Charge-density analysis of green fluorescent protein at ultra-high resolution

Green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* is a light-emitting protein that absorbs blue light and emits green light [1]. The structure consists of an 11-stranded β-barrel plugged by the chromophore (Fig. 1). The chromophore 4-(*p*-hydroxybenzylidene)imidazolin-5-one is formed from three intrinsic residues in the polypeptide chain by post-translational reactions. Two forms of the chromophore, "A" and "B", with different protonated states exist in wild-type GFP. The A and B forms can be individually stabilized by introducing mutations to residues around the chromophore. Many artificial variants with various useful features have been created, and the variants as well as the wild-type GFP are now indispensable as tools in molecular biology. More than 400 structures of related fluorescent proteins, including their variants, have been deposited in the Protein Data Bank (PDB). However, the electron densities of hydrogen atoms and valence electrons, which characterize the chemical states of atoms, are difficult to detect by conventional X-ray analysis. Accordingly, chemical states of atoms constituting the chromophore and the details of the interaction with the surrounding residues in GFP are unclear but indispensable for understanding the fluorescence mechanism. Meanwhile, the charge-density analysis at ultra-high resolution can provide details of the electronic structures of the protein. Recently, we successfully performed the charge density analysis of GFP at ultra-high resolution [2].

The F99S/M153T/V163A/E222Q mutation was introduced to GFP to stabilize the B-form structure of the chromophore. Large (~1.0×0.5×0.5 mm³) crystals were obtained by the seeding method. A diffraction dataset at 0.78 Å resolution was obtained at SPring-8 **BL41XU** beamline using high-energy X-rays ($\lambda = 0.35$ Å, $E = 35.4$ keV) [3]. The structure was initially refined with the parameters of the conventional independent spherical atom model (ISAM). Almost all the hydrogen atoms of the protein residues were detected and included in the structure model. In addition, more than 20 hydrogen atoms of water molecules were included. Residual densities from the valence electrons were observed around each atom after several cycles of refinement (Fig. 2(a)). Further refinements were performed using the parameters of the multipolar atomic model (MAM) to explain the nonspherical features of the valence electrons. The final R_{work} and R_{free} factors converged to 10.8% and 12.5%, respectively. After the MAM refinement, the residual densities were reproduced in the static deformation map, indicating the success of the MAM refinement (Fig. 2(b)).

We analyzed the features of the refined map of GFP by applying the atoms-in-molecules (AIM) theory to detect hydrogen bonding around the chromophore (Fig. 3(a)). We detected some nonconventional (CH···O type) hydrogen bonds between the chromophore and the protein environment in addition to conventional hydrogen bonds. Another nonconventional (CH···N type) hydrogen bond bridging was also detected between two ring moieties of the chromophore. The planarity of the two ring moieties, which is essential to the high fluorescence activity of GFP, seems to be reinforced by the CH···N type interaction. In addition, weaker attractive and repulsive interactions were visualized by the noncovalent interaction (NCI) analysis of the accurate electron density map. Some attractive CH–π and lone pair–π interactions were detected between the chromophore and
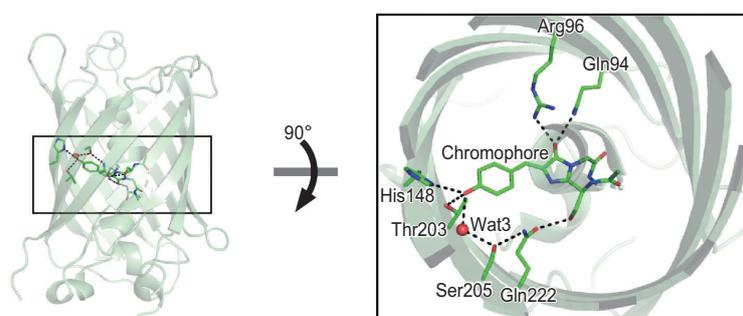

Fig. 1. Structure of GFP. The barrel structure for the F99S/M153T/V163A/E222Q variant is represented as a ribbon model. The chromophore and the surrounding residues are represented as sticks. Conventional hydrogen bonds around the chromophore are indicated as dashed lines.
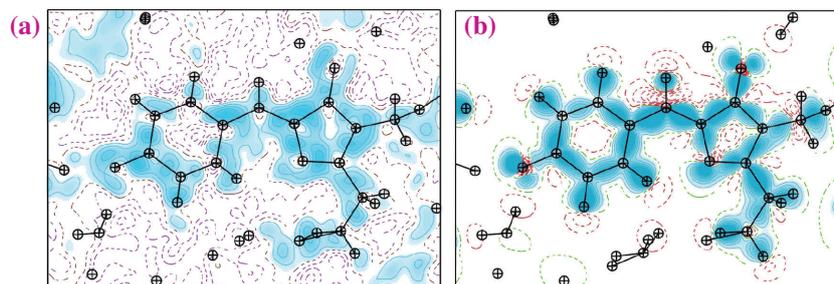
Fig. 2. Contour maps around the chromophore in the GFP. **(a)** Residual ($F_{obs} - F_{calc}$) map after ISAM refinement. The contour intervals are 0.05 e/Å^3. **(b)** Static deformation map after MAM refinement.

Thr62 (Fig. 3(b)). These weak interactions cause electron delocalization in the chromophore from the imidazolinone moiety to the phenolic moiety, which means the phenolate form is more predominant than the quinone form. The resultant electronic structure enhances the transition moment in the excitation reaction for efficient fluorescence.

Interactions in protein molecules have been estimated empirically from geometric parameters such as bond lengths and angles. However, by using the charge density obtained from ultra-high-resolution analysis, intramolecular interactions can be detected quantitatively as features of the experimental electron density. Actually, intramolecular interactions around the active sites have recently been investigated by charge density analyses at ultra-high resolutions for other proteins [4,5]. The information obtained in this work will enable us to perform the rational design of new GFP variants through quantum chemical calculations.
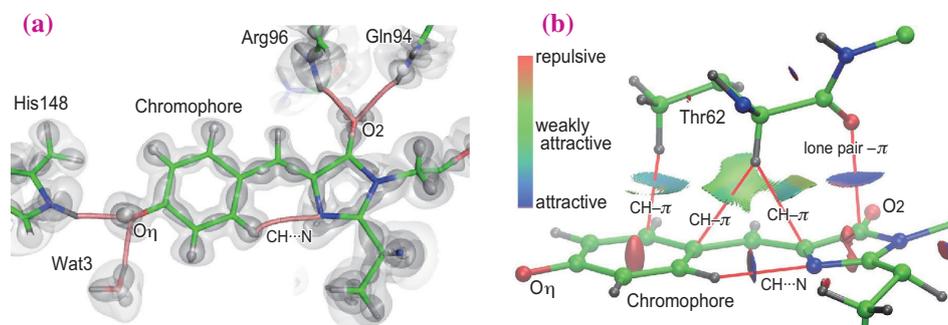

Fig. 3. Interactions between the chromophore and the surrounding residues. **(a)** Hydrogen bonding network revealed by the AIM analysis. The bond paths for conventional and nonconventional hydrogen bonds are represented as pink curves. The gray surfaces represent the deformation electron density at contour levels of +0.01, +0.15 and +0.5 e/Å^3. **(b)** NCI surface around the chromophore. The reduced density gradient isosurface is represented in a blue-green-red scheme. Blue indicates attraction, green indicates very weak attraction and red indicates repulsion.

Kiyofumi Takaba[†] and Kazuki Takeda*

Department of Chemistry, Kyoto University

*Email: ktakeda@kuchem.kyoto-u.ac.jp

[†] Present address: RIKEN SPring-8 Center

References
[1] O. Shimomura *et al.*: Cell. Comp. Physiol. **59** (1962) 223.
[2] *K. Takaba, Y. Tai, H. Eki, H.-A. Dao, Y. Hanazono, K. Hasegawa, K. Miki and K. Takeda: IUCrJ **6** (2019) 387.*
[3] K. Hasegawa *et al.*: J. Synchrotron Rad. **20** (2013) 910.
[4] Y. Hirano *et al.*: Nature **534** (2016) 281.
[5] K. Takaba *et al.*: Sci. Rep. **7** (2017) 43162.