doubly excited states is neglected, allowing the resulting Hamiltonian matrix to be separated into zero, one, and two quantum subspaces. The two-quantum transition dipole moments are expressed in terms of the local transition dipole moments according to

\[
\langle i_j | \hat{W}^{kk}_{ll} | i_j \rangle = \sqrt{2} \langle 0 | \mu_{li_j} \rangle
\]

\[
\langle i_j | \hat{W}^{kk}_{ll} | i_j \rangle = \langle 0 | \mu_{li_j} \rangle
\]

Subsequent diagonalization of the Hamiltonian matrix yields the one- and two-exciton energies and eigenstates in addition to the transition dipoles between the eigenstates.

The key quantities for these calculations are the site frequencies and coupling constants. The site frequencies will be affected by the local environment and will be sensitive to factors such as hydrogen bonding and the electrostatic environment, and recently, multipole and dispersion effects have been shown to be significant. The factors affecting the coupling between sites are more subtle, and the relative orientation of the oscillators is often considered to be important. The amide I mode in polypeptides is one of the most well-developed systems for the simulation of 2DIR. In addition to direct quantum chemical harmonic frequency analysis, the amide I mode and its sensitivity to its local molecular environment have been studied via molecular dynamics simulations, an interaction energy decomposition scheme, a property transfer method, building block model, and empirical fitting. For the amide I mode, so-called maps that link the site frequencies and couplings to the molecular structure have been developed. Factors that can be incorporated in quantifying a shift in the site frequency include the electrostatic field at the carbonyl group and a through-bond shift dependent on the dihedral (\( \psi, \phi \)) angles. Other advances include incorporating different solvents and van der Waals forces in the parametrization of the maps, and the accuracy of different models for the IR spectroscopy of the amide I band of proteins has been assessed. Several approaches have also been proposed to describe the coupling between the vibrational modes. Krimm et al. developed the transition dipole coupling model that describes through-space coupling as the interaction between an oscillator dipole with adjacent dipoles. This model has been used in the simulation of the linear absorption spectra of globular proteins. It was shown that the model was not satisfactory for nearest-neighbor couplings, and Hamm and Woutersen proposed the transition charge coupling model that included higher order multipole contributions. Through-bond coupling can be described using maps that account for the dependence of this coupling on the (\( \psi, \phi \)) angles. These maps are usually based on quantum chemical calculations on small model systems, but once the maps are available, calculations on very large systems or sampling of conformation can be undertaken without the need for further quantum chemical calculations.

Here, we elaborate a more direct method to evaluate the site frequencies and couplings based on harmonic frequency calculations. This approach opens the possibility of computing the necessary inputs for an exciton-based 2DIR calculation that captures all elements of the local environment with through-bond and through-space couplings including dispersion and multipole effects. This approach is not specific to the amide I mode and can, in principle, be applied to any system for which the localized modes have an appropriate form. For very large systems where it is not practical to generate a Hessian matrix or to repeatedly generate a Hessian matrix if sampling dynamics, it can aid in the parametrization of appropriate site frequencies or coupling maps.

The harmonic frequencies and normal modes are obtained through diagonalization of the mass-weighted Hessian matrix

\[
H^{(a)} = Q^T H^{(m)} Q
\]

where the harmonic vibrational frequencies are obtained from the square root of the elements of the diagonal matrix \( H^{(a)} \), \( Q \) contains the normal modes in mass-weighted coordinates, and \( H^{(m)} \) is the mass-weighted Hessian matrix

\[
H^{(m)}_{\alpha \beta} = \frac{1}{\sqrt{m_i m_j}} \frac{\partial^2 E}{\partial R_{\alpha i} \partial R_{\beta j}}
\]

\( R_{\alpha i} \) is the \( \alpha \)-component of atom \( i \) with atomic mass \( m_i \). The normal modes do not form an ideal basis for the calculation of 2DIR because they are usually delocalized over the molecule. For example, the two amide I vibrational modes in a diamide can comprise in-phase and out-of-phase combinations of the two carbonyl stretching modes. Because the Hessian is diagonal, there is also no coupling between the normal modes at the level of the harmonic approximation.

Jacob and Reiher introduced a scheme for localizing normal modes. This is achieved by transforming the normal modes by a unitary transformation (\( U \)) that localizes modes by maximizing a criterion, \( \zeta(Q) \). For a set of \( k \) normal modes for a system of \( n \) atoms, the local modes are given by

\[
Q^{loc} = QU
\]

There is no unique function \( \zeta(Q) \), and two approaches were proposed in the original study. In this work, we use the localization criterion that maximizes the sum of the squares of the atomic contributions to the modes

\[
\zeta_{\alpha i}^{loc}(Q^{loc}) = \sum_{p=1}^{k} \sum_{i=1}^{n} \left( \sum_{\alpha=x,y,z} (Q_{\alpha i p}^{loc})^2 \right)
\]

which is analogous to the Pipk and Mezey scheme for the localization of molecular orbitals. This approach has been applied to polypeptides, with the relationship between the IR and Raman spectroscopies of the amide bands and the secondary structure explored, and has more recently been used to aid in the calculation of linear anharmonic vibrational spectroscopy. In the localized mode basis, the Hessian matrix is no longer diagonal and is given by

\[
H^{loc} = U^T H^{(a)} U
\]

The local modes are much more convenient for the simulation of 2DIR because the diagonal elements of \( H^{loc} \) are the squared angular frequencies of the localized modes and because the off-diagonal elements provide a measure of the coupling between
the vibrational modes, with the sign of the coupling chosen to be consistent with the empirical model. Furthermore, the local modes coincide much more closely with the modes as interpreted in the analysis of experimental spectra. A coupling matrix $\Omega$ is computed as

$$\Omega = U^T H U^{1/2}$$

(15)

In this article, we demonstrate that $\epsilon$ and $\beta_{ij}$ derived from the diagonal and off-diagonal elements of $\Omega_{loc}$ can be used to evaluate the matrix elements in an exciton calculation of 2DIR, yielding results that are consistent with both existing computational methods and experimental observations.

## COMPUTATIONAL DETAILS

Following geometry optimization, vibrational frequencies, normal modes, and the corresponding localized normal modes were computed using density functional theory (DFT) with the B3LYP exchange-correlation functional. All quantum chemical calculations were performed with a development version of the Q-Chem software. For calculations on diimides, the 6-311++G** basis set was used and the harmonic vibrational frequencies were scaled by 0.96. For calculations on hexapeptide $\text{Z-Aib-L-Leu-(Aib)$_2$-Gly-Aib-OtBu}$, a mixed basis set was employed using a combination of the 6-31G basis set for C and H atoms together with 6-31+G* for N and O atoms, and the vibrational frequencies were not scaled.

The parameters and 2DIR spectra derived from the localized normal modes are compared with existing site frequency and coupling maps that are currently used to simulate amide I 2DIR spectra, which we refer to as local-mode and empirical approaches, respectively. More specifically, the site frequencies in the empirical maps are determined from the change in frequency for the amide I mode due to the surrounding electrostatic environment according to

$$\epsilon_i = \epsilon_0 + \sum_{j=1}^{4} l_j \phi_{ij}$$

(16)

where $\epsilon_0$ is the unperturbed frequency that is set to 1680 cm$^{-1}$, the summation is over the four atoms of the peptide group, $\phi_{ij}$ is the electrostatic potential at atom $j$ in peptide group $i$, and $l_j$ are linear expansion coefficients. For neighboring groups, the off-diagonal coupling elements ($\beta_{ij}$) are evaluated using a nearest-neighbor coupling map that consists of force constants derived from DFT calculations on a dipeptide for all combinations of main-chain dihedral angles (in increments of 30°). The remaining off-diagonal coupling elements are given by the transition dipole coupling model, with a transition dipole of magnitude 3.7 D Å$^{-1}$ amu$^{1/2}$ placed 0.868 Å away from the amide carbonyl bond and oriented 20° toward the amide nitrogen in the OCN plane.

The one-quantum Hamiltonian matrix elements were used to construct a scaled two-quantum Hamiltonian, comprising two-quantum local states and couplings. These elements are zero, except for the following cases

$$\delta = + - \Delta H_{ij}$$

(17)

$$\delta = + = - H_{ij}$$

(18)

$$\delta = + \neq - H_{ij}$$

(19)

Indices $i, j$, and $k$ refer to different sites; $\delta$ is the Kronecker delta. The anharmonicity, $\Delta$, was fixed at 16 cm$^{-1}$. A modified version of the peptide C program of Hamm and Zanni was used to construct to the two-quantum Hamiltonians and compute the associated 2DIR spectra. The signal is evaluated as the sum of the rephasing and nonrephasing components with the assumptions that the transition dipoles scale like a harmonic oscillator from the fundamental to the overtone (eq 8) and that the dephasing time for the 1→2 transition is the same as that for the 0→1 transition. Polarization conditions have been described previously by Hochstrasser, and we consider 2D signals computed using the ZZZZ polarization condition.
\[
\langle Z_\alpha Z_\beta X_\gamma X_\delta \rangle = \frac{1}{15} \left( \cos \theta_{\alpha\beta} \cos \theta_{\beta\delta} + \cos \theta_{\gamma\delta} \cos \theta_{\delta\beta} 
\right.
\left. + \cos \theta_{\alpha\delta} \cos \theta_{\beta\gamma} \right)
\]
and the ZZXX polarization condition \(^{61}\)
\[
\langle Z_\alpha Z_\beta X_\gamma X_\delta \rangle = \frac{1}{30} \left( 4 \cos \theta_{\alpha\beta} \cos \theta_{\beta\delta} - \cos \theta_{\gamma\delta} \cos \theta_{\delta\beta} 
\right.
\left. - \cos \theta_{\alpha\delta} \cos \theta_{\beta\gamma} \right)
\]
where \(\theta_{mn}\) is the angle between transition dipoles \(m\) and \(n\). The line width was modeled by a Lorentzian function with a magnitude of 10 cm\(^{-1}\).

The four systems studied are illustrated in Figure 1. In order to probe the influence of dynamics on the spectra, classical molecular dynamics (MD) simulations were performed using the NAMD 2.9 molecular dynamics package \(^{62}\) for glycine-alanine-NHMe (GANHMe), 2,5-dioxopiperazine (DKP), and the hexapeptide. Each molecule was solvated in an octahedral prism box of TIP3P water molecules \(^{63}\) with the atoms of the peptide molecules at least 10 Å away from the edge of the box. Periodic boundary conditions were applied. This gave 375, 635, and 2475 water molecules for DKP, GANHMe, and the hexapeptide, respectively. For GANHMe, a Cl\(^-\) counterion was added to neutralize the system. Long-range interactions were described using the particle mesh Ewald method \(^{64}\) and the Lennard-Jones cutoff was 12 Å. Following energy minimization for 15,000 cycles, an equilibration was performed for 0.25 ns with the NVT ensemble and an integration time step of 2 fs during which all covalent bonds involving hydrogen were constrained using the SHAKE algorithm \(^{65}\). Production dynamics were performed for a period of 2 ns in the NPT ensemble using Langevin dynamics and a damping coefficient of 5 ps\(^{-1}\). The Nosé–Hoover \(^{66}-^{68}\) and Langevin piston \(^{69}\) periods were set to 100 fs, and their time-decay period was set to 50 fs to keep the temperature constant at 310 K while maintaining pressure at 1 atm. Snapshots were sampled uniformly to give a total of 1000 snapshots for each system.

**RESULTS AND DISCUSSION**

The amide I mode arises from the carbonyl stretching mode of the amide group with contributions from out-of-phase CN stretching, CCN deformation modes, and the NH in-plane bend. Figure 2 shows the normal modes and localized vibrational modes for the amide I modes in the linear diamide, NAGNMA. The normal modes corresponding to the amide I modes are delocalized across both amide groups, and the extent to which this delocalization occurs depends on the structure of the molecule and the relative orientation of the amide groups. Following localization, the two modes clearly conform to two amide I modes localized on each of the amide groups. Consequently, within the localized mode framework, it is possible to consider two local amide I modes interacting in a 2DIR experiment.

Table 1 shows the computed site energies and coupling for the three diamide molecules considered here. For DKP, the two amide groups are equivalent, the site energies predicted by the scheme of

<table>
<thead>
<tr>
<th>molecule</th>
<th>method</th>
<th>(\epsilon_1/\text{cm}^{-1})</th>
<th>(\epsilon_2/\text{cm}^{-1})</th>
<th>(\beta_{12}/\text{cm}^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>DKP</td>
<td>Jansen</td>
<td>1694.6</td>
<td>1694.6</td>
<td>-13.2</td>
</tr>
<tr>
<td></td>
<td>local mode</td>
<td>1695.3</td>
<td>1695.3</td>
<td>-2.2</td>
</tr>
<tr>
<td>NAGNMA</td>
<td>Jansen</td>
<td>1718.1</td>
<td>1707.8</td>
<td>+6.2</td>
</tr>
<tr>
<td></td>
<td>local mode</td>
<td>1673.5</td>
<td>1658.5</td>
<td>+8.0</td>
</tr>
<tr>
<td>GANHMe</td>
<td>Jansen</td>
<td>1685.5</td>
<td>1683.1</td>
<td>+6.0</td>
</tr>
<tr>
<td></td>
<td>local mode</td>
<td>1681.1</td>
<td>1665.9</td>
<td>+4.9</td>
</tr>
</tbody>
</table>
Jansen and the local-mode approach give $\varepsilon_1$ and $\varepsilon_2$ to be equal, and both methods give frequencies in good agreement with the experimentally observed value of 1697 cm$^{-1}$. For the two linear diamides, the two sites are no longer equivalent. However, the local-mode approach gives a significantly larger difference between the two frequencies. For these systems, there is no hydrogen bonding present, so this is a through-bond effect. There is also significant variation in the strength of the predicted coupling elements ($\beta_{12}$). Within the local-mode approach, the magnitude of the coupling has the following trend: NAGNMA > GANHMe > DKP. This suggests that additional molecular structure between the two carbonyl groups.

Figure 3. Computed full absorptive 2DIR spectra for the ZZZZ pulse polarization scheme for (a) DKP, empirical; (b) DKP, local mode; (c) NAGNMA, empirical; (d) NAGNMA, local mode; (e) GANHMe, empirical; and (f) GANHMe, local mode. All are based on the minimum energy structure. Regions of zero intensity have been colored white for clarity.
results in a reduction in the strength of the coupling, which is physically intuitive for a through-bond effect. This trend is not observed in the coupling from the Jansen model, which predicts the strongest coupling to occur in DKP. There is a significant difference in the method to compute the site frequencies between the two models. In the Jansen scheme, the frequencies are determined through the perturbation of a reference frequency arising from the electrostatic environment, whereas the local mode frequencies are derived from the DFT calculations within the harmonic approximation. Consequently, differences between the resulting frequencies are not surprising. There is greater similarity between the two models for the description of the coupling strengths. In the Jansen model, the coupling is derived from DFT calculations on dimers using the

Figure 4. Computed 1DIR spectra of (a) DKP, (b) NAGNMA, and (c) GANHMe. Black lines, empirical model; red lines, local-mode model.

Figure 5. Variation of the site energies $\varepsilon_1$ (top), $\varepsilon_2$ (center), and coupling $\beta_{12}$ (bottom) of a dipeptide with $\psi$ and $\phi$ angles. Left, empirical model; right, local-mode model. Blue contours represent positive regions, and red contours represent negative regions. The separation between contours is 5 cm$^{-1}$ for the site energy maps and 2 cm$^{-1}$ for the coupling.
Hessian matrix reconstruction approach. In this approach, the amide I mode is localized on the carbon and oxygen atoms of the carbonyl group, whereas in the local-mode description, the amide I mode also contains significant motion of the nitrogen atom (Figure 2), which is more consistent with the true nature of the amide I mode. This difference in the description of the amide I vibration between the two schemes is likely to result in different values for the computed coupling.

Figure 3 shows the computed 2DIR spectra for the ZZZZ pulse polarization scheme based on the minimum energy structure for the empirical (Jansen) and local-mode models, with the corresponding linear IR spectra shown in Figure 4. For DKP, all of the intensity is retained within one peak, and no cross peaks are evident. In the spectrum based on the empirical maps, the peak is shifted to a higher frequency, 1720 cm$^{-1}$, compared with 1710 cm$^{-1}$, which is a consequence of the larger coupling term. For the linear diamides, two distinct peaks are observed in the spectra. For the empirical model, the positive lobe of the peak at higher frequency is barely evident, whereas it can be clearly distinguished in the local-mode spectra. Cross peaks can also been seen for these molecules. For the empirical spectra, these tend to merge into the main peaks, whereas in the local-mode spectra, they are more clearly separated from the main peaks. This is a consequence of the greater separation between the main peaks in the local-mode spectra, which arises from the greater splitting between the site energies. The linear IR spectra have a similar band profile for the empirical and local-mode models. However, the local-mode model predicts the bands to lie at lower frequencies, and a greater splitting between the two distinct peaks in the linear diamides is evident on closer inspection. These calculations do not include a description of the solvent. The solvent can have a large effect on the computed frequencies, and within a large polypeptide, amide groups can be fully exposed to the solvent or buried in hydrophobic regions. Within the DFT calculations, solvent can be described by the inclusion of specific molecules within the calculation. However, such calculations would be computationally expensive and not practical for many systems. Continuum solvent can provide an efficient alternative that, if necessary, computational approaches are available, does not add significantly to the computational cost, and recent work has shown these solvent models to provide a robust approach for the simulation of the IR spectroscopy of proteins in solution.

At a lower level of computational cost, the electrostatic component of the solvent interaction can be incorporated via eq 16. These spectra are based on a single structure, whereas simulating experimental measurements requires averaging over many structures. These can be achieved through averaging many spectra computed using structural snapshots extracted from MD simulations. For the local-mode approach, this is not straightforward because harmonic frequency analysis requires the structure to be at a minimum. Possible solutions include incorporating empirical-based broadening to the diagonal elements of the Hamiltonian matrix or constrained optimization, whereby the regions of the molecule where the vibrational modes are localized are allowed to relax while the remainder of the molecule is held fixed. Alternatively, the effects of dynamics can be included through the use of the local-mode approach in the generation of maps to describe the site energies and coupling or, as discussed later, through hybrid schemes wherein the local-mode Hamiltonian is combined with fluctuations arising from the dynamics evaluated using empirical approaches.

Table 2. Maxima and Minima in the Site Frequency and Coupling Maps

<table>
<thead>
<tr>
<th>method</th>
<th>quantity</th>
<th>maximum and minimum/cm$^{-1}$</th>
<th>(ψ,φ)/deg</th>
</tr>
</thead>
<tbody>
<tr>
<td>local mode</td>
<td>$\epsilon_1$</td>
<td>1825 (1752)</td>
<td>(0,180), (0,−180)</td>
</tr>
<tr>
<td></td>
<td>$\epsilon_2$</td>
<td>1750 (1680)</td>
<td>(−80,60), (−60,80)</td>
</tr>
<tr>
<td></td>
<td>$\epsilon_3$</td>
<td>1838 (1764)</td>
<td>(0,180), (0,−180)</td>
</tr>
<tr>
<td></td>
<td>$\phi_{12}$</td>
<td>1773 (1702)</td>
<td>(0,−20), (0,20)</td>
</tr>
<tr>
<td></td>
<td>$\beta_{12}$</td>
<td>20.1</td>
<td>(0,180), (0,−180)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−15.0</td>
<td>(0)</td>
</tr>
<tr>
<td>empirical</td>
<td>$\epsilon_1$</td>
<td>1739</td>
<td>(0,180), (0,−180)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1693</td>
<td>(−80,30), (−30,−80)</td>
</tr>
<tr>
<td></td>
<td>$\epsilon_2$</td>
<td>1745</td>
<td>(0,180), (0,−180)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1702</td>
<td>(0,0)</td>
</tr>
<tr>
<td></td>
<td>$\beta_{12}$</td>
<td>21.3</td>
<td>(0,180), (0,−180)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−14.5</td>
<td>(0,0)</td>
</tr>
</tbody>
</table>
there is a reasonable quantitative agreement between the two approaches. The most significant differences occur in the location of the minima for the lower site frequencies, and the local-mode approach predicts a larger difference between the highest and lowest frequencies. Spectra that incorporate structural averaging for DKP and GANHMe are shown in Figure 6, with the linear IR spectra shown in Figure 7. In addition to the inclusion of conformational averaging, these spectra also incorporate a description of the solvent via an electrostatic term added to the diagonal elements (eq 16). The spectra are more typical of experimental 2DIR spectra and are spread along the diagonal, with the spectra for DKP being more compact. This is a result of there being only one peak in the spectrum for the minimum energy structure of this molecule. The similarity of the maps is reflected in the computed spectra, and there is little difference between the spectra computed with the different maps.

For diamide systems, there is a high degree of similarity between the empirical and local-mode approaches. However, for larger systems, the two schemes become more distinct.

Figure 6. Computed full absorptive 2DIR spectra for the ZZZZ pulse polarization scheme incorporating averaging over conformation for (a) DKP, empirical; (b) DKP, local mode; (c) GANHMe, empirical; and (d) GANHMe, local mode. Regions of zero intensity have been colored white for clarity.

Figure 7. Computed 1DIR spectra incorporating averaging over conformation for (a) DKP and (b) GANHMe. Black lines, empirical model; red lines, local-mode model.

Within the local-mode approach, all of the interactions between the amide groups are treated at an equivalent level, whereas in the empirical scheme, non-nearest neighbor interactions are treated at a more approximate level. Furthermore, the site energies incorporate electrostatic, through-bond, and hydrogen bond effects explicitly. To explore this further, we consider
The nature of carbonyl groups in the hexapeptide.

Figure 8. Nature of carbonyl groups in the hexapeptide.

Table 3. Calculated Site Energies and Coupling Constants (in cm$^{-1}$) for the Hexapeptide from Local-Mode and Empirical Models

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>local mode</td>
<td>1718.9</td>
<td>3.8</td>
<td>3.9</td>
<td>1.4</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>1692.8</td>
<td>3.4</td>
<td>2.4</td>
<td>0.8</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>1678.0</td>
<td>5.4</td>
<td>2.5</td>
<td>1.5</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1682.4</td>
<td>2.6</td>
<td>2.1</td>
<td>0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1692.7</td>
<td>1.6</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1686.9</td>
<td>4.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1739.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>empirical</td>
<td>1679.0</td>
<td>0.9</td>
<td>1.5</td>
<td>0.6</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1667.2</td>
<td>1.0</td>
<td>1.8</td>
<td>1.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1669.6</td>
<td>0.2</td>
<td>1.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1668.6</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1653.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

“Within the empirical model, only the peptide carbonyls are considered.
conformational dynamics is not feasible for a molecule as large as a hexapeptide. We also note that these calculations are formally valid only at a minimum energy structure; however, the use of partially constrained optimization is widely used. In order to overcome this limitation, we adopt a hybrid approach where the site energies and coupling constants for a conformation $n$ are expressed as

$$
\epsilon_i(n) = \epsilon_i^{\text{local,min}} + \Delta \epsilon_i^{\text{emp}}(n)
$$

where

$$
\Delta \epsilon_i^{\text{emp}}(n) = \epsilon_i^{\text{emp}}(n) - \epsilon_i^{\text{emp,min}}
$$

and $\epsilon_i^{\text{local,min}}$ is the site energy for the local-mode approach at the minimum energy structure, $\epsilon_i^{\text{emp,min}}$ is the empirical site energy at the minimum energy structure, and $\epsilon_i^{\text{emp}}(n)$ is the empirical site energy at conformation $n$. Similarly, for the coupling constants

$$
\beta_{ij}(n) = \beta_{ij}^{\text{local,min}} + \Delta \beta_{ij}^{\text{emp}}(n)
$$

$$
\Delta \beta_{ij}^{\text{emp}}(n) = \beta_{ij}^{\text{emp}}(n) - \beta_{ij}^{\text{local,min}}
$$

Within this framework, the variation in the matrix elements arising from the presence of the solvent will also be included. Figure 11 shows the simulated spectra for the peptide groups, with the urethane and ester groups excluded, incorporating averaging over molecular conformation. Compared with the spectra for the single minimum energy structure, both spectra show considerable elongation of the signal along the diagonal, and the distinct peak at low frequency observed for the static empirical spectrum is no longer observed and is merged with the main peak. Closer inspection of the spectra shows some significant differences between the two approaches. The hybrid spectrum is shifted to higher frequency and the empirical spectrum shows distinct peaks within the main signal, whereas these are not evident in the hybrid spectrum, which gives a smooth single peak. Both of these effects are likely to be a consequence of the modification of the diagonal elements.

Figure 12 shows absolute magnitude 2DIR spectra for ZZXX polarization. These spectra can be compared to the experimental spectrum reported by Maekawa et al. (see Figure 3a of ref 28). The experimental spectrum has a large signal centered at about 1656 cm$^{-1}$ arising from the peptide groups. Both empirical and local-mode spectra for the peptide groups are qualitatively similar. However, the empirical spectrum is centered at 1643 cm$^{-1}$, whereas the hybrid spectrum is centered at 1658 cm$^{-1}$. The shape of the signal is roughly elliptical, and it is possible to obtain an approximate measure of the shape by considering the ratio of the long and short axes of the ellipse. In the experimental spectrum, this ratio is approximately 1.7, and the corresponding values for the empirical and hybrid spectra are 2.3 and 2.0, respectively. Both of these measures suggest an improvement in the computed spectrum for the local-mode approach. However, we note that direct comparison with experimental is not possible because the experiment spectrum was measured in a different solvent, which can affect the 2DIR signal. Furthermore, other factors, such as the vibrational motional narrowing effect, which describes a phenomenon in which the line width may be overestimated by static structural snapshots, may also influence the computed spectrum. These effects have been shown to be significant for N-methylacetamide in solution, but it has been suggested that these effects become less important for polypeptides because of the spread of the multiple amide modes of the protein. However, in the calculations presented, the broadening of the lines arising from sampling structures from the MD simulations is the same for both the empirical and
hybrid local-mode approaches, so there is no difference between the two approaches in the treatment of these effects.

### CONCLUSIONS

It has been shown that the site frequencies and coupling constants that are required for the simulation of 2DIR spectra with the exciton method can be evaluated directly from quantum chemical harmonic frequency calculations by exploiting a transformation to localized vibrational modes. Such an approach opens the possibility for simulating 2DIR spectra without the need for extensive \textit{a priori} parametrization. This scheme has been demonstrated by considering the 2DIR spectroscopy of the amide I band of polypeptides. For the linear and cyclic diamide systems considered, there are significant differences between the site frequencies and coupling constants computed using the local-mode approach and established empirical models. Within the local-mode approach, there is a greater separation between nonequivalent site frequencies and the coupling constants show a physically reasonable trend of decreasing as the molecular structure between the carbonyl groups increases. For the hexapeptide, the local-mode approach is able to treat all carbonyl groups and their respective interactions at an equivalent level of theory. There is agreement with the available experimental data of a...
large peptide peak at about 1675 cm$^{-1}$ and a peak at higher frequency arising from the urethane and ester carbonyl groups. A challenge for the local-mode approach is its associated computational cost because a harmonic frequency calculation is required for each structure considered. However, the local-mode approach can be incorporated into exciton-based calculations of 2DIR in several different ways. For small systems, the local-mode approach can be used directly, whereas for moderately sized polypeptides, the local-mode approach can be used within a hybrid scheme, wherein the fluctuations arising from conformational dynamics are evaluated using a computationally much cheaper method. For calculations of the 2DIR frequency arising from the urethane and ester carbonyl groups, the local-mode approach can provide a platform for the parametrization of suitable site frequency and coupling maps that specify the value of these parameters with respect to critical geometrical coordinates. The local-mode approach described here is particularly suitable for this because it can be applied directly to different functional groups.

**AUTHOR INFORMATION**

**Corresponding Author**  
*E-mail: nick.besley@nottingham.ac.uk.*

**Funding**  
M.W.D.H.-H. acknowledges the University of Nottingham and the EPSRC (EP/L50502X/1) for funding. N.A.B. thanks the Leverhulme Trust for a Research Fellowship (RF-2014-231) and the EPSRC for funding (Grant No. EP/L021366/1). We also acknowledge the Saudi Cultural Bureau for funding.

**Notes**  
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**  
We thank the University of Nottingham for access to its High Performance Computing and MidPlus facilities and Neil Hunt (Strathclyde) and Tony Parker (RAL) for useful discussions.

**REFERENCES**


(30) Torii, H.; Tasumi, M. Ab initio molecular orbital study of the amide I vibrational interactions between the peptide groups in di- and...
tripeptides and considerations on the conformation of the extended helix. J. Raman Spectrosc. 1998, 29, 81−86.


