doi:10.1246/cl.2011.476

Published on the web April 13, 2011

Molecular Dynamics Simulation for Irreversibility of Green Fluorescent Protein before and after Photoactivation

Jun Koseki, Yukiumi Kita, and Masanori Tachikawa*
Quantum Chemistry Division, Yokohama City University, 22-2 Seto, Kanazawa-ku, Yokohama, Kanagawa 236-0027

(Received February 2, 2011; CL-110095; E-mail: tachi@yokohama-cu.ac.jp)

To elucidate the irreversible photoreaction of green fluorescent protein (GFP), we have theoretically analyzed hydrogen-bonding networks and distributions of water molecules around a chromophore (CRO) of GFP before and after photoactivation. Our molecular dynamics simulation clearly shows that such irreversibility arises from the migration of water molecules to the *hollowed out region* due to the reorientation of amino acid residues together with the disappearance of hydrogen-bonding networks around the CRO.

Green fluorescent protein (GFP) was discovered from jellyfish (*Aequorea victoria*) in the 1960s by Shimomura et al.[1] Nowadays, GFP and many GFP-like proteins are applied as a biological marker in a number of areas such as molecular biology, medicine, and cell biology.[2] On the other hand, in the latter half of the 1990s, Elowitz et al.[3] found the irreversible red shift of absorption and fluorescent spectra of GFP with brief pulses of 488 nm light under the absence of oxygen. The absorption peak for GFP after photoactivation (RsFP) is observed at 525 nm, while major and minor absorption peaks of Wild-type GFP are observed at 398 nm from a neutral state chromophore (CRO) and 475 nm from anionic state CRO, respectively. The emission peak for RsFP is also observed at 600 nm, while the emissions of neutral and anionic states are almost the same peak at 503–508 nm. However, the structure of RsFP and the mechanism for this spectral red shift of GFP have not been revealed.

Recently, we proposed a photoreaction mechanism of GFP, which is based on the formation of Schiff base between CRO and Arg96.[4] The Schiff base reaction proceeds in two steps, a nucleophilic addition and dehydration. Using model structures of GFP and RsFP, we theoretically analyzed the rate-determined step (a nucleophilic addition) with ab initio molecular orbital (MO) calculations, and showed that the photoactivation of GFP can occur in the anionic state of CRO through T_1 state. Although our mechanism can reasonably describe the photoreaction process of GFP and corresponding experimental results, the "irreversible feature" of the photoreaction is still insufficient. To elucidate such irreversibility of the photoreaction, a dynamic analysis of whole protein including water molecules is indispensable because the Schiff base forming reaction is generally known to be reversible by attack of a water molecule.[5]

In this study, we theoretically analyzed structures of GFP and RsFP, focusing on hydrogen-bonding networks and distributions of water molecules around CRO. We performed molecular dynamics (MD) calculations for systems which contain whole fluorescent protein (GFP or RsFP) and about 7000 water molecules in a truncated octahedral cell with a periodic boundary condition. The force fields of AMBER[6,7] and TIP3P[8] were used for the fluorescent protein and water molecules, respectively. A general AMBER force field (GAFF)[7] was also employed for CRO, because the theoretical accuracy of GAFF for analyzing structures of GFP has been well confirmed in our previous paper.[4] The structure of GFP determined by X-ray diffraction (1EMB[9] PDB) was used for an initial structure of GFP, where we added hydrogen atoms to the anionic form of GFP and assumed the proton transfer from CRO to Glu222 through water22 (w22) and Ser205.[2] For the initial structure of RsFP, we used the modified 1EMB structure, where the CRO and R96 were substituted for the Schiff base form and a water molecule. To obtain the thermal equilibrium structure, we have performed MD calculations with NVT ensemble at 300 K after energy minimization and equilibrating 1000000 MD steps using AMBER 9 program package.[10] The time step and total simulation time for the sampling are 1.0 fs and 2.0 ns (2000000 steps), respectively.

Figure 1 shows the thermal equilibrium structures around CRO of (A) GFP and (B) RsFP. In the GFP structure, the hydrogen-bonding network consists of CRO, Ser205, Glu222, w22, etc., while the CRO forms the hydrogen bonding to only Thr203 in the RsFP structure. Such disappearance of the hydrogen-bonding network in the RsFP is attributed to the strong attraction between the CRO and Arg96 belonging to a β-barrel. We note here that the β-barrel has a *robust* geometric stiffness, so that Arg96 cannot change its relative position within RsFP and the CRO moves to Arg96 for the formation of the Schiff base.

In the MD calculation of RsFP structure, we focus on an anionic state of CRO, though there are two possible neutral and anionic CRO states. This is because our preliminary results at HF or CIS/6-31+G* level of ab initio MO calculation for RsFP show that the anionic state of CRO is about 40 kcal mol^{-1} lower than the neutral state. Our results indicate that a proton transfer from a hydroxy group on Thr203 to CRO is less stable in a protein environment, and reasonably agree with the fact that the RsFP has only a single peak at both absorption and emission spectra experimentally.

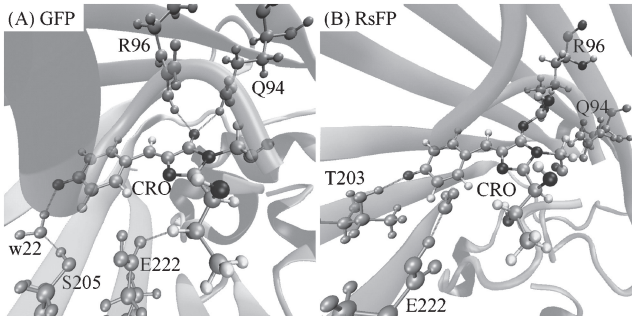

Figure 1. The hydrogen-bonding networks of (A) GFP and (B) RsFP.

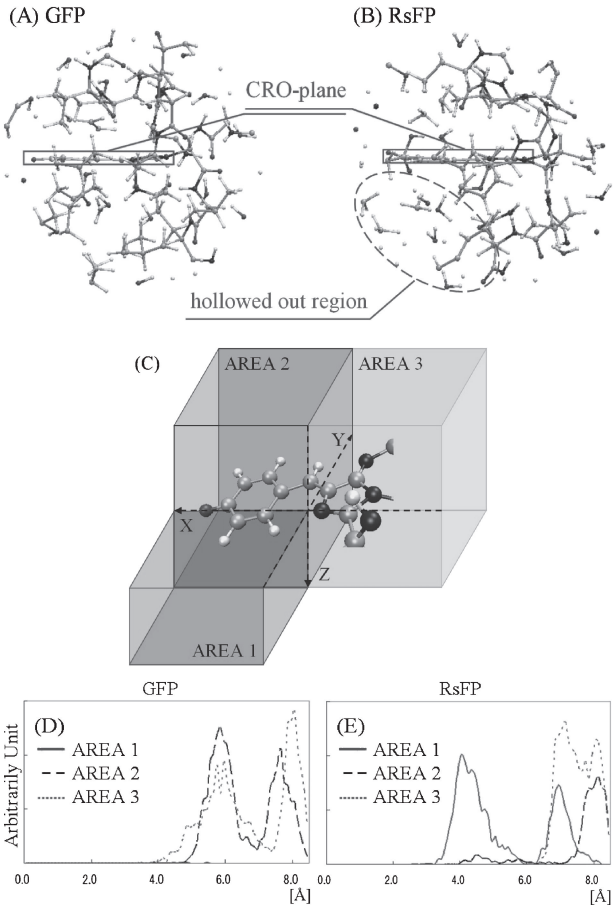

Figure 2. The side view of atoms within 8 Å from center of mass of CRO for (A) GFP and (B) RsFP. (C) Definition of new coordinate axis and areas. The radial distribution functions of oxygen atoms on water molecules from CRO center of mass for (D) GFP and (E) RsFP in AREAs 1, 2, and 3, respectively.

The side view of atoms within 8 Å from a center of mass of CRO for GFP and RsFP are shown in Figures 2A and 2B, respectively. The water molecules are distributed in all directions in GFP, while they tend to be concentrated under the CRO in RsFP. Such large differences in the distributions of water molecules are due to the reorientations of amino acid residues together with above rearrangements of the hydrogen-bonding network around the CRO.

In order to analyze distributions of water molecules around CRO in more detail, we introduced a new coordinate system in which the origin is defined as the center of mass of each CRO as shown in Figure 2C, where the X axis is defined as being toward the oxygen atom on the phenol ring from the origin and the X–Y plane is defined as containing the origin, oxygen atom on the phenol ring, and the oxygen atom on the imidazole ring of GFP (or the corresponding nitrogen atom of RsFP). The most remarkable difference of water distributions between GFP and RsFP is found in the following three areas; [$X \geq 0$, $Y < 0$, $Z \geq 0$], [$X \geq 0$, $Y \geq 0$], and [$X < 0$, $Y \geq 0$], denoted as AREAs 1, 2, and 3 in Figure 2.

Figures 2D and 2E show radial distribution functions (RDF) of oxygen atoms on water molecules from the origin. We found that there are no intensities in the AREA 1 of GFP, while there are two large intensities around 4.0 and 7.0 Å in the AREA 1 of RsFP. These results indicate that the water molecules tend to migrate to the AREA 1 in the RsFP, unlike the case of the GFP. Such condensation of water molecules can be caused by *a hollowing out* of this area due to the Schiff base formation as shown in Figures 2A and 2B. The RDF in AREAs 2 and 3 show that there are few water molecules within 6 Å from the origin in the RsFP, while the water molecules distribute in a range longer than 4 Å in the GFP. We also confirmed that the water molecules around 7 Å of RsFP are on the other side of some residues from the Schiff base connection and water molecules in AREAs 2 and 3 of GFP migrate to the AREA 1 of RsFP. As shown in Figure 2C, the Schiff base connection of RsFP is close to AREAs 2 and 3, and is located at the region from about 4 to 5 Å at AREAs 2 and 3 in Figure 2E. We found that some residues around the Schiff base connection exclude water molecules. These results clearly show that no water molecules can exist around the Schiff base in the RsFP unlike the case in the GFP, and also indicate the reason for the irreversibility of the Schiff base formation of the GFP.

In this study, we have theoretically analyzed the structural difference between the GFP and RsFP with molecular dynamics calculations based on our photoreaction mechanism of GFP in which the Schiff base formation occurs between Arg96 and CRO.[4] We found that a hydrogen-bonding network disappears in the RsFP structure due to the movement of the CRO to Arg96 belonging to a *robust* β-barrel. We also found that water molecules around the CRO are away from the Schiff base connection in the RsFP structure, and migrate to the *hollowed out region* around the phenol ring. Thus, the Schiff base formation is no longer reversible, because no water molecules can exist at the region close to the nitrogen atom on the Schiff base in the RsFP structure. We can conclude that the irreversibility of the photoreaction of GFP arises from an exclusion of water molecules around the Schiff base connection due to the reorientation of amino acid residues with the disappearance of the hydrogen-bonding network in the RsFP structure.

References

1 O. Shimomura, F. H. Johnson, Y. Saiga, *J. Cell. Comp. Physiol.* **1962**, *59*, 223.
2 M. Zimmer, *Chem. Rev.* **2002**, *102*, 759.
3 M. B. Elowitz, M. G. Surette, P.-E. Wolf, J. Stock, S. Leibler, *Curr. Biol.* **1997**, *7*, 809; M. B. Elowitz, M. G. Surette, P.-E. Wolf, J. B. Stock, S. Leibler, *J. Bacteriol.* **1999**, *181*, 197.
4 J. Koseki, Y. Kita, M. Tachikawa, *Biophys. Chem.* **2010**, *147*, 140.
5 J. McMurry, T. Begley, *The Organic Chemistry of Biological Pathways*, Roverts and Company Publishers, **2005**.
6 J. Wang, P. Cieplak, P. A. Kollman, *J. Comput. Chem.* **2000**, *21*, 1049.
7 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157.
8 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
9 K. Brejc, T. K. Sixma, P. A. Kitts, S. R. Kain, R. Y. Tsien, M. Ormö, S. J. Remington, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2306.
10 D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, D. A. Pearlman, M. Crowley, R. C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K. F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D. H. Mathews, C. Schafmeister, W. S. Ross, P. A. Kollman, *AMBER 9*, 9th ed., University of California, San Francisco, **2006**.

