

# A theoretical study on the electronic structure of PYP chromophore in low barrier hydrogen bonding model

Kenji Hirano, Hirofumi Sato\*

Department of Molecular Engineering, Kyoto University, Nishikyo-ku, Kyoto 615-8510, Japan

---

## Abstract

Yamaguchi *et al.* have recently identified positions of hydrogen and deuterium atoms in photoactive yellow protein (PYP) using high-resolution neutron scattering. They reported that the hydrogen bond between the PYP chromophore and Glu46 was not a short ionic hydrogen bond (SIHB) but a low barrier hydrogen bond (LBHB). Furthermore, it was suggested that Arg52 close to the chromophore was deprotonated. In the present study, we investigate the electronic structure of the chromophore in PYP under the condition of protonated or deprotonated Arg52. By analyzing the potential energy curve along the proton migration between Glu46 and the chromophore, we find that a LBHB can be seen only when Arg52 is deprotonated.

**Keywords:** photoactive yellow protein (PYP), low barrier hydrogen bond (LBHB), ONIOM

---

## 1. Introduction

It is well known that hydrogen bond plays a variety of roles in a wide range of chemical and biological phenomena. Among them, low-barrier hydrogen bonding (LBHB) [1] attracts great attention from researchers. It is characterized by short distance between the heavy atoms, a low energy barrier to proton transfer and distinctive features in the NMR spectroscopy. It has also emerged LBHB is important in enzymatic reactions, for example, the mechanism of serine protease function [2]. Photoactive yellow protein (PYP) is one of well-studied photosensors, being responsible for the negative photoaxis to blue light of *Ectothiorhodospira halopira*[3]. Its chromophore, *p*-coumaric acid (pCA) undergoes *trans* to *cis* isomerization upon light absorption, which leading to the arrangement of the the hydrogen-bonding network near the chromophore. This arrangement induced the large structural change of the whole protein through several intermediate states [4, 5].

It has been shown that two unusually short hydrogen bonds are involved in the ground state of wild-type PYP [3, 6, 7]. One is a bond between the chromophore, pCA

and Glu46, and the other is between pCA and Tyr42. Because the positions of the hydrogen atom are absent from X-ray crystallographic or neutron crystallographic structure [6, 7], the properties of these hydrogen bonds had no been clear. Recently, Yamaguchi *et al.* identified about 87 percent of hydrogen positions in PYP using neutron scattering technique [8]. They demonstrated that the hydrogen bond between pCA and Glu46 is a LBHB, whereas the short hydrogen bond between pCA and Tyr42 was not a LBHB. In addition, they reported that Arg52, which has been believed to be protonated as a counter ion of pCA, was deprotonated.

Several theoretical studies on PYP have been reported so far [9, 10, 11]. In these studies, the contribution from neighboring residues was discussed to clarify the protein environmental effect to the spectral shift. However, the effects from LBHB and deprotonated Arg52 were not discussed although the photo-absorption spectra of pCA are greatly changed by the protonation state. It is likely that a LBHB and/or deprotonated Arg52 crucially affect the absorption property of pCA in the protein environment. Apparently, this situation prompts us to elucidate the electronic property of LBHB and the absorption characters of pCA in the realistic conditions. In this article, we study the potential energy curve along the proton migration between Glu46 and pCA to confirm the existence of LBHB using the our own n-layered integrated molecular orbital molecular mechan-

---

\*Corresponding author. Tel: +81 75 3832548. Fax: +81 75 3832799.

Email address: hirofumi@moleng.kyoto-u.ac.jp  
(Hirofumi Sato)

ics (ONIOM) method, which is a powerful tool to compute the electronic structure in protein environment with reasonable computational cost [12, 13, 14]. Additionally, the absorption properties of pCA in the protein environment is computed under the existence of a LBHB and deprotonated Arg52.

## 2. Method

### 2.1. Calculations

In the present work, two-layered ONIOM method in the Gaussian 09 package was employed to compute the electronic structure in the protein environment [12]. The total potential energy in ONIOM scheme is obtained from three independent calculations:

$$E^{\text{ONIOM}} = E^{\text{High,model}} + E^{\text{Low,real}} - E^{\text{Low,model}}, \quad (1)$$

where “real” stands for the full system, which only needs to be calculated at low-level treatment, and “model” denotes the part of the system that needs to be calculated at both the high and low levels. Further computational explanation about ONIOM is described elsewhere [13, 14]. DFT with B3LYP functional and MP2 were adopted to compute the electronic structure of the high-level in ONIOM. The excited states and the excitation energies for the Franck-Condon transitions were investigated using TDDFT. The Amber molecular mechanics force field was adopted as the low-level calculation [15]. Electronic embedding [13] was incorporated in the ONIOM framework to describe electrostatic interaction between the layers more accurately. For the basis sets, we used cc-pVDZ for carbon and hydrogen atoms, and aug-cc-pVDZ for oxygen, nitrogen and sulfur atoms. The *s* and *p* diffuse functions were added on the hydrogen atom located between Glu46 and pCA to better describe the electronic structure around this region.

### 2.2. Model

The geometrical structure of PYP employed for all the calculations was taken from the high-resolution neutron crystallographic structure (PDB data: 2ZOI [8]) including the positions of the hydrogen atom. Because the PDB data does not provide about 13 percent of the hydrogen positions, these structure was supplemented using the TINKER program. The total potential energy curve of the proton migration between Glu46 and pCA was then computed. While the positions of the atom were basically taken from the PDB coordinate, the occupancy of the amino group proton in Arg52 is

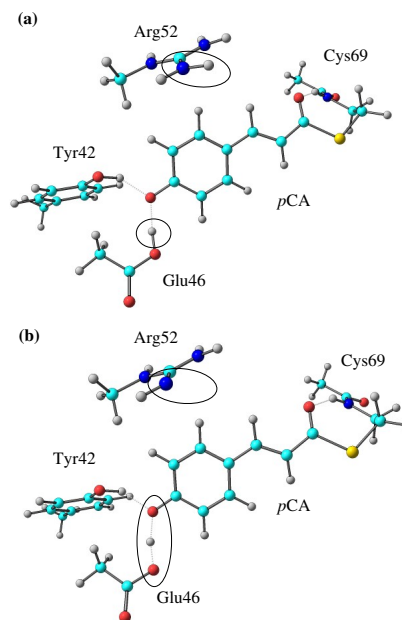


Figure 1: QM part (high-level) in ONIOM computations. (a) DP model and (b) LBHB model

about 0.23. Hence the two models, namely the protonated Arg52 (PA) and deprotonated Arg52 (DA) were prepared to elucidate the effect of the protonation. In these systems, the proton between Glu46 and pCA was assumed to be on the line between the two oxygen of Glu46 and pCA. The potential energy curve was then computed as a function of  $R_c$ , namely the difference between the two oxygen-hydrogen distances.

Two models were also prepared to compare the difference in the excitation energies about the protonation state around chromophore (Figure 1). Figure 1 (a) represents the model which has been so far used in former theoretical studies. In the model, pCA is assumed to be deprotonated and negatively charged, whereas the Glu46 was neutral due to the protonation of its carboxyl group. Arg52 was assumed to be protonated (PA). The position of the hydrogen atoms connected with Glu46 was obtained by the optimization of only quantum chemical (QM) region shown in the figure. All the other atoms were fixed to the experimental coordinates during the optimization. We call this model deprotonated (DP) model after the deprotonation

of pCA moiety. Figure 1 (b) is the other model fully based on the neutron-scattering[8]. In this new theoretical treatment, pCA and Glu46 were assumed to make a LBHB and Arg52 was deprotonated (DA). We call this model LBHB model. The QM regions consisted of 69 atoms for the DP model and 68 atoms for the LBHB model including *p*-coumaric acid, Tyr42, Glu46, Arg52 and Cys69. Then it is embedded into surrounding residues described by MM. The whole protein was neutralized by adding four (five) sodium counter ions for DP (LBHB) models. The ions were added by the Ambertools program.

### 3. Results and Discussion

#### 3.1. Potential energy profile

Figure 2 shows the total potential energy of DA and PA systems obtained by ONIOM method. The relative energies with respect to the energy at  $R_c = +0.34 \text{ \AA}$  are plotted. For PA system, both DFT and MP2 calculations show an asymmetric single-well energy curve along the proton migration. The energy is monotonously increased as  $R_c$  increases, *i. e.* the proton approaches from Glu46 side to pCA. The minimum is found to be placed at about  $R_c = -0.46 \text{ \AA}$ , in which the proton and the Glu46 oxygen distance is  $1.05 \text{ \AA}$ . Hence, the proton is bound to Glu46, presumably corresponding to covalent bond in PA system.

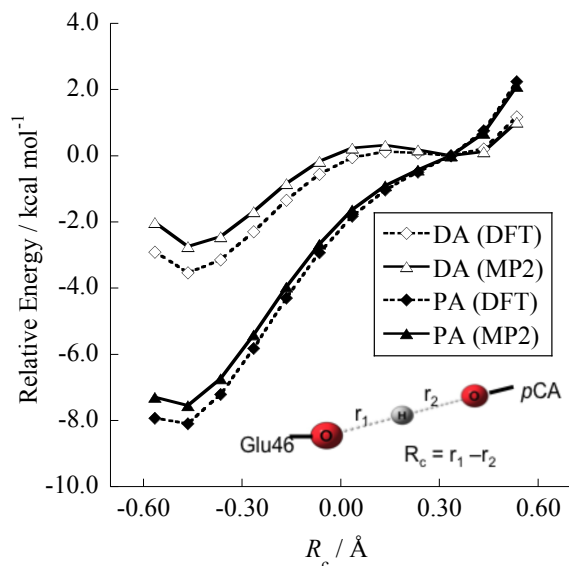


Figure 2: The total potential energy curve of PYP along the proton migration between the Glu46 and the chromophore. Definition of the coordinate,  $R_c$ , and potential energy curves for DA and PA systems.

On the other hand, the potential energy in DA system is remarkably different. Both DFT and MP2 results represent double-well curve that has a low barrier between the two minimum. These are found at about  $R_c = -0.46 \text{ \AA}$  and about  $R_c = 0.34 \text{ \AA}$  both by DFT and MP2. The energy in the former position is lower than the latter one where the proton is located at pCA nearby. The barrier heights from the former minimum (the position of  $R_c = -0.46 \text{ \AA}$ ) is  $3.66 \text{ kcal/mol}$  by DFT and  $3.06 \text{ kcal/mol}$  by MP2 method, corresponding to about  $1100 - 1300 \text{ cm}^{-1}$ . Because typical OH frequency is about  $3000 \text{ cm}^{-1}$ , the zero point energy level may be above the barrier. Namely, the computational results suggest that Glu46 and pCA can form a LBHB. The average position of the hydrogen should be closer to Glu46 rather than pCA due to the asymmetry of the potential curve profile. These are consistent with the experimental report, in which the average distances are respectively  $1.37 \text{ \AA}$  for H - O (pCA) and  $1.21 \text{ \AA}$  for H - O (Glu46). The computational results indicate that the protonation state in Arg52 significantly affects the energy profile.

The difference in the energy profile between the two models is roughly understood in terms of the electrostatic interaction. While the energies in both minima are similar in DA, the total energy in PA is strongly stabilized in  $R_c < 0$  because of the interaction between  $\text{pCA}^-$  and the positively charged Arg. In reality, these two models corresponding to extreme condition are non-exclusivist because the occupancy of the proton is 0.23. The effect of the fluctuation should be also taken into account. The superposition of these two models might be closer to the truth.

#### 3.2. Excitation energies

Undoubtedly, photochemical property is one of the central subjects in PYP research. It is known that the absorption spectrum near  $3 \text{ eV}$  is comprised of many electronic excited states, and TDDFT computations provide more than ten excited states, though many of them exhibit negligibly small oscillator strength. The excitation energy with the largest oscillator strength (1.04) in the low-lying excited states for DP model is  $3.28 \text{ eV}$  using the present ONIOM (TDDFT:Amber) model. This is in good agreement with the previous theoretical study by QM/MM calculation with CAM-B3LYP [11], *i.e.*  $3.34 \text{ eV}$  for the excitation energy with 1.101 for the oscillator strength. The excitation energy is  $3.32 \text{ eV}$  in LBHB model with an oscillator strength of 0.96. Hence, the absorption property is very similar in DP and LBHB models, and characterized as  $\pi\pi^*$  like transition as shown in Figure 3, mainly localized in the chromophore. This electronic character is similar to the ab-

sorption property of the electron transition in  $\text{pCA}^-$  (not shown).

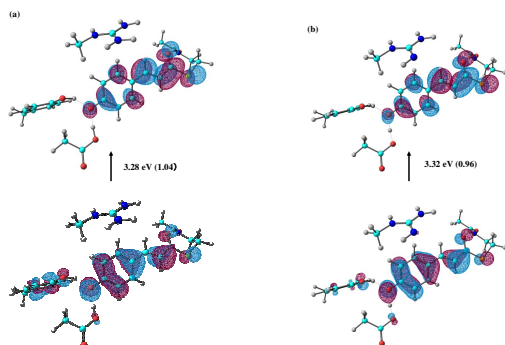


Figure 3: The electron transition character in (a) DP model and (b) LBHB model. The excitation energy (eV) and oscillator strength are also indicated.

It is very likely that the photo absorption wavelength depends on the protonation state of pCA. To investigate further this point, we computed the excitation energy as a function of the proton migration coordinate ( $R_C$ ). Figure 4 represents the change of the excitation energy. Only the transitions with the highest oscillator strength are plotted to avoid complexity. In the region of  $-0.6\text{\AA} < R_C < -0.16\text{\AA}$ , the excitation energies in DA and PA systems are very close to each other. As mentioned above, the transitions are mainly characterized by  $\pi\pi^*$  like transition, and the plots show slight increase near 3.3 eV of excitation. The optimized positions of hydrogen in DA and PA models, respectively  $R_C = -0.54\text{\AA}$  and  $R_C = -0.17\text{\AA}$ , correspond to the both ends of this region. Presumably, the effect of deprotonation of Arg52 is negligible, in other words, the proton is almost attached to Glu46 and the chromophore is regarded as an anion. The absorption property is essentially attributed to that of  $\text{pCA}^-$ .

As the proton approaches to the chromophore, the absorption property is changed. Note that only the state with the highest oscillator strength is plotted in the figure. Several electronic excited states are nearly degenerated and show mixing of other electronic character into the  $\pi\pi^*$  like transition. In  $R_C > 0.24\text{\AA}$ , the excitation energy is about 3.5 eV, and the main contribution to the excited state is again  $\pi\pi^*$  like transition. Because the proton is surely attached to the chromophore, the absorption is essentially attributed to the property of neutral pCA. Actually, the excitation energy of pCA is about 0.5 eV above to that of anion,  $\text{pCA}^-$  [16, 17].

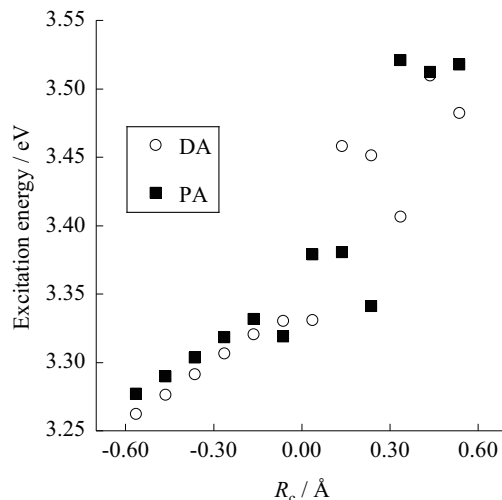


Figure 4: Excitation energies along the proton migration coordinate ( $R_C$ ) between Glu46 and the chromophore.

#### 4. Conclusions

In the present article, we utilized the ONIOM method to study low barrier hydrogen bond (LBHB) between Glu46 and pCA. In addition, the protonation state effect to the absorption property of the chromophore of PYP (pCA) was examined. Based on the neutron scattering crystallographic structure, the energy curve along the proton migration was computed in the protonated and deprotonated Arg52 models. The result revealed that the deprotonation of Arg52 plays a crucial role to make a double well potential curve with a very low barrier (about 3 kcal/mol). This result suggests that a LBHB exists when Arg52 is deprotonated.

The excitation energy and electronic structure were also investigated. The difference in the absorption properties between DP and LBHB models is found to be negligibly small when the proton is attached to Glu46. An analysis on the excited energy indicates several electronic excited states are mixed into the absorption spectra when the proton is shared by Glu46 and pCA.

Of course, the present model is far from the perfect one. For example, the effect of fluctuation of protein, solvent water and counter ions etc. are not taken into account. More sophisticated computation beyond the present ONIOM method is highly desired to further clarify the electronic character.

## Acknowledgment

This work has been financially supported in part by the Grant-in Aid for Scientific Research on Priority Area "Molecular Science of Fluctuations" (2006-21107511) and by the Grant-in Aid for General Research (20550013). All of them were supported by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Japan.

## References

- [1] C. L. Perrin, J. B. Nielson, *Annu. Rev. Phys. Chem.* **1997**, *48*, 511
- [2] P. Frey, S. Whitt, J. Tobin, *Science* **1994**, *264*, 1927
- [3] T. Meyer, *Biochim. Biophys. Acta.* **1985**, *806*, 175
- [4] R. Kort, H. Vonk, X. Xu, W. Hoff, W. Crielaard, K. Hellingwerf, *FEBS Lett.* **1996**, *382*, 73
- [5] Y. Imamoto, M. Kataoka, F. Tokunaga, *Biochemistry* **1996**, *35*, 14047
- [6] S. Anderson, S. Crosson, K. Moffat, *Acta Cryst.*, **2004**, *D60*, 1008
- [7] S. Z. Fisher, S. Anderson, R. Henning, K. Moffat, P. Langan, P. Thiagarajan, A. J. Schultz, *Acta Cryst.*, **2007**, *D63*, 1178
- [8] S. Yamaguchi, H. Kamikubo, K. Kurihara, R. Kuroki, N. Nimura, N. Shimizu, Y. Yamazaki, M. Kataoka, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 440
- [9] E. V. Gromov, I. Burghardt, H. Koppel, L. S. Cederbaum, *J. Am. Chem. Soc.* **2007**, *129*, 6798
- [10] P. B. Coto, S. Marti, M. Oliva, M. Olivucci, M. Merchan, J. Andres, *J. Phys. Chem. B* **2008**, *112*, 7153
- [11] T. Rocha-Rinza, K. Sneskov, O. Christiansen, U. Ryde, J. Kongsted, *Phys. Chem. Chem. Phys.* **2011**, *13*, 1585
- [12] S. Dapprich, I. Komaromi, K. S. Byun, K. Morokuma, M. J. Frisch, *J. Mol. Struct. (THEOCHEM)* **1991**, *461-462*, 1
- [13] T. Vreven, K. Morokuma, *Theor. Chem. Acc.* **2003**, *109*, 125
- [14] T. Vreven, K. S. Byun, I. Komaromi, S. Dapprich, J. A. Montgomery, K. Morokuma, M. J. Frisch, *J. Chem. Theo. Comput.* **2006**, *2*, 815
- [15] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, J., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179
- [16] T. Rocha-Rinza, O. Christiansen, J. Rajput, A. Gopalan, D. B. Rahbek, L. H. Andersen, A. V. Bochenkova, A. A. Granovsky, K. B. Bravaya, A. V. Nemukhin, K. L. Christiansen, M. B. Nielsen, *J. Phys. Chem. A* **2009**, *113*, 9442
- [17] M. Putschogl, P. Zirak, A. Penzkofer, *Chem. Phys.*, **2008**, *343*, 107.