Physicochemical Properties of Blue Fluorescent Protein Determined via Molecular Dynamics Simulation

Veera Krasnenko,1 Alan H. Tkaczyk,1,2 Eric R. Tkaczyk,3,4 Koit Mauring1
1 Institute of Physics, University of Tartu, Riia 142, 51014 Tartu, Estonia
2 Department of Biomedical Engineering, University of Michigan, 2200 Bonisteel Blvd., Ann Arbor, MI 48109-2099
3 Center for Ultrafast Optical Science, University of Michigan, 2200 Bonisteel Blvd., Ann Arbor, MI 48109-2099
4 Michigan Nanotechnology Institute for Medicine and Biological Sciences, 9220 MSRB III, 1150 West Medical Center Drive, SPC 5648, Ann Arbor, MI 48109-5408

Received 28 May 2008; revised 25 July 2008; accepted 25 July 2008
Published online 8 August 2008 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.21065

ABSTRACT:
Blue fluorescent protein (BFP) is a mutant of green fluorescent protein (GFP), where the chromophore has been modified to shift the emitted fluorescence into the blue spectral region. In this study, MD calculations were performed with the GROMACS simulation package and AMBER force field to investigate the dependence of BFPs physicochemical properties on temperature and applied pressure. The MD approach enabled us to calculate the compressibility of protein itself, separately from the nontrivial contribution of the hydration shell, which is difficult to achieve experimentally. The computed compressibility of BFP (3.94 × 10⁻⁵ MPa⁻¹) is in agreement with experimental values of globular proteins. The center-of-mass diffusion coefficient of BFP and its dependence on temperature and pressure, which plays an important role in its application as a probe for intracellular liquid viscosity measurement, was calculated and found to be in good agreement with photobleaching recovery experimental data. We have shown that decreased temperature as well as applied pressure increases the water viscosity, but the concomitant decrease of the BFP diffusion coefficient behaves differently from Stokes-Einstein formula. It is shown that the number of hydrogen bonds around the protein grows with pressure, which explains the aforementioned deviation. Pressure also reduces root mean square (RMS) fluctuations, especially those of the most flexible residues situated in the loops. The analysis of the RMS fluctuations of the backbone Cα atoms also reveals that the most rigid part of BFP is the center of the β-barrel, in accord with temperature B factors obtained from the Protein Data Bank. © 2008 Wiley Periodicals, Inc. Biopolymers 89: 1136–1143, 2008.

Keywords: BFP; MD; hydrogen bonding; compressibility; diffusion coefficient; RMS fluctuations; temperature; pressure; computational chemistry; molecular modeling

INTRODUCTION

Fluorescent proteins have become one of the most widely studied and exploited proteins in cell biology and biochemistry. This popularity owes partly to the existence of multiple natural and mutated forms with widely varying spectral properties. One of these
mutated proteins is blue fluorescent protein (BFP), where the tyrosine ring of the chromophore of wild-type green fluorescent protein (GFP) has been replaced by the histidine ring. BFP contains 227 amino acids, which form an 11-stranded β-barrel wrapped around a central helix. The cylindrical barrel forms a strong skeleton, which protects the chromophore and allows the protein to keep its form even under high applied pressures.

BFP has been used as donor molecule coupled with GFP for fluorescence energy transfer (FRET) imaging. Because the blue fluorescence of BFP can be easily distinguished visually from the emission of other GFP mutants, it can be used for multicolor marking of gene expression or for protein targeting in intact cells. Some protein variants exhibit better quantum yield and photostability compared with BFP, and these brighter and more stable protein variants are now most commonly used in FRET.

The experimental pressure- and temperature-dependence of BFP fluorescence intensity has shown that thermodynamic parameters strongly influence BFP quantum yield. These results were interpreted as an equilibrium shift between protein conformations of different fluorescence quantum yield due to hydrogen bonding and excited state isomerization. Recent studies have shown that the quantum yield of BFP can also be substantially enhanced by mutations, which insert bulkier amino acids surrounding the chromophore, restricting its conformational freedom and the isomerization probability in the excited state.

Previous MD simulations of GFP by McCammon and coworkers (internal dynamics of wild-type GFP), Reuter et al. (protein MD and qualitative analysis of the parameters of the neutral and deprotonated forms of the GFP chromophore), and Nifosi et al. (protein MD and effects of mutation in the environment of GFP chromophore) have shown the importance of the hydrogen bond network around the chromophore and the rigidity of the β-barrel structure for protection of the chromophore neighborhood against solvent molecules.

In this investigation, we calculate the number and lifetime of the intraprotein and protein–water hydrogen bonds, which influence the dynamics of BFP. The dynamics, in turn, are manifested as fluctuations of the coordinates of protein atoms. These calculated root mean square (RMS) fluctuations are compared with experimental temperature factors from X-ray structural study. We also study the pressure dependence of the protein molecule dimensions, because no experimental data exist. We elucidate the effect of two thermodynamic parameters (temperature and pressure) on the diffusion coefficient of BFP. Several variants of GFP, such as cyan, yellow, and gold fluorescent protein, are very similar to BFP, and therefore the results obtained for BFP can be useful for describing diffusion of these mutants as well. These GFP derivatives are used as probes for measuring the physicochemical properties of the surrounding medium, including the densely concentrated cell environment.

**METHODS**

The program GROMACS 3.3 was used for MD simulation of BFP in a water box of TIP3P water models.

For calculation of the potential function, the molecular mechanics force field AMBER96 was used. The potential energy of the system is defined as the sum of six contributions: chemical bonds, bond angles, bond twisting, van der Waals, electrostatic, and hydrogen bonds. For simulations, the crystal structure was immersed in an octahedral box of water, with a buffer of 1.1 nm between the protein atoms and the edge of the box. This resulted in a system with 13,307 water molecules. The protein geometry was optimized to obtain the minimum of energy. Next, position-restrained MD (restraining the atom positions of the macromolecule while letting the solvent move in the simulation) was carried out to “soak” the protein for 20 ps. Then, the simulation of 3 ns MD followed. A separately computed 10 ns trajectory at 300 K proved to yield very similar results (the RMS fluctuations, BFP center-of-mass diffusion coefficient, and average volume of the protein were compared); therefore, 3 ns trajectories were used. The calculations used a 2-fs integration time step and coordinates were written to the output file once every 100 frames. The pair list of neighbors, for which non-bonding interactions were calculated, was updated every 10 steps. The cutoff distances for Lennard-Jones and electrostatic interactions were 1 nm. Long-range electrostatic effects were described with a PME procedure using default parameters. The Berendsen thermostat for temperature coupling with time constant 0.1 ps and the Berendsen exponential relaxation pressure coupling with time constant 0.5 ps were applied. The calculation of 3-ns MD trajectory lasted 34 h on a Dell node of 4 Dual Core AMD Opteron PowerEdge 6950 cluster and 77 h on the Fujitsu Siemens two-Opteron workstation Celsius V810.

The atomic coordinates of BFP were obtained from the Protein Data Bank file 1BFP, and the hydrogen atoms were added using HyperChem 7.5. The BFP contains a chromophore (see Figure 1), which is formed from three amino acids (Ser, His, Gly) during protein folding. The charges of the chromophore atoms (Table 1) were calculated using the HyperChem program.
calculated with Gaussian 03 \textsuperscript{20} using the density functional method up to the B3LYP/6-31G(d) level. The force field parameters for the imidazole ring and propene bridge of the chromophore were adopted from Reuter et al.,\textsuperscript{9} whereas the parameters for histidine ring were adopted from the AMBER force field.

The volume of the protein for different pressures was calculated with the Swiss PDB application,\textsuperscript{21} as the space inside the surface that can be reached with the surface of a water molecule (radius 0.14 nm) rolled over the protein. Protein structures were averaged via the VMD visualization program\textsuperscript{22} over the last 250 frames in a 1-ns trajectory.

The isothermal compressibility $\beta$ of the protein is defined as

$$\beta \equiv - \frac{1}{V} \left( \frac{\partial V}{\partial P} \right),$$

where $V$ is the volume of the protein and $P$ is the pressure. VMD was also used for measuring the diameters $d$ and heights $h$ of the averaged protein barrel-like structure. Each value was calculated as an average of three individual distances between the $C_a$ atoms of opposite backbone strands.

**Diffusion Coefficient**

To determine the self-diffusion coefficient $D_w$ of water, the Einstein relation

$$D_w = \lim_{t \to \infty} \frac{1}{6t} \langle ||r_i(t) - r_i(0)||^2 \rangle$$

was used, where $\langle ||r_i(t) - r_i(0)||^2 \rangle$ represents the mean square displacement (MSD) averaged over the water molecules. $D$ was calculated by GROMACS program g_msd. The viscosity of water was calculated from the self-diffusion coefficient according to the Stokes-Einstein formula

$$\eta = \frac{kT}{6\pi d D}$$

where $k$ is Boltzmann constant, $T$ is temperature, $D$ is diffusion coefficient (calculated according to Eq. (2)), and $d$ is the molecule diameter (for water $d = 0.275$ nm).

A Tcl script (in the VMD program environment) was written to calculate the protein’s center-of-mass diffusion coefficient.

**Hydrogen Bonds**

The average number and lifetime of intraprotein and protein–water hydrogen bonds (H-bonds) were computed by the GROMACS utility g_hbond with cutoff distance 0.35 nm and cutoff angle 30°. The H-bonds lifetimes were calculated from the average over all autocorrelation functions of the bond existence functions $s_i$ of all H-bonds:

$$C(\tau) = \langle s_i(t)s_i(t + \tau) \rangle$$

with $s_i(t) = [0, 1]$ for H-bond $i$ at time $t$. The integral of $C(\tau)$ estimates the average H-bond lifetime $\tau_{HB}$:

$$\tau_{HB} = \int_0^\infty C(\tau)d\tau$$

**Root Mean Square Fluctuations**

The RMS fluctuations were computed by the GROMACS utility g_rmsf. The experimental fluctuations were calculated using the X-ray structure temperature factors $B_i$ of the Protein Data Bank file 1BFP according to Ref. 23:

$$\langle u_i^2 \rangle^{1/2} = \left( \frac{3B_i}{8\pi^2} \right)^{1/2}$$

**RESULTS AND DISCUSSION**

**Compressibility**

In Figure 2, the dependence of BFP diameter, height, and $V^{1/3}$ on applied pressure are depicted. It is notable that the
diameter of the protein cylindrical structure decreases more with applied pressure than does the height. The observed dependence is consistent with the protein’s barrel structure: the linear compressibility along strands is smaller than the linear compressibility perpendicular to the strands. Over the range from 0.1 to 1000 MPa applied pressure, the average compressibility of BFP is $3.9 \times 10^{-5} \text{ MPa}^{-1}$. The error of the calculated compressibility 8% is determined by the uncertainty of the linear regression slope of volume versus pressure data (see Figure 3). It is evident that compressibility does not change in this range of pressure, where the protein has enough free cavities to be compressed or filled with water. In general, at high pressures, the compressibility should start to diminish. The range of compressibility of globular proteins is $3–15 \times 10^{-9} \text{ MPa}^{-1}$, and our calculated value is near the lower limit, which is in accord with the classification of GFP variants in the category for proteins of rigid structure.

**Diffusion Coefficient**

To use BFP as a probe of the viscosity of surrounding medium, it is essential to understand how thermodynamic parameters affect the diffusion of BFP. Both pressure and temperature influence the viscosity of the medium as well as the dynamics of hydrogen bonds between protein and surrounding water.

**Dependence on Temperature.** The diffusion coefficients of water and protein were calculated from the MSDs. Figure 4 depicts the time evolution of MSD of all water molecules at different temperatures. $D$, the diffusion coefficient of water, is estimated from the slopes of these curves in the temperature range 200–375 K (see Figure 5). $D$ is shifted to higher values relative to published experimental data by $2.9 \times 10^{-9} \text{ m}^2/\text{s}$ on average. For example, at 300 K, the diffusion coefficient of water according to our calculations is $5.3 \times 10^{-9} \text{ m}^2/\text{s}$, whereas the experimental value is $2.3 \times 10^{-9} \text{ m}^2/\text{s}$. It is well known that MD is not capable of describing the freezing of
water; therefore, the calculated curve of the self-diffusion coefficient lacks the step-like decrease of $D$ at 273 K.

The viscosity of water at $T = 300$ K, calculated according to formula (3), is $3 \times 10^{-4}$ Pa s, whereas the experimental value is $8.9 \times 10^{-4}$ Pa s. It is known that the TIP3P water model underestimates the viscosity, but we chose it over the improved 4-site water model TIP4P-Ew, which is computationally more demanding.

The time dependence of the computed center-of-mass MSD for the protein also increases with temperature (see Figure 6). The computed center-of-mass diffusion coefficient of BFP in a water box at room temperature, 59 $\mu m^2/s$, compares favorably with the experimental value for GFP ($87 \pm 2 \mu m^2/s$).

The calculated temperature dependence of BFP’s diffusion coefficient (see Figure 7) can be fit with two linear segments at 150–200 K and 200–300 K (with correlation coefficients 0.99968 and 0.99996, respectively) that intersect at $\approx 205$ K, which is in the region of the glass transition temperature of proteins ($T_g$, confer). MD calculations and experimental data from infrared absorption as well as from neutron scattering confirm that the glass transition in the protein–water system is determined by dynamic transition in the water shell around the protein. This is in contrast to an absence of the glass transition at $\approx 200$ K for proteins in vacuum. At temperatures above the glass transition, the activation of solvent translational diffusion takes place, bringing along the radial softening of protein dynamics, primarily increasing the amplitude of side chain motion.

We also performed our simulations in different water box sizes with effective box lengths $V^{1/3}$: 7.23, 7.98, and 8.54 nm, where $V$ is box volume. The protein center-of-mass diffusion coefficient is independent of this: all obtained values of diffusion coefficient at 300 K were within $\pm 1 \mu m^2/s$.

Dependence on Pressure. For all pressures examined, the time dependence of the protein’s center-of-mass MSD remains perfectly linear (correlation coefficient 0.99997), with the slope decreasing with increasing pressure (see Figure 8). This is consistent with increased water viscosity and number of hydrogen bonds between protein and water. The pressure dependence of the center-of-mass diffusion coefficient $D(P)$ of BFP (see Figure 9) is weak at 1–200 MPa and 800–1000 MPa and is almost linear between 400 and 800 MPa. When the Stokes-Einstein formula (3) prerequisites are met, protein diffusion is inversely proportional to the viscosity of water $\eta_w$. Thus, the shape of dependences $D(P)$ for protein and $1/\eta_w(P)$ for water should be similar. As the diffusion coefficient of water $D_w(P)$ is proportional to $1/\eta_w(P)$, we compare the shapes of diffusion coefficients for BFP and water (see Figure 9). It is clear that $D(P)$ for the protein and $D_w(P)$ for water are not parallel. This implies that the interaction between the protein and solvent depends on pressure.
Indeed, the number of hydrogen bonds between protein and water molecules grows with applied pressure (see Figure 10) and is discussed later.

During the calculated 10-ns trajectory, we could observe the diffusion of several water molecules into and out of the protein, though they never reached the immediate neighborhood of the chromophore.

Hydrogen Bonds

One aspect of hydrogen bonding in all GFP-like molecules relates to the stability of the β-barrel, where β-strands are tied to each other by tight hydrogen bonds.

In Figure 10, the dependence of the average number of hydrogen bonds between protein and water on pressure is depicted. The number of hydrogen bonds was calculated for two criteria of hydrogen bond cutoff angle: 30° and 60° (insets in Figure 10). For the weaker hydrogen bonds (cutoff angle 60°), the effect of pressure is more pronounced. At high pressure (1000 MPa), the number of such suboptimal hydrogen bonds between BFP and water increases by 18.9%, whereas for optimal hydrogen bonds (cutoff angle 30°), the number increases only by 6.6%. This result compares favorably with the computational data of Paci.37

Our calculations show that temperature as well as pressure affects the number of hydrogen bonds and hydrogen bond lifetime. When temperature increases from 275 to 375 K, the number of BFP intraprotein hydrogen bonds was found to decrease by 2%, whereas those between water and protein decreased by 16%. Under hydrostatic pressure or decreased temperature, we did not find the formation of additional hydrogen bonds between the chromophore and its surrounding.

The hydrogen bond lifetime ($\tau_{HB}$) dependence on temperature is shown in Figure 11. The mean lifetime of the intraprotein hydrogen bonds and those between the protein and its water shell are found to be $\sim 2$ ps. The curves intersect at 200 K, which is in the region of the glass transition temperature of proteins. The decrease in lifetime between 275 and 375 K was 37% for intraprotein bonds and 54.5% for water–protein bonds. The slight temperature and pressure dependence of chromophore-to-surroundings hydrogen bond lifetime is on the same order of magnitude as the statistical fluctuations of $\tau_{HB}$, as observed when the simulation is repeated with a different set of initial coordinates.

RMS Fluctuations

The RMS fluctuations of the backbone $C_{\alpha}$ atoms are depicted in Figure 12. Almost every peak in the experimental data has a counterpart on the calculated curve, although the magnitude of the calculated fluctuations is bigger. The average of $C_{\alpha}$ fluctuations is 0.285 nm for the X-ray data and 0.185 nm for the calculated data. The most flexible residues are located in the loops connecting the β-strands (Figure 12A), which is

![Figure 9](image1.png)

**FIGURE 9** The diffusion coefficient of the BFP center-of-mass (solid line with stars, left scale) and of water (dash line with bullets, right scale) vs. applied pressure.

![Figure 10](image2.png)

**FIGURE 10** The average number of hydrogen bonds between protein and water with cutoff angle 60° (solid line) and 30° (dashed line–normalized to number of H-bonds with cutoff angle 60°).

![Figure 11](image3.png)

**FIGURE 11** Hydrogen bond lifetime dependence on temperature: intraprotein–solid line with stars, between protein and water–solid line with bullets.
to be expected, because the side chains of these residues point radially outward from the protein, preventing stiffening due to binding. The RMS fluctuations calculated by method of vibrational analysis \(^3\) show similar correlations to experimental data. Also, we can see the expected decrease in RMS fluctuations for almost all flexible parts of the protein at an applied pressure of 1000 MPa. Indeed, the pressure favors a compact structure where the BFP side chains have higher probability for hydrogen bonding. The most efficient increased rigidity of loops is accomplished by the cross-bonding (bridging) of atoms of the legs (branches) (ASP 133-LYP 140, CYS 48-GLY 51) or by connecting the neighboring residues (ASP 197-ASN 198).

CONCLUSIONS

In this article, we have studied how the diffusion coefficient of BFP depends on temperature and pressure. These data may assist the interpretation of the results of intracellular liquid viscosity measurements by GFP-like molecules as probes. The diffusion of GFP is determined by several factors including collisions with obstacles, intrinsic viscosity of the medium, or binding interactions. Our results allow for the differentiation of these factors, as the diffusion coefficient of sole GFP has been revealed. The computed value of the center-of-mass diffusion coefficient for BFP in a water box at room temperature is \( \tilde{5} \, \mu\text{m}^2/\text{s} \) and compares favorably well with the experimental value of \( \tilde{8} \, \mu\text{m}^2/\text{s} \) for GFP in water solution. The temperature dependence of the computed diffusion coefficient of water at room temperature is in accordance with experimental data. However, the value of the water diffusion coefficient in our calculation is shifted by \( 3 \times 10^{-9} \, \text{m}^2/\text{s} \) compared with experimental observation.

Hydrogen bonding is crucial for the conformation and function of proteins. Proteins respond to variations in pressure and temperature by changing the number of hydrogen bonds with water. According to our calculations, the number of hydrogen bonds as well as their lifetime grows with decreased temperature and increased pressure, resulting in stronger interactions between BFP and the solvent. This may explain the violation of Stokes-Einstein formula.

Existing experimental compressibility data on proteins are obtained by sound velocity measurements and contain a contribution due to the protein's compressibility as well as the nontrivial addition from the hydration shell. The task to separate these components is very complicated, and therefore the computational approach, which allows this distinction to be made explicitly at atomic level, can be very informative. The MD simulation data enable the calculation of protein and solvent compressibility. Our computed protein compressibility \( 3.9 \times 10^{-5} \, \text{MPa}^{-1} \) is in agreement with the experimental values for most chromophore-containing proteins.

The analysis of the trajectories highlights structural properties of the protein that are essential for fulfilling its biological function as fluorescent protein: the rigidity of the \( \beta \)-sheet barrel for protection of the chromophore in the protein interior from solvent molecules; the stiff chromophore environment and the flexibility of the loops. The RMS fluctuations of the MD calculation are in accordance with experimentally

---

**FIGURE 12**

A: RMS fluctuations of BFP backbone C\(_\alpha\) atoms. Upper curve: experimental displacements according to temperature factors in the X-ray structure file 1BFP.pdb (black line). Lower curves: calculated displacements at atmospheric pressure (red line) and at an applied pressure of 1000 MPa (blue line). The arrow indicates the location of the chromophore. B: The backbone of BFP, color-coded in the scale of blue-green-red (color sequence in order of fluctuations increase). The most fluctuating residues are labeled and shown as ball and stick models.
observed fluctuations obtained from Protein Data Bank temperature factors.

During a calculated 3-ns trajectory, we could observe the diffusion of several water molecules into and out of the protein, though they never reached the immediate neighborhood of the chromophore. The performed simulations did not reveal the mechanism responsible for pressure-induced enhancement of fluorescence quantum yield, probably because the conformational conversions and/or water diffusion into the chromophore neighborhood take place in millisecond timescales and could not be detected during the nanosecond-time simulation.

Overall, our MD studies with AMBER parameters offer an accurate description of BFP and its interactions with surrounding water molecules. This is reflected by the good agreement between our computed data (compressibility, diffusion coefficients, and B factors) and experimental values.

The authors are grateful to Dr. Artur Suisalu for helpful discussions.

REFERENCES


Reviewing Editor: David Case