Dmitri   Toptygin
8727 Hayshed Ln # 23
Columbia MD 21045
410–884–7068 (home)
410–516–7300 (daytime)
toptygin@jhu.edu

December 19, 2003

Biocomplexity Faculty Search Committee,
c/o Prof. Rob de Ruyter van Steveninck,
Biocomplexity Institute,
Indiana University,
Swain Hall West 117,
Bloomington, IN 47405–7105

Dear Chair of the Search Committee:

I wish to apply for a faculty position created in support of the new interdisciplinary Biocomplexity initiative in the College of Arts and Sciences at Indiana University Bloomington.

I have considerable experience in experimental biophysics, specifically, in experimental molecular dynamics, time–resolved fluorescence spectroscopy, biomolecular measurements and instrumentation. My university training was in the areas of optics, atomic and molecular spectroscopy, quantum radiophysics, and laser technology. Since 1994 I have been working at the Johns Hopkins University, almost exclusively in the areas of biophysics and biochemistry, and coauthored 15 publications (not including conference abstracts) in these areas. I have studied spectrally– and time–resolved fluorescence of Trp, Tyr, and Phe residues in proteins, of rod–like fluorescent probes in model lipid bilayers, of Ca^{2+}–sensitive and pH–sensitive fluorescent dyes in live cells, etc.. I have invented and experimentally tested a new optical technique for assessing orientational order in lipid bilayers. I was among the first scientists who used tryptophan fluorescence to measure protein relaxation dynamics (vibration along soft normal modes). I have also discovered experimentally and explained theoretically the effect of the solvent refractive index on the excited–state lifetime of buried tryptophan residues in proteins.

My curriculum vitae, bibliography, and a brief statement of research interests are enclosed for your review. I have also enclosed a 15–page research proposal as evidence of a first rate research program with impact on molecular dynamics, biomolecular measurements and instrumentation. The names of five references are listed on page 4 of my CV; I have asked three of them to send recommendation letters to you. I am looking forward to hearing from you soon.

Sincerely,

Dmitri Toptygin.

BRIEF STATEMENT OF RESEARCH INTERESTS

DMITRI TOPTYGIN

1. EXPERIMENTAL DYNAMICS OF PROTEIN SOFT VIBRATION MODES.

Normal vibration modes of protein molecules cover a wide frequency range. Infrared and Raman spectroscopies are commonly used to study high–frequency vibrations, which are localized at small chemical groups. Collective motions of domains and secondary–structure elements correspond to the lowest–frequency (soft) vibration modes ($v < 1 cm^{-1}$), and these are the motions that are believed to be directly related the function of enzymes and other proteins. Solvent viscosity damping results in aperiodic relaxation dynamics of protein soft vibrations; the large number of soft vibratoin modes leads to multiexponential dynamics with a wide range of relaxation times. Experimentally–determined relaxation times are directly related to the molecular weight of the moving protein fragments. For protein domains and secondary–structure elements weighing between 1 and 100 kD the relaxation times are in the nanosecond range. Motions of smaller fragments (the size of one aminoacid sidechain) take place on the picosecond time scale. Motions of very small molecules (H_2O) near the surface of the protein take place on the femtosecond time scale. All these relaxation times can be measured using time–resolved fluorescence spectroscopy. I was among the first scientists who made the use of tryptophan fluorescence to study picosecond and nanosecond motions in proteins [*J. Phys. Chem. B* **2001**, *105*, 2043–2055]. One year later A. H. Zewail, a Nobel prize winner, used tryptophan fluorescence in a study of femtosecond motions of water molecules near a protein surface [*J. Phys. Chem. B* **2002**, *106*, 12376–12395]. The potential of time–resolved fluorescence measurements in protein dynamics studies has not been fully utilized yet. I plan to develop new applications using time–resolved fluorescence of tryptophan and other fluorophores to study protein dynamics.

2. ROLE OF NUCLEAR DELOCALIZATION IN PROTEIN FOLDING AND PROTEIN–PROTEIN INTERACTIONS.

In 1954 Gordon, Zieger and Townes built a working NH_3 maser and explained the principle of its operation. According to their theory, the nucleus of every atom in NH_3 is located in two places at the same time (this is true for both states between which the radiative transition takes place). If every nucleus in NH_3 was located in only one place at a time, then it would not be possible to explain the principle of operation of the maser; yet, the maser works. Delocalization of nuclei is common for molecules with Born–Oppenheimer–inseparable wavefunctions. Many Nitrogen–containing molecules fall into this category: NH_3, guanidine hydrochloride, urea, etc. It is very likely that some polypeptides and proteins also belong to the category of molecules with delocalized nuclei. I plan to conduct experimental and theoretical research aimed at understanding the role of nuclear delocalization in protein folding and protein–protein interactions.

DMITRI   TOPTYGIN

This Research Proposal is included as evidence of a first rate research program with impact on molecular dynamics, biomolecular measurements and instrumentation; it corresponds to just one of the two main research interests outlined in the Brief Statement of Research Interests.

Experimental Dynamics of Protein Soft Vibration Modes

1. Background.

1.1. Dielectric relaxation of bulk dielectrics, including solvent relaxation.

Relaxation of bulk dielectrics is commonly studied by electrical and optical methods [1]. The electrical methods measure the characteristics of an electrical capacitor, in which the gap between the plates is filled with the dielectric under study [2]. The optical methods utilize polar fluorescent molecules with a large difference $\Delta\vec{\mu}$ between the permanent dipole moment in the excited state and in the ground state. The instantaneous spectrum of fluorescence emission from a solution of such molecules in a polar solvent varies with time after impulse excitation. The emission spectrum is recorded as a function of the wavenumber, ν. The response function $C_\nu(t)$ represents the dynamics of dielectric relaxation and also the dynamics of solvation [1,3–5]:

$$C_\nu(t)=\frac{\nu_F(t)-\nu_F(\infty)}{\nu_F(0)-\nu_F(\infty)} \tag{1}$$

In Eq.(1) $\nu_F(t)$ represents the time variation of either the peak wavenumber or the center–of–gravity wavenumber of the instantaneous fluorescence emission spectrum.

Time–dependent spectral shifts in the fluorescence emission of polar fluorescent solutes in polar solvents were discovered and attributed to solvent relaxation by Bakhshiev and coworkers in the 1960's [6–8]. Rotational diffusion of solvent molecules is the main contributor to solvent relaxation, however, translational diffusion is also believed to play some role [1]. With the exception of extremely viscous liquids, such as pure glycerol, the solvent relaxation is very fast. For example, relaxation of liquid water takes less than one picosecond [9–11].

1.2. Microscopic dielectric relaxation in a protein matrix.

Microscopic dielectric relaxation in a protein matrix can be studied by incorporating a fluorophore with a large $\Delta\vec{\mu}$ in a protein molecule and recording time–dependent spectral shifts in its fluorescence emission after a short excitation pulse. The mass of a typical protein molecule exceeds that of a water molecule by several orders of magnitude, therefore protein relaxation is several orders of magnitude slower than the relaxation of its solvent. The slowest relaxation modes of a small globular protein may correspond to relaxation times from hundreds of picoseconds to many nanoseconds. Experimental evidence for a nanosecond relaxation process in a protein was reported by Brand and Gohlke in 1971 [12]. Since that report, several research groups published observations of slow relaxation processes in proteins [13–25]. These publications include measurements with femtosecond time resolution, in which sub–picosecond relaxation of the solvent (water) was observed, but in addition to this, the authors also found slow relaxation components that they attributed to protein relaxation [20,21].

1.3. The use of tryptophan fluorescence to study protein relaxation.

The sidechain of a Trp residue represents a polar fluorophore with a large difference $\Delta\bar{\mu}$ between the permanent dipole moment in the excited state and the ground state [26,27]. The presence of Trp residues in proteins reduces the need for unnatural fluorophores to study microscopic dielectric relaxation. The first observation of a relaxation process in a protein using time–resolved fluorescence emission from Trp residues was reported by Grinvald and Steinberg in 1974 [13]. An alternative explanation of time–variant Trp emission spectra attributed the spectral variation to the presence of two or more distinct Trp populations with different emission spectra and different excited–state lifetimes [28]. A heterogeneous ensemble of fluorescence emitters is known to produce an impression of an apparent relaxation process [29], therefore it is extremely important to understand possible causes of heterogeneity in the local environments of Trp residues or other fluorophores that may be used to study protein relaxation. In the case of Trp fluorophores, the main causes of heterogeneity are: (i) the presence of more than one Trp in the aminoacid sequence of the protein [28], (ii) the presence of more than one global conformation of the protein (this is a common situation for unfolded and aggregated proteins), and (iii) the presence of more than one rotational conformation of the Trp sidechain [30,31]. Note, that the latter may produce a false impression of an apparent relaxation process only if the characteristic time of interconversion between different rotational conformations is much longer than the lifetime of the Trp excited state. If the interconversion time is shorter than the excited–state lifetime, then this interconversion represents a real relaxation process that is correctly represented by the time variation of Trp fluorescence emission spectrum.

Since long–lived heterogeneity in the environment of Trp residues (or any other fluorophores for that matter) may result in artificial relaxation dynamics, the use of time–resolved fluorescence to study protein relaxation dynamics is contingent upon our ability to tell the difference between homogeneous and heterogeneous ensembles of fluorescence emitters. Two criteria for long–lived heterogeneity have been proposed and experimentally tested by the authors of this proposal. This is described in preliminary results, see sections 2.1, 2.2 and 2.3.

2. Preliminary Results.

2.1. Corrected instantaneous emission spectra of indole and tryptophan.

Indole represents a fluorophore with a large $\Delta\bar{\mu}$ and it occurs naturally in proteins as the side chain of Trp residues. It is well known that the instantaneous emission spectrum of indole or Trp usually shifts to the red (to the lower energies) with time after a short excitation pulse. In our preliminary studies we asked the questions: what happens to the shape of the instantaneous emission spectrum as it shifts to the red [32] and whether the information contained in the shape can be used to tell the difference between homogeneous and heterogeneous ensembles of fluorescence emitters [25]. A reliable mathematical method was developed for calculating deconvoluted (from the instrument response function) and corrected (for the spectral variation of the instrument sensitivity) instantaneous emission spectra; this method is described briefly in ref. [32] and in greater detail in ref. [25]. Examples of deconvoluted and corrected instantaneous emission spectra are shown in Figs. 1–4.
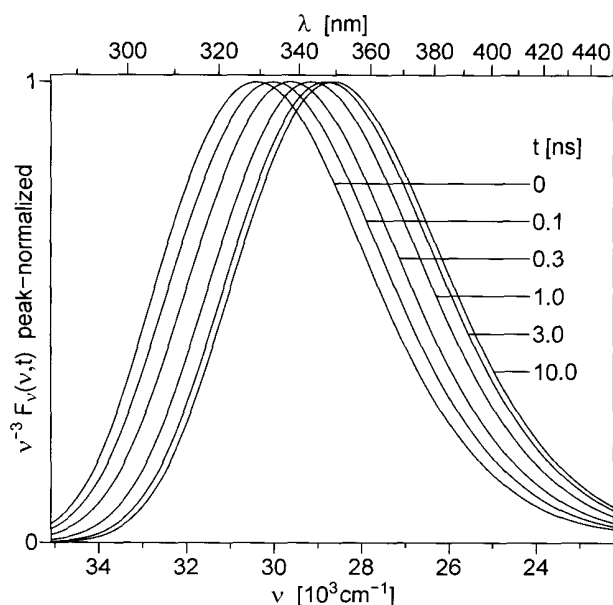
Figure 1. *Instantaneous emission spectra of indole in anhydrous glycerol at +20°C, shown on a linear wavenumber scale.*
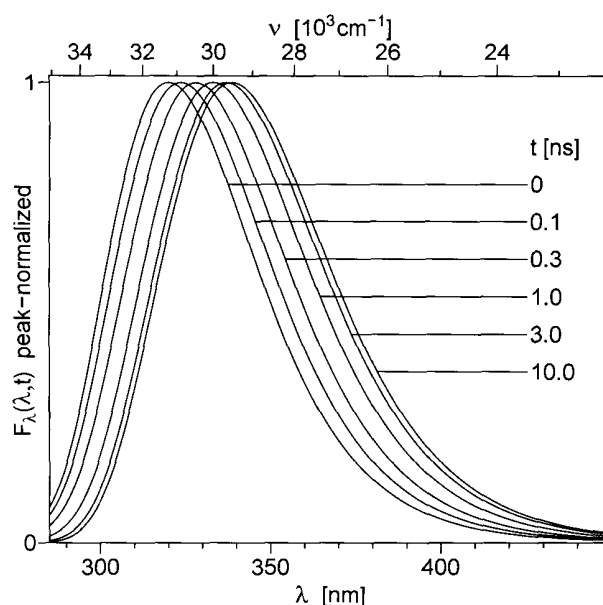

Figure 2. *Instantaneous emission spectra of indole in anhydrous glycerol at +20°C, shown on a linear wavelength scale.*
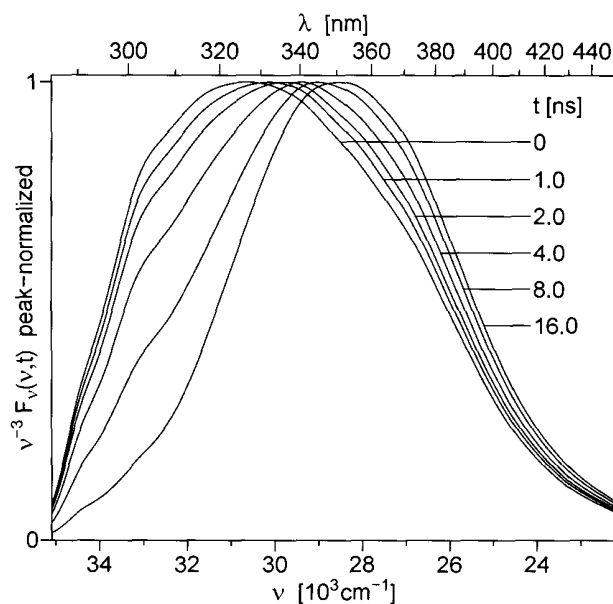

Figure 3. *Instantaneous emission spectra of indole in a heterogeneous, but not relaxing environment.*

The data for Figs. 1 and 2 were obtained with a homogeneous solution of indole in a slow–relaxing polar solvent (anhydrous glycerol). In Fig. 1 the spectra are plotted on the linear wavenumber (ν) scale. Note, that all the spectra in Fig. 1 have the same shape and can be obtained from one another by parallel translation. In Fig. 2 the same data are shown on a linear wavelength (λ) scale, and on this scale the shape is not conserved: the spectra become wider with time. This happens because the energy gap between the excited and ground electronic states is linearly related to the wavenumber, but not to the wavelength [32]. It was also shown, that for the best conservation of the spectral shapes, the emission spectra must be multiplied by v^{-3}, as in Fig.1 [32].

The data for Fig. 3 were obtained using the following model system: fluorescence signals were combined from two cuvettes, one containing a solution of indole in ethyl ether, and the other containing a solution of indole in water [25]. Relaxation of both ethyl ether and water is much faster than the time resolution of the instrument used to collect the data, therefore, the

variation in the shape of the instantaneous emission spectra in Fig. 3 is entirely due to the heterogeneity artifact, and not due to relaxation. It is not difficult to see that the spectra in Fig. 3 become narrower at long times after the excitation pulse. This happens because at early times the emission from both environments is combined, which always makes the spectrum wider; at later times only the emission from the environment where the lifetime is the longest survives.

2.2. Criteria for long–lived heterogeneity.

If the instantaneous emission spectra are represented in the coordinates ν and $\nu^{-3}F_\nu(\nu,t)$, then the narrowing of the emission spectrum at late times after excitation as in Fig. 3 represents evidence of long–lived heterogeneity. This gives one of the two proposed criteria for long–lived heterogeneity [25]. The second criterion is based on the so–called "red–edge excitation effect", for a recent review see [33]. Both excitation and emission spectra of Trp residues are sensitive to their environment. The excitation spectra of Trp residues with red–shifted emission spectra have a red tail in the region of 295nm to 305nm. The excitation spectra of Trp residues with blue–shifted emission spectra are lacking the tail, but they usually have a tall peak near 290nm. Thus, 290nm excitation photoselects Trp residues with blue–shifted emission spectra, whereas 300nm excitation photoselects Trp residues with red–shifted emission spectra. If Trp residues exist in more than one local environment, then the emission spectrum shifts significantly as the exciting wavelength is varied between 290nm and 300nm, unless, of course, the heterogeneity is short–lived, in which case Trp residues "forget" their environment between the instance of excitation and the instance of emission. This gives the second proposed criterion for long–lived heterogeneity. Sensitivity to small spectral variations was enhanced by the use of Singular Value Decomposition, which was applied to study Trp red–edge excitation effect [25].

2.3. Homogeneous Trp fluorescence in proteins.

The above–described criteria for long–lived heterogeneity were applied to several single–Trp proteins. In some cases long–lived heterogeneity was found, however, there was usually a good reason for this: the proteins were either aggregating or thermodynamically unstable. On the other hand, when thermo-dynamically stable small globular single–Trp proteins (E21W and F3W mutants of IIA^{Glc} protein of *Escherichia Coli*) were examined, no signs of long–lived heterogeneity were detected by either criterion. For example, instantaneous emission spectra for the E21W mutant form (see Fig. 4) do not get narrower with time after excitation. In fact, not only the width, but also the shape of the spectra in Fig.4 is conserved except for the parallel translation along the horizontal axis. The same is true for
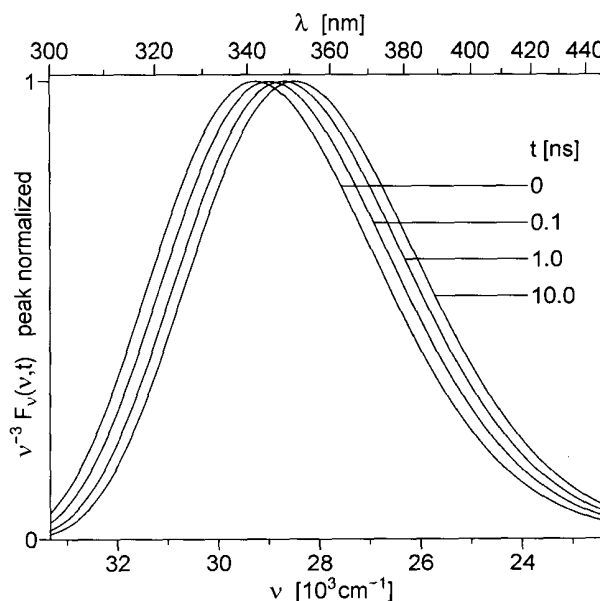

Figure 4. Instantaneous emission spectra of the only Trp residue in E21W mutant of IIA^{Glc} protein of Escherichia Coli. Data from [25].

the homogeneous model system, see Fig. 1, and it is also true for the F3W mutant form (data not shown). The other criterion also revealed no long−lived heterogeneity: there was no variation of the steady−state fluorescence emission spectrum with the exciting wavelength for both single− Trp proteins.

2.4. Protein relaxation on the nanosecond time scale: collective motions and local motions.

The shift of the instantaneous emission spectrum along the wavenumber scale in Fig. 4 represents the relaxation dynamics of a system that consists of the protein and the solvent. The relaxation dynamics is best represented by the time variation of the center of gravity, $\bar{\nu}$, and not by the peak wavenumber, ν_{peak} , which is more affected by the noise in the data. A mathematical definition of $\bar{\nu}$ can be found in refs. [25,32]. Figs. 5, 6 depict the relaxation curves $\bar{\nu}(t)$ for two single−Trp proteins. The choice of the logarithmic time scale in Figs. 5,6 is dictated by the wide range of relaxation rates; on a linear time scale it would not be possible to show what happens between 0.1ns and 0.2ns and also what happens between 10ns and 20ns. The lower and upper limits of the time scale are determined by the time resolution limit of the instrument (0.065ns FWHM) and by the duration of Trp fluorescence (after 20ns the remaining Trp fluorescence intensity is too low to measure the shape of the emission spectrum).
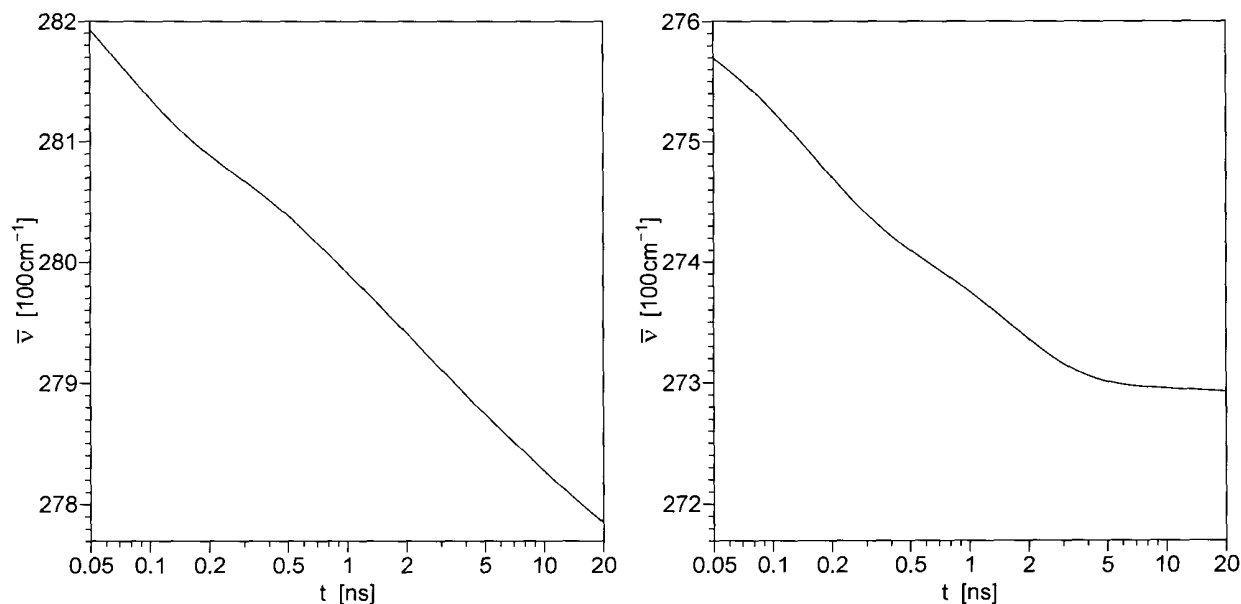

Figure 5. *Relaxation dynamics reported by Trp residue 21 (in the rigid globule) of the E21W mutant of IIA^{Glc} protein from Escherichia Coli.*

Figure 6. *Relaxation dynamics reported by Trp residue 3 (in the flexible tail) of the F3W mutant of IIA^{Glc} protein from Escherichia Coli.*

Figs. 5, 6 depict the relaxation dynamics reported by Trp residues in two different locations within the same protein. The solvent (water) contained no glycerol or other cosolvents of high viscosity. Relaxation of water takes less than 1ps [9−11], therefore, within the time window of Figs. 5, 6 the relaxation of the solvent contributes nothing to $\bar{\nu}(t)$, and the curves represent the dynamics of internal motions in the protein itself. Among these are the motions of the sidechains of individual residues, which usually take place on the time scale of about 100ps

or faster, and the motions of the secondary–structure elements, which may take place on the time scale from a few hundred picoseconds to many nanoseconds. The most remarkable difference between the relaxation curves in Figs. 5 and 6 corresponds to the time window between 3ns and 20ns. The following paragraph explains this difference in terms of the protein structure.

From the X–ray crystal structure [34] and from the NMR structure [35] of the wild type IIA^{Glc} protein of *Escherichia Coli* it is known that residues 19 through 168 form a rigid globule, whereas residues 1 through 18 form a flexible tail, which is invisible both by X–ray and by NMR methods. We have demonstrated that the same is true for the two mutant proteins [25], therefore in the E21W protein the only fluorescent residue is in the rigid globule, whereas in the F3W protein it is in the flexible tail. The fluorescent residue (W21) in E21W mutant protein is a part of a secondary–structure element (β–sheet I), and within 5A° from this residue is the charged sidechain of the Arginine residue (R165), which belongs to a different secondary–structure element (β–sheet XI). Thus, the slow relative motion of the two heavy secondary–structure elements (or even heavier domains containing these secondary–structure elements) contributes to $\bar{v}(t)$ in Fig. 5. This explains why even at 20ns after the excitation pulse the relaxation is not over yet. The Trp residue in F3W is in a flexible part of the protein, therefore $\bar{v}(t)$ in Fig. 6 contains no contribution from the relative motions of heavy domains, and the relaxation is over in about 3ns after the excitation pulse.

3. Theoretical Basis.

In section 2 it was shown on a qualitative level that there is a connection between the specific location of the fluorophore and the experimentally–observed relaxation curve. To make a practical use of the experimental relaxation curves, the effect of fluorophore location needs to be understood on a quantitative level. This is the main aim of the proposed activity. A quantitative model of the observed phenomena can be based on the theory described in this section.

3.1. Normal vibration modes.

It is well known that the motion of a molecule with N nuclei can be represented as a superposition of motions along 3 translational coordinates, 3 rotational coordinates, and $3N–6$ generalized vibrational coordinates [36,37]. For the generalized vibrational coordinates one may use a consistent set of bond lengths, bond angles, and torsion angles. Generalized vibrational coordinates are suitable, but not convenient, for the description of a collective motion of secondary–structure elements or domains. The collective motions are represented much better in terms of the normal coordinates [38]. A transition from the generalized to the normal coordinates requires a Taylor expansion of the potential and kinetic energy of the protein molecule and truncation of all the terms higher than second power [36–38], therefore the description by the normal coordinates is accurate only for vibrations of small amplitude. Collective motions induced by the optical excitation of a Trp residue (or another fluorophore inserted in a protein molecule) are always small in amplitude, therefore the truncation of the terms higher than second power results in insignificant errors.

If the three–dimensional structure of a protein is known, then its normal vibration modes

can be found by one of the forcefield−based simulation engines, which include CHARMm®, Discover®, and the Open Force Field. A normal vibration mode k corresponds to a normal coordinate Q_k , a vibration frequency ω_k , and an eigenvector $\vec{A}_k|$. A linear array of 3N−6 frequencies ω_k and a (3N−6)×(3N−6) matrix A_{jk} are usually calculated by a forcefield−based simulation engine or by solving a generalized eigenvalue−eigenvector problem.

3.2. The connection between normal modes and experimental $\bar{\nu}(t)$ curve.

A small displacement dQ_k along the normal coordinate Q_k changes the distances between atoms; if some atoms carry non−zero net electric charges, then this displacement also changes the local electric field at every point inside and near the protein molecule. $dE_{\Delta\mu}$ will be defined as the change in the projection on the direction of $\Delta\vec{\mu}$ of the local electric field at the location of the fluorophore, where $\Delta\vec{\mu}$ is the difference between the static dipole moments of the fluorophore in the excited and in the ground state. The derivative $dE_{\Delta\mu}/dQ_k$ represents the sensitivity of the local field to the displacements along the normal nuclear coordinate Q_k :

$$S_k = \frac{\partial E_{\Delta\mu}}{\partial Q_k} \tag{2}$$

Note, that in the calculation of the derivatives S_k one has to take into account three contributions to $dE_{\Delta\mu}$: (i) the contribution from the change in the electric field generated by the displacement of charged atoms outside the fluorophore, (ii) the contribution from the rotation of $\Delta\vec{\mu}$, and (iii) the contribution from the translational motion of the fluorophore. In the calculation of the electric field it is necessary to take into account the dielectric properties of the solvent. Thus, calculating the derivatives S_k represents a significant mathematical problem and a considerable computational task. Commercial forcefield−based simulation engines, which include CHARMm®, Discover®, and the Open Force Field, cannot perform this task.

The rigidity R_k of the protein molecule along the normal coordinate Q_k will be defined as the second derivative of the potential energy U with respect to Q_k ,

$$R_k = \frac{\partial^2 U}{\partial Q_k^2} \tag{3}$$

Normal coordinates are often mass−scaled [38]. If this is the case, then

$$R_k = \omega_k^2 \tag{4}$$

If the values of S_k , R_k , and ω_k are known for every normal mode, then it is possible to write a general theoretical expression for $\bar{\nu}(t)$:

$$\bar{\nu}(t) = \bar{\nu}(0) - \frac{|\Delta\vec{\mu}|^2}{hc} \sum_{k=1}^{3N-6} \frac{S_k^2}{R_k} [1-\cos(\omega_k t)] \tag{5}$$

From Eq. (5) it follows that the contribution of each normal mode to $\bar{\nu}(t)$ is directly proportional to the second power of S_k and inversely proportional to the first power of R_k . The latter means that the "soft" modes with low values of R_k have greater contributions to microscopic dielectric relaxation than the "rigid" modes with high values of R_k . Since R_k is directly proportional to ω_k^2 , the "rigid" modes correspond to high vibration frequencies and

can be studied using infrared spectroscopy and/or Raman spectroscopy. On the contrary, the "soft" modes correspond to low frequencies, and this makes it difficult to study these modes using infrared or Raman spectroscopy. In the case of microscopic dielectric relaxation the opposite is true. The "rigid" modes have too small contributions to be detected, while the "soft" modes are responsible for the most part of the observed effect, and therefore can be studied using time–resolved emission from Trp residues and other fluorophores.

Eq. (5) was originally derived for an N–atomic molecule in vacuum. In solution the discrete spectral lines of the vibration frequency spectrum will be broadened. Assuming that the broadened spectral lines have Lorentzian shapes and $\Delta\omega_k$ denotes the half width at half maximum of the frequency distribution for the normal mode k, Eq. (5) can be transformed to

$$\bar{v}(t) = \bar{v}(\infty) + \frac{|\Delta\vec{\mu}|^2}{hc} \sum_{k=1}^{3N-6} \frac{S_k^2}{R_k} \cos(\omega_k t) \exp(-\Delta\omega_k t) \tag{6}$$

For the "soft" modes of large molecules in finite–viscosity solvents the central frequencies ω_k may be smaller than the bandwidths $\Delta\omega_k$. In this case the cosine on the right–hand side of Eq. (6) can be omitted,

$$\bar{v}(t) \approx \bar{v}(\infty) + \frac{|\Delta\mu|^2}{hc} \sum_{soft} \frac{S_k^2}{R_k} \exp(-\Delta\omega_k t) \tag{7}$$

In Eq. (7) the summation is carried out over soft modes only, for which $\omega_k \ll \Delta\omega_k$. The expression on the right–hand side of Eq. (7) represents a linear combination of a constant and a few exponential functions of time; except for the uncommon parameters this is the most popular fitting function for relaxation kinetics of any kind. The most common form of this fitting function is:

$$\bar{v}(t) = \bar{v}(\infty) + \sum \alpha_k \exp(-t/\tau_k) \tag{8}$$

The relaxation amplitudes α_k and relaxation times τ_k that enter in Eq. (8) are directly related to the characteristics of soft normal modes:

$$\alpha_k = \frac{|\Delta\mu|^2 S_n^2}{hc R_n} \quad , \quad \tau_k = 1/\Delta\omega_k \tag{9}$$

3.3. Summary for the theoretical section.

The coefficients S_k defined in Eq. (2) vary with the location of the fluorophore and with the orientation of $\Delta\vec{\mu}$. This explains on the quantitative level why the relaxation curves $\bar{v}(t)$ obtained with the Trp residues inserted in two different locations in the same protein may differ considerably.

The theory also explains why the "soft" modes make significant contributions to the experimental relaxation curves $\bar{v}(t)$, while the "rigid" modes with the frequencies that fall in the spectral range of Infrared and Raman instruments are not significant contributors to $\bar{v}(t)$.

In spite of the positive aspects mentioned above, the present theory is oversimplified and has a number of weaknesses. The parameters $\Delta\omega_k$ were practically taken out of the blue and their

values are not connected to the protein structure in the present theory. Rough approximations were made regarding the dampening effects of a finite–viscosity solvent on the normal mode dynamics: (i) the theory completely ignores the mixing of different normal vibrations by the dissipative terms, and (ii) the solvent effect on the frequencies ω_k is neglected. In fact, the frequencies corresponding to soft vibrations can be significantly reduced or even reduced to zero (aperiodic vibration) due to the presence of the dissipative terms. All these problems can be corrected in the future; this is one of the aims of the proposed activity.

4. Experimental Methods.

4.1. Protein mutagenesis, expression and purification.

IIA^{Glc} protein of *Escherichia Coli* was selected for the current project because of a combination of four important qualities: (i) it is a small (18kD) water–soluble globular protein, (ii) it is exceptionally thermodynamically stable in the temperature range from $-10°C$ to $+69°C$, (iii) the wild–type IIA^{Glc} has no Trp, no Tyr, and no Cys residues, and (iv) the 3D structure of this protein has been determined by X–ray [34] and NMR [35] methods.

As a part of the proposed project several single–Trp mutants of the IIA^{Glc} protein will be made in quantities sufficient for fluorescence measurements. Two mutant proteins, E21W and F3W, have been successfully made and used in the preliminary studies, see pages 4–6 of this Project Description and reference [25]. About 15 or 20 other mutant proteins will be made using similar methods.

Mutagenesis will be done using the QuickChange™ site–directed mutagenesis kit from Straragene. Plasmid pET 21 *crr* as a template and two primers will be used for PCR. The primers will have the Trp codon for the mutated aminoacid. PCR product of the full–size mutated template will be digested with Dpn I enzyme to remove non–mutated template and cloned into XL–1 competent cells.

Each of the mutated IIA^{Glc} proteins will be expressed in *Escherichia Coli* BL21 (DE3) strain where the *crr* gene has been deleted [39] and purified using a previously published technique [40]. The purity of the proteins will be tested using polyacrylamide gels.

Destabilizing effects of mutations will be evaluated by measuring the denaturation temperature, which equals 69.1°C for the wild–type protein. The mutant proteins with denaturation temperatures below 69°C will be tested in terms of activity and biophysical characteristics such as circular dichroism to make sure that the mutations did not result in the loss of the 3D structure of the wild type. The methods for determining the biological activity (phosphotransfer) of the mutant forms of the protein are straightforward and are routine in the laboratory. It is of interest that several mutant forms of the protein already available exhibit full or partial activity. Fluorescence emission from each mutant form will be also tested for heterogeneity, which usually indicates the presence of unfolded protein.

4.2. Phosphorylation of IIA^{Glc} protein.

The protein IIA^{Glc} (formerly known as III^{Glc}) is an important component of the phosphoenolpyruvate:glycose phosphotransferase system (PTS) of *Escherichia Coli*. The PTS

catalyses the transmembrane transport and phosphorylation of a number of simple sugars. Glucose uptake results from a sequential phosphate transfer reaction involving histidine residues of four PTS proteins: Enzyme I, HPr, IIA^{Glc} and IIB^{Glc}. Thus, IIA^{Glc} is available in two states. In a phosphorylated state a negatively–charged phosphate group is covalently attached to the ε–nitrogen atom of His–90. In the dephosphorylated state the negative phoshate group is missing, but every other charged group remains in its place. This presents a unique opportunity to add or remove just one charge in the system under study. pH–dependent titration of carboxyl and amino groups cannot be used for this purpose: for any pH–titratable group with a specific pKa there is always a wide pH range (pKa±2.0 pH units), within which both charged and neutral forms are present. This is known to result in heterogeneity, which hinders the interpretation of experimental relaxation data. If phosphorylation is used to add or remove a charge on a protein, then nearly 100% of protein molecules can be phosphorylated or dephosphorylated at will, and this does not require a pH change.

As a part of the proposed project, in some of the experiments IIA^{Glc} protein will be phosphorylated by adding (prior to fluorescence measurements) an excess of phosphoenol–pyruvate (PEP) and enzymatic quantities of two other proteins: Enzyme I and Hpr. PEP is neither fluorescent nor absorbs in the spectral range where Trp residues are excited. The quantities of Enzyme I and Hpr in solution will be too small to distort the results of fluorescence measurements.

4.3. Time– and spectrally–resolved fluorescence measurements.

Time–correlated single–photon counting data will be obtained using an instrument similar to the prototype that Dmitri Toptygin and his colleagues built at the Jonhs Hopkins University. In that prototype instrument a frequency–doubled output from a cavity–dumped dye laser synchronously pumped by a frequency–doubled output from a mode–locked YAG:Nd laser was used to excite fluorescence. Exciting wavelengths between 286nm and 338nm were generated for different experiments. The exciting radiation was a train of vertically–polarized 12ps laser pulses separated by 245ns intervals. Fluorescence emission was registered by two identical wings in parallel (T–format, 90° geometry). Each emission wing contained a polarizer, a monochromator with 4nm or 8nm spectral resolution, and a microchannel plate photomultiplier R1564U. The polarizers were oriented at 55° to the vertical for the total intensity measurements or at 0° (wing A) and 90° (wing B) for the anisotropy measurements. To remove the strong polarizing effect of the diffraction gratings, the prototype instrument contained a wedge depolarizer before the input slit of the monochromator in each wing. Photon counts were stored in 2048 channels (13.333 ps/channel). The excitation pulse was recorded using a scatterer cell. Typically the excitation pulse was about 5 channels or 65ps FWHM; this represents the time resolution of the prototype instrument. The excitation pulse was recorded quasi–simultaneously by switching between the fluorescence and scatterer cuvettes every 15 seconds; this eliminated systematic errors due to slow drifts in optical and electronic devices.

The prototype instrument has been meticulously tested for systematic errors. After several sources of systematic errors (cross–talk between electronic devices, reflections in coaxial cables) were eliminated, the photon counts strictly follow Poissonian statistics up to

approximately 10^6 photons per channel; at higher photon counts the differential nonlinearity of the analog–to–digit converter can be detected.

To generate instantaneous emission spectra similar to those in Fig.4 and relaxation curves similar to those in Figs.5,6, a set of 31 decay curves is required. The decay curves will be measured at emission wavelengths from 300nm to 450nm, with a 5nm increment. The total number of registered photons in all decay curves must be not less than 150 million; a smaller number of photons results in poor signal–to–noise ratio. Reconstruction of corrected and deconvoluted instantaneous emission spectra from the raw experimental data will be carried out using the method described in ref. [25].

5. General Plan of Work.

Part 1. Computer programming. A computer program will be written based on the theory in Section 3 of this Project Description. The program will not perform normal mode analysis; it will use the array ω_k and the matrix A_{jk} calculated by a commercial forcefield–based simulation engine. During the calculation of the coefficients S_k the protein and the solvent will be treated as homogeneous dielectrics of different permittivities ($\varepsilon \approx 2$ for the protein[1] and $\varepsilon \approx 80$ for water). This will make it possible to calculate theoretical estimates for the amplitude α_k associated with the contribution of specific normal modes to the experimental relaxation curve $\bar{v}(t)$.

Part 2. Computer modeling. A search for prospective single–Trp mutants will start from the known 3D structure of the wild–type IIA^{Glc} protein of *Escherichia Coli*.[2] Each of the residues 19 through 168 will be replaced with a Trp residue (one residue at a time).[3] The resulting 150 single–Trp structures will be energy–minimized. The structures that differ significantly from the wild–type structure will be rejected. For the remaining structures normal mode analysis will be performed and the amplitudes α_k corresponding to soft normal modes will be calculated using the program from Part 1. For each slow normal mode the single–Trp structure that has the highest amplitude α_k associated with this mode and low amplitudes associated with other normal modes will be selected. This must give between 10 and 20 prospective single–Trp mutants.

Part 3. Computer modeling. A search for phosphorylation[4]–sensitive single–Trp mutants of the IIA^{Glc} protein will be undertaken using a subset of the 150 single–Trp structures from Part 2.

---

1   This ε value accounts for the electronic polarizability and not for the nuclear displacements, since the latter are accounted for by the normal modes; including the effect of nuclear displacements in both places would be an error. If the nuclear displacements were included in the dielectric permittivity, then a $\varepsilon \approx 4$ would have to be used for the protein core.

2   As indicated on page 9, IIA^{Glc} protein was selected for a combination of four important qualities: (i) it is a small water–soluble globular protein, (ii) it is exceptionally thermodynamically stable in a wide temperature range, (iii) the wild–type IIA^{Glc} has no Trp, no Tyr, and no Cys, and (iv) the 3D structure of the protein is known.

3   Residues 1 through 18 form a flexible tail of indefinite 3D structure.

4   IIA^{Glc} is a phospho–carrier protein and it can be enzymatically phosphorylated, see page 10. Phosphorylated form of IIA^{Glc} differs from the dephosphorylated form by one charged group. The effect of this charge on the experimentally–recorded relaxation curves will be studied using the phosphorylation–sensitive single–Trp mutants of IIA^{Glc} protein.

Only the mutant proteins with Trp residues within 10Å from the phosphorylation site will be included in this subset. Two or three single–Trp mutants will be selected for which the relaxation curve $\bar{v}(t)$ is expected to be most sensitive to phosphorylation. The phosphate group carries a negative charge; placing a point charge near the fluorophore is expected to alter the amplitudes α_k in a predictable manner. The phosphorylation–sensitive single–Trp mutants will be used to test this hypothesis.

Part 4. Protein production and purification. Mutagenesis, expression, and purification will be carried out to produce the prospective single–Trp mutant proteins selected in Part 2 and 3.

Part 5. Instrument building. The work on this part of the project will start on day one and will be done in parallel with the work in parts 1–4. An instrument for time– and spectrally–resolved fluorescence measurements similar to the prototype described in Section 4.3 will be built. The most important characteristics of the instrument are: not less than 4×10^6 excitation pulses per second, time resolution not worse than 70ps FWHM for the instrument as a whole (including the laser, emission monochromator, microchannel plate photomultiplier, and time–correlated photon–counting electronics), excitation wavelength tunable at least from 285nm to 300nm, emission monochromator tunable at least from 285nm to 480nm, emission monochromator bandwidth 4nm or 8nm, time–correlated photon–counting electronics capable of processing at least 4×10^4 emission photons per second without systematic errors.

Part 6. Picosecond fluorescence measurements. For each single–Trp mutant produced in Part 4, a set of time–correlated single–photon–counting fluorescence decay curves will be collected at 31 emission wavelengths and instantaneous emission spectra will be reconstructed. Both homogeneity criteria will be applied (for the second criterion a set of steady–state emission spectra will be measured using different exciting wavelengths). For the mutant proteins that pass both homogeneity tests the data will be used to generate the relaxation curves $\bar{v}(t)$. Time–resolved fluorescence measurements will be first carried out in a low–viscosity aqueous pH buffer, and subsequently repeated in higher–viscosity solvents containing up to 30% of glycerol as a cosolvent. The data will be used in the study of the solvent viscosity effect on the dynamics of normal vibrations. For every mutant protein in every solvent, a time–resolved anisotropy measurement will be also carried out; this will help to separate the contributions to $\bar{v}(t)$ produced by the Trp sidechain rotation and by the translational motion of the charged groups relative to the fluorophore.

Part 7. Analysis of experimental data. The relaxation data obtained in Part 6 will be fit by the theoretical model equations. As a first approximation, we will use the model based on the simplified theory described in Section 3 of this Project Description. This will give an idea which model parameters differ from the theoretical predictions most of all. The theoretical equations for normal–mode vibrations will be modified to include phenomenological dissipative terms. The effect of the solvent viscosity on the coefficients in front of the dissipative terms will be investigated. The ultimate goal of this is to find general analytical expressions for the coefficients in front of the dissipative terms in terms of the protein shape, mode of vibration, and solvent viscosity. The laws of classical hydrodynamics may be used to account for the effects of protein shape and the mode of vibration.

Part 8. Biophysical theory and computer programming. A new program for the analysis of slow normal vibrations of protein molecules in finite–viscosity solvents will be written based on the results of Part 7. The curves $\bar{v}(t)$ generated by this program will be compared to the experimental curves from Part 6. The general parameters (not specific to the normal modes of IIA^{Glc} protein) in theoretical equations will be adjusted to achieve the best agreement between the theory and the data. The new program is expected to become a powerful tool for the computer analysis of the dynamics of collective motions in proteins. The program will be available to a broad circle of users.

Part 9. Biophysical theory and computer programming. The description of the solvent as a simple homogeneous dielectric will be upgraded by taking into account the motion of free ions (Poisson–Boltzmann equation). This will be also incorporated in the new program from Part 8. The motion of free ions may have a significant effect on the dynamics of proteins with a high density of surface charges. The proposed upgrade will make the program suitable for the studies of such proteins.

6. References Cited.

[1] Maroncelli, M. The dynamics of solvation in polar liquids. *Journal of Molecular Liquids* **1993**, *57*, 1–37.

[2] Kaatze, U.; Giese, K. Dielectric–relaxation spectroscopy of liquids – frequency–domain and time–domain experimental methods. *J. Phys. E Sci. Instrum.* **1980**, *13*, 133–141.

[3] Horng, M.L.; Gardecki, J.A.; Papazyan, A.; Maroncelli, M. Subpicosecond measurements of polar solvation dynamics: coumarin 153 revisited. *J. Phys. Chem.* **1995**, *99*, 17311–17337.

[4] Stratt, R.M.; Maroncelli, M. Nonreactive dynamics in solution: the emerging molecular view of solvation dynamics and vibrational relaxation. *J. Phys. Chem.* **1996**, *100*, 12981–12996.

[5] Gardecki, J.A.; Maroncelli, M. Comparison of the single–wavelength and spectrally–reconstructed methods for determining the solvation–response function. *J. Phys. Chem. A* **1999**, *103*, 1187–1197.

[6] Bakhshiev, N. G.; Mazurenko, Yu. T.; Piterskaya, V. I. Luminescence decay in different portions of the luminescent spectrum of molecules in viscous solutions. *Opt. Spectrosc.* **1966**, *21*, 307–309.

[7] Bakhshiev, N. G.; Mazurenko, Yu. T.; Piterskaya, V. I. Relaxation effects in the luminescence characteristics of viscous solutions. *Izv. Akad. Nauk. SSSR. Ser. Fiz.* **1969**, *32*, 1262–1266.

[8] Mazurenko, Yu. T.; Bakhshiev, N. G. Effect of orientation dipole relaxation on spectral, time, and polarization characteristics of the luminescence of solutions. *Opt. Spectrosc.* **1970**, *28*, 490–494.

[9] Jimenez, R.; Fleming, G.R.; Kumar, P.V.; Maroncelli, M. Femtosecond solvation dynamics of water. *Nature* **1994**, *369*, 471–473.

[10] Vajda, S.; Jimenez, R.; Rosenthal, S.J.; Fidler, V.; Fleming, G.R.; Castner, E.W. Femtosecond to nanosecond solvation dynamics in pure water and inside the gamma–cyclodextrin cavity. *J. Chem. Soc.–Faraday Trans.* **1995**, *91*, 867–873.

[11] Lang, M.J.; Jordanides, X.J.; Song, X.; Fleming, J.R. Aqueous solvation dynamics studied by photon echo spectroscopy. *J. Chem. Phys.* **1999,** *110,* 5884–5892

[12] Brand, L.; Gohlke, J.R. Nanosecond time–resolved fluorescence spectra of a protein–dye complex. *J. Biol. Chem.* **1971,** *246,* 2317–2319.

[13] Grinvald, A.; Steinberg, I.Z. Fast relaxation process in a protein revealed by decay kinetics of tryptophan fluorescence. *Biochemistry* **1974,** *13,* 5170–5178.

[14] Gafni, A.; DeToma, R.P.; Manrow, R.E.; Brand, L. Nanosecond decay studies of a fluorescent probe bound to apomyoglobin. *Biophys. J.* **1977,** *17,* 155–168.

[15] Lakowicz, J.R.; Cherek, H. Dipolar relaxation in proteins on the nanosecond timescale observed by wavelength–resolved phase fluorometry of tryptophan fluorescence. *J. Biol. Chem.* **1980,** *255,* 831–834.

[16] Lakowicz, J.R.; Gratton, E.; Cherek, H.; Maliwal, B.P.; Laczko, G. Determination of time–resolved fluorescence emission spectra and anisotropies of a fluorophore–protein complex using frequency–domain phase–modulation fluorometry. *J. Biol. Chem.* **1984,** *259,* 10967–10972.

[17] Pierce, D.W.; Boxer, S.G. Dielectric relaxation in a protein matrix. *J. Phys. Chem.* **1992,** *96,* 5560–5566.

[18] Wang, R.; Sun, S.; Bekos, E.J.; Bright, F.V. Dynamics surrounding Cys–34 in native, chemically denatured, and silica–adsorbed bovine serum albumin. *Anal. Chem.* **1995,** *67,* 149–159.

[19] Lundgren, J.S.; Bright, F.V. Effects of surfactants on the dynamical behavior of acrylodan–labeled bovine serum albumin. *J. Phys. Chem.* **1996,** *100,* 8580–8586.

[20] Riter, R.R.; Edington, M.D.; Beck, W.F. Protein–matrix solvation dynamics in the α subunit of C–phycocyanin. *J. Phys. Chem.* **1996,** *100,* 14198–14205.

[21] Jordanides, X.J.; Lang, M.J.; Song, X.; Fleming, G.R. Solvation dynamics in protein environments studied by photon echo spectroscopy. *J. Chem. Phys. B* **1999,** *103,* 7995–8005.

[22] Toptygin, D.; Rodgers, M.E.; Brand, L. Time–resolved fluorescence emission spectra of Trp in *E. Coli* galactose repressor exhibit the red shift characteristic of electrostatic relaxation. *Biophys. J.* **1999,** *76,* A260.

[23] Vincent, M.; Gilles, A.M.; de la Sierra, I.M.L.; Briozzo, P.; Barzu, O.; Gallay, J. Nanosecond fluorescence dynamic Stokes shift of tryptophan in a protein matrix. *J. Phys. Chem. B* **2000,** *104,* 11286–11295.

[24] Toptygin, D.; Savtchenko, R.S.; Meadow, N.D.; Brand, L. Homogeneous spectrally– and time–resolved fluorescence emission from single–tryptophan mutants of IIA^{Glc} protein. *Biohys. J.* **2001,** *80,* A1515.

[25] Toptygin, D.; Savtchenko, R.S.; Meadow, N.D.; Brand, L. Homogeneous spectrally– and time–resolved fluorescence emission from single–tryptophan mutants of IIA^{Glc} protein. *J. Phys. Chem. B* **2001,** *105,* 2043–2055.

[26] Weber, G. "Excited states of proteins." In: Light and Life. Edited by McElroy, W. D. and Glass, B. Baltimore: Johns Hopkins Press, 1961. pp. 82–107.

[27] Callis, .R. ^1L_a and ^1L_b transitions of tryptophan: Applications of theory and experimental observations to fluorescence of proteins. *Meth. Enzymology* **1997,** *278,* 113–150.

[28] Ross, J.B.A.; Schmidt, C.J.; Brand, L. Time–resolved fluorescence of the two tryptophans in horse liver alcohol dehydrogenase. *Biochemistry* **1981,** *20,* 4369–4377.

[29] Fee, R.S.; Milsom, J.A.; Maroncelli, M. Inhomogeneous decay kinetics and apparent solvent relaxation at low temperatures. *J. Phys. Chem.* **1991,** *95,* 5170–5181.

[30] Szabo, A. G.; Rayner, D. M. *J. Am. Chem. Soc.* **1980,** *102,* 554.

[31] Chen, R. F.; Knutson, J. R.; Ziffer, H.; Porter, D. *Biochemistry* **1991,** *30,* 5184.

[32] Toptygin, D.; Brand, L. "Spectrally– and time–resolved fluorescence emission of indole during solvent relaxation: a quantitative model." *Chem. Phys. Lett.* **2000,** *322,* 496–502.

[33] Demchenko, A.P. The red–edge effects: 30 years of exploration. *Luminescence* **2002,** *17,* 19–42.

[34] Worthylake, D.; Meadow, N. D.; Roseman, S.; Liao, D.–I.; Herzberg, O.; Remington, S. J. 3–dimensional structure of the Escherichia–coli phosphocarrier protein–III^{Glc}. *Proc. Natl. Acad. Sci. USA* **1991,** *88,* 10382–10386.

[35] Pelton, J. G.; Torchia, D. A.; Meadow, N. D.; Wong, C.–Y.; Roseman, S. Secondary structure of the phosphocarrier protein III^{Glc}, a signal–transducing protein from Escherichia coli, determined by heteronuclear three–dimensional NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* **1991,** *88,* 3479–3483.

[36] Landau, L. D.; Lifshitz, E. M. Mechanics. Vol. 1 of Course of Theoretical Physics, translated from the Russian by J. B. Sykes and J. S. Bell. 3d ed. Oxford; New York: Pergamon Press, 1976. §23, §24.

[37] Landau, L. D.; Lifshitz, E. M. Quantum Mechanics (Non–Relativistic Theory). Vol. 3 of Course of Theoretical Physics, translated from the Russian by J. B. Sykes and J. S. Bell. 3d ed. Oxford; New York: Pergamon Press, 1977. §§100–105.

[38] Levitt, M.; Sander, C.; Stern, P.S. Protein normal–mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *J. Mol. Biol.* **1985,** *181,* 423–447.

[39] Meadow, N. D.; Roseman, S. Rate and equilibrium constants for phosphoryltransfer between active site histidines of Escherichia coli HPr and the signal transducing protein III^{Glc}. *J. Biol. Chem.* **1996,** *271,* 33440–33445.

[40] Pelton, J. G.; Torchia, D. A.; Meadow, N. D.; Wong, C.–Y.; Roseman, S. H^1, N^{15}, and C^{13} NMR signal assignments of III^{Glc}, a signal–transducing protein of *Escherichia Coli,* using 3–dimensional triple–resonance techniques. *Biochemistry* **1991,** *30,* 10043–10057.

[41] Toptygin, D.; Savtchenko, R. S.; Meadow, N. D.; Roseman, S.; Brand, L. Effect of the solvent refractive index on the excited–state lifetime of a single tryptophan residue in a protein. *J. Phys. Chem. B* **2002,** *106,* 3724–3734.