

Trifluoroacetic acid as excipient destabilizes melittin causing the selective aggregation of melittin within the centrin-melittin-trifluoroacetic acid complex^{a)}

Belinda Pastrana-Rios,^{1,2,b)} Liliana del Valle Sosa,¹ and Jorge Santiago¹

¹Department of Chemistry, University of Puerto Rico, Mayagüez Campus,
Call Box 9019, Mayagüez, Puerto Rico 00681, USA

²Protein Research Center University of Puerto Rico, Mayagüez,
Puerto Rico 00681-9019, USA

(Received 22 March 2015; accepted 28 April 2015; published online 15 May 2015)

Trifluoroacetic acid (TFA) may be the cause of the bottleneck in high resolution structure determination for protein-peptide complexes. Fragment based drug design often involves the use of synthetic peptides which contain TFA (excipient). Our goal was to explore the effects of this excipient on a model complex: centrin-melittin-TFA. We performed Fourier transform infrared, two-dimensional infrared correlation spectroscopies and spectral simulations to analyze the amide I/I' band for the components and the ternary complex. Melittin (MLT) was observed to have increased helicity upon its interaction with centrin, followed by the thermally induced aggregation of MLT within the ternary complex in the TFA presence. © 2015 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution 3.0 Unported License. [\[http://dx.doi.org/10.1063/1.4921219\]](http://dx.doi.org/10.1063/1.4921219)

I. INTRODUCTION

Trifluoroacetic acid (TFA) is commonly found as excipient in synthetic peptides. The presence of this excipient is due to the process of solid phase synthesis, whereby this strong counter ion is used to remove the protecting group and direct the synthesis of the peptide. We have chosen Melittin (MLT) as a model peptide for this study to evaluate the effects of this excipient.¹⁻⁸ Melittin is a component of bee venom (*Apis mellifera*) composed of 26 amino acids (Ac-GIGAILKVLATGLPTLISWIKNKRKQ-NH₂) whose C-terminal end is positively charged. The third component in the complex studied is centrin, a highly conserved calcium binding protein (CaBP), belonging to the tenth largest family of proteins within the mammalian genome, the EF-hand superfamily.^{9,10} *Chlamydomonas reinhardtii* centrin (*Cr cen*) is localized to the nucleus-basal body connectors, the distal striated fibers, and the flagellar apparatus in algae.¹¹⁻¹⁴ The protein is comprised of 172 residues (NCBI Accession No. PO5434) and like all EF-hand proteins, centrin responds to the cellular Ca²⁺ influx by selectively binding Ca²⁺ at four helix-loop-helix calcium binding sites. *Cr cen* is highly stable and has been well characterized in the presence and absence of cations using Fourier transform infrared (FT-IR),^{15,16} circular dichroism (CD),^{15,17} and Nuclear Magnetic Resonance (NMR)^{17,18} spectroscopies. The centrin-MLT complex has been studied by our group and its structure has been determined in the absence of TFA.¹⁹ Protein aggregation at the initial state of the sample prior to performing any crystallization screens may well be the cause of the bottleneck observed in high resolution structure determination of protein-peptide complexes. It is possible that the presence of excipients can induce protein aggregation and thus prevents the successful crystallization of

^{a)}Contributed paper, published as part of a Special Topic: Invited Papers of the 2nd International BioXFEL Conference, Ponce, Puerto Rico, 14–16 January 2015.

^{b)}Author to whom correspondence should be addressed. Electronic mail: belinda.pastrana@upr.edu.

protein-peptide complexes, in general. Herein, we explore the effects of TFA on the *Cr* cen-MLT complex as a model protein-peptide complex using FT IR, two-dimensional infrared (2D IR) correlation spectroscopy, and spectral simulations.

II. MATERIALS AND METHODS

Cr cen was bacterially expressed, isolated, and purified as described in Pastrana-Rios *et al.*^{15,19,20} The use of Spectra9 ¹³C-enriched minimal media from Cambridge Isotopes Laboratory, Inc. (Andover, MA) allowed for the expression of recombinant ¹³C-labeled *Cr* cen using *E. coli* BL21 λ DE3 cells. Centrin's subsequent isolation, purification, and characterization were performed as described in the above mentioned procedure. The MLT synthetic peptide was purchased from Abgent, Inc. (San Diego, CA) and used without further purification, i.e., removal of residual TFA. The concentration of each protein was determined using the predicted molar extinction coefficient ($M^{-1} cm^{-1}$): *Cr* cen $\epsilon_{280} = 1490$ and MLT $\epsilon_{280} = 5500$.

A. FT-IR spectroscopy

The ¹³C-*Cr* cen-MLT-TFA complex (1:1:10 and 1:2:20, molar ratios), ¹³C-*Cr* cen, or MLT-TFA complex (w/w) sample was dialyzed under the desired conditions and then lyophilized repeatedly (5 \times), while re-dissolving the sample in D₂O (99 at. % D), thus allowing for a fully H \rightarrow D exchanged protein sample. Typically, 35 μ l aliquot of 60 mg/ml of the protein sample in 50 mM (4-(2-hydroxyethyl)-1-piperazine ethane sulfonic acid) HEPES buffer, 150 mM NaCl, 4 mM MgCl₂, and 4 mM CaCl₂ at pD 6.6 was deposited on a 49 \times 4 mm custom milled CaF₂ window with a fixed path length of 40 μ m. A reference cell was prepared similarly, and both cells were set in a custom dual chamber cell holder. The temperature within the cell was controlled *via* a Neslab RTE-740 refrigerated bath (Thermo Electron Corp., NH) and monitored with a thermocouple positioned in close contact with the sample cell. The temperature accuracy was estimated to be within 1 $^{\circ}$ C. Ten minutes were routinely allowed for thermal equilibrium to be reached before spectral acquisition was begun. The temperature range studied was 5–98 $^{\circ}$ C. The instrument used was a Magna-IR 550 Spectrophotometer (Thermo Electron Corporation, Madison, WI) equipped with an MCT detector, a sample shuttle, and an interface. Typically, 512 scans were co-added, apodized with a triangular function, and Fourier transformed to provide a resolution of 4 cm^{-1} with the data encoded every 2 cm^{-1} .

B. Spectral analysis

The spectral region of interest is the amide I' band has been simplified by the H \rightarrow D exchange of the sample and as a result only the peptide bond carbonyl stretching vibrations are observed along with the arginine side chain modes.^{15,16,21–26} In the situation when the protein sample has also been ¹³C homogeneously labeled along with the H \rightarrow D exchange, then the ¹³C carbonyl stretching modes (*Cr* cen in this case) are termed amide I'* band.^{21,22} The objective of this method is to allow for the simultaneous study of two protein components. Also, TFA absorbs at a discrete wavelength, as a sharp and narrow band, thus allowing for the simultaneous study of all three components. 2D IR correlation spectroscopy is a technique developed by Noda^{27–29} and used extensively by our group.^{15,16,19–21,30–32} The technique uses the FT-IR series of sequential spectra as a function of a perturbation (temperature, H/D exchange, ligand titration, pH, excipient, etc.) to generate synchronous and asynchronous plots. The spectra are collected at regular intervals during the perturbation. In general, the analysis spreads the acquired spectral data into two dimensions, thus enhancing the spectral resolution in the synchronous plot and providing information on the sequence of molecular events that occur in the asynchronous plot.^{15,19,21,28–30} Both plots provide information on the coupling of vibrational modes within the spectral region of interest (1720–1540 cm^{-1}). Consequently, this technique is sensitive to backbone vibrational modes as well as certain side chain modes (i.e., arginine, aspartates, glutamates, and tyrosine) being perturbed and, as a result, one can identify changes in these modes as a function of the perturbation. Furthermore, we have established a method

that has proven useful in the simulation FT-IR spectra and 2D IR correlation plots.^{31,32} The characteristic sub-bands within the amide I'/I* contour were used to simulate the 2D IR correlation plots.

Baseline correction, spectral overlay, peak pick routines, deconvolution, and 2D IR correlation analysis were performed using a kinetics program for MATLAB (MathWorks, Natick, MA) generously provided by Dr. Erik Goormaghtigh from the Free University of Brussels, Belgium.

III. RESULTS

A. Protein and peptide sample

The mass/charge ratio determined by Matrix Assisted Laser Desorption Ionization Time of Flight (MALDI-TOF) mass spectrometry was consistent with the incorporation of the ¹³C-label. In addition, the loss of the first residue (Met) due to proteolytic cleavage at the amino terminal end was obtained (*data not shown*) and varying Ca²⁺ and Mg²⁺ coordination was also evident. Consequently, the protein sample was dialyzed against the appropriate buffer to ensure similar conditions for the pure component as well as the ternary complex.

B. FT-IR spectroscopy

FT-IR spectroscopy is sensitive to conformational changes observed within proteins. Specifically, the amide I'/I* band (1720–1540 cm⁻¹) due to the stretching vibrations of the MLT peptide and the isotopically labeled protein, ¹³C-*Cr* cen, were analyzed. In general, all spectral data sets are arranged in columns and the similar plot types are arranged in rows to facilitate comparison and analysis.

C. Band assignments

Band assignments were done with baseline-corrected spectra collected at 25 °C, and no aggregation was detected for the MLT-TFA or ¹³C-*Cr* cen samples within the temperature range studied (Figs. 1(a) and 1(b)). For **MLT-TFA complex**, the amide I' band comprised TFA (1673.1 cm⁻¹), kink (1664.4 cm⁻¹), random coil (1656.1 cm⁻¹), ₃₁₀-helix (1643 cm⁻¹), β -sheet contributions (1631.3 cm⁻¹), and arginine stretching modes (1606.0 and 1582.4 cm⁻¹) for the single arginine residue (R₂₄) within the MLT sequence. The β -sheet at 1631.3 cm⁻¹ has also been reported previously by our group¹⁹ and by Levin's group³³ for MLT in aqueous solution. For ¹³C-*Cr* cen, the amide I* band rendered the following assignments: CaBS loops (1640.0 cm⁻¹), the hinge loops (1636 cm⁻¹), π -helix (1624.0 cm⁻¹), α -helical contributions (1597.7 cm⁻¹), and arginine stretching modes (1611 cm⁻¹, 1584.7 cm⁻¹, and 1572 cm⁻¹). For the ¹³C-*Cr* cen-MLT-TFA complex (Fig. 1(c)), the amide I'/I* band is composed of TFA (1673.1 cm⁻¹), MLT's kink (1664.4 cm⁻¹), ₃₁₀-helical component (1643 cm⁻¹), and aggregation (1619 cm⁻¹),^{34,35} while for the ¹³C-*Cr* cen within the complex, α -helical (1597 cm⁻¹) and the side chain vibrational modes include the N-D and ¹³C-N guanidinium stretches in the exchanged arginines (1584.7 cm⁻¹ and 1572 cm⁻¹, respectively).

D. Thermal dependence studies

We compared the FT-IR spectra in the spectral region of 1720–1540 cm⁻¹ (the amide I'/I* band) within the temperature range of 5–98 °C in the presence of Ca²⁺ and Mg²⁺, for ¹³C-*Cr* cen, MLT-TFA, and the ¹³C-*Cr* cen-MLT-TFA ternary complex Figs. 1(a)–1(c), respectively. The asymmetrical appearance of the amide I* band for ¹³C-*Cr* cen (Fig. 1(a)) is due to the underlying contributions of the β -sheet and helical components and some overlap by the guanidinium stretching vibrations, which is mainly due to *Cr* cen arginines. The shoulder at 1625 cm⁻¹ is the result of the ¹³C-labeled carbonyl stretching modes associated with the π - and α -helices as the major contributors of the changes observed. The amide I' band for MLT-TFA complex (Fig. 1(b)) lacks the aggregation peak (1619 cm⁻¹); instead, it contains TFA (1673 cm⁻¹), β -sheet (1631.3 cm⁻¹), and the antiparallel β -sheet component (1685 cm⁻¹) also

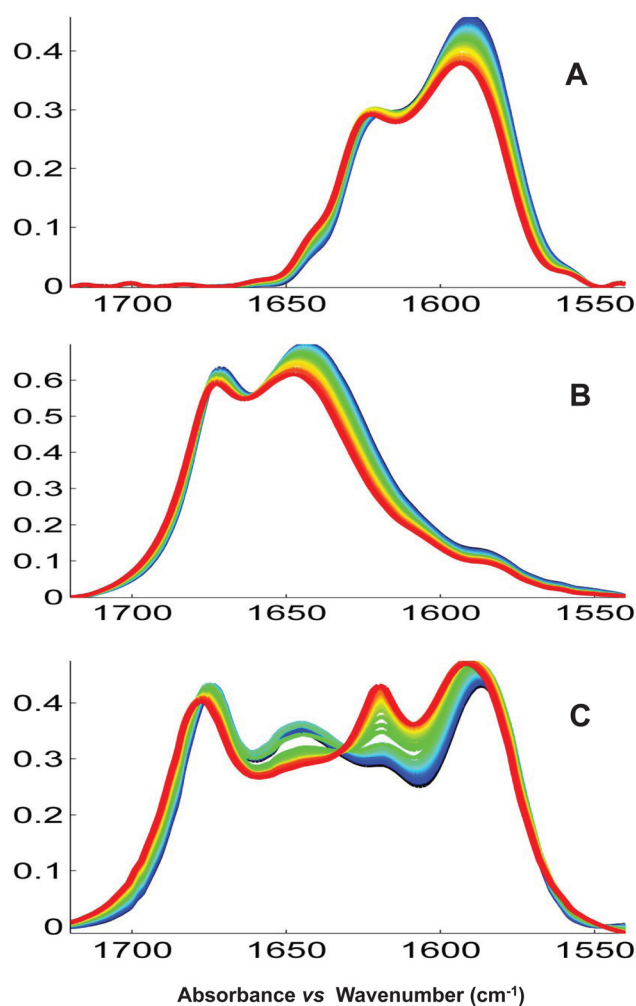


FIG. 1. Overlaid FT-IR spectra of (a) $^{13}\text{C-Crcen}$, (b) The MLT-TFA complex, and (c) $^{13}\text{C-Crcen-MLT-TFA}$ complex. The amide I' and I'* bands for the MLT-TFA complex and labelled centrin are shown along with the changes in the spectral features due to the temperature induced changes in each component as the temperature is increased from 5°C to 98°C within the spectral region of $1740\text{--}1520\text{ cm}^{-1}$.

referred to as β -turn. In the case of the ternary complex (Fig. 1(c)), the aggregation peak at 1619 cm^{-1} is evident along with the decrease in helical composition of the MLT component within the ternary complex.

E. 2D IR correlation analysis

This technique enhances the resolution of the spectral region of interest (Figs. 2(a)–2(c)),^{15,19–21,27–29} which correlates the auto peaks found in the synchronous plots (Figs. 2(d)–2(f)) comprised changes in intensity within peaks that are in phase. While, the asynchronous plots (Figs. 2(g)–2(i)) are comprised of peaks that change in intensity out of phase from each other and establish the order of molecular events involved the aggregation and the role TFA had in this process.³⁶

The synchronous plots for the components ($^{13}\text{C-Cr}$ cen and MLT-TFA) studied within the spectral region $1720\text{--}1540\text{ cm}^{-1}$ are shown in Figs. 2(d) and 2(e), respectively. In each case, a spectral region is observed ($^{13}\text{C-Cr}$ cen $1720\text{--}1650\text{ cm}^{-1}$ and $1600\text{--}1540\text{ cm}^{-1}$ for MLT-TFA) with little or no overlap to allow for the simultaneous study of these protein components. The auto and cross peaks with their band assignments are consistent with the band assignments presented above for the same spectral region. The auto peak with the greatest intensity change for

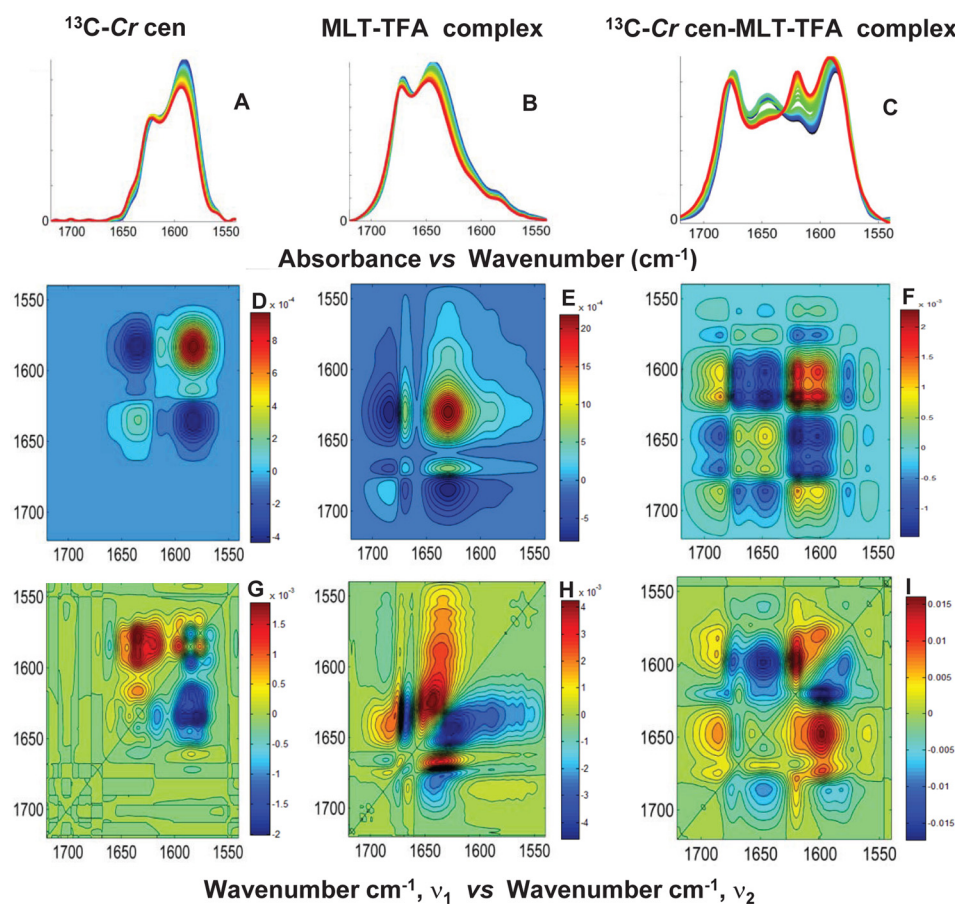


FIG. 2. Overlaid FT-IR spectra ((a)–(c)) and 2D IR correlation spectroscopy: synchronous ((d)–(f)) and asynchronous ((g)–(i)) plots within the spectral region of 1720–1540 cm^{-1} and temperature range from 5 °C to 98 °C for ((a), (d), and (g)) $^{13}\text{C-Cr}$ cen, ((b), (e), and (h)) MLT-TFA complex, and ((c), (f), and (i)) $^{13}\text{C-Cr}$ cen-MLT-TFA complex (1:2:20, molar ratio), respectively. Reprinted with permission from B. Pastrana-Rios and I. Iloro, USPTO Method Patent: Determination of protein, peptide or peptoid aggregation, stability and viability and with the same (2008).³⁶

$^{13}\text{C-Cr}$ cen (Fig. 2(d)) was the β -sheet (1591.8 cm^{-1}) followed by the hinge loops (1636.0 cm^{-1}). For MLT-TFA complex (Fig. 2(e)) was the β -sheet (1637.7 cm^{-1}) followed by the TFA (1673 cm^{-1}).

F. Molecular events

Molecular events for the components (MLT-TFA and $^{13}\text{C-Cr}$ cen) within the spectral region 1720–1540 cm^{-1} were studied using the synchronous plots, shown in Figs. 2(d) and 2(e), and the asynchronous plots, shown in Figs. 2(g) and 2(h), respectively. The asynchronous and synchronous plots are used to establish the sequential order of molecular events leading up to thermal denaturation are shown in Figs. 3(a) and 3(b) and summarized in supplementary material Table 1S.³⁷ More importantly, the information provided can be related to the relative stability of the structural motifs within each protein. For $^{13}\text{C-Cr}$ cen shown in Fig. 3(a), the α -helical contributions (1597.3 cm^{-1}) were the least stable followed by the arginines modes specifically: the guanidinium N-D asymmetric stretch (1611 cm^{-1}), the $^{13}\text{C-N}$ stretch (1572 cm^{-1}), and the N-D asymmetric stretch (1584.7 cm^{-1}), followed by the π -helix (1624.7 cm^{-1}) located at the C-terminal end of the protein. The calcium binding loops (1640 cm^{-1}) along with the hinge loops (1636.0 cm^{-1}) were perturbed at high temperatures, consistent with our previous work.¹⁵ In the case of MLT, within the MLT-TFA complex, the 3_{10} -helix (1643 cm^{-1}) and β -sheet (1631.3 cm^{-1}) are perturbed first followed by the TFA (1671.5 cm^{-1}), and the single arginine

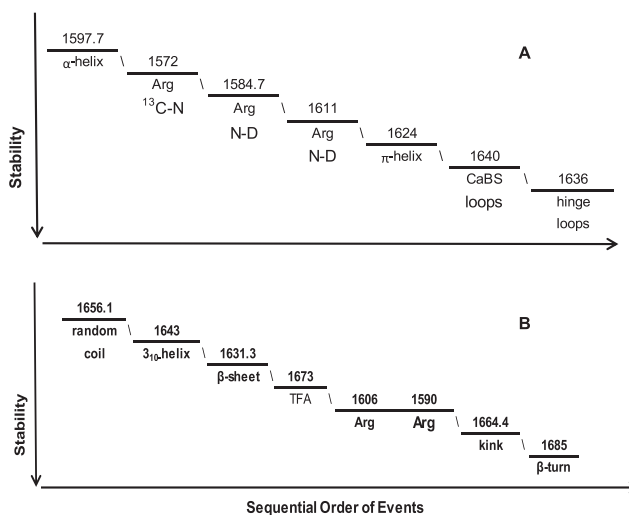


FIG. 3. Schematic representation summarizing the sequential order of events during the thermal perturbation of the components (a) MLT-TFA complex and (b) ^{13}C -Crcen.

(1606 cm^{-1} and 1582.4 cm^{-1}) found at the C-terminal end of the peptide, then the kink (1664.4 cm^{-1}) found near the middle of the peptide, and finally the antiparallel β sheet (1685 cm^{-1}) component due to intermolecular hydrogen bonds with other MLT's in solution is deemed to be the most stable structural motif within the peptide (Fig. 3(b)).

MLT's aggregation within the *Cr* cen-MLT-TFA complex (1:2:20, molar ratio) is shown in the overlaid spectra (Fig. 2(c)) and 2D IR correlation spectroscopy contour plots are shown in Figs. 2(f) and 2(i). The sequential order of events determined by applying Noda's rules^{28,29} for the interpretation of the asynchronous and synchronous plots is shown in Fig. 4 and summarized in supplementary material Table 2S.³⁷ The spectral data for the entire temperature range 5–95 °C were used to understand the interaction within the complex and the thermally induced aggregation process in the presence of TFA. ^{13}C -*Cr* cen and MLT can be observed in the presence of an intense negative cross peak in the asynchronous plot (Fig. 2(i)) at $1643, 1595\text{ cm}^{-1}$ (ν_1, ν_2), suggesting an interaction between MLT's 3_{10} helix and *Cr* cen's helices. Also, observed is the $1643, 1595\text{ cm}^{-1}$ cross peak, suggesting an interaction of MLT's 3_{10} -helix with ^{13}C -*Cr* cen's α -helices. Finally, the existence of cross peak $1673, 1595\text{ cm}^{-1}$ attributed to the interaction between TFA and ^{13}C -*Cr* cen's helical regions suggests an interaction between TFA and ^{13}C -*Cr* cen's helices through the arginine residues located in these regions.

The sequential order of molecular events of ^{13}C -*Cr* cen-MLT-TFA complex (1:2:20, molar ratio) during the thermal perturbation was determined to be the following arginine's ^{13}C -N stretch (1572 cm^{-1}) followed by aggregation (1619 cm^{-1}) and ^{13}C -*Cr* cen helical regions along with arginines' N-D stretching (1584.7 cm^{-1}). The perturbation of centrin by TFA is followed by MLT's and the rc segments adopt a 3_{10} -helical conformation due to its interaction with

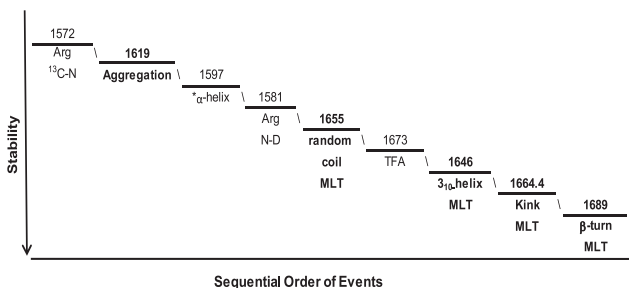


FIG. 4. Schematic representation summarizing the sequential order of events of the thermally induced ^{13}C -Crcen-MLT-TFA complex (1:2:20 molar ratio) and aggregation process.

centrin. TFA is then perturbed followed by the 3_{10} -helix (1646 cm^{-1}), then the kink (1664.4 cm^{-1}) found in the middle of the peptide, and finally the remaining antiparallel β -sheet (1685 cm^{-1}). Furthermore, as MLT progressively loses its interaction with *Cr* cen's helices due to the presence of TFA, then the peptide's 3_{10} -helix begins to aggregate. This event is confirmed by the cross peak splitting at $1673\text{--}1618$, 1584.7 cm^{-1} . This process describes the formation of the complex and TFA induced MLT aggregation during thermal perturbation and the interaction with centrin.

We determined the yield of aggregated species to be 12% MLT. These results have been included in a granted method patent application.³⁶

The ^{13}C -*Cr* cen-MLT-TFA complex (1:1:10, molar ratio) within the spectral region of $1720\text{--}1540\text{ cm}^{-1}$ is shown in supplementary material Fig. 4S.³⁷ These results validate the mechanism proposed for the ternary complex at a higher MLT-TFA molar ratio discussed above. Briefly, the similarities in the peak pattern distribution observed for the ternary complex 1:1:10 with that of 1:2:20 molar ratio, respectively, provide greater confidence for the described mechanism of aggregation. In that, TFA (1673 cm^{-1}), MLT's 3_{10} -helical segment (1643 cm^{-1}), and aggregation (1619 cm^{-1}) peaks substantiate the association with MLT as observed from the negative cross peaks observed in both the synchronous and asynchronous plots. More importantly, the correlation between centrin's arginines (1584 and 1572 cm^{-1}) with TFA as the inducing factor to inhibit the interaction between MLT and centrin which results in MLT's aggregation is also evident.

G. Spectral simulations

Spectral simulations were performed of the amide I* band for the centrin-MLT-TFA complex (1:2:20 molar ratio) shown in Fig. 5, to further understand the aggregation process. Simulation of spectral components was used to simplify the complex amide I'/I* band and generate the 2D IR correlation spectroscopy plots for the temperature range of $5\text{--}95\text{ }^\circ\text{C}$. The spectral components chosen for the simulation were the ^{13}C -*Cr* cen α -helical and β -sheet (1597 and 1588 cm^{-1}) and the TFA (1673 cm^{-1}) as observed within Fig. 5(a). The simulated 2D IR correlation plots (Figs. 5(d) and 5(g)) were able to reproduce splitting pattern of the cross peaks observed between the TFA and centrin main backbone vibrational modes. A second simulation involved MLT's 3_{10} -helix and the aggregation peak (Fig. 5(b)) where the aggregation phenomenon involving MLT's 3_{10} -helix was confirmed by the appearance of the aggregation auto peak (Fig. 5(e)). By considering the sum of the contributing bands considered in the two previous simulations shown in Figs. 5(a) and 5(b), we generated the final set of simulated spectra (Fig. 5(c)), to generate the simulated mechanism of aggregation. Once again the peak pattern (Figs. 5(f) and 5(i)) are consistent with the experimental 2D IR correlation plots observed in Figs. 2(f) and 2(i).

H. Thermal dependence plots

Thermal dependence plots (Fig. 6) were generated using the intensities for the secondary structure contribution of interest within the amide I'/I* band ($1720\text{--}1540\text{ cm}^{-1}$) and normalized using the largest intensity contribution within this same spectral region shown in Fig. 2(c) for the empirical spectral data and for the simulated spectra Fig. 5(c). For each spectral component of interest (i.e., TFA, aggregation, and the MLT α -helical component), the intensity ratio was used to determine the fractional peak intensity. Therefore, a value of 1 is interpreted as the largest contributing peak within the amide I'/I* contour. Determining the fractional intensities establishes the degree of change within the structural motif of the protein in relation to its overall structural components. We determined the fractional loss of MLT's helical component and similarly for MLT aggregation within the ^{13}C -*Cr* cen-MLT complex, shown in Figs. 6(a) and 6(b). In Fig. 6(c), we observed slight intensity change for the TFA component. In addition, the onset of the aggregation process of MLT within the *Cr* cen-MLT-TFA complex was at $\sim 45\text{ }^\circ\text{C}$ (Fig. 6(d)). Finally, the thermal dependence plot for TFA (Fig. 6(e)) and the aggregation peak

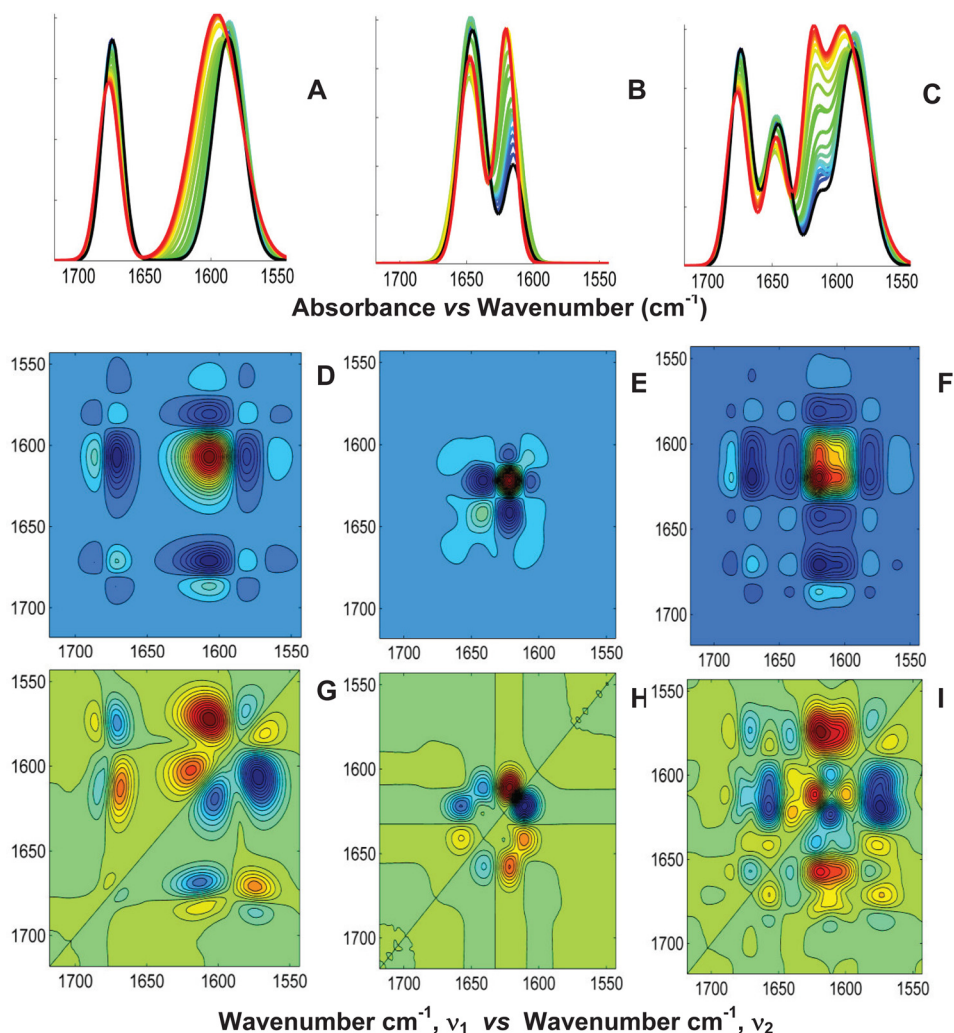


FIG. 5. Simulated spectra and 2D IR correlation plots for the ^{13}C -Cr cen-MLT-TFA complex (1:2:20, molar ratio) for the spectral region of 1720–1540 cm^{-1} . (a) The contribution of TFA (1673 cm^{-1}) and ^{13}C -Cr cen's β -sheet and helical contributions (1588 and 1597 cm^{-1}) and its corresponding (d)–(g) synchronous and asynchronous plots, respectively. (b) MLT's helical contribution (1645 cm^{-1}) and aggregation peak (1619 cm^{-1}) (e) synchronous and (h) asynchronous plots. (c) The simulated spectra, which include all five contributions and the corresponding (f) synchronous and (i) asynchronous plots, respectively.

(Fig. 6(f)) concur at the onset of the temperature when the extent of MLT aggregation event begins to increase at 45 °C.

IV. DISCUSSION

FT-IR spectroscopy allows for the simultaneous study of two proteins and their interaction when one of the protein components is homogeneously ^{13}C -labeled, causing a shift to lower wavenumbers and effectively separating the stretching modes for each protein. The magnitude of the shift is dependent on the mass of the atoms involved in the vibrational mode.²¹ We were able to discern the differences in conformational dynamics of these proteins when compared to the ternary complex. In addition, the N-D Arg stretching modes of centrin account for twelve residues in its sequence, while for MLT there is only one Arg residue. TFA was interacting mainly with centrin's Arg residues and MLT's Lys residues, thus affecting the C-terminal region of the MLT peptide.

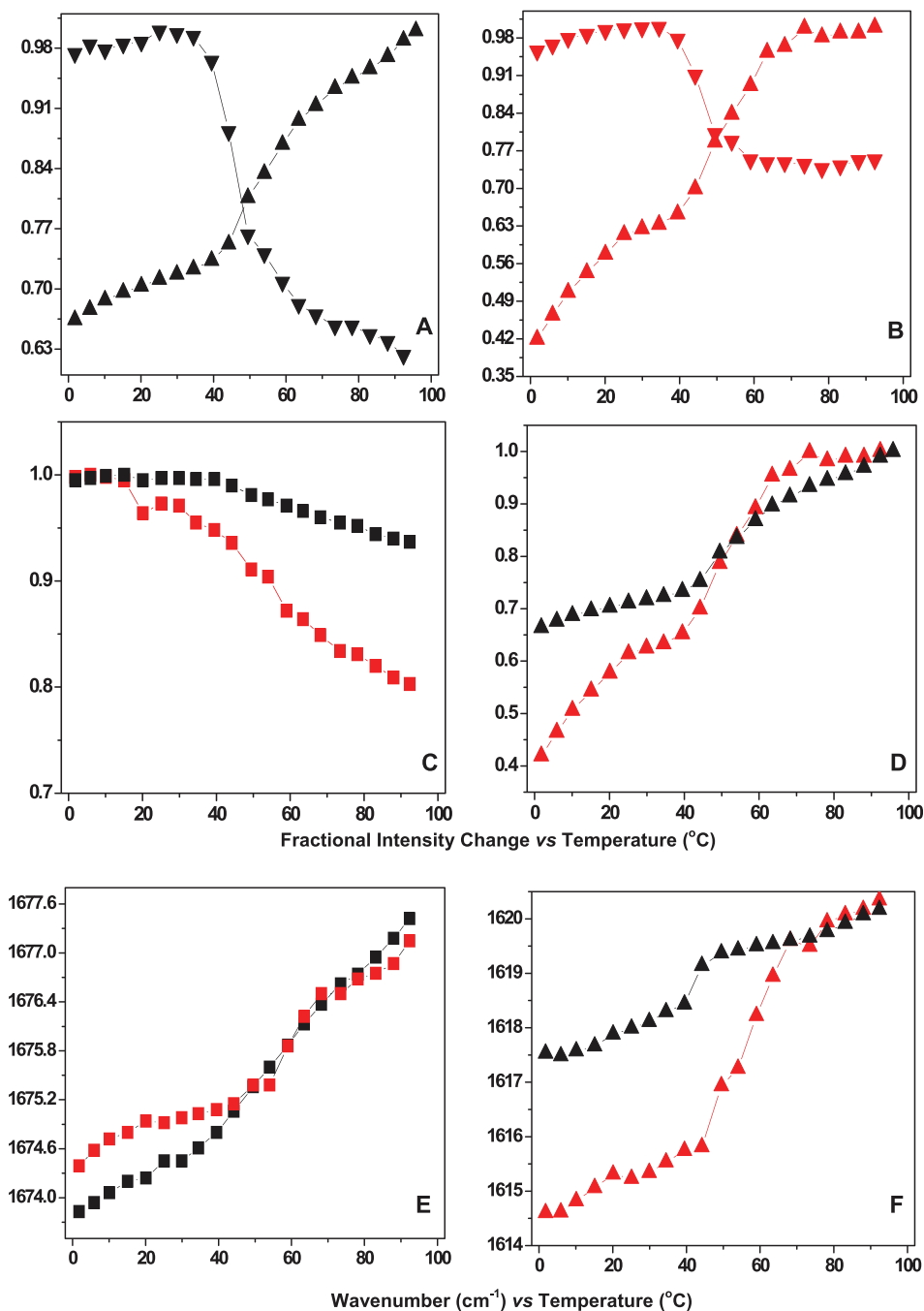


FIG. 6. Temperature dependence plots for ^{13}C -Ccen-MLT-TFA complex (1:2:20, mol ratio). In general, the color scheme is classified as experimental data set (black) and simulated data set (red). The fractional intensity change was plotted for: (a) MLT within the complex (∇) MLT 3_{10} -helical component (1645 cm^{-1}) and (\blacktriangle) aggregation (1619 cm^{-1}); and (b) MLT within the complex (∇) MLT 3_{10} -helical component (1645 cm^{-1}) and (\blacktriangle) aggregation (1619 cm^{-1}). The transition temperature for the aggregation process was reproduced $\sim 45^{\circ}\text{C}$ in the simulated spectral data. (c) TFA contribution (1673 cm^{-1}) (\blacksquare , black) experimental and (\blacksquare , red) simulated. (d) MLT's aggregation peak (1619 cm^{-1}) (\blacktriangle , black) experimental and (\blacktriangle , red) simulated. Wavenumber vs. temperature plots for: (e) TFA (\blacksquare , black) experimental and (\blacksquare , red) simulated; and (f) MLT aggregation (\blacktriangle , black) experimental and (\blacktriangle , red) simulated.

For the MLT-TFA complex, the presence of the β -turn (1685 cm^{-1}) and β -sheet (1631.3 cm^{-1}) contributions in the MLT-TFA complex allows for the assignment of antiparallel pleated sheet between MLTs. This is due to intermolecular hydrogen bonding between chains. The helical component may include the region comprised lysines and the single arginine of the

peptide, while interacting with the TFA. Also, the 2D IR correlation analysis for the components revealed the intrinsic flexibility of MLT in the presence of TFA to be within the β -sheet region (1631.3 cm^{-1}), while for centrin it was determined to be the short β -sheet (1591.8 cm^{-1}) and the hinge loops (1636 cm^{-1}) within the EF-hand domains. In contrast, TFA plays a role in the stability of the centrin protein by increasing its helical contributions by means of interacting with the centrin arginine residues in the ternary complex. The aggregation of MLT involves the newly adopted helical region of the peptide due to its interaction with centrin at low temperatures. As the temperature increases, TFA's role in further perturbing both centrin and MLT continues to elicit MLT's aggregation because the peptide is not able to interact with centrin. The MLT-centrin interaction is being inhibited by the presence of TFA. Direct molecular evidence has been presented to support this mechanism via 2D IR correlation analysis at two different molar ratios of the ternary complex. We also performed spectral simulations to further define the minimal requirements for the aggregation process to occur. This novel approach has proven vital to the understanding of complex interactions and the mechanism of aggregation involving side chain interactions as the key player. In the end, the flexibility of these protein components within the ternary complex was a key for determining the mechanism of aggregation and how TFA played a crucial role in the aggregation. More importantly, the aggregation fraction increased with a concomitant decrease in MLT's helical content, as shown for the ^{13}C -Cr cen-MLT-TFA complex (1:2:20, mole ratio).

In recent years, a tendency among spectroscopist has been to subtract the characteristic peak assigned to TFA at position 1673 cm^{-1} . The alleged justification is the physical removal of such excipient is time consuming and that the TFA has no effect on the synthetic peptide. Our goal was to provide the empirical evidence that disproves such assumption and that it may lead to faulty and erroneous molecular biophysical conclusions. Finally, circular dichroism does not allow for such a study to be performed because only electronic transitions are monitored without the possibility of discerning the contribution of the protein components by using isotope labeling as presented herein.

V. CONCLUSIONS

This approach is essential to understanding the role of excipients in proteins and their potential role in protein aggregation. Moreover, vibrational spectroscopy, 2D IR correlation spectroscopy, and the associated spectral simulations provided a complete molecular description of the aggregation mechanism and the extent of aggregation in a highly complex sample. Although we employed ^{13}C labelled isotope in this study, we envision the application of 2D IR correlation spectroscopy to evaluate protein-peptide complexes, in general. The pure components and mixtures can be evaluated under varying formulation conditions to screen for the presence of aggregate species. This type of evaluation would aid in ensuring the use of high quality proteins and peptides for crystallization.

ACKNOWLEDGMENTS

Dr. Pastrana would like to thank BioXFEL STC and the Hauptman-Woodward Institute for the opportunity to participate in this forum. This project was supported by NIH-COBRE Grant No. P20 RR16439-01 (B.P.-R.), NIH-SCORE Grant No. GM08103-28 (B.P.-R.), and the Camille and Henry Dreyfus Foundation (B.P.-R.) and the University of Puerto Rico.

¹Y. Maulet and J. A. Cox, "Structural changes in melittin and calmodulin upon complex formation and their modulation by calcium," *Biochemistry* **22**(24), 5680 (1983).

²M. Comte, Y. Maulet, and J. A. Cox, "Ca²⁺-dependent high-affinity complex formation between calmodulin and melittin," *Biochem. J.* **209**(1), 269 (1983).

³J. A. Cox, M. Comte, J. E. Fitton, and W. F. DeGrado, "The interaction of calmodulin with amphiphilic peptides," *J. Biol. Chem.* **260**(4), 2527 (1985).

⁴J. A. Cox, F. Tirone, I. Durussel, C. Firanescu, Y. Blouquit, P. Duchambon, and C. T. Craescu, "Calcium and magnesium binding to human centrin 3 and interaction with target peptides," *Biochemistry* **44**(3), 840 (2005).

⁵I. Durussel, Y. Blouquit, S. Middendorp, C. T. Craescu, and J. A. Cox, "Cation- and peptide-binding properties of human centrin 2," *FEBS Lett.* **472**(2-3), 208 (2000).

- ⁶R. A. Newman, W. S. Van Scyoc, B. R. Sorensen, O. R. Jaren, and M. A. Shea, "Inter-domain cooperativity of calmodulin bound to melittin preferentially increases calcium affinity of sites I and II," *Proteins* **71**(4), 1792 (2008).
- ⁷Y. Tanaka, S. Takahashi, Y. Mitsui, S. Itoh, Y. Iitaka, H. Kasai, and T. Okuyama, "X-ray crystallographic and chromatographic characterization of the crystals of Ca²⁺ calmodulin complexed with bee venom melittin," *J. Mol. Biol.* **186**, 675 (1985).
- ⁸M. Kataoka, J. F. Head, B. A. Seaton, and D. M. Engelman, "Melittin binding causes a large calcium-dependent conformational change in calmodulin," *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6944 (1989).
- ⁹R. H. Kretsinger, "Calcium-binding proteins," *Annu. Rev. Biochem.* **45**, 239 (1976).
- ¹⁰S. Nakayama, N. D. Moncrief, and R. H. Kretsinger, "Evolution of EF-hand calcium-modulated proteins. II. Domains of several subfamilies have diverse evolutionary histories," *J. Mol. Evol.* **34**(5), 416 (1992).
- ¹¹J. Laoukili, E. Perret, S. Middendorp, O. Houcine, C. Guennou, F. Marano, M. Bornens, and F. Tournier, "Differential expression and cellular distribution of centrin isoforms during human ciliated cell differentiation in vitro," *J. Cell Sci.* **113**(Pt. 8), 1355 (2000).
- ¹²R. Nishi, Y. Okuda, E. Watanabe, T. Mori, S. Iwai, C. Masutani, K. Sugawara, and F. Hanaoka, "Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein," *Mol. Cell. Biol.* **25**(13), 5664 (2005).
- ¹³J. L. Salisbury, "A mechanistic view on the evolutionary origin for centrin-based control of centriole duplication," *J. Cell. Physiol.* **213**(2), 420 (2007).
- ¹⁴M. A. Sanders and J. L. Salisbury, "Centrin-mediated microtubule severing during flagellar excision in *Chlamydomonas reinhardtii*," *J. Cell Biol.* **108**(5), 1751 (1989).
- ¹⁵B. Pastrana-Ríos, W. Ocaña, M. Ríos, G. L. Vargas, G. Ysa, G. Poynter, J. Tapia, and J. L. Salisbury, "Centrin: Its secondary structure in the presence and absence of cations," *Biochemistry* **41**, 6911 (2002).
- ¹⁶M. Ortiz, H. Hu, W. J. Chazin, and B. Pastrana-Ríos, "Dynamics of hydrogen/deuterium exchange in *Chlamydomonas* centrin," *Biochemistry* **44**(7), 2409 (2005).
- ¹⁷C. Weber, V. D. Lee, W. J. Chazin, and B. Huang, "High-level expression in *Escherichia coli* and characterization of the EF-hand calcium-binding protein caltractin," *J. Biol. Chem.* **269**, 15795 (1994).
- ¹⁸H. Hu and W. J. Chazin, "Unique features in the C-terminal domain provide caltractin with target specificity," *J. Mol. Biol.* **330**, 473 (2003).
- ¹⁹L. D. V. Sosa, E. Alfaro, J. Santiago, D. Narváez, M. C. Rosado, A. Rodríguez, A. M. Gómez, E. R. Schreiter, and B. Pastrana-Ríos, "The structure, molecular dynamics and energetic of centrin-melittin complex," *Proteins: Struct., Funct., Bioinf.* **79**(11), 3132 (2011).
- ²⁰B. Pastrana-Ríos, M. Reyes, J. De Orbata, V. Meza, D. Narváez, A. M. Gómez, A. Rodríguez Nassif, R. Almodovar, A. Díaz Casas, J. Robles, A. M. Ortiz, L. Irizarry, M. Campbell, and M. Colón, "Relative stability of centrins and its relationship with calcium binding," *Biochemistry* **52**(7), 1236 (2013).
- ²¹B. Pastrana-Ríos, "Understanding the mechanism of peptide unfolding," *Biochemistry* **40**, 9074 (2001).
- ²²T. Li, "Investigation of protein-protein interactions by isotope-edited Fourier transformed infrared spectroscopy," *Spectroscopy* **18**, 397 (2004).
- ²³J. L. Arrondo, A. Muga, J. Castresana, and F. M. Goñi, "Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy," *Prog. Biophys. Mol. Biol.* **59**(1), 23 (1993).
- ²⁴A. Barth, "Infrared spectroscopy of proteins," *Biochim. Biophys. Acta* **1767**(9), 1073 (2007).
- ²⁵S. Y. Venyaminov and N. N. Kalnin, "Quantitative IR spectrophotometry of peptide compounds in water (H₂O) solutions. I. Spectral parameters of amino acid residue absorption bands," *Biopolymers* **30**, 1243 (1990).
- ²⁶Y. N. Chirgadze, O. V. Fedorov, and N. P. Trushina, "Estimation of amino acid residue side-chain absorption in the infrared spectra of protein solutions in heavy water," *Biopolymers* **14**, 679 (1975).
- ²⁷I. Noda, A. E. Dowrey, C. Marcott, G. M. Story, and Y. Ozaki, "Generalized two-dimensional correlation spectroscopy," *Appl. Spectrosc.* **54**, 236A (2000).
- ²⁸I. Noda, "Advances in two-dimensional correlation spectroscopy," *Vib. Spectrosc.* **36**, 143 (2004).
- ²⁹I. Noda, "Recent Advancement in the field of two-dimensional correlation spectroscopy," *J. Mol. Struct.* **883/884**, 2 (2008).
- ³⁰Z. Sanoguet-Carrero, M. Campbell, S. Ramos, C. Seda, and B. Pastrana-Ríos, "Effects of phosphorylation in *Chlamydomonas* centrin Ser 167," *Calcium Bind Proteins* **1**(2), 108 (2006).
- ³¹B. Pastrana-Ríos, "Simulation of FT-IR spectra and 2D-COS analysis for the thermal perturbation of apo-centrin," *J. Mol. Struct.* **799**, 163 (2006).
- ³²I. Iloro and B. Pastrana-Ríos, "Simulation of FT-IR spectra and 2D-COS analysis for the H/D exchange of two related ligands," *J. Mol. Struct.* **799**, 153 (2006).
- ³³F. Lavalie, R. G. Adams, and I. W. Levin, "Infrared spectroscopic study of the secondary structure of melittin in water, 2-chloroethanol, and phospholipid bilayer dispersions," *Biochemistry* **21**(10), 2305 (1982).
- ³⁴N. Cerda-Costa, I. De la Arada, F. X. Aviles, J. L. R. Arrondo, and S. Villegas, "Influence of aggregation propensity and stability on amyloid fibril formation as studied by Fourier transform infrared spectroscopy and two-dimensional COS analysis," *Biochemistry* **48**, 10582 (2009).
- ³⁵D. Radovan, V. Smirnovas, and R. Winter, "Effect of pressure on islet amyloid polypeptide aggregation: Revealing the polymorphic nature of fibrillation process," *Biochemistry* **47**, 6352 (2008).
- ³⁶B. Pastrana-Ríos and I. Iloro, Patent USPTO Invention Docket Number 8,268,628, Method Patent: Determination of protein, peptide or peptoid aggregation, stability and viability and with the same (2008).
- ³⁷See supplementary material at <http://dx.doi.org/10.1063/1.4921219> for Table 1S, a summary of the sequential order of events during thermal perturbation for: MLT-TFA and ¹³C-Cr cen. Also, Table 2S the sequential order of events summarizing the temperature perturbation for: ¹³C-Cr cen-MLT-TFA (1:2:20 molar ratio) complex is shown. For Fig. 4S, we include the FT-IR spectral overlay of the ¹³C-Cr cen-MLT-TFA complex (1:1:10 molar ratio) within the spectral region of 1720–1540 cm⁻¹ along with the 2D IR correlation analysis.