Separation and Analysis of Dynamic Stokes Shift with Multiple Fluorescence Environments: Coumarin 153 in Bovine β-Lactoglobulin A

Jeremy Pronchik, Jason T. Giurleo, and David S. Talaga*

Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, 610 Taylor Road, Piscataway, New Jersey 08854

Received: March 27, 2008; Revised Manuscript Received: June 2, 2008

We use time-dependent fluorescence Stokes shift (TDFSS) information to study the fluctuation rates of the lipocalin, β-lactoglobulin A in the vicinity of an encapsulated coumarin 153 molecule. The system has three unique dielectric environments in which the fluorophore binds. We develop a method to decompose the static and dynamic contributions to the spectral heterogeneity. This method is applied to temperature-dependent steady-state fluorescence spectra providing us with site-specific information about thermodynamic transitions in β-lactoglobulin. We confirm previously reported transitions and discuss the presence of an unreported transition of the central calyx at 18 °C. Our method also resolves the contributions to the TDFSS from the coumarin 153 centrally located in the calyx of β-lactoglobulin despite overlapping signals from solvent exposed dyes. Our experiments show dynamics ranging from 3−1200 ps. The analysis shows a decrease in the encapsulated dye's heterogeneity during the relaxation, which is taken as evidence of the breakdown of linear response.

1. Introduction

Solution-phase proteins are in constant motion, with their structures fluctuating throughout their life cycle. Two important examples of protein structural fluctuations are the allosteric effects arising from ligand, substrate, or inhibitor binding and the structural rearrangement of the hydrophobic core during (un)folding.

The ability of an active site of a protein to adapt to different substrates, ligands, or inhibitors is a consequence of the structural lability of proteins in general. These fluctuations occur directly between the atoms of the protein and the small molecule in question, directly affecting its local environment. The time scales and energetic barriers associated with these conformational fluctuations are vital for understanding the dynamic aspects of issues such as enzymatic specificity and drug resistance.[1]

The hydrophobic core of globular soluble proteins is thought to provide the driving force for collapse of the polypeptide chain during protein folding and provides much of the stabilization for the native fold.[2–5] The stability of the hydrophobic core is relatively insensitive to mutations that preserve net hydrophobicity, suggesting that specific interactions between amino acids in the core is less important than its solvation properties. Rearrangements and fluctuations of the core allow for changes in structure to occur at the surface of the protein as it interacts with other species. The time scales and barriers for fluctuations of the hydrophobic core reflect the physical characteristics of its structure.

Recent experiments have attempted to measure such protein conformational dynamics using time-resolved infrared spectroscopy,[6] photon echo,[7,8] transient grating spectroscopy,[7] stimulated emission,[9] as well as time-dependent fluorescence Stokes shift (TDFSS) analysis.[10–12] Attribution of the different spectroscopic signatures to protein motions, water motions, multiple dye locations, or inherent tryptophan/admolecule motions remains a fundamental challenge.

Because one of our motivations is to understand the motions that are involved in protein recognition of small molecules,[13] using a small molecule to probe the fluctuations in its immediate environment is advantageous. For this reason we focus on TDFSS of the commonly used dye coumarin 153 (C153).[14–22]

Coumarin 153. C153 serves as a good reporter of protein dielectric relaxation because of its small size, large solvatochromic shift, well behaved $S_0 \leftrightarrow S_1$ transition (no S_2 crossover),[19] and insolubility in water (upper limit 10 nM).[22] These properties make the dye popular in solvent dynamics studies.[9,15,17,22,23] Specifically, the high hydrophobicity of C153 makes it an ideal probe of the hydrophobic calyx of β-lactoglobulin A (βLG). C153 is also highly sensitive to even small changes in the solvation environment. High resolution, jet-cooled experiments showed the C153 fluorescence spectrum to shift by 1045 cm^{-1} upon gas-phase binding of a single water molecule.[24]

***β*-Lactoglobulin A.** βLG is a 162 residue, 18.4 kDa water soluble protein and is a member of the lipocalin family, which has a highly conserved eight-strand, antiparallel β-barrel (Figure 1). The accessibility of βLG has made it a popular protein to study and much is known of its pH, solvent, and temperature transitions. Our experiment (pH = 2.2) falls in the pH < 3 range where the protein exists as a monomer, but above the conformational transition at pH = 1.5.[25] The EF turn-loop in the β-barrel is folded over the entrance of the calyx[26–29] (Figure 1A) because pH = 2.2 is below the pH = 6.3 Tanford transition.[30]

^{15}N NMR relaxation measurements suggest that at pH = 2 and 45 °C, half of the β-barrel is rigid on the pico- to nanosecond time scale and the other half exhibits greater flexibility similar to that of the major α-helix. Loop EF was among the areas noted as exhibiting a high degree of flexibility.[26]

Alcohol−water binary solvents are known to stabilize α-helices.[31–33] Circular dichroism (CD) data have shown that 50% ethanol solutions induce a transition of βLG from 7% → 56% α-helix (52% → 10% β-sheet). This transition is reversible upon return to aqueous solution.[34] We exploit this feature to both solubilize C153 and open βLG allowing C153 to partition into the hydrophobic calyx.

10.1021/jp802666n CCC: $40.75
Published on Web 08/15/2008

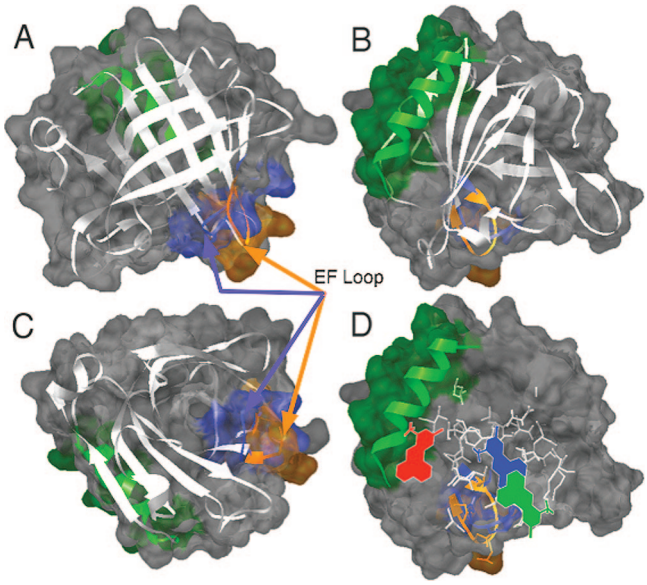

Figure 1. (A) Front view of βLG (PDB ID 2BLG, 3BLG). The β-barrel stands in front of the major α-helix (green). Loop EF is shown in its low (blue) and high (brown) pH conformations. At low pH, the entrance to the calyx is obstructed. (B) Back view. (C) View of the calyx. (D) Duplicate view of B showing the amino acid side chains of the calyx interior along with the three proposed dye binding sites. C153/βLG$_{465}$ (blue) lies in the center of the calyx, C153/βLG$_{500}$ (green) at the calyx mouth, and C153/βLG$_{550}$ (red) on an external hydrophobic patch near the major helix.

βLG experiences a minor structural transition in its major α-helix between 40−55 °C[35,36] and enters a molten-globule state beginning around 60 °C as measured by calorimetry, CD, infrared absorption, and native tryptophan fluorescence.[25,37]

A dye either covalently or noncovalently bound to a protein will usually experience multiple environments.[12,13,38,39] The central calyx of the β-barrel has been shown through X-ray structures and fluorescence energy transfer experiments to bind small, hydrophobic molecules such as retinol,[40] palmitate,[41] vitamin D,[42] cholesterol,[42,43] and the fluorophore ANS.[44,45] These studies typically report binding to one or two external patches as well - one between the β-barrel and the major α-helix[40,45–47] and the other possibly at the exterior of the calyx's closed end.[45,46] Our own docking studies of ANS also suggest three hydrophobic binding locations.[45] Of these, the interior calyx site is of particular interest because it has been shown to be anhydrous in X-ray and proton exchange studies,[26] allowing C153 to report solely the βLG dynamics.

We expect the binding affinities and water exposure of the dye to be affected by thermodynamic transitions in βLG. Dyes bound to the external hydrophobic patch are expected to report the 40−55 °C α-helix transition, whereas the two calyx sites are sensitive to the transition to molten globule.

Binding Heterogeneity. The local environment of proteins is heterogeneous; static and dynamic heterogeneity must be distinguished. Red-edge excitation (REE) fluorescence spectroscopy is a general technique for selectively exciting subpopulations of dyes in a heterogeneous medium. REE preferentially selects molecules lower in energy due to the dielectric of their local environment.[48–50] These studies have resolved the extent of heterogeneity in situations where the relaxation of the solvent is much slower than the fluorescence lifetime, as in a glass or rigid polymer. The temperature dependence of REE fluorescence spectra can help resolve whether a particular fluorescence signal arises because of multiple physical binding sites or because of slow (greater than nanoseconds) local environment fluctuations. In the latter case, as the temperature is raised the magnitude of the REE shift decreases until the local motions are fast enough to erase the memory of transient heterogeneities.

REE has been used to study the heterogeneity of 2-(*p*-toluidinylnaphthalene)-6-sulfonate noncovalently bound to βLG.[48] That study concluded that the observed REE was due to the nanosecond motions of the protein and not to multiple binding locations or conformations. It discounted the possibility of populations of different binding sites shifting at higher temperatures, which would be expected from the greater importance of entropy in the dye's free energy, explaining the REE temperature effects.

In this paper we report the encapsulation of C153 by βLG. Approximately 80% of the dye is located in the hydrophobic calyx/core of the protein. This location is of interest, because of the structural stability attributed to this region in water soluble proteins, which is believed to be important for the development and retention of the native fold.[2–5] Furthermore, the absence of water in the protein's core means the reported TDFSS can be attributed to protein motions without questions of bound or trapped water. We characterize the heterogeneity of binding of C153 using temperature-dependent steady-state fluorescence and REE. We use the TDFSS of C153 to study the fluctuations of βLG in its different binding sites. We adapt our recently developed global fitting approach[51–53] to decompose the steady-state fluorescence spectra and the TDFSS measurements into contributions from three sites. We are able to separate the signals from motions of the protein core from those of the additional binding sites. We use the results of this method to construct a multispectrum model of our system and then perform a traditional model-based fitting procedure. We discuss the dynamics that give rise to the observed behavior of the correlation function.

2. Experimental Methods

2.1. Materials. Laser grade Coumarin 153 was purchased from MP Biomedicals, Inc. (#190357). β-Lactoglobulin A (>90%) was purchased from Sigma-Aldrich (#L7880). Phosphate buffer was made from sodium hydrogen phosphate (Sigma-Aldrich 99.995%), *o*-phosporic acid (Fisher 85%), and HPLC grade water (Sigma-Aldrich). Anhydrous ethanol (99.5+%) came from Sigma-Aldrich. All reagents were used without further purification.

2.2. Encapsulation Procedure. There are two barriers to encapsulation of C153 by the hydrophobic calyx of the protein. First, C153 shows no absorbance or fluorescence when crystals are placed in water, which places its solubility below 10 nM. Second, the EF loop folds over the entrance to the calyx under monomeric pH conditions, blocking access to the hydrophobic core (Figure 1A,C; blue). While coumarin dissolves well in nonpolar solvents such as hexanes, solvents are incompatible with βLG.[54] A 50% ethanol solution solves both problems.

To encapsulate the C153 in βLG, we dissolved 3 mg of βLG into 750 μL of 60 mM sodium phosphate buffer, pH = 2.2. 750 μL of >1.2 mg/mL C153 dissolved in ethanol was combined with the protein solution to yield 1.5 mL of 110 μM βLG and >220 μM C153 in 50% ethanol/aqueous buffer. The solution was incubated at 60 °C for 4 days before the solvent was replaced with the pH = 2.2 phosphate buffer in a series of three, 4× concentrating centrifuge filtrations through a 5 kDa molecular weight cutoff membrane. This partitioned dye into 3−5% of the protein. Refolding of the protein was verified by CD.

2.3. Steady-State Absorption and Fluorescence. UV−vis absorbance of samples prepared as above were measured on a Cary 50 Bio spectrophotometer. Fluorescence spectra were measured on a JY/Horiba photon counting fluorometer with a thermostated sample holder. Fluorometer source and observation slits were matched at 2 nm. Data were collected in 1 nm steps for 0.1 s/step.

2.4. Fluorescence Lifetime. Fluorescence lifetime was obtained by time-correlated single photon counting (TCSPC). A Spectra-Physics Tsunami, mode-locked Ti:sapphire laser was operated in femtosecond mode at 800 nm (12500 cm $^{-1}$). Pulses with ~80 fs temporal width and 12.2 ns separation were passed through a ConOptics electro-optic modulator to select every sixth pulse, providing a 73 ns window for fluorescence. The polarization of second harmonic light at 400 nm (25000 cm $^{-1}$) was set with a half-wave plate and a Glan−Thompson polarizer. Observation polarization was set with a second Glan−Thompson polarizer. Grating polarization artifacts in the SP-150 monochromator were compensated by a polarization scrambler. Photons were detected with a Hamamatsu microchannel plate. Data were relayed to a Pentium PC with a Becker & Hickl SPC-630 photon counting board. The time-to-amplitude collection (TAC) window was 65 ns with gain set to maximum (15×) to create a 4.3 ns TCSPC observation window.

The 4.3 ns window was divided into 4096 data bins with $^1/_8$ dithering, resulting in an effective 3.77 ns window of 3585 bins and 1.1 ps/bin resolution. Instrument response functions (IRFs) for TCSPC experiments were achieved by collecting Raleigh scattering of a turbid, nondairy creamer solution. The fwhm of the IRF was ~47 ps, giving an effective time resolution for our experiment of ~5 ps after deconvolution. TCSPC data were collected at 68 wavenumber points with spectral resolutions of 150 cm^{-1} (from 15069 to 19270 cm^{-1}), 60 cm^{-1} (from 19420 to 20800 cm^{-1}), and 150 cm^{-1} (from 20950 to 23052 cm^{-1}). Monochromator entrance and exit slits passed 5 nm (<184 cm^{-1}), 1.25 nm (<54 cm^{-1}), and 3.75 nm (<200 cm^{-1}) bands of light in each spectral region respectively. Higher spectral resolution was chosen for the region around the spectral mean for increased information. Data were collected in each of the three spectral regions for different lengths of time to achieve 1000−8000 photon counts/bin at maximum. Data collection lasted approximately 12 h for each temperature experiment. Sample temperature was maintained by a Quantum Northwest TLC150 Peltier cooled sample holder (Spokane, WA).

2.5. Spectral Decomposition. We separated the different contributions to the inhomogeneous fluorescence spectra using a totally non-negative linear-least-squares interior point gradient (IPG) fitting method that we previously described.[51] The spectral components for the fitting basis set (*design matrix*) were taken to be pure solvent spectra (Figure 2A) that were reproduced and shifted in energy. The data spectrum was fit to a linear combination of the elements of the design matrix. The contribution of each element to the fit gives the population of each discrete dye environment (Figure 2, top). The energetic position is defined as the mean energy of the shifted basis function.

The IPG fitting algorithm optimizes a set of basis functions to describe data according to a non-negative least-squares model.[51–53] The contribution to the fit of each basis function is reported as a probability that is constrained to be non-negative. The non-negativity requirement forbids nonphysical solutions that imply negative fluorescence intensities.

The extremes of the design matrix are chosen to bracket the data (Figure 2), whereas the interior parts are made dense in areas of interest and sparse in areas of less interest. The length

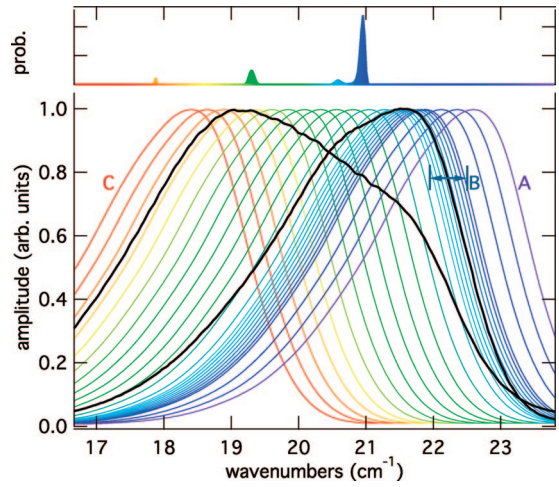

Figure 2. 2D projection of a design matrix (decimated by 10× to facilitate visualization): (A) first basis element and highest energy shift; (B) density of basis spectra increased in areas of interest to get better information; (C) last element and lowest energy shift. The normalized 0 °C (right) and 76 °C (left) steady-state fluorescence spectra are shown in black. Top panel: contributions of each design matrix element reveal the Stokes shift and populations of each C153/βLG species.

of the IPG calculations scales as (~n^2), where n is the number of elements in the design matrix. Reducing the number of elements by widening the energy spacing between bases can drastically decrease the computation time for fitting, but this comes at the cost of clarity. We found that the steady-state data (tens of millions of fluorescence counts) could be described with a grid spacing of 25 cm $^{-1}$, but the TDFSS of C153/βLG$_{465}$ (a few thousand counts) needed a local grid spacing of 5−10 cm^{-1} to get sufficient resolution. In spectral areas where the information was less important, grid spacing was increased to 20−80 cm^{-1} or removed all together. Optimizing the basis like this maximizes information while minimizing computational time.

Our results come from design matrices that describe all fluorescence species with one spectrum that was shifted spectrally to give the set of basis spectra. Basis sets made of different functions for different spectral regions were attempted[55] but did not improve the spectral decomposition.

A chief advantage of IPG fitting is its ability to handle data with distributions of spectra. Small heterogeneities in the local dye environment are expected to cause small variations in fluorescence solvatochromic shifts. This source of spectral broadening can be handled by the algorithm making fitting more robust and the results more comprehensible. Alternately, fitting with several spectra whose widths are allowed to change leads to unnatural, physically unreasonable spectra.

2.6. Normalized Time Slices. The IPG method directly fits the spectra; therefore the different TCSPC traces must be normalized to the same relative fluorescence scale and temporally shifted to correct for variations in laser intensity, monochromator slit widths, instrument spectral sensitivity, and beam path. Multiexponential functions were convoluted with an instrument response function and fit to each transient. From this, a relative time lag was calculated and each transient was temporally shifted into alignment. For TDFSS measurements, the spectra of dyes in molecular solvents are typically normalized to a log-normal distribution,[17] but due to the complexities of our system, such an approach was found to be inappropriate. We performed a modeled fit to our raw data based on our understanding of the system: three fluorescence spectra at different energies (Figure 1) that radiatively decay according

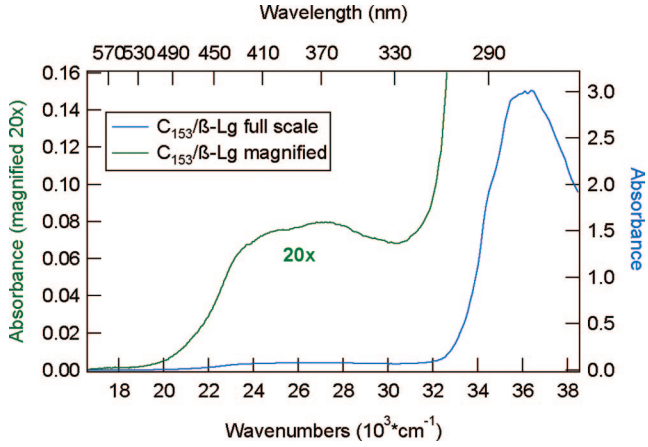

Figure 3. Typical absorbance spectrum of C153/βLG. Extinction coefficients of βLG at 278 nm and C153 at 420 nm are approximately equal; so we can see that only a few percent of proteins have dye bound to them with our sample preparation.

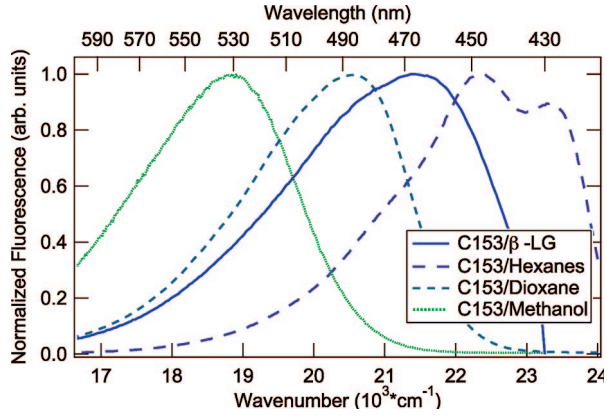

Figure 4. Encapsulating the dye: Fluorescence spectrum of C153/βLG excited at 400 nm with spectra of C153 dissolved in hexanes ($\varepsilon = 1.9$), dioxane ($\varepsilon = 2.2$), and methanol ($\varepsilon = 33$)[56] included as reference points for evaluating the dielectric environments experienced by the dye in βLG.

to single exponential functions, with the calyx-bound species red-shifting in energy according to a multiexponential function. This model was still too simple to be quantitative about TDFSS rates, but it provided us with good normalization factors for the 68 transients.[55] The normalization was similar to that of a C153/βLG steady-state fluorescence spectrum, which is not surprising because the majority of the TDFSS was complete within our observation window. The normalized and shifted data were assembled into a matrix to recreate a 3D time-evolving spectrum. Single-bin, timewise slices of this spectrum were created at 1 ps (−4 to +730 ps) and 10 ps (730 to 2840 ps) intervals for IPG analysis (Figure 12).

3. Results and Discussion

3.1. Steady-State Absorption and Fluorescence.

3.1.1 C153 Is Encapsulated by β-Lactoglobulin. Figure 3 shows the absorption spectrum from a typical preparation of the C153/βLG complex. A small band due to encapsulated C153 appears between 350 and 450 nm. On the basis of molar absorptivities of βLG ($\varepsilon_{280} = 17600$ cm^{-1} M^{-1})[44] and C153 ($\varepsilon_{420} = 18900$ cm^{-1} M^{-1}),[57] the occupancy of C153 was estimated to be 3 ± 2%. Figure 4 compares the fluorescence spectrum from the C153/βLG complex with spectra of C153 in molecular solvents. The reference spectra bracket the Stokes shift induced by the protein environment about the dye: hexanes as a nonpolar solvent, 1,4-dioxane as a solvent similar to an average protein medium,[58] and methanol as a polar, protic solvent whose dielectric constant lies between that of protein and water. The position of the C153/βLG spectrum was consistent with the dye being located in a dielectric environment between that of hexanes and dioxane, consistent with a protein hydrophobic core. Because the C153 Stokes shift is sensitive to even single water molecules,[24] the fluorescence solvatochromic shift of the main peak indicates the dye was completely protected from water. The emission spectrum shape varied slightly with sample preparation/encapsulation procedure, though the peak maximum was always consistent with the same hydrophobic environment.

Magic angle fluorescence decays include contributions from population and energetic relaxation of the excited state but cannot distinguish between relaxation induced by environmental motions and relaxation induced by fluorophore motions. Therefore it is important to establish the motion time scales of the reporting dyes in TDFSS experiments by measuring the fluorescence anisotropy decay because it is specifically sensitive to the dye reorientation. This point is often overlooked in solvation studies that report local motions about native Tryptophan moieties whose reorientation time are commonly reported in the range of the 50−500 ps.[10,59,60]

Time-dependent anisotropy measurements gave single-exponential anisotropy decays with θ(37°C, 460 nm) = 6.9 ns, and θ(37°C, 500 nm) = 6.1 ns indicting that the dye is rotating on the same time scale as a βLG monomer ($\theta_{37°C} = 6.3$ ns)[61] with no faster components observable. The steady-state anisotropy is smaller at wavelengths >530 nm.[55] This is consistent with a small amount of local reorientation for surface-bound C153. We conclude that C153 is rigidly bound by β-lactoglobulin in all but its most red-shifted environments.

3.1.2. C153 Binds to Three Different Sites in βLG. We performed REE experiments to distinguish between static and dynamic heterogeneity of the different dye environments. REE experiments are used in examining inhomogeneous broadening in solvatochromic systems in which the exchange time between solvation environments is much slower than the fluorescence lifetime. We can observe the range of dielectric environments experienced by our dye by increasing the excitation wavelength along the red edge of the absorbance band. Furthermore, we can distinguish between inhomogeneous broadening arising from slow dynamics and broadening from multiple static binding sites by observing the relative REE shifts across a range of temperatures.[50]

Figure 5 shows the results of our REE steady-state fluorescence measurements at 25 °C. The spectrum in the 400 nm excitation panel serves as our starting point for this analysis, because it shares the 400 nm (25000 cm^{-1}) excitation wavelength with our temperature dependent steady-state and time-dependent experiments. Blue-edge excitation (360 nm, 27800 cm^{-1}) showed a reduction in intensity of lower energy emissions, but little movement in the position of the maximum (~460 nm). When using wavelengths longer than 400 nm, we saw the intensity of the bluest species, near 465 nm, decrease sharply beyond 440 nm excitation and the growth of a second species near 500 nm. As the excitation wavelength was shifted to 500 nm, the first two species disappeared and the fluorescence spectrum narrowed with a peak near 550 nm.

We observed no noticeable decrease in the separation of REE species when the temperature was raised from 25 to 50 °C (Figure 6), though the contribution of the 500 and 550 nm species both increased slightly (see section 3.2). We would

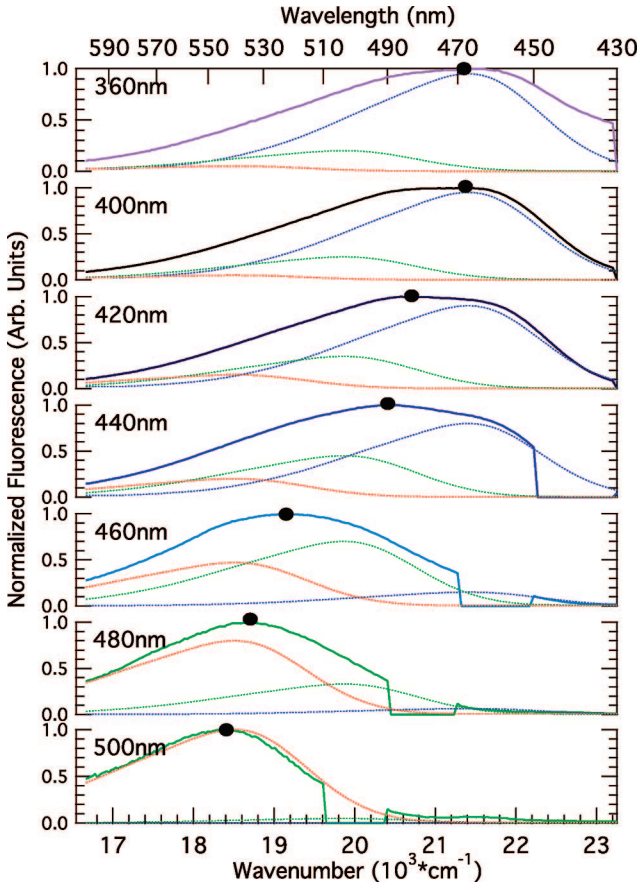

Figure 5. Normalized red-edge excitation series of C153/βLG. Panels are labeled with the excitation wavelength in the upper left corner. Increasing excitation wavelength enhanced fluorescence of longer-wavelength, water-exposed species. The REE spectra can be described with three subspectra whose approximate contributions are shown in dotted traces. Spectra maxima are marked with black circles. Raleigh scattering has been subtracted from the traces.

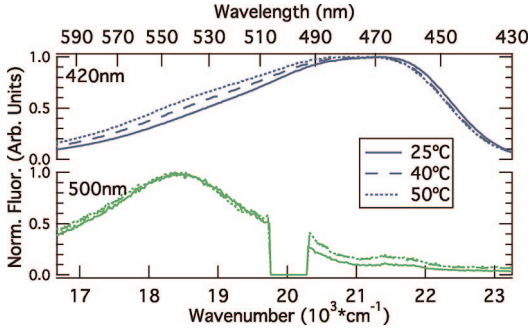

Figure 6. Increasing temperature from 25 to 50 °C yields no change in position of REE species. This indicates the REE effects to arise from several distinct dye binding locations with different dielectric environments. Dyes exchange between these locations at a rate much slower than the fluorescence lifetime. The increase in width is consistent with a thermodynamic change in the partitioning of C153 across different binding sites.

expect that the magnitude of the total shift would decrease with increasing temperature if the REE were due to the heterogeneity of a single binding location. Such a decrease would indicate faster dynamics of the dielectric medium about the single binding site accompanying the increase in temperature.[50] The persistence of the REE shift with temperature indicated our dye was bound to several distinct binding locations whose exchange time was slower than the fluorescence lifetime.

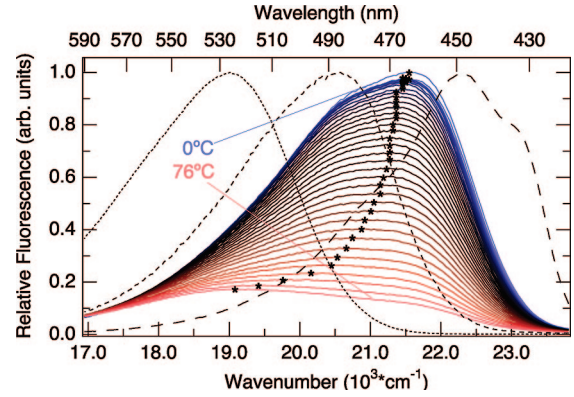

Figure 7. Temperature dependence of C153/βLG. As the solution temperature is raised, the population of dye molecules protected from water decreases relative to those exposed to it. The spectral maximum (*) red-shifts from 463 nm at 0 °C to 525 nm at 76 °C. Reference fluorescence spectra of C153 in hexanes ($\varepsilon = 1.9$), dioxane ($\varepsilon = 2.2$), and methanol ($\varepsilon = 33$) are again included.

The fluorescence solvatochromic shift in the REE data was consistent with a range of environments having dielectric constants spanning ~2−30. Given the environments available to C153 in βLG, we assigned these states as having zero, small, and large exposure to water. The blue, green, and red dotted lines in Figure 5 showed the contributions of each of these environments to the REE spectra. These three contributions adequately reproduced the excitation-wavelength-dependent changes in the spectral envelope. For the remainder of this paper, we refer to the species that give rise to these signals as C153/βLG$_{465}$, C153/βLG$_{500}$, and C153/βLG$_{550}$. This detail will be seen again in section 3.2.2 when we decompose the steady-state fluorescence spectrum over a 76 °C range.

3.2. Decomposition of Temperature-Dependent Fluorescence Spectra. The width of the C153/βLG spectrum was temperature dependent. Figure 7 shows the C153/βLG steady-state fluorescence spectrum from 0 to 76 °C. As temperature increased, the fluorescence intensity decreased, the spectrum broadened, and the band maximum shifted to the red. A temperature effect on C153 independent of the protein can be discounted, because over the same temperature range C153 dissolved in molecular solvents maintained the same spectral width and slightly *blue-shifted* by a solvent-dependent amount (hexanes ~1 nm, ethanol ~6 nm, and dioxane ~6 nm (extrapolated from 12 to 0 °C).[55]

The band maximum wavelength underwent a small jump at ~16 °C. It then underwent a gradual transition that has its midpoint at ~46 °C. The band shifted rapidly near ~68 °C. These data suggested a temperature-dependent partitioning between different protein binding sites. Quantification of these shifts in terms of protein structural transitions required spectral decomposition of the contributions. At low temperatures, C153/βLG$_{465}$ dominated with C153/βLG$_{500}$ and C153/βLG$_{550}$ gaining relative population as the temperature increased. At low temperatures (0−10 °C) the fwhm of the C153/βLG spectrum was 20% wider than spectra of C153 in molecular solvents and 65−85% wider at high temperatures (66−76 °C). This suggested environmental heterogeneity with the relative contribution of the minority species increasing with temperature. To be more precise as to how the C153/βLG spectrum evolves with temperature, we decomposed it into the individual contributions to the fluorescence.

3.2.1. Principal Component Analysis. Principle component analysis (PCA) is often used to decompose contributions to

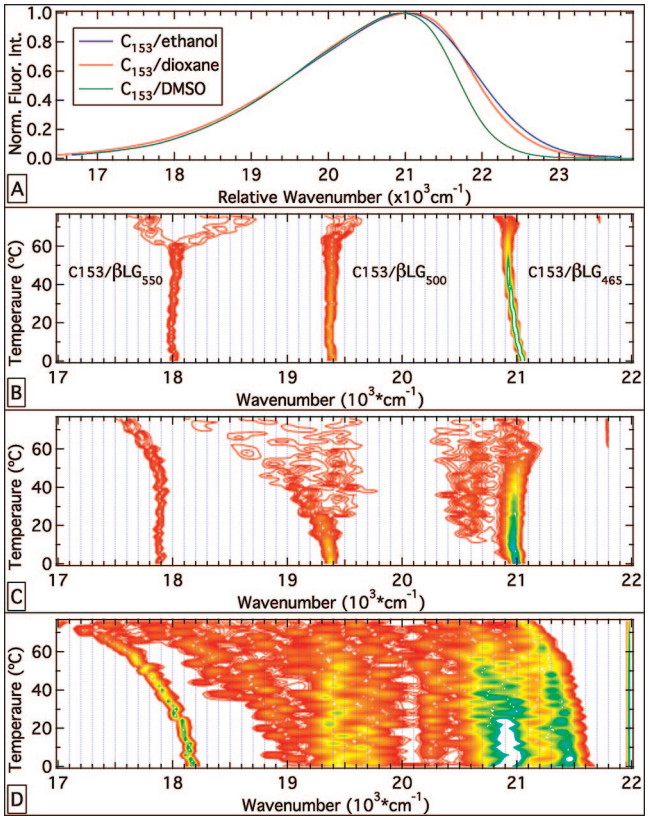

Figure 8. Effect of basis function on IPG decomposition of steady-state data: (A) Basis spectra are aligned at the midpoint of their intensity to show their contrasting shape. (B) Ethanol basis results: poor fits, but narrow distributions. (C) Dioxane: excellent fits and reasonable distributions. Thermodynamic transitions in C153/βLG subspecies are apparent. (D) DMSO: Fits are very good, but distributions are wide, obscuring transitions.

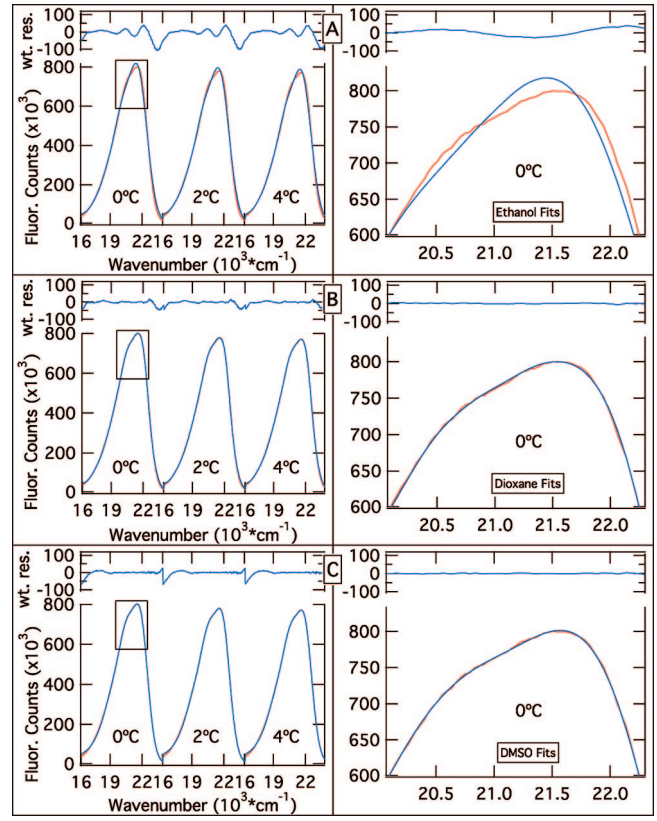

Figure 9. Effects of basis function on fits and residuals (blue) for the more sensitive low-temperature data (red): (A) The C153/ethanol's function has a different shape than the C153/βLG spectra which leads to poor fitting of low temperature spectra. (B) The dioxane basis fits well and retains a simple distribution. (C) Except for the red edge, DMSO's basis functions fit the spectra nearly as well; however, the distributions required were complex.

spectra in systematic studies of spectral evolution.[62–65] For the set of 39 temperature-dependent spectra in Figure 7, PCA produced unacceptably nonphysical results (not shown) by including unreasonably broad spectra with some components contributing negative intensities in the reconstruction. Because the concentration of C153 was ~5 μ M, we expected reabsorption of fluorescence photons to be negligible and all spectral contributions to be positive.

3.2.2. IPG Decomposition of Heterogeneous Spectra. In this section we decompose the heterogeneous C153/βLG spectrum into positive components that have the same spectral homogeneity as C153 spectra taken in molecular solvents. We have previously used an interior point gradient (IPG) method to fit ill-conditioned non-negative least-squares models[13,51] and expand its application to spectral decomposition. In this case we formed the basis set for the IPG-fit and spectral decomposition by shifting a basis spectrum in energy to account for the different Stokes shifts in the different protein environments. The wavelength maximum of C153 fluorescence was strongly dependent on the dielectric constant of the medium; however, the spectral widths and shapes were very similar. Nevertheless, there were discernible differences in their shapes (Figure 8 top panel). C153/DMSO and C153/hexanes[55] showed distinctly narrower fluorescence spectra than other solvents, the latter with partially resolved vibrational structure.

We formed five basis sets, each from a different molecular solvent, and compared their fits (Figure 8 bottom three panels, and Supporting Information). We used two criteria for judging the success of the decomposition. The first was the ability to reproduce the data as judged by the fit residuals. The second was less obvious; among basis sets that satisfied the first requirement, we chose the one that gave the simplest distribution, i.e., required the fewest basis functions to reproduce the data. Narrow distributions provide a clearer image of the transitions in a particular subspecies. Compensation for inadequate basis functions gave rise to gradual trends in the fit distributions and obscured real transitions. We expect that as the protein undergoes structural transitions a given site might change population or width. Protein transitions are expected to show cooperativity and should be relatively abrupt.

Figure 9 shows the fluorescence spectra from the first three temperatures (0, 2, and 4 °C) along with the fits derived from three different basis functions. The 0−4 °C spectra had the largest contribution from one subpopulation of dye, C153/βLG$_{465}$. Fits to this well-defined spectral component were most sensitive to the shape of the basis spectra. The remaining spectra typically fit as well as, or better than, the low temperature spectra.

C153/Ethanol Fits. Fitting consistently required three distributions of spectra to describe the data, and a temperature-wise trend was visible in the position of the C153/βLG$_{465}$ population (Figure 8B). The trend of decreased energy of this population with increased temperature is an artifact of bad fitting that arises from the shape mismatch between the basis and the C153/βLG subspectra (Figure 9A). Although the fit using the C153/ethanol basis gave narrow distributions for three subspecies, we rejected it because of the poor fits. A temperature-dependent shift in C153/βLG$_{465}$ appeared as the fit tried to compensate for a

mismatch in its blue edge shape. The C153/ethanol basis was accurate enough to decompose the data into three subspectra, but finer details were not captured.

C153/Dioxane Fits. C153/dioxane was the best basis spectrum. The fits shown in Figure 9B were very good and eliminate the error seen in the ethanol fits. Figure 8C shows distributions of basis function used in those fits. We saw several sharp transitions in our fit distributions at ~18, 45, and 60 °C, indicating structural transitions in the protein. Three types of dye environments were required to fit the temperature series.

C153/DMSO fits. The C153/DMSO basis set produced fits nearly as good as C153/dioxane except on the red edge (Figure 9C). C153/DMSO does not have the right shape, so to fit the most intense features in the blue portion of the spectrum, the algorithm sacrificed the quality of fit of the weaker red features. The red edge of the data spectra was not fit well and a gradual trend in the C153/βLG$_{550}$ distribution with temperature appeared, similar to C153/βLG$_{465}$ with the ethanol basis. The distributions required for the C153/DMSO fits, however, were much broader than the previous two examples. This stemmed from the narrowness of its profile compared to the C153/βLG subspectra. More shifted spectra must be summed to properly describe the data. The basis function is broad enough that three distributions can still be discerned but is narrower than the underlying C153/βLG widths making the fit distributions diffuse.

C153/Hexanes and C153/Acetonitrile Fits. Two other basis functions were tried whose results can be seen in the Supporting Information of this paper: C153/hexanes and C153/acetonitrile. Both basis functions yielded poor fits at low temperatures. C153/hexanes, which is narrow from lack of inhomogeneous broadening, gave very diffuse distributions similar to C153/DMSO.

A more detailed analysis of what factors determine a good basis function can be found in the Supporting Information.

3.2.3. Thermodynamic Transitions in C153/βLG. *C153 Site Assignments.* We combine the analysis thus far with fluorophore docking studies to hydrophobic pockets on the protein[45,46] to make speculative assignments as to the binding locations that give the different contributions to C153/βLG fluorescence. Because C153/βLG$_{465}$ appears to be entirely sequestered from water we assign this contribution to be from C153 located in the calyx of βLG. C153/βLG$_{500}$ is in a more polar environment, so we assign its location to the next most hydrophobic binding region of βLG, which is near the mouth of the calyx. The C153/βLG$_{550}$ is in a very polar environment, and on the basis of the structure of βLG we speculate that is it packed near the main helical segment. These locations appear in Figure 1. Free C153 in H_2O was not observed.

0−18 °C. Using the C153/dioxane IPG spectral decomposition, and keeping our tentative assignments in mind, we can begin to model thermodynamic behavior of βLG. From 0 to 18 °C there was a gradual red shift of both C153/βLG$_{465}$ and C153/βLG$_{500}$, and no shift in C153/βLG$_{550}$ in contrast to the blue shift typically observed in molecular solvents. This behavior was consistent with a "one-state" transition, that is, a gradual temperature-dependent change in properties.

18 °C. The trend of C153/βLG$_{465}$ to lower energy ended at 18 °C when the C153/βLG$_{465}$ distribution gained a lower energy shoulder with a shift of −375 cm^{-1}, indicating a change in the calyx binding site. The C153/βLG$_{500}$ distribution began to broaden and red-shift at this temperature, which is consistent with an increase in solvent exposure near the mouth of the calyx. There is no structural transition for βLG at 18 °C reported in the literature. One explanation is that the EF loop (Figure 1a) becomes mobile at this point and water is able to transiently

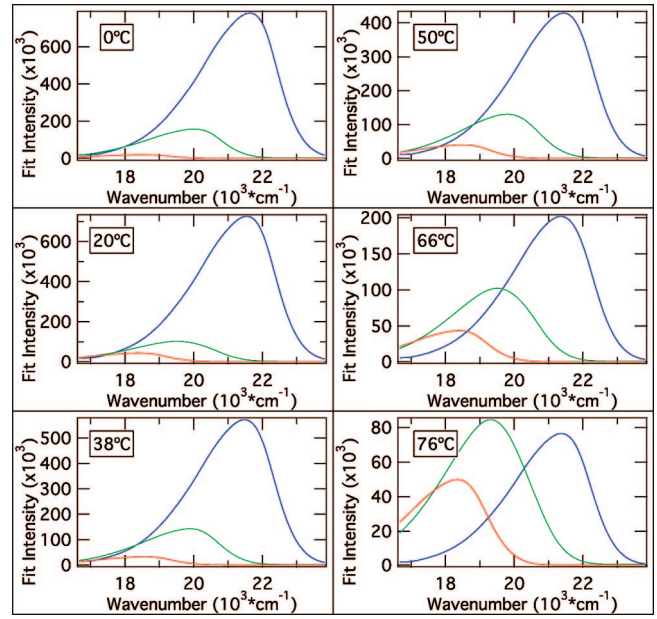

Figure 10. Reconstructed spectra derived from dioxane basis IPG fitting of temperature dependent steady-state spectra: C153/βLG$_{465}$ (blue), C153/βLG$_{500}$ (green), C153/βLG$_{550}$ (red). The relative contribution of C153/βLG$_{465}$ to the total spectrum decreases sharply at high temperatures when the protein transitions to a molten globule. At 76 °C the absolute intensity of C153 was 10% of its value at 0 °C.

access the calyx-bound dye. In the steady state, we observe the contributions from both open and closed states simultaneously. As temperature increases, the relative importance of this new state increases, suggesting that the higher-entropy open state is becoming more favorable than the closed state. The absolute magnitude of the open state remains fairly constant, whereas the contribution of the closed state diminishes.

Because the C153/βLG$_{500}$ distribution also became broader beginning at this temperature the same protein transition was affecting the local environment of it as well. The broadening of the distribution was consistent with increased heterogeneity upon water exposure. C153/βLG$_{550}$ did not appear to be changed by this protein transition.

18−45 °C. From 18 to 45 °C C153/βLG$_{550}$ underwent negligible changes, C153/βLG$_{500}$ increased in width, and C153/βLG$_{465}$ had a consistent profile with the main feature gradually decreasing in population.

45 °C. Near 45 °C there was a break in the C153/βLG$_{550}$ distribution, whereupon it began to systematically redshift with further increases in temperature. This is consistent with our assignment of C153/βLG$_{550}$ as being associated with the α-helix hydrophobic patch, because this is where βLG undergoes a transition in this structural element.[35,36] By 45 °C the main peak of C153/βLG$_{465}$ had become comparable in magnitude to the shoulder.

45−60 °C. Between 45 and 60 °C C153/βLG$_{550}$ red shifts, suggesting increased exposure to water. C153/βLG$_{500}$ is decreasing in absolute amplitude, though its relative contribution is increasing due to the more rapid loss of C153/βLG$_{465}$ population. The bluer component continues to decrease in amplitude and by 60 °C the amount of signal from C153/βLG$_{465}$ has decreased to the point that it is barely resolved from the lower energy shoulder.

60−76 °C. After 60 °C C153/βLG$_{550}$ continues to red shift. C153/βLG$_{500}$ has reduced dramatically in absolute amplitude. The lower energy contribution to C153/βLG$_{465}$ has mostly

disappeared. We understand this to be the collapse and disruption of the calyx structure upon formation of a molten globule state.

Site Binding Affinities. Figure 10 shows the relative populations of the three species as the temperature was increased from 0 to 76 °C. First, C153/βLG$_{465}$ would be expected to have the best enthalpy, but worst entropy of binding, because the calyx of the β-barrel is the most hydrophobic region of the protein, yet the most restrictive of binding geometry. The broad hydrophobic patches at the entrance to the calyx and the exterior of the protein would be expected to bind our hyrophobic dye more poorly, but offer more configurational states corresponding to greater entropy. Therefore, as the temperature increases and the entropy term becomes more important in the system's free energy, we see a relative increase in the redder populations. This is a simplistic picture of our system that does not consider the free energy changes in binding locations accompanying structural transitions.

3.3. TDFSS Analysis. The time-dependent fluorescence Stokes shift is expected to report on the motions of the solvent and protein surrounding the bound C153. Because there are three binding environments, each will contribute to the bulk TDFSS signal. C153/βLG$_{465}$ has the smallest fluorescence Stokes shift ($\sim$450 cm $^{-1}$) but would be expected to have the slowest dynamics because the relaxation is due to protein and amino acid side chain motions. C153/βLG$_{500}$ has a larger shift $\sim$2170 cm^{-1}. The partial exposure to water might result in multiple contributions to the dynamic relaxation. The water in question may be organized at some level by the protein. C153/βLG$_{550}$ has the largest shift $\sim$3550 cm^{-1} and therefore is expected to have the largest exposure to water. The dynamic relaxation should be dominated by the contribution from the water. Again, because the C153 is bound to the protein, any water involved in the solvation of C153 is also part of the protein hydration layer. Ideally one would like to separate each of these contributions to determine the dynamics at each site.

Standard analysis of TDFSS data typically involves fitting individual time-resolved fluorescence transients to multiexpontial decays to obtain the deconvoluted time-resolved intensity as a function of wavelength. This spectrum is normalized and fit to a log-normal distribution. The dynamic behavior is described in terms of the timewise development of the parameters describing the log-normal distribution.[17] This analysis assumes a single environment for the fluorescing species resulting in a simple spectral function. Though this assumption is generally valid for dyes in molecular solvents, it is not adequate for our heterogeneous system with dye molecules responding to disparate dielectric media (Figure 11).

Figure 11 shows the results of the standard analysis for data taken at three different temperatures. If only the 15 °C data were examined and fit to a multiexponential function, one might be tempted to associate the different contributions to the relaxation to the different environments, perhaps invoking some slow dynamics. However, at the higher temperatures, it is clear that artifacts are present in the reduced data. The 37 and 50 °C data both show an apparent recovery of the Stokes shift energy that is thermodynamically forbidden. The decay of the mean of the TDFSS is substantially perturbed by the presence of multiple species in the fluorescence. This is also occurring at 15 °C though it is not obvious because the populations of the minority species is small.

In particular the different environments have different fluorescence lifetimes. Because the different lifetimes have different spectra, the net result is a spectral shift that is primarily sensitive

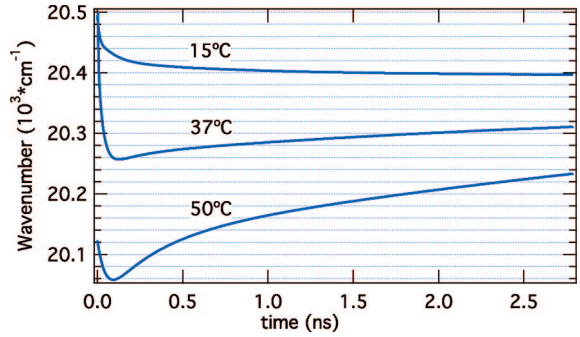

Figure 11. Reduced, deconvoluted data spectral means at 15, 37, and 50 °C. At higher temperatures the contribution of solvated species (lower energy) to the total spectrum is greater and their shorter fluorescence lifetimes gives the appearance of a long time recovery of energy in the mean.

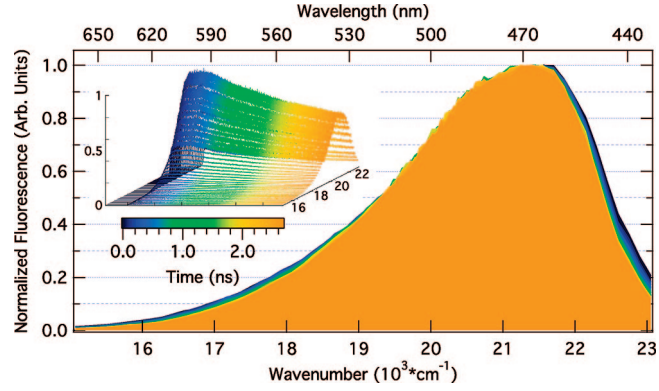

Figure 12. Spectral slices of normalized raw TCSPC data. The inset shows the time evolving spectrum (in wavenumber), which is sliced timewise for IPG fitting (main image). The blue edge shows the TDFSS of the encapsulated species, and the red edge shows the partially quenched solvated species decaying at a faster relative rate.

to these differences, rather than to the dynamics of the medium. The first moment shifts according to the change in the weighted mean of the population, which is governed by lifetimes rather than due to dynamic solvation. We corroborated this by fitting the deconvoluted data with three spectra that have solvent exposure-appropriate rates (not shown).

To isolate the individual contributions to the complex TDFSS signature, we use the same IPG spectral decomposition technique we applied to the temperature-dependent data. Here, the time-evolving fluorescence spectra are reconstructed according to our normalization procedure outlined in section 2.6. The spectrum and time slices for a 37 °C sample are shown in Figure 12. The $t = -4$ ps slice is colored blue and the last slice, $t =$ 2.84 ns, orange. Time zero is set as the maximum of the IRF.

As can be seen on the high energy edge of the graph, the spectrum shifts toward lower energy with time. We explain this as our dye in the central calyx lowering its energy as the protein dielectric relaxes around the excited state, manifested as a TDFSS. The other salient feature is the lack of visible TDFSS on the low energy slope of the spectrum. Here the fluorescence intensities of the lower energy, hydrated species decrease at a faster radiative rate compared to the dye in the central calyx. This is consistent with our expectations of dye molecules exposed to water having greater fluorescence quenching. This change in lifetime between the different species is the cause of the artifacts shown in Figure 11.

On the basis of TDFSS experiments that study small molecular dyes in water, we expect the hydrated species, C153/

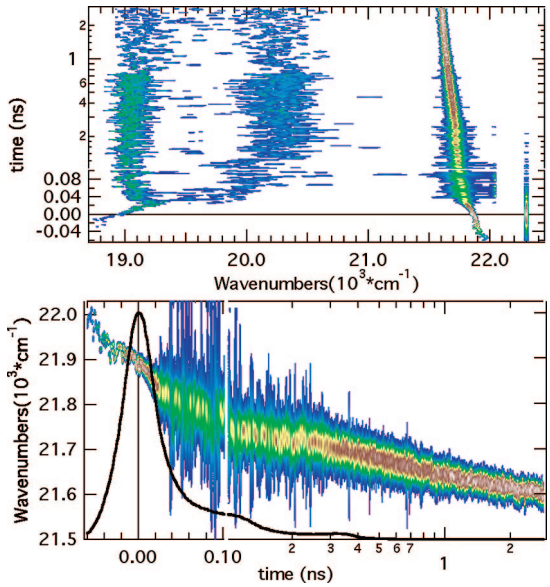

Figure 13. Top: IPG fitting of TDFSS time slices. From left to right the distributions are C153/βLG$_{550}$, C153/βLG$_{500}$, C153/βLG$_{465}$, and Raman scattering. Data spectra are fit each 1 ps up until 730 ps, then every 10 ps. Bottom: the shifting component, C153/βLG$_{465}$, can be isolated for analysis. $t = 0$ is defined as the maximum of the instrument response function (black). The wavenumber axes give the mean of the basis function in the distributions. The temporal axes have a linear/log scale break at 0.1 ns.

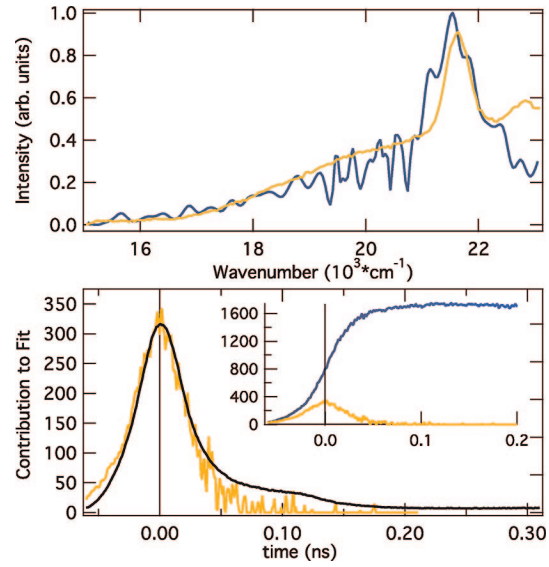

Figure 14. Removing Raman scatter: (top) steady-state Raman spectrum of βLG in pH = 2.2 buffer (orange) along with an early (−74 ps) time slice (blue); (bottom) total contribution of the Raman component of the design matrix (orange), the system's instrument response function (black), and summed contribution of the C153/βLG distribution shown together. Raman scatter interferes with early time slices and needs to be separated from the C153/βLG$_{465}$ distribution before calculating the latter's shift time constants. The inset of the bottom panel compares the Raman contribution (orange) to the summed intensity rise (blue).

βLG$_{500}$ and C153/βLG$_{550}$, to complete their Stokes shifts in less than 1 ps with the majority of the relaxation occurring in tens of femtoseconds.[17,66] These motions are faster than our instrumental time resolution and the species are expected to appear static at their fully relaxed spectral positions. This is consistent with our observations.

The solvent-protected species, C153/βLG$_{465}$, is expected to monotonically decrease in energy as the hydrophobic core relaxes about the excited-state charge distribution. Time resolved infrared measurements of the amide I band place the time scale of protein backbone relaxation at 6−8 ps (amide I, TDIR).[6] The B1 domain of streptococcal protein G (GB1 peptide) probed by the TDFSS of aladan incorporated at different sequence locations showed a range of time scales from 50 fs to 10 ns.[12] Of particular interest they observed time scales for the relaxation of aladan in sequestered environments of GB1 that were substantially different from those observed when aladan was exposed to solvent (∼1, ∼10, ∼100 and ∼1000 ps). Relaxation of exposed dyes was essentially complete within 1 ps with only a small contribution at ∼100 ps.

3.3.1. IPG Decomposition of TDFSS. The results of our IPG decomposition of the spectral time slices can be seen in Figure 13. We again see three species and consider them to be the same C153/βLG$_{465}$, C153/βLG$_{500}$, and C153/βLG$_{550}$ we saw in the steady-state decomposition. The C153/βLG$_{465}$ shoulder species seen in the steady-state decomposition is not visible. At 37 °C, its magnitude is approximately 50% that of C153/βLG$_{550}$, which places it near the limit of detection for the signal-to-noise of our TCSPC data. It is possible that this species has been incorporated into the overall width of the C153/βLG$_{465}$ distribution and may contribute to dynamic changes in width we observe. However, we do not see a corresponding change in our data of the appropriate magnitude or time scale.

3.3.2. Contribution from Raman Scattering. Interpretation of the earliest features in the TDFSS measurement was complicated by a contribution of Raman scattering to the signal centered at 21600 cm^{-1}. To remove this signal from our data, we used a steady-state Raman scattering spectrum obtained from a sample of βLG in pH 2.2 buffer (Figure 14 top, blue).

We placed the contribution from the Raman basis function at 21650 cm^{-1} for display in Figure 13 (top). The bottom panel of Figure 14 shows the Raman contribution to the time-slice fits (orange) along with the instrument response function (IRF, black) of our system. The temporal position of the IRF was previously determined in convoluted, multiexponential fits to both the mean of the C153/βLG$_{465}$ distribution and the summed intensity of all IPG distributions. Figure 14's inset shows the summed intensity of the C153/βLG distribution (blue) along with the Raman peak (orange) to show their relative position and intensities. The IRF and Raman contributions are coincident, showing that the IPG spectral decomposition method accurately projected out the correct temporal evolution of the Raman signal without any input of the instrumental response function, even in the presence many overlapping signals. This shows that IPG decomposition preserves even the fastest (instrument limited) spectral features.

3.3.3. Dynamics in C153 Binding Sites. *C153/βLG$_{550}$ and C153/βLG$_{500}$.* During the initial signal rise (t < 50 ps), C153/βLG$_{500}$, and C153/βLG$_{550}$ have especially weak signals (tens to hundreds of TCSPC counts at max), causing their distributions to coalesce and be otherwise unstable during fitting. We interpret these shifts with time as artifacts of fitting low intensity signals, not true shifts in energy. This arises when the signal-to-noise is such that there is only enough information to define the first moment of the distribution.[51] These fluctuations can also be seen as a result of slight mismatches in the data and the basis spectra as seen in the Raman contribution. As the intensities of C153/βLG$_{550}$ and C153/βLG$_{500}$ increase (t > 60 ps), the fit resolves them at their steady-state positions (Figure 8b), which persist for the rest of the data set. We attribute the early change in position to fitting instability; however, any slow solvation of

these species must be faster than ~40 ps. Therefore we conclude that the TDFSS of C153/βLG$_{500}$ and C153/βLG$_{550}$ is complete within the limits of instrument response and therefore faster than ~10 ps with negligible contributions from slow solvent dynamics.

C153/βLG$_{465}$. The C153/βLG$_{465}$ distribution shifts toward lower energy over several time scales and shows changes in its width. Its clarity relative to the other two species is attributed to both its superior signal-to-noise and our construction of the design matrix, which has been made finer in the area describing it ($\Delta E = 5$ cm^{-1}) relative to C153/βLG$_{500}$ and C153/βLG$_{550}$ ($\Delta E = 40$ cm^{-1}). Concentrating the basis spectra like this allows us to maximize our information from the areas with the best signal-to-noise and interesting dynamics while minimizing computational time (Figure 2).

There is a narrow feature appearing during the rise of the intensity (-40 ps $< t <$ 30 ps) that starts at 21250 cm^{-1} and ends at 21170 cm^{-1}. The temporal profile of this feature is narrower than the instrument response when measured at a single spectral position. However, the overall intensity profile measured at the mean position is essentially identical to the instrument response function. The low intensity of this contribution during the early part of the intensity rise causes the fit to reproduce the mean value and mostly neglect any width present as discussed above. Because the overall temporal evolution appears to be instrument limited, we can propose two sources for this feature. Either an unresolved part of the TDFSS or spectral mismatch between the Raman basis function and the true Raman spectrum could give rise to this signal.

Toward the end of the intensity rise (20 ps $< t <$ 40 ps) the distribution broadens significantly. This increase in width is nearly instrument limited. At this point the width is significantly broader than a neat solvent spectrum. The distribution shifts about 100 cm^{-1} during the first 100 ps and levels out around 200 ps with little change in width. The distribution resumes shifting from 200 to 700 ps with the width decreasing as well during this time. The distribution is shifting and narrowing more slowly from 800 to 2800 ps. The distribution appears to be essentially unimodal for the entire time course. Several dynamic processes must be invoked to explain the presence of four features of the TDFSS evolution on the <100, 290 and 1000 ps time scales.

3.3.4. Modeled Fitting of C153/βLG$_{465}$: Energy. *Description of Model.* The calyx bound dye is energetically separated enough from the other two dyes that we can isolate its behavior (Figure 13, bottom). We fit the C153/βLG$_{465}$ IPG distribution with an evolving Gaussian spectral distribution that is timewise convoluted with an instrument response function (IRF(t)).

$$I_{\text{fit}}(\tilde{\nu}, t) = \text{IRF}(t) \otimes P(\tilde{\nu},t) \tag{1}$$

$$P(\tilde{\nu},t) = \frac{A(t)}{\sigma(t)\sqrt{2\pi}} e^{-(\tilde{\nu}-\langle\tilde{\nu}(t)\rangle)^2/2(\sigma(t))^2} \tag{2}$$

$$A(t) = a_0 e^{-k_r(t-t_0)} \tag{3}$$

$P(\tilde{\nu},t)$ is the normalized Gaussian probability distribution that varies in energy and time, $A(t)$ is the amplitude which decays from an initial value, a_0 according to a single exponential with rate k_r.

$$\langle\tilde{\nu}(t)\rangle = \tilde{\nu}_0 + \sum_{i=1}^{4} \tilde{\nu}_i (1 - e^{-k_{\tilde{\nu}_i}(t-t_0)}) \tag{4}$$

The center of the Gaussian, $\langle\tilde{\nu}(t)\rangle$, decreases in time according to a multiple exponential function with shift amplitudes $\tilde{\nu}_i$ and exponential rates $k_{\tilde{\nu}_i}$.

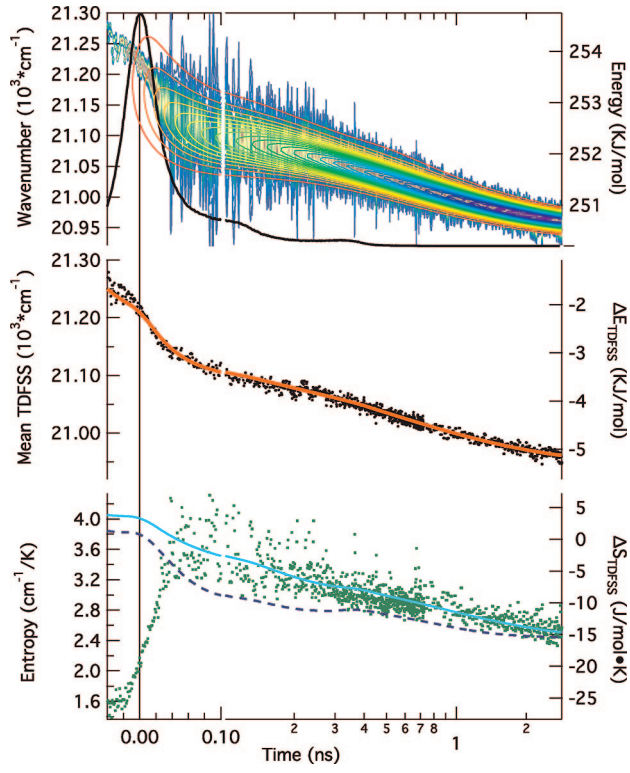

Figure 15. (Top) modeled fit of the C153/βLG$_{465}$ distribution. A Gaussian distribution monotonically decreases in energy and width according to independent multiexponential functions. (Middle) calculated mean of the IPG decomposition (black circles) and mean of the Gaussian fit. We have insufficient data to fit the fastest decay. The temporal axis has a linear/log scale break at 0.1 ns. (Bottom) evolution of the entropy as calculated from the IPG distribution (square dots) and the modeled fit to the IPG distribution (line). The broken line shows the contribution to the entropy from instrument broadening alone.

These magnitudes and rates represent the degree of energetic relaxation of the dye and the time scale of each relaxation process. Finally, the width of the distribution, $\sigma(t)$, also decays from an initial width, σ_0, according to a biexponential with magnitudes and rates independent of those for the energy.

$$\sigma(t) = \sigma_0 + \sum_{i=1}^{2} \sigma_i (1 - e^{-k_{\sigma_i}(t-t_0)}) \tag{5}$$

We also attempted to model the evolution of the distribution by considering exponential processes connecting two or more solvation states with specified widths and positions. However, we were not able to obtain satisfactory fits with the discrete model.

Evaluation of Fit to Model. The convoluted multiexponential evolving Gaussian fit appears as the smooth contours in Figure 15. It succeeds in reproducing the IPG distribution except for the width at $t <$ 40 ps. Prior to 40 ps the convoluted fits are always broader in both time and energy than the IPG distribution. In an ideal situation with no heterogeneity (δ-function spectral distributions), the convolution would broaden the IPG distributions to the same width as the IRF. We see instead a much narrower distribution at early time. We believe this to be the same artifact we see with the early times of C153/βLG$_{550}$ and C153/βLG$_{500}$, in which the IPG distribution is narrow from low signal-to-noise.

The mean TDFSS energy is reproduced by the shifting Gaussian fit as shown in the middle panel of Figure 15. Instrument response limited mean energy shifts have a weak

effect on the $t < 0$ region of the convoluted mean energy evolution. The shifting of the average energy in the convoluted data lags slightly behind the true evolution because of instrument response. Because the shift is to lower energy, this means that the IPG average energy is slightly higher than the deconvolved energy after the rise of the intensity and slightly lower than the deconvolved energy during the rise. On the basis of literature estimates for the total TDFSS of C153 in solvents with similar spectral positions (diisopropyl ether, toluene),[17] we expect a total TDFSS of ~530 cm^{-1} for C153/βLG$_{465}$. We resolve a total shift of 448 ± 32 cm^{-1}, suggesting that we capture about 85% of the total TDFSS in our experiment. More quantitative methods for estimating the magnitude of a system's TDFSS from steady-state spectra exist,[16,17] but our system does not conform to their requirements.

Because the IPG method explicitly includes system heterogeneity, we can quantitatively observe its change with time. We evaluate the width of the distribution in terms of the entropy, ΔS, calculated by

$$\frac{\Delta S}{k_B} = -\int P(\tilde{\nu},t)\log(P(\tilde{\nu},t))\ d\tilde{\nu} = \frac{1}{2}(\log(2\pi\sigma(t)^2)+1) \quad (6)$$

The integral is evaluated for the IPG fit. The far right-hand side of the equation is the analytical expression for a Gaussian distribution of states.

The width of the distribution as measured by its entropy was more perturbed by the IRF than was the average energy. In regions of the distribution evolution where the mean energy was not rapidly evolving, the effect of instrument response on the entropy was small. However, because the energy was evolving through most of the TCSPC window in our experiment, the effect of the IRF was to significantly increase the width during the response region. This effect persisted after the IRF, because the mean energy continued to shift. This resulted in the coupling of the fit of the entropy to that of the mean energy. The dashed line in the bottom panel of Figure 15 shows the contribution to the entropy in the convoluted data that arises just from the shifting mean energy of the distribution with a constant width, corresponding to the asymptotic limit of the width. There is additional width at time zero, and because it persists longer than the instrument response broadening, additional contributions to the width are needed to explain the evolution of the distribution.

The width of the IPG distribution sharply increased at early times followed by a more moderate decrease at longer times. The rise in entropy at $t = 0$ cannot be resolved by the convoluted fit, which always predicts a larger entropy than that observed at $t < 40$ ps. The entropy rise at $t = 0$ is due to the lack of signal resulting in the IPG algorithm fitting the mean value and not the width. The width of the distribution that appears at $t =$ 40 ps could arise either from a fast entropy increase or due to a fast mean energy decay or some combination thereof. We found that better fits were obtained to the distribution as a whole if the initial entropy evolution decreased rather than increased.

Interpretation of Gaussian Model Fits. The evolution of the deconvolved distribution appears in Figure 16. We were able to extract four time components for the shift of the center and two contributions to the evolution of the width. The fit parameters appear in Table 1. Error estimates were based on a sensitivity analysis of the fit parameters.

The fastest shift ($k_{\tilde{\nu}_1}$, $\tilde{\nu}_1$) with a time scale of 3 ps is required to reproduce the evolution of the mean including the sharp feature at $t = 0$ and is instrument response limited. As discussed above, this feature could represent part of the unresolved TDFSS or a small amount of Raman scattering that is not removed by the Raman basis function. This fast contribution to the TDFSS accounts for 31% of the total energy relaxation. The fast time scale is consistent with inertial motions and other processes that are faster than the larger-scale protein motions that are the focus of this work.

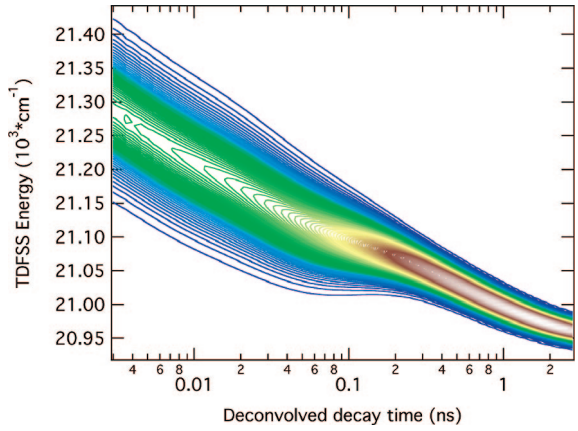

Figure 16. Deconvolved evolution of the TDFSS energy distribution. Because the influence of the IRF has been removed, the early time features are more clearly visible.

TABLE 1: Fit Parameters for Evolution of C153/βLG$_{465}$ Distribution As Defined by Eqs 2–5

rate (ns^{-1})	shift (cm^{-1})	width (cm^{-1})
$k_r = 0.15 \pm 0.02$	$a_0 = 8200 \pm 400$	
	$\tilde{\nu}_0 = 21390 \pm 30$	$\sigma_0 = 52 \pm 8$
$k_{\tilde{\nu}_1} = 350 \pm 150$	$\tilde{\nu}_1 = -140 \pm 30$	
$k_{\tilde{\nu}_2} = 45 \pm 5$	$\tilde{\nu}_2 = -138 \pm 4$	
$k_{\sigma_1} = 10 \pm 1$		$\sigma_1 = -36 \pm 8$
$k_{\tilde{\nu}_3} = 3.4 \pm 0.5$	$\tilde{\nu}_2 = -72 \pm 3$	
$k_{\tilde{\nu}_4} = k_{\sigma_2} = 0.85 \pm 0.15$	$\tilde{\nu}_3 = -98 \pm 8$	$\sigma_2 = -8.2 \pm 0.5$

The mean energy shift observed at 22 ps ($k_{\tilde{\nu}_2}$, $\tilde{\nu}_2$) is in the time regime expected for side chain motions of the protein. This does not appear to allow full relaxation, however, and only accounts for 31% of the total relaxation. Following the energetic portion of the relaxation a decrease in entropy of 9.8 J/(mol K) is observed on a 100 ps time scale. The 290 and 1200 ps decays account for 16% and 22% of the relaxation we resolve. Relaxation occurring at these longer time scales is likely limited by the fluctuations of the protein backbone and local secondary structure allowing additional conformations to be accessed in the calyx. The 1200 ps energy decay is accompanied by a decrease in entropy of 6.0 J/(mol K).

One interpretation of the 100 ps entropy change and the 22 and 290 ps mean energy shift is that there are not discrete rates involved, but rather a distribution of processes active across these time scales. In this case our imposition of discrete exponentials as the fit model would extract values that reproduce the lower moments of these distributions but may not be easily relatable to any particular motion of the system. It would be useful to conduct molecular dynamics simulations to determine the spectrum of fluctuation present in excited-state C153/βLG$_{465}$. Such an investigation is beyond the scope of this paper, however, the observation of Changenet-Barret et al.[9] that linear response theory may not be able to describe the motions of proteins as reflected through the TDFSS of C153 suggests that a comparison of simulations of protein equilibrated around both the ground state and the excited state of C153 would be the minimum necessary to evaluate this phenomenon.

Comparing the evolution of the mean energy and the entropy suggests that the energetics of solvation occur in different phases with a large exothermic energetic component leading to the relaxation followed by a loss of entropy in the later stages. Interpreting the change in entropy in terms of its contribution to the free energy at 37 °C suggests $T\Delta S = 4.9$ kJ/mol of total change as compared to $\Delta E = -5.4$ kJ/mol of negative change in average energy. Note that the distribution of TDFSS energies only reports on the energetics of the C153. The heat that flows out of the dye in the early stages likely contributes to a increase in entropy in the surroundings. Changes in the width reflect changes in the range of solvation energies that modulate the $S_1 \rightarrow S_0$ energy gap. In this representation, a pure solvent basis spectrum would have an entropy of zero. The entropy is reflective of the increase in distribution width in the excited state over that of the pure solvent basis spectrum. The evolution of the entropy is relative to the deconvolved time zero limit of the TDFSS. We discount the possibility of the decrease in variance arising from the intensity decay which has a lifetime of 6.7 ns for the C153/βLG$_{465}$ species.

The electric field of the dye is a driving force for the spatial repositioning of protein charges. Upon excitation, the dipole moment of C153 increases from 6.5 to ~15 D.[17] Because this results in a net increase in the driving force, the net effect should be to increase the range of solvation energies in the excited state. Over time this increased driving force will induce protein motion resulting in a narrower set of conformations, resulting in a decrease in heterogeneity. Therefore at time zero we expect the entropy of the distribution to be at its maximum and this is consistent with our observations. Over time as the TDFSS relaxes the distribution width also narrows. If we interpret the TDFSS distribution in terms of a potential of mean force for the solvation coordinates, then the change in width is analogous to a change in potential force constant. Therefore the evolution of the width parameter suggests a dynamic contribution to a nonlinear response, because linear response would require the force constants to be the same. The evolution of the width implies that as the protein moves in response to the applied electric field, it reduces its conformational freedom. Though the effect is relatively small, it is inherently nonlinear.

4. Summary

We successfully encapsulated C153 inside of the βLG hydrophobic calyx. Red-edge excitation and steady-state temperature-dependent data indicated that C153/βLG has multiple binding environments with different solvatochromic shifts. We developed an IPG method to separate the fluorescence contributions of each binding site and determined there to be two dye locations with access to water and one completely sequestered from it. At least 80% of the dye was encapsulated by the calyx.

When we applied our IPG method to the steady-state temperature-dependent fluorescence spectra, we observed that the populations and spectral properties of the different C153 binding sites changed independently. Their proposed locations in the protein are consistent with their dielectric environment as reported by solvatochromic shifts. Changes in the different binding locations coincided with established transitions in βLG at 45 and 60 °C. We observed a previously unreported thermodynamic transition at 18 °C. We speculated that the EF loop gained conformational freedom at this temperature and allowed transient partial solvation of the calyx bound dye.

The C153/βLG TDFSS evolution was complicated by the presence of three C153 binding environments. The overall mean of the distribution showed a false recovery of TDFSS relaxation energy on the nanosecond time scale at and above 37 °C. This artifact arose from the faster radiative decay of water-exposed (redder) species. The IPG decomposition showed that the encapsulated species C153/βLG$_{465}$ monotonically decreased in energy, whereas the two partially hydrated species maintained a constant energy. We concluded that the fixed positions of the hydrated species essentially result from the completion of solvation within our instrument response. The IPG fitting procedure allowed us to successfully observe and remove the Raman scattering signal from the TDFSS data. This was an important benchmark because it established that the IPG-based spectral decomposition procedure preserves the temporal information in the data. We used the TDFSS of the encapsulated species to determine the characteristic times of fluctuations in βLG.

We reported TDFSS mean energy shift time scales of 3, 22, 290, and 1200 ps for the relaxation of β-lactoglobulin's hydrophobic core responding to the abrupt dipole change in C153. A nonlinear effect was observed in the evolution of the TDFSS distribution; it decreased with time scales of 100 and 1200 ps.

Future work will include a thermodynamic analysis of temperature dependent TDFSS data.

Acknowledgment. We thank Ed Castner for the use of his laser, steady-state absorbance, and steady-state emission instruments as well as for helpful discussions about C153 and TDFSS. We thank Troy Messina for the LabVIEW programming that automated much of our TDFSS data collection as well as helpful conversations. We thank Dave Remeta, Conceicao Minetti, and Ken Breslauer for the use and assistance with their CD spectrometer. This work was supported by a grant from the National Institutes of Health #R01GM071684. J.P. was supported by a DE GAANN grant #P200A03019 to RU CCB.

Supporting Information Available: Textual presentation of optimizing the decomposition basis spectrum and of the hybridmultisolvent fit. Figures of fluorescence spectra, spectral assignments for the dye species, anisotropy, wavelengths vs temperature, and a 3D spectrum model of the transients. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Das, K.; Bauman, J. D.; Clark, A. D.; Frenkel, Y. V.; Lewi, P. J.; Shatkin, A. J.; Hughes, S. H.; Arnold, E. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 1466–1471.
(2) Baum, J.; Dobson, C. M.; Evans, P. A.; Hanley, C. *Biochemistry* **1989**, *28*, 7–13.
(3) Eaton, W. A.; Munoz, V.; Hagen, S. J.; Jas, G. S.; Lapidus, L. J.; Henry, E. R.; Hofrichter, J. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 327–359.
(4) Kuwajima, K. *Proteins: Struct., Funct., Genet.* **1989**, *6*, 87–103.
(5) Tsai, C.-J.; Maizel, Jacob V., J.; Nussinov, R. *Crit. Rev. Biochem. Mol. Biol.* **2002**, *37*, 55–69.
(6) Causgrove, T. P.; Dyer, R. B. *J. Phys. Chem.* **1996**, *100*, 3273–7.
(7) Homoelle, B.; Edington, M.; Diiffey, W.; Beck, W. J. *Phys. Chem. B* **1998**, *102*, 3044–52.
(8) Jordanides, X.; Lang, M.; Song, X.; Fleming, G. J. *Phys. Chem. B* **1999**, *103*, 7995–8005.
(9) Changenet-Barret, P.; Choma, C.; Gooding, E.; De- Grado, W.; Hochstrasser, R. *J. Phys. Chem. B* **2000**, *104*, 9322–9.
(10) Pal, S.; Peon, J.; Zewail, A. *Proc. Natl. Acad. Sci.* **2002**, *99*, 1763–8.
(11) Sen, P.; Mukherjee, S.; Dutta, P.; Hadler, A.; Mandal, D.; Banerjee, R.; Roy, S.; Bhattacharyya, K. *J. Phys. Chem. B* **2003**, *107*, 14563–8.
(12) Abbyad, P.; Shi, X.; Childs, W.; McAnaney, T. B.; Cohen, B. E.; Boxer, S. G. *J. Phys. Chem. B* **2007**, *111*, 8269–8276.
(13) Messina, T. C.; Talaga, D. S. *Biophys. J.* **2007**, *93*, 579–585.
(14) Cave, R.; Castner, E. *J. Phys. Chem. A* **2002**, *106*, 12117–23.

(15) Chapman, C. F.; Fee, R. S.; Maroncelli, M. *J. Phys. Chem.* **1995**, *99*, 4811–19.
(16) Fee, R.; Maroncelli, M. *Chem. Phys.* **1994**, *183*, 235–47.
(17) Horng, M.; Gardecki, J.; Papazyan, A.; Maroncelli, M. *J. Phys. Chem.* **1995**, *99*, 17311–37.
(18) Larsen, D. S. *J. Chem. Phys.* **2004**, *121*, 5039–42.
(19) Lewis, J.; Maroncelli, M. *Chem. Phys. Lett.* **1998**, *282*, 197–203.
(20) Matyushov, D. V.; Newton, M. D. *J. Phys. Chem. A* **2001**, *105*, 8516–8532.
(21) Mühlphordt, A.; Schanz, R.; Ernsting, N.; Fartzdinov, V.; Grimme, S. *Phys. Chem. Chem. Phys.* **1999**, *1*, 3209–18.
(22) Shirota, H.; Castner, Edward W. J. *J. Chem. Phys.* **2000**, *112*, 2367–2376.
(23) Chowdhury, P.; Halder, M.; Sanders, L.; Arnold, R.; Liu, Y.; Armstrong, D.; Kundu, S.; Hargrove, M.; Song, X.; Petrich, J. *Photochem. Photobiol.* **2004**, *79*, 440–6.
(24) Pryor, B.; Palmer, P.; Chen, Y.; Topp, M. *Chem. Phys. Lett.* **1999**, *299*, 536–44.
(25) Kella, N. K. D.; Kinsella, J. E. *Biochem. J.* **1988**, *255*, 113–8.
(26) Kuwata, K.; Hoshino, M.; Forge, V.; Era, S.; Batt, C.; Goto, Y. *Protein Sci.* **1999**, *8*, 2541.
(27) Oliveira, K.; Valente-Mesquita, V.; Botelho, M.; Sawyer, L.; Ferreira, S.; Polikarpov, I. *Eur. J. Biochem.* **2001**, *268*, 477–483.
(28) Ragona, L.; Fogolari, F.; Catalano, M.; Ugolini, R.; Zetta, L.; Molinari, H. *J. Biochem.* **2003**, *278*, 38840–6.
(29) Uhrinova, S.; Smith, M.; Jameson, G.; Uhrin, D.; Sawyer, L.; Barlow, P. *Biochemistry* **2000**, *39*, 3565.
(30) Tanford, C.; Bunville, L.; Nozaki, Y. *J. Am. Chem. Soc.* **1959**, *81*, 4032.
(31) Banik, U.; Saha, R.; Mandal, N. C.; Bhattacharyya, B.; Roy, S. *Eur. J. Biochem.* **1992**, *206*, 15–21.
(32) Thomas, P. D.; Dill, K. A. *Protein Sci.* **1993**, *2*, 2050–65.
(33) Ionescu, R. M.; Matthews, C. R. *Nat. Struct. Biol.* **1999**, *6*, 304–307.
(34) Dalgalarrondo, M.; Bertrand-Harb, C.; Chobert, J. M.; Dufour, E.; Haertle, T. *ACS Symp. Ser.* **1991**, *470*, 86–96.
(35) Qi, X.; C.C.Holt.; McNulty, D.; Brownlow, D. C. S.; Jones, G. *Biochem. J.* **1997**, *324*, 341–6.
(36) Tanford, C.; Taggart, V. J. *Am. Chem. Soc.* **1961**, *83*, 1634–8.
(37) Fessas, D.; Iametti, S.; Schiraldi, A.; Bonomi, F. *Eur. J. Biochem.* **2001**, *268*, 5439–48.
(38) Maglia, G.; Jonckheer, A.; De Maeyer, M.; Frere, J.-M.; Engelborghs, Y. *Protein Sci.* **2008**, *17*, 352–361.
(39) Vivian, J. T.; Callis, P. R. *Biophys. J.* **2001**, *80*, 2093–2109.
(40) Lange, D.; Kothari, R.; Patel, R.; Patel, S. *Biophys. Chem.* **1998**, *74*, 45–51.
(41) Wu, S.; Prez, M.; Puyol, P.; Sawyer, L. *J. Biol. Chem.* **1999**, *274*, 170–4.
(42) Wang, Q.; Allen, J.; Swaisgood, H. *J. Dairy Sci.* **1997**, *80*, 1054–9.
(43) Kontopidis, G.; Holt, C.; Sawyer, L. *J. Dairy Sci.* **2004**, *87*, 785–96.
(44) Collini, M.; DAlfonso, L.; Baldini, G. *Protein Sci.* **2000**, *74*, 1968–74.
(45) Giurleo, J. T.; He, X.; Talaga, D. S. *Protein Sci.*, submitted for publication.
(46) Eberini, I.; Fantucci, P.; Rocco, A. G.; Gianazza, E.; Galluccio, L.; Maggioni, D.; Dal Ben, I.; Galliano, M.; Mazzitello, R.; Gaiji, N.; Beringhelli, T. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 555–567.
(47) Monaco, H. L.; Zanotti, G.; Spadon, P.; Bolognesi, M.; Sawyer, L.; Eliopoulos, E. E. *J. Mol. Biol.* **1987**, *197*, 695–706.
(48) Demchenko, A. *Biophys. Chem.* **1982**, *15*, 101–9.
(49) Demchenko, A. *Eur. Biophys. J.* **1988**, *16*, 121–9.
(50) Demchenko, A. *Luminescence* **2001**, *17*, 19–42.
(51) Giurleo, J. T.; Talaga, D. S. *J. Chem. Phys.* **2008**, 28.
(52) Lawson, C. L.; Hanson, R. J. *Solving Least Squares Problems*; Prentice-Hall, Inc.: Englewood Cliffs, NJ, 1974.
(53) Merritt, M.; Zhang, Y. *J. Optimization Theory Appl.* **2005**, *126*, 191–202.
(54) Lefevre, T.; Subirade, M. *Biochim. Biophys. Acta* **2001**, *1549*, 37–50.
(55) See Supporting Information.
(56) Weast, R. C., Ed. *Handbook of Chemistry and Physics*; CRC Press, Inc.: Boca Raton, FL, 1989−90.
(57) Brackman, U. *Lambdachrome Laser Dyes*; Lambda Physik GmbH: Gttingen, Germany, 1994.
(58) Zhong, D.; Douhal, A.; Zewail, A. H. Proc. *Natl. Acad. Sci. U.S.A.* **2000**, *97*, 14056–14061.
(59) Bucci, E.; Malak, H.; Fronticelli, C.; Gryczynski, I.; Lakowicz, J. R. *J. Biol. Chem.* **1988**, *263*, 6972–7.
(60) Lakowicz, J. R.; Gryczynski, I.; Szmacinski, H.; Cherek, H.; Joshi, N. *Eur. Biophys. J.* **1991**, *19*, 125–40.
(61) Lakowicz, J. *Principles of Fluorescence Spectroscopy*, 2nd ed.; Plenum Publishers: New York, 1999.
(62) de Juan, A.; Tauler, R. *Anal. Chim. Acta* **2003**, *500*, 195–210.
(63) Henrion, R.; Henrion, G.; Bohme, M.; Behrendt, H. *Anal. Bioanal. Chem.* **1997**, *357*, 522–6.
(64) Jhala, E.; Galilee, C.; Reinisch, L. *Appl. Opt.* **2007**, *46*, 5522–5528.
(65) Sakurai, K.; Goto, Y. *Proc. Natl. Acad. Sci.* **2007**, *104*, 15346–15351.
(66) Bursing, H.; Ouw, D.; Kundu, S.; Vohringer, P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 2378–2387.

JP802666N

Separation and Analysis of Dynamic Stokes Shift with Multiple Fluorescence Environments: Coumarin 153 in Bovine β-Lactoglobulin A: Supplemental Materials

Jeremy Pronchik, Jason T. Giurleo, David S. Talaga

June 2, 2008

1 Optimizing the Decomposition Basis Spectrum

Finding an appropriate basis function to describe data can be a challenge, but is very important. As mentioned, we base the goodness of a basis function first by the quality of the fits and second by the least number of parameters required to describe the data. From our experience, the shape and width of the basis are the relevant features requiring attention.

Shape: The data spectra are composed of three overlapping subspectra separated by $\sim$1500 cm^{-1}. As can be seen in Fig. 10, the fluorescence information above 22000 cm^{-1} can be attributed to C153/βLG$_{465}$ alone and this provides us with a good starting point for matching spectral features.Fig. 2 shows the 0°C C153/βLG spectrum in black along with an array of the C153/dioxane basis spectra. In the region labeled 'B,' the similarity in the blue edge slopes of the data and basis spectra can be seen; we attribute the quality of observed fitting to this.

The top panel of Fig. 8 shows the three basis functions discussed in this paper.The difference in blue-edge slopes between the ethanol and dioxane bases can be seen here, indicating the sensitivity of fitting on the basis function's shape. The blue edge slope of the DMSO basis is also very similar to the C153/βLG data and perhaps even better suited as a descriptor of the tail as seen in the residuals of Fig. 9.In principle, a tailored basis function that copied the most successful fitting features of each basis would be the ideal choice, but this level of fine tuning was not necessary for our present work.

Width: The DMSO basis suffers from the narrowness of its width relative to that of the C153/βLG subspectra. Even with no inhomogeneous broadening of each subspecies, a distribution of basis spectra would be required for fitting (Fig. 8).Similar results were found using C153/hexanes as a basis, which is also narrow.

The extreme of this problem was explored by using narrow (σ=500 cm^{-1} and σ=1000 cm^{-1}) Gaussian distributions as our basis. The results were a series of evenly spaced parameter distributions which fit the data well, but had no physically interpretive value. Looking again at the DMSO fitting results in the bottom panel of Fig. 8, we can see three sets of parallel distributions that recall this artifact of having too narrow of a basis function.

2 Hybrid multisolvent fit

It is reasonable to expect that if different molecular solvents yield slightly different spectral shapes then different binding locations also would. The failure of the ethanol basis was chiefly due to its inability to fit the blue region of the spectrum. The dioxane basis fit well, but showed broad distributions at high temperatures. For this reason, we performed fits using a design matrix created from multiple basis spectra to determine if the steady-state temperature analysis could be improved. We chose the dioxane and ethanol bases for constructing the hybrid design matrix, since their spectral positions were most similar to the C153/βLG species. Complicating the design matrix in this fashion did not significantly improve the quality of the fits, however, we did learn about the interdependence of subpopulations in fitting.

Using a design matrix which describes the C153/βLG$_{465}$, C153/βLG$_{500}$, and C153/βLG$_{550}$ regions with dioxane, ethanol, and ethanol respectively (Fig. S1 D-E-E) shows residuals to be slightly worse in the area about C153/βLG$_{465}$ and C153/βLG$_{500}$ and slightly better for C153/βLG$_{550}$. Visually, we still conclude on the same transitions in the C153/βLG$_{465}$ and C153/βLG$_{550}$ distributions, though small qualitative differences can be seen. The C153/βLG$_{500}$ distribution became much narrower with ethanol fitting and the 18°C transition was no longer visible, indicating the loss in fitting precision was too great in that area.

Using the opposite scheme (Fig. S1 E-D-D) showed the effect one fit had on the others more plainly. Despite being fit by dioxane, the C153/βLG$_{500}$ distribution looked nearly identical to that of the purely ethanol design matrix. C153/βLG$_{550}$ was similar to its pure dioxane evolution, though its transition shifted about 5°C cooler. This showed, as one might expect, that the magnitude of the poorly fit species determines the extent of its influence over its neighbors or, roughly speaking, mistakes flow downhill. While we were fortunate to have a

single basis describe all three species well, we believe it to be important to clean up each species' fit before analyzing the results.

We fit using a D-D-E design matrix after seeing improved fitting in the region of C153/βLG$_{550}$. The residuals were slightly smaller, however, the temperature-wise evolution of the distribution was virtually identical to pure dioxane. Since there was essentially no difference in the distributions, we interpret the temperature-dependence of C153/βLG using the fits from the C153/dioxane basis set.

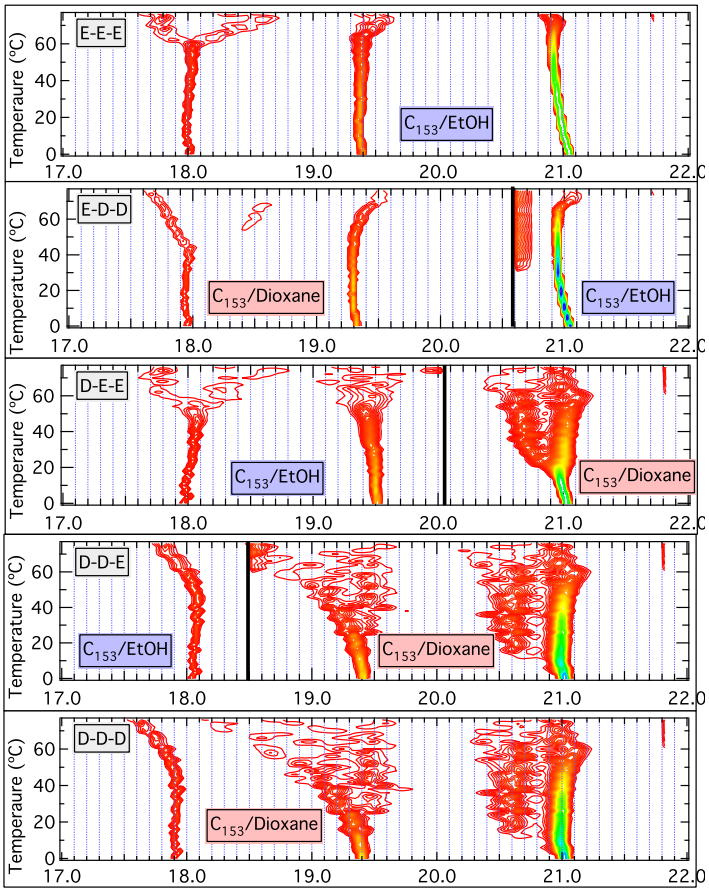

Figure S 1: Deconstruction of the steady-state fluorescence spectra of C153/BLG from 0-76^{0}C. Hybrid design matrices made of segments of C153/dioxane (D) and C153/ethanol (E) basis functions along with the straight C153/dioxane and C153/ethanol results for comparison. Mistakes in fitting the majority, C153/B-LG$_{465}$, species tend to affect shape of the minority species, but less so in the opposite direction. The E-D-D and D-E-E fits were both worse than with the straight D-D-D design matrix and deemed less reliable. The D-D-E design matrix was able to fit the red edge of the spectrum (C153/B-LG$_{550}$) better than D-D-D, but the physical interpretation of transitions was the same.

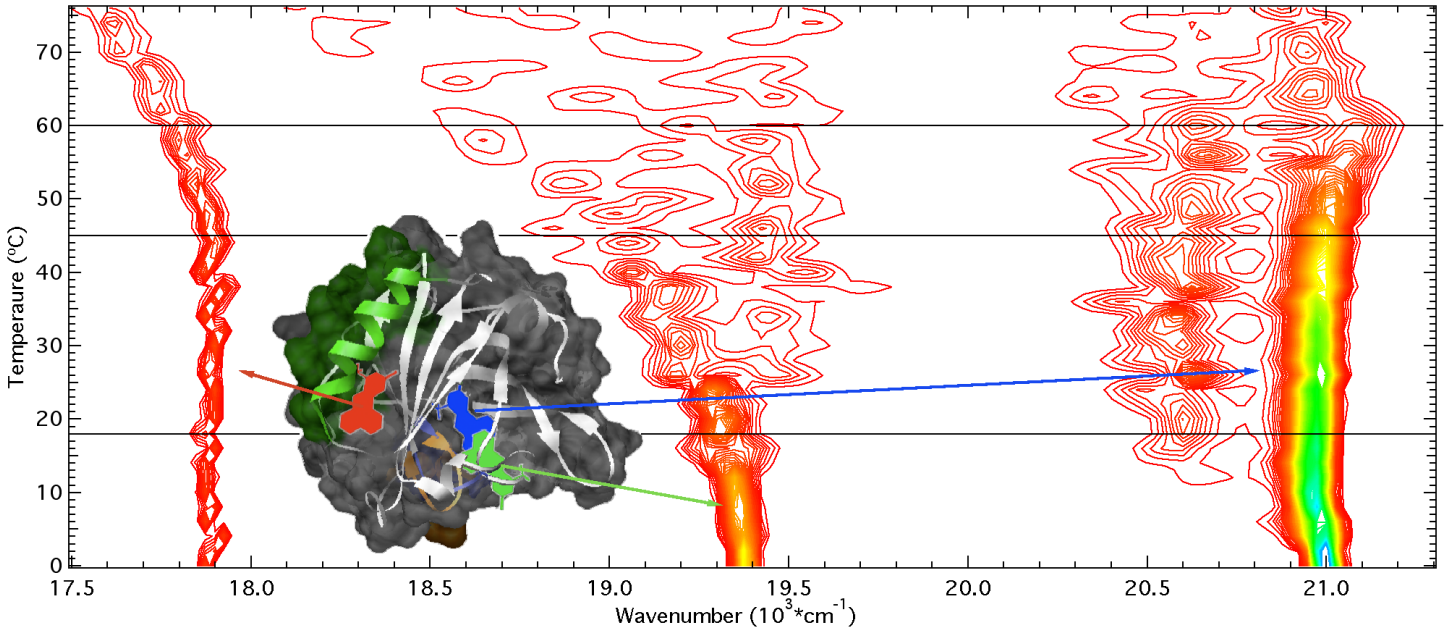

Figure S 2: Our assignments for the locations of the three dye species we have present in our sample.

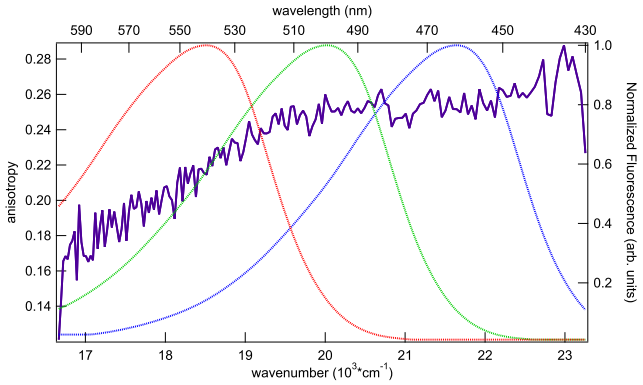

Figure S 4: Steady-state anisotropy for a typical sample at 4^0C. Wavelengths longer than ∼520nm show a decrease in anisotropy which is likely due to the mobility of the C153/BLG_{550} species (red), which influences the fluorescence more at longer wavelengths.

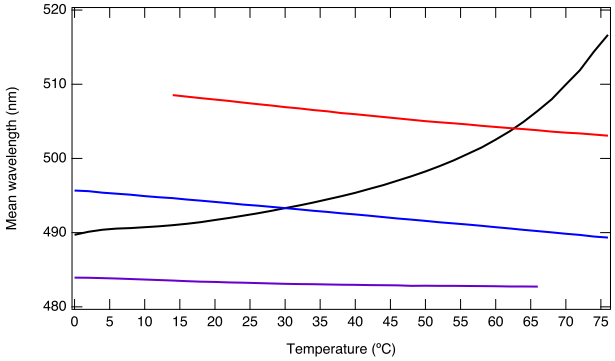

Figure S 5: Mean wavelengths of C153 fluorescence spectra in various environments. The mean of C153/BLG (black) red-shifts with increased temperature contrary to the means of C153 in the molecular solvents dioxane (red), ethanol (blue), and hexanes (purple). Low temperature data for C153/dioxane is excluded; dioxane has a freezing point of 12^0C. High temperature data for C153/hexanes is excluded; the sample evaporated and noisy spectra resulted in a stochastic mean.

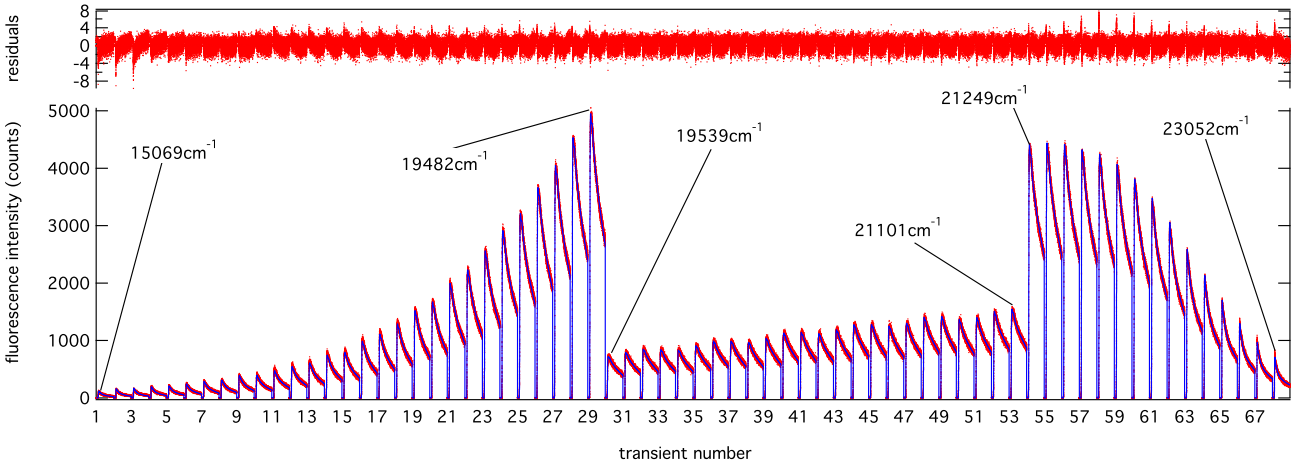

Figure S 3: A 3 spectrum model was used to directly fit the 68 TCSPC transients (fit in blue, raw data and residuals in red). The model used three C153/dioxane fluorescence spectra at different spectral shifts each with its own fluorescence rate. The bluest shifted spectrum was allowed to shift in time according to a multiexponential function. The raw data was taken in three segments having different monochrometer slit widths and for different lengths of time. Fits were good enough to extract normalization values which we used to recreate the time-evolving total spectrum from the TCSPC transients. They were not good enough to capture the time-dependent Stokes shift, especially the fast time components.

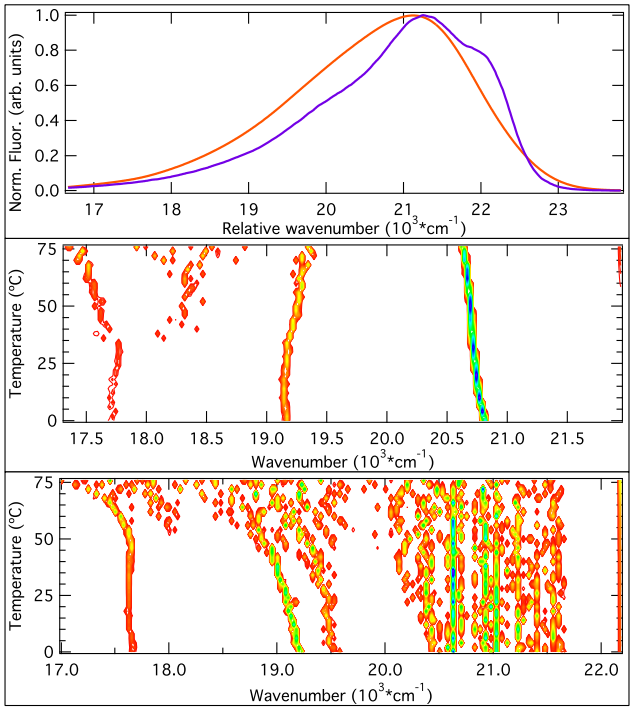

Figure S 6: Deconstruction of the steady-state fluorescence spectra of C153/BLG from 0-76^{0}C. Basis spectra of C153/acetonitrile (orange, middle) and C153/hexanes (purple, bottom) are used. C153/hexanes are similar to C153/DMSO in that both basis spectra are narrow and produce wide distributions. C153/acetonitrile results show a similar mismatch in spectral shape to C153/ethanol and the corresponding systematic error in the C153/BLG$_{465}$ distribution. These results use an active set method not discussed in this paper. It is a similar, non-negative least squares procedure which is faster, but has difficulty with distributions. We found general agreement between the two methods.