PCCP

RSC Publishing

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/c3cp51347e

Two-dimensional stimulated resonance Raman spectroscopy study of the Trp-cage peptide folding†

Hao Ren,[a] Zaizhi Lai,[b] Jason D. Biggs,[a] Jin Wang[b] and Shaul Mukamel*[a]

Received 29th March 2013,
Accepted 11th September 2013

DOI: 10.1039/c3cp51347e

www.rsc.org/pccp

We report a combined molecular dynamics (MD) and *ab initio* simulation study of the ultrafast broadband ultraviolet (UV) stimulated resonance Raman (SRR) spectra of the Trp-cage mini protein. Characteristic two dimensional (2D) SRR features of various folding states are identified. Structural fluctuations erode the cross peaks and the correlation between diagonal peaks is a good indicator of the α-helix formation.

I. Introduction

Protein folding is a fundamental problem in molecular biology because it is one of the key processes in genetic central dogma, and also because some diseases are directly related to misfolding.[1–4] Extensive theoretical and experimental studies have provided considerable insight into the folding process. According to the current view, protein folding proceeds on a moderately rough energy surface,[5–14] whose major features are the local minima and the overall downhill slope toward the native state, forming a rough landscape with funnel-like shape on which the protein folds. Uncovering the folding mechanism in detail requires experimental methods with high temporal and structural resolution. Conventional kinetics and thermodynamics experiments of folding do not usually reveal conventional microscopic structural changes in detail and with sufficient high temporal resolution. Optical spectroscopy techniques can probe such details and transformations of biological molecules on the pico to nanosecond time scale. Ultrafast broadband pulses in the infrared (IR) or ultraviolet (UV) regimes have been used recently to probe the structure of proteins.[15,16] Two-dimensional IR[17–27] or UV[4,28–31] spectra are able to provide detailed information with high temporal and spatial resolution by monitoring the correlations between vibrational or electronic excitations.

Raman spectroscopy is a useful probe, thanks to the sensitivity of vibrations to the local chemical environment. Ultraviolet resonance Raman (UVRR) spectroscopy selectively enhances vibrational modes strongly coupled to the selected electronic chromophore. Water vibrational modes can then be easily filtered out, making UVRR suitable for biological systems *in vivo* or in aqueous environments. Peptides or side-chain vibrational modes can be selectively enhanced by tuning the excitation frequency in resonance with selected electronic transitions, making the technique a useful tool for characterizing protein secondary structure.[32–42] Ultrafast broad-band coherent Raman spectroscopy is widely used to investigate the dynamics of intra- and inter-molecular vibrational modes.[43–49] Tanimura and Mukamel have proposed an off-resonance 2D stimulated fifth-order Raman scattering experiment with two pump pulses followed by a probe.[50] The two-dimensional (2D) stimulated Raman signal is displayed after performing a double Fourier transform with respect to the two time delays. Extending this technique to the resonance regime is not straightforward. In a multi-pulse experiment, resonant pulses can induce considerable excited state populations, especially for species with long excited state lifetimes compared with the pulse duration ($\sim$100 fs), which complicate the analysis of the Raman signals. Recently, we showed that pre-resonance pumps followed by a resonance probe can eliminate the excited state contribution while preserving the resonant selectivity of vibrational modes.[42] This technique provides detailed information about the correlations of different vibrational modes. Pulse polarizations offer additional control parameters for these signals.[51]

Previously it has been shown that when the two-dimensional Raman signal is collected using non-resonant pulses, the sought-after $\chi^{(5)}$ signal is overwhelmed by third-order cascades,[52–57] both of which have the same power dependence. The cascading signal is the result of two $\chi^{(3)}$ processes taking place on different molecules in the illuminated volume, and carries the same information as one-dimensional Raman. In the present work we use pump and probe pulses which are pre-resonant and in resonance with a given electronic transition, respectively. Due to the enhancement of the polarizability on resonance, the signal can be collected using a smaller sample concentration.

[a] Department of Chemistry, University of California, Irvine, California, 92697, USA. E-mail: smukamel@uci.edu
[b] Department of Chemistry, University of New York at Stony Brook, Stony Brook, New York 11794, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3cp51347e

The probability that the third order signal induced by the first two pulses and emitted from one molecule will combine with the third pulse to generate another third-order signal in a second molecule will decrease sharply when the concentration decreases.

In this simulation study, we report the two dimensional stimulated resonance Raman (2DSRR) signals from five states along the nanosecond folding process of the Trp-cage protein, and show that they contain distinct characteristic features of different folding states, and can be used to monitor the folding pathway. Trp-cage is a mini-protein, which contains 20 residues with the sequence "NLYIQWLKDGGPSSGRPPPS". Its fast nanosecond folding makes it a convenient model system for experimental and computational investigations.[21,58–60] Spectra with various pulse polarization configurations were simulated at different folding stages obtained by molecular dynamics (MD) simulations and trajectory analysis. This work complements our early 2DIR[21] investigation of the same system.

II. Theory

2DSRR is a three-pulse pump–pump–probe experiment. The signal is defined as the change in the integrated intensity of the probe pulse due to the presence of both pumps, and is recorded *versus* the two inter-pulse delays t_1 and t_2 (ref. 42) (see Fig. 1). To minimize the excited-state non-Raman contribution, the two pump pulses were tuned to be preresonant with the electronic excitations. The signal is represented by the four loop diagrams shown in Fig. 1.[42,61] In these diagrams, photon absorption and emission are represented by arrows going into and out of the loop, respectively. Moving along the loop clockwise, the left branch shows forward (ket) time evolution, while the right branch shows backward (bra) time evolution. Detailed diagram rules can be found in ref. 42 and 61. The first pulse creates a vibrational excitation in either the ket (S_{i} and S_{iv}) or the bra (S_{ii} and S_{iii}). The second pulse can either create another pure vibrational excitation on the opposite side of the density matrix (S_{i} and S_{iii}) or change the excitation created by the first (S_{ii} and S_{iv}). The probe pulse was tuned to be resonant with the electronic transition of interest. The 2D signal is recorded by varying the two time delays t_1 and t_2 between the three pulses and taking a 2D Fourier transform, $t_1 \rightarrow \Omega_1$, and $t_2 \rightarrow \Omega_2$,

$$S(\Omega_1,\Omega_2) = \sum_{\mathrm{acc'}} \frac{P(\mathrm{a})\alpha^{(1)}_{\mathrm{ca}}}{\Omega_1-\omega_{\mathrm{ca}}+i\Gamma} \times \left(\frac{\left(\alpha^{(2)}_{\mathrm{ac'}}\right)^* \left[\alpha^{(3)}_{\mathrm{c'c}} - \left(\alpha^{(3)}_{\mathrm{c'c}}\right)^*\right]}{\Omega_2-\omega_{\mathrm{cc'}}+i\Gamma} - \frac{\alpha^{(2)}_{\mathrm{c'c}}\left[\alpha^{(3)}_{\mathrm{ac'}} - \left(\alpha^{(3)}_{\mathrm{ac'}}\right)^*\right]}{\Omega_2-\omega_{\mathrm{c'a}}+i\Gamma} \right) \quad (1)$$

where $P(\mathrm{a})$ is the occupation probability of state $|\mathrm{a}\rangle$, and Γ = 10 cm^{-1} is the vibrational lifetime. The effective transition polarizability from state $|\mathrm{c}\rangle$ to $|\mathrm{c}'\rangle$ due to the mth pulse with polarization $\boldsymbol{e}_m$ is given by,

$$\alpha^{(m)}_{\mathrm{c'c}} = \left\langle \mathrm{c}'|\hat{\alpha}^{(m)}|\mathrm{c}\right\rangle = \frac{1}{2\pi}\sum_{\mathrm{b}} \langle \mathrm{c}'|\mathrm{b}\rangle\langle \mathrm{b}|\mathrm{c}\rangle|\boldsymbol{e}_m\cdot\boldsymbol{\mu}_{\mathrm{eg}}|^2 \int_{-\infty}^{\infty} \mathrm{d}\omega \frac{\varepsilon^*_m(\omega)\varepsilon_m(\omega+\omega_{\mathrm{c'c}})}{\omega+\omega_n+\omega_{\mathrm{bc'}}+i\Gamma} \quad (2)$$

here $\hat{\alpha}^{(m)}$ is the transition polarizability operator that corresponds to the mth pulse, $|\mathrm{c}\rangle$ and $|\mathrm{c}'\rangle$ are vibrational states in the ground electronic state, and $|\mathrm{b}\rangle$ is a vibrational state in the excited electronic state. $\boldsymbol{\mu}_{\mathrm{eg}}$ is the transition dipole moment between electronic states $|\mathrm{e}\rangle$ and $|\mathrm{g}\rangle$ and $\mathcal{E}(\omega)$ is the pulse envelop. The Franck–Condon factors $\langle m|n\rangle$ are evaluated using the linearly displaced harmonic oscillator model of molecular vibronic excitations.[41,42] Detailed deviation of eqn (1) and (2) can be found in Appendix A of ref. 42.

Eqn (1) gives peaks for positive Ω_1, since this is a pump–probe measurement, and $S(-\Omega_1, -\Omega_2) = [S(\Omega_1, \Omega_2)]^*$,[62] we can focus on the region $\Omega_1 \geq 0$ without lose of information. The signal is averaged over an isotropic distribution of molecules.[51] $\hat{\alpha}^{(n)}$ is a non-Hermitian operator in the vibrational subspace, and becomes real Hermitian far off-resonance. The diagonal elements of the effective polarizability (α^n_{jj}) correspond to an elastic (*i.e.* Rayleigh) scattering of the nth pulse; the top row elements ($\alpha^{(n)}_{0,j}, j \neq 0$) correspond to a Raman process from the ground vibrational states to a excited state with one vibrational quanta in the jth mode; the off

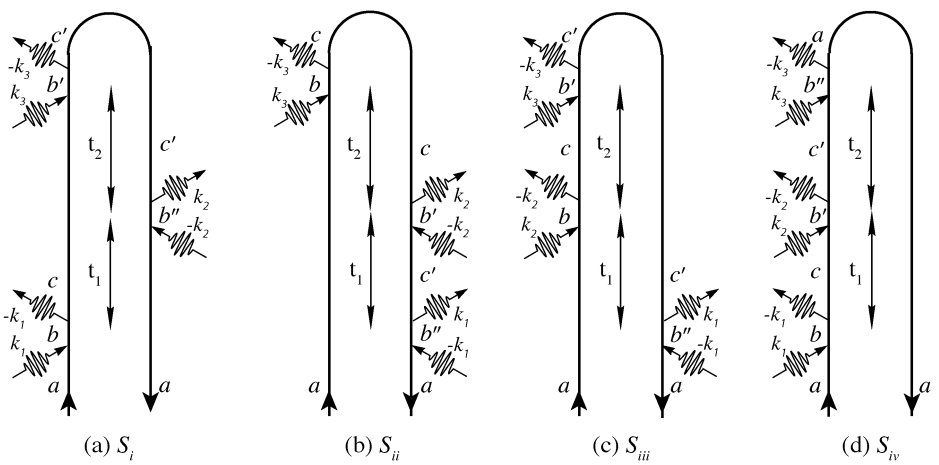

Fig. 1 Loop diagrams for the 2DSRR signal.

diagonal elements ($\alpha^{(n)}_{jk}$, $j \neq k$) represent vibrational quanta transfers from the kth mode to the jth mode *via* a common electronic excitation, which reveals the correlation between these two modes. Typically the off-diagonal elements are one to two orders of magnitude weaker than the diagonal ones.[42]

2DSRR spectra reveal correlations between different vibrational modes.[42] The diagonal peaks come from pathways S_i and S_{iv} and are proportional to $(\alpha^{(2)}_{00})^*(\alpha^{(3)}_{0,j})(\alpha^{(1)}_{j,0})$ and $(\alpha^{(3)}_{0,j})(\alpha^{(2)}_{j,j})(\alpha^{(1)}_{j,0})$, respectively. Both components consist of one elastic and two inelastic Raman contributions corresponding to the jth mode. Diagonal peaks are the strongest features. The off-diagonal peaks contain an off-diagonal element of transition polarizability $\alpha^{(n)}_{j,k}$ and two top row elements $\alpha^{(n)}_{k0}$ and $\alpha^{(n)}_{0j}$, they reflect the correlation between two Raman active vibrational modes when the system interacts with the nth pulse. By focusing on the electronic transitions localized on the tryptophan or tyrosine aromatic groups, we selectively observe cross peaks corresponding to vibrational modes localized on these residues.

III. Computational protocol

A. Molecular dynamics simulations

We carried out all the MD simulations and part of the analysis using AMBER 10 software package[63] with the AMBER ff99SB protein force field.[64] A constant temperature of 315 K was maintained in the MD simulations. An implicit solvation model[65] with a collision frequency of 1 ps^{-1} was used to simulate the solvent environment. The SHAKE algorithm[66] was used to constrain covalent bonds involving hydrogen atoms. The time step was set to 2 fs. 50 trajectories were simulated for 200 ns each. The initial structure of each trajectory was given by a completely extended conformation and the different atom velocities from a Gaussian distribution[67] were assigned to the different trajectories to start the simulations. The total 10 μs simulations provided enough data to construct the free energy landscape (FEL). Five folding states were chosen along the dominant folding pathway from the unfolded state to the folded state on the FEL. For each folding state, 200 snapshots around that location were harvested to calculate the Raman signals.

B. *Ab initio* calculation of electronic transitions and vibronic coupling

Trp-cage has two aromatic residues Tyr3 and Trp6. The electronic transitions in the near UV ($\geq$210 nm, $\leq$47 620 cm^{-1}) region arise from these residues. The geometry and relative positions of Tyr3 and Trp6 residues were extracted from the MD trajectories, all other atoms are treated as electrostatic background with charges taken from the force field parameters.[64] Electronic excitations were calculated by using time-dependent functional theory (TDDFT) with the B3LYP hybrid functional[68,69] and the 6-311++G(d,p) basis set implemented in the Gaussian 09 package.[70] The conductor-like polarizable continuum model (CPCM)[71,72] was used in the self consistent reaction field calculations in the aqueous environment. All the vibrational frequencies were scaled by a factor of 0.97[73,74] to correct for the systematic error in the density functional frequency calculations.

C. Simulations of spectra

The UV absorption spectrum of each folding state was obtained by using the cumulant expression,[42,75] averaging over 200 snapshots, calculated using all the dipole allowed electronic transitions below 47 620 cm^{-1} (210 nm). The 2D Raman signals were obtained using Gaussian pulses with 1768 cm^{-1} (8.3 fs) full width at half maximum (FWHM), and a vibrational linewidth Γ = 10 cm^{-1}. Preresonant pumps with center frequencies $\omega_1 = \omega_2$ = 31 000 cm^{-1} were used to avoid overlap with any absorption peaks (see Fig. 5), allowing excited state non-Raman contributions to be neglected.[42] Three pulse polarization schemes were used: XXX, all parallel polarization; XXY, where the first two pulses are parallel and the third one is perpendicular to the first two; finally we used the super magic angle (SMA)[51] polarization combination to filter out the anisotropic signals.[51] Vibrational modes with frequencies in the range 600–1800 cm^{-1} (50–60 modes for each snapshot) were incorporated by the cumulant expression.[42,75]

IV. Results and discussion

A. The free energy profile

Fig. 2 shows the free energy profile of Trp-cage folding plotted as a function of the root mean square deviation (RMSD) and the radius of gyration (R_g). It was obtained using $F = -\log(P)$, where P is the population obtained from the 10 μs MD simulated data. The C_α-RMSD of the lowest energy structure is less than 2 Å, implying the peptide reaches its folded state.[14,76] Raman spectra were calculated for the five states S1, S25, S50, S75, and S100, along the folding pathway. S1 and S25 are largely linear and random coil structures. After S50 the peptide collapses into a relatively compact conformation with formed orientations of the two strands (N-terminal α-helix and C-terminal coil), dramatically reducing the conformational configuration space, which accelerates the peptide evolution towards the folded state.

B. UV absorption and UVRR

Isolated tyrosine and tryptophan amino acids have six aromatic electronic transitions below 47 620 cm^{-1} ($\geq$210 nm) as listed in Table 1. These TDDFT results have been compared extensively with UV absorption experiments and reasonably good agreement has been achieved.[42] The excitation energies fluctuate in different chemical environments as the protein folds.[31] To avoid overlap with the peptide n–π* transitions around 45 000 cm^{-1}, we focus on the aromatic excitations below 42 000 cm^{-1} dominated by the three transitions Trp-L_b, Tyr-L_b, and Trp-L_a. As shown in Fig. 3, the L_b and L_a transitions in both tyrosine and tryptophan lie in the aromatic ring plane and are perpendicular.

The UV absorption spectra of the five folding states shown in the left column of Fig. 5 are similar. The strong peak above 42 000 cm^{-1} can be assigned to the Trp-B_b and Tyr-L_a transitions, and the weaker peak below 41 000 cm^{-1} is assigned to Trp-L_b, Tyr-L_b and Trp-L_a. By decomposing the dominating configuration interaction (CI)-like coefficients into atomic basis contributions, we found that all transitions considered here are localized either on the

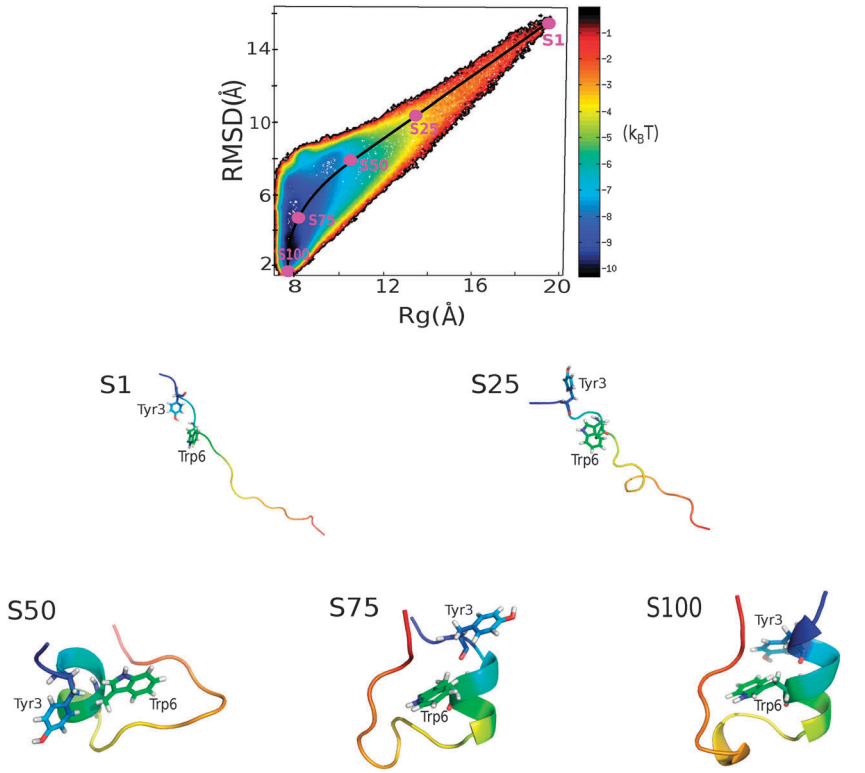

Fig. 2 Free energy profile of Trp-cage folding (top). Five states labeled S1, S25, S50, S75, and S100 were chosen along the folding pathway. These structures are shown in the bottom panel.

Table 1 ≥200 nm UV electronic transitions of isolated tyrosine and tryptophan amino acids

System	Excitation	λ/nm	$\tilde{\nu}$/cm^{-1}	Dipole/Debye				Description
				x	y	z	$\lvert\mu\rvert^2$	
Tyr		251	39 840	0.2591	0.8256	−0.0460	0.7509	L_b
		222	45 063	−1.4171	0.1179	0.4182	2.1970	L_a
Trp		268	37 228	0.5406	1.1506	−0.2523	1.6856	L_b
		253	39 484	−0.6788	0.6651	−0.3767	1.0450	L_a
		222	45 023	1.9159	−1.5794	0.8076	6.8160	B_b
		215	46 351	0.5879	−0.7750	0.5128	1.2092	B_b'

tryptophan or the tyrosine aromatic groups. The absorption spectra are decomposed into tyrosine (green dashed lines) and tryptophan (red dashed lines) contributions in the left column of Fig. 5, showing that tryptophan transitions dominate below 41 000 cm^{-1}.

The 230 nm UV resonance Raman spectra of the five folding states are depicted in Fig. 4 and compared with the previous experiment.[60] There are five strong peaks in all the calculated spectra, according to Harada and Takeuchi's assignment[77,78] of Raman bands of Trp and Tyr, these peaks can be attributed to 1610 (Y8a), 1570 (W3), 1460 (W8a), 1030 (W16) and 860 cm^{-1} (W17). The 1344 (W7) and 766 cm^{-1} (W18) peaks are weaker in our calculation compared with experiments. The UVRR spectra are similar for the five folding states, except the relative intensities of the 1610 and 1570 peaks. By checking the vibrational mode for the ensembles composed of each folding state, we found that there are strong inhomogeneous structural distributions in the

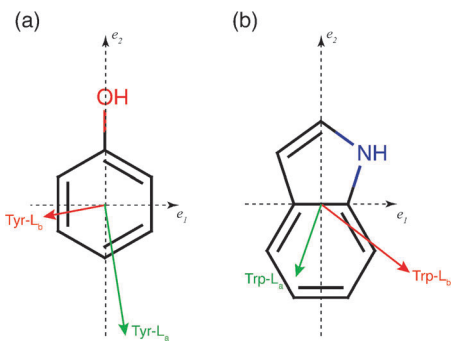

Fig. 3 Molecular structure and dipole moments of the aromatic chromophores of (a) L-tyrosine and (b) L-tryptophan. e_1 and e_2 axes define the aromatic ring plane where the L_b and L_a transitions lie.

S1 states, which result in part of the snapshots having W3 mode with frequencies around 1610 cm^{-1}. In S100, the 1610 cm^{-1}

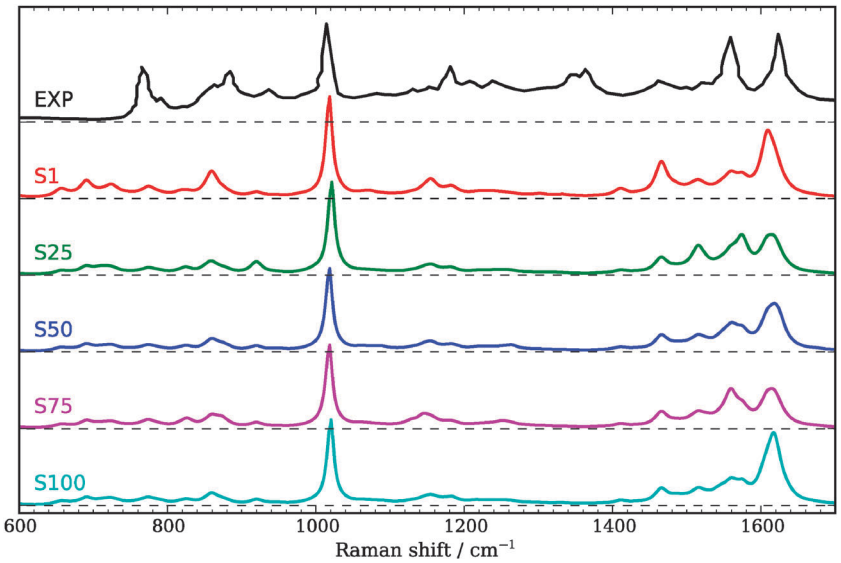

Fig. 4 Calculated 230 nm (43 480 cm^{-1}) UVRR spectra of the five folding states compared with experiments. Experimental spectrum taken from ref. 60.

peaks are much stronger than the 1570 cm^{-1} peaks, since the Trp residue resides in a rigid environment composed of the α-helix and two strand structure, the stronger interaction with other residues blueshifted the W3 mode. Although there are evident differences in the spontaneous resonance Raman spectra of the five folding states, it is hard to characterize these states only using UVRR spectra. We turn to the 2DSRR technique for a much clearer and unambiguous characterization.

C. 2DSRR spectra

The first two pre-resonant pulses (see Fig. 5) pump the system to vibrational excited states coupled to the Trp-L_b excitation, they dominate the 2DSRR spectra even when the probe pulse is resonant with the Tyr-L_b excitation. Since the Trp-L_a and Trp-L_b transitions are perpendicular, the XXY pulse polarization is expected to selectively enhance vibrational modes strongly coupled to both Trp-L_b and Trp-L_a. This can be seen by comparing the 2DSRR spectra with XXY polarization (Fig. 5) and XXX/SMA polarizations [Fig. S1 and S2, ESI†]. The 2DSRR spectra were scaled by the inverse hyperbolic sine function to enhance the visibility of weak peaks,

$$\bar{S} = \mathrm{arcsinh}(CS) = \ln\left(CS + \sqrt{1 + C^2S^2}\right) \qquad (3)$$

where $C = 200 \times S/\max(|S|)$ is a constant chosen to make CS in the range −200 to 200. For $CS < 1$, the scale is linear, $\mathrm{arcsinh}(CS) \approx CS$; for larger CS the scaling becomes logarithmic, where $\mathrm{arcsinh}(CS) \approx (S|S|^{-1})\ln(2|CS|)$, so this scaling is useful for enhancing weak features in the signal.[79]

In Fig. 5 we present the 2DSRR XXY spectra with probe pulse $\omega_3 = \varepsilon_1 = 39\,000$ cm^{-1} (second column) and $\omega_3 = \varepsilon_2 = 40\,000$ cm^{-1} (third column). Characteristic features of the five folding states are given in Table 2. In S1, the spectra have strong diagonal peaks at 1260, 1480 and 1605 cm^{-1}, corresponding to the tryptophan ring C–H/N–H in plane bending (WHB), the W8a,[77,78] and the W3 [77,78] modes, respectively. There are also strong correlations between these modes which appear as cross peaks at (1605, 1260), (1605, 1480) and the transposing positions. The (1480, 1260) cross peaks are weaker and are overwhelmed by the (1605, 1260) peaks. S25 has a strong diagonal peak at 1505 cm^{-1}, accompanied by a weaker diagonal peak at 1565 cm^{-1}, corresponding to the W8a and W3/WPS (tryptophan pyrrole ring stretching) modes, respectively. These two modes are correlated taking the (1565, 1505) and (1505, 1565) cross peaks as evidence. In S50, the strong diagonal peak at 1545 cm^{-1} and the weak diagonal peaks at 1585 and 1480 cm^{-1}, corresponding to the WPS, W3 and W8a modes, respectively, are weakly correlated. Furthermore, the WHB mode at 1190 cm^{-1} correlates with the W3 mode, resulting in cross peaks at (1545, 1190) and (1190, 1545). S75 shows diagonal peaks at 1580 and 1615 cm^{-1}, arising from WPS modes from different snapshots. In folding state S75, the indole N–H atom is hydrogen bonded to the Pro17 peptide oxygen atom, which induces blueshift of the WPS mode compared to 1562 cm^{-1} in free tryptophan. However, the local structure around the Trp6 and Pro17 residues does fluctuate. For some of the snapshots, the hydrogen bond lies in the indole plane results in a blueshift of 60 cm^{-1} displayed as the 1620 cm^{-1} diagonal peak; for the majority of the snapshots, the hydrogen bond tilts out of the indole plane, which implies a weaker interaction that corresponds to a smaller (20 cm^{-1}) blueshift displayed as the 1580 cm^{-1} diagonal peak. S100 has a strong diagonal peak at 1380 cm^{-1}, corresponding to the W7 mode.[77,78] It correlates with the W3 mode at 1545 cm^{-1}. The cross peaks at (1545, 1380), (1270, 1545) and (1545, 1616) indicate the correlation between these modes. These characteristics are easily distinguishable and are shared by different pulse configurations.

Cross peaks in the 2DSRR spectra can help determine whether the frequency splitting comes from structural fluctuation. For instance, in the 2DSRR spectra of S1 state at $\omega_3 = 39\,000$ cm^{-1} (see Fig. 5, 2nd column), there are two strong diagonal peaks at 1605 and 1570 cm^{-1} assigned to the W3 and WPS modes.

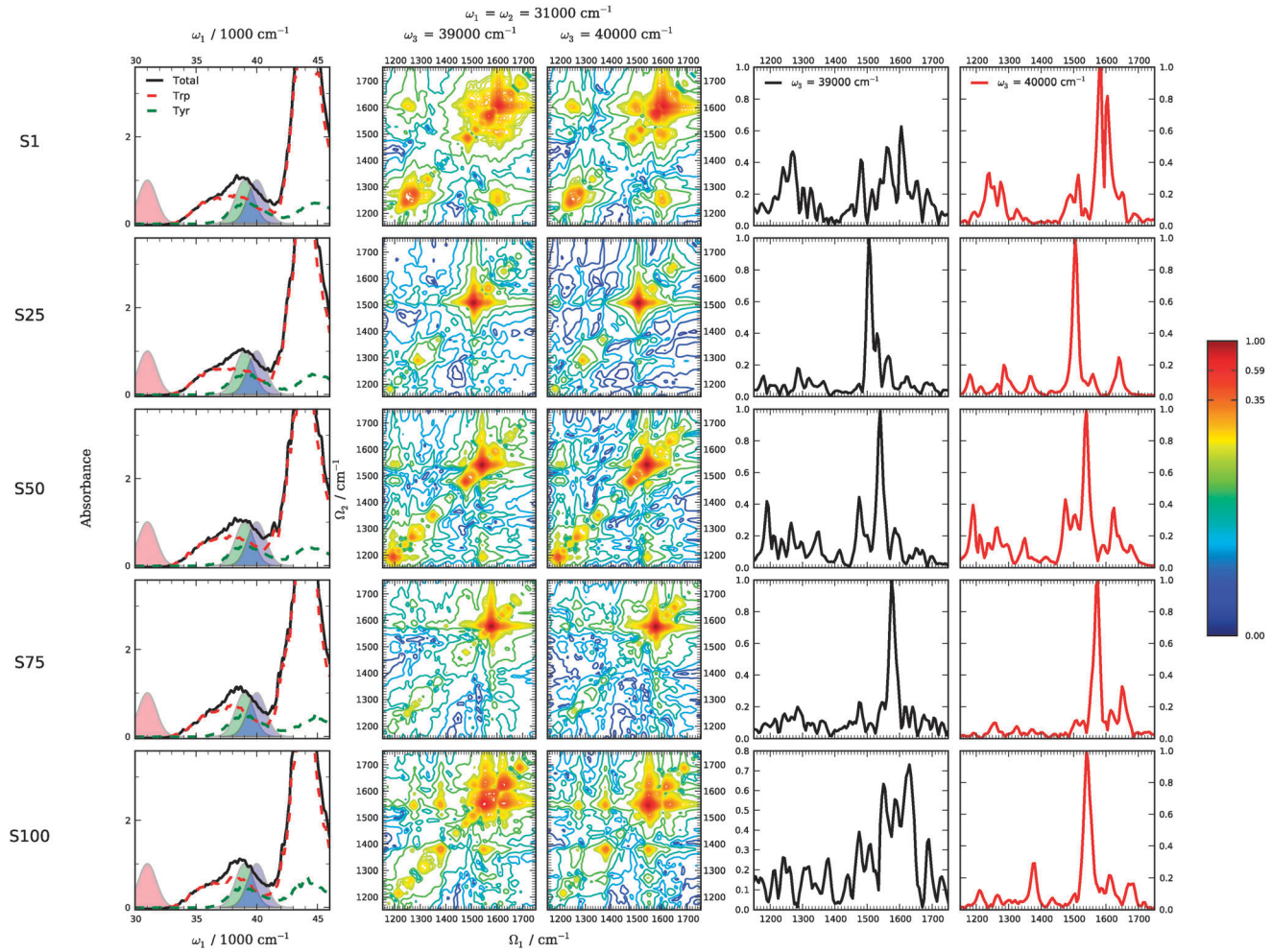

Fig. 5 Left column: near UV absorption of Trp-cage for the various folding states, with contributions from tyrosine (green dashed) and tryptophan (red dashed). Shaded areas represent the pulses with center frequencies 31 000 (red), 39 000 (green), and 40 000 cm^{-1} (blue) used in the 2DSRR simulations. Second column: XXY 2DSRR spectra with $\omega_1 = \omega_2 = 31\,000\ cm^{-1}$, $\omega_3 = 39\,000\ cm^{-1}$. Third column: same except that $\omega_3 = 40\,000\ cm^{-1}$; fourth column: diagonal slices of column 2. Fifth column: diagonal slices of column 3 (red).

Table 2 Characteristic features of 2DSRR spectra

Folding states	Diagonal peaks/cm^{-1}	Cross peaks/cm^{-1}
S1	1605, 1500, 1260	(1605, 1260), (1605, 1500)
S25	1565, 1505	(1565, 1505)
S50	1585, 1545, 1480, 1190	(1545, 1190), (1545, 1480), (1585, 1545)
S75	1615, 1580	(1615, 1580)
S100	1610, 1545, 1380	(1545, 1380), (1610, 1545), (1610, 1380)[a]

[a] The (1610, 1380) cross peaks are weak in the spectra and overwhelmed by the nearby peaks.

It has been proven that the W3 and WPS modes are strongly correlated in both tryptophan monomer and tryptophan–tyrosine dipeptide.[42] The absence of cross peaks corresponding to these diagonal peaks implies that the correlation between these signals is weak. This is the consequence of the fact that the 1605 and 1570 cm^{-1} diagonal peaks come from different snapshots. The splitting of vibrational frequencies can be attributed to structural fluctuations of the folding state S1. The 1605 cm^{-1} peak comes from the W3/WPS modes of snapshots with the Trp6 aromatic group pointing to the N-terminus of the protein (residue 1), and the distances between the sidechains of Trp6 and Ile4 are short. The 1570 cm^{-1} peaks come from the same modes of snapshots with the Trp6 aromatic group pointing to the C-terminus (residue 20), and there are no other residues in its neighborhood with a radius of 4.1 Å. The strong structural fluctuations of S1 is a result of its flexible random coil structure near the N-terminus,[21] where the indole group can freely rotate with respect to the C_β–C_α single bond in Trp6. Similar patterns also appear in the 2DSRR signals of S25 (Fig. 5, column 3), there are strong diagonal peaks at 1640, 1560, and 1505 cm^{-1}, but the (1640, 1560) and (1640, 1505) cross peaks are missing. It is also the consequence of structural fluctuations around Trp6 in the early stage of folding.

The packing density of the Trp6 residue depicted in Fig. 6 is defined as the average number of C_α atoms in a neighborhood of the tryptophan C_α atom. A higher packing density implies

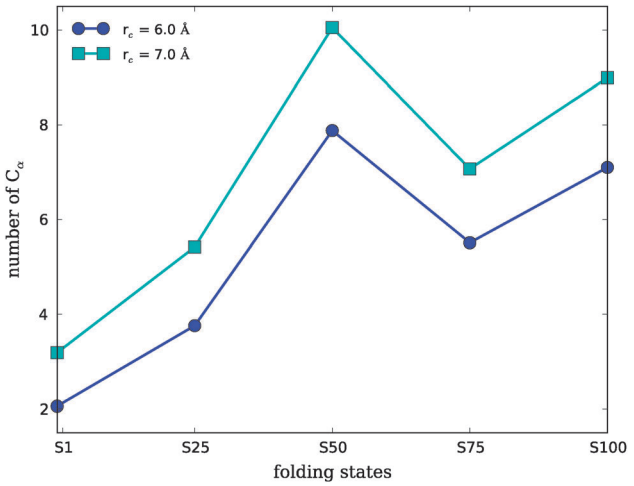

Fig. 6 Packing density of the Trp6 residue for various folding states. This is defined by as the average number of C_α atoms located within a radius r_c from the tryptophan C_α.

stronger inter-residue interaction experienced by Trp6, which induces stronger fluctuations in vibrational frequencies. In S50, an α-helix starts to form near the N-terminus, and the two-strand structure is not complete. This implies that the microenvironment around Trp6 is not rigid. The high packing density indicates strong perturbation to the tryptophan vibration modes, resulting in a broad distribution of diagonal peaks in the 2DSRR spectra. S75 has a much lower Trp6 packing density, which implies weaker fluctuations, and the 2DSRR spectra is much cleaner. In S100, the α-helix and two-strand structure are formed, resulting in a rigid environment around Trp6 that encapsulates it. Most of the diagonal peaks show strong correlations as cross peaks, indicating weak vibrational fluctuations in the rigid environment.

V. Conclusions

By employing a QM/MD protocol to simulate the 2DSRR spectra of the folding of the model mini-protein Trp-cage, we demonstrated that these signals are sensitive to the protein secondary structure and could provide a useful probe for evolving folding states. Frequency shifts of tryptophan modes and their correlations depend on the local chemical environments of the tryptophan residue. Frequency shifts of diagonal peaks come from structural fluctuations when cross peaks are missing; off-diagonal cross peaks reflect strong correlations between the corresponding vibrational modes.

Acknowledgements

We gratefully acknowledge the support of the National Institutes of Health (Grant No. R01 GM-59230), the National Science Foundation (Grant No. CHE-1058791), and the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, US Department of Energy (Grant No. DE-FG02-04ER15571). We acknowledge the computational resource support from the GreenPlanet cluster at UCI (NSF Grant CHE-0840513). Z.Z.L. and J.W. thank the National Science Foundation for support.

References

1 F. Chiti and C. M. Dobson, *Annu. Rev. Biochem.*, 2006, **75**, 333–366.
2 S. M. Johnson, R. L. Wiseman, Y. Sekijima, N. S. Green, S. L. Adamski-Werner and J. W. Kelly, *Acc. Chem. Res.*, 2005, **38**, 911–921.
3 S. Kmiecik and A. Kolinski, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 12330–12335.
4 J. Jiang and S. Mukamel, *Angew. Chem., Int. Ed.*, 2010, **49**, 9666–9669.
5 Y. Levy, P. G. Wolynes and J. N. Onuchic, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 511–516.
6 B. A. Shoemaker, J. J. Portman and P. G. Wolynes, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 8868–8873.
7 Z.-Z. Lai, Q. Lu and J. Wang, *J. Phys. Chem. B*, 2011, **115**, 4147–4159.
8 G. A. Papoian and P. G. Wolynes, *Biopolymers*, 2003, **68**, 333–349.
9 M. Karplus, *Nat. Chem. Biol.*, 2011, **7**, 401–404.
10 J. Wang and G. M. Verkhivker, *Phys. Rev. Lett.*, 2003, **90**, 188101–188104.
11 Z. Lai, J. Su, W. Chen and C. Wang, *Int. J. Mol. Sci.*, 2009, **10**, 1808–1823.
12 W. Xu, Z. Lai, R. J. Oliveira, V. B. P. Leite and J. Wang, *J. Phys. Chem. B*, 2012, **116**, 5152–5159.
13 A. Haabis, W. Żmudziska, A. Liwo and S. Odziej, *J. Phys. Chem. B*, 2012, **116**, 6898–6907.
14 R. Zhou, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 13280–13285.
15 D. Abramavicius, B. Palmieri, D. V. Voronine, F. Sanda and S. Mukamel, *Chem. Rev.*, 2009, **109**, 2350–2408.
16 S. Mukamel, D. Abramavicius, L. Yang, W. Zhuang, I. V. Schweigert and D. V. Voronine, *Acc. Chem. Res.*, 2009, **42**, 553–562.
17 H. S. Chung, M. Khalil, A. W. Smith, Z. Ganim and A. Tokmakoff, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 612–617.
18 H. S. Chung, Z. Ganim, K. C. Jones and A. Tokmakoff, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 14237–14242.
19 D. Abramavicius, W. Zhuang and S. Mukamel, *J. Phys. Chem. B*, 2004, **108**, 18034–18045.
20 W. Zhuang, D. Abramavicius, T. Hayashi and S. Mukamel, *J. Phys. Chem. B*, 2006, **110**, 3362–3374.
21 Z.-Z. Lai, N. Preketes, S. Mukamel and J. Wang, *J. Phys. Chem. B*, 2013, **117**, 4661–4669.
22 J. Bredenbeck, J. Helbing and P. Hamm, *J. Am. Chem. Soc.*, 2004, **126**, 990–991.
23 P. Hamm, J. Helbing and J. Bredenbeck, *Annu. Rev. Phys. Chem.*, 2008, **59**, 291–317.
24 J. Wang, J. Chen and R. M. Hochstrasser, *J. Phys. Chem. B*, 2006, **110**, 7545–7555.
25 Y. S. Kim, L. Liu, P. H. Axelsen and R. M. Hochstrasser, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 7720–7725.

26 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.*, 2009, **106**, 6614–6619.
27 A. Remorino, I. V. Korendovych, Y. Wu, W. F. DeGrado and R. M. Hochstrasser, *Science*, 2011, **332**, 1206–1209.
28 B. A. West and A. M. Moran, *J. Phys. Chem. Lett.*, 2012, **3**, 2575–2581.
29 C.-h. Tseng, P. Sndor, M. Kotur, T. C. Weinacht and S. Matsika, *J. Phys. Chem. A*, 2012, **116**, 2654–2661.
30 J. Jiang and S. Mukamel, *Phys. Chem. Chem. Phys.*, 2011, **13**, 2394–2400.
31 J. Jiang, D. Abramavicius, B. M. Bulheller, J. D. Hirst and S. Mukamel, *J. Phys. Chem. B*, 2010, **114**, 8270–8277.
32 D. E. Schlamadinger, J. E. Gable and J. E. Kim, *J. Phys. Chem. B*, 2009, **113**, 14769–14778.
33 C. Huang, G. Balakrishnan and T. G. Spiro, *J. Raman Spectrosc.*, 2006, **37**, 277–282.
34 A. Soldatova, M. Ibrahim, J. Olson, R. Czernuszewicz and T. Spiro, *J. Am. Chem. Soc.*, 2010, **132**, 4614–4625.
35 A. Ozdemir, I. K. Lednev and S. A. Asher, *Spectrochim. Acta, Part A*, 2005, **61**, 19–26.
36 A. Mikhonin, Z. Ahmed, A. Ianoul and S. Asher, *J. Phys. Chem. B*, 2004, **108**, 19020–19028.
37 L. A. Popova, R. Kodali, R. Wetzel and I. K. Lednev, *J. Am. Chem. Soc.*, 2010, **132**, 6324–6328.
38 A. V. Mikhonin, S. V. Bykov, N. S. Myshakina and S. A. Asher, *J. Phys. Chem. B*, 2006, **110**, 1928–1943.
39 A. V. Mikhonin and S. A. Asher, *J. Phys. Chem. B*, 2005, 3047–3052.
40 S. A. Asher, A. Ianoul, G. Mix, M. N. Boyden, A. Karnoup, M. Diem and R. Schweitzer-Stenner, *J. Am. Chem. Soc.*, 2001, **123**, 11775–11781.
41 H. Ren, J. Jiang and S. Mukamel, *J. Phys. Chem. B*, 2011, **115**, 13955–13962.
42 H. Ren, J. Biggs and S. Mukamel, *J. Raman Spectrosc.*, 2013, **44**, 544–559.
43 M. Kubo, F. Gruia, A. Benabbas, A. Barabanschikov, W. Montfort, E. Maes and P. Champion, *J. Am. Chem. Soc.*, 2008, **130**, 9800–9811.
44 C. Pohling, T. Buckup and M. Motzkus, *J. Biomed. Opt.*, 2011, **16**, 021105.
45 P. Kukura, D. W. McCamant and R. A. Mathies, *Annu. Rev. Phys. Chem.*, 2007, **58**, 461–488.
46 H. Kuramochi, S. Takeuchi and T. Tahara, *J. Phys. Chem. Lett.*, 2012, **3**, 2025–2029.
47 S. Shim, J. Dasgupta and R. Mathies, *J. Am. Chem. Soc.*, 2009, **131**, 7592–7597.
48 B. Mallick, A. Lakhsmanna and S. Umapathy, *J. Raman Spectrosc.*, 2011, **42**, 1883–1890.
49 E. Ploetz, B. Marx and P. Gilch, *J. Raman Spectrosc.*, 2010, **41**, 609–613.
50 Y. Tanimura and S. Mukamel, *J. Chem. Phys.*, 1993, **99**, 9496–9511.
51 D. Healion, J. D. Biggs and S. Mukamel, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2012, **86**, 033429.
52 D. J. Ulness, J. C. Kirkwood and A. C. Albrecht, *J. Chem. Phys.*, 1998, **108**, 3897–3902.
53 D. A. Blank, L. J. Kaufman and G. R. Fleming, *J. Chem. Phys.*, 1999, **111**, 3105–3114.
54 M. Cho, D. A. Blank, J. Sung, K. Park, S. Hahn and G. R. Fleming, *J. Chem. Phys.*, 2000, **112**, 2082–2094.
55 L. J. Kaufman, D. A. Blank and G. R. Fleming, *J. Chem. Phys.*, 2001, **114**, 2312–2331.
56 K. J. Kubarych, C. J. Milne and R. J. D. Miller, *Int. Rev. Phys. Chem.*, 2003, **22**, 497–532.
57 K. C. Wilson, B. Lyons, R. Mehlenbacher, R. Sabatini and D. W. McCamant, *J. Chem. Phys.*, 2009, **131**, 214502.
58 J. W. Neidigh, R. M. Fesinmeyer and N. H. Andersen, *Nat. Struct. Mol. Biol.*, 2002, **9**, 425–430.
59 S. H. Gellman and D. N. Woolfson, *Nat. Struct. Mol. Biol.*, 2002, **9**, 408–410.
60 Z. Ahmed, I. A. Beta, A. V. Mikhonin and S. A. Asher, *J. Am. Chem. Soc.*, 2005, **127**, 10943–10950.
61 J. D. Biggs, J. A. Voll and S. Mukamel, *Philos. Trans. R. Soc., A*, 2012, **370**, 3709–3727.
62 S. Mukamel, A. Piryatinski and V. Chernyak, *Acc. Chem. Res.*, 1999, **32**, 145–154.
63 D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J. Comput. Chem.*, 2005, **26**, 1668–1688.
64 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, **65**, 717–725.
65 A. Onufriev, D. Bashford and D. A. Case, *Proteins: Struct., Funct., Bioinf.*, 2004, **55**, 383–394.
66 J.-P. Ryckaert, G. Ciccotti and H. J. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
67 A. Leach, *Molecular Modelling: Principles and Applications*, Prentice-Hall, Harlow, England, 2nd edn, 2001.
68 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
69 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.
70 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci and G. A. Petersson, *et al.*, *Gaussian 09 Revision C.1*, Gaussian Inc., Wallingford, CT, 2009.
71 V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.
72 M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669–681.
73 A. P. Scott and L. Radom, *J. Phys. Chem.*, 1996, **100**, 16502–16513.
74 J. P. Merrick, D. Moran and L. Radom, *J. Phys. Chem. A*, 2007, **111**, 11683–11700.
75 S. Mukamel, *Principles of Nonlinear Optical Spectroscopy*, Oxford University Press, New York, Oxford, 1st edn, 1995.
76 S. Kannan and M. Zacharias, *Int. J. Mol. Sci.*, 2009, **10**, 1121–1137.
77 T. Miura, H. Takeuchi and I. Harada, *J. Raman Spectrosc.*, 1989, **20**, 667–671.
78 I. Harada, T. Miura and H. Takeuchi, *Spectrochim. Acta, Part A*, 1986, **42**, 307–312.
79 J. Jiang and S. Mukamel, *J. Phys. Chem. B*, 2011, **115**, 6321–6328.

