Novel Two-Dimensional Vibrational Optical Probes for Peptide Fast Folding Investigation

Wei Zhuang[+], Darius Abramavicius[$] and Shaul Mukamel[+*]

Abstract: We demonstrate that early protein folding events may be investigated using a new family of nonlinear infrared techniques which combine the high temporal and spatial resolution of multidimensional spectroscopy with the chirality-specific sensitivity of amide vibrations to structure. We demonstrate how the structural sensitivity of cross-peaks in two-dimensional correlation plots of chiral signals of an α helix and a β hairpin can be used to clearly resolve structural and dynamical details undetectable by circular dichroism, and identify structures indistinguishable by NMR.

Probing the complex energy landscape, pathways and mechanisms of protein folding has been the subject of intense theoretical [1,2,3] and experimental [4] effort. Major advances have been made over the past two decades by fast-triggering the folding process and monitoring the subsequent real-time dynamics by spectral shifts and strengths of optical transitions. Optical techniques allow the investigation of the formation of the basic peptide structural motiffs as well as protein fast-folding in the ps-ns time scale, and address critical questions such as what is the ultimate speed limit of the folding. However, they cannot resolve fine details about correlations between different sections of the peptide (intramolecular) and between the peptide and solvent (intermolecular). NMR techniques have a much higher structural resolution but are limited to longer (μs) timescales. We present simulations of chirality-induced (CI) signals which takes into account structural fluctuations and realistic spectral line broadening by employing a recently developed simulation protocol [5,6]. The cross peaks provide a distinct spectral window for the direct monitoring of dynamical processes such as protein folding.

In 2D optical experiments, three ultrashort laser pulses interact with the protein, and the generated

[+] Department of Chemistry, University of California at Irvine.
[$] Theoretical Physics Department, Faculty of Physics, Vilnius University, Lithuania and Institute of Physics, Lithuania.
[*] Corresponding Author

coherent nonlinear signal, which depends parametrically on three time delays t_1, t_2, t_3, is recorded (figure 1). 2D correlation plots are obtained by a double Fourier transform of the signal with respect to two time delays, holding the third fixed [5]. These femtosecond analogues of multidimensional NMR techniques significantly enhance the spectral resolution by spreading the congested transitions of linear absorption in two dimensions. [5,7,8] The cross peaks (off diagonal peaks in 2D frequency correlation plots) carry signatures of intermode couplings, and their spectral lineshapes probe solvent and conformational fluctuations. Optical 2D techniques have been successfully used to study hydrogen bonding dynamics [9], weak couplings in bistable molecular structures [9], fast chemical exchange in molecular complexes [11,12] and energy transport in photosynthetic antennae [13].

We propose here a new class of techniques which offer a much higher structural sensitivity to peptide secondary structure motifs by making use of peptide chirality. The main idea is as follows: Optical fields vary in both time and space**.** Typical molecules are smaller than the optical wavelength and their response is adequately described by assuming that the field is uniform across the molecule; This is known as the dipole (long wavelength) approximation. Taking the spatial variation of the field into account yields corrections to the response stemming from the variation of the phase of the optical field at different points within the molecule. These corrections, which reflect interferences among contributions from different parts of the molecules, are typically 1000 times weaker than the leading (dipole) contributions. This ratio is determined by the size of the protein divided by the optical wavelength. However, by choosing certain polarization configurations of the laser fields where the dipole term vanishes, the non-dipole corrections generate new background–free signals which can be readily detected. These signals change their signs upon mirror reflection; hence they are finite only in chiral systems and vanish in racemates and nonchiral molecules. The difference in the absorption of left and right-handed circularly polarized light, known as Circular Dichroism (CD), [14,15,16] is an important example of such signals in linear spectroscopy, broadly applied for probing the folding states of proteins and their conformational stability. CD spectra can distinguish between various secondary structures of proteins. We demonstrate that the structural sensitivity can be greatly enhanced by a judicious choice of polarization configurations of 2D techniques aimed at chiral structures. These chirality-induced techniques thus constitute a natural extension of CD to nonlinear spectroscopy.

To appreciate the role of chirality we describe the optical response of proteins using a model of coupled (vibrational or electronic) assembly of chromophores. Each chromophore is localized on a single peptide group. In the infrared these chromophores are the amide I (C=O stretch) vibrational modes whereas in the UV they represent the n—π^* and π —π^* electronic transitions. The system has two types of chirality: the first is associated with the local geometry of individual residues and the second is related to this global (e.g. helical) arrangement. The latter is expected to dominate the response in extended systems[23]. And we will neglect the former in the following arrangement. The absorption spectrum depends on the ensembly-averaged product of two transition dipole vectors $\left\langle \boldsymbol{\mu}_m^{\nu_2} \boldsymbol{\mu}_n^{\nu_1} \right\rangle$ while the CD signal depends on $\left\langle \mathbf{r}_{mn}^{\nu_3} \boldsymbol{\mu}_m^{\nu_2} \boldsymbol{\mu}_n^{\nu_1} \right\rangle$, Here <…> denotes orientational averaging of molecules with respect to the lab frame. $\boldsymbol{\mu}_m^{\nu}$ is the ν cartesian component (ν=x,y,z) of transition dipole of the m'th chromophore, and $\mathbf{r}_{mn}^{\nu}$ is the distance between m'th and n'th chromophores. Similarly, the nonchiral nonlinear response to three pulses depends on the orientationally averaged product of four dipoles $\left\langle \boldsymbol{\mu}_m^{\nu_4} \boldsymbol{\mu}_n^{\nu_3} \boldsymbol{\mu}_k^{\nu_2} \boldsymbol{\mu}_l^{\nu_1} \right\rangle$ whereas its chirality induced component depends on products of the form $\left\langle \mathbf{r}_{mn}^{\nu_5} \boldsymbol{\mu}_m^{\nu_4} \boldsymbol{\mu}_n^{\nu_3} \boldsymbol{\mu}_k^{\nu_2} \boldsymbol{\mu}_l^{\nu_1} \right\rangle$. Non-chiral techniques depend on the structure only implicitly through its effect on the frequencies and transition dipoles which affect peak positions and intensities. The explicit coordinate, ($\mathbf{r}_{mn}^{\nu}$), dependence of the chiral response amplifies the cross peaks and is the reason why these techniques are more sensitive to fine details of the structure.

CD provides a one-dimensional frequency-domain projection of real space chiral configurations; Nonlinear techniques create a multidimensional projection. We shall label a nonlinear experiment, where the 3 incoming pulses are polarized along the ν_1,ν_2,ν_3 directions and the signal is polarized along ν_4 , as $\nu_4\nu_3\nu_2\nu_1$, (Figure 1). Chirality-induced and non-chiral techniques differ by the various choices of ν_j, for example xxxx (all pulses polarized along x) is a non-chiral technique. We have identified 9 independent chirality-induced polarization configurations[24]. The signals further depend on the propagation direction and wavevector of the pulses. The following simulations were made for the infrared chiral reponse of the amide I vibrational (Figure 1) where all beams propagate collinearly along z. For the vibrational spectra we have neglected the local chirality of each peptide unit (which is reasonable for a planar mode) and only included

the global (e.g. helical) chirality. The electronic CD spectra given for comparison were simulated using Woody's standard model [23] which includes the electric and magnetic moments of the chromophores. It thus has local and global chirality.

First we show how CI techniques provide information complementary to Circular Dichroism and NMR for a 15 residue hairpin Trpzip4 (Fig 2A) [17], one of the "Tryptophin Zipper" hairpins. Trpzip4 has a robust, known structure which makes it an excellent model for the characterization of the vibrational states of peptides in aqueous solution, for the investigation of the relations of the vibrational spectra with peptide conformations, and for the evaluation of the distributions of structures[17]. The amide I vibrational band absorption (Fig 2B) has three poorly resolved features; the 1635cm^{-1} peak and the 1675cm^{-1} shoulder are related to the β structure [17], while the 1655cm^{-1} shoulder is related to the turn and coil structures at the two ends. The diagonal peaks of xxxx 2D signals (Fig 2C) resemble the linear absorption. NMR spectra are routinely used to impose constraints on peptide structure; A distance geometry algorithm is then applied to obtain an ensemble of possible conformers consistent with the NMR data [18]. We have selected first two conformers from the reported 20 NMR-determined Trpzip4 structures [19], which have the lowest energy and are thus the best guess of the structure. The RMSD of these two structures is 1.517 A. The calculated electronic (Fig 2D and 2G) and vibrational CD spectra of these conformers (Figure 2E and 2H) are similar while the 2D chirality-induced spectra (Figure 2F and 2I) are different. Conformer I has a strong (1635cm^{-1},1655cm^{-1}) cross peak while conformer II has a cross peaks at (1655cm^{-1},1675cm^{-1}).

We next consider the 30-residue alpha helical peptide SPE4 (Fig 3A), whose sub μs helix-coil transition is difficult to detect using NMR and was investigated by tracking the change of the FTIR around 1630 cm^{-1}[20]. The IR absorption spectrum presented in Fig 3D, has a single peak which only gives a highly averaged signature of the transition dipole arrangement. The diagonal peaks of the nonchiral (Fig. 3G) and the chirality-induced (Fig. 3J) 2D signal show the same optical transitions as in the linear absorption, while the chirality-induced cross peaks are much better resolved.

The sensitivity of CI techniques to structural fluctuations is demonstrated by comparing the CD and the 2D xxxy spectra of two ensembles of a helix with different backbone structure distributions. We chose the configuration at the middle point of the 25 ns trajectory (11000th snapshot) as a reference. The Root Mean

Square Deviations (RMSD) of the central 24 pairs of Ramanchandran angles with respect to the reference structure were calculated (Fig 3B) along the trajectory and their distribution is depicted in Fig 3C. We then selected two 1000-snapshot sub-ensembles with small RMSD (left to the green arrow in Fig 3C) and high RMSD (right to the red arrow). While the simulated electronic CD (Fig 3E and 3F) of amide $n\rightarrow\pi^*$ and $\pi\rightarrow\pi^*$ transitions and the amide I band vibrational CD (Fig 3H and 3I) for these subensembles are very similar, the chirality-induced xxxy signals (Fig 3K and 3L) are markedly different.

We next show that the (-1660, 1630) cm^{-1} cross peak in the CI 2D signal of the helix (Fig 4D) originates from correlations between the helical and the coil part. To that end, we examine the distribution of Ramanchandran angles in Fig. 4A and 4B. Configurations with ϕ within [-55,-90] and ψ within [-35,-75] are considered as the helical domain. Each local vibrational mode i, where any of the 4 angles (ϕ and ψ between i-1, i and i, i+1) lies outside of the helical domain, is considered a coil type, Fig. 4C gives the signal for an isotopically-labeled system where all coil mode frequencies have been shifted by 100 cm^{-1} to the red. Only the pure helix shows in this frequency range and the cross peak of the native system (Fig. 4D) is eliminated. The cross peaks thus provide a unique spectral window for investigating the helix-coil transformation.

Much effort has been devoted to eliminating the strong diagonal peaks in nonchiral 2D techniques of small and rigid molecules in order to better resolve the weaker cross peaks [21]. In the CI 2D signals some diagonal peaks are reduced by interference of contributions from transition dipoles at different parts of the molecule, thereby improving the contrast for the cross peaks. The crosspeaks are induced by correlations between different localized modes and depend on their distance; diagonal peaks have no such distance dependence since they originate from interactions with a single mode. We argue that the diagonal 1660cm^{-1} peak in Fig. 4C is eliminated by the cancellation of signals from different members of the molecular ensemble. To show this, we compare simulations obtained using different numbers of configurations spanning the same, 25ns, time window. Fig. 4D, 4E and 4F present spectra averaged over 22000, 11000 and 5500 configurations respectively. The diagonal 1660cm^{-1} peak is gradually weakened and eventually disappears as the number of configurations is increased.

We have demonstrated that chirality-induced 2D signals can help determine correlations between different parts of a protein by enhancing crosspeak contributions and attribute them to structural features. The

crosspeaks are very sensitive to the secondary structure variation, and the chiral configuration between different chromophores can be determined from the signs of the corresponding crosspeaks (positive vs. negative crosspeak between two transitions correspond to different sense of screw configuration of the corresponding transition dipoles). Moreover, coherent 2D techniques have an intrinsic ultrashort (fs-ns) temporal resolution. Combined with the spatial sensitivity of CI polarization configurations, they offer a new tool, complementary to NMR, for tracking early protein folding events and pinpointing the average structure and its fluctuations along the folding pathways.

The proposed chirality-induced signals are weak. Nevertheless, since they are background-free they may be readily detected using state-of-the-art infrared technology(polarizers and heterodyne detection). The co-linear configuration considered here is the simplest. However non co-linear techniques and combinations of signals in different directions can be used as well to determine the same tensor components. Combinations of several specifically arranged non co-linear experiments may lead to the cancellation of nonchiral terms and only CI terms survive[22]. It is not necessary to combine several experiments for the present xxxy configuration where all laser beams can be arranged in one (yz) plane, the first y-polarized beam propagate along z and the other x-polarized beams can have wavevector along y. All non-chiral contributions will vanish for this configuration.

Method

The simulation protocol for the peptide amide I band lineshape was described in [5]. The helix structure is constructed using MAESTRO package, hairpin structure is obtained from PDB databank (Protein ID 1LE3). NAMD package and CHARMM27 force field are used for the simulation. During the simulation, the backbone dihedral angles of the hairpin are restrained.. The SPECTRON package is used for the modeling of the spectra. An electrostatic model described in reference 5 was used for the construction of the vibrational Hamiltonian. Electronic Hamiltonian in the amide band was constructed using the model described in [23] The NEE module is used for the response function calculation. Double quantum transitions (overtones and combination bands) are introduced through the exciton scattering matrix. A 5.5 cm^{-1} lorentzian is used to simulate the homogeneous broadening. Inhomogeneous broadening is obtained by summing up the calculated

signals of many snapshots of the system. Linear absorption, circular dichroism, and photon echo signals (with t2=0) with polarization configuration xxxx and xxxy were simulated.

Acknowledgement:
This work was supported by the National Institutes of Health Grant 2RO1-GM59230-05 and the National Science Foundation Grant CHE-0446555. We thank Dr. F. Gai at University of Pennylvania for providing the experimental linear absorption signal of spe3 and trz4.

References:

1. Onuchic, J. N., LutheySchulten, Z. & Wolynes, P. G. Annual Review of Physical Chemistry 48, 545-600 (1997)..
2. Elber, R. Curr. Opin. Struct. Biol. 151-156 (2005).
3. Marszalek P.E. et al., Nature 402, 100-103 (1999).
4. Kubelka, J. Hofrichter, J. Eaton, W.A. Current Opinion in Structural Biology 14, 76-88 (2004).
5. Zhuang, W. Abramavicius, D. Hayashi, T. Mukamel, S. J. Phys. Chem. B 108, 18034-18045 (2006).
6. Abramavicius D. and Mukamel, S. J. Chem. Phys. 124, 034113 (2006)
7. Demirdoven, N. et al., Journal of the American Chemical Society 126, 7981-7990 (2004).
8. Asplund, M.C. Zanni, M.T. Hochstrasser, R.M. PNAS 97, 8219-8224 (2000).
9. Cowan, M. L. et al., Nature 434, 199-202 (2005).
10. Larsen, O. F, A.et al., PNAS 102, 13378-13382 (2005).
11. Zheng, J. R. et al., Science 309, 1338-1343 (2005).
12. Y. S. Kim, R. M. Hochstrasser, PNAS 102, 11185-11190 (2005).
13. T. Brixner et al., Nature 434, 625-628 (2005).
14. N. Berova, R. W. Woody, K. Nakanishi, Circular Dichroism. Principles and Applications (Wiley, New York, ed. 2nd, 2000).
15. L. A. Nafie, Annual Review of Physical Chemistry 48, 357-386 (1997).
16. L. D. Barron, Molecular Light Scattering and Optical Activity (Cambridge University Press, 2nd edition, 2004)
17. D. G. Du, Y. J. Zhu, C. Y. Huang, F. Gai, PNAS 101, 15915-15920 (2004).
18. K. Wütrich, NMR of proteins and Nucleic Acids (Wiley, New York, 1995).
19. Cochran, A.G., Skelton, N.J., Starovasnik, M.A. PNAS 98 5578-5583 (2001)
20. T. Wang et al., Journal of Physical Chemistry B 108, 15301-15310 (2004).
21. M. T. Zanni, N.-H. Ge, Y. S. Kim, R. M. Hochstrasser, P. Natl. Acad. Sci. USA 98, 11265-11270 (2001).
22. M. Cho J. Chem. Phys. 119, 7003-7016 (2003).
23. Woody R.W. and Sreerama.N J. Chem. Phys.111, 2844-2845 (1999)
24. Abramavicius D., Zhuang W. and Mukamel S. J. Phys. B. (submitted).

Captions:

Fig. 1: Pulse configuration for femtosecondcoherent infrared spectroscopy. Three laser pulses (light blue) interact with the sample. The fourth pulse (red) is used to detect its nonlinear response. The control parameters are the time intervals between pulses t_1,t_2,t_3. All pulses propagate along z direction (collinear). The nonchiral signal xxxx is generated when all pulses are polarized along x (blue and red). The chirality-induced xxxy signal is obtained by switching the first pulse polarization direction to y (green).

Fig 2: (A): 15-residue beta hairpin peptide Trpzip4. (B): simulated (red) and experimental [17] (green) linear absorption of amide I vibrational band. (C) Simulated xxxx 2D signals for the aide I band. Middle and lower panels compare the simulated spectra for two configurations drawn from the NMR determined hairpin structure ensembles. Electronic CD (D,G) of the amide band, vibrational CD (E,H) of the amide I band and xxxy chirality-induced 2D signals (F,I) for the amide I band. The CD signals are similar for the two configurations. Major differences of the 2D signals in the cross peak region indicate specific couplings among vibrational modes.

Fig. 3: Amide I band vibrational spectra of the 30 residue α helix SPE4 (A) simulated using 22000 snapshots over a 24 ns period. (D): linear absorption (red: simulated, green: experiment [20]) and CD (black) (G): 2D nonchiral xxxx signal. (J): chirality-induced xxxy signal. (B) Simulated RMSD of helix Ramanchandran angles for 22000 snapshots spanning 24ns. (C) the distribution of these RMSD. Left column: simulated spectra for group I (1000 snapshots left to the green arrow). (E) electronic CD of the amide band, (H) vibrational CD of the amide I band, (L) xxxy chirality-induced 2D signals. Right column (F,I,L): same as left column but for the high (right to the red arrow) RMSD group while the ECD and VCD signals are virtually identical. The 2D signals of the two groups are significantly different.

Fig. 4: (A,B): Distribution of the Ramanchandran angles. Regions between the two arrows are identified as “helix”. The (-1660,1630) cm^{-1} cross peaks of the full chirality-induced signal (D) is eliminated when all nonhelical modes are isotopeically labeled by shifting their frequencies $100cm^{-1}$ to red (C). As the number of snapshots in D is decreased to E (11000) and F (5500) we see a strong diagonal peak at (1660,1660) cm^{-1}. The elimination of this diagonal peak in the ensemble provides a clear window for observing the cross peaks.