Determining the Isomerization Mechanism of Photoswitchable Fluorescent Proteins

Jeffrey Chang

(Dated: March 7, 2018)

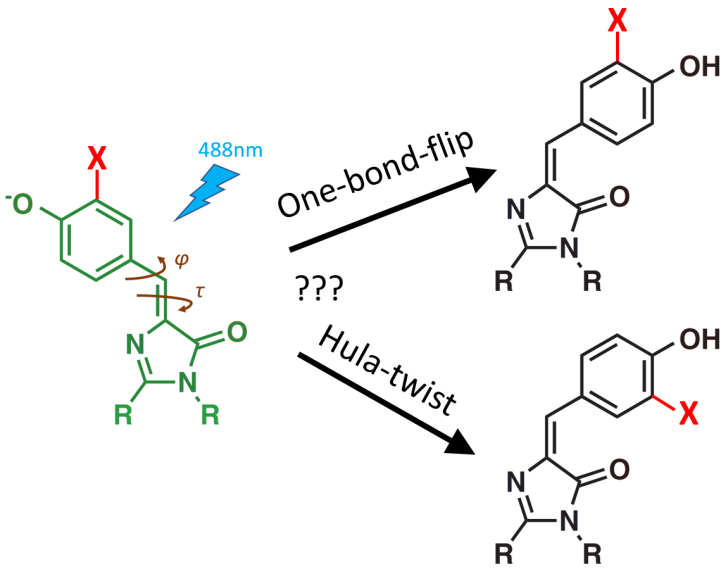

FIG. 1. My proposed experiment will distinguish the one-bond-flip from the hula-twist mechanism for fluorescent protein photoswitching. Figure adopted from [1].

I. OVERVIEW

I propose an experiment that will provide the first direct evidence of how double bonds flip over in photoswitchable fluorescent proteins. These glowing proteins allow biologists visualize cells to remarkable resolutions, and if we wish to design better proteins, we need to understand their photophysics. It is unknown whether these molecular light bulbs turn on and off via a one-bond-flip or a hula-twist. To distinguish between the two mechanisms, I will label one side of the phenolate ring by using amber suppression, and I will observe whether the label switches sides when the double bond flips over by using x-ray crystallography.

This unprecedented experimental evidence will test quantum chemical simulations and inform the design of improved fluorescent proteins.

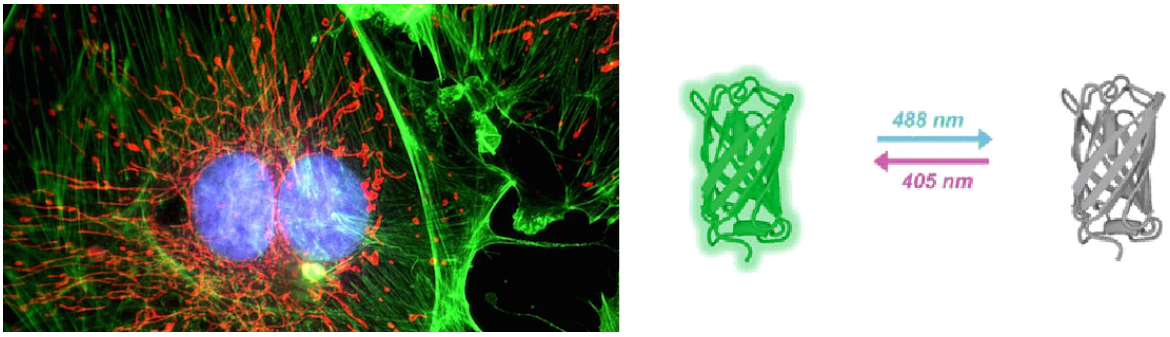

FIG. 2. Photoswitchable dyes allow for incredible cellular imaging. Left, a cellular image made posible by photoswitchable dyes. Right, a photoswitchable dye turns on and off upon laser irradiation of different wavelengths. Figures taken from [3] and [4].

II. MOTIVATION

Recent advances in superresolution microscopy have lead to stunningly beautiful images and remarkable advances in biology (Figure 2) [2]. With clever optical and tricks, new cellular imaging methods can defeat the diffraction limit and visualize live cells at length scales of tens of nanometers. At the heart of these techniques lies 'photoswitchable dyes': fluorescent molecules that we can switch between a bright on-state and a dark off-state (Figure 2). By selectively turning on only a few dyes at a time in separate snapshots, photoswitchable dyes allow us to piece together images of incredible spatial resolution.

Reversibly photoswitchable fluorescent proteins (RSFPs) are incredibly powerful because we can easily control their fluorescence: simply irradiating them with a laser switches them from a fluorescent on-state to a non-fluorescent off-state. Over the past decade, scientists have developed a wide palette of RSFPs, since enhanced dyes allow better microscopy – dyes that photoswitch faster can speed up the imaging framerate, and dyes that shine brighter can increase the signal-to-noise-ratio. For this reason, there is a widespread inerest in improving the photophysical properties of RSFPs.

Despite this demand, we understand rather little about the physics of photoswitching. Currently, RSFPs are not rationally designed from physical principles, but blindly engineered through rounds of directed evolution. In this brute-force approach, scientists randomly mutate the amino acid sequence of an RSFP and select the best candidates from the thousands of random mutants. Even though directed evolution can successfully improve RSFP prop-

erties, it is not based on any physical understanding of photoswitching. If we wish to design better RSFPs via rationale rather than random chance, we must understand the physical factors that underlie photoswitching.

Here, I propose to investigate the mechanism of photoswitch – whether it proceeds through a one-bond-flip or a hula-twist mechanism (see below for details). I expect that this knowledge will inform rational RSFP design. For instance, suppose that isomerization proceeds via a one-bond-flip, and suppose that we observe from a crystal structure that a portion of the protein blocks the path of the one-bond-flip. If we designed a mutant with a wider chromophore pocket to accomodate the spacious one-bond-flip, we may be able to rationally speed up the kinetics or improve the quantum yield of photoswitching.

A true mechanistic understanding of photoswitching should allow us to design better fluorescent dyes for superresolution cellular imaging.

III. BACKGROUND

Now I briefly explain what is currently known about the physical chemistry of photoswitching.

The heart of RSFP photophysics is the chromophore, the light-emitting chemical moiety that lies within the protein matrix. Its chemical structure (Figure 3) consists of a phenolate (P) ring and an imidazolinone (I) ring linked by a single-bond double-bond bridge. The RSFP chromophore can exist in two states: a fluorescent on-state, and a non-fluorescent off-state. As shown in Figure 3, the two states differ in their double-bond geometry – the on-state has a cis orientation, and the off-state has a trans orientation. Remarkably, the two states can interconvert under laser radiation. This remarkable photochemical phenomenon makes RSFPs invaluable for microscopy and endlessly interesting for photochemistry.

The photoisomerization process is summarized in Figure 3. In this schematic, we see the trajectory of the molecule on its potential energy surfaces (PES). The PES describes the energy of the molecule as a function its shape, or its nuclear geometry. As shown in the figure, in the ground-state PES, the on-state is separated from the off-state by a high activation barrier, so the reaction must proceed in an electronic excited state. When the chromophore absorbs a photon (1), the electrons are excited from the ground state into an excited state, and as the nuclei relax on this new PES (2), the molecule moves passes

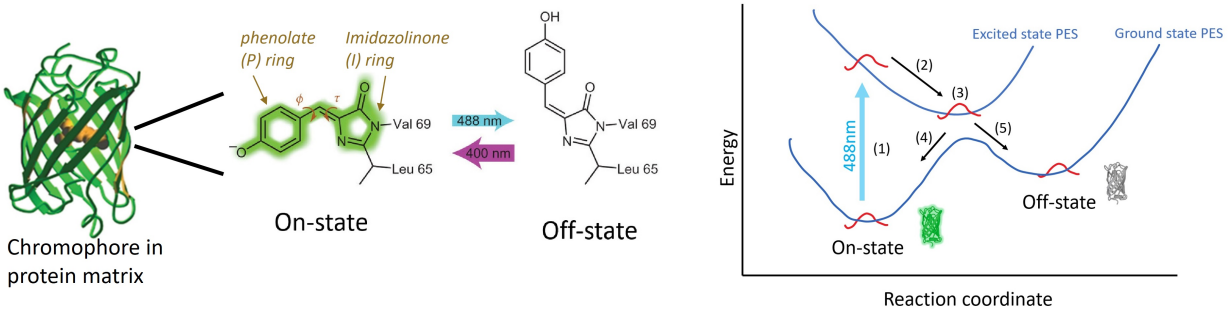

FIG. 3. The RSFP switches between on and off when its chromophore double bond isomerizes from cis to trans. Left, the chromophore is surrounded by the protein matrix. Middle, the two-ring structure of the chromophore can convert between an on and off state. Right, a schematic of the chromophore photoswitching dynamics. I hope to study the trajectory of the red nuclear wavepacket. Figures adopted from [5] and [6].

through an avoided crossing (3), and either returns to the on-state (4) or isomerizes to the off-state (5).

A simple question about photoisomerization is the trajectory that the atoms take as they move along the potential energy surface. The atoms must move around the double bond in one of two ways: a one-bond-flip, or a hula-twist. Shown in Figure 1 are the two conceivable isomerization mechanisms. In the one-bond-flip, only the τ-bond twists around, dragging along the entire P-ring. In the hula-twist, both the ϕ-bond and τ-bond simultaneously twist around, leaving the P-ring mostly stationary. The two processes have profoundly different properties – most notably, the hula-twist takes up less space. Because of this, the one-bond-flip tends to occur in aqueous or flexible media, while the hula-twist tends to occur in more rigid and spatially confied media [7].

Evidently, knowing the mechanism of RSFP photoswitching will tell us not only the shape of its PES, but also the nature of the chromophore pocket and the dynamics of the protein.

IV. LITERATURE REVIEW

Currently, we do not know the molecular trajectory of RSFP photoswitching, experimentally or theoretically.

There is no definitive experimental evidence of which mechanism occurs. Experiments

are tricky because double bonds isomerize on a picosecond time scale [8]. Pump-probe transient-absorption methods are fast enough to resolve the kinetics, but they provide no structural information to distinguish the one-bond-flip from the hula-twist. A recent tour-de-force study [5] has obtained structural snapshots that suggests evidence for hula twist, but it does not rule out a one-bond-flip mechanism.

Theoretical simulations are difficult and unreliable, and they sometimes contradict each other. Some simulations report a hula-twist [8] [9], while others report a one-bond-flip [10] [11]. The problem is inherently very difficult to simulate: it involves tens of atoms in a complex protein environment; it requires computing excited-state energies (which most quantum chemical calculations are not intended for); the nuclear and electronic degrees of freedom couple at the PES crossing; etc. Evidently, chromophore photoisomerization is a theoretical as well as an experimental frontier. To verify the current theoretical methods, we need an experiment to demonstrate how the double bonds in RSFPs flip over.

V. PROPOSED METHODOLOGY

Here I propose an experiment that can directly distinguish the one-bond-flip from the hula-twist mechanism in RSFP photoswitching.

As shown in Figure 1, the two mechanisms result in the same off-state product, but with one key difference – the one-bond-flip process causes the P-ring to end up upside-down compared to the hula-twist process. Normally both mechanisms yield the same product since the P-ring is symmetric. However, if we could somehow differentiate the two sides of the P-ring, the two mechanisms would yield different products.

Figure 1 summarizes the proposed experiment. I will put a label on one side of the P-ring. If I observe that photoswitching flips the label onto the other side, I can deduce that a one-bond-flip occurred; if I observe that the label stays on the same side, I can deduce that a hula-twist occured.

To label the P-ring, I will use amber suppression to add an extra halogen atom on one side of the ring (see appendix for details). Using this genetic code expansion technique, I can incorporate a tyrosine with an extra halogen atom into position 67 of the RSFP amino acid sequence. When amino acids 66-68 transform into the chromophore, the resulting P-ring will have an extra halogen substituent. This halogen atom acts as a label for one side of the

P-ring.

To observe whether this halogen label switches sides during photoswitching, I will use x-ray crystallography. In brief, this biochemical technique allows me to determine the locations of atoms within the protein, both in the on-state and the off-state (see appendix for details). The resulting electron density map will tell me where the halogen atom is before and after the photoswitch.

To summarize, an extra halogen label on one side of the P-ring will allow me to distinguish the one-bond-flip from the hula-twist. By photoswitching the protein and then observing which side the extra halogen atom appears on in X-ray structures, I will be able to deduce which mechanism occurs.

VI. EXPERIMENTAL PLAN AND TIMELINE

I should be able to complete the proposed experiment within half a year of full-time work in the lab. The timeline is summarized in the table below. Note that since protein crystallization is inherently unpredictable [12], it is difficult to estimate how long it will take, but based on the Boxer lab's past experience with crystallizing fluorescent proteins, it should take no longer than a few months.

Experimental Phase	**Estimated time**
Recombinant DNA work	2 weeks - 1 month
Protein expression and purification	2 weeks - 1 month
Protein crystal growth and x-ray diffraction	1 month - 2 months
Structure refinement	1 month - 2 months
Total:	**4 - 6 months**

Here are the detailed experimental plans for my proposal. I will work with the RSFP named rsEGFP2 because it can crystallize and diffract well [13]. A poly-His tag will be attached to the beginning of the gene to allow me to purify the protein with nickel affinity chromatography. To label the P-ring of the chromophore, I will incorporate ortho-chlorinated, ortho-brominated, and ortho-meta-difluorinated tyrosines (Figure 4) into the Y67 position of the protein. I choose these mutants because they have the available amber suppression machinery, and because they minimally perturb the phenolate ring structure.

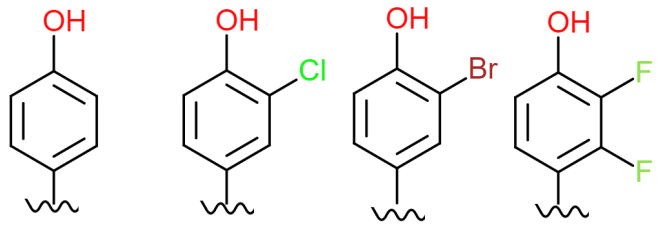

FIG. 4. Shown above are the proposed mutant chromophores in my experiment to label one side of the ring.

To synthesize my proteins, I will use shake-flask protein expression. I will create the genetic materials that encode my protein (ie, the rsEGFP2 gene inserted into a suitable expression vector) using typical DNA recombination techniques (eg, polymerase chain reaction (PCR) and Gibson assembly). I will then transform the DNA into a cell line optimized for expressing protein (BL21), culture the bacteria in liter-sized flasks, and chemically induce the bacteria to produce the protein. Once I obtain the protein, I will sequentially purify it through the standard techniques of nickel affinity chromatography, ion-exchange chromatography, and size-exclusion chromatography.

To crystallize my proteins, I will use the hanging-drop vapor diffusion method [12]. Suitable crystallization conditions will be taken from reference [13]. To obtain crystals in the off-state, I will irradiate the crystals with a 588-nm laser before flash-freezing them in liquid nitrogen [13]. To obtain crystals in the on-state, I will directly flash-freeze them from the crystallization solution. Once I cryocool my crystals, I will diffract x-rays on the crystals using the beamlines at the SLAC linear coherent light source. From the datasets, I will refine atomic structures using the software tool Phenix.

VII. PREPARATION

I have been part of the Boxer group in the chemistry department since September 2017. During my time, I have learned the experimental methods I would use in this proposal: recombinant DNA work, shake-flask protein expression, protein purification, hanging-drop protein crystallization, and x-ray crystallography. I have support from Professor Boxer and the graduate students in his lab for this ambitious investigation. We all hope that I can make good progress.

In general, my broad array of coursework has prepared me well to understand the interdisciplinary material in my proposal – the quantum mechanics of the light emission, the physical chemistry of photoswitching kinetics, the biochemistry of protein structure and dynamics, and more.

VIII. STATEMENT OF WORK (SOW)

I will express a series of rsEGFP2 amber mutants (Wild-type, Chloro, Bromo, 2,3-difluoro), and grow protein crystals in both the on-state and off-state. I will then diffract x-rays on the crystals and refine atomic models of the on and off states to determine whether photoswitching occurs via a one-bond-flip or a hula-twist. Deliverables will be in the form of refined x-ray crystal structures, within three to six months of starting the project.

IX. LINE-ITEM BUDGET

Summer 2018 living expenses (Room and Board)	$4000
Summer 2018 stipend	$3000
Total:	**$7000**

The above funds are requested to support my full-time research (40 hr per week) during the summer on the proposal above.

Appendix A: Experimental Techniques

My proposed experiment relies on two techniques: amber suppression to label the P-ring, and x-ray crystallography to determine the protein structure. Here, I briefly explain how these biochemical techniques work, and why they are suitable for my experiment.

Amber suppression is a technique to incorporate non-canonical amino acids – amino acids outside of Nature's twenty amino acids – into proteins. It works by re-wiring the genetic code of the protein-making cell machinery [14]. Suppose a biochemist wishes to modify a particular amino acid inside a protein. Conventionally, the biochemist would mutate the DNA to make it encode for another one of Nature's twenty amino acids. With amber

suppression, the biochemist can encode a twenty-first amino acid by repurposing a lesser-used codon in the genetic code (the amber stop codon) to encode for this new amino acid. There are two parts to this new machinery. To recognize this new codon, the biochemist must engineer a new tRNA. To attach the new amino acid onto this new tRNA, the biochemist must engineer an enzyme to catalyze this reaction, an amino-acyl-tRNA-synthetase (aaRS). Both parts of this machinery – the tRNA and aaRS – can be obtained from reference [14].

In my proposed experiment, I will use amber suppression to incorporate a labeled P-ring because none of Nature's twenty amino acids includes an asymmetric phenolate ring. I expect successful results because amber suppression has been shown to create fluorescent proteins with substituted chromophores [15].

X-ray crystallography is a standard biochemical technique to determine the structure of proteins – the locations of their atoms. To begin, a crystal of protein is grown from a solution of protein. Crystal growth is incredibly sensitive to the solution conditions – concentrations of salts, pH, temperature, etc, so the crystallographer must search through many possible crystallization conditions until suitable crystals form. Once protein crystals are obtained, they are typically flash-frozen with liquid nitrogen to protect them from radiation damage, and then they are illuminated with x-rays to form a diffraction pattern. By measuring the diffraction angles and intensities, we can determine the mean density of electrons within the crystal. From this density map, we can model the locations of atoms within the crystals.

In my proposed experiment, I will use x-ray crystallography to figure out what side of the P-ring the halogen substituents are on. I expect to successfully obtain crystal structures because my fluorescent protein, rsEGFP2, has been been demonstrated to crystallize under well-documented conditions [13]. The halogenated mutants should readily crystalize because the extra atom minimally perturbs the crystal structure [15].

---

[1] X. X. Zhou and M. Z. Lin, Current opinion in chemical biology **17**, 682 (2013).

[2] S. J. Sahl and W. Moerner, Current opinion in structural biology **23**, 778 (2013).

[3] https://microscopy.stanford.edu/, .

[4] https://basicmedicalkey.com/photoswitchable-fluorescent-proteins-for-superresolution-fluorescence-rcumventing-the-diffraction-limit-of light/, .

[5] N. Coquelle, M. Sliwa, J. Woodhouse, G. Schirò, V. Adam, A. Aquila, T. R. Barends, S. Boutet, M. Byrdin, S. Carbajo, *et al.*, Nature chemistry **10**, 31 (2018).

[6] https://www.biotek.com/resources/white-papers/fluorescent-proteins-filters-mirrors-and wavelengths/, .

[7] R. S. Liu, Accounts of chemical research **34**, 555 (2001).

[8] M. Andresen, M. C. Wahl, A. C. Stiel, F. Gräter, L. V. Schäfer, S. Trowitzsch, G. Weber, C. Eggeling, H. Grubmüller, S. W. Hell, *et al.*, Proceedings of the National Academy of Sciences of the United States of America **102**, 13070 (2005).

[9] S. L. Moors, S. Michielssens, C. Flors, P. Dedecker, J. Hofkens, and A. Ceulemans, Journal of chemical theory and computation **4**, 1012 (2008).

[10] G. Jonasson, J.-M. Teuler, G. Vallverdu, F. Merola, J. Ridard, B. Levy, and I. Demachy, Journal of chemical theory and computation **7**, 1990 (2011).

[11] X. Li, L. W. Chung, H. Mizuno, A. Miyawaki, and K. Morokuma, The Journal of Physical Chemistry Letters **1**, 3328 (2010).

[12] B. Rupp, *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology* (Taylor & Francis Group, 2009).

[13] M. El Khatib, A. Martins, D. Bourgeois, J.-P. Colletier, and V. Adam, Scientific reports **6**, 18459 (2016).

[14] C. C. Liu and P. G. Schultz, Annual review of biochemistry **79**, 413 (2010).

[15] L. M. Oltrogge and S. G. Boxer, ACS central science **1**, 148 (2015).