Studying a reversibly switchable fluorescent protein by time-resolved crystallography using the X-ray free electron lasers

Joyce Woodhouse

To cite this version:

Joyce Woodhouse. Studying a reversibly switchable fluorescent protein by time-resolved crystallography using the X-ray free electron lasers. Biological Physics [physics.bio-ph]. Université Grenoble Alpes, 2018. English. NNT : 2018GREAY026 . tel-02305244

HAL Id: tel-02305244
https://theses.hal.science/tel-02305244

Submitted on 4 Oct 2019

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : **Physique pour les sciences du vivant**

Arrêté ministériel : 25 mai 2016

Présentée par

Joyce WOODHOUSE

Thèse dirigée par **Martin WEIK, Directeur de recherche, Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA)**

préparée au sein de **L'Institut de Biologie Structurale (IBS)**
dans **l'École Doctorale de Physique de Grenoble (EDPG)**

Etude d'une protéine fluorescente photo-commutable par cristallographie résolue en temps en utilisant les lasers à électrons libres

Thèse soutenue publiquement le **3 octobre 2018**,
devant le jury composé de :

Pr. Arwen PEARSON
Center for Free Electron Laser (CFEL), Hambourg, Allemagne, Rapportrice
Pr. Ludovic JULLIEN
Ecole Normale Supérieure (ENS), Paris, France, Rapporteur
Pr. Wim BURMEISTER
Institut de Biologie Structurale (IBS), Grenoble, France, Président
Dr. Daniel PICOT
Institut de Biologie Physico-Chimique (IBPC), Paris, France, Examinateur
Pr. Stephen MEECH
Université de l'Est-Anglie (UEA), Norwich, Royaume-Uni, Examinateur

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : **Physique pour les sciences du vivant**

Arrêté ministériel : 25 mai 2016

Présentée par

Joyce WOODHOUSE

Thèse dirigée par **Martin WEIK, Directeur de recherche, Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA)**

préparée au sein de **L'Institut de Biologie Structurale (IBS)**
dans **l'École Doctorale de Physique de Grenoble (EDPG)**

Studying a reversibly photoswitchable fluorescent protein by time-resolved crystallography using the free electron lasers

Thèse soutenue publiquement le **3 octobre 2018**,
devant le jury composé de :

Pr. Arwen PEARSON
Center for Free Electron Laser (CFEL), Hamburg, Germany, Reporter
Pr. Ludovic JULLIEN
Ecole Normale Supérieure (ENS), Paris, France, Reporter
Pr. Wim BURMEISTER
Institut de Biologie Structurale (IBS), Grenoble, France, President
Dr. Daniel PICOT
Institut de Biologie Physico-Chimique (IBPC), Paris, France, Examiner
Pr. Stephen MEECH
University of East Anglia (UEA), Norwich, United Kingdom, Examiner

Avant-propos

Le rêve d'Arthur

Je suis dans l'espace, avec un vieux… Je vous raconte pas ça au hasard. Quand je me suis réveillé j'ai tout de suite pensé à vous. L'espace… ça a toujours été votre truc ça l'espace. Et les vieux, y en a toujours dans vos histoires à vous. Bref, je flotte dans l'espace… avec les étoiles tout. Et y a un vieux à côté. Alors je sais pas si c'est moi vieux où euh... parce que les rêves c'est toujours le bordel pour ça. Et le vieux me fait :

« Vous êtes prêt à voir le Graal ? »

Alors moi je réponds :

« Oui. »

Alors on se dirige vers une… une grande boule, mais en fait c'est notre Terre à nous. Sauf qu'au lieu d'être bien plate, ben elle est en boule quoi ! Comme je disais les rêves c'est toujours le bordel. On descend, on descend. On atterrit sur un sentier, dans une forêt, sur le territoire du Seigneur Dagonet. Alors me demandez pas pourquoi. D'autant qu'il est même pas dans le rêve ce con-là mais, je sais pas comment vous dire, je sais qu'on est chez Dagonet. Le vieux se retourne et il me fait :

« J'espère que vous avez pas peur de la marche. Parce que je vous préviens c'est pas la porte à côté. »

Moi je lui réponds :

« Je comprends pas pourquoi on a pas atterri directement plus près alors ? »

Et il me répond pas il part devant. Je le suis. Je le suis, je le suis pis au bout d'un moment je me dis :

« Merde… c'est le chemin de Kaamelott ici ! Oh, *que je lui fais au vieux*, c'est pas le chemin de Kaamelott ça ?

- Si pourquoi ?
- Comment pourquoi ? *je lui fais*, le Graal il est pas à Kaamelott quand même ?
- Si, *il me fait*. »

Alors je m'arrête :

« Vous vous foutez de moi ? »

Il se retourne il me dit :

« Vous voulez le voir le Graal où vous voulez pas le voir ? Bon bah bouclez-là et suivez. »

Et il repart. Bon. On arrive à Kaamelott, la baraque vide. Pas un garde à l'entrée, pas un loufiat dans les couloirs, on passe devant la salle de la Table Ronde, pas de Table Ronde, la pièce vide. On continue, on continue et on arrive devant la porte de ma salle de bain :

« Voilà, ouvrez c'est là derrière.

- Mais… là derrière ou ça ? Dans la salle de bain ?
- Oui, dans la salle de bain. »

Alors je le regarde, j'essaie de voir s'il est pas beurré ou quoi, pis je rentre. Là y'a la baignoire vide… enfin vide… y'a de l'eau mais y'a personne dedans. Et y'a du sang partout. Partout, partout… Il me fait :

« Voilà ! C'est le Graal !

- Quoi quel Graal ? La salle de bain ?
- Non pas la salle de bain, la baignoire !
- La baignoire c'est le Graal ?
- Ouais, c'est le récipient qui a reçu le sang du Christ. »

Alors là dans le rêve je lui mets une tarte au vieux… mais la bonne tartine attention, avec la tête qui part sur le côté, les cheveux de travers tout.

« Tu te payes ma gueule ? », *que je lui fais.*

Pis là mon vieux c'est lui qui se retourne, qui vient, il me fout une avoine. J'ai l'impression que le plafond me tombe sur la gueule. Je me ressaisis. Et il me dit :

« Qu'est-ce que c'est que quelqu'un qui souffre ? Et qui fait couler son sang par terre pour que tout le monde soit coupable ? Tous les suicidés sont le Christ. Toutes les baignoires sont le Graal. »

Et vous savez qu'on s'est toujours demandé si y avait pas une inscription gravée au fond du Graal ! Eh ben oui y'en a une.

« Allez voir », *qu'il me fait.*

Alors j'y vais. Et au fond de la baignoire y'a marqué :

« Vous m'avez bien cassé les couilles. »

Et boum je me réveille.

Alexandre Astier, *KAAMELOTT*, Livre VI, Episode 9.

Table of contents

1. Introduction

1.1. Reversibly switchable fluorescent proteins

Fluorescent proteins (FPs) absorb light at one wavelength and emit at another one. The first discovered and the most famous one is the green fluorescent protein (GFP). FPs are valuable tools for fluorescence microscopy, where they are used as genetic markers in cell biology (Adam, 2014). Two families of FPs are known, those originating from Hydrozoa (e.g. GFP) and those from Anthozoa. All FPs share a common fold, consisting in an 11-stranded β-barrel and an α-helix that carries a chromophore formed of three transformed amino-acids. It is the chromophore that absorbs and emits light within an FP. The amino acids forming the chromophore may vary among different FPs, as well as do the amino-acids constituting the chromophore pocket, thus generating a wide range of FPs with a variety of photophysical properties. Variants have been generated by mutation in order to tune or create properties interesting for microscopy applications.

A subgroup of FPs, so-called phototransformable fluorescent proteins (PTFPs), recently enabled the development of super-resolution microscopy (or nanoscopy). PTFPs are FPs that undergo irreversible or reversible transformation upon light absorption. They are divided in three main classes: photoactivatable fluorescent proteins (PAFPs), photoconvertible fluorescent proteins (PCFP) and reversibly switchable fluorescent proteins (RSFPs). Photoactivatable fluorescent proteins are non-fluorescent in their ground state and become fluorescent under light illumination. This *off*-to-*on* activation is an irreversible process involving breakage of a covalent bond and a change in protonation state of the chromophore. The photoactivation mechanism consists in the decarboxylation of a well-conserved glutamate and in the deprotonation of the chromophore that then fluoresces in the anionic form (Adam, 2014). PA-GFP, a variant of GFP is an example of PAFPs used to explore the intracellular traffic (Patterson and Lippincott-Schwartz, 2002). Photoconvertible fluorescent proteins (PCFPs) are fluorescent in their resting state and undergo a photoinduced shift in absorption and emission wavelengths. In other words, they change color. Photoconversion is an irreversible process involving covalent bond breaking. Kaede (Ando et al., 2002), EosFP (Wiedenmann et al., 2004) and the variants of the latter are examples of green-to-red PCFPs, which are used for various multicolor nanoscopy approaches (Adam, 2014). Reversibly switchable fluorescent proteins (RSFPs), or photochromic FPs, undergo a photoinduced change of fluorescent state. They can either switch from a fluorescent to a non-fluorescent state (RSFPs called negative switchers) or from a non-fluorescent to a fluorescent state (RSFPs called positive switchers). Photoswitching is a reversible process involving a chromophore rearrangement (isomerizaton in most cases) and a change in the protonation state of the chromophore. The most well-studied RSFP is the negative switcher Dronpa (Ando et al., 2004) whose molecular mechanism has been extensively studied, as discussed below. Padron (Faro et al., 2011), a variant of Dronpa, and asFP595 (Lukyanov et al., 2000) are examples of positive switchers. RSFPs are exploited for nanoscopy techniques, including REversible Saturable Optical Linear Fluorescence Transitions (RESOLFT) or photochromic Stochastic Optical Fluctuation Imaging (pcSOFI) (Adam, 2014). There is also the special case of biphotochromic FPs, which are able to undergo photoswitching and photoconversion. IrisFP (Adam et al., 2008), a variant of EosFP, is an example of a biphotochromic FP.

The photoswitching mechanism in RSFPs has been extensively investigated to further our understanding of this class of FPs and in order to design improved RSFPs for imaging applications. Dronpa and its variants are the most widely studied. The three-dimensional structure of Dronpa, solved by cryo-crystallography, revealed that the chromophore in the fluorescent *off*-state is a *cis* isomer and that the *on*-to-*off* switching involves a *cis*-to-*trans* isomerization (Andresen et al., 2007). NMR studies at ambient temperature revealed an increased flexibility of the β-barrel and the chromophore in the *off*-state (Mizuno et al., 2008). The structural investigation of the positive switcher asFP595 also revealed a *cis* and a *trans* chromophore in the *on*- and the *off*-state, respectively and molecular dynamics (MD) simulations predicted that isomerization involves a “hula-twist” motion of the chromophore (Andresen et al., 2005). The photoswitching mechanism not only involves structural but also protonation-state changes as evidenced by electronic and vibrational spectroscopy approaches. Until recently, the sequential order of isomerization and (de)protonation was not clear, and if a scenario is now widely accepted, some aspects remain controversially discussed. Spectroscopic characterization of Dronpa suggested that an excited-state proton transfer (ESPT) is involved in the *off*-to-*on* switching (Fron et al., 2007), while time-resolved infrared (TRIR) measurements identified the proton transfer being a ground-state process (Warren et al., 2013). However, another TR-IR study on Dronpa2 (M159T variant) advocated that both isomerization and deprotonation are ground-state processes, attributing the picosecond spectroscopic changes in the excited state to protein conformational changes priming the chromophore for switching (Lukacs et al., 2013), a view corroborated by a recent follow-up TR-IR study on Dronpa2 (Laptenok et al., 2018) that provided evidence for a *cis*-protonated chromophore species being formed with a 91-ns time constant. Results from transient absorption spectroscopy on Dronpa and Dronpa2, from 100 fs to milliseconds (Yadav et al., 2015), are in line with the conclusions of Warren et al., suggesting that during *off*-to-*on* switching, the excited state is formed in a few hundreds of femtoseconds and isomerization occurs concomitantly with its decay on the picosecond time-scale. Deprotonation then takes place on the microsecond time scale (Yadav et al. 2015). In summary, the most widely accepted, but still disputed (Laptenok et al., 2018), scenario for *off*-to-*on* switching in Dronpa and Dronpa2 seems to be excited-state isomerization followed by ground-state deprotonation.

Elucidating the structures of the various photoswitching intermediates is essential to fully understand the mechanism of RSFPs. Time-resolved crystallography, introduced below, is the method of choice to do so. However, as reported in the various spectroscopy studies outlined above, short-lived intermediates are formed in the excited state on a time scale inaccessible for time-resolved crystallography at synchrotron sources (see details below). Luckily, ultra-short pulses produced by a new kind of X-ray sources, so-called X-ray free electron lasers (XFELs), provide the possibility of filling the lack of structural data on the ultra-fast time-scale to eventually understand the photoswitching mechanism. Recently, the first static structure of an FP was solved using an XFEL, *i.e.* that of the biphotochromic FP IrisFP (Colletier et al., 2016b), paving the way for time-resolved crystallography on the ultra-fast time scale at XFELs. The study also presented transient absorption spectroscopy results revealing a switching mechanism similar to Dronpa.

The reversibly switchable enhanced green fluorescent protein (rsEGFP2, Grotjohann et al., 2012) is a variant of rsEGFP (Grotjohann et al., 2011), a reversibly switchable version of the enhanced GFP (EGFP). rsEGFP2 is a marker used in RESOLFT, a super-resolution microscopy method based on the use of RSFPs. rsEGFP2 was designed from rsEGFP in order to obtain a fast switcher that reduces the recording time during a RESOLFT experiment. The protein presents the common β-barrel fold shared by all FPs (Figure 1a). The spectra of the fluorescent *on*-state and the non-fluorescent *off*-state showed maxima at 479 nm and 403 nm respectively, characteristics of a anionic and neutral chromophore respectively (Figure 1c, El Khatib et al., 2016). Belonging to the class of negative switchers, rsEGFP2 is fluorescent in its resting state and undergoes *on*-to-*off* and *off*-to-*on* switching under 488-nm and 405-nm illumination, respectively. The life time of the non-fluorescent state is about 100 min in the crystalline protein (Schirò et al., 2017). The chromophore, hydroxybenzylidene imidazolinone

(HBI) formed from the tripeptide Ser65–Tyr66–Gly67, undergoes *cis*/*trans* isomerization during photoswitching (Figure 1c), accompanied by a side-chain rearrangement of amino acids in the chromophore pocket (El Khatib et al., 2016) and chromophore (de)protonation (Grotjohann et al., 2012), as usually observed in other RSFPs.

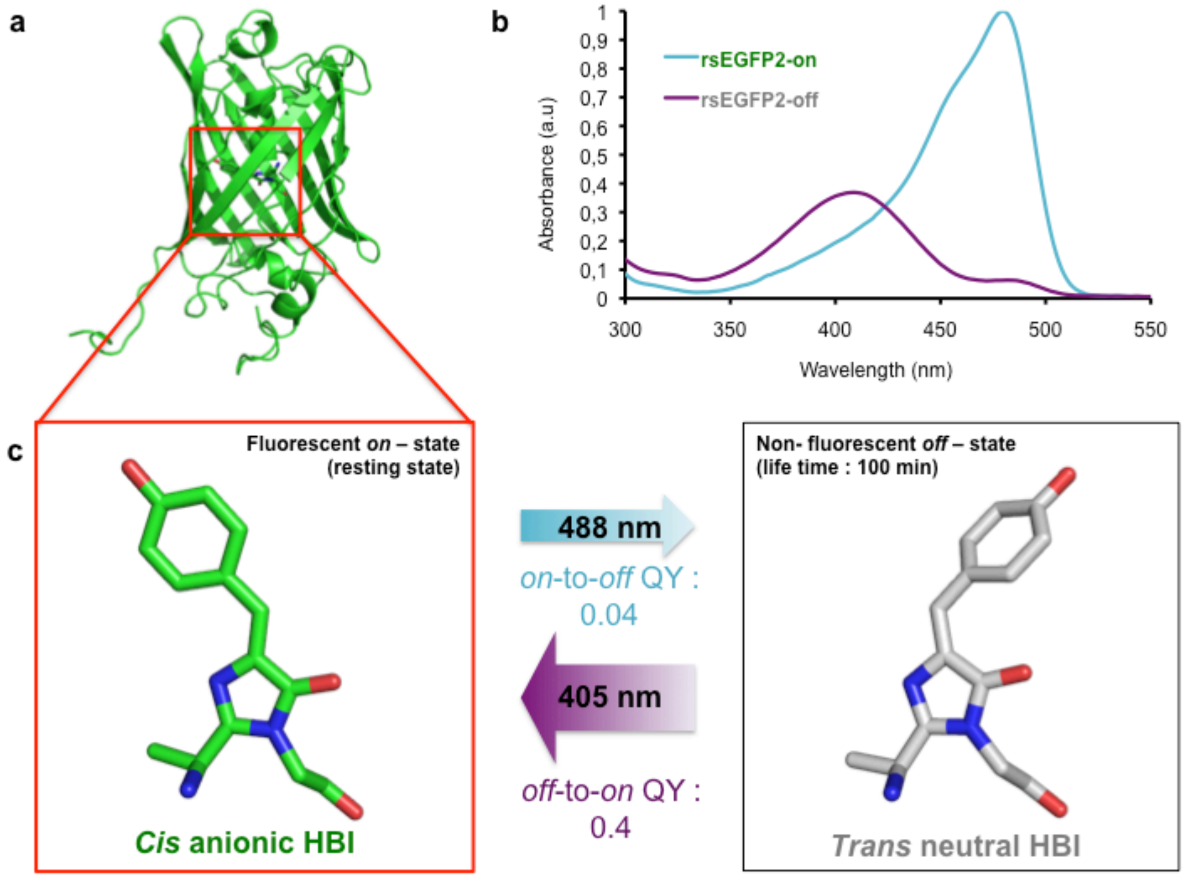


Figure 1: Spectral and structural properties of the reversibly switchable enhanced green fluorescent protein 2 (rsEGFP2). (a) Cryo-crystallography structure of the fluorescent *on*-state (PDB entry: 5DTX, El Khatib et al., 2016). (b) Absorption spectra of the *on*- and *off*-states of rsEGFP2 shown in cyan and magenta, respectively. The absorption maxima are 479 and 403 nm for the *on*- and the *off*-state, respectively (El Khatib et al., 2016). (c) The mechanism of the photoswitching of rsEGFP2. rsEGFP2 is fluorescent in its resting state, carrying an anionic *cis* hydroxybenzedilene imidazolinone (HBI) chromophore. Under 488-nm illumination, it undergoes *cis*-to-*trans* isomerization and a change in protonation state. The non-fluorescent state of the crystalline protein presents a life-time of ~100 min (Schirò et al., 2017) and carries a neutral *trans* chromophore. Under 405-nm illumination, the protein is switched back. The *off*-to-*on* switching quantum yield (QY) is 10 times higher than the one in the reverse *on*-to-*off* direction.

1.2. Time-resolved crystallography

1.2.1. Pump-probe principle and time-resolved crystallography using synchrotron radiation

A protein crystal is not a rigid entity that prohibits protein motions. To the contrary, since a protein crystal is composed on average of 50 % solvent, motions of side chains and even of secondary-structure elements that are not involved in crystal packing are allowed. Consequently, most crystalline proteins remain biologically active (Mozzarelli and Rossi, 1996). It is thus possible to study functionally relevant movements by triggering the reaction in all crystalline molecules simultaneously and by subsequently trapping reaction intermediates with kinetic crystallography approaches. One such an approach is the cryo-trapping strategy consisting in first triggering and then stopping the reaction by flash-cooling the crystal in order to trap a reaction intermediate (Bourgeois and Royant, 2005). Another strategy consists in determining the structure of an intermediate on the fly, in real time. This so-called time-resolved crystallography approach, based on the pump-probe principle, allows to obtain room-temperature snapshots of proteins *at work* that can be assembled to form a molecular movie of the reaction (Levantino et al., 2015b).

The pump-probe principle applied in time-resolved crystallography (Figure 2), mostly uses a laser pulse (*pump*) to initiate the reaction in the crystal (defined as time t=0). The use of a UV-vis laser, for instance, permits to trigger a reaction in a large fraction of photosensitive proteins in a crystal simultaneously. X-rays (*probe*) then interact with the sample after a defined, but variable, time delay (Δt_i). Diffraction patterns are collected, and processed for all Δt_i until a complete dataset of structure factors is formed. The procedure is repeated at a different Δt. Fourier difference maps between the data sets generated at various Δt after the pump and the initial data set (t=0) reveal features indicating the conformational changes that occurred at the different time delays, thereby allowing a molecular movie of the investigated mechanism to be made.

Mainly two model proteins were extensively studied by time-resolved crystallography: the photoactive yellow protein (PYP) and myoglobin. PYP is a photosensitive protein carrying a chromophore that isomerizes after light absorption and undergoes structural changes. Regarding the PYP, the fact that the reaction can be triggered by light makes this protein very convenient for time-resolved studies. The structure of the light-induced state of PYP that had accumulated 10 ms after laser triggering was solved by time-resolved Laue crystallography at a synchrotron sources and revealed *trans*-to-*cis* isomerization of the chromophore and accompanying side-chain rearrangements (Genick et al., 1997). Earlier intermediates were solved after having improved the temporal resolution to the nanosceond time-scale (Perman et al., 1998). The authors argued that observed features in the difference electron density map corresponded to a short-lived intermediate that is a precursor of longer-lived late intermediates. The earliest intermediate of the PYP photocycle was caught with a 100-ps temporal resolution, the limit reachable at a synchrotron source, and allowed to propose a *trans*-to-*cis* isomerization mechanism involving two different pathways controlled by the chromophore environment (Jung et al., 2013). An alternative mechanism was proposed based on a parallel time-resolved crystallography experiment (Schotte et al., 2012).

Myoglobin is another model protein widely chosen for investigating protein dynamics. Photodissociation of carbon monoxide (CO) from the heme that is carried by the protein makes time resolved crystallography experiments on myoglobin feasible. On the nano- to the millisecond time-scale, release and recombination of CO after photodissociation and the structural changes accompanying them have been followed (Srajer et al., 1996) and structural rearrangements during CO migration have been reported (Bourgeois et al., 2003). The temporal resolution was extended to 100 ps and allowed trapping of a short-lived intermediate of CO dissociation from a myoglobin mutant (Schotte et al., 2003).

Time-resolved crystallography at synchrotron sources suffers from several limitations. First, the need to work at or close to room temperature makes the crystal more sensitive to radiation damage that convolutes with the conformational changes to be investigated. Second, large crystals are often required for a sufficient diffraction signal, rendering efficient illumination with the pump laser difficult because of the high optical density. Third, the most-often employed Laue crystallography, using a polychromatic beam in order to recover a larger fraction of the reciprocal space in a time-limited exposure (to prevent radiation damage), generates diffraction patterns that are difficult to analyze, in particular for highly mosaic crystals or for crystals with large unit cells. Fourth, Laue crystallography mostly allows studying reversible processes (e.g. dissociation and recombination of CO in myoglobin, reversible photoisomerization in PYP), whereas investigations of irreversible processes, such as enzyme catalysis, are cumbersome but possible (Schlichting et al., 1990). And last, time-resolved crystallography at synchrotrons is limited to 100-ps resolution by the electron-pulse length, prohibiting the investigation of ultrafast events, such as early intermediates involved in photochemical processes.

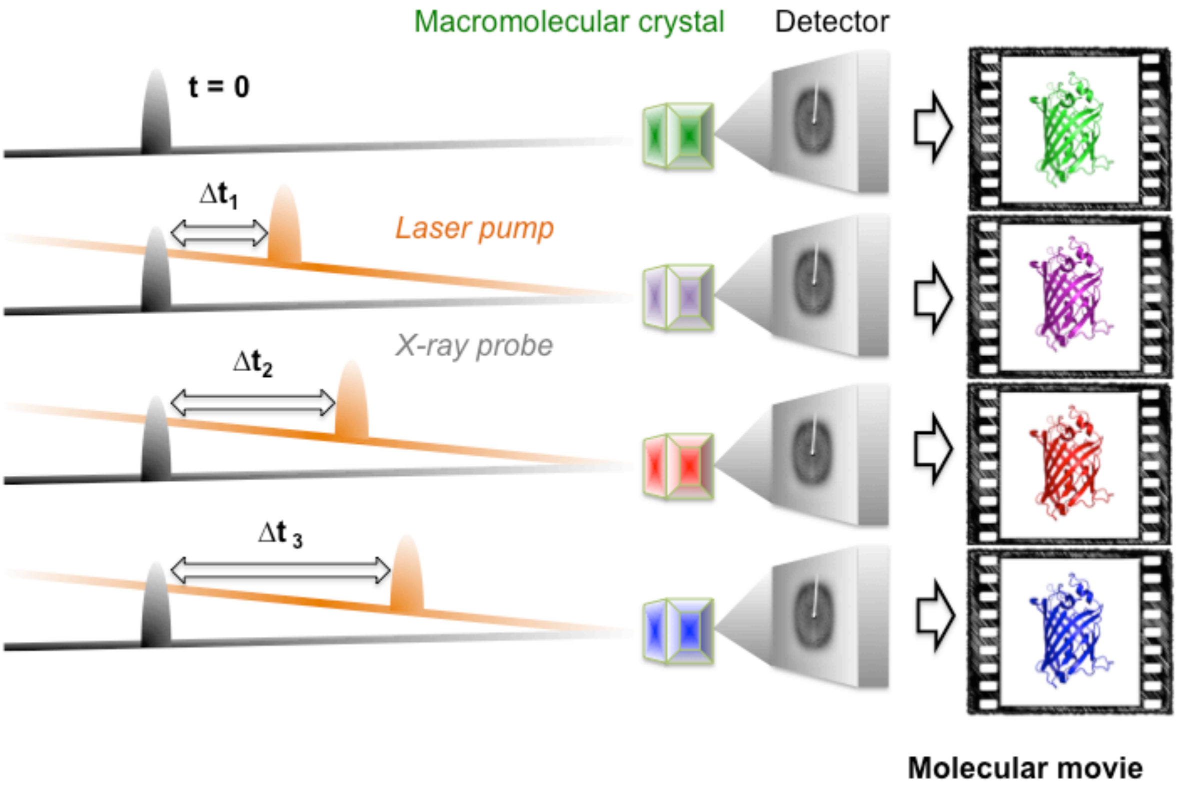


Figure 2: Experimental set-up of a time-resolved crystallography experiment carried out according to the pump-probe principle. At t=0, crystallographic data are collected on a crystal without prior laser pump and the resting-state structure (showed in green) is solved as a reference. The time-delay Δt_i between the laser pump (orange beam) that initiates the reaction and the X-ray probe (black beam) is set. In this example, data are collected for Δt_1, Δt_2 and Δt_3, and structures of the purple, red and blue intermediates are solved, respectively, that are then used as frames in a molecular movie.

1.2.2. Time-resolved crystallography using X-ray free electron lasers (XFELs)

1.2.2.1 X-ray free electron laser and serial crystallography

X-ray free electron lasers are a new kind of X-ray source based on self-amplification of stimulated emission (SASE, Margaritondo and Rebernik Ribic, 2011). Electrons are injected into a linear accelerator and then pass through an undulator where they oscillate under successive magnetic fields of opposite orientations leading to the emission of X-rays. The resonant interaction between X-rays and electrons leads to microbunching of the electron bunches and to an increase in the intensity of the X-ray radiation. Eventually, ultra-intense and coherent X-ray pulses are generated with a peak of brilliance 10 orders of magnitude higher than the one of synchrotron radiation and pulse lengths of 10 to 100 fs. In contrary to synchrotron radiation, X-rays produced by free-electron lasers suffer from pulse to pulse wavelength instability.

In the year 2000, it was proposed based on simulations that with ultra-short and ultra-bright X-ray pulses, it would be possible to recover structural information from a biological sample before it starts to suffer from radiation (Neutze et al., 2000). This so-called “diffract-before-destroy” principle promised that by using XFELs for structural biology, the issue of radiation damage would be overcome. The Linac Coherent Light Source (LCLS, Stanford, California) was the first operational XFEL (Emma et al., 2010), followed by the Spring-8 Angstrom Compact Laser (SACLA, Sayo, Japan) a few years later (Ishikawa et al., 2012). The first macromolecular crystallography experiment using XFELs was carried out at the LCLS. A myriad of randomly oriented nanocrystals of photosystem (PS) I were continuously presented to the X-ray beam, allowing to collect diffraction data that provided the first X-ray structure obtained with XFEL radiation (Chapman et al., 2011), giving birth to serial femtosecond crystallography (SFX). Figure 3 illustrates the set-up for an SFX experiment. The spatial resolution in this first experiment was low (8.5 Å) because of the still limited X-ray wavelength of the instrument (6.7 Å). However, the “diffract-before-destroy” principle was experimentally confirmed, and it was proven that crystallographic data can be collected using ultra-bright and ultra-short pulses, without apparent damage (but see Barty et al., 2012, and Lomb et al., 2011), even at near-atomic resolution (Boutet et al., 2012). By combining *in vivo* crystallization and SFX, the structure of the natively inhibited *Trypanosoma brucei* cystein protease cathepsin B (TbCatB) was solved, revealing the mechanism of its inhibition (Redecke et al., 2013) and providing the first biologically relevant structure solved using an

XFEL. The structure of a G protein-coupled receptor (GPCR) was then solved by SFX from microcrystals grown in lipidic cubic phase (LCP), proving the applicability of the method and of the instruments for membrane protein structural investigation (Liu et al., 2013). The first room-temperature structure of PS II with an undamaged Mn_4CaO_5 cluster was obtained, but at limited resolution (Kern et al., 2012). Femtosecond X-ray emission spectroscopy using XFEL pulses demonstrated that metallic clusters are not damaged by the X-ray pulse on the tens of femtoseconds time scale (Alonso-Mori et al., 2012). X-ray diffraction and emission spectroscopy data were simultaneously collected, yielding an undamaged structure of PS II, still at limited resolution, but with an unperturbed redox state of the Mn_4CaO_5 cluster (Kern et al., 2013). This study evidenced the reliability of the method for metalloprotein studies. Other metalloproteins such as copper nitrite reductase (Fukuda et al., 2016) and cytochrome c oxidase (Andersson et al., 2017) have been studied supporting this reliability. The first high-resolution structure of undamaged PS II was solved from SFX data collected at different positions of several macrocrystals at cryo-temperature (Suga et al., 2015). It allowed to obtain the structure of the native Mn_4CaO_5 cluster of the oxygen-evolving complex that displayed slight differences with the structure solved using synchrotron radiation. The usefulness of XFELs for structural biology was further underlined by the determination of the rhodopsin-arrestin complex (Kang et al., 2015) that unveiled the interaction between rhodopsin and arrestin and provided structural data to understand the arrestin-mediated signaling mechanism. The study was only possible at an XFEL, since microcrystals of the complex only provided low-resolution data at a synchrotron facility. Large protein complexes often imply large unit cells of the crystal that could limit the reachable diffraction resolution, and even more so in the case of nanometric crystals. Limited resolution often also implies potential crystal disorder. Exploiting the property of so-called imperfect crystals, the resolution limit was extended by using continuous diffraction in between Bragg peaks (Ayyer et al., 2016), offering the possibility to reach near atomic resolution for structure determination of large complexes using XFEL radiation.

Experimental phasing of SFX data has been an issue for a long time, making solution of unknown protein structures impossible. The first *de novo* structure was solved using the anomalous signal from heavy atoms, thereby proving that sufficient data quality can be reached in SFX from microcrystals (Barends et al., 2014). However, a huge amount of indexed frames was necessary to reach the data redundancy required for experimental phasing. The authors predicted the possibility of solving *de novo* structures of difficult proteins such as membrane proteins. Indeed, in the following, bacteriorhodopsin (Nakane et al., 2016) and GPCRs (Batyuk et al., 2016) were phased experimentally. The anomalous signal of sulfur atoms present in lysozyme was detected, but was initially too weak to allow for structure solution (Barends et al., 2013). The further development of data processing tools improved data quality and eventually allowed solving a protein structure by exploiting the anomalous signal (Nakane et al., 2015). In the following, an approach was proposed that allowed *de novo* structure determination of thaumatin by single wavelength anomalous diffraction phasing with a reduced number of indexed frames (Nass et al., 2016). The structure of the mosquito larvicide BinAB was the first new structure solved using experimental phasing, with data collected on *in vivo* grown nanocrystals (Colletier et al., 2016a). This study showed that SFX allows collection of data from nanometer-sized crystals that is of sufficient quality to undertake experimental phasing for *de novo* determination of unknown protein structures. Using a two-color mode (Marinelli et al., 2015), in which the XFEL produces two pulses with different wavelengths and with a tunable time-delay, multi-wavelength anomalous diffraction is now to be considered for *de novo* structure determination (Gorel et al., 2017).

Sample preparation is a critical step for an SFX experiment. But there are only a few microcrystallization methods reported to the best of our knowledge. Crystallization using lipidic cubic phase (LCP) is well described (Caffrey and Cherezov, 2009, Liu et al., 2014). The LCP is a pseudo-crystallin material that mimics the membrane environment, making the methods suitable for membrane protein crystallization. The lipidic sponge phase (LSP), a material related to LCP, was also used for crystallization of the photosynthetic reaction center (Johansson et al., 2012) and provided crystals that allowed structure determination at 3.5 Å resolution (Johansson et al., 2013). Microcrystallization of PS II was briefly investigated. Different approaches including batch and free-interface diffusion (FID) methods consisting in triggering the nucleation at the contact zone between highly concentrated protein and precipitant solutions were reported (Kupitz et al., 2014b). Microseeding was applied to improve the morphological aspect and the diffraction quality of PS II microcrystals (Ibrahim et al., 2015). The serendipitous discovery that TbCatB microcrystallisation occurs in Sf9 insect cells (Koopmann et al., 2012) led to the rationalization of *in vivo* crystallization using different expression systems (Duszenko et al., 2015). Marginally, it was reported that supersaturation-controlled microcrystallization reproducibly generates crystals at a desired size by varying the evaporation time of the drop (Lee et al., 2018). The crystals required for an SFX experiment are often micro- or nanometric in size, depending on the type of the experiment. Such small crystals are difficult to detect with optical microscopes. Several other methods can be used for crystal detection, such as second-order nonlinear optical imaging of chiral crystals (SONICC, Wampler et al., 2008) or electron microscopy (Stevenson et al., 2014). The latter was useful in the process of optimizing crystallization (Stevenson et al., 2016). XFELs produce X-ray pulses bright enough to obtain high-resolution diffraction from microcrystals that could not be obtained with synchrotron radiation (see the rhodopsin-arrestin example above). In the case of a time-resolved crystallography experiment, microcrystals are required to ensure sufficient penetration of the pump light to efficiently trigger the reaction. The lack of a method to reproducibly obtain microcrystals of homogeneous size and of good quality has to be filled.

Sample presentation to the XFEL beam is also a critical step in an SFX experiment. Given that a protein crystal is destroyed immediately after interaction with the intense X-ray pulse, the sample has to be continuously refreshed. Various sample delivery methods have been developed, among which the most adequate one needs to be chosen wisely for a given SFX experiment. The first adopted method was liquid injection using a gas dynamic virtual nozzle (GDVN). It consists in focusing a liquid jet with gas under pressure until its diameter reaches a dimension comparable to the X-ray beam diameter (DePonte et al., 2008, Weierstall, 2014). The main drawback of this method is the formidably huge amount of crystalline material that is required and of which about 99.9 % is wasted during the experiment. To fix this issue, other approaches were proposed. An up-graded version of the GDVN consists in using a focusing liquid (non-miscible with the mother liquor), itself focused by a gas stream, resulting in reduced background scattering and sample consumption (Oberthuer et al., 2017). Using an electric field instead of gas dynamics to focus the liquid jet is an alternative solution to limit sample consumption (Sierra et al., 2012, Sierra et al., 2016). Indeed, this electrospun liquid microjet forms droplets at the nozzle exit that reduces sample consumption in the end. Also sound waves are used for droplet ejection (Roessler et al., 2016). The so-called drop-on-demand acoustic ejector has been combined with a conveyer belt drive to present the crystal-containing droplets to the X-ray beam (Fuller et al., 2017). Combined with a pump-laser set-up makes this sample-delivery system suitable for time-resolved X-ray crystallography. Furthermore, an LCP injector was conceived to extrude viscous media containing crystals, thereby slowing down injection (Weierstall et al., 2014, Botha et al., 2015). It presents the advantage to be less sample-consuming and is suitable for studying membrane proteins that crystallized in LCP and that can thus be injected directly in their crystallization medium. Other viscous media were used to inject crystals of soluble proteins. Grease-based media (Sugahara et al., 2015), vaseline (Botha et al., 2015) and agarose (Conrad et al., 2015) have been successfully used for structure determination. As for LCP, the use of viscous media for sample extrusion limits sample consumption, but increases the x-ray scattering background. It was demonstrated that hyaluronic acid-based (Sugahara et al., 2016) or cellulose-based (Sugahara et al., 2017) matrices as alternative viscous media have the advantage to minimize the background issue. Protein crystals are very sensitive and sometimes the mother liquor is not compatible with a given viscous medium. A systematic study reported on chemical compatibility, injection reliability and scattering background of various viscous media (Kovácsová et al., 2017). Solid supports were also developed that, combined with the raster-scanning approach during data collection (Hunter et al., 2014, Roedig et al., 2017), also lead to very low sample consumption. Instead of collecting only one diffraction pattern per microcrystal, as it is the case for all sample delivery methods presented above, several patterns can be collected from several position on a single crystal mounted on a loop at cryogenic temperature (Hirata et al., 2014, Halsted et al., 2018). In this approach called serial femtosecond rotation crystallography (SF-ROX), the crystal is hit at each position after translating and rotating the loop. For a time-resolved SFX experiment (TR-SFX), liquid jet injection is the method of choice (if sufficient sample can be produced) because of stably-running jets and low background scattering. To the best of our knowledge, no TR-SFX experiment using raster-scanning or SF-ROX has been reported. Using LCP or viscous media in TR-SFX can be problematic because of the matrix opacity that potentially limits the pump-efficiency. However, LCP injection has been successfully used in studies of the membrane protein bR (Nango et al., 2016; Nogly et al., 2018, 2016).

The new serial-crystallography approach required new types of crystallographic data processing procedures to be developed. The challenge was to deal with hundreds of thousands of diffraction patterns obtained each from randomly oriented micro- and nanometric crystals of various sizes and shapes. The Monte-Carlo integration method was applied for merging intensities and determination of structure factors. The method was first assessed based on simulated diffraction patterns (Kirian et al., 2010) and then successfully applied to experimental data (Kirian et al., 2011). New tools and pieces of software were developed to implement this new approached of data processing. *CrystFEL* was created to deal with indexing, integration, and merging of diffraction intensities and with evaluation of the data quality (White et al., 2012). The pre-processing software *Cheetah* reduces data volume by finding “hits” (frames containing diffraction patterns), identifying Bragg peaks (facilitating the use of processing software), correcting detector artifacts, subtracting detector background, generating statistics on hit-rates and determining the resolution limit (Barty et al., 2014). The data processing pipeline combining *Cheetah* and *CrystFEL* is described in (White et al., 2013). Other data processing suites are available, such as *cctbx.xfel* (part of the *cctbx* platform, Sauter et al., 2013) and *nXDS* (Kabsch, 2014), a version of *XDS* adapted for serial crystallography (see Sawaya et al., 2014 for a comparison between *CrystFEL* and *cctbx.xfel*, and Botha et al., 2015 for a comparison between *CrystFEL* and *nXDS*). Other pieces of pre-processing software are also available such as *NanoPeakCell* (Coquelle et al., 2015) and *CASS* (Foucar, 2016). Further development of the available software and the addition of new functionalities drastically improved data quality, allowing reducing the number of frames required for a complete dataset and addressing issues inherent to SFX (intensity partialities, wavelength variability, cell parameter distribution, jet instability…). Intensity scaling, the establishment of a partiality model and the post-refinement procedure (including crystal orientation, unit cell dimensions, detector geometry and detector distance) were implemented in the latest versions of *cctbx.xfel* (Sauter, 2015, Ginn et al., 2015) and of *CrystFEL* (White et al., 2016) via the addition of the *partialator* module (White, 2014).

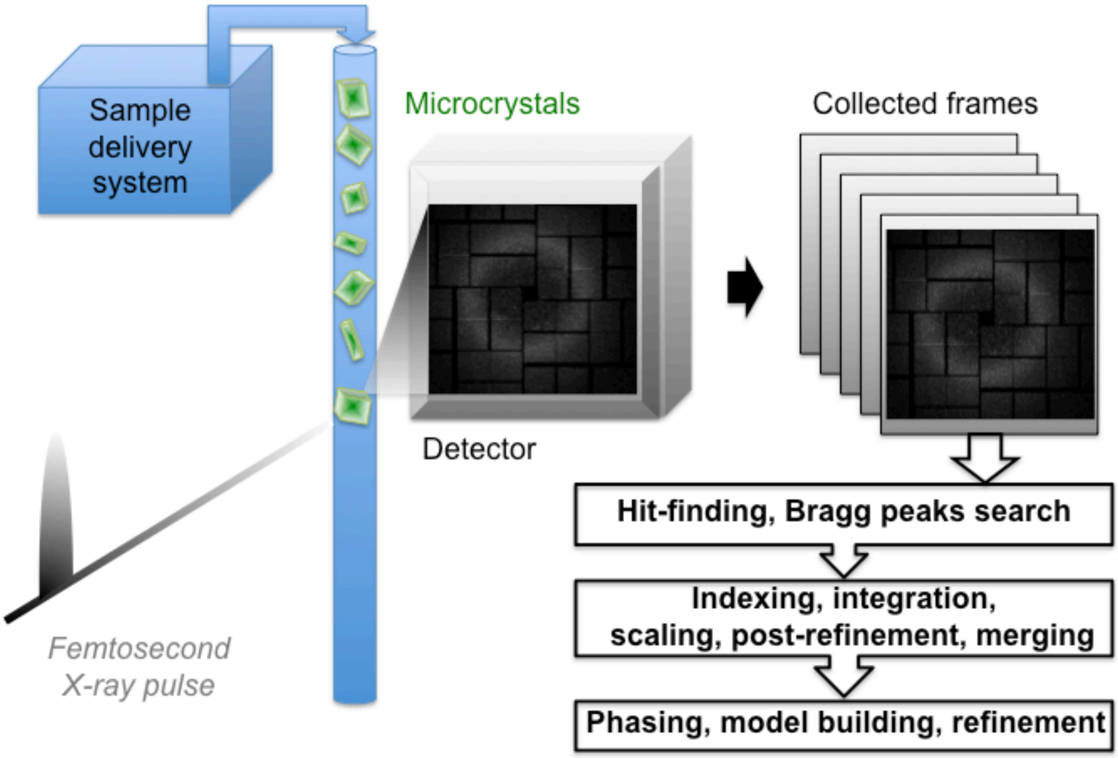


Figure 3: Experimental set-up and data processing pipeline of a serial femtosecond crystallography (SFX) experiment. A suspension of microcrystals (green) is injected across the X-ray beam (shown in black) using a sample delivery system (liquid injector in the example shown). Large amounts of frames are collected from randomly oriented microcrystals and are analysed in the pre-processing (hit-finding, Bragg peak search) and processing (indexing, Monte-Carlo integration, scaling, post-refinement, merging) pipeline. Structure factors are recovered and used for refinement.

Even though serial femtosecond crystallography was successfully developed and applied, access to XFEL sources is still very limited. Indeed, only two facilities are fully operational world-wide (LCLS in California and SACLA in Japan). Others are starting to offer user access, such as the European XFEL (Germany) and PAL-XFEL (South Korea) and the SwissFEL (Switzerland) is being commissioned. However, the fact that the entire XFEL facility produces X-rays for only one (LCLS, SACLA, SwissFEL) or a handful (EuXFEL, PAL-XFEL) of beamlines at a time means beamtime will still remain limited. In parallel, the serial crystallography approach was implemented at synchrotron sources, giving birth to so-called serial synchrotron crystallography (SSX) or serial millisecond crystallography (SMX). With synchrotron radiation the "diffract-before-destroy" principle is not applicable because radicals spread on the time-scale of X-ray exposure and cause chemical and structural damage to macromolecular crystals (Garman and Weik, 2017). However, the serial approach allows spreading the absorbed dose over a multitude of crystals, thereby limiting radiation damage and providing an effective way for room-temperature data collection at synchrotrons. Most sample delivery methods presented above have been adapted to SSX. Data were collected on crystals flowing in their mother liquor through a capillary (Stellato et al., 2014) or embedded in a grease-based viscous medium that is extruded across the X-ray beam (Botha et al., 2015). Solid supports were also used for sample presentation at synchrotrons (Coquelle et al., 2015). However, radiation damage was observed in structures based on data collected using nano- and microfocused beamlines with exposure times of 100 and 200 ms, respectively. High-speed fixed-target sample presentation methods were developed for SSX (and SFX) in order to reduce the time required for recovering a complete dataset (Sherrell et al., 2015) and was then used for proposing a low-dose strategy allowing to solve structure that greatly minimizes radiation damage (Owen et al., 2017). Again exploiting the fixed-target method, devices were designed for protein crystallization and *in situ* X-ray diffraction (Axford et al., 2016, Heymann et al., 2014, Murray et al., 2015). LCP injectors were also used for membrane protein study by SSX (Martin-Garcia et al., 2017). The structure of bR was solved by SSX at room temperature and compared with the cryogenic structure, revealing slight differences, corroborating the interest for room-temperature data collection and the reliability of SSX for room-temperature diffraction experiments. The term "serial crystallography" was initially used for data collection strategies using several randomly oriented crystals. There is a fundamental difference between oscillation methods and serial crystallography that lays in the way data are processed. The Monte-Carlo integration of partial diffraction intensities recorded from randomly oriented crystals with varying size and shape can be considered a fundamental requirement of serial crystallography. Methods inspired by the serial approach have been developed at synchrotron facilities and are abusively considered as “serial crystallography”. For example, the structure of TbCatB has been solved using helicoidal data collection at a synchrotron on microcrystals mounted on a loop (Gati et al., 2014). Because of the rotation, consecutive frames were collected, indexed and merged as is usually done for oscillation data. The same is true for the mesh-and-collect approach (Zander et al., 2015). After scanning the grid, the classical oscillation data collection strategy is applied on each detected crystal, providing partial data sets that are processed and merged as done for standard oscillation data collection. Also the SF-ROX approach (presented above) is rather a multi-crystal than a serial approach, given that diffraction patterns are collected from rotating single crystals and processed as oscillation data.

1.2.2.2 Time-resolved serial femtosecond crystallography

As outlined above, time-resolved crystallography at synchrotron sources is limited by the temporal resolution of 100 ps at best, by the need of large crystals that may be difficult to illuminate, and by the difficulty of investigating irreversible reactions. By providing ultra-short and ultra-bright X-ray pulses, XFELs extend the temporal resolution to the sub-picosecond time-scale and permit working on nano- and microcrystals that are more efficiently illuminated in a pump-probe experiment. The need to continuously refresh the sample anyway makes the study of irreversible reactions, such as enzyme catalysis, feasible. Figure 4 illustrates an SFX set-up adapted for time-resolved crystallography. The proof of principle was provided with time-resolved X-ray diffraction from PS I – ferrodoxin co-crystals collected at time delays of 5 and 10 μs in order to follow the reduction of NADP+, an irreversible reaction (Aquila et al., 2012). Even though not enough data were collected for structure determination, intensity changes indicated that light-induced changes had happened, validating the setup developed for time-resolved experiments at XFELs.

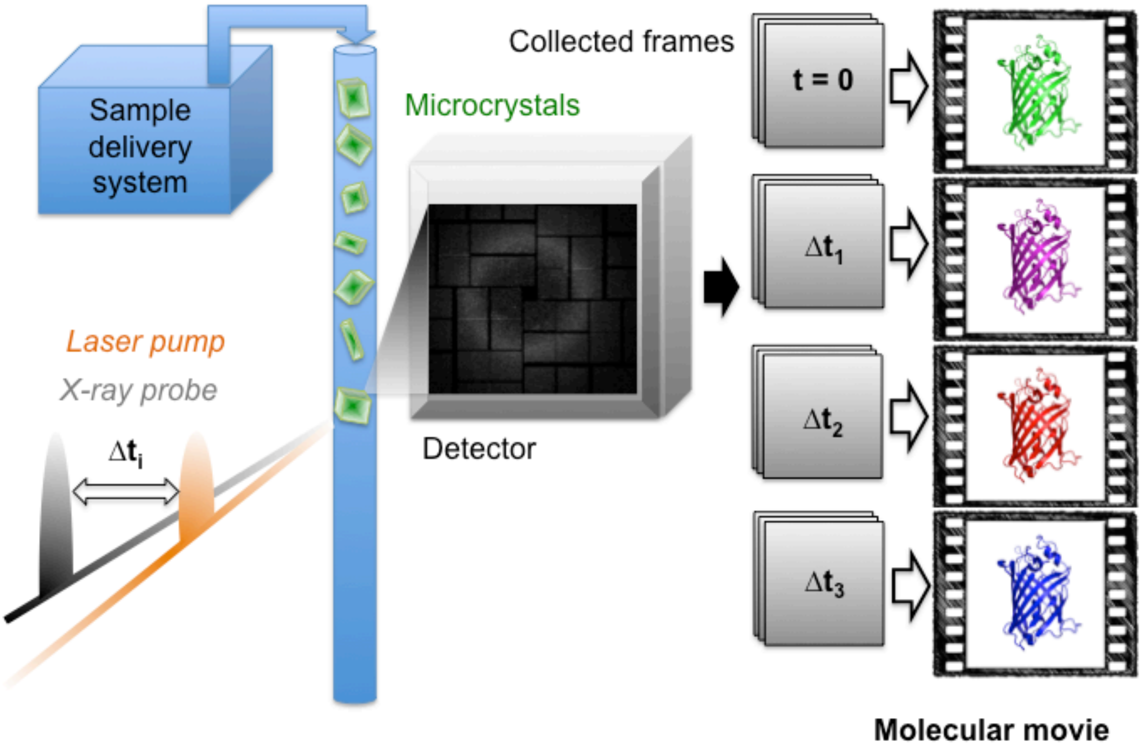


Figure 4: Experimental setup of a time-resolved serial femtoscond crystallography experiment (TR-SFX). The difference with the serial crystallography setup presented in figure 3 is that microcrystals are pumped by the optical laser to initiate the reaction in crystalline proteins. A time delay Δt_i is chosen between the laser pump (orange beam) and the X-ray probe (black beam). Data are collected for each time delay and for the reference state of the protein without laser pump (t=0), and processed following the serial crystallography data processing pipeline. Structures corresponding to each time delay are solved and assembled as a molecular movie.

Photosystem II (PS II) has been extensively studied by SFX with the ultimate goal of solving the mechanism of water oxidation and oxygen production by plants. PS II, being a large membrane protein complex with a metal cluster, is a challenging protein for synchrotron-based structural investigation and thus ideal for being studied by SFX at an XFEL. In a TR-SFX experiment on PS II (Kupitz et al., 2014a), the authors claimed that conformational changes were observed 570 µs after a series of two 527-nm laser flashes. However, it was argued later that no conformational changes were identifiable in the electron density maps (Sauter et al., 2016). Indeed, simultaneous time-resolved X-spectroscopy and SFX, performed by another team, using the same setup than reported in Kern et al., 2013, showed that 250 µs after the second laser pulse, the redox state of the Mn_4CaO_5 cluster remained unchanged and that no structural changes are observable at 4.5 Å resolution (Kern et al., 2014). In the following, a higher-resolution room temperature structure of PS II bond to ammonia, a water analogue, allowed proposing a water-splitting mechanism (Young et al., 2016). In parallel, another team carried out TR-SFX at room temperature and proposed an alternative mechanism of O=O bond formation during water oxidation (Suga et al., 2017).

As summarized above, time-resolved Laue crystallography at synchrotron sources allowed studying PYP after photon absorption and provided clues to describe the photocycle structurally. The structures of the 10-ns and 1-µs photo-intermediates were then also solved using femtosecond pulses provided by an XFEL (Tenboer et al., 2014). Even though the study did not provide any novelty regarding the photocycle undergone by the protein, the presented difference maps displayed clearer features than the maps calculated based on synchrotron Laue data. TR-SFX was later carried out in order to capture early photo-intermediates that could not be studied with the limited temporal resolution of synchrotron sources. The early stages of the *trans*-to-*cis* isomerization were unveiled in an ultra-fast molecular movie ranging from 142 fs to 3 ps (Pande et al., 2016). The TR-SFX study, combined with time-resolved absorption spectroscopy and QM/MM simulations, revealed that at delays shorter than 500 fs, the chromophore is in a twisted *trans* excited-state, a precursor of an early *cis* ground-state.

Being a model for studying protein dynamics, myoglobin was naturally chosen to be investigated by TR-SFX (Barends et al., 2015), representing the first TR-SFX study carried out with sub-picosecond temporal resolution. This molecular movie, ranging from 100 fs to 150 ps, revealed that photo-dissociation of CO started to occur within 100 – 200 fs and is completed after 200 fs. Motions of the heme group are detectable within 300 fs – 3 ps. The later time delays showed increasing displacements of entire α-helices as a function of time. These early events were predicted by QM/MM calculations (Barends et al., 2015) and are in line with the motions observed by time-resolved X-ray scattering performed on myoglobin in solution (Levantino et al., 2015a).

Because of the relevance for understanding molecular processes involved in human vision and because of potential biotechnology applications, bR, the light-driven proton pump, has been widely studied by kinetic crystallography (Wickstrand et al., 2015). Using synchrotron radiation, it was mainly studied by cryo-trapping approaches, allowing to obtain first insight into how isomerization of the retinal chromophore might enable the structural changes leading to proton translocation (Edman et al., 1999). TR-SFX provided the opportunity to study bR dynamics during its photocycle at room temperature. Structural changes that occurred from nano- to milliseconds (Nango et al., 2016) and on the sub-picosecond time scale (Nogly et al., 2018) were uncovered and allowed to propose an overall mechanism of light-induced proton transfer, from isomerization to proton translocation .

Enzymes are of particular biological interest and their catalytic mechanism can be investigated by time-resolved crystallography. Metalloenzymes carry a metal cluster involved in the chemical reaction they catalyze, making them very sensitive to radiation damage. By avoiding radiation damage through SFX experiments, cytochrome c oxidase has been studied and the room-temperature structures of the apo-protein (Andersson et al., 2017) and the CO-bound protein (Ishigami et al., 2017) were solved. Using TR-SFX, the cytochrome c oxydase mechanism was studied by determining intermediate-state structures 20 ns and 100 µs after photolysis (Shimada et al., 2017).

Substrate binding to the active site and turnover are critical steps in enzyme catalysis. Using microcrystals, the diffusion time of substrates through crystals is reduced and a time-resolution of milli- to microseconds can in principle be obtained in a so-called mix-and-inject approach (Schmidt, 2013). This approach has been successfully applied to study riboswitches on the second to minute time scale (Stagno et al., 2017) and for first time-resolved studies on β-lactamase on the second time scale (Kupitz et al., 2017).

Most of the TR-SFX studies undertaken until now involved model proteins that have already been extensively studied by time-resolved/kinetic crystallography at synchrotrons. The field of RSFPs, and by extension PTFPs, misses structural characterization of early intermediates that were predicted by simulation or detected by ultrafast spectroscopy. This PhD thesis presents the planning, the preparation and the conduction of TR-SFX experiments performed with the aim to fill the lack of structural data to understand the mechanism of photoswitching in rsEGFP2. Chapter I addresses the question of sample preparation for such a TR-SFX experiment and proposes a protocol for the reproducible production of microcrystal batches suitable for liquid injection and efficient illumination in a pump-probe TR-SFX experiment. Chapters II and III present the result of TR-SFX experiments on the picosecond and nanosecond time-scale, respectively. Supported by transient absorption spectroscopy data, a mechanism for photoswitching is proposed and discussed. Chapter IV reports on synchrotron structures and photophysical characteristics of the fluorescent and non-fluorescent states of two rsEGFP2 variants. These variants were designed based on TR-SFX results presented in chapter II to test if intermediate-state structures can be used to rationally generate improved variants for nanoscopy applications. Finally, chapter V presents the serendipitous discovery of structural ground-state heterogeneity in rsEGFP2 that is corroborated by the spectroscopic and structural analysis of the two rsEGFP2 variants presented in chapter IV.

References:

Adam, V., 2014. Phototransformable fluorescent proteins: which one for which application? *Histochem. Cell Biol.*, 142, 19–41. https://doi.org/10.1007/s00418-014-1190-5

Adam, V., Lelimousin, M., Boehme, S., Desfonds, G., Nienhaus, K., Field, M.J., Wiedenmann, J., McSweeney, S., Nienhaus, G.U., Bourgeois, D., 2008. Structural characterization of IrisFP, an optical highlighter undergoing multiple photo-induced transformations. *Proc. Natl. Acad. Sci. U. S. A.*, 105, 18343–18348. https://doi.org/10.1073/pnas.0805949105

Alonso-Mori, R., Kern, J., Gildea, R.J., Sokaras, D., Weng, T.-C., Lassalle-Kaiser, B., Tran, R., Hattne, J., Laksmono, H., Hellmich, J., Glöckner, C., Echols, N., Sierra, R.G., Schafer, D.W., Sellberg, J., Kenney, C., Herbst, R., Pines, J., Hart, P., Herrmann, S., Grosse-Kunstleve, R.W., Latimer, M.J., Fry, A.R., Messerschmidt, M.M., Miahnahri, A., Seibert, M.M., Zwart, P.H., White, W.E., Adams, P.D., Bogan, M.J., Boutet, S., Williams, G.J., Zouni, A., Messinger, J., Glatzel, P., Sauter, N.K., Yachandra, V.K., Yano, J., Bergmann, U., 2012. Energy-dispersive X-ray emission spectroscopy using an X-ray free-electron laser in a shot-by-shot mode. *Proc. Natl. Acad. Sci. U. S. A.*, 109, 19103–19107. https://doi.org/10.1073/pnas.1211384109

Andersson, R., Safari, C., Dods, R., Nango, E., Tanaka, R., Yamashita, A., Nakane, T., Tono, K., Joti, Y., Båth, P., Dunevall, E., Bosman, R., Nureki, O., Iwata, S., Neutze, R., Brändén, G., 2017. Serial femtosecond crystallography structure of cytochrome c oxidase at room temperature. *Sci. Rep.*, 7, 4518. https://doi.org/10.1038/s41598-017-04817-z

Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., Miyawaki, A., 2002. An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. Proc. Natl. Acad. Sci. U. S. A. 99, 12651–12656.

https://doi.org/10.1073/pnas.202320599

Ando, R., Mizuno, H., Miyawaki, A., 2004. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. *Science,* 306, 1370–1373. https://doi.org/10.1126/science.1102506

Andresen, M., Stiel, A.C., Trowitzsch, S., Weber, G., Eggeling, C., Wahl, M.C., Hell, S.W., Jakobs, S., 2007. Structural basis for reversible photoswitching in Dronpa. *Proc. Natl. Acad. Sci. U. S. A.*, 104, 13005–13009. https://doi.org/10.1073/pnas.0700629104

Andresen, M., Wahl, M.C., Stiel, A.C., Gräter, F., Schäfer, L.V., Trowitzsch, S., Weber, G., Eggeling, C., Grubmüller, H., Hell, S.W., Jakobs, S., 2005. Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.*, 102, 13070–13074. https://doi.org/10.1073/pnas.0502772102

Aquila, A., Hunter, M.S., Doak, R.B., Kirian, R.A., Fromme, P., White, T.A., Andreasson, J., Arnlund, D., Bajt, S., Barends, T.R.M., Barthelmess, M., Bogan, M.J., Bostedt, C., Bottin, H., Bozek, J.D., Caleman, C., Coppola, N., Davidsson, J., DePonte, D.P., Elser, V., Epp, S.W., Erk, B., Fleckenstein, H., Foucar, L., Frank, M., Fromme, R., Graafsma, H., Grotjohann, I., Gumprecht, L., Hajdu, J., Hampton, C.Y., Hartmann, A., Hartmann, R., Hau-Riege, S., Hauser, G., Hirsemann, H., Holl, P., Holton, J.M., Hömke, A., Johansson, L., Kimmel, N., Kassemeyer, S., Krasniqi, F., Kühnel, K.-U., Liang, M., Lomb, L., Malmerberg, E., Marchesini, S., Martin, A.V., Maia, F.R.N.C., Messerschmidt, M., Nass, K., Reich, C., Neutze, R., Rolles, D., Rudek, B., Rudenko, A., Schlichting, I., Schmidt, C., Schmidt, K.E., Schulz, J., Seibert, M.M., Shoeman, R.L., Sierra, R., Soltau, H., Starodub, D., Stellato, F., Stern, S., Strüder, L., Timneanu, N., Ullrich, J., Wang, X., Williams, G.J., Weidenspointner, G., Weierstall, U., Wunderer, C., Barty, A., Spence, J.C.H., Chapman, H.N., 2012. Time-resolved protein nanocrystallography using an X-ray free-electron laser. *Opt. Express*, 20, 2706–2716. https://doi.org/10.1364/OE.20.002706

Axford, D., Aller, P., Sanchez-Weatherby, J., Sandy, J., 2016. Applications of thin-film sandwich crystallization platforms. Acta Crystallogr. *Sect. F Struct. Biol. Commun.*, 72, 313–319. https://doi.org/10.1107/S2053230X16004386

Ayyer, K., Yefanov, O.M., Oberthür, D., Roy-Chowdhury, S., Galli, L., Mariani, V., Basu, S., Coe, J., Conrad, C.E., Fromme, R., Schaffer, A., Dörner, K., James, D., Kupitz, C., Metz, M., Nelson, G., Xavier, P.L., Beyerlein, K.R., Schmidt, M., Sarrou, I., Spence, J.C.H., Weierstall, U., White, T.A., Yang, J.-H., Zhao, Y., Liang, M., Aquila, A., Hunter, M.S., Robinson, J.S., Koglin, J.E., Boutet, S., Fromme, P., Barty, A., Chapman, H.N., 2016. Macromolecular diffractive imaging using imperfect crystals. *Nature,* 530, 202–206. https://doi.org/10.1038/nature16949

Barends, T.R.M., Foucar, L., Ardevol, A., Nass, K., Aquila, A., Botha, S., Doak, R.B., Falahati, K., Hartmann, E., Hilpert, M., Heinz, M., Hoffmann, M.C., Köfinger, J., Koglin, J.E., Kovacsova, G., Liang, M., Milathianaki, D., Lemke, H.T., Reinstein, J., Roome, C.M., Shoeman, R.L., Williams, G.J., Burghardt, I., Hummer, G., Boutet, S., Schlichting, I., 2015. Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science*, 350, 445–450. https://doi.org/10.1126/science.aac5492

Barends, T.R.M., Foucar, L., Botha, S., Doak, R.B., Shoeman, R.L., Nass, K., Koglin, J.E., Williams, G.J., Boutet, S., Messerschmidt, M., Schlichting, I., 2014. De novo protein crystal structure determination from X-ray free-electron laser data. *Nature*, 505, 244–

247. https://doi.org/10.1038/nature12773

Barends, T.R.M., Foucar, L., Shoeman, R.L., Bari, S., Epp, S.W., Hartmann, R., Hauser, G., Huth, M., Kieser, C., Lomb, L., Motomura, K., Nagaya, K., Schmidt, C., Strecker, R., Anielski, D., Boll, R., Erk, B., Fukuzawa, H., Hartmann, E., Hatsui, T., Holl, P., Inubushi, Y., Ishikawa, T., Kassemeyer, S., Kaiser, C., Koeck, F., Kunishima, N., Kurka, M., Rolles, D., Rudek, B., Rudenko, A., Sato, T., Schroeter, C.D., Soltau, H., Strueder, L., Tanaka, T., Togashi, T., Tono, K., Ullrich, J., Yase, S., Wada, S.I., Yao, M., Yabashi, M., Ueda, K., Schlichting, I., 2013. Anomalous signal from S atoms in protein crystallographic data from an X-ray free-electron laser. *Acta Crystallogr. D Biol. Crystallogr.*, 69, 838–842. https://doi.org/10.1107/S0907444913002448

Barty, A., Caleman, C., Aquila, A., Timneanu, N., Lomb, L., White, T.A., Andreasson, J., Arnlund, D., Bajt, S., Barends, T.R.M., Barthelmess, M., Bogan, M.J., Bostedt, C., Bozek, J.D., Coffee, R., Coppola, N., Davidsson, J., Deponte, D.P., Doak, R.B., Ekeberg, T., Elser, V., Epp, S.W., Erk, B., Fleckenstein, H., Foucar, L., Fromme, P., Graafsma, H., Gumprecht, L., Hajdu, J., Hampton, C.Y., Hartmann, R., Hartmann, A., Hauser, G., Hirsemann, H., Holl, P., Hunter, M.S., Johansson, L., Kassemeyer, S., Kimmel, N., Kirian, R.A., Liang, M., Maia, F.R.N.C., Malmerberg, E., Marchesini, S., Martin, A.V., Nass, K., Neutze, R., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Scott, H., Schlichting, I., Schulz, J., Seibert, M.M., Shoeman, R.L., Sierra, R.G., Soltau, H., Spence, J.C.H., Stellato, F., Stern, S., Strüder, L., Ullrich, J., Wang, X., Weidenspointner, G., Weierstall, U., Wunderer, C.B., Chapman, H.N., 2012. Self-terminating diffraction gates femtosecond X-ray nanocrystallography measurements. *Nat. Photonics*, 6, 35–40. https://doi.org/10.1038/nphoton.2011.297

Barty, A., Kirian, R.A., Maia, F.R.N.C., Hantke, M., Yoon, C.H., White, T.A., Chapman, H., 2014. Cheetah: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J. Appl. Crystallogr.*, 47, 1118–1131. https://doi.org/10.1107/S1600576714007626

Batyuk, A., Galli, L., Ishchenko, A., Han, G.W., Gati, C., Popov, P.A., Lee, M.-Y., Stauch, B., White, T.A., Barty, A., Aquila, A., Hunter, M.S., Liang, M., Boutet, S., Pu, M., Liu, Z.-J., Nelson, G., James, D., Li, C., Zhao, Y., Spence, J.C.H., Liu, W., Fromme, P., Katritch, V., Weierstall, U., Stevens, R.C., Cherezov, V., 2016. Native phasing of x-ray free-electron laser data for a G protein-coupled receptor. *Sci. Adv.*, 2, e1600292. https://doi.org/10.1126/sciadv.1600292

Botha, S., Nass, K., Barends, T.R.M., Kabsch, W., Latz, B., Dworkowski, F., Foucar, L., Panepucci, E., Wang, M., Shoeman, R.L., Schlichting, I., Doak, R.B., 2015. Room-temperature serial crystallography at synchrotron X-ray sources using slowly flowing free-standing high-viscosity microstreams. *Acta Crystallogr. D Biol. Crystallogr.*, 71, 387–397. https://doi.org/10.1107/S1399004714026327

Bourgeois, D., Royant, A., 2005. Advances in kinetic protein crystallography. *Curr. Opin. Struct. Biol.*, 15, 538–547. https://doi.org/10.1016/j.sbi.2005.08.002

Bourgeois, D., Vallone, B., Schotte, F., Arcovito, A., Miele, A.E., Sciara, G., Wulff, M., Anfinrud, P., Brunori, M., 2003. Complex landscape of protein structural dynamics unveiled by nanosecond Laue crystallography. *Proc. Natl. Acad. Sci. U. S. A.*, 100, 8704–8709. https://doi.org/10.1073/pnas.1430900100

Boutet, S., Lomb, L., Williams, G.J., Barends, T.R.M., Aquila, A., Doak, R.B., Weierstall, U., DePonte, D.P., Steinbrener, J., Shoeman, R.L., Messerschmidt, M., Barty, A., White, T.A., Kassemeyer, S., Kirian, R.A., Seibert, M.M., Montanez, P.A., Kenney, C., Herbst, R., Hart, P., Pines, J., Haller, G., Gruner, S.M., Philipp, H.T., Tate, M.W., Hromalik, M., Koerner, L.J., van Bakel, N., Morse, J., Ghonsalves, W., Arnlund, D., Bogan, M.J., Caleman, C., Fromme, R., Hampton, C.Y., Hunter, M.S., Johansson, L.C., Katona, G., Kupitz, C., Liang, M., Martin, A.V., Nass, K., Redecke, L., Stellato, F., Timneanu, N., Wang, D., Zatsepin, N.A., Schafer, D., Defever, J., Neutze, R., Fromme, P., Spence, J.C.H., Chapman, H.N., Schlichting, I., 2012. High-resolution protein structure determination by serial femtosecond crystallography. *Science*, 337, 362–364. https://doi.org/10.1126/science.1217737

Caffrey, M., Cherezov, V., 2009. Crystallizing membrane proteins using lipidic mesophases. *Nat. Protoc.*, 4, 706–731. https://doi.org/10.1038/nprot.2009.31

Chapman, H.N., Fromme, P., Barty, A., White, T.A., Kirian, R.A., Aquila, A., Hunter, M.S., Schulz, J., DePonte, D.P., Weierstall, U., Doak, R.B., Maia, F.R.N.C., Martin, A.V., Schlichting, I., Lomb, L., Coppola, N., Shoeman, R.L., Epp, S.W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmess, M., Caleman, C., Boutet, S., Bogan, M.J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K.-U., Messerschmidt, M., Bozek, J.D., Hau-Riege, S.P., Frank, M., Hampton, C.Y., Sierra, R.G., Starodub, D., Williams, G.J., Hajdu, J., Timneanu, N., Seibert, M.M., Andreasson, J., Rocker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schröter, C.-D., Krasniqi, F., Bott, M., Schmidt, K.E., Wang, X., Grotjohann, I., Holton, J.M., Barends, T.R.M., Neutze, R., Marchesini, S., Fromme, R., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B., Spence, J.C.H., 2011. Femtosecond X-ray protein nanocrystallography. *Nature*, 470, 73–77. https://doi.org/10.1038/nature09750

Colletier, J.-P., Sawaya, M.R., Gingery, M., Rodriguez, J.A., Cascio, D., Brewster, A.S., Michels-Clark, T., Hice, R.H., Coquelle, N., Boutet, S., Williams, G.J., Messerschmidt, M., DePonte, D.P., Sierra, R.G., Laksmono, H., Koglin, J.E., Hunter, M.S., Park, H.-W., Uervirojnangkoorn, M., Bideshi, D.K., Brunger, A.T., Federici, B.A., Sauter, N.K., Eisenberg, D.S., 2016a. De novo phasing with X-ray laser reveals mosquito larvicide BinAB structure. *Nature*, 539, 43–47. https://doi.org/10.1038/nature19825

Colletier, J.-P., Sliwa, M., Gallat, F.-X., Sugahara, M., Guillon, V., Schirò, G., Coquelle, N., Woodhouse, J., Roux, L., Gotthard, G., Royant, A., Uriarte, L.M., Ruckebusch, C., Joti, Y., Byrdin, M., Mizohata, E., Nango, E., Tanaka, T., Tono, K., Yabashi, M., Adam, V., Cammarata, M., Schlichting, I., Bourgeois, D., Weik, M., 2016b. Serial Femtosecond Crystallography and Ultrafast Absorption Spectroscopy of the Photoswitchable Fluorescent Protein IrisFP. *J. Phys. Chem. Lett.*, 7, 882–887. https://doi.org/10.1021/acs.jpclett.5b02789

Conrad, C.E., Basu, S., James, D., Wang, D., Schaffer, A., Roy-Chowdhury, S., Zatsepin, N.A., Aquila, A., Coe, J., Gati, C., Hunter, M.S., Koglin, J.E., Kupitz, C., Nelson, G., Subramanian, G., White, T.A., Zhao, Y., Zook, J., Boutet, S., Cherezov, V., Spence, J.C.H., Fromme, R., Weierstall, U., Fromme, P., 2015. A novel inert crystal delivery medium for serial femtosecond crystallography. *IUCrJ*, 2, 421–430. https://doi.org/10.1107/S2052252515009811

Coquelle, N., Brewster, A.S., Kapp, U., Shilova, A., Weinhausen, B., Burghammer, M., Colletier, J.P., 2015. Raster-scanning serial protein crystallography using micro- and nano-focused synchrotron beams. *Acta Crystallogr. D Biol. Crystallogr.*, 71, 1184–1196. https://doi.org/10.1107/S1399004715004514

DePonte, D.P., Weierstall, U., Schmidt, K., Warner, J., Starodub, D., Spence, J.C.H., Doak, R.B., 2008. Gas dynamic virtual nozzle for generation of microscopic droplet streams. *J. Phys. Appl. Phys.*, 41, 195505. https://doi.org/10.1088/0022-3727/41/19/195505

Duszenko, M., Redecke, L., Mudogo, C.N., Sommer, B.P., Mogk, S., Oberthuer, D., Betzel, C., 2015. In vivo protein crystallization in combination with highly brilliant radiation sources offers novel opportunities for the structural analysis of post-translationally modified eukaryotic proteins. Acta Crystallogr. *Sect. F Struct. Biol. Commun.*, 71, 929–937. https://doi.org/10.1107/S2053230X15011450

Edman, K., Nollert, P., Royant, A., Belrhali, H., Pebay-Peyroula, E., Hajdu, J., Neutze, R., Landau, E.M., 1999. High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle. *Nature*, 401, 822–826. https://doi.org/10.1038/44623

El Khatib, M., Martins, A., Bourgeois, D., Colletier, J.-P., Adam, V., 2016. Rational design of ultrastable and reversibly photoswitchable fluorescent proteins for super-resolution imaging of the bacterial periplasm. *Sci. Rep.*, 6, 18459. https://doi.org/10.1038/srep18459

Emma, P., Akre, R., Arthur, J., Bionta, R., Bostedt, C., Bozek, J., Brachmann, A., Bucksbaum, P., Coffee, R., Decker, F.-J., Ding, Y., Dowell, D., Edstrom, S., Fisher, A., Frisch, J., Gilevich, S., Hastings, J., Hays, G., Hering, P., Huang, Z., Iverson, R., Loos, H., Messerschmidt, M., Miahnahri, A., Moeller, S., Nuhn, H.-D., Pile, G., Ratner, D., Rzepiela, J., Schultz, D., Smith, T., Stefan, P., Tompkins, H., Turner, J., Welch, J., White, W., Wu, J., Yocky, G., Galayda, J., 2010. First lasing and operation of an ångstrom-wavelength free-electron laser. *Nat. Photonics*, 4, 641–647. https://doi.org/10.1038/nphoton.2010.176

Faro, A.R., Carpentier, P., Jonasson, G., Pompidor, G., Arcizet, D., Demachy, I., Bourgeois, D., 2011. Low-temperature chromophore isomerization reveals the photoswitching mechanism of the fluorescent protein Padron. *J. Am. Chem. Soc.*, 133, 16362–16365. https://doi.org/10.1021/ja207001y

Foucar, L., 2016. CFEL-ASG Software Suite (CASS): usage for free-electron laser experiments with biological focus. *J. Appl. Crystallogr.*, 49, 1336–1346. https://doi.org/10.1107/S1600576716009201

Fron, E., Flors, C., Schweitzer, G., Habuchi, S., Mizuno, H., Ando, R., Schryver, F.C.D., Miyawaki, A., Hofkens, J., 2007. Ultrafast excited-state dynamics of the photoswitchable protein Dronpa. *J. Am. Chem. Soc.*, 129, 4870–4871. https://doi.org/10.1021/ja069365v

Fukuda, Y., Tse, K.M., Nakane, T., Nakatsu, T., Suzuki, M., Sugahara, M., Inoue, S., Masuda, T., Yumoto, F., Matsugaki, N., Nango, E., Tono, K., Joti, Y., Kameshima, T., Song, C., Hatsui, T., Yabashi, M., Nureki, O., Murphy, M.E.P., Inoue, T., Iwata, S., Mizohata, E., 2016. Redox-coupled proton transfer mechanism in nitrite reductase revealed by femtosecond crystallography. *Proc. Natl. Acad. Sci. U. S. A.*, 113, 2928–2933. https://doi.org/10.1073/pnas.1517770113

Fuller, F.D., Gul, S., Chatterjee, R., Burgie, E.S., Young, I.D., Lebrette, H., Srinivas, V., Brewster, A.S., Michels-Clark, T., Clinger, J.A., Andi, B., Ibrahim, M., Pastor, E., de Lichtenberg, C., Hussein, R., Pollock, C.J., Zhang, M., Stan, C.A., Kroll, T., Fransson, T., Weninger, C., Kubin, M., Aller, P., Lassalle, L., Bräuer, P., Miller, M.D., Amin, M., Koroidov, S., Roessler, C.G., Allaire, M., Sierra, R.G., Docker, P.T., Glownia, J.M., Nelson, S., Koglin, J.E., Zhu, D., Chollet, M., Song, S., Lemke, H., Liang, M., Sokaras, D., Alonso-Mori, R., Zouni, A., Messinger, J., Bergmann, U., Boal, A.K., Bollinger, J.M., Krebs, C., Högbom, M., Phillips, G.N., Vierstra, R.D., Sauter, N.K., Orville, A.M., Kern, J., Yachandra, V.K., Yano, J., 2017. Drop-on-demand sample delivery for studying biocatalysts in action at X-ray free-electron lasers. *Nat. Methods*, 14, 443–449. https://doi.org/10.1038/nmeth.4195

Garman, E.F., Weik, M., 2017. Radiation Damage in Macromolecular Crystallography. *Methods Mol. Biol. Clifton NJ*, 1607, 467–489. https://doi.org/10.1007/978-1-4939-7000-1_20

Gati, C., Bourenkov, G., Klinge, M., Rehders, D., Stellato, F., Oberthür, D., Yefanov, O., Sommer, B.P., Mogk, S., Duszenko, M., Betzel, C., Schneider, T.R., Chapman, H.N., Redecke, L., 2014. Serial crystallography on in vivo grown microcrystals using synchrotron radiation. *IUCrJ*, 1, 87–94. https://doi.org/10.1107/S2052252513033939

Genick, U.K., Borgstahl, G.E., Ng, K., Ren, Z., Pradervand, C., Burke, P.M., Srajer, V., Teng, T.Y., Schildkamp, W., McRee, D.E., Moffat, K., Getzoff, E.D., 1997. Structure of a protein photocycle intermediate by millisecond time-resolved crystallography. *Science*, 275, 1471–1475.

Ginn, H.M., Brewster, A.S., Hattne, J., Evans, G., Wagner, A., Grimes, J.M., Sauter, N.K., Sutton, G., Stuart, D.I., 2015. A revised partiality model and post-refinement algorithm for X-ray free-electron laser data. *Acta Crystallogr. D Biol. Crystallogr.*, 71, 1400–1410. https://doi.org/10.1107/S1399004715006902

Gorel, A., Motomura, K., Fukuzawa, H., Doak, R.B., Grünbein, M.L., Hilpert, M., Inoue, I., Kloos, M., Kovácsová, G., Nango, E., Nass, K., Roome, C.M., Shoeman, R.L., Tanaka, R., Tono, K., Joti, Y., Yabashi, M., Iwata, S., Foucar, L., Ueda, K., Barends, T.R.M., Schlichting, I., 2017. Multi-wavelength anomalous diffraction de novo phasing using a two-colour X-ray free-electron laser with wide tunability. *Nat. Commun.*, 8, 1170. https://doi.org/10.1038/s41467-017-00754-7

Grotjohann, T., Testa, I., Leutenegger, M., Bock, H., Urban, N.T., Lavoie-Cardinal, F., Willig, K.I., Eggeling, C., Jakobs, S., Hell, S.W., 2011. Diffraction-unlimited all-optical imaging and writing with a photochromic GFP. *Nature*, 478, 204–208. https://doi.org/10.1038/nature10497

Grotjohann, T., Testa, I., Reuss, M., Brakemann, T., Eggeling, C., Hell, S.W., Jakobs, S., 2012. rsEGFP2 enables fast RESOLFT nanoscopy of living cells. *eLife*, 1, e00248. https://doi.org/10.7554/eLife.00248

Halsted, T.P., Yamashita, K., Hirata, K., Ago, H., Ueno, G., Tosha, T., Eady, R.R., Antonyuk, S.V., Yamamoto, M., Hasnain, S.S., 2018. An unprecedented dioxygen species revealed by serial femtosecond rotation crystallography in copper nitrite reductase. *IUCrJ*, 5, 22–31. https://doi.org/10.1107/S2052252517016128

Heymann, M., Opthalage, A., Wierman, J.L., Akella, S., Szebenyi, D.M.E., Gruner, S.M., Fraden, S., 2014. Room-temperature serial crystallography using a kinetically optimized microfluidic device for protein crystallization and on-chip X-ray diffraction. *IUCrJ*, 1, 349–360. https://doi.org/10.1107/S2052252514016960

Hirata, K., Shinzawa-Itoh, K., Yano, N., Takemura, S., Kato, K., Hatanaka, M., Muramoto, K., Kawahara, T., Tsukihara, T., Yamashita, E., Tono, K., Ueno, G., Hikima, T., Murakami, H., Inubushi, Y., Yabashi, M., Ishikawa, T., Yamamoto, M., Ogura, T., Sugimoto, H., Shen, J.-R., Yoshikawa, S., Ago, H., 2014. Determination of damage-free crystal structure of an X-ray-sensitive protein using an XFEL. *Nat. Methods*, 11, 734–736. https://doi.org/10.1038/nmeth.2962

Hunter, M.S., Segelke, B., Messerschmidt, M., Williams, G.J., Zatsepin, N.A., Barty, A., Benner, W.H., Carlson, D.B., Coleman, M., Graf, A., Hau-Riege, S.P., Pardini, T., Seibert, M.M., Evans, J., Boutet, S., Frank, M., 2014. Fixed-target protein serial microcrystallography with an x-ray free electron laser. *Sci. Rep.,* 4, 6026. https://doi.org/10.1038/srep06026

Ibrahim, M., Chatterjee, R., Hellmich, J., Tran, R., Bommer, M., Yachandra, V.K., Yano, J., Kern, J., Zouni, A., 2015. Improvements in serial femtosecond crystallography of photosystem II by optimizing crystal uniformity using microseeding procedures. *Struct. Dyn. Melville N*, 2. https://doi.org/10.1063/1.4919741

Ishigami, I., Zatsepin, N.A., Hikita, M., Conrad, C.E., Nelson, G., Coe, J.D., Basu, S., Grant, T.D., Seaberg, M.H., Sierra, R.G., Hunter, M.S., Fromme, P., Fromme, R., Yeh, S.-R., Rousseau, D.L., 2017. Crystal structure of CO-bound cytochrome c oxidase determined by serial femtosecond X-ray crystallography at room temperature. *Proc. Natl. Acad. Sci. U. S. A.,* 114, 8011–8016. https://doi.org/10.1073/pnas.1705628114

Ishikawa, T., Aoyagi, H., Asaka, T., Asano, Y., Azumi, N., Bizen, T., Ego, H., Fukami, K., Fukui, T., Furukawa, Y., Goto, S., Hanaki, H., Hara, T., Hasegawa, T., Hatsui, T., Higashiya, A., Hirono, T., Hosoda, N., Ishii, M., Inagaki, T., Inubushi, Y., Itoga, T., Joti, Y., Kago, M., Kameshima, T., Kimura, H., Kirihara, Y., Kiyomichi, A., Kobayashi, T., Kondo, C., Kudo, T., Maesaka, H., Maréchal, X.M., Masuda, T., Matsubara, S., Matsumoto, T., Matsushita, T., Matsui, S., Nagasono, M., Nariyama, N., Ohashi, H., Ohata, T., Ohshima, T., Ono, S., Otake, Y., Saji, C., Sakurai, T., Sato, T., Sawada, K., Seike, T., Shirasawa, K., Sugimoto, T., Suzuki, S., Takahashi, S., Takebe, H., Takeshita, K., Tamasaku, K., Tanaka, H., Tanaka, R., Tanaka, T., Togashi, T., Togawa, K., Tokuhisa, A., Tomizawa, H., Tono, K., Wu, S., Yabashi, M., Yamaga, M., Yamashita, A., Yanagida, K., Zhang, C., Shintake, T., Kitamura, H., Kumagai, N., 2012. A compact X-ray free-electron laser emitting in the sub-ångström region. *Nat. Photonics*, 6, 540–544. https://doi.org/10.1038/nphoton.2012.141

Johansson, L.C., Arnlund, D., Katona, G., White, T.A., Barty, A., DePonte, D.P., Shoeman, R.L., Wickstrand, C., Sharma, A., Williams, G.J., Aquila, A., Bogan, M.J., Caleman, C., Davidsson, J., Doak, R.B., Frank, M., Fromme, R., Galli, L., Grotjohann, I., Hunter, M.S., Kassemeyer, S., Kirian, R.A., Kupitz, C., Liang, M., Lomb, L., Malmerberg, E., Martin, A.V., Messerschmidt, M., Nass, K., Redecke, L., Seibert, M.M., Sjöhamn, J., Steinbrener, J., Stellato, F., Wang, D., Wahlgren, W.Y., Weierstall, U., Westenhoff, S., Zatsepin, N.A., Boutet, S., Spence, J.C.H., Schlichting, I., Chapman, H.N., Fromme, P., Neutze, R., 2013. Structure of a photosynthetic reaction centre determined by serial femtosecond crystallography. *Nat. Commun.*, 4, 2911. https://doi.org/10.1038/ncomms3911

Johansson, L.C., Arnlund, D., White, T.A., Katona, G., Deponte, D.P., Weierstall, U., Doak, R.B., Shoeman, R.L., Lomb, L., Malmerberg, E., Davidsson, J., Nass, K., Liang, M., Andreasson, J., Aquila, A., Bajt, S., Barthelmess, M., Barty, A., Bogan, M.J., Bostedt, C., Bozek, J.D., Caleman, C., Coffee, R., Coppola, N., Ekeberg, T., Epp, S.W., Erk, B., Fleckenstein, H., Foucar, L., Graafsma, H., Gumprecht, L., Hajdu, J., Hampton, C.Y., Hartmann, R., Hartmann, A., Hauser, G., Hirsemann, H., Holl, P., Hunter, M.S., Kassemeyer, S., Kimmel, N., Kirian, R.A., Maia, F.R.N.C., Marchesini, S., Martin, A.V., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Schlichting, I., Schulz, J., Seibert, M.M., Sierra, R.G., Soltau, H., Starodub, D., Stellato, F., Stern, S., Strüder, L., Timneanu, N., Ullrich, J., Wahlgren, W.Y., Wang, X., Weidenspointner, G., Wunderer, C., Fromme, P., Chapman, H.N., Spence, J.C.H., Neutze, R., 2012. Lipidic phase membrane protein serial femtosecond crystallography. *Nat. Methods,* 9, 263–265. https://doi.org/10.1038/nmeth.1867

Jung, Y.O., Lee, J.H., Kim, J., Schmidt, M., Moffat, K., Srajer, V., Ihee, H., 2013. Volume-conserving trans-cis isomerization pathways in photoactive yellow protein visualized by picosecond X-ray crystallography. *Nat. Chem.,* 5, 212–220. https://doi.org/10.1038/nchem.1565

Kabsch, W., 2014. Processing of X-ray snapshots from crystals in random orientations. *Acta Crystallogr. D Biol. Crystallogr.,* 70, 2204–2216. https://doi.org/10.1107/S1399004714013534

Kang, Y., Zhou, X.E., Gao, X., He, Y., Liu, W., Ishchenko, A., Barty, A., White, T.A., Yefanov, O., Han, G.W., Xu, Q., de Waal, P.W., Ke, J., Tan, M.H.E., Zhang, C., Moeller, A., West, G.M., Pascal, B.D., Van Eps, N., Caro, L.N., Vishnivetskiy, S.A., Lee, R.J., Suino-Powell, K.M., Gu, X., Pal, K., Ma, J., Zhi, X., Boutet, S., Williams, G.J., Messerschmidt, M., Gati, C., Zatsepin, N.A., Wang, D., James, D., Basu, S., Roy-Chowdhury, S., Conrad, C.E., Coe, J., Liu, H., Lisova, S., Kupitz, C., Grotjohann, I., Fromme, R., Jiang, Y., Tan, M., Yang, H., Li, J., Wang, M., Zheng, Z., Li, D., Howe, N., Zhao, Y., Standfuss, J., Diederichs, K., Dong, Y., Potter, C.S., Carragher, B., Caffrey, M., Jiang, H., Chapman, H.N., Spence, J.C.H., Fromme, P., Weierstall, U., Ernst, O.P., Katritch, V., Gurevich, V.V., Griffin, P.R., Hubbell, W.L., Stevens, R.C., Cherezov, V., Melcher, K., Xu, H.E., 2015. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature,* 523, 561–567. https://doi.org/10.1038/nature14656

Kern, J., Alonso-Mori, R., Hellmich, J., Tran, R., Hattne, J., Laksmono, H., Glöckner, C., Echols, N., Sierra, R.G., Sellberg, J., Lassalle-Kaiser, B., Gildea, R.J., Glatzel, P., Grosse-Kunstleve, R.W., Latimer, M.J., McQueen, T.A., DiFiore, D., Fry, A.R., Messerschmidt, M., Miahnahri, A., Schafer, D.W., Seibert, M.M., Sokaras, D., Weng, T.-C., Zwart, P.H., White, W.E., Adams, P.D., Bogan, M.J., Boutet, S., Williams, G.J., Messinger, J., Sauter, N.K., Zouni, A., Bergmann, U., Yano, J., Yachandra, V.K., 2012. Room temperature femtosecond X-ray diffraction of photosystem II microcrystals. *Proc. Natl. Acad. Sci. U. S. A.*, 109, 9721–9726. https://doi.org/10.1073/pnas.1204598109

Kern, J., Alonso-Mori, R., Tran, R., Hattne, J., Gildea, R.J., Echols, N., Glöckner, C., Hellmich, J., Laksmono, H., Sierra, R.G., Lassalle-Kaiser, B., Koroidov, S., Lampe, A., Han, G., Gul, S., Difiore, D., Milathianaki, D., Fry, A.R., Miahnahri, A., Schafer, D.W., Messerschmidt, M., Seibert, M.M., Koglin, J.E., Sokaras, D., Weng, T.-C., Sellberg, J., Latimer, M.J., Grosse-Kunstleve, R.W., Zwart, P.H., White, W.E., Glatzel, P., Adams, P.D., Bogan, M.J., Williams, G.J., Boutet, S., Messinger, J., Zouni, A., Sauter, N.K., Yachandra, V.K., Bergmann, U., Yano, J., 2013. Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature. *Science,* 340, 491–495. https://doi.org/10.1126/science.1234273

Kern, J., Tran, R., Alonso-Mori, R., Koroidov, S., Echols, N., Hattne, J., Ibrahim, M., Gul, S., Laksmono, H., Sierra, R.G., Gildea, R.J., Han, G., Hellmich, J., Lassalle-Kaiser, B., Chatterjee, R., Brewster, A.S., Stan, C.A., Glöckner, C., Lampe, A., DiFiore, D., Milathianaki, D., Fry, A.R., Seibert, M.M., Koglin, J.E., Gallo, E., Uhlig, J., Sokaras, D., Weng, T.-C., Zwart, P.H., Skinner, D.E., Bogan, M.J., Messerschmidt, M., Glatzel, P., Williams, G.J., Boutet, S., Adams, P.D., Zouni, A., Messinger, J., Sauter, N.K., Bergmann, U., Yano, J., Yachandra, V.K., 2014. Taking snapshots of photosynthetic water oxidation using femtosecond X-ray diffraction and spectroscopy. *Nat. Commun.*, 5, 4371. https://doi.org/10.1038/ncomms5371

Kirian, R.A., Wang, X., Weierstall, U., Schmidt, K.E., Spence, J.C.H., Hunter, M., Fromme, P., White, T., Chapman, H.N., Holton, J., 2010. Femtosecond protein nanocrystallography—data analysis methods. *Opt. Express*, 18, 5713–5723.

Kirian, R.A., White, T.A., Holton, J.M., Chapman, H.N., Fromme, P., Barty, A., Lomb, L., Aquila, A., Maia, F.R.N.C., Martin, A.V., Fromme, R., Wang, X., Hunter, M.S., Schmidt, K.E., Spence, J.C.H., 2011. Structure-factor analysis of femtosecond microdiffraction patterns from protein nanocrystals. *Acta Crystallogr. A*, 67, 131–140. https://doi.org/10.1107/S0108767310050981

Koopmann, R., Cupelli, K., Redecke, L., Nass, K., Deponte, D.P., White, T.A., Stellato, F., Rehders, D., Liang, M., Andreasson, J., Aquila, A., Bajt, S., Barthelmess, M., Barty, A., Bogan, M.J., Bostedt, C., Boutet, S., Bozek, J.D., Caleman, C., Coppola, N., Davidsson, J., Doak, R.B., Ekeberg, T., Epp, S.W., Erk, B., Fleckenstein, H., Foucar, L., Graafsma, H., Gumprecht, L., Hajdu, J., Hampton, C.Y., Hartmann, A., Hartmann, R., Hauser, G., Hirsemann, H., Holl, P., Hunter, M.S., Kassemeyer, S., Kirian, R.A., Lomb, L., Maia, F.R.N.C., Kimmel, N., Martin, A.V., Messerschmidt, M., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Schlichting, I., Schulz, J., Seibert, M.M., Shoeman, R.L., Sierra, R.G., Soltau, H., Stern, S., Strüder, L., Timneanu, N., Ullrich, J., Wang, X., Weidenspointner, G., Weierstall, U., Williams, G.J., Wunderer, C.B., Fromme, P., Spence, J.C.H., Stehle, T., Chapman, H.N., Betzel, C., Duszenko, M., 2012. In vivo protein crystallization opens new routes in structural biology. *Nat. Methods,* 9, 259–262. https://doi.org/10.1038/nmeth.1859

Kovácsová, G., Grünbein, M.L., Kloos, M., Barends, T.R.M., Schlesinger, R., Heberle, J., Kabsch, W., Shoeman, R.L., Doak, R.B., Schlichting, I., 2017. Viscous hydrophilic injection matrices for serial crystallography. *IUCrJ,* 4, 400–410. https://doi.org/10.1107/S2052252517005140

Kupitz, C., Basu, S., Grotjohann, I., Fromme, R., Zatsepin, N.A., Rendek, K.N., Hunter, M.S., Shoeman, R.L., White, T.A., Wang, D., James, D., Yang, J.-H., Cobb, D.E., Reeder, B., Sierra, R.G., Liu, H., Barty, A., Aquila, A.L., Deponte, D., Kirian, R.A., Bari, S., Bergkamp, J.J., Beyerlein, K.R., Bogan, M.J., Caleman, C., Chao, T.-C., Conrad, C.E., Davis, K.M., Fleckenstein, H., Galli, L., Hau-Riege, S.P., Kassemeyer, S., Laksmono, H., Liang, M., Lomb, L., Marchesini, S., Martin, A.V., Messerschmidt, M., Milathianaki, D., Nass, K., Ros, A., Roy-Chowdhury, S., Schmidt, K., Seibert, M., Steinbrener, J., Stellato, F., Yan, L., Yoon, C., Moore, T.A., Moore, A.L., Pushkar, Y., Williams, G.J., Boutet, S., Doak, R.B., Weierstall, U., Frank, M., Chapman, H.N., Spence, J.C.H., Fromme, P., 2014a. Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature,* 513, 261–265. https://doi.org/10.1038/nature13453

Kupitz, C., Grotjohann, I., Conrad, C.E., Roy-Chowdhury, S., Fromme, R., Fromme, P., 2014b. Microcrystallization techniques for serial femtosecond crystallography using photosystem II from Thermosynechococcus elongatus as a model system. Philos. *Trans. R. Soc. Lond. B. Biol. Sci.,* 369, 20130316. https://doi.org/10.1098/rstb.2013.0316

Kupitz, C., Olmos, J.L., Holl, M., Tremblay, L., Pande, K., Pandey, S., Oberthür, D., Hunter, M., Liang, M., Aquila, A., Tenboer, J., Calvey, G., Katz, A., Chen, Y., Wiedorn, M.O., Knoska, J., Meents, A., Majriani, V., Norwood, T., Poudyal, I., Grant, T., Miller, M.D., Xu, W., Tolstikova, A., Morgan, A., Metz, M., Martin-Garcia, J.M., Zook, J.D., Roy-Chowdhury, S., Coe, J., Nagaratnam, N., Meza, D., Fromme, R., Basu, S., Frank, M., White, T., Barty, A., Bajt, S., Yefanov, O., Chapman, H.N., Zatsepin, N., Nelson, G., Weierstall, U., Spence, J., Schwander, P., Pollack, L., Fromme, P., Ourmazd, A., Phillips, G.N., Schmidt, M., 2017. Structural enzymology using X-ray free electron lasers. *Struct. Dyn.*, Melville N 4, 044003. https://doi.org/10.1063/1.4972069

Laptenok, S.P., Gil, A.A., Hall, C.R., Lukacs, A., Iuliano, J.N., Jones, G.A., Greetham, G.M., Donaldson, P., Miyawaki, A., Tonge, P.J., Meech, S.R., 2018. Infrared spectroscopy reveals multi-step multi-timescale photoactivation in the photoconvertible protein archetype dronpa. *Nat. Chem.,* 10, 845–852. https://doi.org/10.1038/s41557-018-0073-0

Lee, D.B., Kim, J.-M., Seok, J.H., Lee, J.-H., Jo, J.D., Mun, J.Y., Conrad, C., Coe, J., Nelson, G., Hogue, B., White, T.A., Zatsepin, N., Weierstall, U., Barty, A., Chapman, H., Fromme, P., Spence, J., Chung, M.S., Oh, C.-H., Kim, K.H., 2018. Supersaturation-controlled microcrystallization and visualization analysis for serial femtosecond crystallography. *Sci. Rep.*, 8, 2541. https://doi.org/10.1038/s41598-018-20899-9

Levantino, M., Schirò, G., Lemke, H.T., Cottone, G., Glownia, J.M., Zhu, D., Chollet, M., Ihee, H., Cupane, A., Cammarata, M., 2015a. Ultrafast myoglobin structural dynamics observed with an X-ray free-electron laser. *Nat. Commun.*, 6, 6772. https://doi.org/10.1038/ncomms7772

Levantino, M., Yorke, B.A., Monteiro, D.C., Cammarata, M., Pearson, A.R., 2015b. Using synchrotrons and XFELs for time-resolved X-ray crystallography and solution scattering experiments on biomolecules. *Curr. Opin. Struct. Biol.*, 35, 41–48. https://doi.org/10.1016/j.sbi.2015.07.017

Liu, W., Ishchenko, A., Cherezov, V., 2014. Preparation of microcrystals in lipidic cubic phase for serial femtosecond crystallography. *Nat. Protoc.*, 9, 2123–2134. https://doi.org/10.1038/nprot.2014.141

Liu, W., Wacker, D., Gati, C., Han, G.W., James, D., Wang, D., Nelson, G., Weierstall, U., Katritch, V., Barty, A., Zatsepin, N.A., Li, D., Messerschmidt, M., Boutet, S., Williams, G.J., Koglin, J.E., Seibert, M.M., Wang, C., Shah, S.T.A., Basu, S., Fromme, R., Kupitz, C., Rendek, K.N., Grotjohann, I., Fromme, P., Kirian, R.A., Beyerlein, K.R., White, T.A., Chapman, H.N., Caffrey, M., Spence, J.C.H., Stevens, R.C., Cherezov, V., 2013. Serial femtosecond crystallography of G protein-coupled receptors. *Science*, 342, 1521–1524. https://doi.org/10.1126/science.1244142

Lomb, L., Barends, T.R.M., Kassemeyer, S., Aquila, A., Epp, S.W., Erk, B., Foucar, L., Hartmann, R., Rudek, B., Rolles, D., Rudenko, A., Shoeman, R.L., Andreasson, J., Bajt, S., Barthelmess, M., Barty, A., Bogan, M.J., Bostedt, C., Bozek, J.D., Caleman, C., Coffee, R., Coppola, N., Deponte, D.P., Doak, R.B., Ekeberg, T., Fleckenstein, H., Fromme, P., Gebhardt, M., Graafsma, H., Gumprecht, L., Hampton, C.Y., Hartmann, A., Hauser, G., Hirsemann, H., Holl, P., Holton, J.M., Hunter, M.S., Kabsch, W., Kimmel, N., Kirian, R.A., Liang, M., Maia, F.R.N.C., Meinhart, A., Marchesini, S., Martin, A.V., Nass, K., Reich, C., Schulz, J., Seibert, M.M., Sierra, R., Soltau, H., Spence, J.C.H., Steinbrener, J., Stellato, F., Stern, S., Timneanu, N., Wang, X., Weidenspointner, G., Weierstall, U., White, T.A., Wunderer, C., Chapman, H.N., Ullrich, J., Strüder, L., Schlichting, I., 2011. Radiation damage in protein serial femtosecond crystallography using an x-ray free-electron laser. *Phys. Rev. B Condens. Matter Mater. Phys.*, 84, 214111. https://doi.org/10.1103/PhysRevB.84.214111

Lukacs, A., Haigney, A., Brust, R., Addison, K., Towrie, M., Greetham, G.M., Jones, G.A., Miyawaki, A., Tonge, P.J., Meech, S.R., 2013. Protein photochromism observed by ultrafast vibrational spectroscopy. *J. Phys. Chem. B*, 117, 11954–11959. https://doi.org/10.1021/jp406142g

Lukyanov, K.A., Fradkov, A.F., Gurskaya, N.G., Matz, M.V., Labas, Y.A., Savitsky, A.P., Markelov, M.L., Zaraisky, A.G., Zhao, X., Fang, Y., Tan, W., Lukyanov, S.A., 2000. Natural animal coloration can Be determined by a nonfluorescent green fluorescent protein homolog. *J. Biol. Chem.*, 275, 25879–25882. https://doi.org/10.1074/jbc.C000338200

Margaritondo, G., Rebernik Ribic, P., 2011. A simplified description of X-ray free-electron lasers. *J. Synchrotron Radiat.*, 18, 101–108. https://doi.org/10.1107/S090904951004896X

Marinelli, A., Ratner, D., Lutman, A.A., Turner, J., Welch, J., Decker, F.-J., Loos, H., Behrens, C., Gilevich, S., Miahnahri, A.A., Vetter, S., Maxwell, T.J., Ding, Y., Coffee, R., Wakatsuki, S., Huang, Z., 2015. High-intensity double-pulse X-ray free-electron laser. *Nat. Commun.*, 6, 6369. https://doi.org/10.1038/ncomms7369

Martin-Garcia, J.M., Conrad, C.E., Nelson, G., Stander, N., Zatsepin, N.A., Zook, J., Zhu, L., Geiger, J., Chun, E., Kissick, D., Hilgart, M.C., Ogata, C., Ishchenko, A., Nagaratnam, N., Roy-Chowdhury, S., Coe, J., Subramanian, G., Schaffer, A., James, D., Ketwala, G., Venugopalan, N., Xu, S., Corcoran, S., Ferguson, D., Weierstall, U., Spence, J.C.H., Cherezov, V., Fromme, P., Fischetti, R.F., Liu, W., 2017. Serial millisecond crystallography of membrane and soluble protein microcrystals using synchrotron radiation. *IUCrJ*, 4, 439–454. https://doi.org/10.1107/S205225251700570X

Mizuno, H., Mal, T.K., Wälchli, M., Kikuchi, A., Fukano, T., Ando, R., Jeyakanthan, J., Taka, J., Shiro, Y., Ikura, M., Miyawaki, A., 2008. Light-dependent regulation of structural flexibility in a photochromic fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.*, 105, 9227–9232. https://doi.org/10.1073/pnas.0709599105

Mozzarelli, A., Rossi, G.L., 1996. Protein function in the crystal. *Annu. Rev. Biophys. Biomol. Struct.*, 25, 343–365. https://doi.org/10.1146/annurev.bb.25.060196.002015

Murray, T.D., Lyubimov, A.Y., Ogata, C.M., Vo, H., Uervirojnangkoorn, M., Brunger, A.T., Berger, J.M., 2015. A high-transparency, micro-patternable chip for X-ray diffraction analysis of microcrystals under native growth conditions. *Acta Crystallogr. D Biol. Crystallogr.*, 71, 1987–1997. https://doi.org/10.1107/S1399004715015011

Nakane, T., Hanashima, S., Suzuki, M., Saiki, H., Hayashi, T., Kakinouchi, K., Sugiyama, S., Kawatake, S., Matsuoka, S., Matsumori, N., Nango, E., Kobayashi, J., Shimamura, T., Kimura, K., Mori, C., Kunishima, N., Sugahara, M., Takakyu, Y., Inoue, S., Masuda, T., Hosaka, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Inoue, T., Nureki, O., Iwata, S., Murata, M., Mizohata, E., 2016. Membrane protein structure determination by SAD, SIR, or SIRAS phasing in serial femtosecond crystallography using an iododetergent. *Proc. Natl. Acad. Sci. U. S. A.*, 113, 13039–13044. https://doi.org/10.1073/pnas.1602531113

Nakane, T., Song, C., Suzuki, M., Nango, E., Kobayashi, J., Masuda, T., Inoue, S., Mizohata, E., Nakatsu, T., Tanaka, T., Tanaka, R., Shimamura, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Nureki, O., Iwata, S., Sugahara, M., 2015. Native sulfur/chlorine SAD phasing for serial femtosecond crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, 71, 2519–2525. https://doi.org/10.1107/S139900471501857X

Nango, E., Royant, A., Kubo, M., Nakane, T., Wickstrand, C., Kimura, T., Tanaka, T., Tono, K., Song, C., Tanaka, R., Arima, T., Yamashita, A., Kobayashi, J., Hosaka, T., Mizohata, E., Nogly, P., Sugahara, M., Nam, D., Nomura, T., Shimamura, T., Im, D., Fujiwara, T., Yamanaka, Y., Jeon, B., Nishizawa, T., Oda, K., Fukuda, M., Andersson, R., Båth, P., Dods, R., Davidsson, J., Matsuoka, S., Kawatake, S., Murata, M., Nureki, O., Owada, S., Kameshima, T., Hatsui, T., Joti, Y., Schertler, G., Yabashi, M., Bondar, A.-N., Standfuss, J., Neutze, R., Iwata, S., 2016. A three-dimensional movie of structural changes in bacteriorhodopsin. *Science*, 354, 1552–1557. https://doi.org/10.1126/science.aah3497

Nass, K., Meinhart, A., Barends, T.R.M., Foucar, L., Gorel, A., Aquila, A., Botha, S., Doak, R.B., Koglin, J., Liang, M., Shoeman, R.L., Williams, G., Boutet, S., Schlichting, I., 2016. Protein structure determination by single-wavelength anomalous diffraction phasing of X-ray free-electron laser data. *IUCrJ*, 3, 180–191. https://doi.org/10.1107/S2052252516002980

Neutze, R., Wouts, R., van der Spoel, D., Weckert, E., Hajdu, J., 2000. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature*, 406, 752–757. https://doi.org/10.1038/35021099

Nogly, P., Panneels, V., Nelson, G., Gati, C., Kimura, T., Milne, C., Milathianaki, D., Kubo, M., Wu, W., Conrad, C., Coe, J., Bean, R., Zhao, Y., Båth, P., Dods, R., Harimoorthy, R., Beyerlein, K.R., Rheinberger, J., James, D., DePonte, D., Li, C., Sala, L., Williams, G.J., Hunter, M.S., Koglin, J.E., Berntsen, P., Nango, E., Iwata, S., Chapman, H.N., Fromme, P., Frank, M., Abela, R., Boutet, S., Barty, A., White, T.A., Weierstall, U., Spence, J., Neutze, R., Schertler, G., Standfuss, J., 2016. Lipidic cubic phase injector is a viable crystal delivery system for time-resolved serial crystallography. *Nat. Commun.*, 7, 12314. https://doi.org/10.1038/ncomms12314

Nogly, P., Weinert, T., James, D., Carbajo, S., Ozerov, D., Furrer, A., Gashi, D., Borin, V., Skopintsev, P., Jaeger, K., Nass, K., Båth, P., Bosman, R., Koglin, J., Seaberg, M., Lane, T., Kekilli, D., Brünle, S., Tanaka, T., Wu, W., Milne, C., White, T., Barty, A., Weierstall, U., Panneels, V., Nango, E., Iwata, S., Hunter, M., Schapiro, I., Schertler, G., Neutze, R., Standfuss, J., 2018. Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser. *Science*, 361. https://doi.org/10.1126/science.aat0094

Oberthuer, D., Knoška, J., Wiedorn, M.O., Beyerlein, K.R., Bushnell, D.A., Kovaleva, E.G., Heymann, M., Gumprecht, L., Kirian, R.A., Barty, A., Mariani, V., Tolstikova, A., Adriano, L., Awel, S., Barthelmess, M., Dörner, K., Xavier, P.L., Yefanov, O., James, D.R., Nelson, G., Wang, D., Calvey, G., Chen, Y., Schmidt, A., Szczepek, M., Frielingsdorf, S., Lenz, O., Snell, E., Robinson, P.J., Šarler, B., Belšak, G., Maček, M., Wilde, F., Aquila, A., Boutet, S., Liang, M., Hunter, M.S., Scheerer, P., Lipscomb, J.D., Weierstall, U., Kornberg, R.D., Spence, J.C.H., Pollack, L., Chapman, H.N., Bajt, S., 2017. Double-flow focused liquid injector for efficient serial femtosecond crystallography. *Sci. Rep.*, 7, 44628. https://doi.org/10.1038/srep44628

Owen, R.L., Axford, D., Sherrell, D.A., Kuo, A., Ernst, O.P., Schulz, E.C., Miller, R.J.D., Mueller-Werkmeister, H.M., 2017. Low-dose fixed-target serial synchrotron crystallography. *Acta Crystallogr. Sect. Struct. Biol.*, 73, 373–378. https://doi.org/10.1107/S2059798317002996

Pande, K., Hutchison, C.D.M., Groenhof, G., Aquila, A., Robinson, J.S., Tenboer, J., Basu, S., Boutet, S., DePonte, D.P., Liang, M., White, T.A., Zatsepin, N.A., Yefanov, O., Morozov, D., Oberthuer, D., Gati, C., Subramanian, G., James, D., Zhao, Y., Koralek, J., Brayshaw, J., Kupitz, C., Conrad, C., Roy-Chowdhury, S., Coe, J.D., Metz, M., Xavier, P.L., Grant, T.D., Koglin, J.E., Ketawala, G., Fromme, R., Šrajer, V., Henning, R., Spence, J.C.H., Ourmazd, A., Schwander, P., Weierstall, U., Frank, M., Fromme, P., Barty, A., Chapman, H.N., Moffat, K., van Thor, J.J., Schmidt, M., 2016. Femtosecond Structural Dynamics Drives the Trans/Cis Isomerization in Photoactive Yellow Protein. *Science*, 352, 725–729. https://doi.org/10.1126/science.aad5081

Patterson, G.H., Lippincott-Schwartz, J., 2002. A photoactivatable GFP for selective photolabeling of proteins and cells. *Science*, 297, 1873–1877. https://doi.org/10.1126/science.1074952

Perman, B., Srajer, V., Ren, Z., Teng, T., Pradervand, C., Ursby, T., Bourgeois, D., Schotte, F., Wulff, M., Kort, R., Hellingwerf, K., Moffat, K., 1998. Energy transduction on the nanosecond time scale: early structural events in a xanthopsin photocycle. *Science*, 279, 1946–1950.

Redecke, L., Nass, K., DePonte, D.P., White, T.A., Rehders, D., Barty, A., Stellato, F., Liang, M., Barends, T.R.M., Boutet, S., Williams, G.J., Messerschmidt, M., Seibert, M.M., Aquila, A., Arnlund, D., Bajt, S., Barth, T., Bogan, M.J., Caleman, C., Chao, T.-C., Doak, R.B., Fleckenstein, H., Frank, M., Fromme, R., Galli, L., Grotjohann, I., Hunter, M.S., Johansson, L.C., Kassemeyer, S., Katona, G., Kirian, R.A., Koopmann, R., Kupitz, C., Lomb, L., Martin, A.V., Mogk, S., Neutze, R., Shoeman, R.L., Steinbrener, J., Timneanu, N., Wang, D., Weierstall, U., Zatsepin, N.A., Spence, J.C.H., Fromme, P., Schlichting, I., Duszenko, M., Betzel, C., Chapman, H.N., 2013. Natively inhibited Trypanosoma brucei cathepsin B structure determined by using an X-ray laser. *Science*, 339, 227–230. https://doi.org/10.1126/science.1229663

Roedig, P., Ginn, H.M., Pakendorf, T., Sutton, G., Harlos, K., Walter, T.S., Meyer, J., Fischer, P., Duman, R., Vartiainen, I., Reime, B., Warmer, M., Brewster, A.S., Young, I.D., Michels-Clark, T., Sauter, N.K., Kotecha, A., Kelly, J., Rowlands, D.J., Sikorsky, M., Nelson, S., Damiani, D.S., Alonso-Mori, R., Ren, J., Fry, E.E., David, C., Stuart, D.I., Wagner, A., Meents, A., 2017. High-speed fixed-target serial virus crystallography. *Nat. Methods*, 14, 805–810. https://doi.org/10.1038/nmeth.4335

Roessler, C.G., Agarwal, R., Allaire, M., Alonso-Mori, R., Andi, B., Bachega, J.F.R., Bommer, M., Brewster, A.S., Browne, M.C., Chatterjee, R., Cho, E., Cohen, A.E., Cowan, M., Datwani, S., Davidson, V.L., Defever, J., Eaton, B., Ellson, R., Feng, Y., Ghislain, L.P., Glownia, J.M., Han, G., Hattne, J., Hellmich, J., Héroux, A., Ibrahim, M., Kern, J., Kuczewski, A., Lemke, H.T., Liu, P., Majlof, L., McClintock, W.M., Myers, S., Nelsen, S., Olechno, J., Orville, A.M., Sauter, N.K., Soares, A.S., Soltis, S.M., Song, H., Stearns, R.G., Tran, R., Tsai, Y., Uervirojnangkoorn, M., Wilmot, C.M., Yachandra, V., Yano, J., Yukl, E.T., Zhu, D., Zouni, A., 2016. Acoustic Injectors for Drop-On-Demand Serial Femtosecond Crystallography. *Struct. Lond. Engl.*, 1993 24, 631–640. https://doi.org/10.1016/j.str.2016.02.007

Sauter, N.K., 2015. XFEL diffraction: developing processing methods to optimize data quality. *J. Synchrotron Radiat.*, 22, 239–248. https://doi.org/10.1107/S1600577514028203

Sauter, N.K., Echols, N., Adams, P.D., Zwart, P.H., Kern, J., Brewster, A.S., Koroidov, S., Alonso-Mori, R., Zouni, A., Messinger, J., Bergmann, U., Yano, J., Yachandra, V.K., 2016. No observable conformational changes in PSII. *Nature*, 533, E1–2. https://doi.org/10.1038/nature17983

Sauter, N.K., Hattne, J., Grosse-Kunstleve, R.W., Echols, N., 2013. New Python-based methods for data processing. *Acta Crystallogr. D Biol. Crystallogr.*, 69, 1274–1282. https://doi.org/10.1107/S0907444913000863

Sawaya, M.R., Cascio, D., Gingery, M., Rodriguez, J., Goldschmidt, L., Colletier, J.-P., Messerschmidt, M.M., Boutet, S., Koglin, J.E., Williams, G.J., Brewster, A.S., Nass, K., Hattne, J., Botha, S., Doak, R.B., Shoeman, R.L., DePonte, D.P., Park, H.-W., Federici, B.A., Sauter, N.K., Schlichting, I., Eisenberg, D.S., 2014. Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam. *Proc. Natl. Acad. Sci. U. S. A.*, 111, 12769–12774. https://doi.org/10.1073/pnas.1413456111

Schirò, G., Woodhouse, J., Weik, M., Schlichting, I., Shoeman, R.L., 2017. Simple and efficient system for photoconverting light-sensitive proteins in serial crystallography experiments. *J. Appl. Crystallogr.*, 50, 932–939. https://doi.org/10.1107/S1600576717006264

Schlichting, I., Almo, S.C., Rapp, G., Wilson, K., Petratos, K., Lentfer, A., Wittinghofer, A., Kabsch, W., Pai, E.F., Petsko, G.A., 1990. Time-resolved X-ray crystallographic study of the conformational change in Ha-Ras p21 protein on GTP hydrolysis. *Nature*, 345, 309–315. https://doi.org/10.1038/345309a0

Schmidt, M., 2013. Mix and Inject: Reaction Initiation by Diffusion for Time-Resolved Macromolecular Crystallography [WWW Document]. *Adv. Condens. Matter Phys.*, URL https://www.hindawi.com/journals/acmp/2013/167276/ (accessed 8.15.18).

Schotte, F., Cho, H.S., Kaila, V.R.I., Kamikubo, H., Dashdorj, N., Henry, E.R., Graber, T.J., Henning, R., Wulff, M., Hummer, G., Kataoka, M., Anfinrud, P.A., 2012. Watching a signaling protein function in real time via 100-ps time-resolved Laue crystallography. *Proc. Natl. Acad. Sci. U. S. A.*, 109, 19256–19261. https://doi.org/10.1073/pnas.1210938109

Schotte, F., Lim, M., Jackson, T.A., Smirnov, A.V., Soman, J., Olson, J.S., Phillips, G.N., Wulff, M., Anfinrud, P.A., 2003. Watching a protein as it functions with 150-ps time-resolved x-ray crystallography. *Science*, 300, 1944–1947. https://doi.org/10.1126/science.1078797

Sherrell, D.A., Foster, A.J., Hudson, L., Nutter, B., O'Hea, J., Nelson, S., Paré-Labrosse, O., Oghbaey, S., Miller, R.J.D., Owen, R.L., 2015. A modular and compact portable mini-endstation for high-precision, high-speed fixed target serial crystallography at FEL and synchrotron sources. *J. Synchrotron Radiat.*, 22, 1372–1378. https://doi.org/10.1107/S1600577515016938

Shimada, A., Kubo, M., Baba, S., Yamashita, K., Hirata, K., Ueno, G., Nomura, T., Kimura, T., Shinzawa-Itoh, K., Baba, J., Hatano, K., Eto, Y., Miyamoto, A., Murakami, H., Kumasaka, T., Owada, S., Tono, K., Yabashi, M., Yamaguchi, Y., Yanagisawa, S., Sakaguchi, M., Ogura, T., Komiya, R., Yan, J., Yamashita, E., Yamamoto, M., Ago, H., Yoshikawa, S., Tsukihara, T., 2017. A nanosecond time-resolved XFEL analysis of structural changes associated with CO release from cytochrome c oxidase. *Sci. Adv.*, 3, e1603042. https://doi.org/10.1126/sciadv.1603042

Sierra, R.G., Gati, C., Laksmono, H., Dao, E.H., Gul, S., Fuller, F., Kern, J., Chatterjee, R., Ibrahim, M., Brewster, A.S., Young, I.D., Michels-Clark, T., Aquila, A., Liang, M., Hunter, M.S., Koglin, J.E., Boutet, S., Junco, E.A., Hayes, B., Bogan, M.J., Hampton, C.Y., Puglisi, E.V., Sauter, N.K., Stan, C.A., Zouni, A., Yano, J., Yachandra, V.K., Soltis, S.M., Puglisi, J.D., DeMirci, H., 2016. Concentric-flow electrokinetic injector enables serial crystallography of ribosome and photosystem II. *Nat. Methods*, 13, 59–62. https://doi.org/10.1038/nmeth.3667

Sierra, R.G., Laksmono, H., Kern, J., Tran, R., Hattne, J., Alonso-Mori, R., Lassalle-Kaiser, B., Glöckner, C., Hellmich, J., Schafer, D.W., Echols, N., Gildea, R.J., Grosse-Kunstleve, R.W., Sellberg, J., McQueen, T.A., Fry, A.R., Messerschmidt, M.M., Miahnahri, A., Seibert, M.M., Hampton, C.Y., Starodub, D., Loh, N.D., Sokaras, D., Weng, T.-C., Zwart, P.H., Glatzel, P., Milathianaki, D., White, W.E., Adams, P.D., Williams, G.J., Boutet, S., Zouni, A., Messinger, J., Sauter, N.K., Bergmann, U., Yano, J., Yachandra, V.K., Bogan, M.J., 2012. Nanoflow electrospinning serial femtosecond crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, 68, 1584–1587. https://doi.org/10.1107/S0907444912038152

Srajer, V., Teng, T., Ursby, T., Pradervand, C., Ren, Z., Adachi, S., Schildkamp, W., Bourgeois, D., Wulff, M., Moffat, K., 1996. Photolysis of the carbon monoxide complex of myoglobin: nanosecond time-resolved crystallography. *Science*, 274, 1726–1729.

Stagno, J.R., Liu, Y., Bhandari, Y.R., Conrad, C.E., Panja, S., Swain, M., Fan, L., Nelson, G., Li, C., Wendel, D.R., White, T.A., Coe, J.D., Wiedorn, M.O., Knoska, J., Oberthuer, D., Tuckey, R.A., Yu, P., Dyba, M., Tarasov, S.G., Weierstall, U., Grant, T.D., Schwieters, C.D., Zhang, J., Ferré-D'Amaré, A.R., Fromme, P., Draper, D.E., Liang, M., Hunter, M.S., Boutet, S., Tan, K., Zuo, X., Ji, X., Barty, A., Zatsepin, N.A., Chapman, H.N., Spence, J.C.H., Woodson, S.A., Wang, Y.-X., 2017. Structures of riboswitch RNA reaction states by mix-and-inject XFEL serial crystallography. *Nature*, 541, 242–246. https://doi.org/10.1038/nature20599

Stellato, F., Oberthür, D., Liang, M., Bean, R., Gati, C., Yefanov, O., Barty, A., Burkhardt, A., Fischer, P., Galli, L., Kirian, R.A., Meyer, J., Panneerselvam, S., Yoon, C.H., Chervinskii, F., Speller, E., White, T.A., Betzel, C., Meents, A., Chapman, H.N., 2014. Room-temperature macromolecular serial crystallography using synchrotron radiation. *IUCrJ*, 1, 204–212. https://doi.org/10.1107/S2052252514010070

Stevenson, H.P., Lin, G., Barnes, C.O., Sutkeviciute, I., Krzysiak, T., Weiss, S.C., Reynolds, S., Wu, Y., Nagarajan, V., Makhov, A.M., Lawrence, R., Lamm, E., Clark, L., Gardella, T.J., Hogue, B.G., Ogata, C.M., Ahn, J., Gronenborn, A.M., Conway, J.F., Vilardaga, J.P., Cohen, A.E., Calero, G., 2016. Transmission electron microscopy for the evaluation and optimization of crystal growth. *Acta Crystallogr. Sect. Struct. Biol.*, 72, 603–615. https://doi.org/10.1107/S2059798316001546

Stevenson, H.P., Makhov, A.M., Calero, M., Edwards, A.L., Zeldin, O.B., Mathews, I.I., Lin, G., Barnes, C.O., Santamaria, H., Ross, T.M., Soltis, S.M., Khosla, C., Nagarajan, V., Conway, J.F., Cohen, A.E., Calero, G., 2014. Use of transmission electron microscopy to identify nanocrystals of challenging protein targets. *Proc. Natl. Acad. Sci. U. S. A.*, 111, 8470–8475. https://doi.org/10.1073/pnas.1400240111

Sugahara, M., Mizohata, E., Nango, E., Suzuki, M., Tanaka, T., Masuda, T., Tanaka, R., Shimamura, T., Tanaka, Y., Suno, C., Ihara, K., Pan, D., Kakinouchi, K., Sugiyama, S., Murata, M., Inoue, T., Tono, K., Song, C., Park, J., Kameshima, T., Hatsui, T., Joti, Y., Yabashi, M., Iwata, S., 2015. Grease matrix as a versatile carrier of proteins for serial crystallography. *Nat. Methods*, 12, 61–63. https://doi.org/10.1038/nmeth.3172

Sugahara, M., Nakane, T., Masuda, T., Suzuki, M., Inoue, S., Song, C., Tanaka, R., Nakatsu, T., Mizohata, E., Yumoto, F., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Nureki, O., Numata, K., Nango, E., Iwata, S., 2017. Hydroxyethyl cellulose matrix applied to serial crystallography. *Sci. Rep.*, 7, 703. https://doi.org/10.1038/s41598-017-00761-0

Sugahara, M., Song, C., Suzuki, M., Masuda, T., Inoue, S., Nakane, T., Yumoto, F., Nango, E., Tanaka, R., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Nureki, O., Numata, K., Iwata, S., 2016. Oil-free hyaluronic acid matrix for serial femtosecond crystallography. *Sci. Rep.*, 6, 24484. https://doi.org/10.1038/srep24484

Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H., Shen, J.-R., 2015. Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature*, 517, 99–103. https://doi.org/10.1038/nature13991

Suga, M., Akita, F., Sugahara, M., Kubo, M., Nakajima, Y., Nakane, T., Yamashita, K., Umena, Y., Nakabayashi, M., Yamane, T., Nakano, T., Suzuki, M., Masuda, T., Inoue, S., Kimura, T., Nomura, T., Yonekura, S., Yu, L.-J., Sakamoto, T., Motomura, T., Chen, J.-H., Kato, Y., Noguchi, T., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Nango, E., Tanaka, R., Naitow, H., Matsuura, Y., Yamashita, A., Yamamoto, M., Nureki, O., Yabashi, M., Ishikawa, T., Iwata, S., Shen, J.-R., 2017. Light-induced structural changes and the site of O=O bond formation in PSII caught by XFEL. *Nature*, 543, 131–135. https://doi.org/10.1038/nature21400

Tenboer, J., Basu, S., Zatsepin, N., Pande, K., Milathianaki, D., Frank, M., Hunter, M., Boutet, S., Williams, G.J., Koglin, J.E., Oberthuer, D., Heymann, M., Kupitz, C., Conrad, C., Coe, J., Roy-Chowdhury, S., Weierstall, U., James, D., Wang, D., Grant, T., Barty, A., Yefanov, O., Scales, J., Gati, C., Seuring, C., Srajer, V., Henning, R., Schwander, P., Fromme, R., Ourmazd, A., Moffat, K., Van Thor, J., Spence, J.H.C., Fromme, P., Chapman, H.N., Schmidt, M., 2014. Time-Resolved Serial Crystallography Captures High Resolution Intermediates of Photoactive Yellow Protein. *Science*, 346, 1242–1246. https://doi.org/10.1126/science.1259357

Wampler, R.D., Kissick, D.J., Dehen, C.J., Gualtieri, E.J., Grey, J.L., Wang, H.-F., Thompson, D.H., Cheng, J.-X., Simpson, G.J., 2008. Selective detection of protein crystals by second harmonic microscopy. *J. Am. Chem. Soc.*, 130, 14076–14077. https://doi.org/10.1021/ja805983b

Warren, M.M., Kaucikas, M., Fitzpatrick, A., Champion, P., Sage, J.T., van Thor, J.J., 2013. Ground-state proton transfer in the photoswitching reactions of the fluorescent protein Dronpa. *Nat. Commun.*, 4, 1461. https://doi.org/10.1038/ncomms2460

Weierstall, U., 2014. Liquid sample delivery techniques for serial femtosecond crystallography. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 369, 20130337. https://doi.org/10.1098/rstb.2013.0337

Weierstall, U., James, D., Wang, C., White, T.A., Wang, D., Liu, W., Spence, J.C.H., Bruce Doak, R., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N.A., Liu, H., Basu, S., Wacker, D., Han, G.W., Katritch, V., Boutet, S., Messerschmidt, M., Williams, G.J., Koglin, J.E., Marvin Seibert, M., Klinker, M., Gati, C., Shoeman, R.L., Barty, A., Chapman, H.N., Kirian, R.A., Beyerlein, K.R., Stevens, R.C., Li, D., Shah, S.T.A., Howe, N., Caffrey, M., Cherezov, V., 2014. Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nat. Commun.*, 5, 3309. https://doi.org/10.1038/ncomms4309

White, T.A., 2014. Post-refinement method for snapshot serial crystallography. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 369, 20130330. https://doi.org/10.1098/rstb.2013.0330

White, T.A., Barty, A., Stellato, F., Holton, J.M., Kirian, R.A., Zatsepin, N.A., Chapman, H.N., 2013. Crystallographic data processing for free-electron laser sources. *Acta Crystallogr. D Biol. Crystallogr.*, 69, 1231–1240. https://doi.org/10.1107/S0907444913013620

White, T.A., Kirian, R.A., Martin, A.V., Aquila, A., Nass, K., Barty, A., Chapman, H.N., 2012. CrystFEL: a software suite for snapshot serial crystallography. *J. Appl. Crystallogr.*, 45, 335–341. https://doi.org/10.1107/S0021889812002312

White, T.A., Mariani, V., Brehm, W., Yefanov, O., Barty, A., Beyerlein, K.R., Chervinskii, F., Galli, L., Gati, C., Nakane, T., Tolstikova, A., Yamashita, K., Yoon, C.H., Diederichs, K., Chapman, H.N., 2016. Recent developments in CrystFEL. *J. Appl. Crystallogr.*, 49, 680–689. https://doi.org/10.1107/S1600576716004751

Wickstrand, C., Dods, R., Royant, A., Neutze, R., 2015. Bacteriorhodopsin: Would the real structural intermediates please stand up? *Biochim. Biophys. Acta*, 1850, 536–553. https://doi.org/10.1016/j.bbagen.2014.05.021

Wiedenmann, J., Ivanchenko, S., Oswald, F., Schmitt, F., Röcker, C., Salih, A., Spindler, K.-D., Nienhaus, G.U., 2004. EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. *Proc. Natl. Acad. Sci. U. S. A.*, 101, 15905–15910. https://doi.org/10.1073/pnas.0403668101

Yadav, D., Lacombat, F., Dozova, N., Rappaport, F., Plaza, P., Espagne, A., 2015. Real-time monitoring of chromophore isomerization and deprotonation during the photoactivation of the fluorescent protein Dronpa. *J. Phys. Chem. B*, 119, 2404–2414. https://doi.org/10.1021/jp507094f

Young, I.D., Ibrahim, M., Chatterjee, R., Gul, S., Fuller, F., Koroidov, S., Brewster, A.S., Tran, R., Alonso-Mori, R., Kroll, T., Michels-Clark, T., Laksmono, H., Sierra, R.G., Stan, C.A., Hussein, R., Zhang, M., Douthit, L., Kubin, M., de Lichtenberg, C., Long Vo, P., Nilsson, H., Cheah, M.H., Shevela, D., Saracini, C., Bean, M.A., Seuffert, I., Sokaras, D., Weng, T.-C., Pastor, E., Weninger, C., Fransson, T., Lassalle, L., Bräuer, P., Aller, P., Docker, P.T., Andi, B., Orville, A.M., Glownia, J.M., Nelson, S., Sikorski, M., Zhu, D., Hunter, M.S., Lane, T.J., Aquila, A., Koglin, J.E., Robinson, J., Liang, M., Boutet, S., Lyubimov, A.Y., Uervirojnangkoorn, M., Moriarty, N.W., Liebschner, D., Afonine, P.V., Waterman, D.G., Evans, G., Wernet, P., Dobbek, H., Weis, W.I., Brunger, A.T., Zwart, P.H., Adams, P.D., Zouni, A., Messinger, J., Bergmann, U., Sauter, N.K., Kern, J., Yachandra, V.K., Yano, J.,

2016. Structure of photosystem II and substrate binding at room temperature. *Nature*, 540, 453–457. https://doi.org/10.1038/nature20161

Zander, U., Bourenkov, G., Popov, A.N., de Sanctis, D., Svensson, O., McCarthy, A.A., Round, E., Gordeliy, V., Mueller-Dieckmann, C., Leonard, G.A., 2015. MeshAndCollect: an automated multi-crystal data-collection workflow for synchrotron macromolecular crystallography beamlines. *Acta Crystallogr. D Biol. Crystallogr.*, 71, 2328–2343. https://doi.org/10.1107/S1399004715017927

2. Chapter I: Microcrystallization of rsEGFP2

Introduction

Sample preparation is one of the major bottlenecks in a serial crystallography experiment. Depending on the kind of X-ray source used (synchrotron or free electron laser) and the way the sample is presented to the X-ray beam (liquid injection, high viscosity extrusion (HVE), raster scanning using solid support, …), crystal requirements in terms of size and quantity are not the same. For example, in a serial femtosecond crystallography (SFX) experiment at an X-ray free electron laser (XFEL) using a gas dynamic virtual nozzle (GDVN), micro- and even nanometer sized crystals are suitable because the peak brilliance of this source allows high-resolution diffraction from such tiny crystals. Moreover, larger crystals could clog the injection system. To give another example, in a serial synchrotron crystallography experiment using a HVE injector, larger crystals are required to obtain high-resolution diffraction at a sufficiently high signal-to-noise ratio, given that synchrotron radiation is less brilliant and that scattering from the viscous medium increases background. Each type of serial crystallography experiment thus requires establishing a specific microcrystallization protocol that allows generating crystals of adequate size and in sufficient amounts.

Only a handful of microcrystallization methods for serial crystallography experiments have been published so far: membrane protein crystallization in lipidic cubic phase (Liu et al., 2014); *in vivo* crystallization (Koopmann et al., 2012, Duszenko et al., 2015); batch crystallization and free interface diffusion (Kupitz et al., 2014). Microseeding has been used to produce monodisperse uniformly-sized microcrystals (Ibrahim et al., 2015). These methods can often only be applied successfully for one particular protein or one particular type of protein. There is still no general approach for protein microcrystallization so far.

This chapter presents the establishment and optimization of a microcrystallization protocol for the reversibly switchable enhanced green fluorescent protein 2 (rsEGFP2). All serial crystallography experiments presented in the following chapters relied on these rsEGFP2 microcrystals. For a time-resolved serial femtosecond crystallography (TR-SFX) experiment based on the pump-probe principle and using liquid injection, microcrystals should fulfill the following requirements. In addition to being tiny enough to prevent injector clogging as mentioned above, their size needs to be small enough to allow full penetration of the pump-laser used to trigger the reaction. Figure 1 (provided by Jacques-Philippe Colletier) displays the estimated light transmission through an rsEGFP2 crystal as a function of the crystal thickness. For rsEGFP2 in the *off*-state, less than 10% of the light intensity is transmitted through a 10-µm-sized crystal. For an rsEGFP2 crystal of the same size in the *on*-state, the transmission is close to zero. rsEGFP2 crystals larger than 10 µm are thus not suitable for pump-probe experiments, because beyond this size, a crystal would not be fully illuminated. Moreover, liquid injection is a very sample-consuming delivery method that requires several grams of crystalline protein for a typical TR-SFX experiment (*i.e.* five 12-hours shifts at the LCLS). Therefore, a protocol was required that allows obtaining large quantities of microcrystals with a homogeneous and tunable size distribution, in a reproducible manner. Guidance and super-vision of rsEGFP2 microcrystallization came from our collaborator Prof. Ilme Schlichting (MPI Heidelberg, Germany) who made her unpublished results and protocols available to us.

Results

The first step consists in exploring the phase diagram of the protein in the known crystallization conditions of rsEGFP2 (El Khatib et al., 2016) as illustrated in Figure 2, in order to identify in an empirical manner the zone where microcrystals are obtained with a high probability. The solubility curve separates the under-saturation zone, where the protein is still soluble, from the super-saturation zone. The latter is constituted of the metastable zone, where crystal growth is favored, the precipitation zone, where the protein is so unstable that it precipitates, and the nucleation zone, between the precipitation and the metastable zones, where *de novo* crystal formation is favored over crystal growth. The separation into zones is not rigorous, but provides an idea of how a given protein behaves in a certain crystallization condition and provides a basis for crystallogenesis optimization.

In order to explore the crystallization phase diagram of rsEGFP2 in a known condition (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 8 was used as buffer, ammonium sulfate (AS) was used as precipitant, 4 µl microbatches were set up as sitting drops at various protein and precipitant concentrations (Figure 3, see Materials and Methods section for details). Immediately after adding the protein to the buffer and precipitant solution, the drops turn cloudy. After one night, 20-50 µm-sized crystals appear in some drops, namely at 1.8 and 2.0 M AS, and at protein concentrations of 10 and 20 mg/ml (Figure 3a and b). At 2 M AS, 20 mg/ml protein concentration, crystals are the tiniest (~ 5 × 2 × 2 μm^3, Figure 3b). As shown in Figure 3c, at 2.2 M AS the drops still look cloudy after 1 day and the protein precipitated apparently. But after 3 days, crystals appeared at 10 mg/ml protein concentration and are still appearing and growing after 5 days. At 20 mg/ml, crystals appeared and already reached their final size after 3 days. This first screening thus identified the most interesting condition to be 2 M AS, 20 mg/ml protein, in which 5-10 µm sized crystals were obtained after 1 day. It was then decided to scale up this condition and reproduce it in an Eppendorf tube. Microcrystals of 5-10 µm were observed in a 40-µl batch as well (Figure 3c). It appeared that the crystals were quite inhomogeneous in size and shape. Microseeding was performed to remedy this issue. Seeds were prepared by crushing larger crystals in their own mother liquor until fragments are sub-micrometric in size (Figure 4, see Materials and Methods section for details).

For microseeding, a 50-µl batch was prepared (2 M AS, 115.4 mM HEPES pH 8, 12.3 mM NaCl, 20 mg/ml protein) and seeds were added. It appears that seeding improves homogeneity in size and shape (Figures 5 and 6). In the following, it was decided to investigate the effect on crystal size and shape of changing various parameters, namely AS concentration, amount of seeds, batch volume and temperature. Seeded 50 µl batches were prepared with AS concentrations ranging from 1.4 to 2.0 M and at protein concentrations of 10, 15 and 20 mg/ml. The amount of seeds was kept constant. Needle-shaped crystals of 10 µm in length on average appeared at low AS concentration, and ball-shaped crystals of 3 µm in diameter on average appeared at higher AS concentration (Figure 5). It seems that, in seeded batches, the precipitant concentration affects crystal shape and thus allows its tuning. In contrast, the protein concentration does neither affect crystal size nor shape.

In order to assess the effect of the added amount of seeds, 50 µl batches were prepared at constant AS (2 M) and protein (20 mg/ml) concentrations and with different seed dilutions in the final volume. With decreasing amounts of seeds, the crystal size increases (Figure 6a). Crystals adopt a ball-like shape, as expected at 2 M AS (Figure 5). Thus, it is possible to tune the crystal size by varying the amount of seeds added to the batch. In order to test if the batch volume affects crystal size and shape, seeded batches were set up with volumes ranging from 100 µl to 2 ml (Figure 6b and 6c). If protein, precipitant and seed solutions are rapidly and well mixed, the batch volume is not expected to affect crystal size and shape. Indeed, the batch volume does not affect crystals size and shape in the range tested (Figure 6b and c). A protocol has thus been established that allows for the large-scale microcrystallization required by TR-SFX using liquid-jet injection.

The effect of the temperature was also tested. Microcrystals that grew at room temperature (20 °C) were compared with those grown in the cold room (4 °C). A low amount of seeds (1/3200 v/v) was added in order to obtain sufficiently large crystals to observe significant differences in size and shape, if any. Unequivocally, crystals are of different shape at different temperatures (Figure 6d). At 4 °C, stick-shaped crystals are observed when ball-shaped crystal are obtained at 20 °C, as expected. Temperature thus appears as another parameter to be considered to tune crystal shape.

Discussion

We have established a protocol that allows reproducible generation of homogeneously-sized rsEGFP2 microcrystals whose size and shape can be tuned. A crystallization condition in which macrocrystals were successfully obtained (El Khatib et al., 2016) represented the starting point for exploration of the protein-crystallization phase diagram. The batch method appeared as the method of choice for large-scale microcrystallization. We showed that microcrystallization conditions identified in a small batch volume can be scaled up to a larger volume. Microseeding is critical for obtaining a homogeneous size and shape distribution in the batch. Indeed, adding sub-micrometer crystalline fragments to the batch is an efficient way to trigger and control nucleation. Modification of various parameters in the seeded batch allowed tuning the crystal aspect: the precipitant concentration and temperature affect the shape of the crystals, the amount of seeds affects the size.

Microseeding is an essential aspect of the protocol presented here. Microseeding protocols have already been reported in the context of macrocrystallization. It was thus known that nucleation can be controlled by adding seeds to the drop and that the number of crystals and their size can be controlled by the amount of seeds (D'Arcy et al., 2014). It is commonly admitted that seeding improves the crystal properties (shape, crystalline order, diffraction quality, …), and allows shape tuning. The microseeding approach was applied here for the batch crystallization method as reported earlier (Schlichting, 2015, Ibrahim et al., 2015). The way the seeds are prepared can be critical. Given that microcrystals are to be produced, seeds have to be sub-micrometer in size. Crushing crystals as finely as possible and filtering the obtained fragments is mandatory. Removing larger fragments by filtering increases the seed stock homogeneity in size, that in turn increases the quality of the crystal batch. The way the seeds are added, and by extension, the order in which the solutions are mixed is important. There is no absolute rule, but it can be advised to first mix buffer, salt, precipitant and additives, if any, then add the seeds and finally the concentrated protein solution. If a small volume of seed suspension is added directly to a large volume of protein solution without precipitant, there is a risk that the seeds dissolve. If highly concentrated protein and precipitant solutions were mixed first, there is a risk that nucleation occurs spontaneously and that subsequent microseeding would not be beneficial.

Following the approach presented here, it has been possible to obtain ten grams of microcrystals that were required for TR-SFX experiments presented in the following chapters. Furthermore, investigations on the same protein with other X-ray sources and sample delivery methods other than liquid-jet injection during SFX were made possible given that the established protocol could easily provide crystals in the size and quantities required, by modifying key parameters such as precipitant concentration and the amount of seeds added. It has to be kept in mind that the established protocol is specific to rsEGFP2 and that it might not be successfully applied to another protein. However, we encourage to empirically vary different parameters in order to optimize the crystallogenesis for a protein in a known macrocrystallization condition.

Once a microcrystallization protocol has successfully been established, another hurdle to be taken is large-scale protein production. Indeed, an SFX experiment using liquid injection is a very sample consuming method, requiring grams of protein, making the purification step challenging, even for well-expressing, easy-to-handle proteins, such as fluorescent proteins. Several months of expression and purification were necessary to produce the needed protein quantity. During this period, a dozen of Erlenmeyer flasks were daily used for bacterial culture and expression. With an efficient expression vector, hundreds of milligrams of protein could be extracted for each 500-ml-culture, requiring a huge amount of Ni-NTA resin (several hundreds of milliliters). After elution, large volumes of protein solution were recovered and needed to be dialysed to exchange the buffer. Hundreds of millilters of protein solution needed to be dialysed against tens of liters of buffer (the largest receptacle found in the laboratory was the garbage can). An automatic system was used for the size-exclusion chromatography step, allowing to perform several successive injections on several columns in parallel. After successful purification, protein concentrating starting from liters of solution is yet another challenge. Amicon falcons are not convenient for such a volume. Rather, a system using nitrogen pressure to push the solution through a porous membrane was used to reduce the protein volume in a first step, making feasible the ultra-filtration in a second step to reach the desired concentration for crystallization experiments.

During the SFX experiments using liquid injection described further in the following chapters, the trend of crystals to sticking one to each other was observed. As a consequence, *families* of *social* crystals traveled together through the tubing and clogged the injector. Using a highly diluted crystal suspension limited the frequency with which clogging occurred, yet at the expense of a hit-rate being limited to 2 - 8 %. As a consequence the vast majority of collected diffraction patterns were empty (without diffraction signal), and it took more time to obtain a complete, high-quality dataset. To prevent this loss of precious beam time, a solution was found that consisted in adding fragments of roughly-crushed crystals to the sample. This so called *sand*, prepared analogously to seeds, but without filtration, prevented the aggregation of crystals, and allowed increasing the crystal-concentration during injection and thus the hit rate to 20 – 30 %.

Materials and Methods

Expression and purification

The reversibly switchable enhanced green fluorescent protein (rsEGFP2 (Grotjohann et al., 2012) was expressed and purified as previously described (El Khatib et al., 2016). Briefly, competent E. coli BL21(DE3) cells were transformed. Bacteria were grown in self-inducible medium at 37 °C. After lysis, the protein fused to a N-terminal polyhistidine tag was purified by affinity chromatography using Ni-NTA beads followed by a size-exclusion chromatography step.

Microcrystallization

For microcrystallization assays, rsEGFP2 was concentrated to either 100 mg/ml in a buffer containing 50 mM HEPES pH 7.5, 50 mM NaCl. Microbatches were set up on sitting-drop plates. The volumes were 4 µl and 500 µl for the drop and the reservoir solution, respectively. The final batch condition was 110 mM HEPES pH ~ 8, 10 mM NaCl, with variable protein and AS concentrations (10 and 20 mg/ml and from 1.8 to 2.8 M for the protein and the precipitant, respectively). The tables below summarize the preparation of the microbatches:

Well (500 µl)	1.8 M AS	2.0 M AS	2.2 M AS	2.4 M AS	2.6 M AS	2.8 M AS
4 M AS	225 µl	250 µl	275 µl	300 µl	325 µl	350 µl
1 M HEPES pH 8	50 µl	50 µl	50 µl	50 µl	50 µl	50 µl
H_2O	225 µl	200 µl	175 µl	150 µl	125 µl	100 µl

Drop (4 µl) **10 mg/ml protein**	1.8 M AS	2.0 M AS	2.2 M AS	2.4 M AS	2.6 M AS	2.8 M AS
4 M AS	1.8 µl	2.0 µl	2.2 µl	2.4 µl	2.6 µl	2.8 µl
1 M HEPES pH 8	0.4 µl	0.4 µl	0.4 µl	0.4 µl	0.4 µl	0.4 µl
50 mg/ml	0.8 µl	0.8 µl	0.8 µl	0.8 µl	0.8 µl	0.8 µl
H_2O	1.0 µl	0.8 µl	0.6 µl	0.4 µl	0.2 µl	0 µl

Drop (4 µl) **20 mg/ml protein**	1.8 M AS	2.0 M AS	2.2 M AS	2.4 M AS	2.6 M AS	2.8 M AS
4 M AS	1.8 µl	2.0 µl	2.2 µl	2.4 µl	2.6 µl	2.8 µl
1 M HEPES pH 8	0.4 µl	0.4 µl	0.4 µl	0.4 µl	0.4 µl	0.4 µl
100 mg/ml	0.8 µl	0.8 µl	0.8 µl	0.8 µl	0.8 µl	0.8 µl
H_2O	1.0 µl	0.8 µl	0.6 µl l	0.4 µl	0.2 µl	0 µl

Microcrystals appeared at 20 °C within 12-24 h in some drops and were still growing during the following days.

The most promising condition was scaled up to 40 µl in an Eppendorf tube. 8 µl of concentrated protein (100 mg/ml) were rapidly mixed with 20 µl of AS at 4 M, 4 µl of HEPES pH 8 at 1 M and 8 µl of water. The final concentrations were 2 M AS, 110 mM HEPES pH ~ 8, 10 mM NaCl, 20 mg/ml protein.

To produce seeds, macro- or microcrystals stored in their respective mother liquor of slightly varying conditions were crushed by 1 – 1.5 mm beads in a tube using a Bead Bug Homogenizer (Sigma – Aldrich) by several 4000-rpm cycles of 30 – 60 seconds duration each. Because of the heating during the crushing procedure, tubes were kept for 30-60 seconds in an iced bath between each cycle to prevent protein denaturation or crystal dissolution. The seeds were sequentially filtered through stainless steel frits with 2, 0.5 and then 0.2 µm pore sizes by Dr. Robert Shoeman (MPI Heidelberg). The filtered seeds were washed and stored in 2 M AS, 100 mM HEPES pH 8. Every further dilution of the seeds stock was carried out using this buffer.

In order to asses the effect of AS concentration and of protein concentration on the aspect of the microcrystals in the seeded batches, 50 µl batches were prepared at 20 °C in 115.4 mM HEPES pH ~ 8, 12.3 mM NaCl with an AS concentration from 1.4 to 2 M and a protein concentration of 10, 15 and 20 mg/ml in the final batch. To do so, different volumes of the 4 M AS stock solution were added and the batch completed with water to conserve the desired volume. Sedimented seeds were diluted twice. 2 µl of diluted seeds were added in the batch to reach a 1/50 dilution (v/v) in the final batch. In order to asses the effect of the amount of seeds in the final batches on the aspect of the microcrystals, trials were carried out in 2 M AS, 115.4 mM HEPES pH 8, 12.3 mM NaCl, 20 mg/ml protein. Diluted seeds stocks were prepared by diluting sedimented seeds. 2 µl from diluted seeds stock were added to the 50 µl batches. The table below summarizes the dilution of the seeds stock and the final seeds dilution in the batches:

Seeds dilution in the stock	1/1	1/2	1/4	1/8
Final dilution in the batch	1/25	1/50	1/100	1/200

In order to study the effect of the final batch volume, microcrystallization was carried out in 2 M AS, 115.4 mM HEPES pH 8, 12.3 mM NaCl, 20 mg/ml protein, 1/50 seeds (v/v) in batch volumes of 100, 150, 200, and 250 µl and 2 ml, at 20 °C. The effect of temperature was explored by microcrystallization at 20 and 4 °C in 2 M AS, 115.4 mM HEPES pH 8, 12.3 mM NaCl, 20 mg/ml protein, 1/3200 seeds (v/v) in a volume of 2 ml. In all microcrystalization trials described, buffer and precipitant solution were mixed first, then seeds were added and the protein was added last.

Figures and tables

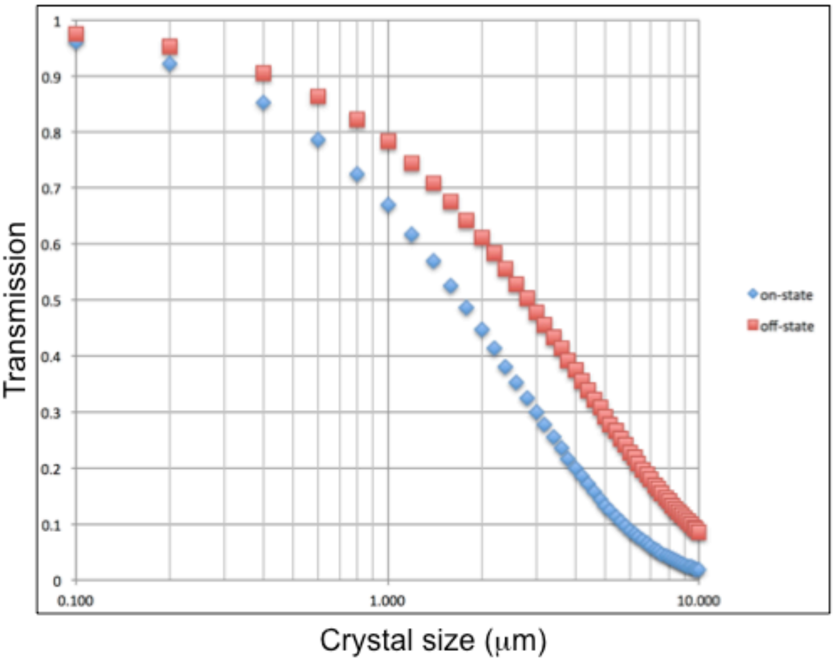


Figure 1: Illumination efficiency as a function of the crystal size. Estimated light transmission as a function of the size of an rsEGFP2 crystal in the *on*-state (blue diamonds) and in the *off*-state (red squares). Figure produced by Jacques-Philippe Colletier.

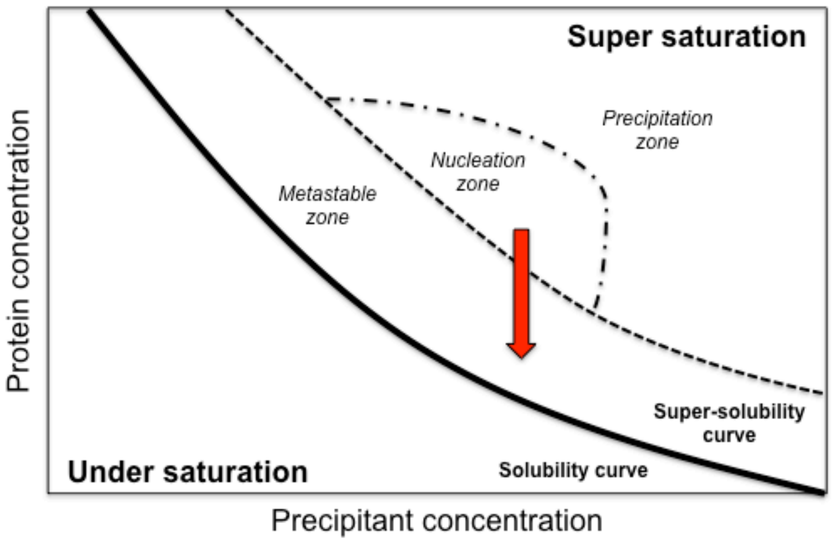


Figure 2: Schematic view of a protein crystallization phase diagram. The under-saturation and super-saturation zones are separated by the solubility curve. The super-saturation zone includes the metastable, the nucleation and the precipitation zones. The red arrow illustrates the change in condition as crystallization in the batch method proceeds. After mixing, precipitant and buffer concentrations do not change, but the protein concentration decreases as crystals appear in the nucleation zone and grow in size. Adapted from Chayen, 1998.

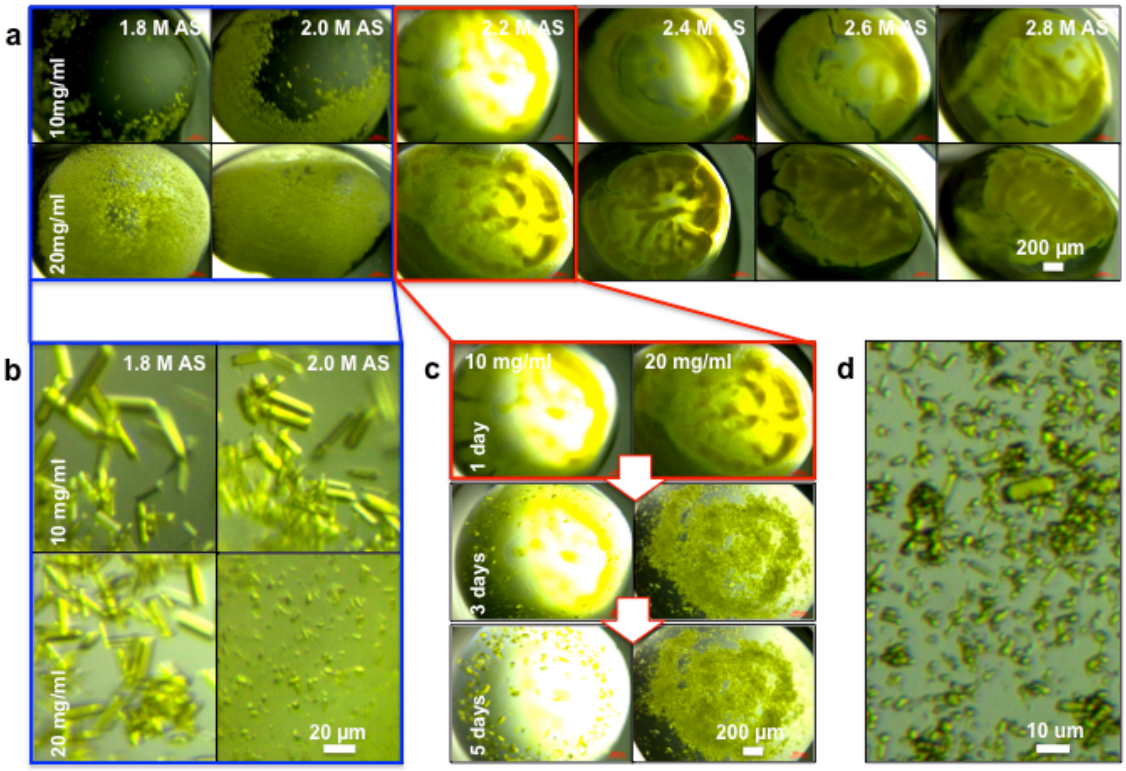


Figure 3: Exploring the phase diagram to find interesting conditions for batch crystallization. (a) Microbatch in 4 µl sitting drops at a protein concentration of 10 mg/ml (top) and 20 mg/ml (bottom) and at a precipitant concentration ranging from 1.8 to 2.8 M AS, 110 mM HEPES pH ~8, 10 mM NaCl, 20 °C, (b) Zoom on microbatch conditions that yield microcrystals after 1 day, (c) Crystal formation followed in time at 2.2 M AS, 10 mM HEPES pH ~8, 10 mM NaCl, 20°C, 10 mg/ml (left) and 20 mg/ml (right) protein concentration. (d) Upscaling the condition with 2 M AS, 110 mM HEPES pH ~8, 10 mM NaCl, protein concentration 20 mg/ml, in a 40 µl batch, at 20°C.

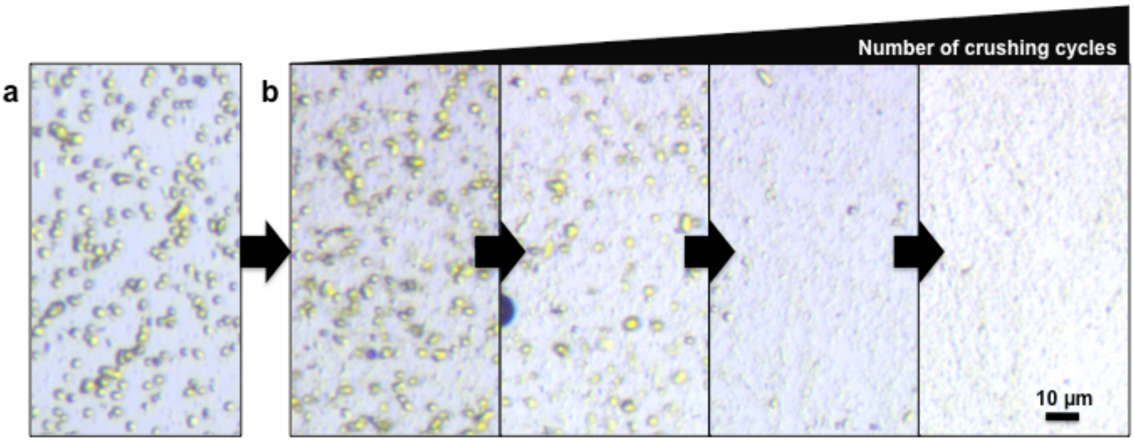


Figure 4: Evolution of the seed size when crushing with a Bead Bug Homogenizer. (a) Microcrystals without crushing. (b) Microcrystalline seeds become smaller with increasing number of crushing cycles.

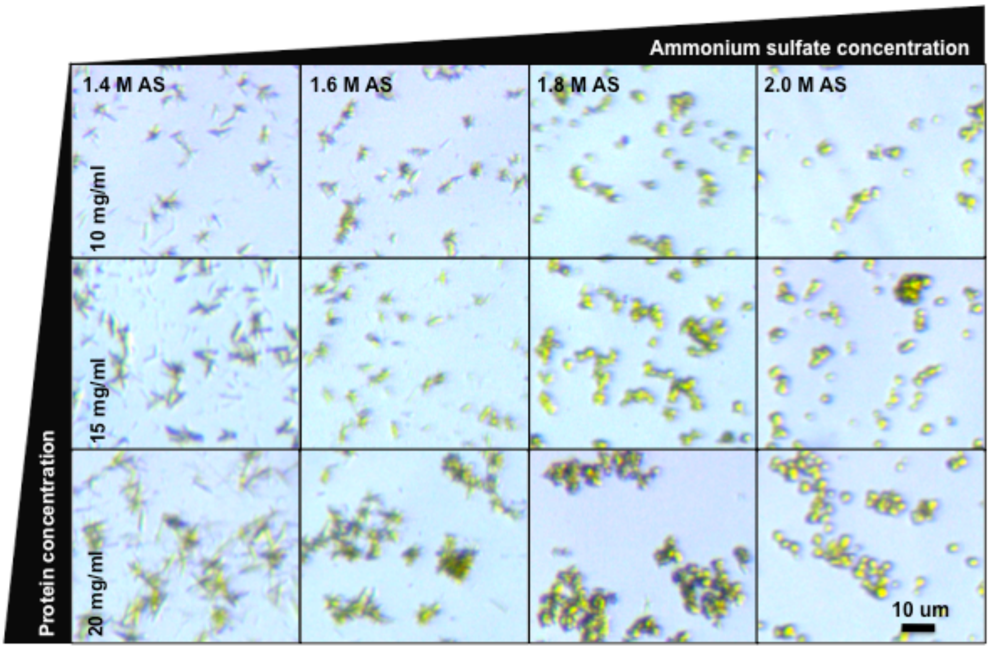


Figure 5 : Effect of the precipitant and protein concentrations on microcrystal shape when using seeding. Crystals obtained at 20 °C in 50 µl batches, 115.4 mM HEPES pH ~8, 12.3 mM NaCl, and in a range of AS concentrations (1.4, 1.6, 1.8 and 2.0 M) and of protein concentrations (10, 15 and 20 mg/ml) with seeds (dilution 1/50 (v/v) in the final batch volume).

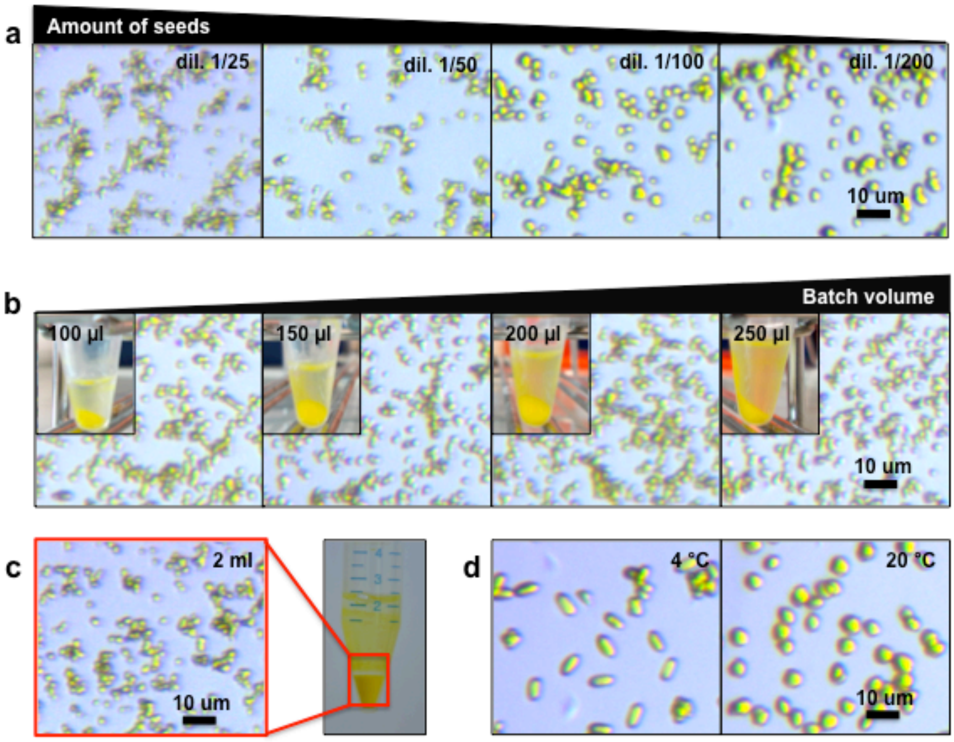


Figure 6 : Seeding allows to reproducibly obtain microcrystals that are homogeneous in size and shape, in a tunable way using the batch method in large volumes (100 µl – 2 ml). (a) Effect of the amount of seeds on microcrystal size with a range of seed dilutions (1/25, 1/50, 1/100, 1/200 v/v) in the final batch volume (50 µl batch, 20 mg/ml protein, 115.4 mM HEPES pH ~8, 12.3 mM NaCl, 2 M AS, 20 °C), (b) and (c) effect of the batch volume (20 mg/ml protein, 115.4 mM HEPES pH ~8, 12.3 mM NaCl, 2 M AS, with seeds at a dilution of 1/50 v/v in the final volume), (d) effect of temperature. Crystals obtained at 20 mg/ml protein, 115.4 mM HEPES pH ~8, 12.3 mM NaCl, 2 M ammonium sulfate, with seeds (dilution 1/3200 (v/v) in the final volume) at 4 and 20 °C.

References

Chayen, N.E., 1998. Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. *Acta Crystallogr. D Biol. Crystallogr.*, 54, 8–15.

D’Arcy, A., Bergfors, T., Cowan-Jacob, S.W., Marsh, M., 2014. Microseed matrix screening for optimization in protein crystallization: what have we learned? *Acta Crystallogr. Sect. F Struct. Biol. Commun.*, 70, 1117–1126. https://doi.org/10.1107/S2053230X14015507

Duszenko, M., Redecke, L., Mudogo, C.N., Sommer, B.P., Mogk, S., Oberthuer, D., Betzel, C., 2015. In vivo protein crystallization in combination with highly brilliant radiation sources offers novel opportunities for the structural analysis of post-translationally modified eukaryotic proteins. *Acta Crystallogr. Sect. F Struct. Biol. Commun.*, 71, 929–937. https://doi.org/10.1107/S2053230X15011450

El Khatib, M., Martins, A., Bourgeois, D., Colletier, J.-P., Adam, V., 2016. Rational design of ultrastable and reversibly photoswitchable fluorescent proteins for super-resolution imaging of the bacterial periplasm. *Sci. Rep.*, 6, 18459. https://doi.org/10.1038/srep18459

Grotjohann, T., Testa, I., Reuss, M., Brakemann, T., Eggeling, C., Hell, S.W., Jakobs, S., 2012. rsEGFP2 enables fast RESOLFT nanoscopy of living cells. *eLife*, 1, e00248. https://doi.org/10.7554/eLife.00248

Ibrahim, M., Chatterjee, R., Hellmich, J., Tran, R., Bommer, M., Yachandra, V.K., Yano, J., Kern, J., Zouni, A., 2015. Improvements in serial femtosecond crystallography of photosystem II by optimizing crystal uniformity using microseeding procedures. *Struct. Dyn.*, Melville N 2. https://doi.org/10.1063/1.4919741

Koopmann, R., Cupelli, K., Redecke, L., Nass, K., Deponte, D.P., White, T.A., Stellato, F., Rehders, D., Liang, M., Andreasson, J., Aquila, A., Bajt, S., Barthelmess, M., Barty, A., Bogan, M.J., Bostedt, C., Boutet, S., Bozek, J.D., Caleman, C., Coppola, N., Davidsson, J., Doak, R.B., Ekeberg, T., Epp, S.W., Erk, B., Fleckenstein, H., Foucar, L., Graafsma, H., Gumprecht, L., Hajdu, J., Hampton, C.Y., Hartmann, A., Hartmann, R., Hauser, G., Hirsemann, H., Holl, P., Hunter, M.S., Kassemeyer, S., Kirian, R.A., Lomb, L., Maia, F.R.N.C., Kimmel, N., Martin, A.V., Messerschmidt, M., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Schlichting, I., Schulz, J., Seibert, M.M., Shoeman, R.L., Sierra, R.G., Soltau, H., Stern, S., Strüder, L., Timneanu, N., Ullrich, J., Wang, X., Weidenspointner, G., Weierstall, U., Williams, G.J., Wunderer, C.B., Fromme, P., Spence, J.C.H., Stehle, T., Chapman, H.N., Betzel, C., Duszenko, M., 2012. In vivo protein crystallization opens new routes in structural biology. *Nat. Methods*, 9, 259–262. https://doi.org/10.1038/nmeth.1859

Kupitz, C., Grotjohann, I., Conrad, C.E., Roy-Chowdhury, S., Fromme, R., Fromme, P., 2014. Microcrystallization techniques for serial femtosecond crystallography using photosystem II from Thermosynechococcus elongatus as a model system. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 369, 20130316. https://doi.org/10.1098/rstb.2013.0316

Liu, W., Ishchenko, A., Cherezov, V., 2014. Preparation of microcrystals in lipidic cubic phase for serial femtosecond crystallography. *Nat. Protoc.*, 9, 2123–2134. https://doi.org/10.1038/nprot.2014.141

Schlichting, I., 2015. Serial femtosecond crystallography: the first five years. *IUCrJ*, 2, 246–255. https://doi.org/10.1107/S205225251402702X

3. Chapter II: Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography

Authors: Nicolas Coquelle[1#], Michel Sliwa[2#], Joyce Woodhouse[1#], Giorgio Schirò[1#], Virgile Adam[1#], Andrew Aquila[3], Thomas R. M. Barends[4], Sébastien Boutet[3], Martin Byrdin[1], Sergio Carbajo[3], Eugenio De la Mora[1], R. Bruce Doak[4], Mikolaj Feliks[1], Franck Fieschi[1], Lutz Foucar[4], Virginia Guillon[1], Mario Hilpert[4], Mark Hunter[3], Stefan Jakobs[5], Jason E. Koglin[3], Gabriela Kovacsova[4], Thomas J. Lane[3], Bernard Lévy[6], Mengning Liang[3], Karol Nass[4], Jacqueline Ridard[6], Joseph S. Robinson[3], Christopher M. Roome[4], Cyril Ruckebusch[2], Matthew Seaberg[3], Michel Thepaut[1], Marco Cammarata[7], Isabelle Demachy[6], Martin Field[1], Robert L. Shoeman[4], Dominique Bourgeois[1], Jacques-Philippe Colletier[1], Ilme Schlichting[4], Martin Weik[1]

My contribution to the work described in this chapter was to express, purify and crystallize the protein, prepare the crystal suspension for injection during XFEL data collection, and take part to the discussion during the writing process (see author contribution p70).

Affiliations:

[1]Institut de Biologie Structurale, Univ. Grenoble Alpes, CEA, CNRS, F-38044 Grenoble, France.
[2] Université de Lille, CNRS, UMR 8516, LASIR, Laboratoire de Spectrochimie Infrarouge et Raman, F59 000 Lille, France.
[3]Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, 2575, Sand Hill Road, Menlo Park, CA 94025, USA.
[4]Max-Planck-Institut für medizinische Forschung, Jahnstrasse 29, 69120 Heidelberg, Germany.
[5]Department of NanoBiophotonics, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany.
[6]Laboratoire de Chimie-Physique, UMR 8000, CNRS/University Paris-Sud, University Paris-Saclay, 91405 Orsay, France.
[7]Department of Physics, UMR UR1-CNRS 6251, University of Rennes 1, Rennes, France.

*Correspondence to: martin.weik@ibs.fr, ilme.schlichting@mpimf-heidelberg.mpg.de, colletier@ibs.fr

#These authors contributed equally to this work.

Abstract:

Chromophores absorb light in photosensitive proteins, thereby initiating fundamental biological processes such as photosynthesis, vision and biofluorescence. A structural description of ultrafast photochemical events, in particular of the excited states that connect chemistry to biological function, has remained elusive. Here, we report on the structures of two excited states in the reversibly photoswitchable fluorescent protein rsEGFP2. We identified the states through femtosecond illumination of rsEGFP2 in its non-fluorescent *off* state and observed their build up (within less than one picosecond) and decay (on the several picosecond time scale). Using an X-ray free-electron laser, we performed picosecond time-resolved crystallography and show that the hydroxybenzylidene imidazolinone chromophore in one of the excited state assumes a near-canonical twisted configuration halfway between the *trans* and *cis* isomers. This is in line with excited-state quantum mechanics/molecular mechanics and classical molecular dynamics simulations. Our new understanding of the structure around the twisted chromophore enabled the design of a mutant that displays a two-fold increase in its *off*-to-*on* photoswitching quantum yield.

Fluorescent proteins (FP) have revolutionized cellular imaging by serving as genetically encoded markers[1]. The propensity of FPs to fluoresce upon visible-light absorption or, alternatively, to relax *via* non-radiative pathways is largely dictated by the excited-state structural dynamics of the protein-embedded chromophore[2,3]. In solution, the isolated chromophore of the green fluorescent protein (GFP[4]) adopts a canonical 90° twisted geometry in the excited state[2] from which it reverts back to the planar ground state solely *via* radiationless decay[5]. The tight β-barrel scaffold of FPs restricts the conformational dynamics of the chromophore, hindering radiationless decay and thus leading to an increase in the fluorescence quantum yield. Modulation of the chromophore twisting has been proposed to affect the photophysical properties of all FPs and has therefore been analysed by numerous theoretical and experimental studies[2,3,6-10].

A subgroup of FPs, so-called reversibly photoswitchable fluorescent proteins (RSFPs)[11], can be toggled back and forth between a fluorescent (*on*) and a non-fluorescent (*off*) state. This property is of considerable interest for advanced fluorescence imaging and biotechnological developments[12]. Photoswitching in RSFPs is a two-step process that involves radiationless decay in the form of *cis*-to-*trans* or *trans*-to-*cis* isomerization and a change in the protonation state of the chromophore, as shown for asFP595[13], Dronpa[10,14,15], Padron[16] and IrisFP[17]. The structural basis of ultrafast events leading to photoswitching, as well as the order in which these proceed, have so far remained elusive[18] as no crystallographic structure of an FP in an electronic excited state has been determined. Here, we report the structural and spectroscopic characterization of excited states in rsEGFP2[19], an efficient tag in super-resolution microscopy, by the combined use of ultrafast transient absorption spectroscopy, time-resolved crystallography using an X-ray laser, and excited-state QM/MM and classical MD simulations. Based on one of the excited-state structures, a mutant was designed that displays a two-fold increase in *off*-to-*on* photoswitching quantum yield.

Results and discussion

Femtosecond transient absorption spectroscopy of rsEGFP2 in solution

The chromophore of rsEGFP2 in its fluorescent *on* state (anionic *cis* isomer) can be photoswitched to the non-fluorescent *off* state (protonated *trans* isomer) by irradiation at 488 nm (Fig. 1 and Supplementary Fig. 13). rsEGFP2 can then be toggled back to the *on* state by 400 nm light. Here, we focus on the *off*-to-*on* photoswitching, whose quantum yield is an order of magnitude higher than that of the *on*-to-*off* photoswitching (supplementary table 7). We used femtosecond transient absorption (TA) spectroscopy to monitor spectral changes induced by excitation at 400 nm of *off*-state rsEGFP2 in solution (Fig. 2 and Supplementary Fig. 1). Spectral changes appear within few hundreds of fs, reach a maximum close to 0.5 ps and decay on a time scale of several ps (Fig. 2). At 0.5 ps two positive (maxima at 355 and 455 nm) and two negative bands (a narrow and a broad band with maxima at 405 and 530 nm, respectively) are observed. The former are characteristic of absorption by excited-state species, while the latter indicate depopulation of the *off* state (405 nm) and stimulated emission of excited-state species (530 nm). A global decay analysis was carried out, where we focused on spectral changes observed at four wavelengths (385, 455, 530 and 650 nm). This analysis yielded three time constants, *i.e.* 90 fs, 0.9 ps and 3.7 ps (Fig. 2b). We attribute the first time constant to the build-up, from the Franck-Condon state, of two electronically-excited intermediates, which respectively decay with 0.9 and 3.7 ps time constants to ground states - that is, either backward to the protonated *trans* isomer or forward to a still protonated *cis* isomer. The decay of the stimulated emission band from 500 to 700 nm needed to be fitted with two exponentials, which implies necessarily that there are two excited-state species. That the *cis* isomer is still protonated is reflected by the appearance of a positive band at 380 nm (see Supplementary Information for details). Similar transient spectra were obtained on similar time scales (Supplementary Fig. 1f-j) when measurements were performed in heavy (D_2O) instead of light water (H_2O), suggesting that excited-state proton-transfer (ESPT) is unlikely. Our transient spectroscopic data are thus consistent with *trans*-to-*cis* isomerization occurring in the excited state within picoseconds, followed by ground-state proton transfer on a slower time-scale (Fig. 3), in line with previous studies on Dronpa[10] and IrisFP[18].

Ground state structures determined by static serial femtosecond crystallography

We investigated the geometry of the excited state experimentally, by time-resolved serial femtosecond crystallography (TR-SFX) at an X-ray free-electron laser (XFEL), a method that can provide structural snapshots on the sub-picosecond time scale[20,21]. First, to determine the *on*-state structure we collected static SFX data at room temperature (RT) on rsEGFP2 in the *on* state (reference dataset) from microcrystals (Supplementary Fig. 12) streamed across the X-ray beam using a liquid microjet. The RT structure of rsEGFP2 in the *off* state was obtained from static SFX (laser-off data set) of microcrystals that were pre-illuminated prior to injection with 488 nm cw laser-light in a custom-made inline device[22].

Careful analysis of electron density maps calculated from the laser-off data set showed that it was necessary to include the *on* state at an occupancy of 10% in the model, the remainder corresponding to the *off* state (see Supplementary Information). Since these occupancies are in line with the 90% pre-illumination efficiency orthogonally assessed by absorption spectroscopy (see Supplementary Information) it serves as an internal control, establishing that our data allow the identification of small subpopulations reliably. The rsEGFP2 *off* and *on* state structures, both refined against 1.7 Å resolution data (Supplementary Table 5), show the hydroxybenzylidene imidazolinone (HBI) chromophore in its *trans* and its *cis* isomer state, respectively, held by a central α-helix within a β-barrel typical of all fluorescent proteins (Supplementary Figs. 16, 17, 18). Conformational changes previously identified by synchrotron methods[23] that accompany HBI isomerization between the *off* and *on* states are clearly visible (Tyr146, Asn147, His149, Val151, Tyr152).

Excited-state structures determined by time-resolved serial femtosecond crystallography

We then set out to visualize the ultrafast structural changes occurring after photoexcitation of the *off* state. We could not perform a pump-laser power titration due to scarcity of XFEL beamtime. We decided to err on the high rather than on the low side and used a very high laser power density of nominally 400 GW/cm^2. Our reasoning for this choice was that previous experiments[20,21] used a similar power density and that the spatial overlap of the pump laser focus with the XFEL beam at the interaction region can decrease due to drifts. This was indeed observed during the experiment, causing a 2 - 5 fold reduction of the nominal value before manual realignment, which resulted in an actual power density range between 80 and 400 GW/cm^2. rsEGFP2 microcrystals were pre-illuminated with 488 nm laser-light to generate the *off* state and TR-SFX data were collected at pump-probe delays of 1 ps (laser-on-Δ1ps data set) and 3 ps (laser-on-Δ3ps data set) after pump-laser excitation at 400 nm to yield intermediate-state structures. These delays were chosen on the basis of the time course of the spectral changes observed in solution (Fig. 2b), which are close to their maximum at 1 ps, and have substantially decayed at 3 ps. A difference Fourier map was calculated at 1.9 Å resolution between the laser-on-Δ1ps and laser-off data sets ($F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$), which revealed strong peaks on the chromophore and surrounding residues (Fig. 4a). The highest peaks are located on the imidazolinone ring of the chromophore, on the methylene bridge, and to a lesser extent on the phenol ring (Fig. 4c, d and Supplementary Fig. 28). A strong negative peak is located on water molecule #17. The strongest positive peaks are compatible with the presence of two new chromophore conformers in the laser-on-Δ1ps dataset (Fig. 4c, d, Supplementary Figs. 23-25). The first conformer (model P (planar)) is very close to the *off*-state conformer, whereas the second (model T (twist)) is midway between the *off*- and *on*-state conformers (Fig. 4c,d, Supplementary Figs. 23 and 25, Supplementary movie). The water molecule #17 is hydrogen bonded to the protonated phenol group of the chromophore both in the *off* state[19] and in model P, but not in model T, where instead the chromophore phenol oxygen forms a hydrogen bond with water #356. A total occupancy of 6 - 7 % was estimated for these two conformers by two different approaches (see Supplementary Information), leading us to perform difference refinement[24,25]. The laser-on-Δ1ps data are best modeled by a combination of models P and T at relative occupancies of 0.6 and 0.4, respectively (see Supplementary Information). A simulated difference Fourier map with the experimentally determined absolute occupancies of 4 and 3% for models P and T, respectively, reproduced the features in the $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ maps (see Supplementary Information).

Structural changes occurring within 3 ps after photoexcitation were assessed by inspecting a $F_{obs}^{laser\text{-}on\text{-}\Delta 3ps}$-$F_{obs}^{laser\text{-}off}$ electron density map (Fig. 4e, f) calculated at 1.9 Å resolution and comparing it to a ($F_{obs}^{laser\text{-}on\text{-}\Delta 1ps\text{-}random}$-$F_{obs}^{laser\text{-}off}$) map calculated from data consisting of the same number of randomly selected indexed images as used for the 1 ps data set (Supplementary Fig. 28b, c). Due to the lower quality of the laser-on-Δ3ps data set, we refrained from performing difference refinement and instead interpreted changes in difference-Fourier map features. At 3 ps, only model T is clearly consistent with the observed peaks (Fig. 4e, f; see Supplementary Information). Integration of peaks associated with model T in the $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps\text{-}random}$-$F_{obs}^{laser\text{-}off}$ and $F_{obs}^{laser\text{-}on\text{-}\Delta 3ps}$-$F_{obs}^{laser\text{-}off}$ maps indicate a 44% decrease between 1 ps and 3 ps, suggesting that the intermediate state corresponding to model T has been partially depleted after 3 ps. There is no clear evidence for the presence of model P at 3 ps, either because the corresponding species has fully decayed or because the lower quality of the laser-on-Δ3ps data set does not allow identification of an intermediate state so close to the *off*-state conformer. Positive peaks overlapping with the chromophore *cis* isomer are present in the $F_{obs}^{laser\text{-}on\text{-}\Delta 3ps}$-$F_{obs}^{laser\text{-}off}$ map (Fig. 4e, f; Supplementary Fig. 28b) but absent in the $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ map (Fig. 4c, d; Supplementary Fig. 28c), suggesting that a fraction of the chromophores has isomerised at 3 ps, but not yet at 1 ps. In addition, integration of negative peaks associated with the *off*-state *trans* chromophore maps reveal a 37% decrease between 1 ps and 3 ps. Thus, not only have intermediate states decayed after 3 ps, but ground states have also populated (*cis* isomer) and repopulated (*off* state).

Excited-state simulations

Photoinduced changes in chromophore conformation on the ultrafast time scale were also explored computationally by means of excited state quantum-mechanics/molecular mechanics (QM/MM) and classical molecular dynamics (MD) simulations. These were carried out starting from the static SFX structure of rsEGFP2 in its *off* state (laser-off data set). During a 100 ps simulation of the rsEGFP2 QM/MM system in its ground state (S_0), the ring-bridging dihedral angles of the chromophore, *tau* and *phi* (for a definition see Fig. 1 and Supplementary Fig 16c), remained close to the ones observed in the crystallographic *off*-state structure. Subsequently, trajectories of 2 ps duration in the S_1 excited state were spawned from the ground state simulation, at 1ps intervals. Supplementary Fig 8 shows the values of the *tau* and *phi* angles as a function of time, averaged over 100 S_1 excited-state trajectories (left panel). Upon excitation the *tau* and *phi* angles change rapidly, from 4 ± 8° and -48 ± 11° in the S_0 state to -86 ± 9° and 16 ± 10° in the S_1 state, respectively (Fig. 4b). This leads to a twisted conformation of the chromophore with perpendicular orientation of the phenol and imidazole groups (Supplementary Fig. 9a), in agreement with previous QM/MM studies on other RSFPs[26]. The change in chromophore conformation is essentially complete after 1 ps, a delay after which model T was observed in the TR-SFX experiments (Fig. 4a-d). Excited-state classical MD simulations were carried out as a complement to QM/MM simulations (see Supplementary Information). For the S_1 state they yielded a major chromophore conformation (*tau* = -73°; *phi* = 11°) similar to the one found by excited-state QM/MM simulations and an additional minor conformation (*tau* = -30°; *phi* = -11°; Fig. 4b). The minor and major MD conformations are accessed on average in less than 30 fs and 1.9 ps, respectively. Both QM/MM and excited-state classical MD simulations thus yield for the electronically excited S_1 state a twisted chromophore conformation very similar to that of model T in the laser-on-Δ1ps TR-SFX structure (Fig. 4b, Supplementary Fig. 9a).

Connection of species and conformations and off-to-on photoswitching model

The high pump laser power density used in the TR-SFX experiments make the possibility that the crystallographically observed structural changes are caused or influenced by ionization-, multi-photon and/or heating effects a legitimate concern, and at this time we cannot exclude this possibility. However, to address these concerns, we performed several additional calculations and experiments (see Supplementary Information) that support our interpretation of the observed structural changes being representative of photoswitching intermediates: a) the effect of multi-photon absorption on structure was addressed by QM/MM calculations on the S_2 and S_3 excited states (see Supplementary Information) which yielded a different conformation than the observed twisted conformation; b) solution studies showed that *off*-to-*on* photoswitching is fully reversible after femtosecond laser excitation at the power density employed in the SFX experiments (Supplementary Fig. 7), making it unlikely that SFX model T corresponds to an irreversibly bleached species. However, it remains possible that multiphoton absorption results in a spectroscopically silent off-pathway intermediate; c) importantly, time-resolved absorption spectroscopy of rsEGFP2 in solution provided circumstantial evidence that a cationic chromophore species indeed forms at 68 GW/cm^2 and 140 GW/cm^2 (see Supplementary Information and Supplementary Fig. 4). However, this species is stable on a time scale of nanoseconds, whereas the species corresponding to SFX model T significantly decays between 1 and 3 ps in TR-SFX; d) DFT calculations of an ionized ground-state cation yielded a planar, not a twisted chromophore. We cannot exclude that SFX model P could result from non-linear excitation and corresponds to the cationic species. In conclusion, experimental and computational data altogether support the notion that the chromophore twisting observed by TR-SFX (model T) was not caused by ionization, irreversible photobleaching or multi-photon absorption.

The intermediate species and conformations detected by TA spectroscopy can be tentatively connected to those determined by TR-SFX and by simulations. In the refined SFX model T (Fig. 4c, d), the two ring-bridging dihedral angles *tau* and *phi* differ by ~90°, resulting in a perpendicular positioning of the phenol and imidazolinone groups in a twisted chromophore (Supplementary movie). This twisted conformation corresponds to the one obtained by QM/MM and classical MD for the S_1 electronic excited state (Fig. 4b; Supplementary Fig. 9). Thus, SFX model T can be attributed to an S_1 intermediate species (denoted I^*_T in Fig. 3), but cannot correspond to either of the excited state species identified by TA spectroscopy (I^*_1, I^*_2 in Fig. 3), as the calculated transition dipole moment for model T is almost zero (not shown). A similar chromophore twisting in the excited state has been reported recently for PYP[21]. The chromophore conformation in the SFX model P (Fig. 4c, d) is similar to the minor conformation that is accessed in less than 30 fs in classical MD simulations of S_1 (Fig. 4b) and could correspond to one of the two TA spectroscopy species (I^*_1, I^*_2) characterized by a broad stimulated emission band at 530 nm (Fig. 2a). QM/MM calculations of the species with the highest transition dipole moment lead to a conformation (red star in Fig. 4b) close to SFX model P.

We present a tentative model for *off*-to-*on* photoswitching of rsEGFP2 (Fig. 3). Excitation of the trans protonated *off* state leads to a Franck-Condon state that relaxes in 90 fs to two excited state species I^*_1, I^*_2 as seen by TA spectroscopy, one of which corresponds to the nearly planar conformation of model P detected in TR-SFX at 1 ps. I^*_1, I^*_2 decay in 0.9 and 3.7 ps, either backward to the *trans* protonated *off* state or forward to a *cis* protonated state as evidenced by TR-SFX at 3 ps. We anticipate that the *cis* protonated state becomes deprotonated on a timescale beyond that of our TA spectroscopy experiments. The spectroscopically invisible twisted intermediate I^*_T observed in TR-SFX, QM/MM and MD simulations is close to the conical intersection (CI) and thus still electronically excited. The scheme illustrates the complementarity of TA spectroscopy and TR-SFX, which are sensitive to electronic and structural changes, respectively.

Chromophore twisting during the transition from the *off*-state to SFX model T involves changes in *tau* and *phi* of -84° and + 43°, respectively. This supports that *trans*-to-*cis* isomerization during *off*-to-*on* photoswitching in rsEGFP2 proceeds *via* a hula-twist motion[27], in line with the mechanism proposed for asFP595[13]. Theoretical studies suggested that upon excitation the protonated HBI chromophore preferentially rotates around *tau*, which would facilitate isomerization[8]. While our data confirm that chromophore twisting involves a large change in *tau*, they also indicate that in rsEGFP2, the protein matrix tunes the relative amount of *tau* and *phi* rotations to minimize the energetic cost of *trans*-to-*cis* isomerisaton through the excited-state potential energy surface.

Protein structural changes accompanying chromophore twisting

Formation of the fully twisted intermediate state I^*_T involves substantial structural changes for residues of the chromophore binding pocket (Supplementary Fig. 23). Three residues stand out, *viz.* Arg97, which is hydrogen-bonded to the chromophore imidazolinone ring in ground state structures (*off* and *on*) but not in the twisted intermediate state, and Val151 and Thr204, which surround the phenol group (Fig. 4c, d). Most strikingly, Thr204 accompanies the isomerization of the chromophore, displaying different rotamers in the *on* state*,* the *off* state, and the twisted intermediate state (Supplementary Fig. 23). In the twisted intermediate the chromophore phenol group and the Val151 side chain are in close proximity (3.1 Å between the OH and CG1 atoms of the former and the latter, respectively). In order to avoid a steric clash of its phenol group with the side chain of Val151, the chromophore imidazolinone ring tilts by 18° in the opposite direction (positive and negative peaks on the chromophore O2 atom in Fig. 4d), which likely causes the observed motion of the α-helix N-terminal to the chromophore downwards along the barrel axis (Fig. 5), and away from beta-strands 7 (residues 146-152) and 10 (residues 200-209) (Supplementary Fig. 27). Chromophore twisting thus seems to be constrained at the phenol, leading to a downward movement of the α-helix carrying the chromophore base (Fig. 5).

Rational protein design based on excited-state structures

The ability of the protein scaffold to accommodate a fully twisted chromophore is likely a key determinant for enabling photoswitching in RSFPs, as compared to non-photoswitchable fluorescent proteins in which full twisting of the confined chromophore is believed to be strongly restrained. The close proximity of Val151 to the twisted chromophore in the excited state (see above; Supplementary movie) suggests that photoswitching of rsEGFP2 could be facilitated by reducing the side chain at position 151 by a mutation to alanine. Indeed, spectroscopic characterization revealed that the V151A variant has an *off*-to-*on* (*on*-to-*off*) photoswitching quantum yield of 0.77 (0.064) compared to 0.40 (0.043) for the WT protein (Supplementary Fig. 30 and Table 7). Efforts to engineer fluorescent proteins have been based so far on ground state structures only[28,29]. The present work provides the first example of rationally improving a photoswitching property based on an excited-state structure. Sub-picosecond time-resolved serial crystallography at XFEL sources is expected to provide novel and comprehensive insight into excited-state intermediates in a variety of other fluorescent proteins, with hereby demonstrated potential for the rational design of enhanced variants.

Methods:

Protein expression, purification and crystallization

The protein was expressed and purified essentially as described[23]. Microcrystals were obtained by seeding, starting from previously established conditions[23]. For details see Supplementary Information.

Ultrafast transient absorption spectroscopy

Ultrafast transient absorption experiments were carried out on rsEGFP2 in 50 mM Hepes, 50 mM NaCl, pH 8, using a similar method as described elsewhere[18]. Briefly, a 1-kHz femtosecond transient absorption set up was used and the pump beam (400 nm, 100 fs pulse length) was obtained by doubling 800 nm fundamental pulses. rsEGFP2 in the *off* state, with an absorbance of about 0.1 at 400 nm (chamber thickness 0.25 mm), was continuously irradiated within a flow cell by 470 nm light to photoswitch the protein to the *off* state after excitation by the pump-laser beam. The instrument response function (110 fs FWHM) was estimated from the stimulated Raman amplification signal (Supplementary Fig. 1a) and spectra were corrected from group velocity dispersion. To avoid multiphoton processes the data shown in Fig. 2 were collected at a low excitation energy density (1.3 mJ cm^{-2}, corresponding to a power density of 13 GW cm^{-2}). In order to address the question if the pump-laser energy density employed in the TR-SFX experiments (see below) leads to either the formation of ionized chromophore species or irreversible formation of photobleached species, time-resolved and steady state absorption spectra[30] of rsEGFP2 in solution were recorded after femtosecond laser excitation at various pump-laser energy densities, respectively (see Supplementary Information).

Pre-illumination of microcrystals, SFX data collection and processing and structure refinement

rsEGFP2 microcrystals (Supplementary Fig. 12) were photoswitched from the *on* to the *off* state within an inline pre-illumination device[22] (cw laser (488 nm); Supplementary Fig. 14) prior to injection with a liquid microjet into the microfocus chamber of the Coherent X-ray Imaging (CXI) Instrument[31] at the Linac Coherent Light Source (LCLS). An optical pump X-ray probe scheme was used for data collection. The crystalline protein in its *off* state was photoexcited (400 nm, 230 fs pulse length, 940 μJ/mm^2, 400 GW/cm^2) 1 ps and 3 ps prior to interaction with the FEL beam (9.5 keV, 35 fs pulse length), yielding SFX data sets with 64,620 (laser-on-Δ1ps) and 22,259 (laser-on-Δ3ps) indexed diffraction patterns, respectively. A data set (laser-off; 65,097 indexed patterns*)* was collected by interleaving images without pump laser excitation (but with pre-illumination) with the 1 ps time-delay light data. A reference data set was collected of rsEGFP2 in the *on* state (34,715 indexed patterns), from crystals that were neither pre-illuminated (488 nm) nor photo-excited with pump-laser illumination (400 nm). NanoPeakCell[32] and CASS[33,34] were used to identify crystal hits, find Bragg peaks and sort laser-on and laser-off images. Diffraction patterns were indexed and integrated with *CrystFEL*[35] (Supplementary Table 5). Room temperature static structures of rsEGFP2 in the *off* state (PDB accession code 5O8A) and the *on* state (wwPDB accession code 5O89) were determined by molecular replacement using phases from the respective cryo-structures determined by synchrotron methods (pdb ID codes 5DTY and 5DTX, respectively[23]). Details on the partial Q-weighted difference refinement procedure[24,25,] used to determine models T and P from the laser-on-Δ1ps data set are described in the Supplementary Information. Extrapolated structure factors (defined by eq. 1 in the Supplementary Information) and the corresponding difference-refined excited-state structure of rsEGFP2 have been deposited in the wwPDB under accession code 5O8B. The experimental laser-on-Δ1ps dataset and the corresponding composite structure of rsEGFP2 (83% *off* state, 10% *on* state, 4% model P, 3% model T) have been deposited in the wwPDB under accession code 5O8C.

Excited state QM/MM simulations

Excited-state simulations were performed using fully solvated models of rsEGFP2 derived from the SFX structure of the *off* state determined in this work. Each model had approximately 34,000 atoms. Excitation was simulated by carrying out classical MD simulations with the neutral chromophore in its S_0 ground state, and then spawning trajectories with the chromophore switched to its S_1 excited state. A QM/MM method was used to describe the chromophore's potential energy surface, with the chromophore (45 atoms) in the QM region, and the remaining atoms of the protein and solvent in the MM region. Overall, 100 and 50 excited state trajectories were generated and analyzed in the QM/MM and MM simulations, respectively. These excited state simulations were complemented by other techniques, including density functional theory and pKa calculations. Full details can be found in the Supplementary Information.

Excited-state classical MD simulations

Classical molecular simulations were performed using specifically designed force fields describing the Potential Energy Surfaces of the ground state and of the first excited singlet-state of the chromophore, accounting for the relation between the two torsion angles *tau* and *phi*[36] in the excited state (see Supplementary Information). All amino acids were taken in their standard protonation state at pH 7, except Glu223 which was protonated to allow hydrogen bonding with the N2 nitrogen atom of the chromophore. The protein was neutralized with Na^+ cations and solvated in a water box. Starting from the SFX structure of the *off* state, a simulation was run for the S_0 state at constant temperature (300 K) and pressure (1 atm). Fifty simulations of the S_1 state of 5 ps duration each were carried out starting from snapshots extracted from the S_0 state simulation, and using a custom-made specific version[36] of the PMEMD module of AMBER[37].

Code availability

The custom-written CNS script used to calculate Q-weighted structure factor amplitude difference Fourier maps is available from Jacques-Philippe Colletier (colletier@ibs.fr) on request.

Data availability

All indexed raw images for the *on*-state reference (doi://**XXXXX**), laser-off (doi://**XXXXX**), laser-on-Δ1ps (doi://**XXXXX**) and laser-on-Δ3ps (doi://**XXXXX**) datasets have been deposited in the Coherent X-ray Imaging Data Bank (CXIDB; http://www.cxidb.org/).

Author contributions: M.W., D.B., J.P.C., and I.S. designed research. S.J. provided rsEGFP2 plasmid. J.W., V.A., V.G., M.T. and F.F. expressed and purified WT protein. V.A. cloned, expressed, purified the protein variant and V.A. and D.B. analyzed it. J.W., V.G. and I.S. carried out microcrystallization. G.S. and M.B. spectroscopically characterized microcrystals. R.L.S and G.S. designed and optimized pre-illumination device. M.Fe. and M.Fi. carried out excited-state QM/MM simulations. J.R., M.L. and I.D. carried out excited-state classical MD simulations. M.S. and C.R. carried out time-resolved absorption spectroscopy. M.C. and G.S. carried out steady state absorption spectroscopy with femtosecond excitation. N.C, M.S., J.W., V.A., A.A., T.R.M.B., S.B., S.C., E.DlM, R.B.D., L.F., M.Hi., M.Hu., J.E.K., G.K., T.J.L., M.L., K.N., J.S.R., C.M.R., M.S., M.C., R.L.S., J.P.C., I.S. and M.W. carried out SFX experiments. R.L.S, R.B.D. and G.K. performed sample injection. J.S.R. and M.C. aligned and controlled the pump laser. A.A., S.B., M.Hu., J.E.K. and M.L. prepared and performed data collection. C.M.R., M.Hi., L.F., N.C., T.R.M.B. and J.P.C. carried out online SFX data processing. N.C. and J.P.C. carried out offline SFX data processing and structure refinement, with input from T.R.M.B., K.N., I.S., D.B. and M.W. J.P.C., G.S., R.L.S., J.W., M.S., M.C., M.Fi., J.R., B.L., wrote the supplementary information with input from M.W., N.C., T.R.M.B., I.S. and D.B. M.W., D.B., I.S. wrote the manuscript with input from T.R.M.B., M.Fi., J.P.C., M.S. and I.D.

Acknowledgments: The study was supported by a grant from the CNRS (PEPS SASLELX) to MW, and an ANR grant to MW, MC, MSl (ANR-15-CE32-0004 BioXFEL). Use of the Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract no. DE-AC02-76SF00515. We acknowledge support from the Max Planck Society. This work used the Multistep Protein Purification Platform (MP3) of the Grenoble Instruct center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). The Chevreul Institute (FR 2638), the Ministère de l'Enseignement Supérieur et de la Recherche, the Région Nord-Pas de Calais and FEDER are acknowledged for financial support. GK acknowledges support from the E.U. under the program FP7-PEOPLE-2011-ITN NanoMem (project number 317079). JW acknowledges support from the CEA through the IRTELIS PhD program. Correspondence and requests for materials should be addressed to Martin Weik, Ilme Schlichting or Jacques-Philippe Colletier.

Figures and tables

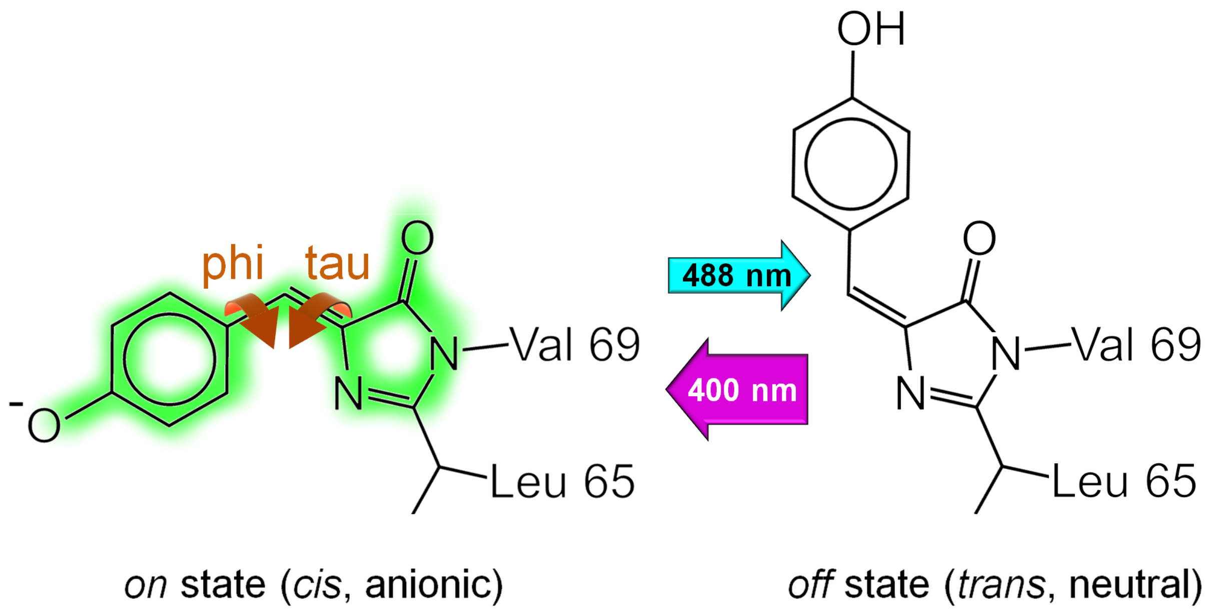


Figure 1: Chromophore isomers and protonation states of rsEGFP2 in the fluorescent *on* (*cis* / anionic; left) and the non-fluorescent *off* (*trans* / neutral; right) states. The two dihedral angles *tau* and *phi* bridging the phenol and imidazolinone rings are indicated by curved arrows (for atom numbering see Supplementary Fig. 16c). *On*-to-*off* and *off*-to-*on* photoswitching is initiated by 488 and 400 nm light, respectively. The size of the colored arrows visualizes the difference in switching quantum yields, which is an order of magnitude higher for the *off*-to-*on* photoswitching than for the reverse direction (supplementary table 7).

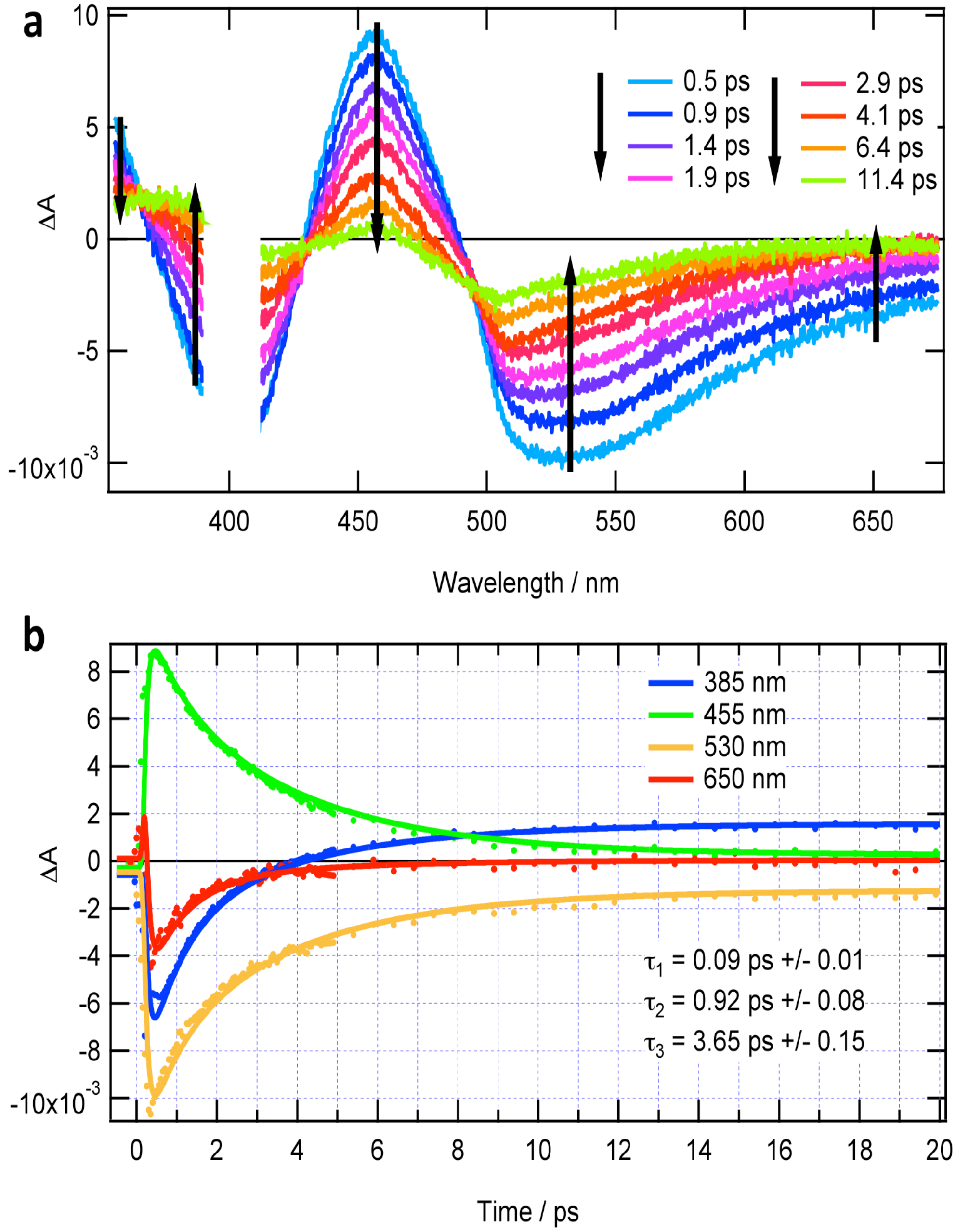


Figure 2: Femtosecond transient absorption spectroscopy of rsEGFP2 in solution. (a) Femtosecond transient difference absorption spectra recorded at different time delays following a femtosecond laser excitation (400 nm) starting from the *trans* neutral *off* state. From these, the spectrum without laser excitation was subtracted to calculate the difference spectrum. (b) Kinetic traces at different wavelengths and the respective fit based on a global analysis with three exponential functions and a constant convoluted with a Gaussian-shaped pulse of 110 fs (FWHM). In these experiments, the pump laser energy density was low enough (1.3 mJ cm^{-2}) to avoid multi-photon processes.

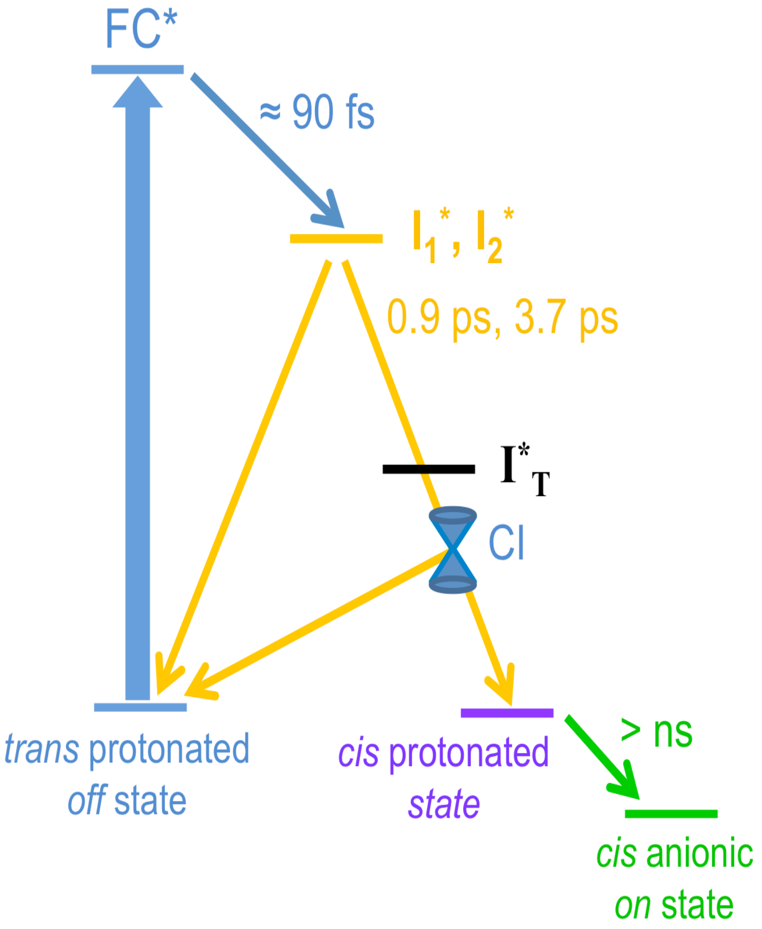


Figure 3: Model for the *off*-to-*on* photoswitching of rsEGFP2. I^*_1, I^*_2 are two excited state species identified by TA spectroscopy. I^*_T corresponds to the twisted intermediate (model T) determined in the 1 ps time-delay SFX data and not detected by TA spectroscopy. CI: conical intersection. FC*: Franck-Condon state. Sub-ps, ps and ns-ms processes are indicated by blue, yellow and green arrows, respectively

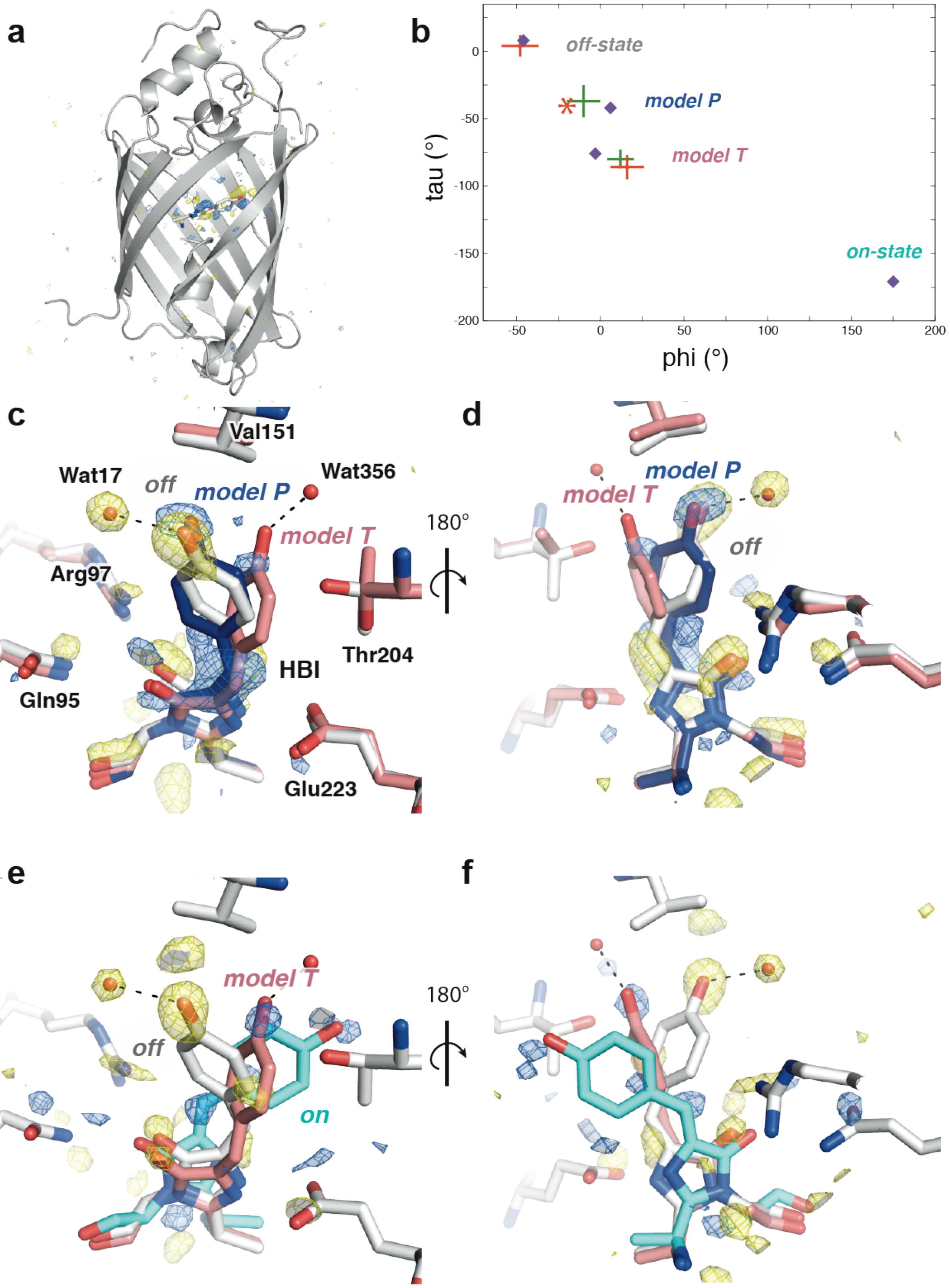


Figure 4: rsEGFP2 in the electronic excited S_1 state. Difference Fourier electron density map $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$, computed between SFX data collected 1 ps after pump laser excitation and without excitation, respectively, contoured at - 3.5σ (yellow) and + 3.5σ (blue) shown for

the entire protein (a) and zoomed into the chromophore region (c, d). In (a) one strand of the β barrel has been omitted for clarity. In (c, d) the chromophore models refined from the laser-off (grey) and laser-on-Δ1ps (model T in pink, and model P in blue) datasets are shown, as well as residues that change conformation upon formation of model T. In the intermediate state model T, the chromophore is twisted, with the phenol and the imidazolinone ring perpendicular to each other. (e, f) Difference Fourier map $F_{obs}^{laser\text{-}on\text{-}\Delta 3ps}$-$F_{obs}^{laser\text{-}off}$, computed between SFX data collected 3 ps after pump laser excitation and without excitation, respectively, contoured at - 3.0σ (yellow) and + 3.0σ (blue). No model was produced from the laser-on-Δ3ps data set. In (e, f) models for the *off* (grey), the *on* (cyan) and the twisted chromophore (pink; determined from the laser-on-Δ1ps data set) are shown. All difference Fourier maps were Q-weighted[38]. (b) Ring-bridging *tau* and *phi* values for the crystallographically determined chromophore models (blue diamonds) and those determined by excited state QM/MM simulations (red crosses) and excited-state MD simulations (green crosses). The size of the crosses corresponds to one standard deviation. The red star corresponds to the conformation with the largest transition dipole moment as calculated by QM/MM.

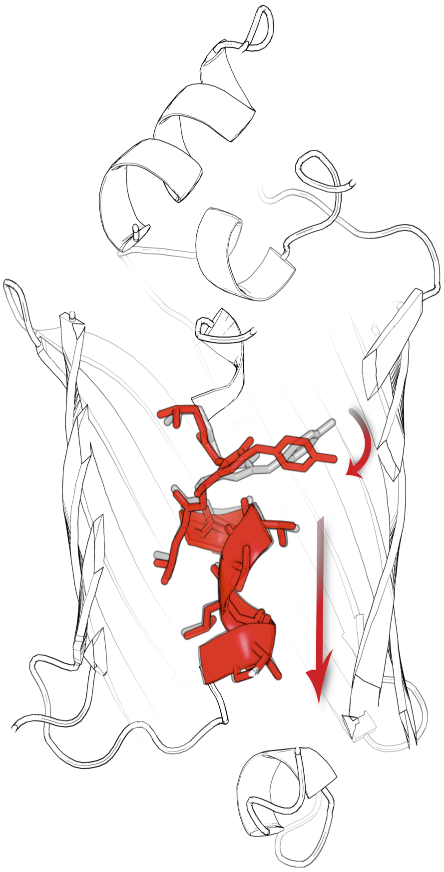


Figure 5: Chromophore twisting is accompanied by a downward movement of the central α-helix along the barrel axis (*off* state in grey, twisted intermediate in red). The corresponding difference electron density is shown in Supplementary Fig. 24.

References

1. Dean, K.M. & Palmer, A.E. Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nat Chem Biol* **10**, 512-23 (2014).
2. Weber, W., Helms, V., McCammon, J.A. & Langhoff, P.W. Shedding light on the dark and weakly fluorescent states of green fluorescent proteins. *Proc Natl Acad Sci U S A* **96**, 6177-82 (1999).
3. Maddalo, S.L. & Zimmer, M. The role of the protein matrix in green fluorescent protein fluorescence. *Photochem Photobiol* **82**, 367-72 (2006).
4. Tsien, R.Y. The green fluorescent protein. *Annual Review of Biochemistry* **67**, 509-544 (1998).
5. Voityuk, A.A., Michel-Beyerle, M.-E. & Rösch, N. Structure and rotation barriers for ground and excited states of the isolated chromophore of the green fluorescent protein. *Chem Phys Lett* **296**, 269-276 (1998).

6. Fang, C., Frontiera, R.R., Tran, R. & Mathies, R.A. Mapping GFP structure evolution during proton transfer with femtosecond Raman spectroscopy. *Nature* **462**, 200-4 (2009).
7. Meech, S.R. Excited state reactions in fluorescent proteins. *Chem Soc Rev* **38**, 2922-34 (2009).
8. Olsen, S., Lamothe, K. & Martinez, T.J. Protonic gating of excited-state twisting and charge localization in GFP chromophores: a mechanistic hypothesis for reversible photoswitching. *J Am Chem Soc* **132**, 1192-3 (2010).
9. Regis Faro, A. et al. Low-temperature chromophore isomerization reveals the photoswitching mechanism of the fluorescent protein padron. *J Am Chem Soc* **133**, 16362-5. (2011).
10. Warren, M.M. et al. Ground-state proton transfer in the photoswitching reactions of the fluorescent protein Dronpa. *Nat Commun* **4**, 1461 (2013).
11. Adam, V., Berardozzi, R., Byrdin, M. & Bourgeois, D. Phototransformable fluorescent proteins: Future challenges. *Curr Opin Chem Biol* **20C**, 92-102 (2014).
12. Nienhaus, K. & Nienhaus, G.U. Fluorescent proteins for live-cell imaging with super-resolution. *Chem Soc Rev* **43**, 1088-106 (2014).
13. Andresen, M. et al. Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proc Natl Acad Sci U S A* **102**, 13070-4 (2005).
14. Ando, R., Mizuno, H. & Miyawaki, A. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. *Science* **306**, 1370-3 (2004).
15. Andresen, M. et al. Structural basis for reversible photoswitching in Dronpa. *Proc Natl Acad Sci U S A* **104**, 13005-9 (2007).
16. Andresen, M. et al. Photoswitchable fluorescent proteins enable monochromatic multilabel imaging and dual color fluorescence nanoscopy. *Nat Biotechnol* **26**, 1035-40 (2008).
17. Adam, V. et al. Structural characterization of IrisFP, an optical highlighter undergoing multiple photo-induced transformations. *Proc Natl Acad Sci U S A* **105**, 18343-8 (2008).
18. Colletier, J.P. et al. Serial Femtosecond Crystallography and Ultrafast Absorption Spectroscopy of the Photoswitchable Fluorescent Protein IrisFP. *J Phys Chem Lett* (2016).
19. Grotjohann, T. et al. rsEGFP2 enables fast RESOLFT nanoscopy of living cells. *eLife* **1**, e00248 (2012).
20. Barends, T.R. et al. Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* **350**, 445-50 (2015).
21. Pande, K. et al. Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein. *Science* **352**, 725-9 (2016).
22. Schiro, G., Woodhouse, J., Weik, M., Schlichting, I. & Shoeman, R.L. Simple and efficient system for photoconverting light-sensitive proteins in serial crystallography experiments. *Journal of Applied Crystallography* **50**, 932-939 (2017).
23. El Khatib, M., Martins, A., Bourgeois, D., Colletier, J.P. & Adam, V. Rational design of ultrastable and reversibly photoswitchable fluorescent proteins for super-resolution imaging of the bacterial periplasm. *Sci Rep* **6**, 18459 (2016).
24. Terwilliger, T.C. & Berendzen, J. Difference refinement: obtaining differences between two related structures. *Acta Crystallogr D Biol Crystallogr.* **51**, 609-18. (1995).
25. Colletier, J.-P. et al. Use of a `caged' analogue to study the traffic of choline within acetylcholinesterase by kinetic crystallography. *Acta Crystallographica Section D* **63**, 1115-1128 (2007).

26. Schafer, L.V., Groenhof, G., Boggio-Pasqua, M., Robb, M.A. & Grubmuller, H. Chromophore protonation state controls photoswitching of the fluoroprotein asFP595. *PLoS Comput Biol* **4**, e1000034 (2008).
27. Liu, R.S. Photoisomerization by hula-twist: a fundamental supramolecular photochemical reaction. *Acc Chem Res* **34**, 555-62 (2001).
28. Goedhart, J. et al. Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nat Commun* **3**, 751 (2012).
29. Pandelieva, A.T. et al. Brighter Red Fluorescent Proteins by Rational Design of Triple-Decker Motif. *ACS Chem Biol* (2016).
30. Hutchison, C.D.M. et al. Photocycle populations with femtosecond excitation of crystalline photoactive yellow protein. *Chemical Physics Letters* **654**, 63-71 (2016).
31. Liang, M. et al. The Coherent X-ray Imaging instrument at the Linac Coherent Light Source. *J. Synchrotron Rad.* **22**, 514-519 (2015).
32. Coquelle, N. et al. Raster-scanning serial protein crystallography using micro- and nano-focused synchrotron beams. *Acta Crystallographica Section D Biological Crystallography* **71**, 1184-1196 (2015).
33. Foucar, L. et al. CASS-CFEL-ASG software suite. *Computer Physics Communications* **183**, 2207-2213 (2012).
34. Foucar, L. CFEL-ASG Software Suite (CASS): usage for free-electron laser experiments with biological focus. *Journal of Applied Crystallography* **49**, 1336-1346 (2016).
35. White, T.A. et al. Crystallographic data processing for free-electron laser sources. *Acta Crystallogr D Biol Crystallogr.* **69**, 1231-1240 (2013).
36. Jonasson, G. et al. Excited State Dynamics of the Green Fluorescent Protein on the Nanosecond Time Scale. *J Chem Theory Comput* **7**, 1990-7 (2011).
37. Case, D.A. et al. AMBER 10. 10 edn (University of California, San Francisco, 2008).
38. Ursby, T. & Bourgeois, D. Improved Estimation of Structure-Factor Difference Amplitudesfrom Poorly Accurate Data. *Acta Crystallographica Section A Foundations of Crystallography* **53**, 564-575 (1997).

4. Chapter III: Photoswitching mechanism of a fluorescent protein revealed by time-resolved serial femtosecond crystallography and transient absorption spectroscopy

Authors: Joyce Woodhouse[1], Gabriela Nass-Kovacs[2,#], Nicolas Coquelle[1,3,#], Lucas M. Uriarte[4,#], Virgile Adam[1,#], Thomas R. M. Barends[2], Martin Byrdin[1], Eugenio de la Mora[1], R. Bruce Doak[2], Mikolaj Feliks[5], Martin Field[1,6], Franck Fieschi[1], Virginia Guillon[1], Stefan Jakobs[7], Yasumasa Joti[8], Pauline Macheboeuf[1], Koji Motomura[9], Karol Nass[2], Shigeki Owada[10], Christopher M. Roome[2], Cyril Ruckebusch[4], Giorgio Schirò[1], Robert L. Shoeman[2], Michel Thepaut[1], Tadashi Togashi[8], Kensuke Tono[8], Makina Yabashi[10], Marco Cammarata[11], Lutz Foucar[2,#], Dominique Bourgeois[1], Michel Sliwa[4,*], Jacques-Philippe Colletier[1], Ilme Schlichting[2,*], Martin Weik[1,*]

I was the main contributor to the work presented in this chapter, except to the transient absorption spectroscopy part.

Affiliations:

[1]Institut de Biologie Structurale, Univ. Grenoble Alpes, CEA, CNRS, F-38044 Grenoble, France.
[2]Max-Planck-Institut für medizinische Forschung, Jahnstrasse 29, 69120 Heidelberg, Germany.
[3]Large-Scale Structures Group, Institut Laue Langevin, 71, avenue des Martyrs, 38042 Grenoble cedex 9 - France
[4]Université de Lille, CNRS, UMR 8516, LASIR, Laboratoire de Spectrochimie Infrarouge et Raman, F59 000 Lille, France
[5]Department of Chemistry, University of Southern California, Los Angeles, USA
[6]Laboratoire Chimie et Biologie des Métaux, BIG, CEA-Grenoble, Grenoble, France
[7]Department of NanoBiophotonics, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany
[8]Japan Synchrotron Radiation Research Institute, 1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5198, Japan
[9]Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Sendai 980-8577, Japan
[10]RIKEN SPring-8 Center, Sayo, Japan
[11]Department of Physics, UMR UR1-CNRS 6251, University of Rennes 1, Rennes, France

*Correspondence

#These authors contributed equally to this work.

Abstract:

Reversibly switchable fluorescent proteins (RSFPs) serve as markers in advanced fluorescence imaging. Photoswitching from a non-fluorescent *off*-state to a fluorescent *on*-state involves *trans*-to-*cis* chromophore isomerization proton transfer. Whereas excited-state events on the ps timescale have been structurally characterized, conformational changes on slower timescales remain elusive. Here we studied the *off*-to-*on* photoswitching mechanism in the RSFP rsEGFP2 by a combination of time-resolved serial crystallography at an X-ray free-electron laser and ns-resolved pump-probe UV-visible spectroscopy. Ten ns after photoexcitation, the crystal structure features a chromophore that isomerized from *trans* to *cis* but the surrounding pocket features conformational differences compared to the final *on*-state. Spectroscopy identifies the chromophore in this first ground-state photo-intermediate as being protonated. Deprotonation then occurs on the μs timescale and correlates with a conformational change of the conserved neighbouring histidine. Together with a previous excited-state study, our data allow establishing a detailed mechanism of *off*-to-*on* photoswitching in rsEGFP2.

Phototransformable fluorescent proteins (PTFPs) represent invaluable tools for advanced fluorescence microscopy, by serving as genetically-encoded markers that change emission colour or intensity when irradiated with visible light at specific wavelengths[1,2]. A subgroup of of PTFPs, so-called reversibly switchable fluorescent proteins (RSFPs), are used in ensemble-based nanoscopy approaches such as RESOLFT (reversible saturable optical fluorescence transition[3]) or NL-SIM (nonlinear structured illumination microscopy[4]), where they are repeatedly toggled back and forth between a fluorescent *on*- and a non-fluorescent *off*-state by irradiation with light at two different wavelengths[5]. Depending on whether irradiation at the excitation peak turns the fluorescent state *off* or *on*, RSFP are coined negative or positive switchers, respectively. It is widely accepted that the molecular basis of photoswitching in RSFPs (with the exception of some engineered variants such as Dreiklang[6]) is a combination of isomerization and change in protonation state of the chromophore phenol moiety, as evidenced by X-ray crystallography[7] and absorption and fluorescence spectroscopies[8]. In all negative RSFP, the fluorescent *on*-state chromophore is an anionic *cis*-isomer (*cis*-phenolate) whereas the non-fluorescent *off*-state is a neutral *trans* isomer (*trans*-phenol). The chronological order of isomerization and protonation-changes has been debated, as well as corresponding time-scales on which they occur.

Over the past decade, several spectroscopic investigations focused on Dronpa, a negative RSFP from Anthozoa (e.g. corals)[9]. Due to the thousand-fold higher switching quantum yield (QY) in the *off*-to-*on* (0.37) direction as compared to *on*-to-*off* (3.2×10^{-4})[8], only *off*-to-*on* photoswitching has been studied in practice. The first investigation of Dronpa by ultrafast optical spectroscopy suggested that the deprotonation of the *trans*-phenol occurs in the excited state on the ps timescale via an excited-state proton transfer (ESPT) mechanism[10]. Shortly after, however, time-resolved infra-red (TR-IR) spectroscopy on Dronpa[11] and its fast-switching M159T mutant[12] indicated that isomerization accompanies the excited-state decay. *Trans*-to-*cis* isomerization is thus followed by chromophore deprotonation, which takes place on the µs timescale as a thermally-driven ground-state process. Further evidence that *trans*-to-*cis* isomerization occurs within ps in the excited state was provided by transient absorption spectroscopy, which attributed the first photoproduct to a *cis*-protonated chromophore[13]. However, another TR-IR study on Dronpa-M159T mutant advocated that both isomerization and deprotonation are ground-state processes, attributing the ps spectroscopic changes in the excited state to protein conformational changes priming the chromophore for switching[14]. Thus, the existence and structure of a *cis*-protonated switching intermediate has remained elusive in Dronpa as well as in RSFPs from Hydrozoa (e.g. jellyfish). To address this issue, we studied *off*-to-*on* photoswitching in rsEGFP2, a negative RSFP, by a combination of time-resolved crystallography and transient absorption spectroscopy.

rsEGFP2 is a bright, photostable, monomeric and photoswitchable variant of the enhanced green fluorescent protein (EGFP). This fast RSFP is extensively used for tagging proteins in mammalian cells for live-cell RESOLFT nanoscopy[15]. Its hydroxybenzylidene imidazolinone (HBI) chromophore, formed autocatalytically from the three residues Ala-Tyr-Gly, is carried by an α-helix entrapped within an 11-stranded β-barrel[16] (Fig. 1a). rsEGFP2 remains photoactive in the crystalline state, so that first insights into the structure of its *on*- and *off*-state chromophores could be obtained by conventional crystallography[16]. Optical spectroscopy measurements and crystal structures established that in rsEGFP2, as in other RSFPs such as Dronpa[17], asFP595[7] mTFP0.7[18] or IrisFP[19], the resting *on*-state is characterized by a *cis*-phenolate chromophore, which absorbs at 479 nm and emits at 503 nm (fluorescence QY of 0.35), whereas the *off*-state features a *trans*-phenol chromophore, which absorbs at 403 nm. As for Dronpa, rsEGFP2 is a negative RSFP, meaning that *on*-to-*off* photoswitching competes with fluorescence[16] and is triggered by illumination of the anionic species at 479 nm with a switching $\mathrm{QY}^{\textit{on}\text{-to-}\textit{off}}$ of 1.65×10^{-2}. Back-switching to the *on*-state is triggered by illumination at 405 nm, and characterized by a $\mathrm{QY}^{\textit{off}\text{-to-}\textit{on}}$ of 0.33[16] (Fig. 1b).

The two rsEGFP2 chromophore states that can be accumulated at room temperature and physiological pH were characterized by conventional X-ray crystallography[16]. In order to characterize the room-temperature structure of photoswitching intermediate states that only form transiently, another method is needed that allows collecting structural data on a time scale from a few ps to μs. This methodology is time-resolved serial femtosecond (fs) crystallography (TR-SFX) using X-ray free electron lasers (XFELs). Recently, we combined ultra-fast optical spectroscopy and TR-SFX to study excited-state structures of rsEGFP2 on the ps timescale during *off*-to-*on* switching[20]. Optical spectroscopy showed that the excited state decays on the time scale of a few ps. TR-SFX revealed that 1 ps after photoexcitation at 400 nm, the chromophore adopts a twisted conformation half-way between the *trans* and *cis* isomers, while at 3 ps, the *cis* chromophore of a presumable ground-state intermediate starts to become populated. The quality of the latter data did not, however, allow building a structural model reliably. Consequently, the structure of the first photoproduct during rsEGFP2 *off*-to-*on* photoswitching has remained elusive and the exact mechanism of photoswitching unresolved.

Here, we describe TR-SFX experiments that allowed capturing an *off*-to-*on* switching intermediate, formed 10 ns after photoexcitation of rsEGFP2. The intermediate structure reveals that the initially *trans*-protonated chromophore has transitioned to the *cis* isomer and evidences conformational changes in the main and side chains of residues in the chromophore pocket. Pump-probe UV-visible spectroscopy indicates that at this time, the chromophore is still protonated, as deprotonation only occurs on the µs to ms time scale in solution. Accordingly, structural features revealed by the 10 ns structure are compatible with a *cis* protonated chromophore only. Thus, our data, together with a ps study published earlier[20], establish that *off*-to-*on* photoswitching in rsEGFP2 involves excited-state isomerization on the ps time scale followed by ground state proton transfer on the µs to ms time scale and allow proposing a detailed switching mecanism.

Results

Time-resolved UV-visible transient absorption spectroscopy of rsEGFP2 in H_2O and D_2O solution and of rsEGFP2 microcrystals

Time-resolved UV-visible transient absorption spectra were recorded on the 100-ns – 10-ms timescale after single-shot ns-laser excitation at 410 nm of rsEGFP2 in its *off*-state in solution (50 mM HEPES pH 8, 50 mM NaCl; Fig. 2a). On the examined time-scale, the excited state had decayed[20] so that the photodynamics of rsEGFP2 in the ground state could be monitored. The transient difference absorbance spectrum at 100 ns shows a broad positive band with a maximum at 420 nm (magenta in Fig. 2a). This band evolves within 5 µs to a spectrum (light blue) with two positive maxima at 390 nm and 460 nm. This spectrum is similar to a static difference absorption spectrum between the *cis* protonated form generated at pH 4 and the *trans* protonated form at pH 8 generated by light irradiation (Fig. 2b, light-blue spectrum), suggesting that at 5 µs the chromophore has isomerized, but is still in the protonated state. Within 180 µs, the first peak (390 nm) vanishes, while a negative band appears at 420 nm. The second peak (460 nm) increases and shifts to 480 nm. Subsequently, and on a timescale from 100 µs up to a few ms, the maximum of the negative band shifts from 420 to 400 nm while the positive band at 480 nm further increases in amplitude. The transient difference spectrum obtained after 3 ms resembles the static difference spectrum between the *cis* anionic form and the *trans* protonated form (Fig. 2b, light-green spectrum). Therefore, the positive band (480 nm) in the transient spectra can be assigned to the absorbance of the *cis* anionic form, whereas the negative band (390 nm) relates to the depopulation of the *trans* protonated form.

Global fit analysis of kinetic traces for all wavelengths identified three time constants of 4.4 ± 0.1 µs, 40.5 ± 0.4 µs and 919.4 ± 1.2 µs, respectively. When similar experiments were carried out in D_2O solution (50 mM HEPES pD 8, 50 mM NaCl ; Supplementary Fig. 1), time evolution also required fitting with three exponential functions, yielding time constants of 5.5 ± 0.1 µs, 73.8 ± 0.2 µs and 2198.0 ± 1.1 µs. Thus, the first time constant is similar in H_2O and D_2O, but a significant isotope effect is observed for the two others (k_H/k_D= 1.80 and 2.40, respectively) which can thus be assigned to proton transfer steps. Decay associated spectra (Fig. 2c) show that the 4.4 µs time constant is mainly characterized by a growth of the positive band at 460 nm. The 40.5 µs time constant has some positive and negative contributions characteristic of the band shift observed, while the 919.4-µs time constant is mainly characterized by the respective decay and growth of the 390-nm and 480-nm bands. Averaging the second and third time constant according to their absorbance at 480 nm in the respective decay associated spectrum (Fig. 2c) yields a single characteristic time constant for chromophore deprotonation of 759.1 µs in H_2O and 1914.3 in D_2O.

Nanosecond transient absorption data were also recorded from a suspension of rsEGFP2 microcrystals (in 100 mM HEPES pH 8, 2.5 M ammonium sulphate), using a modified flash-photolysis setup (see methods). Transient difference spectra similar to those measured in solution were obtained, with positive bands first growing at 460 nm and then at 480 nm (Supplementary Fig. 2). A global fit analysis of kinetic traces for all wavelengths yielded time constants of 2.6 ± 0.1 µs, 36.5 ± 0.5 µs and 310 ± 3 µs, *i.e.* significantly smaller than of rsEGFP2 in H_2O solution (50 mM HEPES pH 8, 50 mM NaCl), with a single characteristic time constant for chromophore deprotonation of 188 µs. When adding ammonium sulphate in a solution experiment (50 mM HEPES, 50 mM NaCl, pH 8, 1.25 M ammonium sulphate) time constants of 3.7 ± 0.1 µs, 41.2 ± 0.5 µs and 310.7 ± 6 µs (Supplementary Fig. 3) were found that are similar to those for microcrystals, with single characteristic time constant for deprotonation of 258 µs. The presence of ammonium sulphate in solution and in the crystal thus accelerates proton transfer.

rsEGFP2 structure solved by TR-SFX 10 ns after photoexcitation

In order to structurally characterize intermediates along the *off*-to-*on* photoswitching reaction of rsEGFP2, TR-SFX was carried out at the Spring-8 Angstrom Compact Laser (SACLA[21]) XFEL, according to an optical pump – X-ray probe scheme. A pump-probe delay of 10 ns was specifically chosen in order to increase chances to trap a ground-state intermediate featuring a *cis*-protonated chromophore. Indeed, the 10-ns delay is much shorter than the µs – ms time-scale on which time-resolved UV-visible transient absorption spectroscopy identified chromophore deprotonation to occur (see above), while much longer than the ps timescale on which a twisted-chromophore isomerization intermediate has been characterized in the excited state[20]. Prior to injection, rsEGFP2 microcrystals were photoswitched from their (resting) *on*-state to their *off*-state (lifetime of 100 min[22]) by illumination with 488-nm light, using a dedicated pre-illumination device[22]. Two interleaved data sets were collected, with (*laser_on_Δ10ns*) and without (*laser_off*) activation by the 400-nm pump-laser (power density 1.4 GW/cm^2). From the *laser_off* data set, a *laser_off* structure was determined at 1.7 Å resolution. The initial model consisted in 90% *off*-state and 10% *on*-state (see Method section for details), with the chromophore being in the *trans* and the *cis* conformation, respectively (Fig. 3a). The reason for including the *on*-state model was because of absorption spectroscopy that indicated a residual amount of ~10% of the molecules remained in the *on*-state after pre-illumination[22]. However, in the resulting $F_{obs}^{laser\_off}$-$F_{calc}^{laser\_off}$ map, negative peaks on the *trans* chromophore and on the side chains of Tyr146 and His149 were observed, suggesting that the *off*-state is less than 90% occupied (Fig. 3a). Moreover, a positive peak at 3.6 σ was observed halfway between the *trans* and *cis* conformers of the chromophore, suggesting the presence of a third, hitherto unobserved conformation (Fig. 3a). An ensemble refinement against the *laser_off* dataset was carried out and also suggested the presence of a third chromophore conformation (oval contours in Fig. 3b and c). Ensemble refinement identified this conformation as a *trans* isomer (called *trans2* hereafter), irrespective of whether the starting model was the *cis* chromophore of the *on*-state (Fig. 3b) or the *trans* chromophore (called *trans1* hereafter) of the *off*-state (Fig. 3c). Thus, the chromophore was modelled in a triple conformation, *i.e.* 70% *trans1*, 20% *trans2*, and 10% *cis* and refined against the *laser_off* dataset (Fig. 3d). Ensemble refinement also suggested a third conformation of the His149 side chain (His149-*supp,* cf. circles in fig. 3b and c), in addition to the ones characteristic of the *on-* (His149-*on*) and the *off-* (His149-*off*) states (Fig. 3b, c), respectively. Likewise, three conformations of His149 (His149-*off*, His149-*supp,* His149-*on*) were included in the *laser off*_model at 70, 20 and 10% occupancy, respectively. Inclusion of these new chromophore and His149 conformers cleared all major peaks in the $F_{obs}^{laser\_off}$-$F_{calc}^{laser\_off}$ map (Fig. 3d). Alternate conformations were also included for other residues (positions 146-148 and 150-152) of β-strand-7, who is known to adopt slightly different conformations in the *on*- and the *off*-states at room-temperature[20], and for Thr204. For those residues, however, only a single additional conformer was needed to account for the presence of both the *cis* and the *trans2* isomers of the chromophore, so that occupancies were refined to 70% (corresponding to the *trans1* chromophore conformation) and 30% (*cis* and *trans2*) (see details in the Methods section).

Following pump-laser irradiation, some protein molecules in the crystal are excited and change structure whereas others are not and thus remain in the *laser_off* structure. The structural features that occur within 10 ns after photoexcitation can be disentangled from the resting *laser_off* structure in a q-weighted[23] difference Fourier map. Such a map was calculated at 1.9 Å resolution, using observed structure factor amplitudes of the *laser_on_Δ10ns* and *laser_off* data sets ($F_{obs}^{laser\_on\_\Delta 10ns}$-$F_{obs}^{laser\_off}$) and phases calculated from the *laser_off* model (Fig. 4a). The strongest features are located at the chromophore and its direct environment. Negative (down to -5.1 σ) and positive (up to 5.7 σ) peaks are observed at the positions of the *trans1* and the *cis* chromophore, respectively, in particular on their phenol group and methylene bridge. No negative peak is observed on the *trans2* chromophore. Additionally, negative peaks are present on the *off*-state conformers of Tyr146 (-5.8 σ) and His149 (-5.1 σ on His149-*off* and -3.8 σ on His149-*supp*), and positive ones on their *on*-state conformers (3.9 σ and 3.6 σ, respectively). The *off*-state has thus been depleted and an intermediate state built up with a chromophore conformation similar to the one observed in the *on*-state. For Tyr146 and His149, the negative peak is larger than the positive one (-5.8 /3.9 σ and -5.1 /3.6 σ for Tyr146 and His149 respectively), suggesting that their side chains display some degree of disorder in the 10-ns intermediate state.

A structural model for the rsEGFP2 intermediate structure 10 ns after photoexcitation (*laser_on_Δ10ns* intermediate structure) was determined by difference refinement[24,25] using extrapolated structure factor amplitudes ($F_{extrapolated}^{laser\_on\_\Delta 10ns}$, see eq. 1 in Methods section)[26]. An occupancy of ~ 50% was estimated for the intermediate state in the *laser_on_Δ10ns* dataset (Supplementary Fig. 4, see details in the Methods section). The 50% occupancy exceeds the primary quantum yield of the *off*-to-*on* photoswitching reaction (0.33[16]) because those molecules that after excitation return to the initial ground state within a few ps[20] can be excited again during the 100-ps pump pulse. The $2F_{extrapolated}^{laser\_on\_\Delta 10ns}$ - F_{calc} map (Fig. 4b) and the corresponding simulated annealing composite omit map (Fig. 4c) feature continuous and well-defined electron density in the intermediate state for the entire chromophore and for the side chains of Tyr146 and Val151. In contrast, the side chain of His149 is not well defined, and positive and negative peaks appear in the $F_{extrapolated}^{laser\_on\_\Delta 10ns}$ - F_{calc} map when a single conformer is included (Supplementary Fig. 5a, b). Therefore, we included in the *laser_on_Δ10ns* intermediate structure two alternate conformations for His149 (see the Methods section for alternative models that were assessed), one being similar to the His149-*on* conformer (*on*-like) with an occupancy of 40% and the other similar to the His149-*off* conformer (*off*-like) with an occupancy of 60% (Fig. 4b and Supplementary Fig. 5). Inclusion of a second conformer for His149 resulted in disappearance of the corresponding peaks in the $F_{extrapolated}^{laser\_on\_\Delta 10ns}$ -F_{calc} map. In the refined model of the *laser_on_Δ10ns* intermediate structure (Supplementary Fig. 5d), the δ (ND1) nitrogen atom of the *on*-like His149 conformer is at 4.3 Å from the phenol-oxygen of the chromophore and H-bonded to the main chain carbonyl oxygen of Asn147 (distance: 2.6 Å). As the latter is an obligate H-bond acceptor, this observation strongly suggests that the δ nitrogen (ND1) of the *on*-like His149 conformer is protonated in the *laser_on_Δ10ns* intermediate structure. Fig. 5 displays an overlay of the *laser_on_Δ10ns* intermediate structure and the *on*-state structure (PDB entry code 5O89[20]) and shows that the main difference between these are the conformations of His149 and Glu223. The δ nitrogen atom (ND1) of His149 H-bonds to the deprotonated chromophore phenol oxygen in the *on*-state (Fig. 5c), but not to the protonated chromophore phenol oxygen of the *on*-like His149 conformation adopted in the *laser-on-Δ10ns* intermediate structure (Fig. 5b). The increased distance between the chromophore phenol oxygen and *on*-like His149ND1 may sign for the protonation of the chromophore, resulting in increased van der Waals exclusion volume. The ε (NE2) nitrogen atom of the *off*-like His149 conformer is at 2.7 Å from the phenol-oxygen of the chromophore, indicating that the two atoms are H-bonded (Fig. 5b). Assuming that the protonation state of His149 does not change between the two conformers observed in the *laser_on_Δ10ns* intermediate structure, *i.e.* that the ε (NE2) nitrogen of the *off*-like His149 conformer is unprotonated, the chromophore phenol would be an obligate donor for this H-bond. Thus, the two His149 conformers observed 10 ns after excitation by the pump laser are compatible with a protonated *cis* isomer of the chromophore. As to Glu223, one of its carboxyl oxygens is in H-bonding distance to N2 of the chromophore in the *laser_on_Δ10ns* intermediate (Fig. 5e) but not in the *on*-state structure (Fig. 5f), and is in H-bonding distance to Ser206OG in the *on*-state structure (Fig. 5f) but not in the *laser_on_Δ10ns* intermediate structure (Fig. 5e).

Discussion

The mechanism of *off*-to-*on* photoswitching in rsEGFP2 involves *trans*-to-*cis* isomerization and deprotonation of the chromophore, and it was here investigated by ns-resolved pump-probe UV-visible spectroscopy and TR-SFX. Spectroscopy in solution provides evidence for the existence of three intermediate states decaying with time constants of 4.4, 40.5, and 919.4 µs, respectively. Only the latter two time constants are markedly affected when experiments are carried out in D_2O instead of H_2O solution, indicating the involvement of proton transfer. The data show that chromophore deprotonation takes place after excited-state decay in the ground state, with an overall time constant of 759.1 µs. The crystal structure obtained 10 ns after phototriggering features a *cis* chromophore that is thus still in the protonated state. In an earlier TR-SFX study[20], chromophore twisting was observed in the excited state 1 ps after photoexciting the *off*-state. Also a low-occupancy population of a *cis* isomer was found 3 ps after photoexcitation that indicated chromophore isomerisation occurs in the excited state. Due to insufficient data quality, however, no structural model could be established of the 3-ps intermediate state[20]. The present spectroscopic and structural evidence unequivocally shows that the excited-state chromophore isomerisation in rsEGFP2[20] is followed by proton transfer in the ground-state, as suggested for Dronpa based on time-resolved IR and UV-vis spectroscopy[11-13]. *Off*-to-*on* switching in rsEGFP2 (hydrozoan origin) and in Dronpa (anthozoan origin) thus involves the same sequential order of steps.

The TR-SFX structure determined from data collected 10 ns after triggering photoswitching from rsEGFP2 in its *off*-state differs from the structure of the stable *cis* anionic *on*-state (Fig. 5) and thus corresponds to that of an intermediate state (*laser_on_Δ10ns*). Spectroscopy (see above) provides evidence that the chromophore is still protonated (neutral) so that the 10-ns intermediate corresponds to a *cis* protonated photoproduct. In the 10-ns intermediate structure, the side chain of His149 is present in two alternate conformations (*on*-like and *off*-like, Fig. 4b). The observation that a significant fraction (60%) of His149 is still in an *off*-like conformation when the chromophore has already isomerized suggests that isomerization precedes the conformational change in the His149 side chain. The latter is thus completed on a time-scale longer than 10 ns and could correspond to the 4.4 µs process (no isotope effect) identified by spectroscopy in solution. Notably, the water molecule being hydrogen bonded to the chromophore phenolate in the *on*-state (W356 in Fig. 5a), and absent in the *off*-state (not shown), is present at 10 ns (Fig. 5b). Interestingly, Thr204 adopts the same rotamer in the 10-

ns intermediate and the *on*-state structure (Fig. 5a), whereas it is in different rotamers in 1-ps intermediate and in the *off*-state structure[20].

The hydrogen bonding network in the chromophore pocket and the protonation state of His149 in the *on*-like conformation can be inferred from inspecting the *laser-on-Δ10ns* structure (Fig. 5b). Double (de-)protonation of His149 was not considered because its probability is negligible at the pH value (pH 8) of the crystalline protein. In the *on*-like conformation, His149ND1 is 2.6 Å from the carbonyl oxygen of Asn147 (Fig. 5b). Given this short distance, the two atoms are likely to be H-bonded. Since Asn147O can only serve as an H-bond acceptor, His149 would be ND1 protonated in the *on*-like conformation. The chromophore phenol group, known from spectroscopy to be protonated at 10 ns, engages in two H-bonds, *i.e.* with W356 and with Thr204OG1 (Fig. 5b).

A possible scenario for the transition from the *laser_on_Δ10ns* intermediate to the stable *cis* anionic *on*-state might be as follows. The occupancy of the *off*-like conformation of His149 decreases as that of the *on*-like conformer progressively increases. The chromophore transfers its proton to an unknown partner. A proton transfer from the chromophore to His149 seems unlikely, given that the closest nitrogen atom of the latter (ND1) is most likely already protonated. His149 moves from the *on-like* conformation in the *laser_on_Δ10ns* intermediate closer to the chromophore in the *cis* anionic *on*-state so that its anionic phenolate group now engages in H-bonds with His149ND1 (2.8 Å), Thr204OG1 (2.7 Å) and W356 (2.8 Å). Asn147O is not H-bonded to His149ND1 anymore in the *on*-state, but rather to W356. Changes in the H-bonding network around the chromophore phenol group between the *laser_on_Δ10ns* intermediate and the *cis* anionic *on*-state structure (Fig. 5) are strikingly similar to the ones between the protonated A form and the anionic B form of wild-type GFP[27]. An exception is the H-bond between the chromophore phenol and Thr204OG1 present in the *laser_on_Δ10ns* intermediate structure but absent in the protonated A form of wild-type GFP (Fig. 2a in [27]). The presence of this H-bond in the *laser_on_Δ10ns* intermediate structure might drive chromophore deprotonation by stabilizing the anionic chromophore of the *on*-state. The pathway taken by the proton upon chromphore deprotonation remains unclear. We speculate that the proton might be transferred from the chromophore via Thr204OG1 and His149O to the solvent outside the protein as suggested for GFP[28]. A proton pathway between the solvent and the chromophore has also been suggested to exist in the positive RSFP Padron[29,30].

The *laser_off* structure, determined from SFX data without pump-laser excitation, features the chromophore in two *trans* conformations, one of which (*trans2*, 20% occupancy) was not previously observed in the *off*-state of rsEGFP2. The third His149 conformation (His149-*supp*), included in the *laser_off* structure in addition to the ones in the *on* (His149-*on*) and the *off* (His149-*off*) states (Fig. 3d) might correspond to the *trans2* conformation of the chromophore to which it would be H-bonded (distance of 3.0 Å between phenol oxygen and ND1). We do not yet have an explanation for this structural *off*-state heterogeneity, but note that a similar chromophore conformation has been observed in rsFolder[16], a reversibly switchable fluorescent protein designed for efficient expression in oxidizing cellular environments. There is no negative peak in the $F_{obs}^{laser\_on\_\Delta 10ns}$ - $F_{obs}^{laser\_off}$ difference Fourier map on the phenol group of the *trans*2 chromophore (Fig. 4a). However, a negative peak on His149-*supp* (Fig. 4a) provides circumstantial evidence that the *trans2* chromophore of the *off*-state efficiently reacts to pump-laser activation at 400 nm. A negative peak on the *trans2* chromophore might have been annihilated by the nearby positive peak on the *cis* chromophore (Fig. 4a). We speculate that rsEGFP2 molecules featuring His149 in the *on*-like conformation 10 ns after excitation mostly correspond to chromophores initially in the *trans2* state, while those displaying His149 in the *off*-like conformation would prominently correspond to chromophores initially in the *trans1* state. Indeed, *trans1* chromophores do not directly H-bond to His149 in the *off*-state, so that the motion of this residue may lag behind.

Based on results obtained by time-resolved pump-probe UV-visible spectroscopy and TR-SFX, a model for the rsEGFP2 *off*-to-*on* photoswitching process can be proposed (Fig. 6). Excitation at 400 nm of the *trans* protonated chromophore in the *off*-state triggers excited-state isomerization that passes through a twisted conformation after 1 ps[20]. After excited-state decay, the first photoproduct in the ground state is a *cis* protonated chromophore that we characterized by TR-SFX after 10 ns (magenta bold bar in Fig. 6). Transient absorption spectroscopy in H_2O solution evidences a first ground-state process with a time constant of 4.4 μs (no isotope effect) that might correspond to a protein rearrangement during which the *off*-like conformation of His149 vanishes and the *on*-like conformation becomes fully occupied. Chromophore deprotonation then occurs with a characteristic time constant of 759.1 μs (isotope effect) to yield the final *cis* anionic *on*-state. Concomitantly or subsequently to chromophore deprotonation, His149 moves from the *on*-like conformation to the one adopted in the final *on*-state and establishes the third H-bond with the anionic chromophore.

The present study illustrates the power of using TR-SFX at an XFEL to study and understand light-induced structural changes in photoswitchable fluorescent proteins. Even though structural changes on the ns time-scale could have in principle been studied by Laue crystallography at synchrotron sources, only the intense XFEL radiation provides a useful diffraction signal from microcrystals whose size is limited by the requirement of sufficient optical activation by the pump laser. TR-SFX is anticipated to provide structural insight into photoswitching of other fluorescent proteins that might be of value for rationally tailoring their photophysical properties for nanoscopy[20]. More generally, the number of photosensitive proteins whose molecular workings have successfully been uncovered by TR-SFX[20,31-40] is likely to increase further[41].

Methods

Protein expression, purification, crystallization, and filtering

rsEGFP2[15] was expressed and purified as described earlier[16]. For TR-SFX, rsEGFP2 microcrystals (3 × 3 × 3 μm^3) were generated by seeding (Woodhouse *et al.*, to be described elsewhere) a solution with final protein, precipitant, salt and buffer concentrations of ~20 mg/ml, ~2 M ammonium sulphate, 20 mM NaCl, 120 mM HEPES pH 8, respectively. Prior to injection, sedimented microcrystals were resuspended in 100 mM HEPES pH 8, 2.5 M ammonium sulphate. A suspension containing 2-8% (v/v) microcrystals was filtered through a 20-μm stainless steel filter using a sample loop and a manually-driven syringe.

Transient absorption spectroscopy of rsEGFP2 in solution and in microcrystals

Details of the conventional ns-ms transient absorption setup (90° geometry between pump and probe beams) for solution measurements were published earlier[42]. Briefly, a ns pump laser at 410 nm (8 ns, 3 mJ) was used and focused onto a 1 cm × 1 mm cell containing the protein solution (absorbance of about 1 at 410 nm in a 1 cm path-length). Before pump-laser excitation, rsEGFP2 was switched *off* with 488-nm light (200 mW) from a continuous wave laser diode (Cobolt 06-MLD 488) for 2 minutes. After pump-laser excitation, the time traces where recorded from 340 to 520 nm every 10 nm to reconstruct the transient absorption spectra. Between each single shot excitation rsEGFP2 in solution was converted to its *off* state by visible-light irradiation with a lamp and a long pass filter at 450 nm and a Hamamatsu Xe lamp (8 mW/cm^2) equipped with a fibre and a band pass filter centred at 480 nm.

For spectroscopy experiments in H_2O solution, rsEGFP2 was either in 50 mM HEPES pH 8, 50 mM NaCl or in 50 mM HEPES pH 8, 50 mM NaCl, 1.25 M ammonium sulphate. For spectroscopy experiments in D_2O solution, H_2O / D_2O exchange was performed by several cycles of dilution/concentration. Thirty ml of 50 mM HEPES pD 8, 50 mM NaCl were added to a tube containing 2 ml of rsEGFP2 at 70 mg/ml in 50 mM HEPES pH 7.5, 50 mM NaCl. Then the protein solution was concentrated to a final volume of 3 ml using 10 kDa concentrators (Amicon 15 ml Centrifugal Filters 10 kDa Millipore). The dilution/concentration cycle was repeated three times, so that the total dilution factor was at least 100, *i.e.* the final D_2O content was ~ 99%.

For spectroscopy experiments on microcrystal suspensions, rsEGFP2 microcrystals (3 × 3 × 3 μm^3) were suspended in 100 mM HEPES pH 8, 2.5 M ammonium sulphate. Suspensions were irradiated for 30 min at 488 nm (Cobolt 06-MLD 488; 200 mW) to achieve *off*-switching. For the subsequent ns – ms transient absorption experiment, the crystals were placed in a cell with a 1-mm spacer between windows. The setup was modified (collinear geometry between pump and probe beams) to minimize light diffusion[34]. The pump laser (410 nm, 5 mJ) was focused using a spherical lens onto the cell containing the colloidal microcrystal suspension. The probe light source was not modified from conventional experiments and the focal spot size of both pump and probe beams was 2 mm. Following single-shot excitation, the time traces were recorded from 350 to 520 nm using 10-nm steps to reconstruct the transient difference absorption spectra. Between each single shot excitation the cell was moved manually to probe a previously unexposed area.

Pre-illumination, crystal-injection, pump-laser excitation, SFX data collection and on-line monitoring

A suspension of rsEGFP2 microcrystals (2 – 8% (v/v) was transferred into a stainless steel syringe that was mounted on an anti-settling device[43] and whose temperature was maintained at 20°C with an adjustable Peltier element. The microcrystal suspension was injected with a gas dynamic virtual nozzle (GDVN[44]) injector into the helium-filled *Diverse Application Platform for hard X-ray Diffraction in SACLA* (DAPHNIS[45]) chamber on beamline BL3-EH4c at SACLA[21] (hutch temperature ca 27 ºC). The GDVN carried a sample capillary of 100 μm inner diameter and operated at flow rates of 30 - 40 µl/min. Both parameters, as well as the low microcrystal concentration, were critical to avoid microcrystal aggregation and thus nozzle-clogging that occurred at lower flow-rates and smaller nozzle diameters. The resulting jet had a diameter of 3-5 μm at a focusing helium pressure of 15 psi.

On their way from the sample syringe to the injector, rsEGFP2 microcrystals were photoswitched from their *on*-state (the resting state) to the *off*-state (see Fig. 1) by irradiation with continuous 488-nm laser-light (cw, 200 mW) within a pre-illumination device[22]. The analysis of absorption spectra (Fig. 8 in[22]) obtained from microcrystals collected at the exit of the device with (200 mW) and without pre-illumination indicated that approximately 90% of microcrystalline rsEGFP2 were switched *off*. The transit time from the pre-illumination device to the interaction zone was less than, or equal to, 1 min. This time interval is short compared to the *off*-to-*on* thermal recovery of microcrystalline rsEGFP2 (100 min[22]) so that 90% of the proteins were in their *off*-state when reaching the interaction zone.

SFX data were collected at 30 Hz using X-rays (nominal photon energy 10 keV, pulse length ≤ 10 fs, 300 μJ per pulse at the sample position, 2×10^{11} photons/pulse) focused to 1.4 μm (horizontal) × 1.6 μm (vertical) (FWHM). Data were recorded on an octal-MPCCD detector with eight sensor modules[46] positioned 52 mm away from the sample. Online monitoring of diffraction data, such as determination of hit-rate and estimation of the fraction of multiple hits, was done with CASS[47].

TR-SFX data were collected using an optical pump – X-ray probe scheme with a pump-probe delay of 10 ns. The *off*-to-*on* transition was triggered by a pump-laser pulse (400 nm, 100 ps pulse length, 180 μm × 180 μm focal spot (FWHM), 46 μJ per pulse, leading to a power density of 1.4 GW/cm^2) generated by a Ti:sapphire laser, aligned perpendicularly to both the liquid jet and the X-ray beam. A sequence of interleaved *laser_off* and *laser_on* diffraction images was collected at 30 Hz.

Off-line hit-finding, sorting of laser-on and laser-off data, data processing

A total of about 609,000 frames were collected for the *laser_on_Δ10ns* and the *laser_off* datasets, respectively. NanoPeakCell[48] was used to perform offline hit-finding and sort frames with and without prior pump-laser excitation into *laser_on_Δ10ns* (13300 hits) and *laser_off* (12487 hits) data sets, respectively. CrystFEL 0.6.2 was used for further data processing (*i.e* indexing and integration), resulting in 9781 and 9997 indexed frames for the *laser_on_Δ10ns* and *laser_off* data sets, respectively. The "ring-nocen" method was used for intensity integration. The sample-detector distance was refined until the distribution of unit-cell parameters was Gaussian[49]. Merging with the *partialator* module of CrystFEL 0.6.2 was optimal when combining scaling and partiality refinement without the *push-res* option as judged by R_{split}, CC* and the Wilson plot. Data collection statistics are given in Supplementary Table 1.

Structure solution and refinement

The structure corresponding to the *laser_off* data set was phased by molecular replacement using *Phaser*[50] with the structure of rsEGFP2 in it *off*-state determined by synchrotron cryo-crystallography (PDB entry 5DTY[16]) as a starting model. Refinement with the *Phenix* suite[51] included positional and isotropic individual B factor refinement in reciprocal space. Model building was carried out using *Coot*[52] and occupancies were set manually. Given that the pre-illumination efficiency was approximately 90%[22], an alternate chromophore conformation corresponding to the *cis* isomer in the *on*-state was included at 10% occupancy. Accordingly, those residues adopting different conformations in the *on*- and the *off*-states, *i.e.* the entire β-strand 7 and Thr204, were also first modelled in double conformations, (90% *off* and 10% *on*). When the remaining 90% were attributed to the *trans* isomer of the *off* state, a negative peak in the $F_{obs}^{laser\_off} - F_{calc}^{laser\_off}$ map indicated that the relative occupancy must be lower than 90% and a positive $F_{obs}^{laser\_off} - F_{calc}^{laser\_off}$ peak halfway between the *cis* and the *trans* (called *trans1* hereafter) isomers indicated the presence of a third chromophore conformer (called *trans2* hereafter) occupied at 20% or less (Fig. 3a). Ensemble refinement against the *laser_off* dataset was carried out with the *Phenix* suite starting from the *on*-state model (*cis* chromophore; Fig. 3b) or the *off*-state model (*trans1* chromophore; Fig. 3c). The values of *pTLS*, *tbath* and *tx* were varied, where *pTLS* is the fraction of the molecules included in the TLS fitting procedure, *tbath* is a parameter that controls the X-ray weight and that is coupled to the thermostat temperature and *tx* is the relaxation time used during the simulation. The tested values were 0.6, 0.8, 0.9 and 1.0 for *pTLS*, 2.5, 5 and 10 K for *tbath*, and 0.35, 0.7 and 1.4 ps for *tx*. R_{free} was plotted as a function of the different sets of parameter values using a custom-made python script. The lowest R_{free} was chosen as a criterion to determine the most appropriate set of values of these three parameters. Ensemble refinement indicated the presence of a third chromophore conformation (*trans2*) and a third rotamer of His149 (His149-*supp*). Consequently, the chromophore and His149 were modeled with three alternate conformations, *i.e. cis*, *trans1*, *trans2* and His149-*on*, His149-*off*, His149-*supp*, at occupancies of 10%, 70% and 20%, respectively (Fig. 3d). Inspecting the $F_{obs}^{laser\_off} - F_{calc}^{laser\_off}$ map indicated that all other residues in β-strand 7 (positions 146-148 and 150-152) and of Thr204 needed to be modelled as occupying the *on*- and *off-state* conformers, at relative occupancies of 30% and 70%, respectively.

Structural changes 10 ns after pump-laser excitation were qualitatively identified in a q-weighted[23] (qW) difference Fourier electron density map ($F_{obs}^{laser\_on\_\Delta 10ns}$-$F_{obs}^{laser\_off}$), calculated with the *laser_off* structure as a phase model using CNS[53] (Fig. 4a). For further analysis, extrapolated structure factors (F_{ext}) were calculated based on the q-weighted structure factor amplitude differences according to the following formula: $F_{extrapolated}^{laser\_on\_\Delta 10ns} = \alpha * q/\langle q\rangle * (F_{obs}^{laser\_on\_\Delta 10ns} - F_{obs}^{laser\_off}) + F_{obs}^{laser\_off}$ (eq. 1), where α is the inverse of the estimated occupancy of the fraction of molecules that changed conformation upon photoexcitation and q is the Bayesian weight retrieved through the q-weighting procedure[23]. To determine α, a procedure reported earlier[20] was applied, *i.e.* $F_{extrapolated}^{laser\_on\_\Delta 10ns}$ were calculated and the ratio of integrated peaks in the $F_{extrapolated}^{laser\_on\_\Delta 10ns}$ – $F_{calc}^{laser\_off}$ map and of integrated peaks in the qW $F_{obs}^{laser\_on\_\Delta 10ns}$ - $F_{obs}^{laser\text{-}off}$ map was plotted as a function of α (Supplementary Fig. 4). Only peaks in the vicinity of the chromophore, Tyr146, His149, Val151 and Thr204 were integrated. The most appropriate value of α is the one at which a plateau is reached, *i*.e. 2 in the present case (Supplementary Fig. 4). An α of 2 corresponds to an occupancy of 50% of the *laser_on_Δ10ns* intermediate structure that exceeds the primary switching quantum yield (0.33) because of multiple consecutive excitations during the 100-ps pump pulse. When F_{ext} were calculated without q-weighting, α was estimated to be 2.5 (40% occupancy), at which similar $2F_{ext} – F_{calc}$ maps were obtained.

The *laser_on_Δ10ns* intermediate structure was refined using the difference refinement procedure, and $2F_{extrapolated}^{laser\text{-}on\text{-}\Delta 10ns}$ – F_{calc} and $F_{extrapolated}^{laser\text{-}on\text{-}\Delta 10ns}$ – F_{calc} electron density maps were used to build the model. The same procedure as for the *laser-off* model (Fig. 4b and Supplementary Fig. 5) was followed. A simulated annealing $2F_{extrapolated}^{laser\_on\_\Delta 10ns}$ - F_{calc} composite omit map was calculated to remove any potential model bias (Fig. 4c and Supplementary Fig. 5d). His149 was modeled either in a conformation (*off*-like) similar to His149-*off* in the *off*-state (Supplementary Fig. 5a), in a conformation (*on*-like) similar to His149-*on* in the *on*-state (Supplementary Fig. 5b) or by a mixture of both (40% and 60%, respectively) (Supplementary Fig. 5c) and coordinates and individual B factors were refined. The various computed maps indicate that the mixture of conformations best fits the data. Asn147, Ser148 and Asn150 were also modelled in alternate conformations allowing to release the strain on the peptide bonds around His149 in the *on*- and *off*-like conformations. Simulated annealing composite omit maps (Fig. 4c and Supplementary Fig. 5d), a polder map with omitted His149 (Supplementary Fig. 5e) and a feature-enhanced map (Supplementary Fig. 5f) were calculated with *Phenix* for validation. Hydrogens were added in the riding positions using *phenix.reduce* in the last step of reciprocal-space refinement, but not kept in the final model. Refinement statistics are given in Supplementary Table 1. Figures were prepared using PyMOL[54].

Figures and tables

Supplementary Table 1: Data collection and refinement statistics

Dataset	laser_off	laser_on_Δ10ns
PDB entry code	6GMJ	6GMK
Pre-illumination (488 nm)	yes	yes
Pump-laser excitation (400 nm)	no	yes
Nominal pump-probe delay	n.a.	10 ns
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters		
a (Å)	51.7	52.0
b (Å)	62.9	62.9
c (Å)	71.7	72.1
Collected frames	~609,000	~609,000
Indexed frames	9,997	9,781
Resolution (Å)	29.46 – 1.7 (1.73 – 1.7)	31.30 – 1.9 (1.93 – 1.9)
R_{split} (%)	21.83 (48.01)	11.67 (41.80)
CC*	0.98 (0.91)	0.99 (0.93)
I / σ(I)	4.74 (2.52)	6.56 (3.03)
Completeness (%)	100 (100)	100 (100)
Multiplicity	76.7 (23.6)	107.34 (13.9)
R_{iso} (with respect to *laser_off* dataset)*	n.a	0.295
Refinement statistics		
Refinement	Classical refinement	Difference-refinement
Resolution (Å)	1.7 (1.73-1.70)	1.9 (1.96-1.90)
Number of reflections	26348 (2736)	18204 (2610)
R_{free}	18.12 (25.45)	23.67 (32.95)
R_{work}	14.91 (22.87)	18.95 (25.81)
Number of protein atoms	2230	2039
Number of water molecules	254	163
B-factor protein (Å^2)	40.6	33.5
r.m.s.d. bond lengths (Å)	0.022	0.006
r.m.s.d. angles (°)	1.999	0.873

Ramachandran favoured	98.51	98.80
Ramachandran allowed	1.49	0.80
Ramachadran outliers	0.00	0.40
Rotamer outliers	0.82	0.88
C-beta outliers	0	0
Clashscore	9.5	5.2

*R_{iso} was calculated using Scaleit up to 1.9 Å resolution.

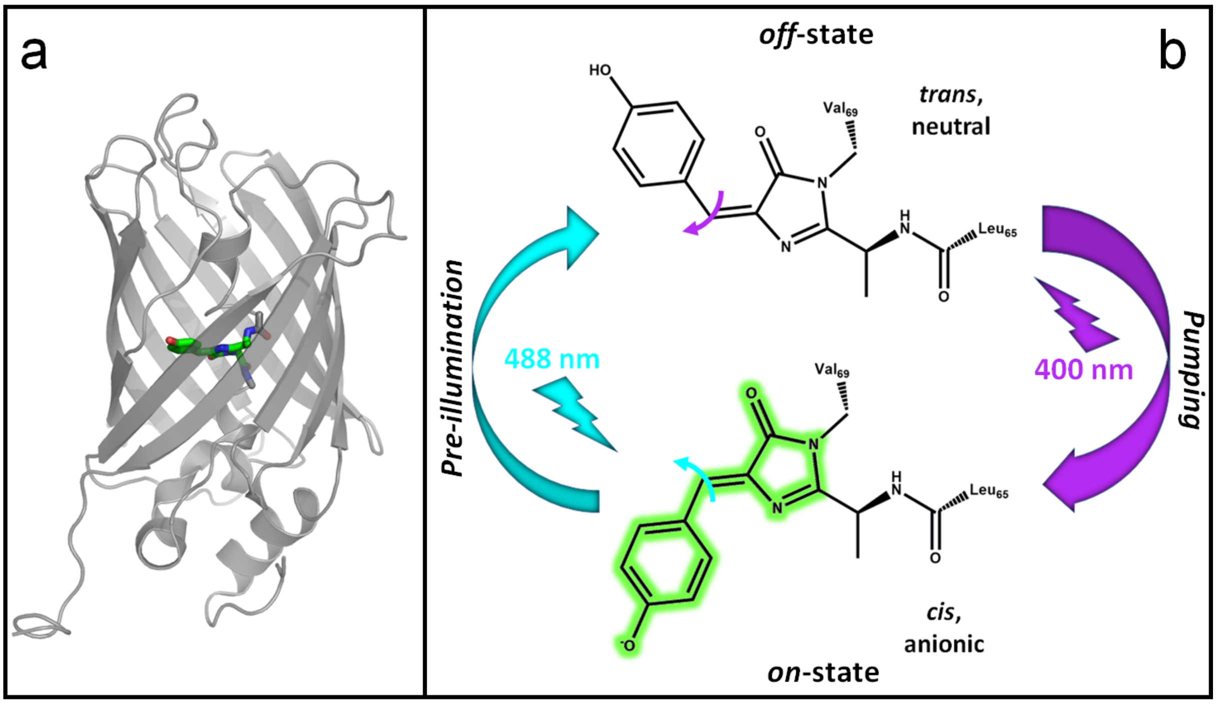


Figure 1: (*a*) Three-dimensional structure of rsEGFP2 showing an 11-stranded β-barrel embedding a chromophore held by a central α-helix. (*b*) Photoswitching scheme. The chromophore of rsEGFP2 in its fluorescent *on* state (bottom) absorbs at 479 nm (laser 488 nm) and emits at 503 nm. The competitive isomerization leads the chromophore to its non-fluorescent *off*-state, which absorbs at 403 nm (laser 400 nm) and ultimately isomerizes back to the initial *on*-state. Flashes and arrows represent the colour of each actinic wavelength. The width of arrows is representative of the efficiency of each isomerization.

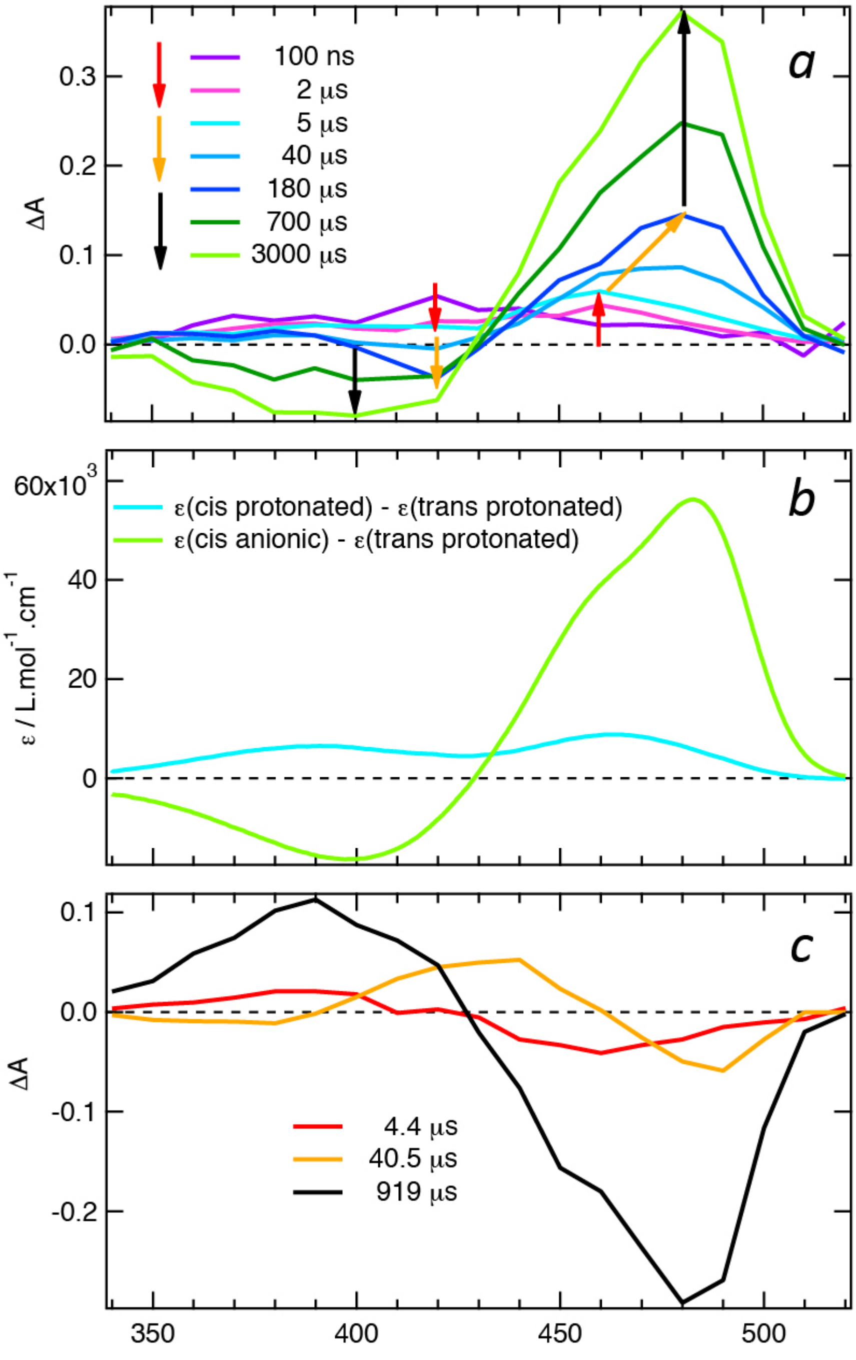


Figure 2: Transient absorption spectroscopy of rsEGFP2 in H_2O solution (50 mM HEPES pH 8, 50 mM NaCl). (*a*) Time-resolved difference absorption spectra recorded after a 410 nm nanosecond excitation of the *trans* protonated *off* state. The spectrum without laser excitation was subtracted to calculate the difference spectrum. The coloured arrows correspond to the three time constants identified by a global fit analysis of kinetic traces for all wavelengths. (*b*) Static difference absorption spectra for *cis* protonated (obtained at pH 4) minus *trans* protonated form (obtained at pH 8). (*c*) Decay associated spectra obtained by fitting the kinetic traces in panel *a* for all wavelengths with three exponential functions.

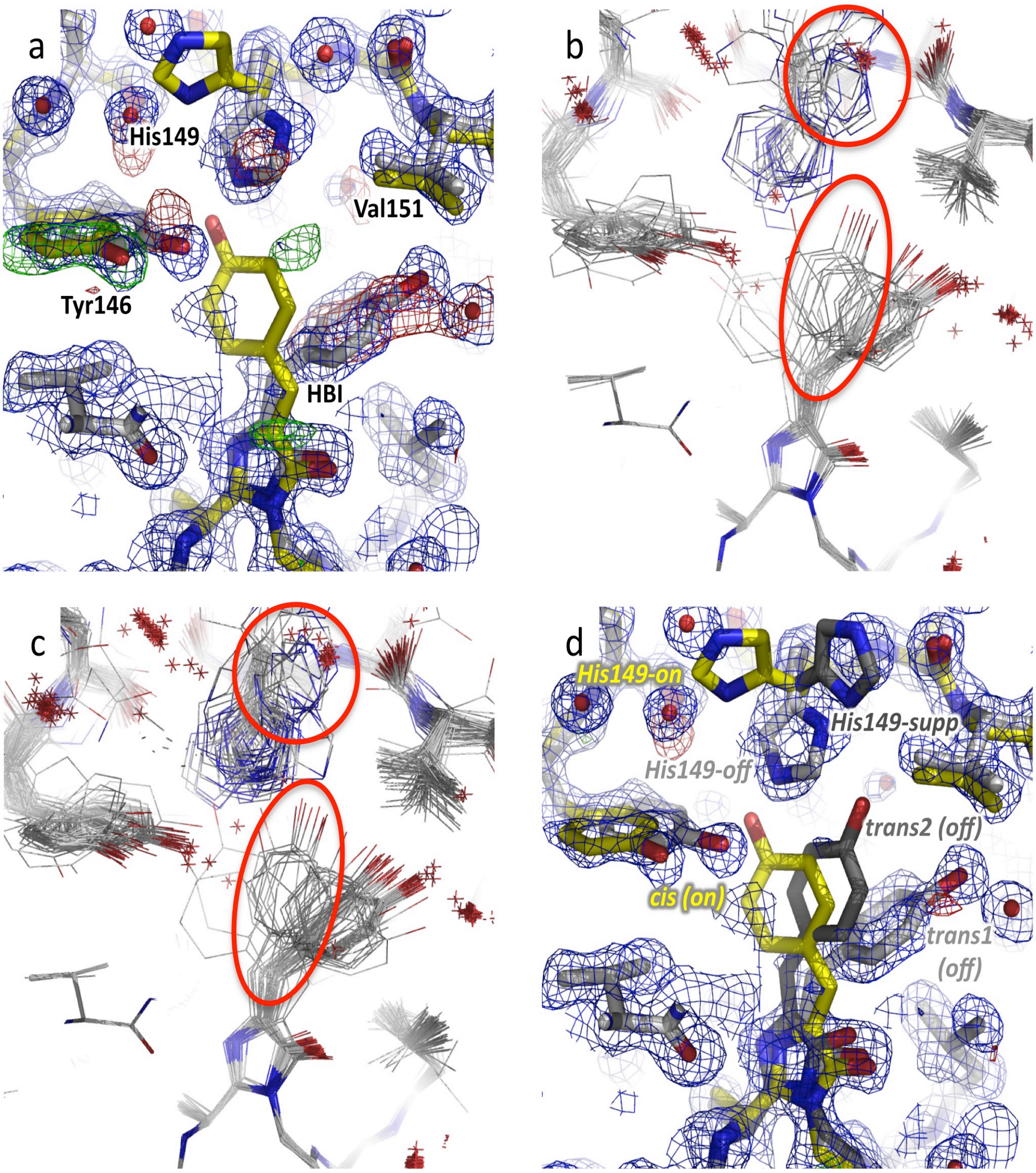


Figure 3: Chromophore (HBI) and its neighbouring residues in rsEGFP2 *laser_off* structure determined by serial femtosecond crystallography without pump-laser activation. (*a*) The initial *laser-off* model consists of a mixture of the *off*-state conformer (light grey carbon trace) and the residual *on*-state conformer (yellow carbon trace) with occupancies of 90% and 10%, respectively. B-sharpened $2F_{obs}^{laser\_off} - F_{calc}^{laser\_off}$ (blue) and $F_{obs}^{laser\_off} - F_{calc}^{laser\_off}$ (green/red) maps at 1.7 Å resolution are displayed at 1σ and $+/- 3\sigma$, respectively. A positive $F_{obs}^{laser\_off}-F_{calc}^{laser\_off}$ peak between the *trans* (grey) and the *cis* (yellow) chromophores suggests the presence of an additional conformer. (*b*) Result of ensemble refinements against the *laser _off* data set starting for the *off*-state model (chromophore 100% *trans1*) or (*c*) the *on*-state model (chromophore 100% *cis*). A third chromophore conformation and a third rotamer of His149 are revealed. (*d*) The final *laser-off* model features triple conformations of His149 and of the chromophore, *i.e*. His149-*off* and *trans1* (light grey), His149-*on* and *cis* (yellow) and the additional His149-*supp* and *trans2* (dark grey) conformations, at 70%, 10% and 20% occupancy, respectively. Alternate conformations of Tyr146 and Val151 are the ones of the *off*- and the *on*-state, at 70% and 30% occupancy, respectively. B-sharpened $2F_{obs}^{laser\_off} - F_{calc}^{laser\_off}$ (blue) and $F_{obs}^{laser\_off}-F_{calc}^{laser\_off}$ (green/red) maps calculated from the *laser-off* data set at 1.7 Å resolution are displayed at 1σ and $+/- 3\sigma$, respectively.

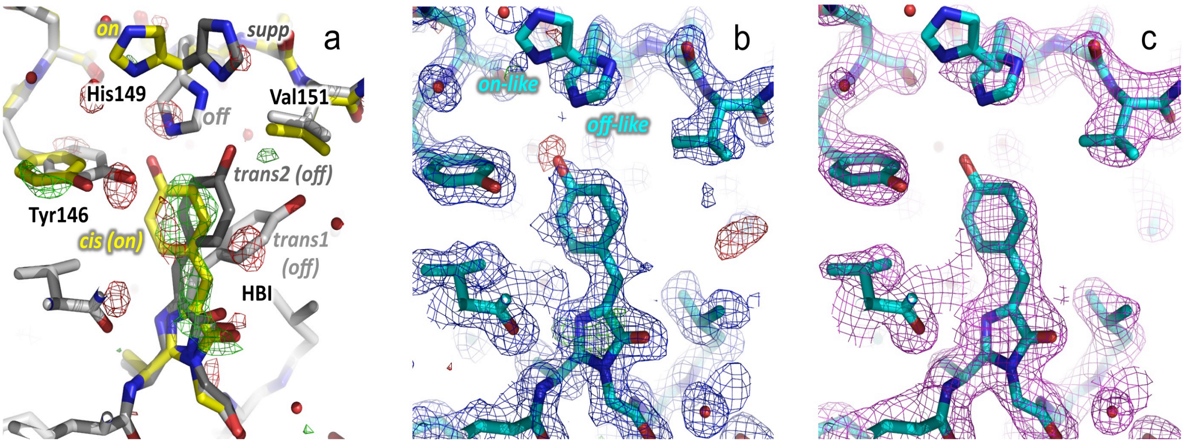


Figure 4: Chromophore (HBI) and its binding pocket 10 ns after pump-laser activation. (*a*) Q-weighted difference electron density map ($F_{obs}^{laser\_on\_\Delta 10ns}$ - $F_{obs}^{laser\_off}$) is contoured at +3σ (green) and -3σ (red) and overlaid onto the model determined from the *laser_off* data set. (*b*) Model of the *laser_on_Δ10ns* intermediate structure (cyan) determined by difference refinement at 1.9 Å resolution. B-sharpened $2F_{extrapolated}^{laser\_on\_\Delta 10ns}$ - F_{calc} (blue, 1σ) and $F_{extrapolated}^{laser\_on\_\Delta 10ns}$ - F_{calc} maps (green/red, +/- 3σ, respectively) are shown. (*c*) Simulated annealing composite omit $2F_{extrapolated}^{laser\_on\_\Delta 10ns}$ - F_{calc} map (magenta, 1σ).

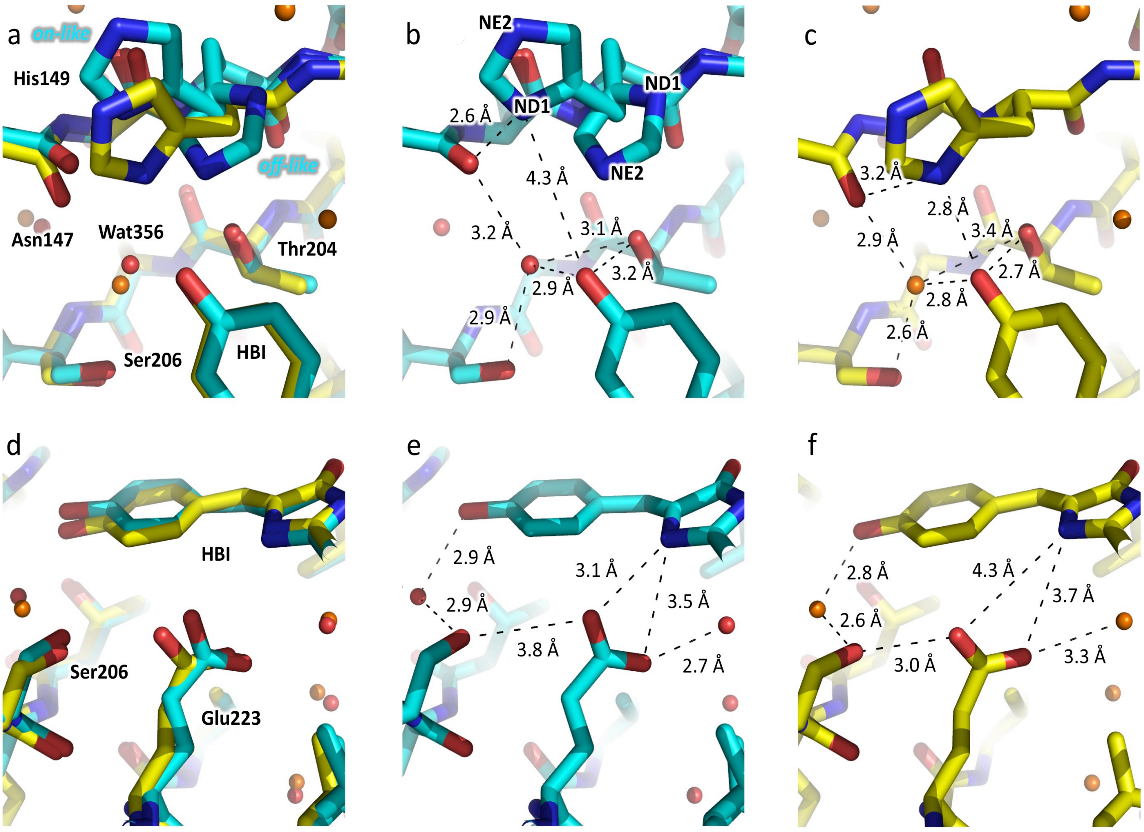


Figure 5: Overlay (*a, d*) of the chromophore region in the 10-ns intermediate structure (cyan, *b, e*) and the *on*–state structure (yellow, *c, f*) determined from SFX at room temperature (PDB entry code 5O89[20]). The focus is on the phenol moiety of the chromophore, His149 and water W356 (*a*, *b, c*) and on the entire chromophore, Ser206 and Glu223 (*d, e, f*). Key distances are indicated.

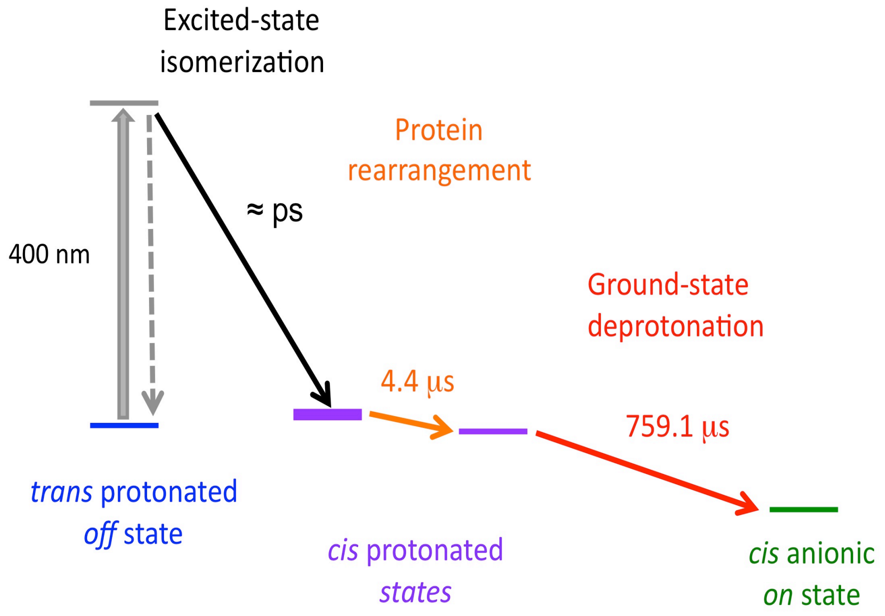
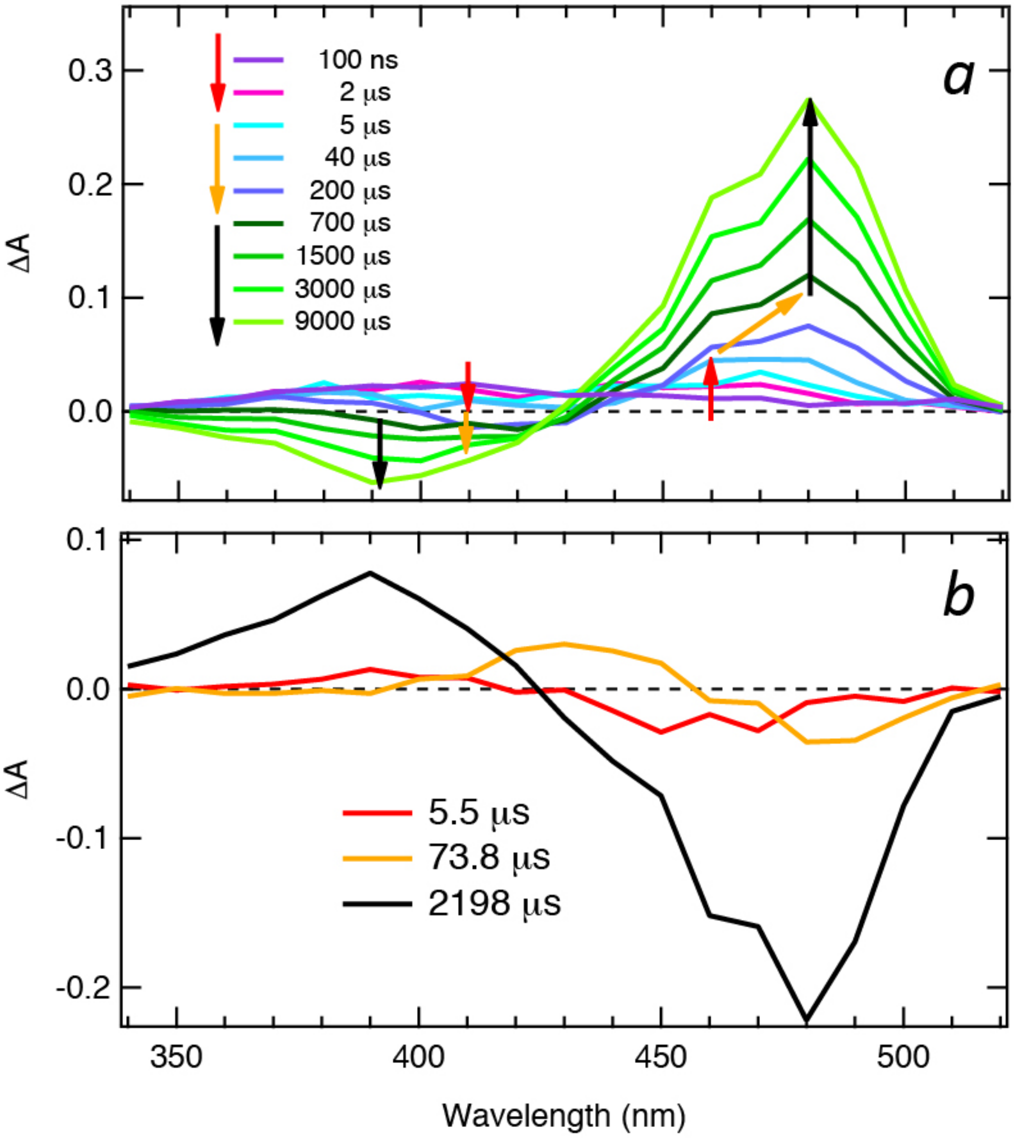
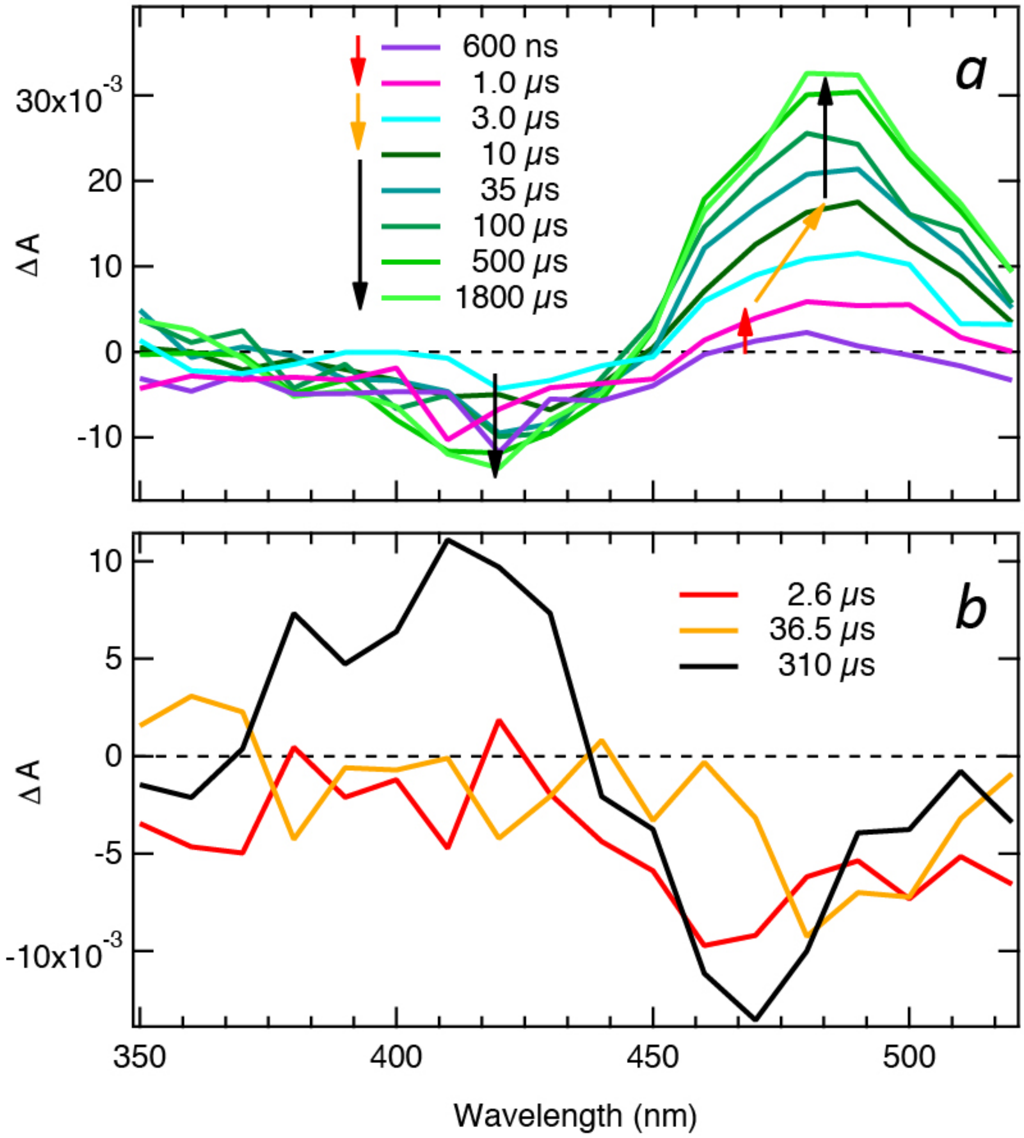


Figure 6: Model for the rsEGFP2 *off*-to-*on* photoswitching process. The bold magenta bar represents the *laser_on_Δ10ns* intermediate structure determined by TR-SFX. Time constants correspond to those determined by nanosecond-resolved pump-probe UV-visible absorption spectroscopy in H_2O solution (50 mM HEPES pH 8, 50 mM NaCl).

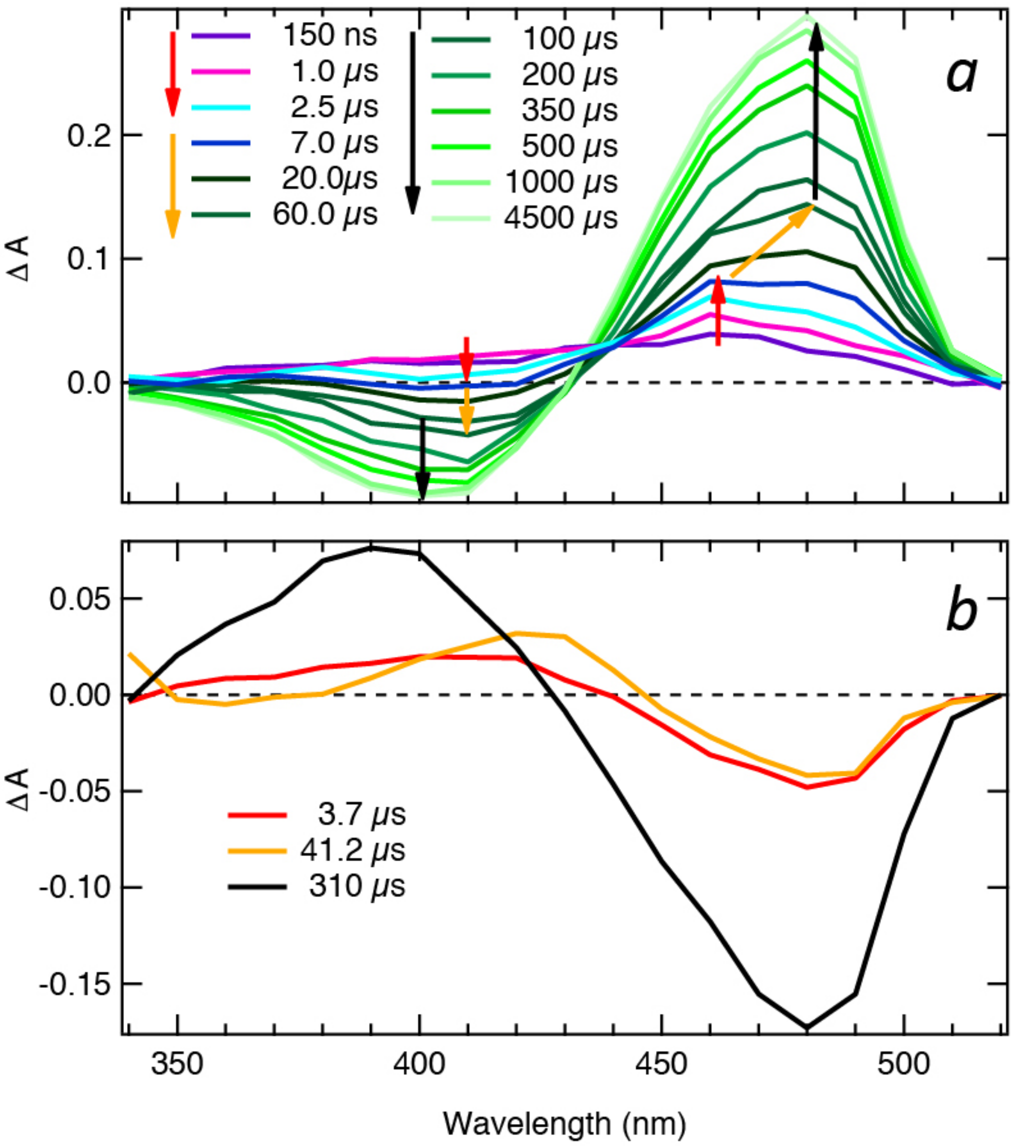
Supplementary figures



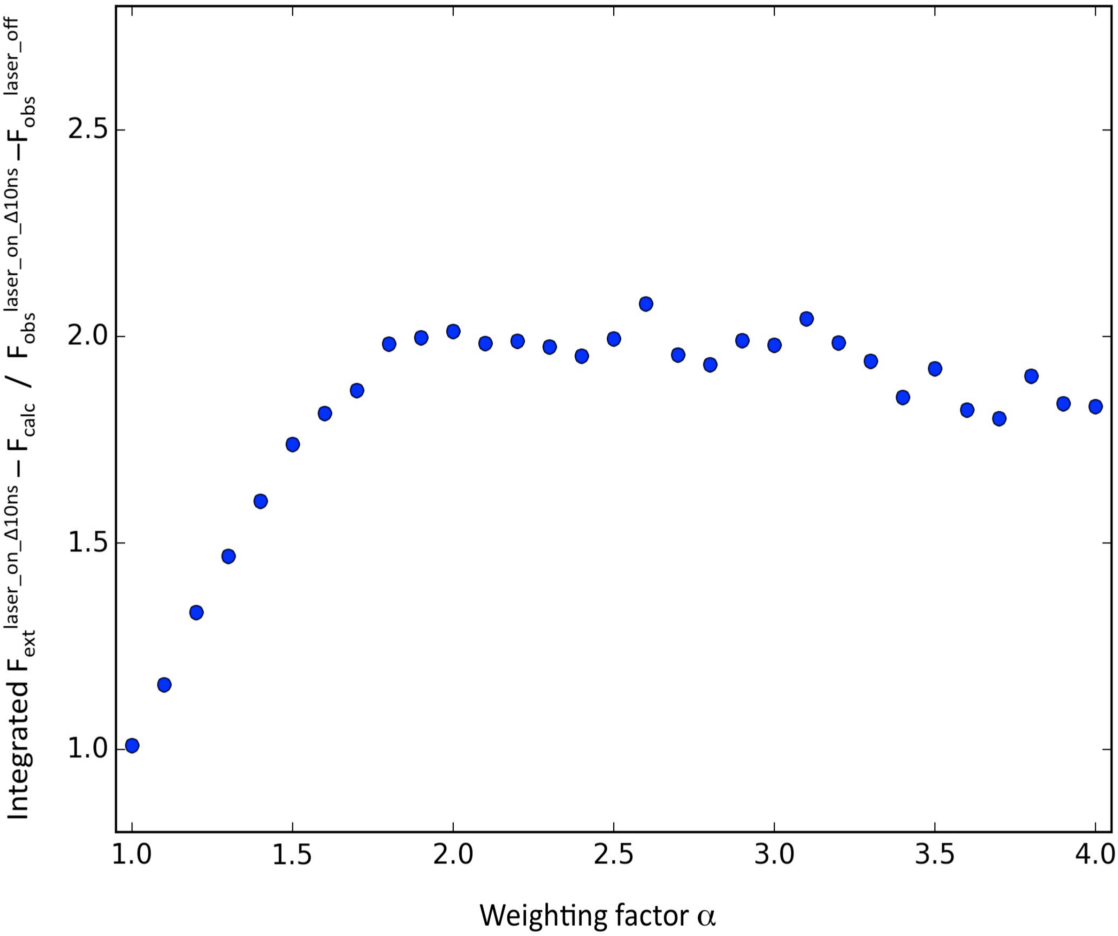
Supplementary Figure 1: Transient absorption spectroscopy of rsEGFP2 in D_2O solution (50 mM HEPES pD 8, 50 mM NaCl). (*a*) Time-resolved difference absorption spectra recorded after nanosecond excitation at 410 nm of the *trans* protonated *off*-state. The spectrum without laser excitation was subtracted to calculate the difference spectrum. The coloured arrows correspond to the three time constants identified by a global fit analysis of kinetic traces at all wavelengths. (*b*) Decay associated spectra obtained by fitting the kinetic traces in panel *a* for all wavelengths with three exponential functions.



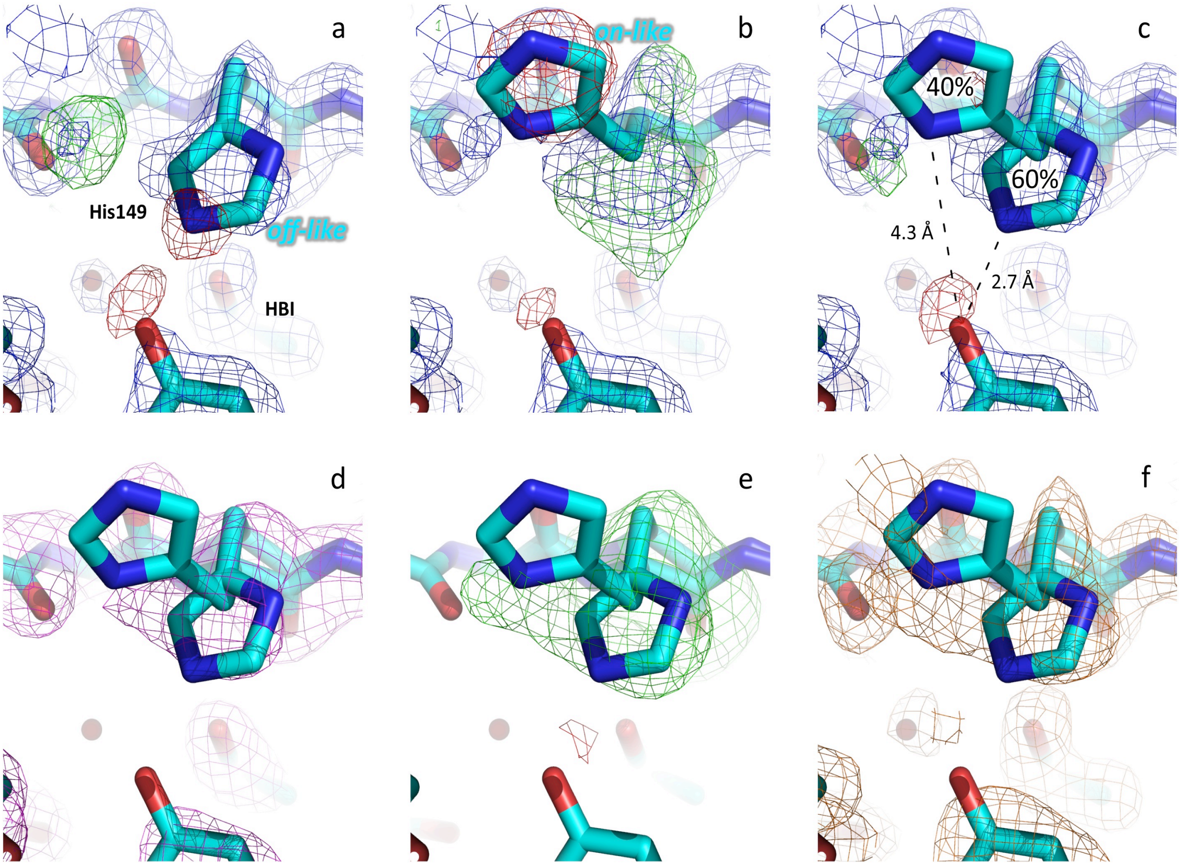
Supplementary Figure 2: Transient absorption spectroscopy of a suspension of rsEGFP2 microcrystals in 100 mM HEPES pH 8, 2.5 M ammonium sulphate. (*a*) Time-resolved difference absorption spectra recorded after a 410 nm nanosecond excitation of the *trans* protonated *off* state. The spectrum without laser excitation was subtracted to calculate the difference spectrum. The coloured arrows correspond to the three time constants identified by a global fit analysis of kinetic traces for all wavelengths. (*b*) Decay associated spectra obtained by fitting the kinetic traces in panel *a* for all wavelengths with three exponential functions.



Supplementary Figure 3: Transient absorption spectroscopy of rsEGFP2 in H_2O solution (50 mM HEPES pH 8, 50 mM NaCl, 1.25 M ammonium sulphate). (*a*) Time-resolved difference absorption spectra recorded after a 410 nm nanosecond excitation of the *trans* protonated *off* state. The coloured arrows correspond to the three time constants identified by a global fit analysis of kinetic traces for all wavelengths. The spectrum without laser excitation was subtracted to calculate the difference spectrum. (*b*) Decay associated spectra obtained by fitting the kinetic traces in panel *a* for all wavelengths with three exponential functions.



Supplementary Figure 4: Determination of the weighting factor α, corresponding to the inverse of the occupancy of the laser-on-Δ10ns model. The ratio of integrated peaks in the $F_{extrapolated}^{laser\_on\_\Delta 10ns} - F_{calc}^{laser\_off}$ map and of integrated peaks in the qW $F_{obs}^{laser\_on\_\Delta 10ns}$ - $F_{obs}^{laser\_off}$ map is plotted as a function of α.



Supplementary Figure 5: Modelling the side chain of His149 in the *laser_on_Δ10ns* intermediate structure (*a*) in a conformation (*off*-like) similar to His149-*off* in the *off*-state (*b*) in a conformation (*on*-like) similar to His149-*on* in the *on*-state (*c*) in a mixture of both at occupancies of 40% and 60%, respectively. B-sharpened $2F_{extrapolated}^{laser\_on\_\Delta10ns}$ - F_{calc} (blue, 1σ) and $F_{extrapolated}^{laser\_on\_\Delta10ns}$ - F_{calc} maps (green/red, +/- 3σ, respectively) obtained using extrapolated structure factors are shown. (d) Simulated annealing composite omit $2F_{extrapolated}^{laser\_on\_\Delta10ns}$ - F_{calc} map (magenta, 1σ). (*e*) Polder map where His149 was omitted (green/red, +/- 3σ, respectively). (*f*) Feature enhanced $2F_{extrapolated}^{laser\_on\_\Delta10ns}$ - F_{calc} map (orange, 1σ). Models are displayed in cyan.

Acknowledgments: We are grateful to So Iwata and Rie Tanaka for valuable comments and technical assistance related to the experiments at SACLA. The XFEL experiments were carried out at BL3-EH4c of SACLA with the approval of the Japan Synchrotron Radiation Research Institute (JASRI; Proposal No. 2015A8031). We warmly thank the SACLA staff for assistance. MSl is grateful to Jean-Julien Dubois for technical support in time-resolved spectroscopy experiments and MW to Marco Kloos and James Remington for discussions. The study was supported by a travel grant from the CEA to JW, NC, VA, EdlM, PM, JPC and MW, a grant from the CNRS (PEPS SASLELX) to MW, an ANR grant to MW, MC, MSl (BioXFEL) and a PhD fellowship from Lille University to LMU. We acknowledge support from the Max Planck Society. This work used the Multistep Protein Purification Platform (MP3) of the Grenoble Instruct centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). The Chevreul Institute (FR 2638), the Ministère de l'Enseignement Supérieur et de la Recherche, the Région Nord-Pas de Calais and FEDER are acknowledged for financial support. GNK acknowledges support from the E.U. under the program FP7-PEOPLE-2011-ITN NanoMem (project number 317079). JW acknowledges support from the CEA through the IRTELIS PhD program.

References

1. Nienhaus, K. & Nienhaus, G.U. Fluorescent proteins for live-cell imaging with super-resolution. *Chem Soc Rev* 43, 1088-106 (2014).
2. Adam, V., Berardozzi, R., Byrdin, M. & Bourgeois, D. Phototransformable fluorescent proteins: Future challenges. *Curr Opin Chem Biol* 20C, 92-102 (2014).
3. Hofmann, M., Eggeling, C., Jakobs, S. & Hell, S.W. Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins. *Proc Natl Acad Sci U S A* 102, 17565-9 (2005).
4. Gustafsson, M.G.L. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences of the United States of America* 102, 13081-13086 (2005).
5. Nienhaus, K. & Nienhaus, G.U. Photoswitchable Fluorescent Proteins: Do Not Always Look on the Bright Side. *ACS Nano* (2016).
6. Brakemann, T. et al. A reversibly photoswitchable GFP-like protein with fluorescence excitation decoupled from switching. *Nature Biotechnology* 29, 942 (2011).
7. Andresen, M. et al. Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proc Natl Acad Sci U S A* 102, 13070-4 (2005).
8. Habuchi, S. et al. Reversible single-molecule photoswitching in the GFP-like fluorescent protein Dronpa. *Proc Natl Acad Sci U S A* 102, 9511-6 (2005).

9. Ando, R., Mizuno, H. & Miyawaki, A. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. *Science* 306, 1370-3 (2004).
10. Fron, E. et al. Ultrafast excited-state dynamics of the photoswitchable protein Dronpa. *J Am Chem Soc* 129, 4870-1 (2007).
11. Warren, M.M. et al. Ground-state proton transfer in the photoswitching reactions of the fluorescent protein Dronpa. *Nat Commun* 4, 1461 (2013).
12. Kaucikas, M., Tros, M. & van Thor, J.J. Photoisomerization and proton transfer in the forward and reverse photoswitching of the fast-switching M159T mutant of the Dronpa fluorescent protein. *J Phys Chem B* 119, 2350-62 (2015).
13. Yadav, D. et al. Real-time monitoring of chromophore isomerization and deprotonation during the photoactivation of the fluorescent protein Dronpa. *J Phys Chem B* 119, 2404-14 (2015).
14. Lukacs, A. et al. Protein photochromism observed by ultrafast vibrational spectroscopy. *J Phys Chem B* 117, 11954-9 (2013).
15. Grotjohann, T. et al. rsEGFP2 enables fast RESOLFT nanoscopy of living cells. *Elife* 1, e00248 (2012).
16. El Khatib, M., Martins, A., Bourgeois, D., Colletier, J.P. & Adam, V. Rational design of ultrastable and reversibly photoswitchable fluorescent proteins for super-resolution imaging of the bacterial periplasm. *Sci Rep* 6, 18459 (2016).
17. Andresen, M. et al. Structural basis for reversible photoswitching in Dronpa. *Proc Natl Acad Sci U S A* 104, 13005-9 (2007).
18. Ai, H.-w., Henderson, J N., Remington, S J. & Campbell, Robert E. Directed evolution of a monomeric, bright and photostable version of Clavularia cyan fluorescent protein: structural characterization and applications in fluorescence imaging. *Biochemical Journal* 400, 531-540 (2006).
19. Adam, V. et al. Structural characterization of IrisFP, an optical highlighter undergoing multiple photo-induced transformations. *Proc Natl Acad Sci U S A* 105, 18343-8 (2008).
20. Coquelle, N. et al. Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nature Chemistry* 10, 31-37 (2018).
21. Yabashi, M., Tanaka, H. & Ishikawa, T. Overview of the SACLA facility. *Journal of Synchrotron Radiation* 22, 477-484 (2015).
22. Schiro, G., Woodhouse, J., Weik, M., Schlichting, I. & Shoeman, R.L. Simple and efficient system for photoconverting light-sensitive proteins in serial crystallography experiments. *Journal of Applied Crystallography* 50, 932-939 (2017).
23. Ursby, T. & Bourgeois, D. Improved Estimation of Structure-Factor Difference Amplitudesfrom Poorly Accurate Data. *Acta Crystallographica Section A* 53, 564-575 (1997).
24. Terwilliger, T.C. & Berendzen, J. Difference refinement: obtaining differences between two related structures. *Acta Crystallogr D Biol Crystallogr.* 51, 609-18. (1995).
25. Colletier, J.-P. et al. Use of a `caged' analogue to study the traffic of choline within acetylcholinesterase by kinetic crystallography. *Acta Crystallographica Section D* 63, 1115-1128 (2007).
26. Genick, U. Structure-factor extrapolation using the scalar approximation: theory, applications and limitations. *Acta Crystallographica Section D* 63, 1029-1041 (2007).
27. Brejc, K. et al. Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. *Proc Natl Acad Sci U S A* 94, 2306-11 (1997).

28. Agmon, N. Proton pathways in green fluorescence protein. *Biophys J* 88, 2452-61 (2005).
29. Regis Faro, A. et al. Low-temperature chromophore isomerization reveals the photoswitching mechanism of the fluorescent protein padron. *J Am Chem Soc* 133, 16362-5. (2011).
30. Brakemann, T. et al. Molecular basis of the light-driven switching of the photochromic fluorescent protein Padron. *J Biol Chem* 285, 14603-9 (2010).
31. Barends, T.R. et al. Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* 350, 445-50 (2015).
32. Pande, K. et al. Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein. *Science* 352, 725-9 (2016).
33. Tenboer, J. et al. Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* 346, 1242-6 (2014).
34. Nango, E. et al. A three-dimensional movie of structural changes in bacteriorhodopsin. *Science* 354, 1552-1557 (2016).
35. Shimada, A. et al. A nanosecond time-resolved XFEL analysis of structural changes associated with CO release from cytochrome c oxidase. *Sci Adv* 3, e1603042 (2017).
36. Suga, M. et al. Light-induced structural changes and the site of O=O bond formation in PSII caught by XFEL. *Nature* 543, 131-135 (2017).
37. Nogly, P. et al. Lipidic cubic phase injector is a viable crystal delivery system for time-resolved serial crystallography. *Nat Commun* 7, 12314 (2016).
38. Tosha, T. et al. Capturing an initial intermediate during the P450nor enzymatic reaction using time-resolved XFEL crystallography and caged-substrate. *Nat Commun* 8, 1585 (2017).
39. Kern, J. et al. Simultaneous femtosecond X-ray spectroscopy and diffraction of photosystem II at room temperature. *Science* 340, 491-5 (2013).
40. Young, I.D. et al. Structure of photosystem II and substrate binding at room temperature. *Nature* 540, 453-457 (2016).
41. Domratcheva, T. & Schlichting, I. Spiers Memorial Lecture. Introductory lecture: the impact of structure on photoinduced processes in nucleic acids and proteins. *Faraday Discuss* 207, 9-26 (2018).
42. Colletier, J.P. et al. Serial Femtosecond Crystallography and Ultrafast Absorption Spectroscopy of the Photoswitchable Fluorescent Protein IrisFP. *J Phys Chem Lett* 7, 882-7 (2016).
43. Lomb, L. et al. An anti-settling sample delivery instrument for serial femtosecond crystallography. *Journal of Applied Crystallography* 45, 674-678 (2012).
44. Weierstall, U., Spence, J.C. & Doak, R.B. Injector for scattering measurements on fully solvated biospecies. *Rev Sci Instrum* 83, 035108 (2012).
45. Tono, K. et al. Diverse application platform for hard X-ray diffraction in SACLA (DAPHNIS): application to serial protein crystallography using an X-ray free-electron laser. *Journal of Synchrotron Radiation* 22, 532-537 (2015).
46. Kameshima, T. et al. Development of an X-ray pixel detector with multi-port charge-coupled device for X-ray free-electron laser experiments. *Rev Sci Instrum* 85, 033110 (2014).
47. Foucar, L. et al. CASS-CFEL-ASG software suite. *Computer Physics Communications* 183, 2207-2213 (2012).
48. Coquelle, N. et al. Raster-scanning serial protein crystallography using micro- and nano-focused synchrotron beams. *Acta Crystallographica Section D* 71, 1184-1196 (2015).

49. Nass, K. et al. Protein structure determination by single-wavelength anomalous diffraction phasing of X-ray free-electron laser data. *IUCrJ* 3, 180-91 (2016).
50. McCoy, A.J. et al. Phaser crystallographic software. *Journal of Applied Crystallography* 40, 658-674 (2007).
51. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica. Section D, Biological Crystallography* 66, 213-221 (2010).
52. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica. Section D, Biological Crystallography* 60, 2126-2132 (2004).
53. Brunger, A.T. Version 1.2 of the Crystallography and NMR system. *Nature Protocols* 2, 2728-2733 (2007).
54. The PyMOL Molecular Graphics System. 1.5.0.4 edn (Schrödinger, LLC.).

5. Chapter IV: Structural characterization of the *on-* and *off*-state of rsEGFP2 V151A and V151L

Introduction

Chapter II reported on a twisted chromophore conformation that is adopted in the excited state 1 ps after illumination of rsEGFP2 in its *off*-state (Coquelle et al., 2018). From this excited-state structure, the hypothesis was formulated that amino acid side chains in the vicinity of the chromophore might affect its motion during photoswitching and thus the photophysical properties of rsEGFP2. In particular, the side chain of Val151 seemed to hinder formation of the twisted chromophore. In order to test the hypothesis that Val151 possibly acts as a gate obstructing chromophore isomerization (Figure 4d in Coquelle et al., 2018, chapter II), variants were generated in which Val151 was substituted by an amino acid with a longer (Leu) or a shorter (Ala) side chain. They were characterized structurally in the framework of this thesis and spectroscopically by Dr. Virgile Adam (IBS), in order to assess the effect of these mutations on protein properties, thereby clarifying the role of Val151 in rsEGFP2.

Figure 1 shows the absorption spectra of the *on*- and *off*-states of rsEGFP2 and its two variants, V151A and V151L (data provided by Dr. Virgile Adam). The *on*-state spectra of the three versions display a peak with a maximum at around 488 nm. The spectra for rsEGFP2 and its variant V151L are almost super-imposable. A shoulder to the peak is observed at ~ 400 nm in rsEGFP2 and its variant V151A, which is missing in the *on*-state spectrum of rsEGFP2 V151L. The *off*-state spectra of the three versions of the protein are quite different. The main peak of the *off*-state spectrum of rsEGFP2 V151A (maximum at 404 nm) is blue-shifted compared to the original protein (maximum at 406 nm), and the peak of the *off*-state spectrum of rsEGFP2 V151L (maximum at 425 nm) is red-shifted compared to the original protein. A secondary peak is observed in the *off*-state spectrum of rsEGFP2 and V151L around 490 nm. The red-shift of the rsEGFP2 V151L *off*-state spectrum increases the overlap of the *on*- and *off*-state spectra, whereas the observed blue-shift of the *off*-state spectrum of V151A decreases this overlap. A decrease and an increase in the overlap of the *on*- and *off*-state spectra could lead to an increase and a decrease of the switching contrast (the ratio between the fluorescence maximum measured under 405-nm-light illumination and the fluorescence minimum measured under 488-nm-light illumination after an equilibrium is reached). Indeed, this contrast is improved in rsEGFP2 V151A and degraded in rsEGFP2 V151L (Figure 2). Table 1 reports various photophysical parameters for rsEGFP2 and its variants (data provided by Dr. Virgile Adam).

Overall, the *off*-to-*on* switching quantum yield is ten times higher than the *on*-to-*off* one for rsEGFP2 and its variants (Table 1). The mutation leads to an increase and a decrease of *off*-to-*on* and *on*-to-*off* switching quantum yields for rsEGFP2 V151A and V151L, respectively, compared to the original version. In other words, V151A is a better switcher, while V151L is less efficient. The shoulder observed in the *on*-state spectrum at 400 nm of rsEGFP2 (cyan trace in Figure 1) and rsEGFP2 V151A (green trace in Figure 1) could correspond to a small population of proteins in the *off*-state generated by ambient light. Such ambient-light conditions would not be sufficient to switch *off* rsEGFP2 V151L which could explain why the shoulder is missing in its *on*-state spectrum (yellow trace in Figure 1). Moreover, the secondary peak observed at around 490 nm in the *off*-state spectrum of rsEGFP2 (magenta trace in Figure 1) and rsEGFP2 V151L (red trace in Figure 1) could result from a residual *on*-state population, illustrating the decreased contrast compared to rsEGFP2 V151A for which this secondary peak is missing in the *off*-state spectrum (blue trace in Figure 1).

Structural investigation of the rsEGFP2 V151A and V151L variants was then undertaken in this thesis to reveal potential structural changes in the chromophore pocket that could help explaining the observed photo-physical differences between variants and the original protein. The current chapter presents X-ray crystallographic structure solution and analysis of rsEGFP2 V151A and V151L variants in their *on*- and *off*-states.

Results

X-ray crystal structures of rsEGFP2 variants

X-ray crystallographic data were collected at the ESRF synchrotron on the obtained macrocrystals to solve the structure of the fluorescent *on*- and the non-fluorescent *off*-states of the rsEGFP2 V151A and V151L variants. To determine the structure of the resting state, *i.e.* the *on*-state, crystals were not illuminated. To generate the *off*-state prior to data collection, crystals were exposed to 488-nm laser light until the crystal turned white, indicating that photoswitching from the *on*- to the *off*-state occurred.

Crystal structures of rsEGFP2 V151A in the on- and off-states

A model of the *on*-state of rsEGFP2 V151A was built in a well defined $2F_{obs}$-F_{calc} electron density map at 1.4 Å resolution with final R_{work}/R_{free} values of 17.76%/14.82% (Figure 3a and Table 2). The mutation at position 151 from valine to alanine is evident. The chromophore appears to be a *cis* isomer, as expected for the *on*-state. The phenol group of the chromophore and the imidazole ring of the His149 were refined with an occupancy of 0.9 indicating that these two groups are either not fully occupied, or partially disordered. A model of the 488-nm illuminated state of rsEGFP2 V151A was built and refined in a $2F_{obs}$-F_{calc} map at 1.55 Å resolution with a final R_{work}/R_{free} of 20.85%/16.04% (Figure 3b and Table 2). In contrast to the *on*-state, the chromophore appears to be in a *trans* isomerization state, as expected for the *off*-state and also His149 is in a different conformation.

Crystal structure of rsEGFP2 V151L in the on- and off-states

A model of the fluorescent *on*-state of rsEGFP2 V151L was built in a $2F_{obs}$-F_{calc} electron density map and refined at 1.8 Å resolution with final R_{work}/R_{free} of 20.91%/16.17% (Figure 3c and Table 2). It can be seen that the valine in position 151 has indeed been substituted by a leucine. The chromophore appears in the *cis* configuration. A model of the 488-nm illuminated state of rsEGFP2 V151L was built in a $2F_{obs}$-F_{calc} map and refined at 1.7 Å resolution with a final R_{work}/R_{free} of 15.28%/19.65% (Figure 3d and Table 2). The chromophore, His149 and Leu151 are displayed in two conformations with occupancies of 50% each. The alternate conformations of the chromophore correspond to a *trans* isomer, characteristic of the *off*-state, and to a *cis* isomer similar to the one in the *on*-state, indicating that *on*-to-*off* photoswitching was not fully efficient. Several datasets were collected on different crystals after 488-nm illumination, and in each case a mixture of *cis* and *trans* isomers was observed. Table 2 reports statistics for crystallographic data collections and processing and of model refinement.

Comparison of rsEGFP2 V151A and rsEGFP2 on- and off-state structures

The *on*-state structures of rsEGFP2 V151A and of rsEGFP2 are very similar, except for the introduced mutation (Figure 4a). The distance between the oxygen of the phenol and His149-ND1 was measured to be 2.7 and 2.8 Å for rsEGFP2 and rsEGFP2 V151A respectively, indicating a possible hydrogen bond in both cases. There is a striking difference between the *off*-state of rsEGFP2 and the *off*-state of rsEGFP2 V151A (Figure 4b). His149 loses the hydrogen bond to the chromophore when switching from the *on*- to the *off*-state, but only in rsEGFP2 does its NE2 atom form another bond with the phenol oxygen of Tyr146 (distances of 2.6 and 4.9 Å for rsEGFP2 and rsEGFP2 V151A, respectively). In both *off*-state structures, the chromophore is hydrogen bonded to a water molecule (distances of 2.8 and 2.9 Å for rsEGFP2 and rsEGFP2 V151A, respectively).

Comparison of rsEGFP2 V151L and rsEGFP2 on- and off-state structures

The *on*-state of the V151L variant is very similar to the one of rsEGFP2, as for V151A, except for the introduced mutation (Figure 4c). The distance between the oxygen of the chromophore phenol group and His149-ND1 is 2.9 Å for V151L, indicating a possible hydrogen bond. The comparison between rsEGFP2 and rsEGFP2 V151L *off*-state structures displays a most striking difference: the chromophore is in two different *trans* conformations (Figure 4d). The chromophore in the *off*-state of rsEGFP2 is not hydrogen bonded to His149 as in the *on*-state, but rather to a water molecule. In V151L, to the contrary, the oxygen atom of the chromophore phenol group forms a hydrogen bond with His149-ND1 (distance of 2.7 Å). His149 in the *off*-state does not adopt the same conformer in the V151L variant as in rsEGFP2. In the latter, Tyr146 moved from its *on*-state position to form a hydrogen bond with His149 in the *off*-state. In the V151L variant (Figure 4d), however, Tyr146 adopts the same conformation in the *on*- and the *off*-states (Figure 3d).

Comparison of on- and off-state structures in V151A and V151L variants

Comparing the *on*- and the *off*-state structures shows that photoswitching involves a movement in His149 and chromophore isomerization in both variants, yet with a smaller amplitude in the V151L than in the V151A variant (Figure 5a and b). The *on*-state structures of the two variants are very similar (Figure 5c) but the *off*-state structures differ (Figure 5d). In the V151L variant, the chromophore is hydrogen bonded to His149, which is in a different conformation than in the V151A variant.

Discussion

The crystal structures of rsEGFP2 V151A and V151L variants in their *on*- and *off*-states were solved and compared to the corresponding structures of the original rsEGFP2 protein. The *on*-state structures of the variants and of the original protein are very similar (Figures 4a and 4c). However, there is an important difference between the *off*-state structures of the two variants. While in the V151A variant the chromophore displays a conformation close to the one observed in rsEGFP2 (Figure 4b), the V151L *off*-state structure reveals a chromophore conformation different from the one of rsEGFP2 (Figure 4d), yet similar to the one in the *off*-state of rsFolder (El Khatib et al., 2016). As a consequence, the isomerization amplitude for the V151L variant (Figure 5b) is much smaller than for the V151A variant and the orignal protein (Figure 4b). Substituting valine in position 151 by a leucine residue thus introduced a steric obstacle limiting the amplitude of chromophore isomerization during photoswitching. Another unique feature of the V151L variant is its hydrogen bond between the chromophore phenol group and His149 in both its *on*- and *off*-states (Figure 5b), which is not present in the *off*-states of the V151A variant and the original protein (Figure 4b).

The degraded switching contrast of V151L compared to rsEGFP2 (Table 1, Figure 2) is manifested in the structure of the 488-nm illuminated state (Figure 3d), where it appears clearly that half of the crystalline proteins remained in the *on*-state as judged from a 50% occupancy of a *cis* chromophore that adds to the 50% of a *trans* chromophore characteristic of the *off*-state. This mixture reveals the difficulty, or even the probable impossibility, to fully switch-off the protein. Furthermore, we reported that the *on*-to-*off* switching quantum yield of V151A is increased compared to rsEGFP2 (Coquelle et al., 2018). In other words, V151A is a better switcher than rsEGFP2. A model of the V151A structure in the *on*-state was refined with a reduced occupancy for the chromophore and for His149, indicating a possible depopulation of the *on*-state due to the exposure to the ambient light of the crystal during harvesting and flash-cooling.

While the *on*-state spectra of rsEGFP2 V151A and V151L are quite similar to the one of rsEGFP2, with a maximum at around 488 nm, the *off*-state spectra are quite different (Figure 1). The *off*-state spectrum of rsEGFP2 shows a peak at 406 nm. The same peak is blue-shifted with a maximum at 404 nm for V151A and red-shifted with a maximum at 425 nm for V151L. The differences in chromophore conformation of rsEGFP2 and its variants in the *off*-state correlates with the observed shift in their respective absorption spectrum. This shift leads to a change in the degree of overlap between *on*- and *off*-state absorption spectra, providing an explanation for the improved and degraded contrast in V151A and V151L, respectively (Figure 1). Moreover, a secondary absorption peak at around 480 nm in the *off*-states of rsEGFP2 and of V151L could be interpreted as resulting from a residual *on*-state fraction that illustrates the lower contrast compared to V151A.

Furthermore, rsEGFP2 displays *on*-to-*off* and *off*-to-*on* photoswitching quantum yields intermediate between those of its two variants (Table 1). Moreover, the *off*-state spectrum of rsEGFP2 appears like a mixture of the *off*-state spectra of V151A and V151L (Figure 2). Together, these observations support the notion that *off*-to-*on* photoswitching behavior of the original rsEGFP2 protein is a mixture of the behaviors of its two variants. The crystal structures of the two variants provide structural clues to understand the photophysical properties of rsEGFP2 and its two variants. Indeed, the side-chain length of the residue in position 151 alters the chromophore environnement and seems to modify chromophore-switching behavior. The longer (shorter) the sidechain is, the lower (higher) are the switching contrast and the switching quantum yields. In the following chapter, experimental evidence will be provided that supports the hypothesis that the rsEGFP2 *off*-state can be decomposed into two structurally different *off*-states.

The results presented in this and the following chapter were the starting point of a new PhD project (Kyprianos Hadjidemetriou) aimed at studying the rsEGFP2 variants by TR-SFX following the same approach than already published (chapter II, Coquelle et al., 2018). In addition to the femtosecond transient difference absorption study by Michel Sliwa (unpublished), TR-SFX data at pump probe delays of 1, 3, and 20 ps were successfully collected on V151A microcrystals during a recent LCLS beamtime (LR38, 22 – 26 February 2018). V151A microcrystals for this experiment were produced according to the protocol presented in chapter I.

Material and methods

Site-directed mutagenesis

The rsEGFP2 variants were generated following the same procedure than described in (Coquelle et al., 2018). 5’- CAACAGCCACAACGCCTATATCATGGCC-3’ and 5’-

CTACAACAGCCACAACCTCTATATCATGGCCG-3' fragments were used as primers for V151A and V151L substitution, respectively.

Expression and purification

The variants (V151A and V151L) of rsEGFP2 were expressed and purified following the same procedure than described in chapter I. Competent *E. coli* BL21 (DE3) cells were transformed. Bacteria were grown in self-inducible medium at 37 °C. After lysis, the protein, fused to a N-terminal polyhistidine tag, was purified by affinity chromatography using Ni-NTA beads followed by a size-exclusion chromatography step.

Crystallization

For V151L, crystallization trials were performed using the vapor diffusion method with hanging drops in 24-well plates. The well solution was composed of 100 mM HEPES pH 7.8 - 8.4 and 1.7 - 2.4 M ammonium sulfate for a total volume of 500 μl. Drops were prepared at 20 °C by mixing 2 μl from the well with 2 μl of protein solution at 10 mg/ml in 50 mM HEPES pH 7.5, 50 mM NaCl. Crystals appeared within a few days and were still growing during a few weeks to reach their final size.

For V151A, crystallization trials were first assessed following the same procedure using the same conditions than for rsEGFP2 V151L but without success (drops remained clear even after several days). A cross-seeding approach was then adopted. First, V151A crystals were generated by vapor diffusion in sitting drops on 24-well plates using seeds of original rsEGFP2 whose preparation is described in chapter I. The well solution was composed of 100 mM HEPES pH 8, 2 M ammonium sulfate. Drops were prepared at 20 °C by mixing 10 μl of protein solution at 10 mg/ml, 50 mM HEPES pH 7.5, 50 mM NaCl with 10 μl of solution at 2 M ammonium sulfate, 100 mM HEPES pH 8 containing rsEGFP2 seeds diluted 1/1000 (v/v). After a few days, crystals appeared and reached their final size. Drops were then collected and pooled in a tube. After centrifugation, the supernatant was removed and the pellet suspended in 2 M ammonium sulfate, 100 mM HEPES pH 8 to a total volume of ~240 μl. Crystals were then crushed using a Bead Bug Homogenizer for four to five 180-second cycles at 4500 rpm The tubes were centrifuged, the supernatant was removed and the pellet formed of crystal fragments was suspended in 2 M ammonium sulfate, 100 mM HEPES pH 8. Seed stocks were constituted with a range of seed dilutions (1/1000, 1/2000, 1/4000, 1/8000, 1/16000, 1/32000 (v/v)) in 1.8 - 2.4 M ammonium sulfate concentration. Finally, rsEGFP2 V151A crystals were generated on 24-well plates using the vapor diffusion method in hanging drops. The well solution was composed of 100 mM HEPES pH 8, and 1.8 - 2.4 M ammonium sulfate concentration. Drops were prepared by mixing 2 μl of protein solution at 10 mg/ml, 50 mM HEPES pH 7.5, 50 mM NaCl with 2 μl of a V151A seed stock at various dilutions. Crystal appeared and reached their final size within a few days. The reason why seeding was necessary to obtain crystals is still unclear. Generation of V151A macrocrystals was later repeated by Maxime Killer who spontaneously obtained crystals without using seeds.

In crystallo photoswitching, X-ray data collection and processing

For structure determination of the rsEGFP2 V151A variant in its *on*-state, a macrocrystal (200 × 20 × 20 μm^3) grown in 2.2 M ammonium sulfate, 100 mM HEPES pH 8 was harvested with a cryo-loop and soaked for a few seconds in a cryo-solution composed of 2 M ammonium sulfate, 100 mM HEPES pH 8, 20 % glycerol (the same cryo-solution was used for all crystal preparations presented here). Then the crystal was flash-cooled in liquid nitrogen. For structure determination of V151A in the *off*-state a macrocrystal (150 × 10 × 10 μm^3) grown in the same condition was exposed to 488-nm laser-light, guided through a fiber, for around one minute at room temperature (RT) directly in the drop where it grew until the crystal turned white. The crystal was then harvested with a cryo-loop and soaked in the cryo-solution before being flash-cooled. For structure determination of the rsEGFP2 V151L variant in its *on*-state, a macrocrystal (300 × 100 × 100 μm^3) grown in 1.7 M ammonium sulfate, 100 mM HEPES pH 7.8, 10 mg/ml was harvested with a cryo-loop and soaked in the cryo-solution before being flash-cooled in liquid nitrogen. For structure determination of V151L in the *off*-state, the same procedure than for V151L was applied. A macrocrystal (150 × 50 × 50 μm^3) grown in 1.8 M ammonium sulfate, 100 mM HEPES pH 8.2, 10 mg/ml protein was exposed to 488-nm laser-light for around one minute at RT as described above and then soaked in the cryo-solution before being flash-cooled.

X-ray crystallographic oscillation data were collected at 100 K on ID29 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) operating in 7/8 multibunch mode at an X-ray wavelength of 0.976 Å (12.7 keV) and with a beamsize of 30 µm × 30 µm. Data were recorded with a Pilatus 6M detector (Dectris). For rsEGFP2 V151A in the *on*-state, 1050 frames were recorded with an oscillation range of 0.10 ° and an exposure time of 0.037 s per frame with a photon flux of 1.6×10^{11} ph/s at 5 % transmission. For rsEGFP2 V151A in the *off*-state, 547 frames were recorded (oscillation range 0.15 °, exposure time 0.039 s per frame, photon flux 2.8×10^{11} ph/s, 9.94 % transmission). For rsEGFP2 V151L in the *on*-state, 794 frames were recorded (oscillation range 0.15 °, exposure time 0.1 s per frame, photon flux 2.2×10^{10} ph/s, 1.27 % transmission). For rsEGFP2 V151L in the *off*-state, 950 frames were recorded (oscillation range 0.10 °, exposure time 0.037 s per frame, photon flux 2.8×10^{10} ph/s, 9.02 % transmission). Images were indexed and intensities were integrated and merged using XDS (Kabsch, 2010).

Molecular replacement and structure refinement

For both states of the two variants, molecular replacement was performed with *Phaser* (McCoy et al., 2007) using the cryo-structure of rsEGFP2 in the fluorescent *on*-state as a starting model (PDB entry: 5DTX, El Khatib et al., 2016). Manual model building was performed in *Coot* (Emsley et al., 2010) and refinement with the *Phenix* suite (Afonine et al., 2012) included positional and isotropic individual B factor refinement in reciprocal and real space. *PyMOL* (DeLano, 2002) was used to produce figures. For the structure of the *on*-state of rsEGFP2 V151A, the occupancy of the phenol of the chromophore and the imidazole of His149 was fixed to 0.9 before final cycles of refinement were carried out. For the structure of rsEGFP2 V151L after 488-nm-light illumination, the chromophore, His149 and Leu151 were modeled in alternate conformations with an occupancy fixed to 0.5 before the final cycles of refinement.

Figures and tables

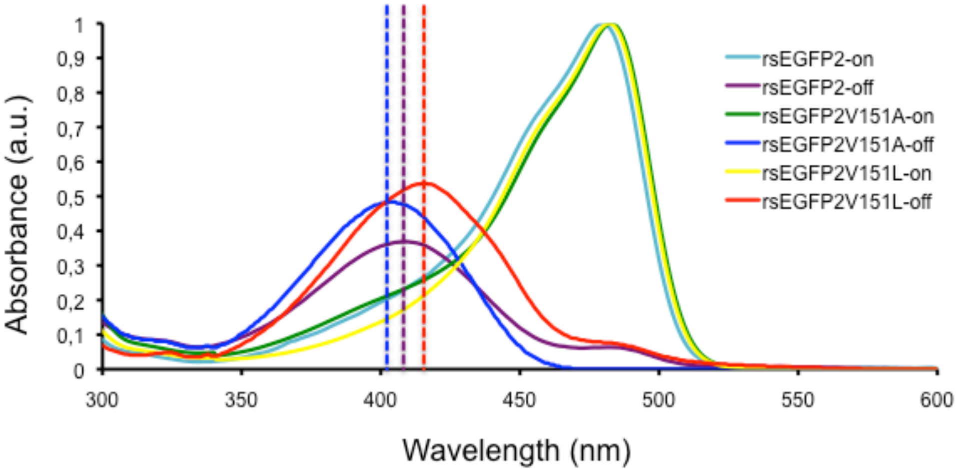


Figure 1: Absorption spectra of rsEGFP2 and its two variants, V151A and V151L in their fluorescent *on*-state and non-fluorescent *off*-state in solution. Spectra of the *on* and *off*-states of rsEGFP2 are represented in cyan and magenta, respectively. Spectra of the *on*- and *off*-states of rsEGFP2 V151A are in green and blue, respectively. Spectra of the *on*- and *off*-states of rsEGFP2 V151L are represented in yellow and red, respectively. Doted vertical lines are indicating the peak maxima in the *off*-state spectra. Spectroscopy measurements, normalization and analyses were performed by Dr. Virgile Adam.

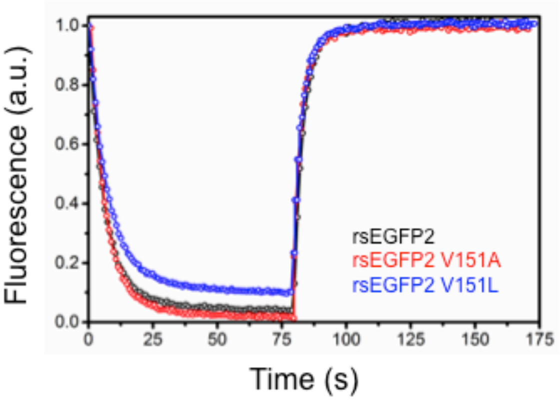


Figure 2: Reversible fluorescence switching cycle of rsEGFP2 (black), rsEGFP2-V151A (red), rsEGFP2-V151L (blue) at low illumination power density (488 nm: 6.7 mW/cm^2, 405 nm: 1 mW/cm^2). Data provided by Dr. Virgile Adam.

	rsEGFP2 WT	rsEGFP2-V151A	rsEGFP2-V151L
Absorption maximum *on* [nm]	479	483	482
Absorption maximum *off* [nm]	406	404	425
Extinction coefficient (ε) *on* (488 nm), pH 7.5 [$M^{-1}.cm^{-1}$]	57150	47660	65847
Extinction coefficient (ε) *off* (404 nm), pH 7.5 [$M^{-1}.cm^{-1}$]	22000	24640	37153
Extinction coefficient (ε) *off* (488 nm), pH 7.5 [$M^{-1}.cm^{-1}$]	87.8	37.4	302.8
Fluorescence quantum yield	0.31	0.25	0.41
Brightness	17 145	11 915	26 997
On-to-*off* switching quantum yield (Φ_{off})	0.0431	0.0639	0.0284
Off-to-*on* switching quantum yield (Φ_{on})	0.402	0.77	0.31
Contrast (based on ε)	~50 (650)	~100 (1270)	~20 (210)
Thermal recovery (h)	4.5	65	10
pKa	5.8	6.3	5.7

Table 1: Photophysical parameters for rsEGFP2 and its V151A and V151L variants. The analysis was carried out by Dr. Virgile Adam.

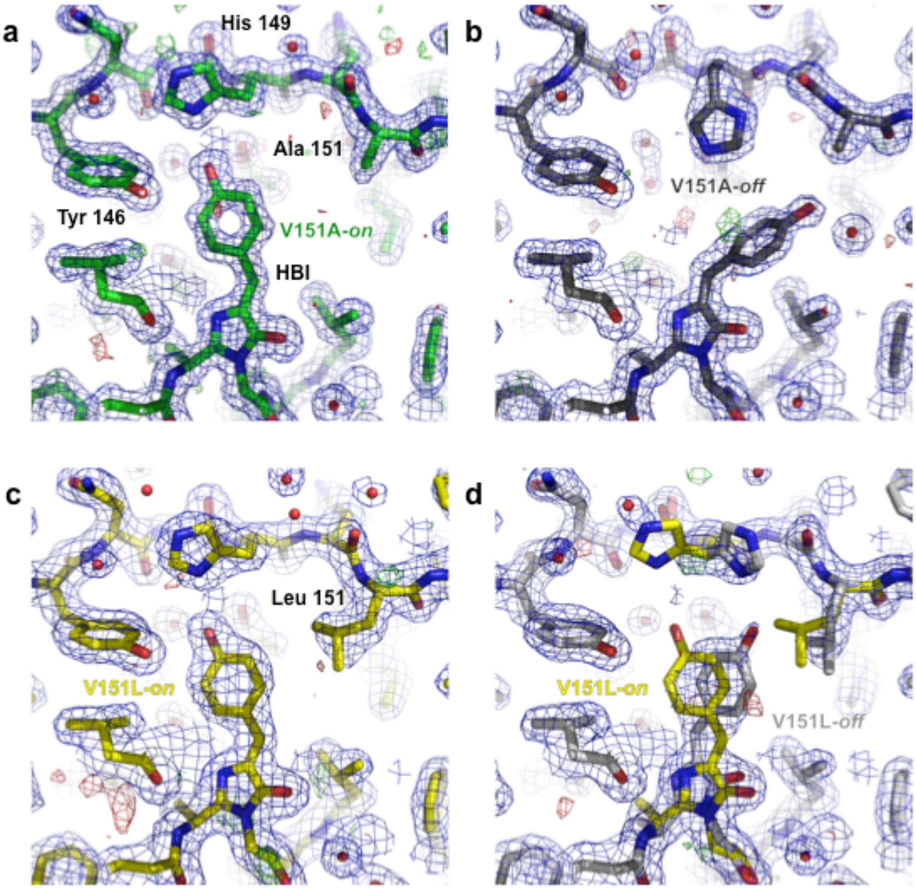


Figure 3 : Zoom on the chromophore (HBI) region in crystal structures of rsEGFP2 variants V151A and V151L in their *on*- and *off*-states. (*a, b*) Electron density maps (mesh) and models of the rsEGFP2 V151A variant in the (*a*) *on*- (green) and (*b*) *off*-state (dark grey). (*c, d*) Electron density maps (mesh) and model of the rsEGFP2 V151L variant in the (*c*) *on*- (yellow) and (*d*) *off*-state (light grey). In (*d*) an alternate conformation of the chromophore corresponding to the *cis* isomer of the *on*-state (yellow) remains in the *off*-state. $2F_o$-F_c (1σ) and F_o-F_c (+/- 3σ) electron density maps are shown in blue and green/red, respectively.

Data collection and processing				
Dataset	rsEGFP2 V151A *on*-state	rsEGFP2 V151A *off*-state	rsEGFP2 V151L *on*-state	rsEGFP2 V151L *off*-state
Illumination (488 nm)	No	Yes	No	Yes
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters				
a (Å)	50.99	51.16	51.14	51.33
b (Å)	62.09	60.29	62.51	61.21
c (Å)	69.41	66.61	70.42	70.29
Collected frames	1050	547	794	950
Observations	166,004 (12,182)*	89,498 (6,540)	91,956 (14,602)	84,740 (13,701)
Resolution (Å)	1.4 (1.44-1.4)	1.55 (1.59-1.55)	1.8 (1.91-1.8)	1.7 (1.81-1.7)
R_{meas} (%)	4.8 (42.5)	4.5 (60.6)	6.5 (46.6)	4.7 (42.8)
CC $_{1/2}$ (%)	99.9 (85.3)	99.9 (70.7)	99.9 (91.6)	99.9 (87.1)
I/σ(I)	17.26 (3.23)	14.82 (2.26)	14.17 (2.72)	4.7 (42.8)
Completeness (%)	97.4 (96.7)	97.0 (98.4)	98.3 (98.3)	97.2 (98.5)
Multiplicity	3.9 (3.9)	3.0 (3.0)	4.3 (4.3)	3.5 (3.5)
Refinement statistics				
Resolution (Å)	1.4 (1.43-1.40)	1.55 (1.60-1.55)	1.8 (1.88-1.80)	1.7 (1.77-1.70)
Number of reflections	42,980 (2,810)	29,616 (2,662)	21,186 (2620)	24,184
R_{free} (%)	17.71 (25.06)	20.85 (28.05)	20.91 (28.90)	19.65 (27.01)
R_{work} (%)	14.74 (21.41)	16.04 (23.71)	16,17 (25.45)	15.28 (21.96)
Number of protein atoms	2,037	2,014	2,064	2,228
Number of solvent atoms	356	278	317	298
B-factor protein (Å^2)	12.84	16.72	24.03	24.51
r.m.s.d bond lengths (Å)	0.014	0.011	0.003	0.007
r.m.s.d angles (°)	1.421	1.261	0.790	1.110

Ramachandran favored (%)	99.21	98.38	98.03	97.78
Ramachandran allowed (%)	0.79	1.62	1.97	1.85
Ramachandran outliers (%)	0.00	0.00	1.74	0.37
Rotamer outliers	0.44	0.45	0.00	2.03
C-beta outliers	1	0	0	0
Clashscore	5.66	6.76	3.18	5.89

Table 2 : Data collection and refinement statistics

[*] Values in brackets correspond to the highest resolution shell.

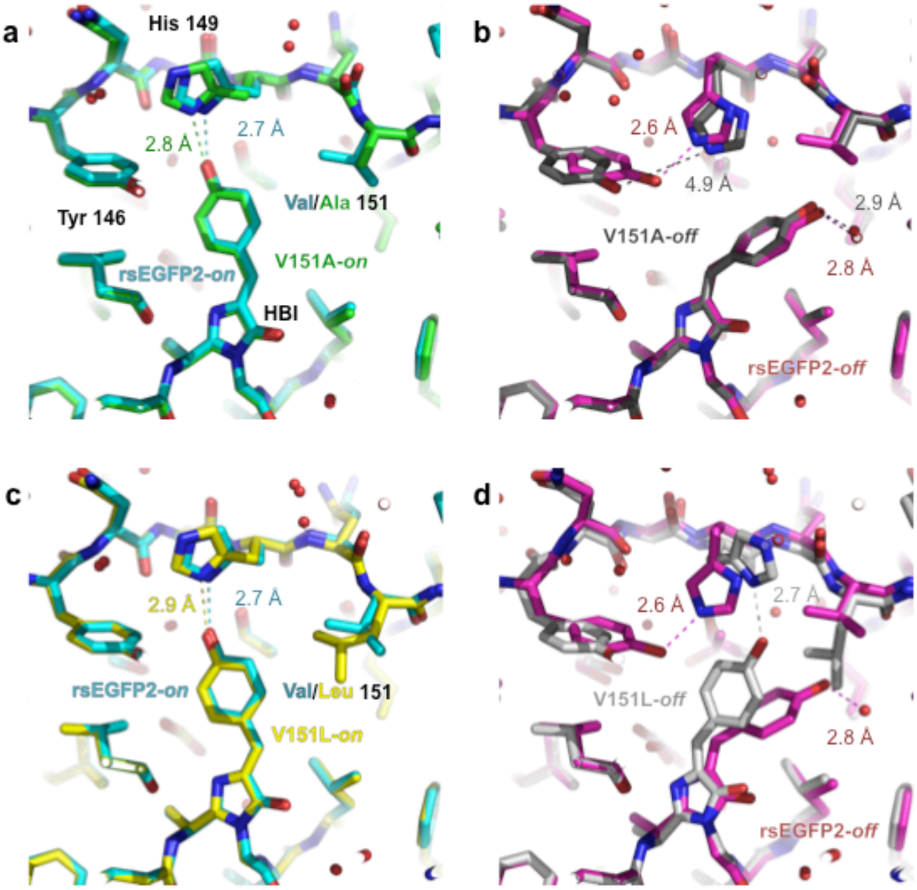


Figure 4: Structural comparison between rsEGFP2 and its V151A and V151L variants. (*a*) Model of rsEGFP2 V151A in the *on*-state (green) superimposed on the model of rsEGFP2 in the *on*-state (cyan, PDB entry: 5DTX, El Khatib et al., 2016) (*b*) Model of rsEGFP2 V151A in the *off*-state (dark grey) superimposed on the model of rsEGFP2 in the *off*-state (magenta, PDB entry: 5DTY, El Khatib et al., 2016). (*c*) Model of rsEGFP2 V151L in the *on*-state (yellow) superimposed on the model of WT rsEGFP2 in the *on*-state (cyan). (*d*) Model of rsEGFP2 V151L in the *off*-state (light grey) superimposed on the model of WT rsEGFP2 in the *off*-state (magenta). The *cis* conformer was removed from rsEGFP2 V151L in the *off*-state for clarity. The length of key hydrogen bonds (dotted lines) are indicated.

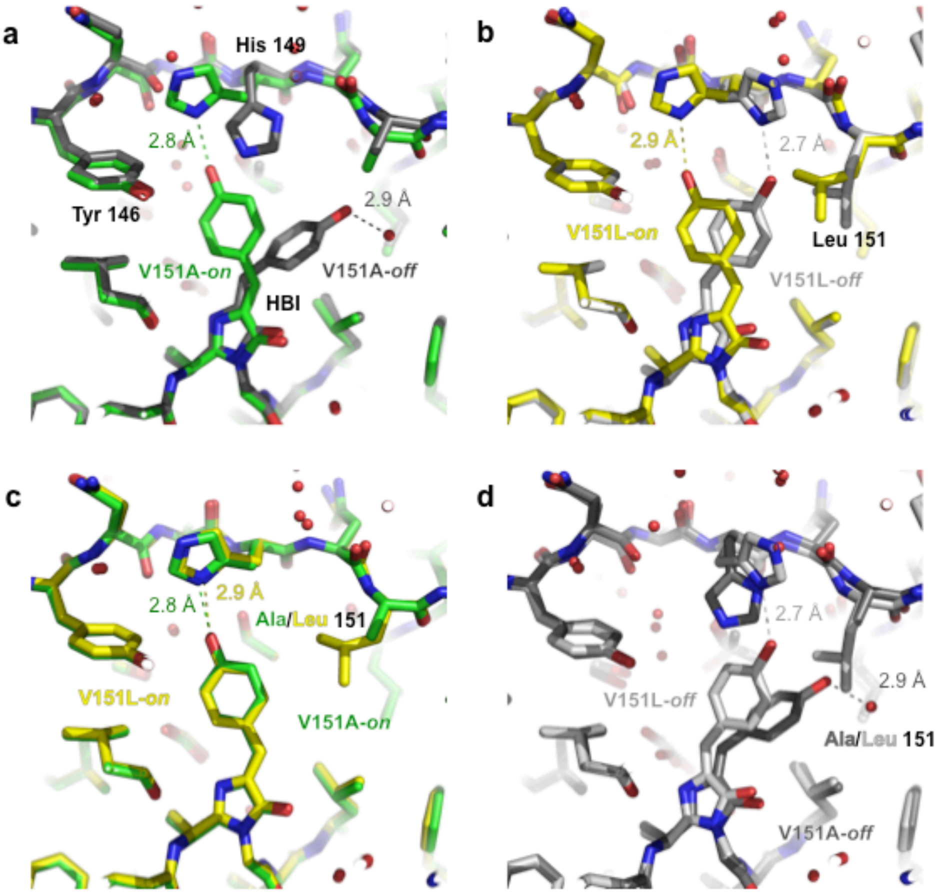


Figure 5 : Comparison of *on*- and *off*-state structures in V151A and V151L variants. (a) Comparison of *on*- (green) and *off*-state structures (dark grey) of rsEGFP2 V151A. (b) Comparison of *on*- (yellow) and *off*-state structures (light grey) of rsEGFP2 V151L. (c) Comparison of *on*-state structures of V151A (green) and V151L variants (yellow). (d) Comparison of *off*-state structures of V151A (dark grey) and rsEGFP2 V151L variants (light grey). The *cis* conformer was removed from rsEGFP2 V151L in the *off*-state for clarity.

References

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D., 2012. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. 68, 352–367. https://doi.org/10.1107/S0907444912001308

Coquelle, N., Sliwa, M., Woodhouse, J., Schirò, G., Adam, V., Aquila, A., Barends, T.R.M., Boutet, S., Byrdin, M., Carbajo, S., De la Mora, E., Doak, R.B., Feliks, M., Fieschi, F., Foucar, L., Guillon, V., Hilpert, M., Hunter, M.S., Jakobs, S., Koglin, J.E., Kovacsova, G., Lane, T.J., Lévy, B., Liang, M., Nass, K., Ridard, J., Robinson, J.S., Roome, C.M., Ruckebusch, C., Seaberg, M., Thepaut, M., Cammarata, M., Demachy, I., Field, M., Shoeman, R.L., Bourgeois, D., Colletier, J.-P., Schlichting, I., Weik, M., 2018. Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. Nat. Chem. 10, 31–37. https://doi.org/10.1038/nchem.2853

DeLano, W., 2002. The PyMOL Molecular Graphics System (2002) DeLano Scientific, Palo Alto, CA, USA. http://www.pymol.org.

El Khatib, M., Martins, A., Bourgeois, D., Colletier, J.-P., Adam, V., 2016. Rational design of ultrastable and reversibly photoswitchable fluorescent proteins for super-resolution imaging of the bacterial periplasm. Sci. Rep. 6, 18459. https://doi.org/10.1038/srep18459

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501. https://doi.org/10.1107/S0907444910007493

Kabsch, W., 2010. XDS. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132. https://doi.org/10.1107/S0907444909047337

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674. https://doi.org/10.1107/S0021889807021206

6. Chapter V: Investigating the *off*-state heterogeneity of rsEGFP2

Introduction

The present chapter is a direct continuation of the previous one and aims at linking various observations reported in the previous chapters. As a reminder, in chapter III, the *laser-off* structure of rsEGFP2 revealed an unexpected second *trans* isomer populated at 20% in one of the TR-SFX experiments (Figure 3 of chapter III) but not in a similar one (Figures S17 and S18 of Coquelle et al., 2018 reported in chapter II). However, no explanation was provided for its observation. In chapter II, a movement of V151 after formation of the excited electronic state during photoswitching was demonstrated. Variants of rsEGFP2 (V151A and V151L) were subsequently generated and characterized, and the *off*- and *on*-state structures of the variants were solved and discussed in chapter IV (Figures 3-5 of chapter IV). If the *on*-state structures of rsEGFP2 and its variants are similar, the *off*-state structures are different, in particular for the V151L variant, in which the *off*-state displays the chromophore in a different conformation than in the V151A variant.

Here is reported a second serendipitous observation of the conformational heterogeneity in the rsEGFP2 *off*-state structure after 488-nm-illumination. Difference Fourier electron density maps are shown and confronted with the *laser-off* model of rsEGFP2 in which this conformational heterogeneity was observed first (Figure 3 of chapter III). Furthermore, the model of the heterogeneous *off*-state of rsEGFP2 is compared to the *off*-state model of its variants V151A and V151L. This comparison reveals that the conformational heterogeneity of rsEGFP2 is controlled by the residue in position 151 and that it can be bisected in the two variants. Our conclusions reinforce the prospect of a structure-based rational design of fluorescent proteins to improve their photophysical and structural characteristics.

Results

Confirming off-state heterogeneity in rsEGFP2 after 488-nm-illumination

During a TR-SFX experiment at the LCLS CXI instrument (proposal number LM47, June 2016), rsEGFP2 microcrystals were pre-illuminated with 488-nm laser light before injection (Schirò et al., 2017; see manuscript in appendix D). As during earlier TR-SFX experiments (*i.e.* LI56 at the LCLS (Coquelle et al., 2018)) and 2015A8031 at SACLA (chapter III, Woodhouse et al., under revision) a pre-illumination device was used to the characterization of which I marginally contributed (Schirò et al., 2017; see manuscript in appendix D). During LM47, it became evident spectroscopically that pre-illumination was not fully efficient, and that 30% of crystalline rsEGFP2 remained in the *on*-state (Figure 1; as a reminder, only 10% remained in the *on*-state after photoswitching during LI56 and SACLA2015A8031). We noticed after LM47 that the SMA fiber within the HPLC tee union of the pre-illumination device (see figure 2 in Schirò et al., 2017) was loosened, leading to a 50% drop in laser power at the sample position (unpublished). $2F_{obs}^{laser\_off\_LM47} - F_{calc}^{laser\_off\_LI56}$ and $F_{obs}^{laser\_off\_LM47} - F_{calc}^{laser\_off\_LI56}$ maps, calculated with the *laser_off* data set collected during LM47 and the phases of the *laser_off* structure (published in Coquelle et al., 2018) features negative density on the *trans1* chromophore (Figure 2a). The positive density that appears in the $F_{obs}^{laser\_off\_LM47} - F_{calc}^{laser\_off\_LI56}$ map looks different from a *cis* isomer of the chromophore. That pre-illumination was inefficient during LM47 was structurally evident from a difference Fourier map calculated between *laser_off* data sets collected during LI56 and LM47, *i.e.* $F_{obs}^{laser\_off\_LM47}$ - $F_{obs}^{laser\_off\_LI56}$ that features negative and positive density on the *cis* and the *trans1* chromophore, respectively (Figure 2b). The maps of figure 2 indicate that the lack of efficiency of the pre-illumination leads to the appearance of another conformation of the chromophore.

To explore structural changes generated by pre-illumination during LM47, a q-weighted (qw) difference Fourier map was calculated between the laser_off_LM47 data set and the on_state_reference_LI56 data set collected during the LCLS beamtime LI56, phased with the *on*-state model (pdb entry: 5O89, Coquelle et al., 2018), *i.e.* $F_{obs}^{laser\_off\_LM47} - F_{obs}^{on\_state\_state\_reference}$ (Figure 3a). This map features strong negative (- 8.4 σ) and positive (6.2 σ) peaks at the chromophore. Couples of negative and positive peaks are present all along the β7–strand and especially around the side chains of key residues in the vicinity of the chromophore, such as Tyr149 (- 4.9 σ/3.7 σ), Val151 (- 5.5 σ/3.7 σ), and His149, for which a relatively strong negative peak (-8.4 σ) seems to be associated with two positive peaks (4.7

and 5.3 σ). These features taken together are evidence for chromophore isomerisation and concomitant motions of the residues in the chromophore pocket as a result of *on*-to-*off* photoswitching during pre-illumination that occurs at least partially. Intriguing are the peculiar shape of the positive electron density associated with the chromophore and the fact that two positive peaks seem to be assignable to His149. Figure 3b displays the same qw-$F_{obs}^{laser\_off\_LM47} - F_{obs}^{on\_state\_state\_reference}$ map, overlaid with the *laser-off* model determined from SACLA2015A8031 data described in chapter III (PDB entry: 6GMJ, from which the *cis* isomer has been omitted). Strikingly, the chromophore in the *trans1* and *trans2* conformations fits into the positive electron density peak. Moreover, His149 in its *off* and *supp* conformations (as named in chapter III) explain the two positive peaks associated to this residue in the map. The Fourier difference map after 488-nm pre-illumination thus reveals two conformations of the protein in the *off*-state, as has already been observed before (see chapter III). In the following, the heterogeneous *off*-state model of rsEGFP2 is compared to the *off*-state structure of its two variants V151A and V151L (Figure 4). In the remainder of the chapter, the protein conformation associated to the *trans1* and *trans2* isomer will be named *off1* and *off2* conformation of rsEGFP2.

Comparison between off-state chromophore conformations in rsEGFP2 and its variant V151A

Figure 4a shows an overlay of the *trans* chromophore conformations observed in the *off*-state models of rsEGFP2 (see chapter III) and of rsEGFP2 V151A. The *trans1* chromophore conformation in rsEGFP2 (light grey) is almost superimposable to the *off*-state conformation of V151A (cyan). Strikingly, the Tyr146 conformation in the V151A *off*-state superimposes with the one corresponding to the *off2* and not the *off1* conformation of rsEGFP2. In rsEGFP2, the *trans1* chromophore forms a hydrogen bond with His149 (2.8 Å), whereas in the *off*-state of V151A, the phenol of Tyr149 is too far from His149 to establish a hydrogen bond (4.9 Å). Another difference is the position of His149 itself, which is not perfectly superimposable in V151A and in the conformation corresponding to the *trans1* chromophore. An explanation for this mismatch could be the wider space in the V151A variant because of the substitution near His149.

Comparison between off-state chromophore conformations in rsEGFP2 and its variant V151L

Figure 4b shows an overlay of the two *trans* conformations observed in the rsEGFP2 *off*-state (see chapter III) and the *off*-state structure of the V151L variant. Unlike V151A, the chromophore conformation in the V151L *off*-state (magenta) is almost superimposable to the *off2* conformation in rsEGFP2 (dark grey). His149 adopts similar conformations in V151L and in the rsEGFP2 *off*-state with the *trans2* chromophore. In *off2*, however, His149 is further from the phenol of the chromophore (3.0 Å) than in the *off*-state of V151L (2.7 Å). The steric hindrance of the Val151 side chain in rsEGFP2 could explain this difference. It both cases, the distance between the imidazole of His149 and the phenol of the chromophore allows hydrogen bonding.

Discussion

The difference Fourier map (Figure 3) calculated between the *laser_off_LM47* data set collected during the LM47 beamtime at the LCLS of the *off*-state and the *on_state_reference_LI56* data set collected during the LCLS beamtime LI56 of the *on*-state (Coquelle et al., 2018) revealed again the second chromophore conformation after 488-nm illumination. It is the second observation of this heterogeneity. Indeed, in chapter III, the *laser_off* model already displayed the chromophore in two *trans* conformations and it was proposed that *off2* could correspond to an alternate *off*-state. The similarity between *off2* and *off1* and the *off*-states of V151L and V151A, respectively, reinforces the notion of an alternate *off*-state. The assumptions that the *off1* and *off2* conformations of rsEGFP2 actually correspond to the *off*-states of its variants V151A and V151L, respectively, and that they can both be generated by 488-nm illumination could explain why rsEGFP2 adopts a behavior which appears as a mixture of the behaviors of its two variants, as suggested by the spectroscopic characterization reported in the previous chapter.

In chapter II (Coquelle et al., 2018) rose the idea that V151 in rsEGFP2 generates steric hindrance in the excited state, acting as a gate during the *trans*-to-*cis* isomerisation of *off*-to-*on* photoswitching. From figure 2 of the present chapter, it appears that V151 is also at the origin of the observed *off*-state heterogeneity after *on*-to-*off* switching in rsEGFP2. Indeed, if the valine in position 151 is replaced by an alanine (shorter side chain) and a leucine (longer side chain), a chromophore isomer similar to *trans1* and *trans2* is generated after 488-nm illumination, respectively. Thus, the intermediate side chain length of a valine in rsEGFP2 allows both *trans* conformations to be formed during *on*-to-*off* switching.

The reason why the second *off*-state conformation, *i.e. off2*, has not been observed before in the rsEGFP2 *off*-state structure determined by traditional cryo-crystallography at a synchrotron (PDB entry: 5DTY, El Khatib et al., 2016) is not clear. A possible explanation is that cryo-cooling altered the conformational energy landscape (Fraser et al., 2011) in a way that the *trans2* conformation disappeared and only *trans1* remained. The *trans2* conformation was not reported in the *off*-state structure of rsEGFP2 solved by room-temperature SFX in Coquelle et al., 2018 (see chapter II) either, but was observed in two subsequent similar SFX experiments (present chapter and chapter III) using the same pre-illumination device (Schirò et al., 2017, see appendix). Since the *trans2* conformation in LM47 was more populated than in SACLA2015A8031 and even more than in LI56 (Coquelle et al., 2018), we suspect, but cannot prove, that the degrading pre-illumination efficiency (see Figure 1 in this chapter) is responsible for the formation of *off2* in RT SFX experiments.

Assuming that a lower 488-nm pre-illumination efficiency is indeed related to the appearance of the *trans2* chromophore conformation, and assuming that the *trans1* and *trans2* chromophore conformations correspond to those adopted in the *off*-states of V151A and V151L, respectively, the following thought experiment can be proposed based on the photophysical properties of rsEGFP2 and its two variants (see Table 1 in chapter IV). Under 488-nm light illumination of the *on*-state (Figure 5a), *cis*-to-*trans* isomerisation occurs in a fraction of rsEGFP2 molecules (e.g. in 1/3) and *trans1* and *trans2* are generated at e.g. equal occupancies (Figure 5b). Because of its lower switching contrast determined *via* the V151L *off*-state (see Figure 2 and Table 1 in chapter IV), *trans2* will more likely switch back to the *cis* conformation if 488-nm illumination is maintained and the newly-generated *cis* conformation will isomerize again and generate *trans1* and *trans2* at the supposed 50/50 ratio (Figure 5c). For simplicity, we assume in Figure 5c that only *trans2* but not *trans1* responds to 488-nm light illumination. The more often the *on*-to-*off* and *off*-to-*on* switching cycles are repeated, the more populated is the *trans1* and the less populated the *trans2* chromophore conformation (Figure 5d).

The tentative explanation for the appearance of a second *off2* conformation is based on two assumptions, *i.e.* that the pre-illumination efficiency was degraded and that the features observed in the difference Fourier map between the laser_off_LM47 data set and the on_state_reference_LI56 experiment correspond to the heterogeneous *off*-state modeled in chapter III. The first assumption could be tested by solving rsEGFP2 structures after varying the 488-nm illumination (either the laser power or the exposure time) and assessing the relative populations of *on*-, *off1*- and *off2*-states. Such a series of experiments was undertaken, in which cryo-crystallographic data were collected on a synchrotron beamline after room-temperature illumination (unpublished data). *Cis* and *trans1* isomers were observed at a ratio depending on illumination conditions, but none of the crystal structures showed a *trans2* isomer. This experiment could be repeated in an XFEL experiment, using the same pre-illumination setup than described in Schirò et al., 2017 operated at various laser powers. To test the second assumption, extrapolated structure factors (Genick, 2007) could be calculated from $F_{obs}^{laser\_off\_LM47} - F_{obs}^{on\_state\_state\_reference}$ and the difference refinement procedure could be applied to extract only the heterogeneous *off*-state, to model it and to compare it to the *laser_off* structure reported in chapter III. Such a difference refinement is currently in progress.

As a perspective, and based on the hypothesis that *off1* and *off2* of rsEGFP2 correspond to V151A and V151L *off*-state, respectively, these two variants could be studied by TR-SFX on similar time-scales than for rsEGFP2 (Chapter II and III) with the fundamental aim to understand the photophysical behavior of rsEGFP2 and the more practical goal to characterize rationally-designed variants. Such TR-SFX experiments have been carried out (at the LCLS (LR38), 22 – 26 February 2018 and at SACLA (2018A8026, 27 – 29 July 2018) and are currently being analysed.

Material and methods

Expression, purification and microcrystallization

rsEGFP2 was expressed and purified as previously described (El Khatib et al., 2016). Microcrystals were obtained as reported (Coquelle et al., 2018, chapter II), following the protocol described in chapter I.

Injection, data collection and processing

Sample injection, data collection and processing of the *on*-state dataset were already described in chapter II (Coquelle et al., 2018). rsEGFP2 microcrystals for collection of a renewed *laser_off* data set were prepared for injection following roughly the same procedure than reported in chapter II. Briefly, a microcrystal suspension was centrifuged in a 15-ml falcon for 5 min at 1000 g. The mother liquor (~2 M ammonium sulfate, 20 mM NaCl, 120 mM HEPES pH 8) was removed and the pellet was resuspended in 2.5 M ammonium sulfate, 100 mM HEPES pH 8. In order to increase the hit rate during data collection (limited to 8 % in LI56 and SACLA20158031, see Discussion section in chapter I), fragments of crushed crystals, so-called *sand*, were added to the suspension of homogeneously size microcrystals to avoid their aggregation. The fragments of crushed crystals were prepared analogously to seeds (see chapter I) but without filtering. By adding 5 – 10 % (v/v) of *sand*, it was observed that the crystal suspension could be injected at a concentration of 20-30% (v/v), allowing to drastically increase the hit-rate and thus reducing the time required to collect a complete data set with sufficient quality. SFX data were collected in the microfocus vacuum chamber of the CXI instrument at the LCLS (proposal LM47, 23 – 27 June 2016) operating at 120 Hz, with X-ray pulses of 35 fs in length and a photon energy of nominally 9.1 keV. The X-ray beam was focused to 1 – 2 μm.

Difference electron density map calculation

A q-weighted difference Fourier map (Ursby and Bourgeois, 1997), $F_{obs}^{laser\_off\_LM47}$ - $F_{obs}^{on\_state\_reference\_LI56}$, was calculated using CNS (Brünger et al., 1998) between the *laser_off_LM47* data set collected during the LM47 beamtime at the LCLS of the *off*-state and the *on_state_reference_LI56* data set collected during the LCLS beamtime LI56 of the *on*-state, phased with the *on*-state model (pdb entry 5O89, Coquelle et al., 2018). Figures were prepared using PyMOL (DeLano, 2002).

Figures and tables

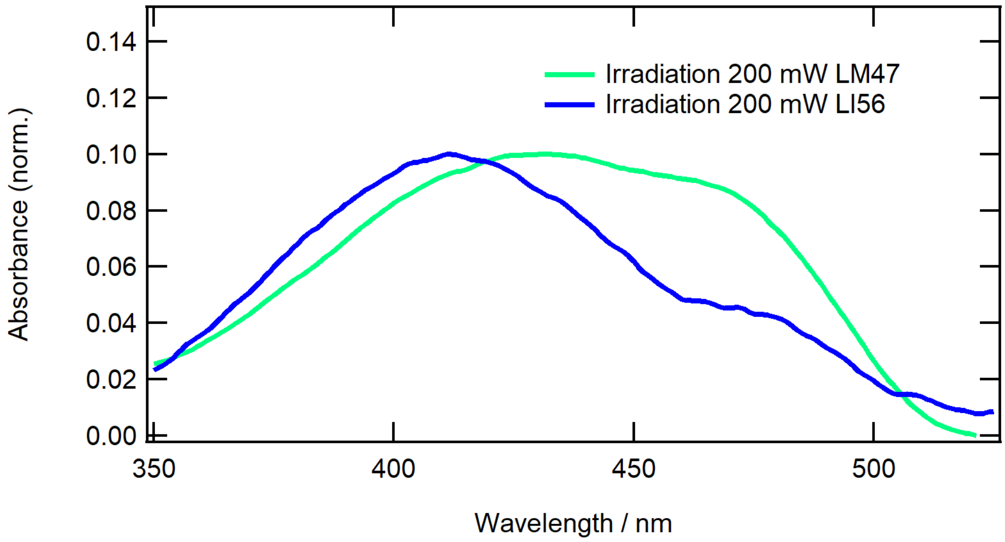


Figure 1: Absorption spectra of rsEGFP2 microcrystals (3 × 3 × 3 µm) after having been pre-illuminated at a laser power of nominally 200 mW within a custom-made device (Schirò et al., 2017) during LCLS experiments LM47 (green; June 2016) and LI56 (blue; May 2015). Whereas during LI56 90% of rsEGFP2 were switched to the *off*-state and only 10% remained in the *on*-state, the efficiency decreased during LM47 where 70% were switched to the *off*-state and 30% remained in the *on*-state.

Dataset	laser_off
Pre-illumination (488 nm)	yes
Space group	$P2_12_12_1$
Unit cell parameters a (Å) b (Å) c (Å)	 51.1 62.4 70.7
Resolution (Å)	8.32 – 1.7 (1.72 – 1.7)
R_{split} (%)	5.69 (59.48)
CC*	0.99 (0.84)
I / σ(I)	13.20 (2.78)
Completeness (%)	100 (100)

Table 1: Data collection statistics calculated after removal of overloads.

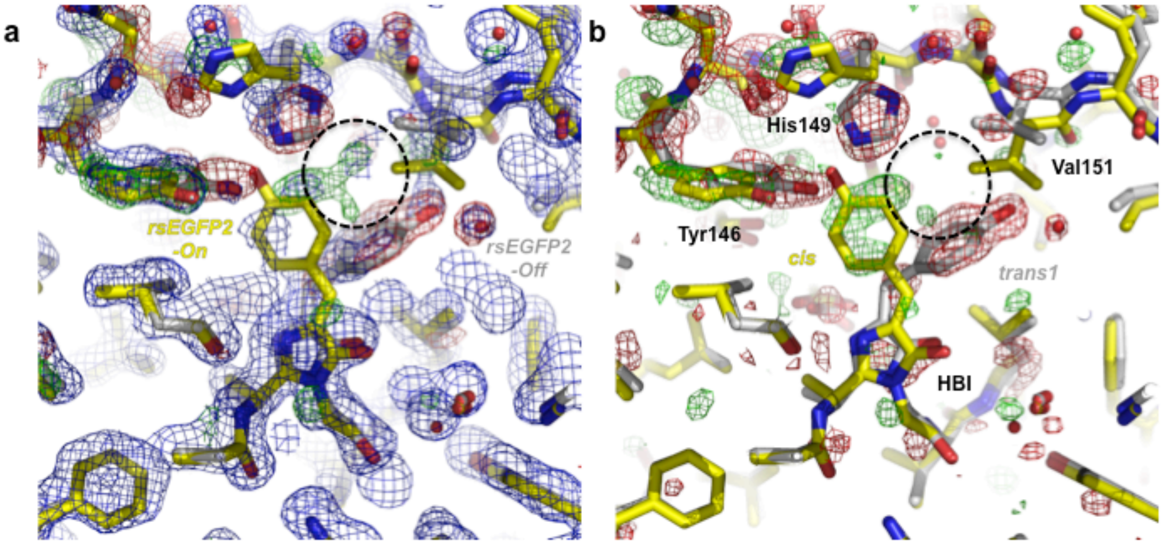


Figure 2: Structural evidence for the lack of efficiency of the 488-nm pre-illumination. (a) $2F_{obs}^{laser\_off\_LM47} - F_{calc}^{laser\_off\_LI56}$ (blue) and $F_{obs}^{laser\_off\_LM47} - F_{calc}^{laser\_off\_LI56}$ (green/red) displayed at 1σ and +/- 3σ respectively. (b) $F_{obs}^{laser\_off\_LM47} - F_{obs}^{laser\_off\_LI56}$ (green/red) displayed at +/- 3σ. Model of the *laser_off_LI56* (as a reminder 90 % *trans* (light grey), 10 % *cis* (yellow), PDB entry: 5O8A, Coquelle et al., 2018) was used for phasing. The dotted black circle points out intriguing positive difference electron density.

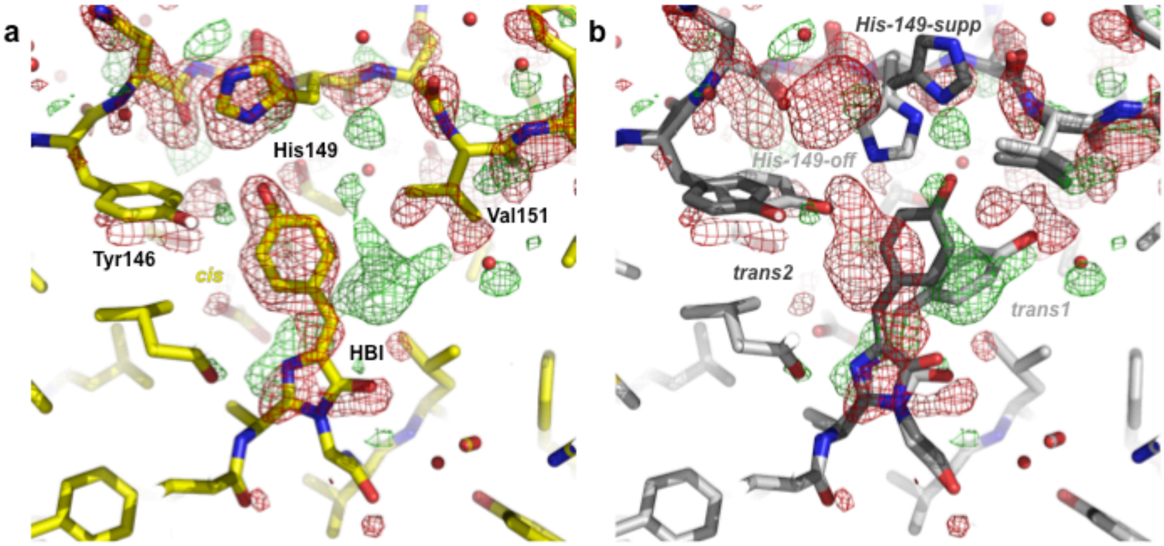


Figure 3 : Investigating the heterogeneity of the *off*-state of rsEGFP2. (*a, b*) A difference Fourier map $F_{obs}^{laser\_off\_LM47}$ - $F_{obs}^{on\_state\_reference\_LI56}$ map calculated between the *laser_off_LM47* data set collected during the LM47 beamtime at the LCLS of the *off*-state and the *on_state_reference_LI56* data set collected during the LCLS beamtime LI56 of the *on*-state, phased with the *on*-state model (yellow, PDB code 5O89, Coquelle et al., 2018) is displayed at. + / - 3 σ (green / red). In (*a*) is shown the *on-state* model and in (b) the *laser-off* model solved from data collected during SACLA2015A8031 (PDB entry: 6GMJ (Woodhouse et al. under revision, chapter III)). The *trans1* and *trans2* conformations are displayed in light and dark grey, respectively. The *cis* conformer was removed for clarity from the *laser-off* model.

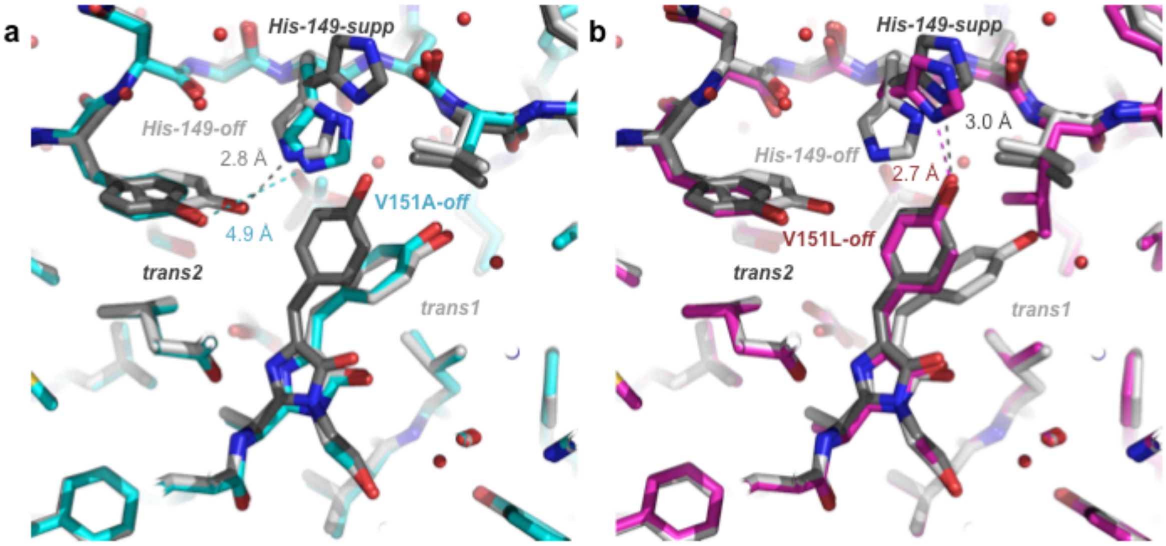


Figure 4: Comparing the heterogeneous *off*-state structures of rsEGFP2 (PDB entry: 6GMJ (Woodhouse et al. under revision, chapter III)) with the *off*-state structures of its variants V151A (a) and V151L (b). The rsEGFP2 conformations corresponding to the *trans1* and *trans2* chromophores are displayed in light and dark grey, respectively. The residual *cis* conformer present in the model was removed for clarity. The *off*-state structures of V151A and V151L are displayed in cyan and magenta, respectively.

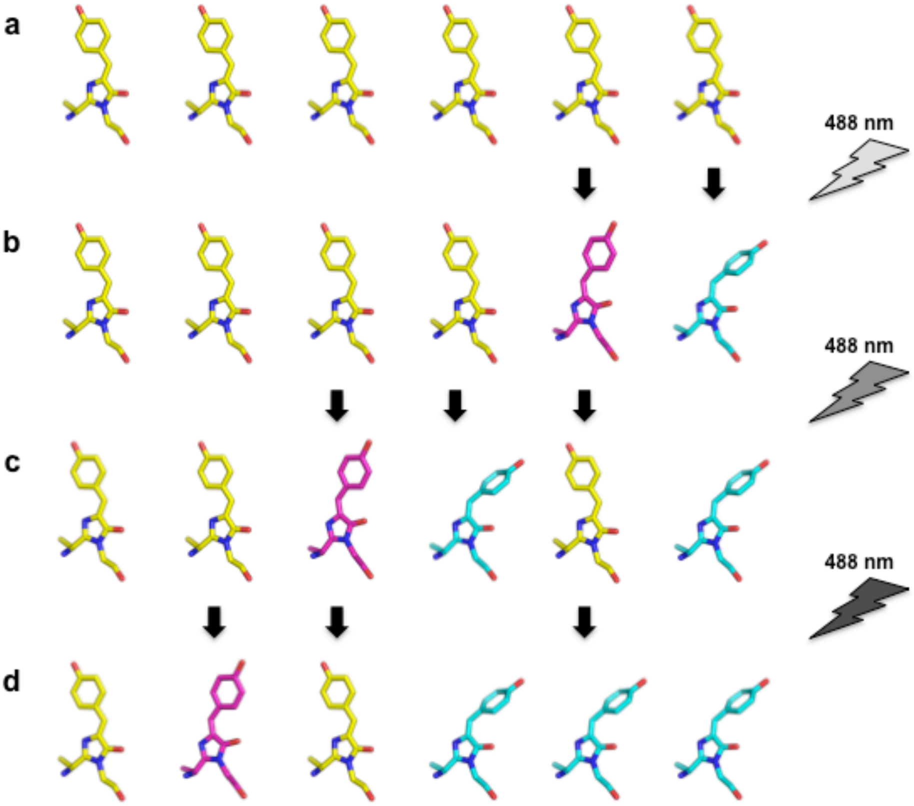


Figure 5: Proposed thought experiment explaining the increasing ratio of *trans1* and *trans2* populations under continuous 488-nm illumination. The grey-scale (from light to dark) of the lightning symbol represents the total accumulated number of photons during illumination. (*a*) Starting point where it is assumed that every molecule is in the fluorescent *on*-state with a *cis* chromophore (yellow). (*b*) Evolution of the ensemble of molecules after illumination. In one third of the molecules, the chromophore has isomerized. For simplicity, it is assumed that the *trans1* (cyan) and *trans2* (magenta) isomers are formed in equal proportions. (*c*) Evolution of the ensemble under maintained 488-nm illumination. The chromophores in two rsEGFP2 molecules that had remained in the *on*-state now isomerized to *trans1* and *trans2* isomers. The *trans2* previously generated in panel (*b*) isomerized back to the *cis* isomer. Again for the sake of simplicity, it is assumed that under 488-nm illumination *trans1* (=V151A-*off*-state) never isomerizes back to *cis* and that *trans2* (=V151L-*off*-state) always isomerizes back to *cis*, representing the improved and degraded contrast of V151A and V151L compared to rsEGFP2, respectively. (*d*) Further evolution upon continued 488-nm illumination. After several more cycles, *i.e.* more accumulated photons, *trans1* would eventually be the only isomer and *trans2* would be fully depleted.

Reference :

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., Warren, G.L., 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, 54, 905–921.

Coquelle, N., Sliwa, M., Woodhouse, J., Schirò, G., Adam, V., Aquila, A., Barends, T.R.M., Boutet, S., Byrdin, M., Carbajo, S., De la Mora, E., Doak, R.B., Feliks, M., Fieschi, F., Foucar, L., Guillon, V., Hilpert, M., Hunter, M.S., Jakobs, S., Koglin, J.E., Kovacsova, G., Lane, T.J., Lévy, B., Liang, M., Nass, K., Ridard, J., Robinson, J.S., Roome, C.M., Ruckebusch, C., Seaberg, M., Thepaut, M., Cammarata, M., Demachy, I., Field, M., Shoeman, R.L., Bourgeois, D., Colletier, J.-P., Schlichting, I., Weik, M., 2018. Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nat. Chem.*, 10, 31–37. https://doi.org/10.1038/nchem.2853

DeLano, W., 2002. The PyMOL Molecular Graphics System (2002) DeLano Scientific, Palo Alto, CA, USA. http://www.pymol.org.

El Khatib, M., Martins, A., Bourgeois, D., Colletier, J.-P., Adam, V., 2016. Rational design of ultrastable and reversibly photoswitchable fluorescent proteins for super-resolution imaging of the bacterial periplasm. *Sci. Rep.*, 6, 18459. https://doi.org/10.1038/srep18459

Fraser, J.S., van den Bedem, H., Samelson, A.J., Lang, P.T., Holton, J.M., Echols, N., Alber, T., 2011. Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proc. Natl. Acad. Sci. U. S. A.*, 108, 16247–16252. https://doi.org/10.1073/pnas.1111325108

Genick, U.K., 2007. Structure-factor extrapolation using the scalar approximation: theory, applications and limitations. *Acta Crystallogr. D Biol. Crystallogr.*, 63, 1029–1041. https://doi.org/10.1107/S0907444907038164

Schirò, G., Woodhouse, J., Weik, M., Schlichting, I., Shoeman, R.L., 2017. Simple and efficient system for photoconverting light-sensitive proteins in serial crystallography experiments. *J. Appl. Crystallogr.*, 50, 932–939. https://doi.org/10.1107/S1600576717006264

Ursby, T., Bourgeois, D., 1997. Improved Estimation of Structure-Factor Difference Amplitudesfrom Poorly Accurate Data. *Acta Crystallogr. A*, 53, 564–575. https://doi.org/10.1107/S0108767397004522

7. Summary, concluding remarks and perspectives

The results presented in this thesis illustrate the pipeline followed for several TR-SFX experiments performed on rsEGFP2, in order to structurally elucidate the photoswitching mechanism of the protein. This achievement has been possible thanks to the use of XFELs and the use of XFEL-based method. As outlined above, sample preparation is a critical step for a successful TR-SFX experiment, but only a handful of methods and protocols have been reported. Here, a microcrystallization protocol for rsEGFP2 based on the batch method is presented that uses microseeding in order to reproducibly obtain crystals that are homogeneous in size and shape. The preparation of seeds and microcrystals is described. Even though this particular protocol is suitable only for rsEGFP2, the general approach consisting in exploring the phase diagram of the protein for a chosen condition, in using microseeding to improve crystallization, and in varying different parameters such as the amount of seeds, protein and precipitant concentration, is valuable for any protein of interest in the context of an SFX experiment. The encountered issues arising from manipulating such crystals are reported and clues are proposed how to fix them. For example, injector clogging, which limits crystal concentration and thus the hit-rate, has been fixed by adding *sand* to the crystal suspension. Thanks to the solid sample preparation protocol, several successful TR-SFX experiments have been carried out at both the LCLS and SACLA, allowing structure determination of reaction intermediates that unveil the photoswitching mechanism of rsEGFP2.

Investigations on the ultra-short time-scale, 1 ps after photon absorption, allowed structure determination of an early photoswitching intermediate carrying a “twisted” chromophore. Corroborated by transient absorption spectroscopy and predicted by QM/MM simulations, the “twisted” chromophore is identified as an excited state intermediate. TR-SFX data collected at 3 ps did not allow structure determination, but provided evidence that isomerization already occurred, revealing that *trans/cis* isomerization is an excited-state process. TR-SFX on a longer time-scale allowed trapping the first photoproduct of *off*-to-*on* photoswitching. The structure solved at 10 ns after 400-nm laser pumping shows a *cis* chromophore. Transient absorption spectroscopy reveals that the change in chromophore protonation state occurs on the μs time-scale, proving that the 10-ns intermediate is still protonated. The 10-ns structure informs about alternate conformations of His149 and raises questions to be answered in future TR-SFX experiments.

Figure 1 illustrates a proposed photocycle of rsEGFP2 based on the structural results presented in this thesis and summarized above. Under 488-nm illumination of the fluorescent *on*-state, photoswitching leads to the formation of a heterogeneous *off*-state, constituted of *off-1* and *off-2* substates, as described in chapter V, with a *trans1* and a *trans2* chromophore, respectively, as modeled in chapter III. Given that *off-2* is not observed in the *laser_off_LI56* structure (chapter II), and is only a minor species observed in the *laser_off_SACLA* structure (chapter III), it is assumed here that *off*-1 is the main substate from which are generated the photo-induced intermediates at 1 ps and 10 ns. The 1-ps intermediate displays a “twisted” chromophore attributed to an excited state based on transient absorption spectroscopy and QM/MM simulations. Evidence is provided in chapter II that isomerization occurs on the picosecond time-scale, in line with the most accepted mechanism in which isomerization is an excited-state process (Kaucikas et al., 2015). The 10-ns intermediate displays a *cis* conformation of the chromophore, similar to the *on*-state, but with alternate conformations of the His149 side chain. Transient absorption spectroscopy provides evidence that at 10 ns, the chromophore is still protonated, in line with the most accepted mechanism in which isomerization precedes ground-state deprotonation during photo-induced *off*-to-*on* switching (Warren et al., 2013, Yadav et al., 2015). Regarding the two conformations of His149 at 10 ns, two scenarios can be envisaged. In the first, the alternate conformation of His149 is explained by a certain disorder being the consequence of the hydrogen bound between the phenol of the chromophore and the imidazole of His149 that is missing because of the neutral state of the former. In the second, the presence of alternate conformations of His149 at 10 ns could be a consequence of the *off*-state heterogeneity as a starting point.

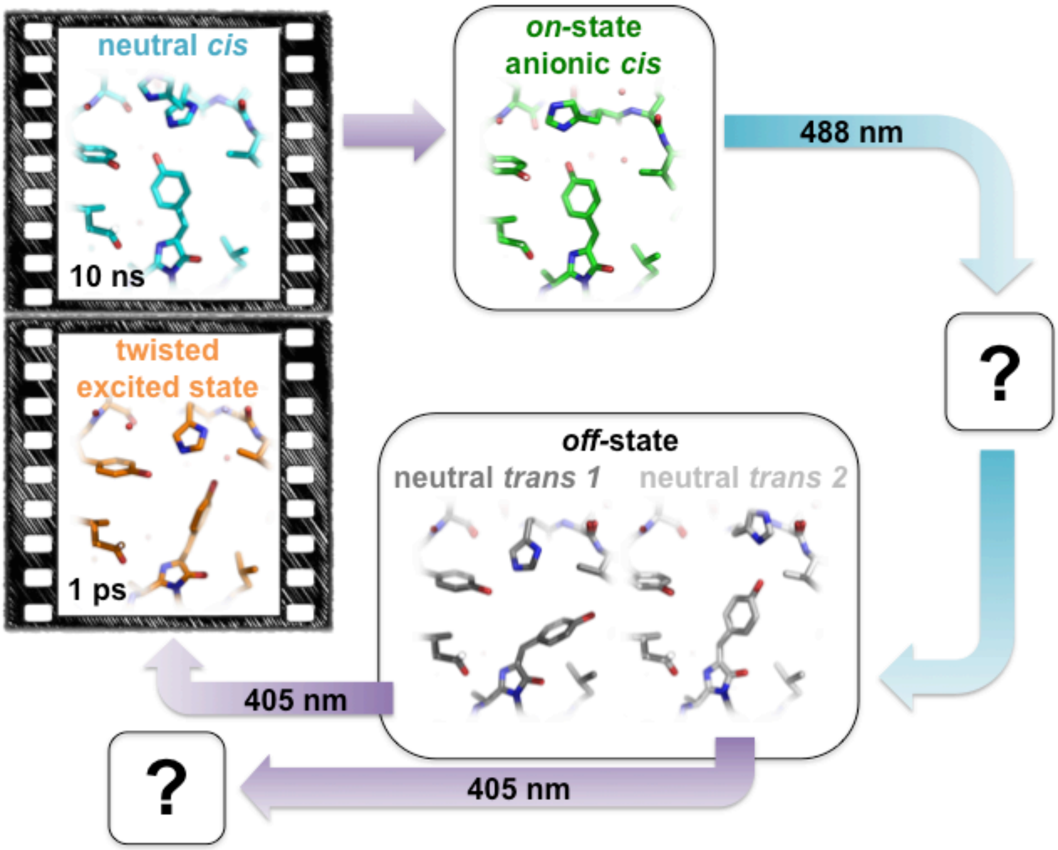


Figure 1: Proposed photocycle of rsEGFP2 based on the structural observations presented in this PhD thesis. Under 488-nm illumination, *on*-to-*off* switching occurs, from the *on*-state with its anionic *cis* chromophore (green) to a heterogeneous *off*-state displaying neutral *trans1* (*off-1* shown in dark grey) and *trans 2* (*off-2* shown in light grey) chromophores. Under 400-nm illumination, the *off-1*-state, undergoes *off-1*-to-*on* switching. The 1-ps intermediate displays a twisted chromophore (orange) in the excited state. The 10-ns intermediate (cyan) carries a neutral *cis* chromophore and revealed an alternate conformation of the His149. *On*-to-*off* switching intermediates and the *off-2*-to-*on* switching mechanism remain unknown, as illustrated by question marks.

The photoswitching mechanism of Dronpa and Dronpa2 has been extensively studied and starts to be well understood. It consists in exited-state isomerization on the picosecond time-scale, followed by ground-state deprotonation on the microsecond time-scale (Kaucikas et al., 2015; Warren et al., 2013; Yadav et al., 2015). The ensemble of results presented in this PhD thesis is in line with this proposed mechanism. However, it has been recently reported that instead of being an excited-state process, isomerization occurs rather in the ground state with conformational changes extending to the nanosecond time-scale (Laptenok et al., 2018). As illustrated by the 10-ns intermediate structure presented in the present manuscript, conformational rearrangements take place on the nanosecond time-scale, but the accepted view that isomerization already occurred is not questioned. However, one should not forget the fact that RSFPs from different families are compared. The chromophores of Dronpa and rsEGFP2, as well as the residues in the chromophore pocket are different. One cannot assume *per se* that the switching mechanism is the same for these two RSFPs .

Several questions remain to be answered. For instance, the fate during photoswitching of *off*-2 is unclear. As a perspective, and assuming that *off*-2 corresponds to the *off*-state of the V151L variant as discussed in chapter V, this variant could be investigated by TR-SFX in order to test the hypothesis that the *off*-state heterogeneity in rsEGFP2 is correlated with the alternate conformations of His149 at 10 ns. Such an experiment is very challenging, since V151L is a *bad* switcher compared to its parent protein, making efficient pre-illumination and pumping difficult. Such an experiment has been carried out on the nanosecond time-scale recently at SACLA (SACLA20188026; 27 – 29 July 2018), using the same set-up than reported in chapter III. Results are currently being examined. Assuming that *off*-1 corresponds to the *off*-state of the V151A variant as discussed in chapter V, this variant is also of great interest. As a relatively *good* switcher compared to its parent and to V151L, it is easier to handle for such a TR-SFX experiment and promises to generate clearer features in Fourier difference electron density maps. TR-SFX experiment on rsEGFP2 V151A were already performed on the picosecond (LCLS, LR38, February 2018) and the nanosecond time scales (SACLA20188026; 27 – 29 July 2018). Since we hypothesize that rsEGFP2's properties are a mixture of its variant's (V151A and V151L) properties, unveiling the photoswitching mechanism of the variants would lead to a better comprehension of the original protein and bears potential for applications in nanoscopy. Based on the photocycle proposed in figure 1, another challenging question remains: what happens structurally during *on*-to-*off* switching ? The *on*-to-*off* switching quantum yield is ten times lower than the one for *off*-to-*on* switching.

Since it is reported in chapter II that even *off*-to-*on* switching intermediates are occupied at a very low level on the ultra-fast time scale, the situation is expected to be worse starting from the *on*-state so that it is probably impossible to structurally characterize *on*-to-*off* switching intermediates. Unveiling the mechanism of fluorescence emission is a better option, since the fluorescence quantum yield is of the same order of magnitude than the *off*-to-*on* switching quantum yield. Excited-state intermediates of the fluorescent *on*-state have been followed structurally over five decades in time (LCLS, LM47, June 2016) and results are currently being analysed.

The work presented here allowed understanding the *off*-to-*on* photoswitching mechanism of rsEGFP2. It is interesting to realize that the rsEGFP2 switching mechanism is similar than the one of Dronpa belonging to another RSFP family. The deeper comprehension of the switching mechanism allows a better use of rsEGFP2 and its variants V151A and V151L in nanoscopy applications. Furthermore, variants other than V151A and V151L have been generated and are currently being characterized and tested as nanoscopy candidates.

References:

Kaucikas, M., Tros, M., van Thor, J.J., 2015. Photoisomerization and proton transfer in the forward and reverse photoswitching of the fast-switching M159T mutant of the Dronpa fluorescent protein. *J. Phys. Chem. B*, 119, 2350–2362. https://doi.org/10.1021/jp506640q

Laptenok, S.P., Gil, A.A., Hall, C.R., Lukacs, A., Iuliano, J.N., Jones, G.A., Greetham, G.M., Donaldson, P., Miyawaki, A., Tonge, P.J., Meech, S.R., 2018. Infrared spectroscopy reveals multi-step multi-timescale photoactivation in the photoconvertible protein archetype dronpa. *Nat. Chem.*, 10, 845–852. https://doi.org/10.1038/s41557-018-0073-0

Warren, M.M., Kaucikas, M., Fitzpatrick, A., Champion, P., Sage, J.T., van Thor, J.J., 2013. Ground-state proton transfer in the photoswitching reactions of the fluorescent protein Dronpa. *Nat. Commun.*, 4, 1461. https://doi.org/10.1038/ncomms2460

Yadav, D., Lacombat, F., Dozova, N., Rappaport, F., Plaza, P., Espagne, A., 2015. Real-time monitoring of chromophore isomerization and deprotonation during the photoactivation of the fluorescent protein Dronpa. *J. Phys. Chem. B*, 119, 2404–2414. https://doi.org/10.1021/jp507094f

Appendix A: Résumé de la thèse en français

Les protéines fluorescentes photo-commutables (RSFPs) ont la propriété de passer d'un état fluorescent à un état non-fluorescent en réponse à la lumière. Cette propriété en fait des outils de marquage pour la microscopie de super-résolution (ou nanoscopie). Le mécanisme de photo-commutation implique l'isomérisation du chromophore ainsi qu'un changement d'état de protonation de ce dernier. Le mécanisme a été très étudié par différentes approches de spectroscopie et de simulation mais reste encore mal compris, l'ordre séquentiel des évènements est notamment encore débattu. Certains de ces évènements de la photo-commutation se déroulent à des échelles de temps très courtes, ce qui rend difficile l'étude structurale par cristallographie des rayons X à l'aide des sources synchrotron actuelles dont la résolution temporelle est encore limitée. Les lasers à électrons libres (XFELs) sont une nouvelle source de rayons X produisant des impulsions suffisamment courtes pour permettre l'étude structurale des intermédiaires précoces ou à courte durée de vie qui se forment ou cours de la photo-commutation, et suffisamment brillantes pour permettre la collecte de données cristallographiques sur des cristaux de tailles nano- et micrométrique. L'utilisation de ce nouveau genre d'instrument a permis l'émergence de la cristallographie sérielle, une nouvelle approche de la cristallographie aux rayons X. Cette approche a depuis été adaptée aux lignes synchrotrons.

Le travail présenté ici se focalise sur l'étude de rsEGFP2, une protéine fluorescente photo-commutable de la famille de la protéine fluorescente verte (GFP). Il y est décrit la mise au point d'un protocole de microcristallisation permettant l'obtention d'échantillons en vue d'une expérience de cristallographie résolue en temps au XFEL. Un mécanisme de photo-commutation y est proposé à travers le résultat de deux expériences sur les deux XFELs actuellement opérationnels, à des échelles de temps différentes, dévoilant un chromophore « twisté » à l'état excité ainsi qu'un état *cis* protoné de ce dernier. La caractérisation structurale des variants de rsEGFP2 par cristallographie d'oscillation « classique » combinée à la découverte fortuite d'une conformation alternée du chromophore dans l'état non-fluorescent, issue d'expérience de cristallographie sérielle, apporte un complément d'explication des propriétés photophysiques de la protéine.

Introduction

Le protéines fluorescentes (FPs) sont des protéines qui absorbent la lumière à une certaine longueur d'onde et en émettent à une autre. La première identifiée et la plus connue est la GFP. Le repliement commun à toutes les FPs est un tonneau de feuillets β, traversé par une hélice α qui porte le chromophore formé de trois acides aminés modifiés lors de la maturation et qui est responsable des propriétés d'absorption et d'émission de la protéine. Les propriétés photophysiques de la protéine dépendent des acides aminés qui constituent le chromophore ainsi que de ceux qui l'entourent. Les FPs se repartissent en trois classes. Les protéines fluorescentes photo-activables (PAFPs) n'émettent pas de lumière à l'état basal. Elles subissent une réaction photo-induite et irréversible pour passer d'un état non-fluorescent à un état fluorescent. Les protéines fluorescentes photo-convertibles (PCFPs) émettent de la lumière à une certaine longueur d'onde à l'état basal et subissent une réaction photo-induite l'amenant à un autre état fluorescent émettant de la lumière à une autre longueur d'onde. Les protéines fluorescentes photo-commutables (RSFPs) peuvent passer réversiblement d'un état fluorescent à un état non-fluorescent. Elles se divisent en deux catégories : les commutateurs positifs, non-fluorescents à l'état basal et les commutateurs négatifs fluorescents à l'état basal. Le mécanisme de photo-commutation implique l'isomérisation du chromophore et un changement d'état de protonation de ce dernier. La protéine fluorescente verte améliorée réversiblement commutable (rsEGFP2) est un commutateur négatif utilisé comme marqueur en microscopie de super-résolution. La compréhension du mécanisme de commutation chez cette protéine (et par extension des autre RSFPs) mènerait à l'élaboration de protéines variantes aux propriétés améliorées ou nouvelles, mieux adaptées aux applications pour lesquelles elles ont été conçues. L'ordre séquentiel de ces évènements et leur nature ont longtemps été débattus et sont encore actuellement très étudiés chez Dronpa et ses variants. Le manque de données structurales concernant les différents intermédiaires qui pourraient se former pendant la photo-commutation constitue un frein à la compréhension du mécanisme du phénomène. La cristallographie résolue en temps, qui permet de résoudre la structure d'intermédiaire formée au cours d'une réaction, est une approche envisageable pour combler ce manque. Les évènements de photo-commutation étant très précoces ou très rapides (de l'ordre de la picoseconde, voir plus court), le synchrotron, limitée en résolution temporelle, n’est plus adapté pour l’étude d’un tel phénomène. D’où la nécessité d’un autre type d’instrument produisant un nouveau genre de rayons-X : les laser à électrons libres (XFELs), capable de produire des impulsions ultra-brilliantes et ultra-courtes, rendant accessible les échelles de temps des évènements précoces de la photo-commutation.

La cristallographie résolue en temps repose sur le principe « pompe-sonde », qui consiste en l’utilisation d’une « pompe » (généralement un laser optique) pour déclencher la réaction dans le cristal de protéine, et d’une « sonde » de rayons-X utilisée pour la collecte de données cristallographiques. Le délai entre la « pompe » et la « sonde » est ajusté pour chaque collecte. Les facteurs de structure correspondant à chaque délai sont récupérés et permettent de détecter les changements de conformations au cours de la réaction et de reconstituer ainsi un film moléculaire. La diffraction de Laue, utilisant un faisceau polychromatique de rayons-X, a initialement été adoptée pour les études résolues en temps au synchrotron à température ambiante. Cependant, les dégâts d’irradiation, exacerbés par la collecte à température ambiante, la possibilité de n’étudier presque exclusivement que des réactions réversibles, la difficulté d’illuminer efficacement un cristal de taille suffisante pour assurer une diffraction à haute résolution, et l’analyse compliquée des images de diffractions obtenus avec un faisceau polychromatique constituent des limitations à la méthodes. Et par-dessus tout, le synchrotron a une limitation temporelle de 100 ps.

Le XFEL est une nouvelle source de rayons-X produisant des impulsions de quelques femtosecondes avec un pic de brillance de dix ordres de grandeur supérieur à ce que le synchrotron peut produire. Le principe « diffraction-avant-destruction » prédit qu’avec de tels impulsions, l’information de diffraction peut être enregistrée avant que l’échantillon ne commence à être détruit, permettant de s’affranchir du problème des dégâts d’irradiation. Le principe a été expérimentalement testé et permet la collecte de données sur des cristaux de taille nano- et micrométrique, injectés en série, et présentés aux rayons-X avec une orientation aléatoire. Cette nouvelle approche de cristallographie sérielle à la femtoseconde (SFX) ouvre la possibilité d’étudier des protéines difficiles à cristalliser, telles que les protéines membranaires, les protéines radiosensibles, telles que les metalloprotéases, et les grands assemblages dont la diffraction est limitée en résolution spatiale au synchrotron. La méthode a très vite évolué. Le développement de nouvelles manières de présenter l’échantillon a permis de limiter la consommation de matériel biologique, l’optimisation des outils de traitements de données a contribué à améliorer la qualité des données, permettent le phasage expérimental d’une part, et réduisant le nombre d’images requis pour un jeu complet de qualité suffisante. L’adaptation de la méthode « sérielle » au synchrotron a permis la d'étendre de la dite méthode, les XFELs étant des instruments difficiles d'accès. La résolution temporelle atteignable au XFEL offre de nouvelles perspectives concernant la cristallographie résolue en temps, permettant d'explorer des échelles de temps jusqu'alors inaccessibles, et donc de s'intéresser aux évènements précoces et aux intermédiaires de courte durée de vie chez les protéines photosensible tels que les RSFPs.

Résultats et discussion

Chapitre I : Microcistallization de rsEGFP2

La préparation des échantillons est une étape critique pour le succès d'une expérience de cristallographie résolue en temps au XFELs. La taille des cristaux est à prendre en considération, ils ne doivent pas excéder une certaine taille de peur, d'une part de limiter l'efficacité de l'illumination avec le laser optique, et d'autre part, de bloquer l'injecteur lors de l'expérience. Un équilibre entre efficacité d'illumination et qualité de diffraction est à trouver pour déterminer la taille idéale en fonction du type d'expérience. Très peu de méthodes de microcristallisation ont été rapportées.

Il est ici proposé un protocole de microcristallisation de la protéine rsEGFP2 en vue d'expériences résolues en temps au XFEL, et de cristallographie sérielle en générale. A partir de condition d'une cristallisation connue, un diagramme de phase est établie, en faisant varier la concentration de protéine et de précipitent. La zone intéressante, zone de nucléation, celle qui permet d'obtenir spontanément des microcristaux, est identifiée et reproduite en grand volume. L'apport du micro-ensemencement (« micro-seeding »), par l'ajout de fragments de cristaux sous-micrométriques permet de contrôler la nucléation et de drastiquement améliorer l'homogénéité en taille et en forme des cristaux. Il est également montré que la quantité de fragments ajoutés à la préparation permet d'ajuster la taille finale des cristaux obtenus : moins il y aura de « graines » ajoutés, plus les cristaux seront de grande taille. La forme des cristaux est également modifiable en faisant varier la concentration de précipitant dans les préparations ensemencées. Pour finir, il est observé que le volume de la préparation ne semble pas affecter l'aspect final des cristaux.

Ainsi, le protocole proposé permet d'obtenir de façon reproductible et en grands volumes des cristaux homogènes en taille et en forme, ces deux aspects étant facilement modifiables en faisant varier quelques paramètres. Quelques ajustements à la « recette » permettent de préparer des échantillons en vue d'un type d'expérience de cristallographie sérielle donné, peu importe la source (synchrotron ou XFEL) ou la méthode de présentation de l’échantillon (injection liquide, milieux visqueux, support solide…). Il est à noter que le protocole présenté n’est valable que pour une protéine donnée, mais que l’approche proposée peut être appliquée de manière exploratoire sur d’autres protéines.

Chapitre II : Chromophore « twisté » à l’état excité d’une protéine fluorescente photo-commutable capturé par cristallographie à la femtoseconde résolue en temps

L’élucidation de la structure d’intermédiaires très précoces, tels que les états excités, est très importante pour comprendre les processus photo-induits tels que la photo-commutation. Jusqu’à présent, ils étaient inaccessibles, mais l’utilisation des lasers à électrons libres permet de révéler leur structure. Grâce à la combinaison de la cristallographie résolue en temps en utilisant le laser à électrons libres, la spectroscopie d’absorption ultra-rapide et des simulations de dynamique moléculaire, deux états excités ont été dévoilés et discutés.

Les résultats de spectroscopie d’absorption révèle que l’isomerisation *trans*/*cis*, lors de la commutation de l’état non-fluorescent à l’état fluorescent, se produit à l’état excité à l’échelle de la picoseconde, et que la déprotonation se déroule sur une échelle de temps plus longue. Les données de cristallographie résolue en temps à 1 ps révèlent deux états du chromophore généré par la pompe laser : l’un, planaire et l’autre twisté. Les calculs de dynamique moléculaire prédisent la formation de ces deux états à l’échelle de la picoseconde. Les données de cristallographie résolue en temps à 3 ps dévoile l’isomère *cis* du chromophore, suggérant que l’isomérisation a donc déjà eu lieu.

Mis en commun, les résultats semblent indiquer que l’isomérisation *trans*/*cis* du chromophore se déroule à l’échelle de la picoseconde et implique la formation d’un état « twisté » à mi-chemin entre les isomères *trans* et *cis*, associé à un état excité. La structure de cet état excité révèle l’importance de l’acide aminé en position 151 dans le mécanisme de commutation. Des variants ont été générés pour lesquels la valine 151 a été substituée par une alanine, à la chaîne latérale plus courte ou une leucine à la chaîne latérale plus longue. Ces variants ont été caractérisés et affichent un contraste amélioré et un rendement quantique de commutation augmenté pour le premier, par opposition à un contraste dégradé et un rendement quantique de commutation abaissé pour le second.

Chapitre III : Mécanisme de photo-commutation élucidé par cristallographie à la femtoseconde et spectroscopie d'absorption résolues en temps.

Si, au vu des résultats précédents, il s'avère que l'isomérisation du chromophore se produit à l'état excité à l'échelle de la picoseconde, la suite des évènements de la photo-commutation reste encore à élucider. La combinaison de spectroscopie d'absorption UV-visible résolue en temps à l'échelle de la micro- milliseconde et de cristallographie résolue en temps à l'échelle de la nanoseconde permet de proposer un mécanisme de photo-commutation.

La spectroscopie d'absorption à l'échelle de la micro- milliseconde révèle que la déprotonation a lieu à l'échelle de la centaine de microsecondes, et qu'un réarrangement conformationnel se produit à l'échelle de la milliseconde. La structure de l'intermédiaire formé à 10 ns dévoile l'isomère *cis* du chromophore. La déprotonation se déroulant plus tard, cet état correspondrait à l'isomère *cis* protoné, le premier photo-produit formé lors de la commutation de l'état non-fluorescent vers l'état fluorescent. La conformation alternée de l'histidine 149 constitue la différence marquante avec l'état *cis* anionique de l'état fluorescent. Cet acide aminé forme une liaison hydrogène avec le chromophore dans l'état fluorescent. Le changement d'état de protonation du chromophore expliquerait cette différence qui implique nécessairement un réarrangement sur une échelle de temps plus longue. Ce résultat, qui complète celui du résultat discuté dans le chapitre précédent, permet de proposer un scénario de photo-commutation dans lequel l'isomérisation se produit à l'état excité à l'échelle de la picoseconde, suivi par la déprotonation à l'état basal à l'échelle de la microseconde et un réarrangement conformationel à l'échelle de la milliseconde.

Il est également rapporté la découverte fortuite d'un deuxième isomère *trans* formé suite à la l'illumination à 488 nm, isomère qui pourrait être assimilé à un état *off* non-fluorescent alternatif.

Chapitre IV : Caractérisation structurale des états fluorescent et non-fluorescent des variants V151A et V151L de rsEGFP2.

Dans le chapitre II, il était discuté que la formation d'un intermédiaire « twisté » à l'état excité était lié à la longueur de la chaine latérale du résidu en position 151, que raccourcir ou rallonger la chaine latérale de ce dernier influerait les propriétés de la protéine. Les variants V151A et V151L ont été générés et caractérisés, et il s'avère que le premier est un meilleur commutateur avec un meilleur contraste, et que le deuxième se comporte de manière opposée. Il est ici présenté et discuté la structure des états fluorescents et non-fluorescents des variants V151A et V151L obtenues par cristallographie aux rayons-X au synchrotron en condition cryogénique.

La structure de l'état fluorescent des deux variants est très similaire à celle du parent d'origine, mise à part la mutation introduite. Cependant, des différences sont notables si on compare les structures des états non-fluorescents. Chez V151A, le chromophore est en *trans*, très similaire à celui de rsEGFP2. La tyrosine 146 et l'histidine 149 ont des positions différentes et ne sont pas liées par liaison hydrogène comme c'est le cas pour rsEGFP2. Chez V151L, le chromophore est aussi en *trans*, mais différent de celui de rsEGFP2 et semble se lier à l'histidine 149, contrairement à rsEGFP2, dont la conformation de l'histidine 149 est différente. Il est également rapporté que les états fluorescent et non-fluorescent coexistent dans le cristal après illumination à 488 nm, conséquence d'une difficulté à éteindre totalement les protéines, probablement due au mauvais contraste affiché par ce variant.

Les résultats structuraux montrent des différences entre les trois versions de la protéine qui peuvent expliquer les différences de comportement photophysique rapportées. Il est également remarqué que rsEGFP2 se comporte spectroscopiquement comme un mélange de ces deux variants.

Chapitre V : Etude de l'hétérogénéité de l'état non-fluorescent de rsEGFP2.

Ce dernier chapitre propose une explication aux différentes observations faites et discutées dans les chapitres précédents. A partir des résultats rapportés dans le chapitre II, les variants V151A et V151L de rsEGFP2 ont été générés, puis caractérisés spectroscopiquement et structuralement dans le chapitre IV. Le chapitre III rapporte une conformation inattendue et intrigante assimilable à un état non-fluorescent alternatif. Suite à une expérience de cristallographie résolue en temps au XFEL, cet état est réapparu.

Sur la base d'une carte de différence calculée entre un état obtenu après illumination à 488 nm et un état non illuminé, donc fluorescent, de la densité électronique apparaît indiquant la présence des deux états *trans* du chromophore déjà observés dans le chapitre III, mais observés ici de manière plus claire. Sur la base d'une carte de différence calculée entre deux états obtenus après illumination à 488 lors de l'expérience courante et une expérience précédente (celle du chapitre II), il semble que l'efficacité insuffisante de l'illumination visant à éteindre la protéine est la raison de l'apparition de ce nouvel état. La comparaison des états *off-1* et *off-2* avec les états non-fluorescents des variants V151A et V151L permet de proposer que le *off-1* est assimilable au *off*-V151A et le *off-2* assimilable au *off*-V151L. En partant de cette dernière hypothèse, et en se basant sur les propriétés des variants, notamment le contraste, un modèle est proposé pour expliquer le peuplement relatif des états *off-1* et *off-2* par une illumination prolongée à 488 nm.

De plus, si il s'avère que les états *off-1* et *off-2* peuvent coexister lorsque la protéine s'éteint, cela expliquerait pourquoi en solution le comportement spectroscopique de rsEGFP2 est un mélange des comportements de ses deux variants. Des expériences de spectroscopie ultra-rapides entreprises par le Dr. Michel Sliwa (non présentées dans le manuscrit) semblent confirmer cet état de fait et posent les bases de futures études de cristallographie résolue en temps au XFEL sur les variants V151A et V151L.

Conclusion

Grâce au protocole de microcristallisation proposé dans le chapitre I, il a été possible de produire de manière reproductible et en grande quantité les échantillons nécessaires aux expériences de cristallographie résolue en temps au XFEL discutées dans les chapitres suivants. Ces expériences ont permis d'élucider la structure d'un intermédiaire « twisté » à l'état excité se formant à l'échelle de la picoseconde (Chapitre II), ainsi que d'un intermédiaire où le chromophore est *cis* protoné à 10 ns (Chapitre III), premier photo-produit de la commutation. Ainsi un mécanisme de photo-commutation est proposé, semblant s'accorder avec le mécanisme déjà élucidé, mais encore discuté de façon controversée, chez Dronpa.

La caractérisation des variants V151A et V151L, la résolution des structure de leurs états fluorescent et non-fluorescent, ainsi que la découverte et l'explication proposée de l'hétérogénéité de l'état non-fluorescent de rsEGFP2, permettent d'une part de comprendre plus finement le mécanisme de photo-commutation, et pavent le chemin vers la conception rationnelle basée sur la structure de nouvelles protéines dérivées de rsEGFP2 et optimisées pour les applications en microscopie de super-résolution.

En guise de perspective, l'étude par cristallographie résolue en temps des variants V151A et V151L a été entreprise récemment.

Appendix B: Supplementary Information: Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography

Supplementary Information

Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography

Nicolas Coquelle, Michel Sliwa, Joyce Woodhouse, Giorgio Schirò, Virgile Adam, Andrew Aquila, Thomas R. M. Barends, Sébastien Boutet, Martin Byrdin, Sergio Carbajo, Eugenio De la Mora, R. Bruce Doak, Mikolaj Feliks, Franck Fieschi, Lutz Foucar, Virginia Guillon, Mario Hilpert, Mark Hunter, Stefan Jakobs, Jason E. Koglin, Gabriela Kovacsova, Thomas J. Lane, Bernard Lévy, Mengning Liang, Karol Nass, Jacqueline Ridard, Joseph S. Robinson, Christopher M. Roome, Cyril Ruckebusch, Matthew Seaberg, Michel Thepaut, Marco Cammarata, Isabelle Demachy, Martin Field, Robert L. Shoeman, Dominique Bourgeois, Jacques-Philippe Colletier, Ilme Schlichting, Martin Weik

correspondence to: martin.weik@ibs.fr, Ilme.Schlichting@mpimf-heidelberg.mpg.de, colletier@ibs.fr

Table- of- contents

1. Ultrafast transient absorption spectroscopy

2. QM/MM and MD simulations

3. Serial femtosecond crystallography

1. Ultrafast transient absorption spectroscopy

Ultrafast transient absorption spectroscopy of rsEGFP2 in solution in the picosecond time-domain carried out at low pump energy

The ultrafast transient absorption experiments were carried out on rsEGFP2 in 50 mM Hepes, 50 mM NaCl, pH 8, using a method similar to that described elsewhere for IrisFP[1]. Briefly, a 1-kHz femtosecond transient absorption set up was used; and the pump beam (400 nm) was obtained by doubling fundamental 800 nm pulses. The solution of rsEGFP2 in the *off* state (absorbance ca. 0.1 at 400 nm) was inside a flow cell (1 mm thick CaF_2 windows, optical path length 250 μm). Before entering the flowcell, the reservoir solution was irradiated continuously with 470 nm light to switch the protein back to the *off* state after excitation by the pump beam (2 μJ / pulse at the sample position; spot diameter 0.4 mm; energy density 1.3 mJ/cm^2). Owing to the low absorbance and the low pump excitation power, multiphoton processes and formation of ionic species were avoided (see next section for higher-power experiments). The instrument response function (110 fs FWHM) was estimated using the stimulated Raman amplification signal (Supplementary Fig. 1a) and spectra were corrected from group velocity dispersion.

Supplementary Fig. 1a-c shows the transient difference absorbance spectra from 355 to 675 nm and from -0.1 to 39.9 ps after excitation of rsEGFP2 in its *off* state by a 110 fs pulse at 400 nm. Between -0.1 and 0.4 ps, four difference bands are growing (Supplementary Fig. 1a): two positive (maxima at 355, 455 nm) and two negative bands (narrow and broad band with minima at 405 nm and 530 nm, respectively). The former are characteristic of absorption by excited-state species, the latter reflect the depopulation of the *off* state (405 nm) and stimulated emission of excited-state species (530 nm). Between 0.5 and 39.9 ps, all bands decay with two time constants (a global decay analysis was carried out where we focused on spectral changes observed at four wavelengths (385, 455, 530 and 650 nm) and this analysis yielded two decay time constants of 0.92 ± 0.08 ps and 3.65 ± 0.15 ps, Supplementary Fig 1d) and a positive band is growing at 380 nm to reach a difference spectrum characterized by this positive band and a small negative band between 465 and 600 nm with a maximum at 502 nm (Supplementary Fig. 1c) that do not change up to upper limit of the pump-probe delay of 1 ns (not shown). The band at 380 nm is most likely due to formation of a protonated *cis* isomer in line with its presence in a spectrum recorded from rsEGFP2 in its fluorescent *on* state at pH 4 where the *cis* isomer of the chromophore is protonated (Supplementary Fig. 2). The negative band with the maximum at 502 nm (Supplementary Fig. 1c, difference spectrum at 39.9 ps) originates from the depopulation and stimulated emission of the anionic *cis* isomer present in a residual population (about 10%) of the fluorescent *on* state (Supplementary Fig. 2).

Since the residual population (10%) of the fluorescent *on* state emits at 502 nm, *i.e.* close to the time-dependent stimulated-emission band with a maximum at 530 nm (Fig. 2a, Supplementary Fig. 1a-c), its time evolution needed to be determined. To this end, spectral changes of a solution with 100% of the rsEGFP2 molecules being in their fluorescent *on* state (anionic *cis* chromophore) were followed upon femtosecond excitation at 400 nm (Supplementary Fig. 3). The expected negative band with a maximum at 502 nm rises within a few hundred femtoseconds (Supplementary Fig. 3a), followed by a vibrational relaxation within a few picoseconds (shift and increase of the band), and then remains constant (Supplementary Fig. 3b) up to 50 ps. Consequently, the decay of the spectra in Fig. 2a on the ps time scale (decay in 0.9 and 3.7 ps) and Supplementary Fig. 1b,c is solely due to photo-dynamics of the protonated *off* state without contributions from the photo-dynamics of the anionic *on* state.

Ultrafast transient absorption spectroscopy of rsEGFP2 in solution as a function of pump energy

In the previous section, transient absorption spectroscopy was carried out at low pump energy (corresponding to 13 GW/cm^2 and less than one photon absorbed per protein) to avoid multi-photon and photobleaching processes. Since TR-SFX experiments were carried out at much higher pump energy (corresponding to nominal 400 GW/cm^2), transient absorption spectroscopy was carried out at two additional pump energies, corresponding to 68 and 140 GW/cm^2 (Supplementary Fig. 4). Excitation at higher pump energies was not possible due to laser scattering and a strong signal from the solvent (spectra with 140 GW/cm^2 were cut below 430 nm). At 13, 68 and 140 GW/cm^2, the 455 nm band (absorption of excited state species) reaches its maximum at 0.6 ps and decays with similar time constants (2.75 +/- 0.03 ps, 2.73 +/- 0.05 ps, 2.47 +/- 0.14 ps, respectively), indicating that the formation and decay of excited state species do not strongly depend on the pump energy. However, at 68 and 140 GW/cm^2 an additional species is formed within a few hundred femtoseconds that absorbs from 500 to 700 nm with a maximum at 700 nm (Supplementary Fig. 4e,f,h,i) and whose spectrum remains constant up to 1 ns (not shown). Similar pump-energy dependent experiments on Dronpa[2] and photoactive yellow protein[3] yielded a similar additional band that has been interpreted as being due to the formation of solvated electrons and a chromophore radical cation.Importantly, the band attributed to a radical cation and solvated electron species does not evolve on the ps time scale. This means that this radical does not correspond to either of the two SFX intermediate states (models P and T, see below) whose occupancies decrease between 1 and 3 ps.

Power titration with femtosecond excitation of rsEGFP2 off-to-on photoswitching in solution

The *off*-to-*on* photoswitching efficiency with femtosecond laser excitation of rsEGFP2 in solution was investigated as a function of the laser energy density, similarly to experiments carried out on crystals of the photoactive yellow protein[4]. The experiments were carried out on a 2.5 mM rsEGFP2 solution in 50 mM Hepes, 50 mM NaCl, pH 8. The sample was sandwiched between the faces of an Infrasil cuvette with an optical path length of 50 μm. The pump beam to trigger *off*-to-*on* photoswitching (400 nm, about 150 fs pulse length, circular polarization) was obtained by using the second harmonic of a Coherent Femtosecond Amplifier (Elite). The white probe beam was obtained by a cold halogen lamp (Schott). Pump and probe beam spatial profiles at the sample position were characterized by a beam profiler camera (Spiricon) and were 380×250 μm^2 FWHM and 68×62 μm^2 FWHM, respectively. The stability of pump-probe spatial overlap was checked several times during the experiments. A LED source at 490 nm (Thorlabs) was collimated and focused on the sample (about 200 mW/cm^2) and used to prepare the *off* state before femtosecond excitation and to check the reversibility of photoswitching after femtosecond excitation.

Before pump-laser excitation, a spectrum of the *on* state (resting state) was collected, followed by illumination with the LED source at 490 nm for about 15 min in order to generate the *off* state (see Supplementary Fig. 5). The *on*- and *off*-state spectra were then used as a reference for interpreting the spectra after fs excitation.

The scheme of the experiment was as follows:

1 - spectrum acquisition (*on* state) during 5 seconds
2 - 490 nm LED illumination during 15 seconds
3 - spectrum acquisition (*off* state) during 5 seconds
4 - 400 nm femtosecond pulse
5 - spectrum acquisition during 5 seconds
6 - 490 nm LED illumination during 15 seconds

The above sequence was repeated five times at each laser energy. It was verified by calculation that protein molecules outside the pump surface cannot reach the probe surface within the time necessary to complete one cycle. The spectra acquired at steps 5 and 1 were used as a measure of the *on* state switching efficiency by fs excitation and of the reversibility of photoswitching after fs excitation, respectively (see Supplementary Fig. 5). The spectra were then fitted in the 380-525 nm wavelength range by a linear combination of the *on* and *off* reference spectra. Supplementary Fig. 5 shows the *on* and *off* reference spectra (blue and green curves, respectively), the spectrum after 400 nm fs excitation at 950 μJ/mm^2 (red curve), the result of the fitting procedure for the spectrum after 400 nm fs excitation (cyan curve), and the spectrum after 490 nm LED re-illumination (magenta curve).

The coefficients of the linear combination obtained with the fitting procedure give an estimate of the *on* and *off* state fractions at every step in the sequence. Supplementary Fig. 6 shows the fraction of *on* state after 400 nm laser excitation as a function of the energy density up to 1 mJ/mm^2. Supplementary Fig. 7 shows the sum of the fractions of *on* and *off* states after 400 nm laser excitation (left panel) and the fraction of *off* state after 490 nm LED re-illumination (right panel) in the same energy density range. Points in Supplementary Fig. 6 are averaged over several (typically five) repetitions. It is worth noting that sample concentration and thickness were chosen such as to have similar numbers of protein molecules in the excited volume and in microcrystals during the SFX experiment. This condition enables a direct comparison of the excitation conditions used during SFX experiments and the spectroscopic measurements described here.

Spectra in Supplementary Fig. 5 (red and cyan curves) and data in Supplementary Fig. 7 left panel demonstrate that the product of fs excitation is spectroscopically interpretable as a linear combination of *on* and *off* states over the entire energy-density range explored, including the one in the SFX experiment (0.94 mJ/mm^2). Data in Supplementary Fig. 6 shows that the *on*-to-*off* photoswitching efficiency is monotonic in the entire energy range, excluding the formation of a photobleached species that would lead to a decrease in *on*-state formation above a certain energy-density threshold. Spectra in Supplementary Fig. 5 (green and magenta curves) and data in Supplementary Fig. 7 right panel demonstrate that 99% on average of the product of femtosecond excitation is reversibly switchable back to the *off* state upon illumination by 490 nm LED radiation.

2. QM/MM and MD simulations

Excited state QM/MM simulations: calculation of protonation states

All QM/MM simulations described were performed starting from the crystallographic *off*-state conformation of the laser-off structure obtained in this work (see below).

Protonation states for the residues in the protein were obtained by performing Poisson-Boltzmann (PB) continuum electrostatic calculations combined with a Monte Carlo titration. These calculations were performed using the Pcetk module[5] of the pDynamo[6] program coupled to the external PB solver, extended-MEAD[7].

The initial setup of the electrostatic model was performed in CHARMM[8,9]. Molecular mechanics parameters, such as atomic charges and radii, were taken from the CHARMM force field[10] for the protein, whereas those for the neutral and anionic forms of the chromophore were adapted from the work of Reuter et al[11]. Hydrogen atoms were added to the X-ray crystallographic structure of the *off*-state structure of rsEGFP2 using the HBUILD routine in CHARMM.

Titratable residues were protonated according to their standard protonation states at pH 7. The coordinates of the added hydrogens were geometry optimized in CHARMM, whereas the coordinates of the heavy atoms were kept fixed. During the subsequent PB calculations in pDynamo, the solvent was assigned an ionic strength of 100 mM and a temperature of 300 K. To calculate the molecular surface of the protein, a solvent probe with a radius of 1.4 Å and an ion exclusion layer of 2.0 Å were used. The electrostatic potential was calculated using four focusing steps at resolutions of 2.0, 1.0, 0.5, and 0.25 Å. Each step employed a grid of 121^3 points. The Metropolis Monte Carlo method, as implemented in pDynamo[5], was used to calculate the protonation states of titratable groups. Each MC run consisted of 500 equilibration scans followed by 20000 production scans. Double moves[12] were applied for the pairs of groups whose absolute electrostatic interaction energy was calculated to be > 2.0 kcal mol^{-1}. Titration curves were calculated for a pH range between 0 and 14 at a resolution of 0.5 pH units.

At pH 7 the calculations indicated that the residues of the protein were all in their expected forms with the exception of Glu223, which was protonated. The histidines were all singly protonated with the delta-protonated forms being most stable for His3, His182, His200 and His218, and the epsilon-protonated forms for the remaining histidines, His5, His4, His26, His78, His82, His140, His149 and His170. The chromophore itself was determined to be predominantly neutral, 85%, with the remaining 15% anionic. The populations of the cationic and zwitterionic forms of the chromophore were negligible for the whole range of pH values.

Excited state QM/MM simulations: molecular dynamics simulations

Once the predominant protonation states of the protein at pH 7 had been determined, a simulation model of the system with the chromophore in its neutral state was prepared using the VMD [13] and NAMD [14] programs. A molecular mechanical (MM) model of the protein was setup using the appropriate protonation states and the same CHARMM force field as indicated in the previous section. Using the solvate module of VMD this system was then solvated in an orthorhombic water box of approximate dimensions 62.0 × 73.7 × 71.3 $Å^3$ to ensure a layer of water of at least 10 Å on each side of the protein. In addition, K^+ cations were included to make the total system neutral. The complete model consisted of 3735 protein atoms, 10232 water molecules and 6 K^+ cations (34437 atoms in total). Standard CHARMM parameters were employed for the K^+ cations, and the TIP3P model[15] was taken for water.

In the next step, a geometry refinement of the model was performed in which the protein heavy atoms were weakly restrained to their experimental crystallographic positions using harmonic positional restraints (force constant 2 kcal mol^{-1} $Å^{-2}$). The solvent atoms and protein hydrogens were unrestrained. This refinement consisted of 1000 steps of geometry relaxation using the steepest descent method, a heating molecular dynamics (MD) simulation in which the temperature was gradually increased from 20 to 300 K at a rate of 10 K ps^{-1}, and a 10 ns MD equilibration run at 300 K. All calculations were done with the application of periodic boundary conditions and determination of the long-range electrostatic interactions using the particle-mesh Ewald method. The MD runs were carried out in the NPT ensemble under the normal pressure of 1 bar and using a time step of 2 fs.

To terminate, a 100 ps MD simulation was carried out (using the same conditions of temperature and pressure as above) in which all atoms were permitted to move freely. The relatively short duration of the simulation was employed to ensure that the protein atoms did not deviate substantially from their crystallographic positions. For the subsequent calculations, structures were saved at 2 ns and 10 ps intervals from the restrained and unrestrained MD simulations, respectively, giving 15 snapshots in all.

Excited state QM/MM simulations: method

All quantum chemical/molecular mechanical (QC/MM) simulations were done with the pDynamo program[6]. Most calculations were carried out with pDynamo's semi-empirical QC methods, although tests were also performed with density functional theory (DFT) and complete active space self-consistent field (CASSCF) molecular orbital methods using pDynamo coupled to the ORCA QC program[16].

We preceded our simulations with tests to calibrate different QC/MM models versus the rsEGFP2 system. For this we benefited from our previous extensive experience of the application of QC/MM methods to similar fluorescent proteins (see, for example[17,18]). Among other things, we examined the QC model, the number of atoms in the QC region, the number of mobile atoms for the QC/MM MD simulations, and the treatment of the QC/MM interactions.

Our final QC/MM model, which we used for the simulations presented in this paper, provided a reasonable compromise between accuracy and cost, and consisted of the following four points:

- the semi-empirical AM1 QC method[19] combined with the same CHARMM MM model for the non-QC atoms as employed for the MD simulations. We have shown previously that the AM1 semi-empirical method, in combination with a suitable configuration interaction (CI) method, provides a reasonable description of the ground and lower energy excited states of the chromophore[17,18]. Here, we employed a FOMO-CASCI approach, which has been shown by others to approximate state-averaged *ab initio* CASSCF methods well[20]. For our production runs, we employed an active space of 4 electrons in 3 orbitals, as we found that this gave results in accord with larger and more expensive CI spaces.
- a QC region comprising the chromophore with QC/MM covalent interactions treated *via* the link atom method[6]. This gave 45 QC atoms in total of which 3 were link atoms. Other QC regions that were tested included either all or various combinations of the additional groups, Gln95, Arg97, Ser164, Gln184, Glu223, and the water hydrogen-bonded to the chromophore's phenol group.
- an approximately spherical mobile region of 15 Å radius around the chromophore containing 2224 atoms. All other atoms were kept fixed at their initial positions during the dynamics although their interactions with the mobile atoms were, of course, computed. Test simulations allowing all atoms in the system to move showed no significant differences with those presented here.

- the inclusion of the full QC (electronic density and nuclear charge)/MM (partial charge) electrostatic interactions in the QC SCF/CI calculation. All the QC/MM and MM/MM non-bonding (electrostatic and Lennard-Jones) interactions were smoothed using the atom-based force switching approximation with inner and outer cutoffs of 8 and 12 Å, respectively[6].

With each QC/MM model, MD simulations were performed using either a velocity Verlet algorithm (all atoms mobile) or a Langevin algorithm with a small collision frequency of 10^{-2} ps^{-1} (immobile atoms 15 Å from the chromophore). The time step for each algorithm was 1 fs and simulations were run at a temperature of 300 K. For each electronic state, equilibration runs of, typically, 100 ps duration were carried out before production runs were started. To be clear, complete QC SCF and CI calculations were performed at each time step so that a 100 ps simulation with a 1 fs timestep required 10^5 SCF/CI energy and force calculations.

Excited state QM/MM simulations : chromophore models in vacuum

To test the QC model employed in the QC/MM simulations, the trans structure of the chromophore was extracted from the X-ray crystallographic structure of the *off*-state conformer and studied in vacuum. The geometry-optimized structure in the ground singlet state, S_0, has *tau* and *phi* angles of -1° and -42°, respectively. The non-planar *phi* angle is due to the unfavorable steric interaction between the phenol and the imidazolinone oxygen in the trans state. The vertical excitation energies to higher states are calculated to be 3.29, 4.01, 5.83 and 1.94 eV for the S_1, S_2, S_3 and T_1 states, respectively. The $S_0 \rightarrow S_1$ energy is in good agreement for that found for the neutral form of the chromophore in previous studies using a variety of semi-empirical and *ab initio* theoretical approaches[21,22].

Geometry optimization of the S_1 excited state leads to a lowest energy minimum with *tau* and *phi* angles of -88° and 0°, respectively, with an energy 2.13 eV higher than the geometry optimized form of S_0. There are also additional, higher energy minima for S_1 with planar values for *tau* and values of ± 90° for *phi* but we do not observe these in our subsequent simulations. The optimized geometry of the S_2 state resembles that of S_0 with values of -24° for both the *tau* and *phi* angles, whereas the geometry of S_3 has *tau* and *phi* angles of 42° and -72°, respectively, and a distorted imidazolinone ring. By contrast, the optimized geometry of T_1 is similar to that of S_1 with *tau* and *phi* angles of -105° and 1°, respectively, and an energy with respect to S_0 of 1.07 eV.

Excited state QM/MM simulations of the chromophore in rsEGFP2

The initial 100 ps simulation of the rsEGFP2 QC/MM system in its S_0 state yielded structures closely resembling the X-ray crystallographic *off*-state structure with *tau* and *phi* values (mean ± one standard deviation) of 2 ± 8° and -46 ± 11°, respectively. The vertical $S_0 \rightarrow S_1$ excitation energy was consistent with that observed for the chromophore model in vacuum with values of 3.30 ± 0.11 eV. This is equivalent to an absorption band maximum of 376 nm, which compares reasonably to the experimental value of 408 nm.

It is in principle possible to carry out quantum dynamical simulations to mimic light-excitation of the type that occurs in the TR-SFX experiments (see, for example[23]). However, these calculations are expensive and also sensitive to the way in which the radiation field is treated. Instead, in this work, we employed a simpler approach in which we performed simulations in the S_0 ground state of rsEGFP2 and then, at given intervals, spawned trajectories in the appropriate excited state. At the end of each excited state trajectory, the ground state simulation was resumed from its configuration at the last spawning point so as to have a continuous ground state trajectory. The switch from the ground to excited states was assumed to occur instantaneously without modification of the atomic positions and velocities.

We start with excitations to the S_1 state. Supplementary Fig. 8 (left panel) shows the change in the values of the *tau* and *phi* angles as a function of time for 100 S_1 excited state trajectories spawned at 1 ps intervals from a S_0 ground state trajectory. Each excited state trajectory was run for 2 ps. The angles change rapidly from their initial ground state values (*tau* = 4 ± 8°, *phi* = -48 ± 11°) to those of the S_1 state (*tau* = -86 ± 9°, *phi* = 16 ± 10°) with the change essentially being complete after 1 ps. Supplementary Fig. 9 shows the average structure at 1 ps calculated from these trajectories, which is in close agreement with the experimental intermediate structure determined by time-resolved serial femtosecond crystallography at the same time point (see below). Although we do not expect to obtain very accurate excited state kinetic constants from our simulations, due to the approximate nature of our dynamical approach, we nevertheless performed exponential fits for the curves in Supplementary Fig. 8 (left panel). We find that the change in *tau* is a mono-exponential process with a time constant of 0.32 ps, whereas the change in *phi* is bi-exponential with fast and slow components of 0.09 and 0.86 ps, respectively.

To estimate how the initial S_0 structures influence the excitation dynamics, we repeated the analyses of the preceding paragraph using various subsets of the excitation trajectories. The results were, in general, relatively insensitive to how the trajectories were selected. To illustrate this, we show in Supplementary Fig. 8 (right panel) the change in the values of the *tau* and *phi* angles as a function of time for the 28 (out of 100) trajectories whose initial vertical $S_0 \rightarrow S_1$ excitation energy was within one third of the standard deviation (0.11 eV) of the average excitation energy (i.e. 3.30 ± 0.037 eV). As can be seen, the dynamics is very similar although the fluctuations at the start of the trajectory are a little larger. Exponential fits of these curves are qualitatively similar to the previous ones, with a time constant for the change in *tau* of 0.36 ps, and time constants for the change in *phi* of 0.10 and 0.60 ps. The average structures after 1 ps derived from the full (100) and reduced (28) sets of excitation trajectories are extremely similar and thus are not presented here. Surface hopping from the S_1 to the S_0 state was not permitted during these trajectories. When surface hopping was permitted by using the approximate AS3 surface hopping algorithm of Fabiano and co-workers[24] a substantial deexcitation to the S_0 state can occur during the 2 ps duration of the excitation trajectories. In contrast to the conclusions of Fabiano and co-workers, we note that the fraction of trajectories that deexcite is sensitive to the values of the parameters employed in the AS3 algorithm. However, with a choice of parameter values similar to theirs, we find that still approximately 15-20% of the excitation trajectories survive in the S_1 state at 1 ps. Observation of a twisted chromophore in Fig. 4 and Supplementary Fig. 9 is thus not an artifact from preventing surface hopping. Deexcitation to the S_0 state is rapid, with the ground state S_0 structure being recovered within about 150 fs.

In addition to the $S_0 \rightarrow S_1$ simulations, we also performed simulations from S_0 to the S_2 and S_3 states. We do not show results for these here as they generate structures which are consistent with model chromophore calculations (see above) that are far from the twisted intermediate structure observed experimentally. Likewise, we also examined the behavior of the T_1 state of the chromophore in the protein, although this is unlikely to be relevant to the experimental observations due to the short time span of the experiments (~1 ps) and the much longer time constants associated with singlet to triplet intersystem crossing (> 1 ns). Interestingly, we find that the average T_1 dynamical structure of the chromophore in the protein is actually quite close to that of the S_0 ground state despite the fact that its vacuum optimized structure resembles that of S_1.

Finally, we also performed switching simulations from S_1 to S_0, *i.e.* by carrying out a QM/MM MD simulation in the S_1 state and then periodically switching to S_0. In agreement with experiment we find that the majority of the simulations revert back to the *trans* configuration of the S_0 ground state, although a small fraction, ~ 4%, leads to the *cis* configuration.

DFT calculations of the cationic ground-state chromophore in rsEGFP2

Quantum mechanical calculations of the cationic chromophore, generated by removal of one electron from the neutral *trans* species of the *off* state, were carried out for all geometries characterized by the values of *tau* and *phi* in the interval from -60° to +60° in steps of 30°. For each geometry, all coordinates except for *tau* and *phi* were optimized using density functional theory, with the B3LYP functional and 6-31G* basis set[25]. The energy landscape resulting from these calculations exhibits a unique minimum-energy geometry, corresponding to the planar *trans* chromophore of the *off*-state.

Excited state classical MD simulations: force field

Starting from the crystallographic *off*-state conformation of the laser-off structure obtained in this work (see below), force field molecular dynamics simulations of the protein with the neutral chromophore in the electronic ground state S_0 were first carried out to sample initial geometries for starting the excited state S_1 chromophore simulations.

The Amber 99 force field of the AMBER suite[26] was used for standard amino-acids. Concerning the chromophore, force constants of the bonded terms and van der Waals parameters were taken by analogy in the Amber force field. Equilibrium values were taken from geometries optimized at the DFT level of calculations for the S_0 state and CIS (Configuration Interaction between Singly excited states) level for the S_1 state, using the 6-31G* basis set[25]. The results are shown in Supplementary tables 1 and 2. Point charges for the S_0 and S_1 states were derived by fitting the *ab initio* electrostatic potential created by the chromophore (including the two peptide links) in the S_0 or S_1 state. Several constraints were imposed in the fitting procedure using a custom made program, as done previously by Jonasson *et al*[27]. The calculated values are shown in Supplementary table 3. Concerning the dihedral potentials of the *tau* and *phi* torsions around the ring bridging bonds, the standard form of the Amber force field was taken for the S_0 state with small force constants equal to 6 kcal/mol and 3 kcal/mol, respectively, to allow the chromophore to keep the large *phi* torsion angle (-46°) observed in the crystallographic *off*-state conformation of the laser-off structure. For the S_1 state of the neutral chromophore, we have to define a coupled potential of the *tau* and *phi* torsions as already done for the anionic form of the chromophore[27]. Quantum mechanical calculations of the excited state of the neutral chromophore has been pointed out to be a difficult task[21,28], and only sparse quantitative information is available in the literature for the neutral excited state. Minima in the *tau*-twisted geometries (*tau* = +/- 90° and *phi* = 0) have been reported, with an energy stabilization with respect to the planar geometry of either -8 kcal/mol[29] [21] or -19 kcal/mol[30]. The minimum energy path from the *cis* planar geometry to the *tau*-twisted geometry calculated by Toniolo *et al* [29] shows a low barrier of about 1 kcal located near *tau* = 45°. No minimum along the *phi* torsion was reported. Olsen *et al*[30] have calculated the energy of a constrained geometry with *phi* = 90° and *tau* = 0°, yielding a destabilization of 13 kcal/mol with respect to the planar geometry. In this work we have calculated the energy landscape of the excited state using simply the CIS method. This was done only on a grid of points with absolute values of *tau* and *phi* smaller than 60°, since CIS in not reliable for absolute values near 90°. In order to obtain a complete grid we have added to the CIS points the two *tau* - twisted minima (*tau* = +/- 90° and *phi* = 0), with an energy stabilization equal to -8 kcal/mol[21,29]. The entire set of values was fitted by a sum of products of sine and cosine of multiples of *tau* and *phi*, as reported earlier[27]. The result is shown in Supplementary Fig. 10 together with the expression of the fitted analytical energy function. In agreement with the above-mentioned theoretical results, we find a strictly repulsive *phi* profile, and a *tau* profile between the *cis* and the *tau*-twisted geometry with a small barrier (0.7 kcal/mol) located at about 40°. The *tau* profile starting from the *trans* geometry is barrier-less and stabilized by -8 kcal/mol in the two *tau*-twisted geometries (*tau* = +/- 90° and *phi* = 0).

Concerning the remaining torsional parameters, those for the phenol ring were taken from the Amber force field. For the imidazolinone ring, no specific parameters are available in the standard force fields. However, these parameters, especially those governing the flexibility at the nitrogen N3 site linked to the rest of the protein, play an important role when simulating the overall dynamics of the chromophore within the protein. In this work we used a custom made specific torsional force field for the imidazolinone ring, comprising nine independent terms[31] : five terms for the intra-cycle dihedral angles (that imply only the five atoms of the ring) and four improper angles (all atoms of the ring, except N2). Usual ring force fields include also exo-cyclic dihedral terms (for instance O2-C2-N3-C1), so that the dynamics at a given atom site results from an inextricable combination of many terms.

The force constants were adjusted on the fluctuations of the nine angles observed in a 10 ps DFT/MD simulation of the S_0 anionic methyl-terminated chromophore (A. de la Lande, private communication, 2011). The DFT/MD was performed using a modified version of deMon2k[32]. The resulting average geometry is planar. The DFT/MD fluctuations (Supplementary table 4) are found in the interval ~ 4 - 8° except the one for the N3 nitrogen improper torsion that takes, as expected, a much larger value (15.1°). The force constants of our specific force field and the resulting fluctuations are shown in the last two columns of Supplementary table 4. It appears that all DFT/MD fluctuations are well reproduced by the force field, particularly the higher flexibility of N3. This force field was initially tailored for the anionic S_0 chromophore. We have assumed it can also be used for the neutral chromophore since the higher flexibility of N3, as compared to the carbon atoms in the imidazolinone ring, is critical and is correctly reproduced here.

Excited state classical MD simulations : MD protocol

All amino acids were in their standard protonation state at pH 7, except for Glu223, which was protonated to allow hydrogen bonding with N2 of the chromophore. The protein was neutralized with Na^+ cations and solvated in a water box, resulting in 34474 atoms. The system was first minimized and equilibrated using the protocol described in[27].

A free dynamics simulation of 2 ns duration was run for the S_0 state at constant temperature (300 K) and pressure (1 atm). Fifty simulations of the S_1 state of 5 ps duration each were carried out starting from snapshots extracted from the second half of the free S_0 state simulation, and using a custom-made version of the PMEMD module of AMBER[27].

Excited state classical MD simulations: protein dynamics with the chromophore in S_0

The average value and the standard deviation of the dihedrals *tau* and *phi* in the ground state simulations are equal to -7° ± 10° and -38° ± 13°, respectively. Among the H-bonds linking the O2 oxygen of the imidazolinone ring of the chromophore to the OH oxygen of the phenol ring of the chromophore (O2-Arg97(NH2) / Arg97(NH2)-Gln184(OE1) / Gln184(OE1)-W17 / W17-Ser164(OG) / W17-OH) three of them appear as strong H-bonds: O2-Arg97(NH2), Arg97(NH2)-Gln184(OE1) and W17-OH. The H-bonds of the water W17 to Ser164 and to Gln184 exhibit a weaker character.

Excited state classical MD simulations: protein dynamics with the chromophore in S_1

The excited chromophore first accesses in less than 30 fs a flat area of the potential energy surface around the planar *trans* geometry (Supplementary Fig. 10), where its average conformation is characterized by *tau* and *phi* values of -37° ± 12° and -10° ± 10°, respectively (conformation 1, Supplementary Fig. 11). Subsequently, full chromophore twisting occurs, leading to a conformation with *tau* and *phi* values of -80° ± 7° and 12° ± 8°, respectively (conformation 2, Supplementary Fig. 11). The fully twisted chromophore conformation is observed in all 50 simulations and forms in less than 1 ps in 18 of them and in less than 4 ps in the remaining 32, with an average value of 1.9 ps. The H-bond between the phenol oxygen of the chromophore and the water molecule W17 is formed in 98% of the S_0 snapshots and in 97% and 56% of the S_1 snapshots in which conformation 1 and conformation 2 are observed, respectively. Full twisting in *tau* is thus accompanied by a weakening of that H-bond.

S_0-S_1 energy gap and oscillator strengths calculated from QM and QM/MM simulations

TDDFT and semi-empirical QM calculations on a chromophore model were performed to determine the energy gaps and oscillator strengths of the principal conformations observed experimentally. Reliable theoretical values of the emission wavelength from the fully *tau*-twisted neutral S_1 chromophore calculated by Toniolo et al[29] and by Polyakov et al[21] are ~1100 nm and 953 nm, respectively, much too far from the observed 530 nm. Since no values for a conformation similar to SFX model P (*tau* and *phi* values of -37° and -10°, respectively) are available in the literature we have performed a calculation at the simple TDDFT level with the B3LYP density functional and the cc-pVTZ basis set, using the Gaussian 03 package. This level of calculation has been shown to provide a value of the absorption energy of the planar neutral chromophore[28] in good agreement with the accurate values of Polyakov et al[21]. The calculated emission wavelength for SFX model P is equal to 570 nm, not far from the observed 530 nm. The corresponding oscillator strength is 0.15, which is smaller than the one associated with the fully planar chromophore (0.66[28]).

The semi-empirical QM calculations provide similar results. Fully planar and SFX model P chromophores are characterized by energy gaps corresponding to 380 and 484 nm, respectively and oscillator strengths of 0.87 and 0.53. The oscillator strength of the fully twisted chromophore (SFX model T) is negligibly small (< 0.001) and its energy gap corresponds to an emission band of ~1500 nm.

We also analyzed the oscillator strengths for the structures in the QM/MM excited state simulations. As in the model chromophore calculations, their values are small for structures around the fully *tau*-twisted conformation of the chromophore (SFX model T). However, the structures with maximum oscillator strengths are clustered quite sharply in the region with *tau* values of ~ -40° and *phi* values of ~ -20° and with excitation energies of ~ 220 kJ/mol. The latter is equivalent to an emission of ~ 540 nm, which is red-shifted with respect to the emission for the model chromophore calculations. Although we only see a transient population of structures of this type, they correspond well to the minor conformation detected by MD simulations and to SFX model P.

In summary, it follows that the stimulated emission band with a maximum at 530 nm seen in the ultrafast transient absorption experiments originates most probably from SFX model P (near planar chromophore), but not from SFX model T (fully twisted chromophore).

3. Serial femtosecond crystallography

Protein expression and purification

The reversibly photoswitchable enhanced green fluorescent protein 2 (rsEGFP2[33]) was expressed and purified as described[34]. After cell lysis, purification of rsEGFP2 fused to an N-terminal polyhistidine tag was achieved using 100 ml of Ni-NTA affinity resin, followed by size exclusion chromatography on a HiLoad 16/600 Superdex 75 prep grade column. To obtain the large amount of protein required for these experiments, 136 cycles of size exclusion chromatography were performed in 9 days using two high-throughput Akta Xpress systems on the MP3 platform (IBS, Grenoble).

Microcrystal growth, preparation and injection

rsEGFP2 microcrystals were obtained by seeding (Woodhouse *et al.*, to be described elsewhere), using crystallization conditions similar to those described in[34] which yield macrocrystals in the hanging drop vapor diffusion geometry. The final protein, precipitant, salt and buffer concentrations were ~20 mg/ml, ~2 M ammonium sulphate, 20 mM NaCl, 120 mM HEPES pH 8, respectively. Just prior to injection, rsEGFP2 microcrystals ($3 \times 3 \times 3\ \mu m^3$ (Supplementary Fig. 12)) were centrifuged for 5 min in a 15 ml tube at 1000 g. The supernatant was removed and the pellet resuspended and diluted 25- to 50-fold with 100 mM HEPES pH 8, 2 M ammonium sulphate (sample from which the *on*-state reference data set was collected) or 2.5 M ammonium sulphate (for samples from which all other data sets were collected). A solution containing 2 - 10% (v/v) crystals was filtered through a 20 μm stainless steel filter using a sample loop and a high-performance liquid chromatography (HPLC) pump. After filtering, the crystal slurry was concentrated by centrifugation and removal of part of the supernatant. Crystal suspensions (typically 2 to 4 ml volume) were transferred into a stainless steel sample syringe installed on an anti-settling device[35] equipped with a regulated, adjustable Peltier element-cooled syringe holder. The sample temperature was 20°C. Crystals were injected into the vacuum (~0.1 millitorr) microfocus chamber of the Coherent X-ray Imaging (CXI) endstation[36] of the Linac Coherent Light Source (LCLS) at SLAC National Accelerator Laboratory using a gas dynamic virtual nozzle (GDVN) injector[37]. The GDVN featured a sample capillary of 75 μm inner diameter and was run at a flow rate of 30 μl/min. The resulting jet had a diameter ranging from 3 to 6 μm (corresponding to linear flow rates of 40 to 25 m/s, respectively), depending on the focusing helium pressure. Attempts to increase the crystal concentration to above 8% (v/v) resulted in frequent clogging of the injector tubing or GDVN with rsEGFP2 microcrystals. To reduce the probability of clogging, tubing and nozzle were rinsed with water every 10 min for 2 min. A total of about 1g rsEGFP2 microcrystals were used during the five 12 hrs shifts during which our TR-SFX data were collected.

Pre-illumination of microcrystals to convert rsEGFP2 from the on to the off state

The absorption maxima of rsEGFP2 in its fluorescent (*on*) and its non-fluorescent (*off*) states are at 479 and 406 nm, respectively (Supplementary table 7). The resting state is the *on* state. Our goal was to study a picosecond timescale intermediate state along the *off*-to-*on* transition, requiring to switch all protein molecules in the crystal to the *off* state by pre-illumination (488 nm), before triggering the reaction with the pump laser (400 nm; see below), and probing the structure with X-ray pulses (Supplementary Fig. 13). The low primary quantum yield of the *on*-to-*off* photoswitching (0.04, Supplementary Table 7) required a large excess of photons per molecule to ensure efficient photoconversion. Pre-illumination was achieved by 488 nm laser light using a pre-illumination device (Supplementary Fig. 14)[38]. Briefly, the output beam of a CW laser (488 nm wavelength, 200 mW nominal power) was guided by a SMA optical fiber (200 μm core) to the tubing (254 μm inner diameter) connecting the sample reservoir with the GDVN. The optical fiber and the tubing were aligned with a stainless steel HPLC Tee union thereby forming an illumination cuvette, at a linear distance of 2 m upstream of the interaction zone. The transit time of a crystal under laser light within the pre-illumination region was 0.025 s. Absorption spectroscopy of microcrystals collected at the exit of the Tee union showed that crystalline rsEGFP2 was ~90% in the *off* state, and the remaining ~10% in the *on* state (Supplementary Fig. 15). At a flow rate of 30 μl/min, the delay between pre-illumination and laser pumping was 40 s, *i.e.* a time far shorter than the characteristic time of the *off*-to-*on* thermal recovery of monomeric rsEGFP2 in microcrystals (~100 min)[38]. Thus, the fraction of *off* state conformers in rsEGFP2 microcrystals presumably remains close to 90% at the interaction zone prior to laser pumping.

Femtosecond pump-laser excitation of rsEGFP2 crystals in TR-SFX experiments

Triggering photoswitching from the *off* towards the *on* state was achieved with pulses (~230 fs FWHM pulse length; 400 nm wavelength; 1.5 μJ/pulse; circularly polarized) generated by the second harmonic of a Ti:sapphire laser focused to a spot size of 40 μm × 40 μm (FWHM;

energy density 940 μJ mm^{-2} ; power density 400 GW cm^{-2}), in nearly collinear geometry with respect to the X-ray beam. The spatial overlap of the pump-laser focus with the XFEL beam at the interaction region was assessed every few hours by using a fine-grinded YAG screen in combination with a high-resolution camera monitoring the screen. Misalignments on the order of 20 – 30 μm were observed and corrected for.

SFX data collection and on-line monitoring

SFX data were collected at the CXI instrument of the LCLS (29 April – 3 May 2015; proposal LI56) operating at a repetition rate of 120 Hz, with a nominal photon energy of 9.5 keV (that is, a nominal wavelength of 1.3 Å) and an X-ray pulse length of 35 fs. The X-ray beam was focused to 1.3 μm (FWHM), with a beamline transmission of roughly 50%. The transmission was further reduced to 25% of the incident beam intensity, resulting in 0.3 mJ at the sample position. Data were acquired on the CSPAD detector[39], which was used in dual-gain mode with a low gain setting in the inner region to compensate for its relatively low dynamic range and obtain accurate measurements of both high and low resolution intensities. CASS[40,41] was used for online monitoring of the diffraction data, determination of the hit rate and resolution cutoff, and to estimate the fraction of multiple hits. Online real-time analysis was crucial to adjust the injector position and crystal concentration. In order to avoid FEL-related systematic effects (since the FEL can behave like two distinct alternating 60 Hz sources), equal amounts of pump laser-on [1] and pump laser-off [0] data were collected with non-uniform interleaving[42] in the following repetitive sequence: [1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0].

Laser-on data were collected at nominal pump-probe delays of ~1 ps (laser-on-Δ1ps) and 3 ps (laser-on-Δ3ps). When collecting reference data (laser-off data), the pump-laser pulse was set to occur 20 ns after the X-ray pulse. A total of four data sets were collected at room temperature:

i) Reference data set: no pre-illumination, no pump-laser data allowed the determination of the structure of the rsEGFP2 resting state (fluorescent *on* state);

ii) *Laser-off* data set: pre-illumination, no pump-laser data allowed the determination of the structure of rsEGFP2 in the non-fluorescent *off* state;

iii) *Laser-on-Δ1ps* data set: pre-illumination, pump-laser data captured rsEGFP2 intermediate states, 1 ps after initiation of the *off*-to-*on* transition; and

iv) *Laser-on-Δ3ps* data set: pre-illumination, pump-laser data captured an rsEGFP2 intermediate state, 3 ps after initiation of the *off*-to-*on* transition.

Note that dataset ii) was in fact collected interleaved with dataset iii). During collection of dataset iv), however, only pump-laser-on data (light) were collected due to time constraints.

Calculation of absorbed X-ray dose

The diffraction-weighted dose[43] calculated for an rsEGFP2 crystal ($3 \times 3 \times 3\ \mu m^3$; containing 2.5 M ammonium sulphate in 39% of its volume) that is fully hit by the 1.3 μm Gaussian beam ($8x10^{10}$ photons/pulse) of CXI, was estimated with the RADDOSE3D server (http://www.raddo.se/)[44] to be 12 MGy. Note that RADDOSE3D does not yet take into account that photoelectrons escaping from a microcrystal diminish the energy deposited in the crystal, so that the absorbed dose was in effect smaller than 12 MGy. On the other hand, photoelectrons generated in the surrounding buffer within the jet may stream into the injected crystals, increasing the nominal dose.

Time-sorting diffraction images based on arrival time diagnostics

Time resolution in SFX experiments is achieved by temporally separating the 400 nm optical laser pulse (pump) from the following X-ray pulse (probe) by a well-defined, nominal pump-probe delay, in our case 1 ps or 3 ps. Due to the inherent time jitter of the FEL X-ray pulses, *i.e.*, the shot-to-shot temporal deviation from the expected periodic arrival time, and because of thermal drifts of the pump-probe synchronization, the actual and nominal pump-probe delays differ. The actual pump-probe delay was monitored by an arrival-time monitor (the so-called *timing tool*) developed at the LCLS X-ray free electron laser[45]. Briefly, the tool uses the ultrafast free-carrier generation induced by X-rays in a Si_3N_4 membrane to encode the relative arrival time of X-ray and pump-laser pulses. A pixel/time calibration was performed during the experiment and used to associate a correct time delay to every indexed diffraction pattern. The diagnostics showed that the delays were distributed around the nominal delay with a spread of about 400 fs for the (nominally 1 ps) laser-on-Δ1ps data and 250 fs for the (nominally 3 ps) laser-on-Δ3ps data. The time resolution of the experiment is in principle also affected by the group velocity mismatch between the X-ray and optical pulses. However, due to the small jet diameter at the X-ray / pump laser intersection point (3-6 μm), this effect is limited to a few fs. By also taking into account the duration of the photolysis pulses (230 fs), we estimate an overall time resolution of ~430 fs.

Hit-finding and sorting of laser-on and laser-off data

We used NanoPeakCell[46] (https://github.com/coquellen/NanoPeakCell) to select hits, find Bragg-peaks and sort laser-on and laser-off images. NanoPeakCell hit-finding parameters were adjusted after visual inspection of the first diffraction patterns. All patterns with more than 10 pixels above the threshold value of 3000 were considered as hits – a deliberately permissive setting that was chosen to avoid missing potential hits. Laser-on-Δ1ps and laser-off hits were sorted based on the laser-state code present in the LCLS raw '.xtc' data stream. Indexing, integration and merging was performed with CrystFEL 0.6.1[47]. As this was the first use of NanoPeakCell in a time-resolved SFX experiment, the more established pre-processing software CASS[40,41] was also used in parallel for data-sorting, hit- and Bragg-peak-finding and automatic feeding to CrystFEL. The two programs yielded similar indexing, scaling and refinement statistics as well as electron density maps. For practical reasons, NanoPeakCell was chosen for the final off-line re-processing.

The overall hit rate was on average 2.5% due to the low crystal concentration required to prevent clogging of injection tubing and the nozzle (see sub-section on *Microcrystal growth, preparation and injection*). Indexing relied on the Bragg peaks found by NanoPeakCell, which on average extended to 2.18 (±0.4) Å resolution. The overall indexing rate was ~50%. Reaching such a high indexing rate was only possible after optimizing the geometry of the tiles of the CSPAD detector[42]. The sample-detector distance was also refined[48], by repeating the indexing task at various distances and selecting the distance yielding the maximal indexing rate. Distance refinement resulted in Gaussian-like distributions of all three unit-cell parameters. Integration was performed using the "*ring-nocen*" method in CrystFEL, and data were scaled and merged using the "*partialator*" module. Final data statistics, including R_{split}, signal-to-noise ratio (SNR) and CC*, are given in Supplementary table 5. The use of the scaling option in CrystFEL produced data with better merging statistics than the simple Monte-Carlo averaging, but neither did partiality refinement nor post-refinement (including partiality refinement; Supplementary table 6). A direct correlation was established between the Gaussian-like distribution of unit-cell parameters, the merging statistics and the content of difference Fourier density maps[48].

On-state and off-state structures of rsEGFP2: Data collection and refinement

Two reference datasets were collected, nominally corresponding to rsEGFP2 in the *on* and in the *off* state, respectively. Both structures were phased by molecular replacement (MR) using Phaser[49] and refinement was performed using the Phenix suite[50]. Graphic operations in real space were performed using Coot[51]. Figures were produced using PyMOL[525051] or custom-made python scripts and the matplotlib library[53]. All structure factor amplitude difference Fourier maps were Q-weighted as described[54] and produced using a custom-written CNS[55] script[56]. Briefly, Q-weighting assigns a weight to structure factor amplitude differences based on measurement errors (*sigma*), using Bayesian statistics. Q-weighting was introduced in the early days of time-resolved Laue crystallography as a means to improve the estimates of difference amplitudes and to reduce the noise in difference maps[54]. Q-weighting also increases the signal-to-noise ratio of amplitude differences from monochromatic datasets, and was shown to improve difference refinement[57] accuracy[58]. More recently, Q-weighting was used to generate extrapolated structure factors for difference refinement of a photobleached chromophore structure from synchrotron data[59] and to compute difference maps from XFEL data[60]. Despite integration of laser-off, laser-on-Δ1ps and laser-on-Δ3ps data sets at 1.7 Å resolution, Q-weighted difference Fourier maps were calculated at 1.9 Å resolution because of poor Q-weight values in the 1.7 – 1.9 Å resolution range.

rsEGFP2 *on*-state structure: the reference dataset that allowed solving the rsEGFP2 *on*-state structure was collected from rsEGFP2 microcrystals that were neither pre-illuminated at 488 nm nor irradiated with the pump laser at 400 nm. 986,178 patterns were collected, of which 105,612 (10.7%) contained hits and 34,715 (3.5%) were indexed (Supplementary table 5). The structure was phased by MR using the synchrotron cryo-crystallographic (100 K) structure of rsEGFP2 in the *on* state (PDB code 5DTX[34]) as a search model. Subsequent refinement of the *on*-state structure included iterative cycles of model adjustment in real space, and coordinates and individual B-factor refinement in reciprocal space. The statistics (R_{free}, R_{work} = 0.195, 0.167 at 1.7 Å resolution) of the final SFX *on*-state structure (Supplementary Fig. 16a, b) are comparable to those of the cryo-crystallographic structure of rsEGFP2 in the *on* state (R_{free}, R_{work} = 0.206, 0.174 [34] at 1.45 Å resolution). The two structures superimpose with an rmsd of 0.17 Å over 243 Cα atoms.

Pre-illuminated (488 nm), laser-off structure: the laser-off dataset was collected over three experimental shifts from rsEGFP2 microcrystals pre-illuminated at 488 nm, but without pump-laser excitation at 400 nm. 3,531,773 patterns were collected, of which 132,264 (3.7%) contained hits and 65,097 (1.8%) were indexed (Supplementary table 5). A structure factor amplitude difference Fourier map was calculated between the laser-off dataset and the *on*-state reference dataset ($F_{obs}^{laser\text{-}off} - F_{obs}^{reference}$), using phases calculated from the *on*-state structure (Supplementary Fig. 17). A clear negative peak (-10.8 σ) on the *on*-state chromophore and a similarly high positive peak (+7.1 σ) around the coordinates of the *off*-state chromophore indicate that, in the laser-off data set, a large fraction (estimated at around 90%, see below) of the chromophores in the crystal have switched from the *on* to the *off* state upon pre-illumination. This observation is in line with the spectroscopic results obtained on pre-illuminated microcrystals collected at the exit of the pre-illumination device (see sub-section on *Pre-illumination of microcrystals*). The $F_o^{laser\text{-}off} - F_o^{reference}$ map illustrates that the *cis* (*on* state) to *trans* (*off* state) isomerization of the chromophore is accompanied by backbone motions from β-strand-7 residues (Tyr146 to Tyr152) and by conformational changes in the side chains of Tyr146, His149, Val151 and Thr204. These motions are qualitatively identical to those reported earlier[34].

The laser-off dataset was phased by molecular replacement using the *off*-state synchrotron structure of rsEGFP2 determined at 100 K (PDB code 5DTY[34]) as a search model. Examination of the initial $mF_{obs}^{laser\text{-}off}\text{-}DF_{calc}$ map revealed positive peaks (higher than 3 σ) on the *on* state conformations of Tyr146, Val151 and Thr204 side chains, and negative peaks on the *off*-state conformation of the chromophore and His149 side chain. Likewise, the $2mF_{obs}^{laser\text{-}off}\text{-}DF_{calc}$ map (Supplementary Fig. 18) indicated the presence of two conformers for the Tyr146 side chain. Hence, both maps were suggestive of a fraction of rsEGFP2 molecules having remained in the *on* state, in agreement with spectroscopic measurement on pre-illuminated microcrystals (Supplementary Fig. 15). The absence of a positive peak on His149 was tentatively explained by the observation that, in the *off*-state structure, two waters bind at its *on*-state position. These could compensate for the difference in electron density at His149.

We further investigated the hypothesis that pre-illuminated rsEGFP2 crystals contain both *on*-state and *off*-state conformers by including the *on*-state conformation of rsEGFP2 (*on* state SFX structure solved from the reference data set, see above) as an alternative conformation in the refinement. The relative occupancies of the *on*- and *off*-state chromophores were varied manually (from 0 to 45 % and from 100 to 55 %, respectively, by 5 % steps), and the refinement statistics and $mF_{obs}^{laser\text{-}off}\text{-}DF_{calc}$ maps inspected and compared. Refinement statistics (namely R_{free} and R_{work}) would suggest that the occupancy of the *on*-state conformation is 25% in the laser-off SFX structure (Supplementary Fig. 19), but integration of $mF_{obs}^{laser\text{-}off}$ -DF_{calc} map (Supplementary Fig. 19b) reveals negative peaks on the *on*-state chromophore at occupancies higher than 10% (Supplementary Fig. 19b). Thus, one or more long–lived intermediate states, generated by the pre-illumination and invisible in the $2mF_{obs}^{laser\text{-}off}$ -DF_{calc} map, could exist that would display a different conformation of the chromophore and yet the same protein environment as the reference *on* state structure. As it was not possible to model these intermediates from electron density maps, we chose to use a starting model consisting of 90% *off* state and 10% *on* state for the refinement of the laser-off structure, in agreement with spectral measurements on pre-illuminated microcrystals collected at the exit of the pre-illumination device (Supplementary Fig. 15). Water 24, which is present in the chromophore pocket in the *on*-state structure but not in that of the *off* state[34], was also treated as an alternate conformer. Reciprocal space refinement included positional and B-factor refinement for all atoms. The resulting observable-to-parameter ratio was 1.52. The final laser-off structure displays refinement statistics (R_{free}, R_{work} = 0.176, 0.145 at 1.7 Å resolution) that are comparable to those reported for the cryo-crystallographic structure of rsEGFP2 in the *off* state (R_{free}, R_{work} = 0.224, 0.178 at 1.5 Å resolution), and the two structures superimpose with an rmsd of 0.12 Å over 233 Cα atoms. Further details regarding the SFX laser-off structure are listed in Supplementary table 5.

Of important note, it was verified that pre-illumination efficiency was constant during the three experimental shifts which were used to collect the laser-off dataset. To this end, we calculated difference Fourier maps between the laser-off dataset consisting of all data and reduced laser-off datasets produced from data collected during individual shifts only. The three maps, phased using the final laser-off structure (90% *off*, 10% *on*), show no features on the chromophore when displayed at ± 3.5 σ, and only minor peaks at ± 3 σ. Likewise, difference Fourier maps calculated between the laser-off dataset and datasets produced from one-to-one combinations of two shifts were featureless (Supplementary Fig. 20). Altogether, these maps confirm that the laser-off data from the three shifts contain similar structural information, and hence justify merging them to yield a single laser-off dataset and structure.

1-ps intermediate-state structure: data collection, modeling and refinement

The laser-on-Δ1ps dataset was collected over three experimental shifts from rsEGFP2 microcrystals that had been pre-illuminated at 488 nm, excited by a pump-laser at 400 nm, and probed after 1 ps by an XFEL pulse. Acquisition of this dataset and of the laser-off dataset was interleaved (see above). 3,531,773 patterns were collected, of which 132,464 (3.7%) contained hits and 64,620 (1.8%) were indexed (Supplementary Table 5). All patterns were merged into a single dataset referred to hereafter as the laser-on-Δ1ps dataset.

A difference Fourier map was calculated between the laser-on-Δ1ps dataset and the laser-off dataset, using phases calculated from the final laser-off model (90% *off*, 10% *on*). The $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ electron density map shows strong features on the chromophore and its vicinity, when displayed at ± 3.5 σ (Fig. 4a; Supplementary Fig. 21). The peaks are located around the chromophore (residue 68) and, with few exceptions that are outlined below, no peaks are observed at above ± 4 σ on the rest of the protein. Difference Fourier maps calculated from laser-on-Δ1ps datasets that were obtained using either all indexed images, images collected during individual shifts or combinations thereof, all revealed the same features (Supplementary Fig. 22). These six maps illustrate that the structural information present in the various datasets collected with a nominal pump-probe delay of 1 ps is similar, at least qualitatively (Supplementary Fig. 21). Hence, the $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ map calculated from all 1 ps light and dark data was chosen to inspect, model and refine structural changes of the protein at 1 ps after the pump-excitation. The clarity of this map can in part be explained by the high isomorphism between the two datasets compared (Supplementary Table 5).

Negative peaks are seen at the position of the chromophore in the *off*-state (peaks at the phenol hydroxyl OH, the imidazolinone carbonyl O2 and the methylene linker CB2 are at -9.1, -7.0 and -7.9 σ, respectively), suggesting that this conformation is depleted (see Supplementary Fig. 16c for a topology of the chromophore). A large positive peak appears nearby the imidazolinone ring of the chromophore, tipping at the linker atom CB2 (+6.0 σ). This peak highlights a reorientation of the chromophore phenol group and, by its bilobal nature, is suggestive of the presence of two states. In line with this observation, two positive peaks are seen between the *on*- and *off*-state positions of the phenol group (Supplementary Figs. 21 and 23). The $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ map also features peaks on residues in the immediate vicinity of the chromophore (Supplementary Figs. 21 and 23). First, a strong negative peak is seen on Wat17 (-5.4 σ), which is H-bonded to the chromophore OH in the *off*-state conformation. Also, a negative peak is seen on the side chain of Arg97 (strongest peak: -4.0 σ) which, together with Glu223, stabilizes the imidazolinone ring of the chromophore, in both the *on*- and the *off*-state structures of rsEGFP2 (3 Å distance from NH2 to chromophore imidazolinone oxygen O3). Likewise, a negative peak is seen on Wat19 (-4.2 σ), at 2.7 Å from Glu223 (Oε1) and Asn69 (N). Negative peaks are observed on the side chains of Val151 and Phe166 (-3.2 and -3.0 σ, respectively) as well as on the main chain of Val151 (-3.2 σ). Gln95, another important residue for the stabilization of the chromophore, is also highlighted in the $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ map. In both the *on*- and the *off*-states, its side chain nitrogen is H-bonded to the chromophore carbonyl oxygen O3 through Wat8 (at 2.7 and 2.7 Å from the latter two). All three atoms feature negative peaks between -4.2 σ (Gln95 Oε1 and Wat8) and -4.8 σ (chromophore O3). The transition of the chromophore from the *off* to the intermediate state is accompanied by structural changes in the protein backbone, as indicated by a pair of positive and negative peaks seen on the backbone atoms of His149 (+3.2 σ and -3.4 σ). At the C-terminal end of the chromophore, pairs of negative and positive peaks are observed on the backbone carbonyl and side chain of Val69 (+4.2 σ and -3.3 σ), and on the sulphur atom of Cys71 (+4.4 σ and -4.2 σ), indicative of local changes. In addition to these, negative peaks are also seen on the backbone carbonyl oxygen atoms of residues that contribute the α-helix N-terminal to the chromophore (Supplementary Fig. 24). The height of these peaks inversly correlates with their distance to chromophore, being -5.9 σ on Leu65 (*n*-1), -3.9 σ on Thr 64 (*n*-2), and -3.3 σ on Val62 (*n*-4) and Leu 61 (n-5). Note that in Fig. 4, not all $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ peaks are visible, given the complex environment of the chromophore. These are yet illustrated in Supplementary Figs. 23 and 24.

The peaks observed in the $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ map suggest that part of the chromophores have moved away from their *off*-state position, and that an intermediate has built up, approximately mid-way between the *on*- and the *off*-state conformations of the chromophore. In order to assess whether these peaks indeed correspond to a new conformer, and not to a changed mixture of *on*- and *off*-state conformers[61], we calculated structure factor amplitudes (F_c) to 1.7 Å resolution from a model consisting of a 1:1 mixture of *on*- and *off*-state conformations and assigned to each F_c the experimental sigma value observed in the laser-on-Δ1ps dataset. From this simulated dataset, a $2mF_o$-DF_c electron density map was calculated, which shows a clear separation of the densities of the two states at 1.7 Å resolution – and not a merged single continuous peak between the two conformers. Hence, the positive $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ peaks observed mid-way between the *on*- and the *off*-state conformations of the chromophore most likely indicate the buildup of one (or two) intermediate(s) state(s).

These two $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ peaks are at similar height, *viz.* +5.4 (peak 1) and +6.3 σ (peak 2), and each is 2.8 Å away from a negative peak, covering either the *off*-state chromophore OH or Wat17 (Supplementary Fig. 25). The distance between peak 2 and the *off*-state chromophore OH is 1.5 Å, while that between peak 1 and the *off*-state chromophore CE2 is 1.1 Å. Hence, we considered two possible intermediate state models. In the first model (model P), the chromophore OH and CE1 were placed into peaks 2 and 1, respectively, but at the cost of a suboptimal fit into the peak on the imidazolinone ring (model P; Supplementary Fig. 25a, b). Wat17 was placed in a minor positive peak at 2.4 and 2.7 Å from Gln184 Nε2 and the chromophore OH, respectively. In the second model (model T; Supplementary Fig. 25c, d), we fitted the chromophore OH in peak 1 and the base of the phenol ring in the bifid peak on the imidazolinone ring. Peak 2 was accounted for by placing Wat17 at 2.7 Å distance from the chromophore OH. In this model, the chromophore displays a twisted geometry in that its imidazolinone and phenol rings are quasi perpendicular to one another – as opposed to being coplanar, as observed in the *on*- and *off*-state structures. Thus, in the two models, we initially accounted for the preservation of the H-bond between Wat17 and the chromophore OH.

After phasing by molecular replacement using the final laser-off structure (consisting of 90% *off*-state conformer and 10% *on*-state conformer) as a search model, no density was visible in the $2mF_o^{laser\text{-}on\text{-}\Delta 1ps}$ - DF_c (contoured at +1 σ) and $mF_o^{laser\text{-}on\text{-}\Delta 1ps}$ - DF_c maps (contoured at ±3 σ) that would have indicated the presence of model P or model T. Simulated maps (using calculated structure factors F_c and experimental sigma values) computed to 1.7 Å resolution from models consisting of a *i:j* mixture of the *off*-state conformation (all protein and chromophore atoms) and either model P or model T (only chromophore atoms) reveal that both intermediates would have been visible in $2mF_o^{laser\text{-}on\text{-}\Delta 1ps}$ - DF_c and $mF_o^{laser\text{-}on\text{-}\Delta 1ps}$ - DF_c electron density maps, if their occupancy had been higher than 25%. Therefore, the putative intermediate state(s) undraped by the $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ map is (are) presumably populated to less than 25%.

It was shown that in cases where insight into small structural changes was needed and where modeling errors for the two structures are correlated – as in the present case, difference refinement should lead to improved estimates of structural differences[57] – as compared to conventional refinement. Bayesian weighting of structure factor amplitude differences was further proposed to improve these estimates[62]. Q-weighting is one such Bayesian statistics approach, which assigns a weight to each amplitude difference based on the experimental error (i.e. *sigma* values) of subtracted structure factors[58,57]. Here, Q-weighted difference amplitudes were used to produce extrapolated structure factor amplitudes for refinement of the laser-on-Δ1ps data set using the following formula[59] :

$$F_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps} = \alpha * Q/\langle Q\rangle * (F_{obs}^{laser\text{-}on\text{-}\Delta 1ps} - F_{obs}^{laser\text{-}off}) + F_{obs}^{laser\text{-}off} \text{ (eq. 1)},$$

where α is inversely proportional to the fraction of molecules that changed conformation and Q is the Bayesian weight retrieved through the Q-weighting procedure[54]. <Q> was 0.91 for the laser-on-Δ1ps *vs*. laser-off data. It is of important note that since Q is solely determined from comparisons of $F_{obs}/\sigma F_{obs}$ in the two datasets, all terms in this equation are experimental or based on experimental measurements, and no calculated structures factors (F_{calc}) are involved. Hence, our extrapolated structure factors are not biased by the laser-off structure. Despite integration of laser-off, laser-on-Δ1ps and laser-on-Δ3ps data sets at 1.7 Å resolution, difference refinement was carried out at 1.9 Å resolution because of poor Q-weight values in the 1.7 – 1.9 Å resolution range.

The first step in the difference refinement was to determine α. To this end, we compared the experimental $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps} - F_{obs}^{laser\text{-}off}$ map with the $mF_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps}$ -DFc maps obtained from difference refinement of the *off*-state structure (*i.e.* the same structure that was used to phase the experimental $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps} - F_{obs}^{laser\text{-}off}$ map) against $F_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps}$ calculated for various α values. At this stage, only rigid-body and B-factor refinement was performed, as our intention was to obtain a replica of $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps} - F_{obs}^{laser\text{-}off}$ features in the $mF_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps}$ -DFc maps. We were expecting that, at the appropriate α value, the latter maps would feature peaks of the same height and at the same position than those observed in the former map (Supplementary Fig. 26) – *i.e.* a negative peak on the *off*-state chromophore, and positive peaks at the positions of models P and T. To estimate the similarity between maps, the ratio between absolute integrated positive and negative peaks (above 3.0 σ) in the $F_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps} - F_c^{laser\text{-}off}$ and the $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps} - F_{obs}^{laser\text{-}off}$ maps was calculated for α-values ranging from 1 to 50 – corresponding to occupancies of 100 to 2 % (Supplementary Fig. 26b, c). We obtained a curve of sigmoidal appearance that started leveling of at α≈15, corresponding to a 6.7 % decrease in occupancy of *off*-state conformers (Supplementary Fig. 26c). The agreement between difference maps continued to improve in the range $15 \leq \alpha \leq 30$. Furthermore, electron density became visible in extrapolated $2F_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps} - F_{calc}^{laser\text{-}off}$ maps that overlaid with phenol atoms of the chromophore in models P and T – although neither was included in the model that was used to phase these maps (Supplementary Fig. 26d, e). Challenged by the ability offered by the difference data to model such low occupancies, we used a second approach to estimate the occupancies of models P and T, whereby calculated structure factors were generated for two hypothetical models, consisting in i) 100 % *off*-state (F_{calc}^{dark}) and ii) a 50:50 mixture of model P and model T (F_{calc}^{light}); the negative peak on the *off*-state chromophore in this simulated F_{calc}^{light}-F_{calc}^{dark} map accounts for 100% of conversion from the dark to a light structure. Comparison of the integrated values of this peak in the simulated F_{calc}^{light}-F_{calc}^{dark} and experimental $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ maps also suggests a 6.9 % decrease in occupancy of *off*-state conformers in the laser-on-Δ1ps dataset.

At this point, we could safely assume that, overall, ~7% of microcrystalline rsEGFP2 monomers had departed from the *off*-state, but it remained unclear whether model P, model T or a combination thereof, offered the best model of the difference data. Therefore, we generated a model with both conformers present, and refined it against extrapolated structures factors generated with $\alpha = 15$ (Supplementary Fig. 26c). Reciprocal space refinement included positional and B-factor refinement for all atoms, and occupancy refinement for model P and model T chromophores. Occupancies of the latter were respectively refined to 0.6 and 0.4, suggesting that, in the difference data, the two conformers co-exist. At $\alpha = 15$, relative occupancies of 0.6 and 0.4 correspond to absolute occupancies of ~4 % and ~3 % for model P and model T, respectively. Note that in control calculations, the occupancies of model P and model T did not vary much as a function of α (Supplementary Fig. 26f). From there on, refinement was straightforward, and included real space fitting of residues in the $2F_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps}$ - F_{calc} electron density map. Standard equilibrium restraints (including bond, angle, torsion angle and separate planarity restraints for the imidazolinone and phenol groups) were used in the crystallographic refinement – *i.e.,* the same restraints that were used for the refinement of the static *off*- and *on*-state structures. Residues that showed conformational changes were modeled in alternate conformation, which we hypothesized to be associated to model T (and not to model P) and whose occupancy were set accordingly, due to the stronger dissimilarity of this model with the *off*-state structure (see below). R_{free}/R_{work} of the final model are 0.293/0.275 at 1.9 Å resolution.

The chromophore in model P displays a conformation close to the *off*-state chromophore, and the C2-CA2-CB2-CG2 (*tau*) and CA2-CB2-CG2-CD2 (*phi*) dihedral angles change by -

50° and +52°, respectively. The chromophore conformation in model T, however, displays an imidazolinone ring that is nearly perpendicular to that of the phenol group (Fig. 4c, d), as illustrated by respective changes in *tau* and *phi* of -84° and +43°, with respect to the *off*-state (Fig. 4c, d). We hypothesize that the similarity of model P to the *off*-state chromophore indicates that no major conformational changes in the protein scaffold accompanies the slight change in chromophore conformation. Consequently, conformational changes in the difference-refined laser-on-Δ1ps structure with respect to the *off*-state structure are interpreted as accompanying chromophore twisting towards model T. The most notable conformational change is that occurring in the side chain of Thr204 (β-strand-10), which rotates by 120° in model T (Supplementary Fig. 23). Hence, the *on*-state, the *off*-state and model T are each associated to a specific rotamer of Thr204. In model T, the chromophore phenol OH is at 2.9 and 3.1 Å from Wat356 and Wat17, respectively. Wat17 is H-bonded to Gln184(Nε1) and Ser164(Oγ) in both the *off*-state structure and in model P. As Ser164Oγ is itself a donor in the H-bond it forms with Tyr52(O), the chromophore OH is an obligate donor in the H-bond it establishes with Wat17 in the *off*-state. The same applies to Wat356, which is H-bonded to His149(O), Val151(N) and Leu202(O) and Thr204(OH) in the *off*-state structure of rsEGFP2. Upon transition of the chromophore to model T, steric hindrance forces Thr204 side chain to rotate by 120°, resulting in the breaking of its H-bond with Wat356, and in the latter becoming available for H-bonding to the chromophore phenol-OH. Upon transition to model T and formation of this H-bond, that with Wat17 is forced to break, since the chromophore OH can only be donor in a single H-bond. Thus, this water was kept in model P, but not in model T (Supplementary Fig. 23). Also due to steric hindrance, His149 side chain draws back from the chromophore phenol, but no rotation of its side chain is suggested by the data. Around the imidazolinone ring of the chromophore, side chains of several residues display conformational changes, including Leu43 (β-strand-3), Gln95, Arg97 (β-strand-5) and Glu223 (β-strand-11). Minor conformational changes are also observed in residues Val151 (β-strand-7) and Phe166 (β-strand-8).

A distance difference matrix (DDM) was calculated between the difference refined laser-on-Δ1ps structure and the *off*-state conformer (Supplementary Fig. 27a). This DDM illustrates that in addition to the above-described local-changes, the β-barrel expands upon the transition from the *off*-state to model T, and thus offers a peek into concerted conformational changes across the protein. Most notably, β-strands-7 and -10 draw away from the rest of the structure and from the chromophore (Supplementary Fig. 27a, b).

Difference Fourier maps calculated with data collected after a 3 ps pump-probe delay

The laser-on-Δ3ps dataset was collected during three hours from rsEGFP2 microcrystals pre-illuminated at 488 nm, excited by a pump-laser at 400 nm, and probed after 3 ps by XFEL pulses. 822,721 patterns were collected, of which 34,750 contained hits and 22,259 were indexed (Supplementary Table 5). Due to time constraints, we did not interleave the collection of this dataset (laser-on-Δ3ps) with that of a laser-off dataset (cf. section *SFX data collection and on-line monitoring*) but instead used as a reference dataset, the laser-off dataset collected along with the laser-on-Δ1ps dataset. This choice was justified because the pre-illumination setup was stable during the three experimental shifts used for collection of the laser-on-Δ1ps dataset, resulting in similar levels of *on*-to-*off* photoswitching efficiency. Indeed, difference Fourier maps calculated between the laser-off dataset and reduced laser-off datasets produced on a per-shift basis or from one-to-one combinations of shifts, were essentially featureless (see Supplementary Fig. 20).

The $F_{obs}^{\text{laser-on-}\Delta\text{3ps}}\text{-}F_{obs}^{\text{laser-off}}$ map was computed using phases calculated from the laser-off structure (consisting of 90% *off*-state and 10% *on*-state conformer). For a valid comparison, we generated a reduced laser-on-Δ1ps dataset, by randomly selecting as many indexed patterns in the laser-on-Δ1ps data, as recorded for the laser-on-Δ3ps dataset (*i.e.* 22,262 patterns). This dataset is hereafter referred to as laser-on-Δ1ps-random dataset (Supplementary Table 5) and was used to compute a $F_{obs}^{\text{laser-on-}\Delta\text{1ps-random}}\text{-}F_{obs}^{\text{laser-off}}$ map. As expected, features in this map are weaker than in the $F_{obs}^{\text{laser-on-}\Delta\text{1ps}}\text{-}F_{obs}^{\text{laser-off}}$ map, and hence better visualized at a contour level of ± 3.0 (instead of ± 3.5 σ for the $F_{obs}^{\text{laser-on-}\Delta\text{1ps}}\text{-}F_{obs}^{\text{laser-off}}$ map). The two maps are nevertheless qualitatively similar, showing negative peaks on the *off*-state chromophore (strongest peaks at -10.9 vs. -6.7 σ, in the $F_{obs}^{\text{laser-on-}\Delta\text{1ps}}\text{-}F_{obs}^{\text{laser-off}}$ and $F_{obs}^{\text{laser-on-}\Delta\text{1ps-random}}\text{-}F_{obs}^{\text{laser-off}}$ maps, respectively) and Wat17 (-6.4 vs. -4.5 σ), and positive peaks on the model T chromophore (5.4 vs. 4.0 σ) and model P chromophore (6.3 vs. 4.5 σ).

The $F_{obs}^{\text{laser-on-}\Delta\text{3ps}}\text{-}F_{obs}^{\text{laser-off}}$ map differs significantly from the two laser-on-Δ1ps maps (Fig. 4; Supplementary Fig. 28 and 29). Both the negative peak observed on the *off* state chromophore and the positive peaks on the chromophore models P and T are lower in the $F_{obs}^{\text{laser-on-}\Delta\text{3ps}}\text{-}F_{obs}^{\text{laser-off}}$ map (Supplementary Fig. 28 and 29). The large positive peak observed on the imidazolinone ring (peaking at 5.9 σ on the CB2 atom) in the $F_{obs}^{\text{laser-on-}\Delta\text{1ps-random}}\text{-}F_{obs}^{\text{laser-off}}$ map is strongly diminished. Likewise, positive peaks on the phenol-OH of model P and model T are respectively at 2.4 and 3.9 σ in the $F_{obs}^{laser-on-\Delta3ps}$-$F_{obs}^{laser-off}$ map, to compare to 4.5 and 4.0 σ in the $F_{obs}^{laser-on-\Delta1ps-random}$-$F_{obs}^{laser-off}$ map. It is thus likely that model P, which was populated to 60% in the difference-refined laser-on-Δ1ps structure, has nearly vanished at 3 ps. To estimate by how much the population of model T evolves between 1 and 3 ps, we integrated peaks around the chromophore in the $F_{obs}^{laser-on-\Delta3ps}$-$F_{obs}^{laser-off}$ map (*i.e.*, all positive peaks on the chromophore were assigned to model T), and compared this value to 40% (occupancy of model T in the difference-refined laser-on-Δ1ps structure) of that obtained from the integration of peaks around model T and model P in the $F_{obs}^{laser-on-\Delta1ps-random}$-$F_{obs}^{laser-off}$ map. Doing so, we found that model T decreases by 44 % (Supplementary Fig. 28b, c). In line with the decrease in occupancy of the intermediate states (-100% and -44 % for model P and model T, respectively), a 37% decrease is observed of the integrated value of negative peaks around the *off*-state chromophore, when the $F_{obs}^{laser-on-\Delta1ps-random}$-$F_{obs}^{laser-off}$ and the $F_{obs}^{laser-on-\Delta3ps}$-$F_{obs}^{laser-off}$ are compared, suggesting that at least a fraction of the intermediate states have fallen back to the *off*-state. In addition, positive peaks are seen in the $F_{obs}^{laser-on-\Delta3ps}$-$F_{obs}^{laser-off}$ map, which overlay with the position of the *on*-state chromophore (peaks at 4.8, 3.7 and 3.2 σ on CB2, OH and CG2) and are absent in the $F_{obs}^{laser-on-\Delta1ps-random}$-$F_{obs}^{laser-off}$ map. Thus, the crystallographic evidence suggests that a fraction of the intermediate states have progressed towards the *cis* conformation – hence completing isomerization.

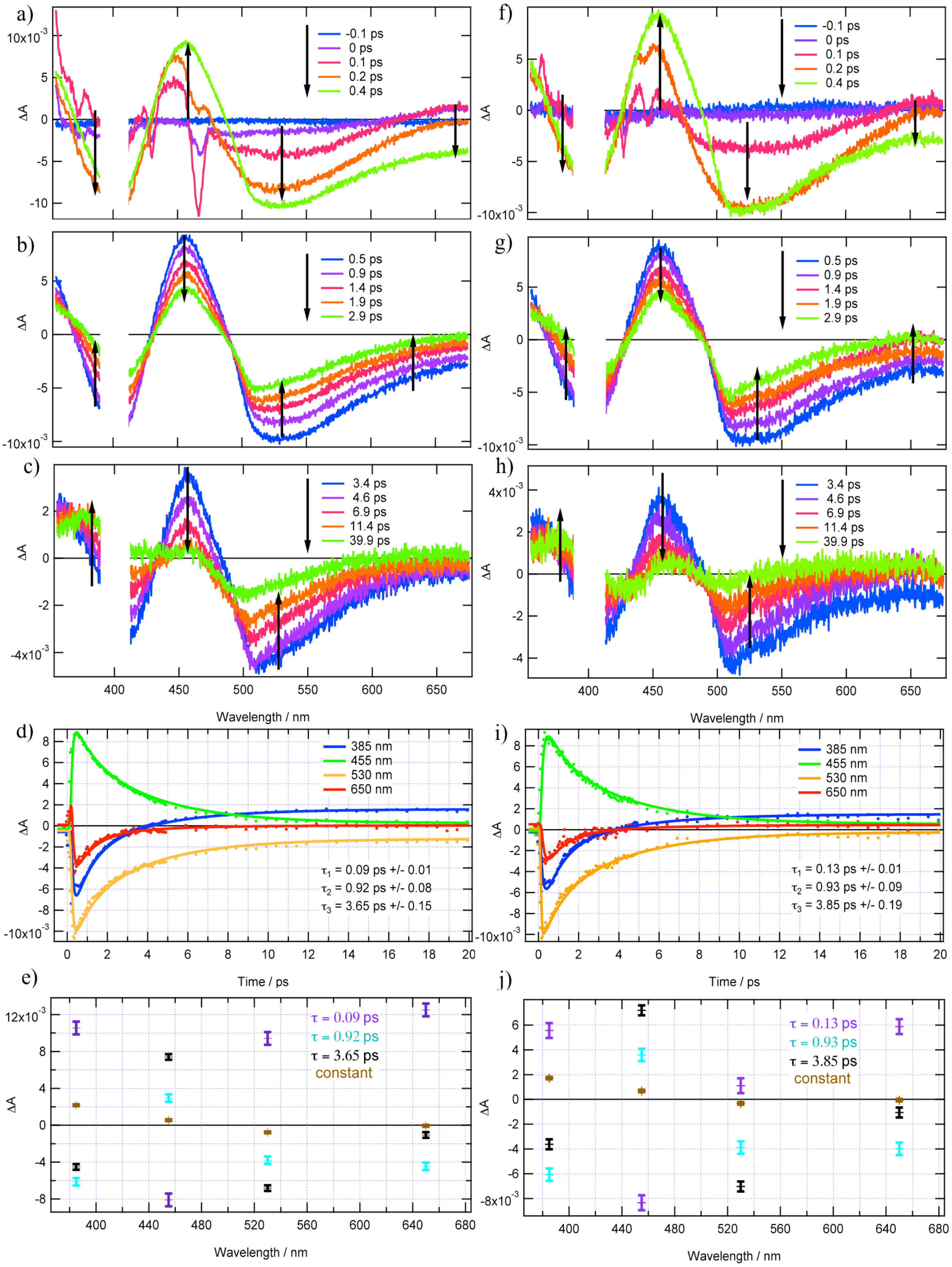
Further insight into intermediate-state conformations and occupancies at 3 ps could have been obtained from difference refinement. Due to the lower quality of the laser-on-Δ3ps dataset, however, the agreement between difference maps is lower and shows a noisier evolution as a function of α, than observed with the laser-on-Δ1ps datasets (Supplementary Fig. 26c). Thus, we refrained from further pursuing difference refinement of the 3ps dataset, and restrained ourselves to a qualitative description of features in the $F_{obs}^{laser-on-\Delta3ps}$-$F_{obs}^{laser-off}$ map (Supplementary Fig. 28 and 29).

4. Site-directed mutagenesis, purification and spectroscopic characterisation of rsEGFP2 and rsEGFP2-V151A

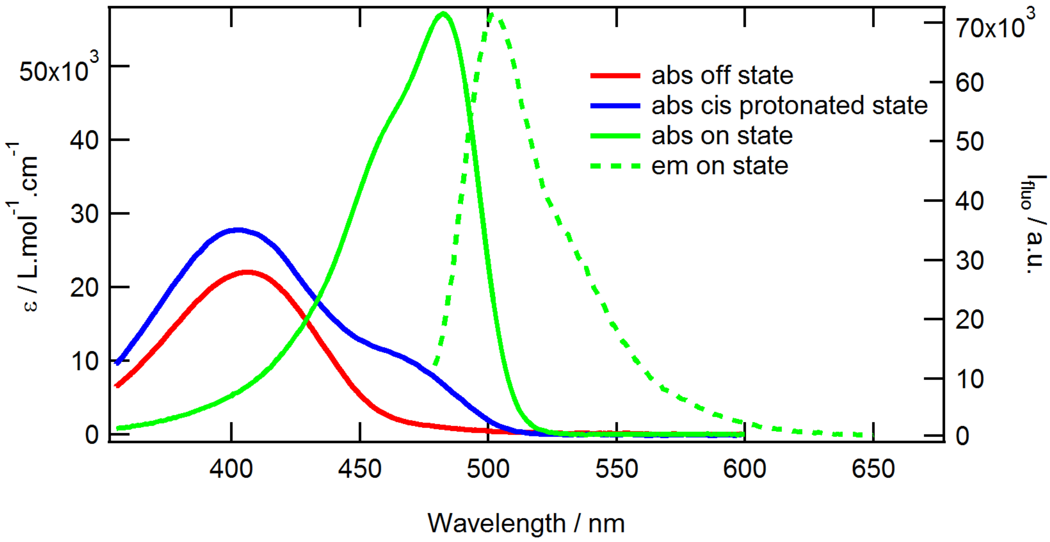
The primer for V151A substitution in rsEGFP2 (5′-CAACAGCCACAACGCCTATATCATGGCC-3′) was 5’-phosphorylated with 10 U of T4 polynucleotide kinase (New England Biolabs, Ipswich, USA). PCR was carried out in a 50-µl mixture using 1 U of Q5 Hot Start High-Fidelity DNA Polymerase (New England Biolabs) and 20 U of 9°N DNA Ligase (New England Biolabs). Digestion of the template methylated DNA was achieved by incubating the PCR product with 10 U *Dpn*I (New England Biolabs) at 37°C for 1 h. Then the sample was subjected to two more PCR cycles, ensuring the polymerisation of the reverse strand from digested fragments as primers. Competent *E. coli* DH5α cells were transformed by the final product (5 µL/50 µL cells) for DNA minipreps and sequencing. Fluorescent proteins (rsEGFP2 and rsEGFP2-V151A) fused to an N-terminal polyhistidine tag were expressed in *E. coli* BL21 (DE3) grown in an auto-inducible medium. After cell lysis, the fluorescent proteins were purified by Ni-NTA affinity chromatography.

Absorption spectra were recorded using a Jasco V-630 UV/VIS photospectrometer (Easton, USA). Emission spectra were measured with a CCD-based spectrometer (AvaSpec-ULS2048, Avantes, Eerbeek, The Netherlands) coupled with optic fibres to a cuvette holder. Photoinduced fluorescence switching cycles were measured as described[34] at low laser power densities: 6.7 mW/cm² at 488 nm (to both excite and switch-off fluorescence) and 1 mW/cm² at 405 nm (to promote *off*-to-*on* photoswitching). Photoswitching quantum yields were calculated by a custom-made routine in MATLAB (The MathWorks Inc., Natick, Massachusetts, USA) following a method reported earlier[63]. Briefly, a kinetic model involving one-to-one reversible conversion between the *on* and the *off* state was used to fit photoswitching raw data (Supplementary Fig. 30). Molar extinction coefficients of *on* and *off* states (Supplementary Table 7) were calculated using the Ward method[64].

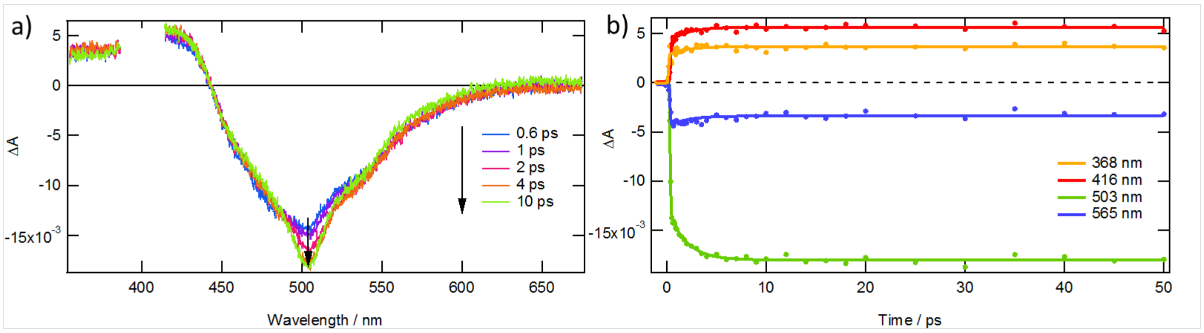
Supplementary figures



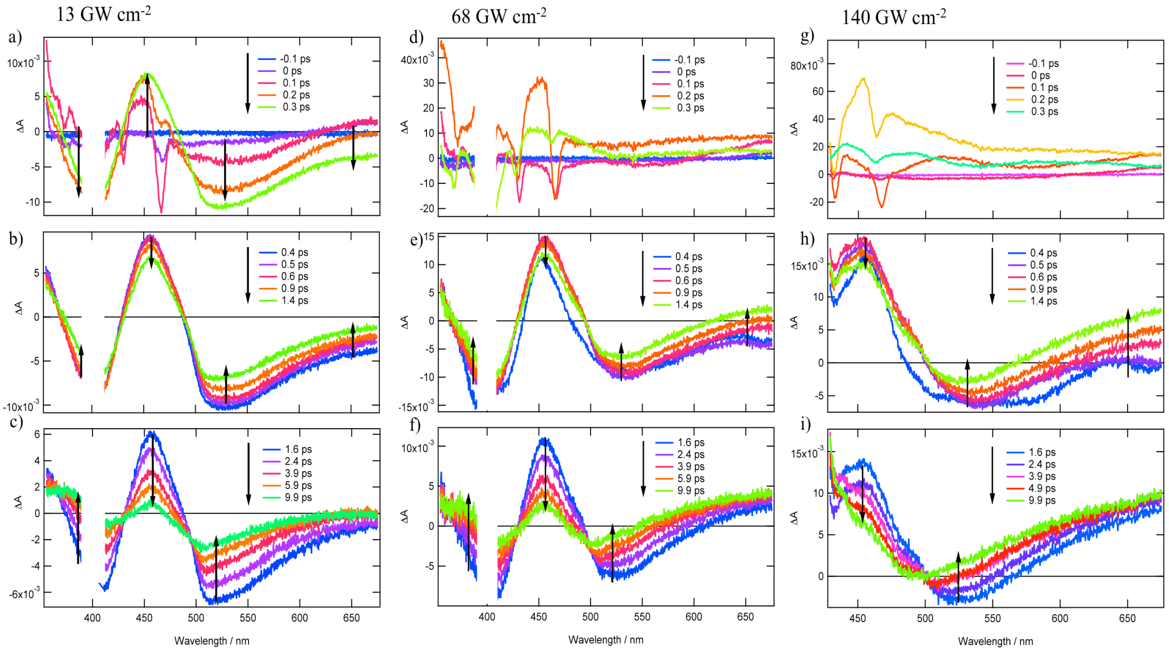
Supplementary Figure 1: Ultrafast transient difference absorption spectra of rsEGFP2 (*trans* protonated *off* state) in H_2O (a - d) and D_2O (f - i) under continuous irradiation at 470 nm at pH 8/pD 8 recorded after excitation by the 400 nm femtosecond pulse (a, f) from -0.1 to 0.4 ps, (b, g) from 0.5 to 2.9 ps, (c, h) from 3.4 to 39.9 ps, (d, i) global fitting with four wavelengths using three exponential functions and a constant convoluted with a Gaussian pulse of 110 fs (FWHM; energy density 1.3 mJ/cm^2) and (e, j) pre-exponential factors associated to the kinetics at 385, 455, 530 and 650 nm.



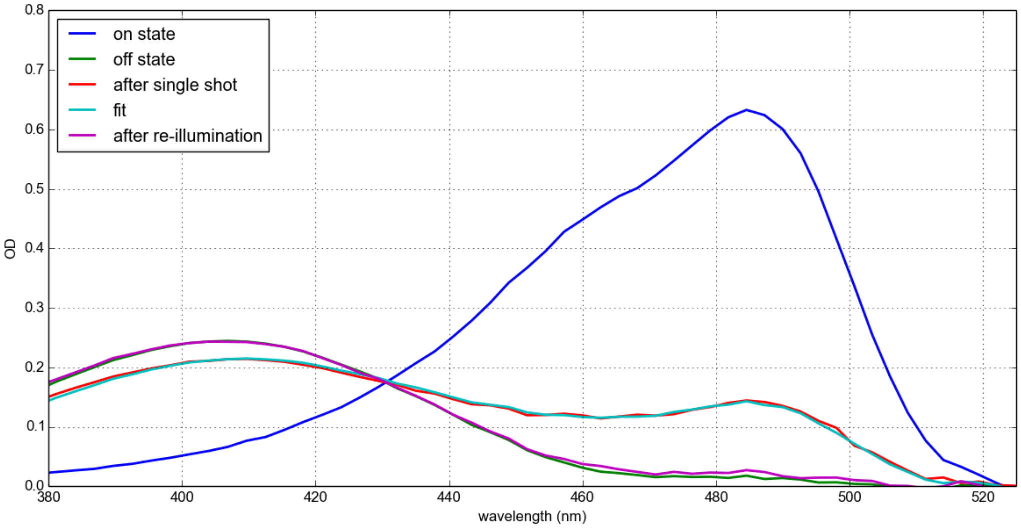
Supplementary Figure 2: Steady state absorption spectra (molar absorptivity) of rsEGFP2 in its *off* state at pH 8 (red), its *on* state at pH 4 (blue), its *on* state at pH 8 (green) and emission spectrum (excitation at 465 nm) of rsEGFP2 *on* state at pH 8 (green dashed line).



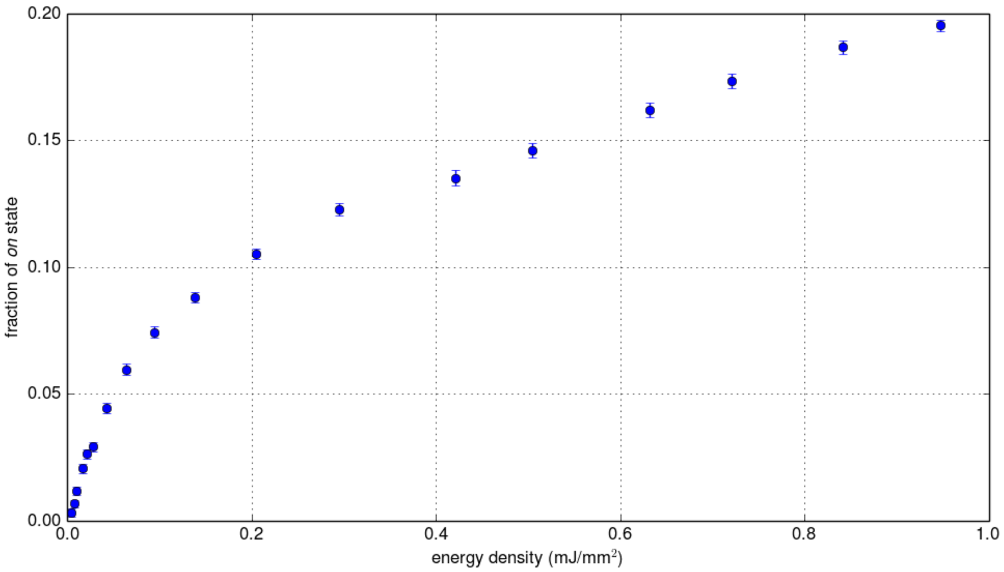
Supplementary Figure 3: (*a*) Ultrafast transient difference absorption spectra of rsEGFP2 (*cis* anionic *on* state) in solution at pH 8 recorded after excitation by a 400 nm femtosecond pulse from 0.6 to 10 ps. (*b*) Examples of the kinetics at four different wavelengths and their respective fit obtained by a global analysis with one exponential function (1.78 ps +/- 0.17 ps, vibrational relaxation) and a constant convoluted with a Gaussian pulse of 166 fs (FWHM).



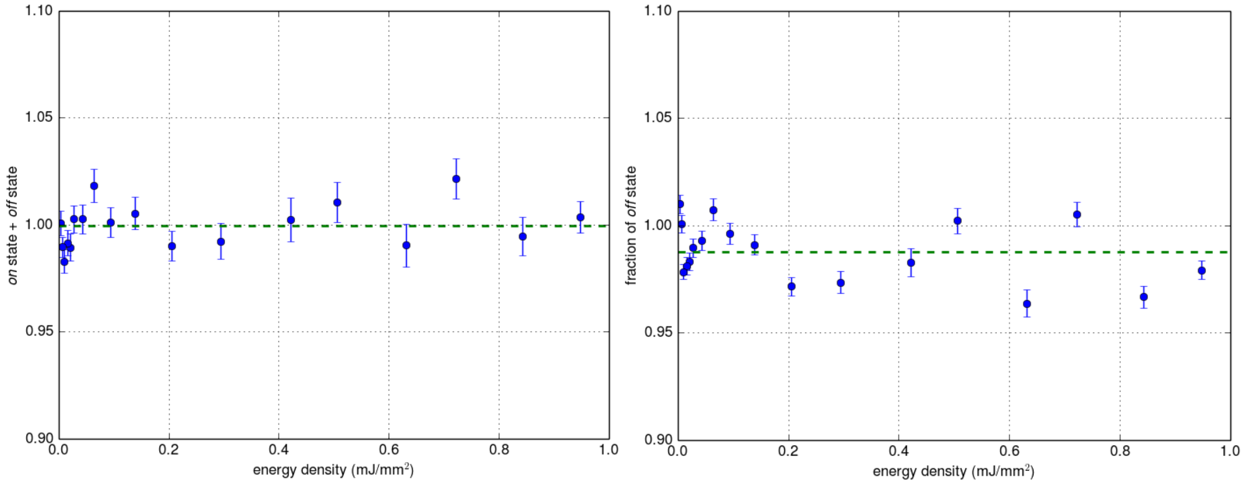
Supplementary Figure 4: Ultrafast transient difference absorption spectra of rsEGFP2 (*off* state) in H_2O under continuous irradiation at 470 nm at pH 8 recorded after excitation by a 400 nm femtosecond pulse from -0.1 to 0.3 ps (a, d, g), from 0.4 to 1.4 ps (b, e, h), and from 1.6 to 9.9 ps (c, f, i) for three different pump power densities (13 GW cm^{-2} (left), 68 GW cm^{-2} (middle), and 140 GW cm^{-2} (right).



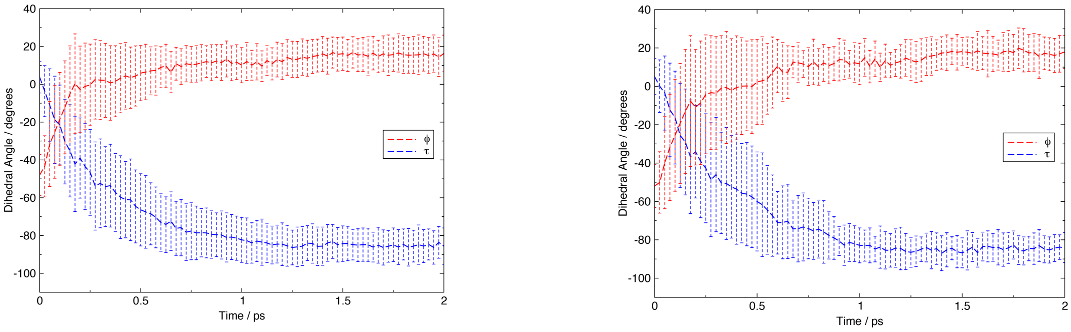
Supplementary Figure 5: *On* and *off* reference spectra (blue and green curves, respectively), spectrum after 400 nm femtosecond excitation of the *off* state at 950 μJ/mm^2 (red curve) and result of the corresponding fit (cyan curve), and spectrum after 490 nm LED re-illumination (magenta curve).



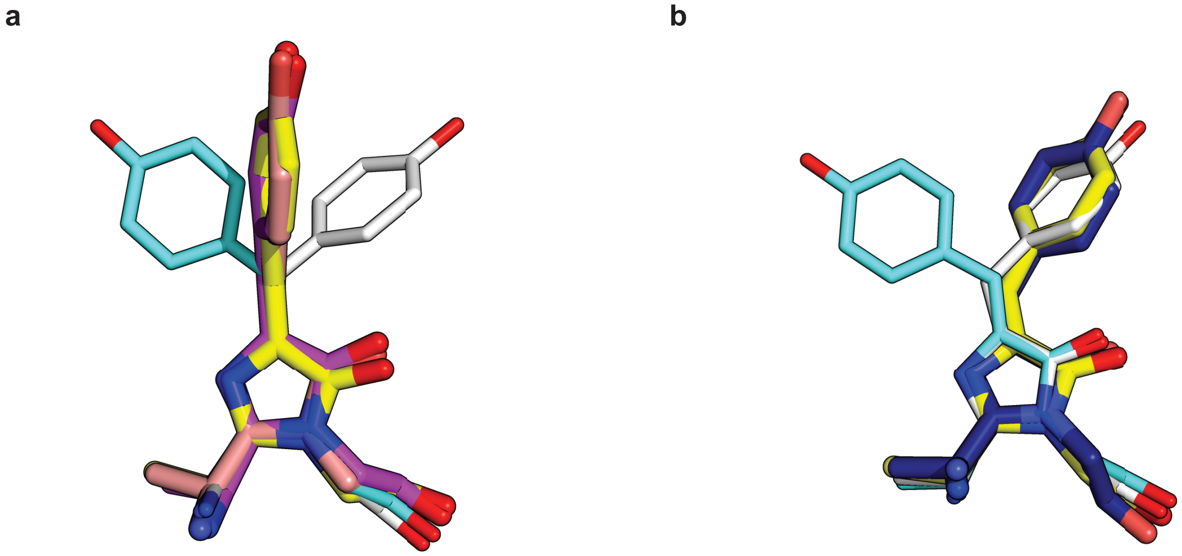
Supplementary Figure 6: Fraction of *on* state after 400 nm laser excitation as a function of the energy density in the range 0-1 mJ/mm^2. Points are averaged over several (typically 5) repetitions.



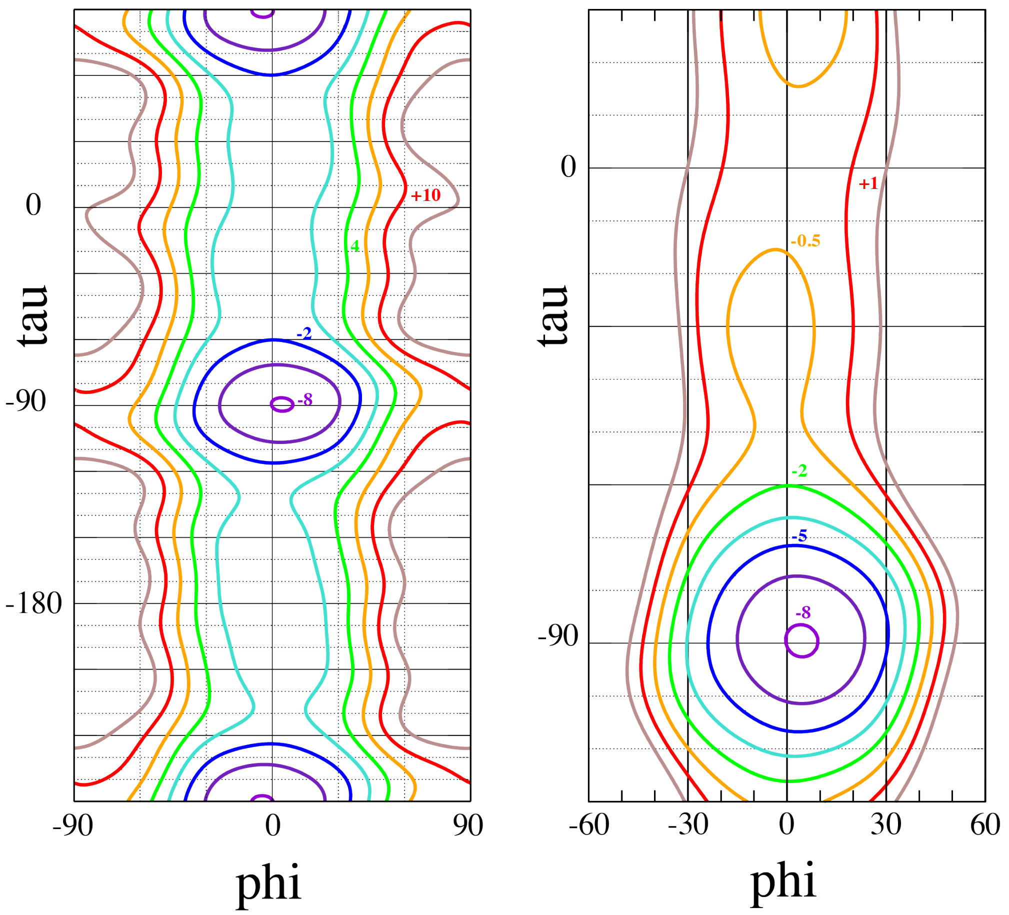
Supplementary Figure 7: Left panel: sum of the fractions of *on* and *off* states after 400 nm laser excitation as a function of the energy density in the range 0-1 mJ/mm^2. Right panel: fraction of *off* state after 490 nm LED re-illumination in the same energy density range. Points in both plots are averaged over several (typically 5) repetitions. Dashed green lines indicate the mean value of all points in every plot.



Supplementary Figure 8: The change in the *tau* (τ, blue) and *phi* (ϕ, red) dihedral angles of the chromophore in rsEGFP2 upon excitation to S_1 as a function of time obtained from QM/MM simulations. The curves correspond to the mean values calculated from the structures of 100 $S_0 \rightarrow S_1$ excitation trajectories; the error bars indicate one standard deviation on either side of the mean (left panel). The right panel shows results from a subset of 28 (out of 100) trajectories. These were selected because they had starting structures whose vertical $S_0 \rightarrow S_1$ excitation energy was within one third of a standard deviation of the average excitation energy. Note that surface hopping from the S_1 to the S_0 state was not permitted during these trajectories.

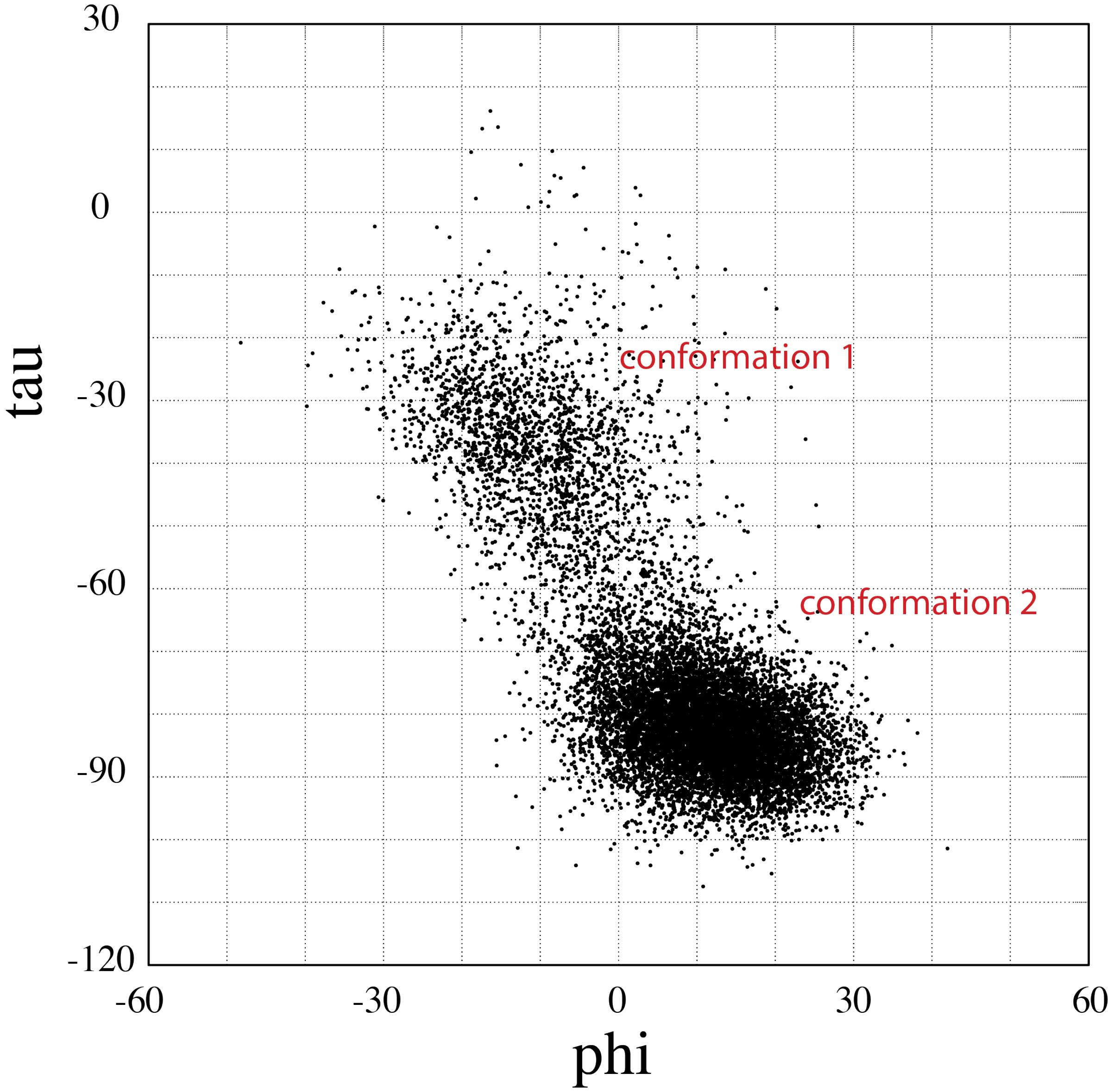


Supplementary Figure 9: Chromophore conformations determined by TR-SFX, excited-state QM/MM and excited-state classical MD simulations. (*a*) Chromophore conformation averaged over 100 S_1 QM/MM excited-state trajectories at 1 ps (magenta), superimposed with the chromophore conformation in the *off* (white), *on* (cyan) and the 1 ps intermediate state corresponding to model T (salmon) determined by TR-SFX. The major conformation found for S_1 by MD is shown in yellow. (*b*) The minor conformation found for S_1 by MD (yellow) is superimposed with the chromophore conformation in the *off* (white), *on* (cyan) and the 1 ps intermediate state corresponding to model P (blue) determined by TR-SFX.

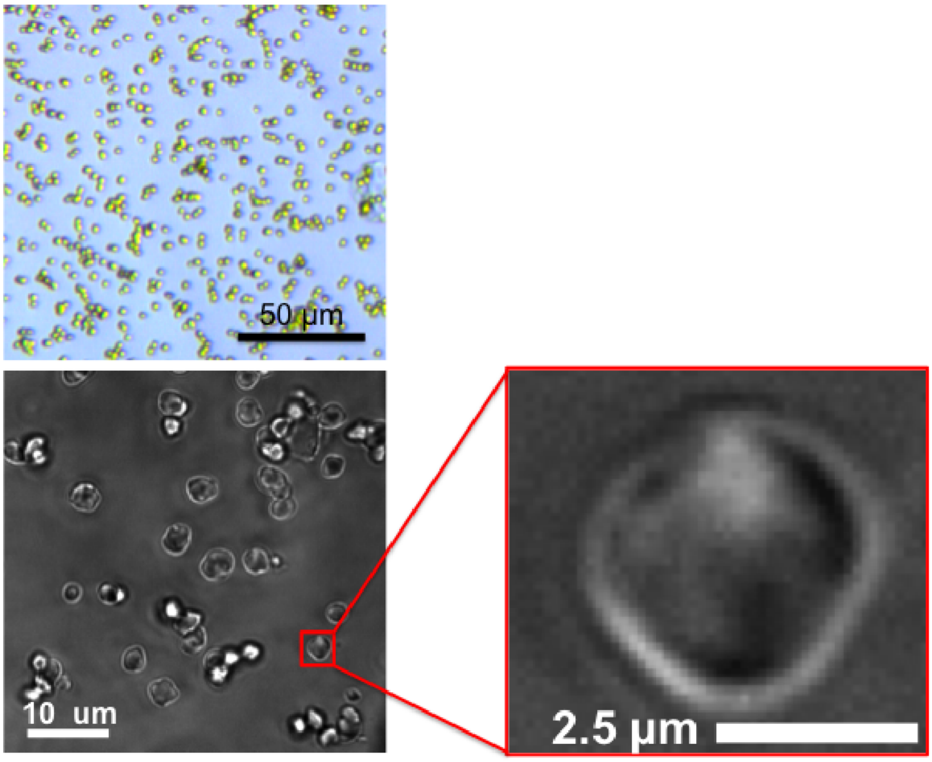


Supplementary Figure 10 : Energy landscape of the excited neutral chromophore in vacuum as a function of *phi* and *tau*. Energy values (kcal/mol) are indicated on the energy contour lines. The energy reference is that of the *trans* planar geometry (*phi* = *tau* = 0°). (left panel) landscape of the entire domain. (right) zoom on the domain explored in the present simulations. Analytical expression of the energy surface : $E = 6.093 - 8.821*\cos(2\varphi) - 0.524*\cos(2\varphi)\ \cos(\tau) + 0.547*\cos(6\varphi) + 2.717*\cos(2\tau) - 2.091*\cos(4\tau) + 0.635*\cos(6\tau) + 0.604*\cos(\tau) + 0.673*\sin(2\varphi)\ \sin(3\tau) - 0.701*\cos(2\varphi)\cos(3\tau) - 0.680*\sin(2\varphi)\ \sin(8\tau) + 0.4188*\sin(4\varphi)\ \sin(8\tau) + 0.370*\cos(2\varphi)\cos(6\tau) + 0.214*\sin(4\varphi)\ \sin(3\tau) - 0.203*\cos(2\varphi)\cos(5\tau) + 0.142*\cos(3\tau)$

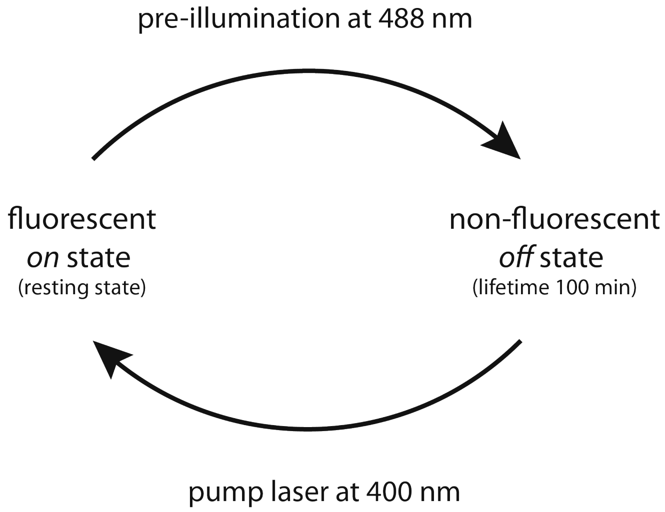
Note that in this expression φ is identical to *phi*, and τ is defined as the dihedral N2-CA2-CB2-CG2 (τ = *tau* + 180°).



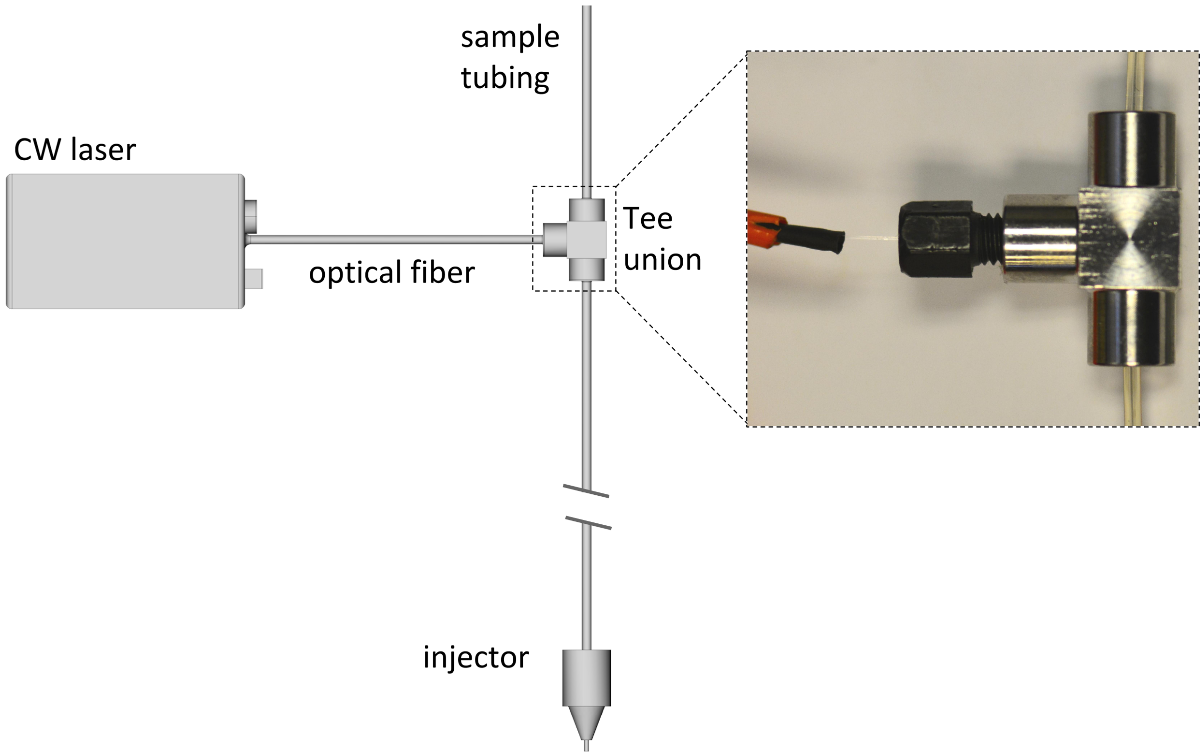
Supplementary Figure 11: Chromophore dihedral angles *tau* and *phi* observed in the first 5 ps of 50 S_1 simulations of rsEGFP2.



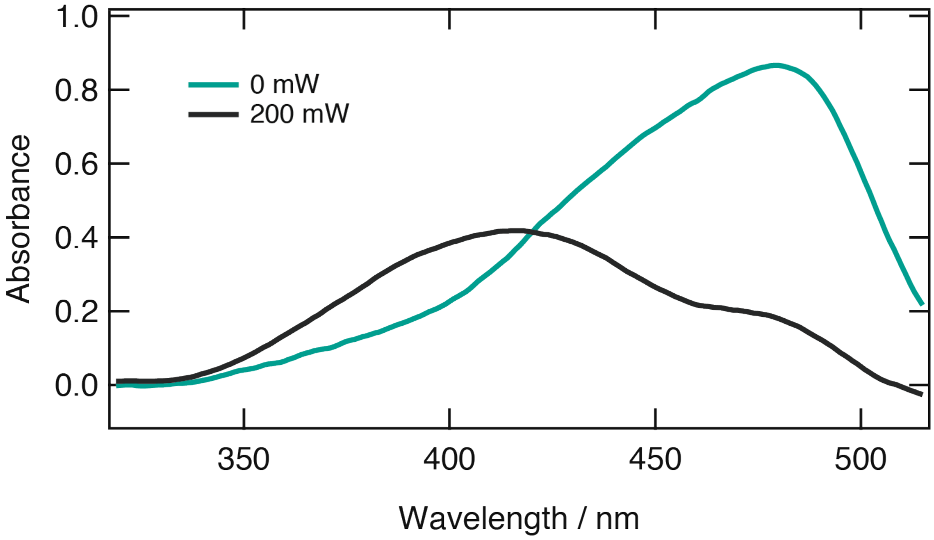
Supplementary Figure 12: Microcrystals ($3 \times 3 \times 3\ \mu m^3$) of rsEGFP2 grown by seeding have a homogeneous size distribution. Micrographs were obtained with an optical microscope (Zeiss SteREO Discovery.V12) carrying an Achromat S 1.5 x objective (upper panel) and an inverted microscope (Olympus IX 81) with oil immersion and a total magnification of 150-fold (lower panels).



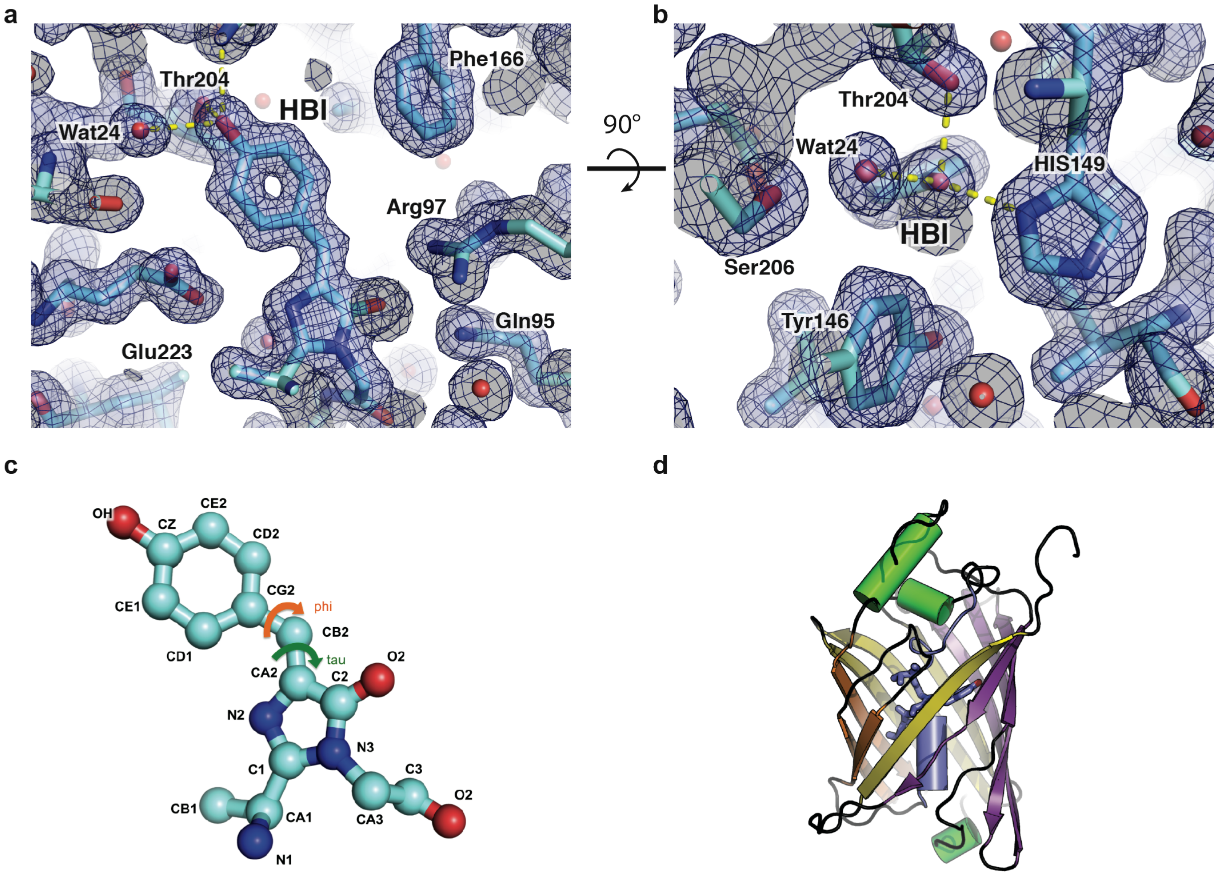
Supplementary Figure 13: rsEGFP2 microcrystals are photoswitched from the fluorescent *on* state (resting state) to the non-fluorescent *off* state (lifetime 100 min in microcrystals[38] by pre-illumination with 488 nm laser light. Pump-laser illumination at 400 nm triggers photoswitching from the *off* to the *on* state.



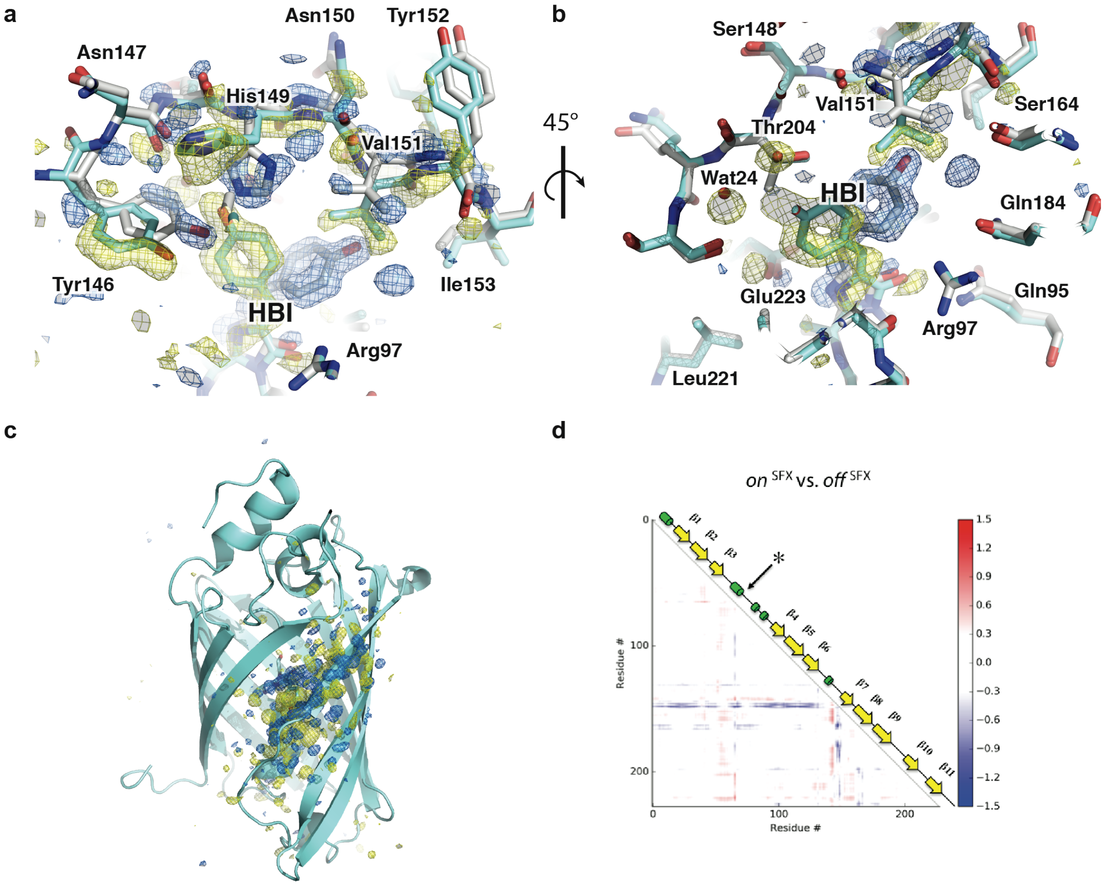
Supplementary Figure 14: Pre-illumination setup used for switching rsEGFP2 from the *on* to the *off* state[38]. A HPLC stainless steel Tee union was used to keep the polished core of a 200 μm optical fiber in place and in direct contact with the side of the tubing (ID 254 μm), through which the protein crystals were pumped. The other end of the optical fiber was connected to a 200 mW CW laser (488 nm). The protective sleeve of the optical fiber was pulled back in the photo to show the fiber core.



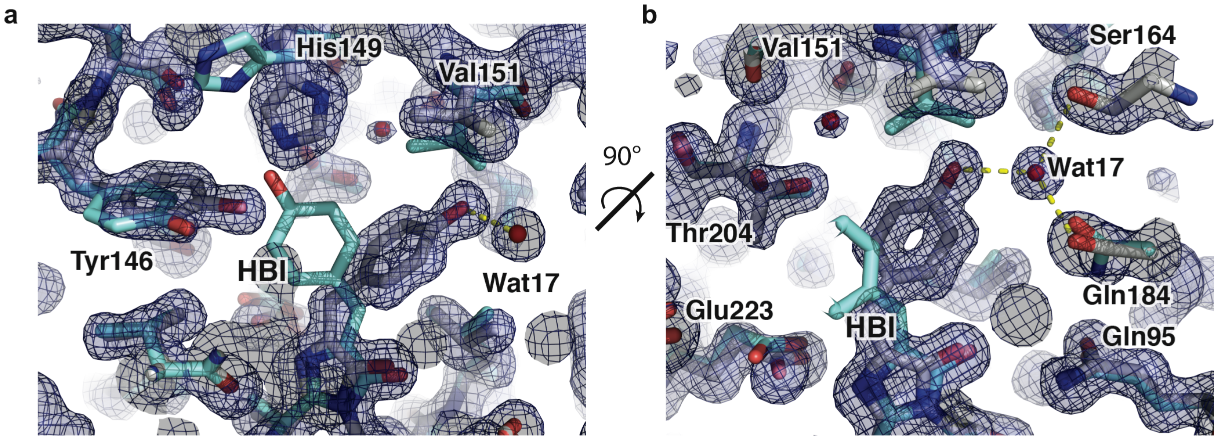
Supplementary Figure 15: Absorption spectra of rsEGFP2 microcrystals without (green) and with 488 nm pre-illumination at 200 mW (black). Calculations based on these data indicate that 90% of crystalline rsEGFP2 has been converted to the non-fluorescent *off* state, and 10% remains in the fluorescent *on* state.



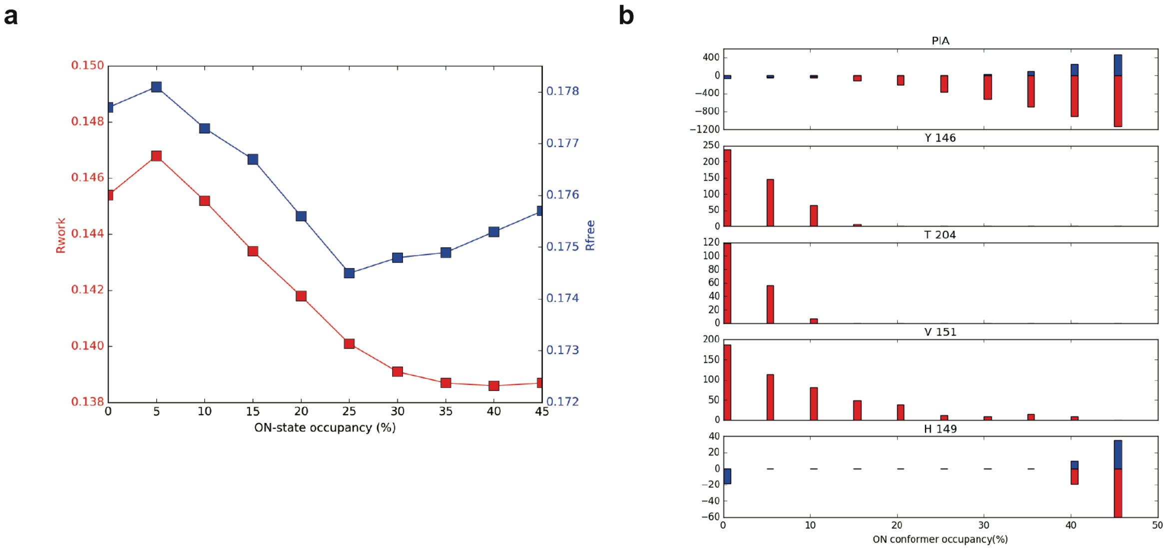
Supplementary Figure 16: SFX structure of rsEGFP2 in the *on* state. The final *on* model (cyan) is shown with a $2mF_{obs}^{reference} - DF_{calc}$ map contoured at 1 σ. Panels (*a*) and (*b*) show two views, 90° apart. In panel (*b*), the H-bonding partners of the chromophore (HBI) phenol group are shown, namely His149, Thr204 and Wat24 (H-bond lengths of 2.9, 2.8 and 2.6 Å, respectively). (*c*) Topology of the chromophore in the *cis* isomer. Two dihedral angles (*tau* (C2-CA2-CB2-CG2) and *phi* (CA2-CB2-CG2-CD2)) determine the isomerization state of the chromophore (green and orange arrows, respectively). (*d*) Ribbon diagram of rsEGFP2. Strands whose residues face the chromophore phenol and imidazolinone groups are highlighted in purple and yellow, respectively. Strands whose residues face the α-helix N-terminal to the chromophore are highlighted in orange. The structural element holding the chromophore in the middle of the β-barrel, including the α-helix N-terminal to it, is shown in blue. Other α-helices are shown in green, and residues in loops are shown in black.



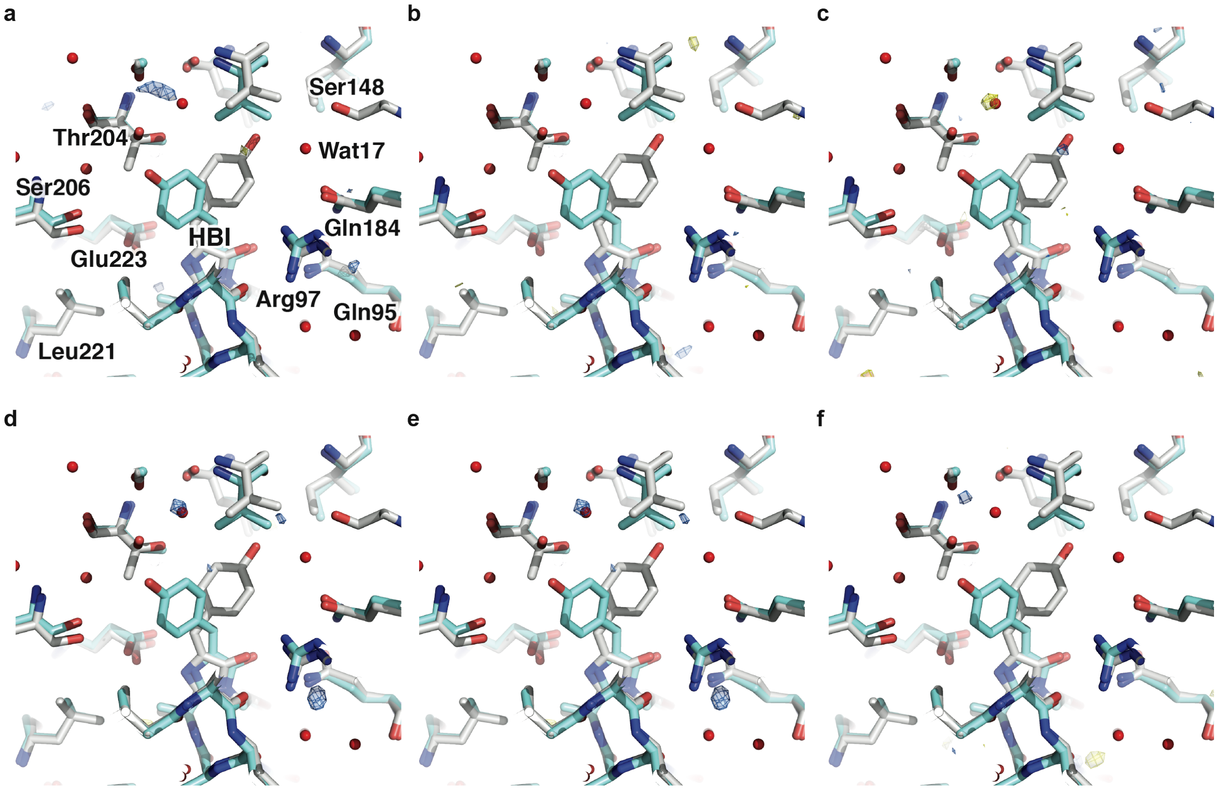
Supplementary Figure 17 : In-line pre-illumination of crystals results in efficient *on*-to-*off* transition. Q-weighted[54] difference Fourier map calculated between the reference (no pre-illumination) and the laser-off (pre-illumination at 488 nm) datasets, using the final *on* model for phasing. In all panels, the *on*-state model appears as a cyan carbon trace, and the *off*-state model as a grey carbon trace. Difference Fourier maps are contoured at ± 4.0 σ, with negative and positive peaks displayed in yellow and blue, respectively. Panels (*a*) and (*b*) show two views of the chromophore pocket, ~45° apart, and allow observing all conformational changes that accompany the *on*-to-*off* transition. This transition involves large conformational changes in residues from β-strand-7 (Tyr146, His149, Val151; mostly visible on panel (*a*)) and β-strand-10 (Thr204; mostly visible on panel (*b*)). These changes include both main and side chain atoms. (*c*) Overview of $F_{obs}^{laser\text{-}off}$-$F_{obs}^{reference}$ map. The ribbon diagram of the *on* model is shown in the same color as above. Only the chromophore is shown as sticks. (*d*) Distance difference matrix (DDM) calculated between structures of rsEGFP2 from SFX *on*- and laser-off structures. The position of the chromophore is indicated by *. The *on*-state structure appears more compact than the laser-off structure. Residues in β-strand-7 (*i.e.* residues 149 to 156) show the largest backbone deviation, moving closer to the chromophore in the *on*-state structure.



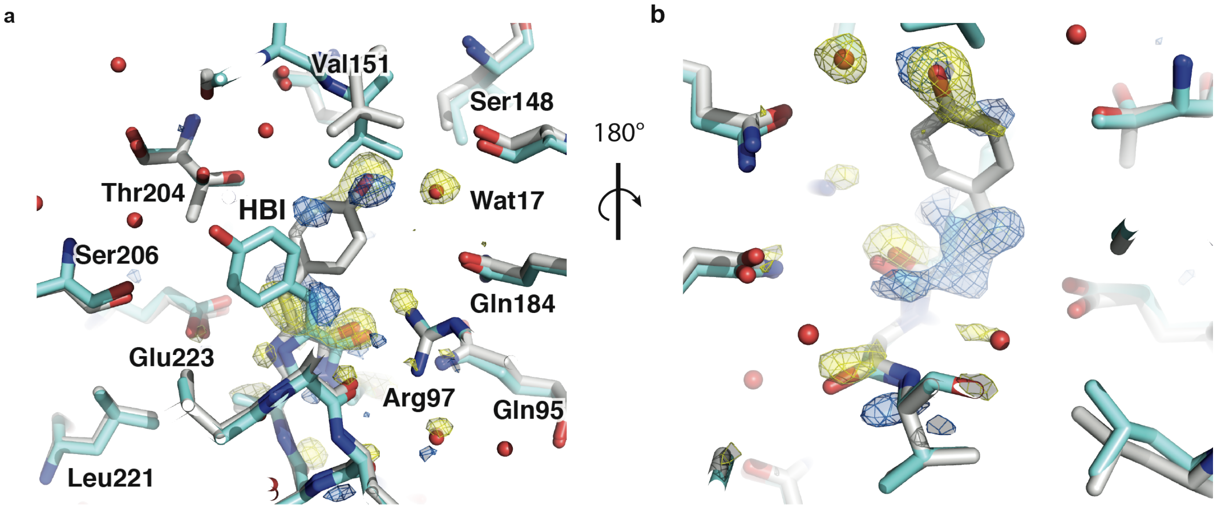
Supplementary Figure 18: Occupancy of the *on*-state conformer in the laser-off dataset. In both panels, the *off*-state conformer of the laser-off structure appears as a grey carbon trace, and its *on*-state conformer as a cyan carbon trace. (*a, b*) $2mF_{obs}^{laser\text{-}off}$ - DF_{calc} map calculated from the laser-off dataset. Views (*a*) and (*b*) are roughly 90° apart. In panel (*b*), Wat17 is shown, which is H-bonded to the phenol hydroxyl group of the *off-s*tate chromophore (distance: 2.7 Å). This water is also H-bonded to Ser164 and Gln184 (2.9 and 2.9 Å distance, respectively).



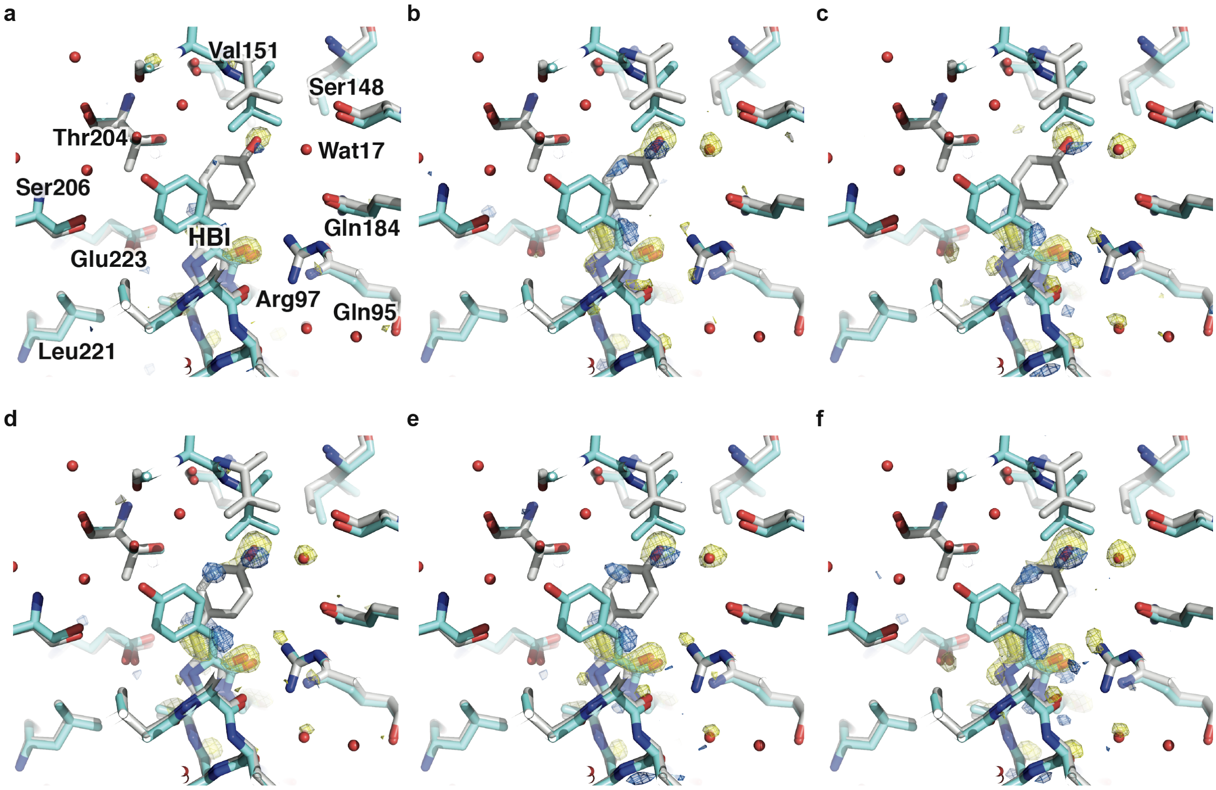
Supplementary Figure 19: Occupancy of the *on*-state conformer in the laser-off dataset (*a*). R_{free} (blue) and R_{work} (red) values obtained when refining the laser-off dataset with 0 to 45% of the *on*-state model and 100 to 55% of the *off* model. The plot suggests that the best model is that composed of 25 and 75% of *on*-state and *off*-state conformers. (*b*) Integration of positive and negative $mF_{obs}^{laser\text{-}off}$ -DF_{calc} peaks on key residues subject to structural rearrangements upon *off*-to-*on* transition. Peaks located on the *off*-state and *on*-state conformers are colored in blue and red, respectively. Strong negative peaks are observed on the *on*-state chromophore in $mF_{obs}^{laser\text{-}off}$ -DF_{calc} maps with an *on*-state occupancy higher than 10%. Therefore, we set the relative occupancy of the *on*-state and *off*–state conformers to 10 and 90%, respectively, in the final *laser-off* model.



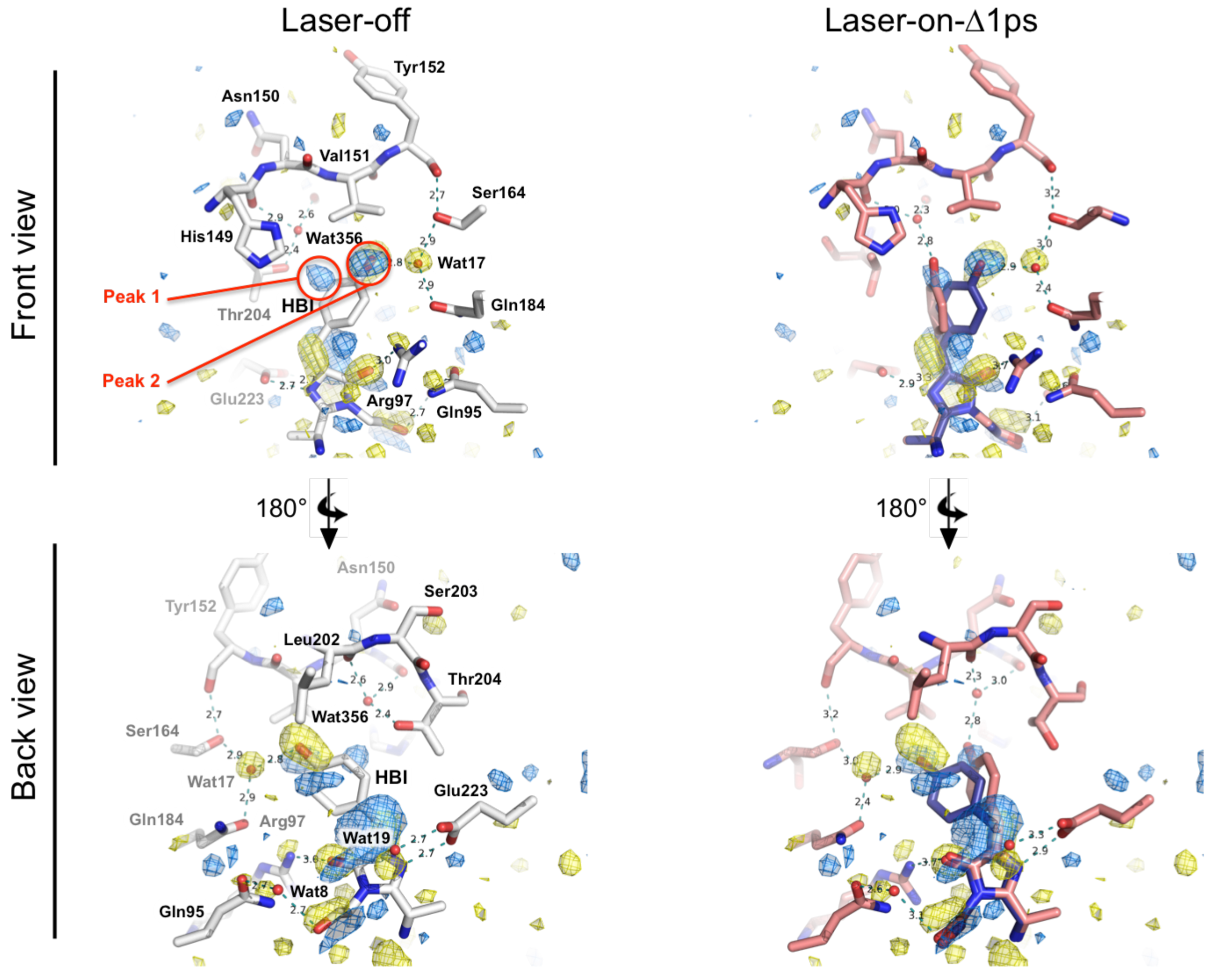
Supplementary Figure 20: All laser-off data collected over three experimental shifts contain similar structural information. Q-weighted[54] difference Fourier maps calculated from reduced laser-off datasets produced on a per-shift basis (1st row) or by a combination of shifts (2nd row). The complete laser-off dataset was used as a reference, and phases were calculated from the final laser-off model, consisting of 90% and 10% of *off*- and *on*-state conformers, respectively. In all panels, the *off*-state model appears as a grey carbon trace, and the on-state model as a cyan carbon trace. Difference Fourier maps are contoured at ± 3.5 σ, with negative and positive peaks displayed in yellow and blue, respectively. (*a*) 1st shift (12 hours); (*b*) 2nd shift; (*c*) 3rd shift; (*d*) 1st and 2nd shifts; (*e*) 1st and 3rd shifts; (*f*) 2nd and 3rd shifts. (*a-f*) Regardless of which reduced laser-off dataset is considered, the difference Fourier maps are virtually featureless, indicating that the structural information collected during the various shifts is both qualitatively and quantitatively highly similar. Thus, the pre-illumination setup and the *on*-to-*off* photoswitching efficiency were stable over the course of the three shifts. The flatness of these maps validates the merging of laser-off data collected during the three shifts.



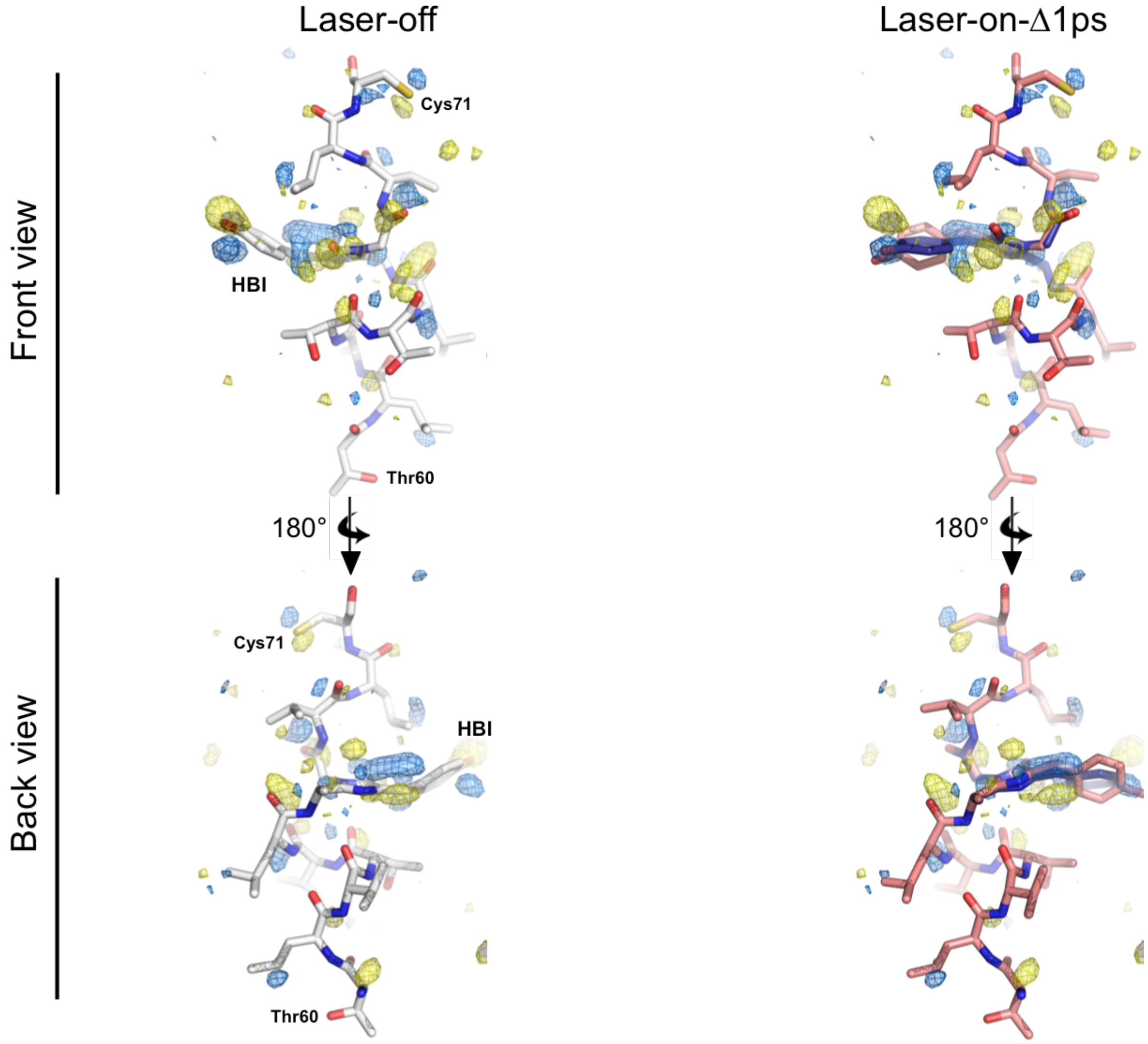
Supplementary Figure 21: Pump-laser illumination at 400 nm results in conformational changes in the chromophore region after 1 ps. Q-weighted[54] difference Fourier map ($F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$) calculated between the laser-on-Δ1ps dataset and the laser-off dataset, using the final laser-off model consisting of 90% and 10% of *off*- and *on*-state conformers, respectively for phasing. The *off*-state model appears as a grey carbon trace, and the *on*-state model as a cyan carbon trace. Difference Fourier maps are contoured at $\pm$ 3.5 σ, with negative and positive peaks displayed in yellow and blue, respectively. Panels (*a*) and (*b*) show two views of the chromophore pocket, rotated by ~180°, allowing to observe peaks on both sides of the chromophore. Strong negative peaks, up to -10 σ, are seen on the phenol hydroxyl, imidazolinone carbonyl oxygen and methylene linker of the *off*-state chromophore, as well as on waters 8, 17, and 19 which are H-bonded to it. These negative peaks indicate that the fraction of crystalline *off*-state chromophore is decreased, 1 ps post-pumping at 400 nm, and that waters in direct interaction with the chromophore are affected at the same time. Negative peaks are also observed on side chain atoms of residues stabilizing the chromophore imidazolinone ring (Gln95, Arg97 and Glu223), and on backbone atoms of residues N-terminal or C-terminal to the chromophore (Leu61 to Cys71, where the chromophore is HBI). In addition, positive peaks are seen mid-way between the *on*- and *off*-state chromophore. Hence, a fraction of the crystalline *off* chromophore has transitioned to a new state. A positive peak is also seen for Wat17, suggesting that it accompanies this transition.



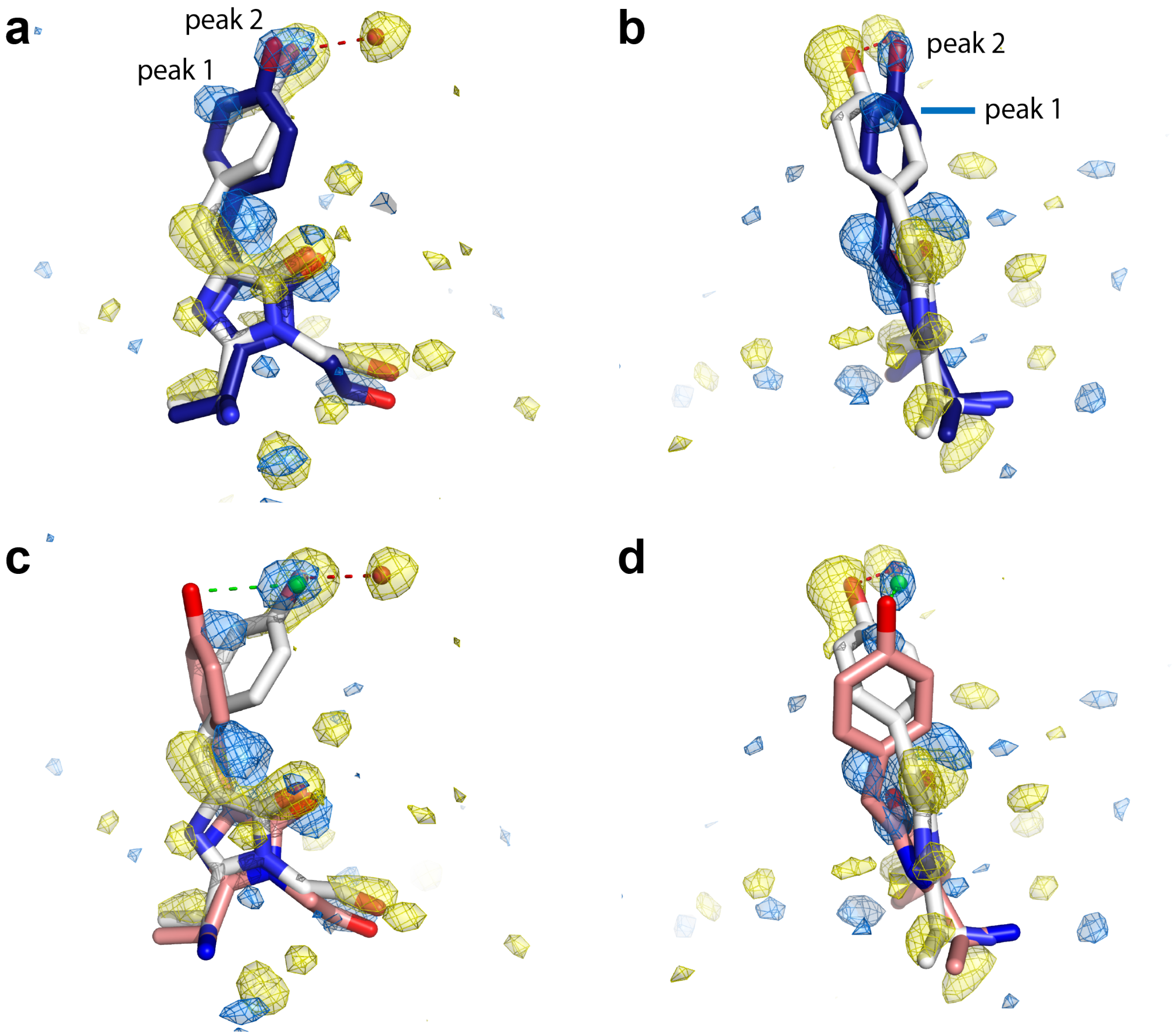
Supplementary Figure 22: All laser-on-Δ1ps data collected over three experimental shifts contain similar structural information. Q-weighted[54] difference Fourier maps calculated from segmented laser-on-Δ1ps datasets produced on a per-shift basis (1st row) or by binary combination of shifts (2nd row). The final laser-off dataset was used as a reference, and phases were calculated from the final laser-off model, consisting of 90 and 10% of *off*- and *on*-state conformers, respectively. In all panels, the *off*-state model appears as a grey carbon trace, and the *on*-state model as a cyan carbon trace. Difference Fourier maps are contoured at ± 3.5 σ, with negative and positive peaks displayed in yellow and blue, respectively. (*a*) 1st shift (12 hours); (*b*) 2nd shift; (*c*) 3rd shift; (*d*) 1st and 2nd shifts; (*e*) 1st and 3rd shifts; (*f*) 2nd and 3rd shifts. (*a-f*) Regardless of which subset of the full laser-on-Δ1ps dataset is considered, the negative features of the maps are comparable, *i.e.* peaks of similar intensities are seen on the phenol hydroxyl, imidazolinone carbonyl oxygen and methylene linker of the chromophore, as well as on waters 8, 17, and 19. As for positive peaks, these only rise above 3.5 σ in difference Fourier maps calculated from data combination from two shifts, illustrating the direct relationship between the clearness of map features and the number of indexed patterns used to produce parent datasets. More importantly, the similarity between the difference Fourier maps shown in panels *a-f* and that computed from the complete laser-on-Δ1ps dataset (shown in Supplementary Fig. 20) retrospectively validates the merging of laser-on-Δ1ps data collected over the three shifts.



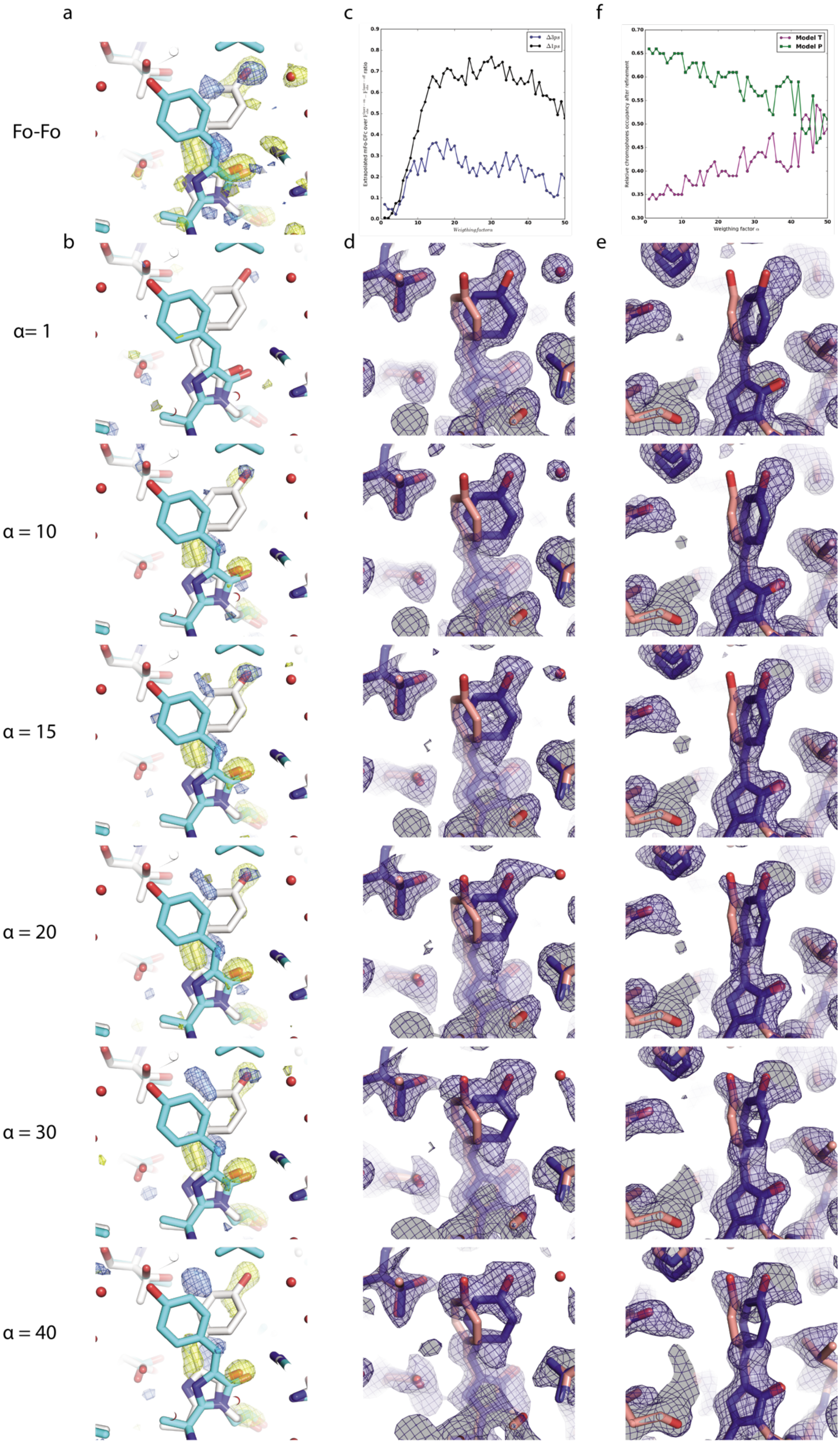
Supplementary Figure 23: Peaks in the Q-weighted[54] $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ map are concentrated around the chromophore region. Upper and lower panels show opposite views of the chromophore region. The *off*-state model (grey) accounting for 90% of the laser-off data set is displayed in the left panels, and the difference-refined laser-on-Δ1ps model is shown in the right panels (model T in pink, and model P in blue). In all panels, the $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ map is contoured at ± 3σ. Red circles highlight peaks 1 and 2 into which model T and model P were built, respectively. Residues visible in the panels are those that display peaks higher than ± 3σ in the vicinity of the chromophore, and they accordingly display dynamics in the difference-refined structure (right panels). Structural waters interacting with the chromophore are also shown. Hydrogen-bonds are highlighted by dashed lines, and the bond-length (in Å) overlaid on these. Wat17 and Wat356 are obligate acceptors in the H-bond they establish with the chromophore, implying that the phenol OH of the chromophore can only interact with one of them at the time. H-bonding to Wat17 is thus exclusive to the *off*-state chromophore and to model P, while H-bonding to Wat356 is the preserve of model T. Formation of model T is accompanied by a 120° rotation in the side chain of Thr204, hitherto H-bonded in a similar fashion (obligate donor) to Wat356.



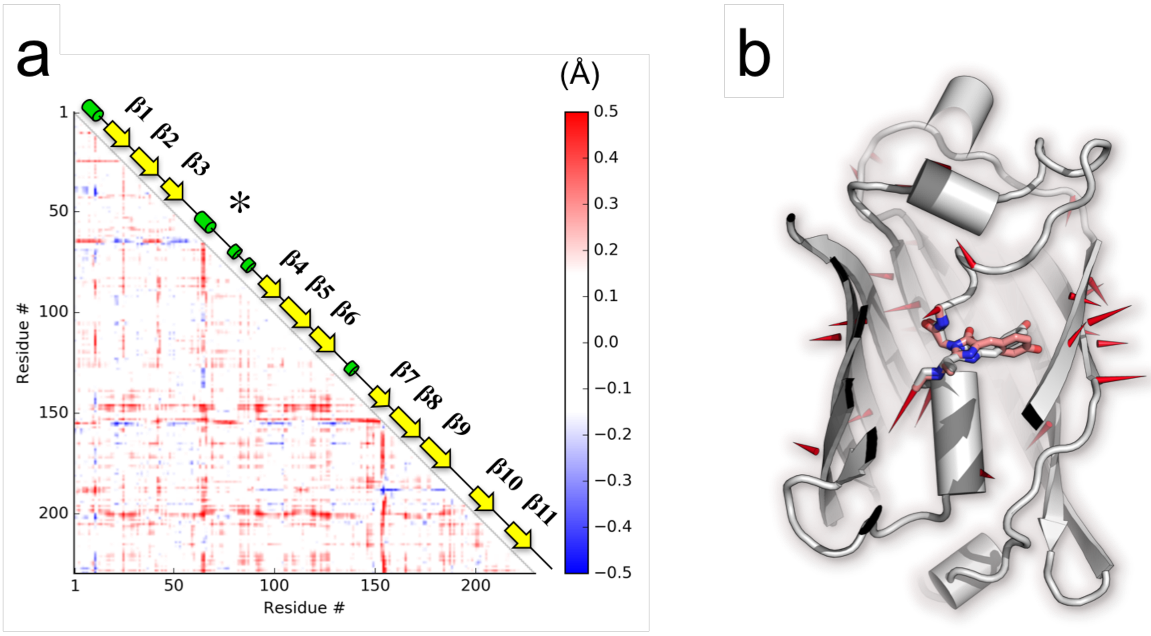
Supplementary Figure 24: The α-helix N-terminal to the chromophore features a succession of peaks in the Q-weighted[54] $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ map. Upper and lower panels show opposite views of the α-helix N-terminal to the chromophore. The *off*-state model accounting for 90% of the laser-off data set is displayed in the left panels (grey), and the difference-refined laser-on-Δ1ps model is shown in the right panels (model T in pink, and model P in blue). In all panels, the $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ map is contoured at ± 3σ. All residues in the α-helix N-terminal to the chromophore show peaks in the difference map and likewise for the three residues C-terminal to it.



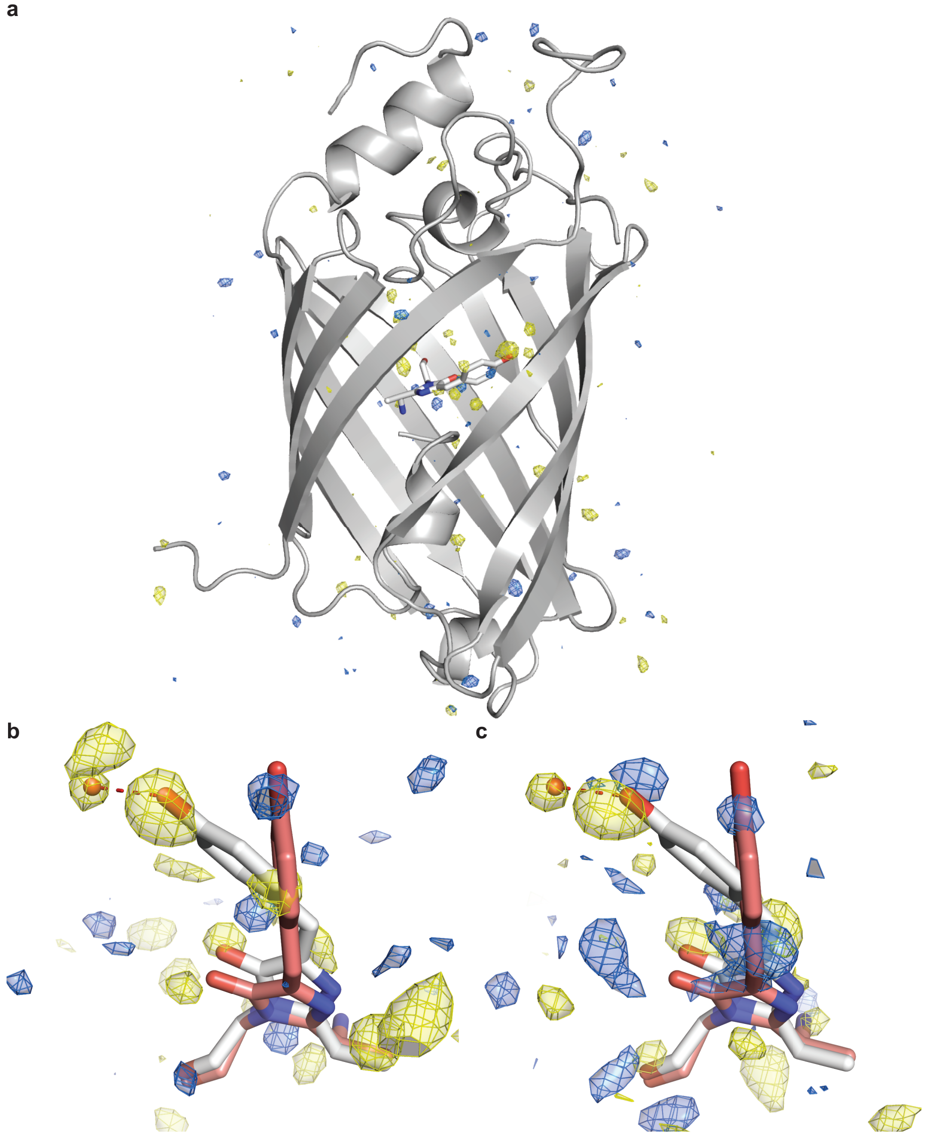
Supplementary Figure 25: Modeling of the *intermediate* state models at 1 ps. In all panels (*a-d*), the Q-weighted[54] difference Fourier map ($F_{obs}^{laser-on-\Delta 1ps}$-$F_{obs}^{laser-off}$) calculated between the laser-on-Δ1ps dataset and the laser-off dataset, is contoured at ± 3.5σ, with negative and positive peaks displayed in yellow and blue, respectively. The *off*-state model appears as a grey carbon trace, and the *intermediate*-state models P (*a, b*) and T (*c, d*) are shown as blue and pink carbon traces, respectively. Negative peaks are observed on the *off*-state chromophore, suggesting that upon pumping at 400 nm and at a pump-probe delay of 1ps, the latter has been depleted. Two positive peaks appear midway between the *on* (not shown; see Supplementary Fig. 21 and 22) and *off*-state conformations of the chromophore, suggesting the buildup of an intermediate state: two conformations of the chromophore could explain these two peaks. (*a, b*) In model P, the chromophore OH and CE1 fit in the two peaks. However, the large bilobal peak observed on the imidazolinone is only partially accounted for. (*c, d*) In model T, the chromophore OH fits in peak 1, whereas peak 2 accommodates water 17, H-bonded to the chromophore OH in the *off*-state conformation. This model provides a better fit for the peak observed on the imidazolinone but implies that the H-bond between the chromophore OH and water 17 is preserved after the transition to the intermediate state.

a
c
f
Fo-Fo
b
d
e
α= 1
α = 10
α = 15
α = 20
α = 30
α = 40
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Extrapolated mFo-DFc over $F_{obs}^{laser-on} - F_{obs}^{laser-off}$ ratio
0
10
20
30
40
50
$Weigthing factor α$
Δ3ps
Δ1ps
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
Relative chromophores occupancy after refinement
Weigthing factor α
Model T
Model P

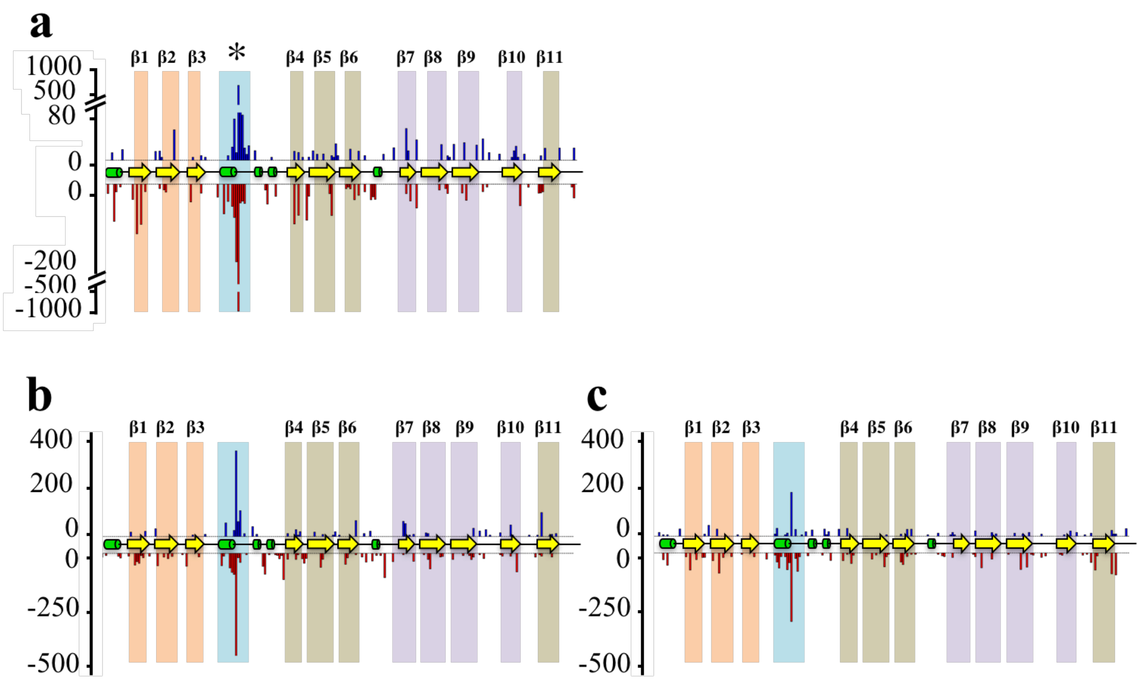
Supplementary Figure 26: **Determination of absolute and relative occupancy of models T and P at 1 ps.** (*a*) Experimental $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$ - $F_{obs}^{laser\text{-}off}$ map contoured at 3σ. The laser-off model is displayed, with the chromophore *off*- and *on*-conformers represented as white and cyan sticks, respectively. (*b*) $F_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps}$ - $F_{calc}^{laser\text{-}off}$ maps obtained from extrapolated structure factors calculated using increasing α values and phased using the laser-*off* model. All maps are contoured at ± 3σ. At α=1, which would correspond to 100% occupancy of the intermediate state, the $F_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps}$ - $F_{calc}^{laser\text{-}off}$ map is featureless. When α reaches 15, the resulting $F_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps}$ - $F_{calc}^{laser\text{-}off}$ map reveals peaks qualitatively and quantitatively similar to those observed in the experimental $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$ - $F_{obs}^{laser\text{-}off}$ map. (*c*) Determination of α, *i.e.* the inverse of the intermediate state occupancy. $F_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps}$ - $F_{calc}^{laser\text{-}off}$ peaks higher than ± 3σ (b) observed in a radius of 1.5 Å around the chromophore atoms in all its conformations (*off* and *on* conformers, as well as models T and P) were integrated and the ratio of their absolute sum over the overall signal of the experimental $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$ - $F_{obs}^{laser\text{-}off}$ (black) was plotted. We attempted to use the same approach to determine α for the laser-on-Δ3ps (blue). For the laser-on-Δ1ps dataset, this ratio increases with α values up to ~15, and then reaches a plateau at ~0.8. For the laser-on-Δ3ps dataset, the curve shape is similar but the plateau only reaches 0.3. Therefore, we decided not to perform difference refinement with the laser-on-Δ3ps dataset. (*d, e*) $2mF_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps}$ - $DF_{calc}^{laser\text{-}off}$ maps obtained from extrapolated structure factors calculated using increasing α values, and phased using the laser-*off* model. All maps are contoured at 1σ. The chromophore models T and P are shown in salmon and blue sticks, respectively (*d*, *e*). α values are identical to those used to generate the $F_{extrapolated}^{laser\text{-}on\text{-}\Delta 1ps}$ - $F_{calc}^{laser\text{-}off}$ maps shown in (*b*). For α=1, the density is surrounding the *off*-state chromophore (model not shown), but with increasing α values, electron density builds up around model T and model P. Recall that no phase information from the latter was introduced during map computation. (*f*) Refined occupancies of model T (magenta) and model P (green) obtained at increasing α values in the laser-on-Δ1ps dataset.



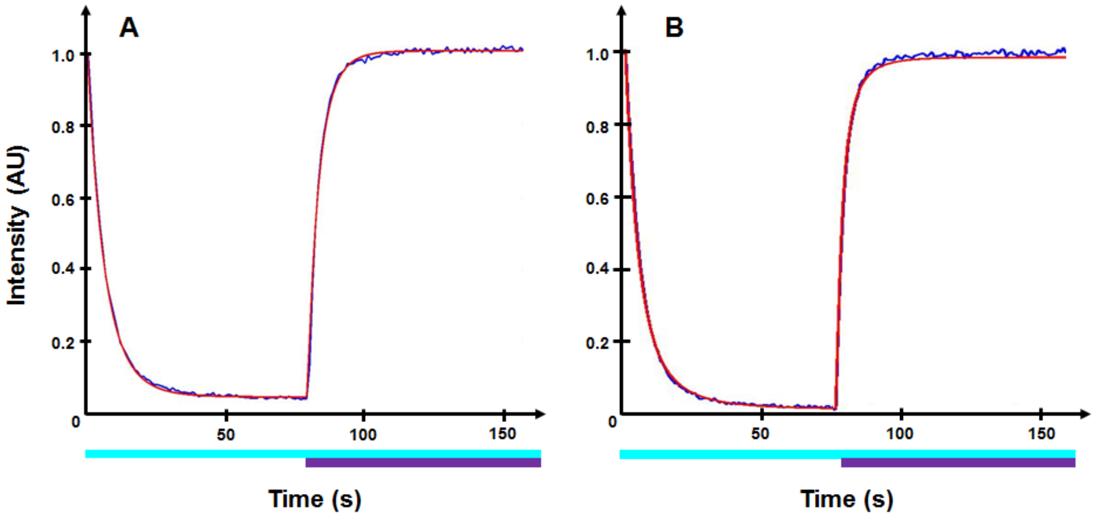
Supplementary Figure 27: The β-barrel expands upon the transition from the *off* state to model T. (*a*) Difference distance matrix calculated between the *off*-state structure and the difference-refined laser-on-Δ1ps structure. Red and blue indicate increasing and decreasing distances, respectively, in model T with respect to the *off*-state structure. (*b*) Porcupine plot of the differences between the *off*-state structure and the difference-refined laser-on-Δ1ps structure. The *off*-state (cyan) and model T (pink) chromophores are shown. The movement of Cα atoms by more than 0.2 Å is indicated by arrows on the ribbon representation, with the magnitude of motions illustrated by the length of arrows exaggerated seven fold to increase visibility.



Supplementary Figure 28: After pumping at 400 nm, the crystalline system further evolves between 1 ps and 3 ps. (*a*) Overview of the Q-weighted[54] difference Fourier map ($F_{obs}^{laser\text{-}on\text{-}\Delta 3ps}$-$F_{obs}^{laser\text{-}off}$; contoured at ± 3.5σand colored as in Supplementary Fig. 23), overlaid on the *off*-state model (grey). (*b, c*) Close-up view of the *off*-state (grey) and *intermediate* chromophore model T (pink), overlaid with the $F_{obs}^{laser\text{-}on\text{-}\Delta 3ps}$-$F_{obs}^{laser\text{-}off}$ (*b*) and $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps\text{-}random}$-$F_{obs}^{laser\text{-}off}$ (*c*) maps, both contoured at ± 3σ. Comparison of (*b*) and (*c*) suggests that the *intermediate* state corresponding to model T has decayed on the 1 to 3 ps timescale.



Supplementary Figure 29: Sequence-wise integration of peaks from Q-weighted[54] difference Fourier map $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$-$F_{obs}^{laser\text{-}off}$ (*a*), $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps\text{-}random}$-$F_{obs}^{laser\text{-}off}$ (*b*), and $F_{obs}^{laser\text{-}on\text{-}\Delta 3ps}$-$F_{obs}^{laser\text{-}off}$ (*c*). In all three panels, the x-axis corresponds to the protein sequence, with the secondary structure illustrated by a cartoon. The y-axis reports the integrated values of difference Fourier map peaks, with negative and positive peaks colored in red and blue, respectively. The α-helix central to the β-barrel and which supports the chromophore is highlighted by a blue shaded band. β-strands whose residues face the phenol and imidazolinone groups of the chromophore are highlighted by purple and yellow shaded bands, respectively. β-strands whose residues face the α-helix N-terminal to the chromophore are highlighted by orange shaded bands. All peaks higher than ± 3σ were integrated and assigned to the closest residue within a 1.5 Å radius. In panel (*a*), * indicates the peak corresponding to the chromophore. (*b, c*) For the sake of comparison, a laser-on-Δ1ps-random dataset was generated (with as many indexed images as in the laser-on-Δ3ps dataset). The integrated peak values for the chromophore, negative and positive alike, are decreased in the difference Fourier map calculated with the laser-on-Δ3ps dataset. Thus, the combined occupancy of model T and model P decreases on the 1-3 ps timescale, whereas the *off*-state chromophore appears to repopulate.



Supplementary Figure 30: Reversible switching cycle of rsEGFP2 (*a*) and rsEGFP2-V151A (*b*). Raw data (blue) were fitted (red) using a kinetic model involving one-to-one reversible conversion between the *on*- and the *off*-state. Cyan and violet lines represent the illumination times at 488- and 405-nm, respectively.

Supplementary Table 1 : Equilibrium bond lengths (Å) used in the excited-state MD simulations. All other values were taken from standard residues in the Amber force field. Atom names are according to Supplementary Fig. 16c.

	S_0	S_1
C1-N2	1.296	1.302
N2-CA2	1.410	1.359
CA2-CB2	1.367	1.419
CB2-CG2	1.445	1.399
CA2-C2	1.486	1.462
C2-O2	1.225	1.211
C2-N3	1.406	1.403
N3-C1	1.394	1.356
N3-CA3	1.458	1.448
C1-CA1	1.508	1.506

Supplementary Table 2 : Equilibrium angles (°). All other values were taken from standard residues in the Amber force field. Atom names are according to Supplementary Fig. 16c.

	S_0	S_1
C1-N2-CA2	107	108
N2-CA2-C2	108	108
CA2-C2-N3	103	103
C2-N3-C1	108	108
N3-C1-N2	114	112
N2-CA2-CB2	119	118
C2-CA2-CB2	133	134
CA2-CB2-CG2	134	132
CB2-CG2-CD1/CD2	120	120
N3-C2-O2	123	122
CA2-C2-O2	134	135

Supplementary Table 3 : Atomic charges (in a.u.). Atom names are according to Supplementary Fig. 16c.

	S_0	S_1
N1	-0.4157	-0.4157
H1	0.2719	0.2719
CA1	0.0236	0.0018
HA1	0.1301	0.1407
CB1	-0.1712	-0.1841
HB1 (3 atoms)	0.0395	0.0720
C1	0.5929	0.1047
N2	-0.6255	-0.4626
CA2	0.0691	0.3487
CB2	-0.1160	-0.4775
HB2	0.1605	0.1832
CG2	0.0484	0.2545
CD1 / CD2	-0.1139	-0.1480
HD1 /HD2	0.1570	0.1651
CE1 (CE2)	-0.3213	-0.3385
HE1 (HE2)	0.1875	0.1922
CZ	0.4182	0.4685
OH	-0.5536	-0.5457

HH	0.4007	0.3971
C2	0.5082	0.3331
O2	-0.4743	-0.5098
N3	-0.3261	-0.0604
CA3	-0.1529	-0.0370
HA3 (2 atoms)	0.1226	0.1009
C3	0.5973	0.5973
O3	-0.5679	-0.5679

Supplementary Table 4: Fluctuations (°) of the imidazolinone dihedral and improper angles during dynamics of the S_0 methyl-terminated anionic chromophore, obtained either by DFT/MD or by classical dynamics with the force field used in this work. Force constants k used in this work (kcal/mol for intra-cycle dihedrals, kcal/mol/radian^2 for improper dihedrals).

angle	Fluctuations		Force constants
	DFT/MD	Force field	*k*
Intra-cycle dihedrals			
C1-N2-CA2-C2	4.5	4.4	12.0
N2-CA2-C2-N3	5.6	5.0	5.5
CA2-C2-N3-C1	5.5	4.9	5.5
C2-N3-C1-N2	5.0	4.7	8.0
N3-C1-N2-CA2	3.8	4.0	16.0
Improper torsions			
CA2	7.3	6.9	6.0
C2	4.4	4.4	30.0
N3	15.1	11.9	4.0
C1	7.9	7.7	10.0

Supplementary Table 5: Data collection and refinement statistics

Dataset	*on*-state reference	laser-off	laser-on-Δ1ps	laser-on-Δ3ps	laser-on-Δ1ps-random
Pre-illumination (488 nm)	no	yes	yes	yes	yes
Pump-laser excitation (400 nm)	no	no	yes	yes	yes
Nominal pump-probe delay	n/a	n/a	1 ps	3 ps	1 ps
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters					
a (Å)	51.7 ± 0.1	51.5 ± 0.1	51.5 ± 0.1	51.5 ± 0.1	51.5 ± 0.1
b (Å)	62.3 ± 0.1	62.9 ± 0.1	62.9 ± 0.1	62.9 ± 0.1	62.9 ± 0.1
c (Å)	71.9 ± 0.1	71.8 ±0.1	71.8 ±0.1	71.8 ±0.1	71.8 ±0.1
Collected frames	986,178	3,531,773	3,531,773	822,721	n/a
Indexed frames	34,731	65,102	64,629	22,262	22,262
Indexed frames (after outlier rejections)	34,715	65,097	64,620	22,259	22,260
Observations	6,655,463	13,199,336	13,080,130	4,176,799	4,698,783
Resolution (Å)	17.3 – 1.7 (1.76 - 1.70)	17.8 – 1.7 (1.76 – 1.70)	17.8 – 1.7 (1.76 – 1.70)	17.4 – 1.7 (1.76 – 1.70)	17.3 – 1.7 (1.76 – 1.70)
Rsplit (%)	9.1 (59.1)	6.8 (34.0)	6.8 (32.1)	11.4 (64.2)	11.8 (56.9)
CC*	0.996 (0.870)	0.998(0.962)	0.998(0.958)	0.995 (0.883)	0.995 (0.892)
I / σ(I)	9.2 (2.2)	11.5 (2.9)	11.4 (3.0)	7.3 (2.2)	6.8 (1.9)
Completeness (%)	99.8 (98.7)	99.8 (100)	99.9 (100)	98.0 (73.2)	99.9 (100.0)
multiplicity	254 (9)	665 (35)	516 (34)	162 (4)	179 (8)
Riso (with respect to dark dataset)*	22.0	n.a	5.0	7.5	6.8
Refinement statistics					
Refinement strategy	Classical refinement	Classical refinement	Difference-refinement		
PDB ID	5O89	5O8A	5O8B(#)		
Resolution (Å)	1.7 (1.76-1.70)	1.7 (1.76-1.70)	1.9 (1.97-1.90)		
Number of reflections	25573 (2287)	25880 (2384)	18792 (1708)		
Rfree	0.195 (0.308)	0.184 (0..263)	0.293 (0.379)		
Rwork	0.167 (0.244)	0.146 (0.196)	0.275 (0.344)		
Number of protein atoms	1960	3802	1936		
Number of water atoms	315	329	312		
B-factor protein (Å^2)	25.15	20.8	27.2		
r.m.s.d. bond lengths (Å)	0.006	0.006	0.003		
r.m.s.d. angles (°)	0.97	0.85	0.51		

Ramachandran favored	98.7	96.5	96.4		
Ramachandran allowed	1.3	2.6	3.6		
Ramachadran outliers	0	0.9	0.0		
Rotamer outliers	0	0.48	0.45		
C-beta outliers	0	0	0		
Clashscore	3.63	6.81	6.49		

*R_{iso} was calculated using Scaleit[65]

[#] 5O8B corresponds to the difference-refined excited-state structure of rsEGFP2 1 ps following 400 nm-laser irradiation of the *off* state and contains extrapolated structure factors defined by eq. 1 in the Supplementary section entitled '*1-ps intermediate-state structure: data collection, modeling and refinement'.* Also, a composite structure based on 5O8A and 5O8B (83% *off* state, 10% *on* state, 4% model P, 3% model T) and containing observed structure facture amplitudes $F_{obs}^{laser\text{-}on\text{-}\Delta 1ps}$ has been deposited under accession code 5O8C.

Supplementary Table 6 : Effect of various merging schemes on overall data statistics

Dataset	# indexed patterns	Overall Signal to noise ratio (Monte-Carlo and scaling)	Rsplit (%)				
			Monte-Carlo	Monte-Carlo and scaling	Monte-Carlo and partiality	Monte-Carlo Partiality Scaling	Monte-Carlo Partiality Scaling Post-refinement
reference	34731	9.2	10.8	9.1	11.4	9.9	9.8
laser-on-Δ1ps	64629	10.1	9.5	7.3	11.1	8.7	9.1
laser-off	65102	10.0	9.6	7.3	11.2	8.78	9.1
laser-on-Δ3ps	22262	7.3	13.4	11.4	14.5	12.9	12.6
Shift 3 – pump on	16333	5.5	16.4	14.1	18.0	15.6	15.8
Shift 3 – pump off	16415	5.5	16.3	14.0	17.7	15.4	16.3
Shift 4 – pump on	24072	6.5	14.5	11.4	15.8	13.2	13.1
Shift 4 – pump off	24260	6.5	14.5	11.5	15.7	13.1	13.0
Shift 5 – pump on	24224	7.1	13.5	11.1	14.6	12.6	12.3
Shift 5 – pump off	24427	7.1	13.2	10.9	14.5	12.6	12.4
Shifts 3 & 4 – pump on	40405	7.9	11.7	9.2	13.3	10.7	11.1
Shifts 3 & 4 – pump off	40675	8.0	11.7	9.2	13.2	10.7	11.1
Shifts 3 & 5 – pump on	40557	8.1	11.6	9.2	13.3	10.6	10.9
Shifts 3 & 5 – pump off	40842	8.1	11.8	9.3	13.4	10.7	11.1
Shifts 4 & 5 – pump on	48296	9.0	10.3	8.2	11.7	9.7	9.7
Shifts 4 & 5 – pump off	48687	9.0	10.4	8.2	11.8	9.7	9.9

Merging was performed at 1.6 Å resolution with 3 iterations when needed, and using the scsphere model for estimating partiality [66].

Supplementary Table 7 : Photophysical parameters of rsEGFP2 WT and rsEGFP2-V151A at pH 7.5

	rsEGFP2	rsEGFP2-V151A
Absorption maximum *on* [nm]	479	483
Absorption maximum *off* [nm]	406	404
On-to-*off* photoswitching quantum yield (Φ_{off})	0.04	0.06
Off-to-*on* photoswitching quantum yield (Φ_{on})	0.40	0.77
Extinction coefficient (ε) *on* [$M^{-1}.cm^{-1}$]	57 150	47 660
Extinction coefficient (ε) *off* [$M^{-1}.cm^{-1}$]	22 000	24 640

References

1. Colletier, J.P. et al. Serial Femtosecond Crystallography and Ultrafast Absorption Spectroscopy of the Photoswitchable Fluorescent Protein IrisFP. *J Phys Chem Lett* (2016).
2. Yadav, D. et al. Real-time monitoring of chromophore isomerization and deprotonation during the photoactivation of the fluorescent protein Dronpa. *J Phys Chem B* **119**, 2404-14 (2015).
3. Lincoln, C.N., Fitzpatrick, A.E. & van Thor, J.J. Photoisomerisation quantum yield and non-linear cross-sections with femtosecond excitation of the photoactive yellow protein. *Phys Chem Chem Phys* **14**, 15752-64 (2012).
4. Hutchison, C.D.M. et al. Photocycle populations with femtosecond excitation of crystalline photoactive yellow protein. *Chemical Physics Letters* **654**, 63-71 (2016).
5. Feliks, M. & Field, M.J. Pcetk: A pDynamo-based Toolkit for Protonation State Calculations in Proteins. *J Chem Inf Model* **55**, 2288-96 (2015).
6. Field, M.J. The pDynamo Program for Molecular Simulations using Hybrid Quantum Chemical and Molecular Mechanical Potentials. *J Chem Theory Comput* **4**, 1151-61 (2008).
7. Bashford, D. An Object-Oriented Programming Suite for Electrostatic Effects in Biological Molecules. *Lecture Notes in Computer Science* **1343**, 233-240 (1997).
8. Brooks, B.R. et al. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry* **4**, 187-217 (1983).
9. Brooks, B.R. et al. CHARMM: the biomolecular simulation program. *J Comput Chem* **30**, 1545-614 (2009).
10. MacKerell, A.D. et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* **102**, 3586-616 (1998).
11. Reuter, N., Lin, H. & Thiel, W. Green fluorescent proteins: empirical force field for the neutral and deprotonated forms of the chromophore. Molecular dynamics simulations of the wild type and S65T mutant. *Journal of Physical Chemistry B* **106**, 6310-6321 (2002).
12. Beroza, P., Fredkin, D.R., Okamura, M.Y. & Feher, G. Protonation of interacting residues in a protein by a Monte Carlo method: application to lysozyme and the photosynthetic reaction center of Rhodobacter sphaeroides. *Proc Natl Acad Sci U S A* **88**, 5804-8 (1991).
13. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-8, 27-8 (1996).
14. Phillips, J.C. et al. Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781-802 (2005).
15. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. Comparison of simple potential functions for simulating liquid water. . *Journal of Chemical Physics* **79**, 926-935 (1983).
16. Neese, F. & The ORCA program system. Wiley Interdisciplinary Review of Computational Molecular Science. **2**, 73-78 (2011).
17. Lelimousin, M., Adam, V., Nienhaus, G.U., Bourgeois, D. & Field, M.J. Photoconversion of the fluorescent protein EosFP: a hybrid potential simulation study reveals intersystem crossings. *J Am Chem Soc* **131**, 16814-23 (2009).
18. Roy, A., Field, M.J., Adam, V. & Bourgeois, D. The nature of transient dark states in a photoactivatable fluorescent protein. *J Am Chem Soc* **133**, 18586-9 (2011).

19. Dewar, M.J.S., Zoebisch, E.G., Healy, E.F., Stewart, J.J.P. & AM1: a new general purpose quantum mechanical molecular model. *Journal of the American Chemical Society* **107**, 3902-3909 (1985).
20. Slavicek, P. & Martinez, T.J. Ab initio floating occupation molecular orbital-complete active space configuration interaction: an efficient approximation to CASSCF. *J Chem Phys* **132**, 234102 (2010).
21. Polyakov, I.V., Grigorenko, B.L., Epifanovsky, E.M., Krylov, A.I. & Nemukhin, A.V. Potential Energy Landscape of the Electronic States of the GFP Chromophore in Different Protonation Forms: Electronic Transition Energies and Conical Intersections. *J Chem Theory Comput* **6**, 2377-87 (2010).
22. Wanko, M., Garcia-Risueno, P., Rubio, A. & Excited states of the green fluorescent protein chromophore: Performance of ab initio and semi-empirical methods. *Physica Status Solidi B* **249**, 392-400 (2012).
23. Field, M.J. A molecular dynamics simulation of photo-induced electron transfer in an organic donor-acceptor compound. *Chemical Physics Letters* **195**, 388-392 (1992).
24. Fabiano, E., Groenhof, G. & Thiel, W. Approximate switching algorithms for trajectory surface hopping. *Chemical Physics* **351**, 111-116 (2008).
25. Frisch, M.J. et al. Gaussian 03, Revision C.02. (2003).
26. Case, D.A. et al. AMBER 10. 10 edn (University of California, San Francisco, 2008).
27. Jonasson, G. et al. Excited State Dynamics of the Green Fluorescent Protein on the Nanosecond Time Scale. *J Chem Theory Comput* **7**, 1990-7 (2011).
28. Filippi, C., Zaccheddu, M. & Buda, F. Absorption Spectrum of the Green Fluorescent Protein Chromophore: A Difficult Case for ab Initio Methods? *Journal of Chemical Theory and Computation* **5**, 2074-2087 (2009).
29. Toniolo, A., Olsen, S., Manohar, L. & Martinez, T.J. Conical intersection dynamics in solution: the chromophore of Green Fluorescent Protein. *Faraday Discuss* **127**, 149-63 (2004).
30. Olsen, S., Lamothe, K. & Martinez, T.J. Protonic gating of excited-state twisting and charge localization in GFP chromophores: a mechanistic hypothesis for reversible photoswitching. *J Am Chem Soc* **132**, 1192-3 (2010).
31. Jonasson, G., PhD thesis, Etude théorique de l'extinction de fluorescence des protéines fluorescentes : champ de forces, mécanisme moléculaire et modèle cinétique, Université Paris-Sud (2012).
32. A.M. Koster, P. Calaminici, M.E. Casida, V.D. Dominguez, Flores-Moreno, G. Geudtner, A. Goursot, T. Heine, A. Ipatov, F. Janetzko, J.M. del Campo, J.U. Reveles, A. Vela, B. Zuniga and D.R. Salahub, deMon2k, Version 2, The deMon developers, Cinvestav, Mexico City (2006).
33. Grotjohann, T. et al. rsEGFP2 enables fast RESOLFT nanoscopy of living cells. *eLife* **1**, e00248 (2012).
34. El Khatib, M., Martins, A., Bourgeois, D., Colletier, J.P. & Adam, V. Rational design of ultrastable and reversibly photoswitchable fluorescent proteins for super-resolution imaging of the bacterial periplasm. *Sci Rep* **6**, 18459 (2016).
35. Lomb, L. et al. An anti-settling sample delivery instrument for serial femtosecond crystallography. *Journal of Applied Crystallography* **45**, 674-678 (2012).
36. Liang, M. et al. The Coherent X-ray Imaging instrument at the Linac Coherent Light Source. *J. Synchrotron Rad.* **22**, 514-519 (2015).
37. Weierstall, U., Spence, J.C. & Doak, R.B. Injector for scattering measurements on fully solvated biospecies. *Rev Sci Instrum* **83**, 035108 (2012).

38. Schiro, G., Woodhouse, J., Weik, M., Schlichting, I. & Shoeman, R.L. Simple and efficient system for photoconverting light-sensitive proteins in serial crystallography experiments. *Journal of Applied Crystallography* **50**, 932-939 (2017).
39. Blaj, G. et al. X-ray detectors at the Linac Coherent Light Source. *Journal of Synchrotron Radiation* **22**, 577-583 (2015).
40. Foucar, L. et al. CASS-CFEL-ASG software suite. *Computer Physics Communications* **183**, 2207-2213 (2012).
41. Foucar, L. CFEL-ASG Software Suite (CASS): usage for free-electron laser experiments with biological focus. *Journal of Applied Crystallography* **49**, 1336-1346 (2016).
42. Barends, T.R. et al. Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* **350**, 445-50 (2015).
43. Zeldin, O.B., Brockhauser, S., Bremridge, J., Holton, J.M. & Garman, E.F. Predicting the X-ray lifetime of protein crystals. *Proc Natl Acad Sci U S A* **110**, 20551-6 (2013).
44. Zeldin, O.B., Gerstel, M. & Garman, E.F. RADDOSE-3D: time- and space-resolved modelling of dose in macromolecular crystallography. *Journal of Applied Crystallography* **46**, 1225-1230 (2013).
45. Harmand, M. et al. Achieving few-femtosecond time-sorting at hard X-ray free-electron lasers. *Nature Photonics* **7**, 215-218 (2013).
46. Coquelle, N. et al. Raster-scanning serial protein crystallography using micro- and nano-focused synchrotron beams. *Acta Crystallographica Section D Biological Crystallography* **71**, 1184-1196 (2015).
47. White, T.A. et al. Crystallographic data processing for free-electron laser sources. *Acta Crystallogr D Biol Crystallogr.* **69**, 1231-1240 (2013).
48. Nass, K. et al. Protein structure determination by single-wavelength anomalous diffraction phasing of X-ray free-electron laser data. *IUCrJ* **3**, 180-91 (2016).
49. McCoy, A.J. et al. Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674 (2007).
50. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica. Section D, Biological Crystallography* **66**, 213-221 (2010).
51. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica. Section D, Biological Crystallography* **60**, 2126-2132 (2004).
52. The PyMOL Molecular Graphics System. 1.5.0.4 edn (Schrödinger, LLC.).
53. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **9**, 90-95 (2007).
54. Ursby, T. & Bourgeois, D. Improved Estimation of Structure-Factor Difference Amplitudesfrom Poorly Accurate Data. *Acta Crystallographica Section A Foundations of Crystallography* **53**, 564-575 (1997).
55. Brunger, A.T. Version 1.2 of the Crystallography and NMR system. *Nature Protocols* **2**, 2728-2733 (2007).
56. Colletier, J.-P. et al. Shoot-and-Trap: Use of specific x-ray damage to study structural protein dynamics by temperature-controlled cryo-crystallography. *Proceedings of the National Academy of Sciences* **105**, 11742-11747 (2008).
57. Terwilliger, T.C. & Berendzen, J. Difference refinement: obtaining differences between two related structures. *Acta Crystallogr D Biol Crystallogr.* **51**, 609-18. (1995).
58. Colletier, J.-P. et al. Use of a `caged' analogue to study the traffic of choline within acetylcholinesterase by kinetic crystallography. *Acta Crystallographica Section D* **63**, 1115-1128 (2007).

59. Duan, C. et al. Structural evidence for a two-regime photobleaching mechanism in a reversibly switchable fluorescent protein. *J Am Chem Soc* **135**, 15841-50 (2013).
60. Colletier, J.P. et al. De novo phasing with X-ray laser reveals mosquito larvicide BinAB structure. *Nature* **539**, 43-47 (2016).
61. Schlichting, I. & Chu, K. Trapping intermediates in the crystal: ligand binding to myoglobin. *Curr Opin Struct Biol* **10**, 744-52 (2000).
62. Terwilliger, T.C. & Berendzen, J. Bayesian difference refinement. *Acta Crystallogr D Biol Crystallogr* **52**, 1004-11 (1996).
63. Duan, C. et al. Rational design of enhanced photoresistance in a photoswitchable fluorescent protein. *Methods and Applications in Fluorescence* **3**, 014004 (2015).
64. Ward, W.W. Properties of the coelenterate green-fluorescent proteins. in *Bioluminescence and chemiluminescence: basic chemistry and analytical applications* (ed. DeLuca M. A., M.W.D.) (Academic Press 1981).
65. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-42 (2011).
66. White, T.A. Post-refinement method for snapshot serial crystallography. *Philos Trans R Soc Lond B Biol Sci* **369**, 20130330 (2014).

Appendix C: Serial femtosecond crystallography and ultrafast absorption spectroscopy of the photoswitchable fluorescent protein IrisFP

The Journal of Physical Chemistry Letters
Letter
pubs.acs.org/JPCL

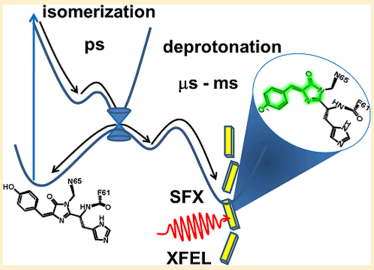
Serial Femtosecond Crystallography and Ultrafast Absorption Spectroscopy of the Photoswitchable Fluorescent Protein IrisFP

Jacques-Philippe Colletier,[*,†] Michel Sliwa,[*,‡] François-Xavier Gallat,[†] Michihiro Sugahara,[§] Virginia Guillon,[†] Giorgio Schirò,[†] Nicolas Coquelle,[†] Joyce Woodhouse,[†] Laure Roux,[†] Guillaume Gotthard,[†,∥] Antoine Royant,[†,∥] Lucas Martinez Uriarte,[‡] Cyril Ruckebusch,[‡] Yasumasa Joti,[⊥] Martin Byrdin,[†] Eiichi Mizohata,[∇] Eriko Nango,[§] Tomoyuki Tanaka,[§] Kensuke Tono,[⊥] Makina Yabashi,[§] Virgile Adam,[†] Marco Cammarata,[¶] Ilme Schlichting,[○] Dominique Bourgeois,[†] and Martin Weik[*,†]

[†]Institut de Biologie Structurale, Université de Grenoble Alpes, CEA, CNRS, F-38044 Grenoble, France
[‡]Université de Lille, CNRS, UMR 8516, LASIR, Laboratoire de Spectrochimie Infrarouge et Raman, F59 000 Lille, France
[§]RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5148, Japan
[∥]The European Synchrotron Radiation Facility (ESRF), 6 rue Jules Horowitz, BP 220, 38043 Grenoble Cedex, France
[⊥]Japan Synchrotron Radiation Research Institute, 1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5198, Japan
[∇]Department of Applied Chemistry, Graduate School of Engineering, Osaka University, Osaka 565-0871, Japan
[¶]Department of Physics, UMR UR1-CNRS 6251, University of Rennes 1, Rennes, France
[○]Max-Planck-Institut für medizinische Forschung, Jahnstrasse 29, 69120 Heidelberg, Germany

Ⓢ *Supporting Information*

ABSTRACT: Reversibly photoswitchable fluorescent proteins find growing applications in cell biology, yet mechanistic details, in particular on the ultrafast photochemical time scale, remain unknown. We employed time-resolved pump−probe absorption spectroscopy on the reversibly photoswitchable fluorescent protein IrisFP in solution to study photoswitching from the nonfluorescent (*off*) to the fluorescent (*on*) state. Evidence is provided for the existence of several intermediate states on the pico- and microsecond time scales that are attributed to chromophore isomerization and proton transfer, respectively. Kinetic modeling favors a sequential mechanism with the existence of two excited state intermediates with lifetimes of 2 and 15 ps, the second of which controls the photoswitching quantum yield. In order to support that IrisFP is suited for time-resolved experiments aiming at a structural characterization of these ps intermediates, we used serial femtosecond crystallography at an X-ray free electron laser and solved the structure of IrisFP in its *on* state. Sample consumption was minimized by embedding crystals in mineral grease, in which they remain photoswitchable. Our spectroscopic and structural results pave the way for time-resolved serial femtosecond crystallography aiming at characterizing the structure of ultrafast intermediates in reversibly photoswitchable fluorescent proteins.



Fluorescent proteins are widely used genetically encoded markers in live biological cells.[1] A subgroup of them, so-called reversibly switchable fluorescent proteins (RSFPs),[2] provide essential tools in advanced fluorescence nanoscopy of live cells (PALM, RESOLFT).[3] RSFPs can be repeatedly toggled back and forth between a fluorescent (*on*) and a nonfluorescent (*off*) state by irradiation with light at two different wavelengths. The most extensively studied RSFP is Dronpa,[4] whose photoswitching mechanism involves chromophore isomerization between *cis* (*on* state) and *trans* (*off* state) isomers along with proton transfer (*off* state neutral; *on* state anionic) as shown by X-ray crystallography[5,6] and absorption spectroscopy.[7] Time-resolved ultrafast spectroscopic experiments on Dronpa and its mutants showed that several photoswitching intermediates exist with lifetimes on the picosecond (ps) and the microsecond (µs) time scales.[8−13] Whether isomerization and proton transfer occur sequentially or in a concerted manner yet remains a matter of debate. On the one hand, consensus has emerged recently that ps intermediates along the *off*-to-*on* switching pathway reflect ultrafast chromophore isomerization in the excited state, whereas µs intermediates result in chromophore deprotonation as a thermally driven ground-state process.[9,12] On the other hand, a recent study[13] argues that ps-intermediate states do not originate from isomerization, which instead would happen like

Received: December 15, 2015
Accepted: February 11, 2016

DOI: 10.1021/acs.jpclett.5b02789
J. Phys. Chem. Lett. 2016, 7, 882−887

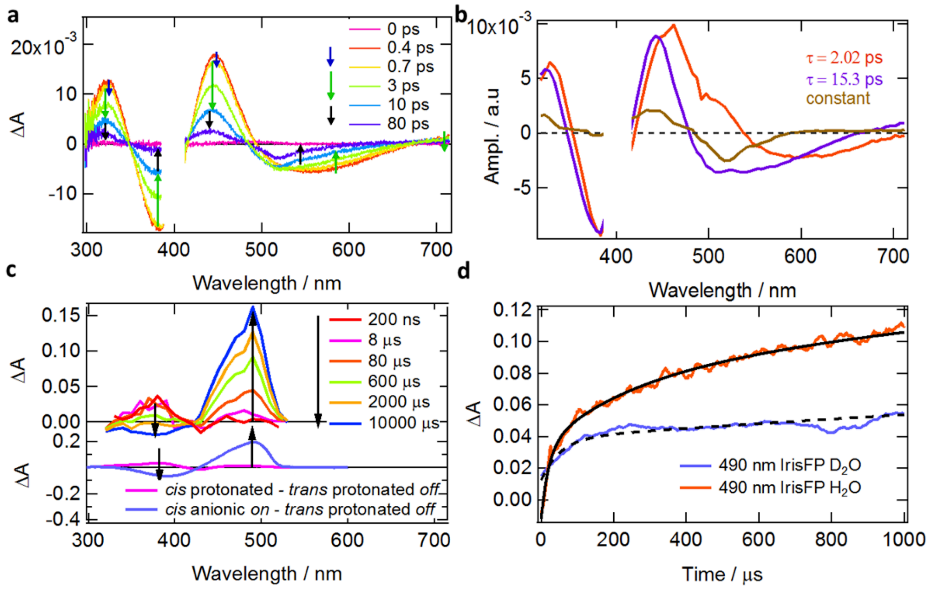


Figure 1. (a) Ultrafast transient absorption spectra of IrisFP in the green *trans* protonated *off* state at pH 7. Spectra were recorded at different time delays following a 400 nm femtosecond laser excitation; the region near 400 nm is a blind zone due to strong laser scattering. (b) Decay associated spectra obtained by global fitting with three exponential functions (96 fs ± 20 fs, 2.02 ps ± 0.05 ps, 15.3 ps ± 0.3 ps) and a constant. (c) Upper panel: microsecond transient absorption spectra of IrisFP in the green *trans* protonated *off* state recorded at different time delays following a 410 nm laser excitation. For a comparison, the lower panel shows two steady state difference spectra calculated between the green *cis* anionic *on* state at pH 7 and the green *trans* protonated *off* state at pH 7, on the one hand, and the green *cis* protonated state at pH 5 and the green *trans* protonated *off* state at pH 7, on the other hand. (d) Kinetics traces and multiexponential fit at 490 nm for IrisFP in D_2O and H_2O on the microsecond time scale.

proton transfer on the slower time scale. Ambiguities and open questions also exist in theoretical calculations,[14,15] for example, concerning the existence of a twisted intramolecular charge transfer state[14] and whether isomerization during photoswitching involves one-bond rotation[14] or a hula twist mechanism.[15] So far, no three-dimensional structures of picosecond intermediates in either Dronpa or any other switchable fluorescent protein has been determined. If available, they would provide direct structural insight into early events of photoswitching and could validate one or several of the photoswitching mechanisms hypothesized based on spectroscopic experiments.

Here, we focused our attention on IrisFP,[16] a peculiar RSFP that is both photoconvertible from green (emission maximum: 516 nm) to red (emission maximum: 581 nm) and, in each of these states, reversibly photoswitchable between an *on* and an *off* state[16] (Figure S1). The green *off*-to-*on* photoswitching mechanism of IrisFP was studied on both the fs−ps and the μs−ms time scales using time-resolved absorption spectroscopy in H_2O and D_2O. Ultrafast and slow processes were detected that we tentatively attribute to excited-state chromophore isomerization and ground-state proton transfer, respectively. However, only structural information may ultimately confirm or invalidate this premise. Structural elucidation of ultrafast intermediates is not feasible using synchrotron sources, where the time resolution is limited to 100 ps. Rather, an X-ray free electron laser (XFEL) needs to be used, with pulse lengths of 5 to 50 fs, that is, a duration well below the lifetime of these intermediate states. In order to prove the feasibility of such experiments, we employed serial femtosecond crystallography (SFX[17,18]) at an XFEL to determine the static structure of IrisFP in its green fluorescent *on* state. Together, these photophysical and structural data lay a solid ground for carrying out time-resolved SFX experiments aiming at the determination of ultrashort photoswitching intermediates in RSFPs.

As in Dronpa, photoswitching in IrisFP involves *cis*−*trans* isomerization of the chromophore and substantial rearrangements of residues in the chromophore cavity as demonstrated by the 3D structures of the various dark (*off*) and fluorescent (*on*) states solved at a resolution between 1.8 and 2.2 Å by cryo-crystallography at a synchrotron.[16] In addition, switching is accompanied by a change in the chromophore protonation state as shown by UV−vis spectroscopy, that is, the green anionic *cis* species absorbs at 488 nm and is fluorescent, whereas the green protonated *trans* state absorbs at 390 nm and does not fluoresce.[16] Compared to Dronpa, IrisFP is characterized by a higher *off*-to-*on* switching quantum yield (15% compared to 7%[19]), making it a priori more suitable for time-resolved SFX experiments.

Ultrafast photodynamics of IrisFP in solution were monitored on the 100 fs to 1 ns time scale by means of femtosecond transient absorption spectroscopy upon excitation of the protein in its green *off* state (100 fs pulse, 400 nm; see

Supporting Information and Methods). The transient spectrum features two negative and two positive bands (Figure 1a and Figure S2). The positive bands (maxima at 320 and 445 nm) are characteristic of the absorption by excited state species. The negative band in the 350−415 nm region with a maximum at 382 nm corresponds to the depopulation of the green *trans* protonated *off* ground state (Figure S3), whereas the broad one in the 490−690 nm region is assigned to the stimulated emission of excited state species. The latter features a maximum at 565 nm, a long tail up to 700 nm and a shoulder at 518 nm. The shoulder originates from those molecules that are in the green *cis* anionic *on* state (about 10%) in the photostationary state of the solution (Figure S3). A global decay analysis of the transient spectra reveals components with characteristic times of 96 fs, 2.02 ps, and 15.3 ps (Figure 1b, Figure S4). The fastest component characterizes the slight decay of the two positive bands within a few hundred femtoseconds (Figure 1a) and can be assigned to ultrafast intramolecular vibrational relaxation within the chromophore. The 2.02 and 15.3 ps components characterize the global reduction of the transient signals, with stimulated emission maxima at 600 and 540 nm, respectively (Figure 1b). These two ps components can be assigned to two excited-state species (I_1*, I_2*) that are formed and decay either sequentially or in parallel.

Time-resolved UV−visible transient absorption was then followed from 10 ns up to the ms time scale in laser flash-photolysis experiments. The spectrum obtained 200 ns postexcitation (Figure 1c, upper panel) matches that previously reported for a green *cis* protonated state at pH 5[16] (Figure 1c, lower panel). The spectrum remains unchanged up to a few μs, and then slowly (10 ms) evolves (Figure 1c, upper panel) into that characteristic of the green *cis* anionic *on* state (Figure 1c, lower panel). In D_2O solution the same bands are observed, but the recovery of the green *cis* anionic *on* state on the μs−ms time scale shows a pronounced H/D exchange effect (Figure 1d and Figure S5), demonstrating the involvement of a proton transfer. Indeed, time constants of the three exponential functions used to fit the temporal evolution at 490 nm were 51 μs (38%), 2.7 ms (17%), and 16 ms (45%) in heavy water, to compare with 21 μs (33%), 227 μs (21%), and 2.1 ms (46%) in light water (Figure 1d and Figure S5). To the contrary, femtosecond transient absorption spectra and kinetics in D_2O (49.5 fs, 2.1 ps, and 17.9 ps) are similar to those in H_2O (96 fs, 2.02 ps, and 15.3 ps; Figures S6 and S7). Therefore, we propose that spectral changes on the μs−ms time scale (Figure 1c) reflect proton transfer during the transition from a green *cis* protonated state to the final green *cis* anionic *on* state (Figure 2) as proposed for Dronpa.[9,12] Consequently, ps spectral changes (Figure 1a) are assigned to *trans*−*cis* isomerization characterizing the transition from the *trans* protonated *off* state to a green *cis* protonated state (Figure 2).

In order to determine whether the two ps excited states, I_1* and I_2*, decay sequentially or in parallel, two kinetic models were tested using the hybrid hard- and soft-multivariate curve resolution method.[20] In the case of Dronpa, a parallel decay from several excited states has been put forward,[12] whereas for IrisFP, only a sequential mechanism fitted the fs data well (Figures 2 and S8). The relaxation to the ground state (either to the green *trans* protonated *off* state or to a green *cis* protonated state; Figures 2 and S8) thus occurs solely from the second excited state species (I_2*). This species is crucial as it controls the quantum yield of *trans*-to-*cis* isomerization, which was determined to be 17% using a comparative method with a solution of $[Ru(bpy)_3]^{2+}$.[21] I_1* and I_2* might either be two distinct excited states with different chromophore geometries, or represent *hot* and *cold* intermediates of the same excited state with the same chromophore geometry. Twisting around the ring-bridging bonds has been reported to occur in the excited GFP chromophore in solution.[22] Even though the structural nature of I_1* and I_2* remains unknown, we speculate that such a twisting might also characterize the IrisFP chromophore in I_2* (and possibly I_1*), as a possible pathway for radiationless decay explaining the nonfluorescent character of this state. Only time-resolved SFX experiments[23,24] revealing the three-dimensional structures of I_1* and I_2* will be able to confirm the exactness of the proposed photomechanism.

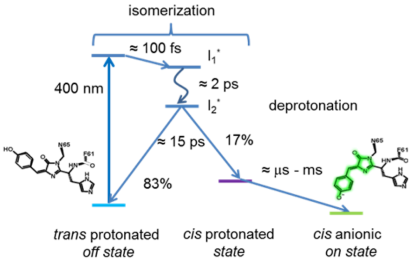


Figure 2. General photomechanism of *off*-to-*on* photoswitching for green IrisFP in solution.

In order to indicate the feasibility of time-resolved SFX experiments, we ambitioned carrying out static SFX on crystalline IrisFP in its *on* state. However, because of the small amount of available sample, liquid injection using a gas dynamic virtual nozzle could not be considered, because it typically requires tens of milligrams of crystalline sample. The only viable option was resorting to a recently developed injection technology, whereby microcrystals are embedded in mineral grease for presentation to the X-ray beam.[25] Thus, a prerequisite was to verify that crystalline IrisFP embedded in the grease matrix indeed remains photoswitchable (Figure S11a). The thermal recovery to the *on* from the *off* state occurs with a half-life of 220 and 65 min in IrisFP crystals with and without grease embedding, respectively (see Supporting Information and Figure S16). The more than 3-fold slowdown of IrisFP thermal recovery in grease is in line with a recent study on Dronpa in solution, whose photoswitching kinetics has been reported to slow down when the viscosity of the surrounding solution was increased.[26] Using in crystallo microspectrophotometry, we observed slow spectral changes (time scale of tens of minutes) after grease-embedding (Figure S11a) that are absent in a control experiment carried out on an IrisFP crystal without grease (Figure S11b). These changes are reverted upon restoring the initial humidity of crystalline IrisFP (Figure S11a). In order to assess a potential effect of grease embedding on the crystal structure, two data sets were collected at a synchrotron beamline by the standard oscillation method, from two IrisFP macrocrystals harvested from the same drop and flash-cooled at 100 K, either after cryoprotection and grease embedding (control 1), or after cryoprotection only (control 2). The Fourier difference electron density ($F_o^{\ 1} - F_o^{\ 2}$) map calculated between controls 1 and 2 (R_{iso} = 15.2%) shows no features around the chromophore or elsewhere in the protein above a contour level of ±3.5σ (not shown), indicating

DOI: 10.1021/acs.jpclett.5b02789
J. Phys. Chem. Lett. 2016, 7, 882−887

that there is no major effect of grease embedding on the structure of IrisFP at 100 K.

For SFX experiments, microcrystals (Figure S9; 10 × 10 × 50−80 μm^3) of IrisFP in its green *on* resting state were thus mixed with the grease and the mix containing randomly oriented microcrystals was streamed across the XFEL beam (pulse length <10 fs; 7 keV photon energy; 30 Hz repetition rate) by means of syringe-injection at BL3[27] of SACLA[28] (Figure S10). For 0.5 mg of crystalline IrisFP, 66300 diffraction patterns were collected within 37 min, of which 10899 were indexed using *CrystFEL*[29] (overall redundancy 159, CC*: 0.98, R_{split}: 22%; see Table S1 for details). The structure was solved by molecular replacement at 2.75 Å resolution (CC* in 2.90−2.75 Å resolution range: 0. 89; R_{work}/R_{free} 0.174/0.232; Table S1), using the structure of IrisFP determined at 100 K[16] as a starting model.

The structure of IrisFP (tetrameric; 226 amino acids per monomer,[16] of which 220−222 are visible in the electron density maps) determined by SFX at room temperature is very similar to the one determined by traditional cryocrystallography at the synchrotron (RMSD of 0.565 Å over the four monomers). In particular, the chromophore isomeric state can be determined as being *cis* (Figure 3a), as expected for the green fluorescent *on* state, and not *trans* (Figure 3b). Each of the three residues forming the chromophore (His62, Tyr63 and Gly64) are well defined in an electron density map calculated with a model from which the chromophore had been omitted (Figure 3a). Yet, the SFX structure displays minor changes compared to the synchrotron structure determined from a crystal embedded in grease and flash-cooled at 100 K (control 1). A decrease in unit cell volume is observed at 100 K (−6%) that is reflected in a reduction of the chromophore-cavity volume by 25% (for details see Supporting Information and Table S1). At the atomic level, the most striking difference between the SFX and the synchrotron structures is that Arg66 side chain, which stabilizes the chromophore hydrogen bonding, adopts two alternate conformations in the former (Figure 3a). This difference and the change in the cavity volume might be due to the difference in data collection temperature. We suggest that the SFX structure determined from room-temperature data is physiologically more meaningful.

Our time-resolved absorption spectroscopy experiments of IrisFP in solution provided evidence for chromophore isomerization with a yield of 17%, involving two excited states on the ps time scale. Kinetic modeling is in favor of a sequential mechanism in which the excited state I_2* with a lifetime of 15.3 ps controls the *off*-to-*on* switching quantum yield. The structural nature of the spectroscopically resolved intermediate state remains elusive. In view of future crystallographic experiments aiming at a structural characterization of this crucial intermediate, we used serial femtosecond crystallography (SFX) at an X-ray free electron laser to solve the structure of IrisFP in its *on* state. The structure of IrisFP in its green *on* state presented here can be envisioned as the first time point of a molecular movie. Time-resolved SFX on the ps time scale[23] has now the potential to structurally resolve details of the isomerization mechanism, such as the putative existence of twisted chromophore intermediates. Such structural snapshots will provide essential insights into the structural dynamics of a fluorescent protein during the early events of photoswitching. This, in turn, could prove essential for the rational design of more efficient fluorescent proteins for applications in biological nanoscopy.

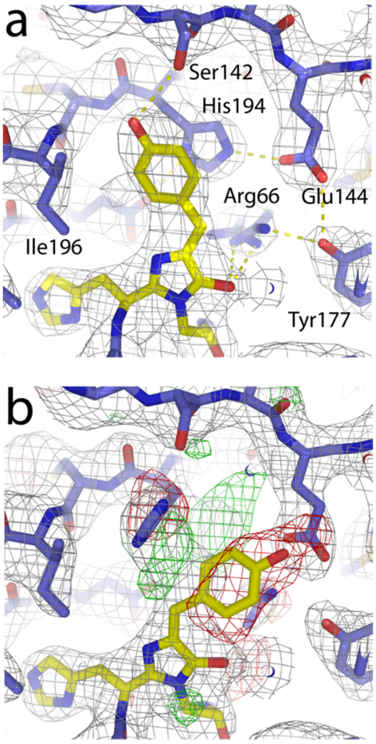


Figure 3. The IrisFP chromophore in the SFX structure is in the *cis* conformation. The refined *cis* model is overlaid in (a), featuring Arg66 in two alternate conformations, whereas the *trans* model[16] is overlaid in (b). Unbiased $2F_{obs} - F_{calc}$ (contoured in gray at 1 σ) and $F_o - F_c$ (contoured in red and green at −3 and +3 σ, respectively) electron density maps obtained from SFX data following molecular replacement starting with the *cis* (a) or the *trans* model (b) of IrisFP. Strong $F_o - F_c$ peaks appear on the *trans* chromophore and on His194 and Glu144 (b). Thus, only the *cis* model fits the $2F_{obs} - F_{calc}$ electron density map well.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.5b02789.

SFX and synchrotron data collection, processing and statistics, structure solution and refinement, absorption spectroscopy of IrisFP crystals, time-resolved absorption spectroscopy. (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

* E-mail: weik@ibs.fr.
* E-mail: colletier@ibs.fr.
* E-mail: michel.sliwa@univ-lille1.fr.

Author Contributions

(J.-P.C. and M.Sl.) These authors contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to Leonard Chavas and So Iwata for technical support at SACLA. J.P.C. thanks Duilio Cascio, Michael Sawaya, and David Eisenberg for stimulating discussions and continuing support. M.Sl. is grateful to Jean-Pierre Verwaerde and Julien Dubois for technical support in the rotating cell design and time-resolved experiments. The XFEL experiments were carried out at BL3 of SACLA with the approval of the Japan Synchrotron Radiation Research Institute (JASRI; Proposal No. 2013B8046), and the synchrotron experiments at ID29 of ESRF, under long-term projects MX1464, MX1583, and MX1676 (IBS BAG). We warmly thank the SACLA and ESRF staff for assistance.

The study was supported by a grant from the CNRS (PEPS SASLELX) to M.W., and ANR grants to M.W., M.C., M.Sl. (BioXFEL) and to D.B. (NoBleach). The Cryobench is a platform of the Grenoble Instruct center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from ESRF, FRISBI (ANR-10-INSB-05-02), and GRAL (ANR-10-LABX-49-01). The Chevreul Institute (FR 2638), the Ministère de l'Enseignement Supérieur et de la Recherche, the Région Nord—Pas de Calais and FEDER are acknowledged for financial support to M.Sl.

■ REFERENCES

(1) Dean, K. M.; Palmer, A. E. Advances in Fluorescence Labeling Strategies for Dynamic Cellular Imaging. *Nat. Chem. Biol.* **2014**, *10*, 512−523.

(2) Adam, V.; Berardozzi, R.; Byrdin, M.; Bourgeois, D. Phototransformable Fluorescent Proteins: Future Challenges. *Curr. Opin. Chem. Biol.* **2014**, *20*, 92−102.

(3) Nienhaus, K.; Ulrich Nienhaus, G. Fluorescent Proteins for Live-cell Imaging with Super-resolution. *Chem. Soc. Rev.* **2014**, *43*, 1088−1106.

(4) Ando, R.; Mizuno, H.; Miyawaki, A. Regulated fast Nucleocytoplasmic Shuttling Observed by Reversible Protein Highlighting. *Science* **2004**, *306*, 1370−1373.

(5) Wilmann, P. G.; Turcic, K.; Battad, J. M.; Wilce, M. C.; Devenish, R. J.; Prescott, M.; Rossjohn, J. The 1.7 A Crystal Structure of Dronpa: a Photoswitchable Green Fluorescent Protein. *J. Mol. Biol.* **2006**, *364*, 213−224.

(6) Andresen, M.; Stiel, A. C.; Trowitzsch, S.; Weber, G.; Eggeling, C.; Wahl, M. C.; Hell, S. W.; Jakobs, S. Structural Basis for Reversible Photoswitching in Dronpa. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 13005−13009.

(7) Habuchi, S.; Ando, R.; Dedecker, P.; Verheijen, W.; Mizuno, H.; Miyawaki, A.; Hofkens, J. Reversible Single-molecule Photoswitching in the GFP-like Fluorescent Protein Dronpa. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 9511−9516.

(8) Fron, E.; Flors, C.; Schweitzer, G.; Habuchi, S.; Mizuno, H.; Ando, R.; De Schryver, F. C.; Miyawaki, A.; Hofkens, J. Ultrafast Excited-state Dynamics of the Photoswitchable Protein Dronpa. *J. Am. Chem. Soc.* **2007**, *129*, 4870−4871.

(9) Warren, M. M.; Kaucikas, M.; Fitzpatrick, A.; Champion, P.; Sage, J. T.; van Thor, J. J. Ground-state Proton Transfer in the Photoswitching Reactions of the Fluorescent Protein Dronpa. *Nat. Commun.* **2013**, *4*, 1461.

(10) Walter, A.; Andresen, M.; Jakobs, S.; Schroeder, J.; Schwarzer, D. Primary Light-induced Reaction Steps of Reversibly Photoswitchable Fluorescent Protein Padron0.9 Investigated by Femtosecond Spectroscopy. *J. Phys. Chem. B* **2015**, *119*, 5136−5144.

(11) Kaucikas, M.; Tros, M.; van Thor, J. J. Photoisomerization and Proton Transfer in the Forward and Reverse Photoswitching of the Fast-switching M159T Mutant of the Dronpa Fluorescent Protein. *J. Phys. Chem. B* **2015**, *119*, 2350−2362.

(12) Yadav, D.; Lacombat, F.; Dozova, N.; Rappaport, F.; Plaza, P.; Espagne, A. Real-time Monitoring of Chromophore Isomerization and Deprotonation during the Photoactivation of the Fluorescent Protein Dronpa. *J. Phys. Chem. B* **2015**, *119*, 2404−2414.

(13) Lukacs, A.; Haigney, A.; Brust, R.; Addison, K.; Towrie, M.; Greetham, G. M.; Jones, G. A.; Miyawaki, A.; Tonge, P. J.; Meech, S. R. Protein Photochromism Observed by Ultrafast Vibrational Spectroscopy. *J. Phys. Chem. B* **2013**, *117*, 11954−11959.

(14) Li, X.; Chung, L. W.; Mizuno, H.; Miyawaki, A.; Morokuma, K. Primary Events of Photodynamics in Reversible Photoswitching Fluorescent Protein Dronpa. *J. Phys. Chem. Lett.* **2010**, *1*, 3328−3333.

(15) Moors, S. L. C.; Michielssens, S.; Flors, C.; Dedecker, P.; Hofkens, J.; Ceulemans, A. How Is cis−trans Isomerization Controlled in Dronpa Mutants? A Replica Exchange Molecular Dynamics Study. *J. Chem. Theory Comput.* **2008**, *4*, 1012−1020.

(16) Adam, V.; Lelimousin, M.; Boehme, S.; Desfonds, G.; Nienhaus, K.; Field, M. J.; Wiedenmann, J.; McSweeney, S.; Nienhaus, G. U.; Bourgeois, D. Structural Characterization of IrisFP, an Optical Highlighter Undergoing Multiple Photo-induced Transformations. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 18343−18348.

(17) Chapman, H. N.; Fromme, P.; Barty, A.; White, T. A.; Kirian, R. A.; Aquila, A.; Hunter, M. S.; Schulz, J.; DePonte, D. P.; Weierstall, U.; et al. Femtosecond X-ray Protein Nanocrystallography. *Nature* **2011**, *470*, 73−77.

(18) Schlichting, I. Serial Femtosecond Crystallography: the First Five Years. *IUCrJ* **2015**, *2*, 246−255.

(19) Duan, C.; Byrdin, M.; El Khatib, M.; Henry, X.; Adam, V.; Bourgeois, D. Rational Design of Enhanced Photoresistance in a Photoswitchable Fluorescent Protein. *Methods Appl. Fluoresc.* **2015**, *3*, 014004.

(20) Ruckebusch, C.; Sliwa, M.; Pernot, P.; de Juan, A.; Tauler, R. Comprehensive Data Analysis of Femtosecond Transient Absorption Spectra: A Review. *J. Photochem. Photobiol., C* **2012**, *13*, 1−27.

(21) Gueret, R.; Castillo, C. E.; Rebarz, M.; Thomas, F.; Hargrove, A. A.; Pecaut, J.; Sliwa, M.; Fortage, J.; Collomb, M. N. Cobalt(III) tetraaza-macrocyclic Complexes as Efficient Catalyst for Photoinduced Hydrogen Production in Water: Theoretical Investigation of the Electronic Structure of the Reduced Species and Mechanistic Insight. *J. Photochem. Photobiol., B* **2015**, *152*, 82−94.

(22) Meech, S. R. Excited State Reactions in Fluorescent Proteins. *Chem. Soc. Rev.* **2009**, *38*, 2922−2934.

(23) Barends, T. R.; Foucar, L.; Ardevol, A.; Nass, K.; Aquila, A.; Botha, S.; Doak, R. B.; Falahati, K.; Hartmann, E.; Hilpert, M.; et al. Direct Observation of Ultrafast Collective Motions in CO Myoglobin upon Ligand Dissociation. *Science* **2015**, *350*, 445−450.

(24) Tenboer, J.; Basu, S.; Zatsepin, N.; Pande, K.; Milathianaki, D.; Frank, M.; Hunter, M.; Boutet, S.; Williams, G. J.; Koglin.; et al. Time-resolved Serial Crystallography Captures High-resolution Intermediates of Photoactive Yellow Protein. *Science* **2014**, *346*, 1242−1246.

(25) Sugahara, M.; Mizohata, E.; Nango, E.; Suzuki, M.; Tanaka, T.; Masuda, T.; Tanaka, R.; Shimamura, T.; Tanaka, Y.; Suno.; et al. Grease Matrix as a Versatile Carrier of Proteins for Serial Crystallography. *Nat. Methods* **2015**, *12*, 61−63.

(26) Kao, Y. T.; Zhu, X.; Min, W. Protein-flexibility Mediated Coupling Between Photoswitching Kinetics and Surrounding Viscosity of a Photochromic Fluorescent Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 3220−3225.

(27) Tono, K.; Togashi, T.; Inubushi, Y.; Sato, T.; Katayama, T.; Ogawa, K.; Ohashi, H.; Kimura, H.; Takahashi, S.; Takeshita, K.; et al. Beamline, Experimental Stations and Photon Beam Diagnostics for the Hard X-ray Free Electron Laser of SACLA. *New J. Phys.* **2013**, *15*, 083035.

(28) Ishikawa, T.; Aoyagi, H.; Asaka, T.; Asano, Y.; Azumi, N.; Bizen, T.; Ego, H.; Fukami, K.; Fukui, T.; Furukawa, Y.; et al. A compact X-

ray Free-electron Laser Emitting in the Sub-angstrom Region. *Nat. Photonics* **2012**, *6*, 540−544.

(29) White, T. A.; Barty, A.; Stellato, F.; Holton, J. M.; Kirian, R. A.; Zatsepin, N. A.; Chapman, H. N. Crystallographic Data Processing for Free-electron Laser Sources. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1231−1240.

Appendix D: Simple and efficient system for photoconverting light-sensitive proteins in serial crystallography experiments

JOURNAL OF APPLIED CRYSTALLOGRAPHY

ISSN 1600-5767

Simple and efficient system for photoconverting light-sensitive proteins in serial crystallography experiments

Giorgio Schirò,[a] Joyce Woodhouse,[a] Martin Weik,[a] Ilme Schlichting[b] and Robert L. Shoeman[b]*

[a]Institut de Biologie Structurale, Université Grenoble Alpes, CEA, CNRS, Grenoble, 38044, France, and [b]Max Planck Institute for Medical Research, Jahnstrasse 29, Heidelberg, 69120, Germany. *Correspondence e-mail: robert.shoeman@mpimf-heidelberg.mpg.de

Received 13 December 2016
Accepted 25 April 2017

Edited by S. Boutet, SLAC National Accelerator Laboratory, Menlo Park, USA

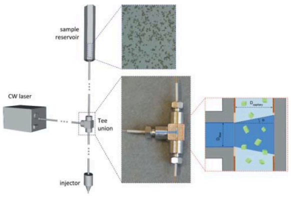
Keywords: serial femtosecond crystallography; light-sensitive proteins; pre-illumination; time resolution; instrumentation.

Supporting information: this article has supporting information at journals.iucr.org/j

Proteins that change their structure in response to light absorption regulate many functional processes in living cells. Moreover, biotechnological approaches like optogenetics and super-resolution fluorescence microscopy recently triggered the generation of new genetically modified photosensitive proteins. Light-induced structural changes in photosensitive proteins can be studied by time-resolved serial femtosecond crystallography (SFX), an X-ray diffraction technique that allows the determination of macromolecular structures at X-ray free-electron lasers from a large number of nano- to micro-sized crystals. This article describes a simple and efficient system for converting photosensitive proteins into light-induced semi-stationary states by inline laser illumination prior to sample injection with a gas-focused liquid jet and subsequent optical pump–X-ray probe exposure. The simple setup of this device makes it suitable for integration into other liquid injectors (like electro-spinning and electro-kinetic injectors) and potentially also in high-viscosity extruders, provided that embedding microcrystals in viscous media does not alter protein photophysical properties. The functioning of the device is demonstrated with an example of a photoswitchable fluorescent protein pre-illuminated (photoactivated) for time-resolved SFX experiments. The device can be easily adapted for the conversion in time-resolved SFX experiments of other microcrystalline proteins, such as photosystems, phytochromes and rhodopsins.

1. Introduction

Light-responsive proteins initiate and regulate many fundamental processes in living systems, such as vision, photosynthesis, phototaxis and the circadian rhythm. They contain endogeneous or prosthetic groups (chromophores) that are sensitive to a narrow band of the visible spectrum. Photon absorption triggers a change in the photophysical state of the chromophore that, in turn, induces a photocycle consisting of various consecutive intermediates, involving more or less extensive protein structural changes. Some photoactive proteins have semi-stable light-inducible states such as the fully light-adapted ground state in bacteriorhodopsin that has full occupancy of all-*trans* retinal (Vodyanoy *et al.*, 1985). Similarly, reversibly photoswitchable proteins such as some fluorescent proteins (Grotjohann *et al.*, 2012) or phytochromes (Rockwell *et al.*, 2006) can be photoconverted from one state to another. These photoconverted states can then serve as starting points for time-resolved (TR) optical pump–X-ray probe experiments. Studying the photoinduced structural changes connecting intermediates in a photocycle or



accompanying photoconversion is crucial for understanding the physical basis of light-responsive protein function.

The structural dynamics of light-responsive proteins can be studied down to the femtosecond time scale by time-resolved serial femtosecond crystallography (SFX). SFX (Chapman *et al.*, 2011; Schlichting, 2015) uses the femtosecond X-ray pulses produced by free-electron lasers (XFELs) (Emma *et al.*, 2010; Ishikawa *et al.*, 2012) to collect single-shot diffraction patterns of nano- or micro-sized crystals delivered for example in a liquid microjet. Tens of thousands of patterns are collected in a serial manner, from which a dataset is constructed that then allows structure determination with standard tools used in macromolecular crystallography. The introduction of SFX allows data collection from crystals too small for synchrotron-based crystallography (Kang *et al.*, 2015; Zhang *et al.*, 2015), room-temperature radiation-damage-free structure determination (Boutet *et al.*, 2012; Hirata *et al.*, 2014), *in vivo* crystallography (Sawaya *et al.*, 2014) and, of interest here, time-resolved studies on the sub-picosecond time scale (Barends *et al.*, 2015; Pande *et al.*, 2016).

In time-resolved crystallography experiments of light-responsive proteins, reaction initiation is carried out by a visible laser pulse (pump) and structural changes are monitored at distinct and variable time delays by X-ray diffraction (probe). Pump–probe experiments at synchrotron sources in the form of Laue crystallography have a long history (Moffat *et al.*, 1984; Schlichting *et al.*, 1990; Srajer *et al.*, 1996). However, they have been limited by (1) a temporal resolution restricted to 100 ps, (2) only partial reaction initiation due to limited laser penetration into macrocrystals of enormous optical density (the protein concentration in crystals is typically more than 10 m*M*) and (3) the very high crystal quality required by the Laue approach. TR-SFX using femtosecond pulses from XFELs overcomes all three limitations because (1) sub-ps temporal resolution can be attained, (2) microcrystals can be utilized, which (because of their smaller dimensions) means that all of, or at least a large fraction of, the molecules in a microcrystal can be light activated, and (3) crystal quality requirements are less stringent than for Laue experiments. However, since SFX datasets are based on a large number of diffraction events on different crystals, an important requirement is that in each crystal the light-responsive proteins are in the same photophysical state before they are activated with a pump-laser pulse. If the starting state is the resting state, crystals can be subjected to TR-SFX without prior illumination. However, if the starting state is a light-activated or photoconverted state, pre-illumination of the crystals prior to pump-laser triggering needs to be carried out. Such a pre-illumination requires that conditions such as the number of photons per molecule, the laser power density on the sample *etc.* are the same for every crystal. An ideal pre-illumination should fulfil the following requirements:

(1) Full crystal photoconversion or, if not achievable, homogeneity of the conversion fraction for all crystals.

(2) Depending on the dark reversion time, minimization of the transit time between laser illumination and X-ray diffraction.

(3) The absence or minimization of light-induced damage on chromophore, protein structure and crystal quality.

(4) Easy integration in serial crystallography environments with stable connection and alignment.

Here we describe the rationale and the design of a simple pre-illumination device to photoconvert crystalline proteins prior to pump-laser triggering in TR-SFX experiments that fulfils the above requirements. We then report on its successful operation in the case of the reversibly photoswitchable fluorescent protein rsEGFP2 (Grotjohann *et al.*, 2012), where the fluorescent resting state is photoconverted to a metastable non-fluorescent state by blue light (Fig. 1). rsEGFP2 is the most widely used molecular marker in reversible saturable optical fluorescence transition (RESOLFT) super-resolution fluorescence microscopy (Hofmann *et al.*, 2005; Grotjohann *et al.*, 2012). We also demonstrate implementation during SFX experiments at the two XFELs currently operational, *i.e.* the LINAC Coherent Light Source (LCLS) and the SPring-8 Angstrom Compact Linear Accelerator (SACLA). Possible applications of the device for exploring light-induced states in other photosensitive proteins are also briefly discussed.

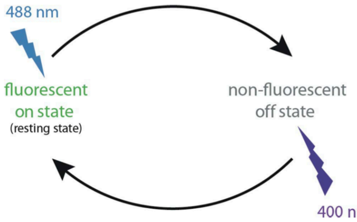


Figure 1
Reversible photoswitching in rsEGFP2 from the fluorescent on state (resting state) with 488 nm light to a non-fluorescent off state (metastable) and back with 400 nm light to the on state.

2. Material and methods

2.1. Methods

2.1.1. General methods. Unless otherwise stated, optical components were purchased from ThorLabs GmbH (Dachau/Munich, Germany) and the high-pressure liquid chromatography (HPLC) components were Upchurch Scientific (IDEX Health & Science) products (purchased from VWR International, http://de.vwr.com). An Oxxius LBX-488-200-CSB-PP laser, with SMA coupler and controller, was purchased from Laser 2000 GmbH (Wessling, Germany). A Gentec Maestro power monitor with a PH100 photo detector, SMA adapter and OD 1 filter were purchased from Laser Components GmbH (http://www.lasercomponents.com). Samples were pumped through polyether ether ketone (PEEK), Radel (Upchurch Scientific) or fused silica capillary tubing [Polymicro TSP (round bore) or WWP (square bore) tubing, both with a brown polyimide coating; purchased from Optronis GmbH, Kehl, Germany] using a Shimadzu Biotech LC-20AD pump and, for experiments at LCLS or SACLA, with an anti-settling sample delivery instrument as previously described (Lomb *et al.*, 2012). Further details are provided in the methods section of the supporting information. After the completion of these experiments, we learned that Polymicro TSU tubing with round bore inner diameters of 50, 75 or 100 µm and a UV/vis light-transparent coating (*ca* 90% transmission at > 400 nm) is also available and believe it might be a better alternative (less sensitive to breakage) than the heat-treated TSP or WWP tubing. Protein and protein crystal absorption spectra were measured with a Nanodrop 2000c spectrometer (Thermo Scientific, http://www.nanodrop.com) and treated for smoothing, background subtraction and normalization with a *Gnuplot*-based (http://www.gnuplot.info/) script.

2.1.2. Optical methods. As shown in Fig. 2, an HPLC stainless steel tee union (Upchurch #U-428, bore diameter 0.5 mm) was used to hold an optical fibre (polished, clear FG200LCC core from a cut Thorlabs M25L05 SMA fibre) in place on and in direct contact with the side of a capillary through which the protein crystals were pumped. The brown polyimide coating on the sample capillary was removed by heating to produce an optically clear window approximately 5 mm long. This clear window was centred in the tee union such that light from the SMA optical fibre directly illuminated the exposed window. As noted by the manufacturer, tubing with an optically clear window produced by heating was sensitive to breakage during assembly of the illumination tee union. The precisely machined interior of the tee union and the shape of the fused silica capillary effectively constitute an illumination cuvette. Several configurations were tested, with WWP fused silica tubing with a square cross section outer diameter (OD) of 300 µm and an inner square cross section of 50 or 100 µm on a side or with round cross section 360 µm OD TSP tubing with inner diameters (ID) up to 250 µm. In some experiments, the straight sample path through the tee union was drilled to accept 1/16″ (1″ = 25.4 mm) OD tubing, and clear 1/16″ Radel tubing with an ID of 254 µm was used instead of the fused silica tubing (see Fig. S1). The sample tubing assembled in the illumination tee union was generally about 25 cm in length and was one continuous piece (there was no liquid/liquid junction within the tee connector). The SMA optical fibre employed for all experiments was prepared by cutting off the SMA connector and sheath and stripping the cladding from one end of a regular Thorlabs M25L05 fibre. The exposed core was end-polished on an Allied Systems wet polishing machine with a 5 µm diamond abrasive. The end of the SMA fibre core could be easily positioned to touch the surface of the sample capillary or tubing and was readily fixed in place by tightening the tee union fingertight fittings (or compression nuts and ferrules) to secure the tubing sleeves used to hold the capillary. The grey sleeves were cut from 1/16″ PEEK tubing with an ID of 400 µm (Upchurch Scientific #1565). Heat shrink tubing was placed over the grey sleeves and inside the orange protective covering to eliminate stray light radiating from the exposed core (visible in Fig. S1). According to the manufacturer's specifications, M25L05 fibre has a numerical aperture of 0.22 (full acceptance angle of 25°) and a transmission of 488 nm light in excess of 99% (attenuation $\simeq$ 30 dB km^{-1}). As indicated, SMA-terminated fibres were coupled using the Thorlabs ADASMA mating sleeve, which has a nominal air gap of 0.056 ± 0.046 mm.

2.1.3. rsEGFP2 microcrystals. rsEGFP2 (reversible switchable enhanced green fluorescent protein 2; Grotjohann *et al.*, 2012) microcrystals were grown in micro-batches using seeding (details to be published elsewhere), starting from conditions reported earlier (El Khatib *et al.*, 2016). The microcrystals (∼3–5 µm; Fig. S2) were stored in 2 *M* ammonium sulfate, 20 m*M* NaCl, 120 m*M* HEPES pH 8. For the laboratory characterization of the photoconversion of the rsEGFP2 microcrystals, the microcrystals were kept in plastic microcentrifuge tubes loosely covered with aluminium foil and the laboratory window shades were closed to minimize exposure to daylight. The experiments were performed with regular laboratory illumination (not in the dark).

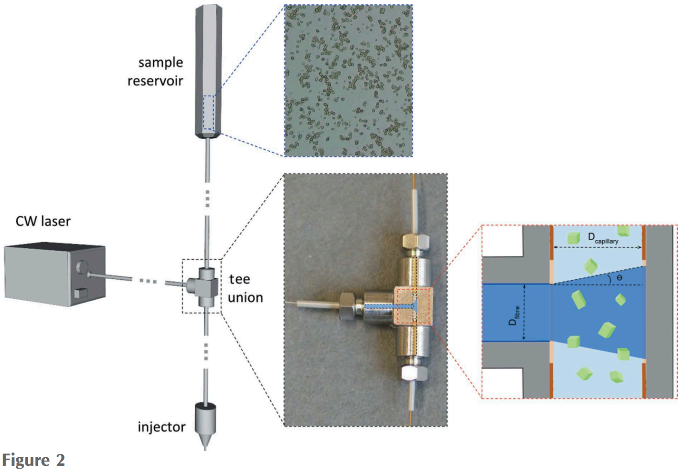


Figure 2
An overview of the HPLC and optical hardware. An Upchurch HPLC tee union (500 µm thru-passages) with a 360 µm OD sample capillary held in the central bore of the tee by grey sleeves, ferrules and compression nuts. The SMA fibre core is held in position in the centre of the side port by a grey sleeve and is at 90° to the sample capillary and is touching the capillary. The upper photograph shows the rsEGFP microcrystals used in these experiments (see also Fig. S2), and the artist's rendition of the cross section of the tee union illustrates the relationship of the optical fibre to the sample capillary, with a schematic representation of the rsEGFP microcrystals (green boxes) passing through the illumination region (blue cone of light)

3. Results

The pre-illumination device (Fig. 2) allows photoconversion of microcrystals of light-responsive proteins from one state to another 'on the fly' in SFX experiments that use liquid injectors. A tee union is inserted in the injection pipeline, within which microcrystals are illuminated with light from a continuous wave (CW) laser as they flow by. We tested the device with the fluorescent protein rsEGFP2 (Grotjohann *et al.*, 2012), which can be reversibly and repeatedly photoswitched from the resting state, the fluorescent on state, to an off state and back, by illumination at 488 and 400 nm, respectively (Fig. 1). We first present a formula expressing the photoconversion efficiency as a function of parameters linked to crystal properties (size, protein concentration, extinction coefficient *etc.*), pre-illumination (laser power, geometry *etc.*) and liquid injection (flow speed). Subsequently, we assessed the efficiency of the device by a laser-power titration in a home laboratory that switches the resting state (on state) to the off state. Next, we determined the lifetime of the off state in the dark, defining the available time span between pre-illumination and pump-laser activation in TR-SFX experiments that aim at studying photoswitching from the off to the on state. We addressed potential laser damage by determining the fraction of molecules that revert back in the dark from the off to the on state. Finally, successful implementation and operation of the device in an SFX pipeline at the LCLS and SACLA is described.

3.1. Definition of photoconversion efficiency

The photoconversion efficiency of a given sample, *i.e.* the fraction of protein molecules photoactivated by laser light, is a function of the liquid flow rate (how long is the sample exposed?), the laser power (how many photons are available for absorption?), the dimensions of the inner diameter of the capillary or tubing (what is the light/photon density and sample layer thickness?), the alignment of the components (how many photons actually strike the sample?), the concentration of the sample (are there enough photons available to convert every molecule?) and the quantum efficiency of the illumination process. The actual photoconversion efficiency can be calculated as a function of the experimental parameters by the following relation:

$$\text{photoconversion efficiency} = \frac{P\alpha\Sigma_{\rm Xtal}\big(D_{\rm fibre} + D_{\rm capillary}\tan\theta\big)\big(1 - 10^{-\varepsilon[c]\delta_{\rm Xtal}}\big)}{(hc/\lambda)\,[c]\,N_{\rm A}V_{\rm Xtal}\,Q/\Sigma_{\rm capillary}}\,\Phi, \qquad (1)$$

where P is the laser power density, α is the light transmission coefficient of the capillary wall, $\Sigma_{\rm Xtal}$ is the mean crystal surface normal to the laser light direction, $D_{\rm fibre}$ and $D_{\rm capillary}$ are the diameters of the optical fibre and of the capillary, respectively, θ is the acceptance angle of the optical fibre, ε is the protein extinction coefficient, $[c]$ is the protein concentration inside the crystal, $\delta_{\rm Xtal} = D_{\rm capillary}\rho$, with ρ the crystal volume fraction, is the 'effective' crystal thickness taking into account the possibility that more than one crystal is in the laser path, as also illustrated in Fig. 2, h is the Planck constant, c is the speed of light, λ is the laser wavelength, $N_{\rm A}$ is the Avogadro number, $V_{\rm Xtal} = \Sigma_{\rm Xtal}\delta_{\rm Xtal}$ is the 'effective' mean crystal volume, Q is the flow rate, $\Sigma_{\rm capillary}$ is the capillary cross section area, and Φ is the quantum efficiency. Solving equation (1) for our test conditions predicts that 100 mW of laser power should be the minimum power needed to saturate the conversion of the rsEGFP sample from the off state to the on state. Given the simple and rigid design of the system shown in Fig. 2, in which the alignment is provided by the high precision of the tee union, there are only four experimental variables that need to be investigated for the rsEGFP2 microcrystals: laser power, capillary ID, sample concentration and sample flow rate.

3.2. Laser power titration and device efficiency

We tested the device by assessing the photoconversion efficiency of rsEGFP2 on to off photoswitching (Fig. 1) triggered by illumination at 488 nm. To this end, spectra of the crystal suspension were collected following passage through the tee union after illumination as a function of nominal laser power. The time necessary for collecting a sample droplet, putting it on the Nanodrop sample holder and acquiring the spectrum was of the order of 15 s, setting the lower limit of time resolution for our tests. The experimental parameters were 30 µl min^{-1} flow rate (typically used in SFX experiments with these rsEGFP2 microcrystals), 250 µm ID round quartz capillary and 33%(*v*/*v*) crystal concentration (the maximum typically used in our various SFX experiments). The spectra are reported in Fig. 3(*a*), with the absorbance at the maximum of the on-state spectrum (490 nm) reported in Fig. 3(*b*). Assuming that the absorbance at 490 nm, A(490 nm), is proportional to the number of protein molecules in the on state $n_{\rm ON}$ (after correction for the background) we can write

$$\begin{aligned} {\rm d}n_{\rm ON} &= -\sigma\,\Phi\,P\,n_{\rm ON}(t)\,{\rm d}t \Rightarrow n_{\rm ON} \\ &= n_{\rm ON}(0)\exp(-\sigma\,\Phi\,\Delta t\,P) \propto A(490\,{\rm nm}), \end{aligned} \qquad (2)$$

where the symbol d refers to infinitesimal variation, t is the time, σ is the rsEGFP2 absorption cross section, Φ (= 0.016; El Khatib *et al.*, 2016) is the quantum efficiency (*i.e.* the probability that an absorbed photon leads to photoswitching from on to off), Δt is the transit time of a crystal under laser light (here 25 ms) and P is the laser power density. Considering that σ, Φ and Δt are constant once the experimental conditions are finalized, equation (2) defines an exponential dependence of A(490 nm) on laser power. The continuous line in Fig. 3(*b*) is a fit with an exponential decay and a constant term, accounting for the background and perfectly approximating the experimental trend. Comparing the absorbance at maximum nominal power (200 mW) with the asymptotic value obtained by the fit procedure, we estimate that about 97% of the saturation value is attained when using 200 mW. Owing to the overlap of on and off spectra at 488 nm, the saturation condition corresponds to about 90% of on → off switching. We checked that pre-illumination works efficiently even at the highest flow rate used during our SFX experiments

($40\ \mu l\ min^{-1}$) with the rsEGFP microcrystals and with a different capillary geometry (100 µm square *versus* 250 µm round ID; data not shown).

3.3. Estimation of laser power loss

In view of the remote placement of the laser head from the illumination point within the tee union, necessary when integrating the device in an XFEL beamline setup, we first tested the power loss due to the use of long SMA fibres to guide laser light to the sample. The Oxxius 488 nm laser has a nominal CW maximum power rating of 200 mW. With the factory SMA collimator/connector and an intact M25L05 fibre (core diameter of 200 µm) installed, the maximum output observed from the fibre was 184 mW (Table S1). Other lengths and types of fibres with two SMA connectors, with one bare end, or two intact SMA fibres joined with an air gap optical union were also tested (Table S1). The electronic Oxxius laser controller permits both local and remote modulation of laser power output from 0 to nominal 200 mW. As expected, reducing the power output of the laser resulted in a corresponding reduction in the power measured at the output of the SMA fibre (Table S1). The quality of the polishing of the naked fibre end also had an influence on the output of the SMA fibre (Table S1). Losses of about 10% or more were observed when multiple SMA fibres were joined with air gap SMA connectors (Table S1 and data not shown), which is not surprising given the fibre acceptance full angle of 25°. To permit remote placement of the laser head from the illumination tee union, a 5 m long SMA fibre with one exposed core and a highly polished end (Table S1) was used for two experiments at the LCLS and SACLA (see below).

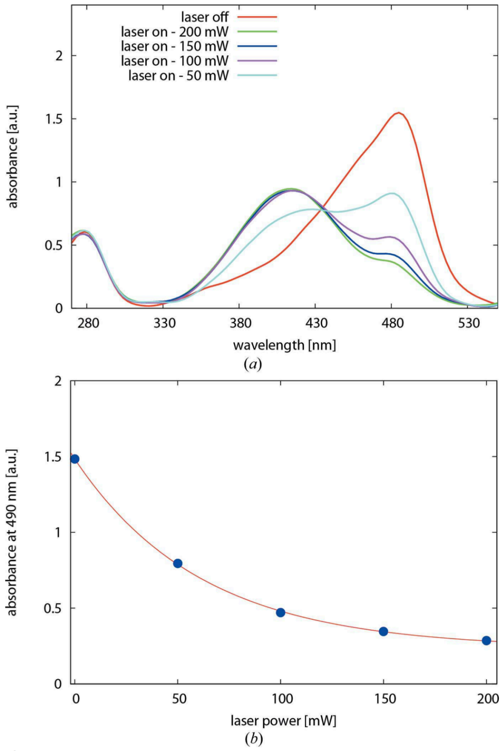


Figure 3
The photoconversion efficiency of rsEGFP2 as a function of laser power. (*a*) Photo-induced spectral changes as a function of laser nominal power, from 0 to 200 mW. (*b*) The absorbance at 490 nm as a function of nominal laser power reveals that at 200 mW photoswitching is in the saturation regime (about 97% of the saturation value). The continuous line is a fit with an exponential function (see main text).

3.4. Dark reversion kinetics

Once the microcrystalline photoresponsive proteins have been photoconverted on the fly with the pre-illumination device, the lifetime of the generated states needs to be much longer than the time it takes for the crystals to reach the point where they will be illuminated by the pump laser in a TR-SFX experiment. In order to determine the lifetime of the off state in rsEGFP2 microcrystals, the latter were concentrated and placed in a 100 µm thick Infrasil cuvette and illuminated with light from a high-power 490 nm LED (Thorlabs M490L3) for about 20 min, resulting in an efficient conversion of the rsEGFP2 spectrum to that characteristic of the off state (not shown). Using these off-state microcrystals as a starting point, we determined the kinetics of dark thermal recovery of the on state by collecting absorption spectra as a function of time with a JASCO V-670 spectrometer. The fraction of on-state recovery was calculated by the following quantity:

$$\text{fraction of on-state recovery} = \frac{A(490\,\text{nm}, t) - A(490\,\text{nm, off})}{A(490\,\text{nm, on}) - A(490\,\text{nm, off})}, \qquad (3)$$

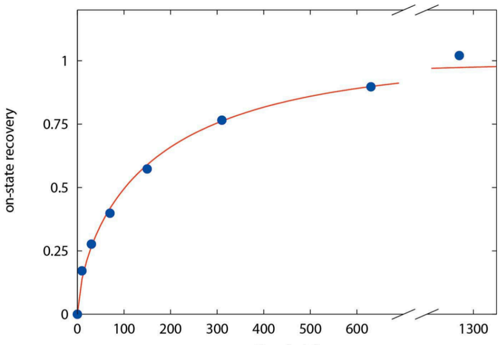


Figure 4
Kinetics of on-state thermal recovery after fluorescence photoswitching with 490 nm LED light. The continuous line is a fit with a stretched exponential $1 - \exp[-(kt)^{\alpha}]$, with best values for k and α of $0.0056\ min^{-1}$ and 0.66, respectively. The time for half of the molecules to recover is about 100 min, while the process is over after about 1000 min.

where A(490 nm, on) and A(490 nm, off) are the absorbance values at 490 nm recorded just before and just after LED photoswitching, respectively, while A(490 nm, t) is the same quantity at a generic time t after photoswitching. This quantity is 0 and 1 when 0 and 100% of rsEGFP2 microcrystals are in the on state, respectively. The fraction of on-state recovery is plotted in Fig. 4. Crystalline rsEGFP2 protein (in 2 *M* ammonium sulfate, 20 m*M* NaCl, 120 m*M* HEPES pH 8) in the off state thermally decays to the on state with a half-life of about 100 min (see Fig. 4), compared to about 278 min in solution at lower ionic strength and pH (50 m*M* HEPES pH 7.5; El Khatib *et al.*, 2016). Boundary conditions for successful sample injection of crystal suspensions in high-salt buffers employing our gas dynamic virtual nozzle (GDVN) injection system include flow rates with maxima of the order of 30–40 µl min^{-1} (Boutet *et al.*, 2012). Transit times for delivery of the sample from the tee union inserted into our GDVN injector system in the experiments performed at LCLS and SACLA are of the order of $\leq$1 min, *i.e.* much shorter than the thermal back switching time (see Fig. 4), making it possible to position the inline photo-illumination assembly outside of the vacuum/helium chamber, between the sample reservoir and the exit of the injector nozzle itself. It should be noted that, since the sample capillary is continuous through the pre-illumination tee union, there is no interruption of sample flow and no influence of this pre-illumination hardware on the sample delivery.

3.5. Assessing laser radiation damage during pre-illumination

Mandatory requirements for the device are the absence of any laser-induced modification of chromophore photophysical and photochemical properties other than photoswitching, of protein structure and of crystal quality. We checked the rsEGFP chromophore integrity by measuring the absorption spectrum after complete on-state thermal recovery. Fig. 5 shows the comparison of absorption spectra collected before laser illumination, directly following laser illumination and after 48 h of thermal recovery in the dark. Interestingly, the recovery after 48 h in the dark resulted in a slightly higher recovery of the on state compared to the starting material, which had been handled in relatively low level illumination in the laboratory. Since the spectral line shape is highly sensitive to structural and chemical modifications of the chromophore and its environment, the full recovery of the on-state spectrum confirms that laser illumination prior to injection does not perturb the properties of the rsEGFP2 chromophore. SFX experiments revealed that rsEGFP2 microcrystals diffracted to 1.7 Å resolution both with and without pre-illumination (Coquelle *et al.*, 2017).

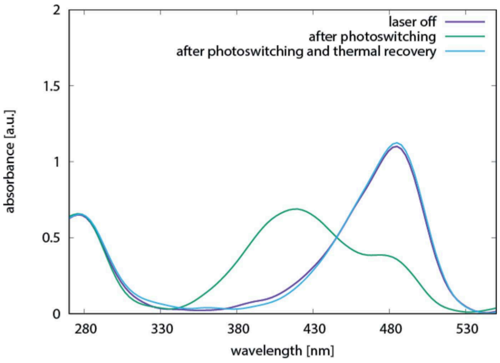


Figure 5
Comparison of spectra obtained before pre-illumination (purple), after photoswitching to the off state (green) and after 48 h of thermal recovery of the on state (blue). The recovery of the resting (on) state (purple) demonstrates that the pre-illumination protocol does not alter the spectral responses of the rsEGFP2 protein.

3.6. Integrating the pre-illumination device in an SFX experiment

We demonstrated successful pre-illumination in actual SFX experiments, during two experimental sessions at the LCLS (April–May 2015; CXI end station; proposal No. LI56) and SACLA (July 2015; end station BL3-EH4c; proposal No. 2015A8031). We collected absorption spectra as a function of pre-illumination laser power; the spectra obtained during the SACLA SFX experiment (analogous data obtained during the LCLS experiment are not shown) are plotted in Fig. 6 and demonstrate the efficiency of the device when integrated in the setup for SFX experiments. The only change we made to the device during the beamtimes at SACLA and LCLS was the use of clear Radel tubing instead of the glass capillary (see Figs. 2 and S1). During the experiment we noticed that the presence of connections increases the risk of clogging (Fig. S3), which entails a considerable loss of sample and time. Clogging is probably due to the increase of impedance at the

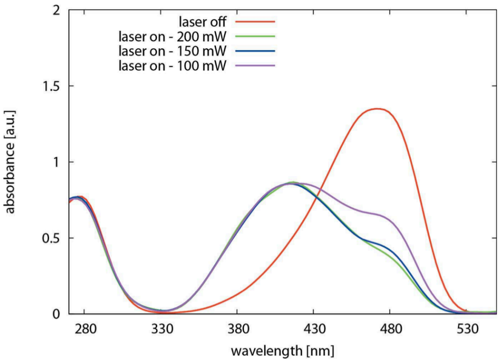


Figure 6
Spectra of rsEGFP2 microcrystals after having passed through the pre-illumination device in the tubing on their way to the GDVN sample injector during an experiment at SACLA XFEL. The power of the pre-illumination laser (488 nm) was 0 mW (red), 100 mW (magenta), 150 mW (blue) and 200 mW (green). The spectrum at 200 mW corresponds to a mixture of 90% rsEGFP2 molecules in the off state and 10% in the on state.

connections between tubing of different diameter, shape and material. More detail is presented in the methods section of the supporting information. As expected, the occurrence of clogging decreases with lowering crystal concentration. It is reasonable that the possibility of clogging markedly depends also on the chemical/hydrodynamic details of the crystal suspension (viscosity, salt concentration, crystal shape *etc.*). In general, to take into account all possible sources of clogging, we suggest to exhaustively test the device efficiency in-house before any XFEL experiment by integrating it into the entire injection system.

4. Discussion

We have designed a pre-illumination device that enables photoconversion of microcrystalline photoresponsive proteins on the fly during liquid-jet injection in SFX experiments. The two major features that distinguish this device from others (see later discussion) are (1) the high-intensity coherent light from the laser light source is completely contained within the device and (2) the alignment of the illumination on the sample by the tee union is simple and precise. The device was tested both offline in the home laboratory and online during two XFEL beamtimes, during which microcrystals of the reversibly photoswitchable protein rsEGFP2 were photoconverted from the on to the off state with 90% efficiency (100% is not attainable for rsEGFP2 owing to the overlap of on and off spectra). In the online XFEL beamtimes, the rsEGFP2 microcrystals in the off state were subjected to TR-SFX with picosecond to multiple nanosecond time delays to study the structural details and kinetics of the off- to on-state transition. The description of these experiments and the results will be reported elsewhere (Coquelle *et al.*, 2017).

While not tested for our experiments, it is obvious that the HPLC tee union within which flowing microcrystals are being illuminated with laser light (Fig. 2) could be replaced by, or supplemented with, a four-way cross (Upchurch Scientific U-430) or even a multiport manifold (such as the Upchurch Scientific P-170). Adding additional ports would enable the connection of a power meter and/or a modular spectrometer and/or permit illumination with multiple lasers. The higher light intensities that can be produced with more lasers were not needed for the experiments with rsEGFP2 but could be necessary when studying larger crystals and/or proteins with higher extinction coefficients and/or higher concentration and/or higher flow rates. TR-SFX experiments at SACLA on bacteriorhodopsin (Nango *et al.*, 2016) employed two focused laser beams.

Although we used the device only with a liquid-jet injector during SFX experiments at an XFEL, high-viscosity injection (Weierstall *et al.*, 2014; Botha *et al.*, 2015) at both XFEL and synchrotron sources can also be combined with inline pre-illumination of photoresponsive proteins, where microcrystals of either soluble or membrane proteins are embedded in high-viscosity media such as lipidic cubic phase (LCP) (Weierstall *et al.*, 2014), grease (Sugahara *et al.*, 2015), vaseline (Botha *et al.*, 2015), agarose (Conrad *et al.*, 2015), hyaluronic acid (Sugahara *et al.*, 2015), sodium carboxymethyl cellulose or the thermo-reversible block copolymer Pluronic F-127 (Kovácsová *et al.*, 2017). Prior to high-viscosity injection, the photostability and kinetics in the viscous medium need to be carefully addressed since the occurrence of alterations has been reported (Colletier *et al.*, 2016). Likewise, absorbance or scattering of the illumination light by the viscous medium needs to be considered in the final pre-illumination protocol. The pre-illumination device described in this study could be readily attached to a high-viscosity extrusion injector (Botha *et al.*, 2015), modified with a longer version of the stainless steel tubing holding the borosilicate glass capillary and internal sample capillary (in which case, the SMA light fibre terminates on the glass capillary rather than the sample capillary).

So far, few pre-illumination schemes have been described that result in efficient formation of a semi-stable state that is subsequently analysed by a time-resolved pump–probe scheme. Nango *et al.* (2016) illuminated their bacteriorhodopsin (bR) microcrystal sample prior to SFX data collection by a tungsten lamp [12 V, 100 W halogen lamp in conjunction with a Thorlabs FGL515M coloured glass filter (to cut light with wavelengths shorter than 515 nm)] for 5 min to light adapt the protein to induce full occupancy of the all-*trans* retinal. Similarly, Nogly *et al.* (2016) illuminated bR microcrystals for 1 min using intense microscope light with a yellow >515 nm long-pass filter before loading it into the LCP injector. The efficiency of neither of these approaches was reported. It was also not mentioned whether the lifetime of the light-adapted state exceeded the time the sample spent in the LCP injector before being measured. Clearly, our approach can circumvent these issues.

The main difference between our pre-illumination approach described here and previously described pump–probe illumination setups is in the lifetime of the intermediates. Pre-illumination induces a semi-stable state that is then either analysed as is or used as a starting point for time-resolved SFX experiments. To this end, a number of setups have been described for pump–probe experiments at XFEL facilities, in which laser light has been directly focused within the beamline vacuum (or helium) chamber either onto the glass capillary of a GDVN (Kupitz *et al.*, 2014) or onto an electrospin liquid injector (Sierra *et al.*, 2012; Kern *et al.*, 2014) or onto the free sample jet (Aquila *et al.*, 2012; Kern *et al.*, 2013; Barends *et al.*, 2015; Pande *et al.*, 2016; Nango *et al.*, 2016). These setups have been generally used for short (around 1 s or less) to very short (few femtosecond, picosecond, nanosecond or microsecond) delay times between the optical pump and X-ray probe. All of these experiments utilize laser light focused onto the sample inside the vacuum or helium sample chamber. However, the alignment and characterization of this illumination can be (very) complicated and (very) time consuming, since the chamber must be opened, the X-ray interaction zone must be accessible to diagnostic equipment, and the coincidence of the sample jet and optical pump must be verified after each nozzle change *etc.* Thus, while pre-illumination could in principle also take place inside the vacuum/helium chamber, the advantages of pre-illumination taking place outside and within sample-delivery tubing, as described here, are the robust alignment of the sample and optical probe within the tee connector, the safety shielding of the laser light by the tee connector, and the simple and rapid collection of a small sample volume for spectroscopic characterization with minimal interruption of the SFX experiment. Regular spectroscopic characterization turned out to be critical in the case of TR-SFX on rsEGFP2 (Coquelle *et al.*, 2017).

The pre-illumination device described here can be easily employed to study light-adapted states of other photosensitive proteins by serial crystallography. Possible candidates include other photoconvertible fluorescent proteins such as Dronpa (Ando *et al.*, 2004), light-oxygen-voltage sensory domains (Christie *et al.*, 1999), phytochromes, cyanobacteriochromes, cryptochromes, (bacterio)rhodopsin, photosystems *etc.* These proteins are characterized by light-induced stationary states with dark reversion half-times ranging from tens of seconds to tens of minutes or even hours (Tischer & Weiner, 2014, and references therein). Dark reversion of light-adapted bacteriorhodopsin, for example, proceeds with a relaxation time of about 30 min (Hofrichter *et al.*, 1989). Photosystem II undergoes a reaction cycle (Kok cycle) involving five states (S_0–S_4), the progression of which requires light illumination at each step. S_0 and S_1 are basically dark stable, whereas S_2 and S_3 dark revert in seconds and minutes, respectively, back to S_1 (Dau & Haumann, 2008). In addition to photoconverting endogenous light-responsive proteins, our pre-illumination device can be used for optical control of proteins engineered with photoswitches such as azobenzene (Broichhagen *et al.*, 2015).

Acknowledgements

We thank the Heidelberg FEL group for stimulating and critical discussions and the MPI machine shop for their help. We also thank the scientific staff and support groups at LCLS and SACLA for their support during our beamtimes at their respective facilities. JW acknowledges support from the CEA through the IRTELIS PhD programme.

Funding information

Funding for this research was provided by: Max-Planck-Gesellschaft; Centre National de la Recherche Scientifique (award No. PEPS SASLELX); Agence Nationale de la Recherche (award No. ANR-15-CE32-0004BioXFEL).

References

Ando, R., Mizuno, H. & Miyawaki, A. (2004). *Science*, **306**, 1370–1373.
Aquila, A. *et al.* (2012). *Opt. Express*, **20**, 2706–2716.
Barends, T. R. *et al.* (2015). *Science*, **350**, 445–450.
Botha, S., Nass, K., Barends, T. R. M., Kabsch, W., Latz, B., Dworkowski, F., Foucar, L., Panepucci, E., Wang, M., Shoeman, R. L., Schlichting, I. & Doak, R. B. (2015). *Acta Cryst.* D**71**, 387–397.
Boutet, S. *et al.* (2012). *Science*, **337**, 362–364.
Broichhagen, J., Frank, J. A. & Trauner, D. (2015). *Acc. Chem. Res.* **48**, 1947–1960.
Chapman, H. N. *et al.* (2011). *Nature*, **470**, 73–77.
Christie, J., Salomon, M., Nozue, K., Wada, M. & Briggs, W. R. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 8779–8783.
Colletier, J. P. *et al.* (2016). *J. Phys. Chem. Lett.* **7**, 882–887.
Conrad, C. E. *et al.* (2015). *IUCrJ*, **2**, 421–430.
Coquelle, N. *et al.* (2017). *Nat. Chem.* In the press.
Dau, H. & Haumann, M. (2008). *Coord. Chem. Rev.* **252**, 273–295.
El Khatib, M., Martins, A., Bourgeois, D., Colletier, J. P. & Adam, V. (2016). *Sci. Rep.* **6**, 18459.
Emma, P. *et al.* (2010). *Nat. Photon.* **4**, 641–647.
Grotjohann, T., Testa, I., Reuss, M., Brakemann, T., Eggeling, C., Hell, S. W. & Jakobs, S. (2012). *eLife*, **1**, e00248.
Hirata, K. *et al.* (2014). *Nat. Methods*, **11**, 734–736.
Hofmann, M., Eggeling, C., Jakobs, S. & Hell, S. W. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 17565–17569.
Hofrichter, J., Henry, E. R. & Lozier, R. H. (1989). *Biophys. J.* **56**, 693–706.
Ishikawa, T. *et al.* (2012). *Nat. Photon.* **6**, 540–544.
Kang, Y. *et al.* (2015). *Nature*, **523**, 561–567.
Kern, J., Alonso-Mori, R. *et al.* (2013). *Science*, **340**, 491–495.
Kern, J., Tran, R. *et al.* (2014). *Nat. Commun.* **5**, 4371.
Kovácsová, G., Grünbein, M. L., Kloos, M., Barends, T. R. M., Schlesinger, R., Heberle, J., Kabsch, W., Shoeman, R. L., Doak, R. B. & Schlichting, I. (2017). *IUCrJ*, **4**, https://doi.org/10.1107/S2052252517005140.
Kupitz, C. *et al.* (2014). *Nature*, **513**, 261–265.
Lomb, L., Steinbrener, J., Bari, S., Beisel, D., Berndt, D., Kieser, C., Lukat, M., Neef, N. & Shoeman, R. L. (2012). *J. Appl. Cryst.* **45**, 674–678.
Moffat, K., Szebenyi, D. & Bilderback, D. (1984). *Science*, **223**, 1423–1425.
Nango, E. *et al.* (2016). *Science*, **354**, 1552–1557.
Nogly, P. *et al.* (2016). *Nat. Commun.* **7**, 12314.
Pande, K. *et al.* (2016). *Science*, **352**, 725–729.
Rockwell, N. C., Su, Y. S. & Lagarias, J. C. (2006). *Annu. Rev. Plant Biol.* **57**, 837–858.
Sawaya, M. R. *et al.* (2014). *Proc. Natl Acad. Sci. USA*, **111**, 12769–12774.
Schlichting, I. (2015). *IUCrJ*, **2**, 246–255.
Schlichting, I., Almo, S. C., Rapp, G., Wilson, K., Petratos, K., Lentfer, A., Wittinghofer, A., Kabsch, W., Pai, E. F., Petsko, G. A. & Goody, R. S. (1990). *Nature*, **345**, 309–315.
Sierra, R. G. *et al.* (2012). *Acta Cryst.* D**68**, 1584–1587.
Srajer, V., Teng, T., Ursby, T., Pradervand, C., Ren, Z., Adachi, S., Schildkamp, W., Bourgeois, D., Wulff, M. & Moffat, K. (1996). *Science*, **274**, 1726–1729.
Sugahara, M. *et al.* (2015). *Nat. Methods*, **12**, 61–63.
Tischer, D. & Weiner, O. D. (2014). *Nat. Rev. Mol. Cell Biol.* **15**, 551–558.
Vodyanoy, V., Karvaly, B. & Lanyi, J. K. (1985). *Photochem. Photobiol.* **42**, 413–421.
Weierstall, U. *et al.* (2014). *Nat. Commun.* **5**, 3309.
Zhang, H. *et al.* (2015). *Cell*, **161**, 833–844.

Appendix E : Slow conformational exchange and overall rocking motion in protein crystals

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-017-00165-8 OPEN

Slow conformational exchange and overall rocking motion in ubiquitin protein crystals

Vilius Kurauskas[1,2,3], Sergei A. Izmailov[4], Olga N. Rogacheva[4], Audrey Hessel[1,2,3], Isabel Ayala[1,2,3], Joyce Woodhouse[1,2,3], Anastasya Shilova[5], Yi Xue[6], Tairan Yuwen[6], Nicolas Coquelle[1,2,3], Jacques-Philippe Colletier[1,2,3], Nikolai R. Skrynnikov[4,6] & Paul Schanda[1,2,3]

Proteins perform their functions in solution but their structures are most frequently studied inside crystals. Here we probe how the crystal packing alters microsecond dynamics, using solid-state NMR measurements and multi-microsecond MD simulations of different crystal forms of ubiquitin. In particular, near-rotary-resonance relaxation dispersion (NERRD) experiments probe angular backbone motion, while Bloch–McConnell relaxation dispersion data report on fluctuations of the local electronic environment. These experiments and simulations reveal that the packing of the protein can significantly alter the thermodynamics and kinetics of local conformational exchange. Moreover, we report small-amplitude reorientational motion of protein molecules in the crystal lattice with an ~3–5° amplitude on a tens-of-microseconds time scale in one of the crystals, but not in others. An intriguing possibility arises that overall motion is to some extent coupled to local dynamics. Our study highlights the importance of considering the packing when analyzing dynamics of crystalline proteins.

---

[1] Université Grenoble Alpes, 71 avenue des martyrs, F-38044 Grenoble, France. [2] CEA, Institut de Biologie Structurale, 71 avenue des martyrs, F-38044 Grenoble, France. [3] CNRS, Institut de Biologie Structurale, 71 avenue des martyrs, F-38044 Grenoble, France. [4] Laboratory of Biomolecular NMR, St. Petersburg State University, St. Petersburg, 199034, Russia. [5] European Synchrotron Radiation Facility, 71 avenue des martyrs, F-38044 Grenoble, France. [6] Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, IN 47907-2084, USA. Vilius Kurauskas, Sergei A. Izmailov, and Olga N. Rogacheva authors contributed equally to this work. Correspondence and requests for materials should be addressed to N.R.S. (email: nikolai@purdue.edu) or to P.S. (email: paul.schanda@ibs.fr)

Proteins perform their functions in aqueous solution at ambient temperatures by sampling a multitude of conformational states involved, for example, in enzymatic turnover or binding to different partners. Understanding protein function requires characterizing the three-dimensional structures, ideally of all these thermally accessible conformations, and deciphering how the different conformations interconvert. The structural basis for understanding protein function have been laid primarily through X-ray diffraction (XRD) of proteins embedded in crystals. Recent methodological advances in XRD aim at characterizing also the structural heterogeneity by fitting multiple conformers—rather than a single structure—into the electron density obtained from Bragg diffraction peaks, or by additionally making use of diffuse scattering intensity[1–6]. The direct structural insight into protein dynamics that one may obtain from such ensemble crystallographic approaches may shed light on molecular mechanisms of functional processes; in particular, when combined with techniques that provide access to kinetics and thermodynamics, such as solution-state NMR spectroscopy[7], one may obtain a fairly complete spatio-temporal picture, which may ultimately lead to a more detailed understanding of protein function.

The densely packed environment of crystals differs substantially from solution, raising the question how packing may alter the properties of proteins. Structures obtained from crystalline proteins are generally in excellent agreement with state-of-the-art solution-NMR structures, and are generally as good in predicting solution-NMR structural parameters as the NMR structures themselves[8, 9]. However, whether dynamics, i.e., the ensemble of co-existing conformers and their pairwise interconversion rates, are also conserved in crystals is less well studied. Evidence that the dynamics required for function are retained in crystals, at least in some cases, comes from the observation of functional activity, such as ligand binding and enzymatic catalysis, in crystals[10–14]. Direct experimental investigations of the dynamics of proteins in crystals at the atomic level have turned out to be difficult. Early studies focused on comparing XRD temperature factors (B-factors) with solution-state NMR order parameters (S^2)[15–17]. Such comparisons only had limited success and remained ambiguous, because of the very different nature and information content of these observables (see ref. [18] and references therein). Recent developments in magic-angle spinning solid-state NMR (MAS ssNMR) of proteins have allowed a more direct comparison of NMR parameters in solution and crystals[19–21]. Furthermore, comparative molecular dynamics (MD) simulations of proteins embedded in an explicit crystal lattice and in solution also produced useful insights, at least over the time scales accessible to MD (typically sub-microsecond)[22–25]. The picture arising from these NMR and MD studies suggests that the crystal lattice has indeed only little impact on dynamics on sub-microsecond time scales: changes are primarily found in loops engaged in crystal contacts, whereas secondary structure elements and loops remote from neighboring molecules showed similar order parameters. Many biological processes, however, occur on longer time scales, often micro-to-milliseconds (μs–ms), and they may involve rare excursions to low-populated "minor states". In many cases, such rare conformations seem to be crucial for functional activity[26], and they may involve larger-scale rearrangements affecting a bigger portion of the structure. Whether such slow dynamic processes are sustained in crystals, and how the packing alters these motions is poorly understood. MD simulation methods only recently attained the ability to model those time scales, and experimental NMR methods likewise became available only recently.

Here we use a combination of ssNMR experimental methods with microsecond-long MD simulations of explicit crystal lattices, in order to investigate how the crystalline environment impacts slow dynamics of the 8.6 kDa regulatory protein ubiquitin. The dynamics of this protein in solution have been extensively characterized experimentally and computationally. In particular, microsecond motion has been shown for the loop connecting ubiquitin's β4 strand and the 3_{10} helix, comprising residues E51 to R54, which exists in two distinct states that differ by a flip of the D52/G53 peptide plane, along with an alteration of hydrogen bonds within this β-turn and to the adjacent helix[27–32]. These two states correspond to so-called type-I and type-II β-turns, respectively, herein referred to as βI and βII states. The dynamic exchange process between these two conformations has been proposed to result in small (within 1 Å) but discernible reshaping of the entire molecule[30], and perturbation of the β-turn equilibrium has been shown to allosterically alter the binding of proteins to a surface located on the opposite side of the ubiquitin molecule[30]. This dynamic process is thus thought to be of considerable functional relevance for ubiquitin. Different crystal structures are available that are in either of these two states, and millisecond-long MD simulations in solution have detected the corresponding microsecond exchange process[32]. Previous ssNMR data revealed that this μs dynamics is present in one of ubiquitin's crystal forms[33, 34].

In addition to this internal motional process, we have recently provided evidence that ubiquitin molecules undergo overall "rocking" motion in the crystal[35]. This motion was detected by a comparative ssNMR $R_{1\rho}$ relaxation measurement of different crystal forms, which revealed that in one of the crystal forms, termed cubic-PEG-ub, elevated relaxation rates throughout the molecule can be explained by an overall reorientational motion with an amplitude of several degrees. The presence of this motion was also confirmed by MD simulations. It was furthermore shown that the observed rocking has direct implications for X-ray crystallography: the crystal form that features rocking motion gives rise to low-resolution diffraction data and high Wilson B-factors in crystallographic experiments at 100 K. However, our previous study did not provide a precise estimate of the rocking motion time scale, and only a range from ~100 ns to 100 μs was indicated[35].

Here we investigate in parallel two different crystal forms, thus avoiding any bias that would arise from comparing dynamics measured by different methods (i.e., solution and solid-state methods). We use Bloch–McConnell-type $R_{1\rho}$ relaxation dispersion experiments, which were recently introduced to solid-state NMR[34], to probe on a per-residue basis the dynamics-induced modulation of the chemical shift of ^{15}N nuclei. In addition, we use $R_{1\rho}$ experiments at higher radio-frequency field strengths, near the rotary-resonance condition $\omega_{\mathrm{RF}}^{\mathrm{eff}} = \omega_{\mathrm{MAS}}$[34, 36, 37], which we refer to here as near-rotary-resonance relaxation dispersion (NERRD) experiment. This experiment is sensitive to the angular fluctuations of the H–N bond vectors. A systematic analysis of experimental NERRD data against numerical simulations provides information about the angular motional amplitude of μs motions.

The combined experimental data establish that crystal contacts indeed alter the kinetics (exchange rates) and thermodynamics (relative populations of states) of μs motions. We find that the relative populations of the exchanging states, βI and βII, can be inverted in going from one crystal to another, and that in both analyzed crystal forms the exchange rates are slowed down significantly with respect to solution. Multi-μs long MD trajectories of WT and mutant proteins confirm these experimental findings, and identify the molecular origins of these altered dynamics. Using NERRD data, we establish that overall rocking occurs on the tens-of-microseconds time scale, with an amplitude of about 3–5°. The crystal form that is amenable to rocking also features

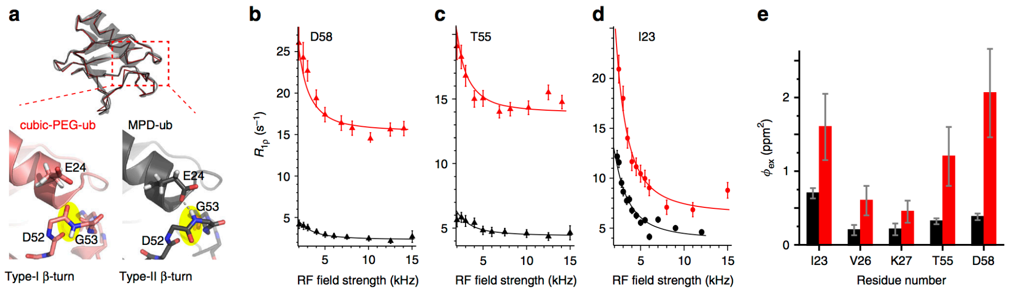


Fig. 1 Evidence for β-turn exchange in MPD-ub (*black*) and cubic-PEG-ub (*red*) crystals. **a** Structures of ubiquitin in the two crystal forms. *Top*: superposition of the main chains in the two crystals, shown as C^{α} backbone traces; for cubic-PEG-ub only one of the two non-equivalent chains is shown. For clarity, the *semi-transparent* cartoon representation is shown only for MPD-ub. *Bottom*: zoom into the β-turn region. The peptide bond undergoing a flip is highlighted in yellow. **b**–**d** BMCRD profiles of ^{15}N sites in the two ubiquitin crystal forms, obtained at a sample temperature of 300 K. Data shown here for D58 and T55 were obtained at B_0 field strengths corresponding to ^{1}H Larmor frequencies of 600 MHz, and those of I23 at 950 MHz, respectively. Error bars of $R_{1\rho}$ values were obtained by standard Monte Carlo analyses, described in the Methods section. *Solid lines* correspond to a global two-site exchange model fit of data from residues I23, V26, K27, T55, D58 within each of the crystal forms (*solid lines*). The fits were performed separately for each of the two crystal forms using in each case the same cluster of the aforementioned five residues. Note that the resonances of E24 and N25 are not visible, presumably as a consequence of the conformational exchange. The comprehensive set of BMCRD data and fitting results can be found in Supplementary Figs. 2, 3, and 5. **e** ϕ_{ex} values, $\phi_{ex} = p_A p_B (2\pi\Delta\delta)^2$ in cubic-PEG-ub (*red*) and MPD-ub (*black*), obtained from the joint fit of residues I23, V26, K27, T55, and D58, exemplified by curves in **b**–**d**. Here p_A and p_B are the populations of the two conformers, while $\Delta\delta$ is the chemical shift difference expressed in the units of ppm. Fitted exchange time constants are reported in Table 1

extensive βI/βII exchange. Furthermore, both rocking and exchange dynamics occur on a similar time scale, suggesting some degree of cooperativity between local and overall motions. Although our attempts to capture this effect by means of the rationally designed mutations proved unsuccessful, we believe that this scenario remains a possibility in this and other protein crystals. Our results highlight the importance of crystal packing effects in the context of crystallography-based investigation of protein dynamics.

Results

Theory. In order to probe μs dynamics we employ ssNMR ^{15}N $R_{1\rho}$ relaxation dispersion (RD) experiments. RD experiments measure the transverse magnetization decay rate constant, $R_{1\rho}$, under an applied radio-frequency field for each of the ^{15}N backbone spins across the protein, and investigate how this decay rate constant depends on the field-strength, ν_{RF}, of the applied RF field. In MAS ssNMR, two different physical mechanisms can give rise to (non-flat) RD profiles:

(i) If the exchanging states differ in their local electronic environment around a given ^{15}N nucleus, and if the exchange process occurs on a μs–ms time scale, the corresponding chemical-shift modulation contributes to the $R_{1\rho}$ decay rate constant. This enhancement of $R_{1\rho}$ can be scaled down by applying a sufficiently strong RF field, i.e., $R_{1\rho}$ decreases with increasing ν_{RF}. This effect, known as "Bloch–McConnell RD" in solution-state NMR[38], depends on the chemical shift difference between the exchanging sites, as well as the time scale of the exchange, the relative populations of the exchanging states and the amplitude of the RF field.

(ii) Fluctuation of the dipolar and CSA interactions relative to the reference frame of the crystallite also induces enhanced $R_{1\rho}$ relaxation, in addition to the above effect that fluctuation of the isotropic chemical shift has[34, 37]. If the relevant motions are in the μs–ms range then this effect can be scaled down not only by application of the RF field, but also by magic angle sample spinning. Mathematically this is expressed by the formula derived by Kurbanov et al.[39] which (in the case of sufficiently strong RF field surpassing the RF carrier offset $\omega_{RF}^{eff} \gg \Omega$) can be rendered in the following familiar form:

$$R_{1\rho} = \frac{1}{8} D_{NH}^2 [4j_0 + J(\omega_N - \omega_H) + 3J(\omega_N) + 6J(\omega_H) \\ +6J(\omega_N + \omega_H)] + \frac{1}{18} C_N^2 [4j_0 + 3J(\omega_N)] \quad (1)$$

$$j_0 = \frac{1}{6} [J(\omega_{RF} - 2\omega_{MAS}) + 2J(\omega_{RF} - \omega_{MAS}) \\ +2J(\omega_{RF} + \omega_{MAS}) + J(\omega_{RF} + 2\omega_{MAS})]. \quad (2)$$

Here $D_{NH} = -(\mu_0/4\pi)\left(\hbar\gamma_N\gamma_H/r_{NH}^3\right)$ is the dipolar coupling strength, $C_N = \omega_N\left(\sigma_{\perp}^N - \sigma_{||}^N\right)$ is the anisotropy of the chemical-shift tensor, $\omega_{RF} = 2\pi\nu_{RF}$ is the spin-lock field strength, and $\omega_{MAS} = 2\pi\nu_{MAS}$. The explicit form of spectral densities $J(\omega)$ depends on the adopted model of motion; the empirical Lipari–Szabo-style expression is $J(\omega) = \frac{2}{5}(1 - S^2)\tau_c/\left(1 + \omega^2\tau_c^2\right)$, where S^2 is the generalized order parameter, expressing the motional amplitude of the bond vector, and τ_c is the correlation time of the motion.

Of particular importance in Eqs. (1) and (2) are spectral densities $J(\omega_{RF} \pm n\omega_{MAS})(n = 1, 2)$. The meaning of these terms is as follows: if the motional process is sufficiently slow, $\omega_{RF}\tau_c \geq 1$ or $\omega_{MAS}\tau_c \geq 1$, then relaxation can be partially suppressed due to refocusing of the dipolar interaction by RF field or MAS, respectively. However, it may also happen that refocusing due to RF field and due to MAS in part cancel each other out (reminiscent of dipolar recoupling). This phenomenon occurs near rotary-resonance conditions, $\omega_{RF} = n\omega_{MAS}$ or, equivalently $\nu_{RF} = n\nu_{MAS}$ $(n = 1, 2)$, and is accompanied by increase in the $R_{1\rho}$ decay rate constant. Such behavior (illustrated by means of numerical simulations, Supplementary Fig. 1) is the basis of NERRD experiment. We note NERRD-type experiments have been used in previous studies[40–44].

The observation of an increase of $R_{1\rho}$ when approaching the rotary-resonance condition is an unambiguous signature of µs motions. Assuming that dynamics can be modeled as an exchange process, the exact shape of such NERRD profiles depends on the time scale of motion, the relative populations of the involved states, as well as the difference in orientation of the NH bond vector between the exchanging states[36, 37]. NERRD thus provides insight into the angular amplitude of the motion, and complements Bloch–McConnell RD data, which probe motions by their isotropic chemical-shift modulation. Importantly, in solution-state NMR this type of information is unavailable, because the rapid (ns) molecular tumbling averages the anisotropic interactions such that they are not informative of slow dynamics.

Both of the above relaxation-dispersion mechanisms, (i) and (ii), simultaneously lead to RF-field-dependent $R_{1\rho}$ decay rate profiles whenever µs–ms motions are present. However, experimental parameters can be chosen such as to separate the two effects. Provided that a fairly large sample-spinning frequency ν_{MAS} is used, the $R_{1\rho}$ RD profiles in the low RF-field range, typically $\nu_{\mathrm{RF}} \leq 10$ kHz, are dominated by isotropic chemical-shift mechanism. For simplicity, we employ the term "Bloch–McConnell relaxation dispersion" (BMCRD) for this situation. On the other hand, in the vicinity of the rotary-resonance conditions, which in our NERRD experiments is at $\nu_{\mathrm{RF}} = \nu_{\mathrm{MAS}} = 44$ kHz and thus well separated from the BMCRD regime, it is the angular fluctuation of the NH bond that dominates the RD profiles[34, 37]. For simplicity, we will reserve the term NERRD for the situation where relaxation due to modulation of isotropic chemical shift is fully suppressed by the strong RF field and the near-rotary-resonance dispersion effect arises solely from the anisotropic interactions.

We note, however, that the two described relaxation dispersion phenomena (BMCRD, NERRD) are not fundamentally different but rather reflect two mechanisms of line broadening induced by µs motion. The two respective data sets can be fitted jointly or independently, and such fits can in ideal cases provide exchange rates, population levels of involved states, chemical-shift differences and jump angles. In systems where the exchange process is fast (less than ~100 µs), which is the case here, it is not possible to disentangle populations and chemical-shift changes, and only their product can be obtained; likewise, populations and jump angles also become entangled (see further discussion below and Supplementary Fig. 4).

Crystal contacts slow down type-I/type-II β-turn exchange. In this study, we investigate dynamics in two different crystal forms of ubiquitin that are obtained in the presence of two different precipitation agents (2-methyl-2,4-pentanediol (MPD) and polyethylene glycol (PEG), see Methods) and are referred to as MPD-ub and cubic-PEG-ub, respectively. XRD-based crystal structures of these different forms have been previously solved at cryo-temperatures (100 K)[35, 45, 46]. While these X-ray data reveal that ubiquitin adopts essentially the same overall structure (backbone RMSD of 0.5 Å), they differ in the conformation of the loop region encompassing residues E51 to R54: in MPD-ub this region forms a so-called type-II β-turn, while it is in βI conformation in cubic-PEG-ub (Fig. 1a). Note that the predominant state in solution is βI[35, 45–47] with an estimated >95% population, in exchange with an alternate state which is believed to be βII[30].

Figure 1b–d shows Bloch–McConnell RD data for three representative amide sites in MPD-ub and cubic-PEG-ub. The non-flat RD profiles for residues located in the β-turn region and the adjacent N-terminal part of the α-helix unambiguously show that this part of the molecule undergoes µs dynamics in both crystals. A joint fit of a two-state exchange model to data from the five residues located in this region is shown as solid lines (see Methods). The use of a two-state exchange model appears justified in the light of the two distinct states observed in different crystal structures and MD simulations (ref. [32] and MD data below). The exchange time constant obtained from these fits is $\tau_{\mathrm{ex}} = 90 \pm 25$ µs for MPD-ub and 125 ± 45 µs for cubic-PEG-ub. We want to stress the fact that the absolute "plateau" levels of these curves are significantly higher in cubic-PEG-ub than in MPD-ub throughout the molecule, as can be readily seen in the three examples in Fig. 1b–d. While the plateau is not relevant for fits of BMCRD data, this finding of elevated $R_{1\rho}$ rates points to the presence of overall angular motion of ubiquitin molecules in cubic-PEG-ub, as we will outline below.

In order to gain insight into the angular motions of the NH bond vectors, we measured NERRD profiles. Figure 2 shows the largest NERRD effects observed in MPD-ub, which are found for the amide sites of D52 and R54, which are adjacent to the peptide plane D52/G53 that undergoes a ca. 180° flip when exchanging between βI and βII conformations. (The amide signal of G53 is not resolved in the spectra of MPD-ub.) Other residues do not show significantly rising NERRD profiles, as exemplified by the data of three residues depicted in black in Fig. 3. An independent experiment at a different MAS frequency (20 vs. 44 kHz in the current experiment) also exhibited the largest NERRD effect for

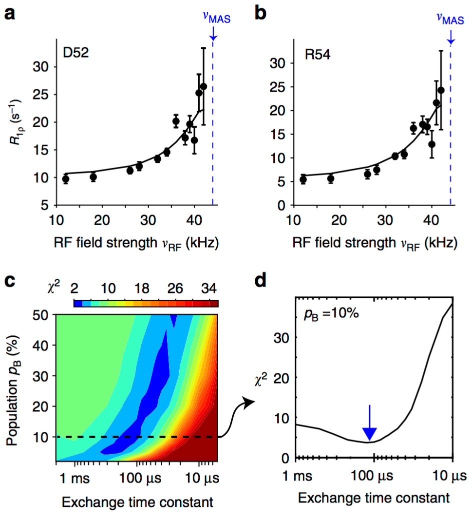


Fig. 2 Angular motion in the β-turn of MPD-ub crystals probed by NERRD NMR. **a**, **b** NERRD dispersion data of the two amide sites in MPD-ub that are adjacent to the flipping peptide plane D52/G53, recorded at 44.053 kHz MAS frequency. Error bars of $R_{1\rho}$ values were obtained by standard Monte Carlo analyses, described in the Methods section. *Solid lines* show fit curves from a two-site jump model that assumes that the minor conformation has a population of $p_B = 10\%$. **c** χ^2 surface of a grid-search of simulated NERRD curves against the experimental data in **a**, **b**. Hereby, a common minor-state population p_B and exchange correlation time $\tau_{\mathrm{ex}} = 1/(k_{AB} + k_{BA})$ were assumed for the two residues, and the jump angle was separately fitted for each residue. More details of this NERRD fit are reported in Supplementary Fig. 4. **d** Cross-section across the χ^2 surface along the τ_{ex} dimension corresponding to a minor-state population of 10%. This minor-state population level is suggested by analyses of BMCRD data, as explained in the text. The best-fit exchange time constant with this assumption is $\tau_{\mathrm{ex}} \sim 100$ µs

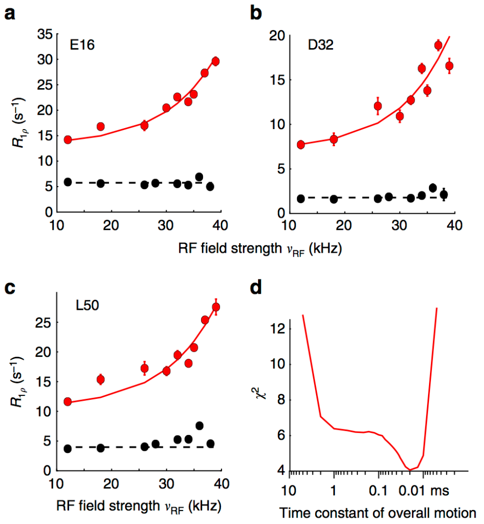


Fig. 3 Evidence for rocking motion in cubic-PEG-ub crystals from NERRD NMR measurements. **a–c** NERRD data of representative amide sites outside the dynamic β-turn region for MPD-ub (*black*) and cubic-PEG-ub (*red*). Error bars were obtained from Monte Carlo analysis, described in the Methods section. The *solid red lines* correspond to a two-state fit of data from 22 amide sites with a common exchange time constant and per-residue-adjustable motional amplitude (all shown in Supplementary Fig. 6). **d** χ^2 value of this fit as a function of the jump time constant. In fitting these NERRD data, we have assumed that protein rocking can be approximated as two-site jump process, which is a crude but justifiable model[70]

residues D52 and R54 in MPD-ub[35], whereas all other residues showed either small NERRD profiles (residues I23, E51, T55) or no significant NERRD effect, mirroring the present data.

To quantitatively rationalize these data, we used numerical spin-dynamics simulations to obtain theoretical NERRD profiles over a grid of exchange parameters, p_B and τ_{ex}, and fitted the experimental data against this grid of simulated curves. *Solid lines* in Fig 2a, b show the best-fit curves for such a joint fit of data from D52 and R54, assuming that they have one common excited-state population and exchange time constant. Figure 2c shows the corresponding plot of the χ^2 surface. Although the population levels obtained from these fits have significant uncertainty, the time scale converges to ~50–200 µs. This is in excellent agreement with the independent Bloch–McConnell RD data, as illustrated in Fig. 1b–d. In addition to confirming the time scale of the process, NERRD data provide information about the amplitude of the NH motion of D52 and R54, which is found to be on the order of 7–11°, depending on the assumed populations (Supplementary Fig. 4). Interestingly, this angle is similar to the difference between the orientations that these two bonds assume in βI and βII states, which is 12 and 14° for the two respective residues (comparing PDB entries 1UBQ and 3ONS).

Taken together, both Bloch–McConnell RD and NERRD data show that the β-turn dynamics in MPD-ub occurs on a time scale of ca. 100 µs; this exchange process causes chemical shift modulation in the β-turn region (Fig. 1) and angular jumps of the amides of D52 and R54, as revealed by NERRD (Fig. 2).

It is interesting to note that the two crystals, which have very different packing environment around the exchanging region, have similar exchange kinetics, but that these kinetics are rather different to the case of solution, where the exchange process has a time constant of $\tau_{ex} = 6 \pm 2$ µs at the same temperature, i.e., is about 15 times faster[30]. The significant slowdown of the exchange process in the crystals is not due to the increased viscosity induced by the precipitation agent: it has been shown[33] that even in the presence of MPD concentrations as high as 45% in solution all CPMG RD profiles were flat, i.e., the underlying process was too fast to induce RD, whereas in MPD-ub crystals very large CPMG RDs were detected, thus confirming that MPD is not the principal cause for slowing of the exchange dynamics[33]. As noted above, the "ground state" is different between the two crystals, being βI in cubic-PEG-ub (as well as in solution), and βII in MPD-ub. Our findings show that the nature of the ground state does not dictate the kinetics, because in solution and cubic-PEG-ub, which have both βI as major state, the kinetics are very different.

βI/βII populations differ in the different crystals. Inspection of the BMCRD data in Fig. 1 reveals a noticeable difference between the two crystals: the amplitude of these RD profiles for residues in cubic-PEG-ub is considerably larger than those of their counterparts in MPD-ub, on average by a factor 3.3 ± 1.0 with regard to the values of ϕ_{ex} in Fig. 1e.

While precise determination of populations is beyond reach in the case of fast exchange, semi-quantitative estimates are nevertheless possible. ϕ_{ex} values in MPD-ub are consistent with $p_{\beta I} \approx 5$–10% and $\Delta\delta$ on the order of a fraction of ppm, as can be expected for the analyzed residues that are located at the periphery of the β-turn. This $p_{\beta I}$ estimate is supported by our MD simulations (see below). It is further reasonable to assume that $\Delta\delta$ values are approximately the same for the MPD-ub and cubic-PEG-ub experiencing the same exchange process (interconversion between βI and βII). If we make this assumption, then based on ϕ_{ex} results we can estimate that the populations of minor species in the cubic-PEG-ub crystal is $p_{\beta II} \approx 20$–50%. Note that the population of minor species cannot exceed 50%, which provides in this context a useful boundary. Supplementary Table 1 provides some example calculations relating the minor-state populations in the two crystals.

MD simulations using explicit crystal lattices have previously been shown to successfully reproduce experimental ssNMR and XRD data, and they may contribute valuable information about mechanistic details of motion that are difficult to access experimentally[6, 35, 48–50]. We simulated an explicit arrangement of ubiquitin molecules according to the crystal lattices of MPD-ub and cubic-PEG-ub, built of 24 and 48 ubiquitin molecules, respectively, in the presence of interstitial water over a time interval of 2 µs. The explicit simulation of a large number of ubiquitin molecules improves the sampling statistics considerably. In addition, we performed a 10 µs simulation of ubiquitin in solution, which we find to be in excellent agreement with a recent 1 ms trajectory[32]. In all these trajectories, we observed multiple βI ↔ βII transitions, enabling us to extract the relative populations of the two conformations, which are reported in Table 1. The minor-state populations are qualitatively in good agreement with the estimates derived from BMCRD data, i.e., simulated MPD-ub has significantly lower minor-state population than cubic-PEG-ub (chain B), mirroring the lower ϕ_{ex} values. It should be noted that in the ssNMR spectra of cubic-PEG-ub the signals from the two inequivalent chains, A and B, overlap in most cases, making it impossible to collect two separate data sets. The experimental data, therefore, reflect an effective average of the dynamics in the two chains. With this caveat the $p_{\beta II}$ populations found in the MD trajectory of cubic-PEG-ub are consistent with the experimental observations.

Table 1 Populations and time scales of conformational exchange dynamics in the region encompassing residues L50–L56 and the adjacent N-terminal part of the α-helix

	MD		NMR (Bloch-McConnell relaxation dispersion data)	
	βII population (%)	Time constant τ_{ex} (μs)	βII population (%)	Time constant τ_{ex} (μs)
Solution	3	2.5	3[a]	6[a]
MPD-ub	90	3.7	~90–95[b]	90 ± 25
Cubic-PEG-ub				
Chain A	9	1.1	~20–50[b]	125 ± 45
Chain B	43	0.4		

The population of the βI state is $p_{\beta I} = 100\% - p_{\beta II}$

[a] Solution-state NMR data are from Smith et al.[30] Note that the exchange rate constant is largely in agreement with other studies, which, however, were only able to provide upper limit estimates, and at lower temperature[28, 29, 31]

[b] The ratio of ϕ_{ex} values from BMCRD fits were used to estimate population levels. See Supplementary Table 1 for details

The time scale of the exchange process in the MD simulations tends to be faster than observed experimentally (Table 1), suggesting that the force field employed in this study does not accurately reflect the transition-state free-energy barrier. Similar findings have been reported before[51]. This observation is perhaps not surprising, given the limited accuracy of MD models (note that force fields have been parameterized primarily against experimental equilibrium data, and not against kinetic data). From the perspective of this study the elevated exchange rates are to some degree advantageous since this improves the statistical sampling of the βI/βII equilibrium.

Thus, NMR and MD techniques independently indicate that the relative populations of βI and βII conformers differ considerably in different crystals and in solution. The βI:βII populations span the wide range from 10:90 (MPD-ub) to 57:43/91:9 (cubic-PEG-ub) and 97:3 (solution).

Molecules in cubic-PEG-ub undergo tens-of-microseconds rocking motion. We also recorded NERRD experiments on cubic-PEG-ub crystals, in order to obtain additional insight into μs dynamics in this crystal form. It shall be reminded that in MPD-ub only few residues (in particular D52 and R54; see Fig. 2a, b) show increased $R_{1\rho}$ close to the rotary resonance condition (i.e., non-flat NERRD profiles), and that the time scale obtained from a fit of these data is in excellent agreement with Bloch–McConnell RD data. In contrast to this finding, essentially all residues in cubic-PEG-ub exhibit non-flat NERRD profiles, as shown in Fig. 3 and Supplementary Fig. 6, and the $R_{1\rho}$ values are generally higher than those observed in MPD-ub. Given that this trend is observed for all residues, we ascribe the underlying motion to a global process involving the entire molecule, such as an overall "rocking" motion of the ubiquitin molecules in the crystal. Overall rocking motion is visible not only in $R_{1\rho}$ measurements, but is also expected to contribute to line widths in ssNMR experiments. This expectation is indeed confirmed by ^{1}H and ^{15}N coherences life times, which differ markedly between the two crystal forms studied here (Supplementary Fig. 7). We have proposed the existence of such overall motions in cubic-PEG-ub recently, based on the observation of elevated $R_{1\rho}$ rate constants in cubic-PEG-ub compared to MPD-ub[35], although the time scale could be estimated only very roughly (ca. 100 ns to 100 μs).

In order to obtain quantitative insight into rocking motion, we have performed a common fit of a dynamic model with a global time scale to NERRD data from 22 well-resolved residues in cubic-PEG-ub, excluding residues in the β-turn region, since for those residues one expects superposition of local and global motion. The best-fit time scale is of the order of tens of microseconds (Fig. 3d). Residue-wise motional amplitudes, expressed in terms of a two-site jump model, are in the range 3–5° (Supplementary Fig. 6). Such a relatively small amplitude is expected for an overall-rocking motion within the confines of the lattice; MD simulations revealed an amplitude of overall motion in cubic-PEG-ub of ~6–8°, although this amplitude is likely exaggerated due to simultaneous small drift (melting) of the crystal lattice[35].

It is interesting to note that the tens-of-microseconds time scale of overall motion, found experimentally, is similar to the time scale of the β-turn conformational exchange in cubic-PEG-ub (125 ± 45 μs). This similarity may be fortuitous, but it might also point to some correlation between the two processes, i.e., the local and the overall motion. Such possibility is discussed in what follows.

Intermolecular contacts that alter conformational equilibria. As noted above, the βI conformation in ubiquitin represents a "default" state, which is found in solution and in the majority of crystal structures. At the same time the βII conformation has also been observed in a number of crystals, such as MPD-ub[46] and complexes of ubiquitin (see references in Sidhu et al.[52]). This leads us to conclude that the crystalline environment can act as a switch converting the protein from βI to βII. Our MD simulations of ubiquitin crystals are well suited to identify intermolecular interactions that are responsible for this transformation.

In our initial analysis of the MPD-ub trajectory we have identified two such putative intermolecular contacts. One is the four-way junction involving a pair of D52 side chains and a pair of K63 side chains, all of which belong to different ubiquitin molecules but come together in the crystal lattice (Fig. 4a). The other is the intermolecular hydrogen bond between the (neutral) carboxyl groups of E64 and E24, which appears to stabilize the coupling between E24 and G53 (Fig. 4b). According to the MD simulation data, both of these interactions show clear preference for the βII over the βI form.

However, additional simulations convinced us that this is the case of correlation rather than causation. Specifically, we have recorded two separate 2-μs-long trajectories of MPD-ub crystals, one of which carried K63A mutation and the other E64A mutation. As it turned out, these trajectories produced exactly the same proportion of βI to βII species as the wild-type crystal simulation. This led us to conclude that the intermolecular contacts involving K63 and E64 side chains are not the key to the β-turn conformation.

At the same time, we have found that an E24A mutation produced a complete reversal of βII to βI ratio compared to the wild-type simulation, from 90%:10% to 9%:91%. In other words, E24A mutation has restored the "default" state of the system, which is dominated by βI. Apparently it is the hydrogen bond between the carboxyl group of E24 and the backbone amide group in G53 which is directly responsible for prevalence of the βII form in the wild-type MPD-ub crystal. But this hydrogen

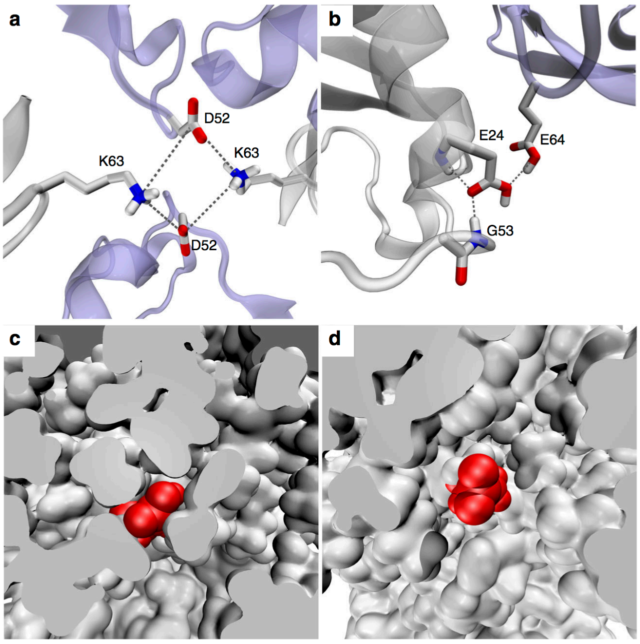


Fig. 4 Intermolecular and intramolecular interactions that affect β-turn motion. **a**, **b** Intermolecular interactions involving the β-turn region in MPD-ub. Neighboring molecules are shown in *light blue*. Simulations of ubiquitin mutants suggest that these interactions are not responsible for causing the preference for βII, see text. **c**, **d** The surrounding of E24 side chain in different crystal forms of ubiquitin. **c** A rare frame from the MPD-ub trajectory where E24 side chain (painted *red*) is projected toward solvent. **d** A typical frame from cubic-PEG-ub trajectory where E24 from ubiquitin molecule chain A is immersed into solvent

bond is an intramolecular hydrogen bond—so how is it related to crystal packing?

Inspection of the trajectories of ubiquitin in solution and in the cubic-PEG-ub crystal revealed that in these simulations the side chain of E24 is mostly immersed in solvent. In solution, E24 side chain is solvated 65% of the entire simulation time. In cubic-PEG-ub, it is solvated 62% of the time for chain A and 52% of the time for chain B. In contrast, it is solvated at the level of only 11% in the MPD-ub simulation (Table 2).

As it turns out, the packing of ubiquitin molecules in the MPD-ub crystal lattice is such that E24 side chain does not have the space to extend outwards. This is illustrated in Fig. 4c, d, which compare the volume of space available to E24 side chain in MPD-ub and in cubic-PEG-ub crystals. In particular, the intrinsically favorable rotameric state ($\chi_1 = 180°$, $\chi_2 = 180°$) that is prominent in other simulations, is completely obstructed in MPD-ub. We conclude that typically the E24 side chain enjoys the entropic benefit of being immersed in solvent, but in MPD-ub it is denied this opportunity and therefore opts for the hydrogen bond with G53.

The propensity of the carboxylic group in E24 to form a hydrogen bond with the G53 amide group is also influenced by its protonation state. The MPD-ub crystal has been obtained from a mother liquor at pH 4.2. Under these conditions the carboxylic group of E24, as positioned in the crystallographic structure, is expected to be mostly protonated[53]. To elucidate the role of this factor, we have altered E24's status from neutral to anionic and then recorded an additional 2-µs-long simulation of the MPD-ub crystal. As it turns out, in this trajectory E24 is partially recruited into both intramolecular and intermolecular salt bridges (Table 2). This happens at the expense of the E24-G53 hydrogen bond, which drops to the level of 28%. Accordingly, the ratio of βII to βI in this trajectory changes from 90%:10% to 71%:27%. This finding supports the hypothesis that the protonation state of E24 also plays a significant role in the stabilization of the βII species in MPD-ub crystal.

Potential coupling between rocking and βI/βII exchange. As discussed above, cubic-PEG-ub crystal is characterized by significant amount of rocking and, at the same time, extensive conformational exchange in the βI/βII region. Furthermore, the characteristic time constants of these two motional processes turn out to be similar. Is this a mere coincidence, or does it point toward a certain common underlying mechanism?

Inspection of the cubic-PEG-ub trajectory led us to identify one potential coupling mechanism. This mechanism involves an ion pair between K11 (chain A) and D52 (chain B). In the crystal structure these two side chains are positioned far apart (Fig. 5a). However, during the MD simulation they frequently form a salt bridge (Fig. 5b). The formation of this salt bridge is associated with (i) slight reorientation of protein molecules in the crystal lattice and (ii) βI → βII transition mediated by residue D52 in the β-turn. Thus the reorientational dynamics of the protein (rocking motion) may be correlated with conformational exchange. This hypothetical scenario is illustrated in the Supplementary Movie 1.

We have yet been able to obtain only partial confirmation for this scenario. The MD data indicate that $\mathrm{K11_{chainA}}$–$\mathrm{D52_{chainB}}$ salt bridge in the cubic-PEG-ub crystal favors βII conformation over

Table 2 Important interactions involving E24 side-chain carboxylic group as characterized by their presence in the MD simulations

E24 side chain contacts	Solution (%)	MPD-ub		Cubic-PEG-ub	
		E24 neutral (%)	E24 charged (%)	Chain A (%)	Chain B (%)
E24-water	65	11	4	62	52
E24-K27(NH_3^+)	34	2	20	26	27
E24-G53(NH)	0.5	71	28	1	2
E24-K63(NH_3^+)	0	1	64	0	0

The contacts involving K27 and G53 are intramolecular, the contact involving K63 is intermolecular. The numbers in each column do not always add up to 100% because E24 also forms certain other contacts (both intermolecular and intramolecular) or, otherwise, some of the listed interactions can occur simultaneously

βI (by a factor of 4.1 considering the trajectory at hand). In order to probe this effect experimentally, we have prepared cubic-PEG-ub crystals of K11A ubiquitin, and performed BMCRD experiments. We saw a small but detectable decrease in RD amplitudes for residues in the β-turn region in the K11A ubiquitin crystal compared to WT crystal (Supplementary Fig. 10). This reduced RD effect may indeed point to a decreased population of minor βII state in K11A.

On the other hand, our attempts to detect the link between K11$_{\mathrm{chainA}}$–D52$_{\mathrm{chainB}}$ and rocking motion proved to be unsuccessful. The MD simulation of K11A crystal failed to produce any significant evidence that on–off K11$_{\mathrm{chainA}}$–D52$_{\mathrm{chainB}}$ salt bridge drives rocking dynamics. In part, this negative result can be attributed to poor convergence of the MD trajectories, which are affected by gradual “melting” of the crystal lattice (discussed by us previously[35]; see Supplementary Fig. 12). One should also bear in mind that ubiquitin molecules in the crystals are held together by a complex network of intermolecular interactions involving several dozen hydrogen bonds and salt bridges (Supplementary Figs. 8 and 9). In this situation the effect from the single-point mutation, K11A, may be masked; moreover, the consequences of the mutation may be more complex and differ from what has been expected. The experimental NERRD data from the sample of K11A ubiquitin also could not detect the expected decrease in rocking within the experimental precision (Supplementary Fig. 11).

Although in this specific case we could not confirm that conformational exchange is coupled to rocking, this hypothesis deserves further investigation. It is envisioned that local conformational rearrangements may lead to “repacking” of protein molecules in the crystal lattice, entailing small overall translation and reorientation (i.e., causing rocking motion). Conversely, rocking motion may lead to subtle changes in the pattern of crystal contacts, triggering local conformational transitions. In general, it should not be surprising that a molecule that is conformationally labile produces a poor crystal which suffers, inter alia, from rocking dynamics.

Probing relaxation dispersion effects. MD simulations on a microsecond time scale can be used to model BMCRD effects. For this purpose MD coordinates are fed, frame by frame, into chemical shift prediction programs such as SHIFTX[54]. In a situation when the calculated chemical shift $\delta(^{15}\mathrm{N})$ shows a pattern of modulation on the μs time scale, one may expect to experimentally detect RD effects at the respective site[55]. Ultimately, this approach makes it possible to uncover the dynamic mechanism behind the observable RD profiles.

We have used this method to probe the backbone ^{15}N sites in MPD-ub and cubic-PEG-ub trajectories. In MPD-ub, relaxation dispersion effects have been predicted for sites in the β-turn experiencing βI ↔ βII exchange, as well as sites at the N-terminal end of the first α helix that are hydrogen-bonded to this β-turn. One example of such behavior, residue R54 from MPD-ub, is illustrated in the Supplementary Movie 2, demonstrating that the βI ↔ βII transition is accompanied by a ^{15}N chemical-shift change in R54.

Similarly, the analysis of the cubic-PEG-ub trajectory suggests that Bloch–McConnell dispersions can occur in the β-turn and its coupled sites (particularly, in the chain B molecules). In addition, we have also found several sites where the origin of predicted RD effects is different. One example of such distinctive behavior is residue H68. Focusing on one specific ubiquitin molecule in cubic-PEG-ub trajectory, we observe H68 side chain engaged in

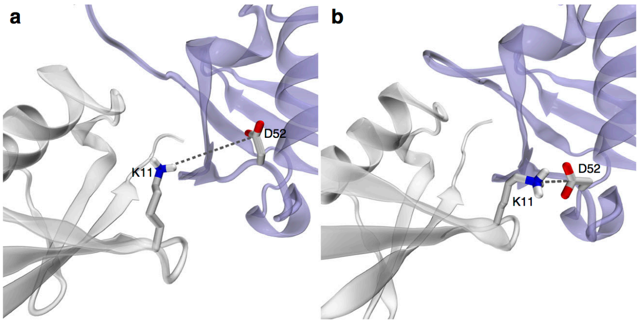


Fig. 5 Dynamic ion pair K11$_{\mathrm{chainA}}$–D52$_{\mathrm{chainB}}$ interaction in cubic-PEG-ub crystals. **a** Crystallographic structure of cubic-PEG-ub (PDB entry 3N30[45]) and **b** a selected frame from cubic-PEG-ub trajectory. The dynamic disorder in residue K11$_{\mathrm{chainA}}$ is reflected in its crystallographic temperature factors; this residue's side chain has not been modeled in our recent structure 4XOL

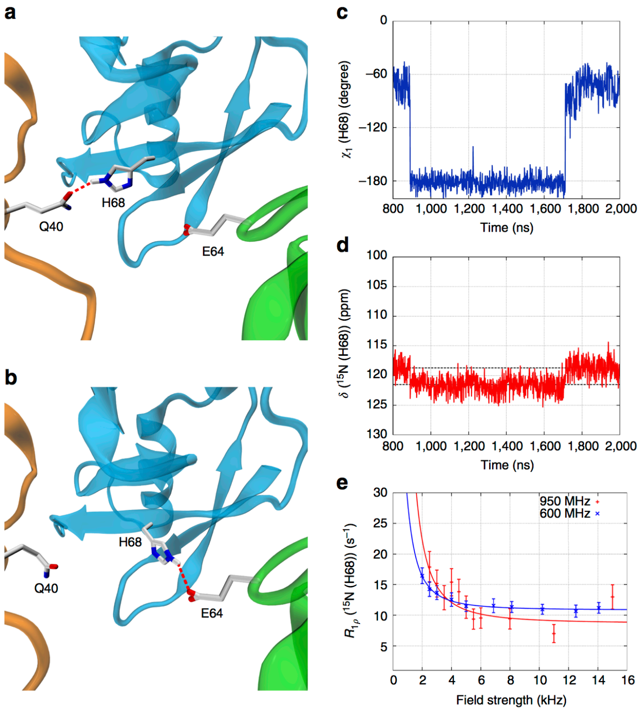


Fig. 6 Microsecond time scale dynamics of residue H68 (chain A) in cubic-PEG-ub crystals. **a**, **b** Intermolecular hydrogen bonds alternatively formed between H68 and Q40 or E64 as seen in the MD simulation of cubic-PEG-ub. **c**, **d** Time variation of side-chain torsional angle $\chi_1(t)$ and main-chain ^{15}N chemical shift $\delta(t)$ for residue H68 in the selected ubiquitin molecule from the simulated unit crystal cell. **e** Experimental relaxation dispersion data for residue H68 (chain unassigned) in the cubic-PEG-ub crystal; also shown are the results of fitting using fast exchange model (*continuous lines*). Error bars of $R_{1\rho}$ values were determined from Monte Carlo analysis. The motional time scale according to the experimental data, 56 µs, is one-to-two orders of magnitude slower than suggested by the MD simulations (the same is true for other manifestations of µs dynamics in ubiquitin crystals). The animated version of this graph is available as Supplementary Movie 3

hydrogen bond with Q40 side chain from a neighboring molecule (Fig. 6a). Later during the simulation H68 side chain swings sideways and forms hydrogen bond with side chain E64 that belongs to another neighboring molecule (Fig. 6b). After ca. 0.8 µs it makes a reverse transition, reestablishing hydrogen bond with Q40, and remains in this conformation until the end of the trajectory.

The rotameric transitions of H68 (from $\chi_1 = -60°$ to 180° and back) cause modulation of the backbone chemical shifts in this residue. For the ^{15}N spin the amplitude of modulation is calculated to be 2.1 ppm (corresponding to the gap between the two *dashed horizontal lines* in Fig. 6d). This is a sizable effect, which can be reliably reproduced by chemical shift prediction software. Therefore the computational analysis points toward possible RD effect at the amide site in residue H68. This prediction is borne out by the experimental data illustrated in Fig. 6e.

Of interest, there is a theoretical possibility that conformational exchange in H68 is influenced by rocking motion (cf. previous section). Indeed, slight reorientation of the central protein molecule relative to its neighbors should favor either H68-Q40 or H68-E64 interaction depending on the direction of reorientational fluctuation (Fig. 6a, b). In this connection, it is noteworthy that RD effects have been experimentally detected at this site in the cubic-PEG-ub crystal form that is affected by rocking, but not in the MPD-ub crystal that is immune to rocking (Supplementary Fig. 3).

Discussion

The advent of novel experimental and theoretical tools to decipher the heterogeneity of structures co-existing in crystals has sparked increased interest in molecular motions of crystalline proteins and the role of crystal packing in the context of dynamics. Here we have obtained a comprehensive portrayal of µs dynamics in different crystal forms of ubiquitin. Experimental ^{15}N $R_{1\rho}$ relaxation data, specifically the combination of BMCRD and NERRD analyses, produced evidence of two major motional modes: conformational exchange in the β-turn formed by residues E51-R54 and overall rocking. The characteristic time scale of a local β-turn interconversion motion in the two investigated crystals is ca. 100 µs, more than one order of magnitude slower than in solution. Overall rocking motion entails small reorientational fluctuations of the protein molecules within the confines of the crystal lattice and affects molecules in one of the two crystal forms, cubic-PEG-ub, much more significantly than in the other, MPD-ub. The amplitude of the rocking motion in cubic-

PEG-ub was determined to be ca. 4°, and the time constant was found to be on the scale of tens of microseconds. These results, as well as other findings reported in our study, show that μs dynamics is dependent on the environment.

MD simulations, including mutant simulations, can rationalize these effects. In particular, we can tentatively explain why MPD-ub prefers the type-II β-turn conformation. The acquired experimental evidence together with MD data hint at the possibility that internal conformational dynamics may be coupled to the rocking motion. Elucidating this connection in protein crystals is a daunting task. In particular, accurately modeling something as subtle and fickle as crystal contacts is a major challenge for the existing MD methodology. Nevertheless, we expect that future advances in the field of MD simulations should eventually make it possible. The identification of intermolecular interactions that control overall motion, either directly or through their involvement in internal conformational exchange, might open future possibilities for rational design of mutant proteins that could reduce the effect of rocking and thus improve diffraction quality.

It is known that internal dynamics undermine the quality of X-ray diffraction data. Crystallographers go a great length in order to reduce the level of dynamics in their samples (e.g., by engineering point mutations or by introducing ligands). It has not been fully appreciated, however, that internal conformational exchange may be connected to rocking motion, which can be even more detrimental for crystallographic analyses. We have shown here that both internal and overall motions can be probed by ssNMR spectroscopy and MD simulations. It is envisaged that both methods will be successfully exploited in future to either help in designing of stable crystals or, conversely, to assist in X-ray diffraction-based studies of protein motions.

Methods

Sample preparation. Uniformly ^{2}H,^{15}N-labeled human ubiquitin (without any affinity tag) was overexpressed in *E. coli*, and purified using anion exchange and size exclusion chromatographies. In order to have a dilute ^{1}H spin network and thus minimize the possible perturbing effect of remote ^{1}H spins on the relaxation of amide ^{15}N, the protein was dissolved in H_2O:D_2O (20:80) mixture at pH 7, resulting in isotope incorporation of ^{1}H/^{2}H at exchangeable sites at a ca. 20:80 ratio, and then lyophilized. Solutions used for crystallization were prepared by dissolving this lyophilized protein in buffers containing the same H_2O:D_2O ratio. All specified pH values below are corrected for the glass electrode isotope effect[56]. For generating MPD-ub crystals, lyophilized ubiquitin was dissolved at 20 mg ml^{-1} concentration in buffer A (20 mM ammonium acetate, pH 4.2). Buffer B (50 mM citrate, pH 4.2) was mixed with MPD at a volume ratio of 40:60. A volume of 500 μl of this solution was deposited in the wells of sitting-drop crystallization plates. The protein solution (37 μl) was mixed with the buffer B/MPD solution (10 μl) and deposited as sitting-drop in the crystallization plate (24-well plates with 1.5 ml reservoir volume). The plate was covered with sealing tape and kept at 4 °C; crystals appeared after 1–2 weeks, and were filled into a 1.6 mm Agilent or 1.3 mm Bruker rotor using an ultracentrifuge device; caps were glued with epoxy glue to avoid dehydration. Cubic-PEG-ub crystals were prepared with the same equipment, but the protein was dissolved in buffer C (20 mM ammonium acetate pH 4.2) and mixed with buffer D (200 mM zinc acetate, 100 mM MES, pH 6.3, 20% w/v PEG 3350) and deposited as sitting-drop. The mother liquor reservoir was filled with 500 μl of buffer D. Crystals were obtained at 20 °C after circa 3–4 weeks.

Solid-state NMR and data analysis. ^{15}N $R_{1\rho}$ relaxation measurements were performed on (i) a 600 MHz (14.1 T) Agilent VNMRS spectrometer (only BMCRD experiment) equipped with a 1.6 mm triple-resonance MAS probe tuned to ^{1}H,^{13}C,^{15}N frequencies and (ii) a 600 MHz and (iii) a 950 MHz (22.3 T) Bruker Avance 2 spectrometer; the latter two were equipped with 1.3 mm MAS probes tuned to ^{1}H,^{13}C,^{15}N frequencies with an auxiliary ^{2}H coil. Resonance assignments of MPD-ub and cubic-PEG-ub have been reported elsewhere[35]. BMCRD experiments at 600 MHz were performed at 39.5 kHz MAS, while BMCRD data at 950 MHz were collected at 50 kHz MAS. BMCRD data of MPD-ub at 600 MHz have been reported before[34]. NERRD data on both MPD-ub and cubic-PEG-ub were recorded on the 600 MHz (Bruker) spectrometer at a MAS frequency of 44.053 kHz. In order to enhance resolution, the BMCRD data on cubic-PEG-ub at 950 MHz were recorded as a series of proton-detected 3D hCONH-based experiments using a two-point approach with spin-lock periods of 1 and 50 ms. The 3D approach enabled us to resolve the otherwise overlapped I23 resonance. All other data were collected with 2D proton-detected hNH correlation experiments, shown in ref. [57], using a series of relaxation delays (typically 8 points up to 125 ms). In all experiments, cross-polarization steps were used for transfer, with ~85 kHz ^{1}H RF field, and a ^{15}N RF field adjusted to the $n = 1$ Hartmann–Hahn condition. Peak volumes in the individual 2D (3D) spectra were measured in NMRView (One-Moon Scientific), and relaxation rate constants were derived from mono-exponential fits, using fitting routines written in python/numpy/scipy language. On-resonance $R_{1\rho}$ relaxation rate constants were derived from the measured rate constants taking into account the resonance offset of the given cross-peak from the RF-field carrier Ω and the longitudinal relaxation-rate constant R_1 as described elsewhere[34]. This correction is fairly small, and visible primarily for low (less than ~3 kHz) RF fields. Experiments at 950 MHz were measured twice, with two different carrier settings (112 and 125 ppm), and the data reported for each residue correspond to the experiment with the carrier position closest to the resonance frequency. Error bars of relaxation rate constants were estimated from standard Monte Carlo simulations[58]: the parameters obtained from fitting the experimental data set were used to generate a set of ca. 500 synthetic data sets, in which the peak intensities were randomly varied, relative to the back-calculated point, within three times the spectral noise level. These synthetic data sets were subsequently fitted in an identical way as the original data set, resulting in 500 fitted decay rate constants. The reported error bars correspond to the standard deviation of these rate constants.

Fits of a two-state exchange model as reported by Meiboom[59], to the BMCRD data were obtained from the program relax, version 4.0[60]. For the fits shown in Fig. 1, we chose five residues for which data of both crystals were available: I23 (only 950 MHz, as no 600 MHz data were available for cubic-PEG-ub) and V26, K27, T55, D58 (600 and 950 MHz). We have investigated how the choice of residues, and the choice of static magnetic field strengths influences the outcome of these fits; we find that using different subgroups of residues in the β-turn region, or including other residues in the fits where available (e.g., T22 in MPD-ub) does not significantly alter the results (Supplementary Fig. 5).

For the analysis of NERRD data, we used numerical simulations, rather than the analytical equations (Eqs. (1) and (2)). We prefer numerical simulations, because Eqs. (1) and (2) report accurately on the initial slope of relaxation decays, but fail to describe the multi-exponential behavior of the relaxation due to crystallite averaging. It has been proposed to determine the initial slope experimentally[61], but we instead choose (i) to use monoexponential fits of the experimental data and (ii) to interpret the resulting rate constants by using numerical simulations of decays, imitating the experimental setup as closely as possible. A similar type of data analysis, involving numerical simulations of spin relaxation, has been used previously[40]. Details about the numerical simulations of the spin evolution and numerical fits are described in the Supporting Information.

MD simulations. The 2-μs-long MD simulations of MPD-ub crystal (24 ubiquitin molecules) and cubic-PEG-ub crystal (48 ubiquitin molecules) recorded under Amber ff99SB*-ILDN force field[62–64] in the presence of SPC/E water[65] are the same as reported by us previously[35], and described in detail in the Supplementary Methods. Analogous 2-μs-long trajectories have been recorded for MPD-ub crystals containing single-residue mutations E24A, G53A, K63A, E64A, or a negatively charged form of E24. For cubic-PEG-ub form, we have recorded the duplicate 2-μs trajectory of the wild-type ubiquitin crystal and the additional trajectories for crystals with K11A, K11A (chain A only), and G53A mutations. All alanine substitutions did not create any steric conflicts in the initial crystal coordinates. The compound statistics from each crystal trajectory is equivalent to $2 \times 24 = 48$ or $2 \times 48 = 96$ μs of ubiquitin dynamics. (Strictly speaking, many short trajectories are not equivalent to one long one, but in our case the simulation length is sufficient to sample the events of interest, i.e., conformational exchange and rocking). The trajectories exhibit a certain amount of drift, as has been discussed by us earlier;[35] the drift is essentially a reflection of rocking motion that becomes progressively worse during the course of the simulations. In addition, we have also recorded a 1-μs-long trajectory of cubic-PEG-ub crystal, where C^α atoms within the secondary structure of ubiquitin have been restrained to their respective positions in the crystallographic structure by means of harmonic restraints with force constant 5 kcal mol^{-1} Å^{-2}. The rocking motion was largely suppressed in this trajectory. However, this trajectory featured a significant proportion of distinctive βII' species, which appears to be an artifact. Therefore the restrained simulation was discontinued. As a control, we have also recorded 2-μs-long trajectories of MPD-ub crystal and cubic-PEG-ub crystal (wild-type, K11A, K11A chain A only, and G53A) using newer ff14SB force field[66]. Finally, the previously reported solution trajectory of ubiquitin has been extended to 10 μs.

In processing the MD trajectories, we used geometric criteria to identify hydrogen bonds (nitrogen–oxygen distance <3.2 Å, angle from 130° to 180°) and salt bridges (at least one nitrogen–oxygen distance <4 Å, centroid distance 5 Å)[67, 68]. Chemical shifts have been calculated by processing the MD coordinates using the program SHIFTX[54]. This program has essentially the same accuracy as SHIFTX + module from SHIFTX2[69], while relying on a simpler parameterization that nicely captures the dependence of ^{15}N chemical shift on several essential structural variables. Given the limited accuracy of chemical shift prediction software, only the more pronounced modulation effects with $\Delta\delta(^{15}\text{N})$ on the order of 1 ppm can be reliably identified. Special tests have been performed to evaluate intermolecular contributions into chemical shifts; we have found that these contributions can be safely neglected for the problem at hand.

Data availability. Solid-state NMR relaxation data and MD simulation frames can be obtained from the corresponding authors upon request.

Received: 19 December 2016 Accepted: 7 June 2017

References

1. Burnley, B. T., Afonine, P. V., Adams, P. D. & Gros, P. Modelling dynamics in protein crystal structures by ensemble refinement. *Elife* **1**, e00311 (2012).
2. de Bakker, P. I. W., Furnham, N., Blundell, T. L. & DePristo, M. Conformer generation under restraints. *Curr. Opin. Struct. Biol.* **16**, 160–165 (2006).
3. Lang, P. T., Holton, J. M., Fraser, J. S. & Alber, T. Protein structural ensembles are revealed by redefining X-ray electron density noise. *Proc. Natl Acad. Sci. USA* **111**, 237–242 (2014).
4. Fraser, J. S. et al. Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proc. Natl Acad. Sci. USA* **108**, 16247–16252 (2011).
5. Van Benschoten, A. H. et al. Measuring and modeling diffuse scattering in protein X-ray crystallography. *Proc. Natl Acad. Sci. USA* **113**, 4069–4074 (2016).
6. Wall, M. E. et al. Conformational dynamics of a crystalline protein from microsecond-scale molecular dynamics simulations and diffuse X-ray scattering. *Proc. Natl Acad. Sci. USA* **111**, 17887–17892 (2014).
7. Mittermaier, A. & Kay, L. E. New tools provide new insights in NMR studies of protein dynamics. *Science* **312**, 224–228 (2006).
8. Mao, B., Tejero, R., Baker, D. & Montelione, G. T. Protein NMR structures refined with Rosetta have higher accuracy relative to corresponding X-ray crystal structures. *J. Am. Chem. Soc.* **136**, 1893–1906 (2014).
9. Bax, A. Weak alignment offers new NMR opportunities to study protein structure and dynamics. *Protein Sci.* **12**, 1–16 (2003).
10. Ferrari, D. et al. Catalysis and electron transfer in protein crystals: the binary and ternary complexes of methylamine dehydrogenase with electron acceptors. *Biochim. Biophys. Acta* **1647**, 337–342 (2003).
11. Pearson, A. R. & Wilmot, C. M. Catching catalysis in the act: using single crystal kinetics to trap methylamine dehydrogenase reaction intermediates. *Biochim. Biophys. Acta* **1647**, 381–389 (2003).
12. Ronda, L., Bruno, S., Bettati, S. & Mozzarelli, A. Protein crystal microspectrophotometry. *Biochim. Biophys. Acta* **1814**, 734–741 (2011).
13. Merli, A. et al. Enzymatic and electron transfer activities in crystalline protein complexes. *J. Biol. Chem.* **271**, 9177–9180 (1996).
14. Mozzarelli, A. & Rossi, G. L. Protein function in the crystal. *Annu. Rev. Biophys. Biomol. Struct.* **25**, 343–365 (1996).
15. Sahu, S., Bhuyan, A., Majumdar, A. & Udgaonkar, J. Backbone dynamics of barstar: a ^{15}N NMR relaxation study. *Proteins* **41**, 460–474 (2000).
16. Kordel, J., Skelton, N., Akke, M., Palmer, A. G. & Chazin, W. Backbone dynamics of calcium-loaded calbinding-D$_{9k}$ studied by two-dimensional proton-detected ^{15}N NMR spectroscopy. *Biochemistry* **31**, 4856–4866 (1992).
17. Powers, R., Clore, G., Garrett, D. & Gronenborn, A. Relationships between the precision of high-resolution protein NMR structures, solution order parameters and crystallographic B factors. *J. Magn. Reson. B* **101**, 325–327 (1993).
18. Fenwick, R. B., van den Bedem, H., Fraser, J. S. & Wright, P. E. Integrated description of protein dynamics from room-temperature X-ray crystallography and NMR. *Proc. Natl Acad. Sci. USA* **111**, E445–E454 (2014).
19. Agarwal, V., Xue, Y., Reif, B. & Skrynnikov, N. R. Protein side-chain dynamics as observed by solution- and solid-state NMR spectroscopy: a similarity revealed. *J. Am. Chem. Soc.* **130**, 16611–16621 (2008).
20. Haller, J. D. & Schanda, P. Amplitudes and time scales of picosecond-to-microsecond motion in proteins studied by solid-state NMR: a critical evaluation of experimental approaches and application to crystalline ubiquitin. *J. Biomol. NMR* **57**, 263–280 (2013).
21. Yang, J., Tasayco, M. L. & Polenova, T. Dynamics of reassembled thioredoxin studied by magic angle spinning NMR: snapshots from different time scales. *J. Am. Chem. Soc.* **131**, 13690–13702 (2009).
22. Halle, B. Flexibility and packing in proteins. *Proc. Natl Acad. Sci. USA* **99**, 1274–1279 (2002).
23. Eastman, P., Pellegrini, M. & Doniach, S. Protein flexibility in solution and in crystals. *J. Chem. Phys.* **110**, 10141–10152 (1999).
24. Stocker, U., Spiegel, K. & van Gunsteren, W. On the similarity of properties in solution or in the crystalline state: a molecular dynamics study of hen lysozyme. *J. Biomol. NMR* **18**, 1–12 (2000).
25. Rueda, M. et al. A consensus view of protein dynamics. *Proc. Natl Acad. Sci. USA* **104**, 796–801 (2007).
26. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).
27. Massi, F. Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR R$_{1\rho}$ relaxation experiments. *Protein Sci.* **14**, 735–742 (2005).
28. Hansen, D. F., Feng, H., Zhou, Z., Bai, Y. & Kay, L. E. Selective characterization of microsecond motions in proteins by NMR relaxation. *J. Am. Chem. Soc.* **131**, 16257–16265 (2009).
29. Salvi, N., Ulzega, S., Ferrage, F. & Bodenhausen, G. Time scales of slow motions in ubiquitin explored by heteronuclear double resonance. *J. Am. Chem. Soc.* **134**, 2481–2484 (2012).
30. Smith, C. A. et al. Allosteric switch regulates protein–protein binding through collective motion. *Proc. Natl Acad. Sci. USA* **113**, 3269–3274 (2016).
31. Mills, J. L. & Szyperski, T. Protein dynamics in supercooled water: the search for slow motional modes. *J. Biomol. NMR* **23**, 63–67 (2002).
32. Lindorff-Larsen, K., Maragakis, P., Piana, S. & Shaw, D. E. Picosecond to millisecond structural dynamics in human ubiquitin. *J. Phys. Chem. B* **120**, 8313–8320 (2016).
33. Tollinger, M., Sivertsen, A. C., Meier, B. H., Ernst, M. & Schanda, P. Site-resolved measurement of microsecond-to-millisecond conformational-exchange processes in proteins by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **134**, 14800–14807 (2012).
34. Ma, P. et al. Probing transient conformational states of proteins by solid-state R$_{1\rho}$ relaxation-dispersion NMR spectroscopy. *Angew. Chem. Int. Ed.* **53**, 4312–4317 (2014).
35. Ma, P. et al. Observing the overall rocking motion of a protein in a crystal. *Nat. Commun.* **6**, 8361 (2015).
36. Quinn, C. M. & McDermott, A. E. Monitoring conformational dynamics with solid-state R$_{1\rho}$ experiments. *J. Biomol. NMR* **45**, 5–8 (2009).
37. Quinn, C. M. & McDermott, A. E. Quantifying conformational dynamics using solid-state R$_{1\rho}$ experiments. *J. Magn. Reson.* **222**, 1–7 (2012).
38. Palmer, A. G. & Massi, F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chem. Rev.* **106**, 1700–1719 (2006).
39. Kurbanov, R., Zinkevich, T. & Krushelnitsky, A. The nuclear magnetic resonance relaxation data analysis in solids: general R$_1$/R$_{1\rho}$ equations and the model-free approach. *J. Chem. Phys.* **135**, 184104 (2011).
40. Smith, A. A., Testori, E., Cadalbert, R., Meier, B. H. & Ernst, M. Characterization of fibril dynamics on three timescales by solid-state NMR. *J. Biomol. NMR* **65**, 171–191 (2016).
41. Lakomek, N.-A. et al. Microsecond Dynamics in Ubiquitin Probed by Solid-State 15N NMR Spectroscopy R1ρ Relaxation Experiments under Fast MAS (60–110 kHz). *Chem. Eur. J.* doi:10.1002/chem.201701738 (2011).
42. Zinkevich, T., Chevelkov, V., Reif, B., Saalwächter, K. & Krushelnitsky, A. Internal protein dynamics on ps to μs timescales as studied by multifrequency 15N solid-state NMR relaxation. *J. Biomol. NMR* **57**, 219–235 (2013).
43. Lamley, J. M., C. Öster, C., Stevens, R. A. & Lewandowski, J. R. Intermolecular Interactions and Protein Dynamics by Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* **54**, 15374–15378 (2015).
44. Ma, P. et al. Probing transient conformational states of proteins by solid-state R1ρ relaxation dispersion NMR spectroscopy. *Angew. Chem. Int. Ed.* **53**, 4312–4317 (2014).
45. Arnesano, F. et al. Crystallographic analysis of metal-ion binding to human Ubiquitin. *Chem Eur. J.* **17**, 1569–1578 (2011).
46. Huang, K.-Y., Amodeo, G. A., Tong, L. & McDermott, A. The structure of human ubiquitin in 2-methyl-2,4-pentanediol: a new conformational switch. *Protein Sci.* **20**, 630–639 (2011).
47. Falini, G. et al. Structural probing of Zn(II), Cd(II) and Hg(II) binding to human ubiquitin. *Chem. Commun.* **45**, 5960–5962 (2008).
48. Kuzmanic, A., Pannu, N. S. & Zagrovic, B. X-ray refinement significantly underestimates the level of microscopic heterogeneity in biomolecular crystals. *Nat. Commun.* **5**, 3220 (2014).
49. Mollica, L. et al. Atomic-resolution structural dynamics in crystalline proteins from NMR and molecular simulation. *J. Phys. Chem. Lett.* **3**, 3657–3662 (2012).
50. Xue, Y. & Skrynnikov, N. R. Ensemble MD simulations restrained via crystallographic data: accurate structure leads to accurate dynamics. *Protein Sci.* **23**, 488–507 (2014).
51. Stafford, K. A., Robustelli, P. & Palmer, A. G. Thermal adaptation of conformational dynamics in ribonuclease H. *PLoS Comput. Biol.* **9**, e1003218 (2013).
52. Sidhu, A., Surolia, A., Robertson, A. D. & Sundd, M. A hydrogen bond regulates slow motions in ubiquitin by modulating a β-turn flip. *J. Mol. Biol.* **411**, 1037–1048 (2011).
53. Olsson, M. H. M., Sondergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011).

54. Neal, S., Nip, A. M., Zhang, H. & Wishart, D. S. Rapid and accurate calculation of protein ^{1}H, ^{13}C and ^{15}N chemical shifts. *J. Biomol. NMR* **26**, 215–240 (2003).
55. Xue, Y., Ward, J. M., Yuwen, T., Podkorytov, I. S. & Skrynnikov, N. R. Microsecond time-scale conformational exchange in proteins: using long molecular dynamics trajectory to simulate NMR relaxation dispersion data. *J. Am. Chem. Soc.* **134**, 2555–2562 (2012).
56. Glasoe, P., Long, F., McConnell, H. M. & Redfield, A. G. Use of glass electrodes to measure acidities in deuterium oxide. *J. Chem. Phys.* **28**, 19–31 (1957).
57. Kurauskas, V. et al. Cross-correlated relaxation of dipolar coupling and chemical-shift anisotropy in magic-angle spinning $R_{1\rho}$ NMR measurements: application to protein backbone dynamics measurements. *J. Phys. Chem. B* **120**, 8905–8913 (2016).
58. Motulsky, H. & Christopoulos, A. *Fitting Models to Biological Data Using Linear and Nonlinear Regression* (GraphPad Software Inc., 2003).
59. Meiboom, S. Nuclear magnetic resonance study of the proton transfer in water. *J. Chem. Phys.* **34**, 375–388 (1961).
60. Morin, S. et al. Relax: the analysis of biomolecular kinetics and thermodynamics using NMR relaxation dispersion data. *Bioinformatics* **30**, 2219–2220 (2014).
61. Krushelnitsky, A., Zinkevich, T., Reif, B. & Saalwächter, K. Slow motions in microcrystalline proteins as observed by MAS-dependent ^{15}N rotating-frame NMR relaxation. *J. Magn. Reson.* **248**, 8–12 (2014).
62. Hornak, V. et al. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725 (2006).
63. Best, R. B. & Hummer, G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B* **113**, 9004–9015 (2009).
64. Lindorff-Larsen, K. et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinf.* **78**, 1950–1958 (2010).
65. Cerutti, D. S., Le Trong, I., Stenkamp, R. E. & Lybrand, T. P. Simulations of a protein crystal: explicit treatment of crystallization conditions links theory and experiment in the streptavidin-biotin complex. *Biochemistry* **47**, 12065–12077 (2008).
66. Maier, J. A. et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
67. Kumar, S. & Nussinov, R. Relationship between ion pair geometries and electrostatic strengths in proteins. *Biophys. J.* **83**, 1595–1612 (2002).
68. Xue, Y., Yuwen, T., Zhu, F. & Skrynnikov, N. R. Role of electrostatic interactions in binding of peptides and intrinsically disordered proteins to their folded targets. 1. NMR and MD characterization of the complex between the c-Crk N-SH3 domain and the peptide Sos. *Biochemistry* **53**, 6473–6495 (2014).
69. Han, B., Liu, Y., Ginzinger, S. W. & Wishart, D. S. SHIFTX2: significantly improved protein chemical shift prediction. *J. Biomol. NMR* **50**, 43–57 (2011).
70. Clore, G. M., Driscoll, P. C., Wingfield, P. T. & Gronenborn, A. M. Analysis of the backbone dynamics of interleukin-1 beta using two-dimensional inverse detected heteronuclear ^{15}N-^{1}H NMR spectroscopy. *Biochemistry* **29**, 7387–7401 (1990).

Acknowledgements

We thank Troels Schwarz-Linnet (University of Copenhagen) for helpful advice about the use of the program relax for data analysis, as well as Dr. Peixiang Ma, Jens D. Haller, Oleg Mikhailovskii, and Manfred Burghammer for insightful discussions. This work was financially supported by the European Research Council (ERC-Stg-2012-311318-Prot-Dyn2Function), the French Research Agency ANR (ANR 10-PDOC-011-01), as well as Commissariat à l'énérgie atomique et aux énérgies alternatives (CEA), Centre National de la Recherche Scientifique (CNRS) and Université Grenoble Alpes. This work used the platforms of the Grenoble Instruct Center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). S.A.I., O.N.R., and N.R.S. acknowledge Russian Science Foundation for financial support of the MD simulation component of this study (RSF grant 15-14-20038). N.C. is supported by a fellowship from the Fondation France Alzheimer.

Author contributions

V.K. and P.S. performed and analyzed NMR experiments. A.H. and I.A. prepared protein samples and crystals. P.S., A.H., J.W., N.C., J.-P.C., and A.S. crystallized ubiquitin in different crystal forms or characterized crystal samples. S.A.I., O.N.R., Y.X., T.Y., and N.R.S. performed and analyzed MD simulations. P.S. and N.R.S. designed the research. P. S. and N.R.S. wrote the manuscript with contributions from other authors.

Additional information

Supplementary Information accompanies this paper at doi:10.1038/s41467-017-00165-8.

Competing interest: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Appendix F: Photochemical mechanism of an atypical algal phytochrome

ChemPubSoc Europe
DOI: 10.1002/cbic.201800016
CHEMBIOCHEM
Full Papers

Photochemical Mechanism of an Atypical Algal Phytochrome

Uzma Choudry+,[a] Derren J. Heyes+,[a] Samantha J. O. Hardman+,[a] Michiyo Sakuma,[a]
Igor V. Sazanovich,[b] Joyce Woodhouse,[c] Eugenio De La Mora,[c] Martin N. Pedersen,[d]
Michael Wulff,[d] Martin Weik,[c] Giorgio Schirò,[c] and Nigel S. Scrutton*[a]

Phytochromes are bilin-containing photoreceptors that are typically sensitive to the red/far-red region of the visible spectrum. Recently, phytochromes from certain eukaryotic algae have become attractive targets for optogenetic applications because of their unique ability to respond to multiple wavelengths of light. Herein, a combination of time-resolved spectroscopy and structural approaches across picosecond to second timescales have been used to map photochemical mechanisms and structural changes in this atypical group of phytochromes. The photochemistry of an orange/far-red light-sensitive algal phytochrome from *Dolihomastix tenuilepis* has been investigated by using a combination of visible, IR and X-ray scattering probes. The entire photocycle, correlated with accompanying structural changes in the cofactor/protein, are reported. This study identifies a complex photocycle for this atypical phytochrome. It also highlights a need to combine outcomes from a range of biophysical approaches to unravel complex photochemical and macromolecular processes in multi-domain photoreceptor proteins that are the basis of biological light-mediated signalling.

Introduction

The red/far-red light-sensitive phytochromes (Phys) are essential photoreceptors that regulate a myriad of photo-morphological processes in plants, bacteria, fungi and algae (e.g., seed germination, leaf-opening).[1,2] Phys contain a covalently bound bilin cofactor and typically photoconvert between a red-light-absorbing P_R state and a far-red light-absorbing P_{FR} state.[1] Photoconversion between the two states is triggered by an ultrafast *E*/*Z* isomerisation of the bilin tetrapyrrole, leading to slower protein structural change (micro- to millisecond timescales) to form an active signalling conformation. The initial photoisomerisation has been studied in numerous Phys from different organisms by means of time-resolved UV/Vis, IR and Raman spectroscopy and is known to involve isomerisation around the C15 and C16 double bond between rings C and D of the bilin.[1,3–9] X-ray crystallography studies of the photosensory region of a bacterial Phy in conjunction with time-resolved X-ray scattering data have shown that this causes a subsequent refolding of a conserved tongue region, leading to large-scale conformational change in the remainder of the protein.[10,11]

Although Phys have proven to be attractive targets for new applications in the field of optogenetics,[12,13] their use has been limited due to their sensitivity to only a narrow range of wavelengths (red and far-red light). However, it was recently discovered that Phys from certain eukaryotic algae exhibited diverse spectral properties capable of spanning most of the visible region.[14] The origin of this spectral diversity is currently unknown and it is unclear whether these algal Phys undergo similar photocycles to those found in typical Phys. To exploit fully the potential of algal Phys as new optogenetic tools, it is important to develop an understanding of the photochemistry and photodynamics that underpin biological signalling.

We have investigated, in detail, the kinetics and structural changes of both the bilin cofactor and protein throughout the entire photocycle of the photosensory region of an algal Phy by using a combination of time-resolved laser photoexcitation spectroscopy across multiple timescales (continuously from picoseconds to seconds). Spectral changes in the visible region were monitored to report on significant chromophore changes, whereas the mid-IR spectral region was used to report on finer structural changes in the bilin cofactor. These measurements have been combined with time-resolved small- and wide-angle X-ray scattering (S/WAXS) measurements to understand how these spectral changes are coupled to larger conformational changes in the protein.

---

[a] U. Choudry,+ Dr. D. J. Heyes,+ Dr. S. J. O. Hardman,+ M. Sakuma,
Prof. N. S. Scrutton
Manchester Institute of Biotechnology, University of Manchester
131 Princess Street, Manchester M1 7DN (UK)
E-mail: nigel.scrutton@manchester.ac.uk

[b] Dr. I. V. Sazanovich
Central Laser Facility, Research Complex at Harwell
Science and Technology Facilities Council, Harwell Oxford
Didcot, OX11 0QX (UK)

[c] J. Woodhouse, Dr. E. De La Mora, Dr. M. Weik, Dr. G. Schirò
CNRS, Université Grenoble Alpes, CEA–Institut de Biologie Structurale
Grenoble 38044 (France)

[d] Dr. M. N. Pedersen, Dr. M. Wulff
European Synchrotron Radiation Facility
Grenoble 38044 (France)

[+] These authors contributed equally to this work.

The ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201800016.

Results and Discussion

The eukaryotic algal Phy from *Dolihomastix tenuilepis* (Dten) has a far-red light-absorbing, P_{FR}, state, with an absorption maximum at $\lambda = 714$ nm and lower intensity features at $\lambda = 375$ and 614 nm, and an orange light-absorbing, P_O, state with an absorption maximum at $\lambda = 598$ nm and a lower intensity feature at $\lambda = 354$ nm (Figure 1).[14] Time-resolved spectroscopy has been used to measure optical changes in the visible and IR regions that occur upon photoconversion of both states of the protein. Measurements in the visible region spanned the complete time range of the photoreaction (ps–ms). These measurements were repeated in H_2O and D_2O buffers to identify any kinetic isotope effects associated with each of the individual steps of the photocycle, and to provide direct comparisons to the IR data, which were collected on samples in D_2O buffer. The IR measurements were recorded over the first 400 μs of the reaction to monitor the initial structural changes of the reaction.

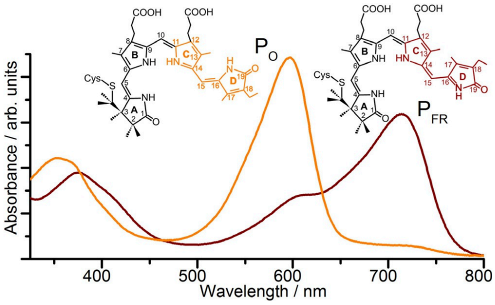


Figure 1. Absorption spectra of Dten Phy in the orange- and far-red-absorbing states, and structures of the bilin chromophore.

Global analysis has been used to model the time-resolved absorption difference spectra from the visible and IR measurements by fitting the data to a sequential, unbranched model to yield evolution-associated difference spectra (EADS). Previous studies of similar photoreactions suggest it is unlikely that any significant reaction branching will occur and that a simple sequential reaction scheme will fit the data adequately.[15,16] Major features in the IR spectral region monitored here include stretching of the C=O bonds on rings A and D at $\tilde{\nu} \approx 1730$ and $1720\ \text{cm}^{-1}$, respectively, and C=C stretching at $\tilde{\nu} \approx 1595$, 1620, 1627 and $1648\ \text{cm}^{-1}$, which represent the bonds between rings B and C, A and B, C and D, and the vinyl group attached to ring D, respectively.[17] Features from the protein amide I group and amino acid side chains occur at $\tilde{\nu} \approx 1650$ and $1450\ \text{cm}^{-1}$.[18] Contour maps of the raw data, along with the corresponding EADS data, are shown in Figures 2 and 3 (below); more extensive data and analysis can be found in the Supporting Information (Figures S1–S6).

Photoconversion of the P_O state gives rise to three resolvable steps in the visible region (represented by EADS 1–4 in Figure 2). There is an initial bleaching of the ground-state absorption feature together with the appearance of a new absorption band at $\lambda = 670$ nm and broad excited-state absorption features between $\lambda = 400$ and 500 nm (Figure 2B, D, EADS1). EADS1 then evolves into EADS2, which has an absorbance maximum at $\lambda = 660$ nm and a proportionally similarly sized excited-state absorption band, with a lifetime of 116 ps (112 ps in D_2O).

In the IR region, although there are differences in the absolute intensities, the shape and positions of the spectral features are very similar between EADS1 and EADS2 (Figure 2F). Hence, it seems unlikely that this transition corresponds to *E*/*Z* bond isomerisation, which would be expected to produce noticeable changes in this spectral region. Instead, this kinetic component is likely to represent excited-state relaxation, potentially of the excited-state P_O, which has not been isomerised. The bleach feature at $\tilde{\nu} = 1723\ \text{cm}^{-1}$ and the transient feature at $\tilde{\nu} = 1709\ \text{cm}^{-1}$, which is already present in EADS1, suggests a change in the C=O bond on ring D; hence, it is possible that *E*/*Z* isomerisation of the bilin cofactor may have already occurred before the first measurement 1 ps after excitation. This assignment is supported by the residuals of the analysis of the visible data (Figures S1 and S3), for which it is possible to observe a very short lived component that cannot be resolved kinetically due to the lack of data points on this timescale and proximity to the time resolution of the system (≈0.2 ps). The next reaction intermediate, which has a visible absorption maximum at $\lambda = 640$ nm (EADS3), is formed with a lifetime of 581 ps in H_2O (862 ps in D_2O). The IR spectra also display significant changes for this step, in both intensity and position of dominant spectral features. There are changes across the full spectral range monitored, with apparently new features observed in regions assigned to C=C stretches of the bilin, as well as amide bands from the protein; these are likely to represent small-scale structural changes and rearrangements in and around the bilin cofactor. The final reaction step takes place beyond the time range of the IR measurements, although a comparison of the final time-resolved difference spectrum with a steady-state difference spectrum of the P_O to P_{FR} conversion (Figure 3A) shows that further spectral evolution in the IR region occurs prior to the final P_{FR} formation. Hence, there are almost certainly some structural changes and rearrangement of the bilin environment beyond the 400 μs of these measurements.

The absorbance feature with a maximum at $\lambda = 725$ nm (EADS4), which is assigned to the final P_{FR} state, is then formed with a lifetime of 57.1 ms (89.3 ms in D_2O). Although all reaction lifetimes, except the first, are slower in D_2O compared with those in H_2O, the spectral shapes are virtually identical under both conditions. This implies that proton exchange does not affect the excited-state properties or isomerisation, but may influence some of the slower steps that are likely to involve changes to the conformation of the bilin cofactor and the protein itself.

In the reverse reaction, photoexcitation of the P_{FR} state (Figure 4) results in quite different dynamics than those of the forward reaction, with five resolvable kinetic steps detected. In the visible region, a fast, 0.3 ps lifetime, component corre-

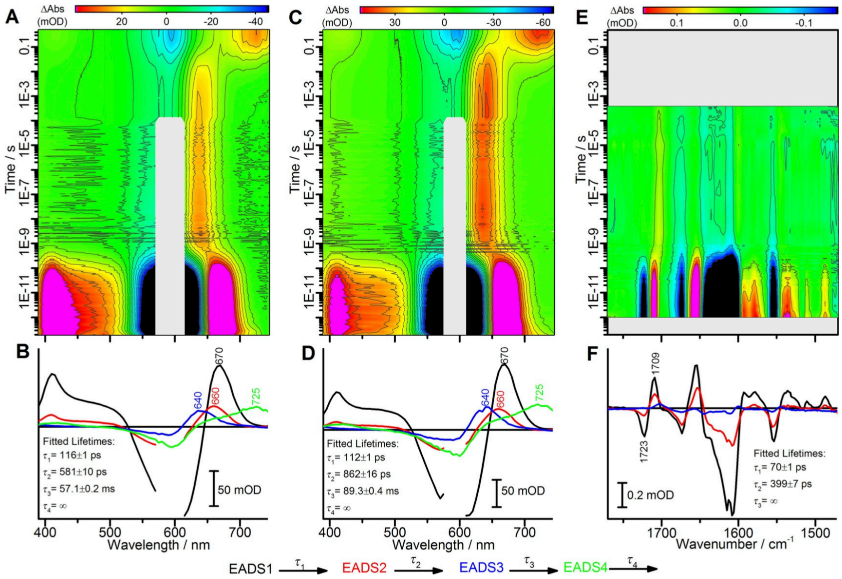


Figure 2. Contour plots of the raw data and EADS data resulting from a global analysis of time-resolved visible and IR transient absorption data, collected after excitation of the P_O state. A) and B) Visible data collected on Dten Phy in H_2O buffer. C) and D) Visible data collected on Dten Phy in D_2O buffer. E) and F) IR data collected on Dten Phy in D_2O buffer. Visible data were recorded from 0.2 ps until no further spectral changes were observed (495 ms); IR data were recorded over a 1 ps to 390 μs window. The gap in the $\lambda = 580$–605 nm range is due to scattered pump light obscuring the transient absorption signal.

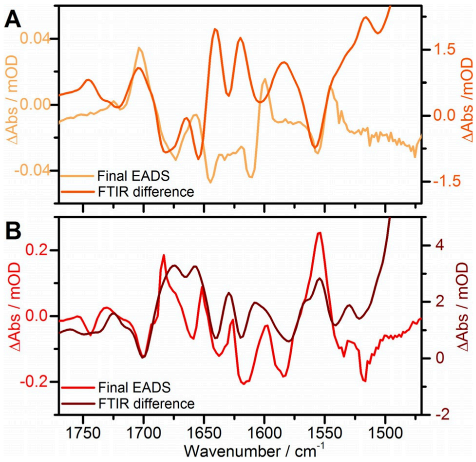


Figure 3. Comparison of the final EADS from global analysis of the time-resolved IR measurements (light-coloured line, left axis) with the static IR difference spectra (dark-coloured line, right axis) for A) P_O to P_{FR} and B) P_{FR} to P_O photoconversions.

sponds to the relaxation of a highly electronically excited state (e.g., S_2) with intense ground-state bleaching and excited-state absorption (EADS1) to a less electronically excited state (e.g., S_1), the spectra of which retains some ground-state bleaching, but displays virtually no excited-state features (EADS2). In the time region covered by both visible and IR measurements, three transitions are observed. EADS2 converts into EADS3 with a lifetime of 5–6 ps. In both the visible and IR regions, the changes result in a significant loss of spectral intensity. In the visible region, a very small positive feature is observed at $\lambda \approx 670$ nm, and the intensity of the ground-state bleaching is again reduced. This is likely to be due to an intermediate with a very slightly blue-shifted absorption feature. In the IR region, the transition corresponds to a change from an initially excited state to a spectrum with few spectral features. The feature arising from the C=O bond of ring D again provides a useful marker of the reaction progress, and the strong bleach feature at $\tilde{\nu} \approx 1700$ cm^{-1} persists for the full 400 μs IR time course. In EADS3, this bleach feature appears to be coupled to a transient feature at $\tilde{\nu} \approx 1717$ cm^{-1}; a shift in the opposite direction to that observed in the P_O to P_{FR} reaction. We therefore assign this 5 ps step to the photoisomerisation reaction. No spectral changes in either the visible or IR regions are then observed for several hundred nanoseconds. The transition from EADS3 to EADS4 occurs with a lifetime of 280 ns in the visible data in H_2O (317 ns in D_2O in the visible and 549 ns in the IR). In the visible region, the intermediate absorption band shifts to $\lambda = 625$ nm and in the IR region a new bleach/transient feature appears at $\tilde{\nu} = 1744/1730$ cm^{-1}. Based on previous findings[19] and the small kinetic isotope effect associated with this step, it is

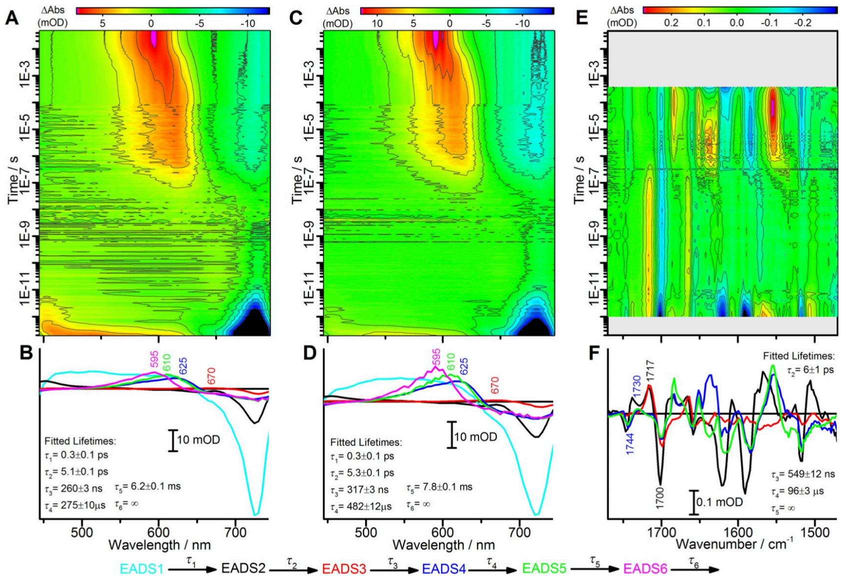


Figure 4. Contour maps of the raw and EADS data, resulting from a global analysis of visible and IR transient absorption data, which were collected after excitation of the P_{FR} state. A) and B) Time-resolved visible data collected on Dten Phy in H_2O-based solutions. C) and D) Time-resolved visible data collected on Dten Phy in D_2O-based solutions. E) and F) Time-resolved IR data collected on Dten Phy in D_2O-based solutions. Visible data were recorded from 0.2 ps until no further spectral changes were observed (22.5 ms); IR data were recorded over a 1 ps to 390 µs window.

likely that this feature is due to the deprotonation of one of the propionic side chains. The disappearance of the positive feature at $\tilde{\nu}=1717\ \mathrm{cm}^{-1}$ is assigned to the C=O bond of ring D, in accordance with previous studies of similar systems.[20] The $\lambda=625$ nm absorbing intermediate described by EADS4 is then converted into an intermediate with an absorption maximum at $\lambda=610$ nm (EADS5) with a lifetime of 275 µs (482 µs in D_2O). Due to the kinetic isotope effect associated with this step, it is likely to involve re-protonation of the bilin factor. The IR data were only collected up to 400 µs, which was close to the lifetime found in the visible datasets for this step. Consequently, it is unlikely that this step will be kinetically well resolved in the IR region; hence the 96 µs lifetime.

A comparison of the final component from the global analysis of the time-resolved IR data with the steady-state difference spectra (Figure 3 B) reveals that, although there are many similarities between the spectra, there are still a number of spectral features that evolve after 400 µs. The majority of these changes occur at $\tilde{\nu}<1650\ \mathrm{cm}^{-1}$, which is in the region that protein absorption is expected to make significant contributions. Hence, the slower steps are only likely to involve large protein conformational changes and not significant changes in the bilin environment. The final step resolved in the visible dataset shows the formation of the final $\lambda=595$ nm absorbing P_O state (EADS6) with a lifetime of 6.2 ms (7.8 ms in D_2O).

Investigations of Phy systems with the well-studied red/far-red transitions suggest that in one direction the reaction consists of a picosecond isomerisation of the bilin cofactor. This is then followed by single-bond rotations, leading to the breakage and subsequent re-formation of some of the hydrogen-bond interactions with the protein environment. The other reaction direction also involves a picosecond isomerisation step, but this is then followed by deprotonation of the bilin chromophore, other minor structural changes, and then re-protonation to form the final stable state.[21] It appears that, despite the differences in wavelength sensitivity of Dten Phy, a similar set of reactions occurs here. It is also known that the red/far-red Phys are capable of undergoing a much slower dark reversion process, which can occur over minutes or days.[22–24] Therefore, measurements over the course of several hours were undertaken to identify any dark reactions from either of the two states of Dten Phy (Figures S7–S8). The P_{FR} state fully converts into the P_O state within a few hours (lifetime of around 2 h) in the dark. However, the P_O state does not undergo any significant dark reversion back to the P_{FR} state over a 24 h time period.

In typical red/far-red Phys, it is known that the slower step(s), which follows the initial photoisomerisation, involves a series of large-scale conformational changes or domain motions in the protein.[10,11,25,26] We have used time-resolved solution X-ray scattering measurements to characterise the nature and kinetics of the structural dynamics during the P_O to P_{FR} transition in Dten Phy. A large difference signal in the small-/wide-angle region ($q<0.25\ \text{Å}^{-1}$) was observed between the two states, which indicated significant structural changes on the nanometre scale. The difference signal develops in both intensity and shape on the microsecond to second timescale (Figure 5 A).

A SVD analysis revealed that all structural data could be described by three main time-independent basis patterns (Figure 5 C–E), which formed a time-independent orthonormal basis on which the full set of time-dependent data could be decomposed. The first, which represents the shape of the signal developing on the micro- to millisecond timescale, shows multiple structural changes across all regions. The second basis pattern describes the signal developing in the tens of milliseconds and contains an oscillation that is similar to the signature of the quaternary opening motion, which has been previously described for the bacterial red/far-red Phy family.[11,27] The third basis pattern is the signal of the final state, as revealed by comparison with a static difference between P_{FR} and P_O stationary states (Figure 5 B). A kinetic analysis of the whole dataset, in terms of the three main basis patterns reveals that the structural dynamics of Dten Phy evolve with three time constants of $\tau_1 = 8.7$ ms, $\tau_2 = 60$ ms and $\tau_3 >$ 500 ms. The second of these time constants correlates well with the 57.1 ms lifetime from the visible time-resolved data. Previous measurements have proposed that the X-ray scattering signal obtained for bacterial Phys represents an opening of the homodimer in the P_{FR} state, which is caused by an increase in the separation of the opposing Phy domains.[11,27] The similar signal obtained for Dten Phy in the second basis pattern suggests that this motion is conserved in algal Phys on the millisecond timescale. However, additional structural dynamics are observed for Dten Phy on the microsecond timescale, which do not appear to be present in the bacterial Phys. Moreover, because further structural changes associated with the P_O to P_{FR} transition occur on slower timescales than that of the optical changes, it is likely that these final conformational changes in the protein do not significantly affect the environment of the bilin chromophore.

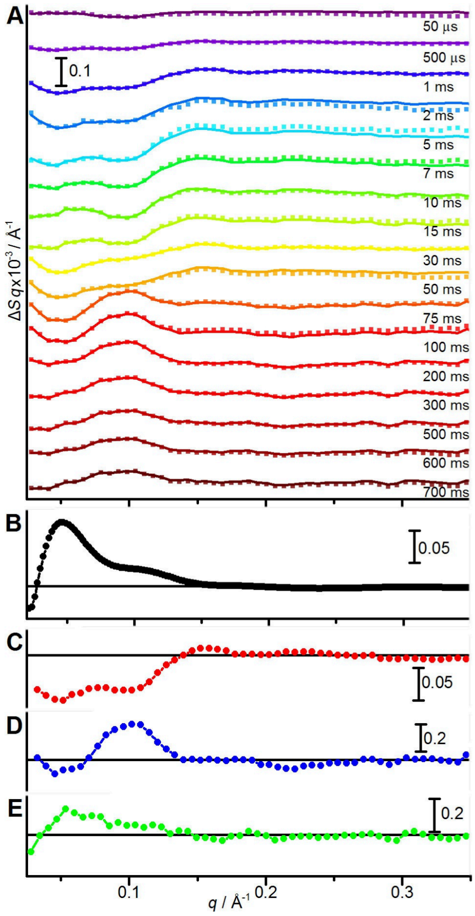


Figure 5. A) Laser-induced time-resolved X-ray scattering difference patterns for Dten Phy in solution, vertically offset for clarity; points are experimental data and lines are linear combinations of the first three basis patterns obtained by singular value decomposition (SVD). Notably, scattering differences are multiplied by the scattering vector *q* to reduce differences in the amplitude scale in S/WAXS regions. B) Static X-ray scattering difference between P_O and P_{FR} states, and C)–E) the three main basis patterns obtained by SVD analysis of the entire time-resolved X-ray scattering dataset.

The presence of these large-scale structural changes in the protein have also been confirmed by viscosity-dependence studies. Visible time-resolved experiments at a range of solvent viscosities (0–60 % glycerol) revealed that the rates of the slower micro- to millisecond steps are strongly dependent on the solvent viscosity (Figure 6). The lifetime of the final step observed in the P_O reaction increases 44-fold in 60 % glycerol, and that of the P_{FR} final lifetime by eightfold. This strongly indicates that these slower steps in the photocycle involve a series of solvent-coupled protein motions. Conversely, no such dependence on solvent viscosity could be observed on the sub-nanosecond dynamics (Figure S9).

Conclusion

Taken together, the time-resolved approaches described herein have allowed us to investigate the entire photocycle of an atypical algal Phy sensitive to orange/far-red light (Scheme 1). It is clear that the photochemical mechanism of Dten Phy is much more complex than that initially observed from time-resolved visible spectroscopy measurements. It has only been possible to unravel these hidden complexities by using additional IR and X-ray scattering probes to study the initial photoisomerisation and slower structural changes in the protein, respectively. Similar to the red/far-red photocycle in typical Phys, it appears as though the orange/far-red photocycle of Dten Phy involves an initial photoisomerisation around the C15=C16 bond between rings C and D of the bilin on the picosecond timescale. Following photoisomerisation, the formation of the final P_{FR} or P_O state involves multiple small- and large-scale structural changes in the protein, not all of which are linked to spectral changes associated with the bilin cofactor. Time-resolved X-ray scattering measurements have identified three kinetically distinct structural transitions, one of which is similar to the opening of the dimer interface that has been observed previously in bacterial Phys.[11,27] This study illustrates how crucial it is to use multiple biophysical tools to obtain a complete mechanistic understanding of these complex light-activated proteins.

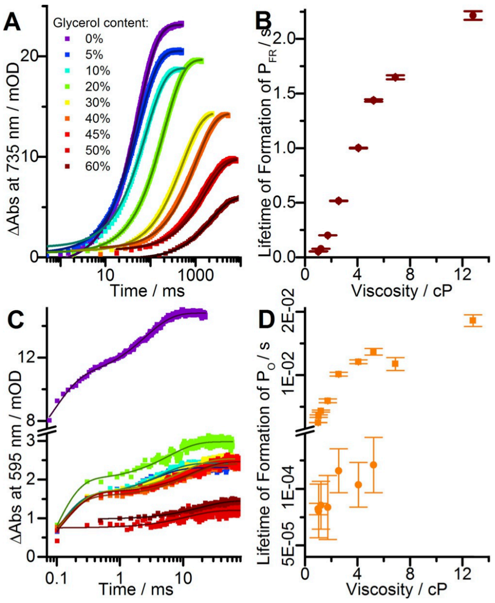
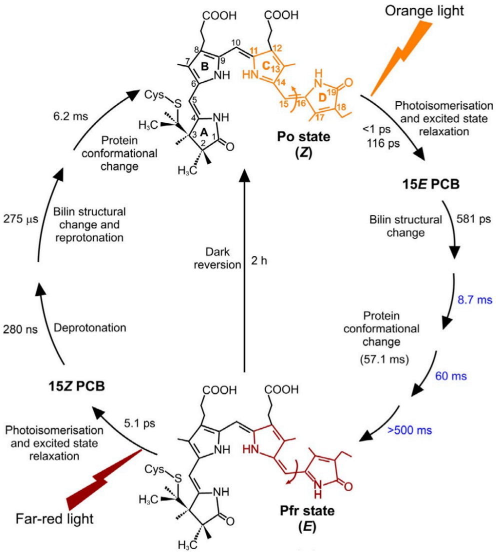


Figure 6. A) and C) The formation of P_{FR} and P_O states, following photoexcitation of P_O and P_{FR} states, respectively at a range of glycerol concentrations. Data (points) were fitted to a single- or double-exponential function (lines). B) and D) The resulting lifetimes, with fitting errors, are shown as a function of viscosity.



Scheme 1. Schematic overview of the Dten Phy photocycle, which involves changes to the phycocyanobilin (PCB) chromophore. Lifetimes derived from visible transient absorption (ps–ms) in H_2O are shown in black, and the protein structural changes identified from S/WAXS (ms–s) are shown in blue.

Experimental Section

Sample preparation and basic characterisation: The Dten gene encoding the photosensory regions PAS-GAF and PHY was ordered from GeneArt, subcloned into pET21a. The vector was then co-transformed, along with pCola Duet vector containing pcyA and HO1 genes, into HMS 174 cells. Cultures were grown at 30°C, 190 rpm, until an $OD_{600\,nm}$ of 0.1 was achieved. The temperature was then dropped to 25°C. Cultures were induced by using 100 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) at an $OD_{600\,nm}$ of between 0.4 and 0.6, and grown overnight. Nickel-IDA resin was used as the first step of purification. The resin was equilibrated with five column volumes of buffer A (50 mM Tris, pH 7.5, 300 mM NaCl, 5 mM imidazole, 7.5% glycerol, 1% β-mercaptoethanol), and incubated with the protein for 3 to 4 h. The protein flow through was collected and the resin washed with four column volumes of buffer B (50 mM Tris, pH 7.5, 300 mM NaCl, 30 mM imidazole, 7.5% glycerol, 1% β-mercaptoethanol). The protein was eluted with buffer C (50 mM Tris, pH 7.5, 300 mM NaCl, 250 mM imidazole, 7.5% glycerol, 1% β-mercaptoethanol). Size-exclusion chromatography was performed on ÄKTA Pure (GE Healthcare) by using a Hi Load 26/60 200 (GE Healthcare) column. The gel filtration column was equilibrated with two column volumes of gel filtration equilibration buffer D (50 mM Tris, pH 7.5, 300 mM NaCl, 7.5% glycerol, 1% β-mercaptoethanol) before the protein was injected. After injection, the protein was eluted in two-column volume gel filtration equilibration buffer. The relevant fractions were analysed by means of SDS-PAGE to pool together the fractions containing the purest protein. The protein was then concentrated by using a Sartorius Stedim Biotech vivaspin column with a 30 kDa cut-off, flash-frozen and stored at −80°C. UV/Vis absorption spectroscopy was performed by using a Cary 50 UV/Vis spectrophotometer, Varian, Agilent Technologies. All spectra were recorded from $\lambda = 300$ to 800 nm. Thorlabs mounted high-power light-emitting diodes (LEDs; of either $\lambda = 530$ or 735 nm) were used in the dark to convert Phys into different isomeric states. For studies of the dark thermal conversion over time, the cuvette was left inside a fully covered spectrometer. Steady-state difference spectra in the IR region were collected by using a Brucker Vertex 80 FTIR spectrometer on samples (≈100 µM) in D_2O buffer (50 mM Tris pD 7.5, 300 mM NaCl, 7.5% glycerol, 1% β-mercaptoethanol) contained between CaF_2 windows separated by a 100 µm Teflon spacer. Difference spectra were recorded for each state (P_O or P_{FR}) by using the opposite state (P_{FR} or P_O, respectively) as the background measurement. For the viscosity studies, protein (200 µL, 10 mg mL^{-1}) was added to glycerol (800 µL; 5, 10, 20, 30, 40, 45, 50, and 60%). Data from the photoconversion measurements were used as the 0% glycerol sample. Data were fitted with a single- (P_{FR} and P_O formation at 50 and 60% glycerol) or double-exponential (P_O formation at $<45\%$ glycerol) function.

Nanosecond laser flash photolysis: The Dten Phy photodynamics were studied by using a laser system consisting of an optical parameter oscillator (OPO) pumped by a Q-switched Nd:YAG laser

(Brilliant B, Quantel). Excitation wavelengths were generated by using the OPO pumped by the third harmonic of the laser. A quartz cuvette with a 2 mm path length for pump excitation and 10 mm path length for probing was used. Samples had an $OD_{735\,nm}$ of 0.8 in the 10 mm path. Transient absorption spectra were recorded by using an LKS-60 flash-photolysis instrument (Applied Photophysics Ltd.). Kinetics were recorded in 5 nm increments across the visible spectral region. Samples illuminated with a $\lambda=530$ nm LED to generate the P_{FR} state were then photoexcited at $\lambda=430$ nm with 15 mJ pulse energy (this excitation wavelength was selected due to the limited pump wavelength range of the laser system); kinetics were recorded over a timescale of 0.05–47.5 ms. Samples were illuminated with a $\lambda=735$ nm LED to generate the P_O state and then photoexcited at $\lambda=595$ nm with 9 mJ pulse energy; kinetics were recorded over a timescale of 0.1–495 ms.

Ultrafast spectroscopy: Time-resolved visible spectroscopy was performed by using a Ti:sapphire amplifier system (Newport Spectra Physics, Solstice Ace) to produce $\lambda=800$ nm pulses at 1 kHz with 100 fs pulse duration. A TOPAS Prime optical parameter amplifier (OPA) with associated NirUVis unit was pumped with a portion of the output of the amplifier to generate the pump beams. A broadband ultrafast pump–probe transient absorbance spectrometer "Helios" (Ultrafast Systems LLC) was used to collect data over a 3 ns time window, with a time resolution of around 250 fs. The probe beam consisted of a white-light continuum generated in a CaF_2 crystal and absorbance changes were monitored between $\lambda=390$ and 745 nm. A broadband sub-nanosecond pump–probe transient absorption spectrometer "Eos" (Ultrafast Systems LLC) was used to collect data on longer timescales. A 2 kHz white-light continuum fibre laser was used to generate the probe pulses; the delay between pump and probe was controlled electronically. Samples reservoirs were illuminated with $\lambda=530$ or 735 nm LEDs and flowed through a 0.2 mm path length quartz cuvette. Samples had an OD of around 0.5 at $\lambda=735$ nm in the 0.2 mm path. Excitation wavelengths of $\lambda=430$ and 595 nm, with 0.5 μJ pulse energy, were used over timescales up to 82 μs. The pre-excitation data were subtracted, and spectral chirp was corrected for, before analysis.

Time-resolved IR spectroscopy was performed by using the TRMPS setup of the ULTRA-LIFEtime system[28] at the Central Laser Facility, STFC, Rutherford Appleton Laboratory, UK. This used a 100 kHz ultrafast laser based on a custom dual Yb:KGW system (Pharos, Light Conversion). Samples in D_2O buffer (50 mM Tris pD 7.5, 300 mM NaCl, 7.5 % glycerol, 1 % β-mercaptoethanol) were contained between two CaF_2 windows, separated by a Teflon spacer to give a path length of approximately 50 μm. Samples reservoirs were illuminated with $\lambda=530$ or 735 nm LEDs and flowed through the cell; the sample holder was rastered to avoid sample damage. Excitation wavelengths of $\lambda=575$ nm at 1 kHz repetition rate and 0.6 μJ pump energy, and $\lambda=735$ nm at 500 Hz repetition rate and 2 μJ pump energy, were used. Data were collected for approximately 35 and 100 min, respectively, per dataset The excitation beam was set at the magic angle with respect to the IR probe beam. Difference spectra were generated relative to the ground state in the spectral window $\tilde{\nu}=1470$–$1770\ cm^{-1}$ at time delays ranging from 1 ps to 390 μs. Pixel to wavenumber calibration was performed by using a polystyrene standard. IR pump–probe transient absorption data from the TRMPS setup were averaged 0.1 μs on either side of each the 10 μs "TRMP" step after the first 10 μs of data. Thus, 105 (P_{FR} reaction) or 57 (P_O reaction) time points were spread over the first 10 μs time range, and thereafter timepoints were every 10 μs up to 390 μs (39 in total)

Global analysis: Global analysis was performed by using the open-source software Glotaran.[29] This procedure reduces the matrix of change in absorbance as a function of time and wavelength, to a model of one or more exponentially decaying time components, each with a corresponding evolution-associated difference spectrum (EADS). All datasets were fitted to a sequential, unbranched model. Errors quoted with the lifetime values were the standard errors calculated during the global analysis. Visible-spectrum datasets from nanosecond laser flash photolysis and ultrafast transient absorption techniques were combined into one dataset. The wavelength points from the ultrafast data were removed, so that the data matched those of the laser flash photolysis data (5 nm steps). The datasets were then scaled to match intensities of the major spectral features in overlapping time regions.

Time-resolved S/WAXS: Time-resolved S/WAXS experiments were performed at the beamline ID09 of the European Synchrotron Radiation facility. A sample of Dten Phy (250 μM) was continuously illuminated by a $\lambda=730$ nm LED to stabilise the P_O form, flowing by a peristaltic pump (Gilson Minipulse 3) into a 2 mm quartz capillary and excited by a 5 ns laser pulse (EKSPLA) at $\lambda=595$ nm ($1\ mJ\,mm^{-2}$). The X-ray scattering signal was monitored at different time delays after laser excitation by 20 μs X-ray pulses. Scattering patterns were normalised the water peak (about $2.2\ Å^{-1}$) before making differences between laser-on and -off patterns. The time evolution of the solvent heating peak at $1.8\ Å^{-1}$ was checked, and did not affect the scattering patterns in the q range over which the protein difference signal used for analysis was present (0–$0.3\ Å^{-1}$). SVD analysis of the time-resolved scattering dataset was performed by using an in-house-developed Python-based code. Visual inspection and autocorrelation analysis of both basis patterns and amplitude vectors[30] indicated that the first three basis patterns contained significant information; the rest contained essentially only random noise. A linear combination of the first three basis patterns was then used to fit the time evolution of the scattering signal and to extract time constants.

Acknowledgements

Time-resolved visible absorption measurements were performed at the Ultrafast Biophysics Facility, Manchester Institute of Biotechnology, as funded by BBSRC Alert14 Award BB/M011658/1. Time-resolved infrared measurements were performed through program access support of the UK Science and Technology Facilities Council (STFC). Studentship funding (UC) from Newport is gratefully acknowledged.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: biophysics · photochemistry · photoreceptor · phytochrome · time-resolved spectroscopy

[1] N. C. Rockwell, Y.-S. Su, Lagarias, *Annu. Rev. Plant Biol.* **2006**, *57*, 837–858.
[2] N. C. Rockwell, Lagarias, *Chem. Eur. J. Chem. Phys.* **2010**, *11*, 1172–1180.
[3] F. Andel, Lagarias, R. A. Mathies, *Biochemistry* **1996**, *35*, 15997–16008.
[4] H. Foerstendorf, T. Lamparter, J. Hughes, W. Gärtner, F. Siebert, *Photochem. Photobiol.* **2000**, *71*, 655–661.

[5] A. Remberg, I. Lindner, T. Lamparter, J. Hughes, C. Kneip, P. Hildebrandt, S. E. Braslavsky, W. Gärtner, K. Schaffner, *Biochemistry* **1997**, *36*, 13389–13395.

[6] C. Kneip, P. Hildebrandt, W. Schlamann, S. E. Braslavsky, F. Mark, K. Schaffner, *Biochemistry* **1999**, *38*, 15185–15192.

[7] W. Rüdiger, F. Thümmler, E. Cmiel, S. Schneider, *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 6244–6248.

[8] C. Schumann, R. Gross, M. M. N. Wolf, R. Diller, N. Michael, T. Lamparter, *Biophys. J.* **2008**, *94*, 3189–3197.

[9] P. Schmidt, T. Gertsch, A. Remberg, W. Gärtner, S. E. Braslavsky, K. Schaffner, *Photochem. Photobiol.* **1998**, *68*, 754–761.

[10] E. S. Burgie, T. Wang, A. N. Bussell, J. M. Walker, H. Li, R. D. Vierstra, *J. Biol. Chem.* **2014**, *289*, 24573–24587.

[11] H. Takala, A. Björling, O. Berntsson, H. Lehtivuori, S. Niebling, M. Hoernke, I. Kosheleva, R. Henning, A. Menzel, J. A. Ihalainen, S. Westenhoff, *Nature* **2014**, *509*, 245–248.

[12] J. J. Tabor, A. Levskaya, C. A. Voigt, *J. Mol. Biol.* **2011**, *405*, 315–324.

[13] S. R. Schmidl, R. U. Sheth, A. Wu, J. J. Tabor, *ACS Synth. Biol.* **2014**, *3*, 820–831.

[14] N. C. Rockwell, D. Duanmu, S. S. Martin, C. Bachy, D. C. Price, D. Bhattacharya, A. Z. Worden, Lagarias, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 3871–3876.

[15] A. F. E. Hauck, S. J. O. Hardman, R. J. Kutta, G. M. Greetham, D. J. Heyes, N. S. Scrutton, *J. Biol. Chem.* **2014**, *289*, 17747–17757.

[16] S. J. O. Hardman, A. F. E. Hauck, I. P. Clark, D. J. Heyes, N. S. Scrutton, *Biophys. J.* **2014**, *107*, 2195–2203.

[17] A. Takiden, F. Velazquez-Escobar, J. Dragelj, A. L. Woelke, E.-W. Knapp, P. Piwowarski, F. Bart, P. Hildebrandt, M. A. Mroginski, *Photochem. Photobiol.* **2017**, *93*, 713–723.

[18] F. Velazquez Escobar, T. Utesch, R. Narikawa, M. Ikeuchi, M. A. Mroginski, W. Gärtner, P. Hildebrandt, *Biochemistry* **2013**, *52*, 4871–4880.

[19] F. Velazquez Escobar, P. Piwowarski, J. Salewski, N. Michael, M. Fernandez Lopez, A. Rupp, B. M. Qureshi, P. Scheerer, F. Bartl, N. Frankenberg-Dinkel, F. Siebert, M. A. Mroginski, P. Hildebrandt, *Nat. Chem.* **2015**, *7*, 423–430.

[20] T. Stensitzki, Y. Yang, A. L. Wölke, E.-W. Knapp, J. Hughes, M. A. Mroginski, K. Heyne, *Photochem. Photobiol.* **2017**, *93*, 703–712.

[21] C. Song, T. Rohmer, M. Tiersch, J. Zaanen, J. Hughes, J. Matysik, *Photochem. Photobiol.* **2013**, *89*, 259–273.

[22] E. Giraud, J. Fardoux, N. Fourrier, L. Hannibal, B. Genty, P. Bouyer, B. Dreyfus, A. Verméglio, *Nature* **2002**, *417*, 202–205.

[23] B. Karniol, R. D. Vierstra, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2807–2812.

[24] R. Tasler, T. Moises, N. Frankenberg-Dinkel, *FEBS J.* **2005**, *272*, 1927–1936.

[25] A. T. Ulijasz, G. Cornilescu, C. C. Cornilescu, J. Zhang, M. Rivera, J. L. Markley, R. D. Vierstra, *Nature* **2010**, *463*, 250–254.

[26] D. J. Heyes, B. Khara, M. Sakuma, S. J. O. Hardman, R. O'Cualain, S. E. J. Rigby, N. S. Scrutton, *PLoS ONE* **2012**, *7*, e52418.

[27] A. Bjorling, O. Berntsson, H. Lehtivuori, H. Takala, A. J. Hughes, M. Panman, M. Hoernke, S. Niebling, L. Henry, R. Henning, I. Kosheleva, V. Chukharev, N. V. Tkachenko, A. Menzel, G. Newby, D. Khakhulin, M. Wulff, J. A. Ihalainen, S. Westenhoff, *Sci. Adv.* **2016**, *2*, e1600920.

[28] G. M. Greetham, P. M. Donaldson, C. Nation, I. V. Sazanovich, I. P. Clark, D. J. Shaw, A. W. Parker, M. Towrie, *Appl. Spectrosc.* **2016**, *70*, 645–653.

[29] J. J. Snellenburg, S. P. Laptenok, R. Seger, K. M. Mullen, I. H. M. van Stokkum, *J. Stat. Softw.* **2012**, *49*, 1–22.

[30] E. R. Henry, J. Hofrichter, *Methods Enzymol.* **1992**, *210*, 129–192.

Manuscript received: January 9, 2018
Accepted manuscript online: March 15, 2018
Version of record online: April 20, 2018

Abstract

Reversibly switchable fluorescent proteins (RSFPs) are able to reversibly toggle between a fluorescent *on*-state and a non-fluorescent *off*-state under visible light irradiation. This property makes them a suitable marker used in super-resolution microscopy (or nanoscopy). The photo-switching mechanism involves isomerisation of the chromophore and a change of its protonation state. This mechanism has been well studied but remains poorly understood. The structural nature and the sequential order of atomistic events are still under debate. Some of them take place on the ultra-fast time scale and make structural investigation by X-ray crystallography impossible using current synchrotron radiation sources whose temporal resolution they offer is limited. X-ray free electron lasers (XFELs) are a new kind of X-ray source producing femtosecond pulses that allow structural investigation of ultra-fast intermediates during photoswitching. They are also so bright that crystallographic data collection from micro- and nanometer-sized crystals became possible. The bright and short XFEL pulses required a new methodology to be developed, the so-called serial crystallography methodology. This method is now being adapted to synchrotron radiation facilities.

Here is presented a time-resolved crystallography study of the reversibly switchable green fluorescent protein 2 (rsEGFP2). A microcrystallization protocol is described allowing the preparation of suitable samples in large amounts for time-resolved serial crystallography experiments. A photoswitching mechanism of rsEGFP2 is proposed based on crystallographic results obtained from data collected at the two XFEL facilities currently fully operational, *i.e.* the LCLS in the USA and SACLA in Japan. In particular, the structure of two photoswitching intermediates have been determined, one featuring a twisted chromophore in the excited state and the other displaying a protonated *cis* isomer of the chromophore in the ground state. The structural characterization of rsEGFP2 variants by traditional oscillation crystallography combined with the serendipitous discovery of an alternate chromophore conformation in the *off*-state during an XFEL experiment provided unique insight into the photophysical behavior of the protein.

Résumé

Les protéines fluorescentes photocommutables (RSFPs) ont la propriété de passer d'un état fluorescent à un état non-fluorescent en réponse à la lumière. Cette propriété en fait des outils de marquage pour la microscopie de super-résolution (ou nanoscopie). Le mécanisme de photocommutation implique l'isomérisation du chromophore ainsi qu'un changement d'état de protonation de ce dernier. Le mécanisme a été très étudié par différentes approches de spectroscopie et de simulation mais reste encore mal compris, l'ordre séquentiel des évènements est notamment encore débattu. Certains de ces évènements de la photocommutation se déroulent à des échelles de temps très courtes, ce qui rend difficile l'étude structurale par cristallographie des rayons X à l'aide des sources synchrotron actuelles dont la résolution temporelle est encore limitée. Les lasers à électrons libres (XFELs) sont une nouvelle source de rayons X produisant des impulsions suffisamment courtes pour permettre l'étude structurale des intermédiaires précoces ou à courte durée de vie qui se forment ou cours de la photocommutation, et suffisamment brillantes pour permettre la collecte de données cristallographiques sur des cristaux de tailles nano- et micrométrique. L'utilisation de ce nouveau genre d'instrument a permis l'émergence de la cristallographie sérielle, une nouvelle approche de la cristallographie des rayons X. Cette approche a depuis été adaptée aux lignes synchrotrons.

Le travail présenté ici se focalise sur l'étude de rsEGFP2, une protéine fluorescente photocommutable de la famille de la GFP. Il y est décrit la mise au point d'un protocole de microcristallisation permettant l'obtention d'échantillons en vue d'une expérience de cristallographie résolue en temps au XFEL. Un mécanisme de photocommutation y est proposé à travers le résultat de deux expériences sur les deux XFELs actuellement opérationnels, à des échelles de temps différentes, dévoilant un chromophore « twisté » à l'état excité ainsi qu'un état *cis* protoné de ce dernier. La caractérisation structurale des variants de rsEGFP2 par cristallographie d'oscillation « classique » combinée à la découverte fortuite d'une conformation alternée du chromophore dans l'état non-fluorescent, issue d'expérience de cristallographie sérielle, apporte un complément d'explication des propriétés photophysiques de la protéine.