THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Study of the γD-Crystallin Protein Using Two-Dimensional Infrared (2DIR) Spectroscopy: Experiment and Simulation

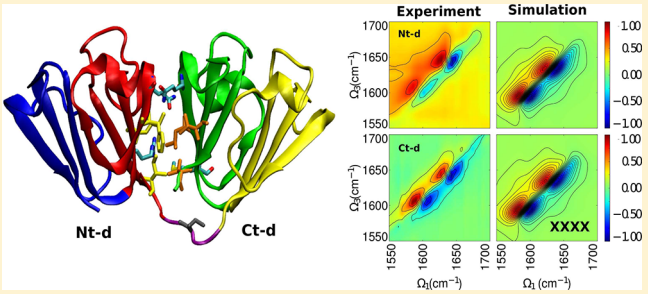
A. R. Lam,*[,†] S. D. Moran,[‡] N. K. Preketes,[†] T. O. Zhang,[‡] M. T. Zanni,[‡] and S. Mukamel[†]

[†]Department of Chemistry, University of California, Irvine, Irvine, California 92697-2025, United States

[‡]Department of Chemistry, University of Wisconsin, Madison, 1101 University Avenue, Madison, Wisconsin 53706, United States

Supporting Information

ABSTRACT: Cataracts is a misfolding protein disease in which one of the major components is the γD-crystallin protein. The conformational structure of the aggregated γD-crystallin and the interactions that cause aggregation are largely unknown. A recent experimental two-dimensional infrared (2DIR) spectroscopy study determined that the C-terminal domain has a high propensity to form β-sheets whereas the N-terminal domain forms a disordered structure in the fiber state. We present a combined computational molecular dynamics and infrared spectroscopy study of the local dynamics of these domains. The computed 2DIR signals agree remarkably well with experiment. We show that the two domains, both of which have a Greek key structural fold, experience different electrostatic environments, which may be related to the fact that the C-terminal domain is more structurally stable than the N-terminal domain. We correlate the vibrational couplings to known energy dissipation mechanisms and reveal their origin.



INTRODUCTION

Many NMR, FTIR, and X-ray crystallography studies have been used to characterize the structures of crystallin protein aggregates.[1−5] Cataracts is a misfolded protein disease that causes blindness and affects a large proportion of the elderly population over the age of 65 (National Eye Institute). Its treatment is invasive, expensive, and performed only when the disease has severely damaged the eye lens. Pathology of human cataractous lenses reveals that cataracts are composed of protein aggregates made up of crystallin proteins. These proteins are degenerative during adulthood and are damaged by radiation exposure and other chemical modifications that obstruct the passage of light through the lens, causing blurred vision that worsens with time.[6−9] Even though the causes of cataracts are known, the aggregation mechanism and the three-dimensional structure of these aggregates remain unknown.

There are three classes of crystallins: α, β, and γ. Their solubilities and stabilities in their native states are critical for keeping the transparency of the lens for a lifetime, since crystallins are not regenerated.[7,9,10] α-Crystallins are oligomers that exhibit chaperone activity, interact with unfolded proteins, and prevent aggregation.[11−15] The β- and γ-crystallins are believed to be structural proteins involved in aggregation and share similar structures composed of antiparallel β-sheets. Their secondary structures are separated into two domains, labeled as the N-terminal domain (N-td) and the C-terminal domain (C-td), which are further clustered into four subdomains called *Greek key motifs* that contain antiparallel β-strands (see Figure 1A). A key biological difference between β- and γ-crystallin proteins is that β-crystallins form domain-swapped dimers while the γ-crystallins are monomeric in solution.[1−3,16]

Wild-type γ-crystallins are stable eye lens proteins in vitro with higher free energies of unfolding than the other crystallin proteins. The two domains of a γ-crystallin protein interact intermolecularly across a hydrophobic domain interface that helps stabilize the protein.[17,18] Human γD-crystallin (HγD-Crys) is the third most abundant γ-crystallin found in young human eyes. HγD-Crys is a 173 amino acid protein exhibiting high structural homology with other γ-crystallins.[10,19] Its secondary structure is composed of antiparallel β-sheets and loops that are grouped in two domains connected by a loop linker. Mutagenesis studies have revealed that single amino acid substitutions in the HγD-Crys sequence are linked to aggregation and juvenile-onset cataracts.[20−26] The aggregation mechanism in the aging lens differs from that of the juvenile-onset cataracts because aging lenses present a large number of modifications such as methionine and cysteine residue oxidation, deamidation, and glycation[27−30]

Fluorescence studies in triple tryptophan mutants of HγD-Crys by Kosinski-Collins et al.[31] found that the C-td is more structurally stable than the N-td. Unfolding/folding experiments under physiological conditions demonstrated that the protein refolds through sequential structuring of its domains,

Special Issue: Michael D. Fayer Festschrift

Received: May 25, 2013
Revised: August 20, 2013

   dx.doi.org/10.1021/jp405159v | J. Phys. Chem. B XXXX, XXX, XXX−XXX

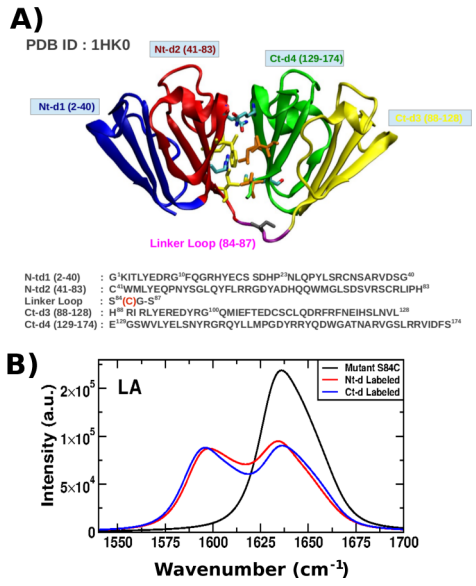


Figure 1. (A) Structure of the γD-crystallin protein showing its four Greek key motifs grouped into two domains, N-t (blue and red) and C-t (green and yellow), which are connected by a loop (purple). Their respective sequences in the crystal structure (PDB ID 1HK0) are shown, with the S84C mutation indicated in red. Residues located in the interdomains (in licorice) are responsible for the stability of the entire structure. (B) Linear absorption (LA) in the IR regime of the S84C-HγD-crystallin mutant for the full unlabeled sequence (black), the protein with the N-td labeled (red), and the protein with the C-td labeled (blue).

where the C-td folds first and the domain interface residues subsequently mediate the refolding of the N-td. Unfolding/folding kinetics studies by Flaugh et al.[17,18] revealed the formation of a partially folded intermediate structure containing the C-td in its native state and the N-td in a random coil state. They further suggested that hydrophobic residues in the intercore domain are critical in the unfolding/refolding process, wherein the refolded C-td plays the role of a nucleating center for the N-td.

Two-dimensional (2D) ultrafast infrared spectroscopy has been used to characterize geometrically ordered structures such as hairpins, amyloid fibrils, and helical dimers for many proteins.[32−38] One of the advantages of this technique is its high temporal resolution: detailed information about the kinetics of the proteins can be obtained on time scale regimes inaccessible by other techniques. It can also be applied continuously to study amyloid and protein folding kinetics on time scales of hours to months. The QM/MM computation of 2D infrared spectra has become a valuable tool in the study of the internal dynamics occurring in biological complexes such as amyloid fibrils.[37,39−43] Computations are often necessary to fully interpret experimental 2D infrared (2DIR) spectra. A recent study of γD-crystallin using isotope labeling of domains and 2DIR spectroscopy by Moran et al.[44] determined that fibril nucleation of the λ-crystallin occurs exclusively in the C-td. The authors further proposed a structural model of the fibrils composed of a segment of β-sheets formed by the C-td and another segment of random structure formed by the N-td.

In this article, we present a study involving simulation of 2DIR spectra of natively folded γD-crystallin and comparison with the experimental results of Moran et al.[44] This comparison of experiment and simulation provides new insight into differences in how the native domains are predisposed to misfolding and aggregation.

To the best of our knowledge, this is the first successful comparison of simulated and experimental 2DIR results for a protein of this size. We found excellent agreement between the two. Details of the experimental conditions may be found in ref 44.

MATERIALS AND METHODS

Protein Model. The HγD-Crys crystal structure at 1.25 Å resolution has been solved[19] and is available in the Protein Data Bank (PDB ID 1HK0). Its sequence consists of 173 amino acids that are grouped into two domains, the N-td and C-td, each of which has two subgroups called Greek key motifs. The four subgroups are labeled as N-td1(2−40), N-td2(41−83), C-td3(88−128), and C-td4 (129−171) as shown in Figure 1A. In order to mirror the experimental studies, the S84C mutant of HγD-Crys (S84C-HγD-Crys) was created and a Met residue was added in the N-Terminus using the Modeller package.[45]

Simulation Protocol. Molecular dynamics (MD) trajectories of the S84C-HγD-Crys mutant were obtained by using the NAMD package (version 2.6)[46] with the CHARMM27 force field.[47] Simulations were carried out in aqueous solutions (TIP3P water model)[48] with cubic periodic boundary conditions. The particle-mesh Ewald protocol was used to calculate the long-range electrostatic interactions, and a cutoff distance of 12 Å was used for nonbonded interactions. The 20 ns MD simulations were performed after 2 ns of equilibration with a time step of 1 fs at T = 300 K. To compute the signals, 100 MD snapshots with a time interval of 2 ps from the last 2 ns were harvested. The simulated linear absorption (LA) and 2DIR signals were computed using the SPECTRON package.[49,50]

A 2DIR experiment uses three input pulses, $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$, and $\mathbf{k}_s$ (where $\mathbf{k}_s$ satisfies $\mathbf{k}_s = \pm\mathbf{k}_1 \pm \mathbf{k}_2 \pm \mathbf{k}_3$), that interact with the protein sample with delays t_1, t_2, and t_3. The signal, defined as the transmission change of the $\mathbf{k}_s$ beam, is recorded. The signals were computed using the fluctuating vibrational exciton (FVE) Hamiltonian,[51] which is based on the Frenkel exciton model. The FVE Hamiltonian, $\hat{H}_{\mathrm{FVE}}$, has the form

$$\hat{H}_{\mathrm{FVE}} = \hat{H}_{\mathrm{S}} + \hat{H}_{\mathrm{F}} \tag{1}$$

where $\hat{H}_{\mathrm{S}}$ is the Hamiltonian of the system, given by

$$\hat{H}_{\mathrm{S}} = \sum_m \varepsilon_m \hat{B}_m^{\dagger}\hat{B}_m + \sum_{m\neq n} J_{mn}\hat{B}_m^{\dagger}\hat{B}_n - \frac{1}{2}\sum_m \Delta_m \hat{B}_m^{\dagger}\hat{B}_m^{\dagger}\hat{B}_m\hat{B}_m \tag{2}$$

and $\hat{H}_{\mathrm{F}}$ is the interaction with the optical field, $\mathbf{E}(t)$, expressed as

$$\hat{H}_{\mathrm{F}} = -\mathbf{E}(t)\cdot\sum_m \boldsymbol{\mu}_m(\hat{B}_m^{\dagger} + \hat{B}_m) \tag{3}$$

$\hat{B}_m^{\dagger}$ and $\hat{B}_m$ are the bosonic annihilation and creation operators for the mth amide-I mode, ε_m is the frequency of mode m, Δ_m is the anharmonicity of mode m, J_{mn} are the harmonic intermode couplings between modes m and n, and $\boldsymbol{\mu}_m$ is the transition dipole moment of mode m. J_{mn} remains the same for all isotopomers, as known from experiments on HγD-Crys.[44]

$\hat{H}_{\mathrm{S}}$ and $\hat{H}_{\mathrm{F}}$ were evaluated for each snapshot with the fluctuating parameters for each amide-I unit using the vibrational map of Hayashi et al.[52] Vibrational maps connect the frequencies and transition dipoles of molecular vibrations

dx.doi.org/10.1021/jp405159v | J. Phys. Chem. B XXXX, XXX, XXX−XXX

with the fluctuating local electrostatic environment. The value $\Delta_m = 16\ \text{cm}^{-1}$ was used for the anharmonicities.[53] The intermode coupling J_{mn} was calculated using the transition dipole coupling scheme.[54] Once the Hamiltonian was constructed and diagonalized, the $S^{\mathbf{k}_s}_{\nu_4\nu_3\nu_2\nu_1}$ signal was computed using eqs 115, 119, and 121 in ref 50:

$$S^{\mathbf{k}_s}_{\nu_4\nu_3\nu_2\nu_1}(\Omega_3, t_2 = 0, \Omega_1) = 2\mathrm{i} \sum_{e_4 \ldots e_1} \langle \mu^{\nu_4}_{e_4} \mu^{\nu_3}_{e_3} \mu^{\nu_2}_{e_2} \mu^{\nu_1}_{e_1} \rangle_0 I^*_{e_1}(-\Omega_1) I_{e_4}(\Omega_3) \Gamma_{e_4e_3e_2e_1}(\Omega_3 + \varepsilon_{e_1} + i\gamma_{e_1}) \mathcal{T}_{e_3e_2}(\Omega_3 + \varepsilon_{e_1} + i\gamma_{e1})$$

$$I_e(\Omega) = \frac{\mathrm{i}}{\Omega - \varepsilon_e + \mathrm{i}\gamma_e}$$

$$\mathcal{T}_{ee'}(\omega) = \frac{\mathrm{i}}{\omega - \varepsilon_e - \varepsilon_{e'} + \mathrm{i}(\gamma_e + \gamma_{e'})}$$

where $I_e(\Omega)$ and $\mathcal{T}_{ee'}$ are the one- and two-exciton Green's functions, respectively. In this study, we set the broadening factor γ_e to 5.5 cm^{-1}. $\Gamma_{e_4e_3e_2e_1}(\Omega)$ is the exciton scattering matrix with two incoming (e_1 and e_2) and two outgoing excitons (e_3 and e_4), as defined in eq 325 in ref 50.

The signals were plotted using the 2D Fourier transform of the times $t_1 \rightarrow \Omega_1$ and $t_3 \rightarrow \Omega_3$. The signals can be either nonchiral or chiral depending on the polarization configuration of the laser pulses. In the dipole approximation, there are three independent signals to leading zeroth order in $\mathbf{k}$: *xyxy*, *xyyx*, and *xxyy* (where *x* and *y* are the polarization of the laser pulses). The absorptive all-parallel signal, *xxxx*, is given by the linear combination *xxxx* = *xyxy* + *xyyx* + *xxyy*. Some combinations of nonchiral induced signals can be used to enhance desired features. The *xyxy* − *xyyx* combination suppresses the population pathways and reveals the coherent quantum dynamics through cross-peaks.[50] The combination *xxyy* − *xyxy* highlights the population dynamics and eliminates the coherent pathways in stimulated emission. For $t_2 = 0$, we have *xyxy* = *xxyy*, and the combination *xyxy* − *xyyx* would reveal only excited-state absorption.

We simulated the absorptive 2DIR signals $S^{\mathbf{k}_\text{I}} + S^{\mathbf{k}_\text{II}}$, where $\mathbf{k}_\text{I} = -\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3$ and $\mathbf{k}_\text{II} = \mathbf{k}_1 - \mathbf{k}_2 + \mathbf{k}_3$. We focused on the amide-I band, which corresponds to the backbone carbonyl vibrational motions. The *xxxx* and *xxyy* signals were compared with experiment. We further predicted the combination *xyxy* − *xyyx* signal. All spectra have been plotted using a linear scale.

RESULTS

Linear Absorption Spectra of the γD-Crys Protein. The linear absorption (LA) signals in the IR regime for the γD-Crys protein with its respective labeled domains are displayed in Figure 1B. The labeling protocol implemented in this study is that described by Moran et al.:[44] one domain is isotopically labeled with ^{13}C, and the other domain remains intact with ^{12}C. In the LA signal, the unlabeled protein displays a peak centered at 1638 cm^{-1}. When one of the domains is isotope-labeled, the signal acquires a bimodal shape with two peaks at 1595 and 1638 cm^{-1} (Figure 1B). The LA signal for the labeled C-td shows that the peak at 1638 cm^{-1} moves slightly to higher frequency, making the bimodal shape more pronounced than for the labeled N-td.

2DIR Signals. The 2DIR signals for the S84C-HγD-Crys mutant and its respective labeled domain samples were computed at $t_2 = 0$. The experimental and simulated all-parallel *xxxx* spectra of the mutant protein and its respective labeled domains are compared in Figure 2. For the unlabeled and

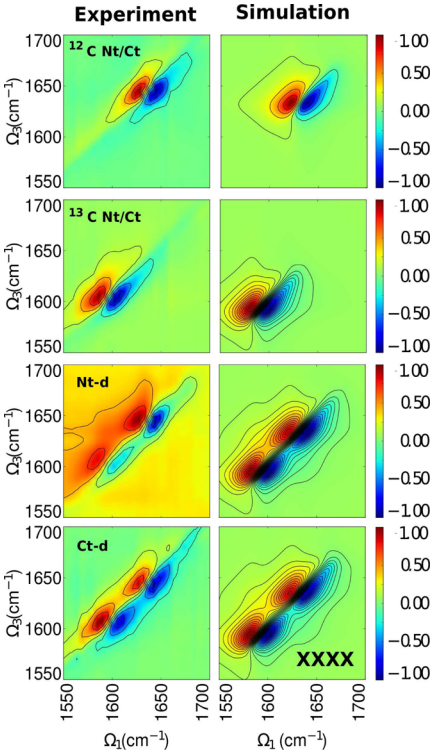


Figure 2. All-parallel *xxxx* signals $S^{\mathbf{k}_\text{I}} + S^{\mathbf{k}_\text{II}}$ from experiment (left) and simulation (right) for (top to bottom) the unlabeled full mutant protein (^{12}C Nt/Ct), the labeled full mutant protein (^{13}C Nt/Ct), the protein with the N-td labeled (Nt-d), and the protein with the C-td labeled (Ct-d).

labeled full protein (Figure 2, ^{12}C Nt/Ct and ^{13}C Nt/Ct, respectively), a pair of peaks, one negative and the other positive, are observed in each spectrum, at $\Omega_3 = 1638\ \text{cm}^{-1}$ for the unlabeled protein and $\Omega_3 = 1595\ \text{cm}^{-1}$ for the labeled protein. These frequencies correspond to a structure with antiparallel β-sheets.[55] The negative peak results from ground-state bleach and excited-state emission pathways, while the positive peak results from excited-state absorption. The broadening of these peaks is due to contributions of other secondary structures formed in the protein. These peaks were reproduced in the simulations (Figure 2, right panels). These pairs are shifted around by 40 cm^{-1}, which is consistent with the effective mass replacement of the carbon. The node slopes[44] were calculated for both of these states to compare the contributions of inhomogeneous versus homogeneous broadening (Table S1 in the Supporting Information). A node slope of 1 represents a perfectly inhomogeneously broadened peak, while a value of infinity represents a perfectly homogeneously broadened peak. The node slope of the unlabeled peak pair is in good agreement with experiment, while the node slope for the labeled full protein shows a 15% deviation from experiment. This deviation may arise because the Hamiltonian (eq 2) accounted for the isotope labeling but the MD simulations used for conformational sampling did not use a heavy carbon, which may have a slight influence on the dynamics.

dx.doi.org/10.1021/jp405159v | J. Phys. Chem. B XXXX, XXX, XXX−XXX

For the proteins with a labeled domain, two pairs of positive and negative peaks appear in the spectra. The experimental spectra of the proteins with labeled domains (Figure 2, Nt-d and Ct-d) show two pairs of peaks near Ω_3 = 1640 cm^{-1} and Ω_3 = 1605 cm^{-1} in each spectrum. Differences between the experimental signals can be seen. The peak pairs in the signal for the N-td-labeled protein are less intense and more separated than those of the C-td-labeled protein (see Table S1), while the high-frequency pair in the C-td-labeled signal are elongated and broader than those in the N-td-labeled signal. For the low-frequency pair, the N-td-labeled signal shows diffused peaks, especially the negative peak, while the positive peak is well-defined and centered at Ω_3 = 1605 cm^{-1}. In the signal for the C-td-labeled protein, the low-frequency peak pair is intense and centered at Ω_3 = 1605 cm^{-1}.

Some of these features are reproduced in the simulated signals for the labeled domains (Figure 2, right panels, third and fourth rows). The two show similar profiles with two pairs of peaks separated by 40 cm^{-1}. In the N-td-labeled signal, the negative peaks are less broad than those in the C-td-labeled signal. The peak pairs in the labeled N-td signal are more separated than those of the labeled C-td signal, as observed in the experiment and the simulated LA spectra. Comparing the simulated and experimental node slopes, we see that the node slopes attributed to the unlabeled domains are in good agreement with experiment while those attributed to the labeled domains show a larger deviation from experiment (see Table S1). Again, we believe this larger deviation from experiment may arise because all of the MD simulations used for conformational sampling used ^{12}C in the amide backbone.

We next turn to the *xxyy* signals. All of the experimental spectra (Figure 3, left panels) have a similar profile as in the *xxxx* spectra but with less intense peaks at Ω_3 = 1640 cm^{-1} and Ω_3 = 1605 cm^{-1}. This pattern is observed in the simulated data as well. However, the peak pairs in the signal for the labeled full protein are stronger and broader than those for the unlabeled full protein, which indicates that the N-td follows different dynamics in which its eigenstates exhibit a wider frequency range despite the high homology in the backbone geometry. We note that in the simulated spectra, we are able to resolve the low intensity cross-peaks better than in the experiment because of the lack of experimental (white) noise in the simulations.

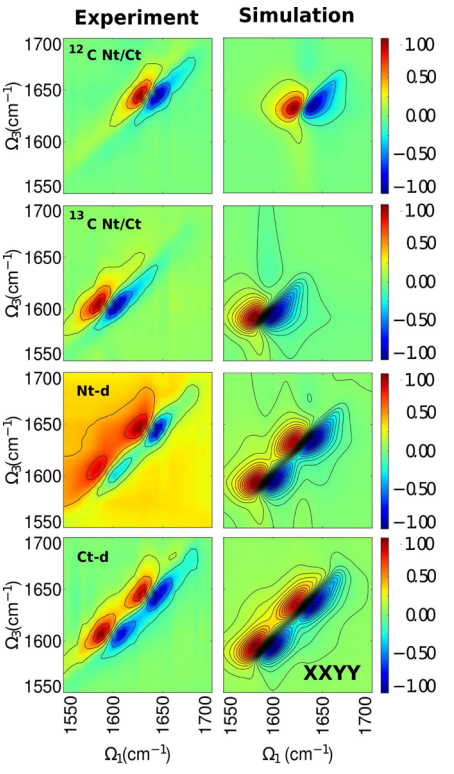


Figure 3. Same as in Figure 2 but for the *xxyy* polarization configuration.

The simulated combination *xyxy* − *xyyx* spectra for proteins with labeled domains (Figure 4) share the same pairs of peaks as observed in the *xxxx* combination. However, the N-td-labeled signal shows more elongated positive peaks along the diagonal than the positive peaks in the C-td. The spectra also show cross-peaks at Ω_1 = 1610 and 1660 cm^{-1} that are more diffuse for the labeled C-td than the labeled N-td.

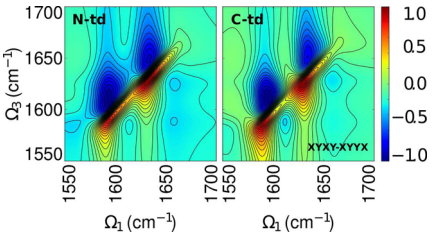


Figure 4. Simulated *xyxy* − *xyyx* signals for the isotopically labeled domains.

Both the experiments and simulations show clear differences in the 2DIR spectra when the N-td or C-td is labeled, despite their high structural homology. The experimental and simulated spectra show some weak cross-peaks that suggest couplings between the backbone amide states. These couplings depend strongly on the geometry of the protein. The domains are connected by a flexible linker loop, and one domain is essentially the reflection of the other domain. The secondary structure propensities for the protein (Figure 5) show reflection between the domains, especially where the propensity to form β-strands is high, except in the residues 40−41 and 60−61 from the N-td2 motif and their corresponding reflections in the C-td2 motif (residues 129−130 and 147−148). Other differences are observed in the turn propensities, where the loop between the N-td1 and N-td2 motifs is higher than that between the C-

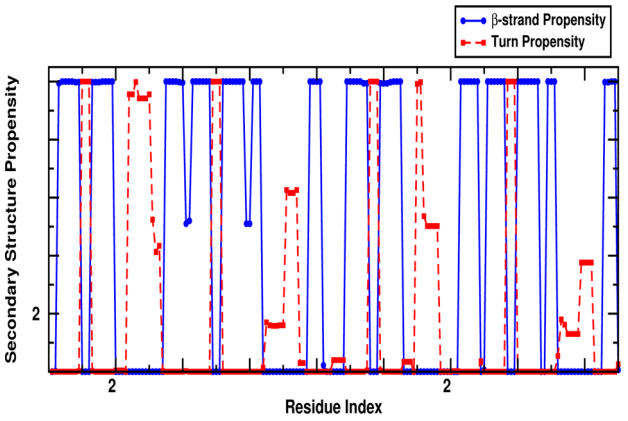


Figure 5. Secondary structure propensities of the domains of γD-crystallin indicating the segments of the sequence that have the probability to form β-strands or turns.

dx.doi.org/10.1021/jp405159v | J. Phys. Chem. B XXXX, XXX, XXX−XXX

td1 and C-td2 motifs. These differences in secondary structure indicate that the interactions between the subdomains may affect the overall stability of each domain.

The coupling matrix (Figure 6) may help reveal how the various residues are relevant in the folding and stability of the

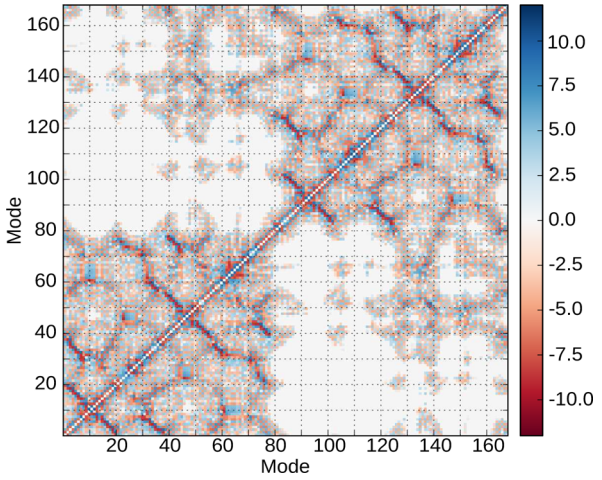


Figure 6. Coupling matrix of the γD-crystallin protein. Interdomain interactions are represented by a red-blue color scale indicating their strengths.

domains because the couplings are sensitive to distance and orientation. Couplings are observed from intramolecular interactions between residues that form β-sheet structures (antiparallel to the diagonal). Residues in the N-td (residue indexes 1−83) have weaker couplings than the corresponding section in the C-td (residue indexes 88−168). In the N-td, R36 is strongly coupled to T4 and K2. These surface residues are solvent-exposed, and it is known that mutant R36S is linked to congenital juvenile-onset cataracts.[56] Interestingly, Q54 and R58 show a strong coupling with those residues in the segment where R36 is located. The image residue for Q54 in the C-td is Q143. Q54 and R58 are solvent-exposed residues that form one of the β-strands in their respective domains and belong to the set of residues in the interface core. Deamidation of Q54 and Q143 introduces a negative charge at physiological pH into the domain interface and destabilizes HγD-Crys, lowers the unfolding kinetic barrier, and leads to the formation of a partially unfolded intermediate with a structured C-td and an unstructured N-td.[29] The R58H mutant, which results in aculeiform cataracts,[20] and the R36S mutant have lower solubilities and crystallize more readily than the wild-type protein.[19]

A cluster of couplings seen in the map corresponds to the hydrophobic interdomain region located between N-td2 (residues M43, Q54, F56, R79, I81) and C-td4 (V132, Q143, L145, and M147). These residues play a role in stabilization of the domains and a decrease in the refolding rate of the N-td, as indicated in alanine screening experiments.[17]

HγD-Crys has, on the other hand, a large number of aromatic residues (14 tyrosine, six phenylalanine, and four tryptophan); a few of them are lone residues, and the rest are distributed in pairs or sets of three residues that maintain the protein structure.[57] In the coupling map, we indicate two tryptophan pairs of interest: W42 and W68 in the N-td and their image residues W131 and W157 in the C-td. W42 and its image W131 are buried at the bottom of their respective domains. The map indicates that W42 couples with Y28 and W131 couples with D114, but the (W42, Y28) pair displays a weaker coupling than the (W131, D114) pair. W42 and its image W131 couple with W68 and W157, respectively; (W42, W68) has a weaker coupling intensity than (W131, W157), where the latter is surrounded by strong couplings generated by other residues. The other two tryptophans, W68 and its image W157, are located at the top of the structure and exhibit some couplings. For W68, we have the pairs (W68, Y6), (W68, Y28), and (W68, Y50), and for W157 we have (W157, H88), (W157, F117), and (W157, W132). It is well-established that UV radiation may contribute to the origin of cataracts as a result of absorption by aromatic residues.[57] These tryptophan residues W42, W68, W131, and W157 were studied using fluorescence quenching,[31] and it was determined that the local environment around W68 and W157 originates the native-state quenching of the protein fluorescence emission. During the refolding/unfolding process, W42 and W68 unfold at lower concentrations of GdnHCl than their images residues W131 and W157, respectively. Further analysis of unfolding and refolding of the tryptophan mutants identified an intermediate state along the wild-type HγD-Crys folding pathway with a structured C-td and unstructured N-td that could be a suitable conformational model for cataract formation.

Even though the two domains are geometrically homologous, they make different contributions to the 2DIR signals. The large separation between densities of states for the pair Q54 and Q143 could be due to their exposure to the solvent and their location in the interface domain, which could change the strengths of the interactions with other residues. The tryptophan residues considered here are buried in their respective domains. Although the two domains share basically identical backbone geometries, the frequency shifts upon domain labeling suggest that the kinetics of these domains are different, where the folding in the N-td depends on the transitions occurring in the C-td, which agrees well with experimental reports.[17,30,59]

DISCUSSION AND CONCLUSIONS

Moran et al.[44] used a combined approach of isotope labeling and 2DIR spectroscopy to study the γD-crystallin protein. They demonstrated that the use of this combined tool is valuable for the characterization of the kinetics and structure in proteins, helping to obtain additional information not available from standard methods. In their studies, they expressed labeled and unlabeled domains of the protein separately, and the full-length γD-crystallin protein was constituted by protein ligation using a cystine residue located in the connecting loop. They determined that C-td of the γD-crystallin protein is the leading domain to form amyloid fibrils and that the use of this covalent linkage between domains plays a role in determining the fibril structure, which is a striking result considering that each individual domain can form fibrils on its own.[58]

We note that our simulations are limited to the computation of 2DIR spectra of the monomeric form of HγD-Crys and do not offer a direct comparison with the 2DIR kinetics experiments.

We used a computational approach to complement these results at the level of the native state of the mutant S84C-γD-crystallin monomers. The computed spectra are in excellent agreement with experiment. Although the N-td and C-td are geometrically similar with highly conserved residues in their sequences, the average coupling map analysis shows an

dx.doi.org/10.1021/jp405159v | J. Phys. Chem. B XXXX, XXX, XXX−XXX

asymmetric contribution of some relevant residues and their corresponding image residues to the signals when one of the domains is isotopically labeled. Experiments by Mills et al.[59] concluded that the isolated N-td is less stable than the full-length protein and less stable than the C-td. A computational study by Das et al.[60] confirmed that the stability and the folding mechanism of the N-td are regulated by the interdomain interactions.

Our findings are consistent with the conclusions that Moran et al. stated about the role of the covalent linkage between domains in determinining the fibril structure and the prevalent dynamics that the C-td has over the N-td in fibril formation. It will be interesting to use 2DIR spectroscopy to analyze the coupling between domains under different environmental conditions and with various mutations.

As stated in the Introduction, we can consider HγD-Crys in terms of two coupled domains. In a very simple approximation, it is possible to think of this as a coupled two-mode exciton system:

$$H \approx \begin{pmatrix} \varepsilon_i & J_{ij} \\ J_{ij} & \varepsilon_j \end{pmatrix}$$

where ε_i is the mean site energy of domain i and J_{ij} is the coupling between modes i and j. The eigenvalues are

$$\varepsilon_{\pm} = \frac{1}{2}[\varepsilon_i + \varepsilon_j \pm \sqrt{4J_{ij}^{\ 2} + (\varepsilon_i - \varepsilon_j)^2}\,]$$

When neither domain is isotope-labeled or both domains are labeled, $\varepsilon_i \approx \varepsilon_j$, and we have (roughly) a linearly coupled homodimer. When either the N-td or the C-td is isotope-labeled, we have a linearly coupled heterodimer.

In this approximation, we expect to see two peaks at ε_+ and ε_- in the linear absorption spectrum. These peaks are separated by $\Delta\omega = [4J_{ij}^{\ 2} + (\varepsilon_i - \varepsilon_j)^2]^{1/2}$. If the system is symmetric, then $(\varepsilon_i - \varepsilon_j)^2 = (\varepsilon_j - \varepsilon_i)^2$ and the N-td- and C-td-labeled isotopomers will produce equivalent linear absorption spectra.[61] However, as seen from the calculated linear absorption spectra, this is not the case, as the peaks for the N-td-labeled isotopomer are less separated than the peaks for the C-td-labeled isotopomer. Since J_{ij} is the same for the two isotopomers, this indicates that $(\varepsilon_i - \varepsilon_j)^2$ is greater for the C-td labeled isotopomer. This is likely related to the fact that the C-td has a better-defined structure, which leads to a lower approximate site energy

While the diagonal elements of the Hamiltonian are decreased by isotopic substitution, the off-diagonal elements remain the same because the transition dipoles of the individual amide-I modes, which determine the coupling, do not change upon isotopic substitution. This was verified experimentally by comparing the diagonal cross sections of the 2DIR spectra of the unlabeled and uniformly labeled ^{13}C proteins.[44]

However, as noted by Fang et al.,[62] the cross-peaks are a direct measure of the off-diagonal anharmonic coupling. Since we assume zero mechanical off-diagonal anharmonicity, all of the off-diagonal anharmonicity must result from electrostatic coupling. In the weak-coupling limit, which is appropriate here, this off-diagonal anharmonicity can be approximated as

$$\Delta_{ij} = -\frac{4\Delta_{ii}}{(\varepsilon_i - \varepsilon_j)^2}$$

where Δ_{ij} is the off-diagonal anharmonicity and Δ_{ii} is the diagonal anharmonicity (16 cm^{-1}). Therefore, systems with a smaller $(\varepsilon_i - \varepsilon_j)^2$ will have stronger cross-peaks. In our case, the N-td-labeled isotopomer has the smaller value and therefore has stronger cross-peaks.

Using this very simple model, we conclude that the asymmetry in the linear absorption and 2DIR spectra is caused by the differences in the site energies in the two domains. The differences in these site energies are caused by the differing local electrostatic environments of residues within these two domains and likely is related to the secondary structure propensity in each domain.

■ ASSOCIATED CONTENT

Supporting Information

Diagonal peak frequencies and FWHM node slopes for the simulated and experimental *xxxx* 2DIR spectra (Table S1). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: alamng@uci.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge the support of the National Institutes of Health (Grants GM059230 and GM091364) and the National Science Foundation (Grant CHE-1058791). N.K.P. was supported by a National Science Graduate Research Fellowship. M.T.Z. acknowledges NIH DK79895.

■ REFERENCES

(1) Wistow, G.; Turnell, B.; Summers, L.; Slingsby, C.; Moss, D.; Miller, L.; Lindley, P.; Blundell, T. X-ray Analysis of the Eye Lens Protein γ-II Crystallin at 1.9 Å Resolution. *J. Mol. Biol.* **1983**, *170* (1), 175−202.

(2) Cooper, P. G.; Carver, J. A.; Aquilina, J. A.; Ralston, G. B.; Truscott, R. J. A ^1H NMR Spectroscopic Comparison of γS- and γB-Crystallins. *Exp. Eye Res.* **1994**, *59* (2), 211−220.

(3) Slingsby, C.; Norledge, B.; Simpson, A.; Bateman, O. A.; Wright, G.; Driessen, H. P. C.; Lindley, P. F.; Moss, D. S.; Bax, B. X-ray Diffraction and Structure of Crystallins. *Prog. Retinal Eye Res.* **1997**, *16* (1), 3−29.

(4) Carver, J. A. Probing the Structure and Interactions of Crystallin Proteins by NMR Spectroscopy. *Prog. Retinal Eye Res.* **1999**, *18* (4), 431−462.

(5) Purkiss, A. G.; Bateman, O. A.; Goodfellow, J. M.; Lubsen, N. H.; Slingsby, C. The X-ray Crystal Structure of Human γS-Crystallin C-Terminal Domain. *J. Biol. Chem.* **2002**, *277* (6), 4199−4205.

(6) Hoenders, H. J.; Bloemendal, H. Lens Proteins and Aging. *J. Gerontol.* **1983**, *38* (3), 278−286.

(7) Slingsby, C.; Clout, N. J. Structure of the Crystallins. *Eye* **1999**, *13* (3b), 395−402.

(8) Bloemendal, H.; de Jong, W.; Jaenicke, R.; Lubsen, N. H.; Slingsby, C.; Tardieu, A. Ageing and Vision: Structure, Stability and Function of Lens Crystallins. *Prog. Biophys. Mol. Biol.* **2004**, *86* (3), 407−485.

(9) Graw, J. Genetics of Crystallins: Cataract and Beyond. *Exp. Eye Res.* **2009**, *88* (2), 173−189.

(10) Jaenicke, R.; Slingsby, C. Lens Crystallins and Their Microbial Homologs: Structure, Stability, and Function. *Crit. Rev. Biochem. Mol. Biol.* **2001**, *36* (5), 435−499.

(11) Horwitz, J. α-Crystallin Can Function as a Molecular Chaperone. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89* (21), 10449−10453.
(12) Groenen, P. J.; Merck, K. B.; de Jong, W. W.; Bloemendal, H. Structure and Modifications of the Junior Chaperone α-Crystallin. From Lens Transparency to Molecular Pathology. *Eur. J. Biochem.* **1994**, *225* (1), 1−19.
(13) Clark, J. I.; Muchowski, P. J. Small Heat-Shock Proteins and Their Potential Role in Human Disease. *Curr. Opin. Struct. Biol.* **2000**, *10* (1), 52−59.
(14) Andley, U. P. Effects of α-Crystallin on Lens Cell Function and Cataract Pathology. *Curr. Mol. Med.* **2009**, *9* (7), 887−892.
(15) Kumar, P. A.; Reddy, G. B. Modulation of α-Crystallin Chaperone Activity: A Target to Prevent or Delay Cataract? *IUBMB Life* **2009**, *61* (5), 485−495.
(16) Cooper, P. G.; Carver, J. A.; Truscott, R. J. ^{1}H-NMR Spectroscopy of Bovine Lens β-Crystallin. The Role of the β B2-Crystallin C-Terminal Extension in Aggregation. *Eur. J. Biochem.* **1993**, *213* (1), 321−328.
(17) Flaugh, S. L.; Kosinski-Collins, M. S.; King, J. Interdomain Side-Chain Interactions in Human γD Crystallin Influencing Folding and Stability. *Protein Sci.* **2005**, *14* (8), 2030−2043.
(18) Flaugh, S. L.; Kosinski-Collins, M. S.; King, J. Contributions of Hydrophobic Domain Interface Interactions to the Folding and Stability of Human γD-Crystallin. *Protein Sci.* **2005**, *14* (3), 569−581.
(19) Basak, A.; Bateman, O.; Slingsby, C.; Pande, A.; Asherie, N.; Ogun, O.; Benedek, G. B.; Pande, J. High-Resolution X-Ray Crystal Structures of Human γD Crystallin (1.25 Å) and the R58H Mutant (1.15 Å) Associated with Aculeiform Cataract. *J. Mol. Biol.* **2003**, *328* (5), 1137−1147.
(20) Heon, E.; Priston, M.; Schorderet, D. F.; Billingsley, G. D.; Girard, P. O.; Lubsen, N.; Munier, F. L. The γ-Crystallins and Human Cataracts: A Puzzle Made Clearer. *Am. J. Hum. Genet.* **1999**, *65* (5), 1261−1267.
(21) Nandrot, E.; Slingsby, C.; Basak, A.; Cherif-Chefchaouni, M.; Benazzouz, B.; Hajaji, Y.; Boutayeb, S.; Gribouval, O.; Arbogast, L.; et al. γD Crystallin Gene (CRYGD) Mutation Causes Autosomal Dominant Congenital Cerulean Cataracts. *J. Med. Genet.* **2003**, *40*, 262.
(22) Santhiya, S. T.; Manohar, M. S.; Rawlley, D.; Vijayalakshmi, P.; Namperumalsamy, P.; Gopinath, P. M.; Löster, J.; Graw, J. Novel Mutations in the γ-Crystallin Genes Cause Autosomal Dominant Congenital Cataracts. *J. Med. Genet.* **2002**, *9* (5), 352−358.
(23) Sandilands, A.; Hutcheson, A. M.; Long, H. A.; Prescott, A. R.; Vrensen, G.; Löster, J.; Klopp, N.; Lutz, R. B.; Graw, J.; Masaki, S.; et al. Altered Aggregation Properties of Mutant γ-Crystallins Cause Inherited Cataract. *EMBO J.* **2002**, *21* (22), 6005−6014.
(24) Evans, P.; Wyatt, K.; Wistow, G. J.; Bateman, O. A.; Wallace, B. A.; Slingsby, C. The P23T Cataract Mutation Causes Loss of Solubility of Folded γD-Crystallin. *J. Mol. Biol.* **2004**, *343* (2), 435−444.
(25) Kumar, M.; Agarwal, T.; Khokhar, S.; Kumar, M.; Kaur, P.; Roy, T. S.; Dada, R. Mutation Screening and Genotype Phenotype Correlation of α-Crystallin, γ-Crystallin and GJA8 Gene in Congenital Cataract. *Mol. Vis.* **2011**, *17*, 693−707.
(26) Zhang, W.; Cai, H. C.; Li, F. F.; Xi, Y. B.; Ma, X.; Yan, Y. B. The Congenital Cataract-Linked G61C Mutation Destabilizes γD-Crystallin and Promotes Non-Native Aggregation. *PLoS One* **2011**, *6* (5), No. e20564.
(27) Stevens, A. The Contribution of Glycation to Cataract Formation in Diabetes. *J. Am. Optom. Assoc.* **1998**, *69* (8), 519−530.
(28) Hanson, S. R.; Hasan, A.; Smith, D. L.; Smith, J. B. The Major in Vivo Modifications of the Human Water-Insoluble Lens Crystallins are Disulfide Bonds, Deamidation, Methionine Oxidation and Backbone Cleavage. *Exp. Eye Res.* **2000**, *71* (2), 195−207.
(29) Flaugh, S. L.; Mills, I. A.; King, J. Glutamine Deamidation Destabilizes Human γD-Crystallin and Lowers the Kinetic Barrier to Unfolding. *J. Biol. Chem.* **2006**, *281* (41), 30782−30792.
(30) Goulet, D. R.; Knee, K. M.; King, J. A. Inhibition of Unfolding and Aggregation of Lens Protein Human γD Crystallin by Sodium Citrate. *Exp. Eye Res.* **2011**, *93* (4), 371−381.
(31) Kosinski-Collins, M. S.; Flaugh, S. L.; King, J. Probing Folding and Fluorescence Quenching in Human γD Crystallin Greek Key Domains using Triple Tryptophan Mutant Proteins. *Protein Sci.* **2004**, *13* (8), 2223−2235.
(32) Smith, A. W.; Tokmakoff, A. Amide I Two-Dimensional Infrared Spectroscopy of β-Hairpin Peptides. *J. Chem. Phys.* **2007**, *126* (4), No. 045109.
(33) Kim, Y. S.; Liu, L.; Axelsen, P. H.; Hochstrasser, R. M. Two-Dimensional Infrared Spectra of Isotopically Diluted Amyloid Fibrils from Aβ40. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (22), 7720−7725.
(34) Strasfeld, D. B.; Ling, Y. L.; Shim, S.-H.; Zanni, M. T. Tracking Fiber Formation in Human Islet Amyloid Polypeptide with Automated 2D-IR Spectroscopy. *J. Am. Chem. Soc.* **2008**, *130* (21), 6698−6699.
(35) Strasfeld, D. B.; Ling, Y. L.; Gupta, R.; Raleigh, D. P.; Zanni, M. T. Strategies for Extracting Structural Information from 2DIR Spectroscopy of Amyloid: Application to Islet Amyloid Polypeptide. *J. Phys. Chem. B* **2009**, *113* (47), 15679−15691.
(36) Shim, S.-H.; Gupta, R.; Ling, Y. L.; Strasfeld, D. B.; Raleigh, D. P.; Zanni, M. T. Two-Dimensional IR Spectroscopy and Isotope Labeling Defines the Pathway of Amyloid Formation with Residue-Specific Resolution. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (16), 6614−6619.
(37) Reddy, A. S.; Wang, L.; Lin, Y. S.; Ling, Y.; Chopra, M.; Zanni, M. T.; Skinner, J. L.; Pablo, J. J. D. Solution Structures of Rat Amylin Peptide: Simulation, Theory, and Experiment. *Biophys. J.* **2010**, *98* (3), 443−451.
(38) Remorino, A.; Korendovych, I. V.; Wu, Y. B.; DeGrado, W. F.; Hochstrasser, R. M. Residue-Specific Vibrational Echoes Yield 3D Structures of a Transmembrane Helix Dimer. *Science* **2011**, *332* (6034), 1206−1209.
(39) Zhuang, W.; Abramavicius, D.; Voronine, D. V.; Mukamel, S. Simulation of Two-Dimensional Infrared Spectroscopy of Amyloid Fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (36), 14233−14236.
(40) Jiang, J.; Abramavicius, D.; Falvo, C.; Bulheller, B. M.; Hirst, J. D.; Mukamel, S. Simulation of Two Dimensional Ultraviolet (2DUV) Spectroscopy of Amyloid Fibrils. *J. Phys. Chem. B* **2010**, *114* (37), 12150−12156.
(41) Zhuang, W.; Sgourakis, N. K.; Zhenyu, L.; Garcia, A.; Mukamel, S. Discriminating Early Stage Aβ42 Monomer Structure Using Chirality-Induced 2DIR Spectroscopy in a Simulation Study. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107* (36), 15687−15692.
(42) Lam, A.; Jiang, J.; Mukamel, S. Distinguishing Amyloid Fibril Structures in Alzheimer's Disease (AD) by Two-Dimensional Ultraviolet (2DUV) Spectroscopy. *Biochemistry* **2011**, *50* (45), 9809−9816.
(43) Wang, L.; Middleton, C. T.; Singh, S.; Reddy, A. S.; Woys, A. M.; Strasfeld, D. B.; Marek, P.; Raleigh, D. P.; de Pablo, J. J.; Zanni, M. T.; et al. 2DIR Spectroscopy of Human Amylin Fibrils Reflects Stable β-Sheet Structure. *J. Am. Chem. Soc.* **2011**, *133* (40), 16062−16071.
(44) Moran, S.; Woys, A.; Buchanan, L.; Bixby, E.; Decatur, S.; Zanni, M. Two-Dimensional IR Spectroscopy and Segmental ^{13}C Labeling Reveals the Domain Structure of Human γD-Crystallin Amyloid Fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (9), 3329−3334.
(45) Fiser, A.; Šali, A. Modeller: Generation and Refinement of Homology-Based Protein Structure Models. *Methods Enzymol.* **2003**, *374*, 461−491.
(46) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781−1802.
(47) MacKerell, A. D.; et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Comput. Chem.* **1998**, *102*, 3586−3616.
(48) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(49) Zhuang, W.; Abramavicius, D.; Mukamel, S. Dissecting Coherent Vibrational Spectra of Small Proteins into Secondary

Structural Elements by Sensitivity Analysis. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (21), 7443−7448.
(50) Abramavicius, D.; Palmieri, B.; Voronine, D. V.; Sanda, F.; Mukamel, S. Coherent Multidimensional Optical Spectroscopy of Excitons in Molecular Aggregates: Quasiparticle versus Supermolecule Perspectives. *Chem. Rev.* **2009**, *109* (6), 2350−2408.
(51) Zhuang, W.; Abramavicius, D.; Hayashi, T.; Mukamel, S. Simulation Protocols for Coherent Femtosecond Vibrational Spectra of Peptides. *J. Phys. Chem. B* **2006**, *110* (7), 3362−3374.
(52) Hayashi, T.; Zhuang, W.; Mukamel, S. Electrostatic DFT Map for the Complete Vibrational Amide Band of NMA. *J. Phys. Chem. A* **2005**, *109* (43), 9747−9759.
(53) Hamm, P.; Lim, M.; Hochstrasser, R. M. Structure of the Amide I Band of Peptides Measured by Femtosecond Nonlinear-Infrared Spectroscopy. *J. Phys. Chem. B* **1998**, *102*, 6123−6138.
(54) Torii, H.; Tasumi, M. Model Calculation on the Amide-I Infrared Bands of Globular Proteins. *J. Chem. Phys.* **1992**, *96* (5), 3379−3387.
(55) Ganim, Z.; Chung, H. S.; Smith, A. W.; DeFlores, L. P.; Jones, K. C.; Tokmakoff, A. Amide I Two-Dimensional Infrared Spectroscopy of Proteins. *Acc. Chem. Res.* **2008**, *41* (3), 432−441.
(56) Kmoch, S.; Brynda, J.; Asfaw, B.; Bezouška, K.; Novák, P.; Řezáčová, P.; Ondrová, L.; Filipec, M.; Sedláček, J.; Elleder, M. Link Between a Novel Human γD-Crystallin Allele and a Unique Cataract Phenotype Explained by Protein Crystallography. *Hum. Mol. Genet.* **2000**, *9* (2), 1779−1786.
(57) Kong, F.; King, J. Contributions of Aromatic Pairs to the Folding and Stability of Long-Lived Human γD-Crystallin. *Protein Sci.* **2011**, *20* (3), 513−528.
(58) Papanikolopoulou, K.; Mills-Henry, I.; Thol, S. L.; Wang, Y.; Gross, A. A. R.; Kirschner, D. A.; Decatur, S. M.; King, J. Formation of Amyloid Fibrils in Vitro by Human γD-Crystallin and Its Isolated Domains. *Mol. Vis.* **2008**, *14*, 81−89.
(59) Mills, I. A.; Flaugh, S. L.; Kosinski-Collins, M. S.; King, J. A. Folding and Stability of the Isolated Greek Key Domains of the Long-Lived Human Lens Proteins γD-Crystallin and γS-Crystallin. *Protein Sci.* **2007**, *16* (11), 2427−2444.
(60) Das, P.; King, J. A.; Zhou, R. β-Strand Interactions at the Domain Interface Critical for the Stability of Human Lens γD-Crystallin. *Protein Sci.* **2010**, *19* (1), 131−140.
(61) Zanni, M. T.; Hamm, P. *Concepts and Methods of 2D Infrared Spectroscopy*; Cambridge University Press: Cambridge, U.K., 2011.
(62) Fang, C.; Senes, A.; Cristian, L.; DeGrado, W. F.; Hochstrasser, R. M. Amide Vibrations Are Delocalized across the Hydrophobic Interface of a Transmembrane Helix Dimer. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16740−16745.