Tracking Conformational Dynamics of Polypeptides by Nonlinear Electronic Spectroscopy of Aromatic Residues: A First-Principles Simulation Study

Artur Nenov,*[a] Silvio a Beccara,[d] Ivan Rivalta,[c] Giulio Cerullo,[e] Shaul Mukamel,*[b] and Marco Garavelli*[a, c]

Received: May 28, 2014
Published online on August 21, 2014

1. Introduction

Proteins are complex systems, the functions of which include ligand binding, enzymatic reactivity, and allosteric signaling that are regulated by their thermal fluctuations. The characterization of the dynamic properties is, therefore, fundamental for understanding protein function and for numerous applications such as rational drug design and protein engineering. Two-dimensional (2D) optical spectroscopy using ultrashort laser pulses[1–3] can be used to investigate dynamics in complex biological systems with high temporal and spectral resolution. Extending the technique to electronic transitions in the ultraviolet (UV) spectral regime (2D UV) has only been achieved recently.[4–8] In particular, the π–π^* transitions of protein residues with aromatic side chains [phenylalanine (Phe), tyrosine (Tyr), and tryptophan (Trp)] generate signals in the near-UV (NUV) range (250–300 nm),[9] thus providing native local probes for tracking protein dynamics in solution with femtosecond time resolution. 2D UV electronic spectroscopy can target the π–π^* transitions of aromatic residues in oligopeptides and proteins without isotopic labeling. We have recently shown how 2D electronic spectra in the NUV distinguish between two different configurations of a model peptide, that is, Cys-Tyr-Phe-Cys (CYFC, in which Cys = cysteine, Figure 1), with distant and vicinal aromatic side chains.[10] Although our earlier simulations of 2D spectra in solution have clearly indicated that 2D UV spectroscopy is able to distinguish between a folded (cyclic peptide with a disulfide bridge; Figure 1, Structure I) and an unfolded (open peptide with broken disulfide bridge; Figure 1, Structure III) structure of the small oligopeptide, the capacity of this technique to resolve aromatic interactions during the dynamics of a folding/unfolding process remains unexplored. In this study, we investigated the sensitivity of 2D UV spectroscopy to aromatic interactions in proteins by following the unfolding/unstacking of the cyclic CYFC tetrapeptide by using simulations based on the dominant reaction pathway (DRP) approach.[11–13] The DRP allows for an efficient exploration of the reactive path space and offers a variational basis for the determination of the most probable pathway in the unbiased dynamics. Simulations of linear absorption (LA) and time-resolved nonlinear spectroscopy are used to correlate the relative posi-

[a] *Dr. A. Nenov, Prof. M. Garavelli*
Dipartimento di Chimica "G. Ciamician"
Università di Bologna, V. F. Selmi 2, 40126 Bologna (Italy)
E-mail: artur.nenov@unibo.it
marco.garavelli@unibo.it

[b] *Prof. S. Mukamel*
Department of Chemistry
University of California, Irvine, California 92697-2025 (USA)
E-mail: smukamel@uci.edu

[c] *Dr. I. Rivalta, Prof. M. Garavelli*
Laboratoire de Chimie
École Normale Supérieure de Lyon
46 Allée d'Italie, 69364 Lyon Cedex 07 (France)

[d] *Dr. S. a Beccara*
Interdisciplinary Laboratory for Computational Science
FBK-CMM and University of Trento, 38123 Trento (Italy)

[e] *Prof. G. Cerullo*
IFN-CNR, Dipartimento di Fisica
Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milano (Italy)

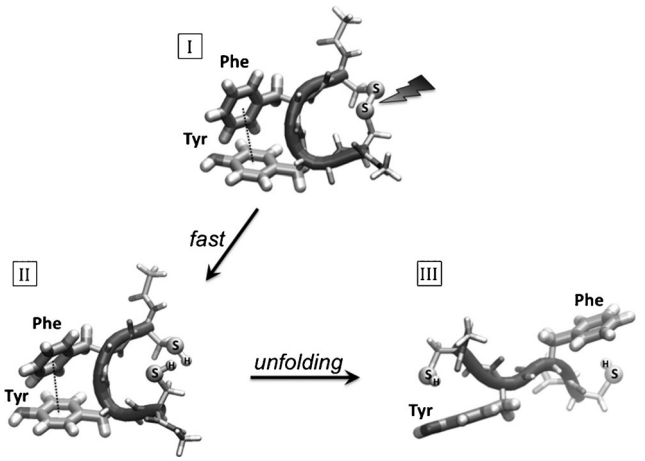

Figure 1. Conformational dynamics of CYFC: after the photoinduced cleavage of the disulfide bond, the sulfur radicals get saturated on an ultrashort timescale. Subsequently, the peptide unfolds. The simultaneous unstacking of the aromatic side chains of Tyr and Phe is probed with different spectroscopic techniques. The color image is in the Supporting Information.

tions of the aromatic side chains of Tyr (Y) and Phe (F) amino acids during the unfolding process with the corresponding electronic spectra. The electronic spectra were obtained by using our sum-over-states and quantum mechanics/molecular mechanics (SOS//QM/MM) method,[14] which combines a hybrid QM/MM scheme based on a wavefunction method that explicitly accounts for environmental effects and interchromophore coupling with nonlinear response theory. This approach allows one to characterize the solvent effect and to directly associate the time-dependent fluctuations of the electronic spectra with the atomistic details of the conformational changes occurring during the unfolding process. The capability of ultrafast electronic spectroscopy to follow folding/unfolding processes of proteins in solution is demonstrated.

Methodology

The time-dependent spectroscopic signals were obtained using our recently developed SOS//QM/MM[14] protocol, which consists of the following steps: 1) configurational space sampling using molecular dynamics simulations, 2) snapshot selection, 3) refinement of the selected geometries at mixed QM/MM level, 4) calculation of excitation energies and transition dipole moments (TDMs), and 5) generation of linear and nonlinear electronic spectra. The application of the individual steps to CYFC is discussed below.

Molecular Dynamics Simulations and Snapshot Selection

Trajectories for the folded to unfolded transition were generated according to the DRP approach.[11–13] The reaction pathway connecting given initial and final molecular configurations X_i and X_f is chosen by maximizing the Onsager–Machlup functional S_{OM}, which can be shown to be proportional to the probability of the pathway [Eq. (1)]:

$$S_{\mathrm{OM}}[X,t] = \int_o^t d\tau \sum_{i=1}^{N_m} \frac{\beta}{4\gamma_i m_i}[m_i\ddot{x}_i + m_i\gamma_i\dot{x}_i + \nabla_i U(X)]^2 \qquad (1)$$

in which $\beta=(k_BT)^{-1}$, $X=(x_i \ldots x_N)$ is the configuration vector of the N-atom molecule, $U(X)$ is the molecular potential energy, and m_i and γ_i are the mass and friction constant of the i-th atom. This functional strictly applies to systems treated in the implicit solvent approximation. In the present work a generalization of the approach described in Ref. [13] to explicit solvent was used. Details are available in the Supporting Information.

The initial X_i and the final X_f geometries were obtained as the global minima of 50 ns classical molecular dynamics simulations of the folded cyclic (i.e. with a –S–S– bond) CYFC with a T-shaped arrangement of the aromatic side chains (Figure 1, structure I) and the unfolded (i.e. with saturated S atoms) CYFC with distant noninteracting aromatic residues (Figure 1, structure III), as described in Ref. [10]. We assume that the unfolding is initiated by the photoinduced cleavage of the –S–S– bond.[15] Multiscale simulations by Nieber et al.[16] indicate that ultrafast intramolecular recombination in the sub-picosecond range[17] may be the dominant path in solution, which would prevent the unfolding. On the other hand, successfully dissociated S–S bonds can participate in intermolecular H transfer with the solvent,[16] which would facilitate the unfolding.[15] In this work we assume that both sulfur radicals are immediately saturated by hydrogen atoms from the water molecules in their close vicinity and the saturation is fast compared with the unfolding/unstacking dynamics. To achieve this we broke the disulfide bond in the folded CYFC geometry and saturated the sulfur atoms with hydrogen atoms. Subsequently, a loose geometry relaxation at the molecular dynamics level was conducted to reduce the repulsion between the sulfhydrylic groups. This geometry was then used as initial molecular configuration X_i. It must be pointed out, though, that intramolecular recombination of the sulfur radicals will aggravate the spectroscopic detection of the unstacking/unfolding. Fingerprint signals of the T-shaped CYFC coming from recombined (i.e. still folded) species will obscure the unfolding.

To find the optimal pathway, an ensemble of trajectories compatible with the boundary conditions is generated by means of ratchet-and-pawl biased molecular dynamics (rMD).[18] By using the rMD algorithm we produced an ensemble of 48 trial unfolding trajectories. The least-action condition was then used to select a single unfolding pathway. Thirty-one snapshots were extracted out of the unfolding pathway covering the dynamics up to an interchromophore separation of 8.5 Å. We note that due to the nature of the DRP the time interval between consecutive snapshots is not known.

QM/MM Refinement

A two-step QM/MM refinement was performed with our software package Cobramm:[19] an initial Hartree–Fock energy minimization run for 30 steps was followed by a complete active space self-consistent field (CASSCF)[20] minimization. An active space of eight electrons and eight orbitals (i.e. CAS(8,8)) was selected for the CASSCF optimization. The software MOLPRO 2010[21] was employed through an interface with Cobramm. The 6-31G* basis set[22] was used for the optimizations. Electrostatic QM/MM interactions were treated through an electrostatic embedding scheme. The link-atom technique and redistribution of residual charges among nearest neighbors were used[23] with both aromatic side chains included in the QM layer and the remaining atoms treated classically. The H-atom link was located along the C_α–C_β bond axis of the aromatic side chains. To preserve the intrinsic conformation of each snapshot, the peptide backbone and the solvent were kept frozen during optimization. Only the water molecules involved in hydrogen bonds with the hydroxyl group of the phenol were allowed to

move. The suitability of the QM/MM refinement for capturing the instantaneous conformation coded in each snapshot is discussed in the Results section.

Excited-State Calculations

Excited-state calculations of the refined snapshots were performed with Molcas 7.7[24] at the state-averaged (SA)-CASSCF level including all valence π electrons and π orbitals of both chromophores in the active space (i.e. CAS(14,13)). The generally contracted ANO-L basis set was utilized[25] and the following contraction scheme was adopted: C,O/[3s2p1d] and H/[2s1p]. Subsequent energy refinement was done perturbationally with the multiconfigurational counterpart of the Møller–Plesset method, denoted as CASPT2,[26] in its single-state (SS) version. An imaginary shift of 0.2 was used.[27] An ionization potential–electron affinity (IPEA) shift[28] of 0.0 was found to give better agreement with experimental data than the default value of 0.25.[10] A set of 13 extravalence orbitals with higher angular momentum (six π^* orbitals in benzene, seven π^* orbitals in phenol) were localized among the virtual orbitals through a procedure described elsewhere[10] and discarded in the perturbation treatment. This approach was found to reproduce experimental data for valence transitions and was also validated against state-of-the-art calculations for benzene and phenol.[10] Seventy states were included in the state-averaging procedure. The number of roots was chosen to include excitations that, upon CASPT2 correction, lie in the energy ranges reported in the electronic spectra. Only states with significant TDM magnitude were selected for the CASPT2 correction, and the threshold was set to 0.03 a.u. This is possible as TDMs are calculated at the CASSCF level. The preselection reduced by half the number of states to which CASPT2 was applied. The MM part of each snapshot was treated as external point charges in both CASSCF and CASPT2 calculations. The Cholesky decomposition approximation was used to speed up the calculation of two-electron integrals.[24]

Electronic Spectroscopy

Using the computed SS-CASPT2 energies and the SA-CASSCF TDMs, LA spectra, pump–probe (PP) spectra, and quasi-absorptive 2D UV spectra were computed for each snapshot by the SOS approach[29] in the dipole approximation with Spectron 2.7.[30] The experimental setups for PP and 2D UV spectroscopy use three pulses with wavevectors k_1, k_2, and k_3 and a local oscillator k_{LO}. Herein, we report quasi-absorptive spectra obtained by summing signals recorded for the rephasing k_I ($k_{LO} = -k_1 + k_2 + k_3$) and non-rephasing k_{II} ($k_{LO} = +k_1 - k_2 + k_3$) conditions.[30–32] This can be achieved experimentally through the partially collinear PP setup with two collinear pump pulses and a noncollinear self-heterodyne probe pulse (i.e. $k_1 = k_2$ and $k_{LO} = k_3$). A constant line broadening of 200 cm^{-1} is used throughout. All calculated signals use the nonchiral *xxxx* pulse polarization configurations and are plotted on a logarithmic scale. Ground-state bleaching (GSB) and stimulated emission (SE) contributions appear as negative (white) peaks, and excited-state absorptions (ESAs) appear as positive (black) peaks in the 2D spectra. A compendium on LA, PP, and 2D UV spectroscopies is provided in the Supporting Information.

2. Results and Discussion

2.1. Analysis of the Unfolding Pathway

Two representative coordinates are chosen to visualize the most characteristic deformations during the unstacking of the Phe-Tyr aggregate: the distance *d* between the centers of the aromatic rings and the angle α between the normals to the planes of the rings (Figure 2). These coordinates have been used in previous studies to characterize aromatic–aromatic amino acid interactions.[33]

Starting from a slightly twisted T-shaped ($\alpha = 65^\circ$) stacked conformation with an interchromophore distance of approximately 5.0 Å, we observe variations in both *d* (5.0–6.0 Å) and α (60–90°; Snapshots 1–10). α finally drops down in the range 40–60°. The twisting is accompanied by a parallel displacement of the two rings, which leads to an increase of *d* (5.5–6.5 Å;

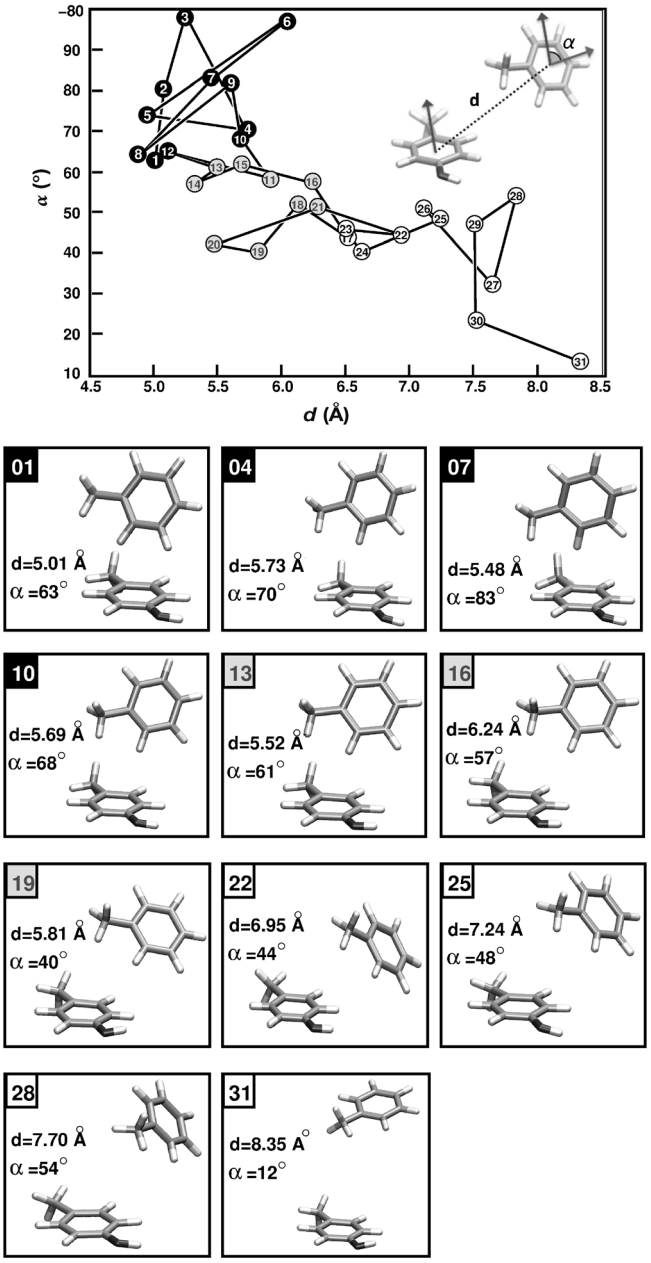

Figure 2. Representation of the conformational dynamics of the CYFC tetrapeptide in the space of the interchromophore distance *d* and the angle between the normals to the aromatic planes α. Color code labeling according to the three-step unfolding/unstacking process (T-shaped, twisted offset stacked, unstacked) is shown in the Supporting Information

Snapshots 11–21). Subsequently, the rings continue to drift away ($d > 6.5$ Å), which allows for α to decrease down to approximately 0° (Snapshots 22–31). Eventually, the rings separate completely and the distance/angle correlation is lost (not shown). Thus, the unstacking dynamics can be summarized as a three-step process: T-shaped→twisted offset stacked→unstacked. An interesting observation is the recurrence of close-contact interactions (Snapshots 5, 8, 14, and 20). It is, therefore, of particular interest to find out if our electronic spectroscopy simulations can reveal these recurrences.

An important issue we need to address before discussing the electronic spectra is the capability of the QM/MM refinement procedure to preserve the intrinsic conformation of the rMD snapshots. Figure 3 compares the rMD (black) and QM/MM (gray) structures with respect to *d*. Although the distance changes by 0.1–0.2 Å upon energy minimization, the global tendency is retained. Only Snapshots 4, 5, and 7 have deviations larger than 0.2 Å, but also in these cases the trend is conserved. It appears that freezing the environment is a strong enough constraint, which allows the retention of characteristic information about the configurational dynamics during the structural refinement.

2.2. Electronic Spectroscopy

Several spectroscopic signals were simulated for the selected snapshots: LA, PP, and quasi-absorptive 2D UV spectroscopy. These techniques characterize the electronic structure of the aromatic residues. The relevant states detectable with the simulated experiments are shown in Figure 4. The spectra are analyzed in this section. Although we know a priori the nature of the different bands observed spectroscopically thanks to the underlying CASSCF/CASPT2 calculations, we begin with a phenomenological analysis of the relative spectral shifts and intensities and only later discuss the origin of the bands in terms of wavefunctions.

The first excited states of Phe and Tyr (i.e. the L_b states), which absorb in the NUV, are selected as targets for the pump pulses. These states have lower oscillator strengths than higher-lying states. However, they are spectrally well separated from the deep-UV (DUV) electronic transitions of the backbone peptide bonds. As demonstrated in our previous study of the CYFC tetrapeptide, setting the incoming pulses in resonance with the L_b transition of Tyr will cover any Phe signal.[10] It is therefore essential to use narrowband pump k_1 and k_2 pulses centered near the frequency of the L_b transition of Phe in order to enhance it. On the other hand, broadband probe k_3 and local oscillator (LO) pulses are required to cover a meaningful fraction of the transient absorption spectrum. Therefore, we used k_1 and k_2 pulses centered on Phe absorption (i.e. 39 000 cm^{-1}) with a full width at half maximum (FWHM) of 733 cm^{-1} (corresponding to a Fourier limited pulse of 20 fs), and a FWHM of 2932 cm^{-1} (corresponding to a 5 fs pulse) for pulse k_3 and the LO, unless specified otherwise. Probing was performed in two spectral windows: NUV (i.e. 35 000–43 000 cm^{-1}) and visible (Vis, that is, 24 000–32 000 cm^{-1}). Probing in the NUV resolves correlated transitions and weak "quartic" couplings in coupled aggregates.[31,32] Probing in the Vis covers the spectral region below the ionization potential and permits the collection of background-free signals of charge-transfer (CT) transitions in coupled chromophore aggregates,[10] which are relatively weak

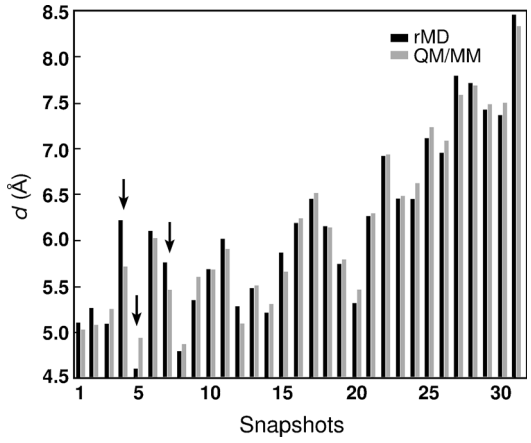

Figure 3. Comparison of the rMD snapshot (black) and the QM/MM refined (gray) interchromophore distance *d*.

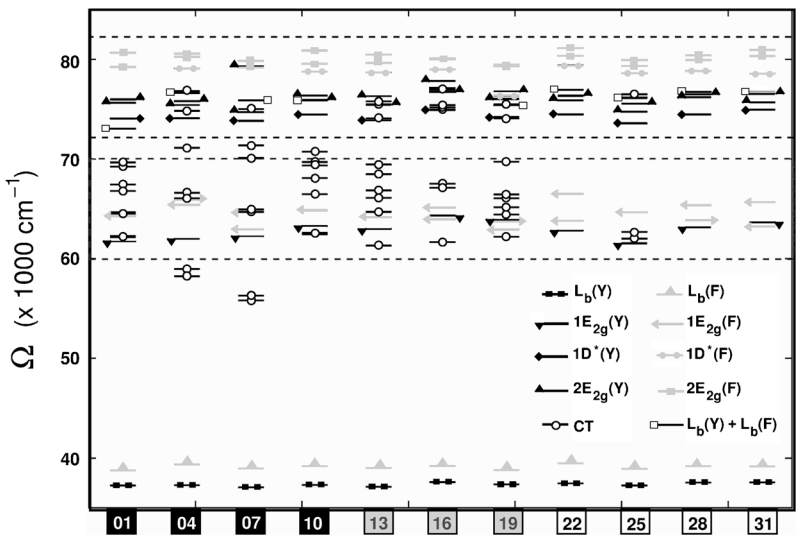

Figure 4. Transition frequencies from the ground state for the singly excited manifold (L_b(Y), L_b(F)) and the bright excited states in the Vis (24 000–32 000 cm^{-1}) and UV (35 000–43 000 cm^{-1}) probing windows for selected snapshots. States with TDM magnitude below the threshold of 0.03 a.u., as well as local states in the range 40 000–60 000 cm^{-1} (e.g. L_a, B_a, B_b), are not shown, as they do not appear in the presented nonlinear spectra. Snapshot labeling according to Figure 2.

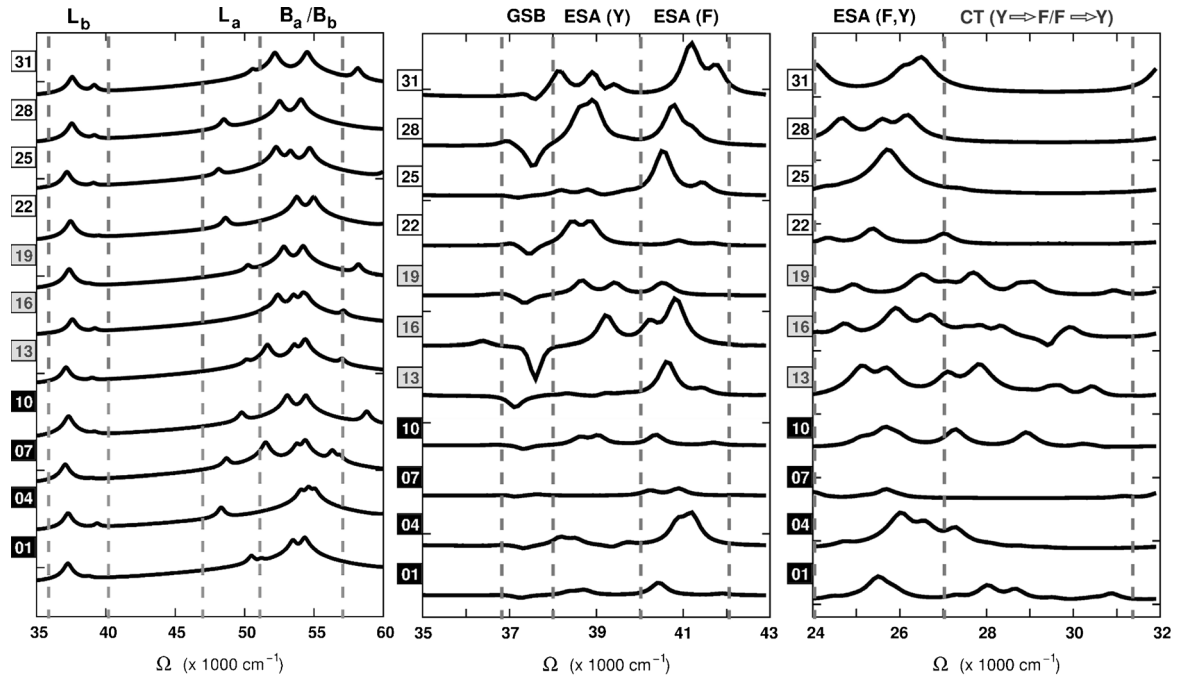

Figure 5. Left panel: Linear absorption spectra for selected snapshots. The L_b, L_a, and B_a/B_b bands are labeled. Middle panel: Time-resolved pump-UV–probe-UV (i.e. one-color PP) spectra for selected snapshots. The GSB and ESA bands are labeled. Right panel: Time-resolved pump-UV–probe-Vis (i.e. two-color PP) spectra for selected snapshots. The ESA and CT bands are labeled. Snapshot labeling according to Figure 2.

and require regions free of the background given by intense local absorptions.

The LA signal reveals the singly excited states, which carry oscillator strength from the ground state (GS). Figure 5, left panel, shows the LA spectra at selected snapshots along the representative unfolding pathway in the NUV to DUV range (i.e. 35 000–60 000 cm^{-1}). Two characteristic bands are seen, a weak band in the NUV between 37 000 and 38 000 cm^{-1} and a strong band between 47 000 and 56 000 cm^{-1}. A closer look at the NUV band reveals that it actually consists of a stronger band around 37 000 cm^{-1} and a weaker band around 39 000 cm^{-1}. These bands correspond to the L_b absorptions of phenol and benzene (Figure 4). The strong feature between 47 000 and 56 000 cm^{-1} includes several intense absorptions of both chromophores. The most redshifted band (47 000–51 000 cm^{-1}) belongs to the GS→L_a absorption in phenol (not shown in Figure 4). The corresponding band for benzene is covered by the more intense GS→B_a and GS→B_b transitions above 51 000 cm^{-1}.[10,34,35]

It is evident from Figure 5, left panel, that the LA is rather insensitive to chromophore stacking. No correlation between the spectra and the conformational dynamics is observed, neither for the L_b bands nor for the higher-lying L_a or B_a/B_b bands. LA is not a good marker for following the conformational dynamics of the aromatic residues.

Figure 5, middle panel, shows the time-dependent PP spectra for selected snapshots. We used a degenerate PP experiment (the marginal of a 2D experiment, see Equation (10) in the Supporting Information) with both pump and probe pulses centered at 39 000 cm^{-1}. This one-color experiment resolves ESA to states at twice the energy of the L_b absorption band (i.e. around 75 000 cm^{-1}, see Figure 4) together with the GSB and SE of opposite sign, which coincide when the delay time t_2 is set to 0. Around 37 000–38 000 cm^{-1} the GSB of phenol is observed. The corresponding GSB signal of benzene, observed as a weak band around 39 000 cm^{-1}, is covered by ESA bands that show fluctuations between 38 000 and 40 000 cm^{-1}. A second band appears between 40 000 and 42 000 cm^{-1}. Each band is a doublet, which shows a conformation-dependent splitting. The signals are ubiquitous in all snapshots, a clear indication of local transitions belonging to either benzene or phenol. A comparison with the individual chromophore spectra allows bands to be assigned to each chromophore. Indeed, the 38 000–40 000 cm^{-1} band is a transition from the L_b state of phenol to a doublet of superexcited states ($2E_{2g}$(Y), Figure 4), whereas the 40 000–42 000 cm^{-1} band is the corresponding transition in benzene ($2E_{2g}$(F), Figure 4).[10] Experimental findings confirm the existence of these superexcited states in pure phenol[36,37] and benzene.[38] We note the weak transition to a doubly excited state 1D*(Y) and 1D*(F) to each band (Figure 4). We do not observe a correlation between the relative shifts and splitting of the local signals and the chromophore stacking.

Figure 5, middle panel, shows that the relative intensity of the bands varies strongly and generally increases for noninteracting chromophores. To understand this we repeated the excited-state calculations in the absence of the environment (i.e. backbone and solvent, hereafter referred to as "gas-phase" calculations). Figure 6 shows a comparison of the TDM magnitudes of the L_b transitions with (black) and without (gray) the environment for the first 25 snapshots of phenol (Figure 6, left) and benzene (Figure 6, right). Indeed, the gas-phase TDM mag-

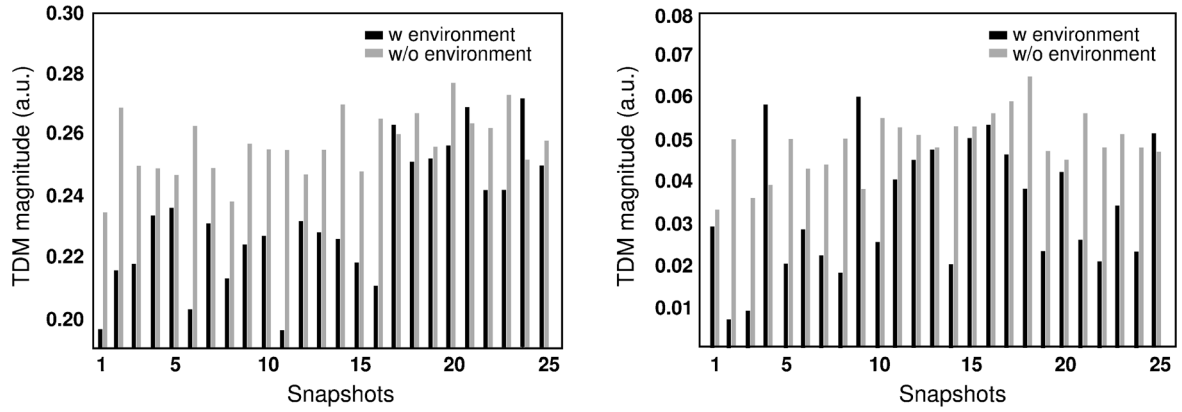

Figure 6. Comparison of the TDM magnitude for the ground state to L_b transition in phenol (left) and benzene (right) in the presence (black) and absence (gray) of the backbone and solvent.

nitude is rather constant, regardless of stacking or unstacking, whereas it decreases for stacked conformations when including the environment effect. The decrease of signal intensity is an environmental effect related to peptide conformation and solvent arrangement. As the GS$\rightarrow L_b$ transition of benzene is weak (symmetry forbidden), the fluctuations of its TDM magnitude result in strong variations of the spectral intensity. The effect should be less pronounced when vibrationally induced increase of the TDM magnitude is taken into account. We note that currently available parameterized Frenkel exciton Hamiltonian models neglect the effect of the environment on the TDM.[9]

Multipulse spectroscopic techniques can utilize pulses with different spectral properties. Probe pulses tunable from the near-IR to the far-UV region can be used to cover a broad spectral range (two-color experiments). Figure 5, right panel, shows the time-resolved two-color PP spectra with NUV pump and Vis probe at selected snapshots. As in the one-color experiment the pump pulse is centered at 39 000 cm^{-1}, whereas the probe pulse is at 28 000 cm^{-1}. All signals are of ESA type since the probe pulse is not in resonance with GS$\rightarrow L_b$ (GSB) and $L_b\rightarrow$GS (SE) transitions. In all snapshots we observe an intense band between 24 000 and 27 000 cm^{-1}. Clearly, this band consists of local absorptions. A comparison with the reference spectra of the monomers,[38] as well as with theoretical calculations,[34,39,40] reveals that one state with E_{2g} symmetry in phenol ($1E_{2g}$, Figure 4) and two states with E_{2g} symmetry in benzene ($1E_{2g}$, Figure 4) absorb in this range ($\approx$7.7–7.8 eV). Valence-bond considerations[41] had already predicted that the $1E_{2g}$ band is the highest band below the ionization limit, which lies at about 8.5 eV for phenol[34] and at about 9.0 eV for benzene.[34] This was later confirmed by multiconfiguration calculations and by experiment. Indeed, no features are observed above 27 000 cm^{-1} for snapshots 22–31, at which the two chromophores are separated by more than 6.5 Å.

Snapshots 11–21, which correspond to the twisted offset stacked conformation of the dimer, exhibit a number of peaks above 27 000 cm^{-1} (Figure 5, right panel; see the Supporting Information for remaining PP spectra). The signals above 27 000 cm^{-1} (i.e. above $\approx$65 000 cm^{-1}/$\approx$8.00 eV from the GS) provide evidence of chromophore–chromophore interactions. T-shaped geometries (Snapshots 1–10) exhibit fewer peaks in the range 27 000–32 000 cm^{-1} (Snapshots 1, 2, 5, and 10 in Figure 4 and in the Supporting Information).

Wavefunction analysis shows that the new signals in the 27 000–32 000 cm^{-1} regime arise from single electron transitions from the L_b bands to two pairs of CT states with configurations H-1(Y)$\rightarrow$L/L+1(F) and H/H-1(F)$\rightarrow$L(Y) (see also the Supporting Information). Each CT state can be reached from the L_b band of both chromophores by either electron transfer (transition within virtual orbitals) or hole transfer (transition within occupied orbitals). In noninteracting aggregates (snapshots 22–31) the CT states are dark and often lie above the ionization potential, which renders them inaccessible. Upon stacking, the overlap of the π orbitals may stabilize the energies by more than 10 000 cm^{-1} and enhances the TDM both from the GS and from the L_b band. The CT states show a complex dependence on interchromophore distance, the orbital overlap, and the environment arrangement.

The two-color PP setup with probing in the Vis region can be used to resolve chromophore–chromophore stacking, albeit not unambiguously, as for particular dimer and environment arrangements the CT signals may lie outside of the 27 000–32 000 cm^{-1} window (e.g. Snapshots 3, 7, and 9; see also the Supporting Information). T-shaped conformations seem to be more sensitive to the surroundings than stacked offset ones. Above 6.5 Å the CT transitions are too weak to be detected.

The strength of 2D spectroscopy lies in the ability to disentangle contributions coming from different chromophores. This is demonstrated in Figure 7, which shows the 2D spectra for snapshots 1, 10, 16, and 19, which collapse to the PP spectra in Figure 5, middle and right panels, when integrated over Ω_1 (see the Supporting Information for remaining 2D UV spectra). We observe two distinct bands corresponding to transitions associated with the GS$\rightarrow L_b$ absorptions in phenol ($\Omega_1 \approx$37 000 cm^{-1}) and benzene ($\Omega_1 \approx$39 000 cm^{-1}). The 2D spectra confirm the conclusion of the one-color PP experiment that the ESA bands at 38 000–40 000 and 40 000–42 000 cm^{-1} are local excitations of phenol and benzene, respectively (Figure 7, left column). As already demonstrated for the PP spectrum, the local signals show no correlation to conformational dynamics.

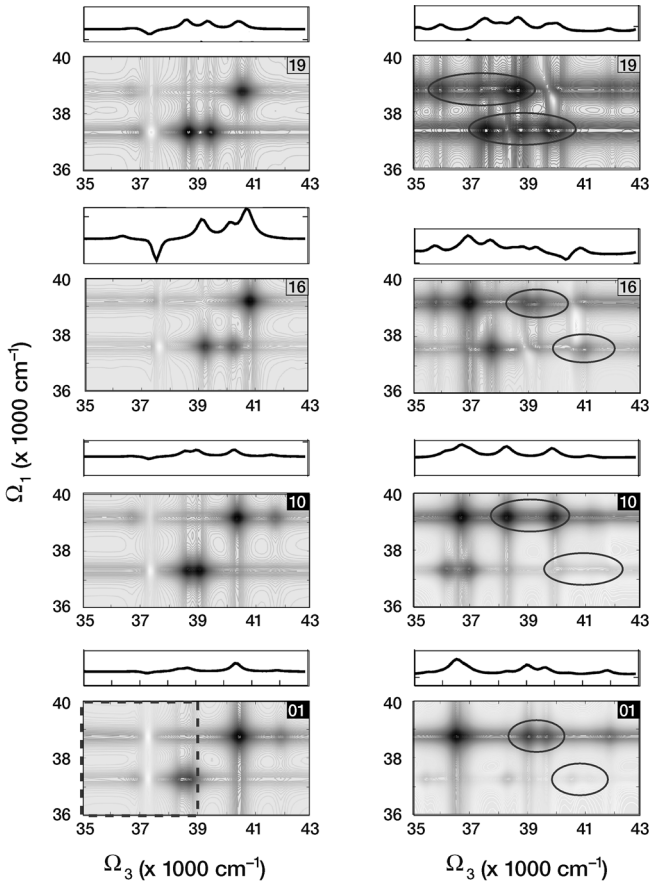

Figure 7. Pump-UV–probe-UV 2D UV (left column) and pump-UV–probe-Vis 2D UV (right column) spectra for selected snapshots together with the marginals (i.e. PP spectra), obtained by integrating over Ω_1. 2D UV signal intensities are normalized [−1(white):1(black)]. GSB/SE and ESA appear as negative and positive contributions, respectively. CT signals are highlighted. Snapshot labeling according to Figure 2. The color image is in the Supporting Information.

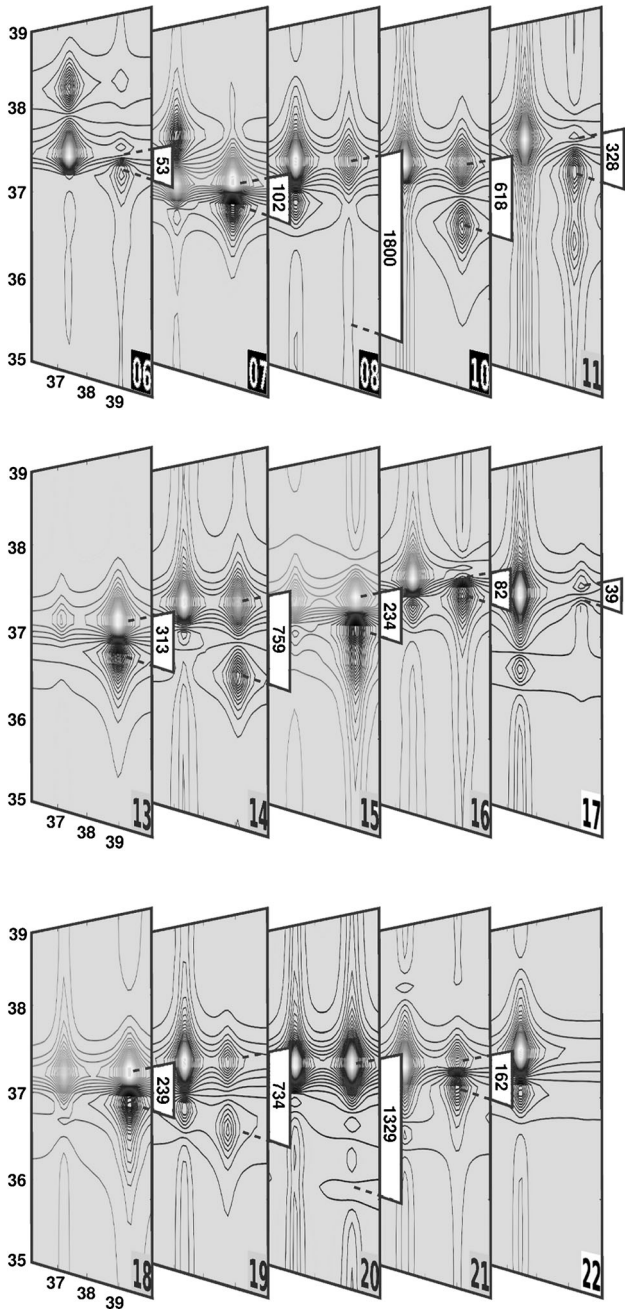

Figure 8. Pump-UV–probe-UV 2D UV spectra for three sequences: 6–7–8–10–11 (upper panel), 13–14–15–16–17 (middle panel), and 18–19–20–21–22 (lower panel). Signal intensities are normalized [−1(white):1(black)]. GSB/SE and ESA appear as negative and positive contributions, respectively. The quartic splittings are highlighted (unit: cm^{-1}). Snapshot labeling according to Figure 2. The color image is in the Supporting Information.

The 2D UV spectra reveal additional signals unresolved in the PP spectra. For most stacked geometries (e.g. snapshots 1 and 10 in Figure 7, left column), a pair of signals emerges along the benzene band: an off-diagonal bleach signal at $\Omega_1 \approx 39\,000\ cm^{-1}/\Omega_3 \approx 37\,000\ cm^{-1}$ and a redshifted ESA, belonging to a transition out of the L_b of benzene to a mixed doubly excited state with a wavefunction $L_b^{benz} + L_b^{phen}$. These two signals are characteristic signatures of the chromophore–chromophore interactions. The splitting between the GSB and the ESA is known as quartic splitting and measures the strength of the interaction.[30]

To better resolve the quartic splitting we focus on the spectral region between 35 000 and 39 000 cm^{-1} (Figure 8) by using the following pulse parameters: a 733 cm^{-1} pump pulse-pair centered at 39 000 cm^{-1} and a 733 cm^{-1} probe pulse centered at 37 000 cm^{-1}. Figure 8 shows three snapshot sequences, which represent recurrences of chromophore–chromophore close contact interactions: 6-7-**8**-10-11, 13-**14**-15-16-17. and 18-19-**20**-21-22. The highlighted snapshots exhibit the shortest interchromophore distance **d** in each sequence. We observe qualitative correlation between the conformational dynamics and the quartic splitting. This is also confirmed by the scatter plot in Figure 9, which correlates the quartic splittings for the T-shaped (snapshots 1–10) and twisted offset stacked (Snapshots 11–21) conformations to the distance **d**. The quartic splittings are generally weak (less than 2000 cm^{-1}): between 6.0 and 6.5 Å they are less than 100 cm^{-1} and vanish completely above 6.5 Å.

As pointed out earlier, the CT states emerging in stacked conformations are common to both chromophores, that is, are

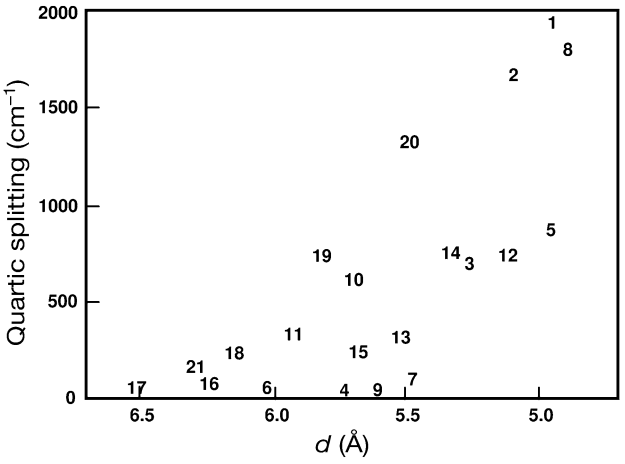

Figure 9. Scatter plot correlating the interchromophore distance *d* and the quartic splitting for Snapshots 1–21.

accessible through a single electron excitation from both L_b states. Thus, transitions to bright CT states result in a characteristic signature in the 2D UV spectra: both L_b bands exhibit identical band patterns with CT peaks in the benzene band redshifted by the energy difference $L_b^{phen} - L_b^{benz}$ (marked with circles in Figure 7, right column). UV-pump–Vis-probe 2D UV spectroscopy with cross-polarized pulses[10] can selectively detect the common states and, thus, unambiguously resolve whether the ESA peaks observed in the corresponding PP spectra (Figure 5, right panel) are associated with transitions to local or CT states.

3. Conclusions

By using the SOS//QM/MM approach for ab initio simulations of electronic spectra, we have demonstrated the ability of different linear and nonlinear spectroscopic techniques to probe peptide conformation dynamics on a model tetrapeptide by using the π–π^* electronic transitions in the aromatic residues Phe and Tyr. The trajectories were obtained in an explicit solvent, by means of the DRP approach. The DRP made it possible to concentrate the computational effort on the reactive part of the trajectories, while reducing the effect of the bias to a minimum. Different stacking arrangements, adopted by the aromatic side chains during peptide unfolding, are found to exhibit characteristic spectral signatures, detectable with tailored experimental setups.

Linear absorption spectroscopy, as well as 1D one-color PP spectroscopy in the NUV show no correlation to the geometrical changes as neither the GS$\rightarrow L_b$ absorptions nor the local ESA signals show correlated shifts. In contrast, two-color PP and 2D UV spectroscopies are very sensitive to stacking interactions, which manifest through off-diagonal bleach signals and bright CT states. The former are correlated to the ESA from the L_b band to the mixed doubly excited $L_b^{benz} + L_b^{phen}$ state present in the aggregate and give rise to the weak quartic splitting (up to 2000 cm^{-1}). The quartic splitting correlates nicely to the interchromophore distance along the unfolding pathway and seems less sensitive to the relative orientation of the chromophores. On the contrary, the energetic position and brightness of CT states depend on the distance, π-orbital overlap, and the environment. This complex dependence may only be resolved by accurate quantum mechanical calculations that take the environment explicitly into account. Intense CT signals can still be observed between 6.0 and 6.5 Å, at which the quartic splitting is very small (< 100 cm^{-1}). Above a separation of 6.5 Å no signature of the chromophore–chromophore interaction is observed.

Nonlinear 2D UV electronic spectroscopy and its marginal PP spectroscopy were proven to be excellent tools for following coherent dynamics. The former provided a combination of high spectral and temporal resolution, whereas the latter represents a compromise between experimental complexity and informative value. The experiments proposed in this study are within reach of current state-of-the-art ultrafast spectroscopy instrumentation.

Acknowledgements

G.C. and M.G. acknowledge support by the European Research Council Advanced Grant STRATUS (ERC-2011-AdG No. 291198). S.M. gratefully acknowledges the support of the National Institute of Health Grant No. GM-59230 and the National Science Foundation through Grant No. CHE-1058791. A.N. wishes to acknowledge the help of the "Hamiltonian of Design" studio (http://hamiltonianofdesign.worldpress.com) for creating the graphics for this publication.

[1] S. Mukamel, *Annu. Rev. Phys. Chem.* **1990**, *41*, 647–681.
[2] S. Mukamel, *Annu. Rev. Phys. Chem.* **2000**, *51*, 691–729.
[3] D. M. Jonas, *Annu. Rev. Phys. Chem.* **2003**, *54*, 425–463.
[4] U. Selig, C.-F. Schleussner, M. Foerster, F. Langhojer, P. Nuernberger, T. Brixner, *Opt. Lett.* **2010**, *35*, 4178–4180.
[5] C.-h. Tseng, S. Matsika, T. C. Weinacht, *Opt. Express* **2009**, *17*, 18788–18793.
[6] C.-h. Tseng, P. Sandor, M. Kotur, T. C. Weinacht, S. Matsika, *J. Phys. Chem. A* **2012**, *116*, 2654–2661.
[7] B. A. West, J. M. Womick, A. M. Moran, *J. Phys. Chem. A* **2011**, *115*, 8630–8637.
[8] B. A. West, A. M. Moran, *J. Phys. Chem. Lett.* **2012**, *3*, 2575–2581.
[9] J. Jiang, S. Mukamel, *Phys. Chem. Chem. Phys.* **2011**, *13*, 2394–2400.
[10] A. Nenov, I. Rivalta, G. Cerullo, S. Mukamel, M. Garavelli, *J. Phys. Chem. Lett.* **2014**, *5*, 767–771.
[11] P. Faccioli, M. Sega, F. Pederiva, H. Orland, *Phys. Rev. Lett.* **2006**, *97*, 108101.
[12] M. Sega, P. Faccioli, F. Pederiva, G. Garberoglio, H. Orland, *Phys. Rev. Lett.* **2007**, *99*, 118102.
[13] S. a Beccara, T. Škrbić, R. Covino, P. Faccioli, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 2330–2335.
[14] I. Rivalta, A. Nenov, G. Cerullo, S. Mukamel, M. Garavelli, *Int. J. Quantum Chem.* **2014**, *114*, 85–93.
[15] C. Kolano, J. Helbing, M. Kozinski, W. Sander, P. Hamm, *Nature* **2006**, *444*, 469–472.
[16] H. Nieber, A. Hellweg, N. L. Doltsinis, *J. Am. Chem. Soc.* **2010**, *132*, 1778–1779.
[17] M. Volk, Y. Kholodenko, H. S. M. Lu, E. A. Gooding, W. F. DeGrado, R. M. Hochstrasser, *J. Phys. Chem. B* **1997**, *101*, 8607–8616.

[18] C. Camilloni, R. Broglia, G. Tiana, *J. Chem. Phys.* **2011**, *134*, 045105.
[19] P. Altoè, M. Stenta, A. Bottoni, M. Garavelli, *Theor. Chem. Acc.* **2007**, *118*, 219 – 240.
[20] B. O. Roos, *Ab Initio Methods in Quantum Chemistry: Part II*, Wiley, Chichester, **1987**.
[21] H. J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schutz, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 242 – 253.
[22] M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, D. J. DeFrees, J. A. Pople, M. S. Gordon, *J. Chem. Phys.* **1982**, *77*, 3654 – 3665.
[23] U. C. Singh, P. A. Kollman, *J. Comput. Chem.* **1986**, *7*, 718 – 730.
[24] F. Aquilante, L. De Vico, N. Ferre, G. Ghigo, P. A. Malmqvist, P. Neogrady, T. B. Pedersen, M. Pitonak, M. Reiher, B. O. Roos, L. Serrano-Andres, M. Urban, V. Veryazov, R. Lindh, *J. Comput. Chem.* **2010**, *31*, 224 – 247.
[25] P. O. Widmark, P. A. Malmqvist, B. O. Roos, *Theor. Chim. Acta* **1990**, *77*, 291 – 306.
[26] K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, K. Wolinski, *J. Phys. Chem.* **1990**, *94*, 5483 – 5488.
[27] G. Ghigo, B. O. Roos, P. A. Malmqvist, *Chem. Phys. Lett.* **2004**, *396*, 142 – 149.
[28] N. Forsberg, P. A. Malmqvist, *Chem. Phys. Lett.* **1997**, *274*, 196 – 204.
[29] G. H. Chen, S. Mukamel, D. Beljonne, J. L. Bredas, *J. Chem. Phys.* **1996**, *104*, 5406 – 5414.
[30] D. Abramavicius, B. Palmieri, D. V. Voronine, F. Sanda, S. Mukamel, *Chem. Rev.* **2009**, *109*, 2350 – 2408.
[31] S. Mukamel, *Principles of Nonlinear Optics and Spectroscopy*, Oxford University Press, Cambridge, **1999**.
[32] P. Hamm, M. Zanni, *Concepts and Methods of 2D Infrared Spectroscopy*, Cambridge University Press, Cambridge, **2011**.
[33] R. Chelli, F. L. Gervasio, P. Procacci, V. Schettino, *J. Am. Chem. Soc.* **2002**, *124*, 6133 – 6143.
[34] J. Lorentzon, P. A. Malmqvist, M. Fulscher, B. O. Roos, *Theor. Chim. Acta* **1995**, *91*, 91 – 108.
[35] D. M. Rogers, J. D. Hirst, *J. Phys. Chem. A* **2003**, *107*, 11191 – 11200.
[36] C. P. Schick, P. M. Weber, *J. Phys. Chem. A* **2001**, *105*, 3725 – 3734.
[37] C. P. Schick, P. M. Weber, *J. Phys. Chem. A* **2001**, *105*, 3735 – 3740.
[38] N. Nakashima, M. Sumitani, I. Ohmine, K. Yoshihara, *J. Chem. Phys.* **1980**, *72*, 2226 – 2230.
[39] T. Hashimoto, H. Nakano, K. Hirao, *J. Chem. Phys.* **1996**, *104*, 6244 – 6258.
[40] T. Hashimoto, H. Nakano, K. Hirao, *J. Mol. Struct. THEOCHEM* **1998**, *451*, 25 – 33.
[41] K. Hirao, H. Nakano, K. Nakayama, M. Dupius, *J. Chem. Phys.* **1996**, *105*, 9227 – 9239.