TARTU UNIVERSITY
Faculty of Physics and Chemistry
Institute of Material Science

VEERA KRASNENKO

MASTER THESIS

PHOTOPHYSICS OF GREEN AND BLUE FLUORESCENT PROTEINS

SPECIALTY: Optics and Spectroscopy

Supervisor: Senior Research Associate,
candidate of physical and
mathematical sciences, Koit Mauring

Tartu 2004

TABLE OF CONTENTS

1. INTRODUCTION

Figure 1. The jellyfish *Aequorea victoria*

In spite of the fact that only decade has passed since discovery of the green fluorescent protein (GFP) present in the jellyfish *Aequorea Victoria* (Fig. 1), it already has become one of the most widely studied and exploited proteins in biochemistry and cell biology. This protein is capable of synthesizing an effectively emissive fluorophore, which promotes its wide use. High-resolution crystal structures of GFP offer good basis to understand and computationally model the relation between protein structure and spectroscopic properties.

Since 1992, the amount of publications applied to GFP increases exponentially and currently approaches 9000!

2. APPLICATION OF THE GFP AND ITS MUTANTS

Amazing property of all GFP-like proteins is their striking steadiness to the change of temperature, pH factor, protein-destroying ferments – proteases and different chemical agents that cause denaturation of proteins.

GFP of *Aequorea Victoria* was cloned in 1992 and was used as a genetic marker by connecting the gene of GFP with the genes of other proteins or by injecting GFP's RNA into various cells (Fig. 2).

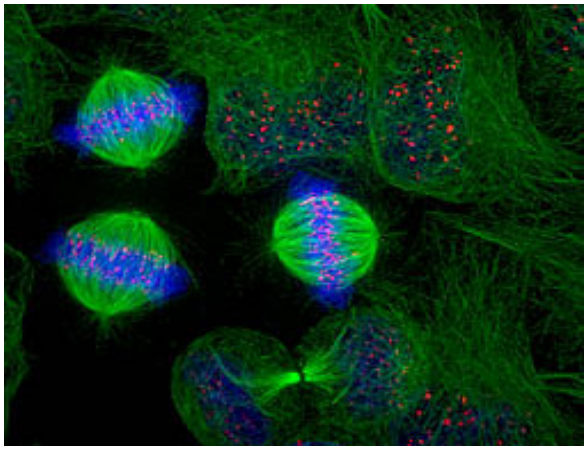

Figure 2. Human epithelial cells: microtubules (labeled green), kinetochores (red), and DNA (blue). (Courtesy of Dr. Jennifer Waters Shuler and Dr. Adrian Salic, Nikon Imaging Center, Harvard Medical School)

GFP has proven its value as the marker of gene expression and protein targeting in intact cells and organisms. Mutagenesis and engineering of GFP into chimeric proteins are opening new vistas in physiological indicators, biosensors and photochemical memory [1], e.g.:

- in immunology – it is possible to bind together antigen with protein-marker and another marker with antibody and to analyze changes in spectra for indication of the antigen-antibody interaction;

- in embryology - we can mark predecessor cells and watch their further behavior: which tissues they produce, where they migrate (some cells make whole journeys inside the body);

- in biology, it is often necessary to track individual proteins, cellular proteins, bacteria and viruses. Most appropriate for this are fluorescent markers (on the intracellular level) and dyes (on the level of the whole organism) [2];

- creation of analogs for transistors (biocomputers). The chromophore of GFP, which is normally mostly anionic, can eventually, be converted into a

protonated state. Excitation of this form restores the normal anionic state. Such cycling can be repeated many times with apparently no fatigue, so that it potentially represents a basis for an optical memory at the single molecule level. It might also be particularly advantageous for multiple determinations of diffusion or trafficking on the same region of interest [1].

That way, it is possible to make (in fluorescent microscope) locations and speed of formation of proteins coded in any other genes visible, to watch the growth of cellular clones, including pathogenic bacteria and cancer cells.

Also, using this technology, we can look into the organism after the propagation of different viruses, including those that contain not DNA, but RNA, i.e. retroviruses, e.g. AIDS [2].

Recently GFP has been applied as a marker in combination with nestin. Nestin is a protein, which is typical for neural stem and some other cells. This method has helped to disprove a dogma about irreplaceability of brain neurons in adulthood [2].

After GFP was cloned from *Aequorea Victoria*, a new problem needed its solution – to mutate GFP to produce markers of different color, to tag at once several genes with fluorescent markers in single experiment.

The application of multicolored markers is especially perspective with fluorescence resonance energy transfer (FRET) (Fig. 3).

This method allows the monitoring of interactions between protein molecules. Mutants from GFP from *Aequorea Victoria* were the first fluorescent markers used in FRET.

FRET is a quantum-mechanical phenomenon that occurs when two fluorophores are in mole-cular proximity

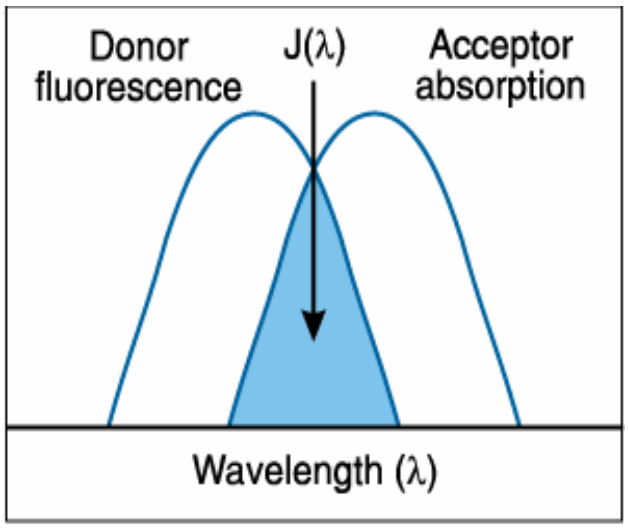

Figure 3 The principle of spectral overlap in FRET

(<100 ° A apart) and the emission spectrum of one fluorophore, the donor, overlaps the excitation spectrum of the second fluorophore, the acceptor. Under these conditions, excitation of the donor can produce emission from the acceptor. Any biochemical transformation that changes the distance between the fluorophores or relative orientation of their transition dipoles will modulate the efficiency of FRET [3-5]. Because FRET is a through-space effect, it is not necessary to perturb either GFP alone but rather only the linkage or spatial relationship between them. The change in ratio of acceptor to donor emissions is nearly ideal for cellular imaging and flow cytometry because the two emissions can be obtained simultaneously and their ratio cancels out variations in the absolute concentration of the GFPs, the thickness of the cell, the brightness of the excitation source, and the absolute efficiency of detection. Because the sample need be excited at only one wavelength, which should preferentially excite the donor, FRET is ideal for laser-scanning confocal microscopy and FACS [5]. With this technique, it is possible to study interaction between antibody and antigen, virus protein and the protein of the host cell, molecule receptor and protein hormone, etc. Also, this technique may be used in medical research to monitor destruction of protein by ferments and the separation of protein complexes by drug molecules [2].

3. STRUCTURE OF GFP MOLECULE

The protein database currently includes 22 structures of GFP, its mutants and two analogs of GFP. GFP-like proteins, evolved from corals and other sea animals, differ very much in absorption and fluorescence spectra. All of them are nonetheless very similar in their structure.

In 1992, the gene of green fluorescent protein was cloned and infused into the gene of *Esherichia coli* [2]. As a result, GFP was obtained in considerable quantity - enough to decrypt the sequence of amino acids and unique characteristics of GFP were discovered. Interestingly, only molecular oxygen is needed for GFP-folding from the outside (in order for bright green fluorescence to appear). Detected fluorescence appears after half an hour after synthesis and reaches its maximum only after several hours [6].

Shimomura O., Jonson F.H. and Saiga Y. ascertained amino acid sequence of green protein (its molecular mass is 28 kDa) and found that it consists of 238 residues [7].

Quantum yield of GFP is remarkably high – about 0.8. In other words, only 20% of absorbed photons transform into heat, while the other 80% transform into emitted light.

Tertiary structure of GFP is very complicated. Protein has a unique structure of the hollow cylinder. The diameter of the cylinder is about 24 Å, its height is 42 Å [1]. GFP is an 11-stranded β–barrel threaded by an α-helix running up the axis of the cylinder (Figure 4). The strands form an almost seamless symmetrical β-sheet structure, and only one irregularity may be found

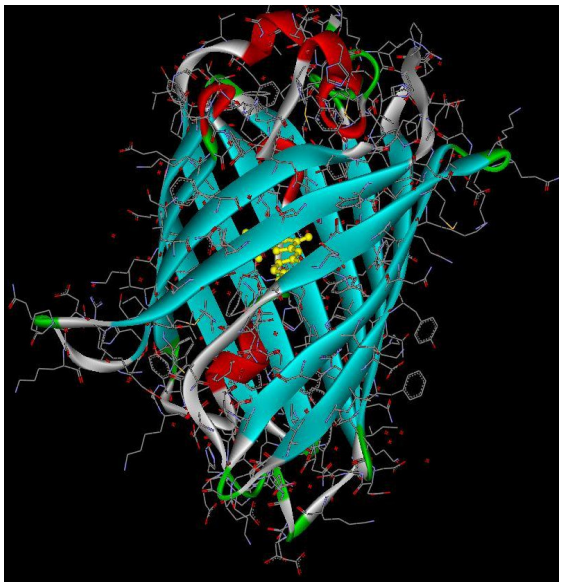

Figure 4. Stereoview of the three-dimensional structure of Green Fluorescent Protein, showing 11 β-strands forming a hollow cylinder through which is threaded a helix bearing the chromophore, shown in ball-and-stick representation.

between the two of them [6]. The chromophore is attached to the α-helix and is buried almost perfectly in the center of the cylinder, which has been named a *β*-can [8,9]. Short fragments of the α-helix and loops also form "covers" that close cylinder from the top and the bottom.

Structured water molecules make up a pellicle around the cylinder's surface; a number of water molecules are also located inside this cylinder [6]. Supposedly, such tightly structured barrel is needed for chromophore protection. It provides stability and durability against heating and action of denaturing agents.

3.1. Structure of GFP chromophore

The chromophore is a 4-(p-hydroxybenzylidene)imidazolidin-5-one [20], which is attached to the peptide backbone through the 1- and 2-positions of the ring and has a cis-geometry [1]. It is located inside of the cylinder and occupies central location. Amino acid removal experiments revealed that for the formation of fluorescent chromophore, an almost complete protein structure is required (residues 2-234 or 7-229 [6]). Fluorescence is preserved by amino and carboxyl terminal removals [10].

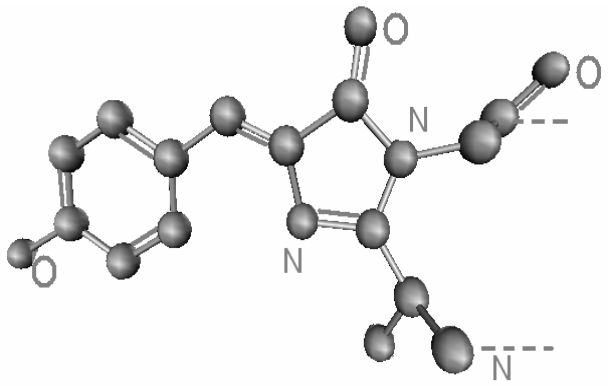

Figure 5. Structure of GFP chromophore

The chromophore is formed from residues 65-67, which are Ser-Tyr-Gly in the native protein. First, GFP folds into a nearly native conformation, then the imidazolinone is formed by nucleophilic attack of the amide of Gly67 on the carbonyl of residue Ser65, followed by dehydration (Tyr66 releases two protons). Finally, molecular oxygen dehydrogenates the *α-β* bond of residue Tyr66 to put its aromatic group into conjugation with the imidazolinone. Only at this stage does the chromophore acquire visible absorbance and fluorescence [11-13].

A number of polar groups and structured water molecules form a network of hydrogen bonds around the chromophore. Particularly important in chromophore surroundings are residues Gln69 (glutamine), Arg96 (arginine), His148 (histidine), Thr203 (threonine), Ser205 (serine) and Glu222 (glutamate) [1].

4. PHOTOPHYSICS OF GFP

From quantum mechanics calculations we know [14] that a free chromophore in excited electronic state can isomerize, i.e. it has possibility to twist around the chemical bonds that bind aromatic rings of the chromophore, which causes nonradiative transitions.

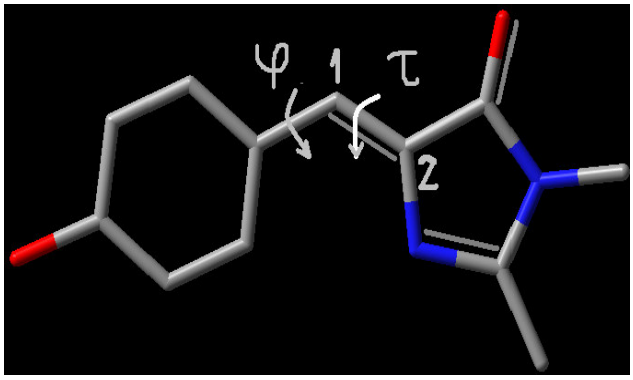

Figure 6. Dependence of chemical bonds spinning energy: φ - around single bond; τ - around double bond.

The neutral chromophore has minimal energy in the ground state when it is flat, i.e. when torsion angles $\tau = 0^\circ$ and $\varphi = 0^\circ$ (Fig. 6). In the excited state position $\tau = 0^\circ$ corresponds to the maximum energy (Fig. 7), while minimal energy relates to $\tau = 90^\circ$. Accordingly, after electronic excitation the chromophore turns out to be in the state with maximum energy and has to twist around chemical bond to minimize its energy.

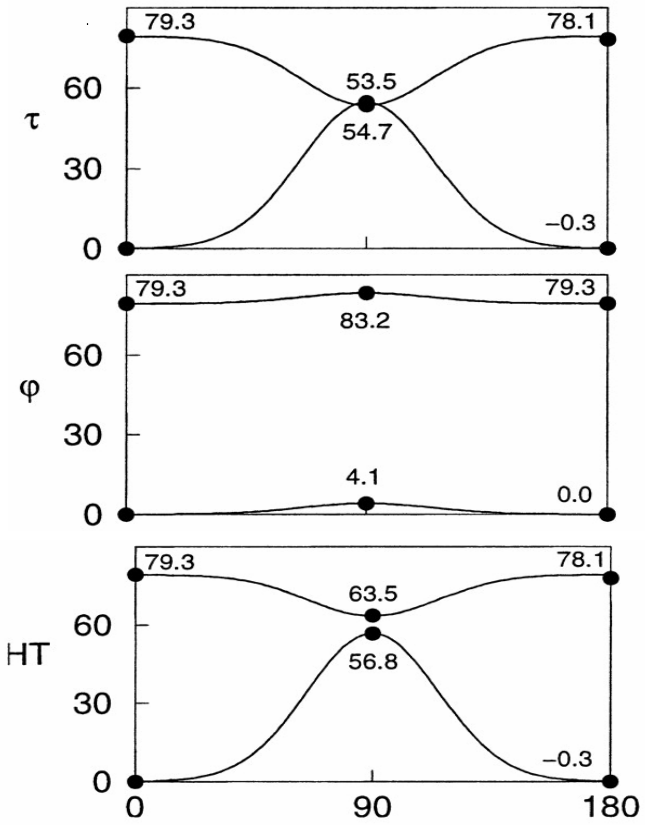

Figure 7. Energy dependence on twisting of chemical bonds [14]: φ - around a single bond; τ - around a double bond; HT – hula twist

In Figure 8, the π - electronic system is shown, calculated by us with CAS method (Complete Active Space implemented in program package of Gaussian'03). We can see that in the ground state, twisting is possible around the single bond (φ angle). In the

excited state, (LUMO) π -electronic cloud doesn't occupy the double bond, its character has been changed to the single bond and twisting most likely will be

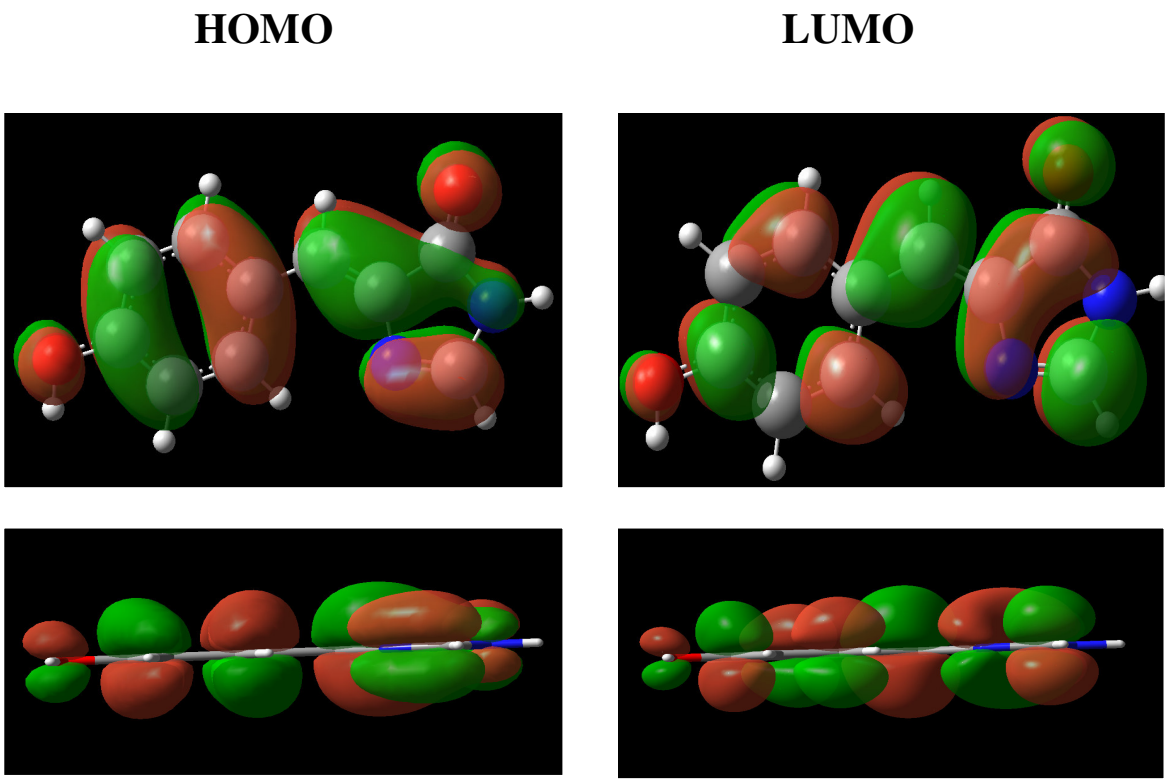

Figure 8. Molecular orbital of chromophore in ground (HOMO) and excited (LUMO) states.

around it (bond 1-2, see Fig. 6) or simultaneously around double and single bonds of propene bridge - hula twist.

The binding of the chromophore aromatic ring with surrounding aminoacids by hydrogen bond hampers the twisting, therefore probability of nonradiative transition decreases.

In the case of nonbonded chromophores, the energy of electronic excitation is spent on twisting, and emitting transition is possible only in long-wavelength area, or transition is nonradiative, since both potential surfaces of ground and excited states (Fig. 7) are located too close to each other [14].

4.1. Temperature dependence of fluorescent GFP mutant

4.1.1. Blue fluorescent protein

Blue fluorescent protein (BFP) is an excellent donor molecule for resonant energy transfer to GFP. It is cloned and well characterized. The BFP is a derivative of the GFP; it contains amino acid substitutions that allow the protein to emit blue light. This difference arose from substituting the tyrosine residue in the chromophore with histidine (Figure 9). BFP has a single excitation maximum at 382 nm; its emission at 448 nm results in strong blue fluorescence. Interestingly, its fluorescence decay is highly dispersive and its lifetime extends from 15 ps to 1,1 ns. This implies that the chromophore has several forms with different probabilities of nonradiative decay and respective fluorescence quantum yield.

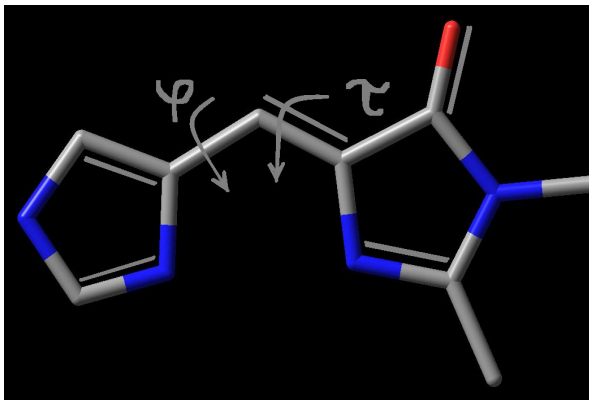

Figure 9. Structure of BFP chromophore

4.1.2. Description of the experiment

We determined the balance between fluorescent and nonfluorescent BFP forms by measuring the dependence of fluorescence intensity on temperature. Fluorescence spectra were measured on dual spectrometer DFS-24 (relative aperture F=1:5.3, inverse dispersion D=0,45 nm/mm). Photodetector was PMT HAMAMATSU in photon counting mode, cooled with Peltier system. Pulses were accumulated into the multichannel analyzer LP4900 B. Absorption spectra were measured with the spectrophotometer Jasco.

Argon laser Coherent CR2000 was applied as a source of fluorescence excitation in ultra-violet. Laser line 351 nm was selected with diffraction grating 600 lines/mm. For spectral measurements in large temperature interval helium cryostat was used. Temperature was adjusted by changing the flow of cold helium gases. Heater power was changed by regulating supply voltage [15].

4.1.3. Results of the experiment

Figure 10 shows the dependence of fluorescence spectrum on temperature.

When temperature decreases, fluorescence intensity increases and absorption maximum shifts in the long-wavelength direction by 2,3 nm. At the same time, spectrum became more structured. Analogously to GFP, we expect that the chromophore in BFP can exist in two forms: hydrogen bonded and nonbonded to neighboring amino acids. The lowering of the temperature increases the contribution of hydrogen-bonded chromophore at the expense of nonbonded ones.

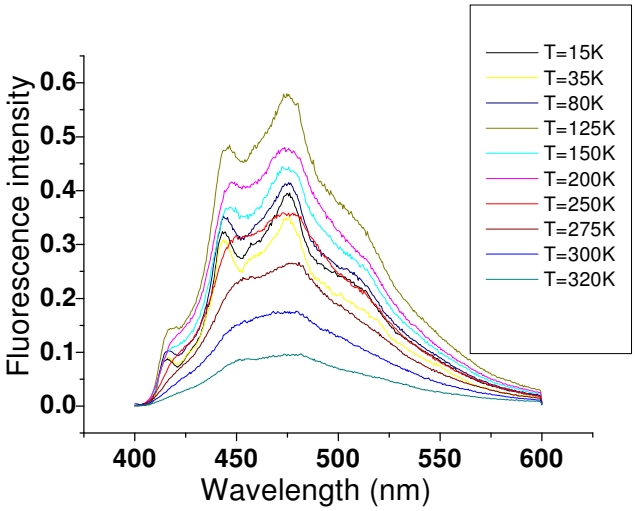

Figure 10. Dependence of fluorescence spectrum on temperature

The dependence of integral fluorescence intensity on temperature is displayed in figure 11. In the range from 320 K to 225 K, fluorescence intensity depends linearly on temperature (correlation coefficient 0.99). As the temperature lowers, the amplitude of the fluctuations between atomic groups in protein decreases and the viscosity of the solvent increases. As a result, the

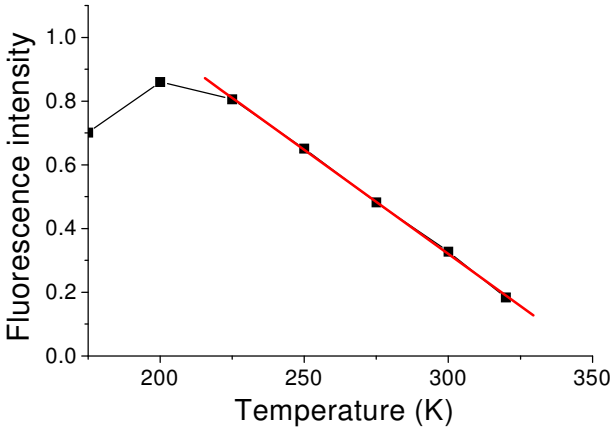

Figure 11. Temperature dependence of fluorescence integral intensity BFP in 50% TRIS-buffer glycerin solution with pH=8

balance between hydrogen-bonded and nonbonded chromophores shifts towards hydrogen-bonded species. This increases the quantum yield of fluorescence.

Figure 12 shows absorption spectra that were obtained after base line subtraction. Since our sample is a scattering medium and has wavelength dependence λ^{-4}, the baseline was approximated as the sum of linear and nonlinear λ^{-4} components. The shape of the spectrum and the absorption do not change with temperature, thus the fluorescence change is not caused by chemical transformations in protein [15].

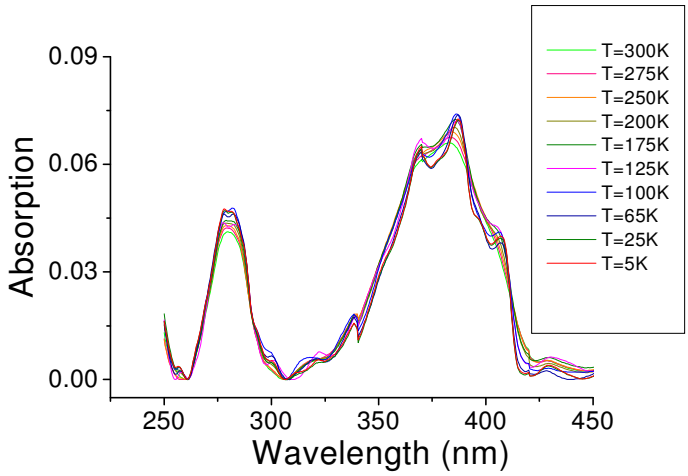

Figure 12. Dependence of absorption on temperature

4.1.4. Analysis of experimental data

Twisting of chemical bonds in BFP and GFP chromophores (cf 4. Photophysics of GFP) slightly differs. In the case of BFP, when chromophore contains an imidazole ring (pentagon) instead of a phenol ring (hexagon) - twisting around a single bond causes a markedly greater energy change, because, in this case, rotation around this bond by $\varphi = 180^\circ$ is not equivalent to configuration $\varphi = 0^\circ$. Moreover, the hydrogen bond with

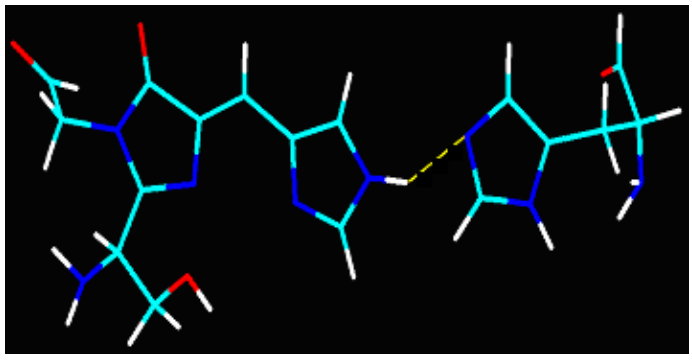

Figure 13. Hydrogen bond between chromophore of the BFP and nearby residue His148. Hydrogen bond is shown as dotted line.

imidazole is formed with its nitrogen atom located away from the rotation axis, which effectively prevents twisting (Fig. 13).

With temperature growth hydrogen bonds become weaker [16]. We can estimate the energy change of hydrogen bond in case of thermal expansion, using dependence of hydrogen bond energy on the distance between atoms [17].

The formula of volumetric expansion:

$$V = V_0 \cdot (1 + \alpha T), \tag{1}$$

where α – volumetric expansion coefficient (in this case $\alpha = 3 \cdot 10^{-4} K^{-1}$) [18, p.516]).

Linear dependence on temperature can be expressed as

$$l = l_0 \cdot (1 + \beta T), \tag{2}$$

where $\frac{l}{l_0} = \sqrt[3]{\frac{V}{V_0}}$, $\beta = \frac{1}{3}\alpha$.

The length of hydrogen bond $l_0 = 3.1$ Å increases by Δl=0,031 Å, when temperature rises 100 K. In [17] the respective bond energy weakening by $\Delta E = 0{,}775 Kcal/mol = 3{,}2 kJ/mol$.

Thermal expansion mechanism itself cannot explain the 4.4-fold increase in population of hydrogen bonded species, as revealed in experiment.

Indeed, equilibrium constant between hydrogen bonded n_A and nonbonded chromophores n_B is given by expression:

$$K = e^{-\frac{\Delta G}{k_b \cdot T}} = \frac{n_B}{n_A}, \tag{3}$$

where ΔG is a change of Gibbs' free energy by hydrogen bond disruption. In this case we do not consider change of other parameters except energy, and because of that, we replace ΔG on energy of dissociation (it is known that energy of hydrogen bond for case N-H---N ΔE=19 *kJ/mol*). At room temperature number of hydrogen bonded chromophores is much less than hydrogen nonbonded, therefore a small change of dissociation energy 3,2 *kJ/mol* cannot lead to a large change of population n_A.

In our preliminary experiments, BFP was inserted into the matrix of trehalose (disaharid), which has glass transition temperature at 352 K. Below this temperature, it has very high viscosity. For this reason the positions of amino acids surrounding chromophore are fixed and transitions between different conformational states are impossible.

Temperature dependence of fluorescence intensity of BFP surrounded by trehalose is described by the Arrhenius equation. Since trehalose restricts conformational transitions of the cylindrical part of the protein, activation energy determined from Arrhenius equation ΔE=19 kJ/mol corresponds to the association (or dissociation) energy of hydrogen bond between chromophore and nearby amino acid residue – histidine (His148).

When the BFP is dissolved in mixture of glycerin and TRIS-buffer (maintains pH=8), large-amplitude oscillations of protein parts are possible and transitions between different conformations take place at room temperature. As a reflection of this situation we got a linear dependence of fluorescence on temperature instead of Arrhenius curve.

After breaking of chromophore's hydrogen bond the number degrees of freedom is increased and instead of oscillation along hydrogen bond new type of movement of five-membered ring appears – two dimensional movement within a half-sphere (lever on a joint). Probability of hydrogen bond formation sharply decreases. Moreover, the twisting of histidine rings (one belonging to chromophore, the other to residue His 148), which were released from hydrogen bonding, appears. The density of states in a given energy interval increases, bringing along the growth of entropy, according to well known Boltzmann thermodynamic equation

$$S = k_b \cdot \ln \Omega , \tag{4}$$

where S is entropy, Ω is number of available states.

To take into account entropy change in the protein we will insert enthalpy instead of energy into Arrhenius equation:

$$\Delta H = \Delta G + T\Delta S, \tag{5}$$

where ΔG – Gibbs' free energy change, by transition from bound state to unbound, ΔH – change of enthalpy, ΔS – change of enthropy.

Since enthalpy determines difference of energies by transition from one conformation to another, consequently we will get Arrhenius equation in the form:

$$k = k_0 \cdot e^{-\frac{\Delta H}{k_b \cdot T}}, \tag{6}$$

where k is reaction rate, k_0 is attempt rate.

We denote B as a state without hydrogen bond, A- a bonded state, transition A→B – bond breaking, B→A – bond formation:

$$A \underset{k_{BA}}{\overset{k_{AB}}{\Longleftrightarrow}} B .$$

The whole amount of chromophores

$$n_0 = n_A + n_B . \tag{7}$$

Then relationship between reaction rate and enthalpy may be written as:

$$\frac{k_{AB}}{k_{BA}} = \frac{k_0^{AB}}{k_0^{BA}} \cdot e^{-\frac{H_B - H_A}{k_b \cdot T}} . \tag{8}$$

Time differentiating number of bound chromophores we get:

$$\dot{n}_A = k_{BA} \cdot n_B - k_{AB} \cdot n_A . \tag{9}$$

In equilibrium $\dot{n}_A = 0$, and it follows that

$$\frac{n_B}{n_A} = \frac{k_{AB}}{k_{BA}} = K , \tag{10}$$

where K – equilibrium constant.

Equilibrium constant can be calculated from equation (3), therefore taking into account (8) and (10) we can write

$$K = \frac{k_{AB}}{k_{BA}} = e^{-\frac{G_B - G_A}{k_b \cdot T}} = \frac{k_0^{AB}}{k_0^{BA}} \cdot e^{-\frac{H_B - H_A}{k_b \cdot T}} . \tag{11}$$

Using equations (4), (5), (11), we get relationship for preexponential factors for transitions A→B and B→A

$$\frac{k_0^{AB}}{k_0^{BA}} = e^{\frac{S_B - S_A}{k_b}} = \frac{\Omega_B}{\Omega_A} \tag{12}$$

and get the expression for the ratio of reaction rates using entropy and enthalpy:

$$\frac{k_{AB}}{k_{BA}} = e^{\frac{S_B - S_A}{k_b}} \cdot e^{-\frac{H_B - H_A}{k_b \cdot T}} . \tag{13}$$

Entropy factor in (13) is $\approx 10^6$ [19], and that explains 4.4-fold change of population between hydrogen bonded and nonbonded chromophores.

Relationship between hydrogen bonded chromophores and its total amount we get from equations (7) and (10)

$$\frac{n_A}{n_0} = \frac{1}{1 + \frac{k_{AB}}{k_{BA}}} \tag{14}$$

or, taking into account (11):

$$\frac{n_A}{n_0} = \frac{1}{1 + \frac{k_0^{AB} \cdot e^{-\frac{H_2}{k_b \cdot T}}}{k_0^{BA} \cdot e^{-\frac{H_1}{k_b \cdot T}}}} = \frac{1}{1 + \frac{k_{AB}}{k_{BA}} \cdot e^{-\frac{\Delta H}{k_b \cdot T}}} \tag{15}$$

5. COMPUTATIONAL CHEMISTRY OF GFP

The aforementioned hypotheses that the chromophore with no hydrogen bonds to surrounding residues may be a state with large value of entropy can be checked by computer simulations. Protein dynamics can be treated as a transition between minima of the potential energy landscape. Since proteins have a high number of degrees of freedom, the potential energy landscape is highly multi-dimensional. It is possible to investigate the transitions between minima by molecular dynamics (MD) methods. But this approach is limited, because many of the processes occurring in proteins have time scales inaccessible to such MD calculations, since current computers can model a protein for only a few nanoseconds. Another approach to get insight into the local dynamics of the protein is to calculate vibrations in harmonic approximation and to get detailed information about fluctuations of every single atom in the protein.

Accessible computational powers allows the calculation of normal modes of GFP only by molecular mechanics (MM) methods and are not sufficient for full quantum mechanical modeling. On the other hand, semi-empirical method AM1 is sufficiently accurate and efficient for calculation of the chromophore.

Calculation of the thermodynamic parameters needs the normal mode analysis of the protein, which requires at least 4 Gbyte of RAM. This calculation for GFP with hydrogen bonded chromophore has been performed on a 64-bit computer by Dr. Farkas in the Eötvös Lorand University in Budapest. The input file for Gaussian 03 was successfully provided by us.

5.1. Torsion potential of a phenol ring in chromophore of GFP

In physiological processes, protein functioning is based on the transition between different conformational states, which is usually accompanied by rearrangement of hydrogen bond network.

It is known [1, 6] that denaturized protein does not fluoresce; synthesized chromophore *in vitro* is also nonfluorescent. As we have mentioned above, the most probable mechanism of chromophore photoisomerisation is simultaneous twisting around two central bonds by 180°. In this case the energy of electronic excitation is transformed into twisting energy, but not into fluorescence. In native GFP,

chromophore is bound with nearest amino acid residues and therefore quantum yield of fluorescence is high (0.8). On the other hand the fluorescence of GFP mutants may be weak or absent at all if chromophore has enough freedom for movement.

Experiments on single GFP molecules, performed in professor Moerner's laboratory in Stanford University, have shown that excited by continuous laser beam molecule of GFP blinked [20]. The fluorescense lasts several seconds, followed by nonfluorescent (dark) state of similar duration. Our experiments with the bulk GFP sample have shown that dark nonfluorescent state is a neutral form of chromophore. Its spectrum is shifted into the short-wavelength direction and cannot be excited at 488 nm. Excitation of the anionic form of chromophore by green light causes its protonization.

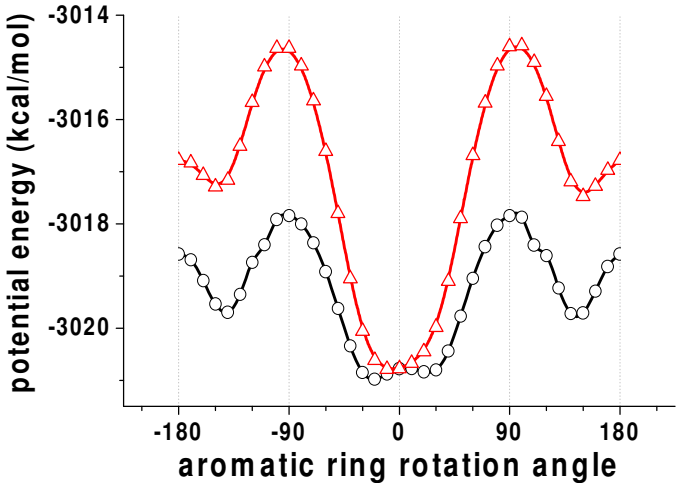

Figure 14. Dependence of potential energy on twisting angle of aromatic ring of GFP chromophore

Using a semi-empirical approximation method AM1 of quantum-mechanic calculation program HYPERCHEM we have calculated potential barriers related to twisting. As we can see from the diagram (Fig. 14), the twisting potential barrier of hydrogen bonded chromophore is two times larger than the barrier of hydrogen nonbonded chromophore. Also, the shape of the potential at 0 degrees changes from two-minimum for hydrogen nonbonded chromophore to one-minimum for hydrogen bonded species. This corresponds to increase of frequency of hindered rotation and causes energetically unfavorable decrease of entropy. Enthalpy of hydrogen bonds compensates it partially; therefore Gibbs' free energy does not change noticeably. In terms of our calculations it is possible to say that non-fluorescent GFP state found in experiment with single molecules is characterized by absence of hydrogen bonds with chromophore and may be a state with large value of entropy (entropy trap).

In order to revive the fluorescent state, hydrogen bonds have to be reestablished, and the entropy must decrease correspondingly. But processes, which decrease entropy, have a low probability. That explains the long lifetime of nonfluorescent GFP state.

5.2. Vibrations-determined properties of Green Fluorescent Protein

A prerequisite to understand protein function is knowledge of the driving forces which induce transitions between protein conformations. Proteins are flexible molecules where these transitions are controlled not only by enthalpic factors, but also by entropy. These two parameters determine the Gibbs' free energy, which governs the equilibrium constants as it was outlined in part 4. The main part of Gibbs' free energy comes from vibrations. Quantum chemical calculations can be used to determine the vibrational frequencies (normal modes) of the protein molecule, but for big molecules this is computationally prohibitively expensive. Therefore the force field methods are appropriate. We used the AMBER force field which has been designed for biological macromolecules [21]. The parameters for the chromophore were adopted from Reuter et al [22]. They used the CHARMM program where the energy is calculated analogously with the AMBER force field. The energy consists from six terms: chemical bonds, bond angles, bond twisting, Van der Waals, electrostatic and hydrogen bonds:

$$E_{total} = \sum_{bonds} K_r \left(r_i - r_{eq}\right)^2 + \sum_{angles} K_\theta \left(\theta_i - \theta_{eq}\right)^2 +$$

$$+ \sum_{dihedrals} \frac{V_n}{2}\left[1 + \cos\left(n\phi - \gamma\right)\right] + \sum_{\substack{nonbond \\ \_VDW}} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right] + \qquad (16)$$

$$\sum_{\substack{nonbond \\ \_el.}} \frac{q_i q_j}{\varepsilon R_{ij}} + \sum_{H-bonds} \left[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}}\right],$$

where K_r, K_θ, V_n, A_{ij}, B_{ij}, C_{ij}, D_{ij} are constants, r_i is the length of the bond and r_{eq} is the equilibrium distance, θ_i is the bond angle value, θ_{eq} is the

equilibrium angle value, γ and ϕ are the dihedral angles, n is the periodicity, R_{ij} is distance between the two atoms, $q_i q_j$ are the charges of the atoms, ε is effective dielectric constant.

The GAUSSIAN'03 program was used for vibrational frequency calculations. The coordinates of atoms were taken from PDB database file 1emb.pdb. Geometry of the molecule was optimized with ultratight option to avoid the imaginary eigenvalues.

The GFP contains 226 amino acids with 3799 atoms totaling 11391 normal modes. Vibrational modes were calculated in harmonic approximation and hindered rotations were treated as harmonic oscillators. The torsional barriers are high in comparison with room temperature k_bT and the discrepancy introduced by harmonic approximation is not big.

5.2.1. Distribution of normal modes of GFP molecule

To obtain a smooth spectrum of vibrations, normal modes were sorted into

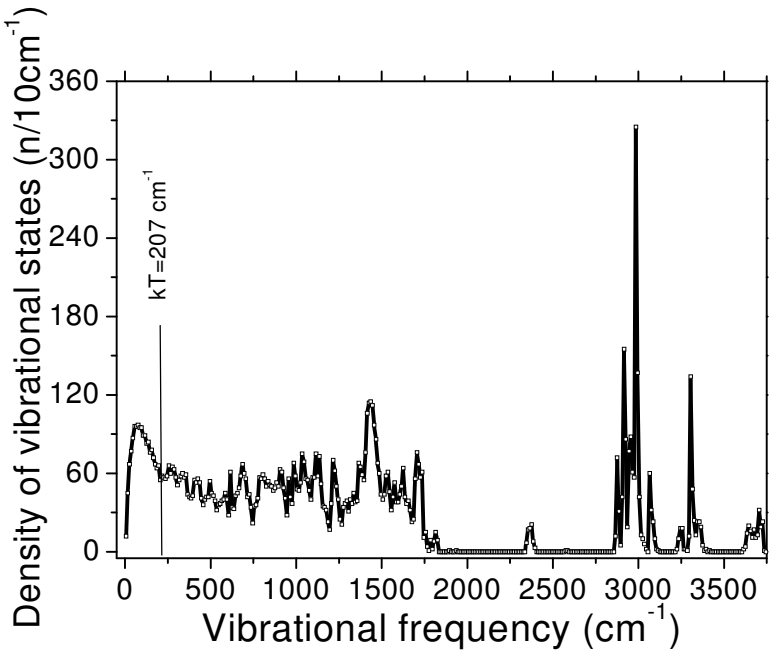

Figure 15. Distribution of 11391 normal modes of GFP molecule. Number of normal modes in each interval of 10 cm^{-1} is shown.

bins by 10 cm^{-1}. Low frequency maximum of the distribution of vibrational states is situated at 75 cm^{-1} (Fig. 15) whereas between 250 and 1250 cm^{-1} the density is almost flat.

The fingerprint region of vibrations has two peaks at 1430 and 1700 cm^{-1}. An amide band of protein backbone has three groups of vibrations in the range of 1300-1650 cm^{-1}. There are almost no vibrations in the range of 1800-2800 cm^{-1}. The peaks between 2800 and 3750 cm^{-1} belong to hydrogen vibrations: C-H modes have frequency of 3000-3100 cm^{-1} whereas O-H is situated 3400-3500 cm^{-1}.

5.2.2. RMS fluctuations

The same method of molecular mechanics has been applied to proteins for calculating normal vibrational modes such as the hinge-bending mode, which is crucial for the enzymes to operate [23]

The RMS (Root Mean Square) fluctuations of atoms are described by [24]:

$$\left\langle \Delta R^2 \right\rangle_{har} = \frac{k_b T}{m} \sum_{j=1}^{3N-6} \frac{\left|\vec{a}_j\right|^2}{\omega_j^2}, \tag{17}$$

where k_b is the Boltzmann constant, T is the absolute temperature, N is the number of atoms, ω_j is the frequency of the j-th normal mode, and $\vec{a}_j$ is the vector of the projections of the j-th normal mode onto the Cartesian components of the vector for the atom of interest.

The experimental fluctuations can be determined from the X-ray structure temperature factors B_i in Protein Data Bank file according to [25]

$$\left\langle u_i^2 \right\rangle_{\exp}^{1/2} = \left(3B_i / 8\pi^2 \right)^{1/2}. \tag{18}$$

The RMS fluctuations of the backbone C_α atoms are depicted in the Fig. 16. The agreement between the experimental and calculated values is fair. Almost every peak in the experimental data has a counterpart on the calculated curve, although the magnitude of the calculated fluctuations is relatively smaller.

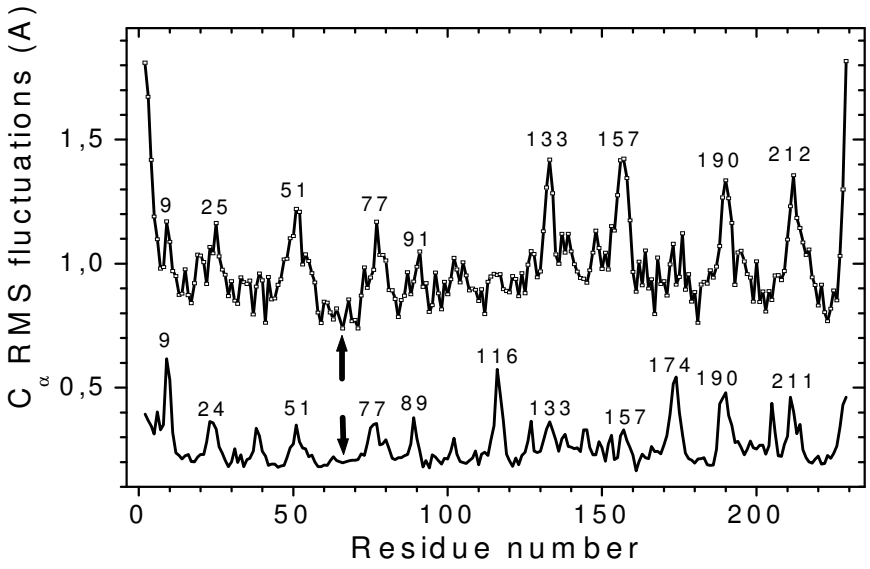

Figure 16. Root mean square fluctuations of backbone C_α atoms. Upper curve: experimental displacements according to temperature factors in the X-ray structure file 1emb.pdb. Lower curve: fluctuations calculated from eigenvectors of normal modes. The residue numbers are indicated, the arrows show the location of the chromophore.

The average of C_α fluctuations is 0.986 Å for the X-ray data and 0.266 Å for the normal mode data.

The most flexible residues are located in the beta hairpins which are shown in the Figure 17. This is to be expected, because the side chains of these residues point radially outward from the protein, preventing stiffening due to binding. In contrast, the residues involved in the beta sheets of the GFP cylinder

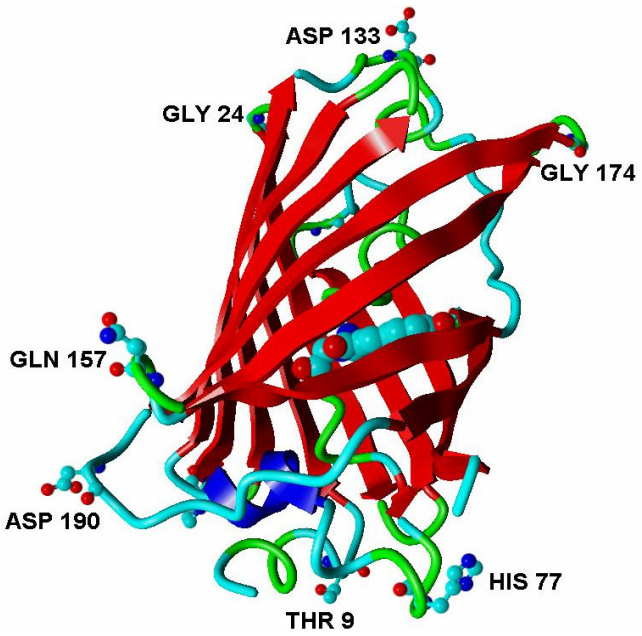

Figure 17. The backbone of Green fluorescent protein. The most fluctuating residues are labelled and shown as ball-and-stick models.

have small fluctuations due to tight hydrogen-bonding between the strands of beta-can. The chromophore and the alpha helix carrying it are well- fastened to the surrounding amino acids and show considerable rigidity. The rigidity of the chromophore is crucial for it to emit luminescence, which is directly related to GFP's function as a fluorescent protein in jellyfish.

5.2.3. Thermodynamic parameters

All main thermodynamic parameters heat capacity C_p, Gibbs' free energy G, enthalpy H and entropy S were calculated from the normal mode frequencies ν_i . It can be expressed as [26]:

$$C_{p\_vibr} = k_b \sum_{i=1}^{3N-6} \frac{\left(\frac{h\nu_i}{k_b T}\right)^2}{2\left(ch\frac{h\nu_i}{k_b T} - 1\right)}, \tag{19}$$

$$G_{vibr} = k_b T \sum \ln\left(1 - e^{-\frac{h\nu_i}{k_b T}}\right), \tag{20}$$

$$S_{vibr} = k_b \sum \left\{ -\ln\left(1 - e^{-\frac{h\nu_i}{k_b T}}\right) + \frac{\frac{h\nu_i}{k_b T} e^{-\frac{h\nu_i}{k_b T}}}{1 - e^{-\frac{h\nu_i}{k_b T}}} \right\}, \tag{21}$$

$$H_{vibr} = k_b T \sum \frac{\frac{h\nu_i}{k_b T} e^{-\frac{h\nu_i}{k_b T}}}{1 - e^{-\frac{h\nu_i}{k_b T}}}. \tag{22}$$

As we can see from the Fig. 15 the majority of the vibrational frequencies have energy larger than the thermal energy at room temperature. By virtue of the Boltzman factor, the vibrational levels above $\nu = 0$ are populated only for lower

frequency modes, which give the main part of the heat capacity, entropy and enthalpy. In Fig. 18 the heat capacities of individual vibrational modes at room temperature are depicted. Modes with frequencies < 375 cm^{-1} contribute half of the total heat capacity while those with frequencies > 1500 cm^{-1} give only 1.5%.

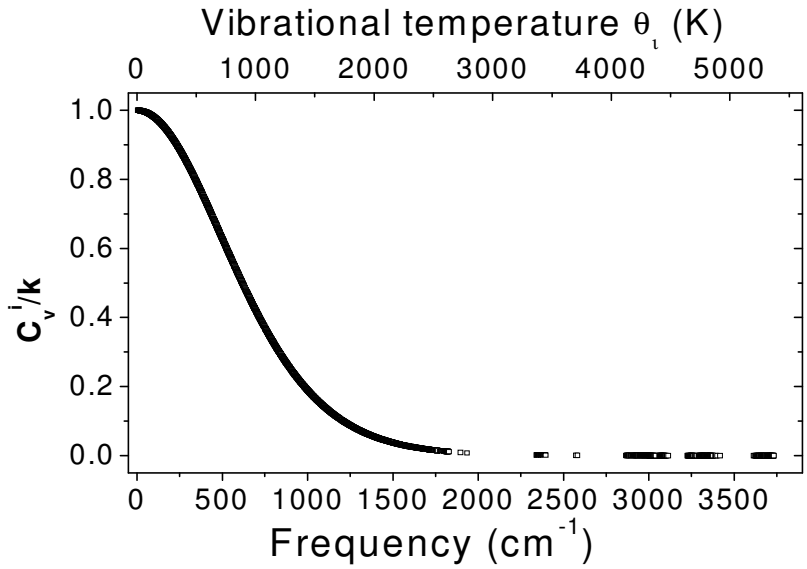

Figure 18. Contribution of vibrational modes to GFP heat capacity at room temperature.

The temperature dependence of heat capacity (Fig.19) is almost linear starting from 100 K. The experimental data of Di Lorenzo et al [27] have similar shape, although around 300 K their heat capacity curve has slight concavity while our calculation results in convex form.

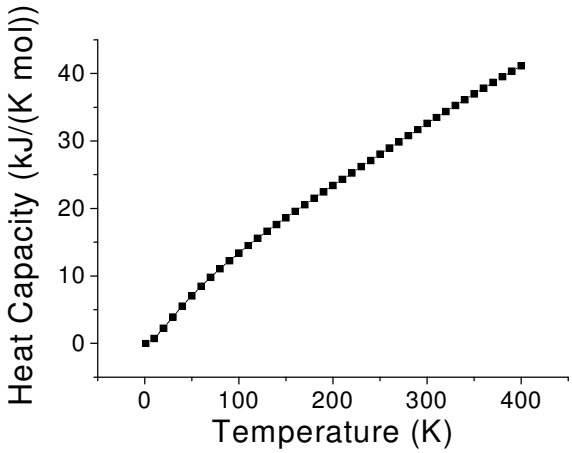

Figure 19. The dependence of heat capacity on temperature

The vibrational spectrum gives the opportunity to calculate all thermodynamic parameters, which have

prime importance for understanding of protein thermal stability. Gibbs' free energy is the state function which governs the transitions between protein conformational states.

At the low temperature $T \leq 10$ K the capacity is perfectly fitted with parabola T^2 (Fig. 19a). That is consistent with the linear dependence of the density of states on frequency. Indeed, the detailed inspection of the normal mode distribution of 1 cm^{-1} resolution (data not shown) reveals quasilinear behavior in the frequency range.

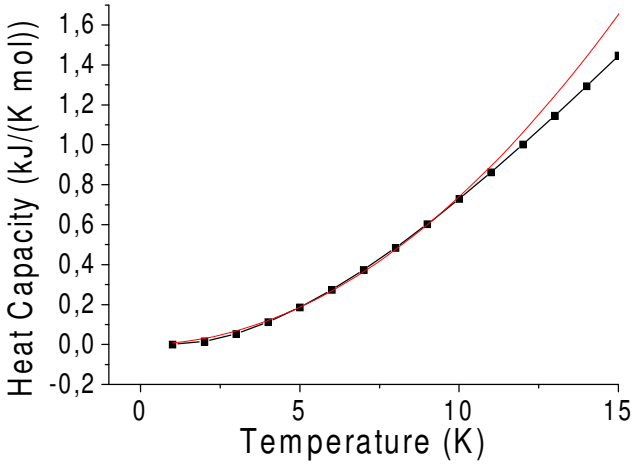

Figure 19a. The capacity is perfectly fitted with parabola

As seen from the Fig. 20, the entropy component of G and enthalpy both grow with temperature. Beneficial for stability entropic component increases steeper than

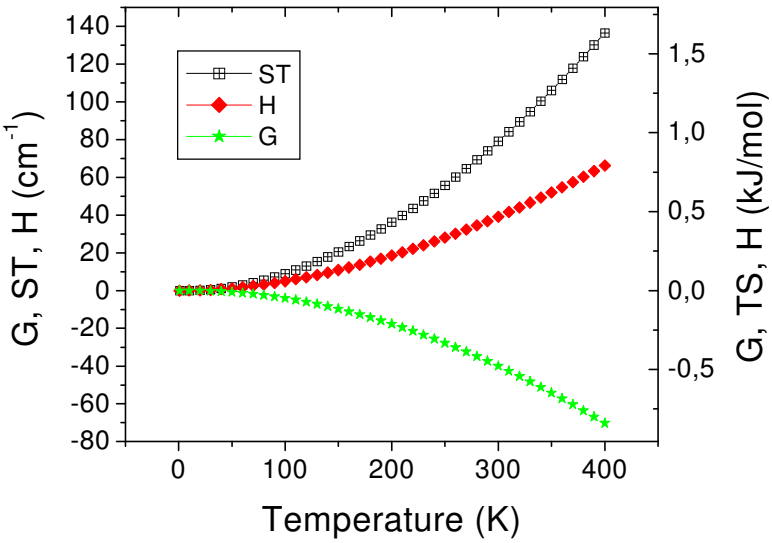

Figure 20. Temperature dependence of thermodynamic parameters of GFP (divided by number of vibrations)

enthalpic, although the exponents are similar ($ST \sim T^{1.92}$ and $H \sim T^{1.84}$). The graph for G decreases monotonically indicating the stabilization of protein at room temperature in respect of cryogenic temperatures. It is well known that proteins denature at low temperatures (at $T<273$ K) but also at high temperatures. Our graph reveals the entropic driving force of the former process whereas it does not describe the denaturation on heating, since at higher temperatures the unharmonicity of vibrations should be taken into account.

SUMMARY

In the present work we have investigated the physicochemical characteristics of Green Fluorescent Protein and its mutant which are important for its functioning as a fluorescent marker and for its thermal stability.

The dependence of potential energy on torsional deformation of the GFP's chromophore has been calculated by semiempirical computational chemistry method AM1.

The experimental data and computer calculations allow asserting that the presence of hydrogen bonds between the chromophore and its surrounding prevent its twisting. Also, hydrogen bonds with chromophore are necessary for preventing radiationless transitions in GFP-like proteins. Our calculations and experiment allow to conclude that nonfluorescent GFP state found in single molecule experiment in Stanford University is characterized by absence of hydrogen bonds with chromophore and it may be a state with large value of entropy (entropy trap). For restoring of the fluorescent state hydrogen bonds have to be reestablished and entropy correspondingly must decrease. It may explain long lifetime of nonfluorescent GFP state, because processes with decreasing of entropy have low probability.

The spectra of vibrational normal modes were determined in the AMBER force field approximation by GAUSSIAN'03 program. RMS fluctuations of the backbone were computed. We also calculated the thermodynamic parameters of isolated GFP – heat capacity, entropy, enthalpy and Gibbs' free energy and their temperature dependencies. The Gibbs' free energy increases on temperature lowering. This hints to destabilization of GFP on temperature lowering and is in accord with cold denaturation of isolated protein.

REFERENCES:

1. Tsien Roger Y, The Green Fluorescent protein, *Annu. Rev. Biochem*. 1998. **67**, 509-544
2. Labas J., Gordejeva A, Fradkov A. Fluorestsirujustsije i cvetnyje belki. *Priroda*, 2003, 3.
3. Stryer, L. *Annu. Rev. Biochem*. 47, 819-846
4. Lakowicz, JR. *Principles of Fluorescence Spectroscopy*. New York: Plenum, 1983.
5. Tsien RY, Basckai BJ, Adams SR. *Trends Cell Biol*. 1993, 3, 242-245
6. Zubova N, Bulavina A, Savitski A. Spektralnyje i fiziko-himitseskije svoistva zelenogo (GFP) i krasnogo (drFP583) fluorestsirujustsih belkov. *Uspehi biologicheskoj himii*, T. 43, 2003, 163-224
7. Shimomura O, Johnson FH, Saiga Y. *J. Cell. Comp. Physiol.* 1962, **59**:223-239
8. Yang F, Moss LG, Phillips GN Jr. *Nat. Biotehnol*. 1996, **14**, 1246-1251
9. Phillips GN Jr. *Curr. Opin. Struct*. Biol. 1997, **7**, 821-827
10. Dopf J, Horiagon TM. *Gene.* 1996, **173**, 39-44
11. Heim R, Prasher DC, Tsien RY. *Proc. Natl. Acad. Sci. USA* 1994, **91**, 12501-12504
12. Cubitt AB, Heim R, Adams SR, Boyd AE, Gross LA, Tsien RY. 1995, *Trends Biochem. Sci.* **20,** 448-455
13. Reid BG, Flynn GC. *Biochemistry*, 1997, **36**, 6786-6791
14. Weber W, Helms V, McCammon JA, Langhoff PW. Shedding light on the dark and weakly fluorescent states of green fluorescent proteins. *Proc. Natl. Acad. Sci. USA*. May 1999, **96**, 6177-6182
15. Krasnenko, V. Bakalaurusetöö, GFP mutandi kromofoori fluorestseeruvad ja mittefluorestseeruvad vormid, Tartu 2002.
16. Czeslik C, Jonas J.Pressure and temperature dependence of hydrogen-bond strength in methanol clusters. *Chemical Physics Letters* 1999, **302**, 633-638
17. Lippincott ER and Schroeder. *The Journal of Chemical Physics*, June, 1955, **23**, Number 6

18. Kharakoz DP. *Biophysical Journal*, July, 2000, **79**, 511-525
19. Frauenfelder H and Wolynes PG. Rate Theories and Puzzles of Hemeprotein Kinetics. *Science*, July 1985, **229**, Number 4711
20. Dickson R, Cubbit A, Tsien R, Moerner W.On/off blinking and switching behaviour of single molecules of green fluorescent protein, *Nature*, 24 July 1997, **388**, 355-358.The investigation of the effects of counterions in protein dynamics simulations
21. Kollman P, Dixon R, Cornell W, Fox T, Chipot C, Pohorille A. In *Computer Simulation of Biomolecular Systems*; Editor: Wilkinson A.; Weiner P.; Van Gunsteren W.; Elsevier 1997, Vol. 3 A, pp 83-96.
22. Reuter N, Lin H, Thiel W. *J Phys Chem* B 2002, **106**, 6310-6321
23. Brooks B, Karplus M. *Proc Natl Acad Sci USA* 1985, **82**, 4995-4999
24. Brooks B, Karplus M. *Proc Natl Acad Sci USA* 1983, **80**, 6571-6575
25. Dauber-Osguthorpe P, Osguthorpe D.J, Stern P.S, Moult J. *Journal of Computational Physics* 1999, **151**, 169-189
26. Volkenstein M, Eljashevits M, Stepanov B. Kolebanija molekul. Moskva, 1949, T.2
27. Di Lorenzo M, Zhang G, Pyda M, Lebedev B, Wunderlich B. Heat capacity of solid-state biopolymers by thermal analysis. *Journal of Polymer Science*. Part B: Polymer Physics, 1999, **37**, 2093-2102

KOKKUVÕTE

Rohelise ja sinise fluorestseeruva proteiini fotofüüsika

Käesolevas töös me uurisime rohelise fluorestseeruva proteiini (GFP) ja ta mutandi füüsikalis-keemilisi karakteristikuid, mis määravad GFP termilist stabiilsust ja funktsioneerimist fluorestseeruva markerina.

GFP kromofoori potentsiaalse energia sõltuvus torsioondeformatsioonist arvutati poolempiirilise meetodi AM1 abil.

Eksperimentaal-andmed ja arvutused lubavad järeldada, et vesiniksidemete olemasolu kromofoori ja teda ümbritsevate aminohapete vahel takistavad väände deformatsioone. Vesiniksidemed on ka vajalikud, et hoida ära fluorestsentsi kustutamist GFP-sarnastes proteiinides. Meie poolt teostatud eksperiment ja arvutused võimaldavad järeldada, et Stanfordi ülikoolis ainumolekuli eksperimendis avastatud GFP mittefluorestseeruvat pime olekut põhjustab kromofoori vesiniksidemete katkemine fotoergastusel ja nimetatud pimeolekut võib kirjeldada kõrge entroopiaga seisundina (entroopialõksuna). Fluorestseeruva seisundi taastumiseks peavad vesinikusidemed taastekkima ja vastavalt sellele peab entroopia vähenema. See võib seletada GFP mittefluorestseeruva oleku pikka eluiga, sest protsessid, kus energia väheneb, omavad madalat tõenäosust.

GFP võnkumiste sagedused arvutasime molekulaar-mehhaanika lähenduses kasutades GAUSSIAN'03 programmi ning jõuvälja AMBER. Arvutasime GFP peaahela C_α ruutkeskmised hälbed. Määrasime isoleeritud GFP jaoks termodünaamilisi parameetreid – soojusmahtuvust, entroopiat, entalpiat ja Gibbsi vaba energiat ja nende temperatuuri sõltuvusi. Leidsime, et temperatuuri alandamisega Gibbsi energia kasvab. See viitab GFP destabiliseerumisele temperatuuri alandamisel ja on kooskõlas isoleeritud proteiini denaturatsiooniga külmutamisel.

Tänuavaldused

Soovin tänada oma juhendajat Dr. K. Mauringut. Teoreetilist küsimuse lahendamisel on olnud suureks abiks prof. V. Hiznjakov. Tänan Dr. Farkast (Eötvös Lorand Ülikool - Budapest) tehnilise abi ja nõuannete eest.

Uurmistööd on toetanud Eesti Teadusfond grandiga number 5546.

Veel tänan sõpru ja meeldivat laser ja spektroskoopia labori kollektiivi moraalse toetuse eest.