A detailed study of the rsEGFP2 photodynamics and its variants using time-resolved optical spectroscopies

Lucas Martinez Uriarte

Université de Lille

Ecole doctorale Sciences de la Matière, du Rayonnement et de l'environnement

LAboratoire de Spectroscopie pour les Interactions, la Réactivité et l'Environnement

Thèse préparée pour obtenir le grade de:

Docteur par l'université de Lille

Discipline : Optique et Lasers, Physico-Chimie, Atmosphère

Lucas Martinez Uriarte

A detailed study of the rsEGFP2 photodynamics and its variants using time-resolved optical spectroscopies.

Public defense *29 January 2021*

Raporteur:

Dr. Ilaria Testa	Professeur, KTH Royal Institute of Technology (Stockholm, Suède)
Dr. Pascale Changenet	Chargée de recherche CNRS, Ecole Polytechnique (Palaiseau, France)

Examinateurs:

Dr. Peter Dedecker	Professeur, KU Leuven (Leuven, Belgium)
Dr. Steve R. Meech	Professeur, University of east Anglia (Norwich, UK)
Dr. Jacques-Philippe Colletier	Directeur de recherche CNRS, CEA-Grenoble (Grenoble, France)
Dr. Cyril Ruckebusch	Professeur, Univ. Lille (Lille, France) **(President of the Jury)**

Directeur de thèse:

Dr. Michel Sliwa	Directeur de recherche CNRS, Univ. Lille (Lille, France)

Co-Directeur de thèse:

Dr. Martin Weik	Directeur de recherche CEA, CEA-Grenoble (Grenoble, France)

Université de Lille

Ecole doctorale Sciences de la Matière, du Rayonnement et de l'environnement

LAboratoire de Spectroscopie pour les Interactions, la Réactivité et l'Environnement

Thèse préparée pour obtenir le grade de:

Docteur par l'université de Lille

Discipline : Optique et Lasers, Physico-Chimie, Atmosphère

Lucas Martinez Uriarte

Etude de la photodynamique de mutants de la protéine fluorescente photo-commutable rsEGFP2 en utilisant des spectroscopies optiques résolues dans temps.

Soutenance publique le *29 janvier 2021*

Raporteur:

Dr. Ilaria Testa	Professeur, KTH Royal Institute of Technology (Stockholm, Suède)
Dr. Pascale Changenet	Chargée de recherche CNRS, Ecole Polytechnique (Palaiseau, France)

Examinateurs:

Dr. Peter Dedecker	Professeur, KU Leuven (Leuven, Belgium)
Dr. Steve R. Meech	Professeur, University of east Anglia (Norwich, UK)
Dr. Jacques-Philippe Colletier	Directeur de recherche CNRS, CEA-Grenoble (Grenoble, France)
Dr. Cyril Ruckebusch	Professeur, Univ. Lille (Lille, France) **(President de Jury)**

Directeur de thèse:

Dr. Michel Sliwa	Directeur de recherche CNRS, Univ. Lille (Lille, France)

Co-Directeur de thèse:

Dr. Martin Weik	Directeur de recherche CEA, CEA-Grenoble (Grenoble, France)

Acknowledgements

Working on this PhD has been an overall enjoyable experience, and the past three years have been very profitable in many aspects. Therefore, I would like to thank all the people that had directly or indirectly contributed to make it possible to finish this project in a comfortable and fun manner.

Firstly, I would like to thank Michel Sliwa, who over the past three years has managed to understand my weaknesses, fears and strengths and therefore made me work in sometimes stressful conditions but always enjoyable. I'm very thankful also for his knowledge and scientific discussions of the project at any time or period of the day. Finally, I need to thank him for his patience and assistance with many administrative aspects, and his effort correcting the manuscript, which have definitively improved a lot with his inputs. I will never forget the first time I met him, and I am pleased to have done the PhD under his supervision and to have him as a friend.

I would also like to thank Martin Weik and all the collaborators from the IBS for their tremendous contributions. First, without their work in developing the new protein variants, this work could not have been possible. Second, I would like to thanks them for the inputs and TR-SFX data. From the IBS, I especially want to thanks Kyprianos Hadjidemetriou and Jacques-Philippe Colletier (member of the Jury).

From the LASIRE, I need to thank Cyril Ruckebusch and Rafaelle Vitale for sharing the office and discussions in data analysis, and Rafaelle Vitale for verifying Chapter 1. To Olivier Devos for the help with TCSPC measurements and fluorescent measurements, and to all of them for the pleasant moments in the lab. I would also like to thanks Aude Bouchet for her friendship and time both in and outside the lab and verifying Chapter 2. I would not want to forget all the PhD students and members of the LASIRE lab that had helped me and accompanied me all over the past years.

A special thanks to all the people that have hosted me in their labs. Amongst them all, I especially want to thank Jiro Abe and Hiroshi Miyasaka, and their lab members. A specific sincere thanks to Ayako Tokunaga and Katsuya Mutoh from Tokyo and Hikaru Sotome,

Tatsuhiro Nagasaka, and Masafumi Koga from Osaka, who I am lucky to consider my friends. Thanks to all of you for showing me your culture and country. Also, thanks to Remi Metivier from the PPSM lab in Paris for his time and setups. I must say that going to Paris was always the best excuse to visit my lovely brother Quique, I thank you for your support and love.

I would also like to thanks Steve R. Meech, and Peter Dedecker for being in the "thesis monitoring committee" which helped to find essential steps and advances in the project, including the experiments during the beamtime of Steve R Meech in the RAL and his help in the interpretation of the data. Finally, to the booth of then and Ilaria Testa and Pascale Changenet for being part of the Jury.

Personally, I wouldn't like to forget my many flatmates who have supported me, and my stinky smell from cycling 13 km in my way back from the lab. From my first share flat, thanks to Marin Copper someone with an immense heart, to Anthony Campisano for his friendship, time and for introducing me to his friends, Marie Douilliez for all the fun times together and Celia Brus for the lovely spirit and her smile. From my second shared flat, I would especially want to thank Herrmann Zimmerman, someone whom I can consider almost a brother. Also thanks to Ornella Candusso, Isabelle Reissberg and Nicolas Hollebecq. From my current flat thanks to Marina Abarquero and Leslie Martin. A special thanks to all my friends, especially to Patricia Schulze and Javier Jaspe. Finally, thanks to my parents Emilio and Estrella, I love you both. Thanks to the Sheriff's and Certucha's families. To end, a special mention to Laura Mahoney for her support, for verifying Chapter 3 and 7, for her time and love.

A mis seres queridos.

Aunque no entiendan de fotodinámica química.

Abstract

Reversible photoswitchable fluorescent proteins (RSFPs) are fluorescent proteins (FP) that can be reversibly toggled back and forward between a fluorescent on-state and a non-fluorescent off-state and thus allow to achieve super-resolution in fluorescence microscopy (e.g., in reversible saturable optical fluorescence transition – RESOLFT – microscopy). Even though their photo-physical parameters (switching and fluorescence quantum yield…) are linked to the image resolution and the image acquisition speed, the switching mechanism that controls these parameters is still a matter of debate. This thesis is focused on elucidating the photodynamics of rsEGFP2, a negative RSFP variant of the aqua victoria green fluorescent protein (avGFP). The rsEGFP2 is currently the reference fluorescent marker in RESOLFT microscopy. The *Off* to *On* switching involves a *trans*-to-*cis* isomerization and a proton transfer. It was previously highlighted that isomerization dynamics is characterized by a twisted chromophore which is formed at the picosecond time scale and restricted by the close proximity to the Valine 151. The mutation of Valine 151 into alanine (V151A) and leucine (V151L) showed that two different off-conformers exist. Their origin is presumably from a hula-twist and a one-bond-flip *On* to *Off* switching mechanism for V151L and V151A, respectively. In this thesis, we employed electronic and vibrational time-resolved absorption spectroscopy from the femtosecond to the minute time scales to study the photodynamics of wild-type rsEGPF2, V151A and V151L. These experiments were combined with the results of time-resolved crystallography obtained by collaborative groups. These two approaches permitted to infer the photo-switching mechanism of rsEGFP2 and its variants. *Off* to *On* photo-switching quantum yields of 11, 12 and 14% were estimated for WT, V151L and V151A, respectively. Such small differences were rationalized hypothesizing a common trans-to-cis isomerization via a sub-picosecond hula-twist mechanism, followed by microsecond preceding a sub-millisecond-scale multi-step deprotonation. Besides, the thesis also coped with the on-to-off dynamics of 20 other variants of rsEGFP2. From this study, it resulted that fluorescence and switching yield are controlled by the existence of at least two different ground states exhibiting a difference in the fluorescence lifetime of one order of magnitude (150 ps vs 2.3 ns). Overall, the outcomes of these studies will not only contribute to a better understanding of the photophysics of RSFPs but will also open newer perspectives towards the design of optimized RSFPs for advanced bio-imaging application.

Résumé

Les protéines fluorescentes photocommutables réversibles (RSFP), qui peuvent être commutées de manière réversible entre un état fluorescent (forme *On*) et un état non fluorescent (forme *Off*), sont maintenant couramment utilisées dans les microscopies de fluorescence super-résolues. Leurs propriétés et caractéristiques photo-physiques (brillance, rendements quantiques de commutations et de fluorescence…) sont liées aux paramètres tels que la résolution et la vitesse d'acquisition de l'image. Cependant le mécanisme et la dynamique de commutation qui sont à l'origine de ces paramètres font toujours débat. Cette thèse porte sur l'élucidation de la photo-dynamique d'une RSFP négative, la protéine rsEGFP2, une protéine couramment utilisée dans les techniques de microscopie de fluorescence super-résolue. Des études antérieures ont montré que le processus de commutation *Off* vers *On* est un processus séquentiel d'isomérisation trans-vers-cis du chromophore à l'état excité suivi d'un transfert de proton à l'état fondamental. De plus un chromophore « twisté » se forme à l'échelle de temps de la picoseconde avec une dynamique contrainte par la proximité de la valine 151. La mutation de cette dernière en alanine (V151A) et leucine (V151L) conduit à l'existence de deux conformères différents pour les formes *Off*. Au cours de la thèse nous avons utilisé la spectroscopie d'absorption transitoire électronique et vibrationnelle de la femtoseconde jusqu'à la minute pour l'étude de la protéine sauvage, V151A et V151L. Nos résultats, combinés à ceux obtenus par cristallographie par nos collaborateurs, ont permis de proposer un mécanisme de photo-commutation *Off* vers *On* pour ces trois protéines. Plus particulièrement il a été montré que le rendement quantique de photo-commutation *Off* vers *On* est similaire, 11, 12 et 14% pour la protéine sauvage, V151L et V151A. Cette faible différence pour des formes off différentes a été rationalisée par l'existence d'un mécanisme d'isomérisation identique de type « hula-twist » avec un temps caractéristique sub-picoseconde suivie d'un réarrangement structural microseconde de la protéine à l'état fondamental et d'une déprotonation milliseconde. La dynamique de commutation *On* vers *Off* a été aussi étudiée pour 20 variants de la protéine rsEGP2. Les résultats montrent que les rendements quantiques de fluorescence et de commutation sont contrôlés par l'existence d'au moins deux états dans l'état fondamental qui sont caractérisés par des temps de vie de fluorescence très différents (150 ps et 2,3 ns). Les résultats de cette thèse devraient non seulement contribuer à la compréhension de la photo-dynamique des RSFPs mais aussi permettre de concevoir de nouvelles protéines optimisées pour la bio-imagerie.

Glossary and abbreviations

ANR: Agence National de la Recherche

BP: Bicycle Pedal isomerization

CCD: Coupled Charge Device

CI: Conical Intersection

DAS: Decay Associated Spectra

DMSO: Dimethyl sulfoxide

EGFP: Enhanced Green Fluorescent Protein

ES: Excited State

ESA: Excited State Absorption

ESIPT: Excited-State Intramolecular Proton Transfer

ESPT: Excited State Proton Transfer

ESRF: European Synchrotron Radiation Facility

FLCS: Fluorescence Lifetime Correlation Spectroscopy

FP: Fluorescent protein

FPDB: Fluorescent Protein Data Base

FTIR: Fourier Transform Infrared Spectroscopy

GS: Ground State

GSB: Ground State Bleaching

GSC: Ground State Conformer

GSPT: Ground State Proton Transfer

GUI: Grafical User Interface

GVD: Group Velocity Dispersion

HBDI: 4-(p-hydroxybenzylidene)5-imidazolinone, rsEGFP2 chromophore

HEPES: 2-[4-(2-hydroxyethyl)piperazin-1-yl] ethanesulfonic acid

HOMO: Highest Occupied Molecular Orbital

HT: Hula Twist

IBS: Institute de Biologie Structural, Grenoble

IC: Internal Conversion

IR: Infrared

ISC: Intersystem Crossing

IVR: Internal Vibrational Redistribution

LCLS: Linac Coherent Light Source, XFEL beam

LUMO: Lowest Unoccupied Molecular Orbital

MCR-ALS: Multi Curve Resolution Alternating Least Squares

NA: Numerical Aperture

NIR: Near Infrared

NMR: Nuclear Magnetic Resonance

OBF: One Bond Flip

***Off*-form:** trans-neutral chromophore conformation for rsEGFP2 and its variants

***Off*-state:** non-fluorescent state for rsEGFP2 and its variants

***On*-form:** cis-anionic chromophore conformation for rsEGFP2 and its variants

***On*-state:** fluorescent state for rsEGFP2 and its variants

PAFP: Photo-activatable Fluorescent Protein

PAINT: Points Accumulation for Imaging in Nanoscale Topography

PALM: Photo Activated Localization Microscopy

PC: Principal Component

PCA: Principal Component Analysis

PCB: Printed Circuit Board

PCFP: Photoconvertible Fluorescent Protein

PI: Principal Investigator

PYP: Photoactive Yellow Protein

RESOLFT: REversible Saturable OpticaL Fluorescence Transitions

RSFPs: Reversible Switchable Fluorescent Proteins

S_1: First singlet Excited State

SACLA: SPring-8 Angstrom Compact free electron LAser, XFEL beam

SBR: Single Bond rotation

SE: Stimulated Emission

SFX: Serial Femtosecond X-ray crystallography

SOFI: Super-resolution Optical Fluctuation Imaging

SPIDER: SParse Image DEconvolution and Reconstruction

SR: Super Resolution

STED: STimulated Emission Depletion

STORM: Stochastic Optical Reconstruction Microscopy

SVD: Singular Value Decomposition

T_1: First triplet Excited State

TCSPC: Time-Correlated Single Photon Counting

TR: Time-Resolved

TR-SFX: Time-Resolved Serial Femtosecond X-ray crystallography

TRMPS: Time Resolved Multiple Probe Spectroscopy

TRIR: Time resolved infrared spectroscopy

TRSR: Time Resolved Stimulated Raman

TRUV-Vis: Time Resolved Ultraviolet Visible spectroscopy

TSLIF: Two-step Laser-Induced Fluorescence

UV: UltraViolet

UV-Vis: UltraViolet Visible

V151A: rsEGFP2 with valine 151 mutated to alanine

V151L: rsEGFP2 with valine 151 mutated to leucine

VR: Vibrational Relaxation

WT: Wild Type rsEGFP2

XFEL: X-ray Free-Electron Lasers

avGFP: Aqua-Victoria Green Fluorescent Protein

bR: bacterioRhodopsin

fs: femtosecond

ms: millisecond

ns: nanosecond

ps: picosecond

rsEGFP2: Reversible Enhanced Green Fluorescent Protein two

µs: microsecond

Table of contents

1 Introduction

The interest of this thesis in the "reversible enhanced green fluorescent protein two" (rsEGFP2) is beyond the elucidation of the ultrafast photodynamics of another photoswitchable protein. rsEGFP2 is a fluorescent protein (FP) which is nowadays widely used as probe for bio-imaging in fluorescence microscopy, especially for nanoscopy. At first glance, the relation between the ultrafast photodynamics of the protein and its use as a marker for fluorescence imaging might not be obvious. Yet, these two aspects are correlated. To understand this relation, it is necessary to first explain the interest of the scientific community in FPs. This interest mainly results from the fact that FPs allow both live-cell imaging and single-molecule microscopy, which bring us into a new era of direct visualization of biological processes and molecular dynamics in real-time.

Throughout the 19th century, several important findings, such as the discovery of the diastase enzyme[1], the first explanation of a biochemical process (i.e. the alcoholic fermentation[2]), or the synthesis of "urea" (the first organic compound being synthesized outside of a living body[3]) have set the basis of biochemistry. New explanations for the different processes ongoing in living cells started to rise during the first half of the 20th century. Some of the most outstanding discoveries include those of DNA and RNA[4,5], of the first antibiotic – penicillin – by Alexander Fleming, and of proteins for which Theodor Svedberg received the Nobel prize in 1926[6]. By the early '50s, several complex biological processes such as the photosynthesis and the generation of ATP in mitochondria were explained. During the second half of the 20th century, the development of several basic techniques like chromatography, electrophoresis or centrifugation, together with more complex analytical spectroscopic techniques such as electron microscopy, nuclear magnetic resonance and x-ray crystallography – all supported by a massive development in computer science – gave rise to modern biochemistry. Thanks to all of these techniques, several breakthroughs in biochemistry were achieved. The most remarkable ones might be considered i. the revolution in gene expression through the work of Francis Crick, Severo Ochoa or Har Gobind Khorana amongst others[6], ii. the elucidation of previously unimaginable aspects of cellular metabolism like the Krebs cycle[7], and iii. the cloning of the first living being[8]. By the late

‘80s, thousands of genes were sequenced and stored in the gene data bank GenBank[9], several living organisms were cloned, hundreds of metabolic routes were described, and thousands of protein structures were elucidated and stored in protein the data bank (PDB)[10]. Nowadays, all these milestones represent biochemistry subfields known as genomics, genetic engineering, metabolomics and proteomics, respectively. Despite the enormous advances in these different biochemistry subfields, several biological processes occurring inside cells that are nowadays well studied, such as gene expression or cell division were still unclear or poorly described. This was mainly due to the absence of a tool that would allow scientists to monitor living organisms with sufficient spatial and temporal resolution.

During the biochemistry revolution which took place throughout the second half of the 20th century, in the early ‘60s, a fluorescent protein was discovered in jellyfish[11]. As found out later, this protein is able to emit a photon thanks to an energy transfer from another protein[12] whose excitation is produced by a Ca^{2+} reaction (bioluminescence). This protein was named the Aqua-Victoria green fluorescent protein (avGFP or GFP). Thanks to the advances achieved in genetics, the protein sequence was determined[13], and it was demonstrated that avGFP could be cloned[14] and that other organisms could express it, preserving its fluorescence properties[15,16]. This observation marked the beginning of its use in molecular biology and biochemistry. The strong interest of the scientific community in avGFP and the advances in genetic engineering and mutagenesis significantly contributed to the design and development of new types of fluorescent proteins[17]. Today, many other fluorescent and non-fluorescent GFP homologues are reported. In Figure 1.1, the most important avGFP mutagenesis-derived proteins are represented. The newly produced fluorescent proteins obtained through mutagenesis, together with the discoveries shortly after of new ones in a variety of sea organisms, such as reef corals[18] and sea anemones[19], have rendered GFPs some of the most useful tools in cell biology. All this success culminated with the 2008 Nobel Prize in chemistry[20] awarded to Osamu Shimomura, Martin Chalfie and Roger Y. Tsien for the discovery of avGFP, the demonstration of its expression in other living organisms and the development and understanding of the GFP proteins respectively.

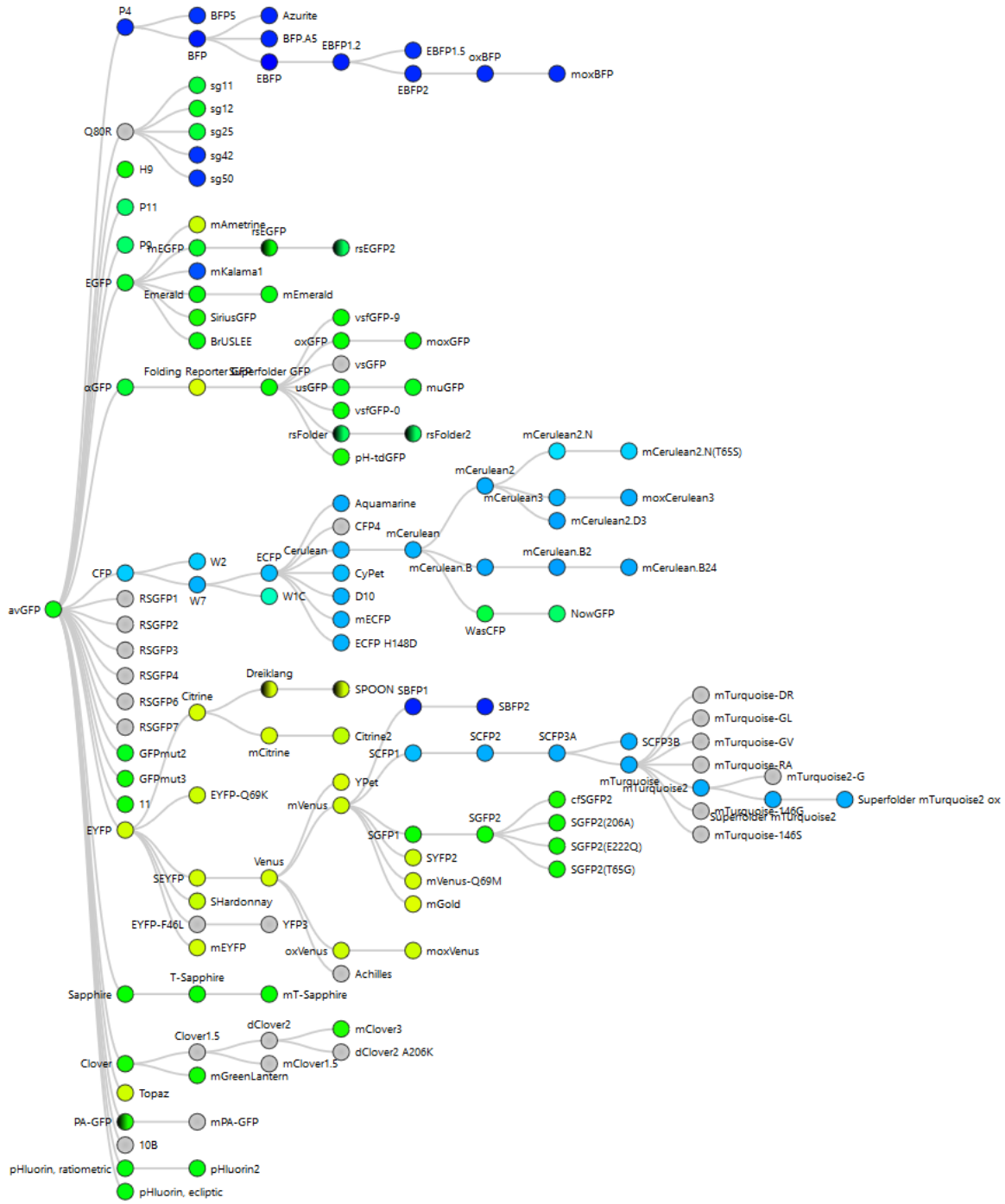


Figure 1.1. Chart representing the main fluorescent proteins that have been genetically engineered from avGFP. The colour of the corresponding circles (Blue, Cyan, Green and Yellow) represents the range of fluorescent emission wavelengths. A grey circle represents a non-fluorescent protein, and a half-coloured circle represents a reversible switchable fluorescent protein (RSFP). Figure reproduced from the fluorescent protein data base.

GFP-like proteins are nowadays extensively used in biochemistry[21]. These proteins are usually utilized in sensors for the determination of pH and Ca^{2+} concentration, crucial features in living cells. They also allow spatial and temporal monitoring of an increasing number of phenomena, such as cellular transport, cell division, gene expression, protein localization, protein dynamics, protein-protein interactions, organelle inheritance and biogenesis. Even the molecular infection dynamics of viruses can be followed using GFP labeling[22]. Among the wide variety of applications of GFP-like proteins, the most important one is undoubtedly their use as markers in fluorescence microscopy[23]. While most small fluorescent molecules are toxic for living cells, fluorescent proteins such as GFPs are usually less harmful when tagging living cells. This distinctive property makes GFPs ideal markers for living cell imaging[24].

Furthermore, the fluorescence from single GFPs, together with the development of reversible switchable fluorescent proteins (RSFPs)[25], has permitted to monitor biological processes and molecular dynamics not only in real-time but also at a spatial resolution higher than the optical diffraction limit in what is generally known as super-resolution fluorescence microscopy.

Towards the end of the nineteenth century, Ernst Abbe[26] and Lord Rayleigh[27] formulated what is known as the optical "diffraction limit" in microscopy (Equation 1.1). The diffraction limit describes the impossibility to resolve two adjacent elements closer than about half the wavelength (λ) in the lateral (x,y) plane and even further apart in the longitudinal plane (z):

$$d = \frac{\lambda}{2n\,Sin\theta} = \frac{\lambda}{2NA} \qquad \text{Equation 1.1}$$

where λ is the wavelength of the incident light, *n* the refractive index of the medium being imaged and Θ the incident angle. $n\ Sin\ \Theta$ is also known as numerical aperture (NA). Assuming NA = 1 and considering an emission wavelength of 400 nm (blue) and 750 nm (red), the maximum lateral resolution achievable by a fluorescence microscope would be between 200 nm and 325 nm. In Figure 1.2, a comparison scheme between the diffraction limit and some small organic and biological structures is represented. In the 20th century, the resolution in microscopy was increased by developing improved objectives and microscopes (featuring, for example, higher NAs). In the 21st century and 100 years after the formulation

of Abbe's theory, on October 8th 2014, Eric Betzig, W.E. Moerner and Stefan Hell were awarded the Nobel Prize in chemistry for "the development of super-resolved fluorescence microscopy"[28]. These techniques exploit specific fluorescent markers and allow taking images with a resolution exceeding the aforementioned diffraction limit. Super-resolution fluorescence microscopy (also called "nanoscopy") is one of the most powerful tools for imaging biological structures[29], as it allows to visualize ongoing process within living cells with enough spatial resolution[30]. Together with electron microscopy, super-resolution fluorescence microscopy approaches are essential tools for bioimaging and exhibit the advantage (compared to electron microscopy) of enabling life imaging.

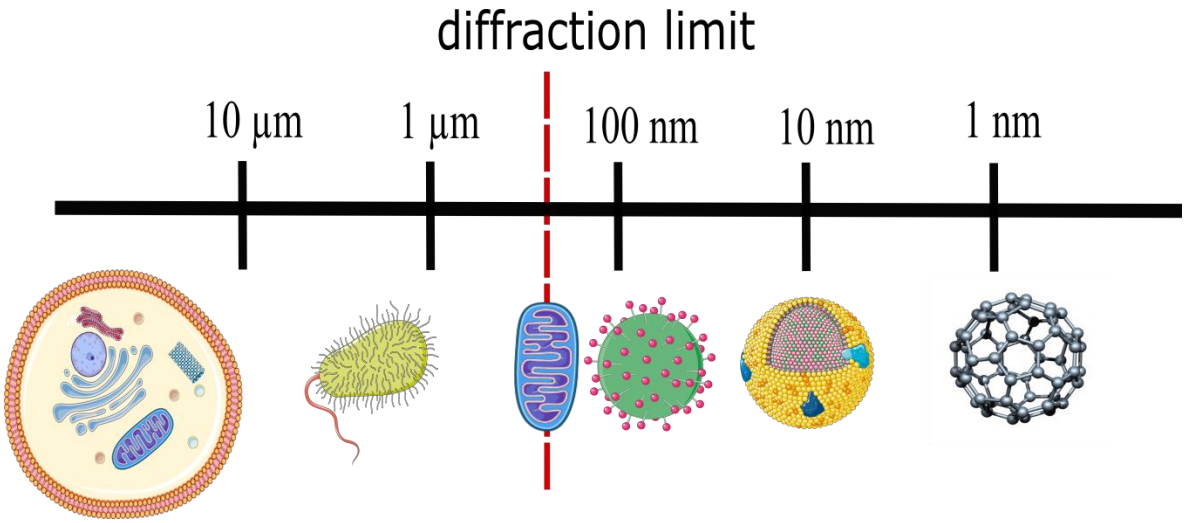


Figure 1.2 Lateral resolution diffraction limit for optical microscopy compared to some small organic and biological structures. From left to right: mammalian cell, a bacteria, a mitochondrion, a coronavirus unit, a lipoprotein, and fullerene C60. This scheme has been designed with ink scape, with some figures retrieved from https://smart.servier.com.

The super-resolution (SR) in fluorescent microscopy is based on fluorophores having bright (*On*) and dark (*Off*) emissive states and can be achieved mainly by two different approaches. The first method uses widefield fluorescence microscopes and is based on the localization of single fluorophores[31] known as super-resolved single-fluorophore microscopy. Some of these techniques are PALM[32], STORM[33] and PAINT[34]. The basic principle of these techniques is to collect a large number of images over time and isolate the emission of individual

fluorophores active within a diffraction-limited volume (1 emitter per μm^3) at different acquisition times which will be otherwise spatially overlapping fluorophores. The final super-resolved images are obtained by reconstruction of all the localized emitters. The strong limitation of such approaches is that they require the presence of only one active emitter in a diffraction-limited volume per image. This results in the need of acquiring large numbers of images to obtain enough fluorophore switching events for a detailed reconstruction (i.e. long acquisition time). The development of new algorithms and methods, such as SOFI[35] or SPIDER[36], that enable localization also in high emitter density conditions, has recently permitted to overcome this issue.

The second type of methods uses confocal microscopes. These techniques consist of physically deactivating the fluorescence near the focal point to obtain a subdiffraction limited spot. These methods are known as super-resolved ensemble fluorophore microscopy, with the most popular technique being STimulated Emission Depletion (STED[37,38]). Basically, the fluorophores are deactivated by a stimulated emission donut beam (Figure 1.3) before the spontaneous emission of a photon, and therefore high energy light sources are needed. Nowadays, thanks to the most advanced methods of this type of nanoscopy, it is possible to achieve up to 1-3 nm 3D multicolour resolution in cells[39].

More specifically, STED belongs to the category of REversible Saturable OpticaL Fluorescence Transitions (RESOLFT) techniques that are based on the ability of individual fluorophores to reversibly switch between an *On* (bright) and an *Off* (dark) state[25]. In STED, the excited and ground states are the bright and dark states. This is universal for all fluorophores, and the only disadvantage of STED is the use of high energy switching beams (stimulated emission donut-beam). However, there exist several types of markers that show intrinsic *On* and *Off* states. These are known as photochromic/photoswitchable fluorescent markers. With these markers, there is no competition of processes (stimulated vs spontaneous emission) when the fluorophore is photoswitched to its dark state, thus reducing the required energy to deactivate the fluorophore (Figure 1.3). These lower excitation energies yield the advantage of lowering the photo-damage of markers and biological tissues that highly intense STED beams can produce[25]. Reversibly photoswitchable fluorescent proteins (RSFPs), and, in particular, rsEGFP2 are the main markers used in RESOLFT. A RESOLFT microscope

setup is schematically depicted in Figure 1.3 and Figure 1.4 contains several RESOLFT images obtained with rsEGFP2 as a fluorescent marker.

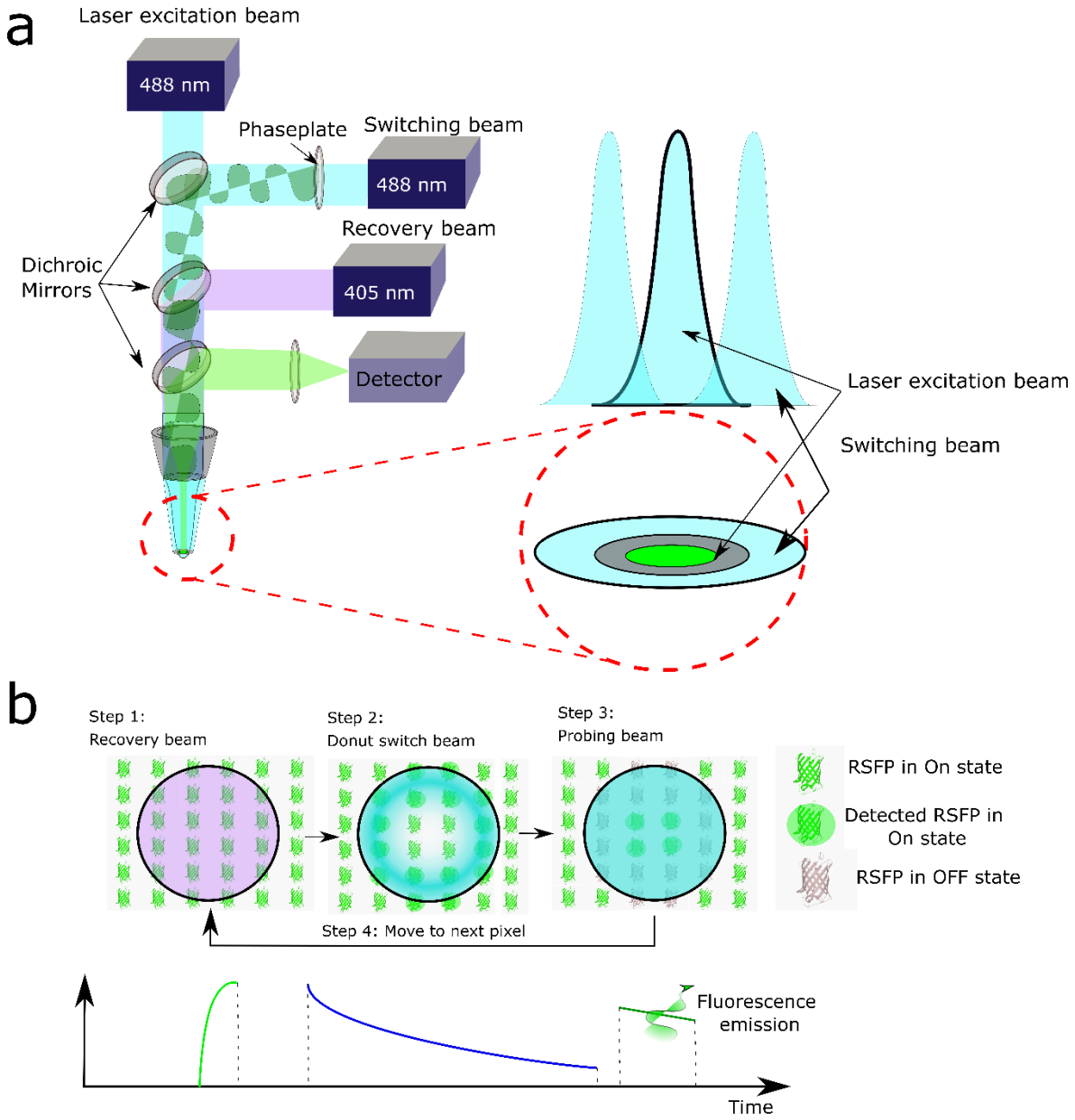


Figure 1.3. a) Example of RESOLFT experiments for a negative photoswitch absorbing at 488 nm b) The typical beam sequence. The duration of each step depends on the photophysical properties of the chosen protein/markers. Figure made with Inkscape.

RSFPs are conventionally classified according to their switching mode[40].For positive RSFPs the light used to get fluorescence induces the switching from the *Off* to the *On* state. There are only four RSFPs of this type; the best-known is Padron. For negative RSFPs, the same

light is used to switch from the *On* to *Off* state and produces fluorescence. This category includes the majority of RSFPs, including rsEGFP2, the protein studied in this thesis. Apart from these two classes of RSFPs, there exist Dreiklang and Spoon that show switching separate from excitation.

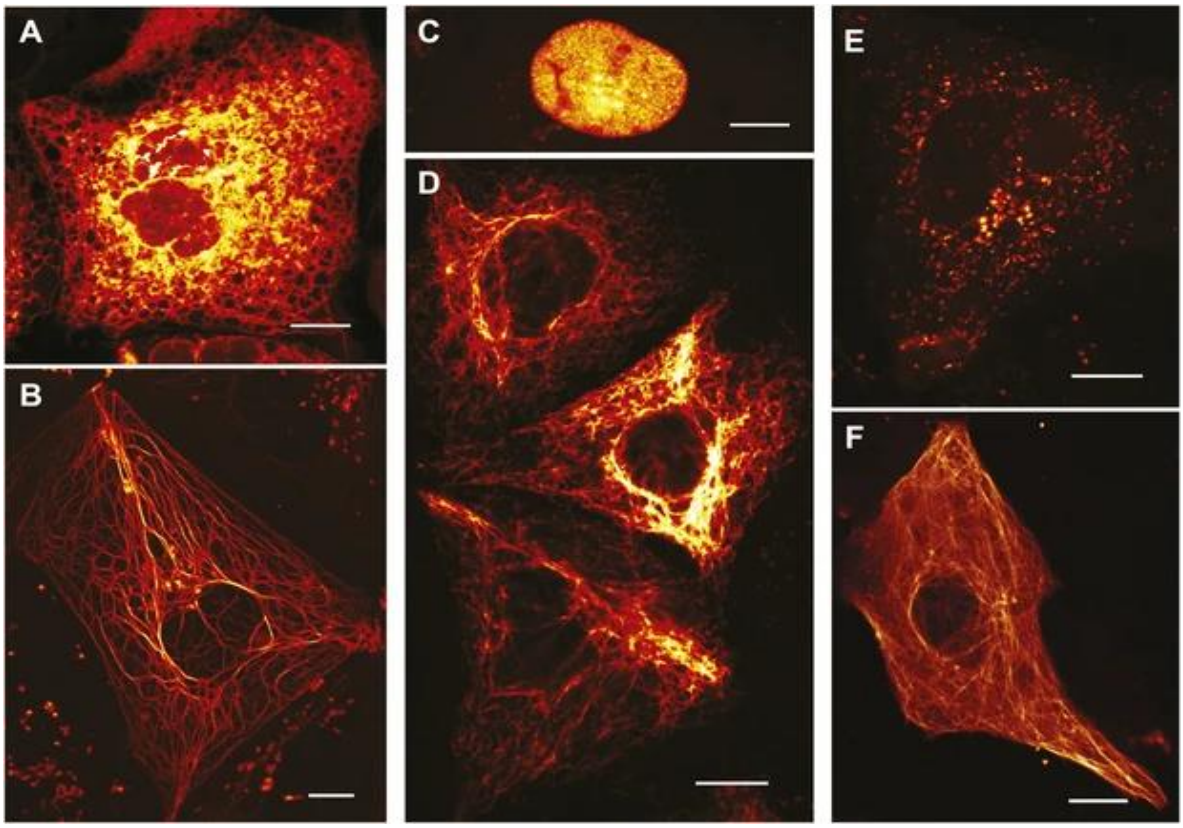


Figure1.4. Expression of various functional rsEGFP2 fusion proteins in mammalian cells. Scale bars: 10 μm. Reproduced from Grotjohann et al. [41]. (A–E): PtK2 cells; (F): Vero cell.

The different types of RSFPs will influence the image acquisition and the RESOLFT experiments[40,42]. As already mentioned, the photophysical parameters of these proteins control the resolution in RESOLFT microscopy. The most important ones are the brightness, the fatigue, the switching contrast and the switching speed which is linked to the switching yield and the cross-section of the *On* and *Off* state at the excitation and switching light. Some theoretical models based on Dronpa, which take into account the experimental acquisition parameters[43], have shown that one of the parameters that mostly affect the final resolution is the *On* to *Off* switching quantum yield: in general, higher quantum yields reduce the switching time, which is beneficial. On the contrary, high *On* to *Off* quantum yield values

lead to very low fluorescence counts since during the readout time the protein also switches to the *Off* state. *On* to *Off* quantum yields around 1% were determined to be optimal[43] for negative RSFPs. One of the advantages of using positive RSFPs is that the problem of the protein being turned *Off* during the readout phase is overcome since positive RSFPs are activated[42] (Figure 1.3).

In scanning microscopy techniques such as RESOLFT, proteins are switched between *On* to *Off* states several times. Therefore, fatigue and photobleaching are also fundamental parameters to consider[43]. As detailed in reference[40], one of the most critical parameters for RESOLFT is fatigue, remarkably rsEGFP2 displays excellent resistance to fatigue at high and low irradiation intensities, combined with excellent tagging capabilities, typical of EGFP derivates. This renders this protein to be one of the preferend fluorescent tags for RESOLFT nanoscopy[40,41]. In a few seconds, tens of images with a resolution below 50 nm can be collected for small fields of view[40]. Some very similar variants of rsEGFP2 with improved characteristics that had also shown promising results are rsFolder[44] and rsGreen[45], as well as the fast photoswitching mutant of Dronpa, Dronpa-M159T[46,47] (negative RSFPs). Despite the possibility of using different fluorescent proteins in RESOLFT, rsEGFP2 and rsEGFP[48] plasmid remain the only products sold by Aberrior instruments as markers for RESOLFT[49]. Some super-resolved images of different cells tagged with rsEGFP2 are displayed in Figure 1.4 (reproduced from Grotjohann et al. [41]).

In general, resolution in RESOLFT has many technical and challenging aspects, but it has been demonstrated that it is directly correlated with different photophysical parameters of the fluorophore used. And this is also the case for other super-resolved microscopy techniques. Indeed, both methods (SR single-fluorophore and SR ensemble fluorophore microscopy) are based on the dynamics of the fluorophores to deactivate into non-fluorescent states. Therefore, the photophysical properties of the markers (e.g. absorption and emission spectra, brightness, switching yield) will determine the maximal image resolution and will affect the acquisition parameters that should be used to get the desired super-resolved image (excitation light, acquisition time, number of pixels…). It is then clear that many aspects of nanoscopy techniques, including the spatial resolution that can be finally achieved, are directly related to the nature of the fluorophore used[43]. These photophysical properties can be tuned for FPs

by mutagenesis: even a single mutation can have huge effects on their photophysical characteristics[50]. The avGFP is constituted by more than 200 amino acids[13], and, thus, a rationalizing the design of the possible mutations is essential to finely tune such features. Therefore, it is fundamental to investigate the photophysical behavior of FPs, as this knowledge can provide essential hints on how mutants will behave and can deeply aid the design of new protein variants. The switching dynamics of RSFPs (except for Dreiklang and Spoon), mainly involves a *cis-trans* isomerization and a proton transfer. These photoreactions take place from the femtosecond to the millisecond time range after excitation and encompass different chromophore intermediates stabilized by the protein cage. The understanding of the role of the chromophore-protein interactions in the photodynamics of RSFPs can yield useful insights into possible new mutations and help with the design of novel optimized RSFPs for nanoimaging.

To this end, it is then needed to follow the protein photoreaction dynamics with femtosecond time resolution. Time-resolved spectroscopies are essential techniques for this kind of studies. However, monitoring in "slow-motion" how chemical bonds are broken, formed or isomerized during a photochemical reaction and how molecules behave after the interaction with light is not trivial as the primary events of a photochemical reaction occur in the femtosecond (fs) time range – 1fs is equivalent to 10^{-15} seconds. The area studying reactions at extremely short timescales is known as "*femtochemistry"*.The fundamental ideas of molecules adopting particular configurations throughout a reaction at very short times were introduced in 1935 by Evans, Polanyi[51], and Eyring[52] as transition states. They hypothesized that transition states might have lifetimes similar to the time scale of molecular vibrations, and thus, it would ever be possible to perform experiments over such short times to confirm them. Several years after, in 1946, George Porter came to the idea of using light pulses to generate free radicals and follow their underlying reactions with another continuous lamp spectroscopically. These techniques are known as pump- probe spectroscopy, the light pulse is the pump which excites the molecules to a higher energy state triggering the reaction, which is follow by the continuous lamp, the probe. The first flash photolysis apparatus was constructed in the following year by George Porter and Ronald Norrish, both published in 1949 the article "Chemical Reactions Produced by Very High Light Intensities" reporting the first-ever flashphotolysis experiment[53] by which reactions were monitored at the

microsecond scale resolution. In the same years, Manfred Eigen developed techniques capable of following chemical reactions at the nanosecond time scale[54]. Porter, Norrish and Eigen received the Nobel prize in Chemistry in 1967 "for their studies of extremely fast chemical reactions, affected by disturbing the equilibrium by means of very short pulses of energy"[55]. Nevertheless, all these techniques were then limited by the detector, i.e. the time of conversion of a photon into an electric signal (few nanoseconds). The maximum time scales resolution that can be achieved with these techniques are still far from the hypothetic lifetimes of transition states (except singlet-triplet state transitions).

In the late 1980s, a series of experiments performed by Ahmed H. Zewail and coworkers[56–60], gave birth to ultrafast spectroscopy and femtochemistry. They were able to capture for the first time molecular signals just in the transition state[56]. Their idea was to use a stroboscope probe light like the one used in a discotheque. Our eyes need about 30 ms to acquire an image. However, if a 1 millisecond short pulse is emitted by a stroboscope every 30 ms, the image captured by our eyes has a temporal resolution of 1 millisecond. This same principle is used in femtochemistry. Basically, two femtoseconds pulses which are in the time scale of molecular vibrations taking place between 10 and100 fs are exploited. Similar to the pump probe experiments introduced by Porter, Norrish and Eigen, there is flash pulse which excites the molecules to a higher energy state triggering the reaction (a femtosecond laser pulse), and a second femtosecond probe pulse (stroboscope pulse light) that arrives at a precise interval after the pump-pulse. The probe pulse captures a "picture" of the reaction with femtoseconds resolution. Using this approach, even if the detector takes few milliseconds to convert the signal (for example by a CCD camera) measured from the fs-probe, the final measurement has still a femtosecond resolution[58,60]. The crucial point is controlling with femtosecond precision, the arrival time of the probe and pump pulses. This can be done by an optical delay line. Indeed, a micrometer difference between the optical pathway of two beams translates into about ~3.3 fs time difference. The first reaction studied using this approach was the unimolecular disintegration of iodine cyanide into iodine atom and cyano radical[59]. Ahmed H. Zewail received the Nobel prize in Chemistry in 1999 "for his studies of the transition states of chemical reactions using femtosecond spectroscopy." [61]

The ensemble of techniques that use a laser pulse to excite a sample and induce changes which are then measured by a subsequent light pulse are known as pump-probe spectroscopic techniques and the most common one is time-resolved femtosecond transient absorption. This is the main technique that will be used in this thesis. The principles of these techniques will be review in Chapter 2. These approaches are generally classified according to the type of probe used (UV-Vis, IR), which, depending on its nature, will yield different information on the intermediate species. Only the combination of several of these techniques (UV-Vis, Raman, IR) allows building a precise scheme of the photoreactions that a molecule undergoes after light excitation. The example of bacteriorhodopsin (bR), which after excitation results in a complex photocycle, illustrates the aforementioned need of multiple time-resolved techniques. The photocycle, in fact, consists of at least seven intermediates of bR characterized by different visible transient absorption spectra and lifetimes ranging from femtoseconds to milliseconds. The longer-lifetime intermediates were studied using flashphotolysis techniques, which highlighted that the last reactive event occurs in a few ms[62]. The primary reaction steps of bR were investigated in the 1970s and to explain them, the bicycle pedal isomerization mechanism occurring in the ps time scale was proposed[63]. These studies were lately revised, and alternative isomerization pathways were proposed in 1985 by Liu et al. [64]. At the end of the 20th century, new results came out revealing the existence of some intermediates in the fs time scale[65]. However, the nature of some of these intermediates remained elusive since the interpretation of the measured time-resolved spectra and the assignment of their characteristic peaks to a specific structure is not always straightforward. It was only with the development of synchrotron and free-electron lasers (FELs) that the primary structure of the fs intermediates of the bR photocycle became observable through X-ray diffraction experiments with sub-nanosecond resolution[66,67]. Nevertheless, some questions have been recently raised since such experiments are conducted on microcrystals using high energy pulses to excite the samples[68]. Only the comparison of the outcomes from time-resolved spectroscopy in solution and crystallography for crystals, allows building precise photo-dynamical schemes for complex systems such as proteins.

Elucidating the structure and dynamics of RSFPs is a prerequisite for understanding their functioning[69]. As mentioned before, the most used techniques for investigating the RSFP

photodynamics are electronic time-resolved spectroscopy and, more specifically, femtosecond transient absorption UV-Vis-NIR (300 nm – 2000 nm) spectroscopy and nanosecond flash photolysis combined with emission decay measurements. The analysis of time-resolved spectroscopy data with advanced analysis tools yields the spectra and the kinetic time constants of a photodynamic system. This information is afterwards linked to atomic-scale structures that defined the molecular properties. Optimizing a property for a specific application is then an easier task since the interconnection between property and structure is known. This was the original idea of Dr. M. Weik when in 2013 he contacted Dr. M. Sliwa. His main objective was to create a French consortium promoting the combined use of time-resolved XFELs and ultrafast optical spectroscopy to determine (at the atomic scale and with temporal resolutions ranging from femtosecond to millisecond) the structures and kinetics of transient species with biological relevance (metastable photoproducts, excited states to name a few). They succeeded to obtain an ANR project, and I have been working within this project: ANR BioXFEL 2015-2022 (PI M. Weik IBS Grenoble, Co-PI M. Sliwa, Co-PI M. Cammarata U. Rennes, S. Boutet LCLS Stanford) elucidating structural intermediates of fluorescent proteins using X-ray free-electron lasers and ultra-fast UV-visible and infrared spectroscopies. Before the beginning of my PhD work, measurements at LCLS combined with transient UV-Visible spectroscopy made it possible to characterize for the first time the geometry of the precursor of the *cis-trans* isomerization of rsEGPF2[70]. These results opened the possibility of designing new RSFPs optimized for super-resolved fluorescence microscopy, and the photodynamics of these new variants was the principal topic of my thesis. The reasons behind the choice of rsEGPF2 and its variants will be detailed in Chapter 3, which positions this thesis with respect to previous findings. Throughout my PhD, I characterized their photo-stationary properties (Chapter 4) and inferred a photo-dynamical scheme ranging from the femtosecond to the millisecond time scale using time-resolved UV-Vis (LASIRE, Miyasaka Laboratory) and IR transient absorption (Rutherford Appleton Laboratory) for both the *Off* to *On* (chapter 5, article in preparation) and *On* to *Off* photodynamics (chapter 6, article in preparation). The details of the experiments are given in Chapter 2 and the materials and methods in Appendix 1. The combination of TR-SFX results (Grenoble) with those obtained by ultrafast optical spectroscopy experiments allowed to clarify the RSFP *Off* to *On* switching mechanism. During my PhD studies, I contributed to

the characterization of 19 rsEGFP2 variants that were designed by Dr. J.P. Colletier based on results from a TR-SFX experiment conducted on rsEGFP2 with the aim on understanding the fluorescence molecular bases (article in preparation). Additionally, I participated in an XFEL beam time in Japan (SACLA), where my main contribution was to in advance characterize the rsEGFP2 variants studied, allowing to propose time delays at which the different excited states would become visible to a TR-SFX delay allowing to build a rationalized photo-mechanical scheme for these new variants. Finally, the photodynamic studies of rsEGFP2 microcrystals, with different excitation powers, and robust analysis validate the TR-SFX measurements. I showed that the deprotonation time constants are faster in crystals than in solution (Nature Communcation 2020[71], second equal author). Finally, I also focused on the development of novel data analysis approaches. I have designed a new software named Ultra Pyfit to analyze time-resolved data in a robust fashion and interpret them also in the light of their intrinsic uncertainty (Appendix 2), with suggestions from Prof. C. Ruckebusch, a specialist in chemometrics. In summary, in my PhD thesis, I tried to combine state-of-the-art characterization approaches and advanced data analysis techniques in the attempt of achieving unprecedented insights into complex systems like RSFP.

1.1 References

(1) Payen, A., Persoz, J. F. Mémoire su la Diastase, les Principaux Produits de ses Réactions et leurs Applications Aux Arts Industrielles. *Ann. Chim.* **1833**, *53*, 73–92.

(2) Louis Pasteur: *Etudes Sur La Bière: Ses Maladies, Causes Qui Les Provoquent, Procédé Pour La Rendre Inaltérable, Avec Une Théorie Nouvelle de la Fermentation*; Masson et cie Editeurs: Paris, 1876.

(3) Wöhler, F. Ueber künstliche Bildung des Harnstoffs. *Ann. Phys. Chem.* **1828**, *88*, 253–256.

(4) Levene, P. A.; La Forge, F. B. On Chondrosamine. *Proceedings of the National Academy of Sciences of the United States of America* **1915**, *1*, 190–191.

(5) Levene P. A. The Structure of Yeast Nucleic Acid. IV. Ammonia Hydrolysis. *Journal of Biological Chemistry* **1919**, *40*.

(6) *From Nobel Lectures, Chemistry 1922-1941,* The Nobel Prize; Elsevier Publishing Company: Amsterdam, 1966.

(7) Krebs, H. A.; Johnson, W. A. Metabolism of ketonic acids in animal tissues. *The Biochemical journal* **1937**, *31*, 645–660.

(8) Cohen, S. N.; Chang, A. C.; Boyer, H. W.; Helling, R. B. Construction of biologically functional bacterial plasmids in vitro. *Proceedings of the National Academy of Sciences of the United States of America* **1973**, *70*, 3240–3244.
(9) Benson, D. A.; Cavanaugh, M.; Clark, K.; Karsch-Mizrachi, I.; Lipman, D. J.; Ostell, J.; Sayers, E. W. GenBank. *Nucleic acids research* **2013**, *41*, D36-42.
(10) https://www.rcsb.org/.
(11) SHIMOMURA, O.; JOHNSON, F. H.; SAIGA, Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *Journal of cellular and comparative physiology* **1962**, *59*, 223–239.
(12) Morin, J. G.; Hastings, J. W. Energy transfer in a bioluminescent system. *Journal of cellular physiology* **1971**, *77*, 313–318.
(13) Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. Primary structure of the Aequorea victoria green-fluorescent protein. *Gene* **1992**, *111*, 229–233.
(14) Heim, R.; Prasher, D. C.; Tsien, R. Y. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **1994**, *91*, 12501–12504.
(15) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science (New York, N.Y.)* **1994**, *263*, 802–805.
(16) Inouye, S.; Tsuji, F. I. Aequorea green fluorescent protein. *FEBS Letters* **1994**, *341*, 277–280.
(17) Tsien, R. Y. The green fluorescent protein. *Annual review of biochemistry* **1998**, *67*, 509–544.
(18) Matz, M. V.; Fradkov, A. F.; Labas, Y. A.; Savitsky, A. P.; Zaraisky, A. G.; Markelov, M. L.; Lukyanov, S. A. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature biotechnology* **1999**, *17*, 969–973.
(19) Wiedenmann, J.; Schenk, A.; Röcker, C.; Girod, A.; Spindler, K.-D.; Nienhaus, G. U. A far-red fluorescent protein with fast maturation and reduced oligomerization tendency from Entacmaea quadricolor (Anthozoa, Actinaria). *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 11646–11651.
(20) Måns Ehrenberg. *The green fluorescent protein: discovery, expression and development*, 2008, Uppsala.
(21) Zimmer, M. Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. *Chemical reviews* **2002**, *102*, 759–781.
(22) Oparka, K. J.; Roberts, A. G.; Cruz, S. S.; Boevink, P.; Prior, D. A.; Smallcombe, A. Using GFP to study virus invasion and spread in plant tissues. *Nature* **1997**, *388*, 401–402.
(23) SHIMOMURA, O. The discovery of aequorin and green fluorescent protein. *Journal of microscopy* **2005**, *217*, 1–15.
(24) Xie, X. S.; Choi, P. J.; Li, G.-W.; Lee, N. K.; Lia, G. Single-molecule approach to molecular biology in living bacterial cells. *Annual review of biophysics* **2008**, *37*, 417–444.
(25) 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. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 17565–17569.

(26) Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv f. mikrosk. Anatomie* **1873**, *9*, 418–440.
(27) Rayleigh. XV. On the theory of optical images, with special reference to the microscope. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1896**, *42*, 167–195.
(28) The Nobel Prize. *the development of super-resolved fluorescence microscopy*, 2014.
(29) Sahl, S. J.; Hell, S. W.; Jakobs, S. Fluorescence nanoscopy in cell biology. *Nature reviews. Molecular cell biology* **2017**, *18*, 685–701.
(30) Huang, B.; Babcock, H.; Zhuang, X. Breaking the diffraction barrier: super-resolution imaging of cells. *Cell* **2010**, *143*, 1047–1058.
(31) Moerner; Kador. Optical detection and spectroscopy of single molecules in a solid. *Physical review letters* **1989**, *62*, 2535–2538.
(32) Betzig, E.; Patterson, G. H.; Sougrat, R.; Lindwasser, O. W.; Olenych, S.; Bonifacino, J. S.; Davidson, M. W.; Lippincott-Schwartz, J.; Hess, H. F. Imaging intracellular fluorescent proteins at nanometer resolution. *Science (New York, N.Y.)* **2006**, *313*, 1642–1645.
(33) Rust, M. J.; Bates, M.; Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature methods* **2006**, *3*, 793–795.
(34) Sharonov, A.; Hochstrasser, R. M. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 18911–18916.
(35) Dertinger, T.; Colyer, R.; Iyer, G.; Weiss, S.; Enderlein, J. Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI). *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 22287–22292.
(36) Hugelier, S.; Rooi, J. J. de; Bernex, R.; Duwé, S.; Devos, O.; Sliwa, M.; Dedecker, P.; Eilers, P. H. C.; Ruckebusch, C. Sparse deconvolution of high-density super-resolution images. *Scientific reports* **2016**, *6*, 21413- 24975.
(37) Klar, T. A.; Jakobs, S.; Dyba, M.; Egner, A.; Hell, S. W. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 8206–8210.
(38) Hell, S. W.; Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics letters* **1994**, *19*, 780–782.
(39) Gwosch, K. C.; Pape, J. K.; Balzarotti, F.; Hoess, P.; Ellenberg, J.; Ries, J.; Hell, S. W. MINFLUX nanoscopy delivers 3D multicolor nanometer resolution in cells. *Nature methods* **2020**, *17*, 217–224.
(40) Jensen, N. A.; Jansen, I.; Kamper, M.; Jakobs, S. Reversibly Switchable Fluorescent Proteins for RESOLFT Nanoscopy. In *Nanoscale Photonic Imaging*; Salditt, T., Egner, A., Luke, D. R., Eds.; Topics in Applied Physics; Springer International Publishing: Cham, 2020; 241–261.
(41) Grotjohann, T.; Testa, I.; Reuss, M.; Brakemann, T.; Eggeling, C.; Hell, S. W.; Jakobs, S. rsEGFP2 enables fast RESOLFT nanoscopy of living cells. *eLife* **2012**, *1*, e00248.
(42) Konen, T.; Grotjohann, T.; Jansen, I.; Jensen, N.; Hell, S. W.; Jakobs, S. *The Positive Switching RSFP Padron2 Enables Live-Cell RESOLFT Nanoscopy Without Sequential Irradiation Steps*, 2020.
(43) Dedecker, P.; Hotta, J.; Flors, C.; Sliwa, M.; Uji-i, H.; Roeffaers, M. B. J.; Ando, R.; Mizuno, H.; Miyawaki, A.; Hofkens, J. Subdiffraction imaging through the selective donut-mode depletion of

thermally stable photoswitchable fluorophores: numerical analysis and application to the fluorescent protein Dronpa. *Journal of the American Chemical Society* **2007**, *129*, 16132–16141.
(44) 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. *Scientific reports* **2016**, *6*, 18459.
(45) Duwé, S.; Zitter, E. de; Gielen, V.; Moeyaert, B.; Vandenberg, W.; Grotjohann, T.; Clays, K.; Jakobs, S.; van Meervelt, L.; Dedecker, P. Expression-Enhanced Fluorescent Proteins Based on Enhanced Green Fluorescent Protein for Super-resolution Microscopy. *ACS nano* **2015**, *9*, 9528–9541.
(46) Böhm, U.; Hell, S. W.; Schmidt, R. 4Pi-RESOLFT nanoscopy. *Nature communications* **2016**, *7*, 10504.
(47) Testa, I.; Urban, N. T.; Jakobs, S.; Eggeling, C.; Willig, K. I.; Hell, S. W. Nanoscopy of living brain slices with low light levels. *Neuron* **2012**, *75*, 992–1000.
(48) Grotjohann, T.; Testa, I.; Leutenegger, M.; Bock, H.; Urban, N. T.; Lavoie-Cardinal, F.; Willig, K. I.; Eggeling, C.; Jakobs, S.; Hell, S. W. Diffraction-unlimited all-optical imaging and writing with a photochromic GFP. *Nature* **2011**, *478*, 204–208.
(49) https://www.abberior.com/jtl-shop/Abberior-expression-plasmids-pQE31-rsEGFP2-25-lg.
(50) Heim, R.; Tsien, R. Y. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology* **1996**, *6*, 178–182.
(51) Evans, M. G.; Polanyi, M. Some applications of the transition state method to the calculation of reaction velocities, especially in solution. *Transactions of the Faraday Society.* **1935**, *31*, 875.
(52) Eyring, H. The Activated Complex in Chemical Reactions. *The Journal of Chemical Physics* **1935**, *3*, 107–115.
(53) Norrish, R. G. W.; Porter, G. Chemical Reactions Produced by Very High Light Intensities. *Nature* **1949**, *164*, 658.
(54) Eigen, M. Methods for investigation of ionic reactions in aqueous solutions with half-times as short as 10 –9 sec. Application to neutralization and hydrolysis reactions. *Discussions of the Faraday Society.* **1954**, *17*, 194–205.
(55) The Nobel Prize in Chemistry 1967. https://www.nobelprize.org/prizes/chemistry/1967/summary/.
(56) Dantus, M.; Rosker, M. J.; Zewail, A. H. Real-time femtosecond probing of ‘‘transition states’’ in chemical reactions. *The Journal of Chemical Physics* **1987**, *87*, 2395–2397.
(57) Knee, J. L.; Khundkar, L. R.; Zewail, A. H. Picosecond photofragment spectroscopy. III. Vibrational predissociation of van der Waals’ clusters. *The Journal of Chemical Physics* **1987**, *87*, 115–127.
(58) Rosker, M. J.; Dantus, M.; Zewail, A. H. Femtosecond real-time probing of reactions. I. The technique. *The Journal of Chemical Physics* **1988**, *89*, 6113–6127.
(59) Khundkar, L. R.; Knee, J. L.; Zewail, A. H. Picosecond photofragment spectroscopy. I. Microcanonical state-to-state rates of the reaction NCNO→CN+NO. *The Journal of Chemical Physics* **1987**, *87*, 77–96.
(60) Zewail, A. H. Laser femtochemistry. *Science (New York, N.Y.)* **1988**, *242*, 1645–1653.

(61) The Nobel Prize in Chemistry 1999. https://www.nobelprize.org/prizes/chemistry/1999/summary/.
(62) Gillbro, T. Flash kinetic study of the last steps in the photoinduced reaction cycle of bacteriorhodopsin. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1978**, *504*, 175–186.
(63) Warshel, A. Bicycle-pedal model for the first step in the vision process. *Nature* **1976**, *260*, 679–683.
(64) Liu, R. S.; Asato, A. E. The primary process of vision and the structure of bathorhodopsin: a mechanism for photoisomerization of polyenes. *Proceedings of the National Academy of Sciences of the United States of America* **1985**, *82*, 259–263.
(65) Mathies, R. A. Photons, femtoseconds and dipolar interactions: a molecular picture of the primary events in vision. *Novartis Foundation symposium* **1999**, *224*, 70-84; discussion 84-101.
(66) Nass Kovacs, G.; Colletier, J.-P.; Grünbein, M. L.; Yang, Y.; Stensitzki, T.; Batyuk, A.; Carbajo, S.; Doak, R. B.; Ehrenberg, D.; Foucar, L.*; et al.* Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin. *Nature communications* **2019**, *10*, 3177.
(67) Nogly, P.; Weinert, T.; James, D.; Carbajo, S.; Ozerov, D.; Furrer, A.; Gashi, D.; Borin, V.; Skopintsev, P.; Jaeger, K.*; et al.* Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser. *Science (New York, N.Y.)* **2018**, *361*.
(68) Miller, R. J. D.; Paré-Labrosse, O.; Sarracini, A.; Besaw, J. E. Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin in the multiphoton regime and biological relevance. *Nature communications* **2020**, *11*, 1240.
(69) Lin, C.-Y.; Romei, M. G.; Oltrogge, L. M.; Mathews, I. I.; Boxer, S. G. Unified Model for Photophysical and Electro-Optical Properties of Green Fluorescent Proteins. *Journal of the American Chemical Society* **2019**, *141*, 15250–15265.
(70) Coquelle, N.; Sliwa, M.; Woodhouse, J.; Schirò, G.; Adam, V.; Aquila, A.; Barends, T. R. M.; Boutet, S.; Byrdin, M.; Carbajo, S.*; et al.* Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nature chemistry* **2018**, *10*, 31–37.
(71) Woodhouse, J.; Nass Kovacs, G.; Coquelle, N.; Uriarte, L. M.; Adam, V.; Barends, T. R. M.; Byrdin, M.; La Mora, E. de; Bruce Doak, R.; Feliks, M.*; et al.* Photoswitching mechanism of a fluorescent protein revealed by time-resolved crystallography and transient absorption spectroscopy. *Nature communications* **2020**, *11*, 741.

2 Photodynamics of photoswitchable proteins: definitions and methodologies

2.1. Outline

The study of photoactive systems is usually done using different electromagnetic radiations. How these systems interact with or produce electromagnetic radiation is known as spectroscopy. Depending on the frequency domain of light, the different spectroscopies that exist will measure spectra (spectrometry) after light-system interaction. Four main phenomena are usually observed: (i). Absorption (which can lead to emission) (ii). Reflection (iii). Scattering and (iv) Diffraction[1]. Therefore, scientists have developed experimental techniques based on these phenomena to study the photo-dynamics of photoactive systems. One of the most common technique (also used in this thesis) is the use of UV-visible absorption spectroscopy. Reversible photoswitchable fluorescent proteins (RSFPs) interact with light by absorbance that leads to emission of light (fluorescence) and chemical reactions. Therefore, in the first section, we will give some basic principles of light absorption (further details can be found in more specialized books) and deactivation processes that occur after the absorption of a UV-Visible photon. This part is followed by some specific definitions regarding RSFPs. In the second part of this chapter, technical details on the different time-resolved methods that allow to measure the photo-dynamics of RSFP will be given. This part will focus on the description of the different methods, which are important to understand the results presented in the following chapters. All the details of the time-resolved experiments and data analysis methods are given in Appendix 1. Then, as the switching mechanism of photoswitchable fluorescent proteins is based on *cis-trans* isomerization and proton transfer, the third part of the chapter will be devoted to giving some general definitions and characteristics of these two processes.

2.2. Absorption of UV-Visible light and de-activation processes

2.2.1 Absorption of UV-Visible light

The absorption of electromagnetic radiation by chemical compounds or atoms is a physical process that brings the systems into an excited state. Depending on the electromagnetic radiation energy, the excited state produced in the atom or molecule differs (e.g. UV-visible, electronic excited state; Infrared, vibrational excited state), and thus the information retrieved from the system varies. There are different spectroscopic techniques used to study the different domains of the electromagnetic spectrum, from X-ray absorption spectroscopy dealing with X-ray to nuclear magnetic resonance or electron spin resonance dealing with radio-waves. In this chapter, we will focus on interactions with UV-Visible light.

When a chemical compound absorbs a visible or UV photon, an electron from a bonding or a non-bonding orbital is promoted to an unoccupied anti-bonding orbital. Generally, an electron from the first vibrational (ν=0) singlet spin multiplicity electronic ground state (S_0) is excited to an electronic excited state, most commonly the first electronic state (S_1). This transition occurs in 10^{-15} seconds without any movement of the nuclei[2]. This principle is known as the Franck-Condon principle.[2–4] For an electronic transition to be possible, the final excited state must have the same vibrational momentum as the ground state, and since nuclei movement is much slower, the S_0 and S_1 energy surfaces should have the same nuclear coordinate[3]. In other words, the final nuclei position in the excited state must be compatible with the original nuclei position in the ground state. From the quantum mechanics point of view, the Franck-Condon principle allows separating the wave function of a molecular state into an electronic component and a nuclear (vibrational) component. This is known as the Born-Oppenheimer approximation[5]. The probability of a molecule to end up in any particular vibrational level is proportional to the square overlap of the vibrational wavefunctions of the states involved in the transition.

Experimentally, the efficiency of light absorption at a wavelength by an absorbing species is characterized by the absorbance (A) that follows the Beer-Lambert law:

$$A = \varepsilon l C \qquad \text{Equation 2.1}$$

Where (ε) is the molar absorption coefficient, (l) is the optical path and (C) the concentration of absorbing species. The molar absorption coefficient quantifies the ability of a molecule to absorb light in a given medium and at a specific wavelength. Classically, molecular absorption of light can be described by considering the molecule as an oscillating dipole. A quantity called oscillator strength (f) can be used, which is related to the integral of the absorption band as follows:

$$f = 2303 \frac{mc_0^2}{\mathcal{N}_a \pi e^2 n} \int \varepsilon(\bar{\nu}) d\bar{\nu} = \frac{4.32 \times 10^{-9}}{n} \int \varepsilon(\bar{\nu}) d\bar{\nu} \qquad \text{Equation 2.2}$$

Where (m) is the electron mass, (c) the speed of light, ($\mathcal{N}_a$) is the Avogadro's number, (e) the electron charge (n) the diffraction index and (ε) is the molar absorption coefficient and ($\bar{\nu}$) is the wavenumber in cm^{-1}. In theoretical chemistry, the transition moment is defined to characterize the transition between an initial state and a final state. The transition moment represents the transient dipole resulting from the displacement of charges during the transition, represented by $\vec{\mu}$ in Equation 2.3, where Ψ_1 and Ψ_2 represent the wavefunctions of initial and final states. The oscillator strength (classical approach) and transition moment (quantum chemistry approach) are linked through the Einstein coefficient B_{12} (Equation 2.4).

$$B_{12} = \frac{2\pi}{3h^2} |\langle \Psi_1 | \vec{\mu} | \Psi_2 \rangle|^2 \qquad \text{Equation 2.3}$$

$$f = \frac{8\pi^2 m\nu}{3he^2} |\langle \Psi_1 | \vec{\mu} | \Psi_2 \rangle|^2 = \frac{mh\nu}{\pi e^2} B_{12} \qquad \text{Equation 2.4}$$

Qualitatively, the transition occurs from the lowest vibrational state of the ground state to the vibrational state of the excited state that resembles the most in terms of vibrational wavefunction.

Figure 2.1.a and 2.1.b represent the electronic transitions obtained by the absorption of UV-Visible photons in a diatomic molecule, where both the ground and first excited electronic states are represented using a Morse-like potential energy function.[6] At room temperature, the most populated vibrational level of the electronic ground state is ν=0 (Boltzmann distribution). Thus, the majority of absorption will occur from this level. For figure 2.1.a, the energy surface of the excited state is identical to the ground state. In this situation, the

electronic transitions do not generate any shift in the nuclei positions of the molecule. According to this diagram, the most probable transition is the S_0 (v=0) →S_1(v=0). On the contrary, for the case pictured in Figure 2.1.b, the potential energy surface of the excited state is quite different from the ground state. The electronic transitions from S_0 to S_1 arising from the Born-Oppenheimer approximation are "vertical" (Franck-Condon transition), thus the S_1(v=0) is not anymore the most populated vibrational level of the excited state in this case. The S_0 (v=0)→S_1(v=4) is now the most probable transition[2].

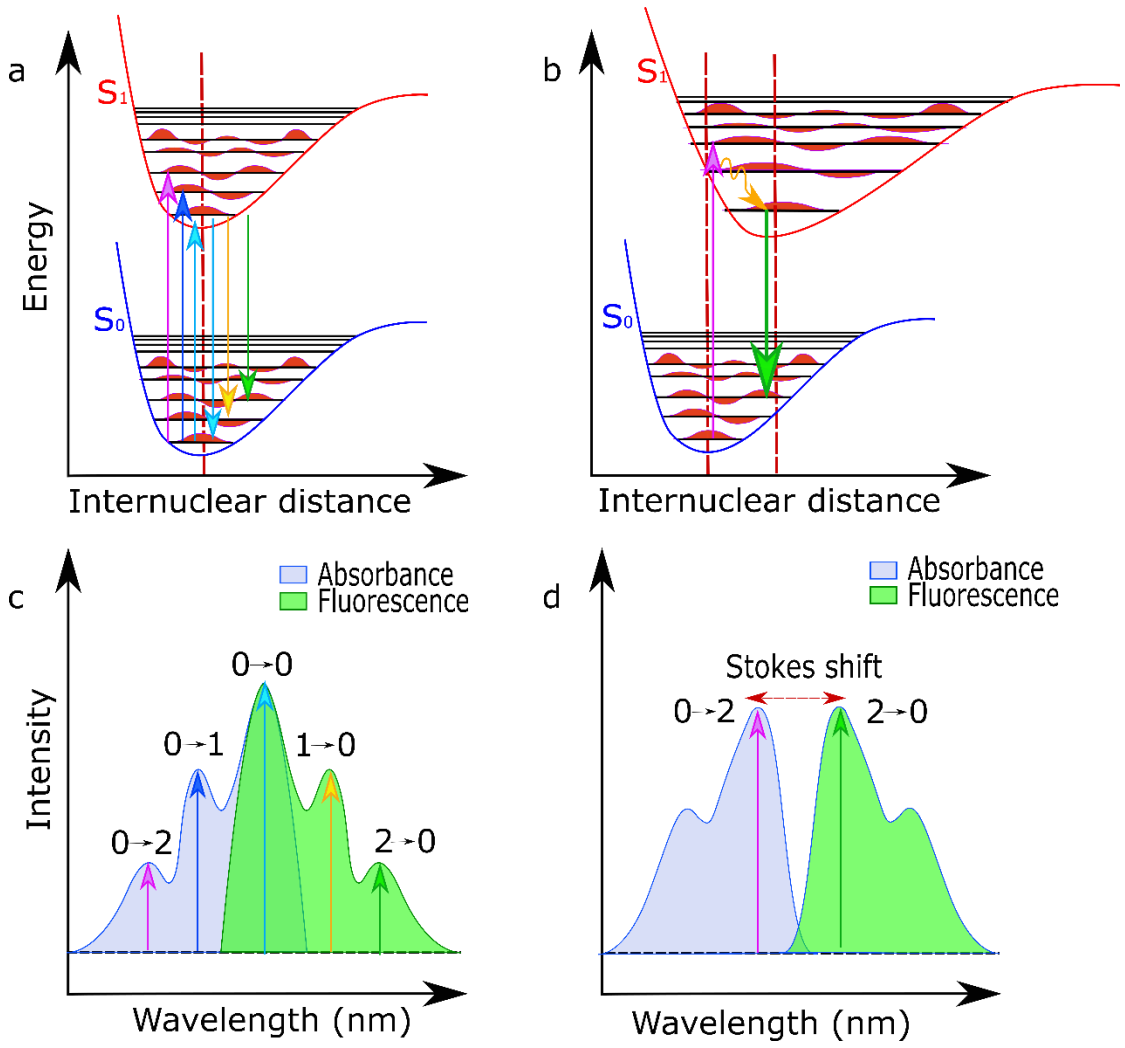


Figure 2.1. a and b. Internuclear coordinate energy diagrams for molecules with similar and different potential energy functions of the ground and excited states, represented by Morse potentials. c and d. Possible absorption (blue) and fluorescence (green) spectra that can be obtained from the systems represented in a and b. Figure made with Inkscape.

2.2.2 De-excitation processes without chemical reaction

The absorption of a UV-visible photon brings the molecule from the electronic ground state S_0 (v=0) to the first, second or higher electronic excited state that can be vibrationally excited (v>0), the molecule will then dissipate this extra energy. The de-activation processes (without any change of conformation or chemical reaction) that may occur can be described simply by the Perrin Jablonski diagram (Figure 2.2)[7,8].

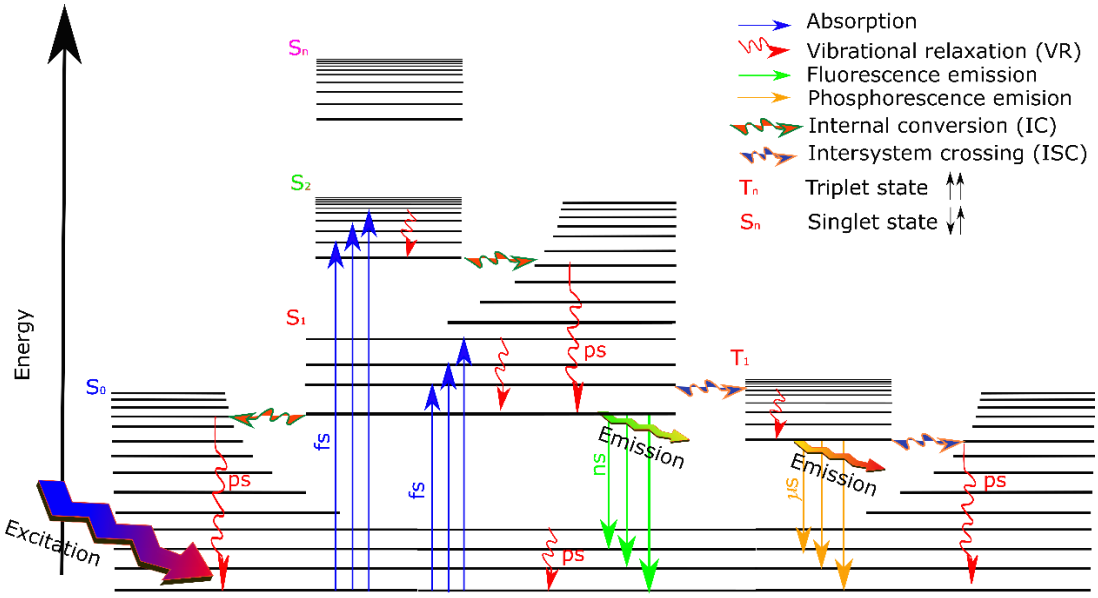


Figure 2.2. Perrin-Jablonski diagram with the approximate time constants associated to the different processes. Figure made with Inkscape.

The non-radiative deactivation processes are:

- Internal vibrational redistribution (IVR): this process (10^{-13}-10^{-12} s) occurs just after excitation where the extra vibrational energy can be redistributed along with the other vibrational modes of the molecule in the same electronic level.
- Vibrational relaxation (VR): this process is a relaxation of the vibrational energy (10^{-12}-10^{-10} s) by interaction with the solvent molecules, where the energy is transmitted to the surrounding molecules. This mechanism is also known as vibrational cooling.
- Internal conversion (IC): this process occurs when there is a vibronic coupling (nuclear motions) between two iso-energetically electronic levels with same spin multiplicity. This is a non-radiative process (10^{-11}-10^{-9} s) occurring within the same

time frame as vibrational relaxation. The time constant of IC is usually inversely dependent on the energy gap between the involved electronic levels (Kasha rule[9]).

- Intersystem crossing (ISC): this process occurs when there is a vibronic coupling (nuclear motions) between two iso-energetically electronic levels with different spin multiplicities. Such a transition involves a change in the spin of the electrons. The transition from the singlet excited state S_1 to a triplet excited state T_1 is a typical intersystem crossing. Such transitions are considered forbidden but allowed by spin-orbital coupling. A time constant about 10^{-10} to 10^{-8} s is usually found for organic molecules while for inorganic molecules, ISC time constant will be faster than IC.

Once the molecule has relaxed to the lowest vibrational level of S_1 (T_1), the molecules can have two different radiative deactivation processes that compete with those mentioned above.

- Spontaneous emission: this radiative process is a mechanism which dissipates the extra energy by emission of a photon. The emission of a photon occurs from S_1 to S_0 states is called fluorescence (same multiplicity of spin states) while from T_1 to S_0, the phosphorescence term should be used. Fluorescence is an allowed transition (10^{-10} to 10^{-7}s) whereas phosphorescence is a forbidden transition (10^{-6} to 1 s). The energy of the emitted photon is smaller than the energy of the photon absorbed by the molecule. This energy shift is known as Stokes shift (expressed in wavenumber)[10]. Besides the energy shift, the absorption and fluorescent emission spectra will look like mirror images because the differences between the vibrational level are the same in the excited and ground states (Figure 2.2). This is called the mirror image rule (Franck & Livingston 1941).[11] A difference of the fluorescence spectrum in comparison to the absorbance one can be assigned to different geometries and thus energy potential functions in the ground and excited states. This leads to a fluorescence spectrum that does not have the same vibrational structure.

  Regarding phosphorescence, we can note that the maximum of the phosphorescence emission spectrum is red-shifted in comparison to the fluorescence. The explanation is the VR occurring after the ISC to the triplet excited state. Therefore, in solution at room temperature, non-radiative de-excitation from the triplet state (small energy gap) is

predominant over phosphorescence. For this reason, phosphorescence is usually observed at low temperatures or in a rigid medium.

- Stimulated emission: the emission of fluorescence photons just described above is a spontaneous process. However, under certain conditions, induced emission can occur (e.g. LASER) and should also be described here as it is observed in femtosecond transient absorption spectroscopy. The comparison of induced and spontaneous emission can be made using the Einstein coefficient for a two-level system. We already introduced above the induced absorption coefficient B_{12} (Equation 2.3) to discuss the probability of absorbance of light[12]. Stimulated emission occurs when a molecule in the S_1 state is hit by a photon with an energy matching the difference between S_1 and S_0 states. The molecule then deactivates emitting one photon with the same energy that the one that hit it (same direction and wavelength, coherent photon). Stimulated emission plays an essential role in LASER[13], pump-probe absorption spectroscopy[14], as well as STED-microscopy[15,16,17] (see chapter 1). Stimulated emission process is not observed in conventional steady-state optical spectroscopy since the excited state is, at room temperature, much less populated than the electronic ground state (Boltzmann distribution). Considering the same degeneracy for the S_0 and S_1, Boltzmann statistics, and a radiation density that is expressed by Planck's black body law, there is a relation between stimulated (B_{21}) and spontaneous (A_{21}) coefficients that is given in Equation 2.5. In this equation, the frequency cube factor entails spontaneous emission to be the primary radiative process in the visible frequency domain.

$$A_{21} = \left(\frac{8\pi h}{c^3}\right) \cdot \upsilon^3 B_{21} \qquad \text{Equation 2.5}$$

2.2.3 Excited state lifetime and quantum yield

Different time constants were given above for the different processes (IC, ISC, fluorescence). Since we are dealing with fluorescent proteins, several definitions concerning the fluorescence process need to be clarified (lifetime and quantum yield). It should be noted that the emission of a photon is as fast as the absorption of a photon (10^{-15} s). However, the detection of photons is throughout the lifetime of the relaxed first singlet excited state $S_1(\nu=0)$

(similarly to the lifetime of radioactive species) which depends on all processes that affect the lifetime (IC, ISC). Considering a first-order monomolecular kinetic model for the excited state relaxation, we can formalize the above processes, in a first approximation, with the rate constant:

k_r^s: rate constant for radiative deactivation $S_1 \rightarrow S_0$ with fluorescence emission

k_{ic}^s: rate constant for internal conversion $S_1 \rightarrow S_0$

k_{isc}: rate constant for intersystem crossing

Therefore, the lifetime of an excited state A^* can be expressed by Equation 2.6:

$$-\frac{d[A^*]}{dt} = (k_r^s + k_{ic}^s + k_{isc})[A^*] = (k_r^s + k_{nr}^s)[A^*] \qquad \text{Equation 2.6}$$

where k_{nr}^s takes into account all non-radiative processes. After integration, the result is given by Equation 2.7:

$$[A^*] = [A^*]_0 exp\left(-\frac{t}{\tau_S}\right) \qquad \text{Equation 2.7}$$

where τ_s is the lifetime of A^* and is related to the rate constants by Equation 2.8:

$$\tau_S = \frac{1}{k_r^s + k_{nr}^s} \qquad \text{Equation 2.8}$$

An important parameter for fluorescent proteins is the fluorescent quantum yield. In general, any photo-induced quantum yield is defined as the number of defined events occurring per photon absorbed[18] (equation 2.9).

$$\Phi_f = \frac{Number\ of\ events}{Number\ of\ photons\ absorbed} \qquad \text{Equation 2.9}$$

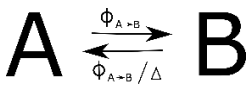
In fluorescence, the quantum yield (ϕ_f) the number of emitted photons divided by the total number of photons absorbed. The number of emitted photons is directly correlated to the number of molecules in the excited state that deactivate via fluorescence (Equation 2.10).

$$\Phi_f = \frac{k_r^s}{k_r^s + k_{nr}^s} \qquad \text{Equation 2.10}$$

2.2.4 Photoswitching quantum yield

Some definitions should also be given to characterize the switching properties of RSFPs. In the literature, the equilibrium between two different chemical forms (A and B) of the same compound that have two different absorption spectra is known as photochromic molecules. For a photochromic system with A and B states, the switching quantum yield from A to B is defined as the number of molecules that transform from A to B divided by the number of photons absorbed by A. Furthermore, the thermal equilibrium between A and B states implies that a non-radiative process for the transformation between A and B should be involved. Such a system has been schematically represented in the scheme 2.1. These parameters are essential for RSFPs.

Scheme 2.1. Simple representation of a photochromic system.



A crucial parameter is the so-called switching contrast which is defined as the ratio between the fluorescence signal in the *On*-state and the residual signal after *Off* switching. Such residual fluorescence after *Off* switching mainly originates from back switching of the O*ff* state chromophore by the *Off* switching light and the thermal back-reaction Equation 2.12 (More details regarding the intensity of absorption are given in chapter 4).

$$\frac{dC_{\text{off}}}{dt} = -\frac{dC_{on}}{dt} = \phi_{\text{on}\to\text{off}} I_{\text{on}}^{\text{abs}} - \phi_{\text{off}\to\text{on}} I_{C\text{off}}^{\text{abs}} - k_{\text{off}} C_{\text{off}} \qquad \text{Equation 2.12}$$

The switching contrast is directly linked to the photo-stationary state obtained after a certain irradiation time (∞), i.e. the concentration of *On* and *Off* state ($[\text{on}]^{\infty}$, $[\text{off}]^{\infty}$) at the thermodynamic equilibrium under specific irradiation wavelength. Generally, the thermal back-reaction for RSFPs (few minutes to hours) can be neglected in nanoscopy. The photo-stationary state or overall conversion efficiency (x^{∞}, defined in Equation 2.13) are then depending solely on the absorption cross-section (σ or molar absorption coefficient, ε) and switching quantum yield ($\Phi_{\text{on}\to\text{off}}$ and $\Phi_{\text{off}\to\text{on}}$) at the irradiation wavelength which follows Equation 2.14.

$$\chi^{\infty}_{\text{on}\rightarrow\text{off}} = \frac{[\text{on}]^{\infty}}{[\text{on}]^{\infty}+[\text{off}]^{\infty}} = \frac{\varepsilon_{\text{on}}\Phi_{\text{on}\rightarrow\text{off}}}{\varepsilon_{\text{off}}\Phi_{\text{off}\rightarrow on}+ \varepsilon_{\text{on}}\Phi_{\text{on}\rightarrow\text{off}}} \qquad \text{Equation 2.13}$$

$$\frac{[\text{on}]^{\infty}}{[\text{off}]^{\infty}} = \frac{\sigma_{\text{off}}\Phi_{\text{off}\rightarrow\text{on}}}{\sigma_{\text{on}}\Phi_{\text{on}\rightarrow\text{off}}} = \frac{\varepsilon_{\text{off}}\Phi_{\text{off}\rightarrow\text{on}}}{\varepsilon_{\text{on}}\Phi_{\text{on}\rightarrow\text{off}}} \qquad \text{Equation 2.14}$$

Therefore, one of the parameters that have been determined for rsEGFP2 and its variants are the switching yield and the photo-stationary state obtained after a certain irradiation time, i.e. the concentration of *On* and *Off* state ($[\text{on}]^{\infty}$, $[\text{off}]^{\infty}$) at the thermodynamic equilibrium under specific irradiation wavelength.

Other deactivation processes can occur, such as energy transfer, electron transfer or conformational changes. For each of these processes, one can define a rate kinetic constant and also a quantum yield. In Figure 2.3, the time scales of the different photophysical processes presented in this section are plotted together with the time resolution of the different time resolved (TR) techniques that allow studying the photodynamic of photoactive molecules. The techniques will be described in the next section of this chapter.

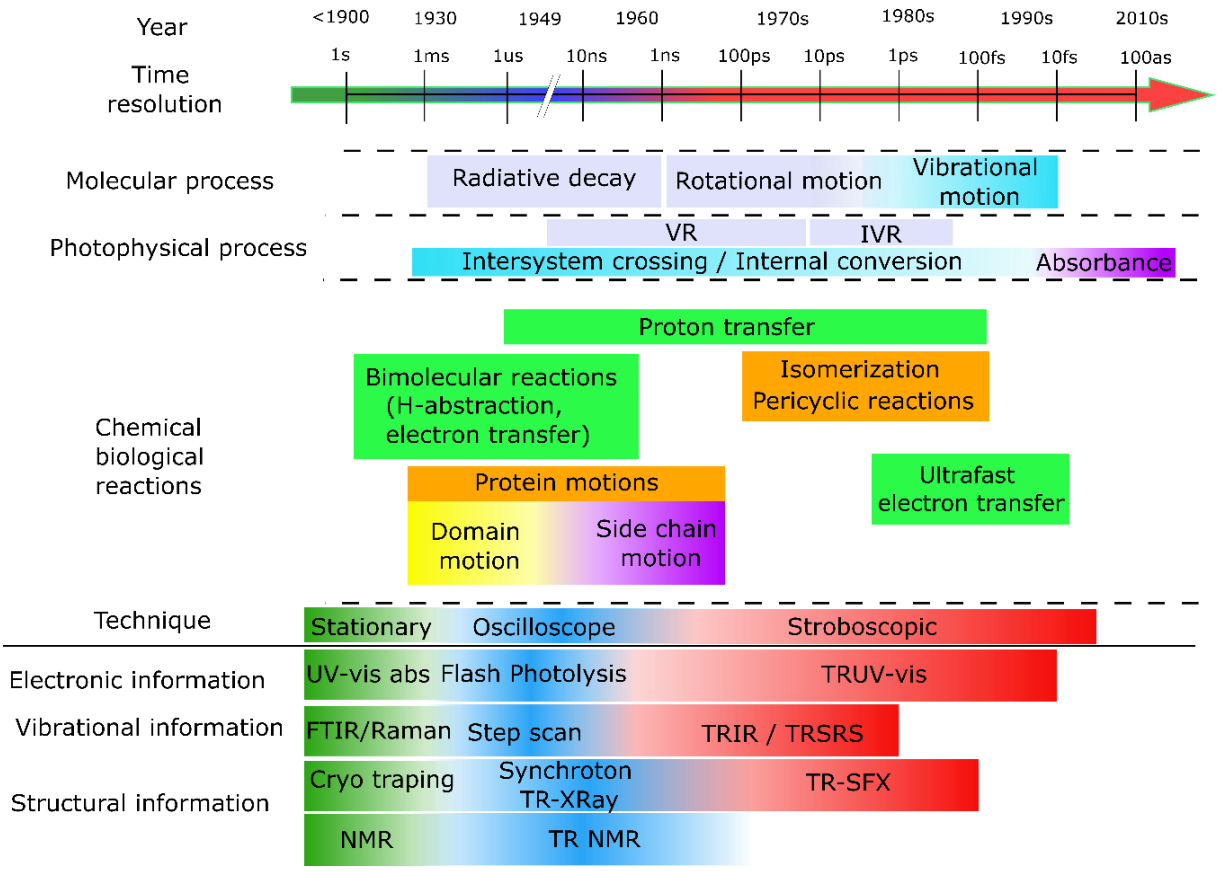


Figure 2.3. Time domains of photophysical processes and chemical reactions plotted together with the main time-resolved techniques used to study them. Figure made with Inkscape.

2.3 Time-resolved techniques for studying the photodynamics of photoswitchable proteins

The study of photochemical reactions has been an intense field of research where two main questions can be pointed out: how to follow ultrafast dynamics and how to characterize precisely the nature of intermediates. This last question is even more critical for proteins due to the influence of the protein cage (amino-acids) on the geometry of the photoactive chromophore. Chemical reactions occur via transient excited states or reaction intermediates which are short-living species that play a major role in the reactions. The goal is to build a

photo-dynamical scheme, i.e. to determine the nature of the different short-living transient species, their lifetime, their rate constant and the different pathways with their formation yield. The final target is then to find which parameters control the properties of interest (for example for rhodopsin the switching yield) in order to design efficient new bio-materials. The classical method is to use time-resolved spectroscopy which consists in measuring spectra while the photoreaction is occurring. Time-resolved spectroscopy techniques are widely used to investigate the properties of complex systems from proteins[19,20] to photochromic molecules[21] or metals complexes[22], among others. Two main milestones can be pointed out in time-resolved spectroscopies: the development of techniques that can follow dynamics in the nanosecond time scale[23,24], the limit of fast detectors and oscilloscopes[24], and the development of stroboscopic optical methods to probe sub-nanosecond dynamics (Figure 2.3). Both approaches will be explained in the following section.

Time-resolved techniques are often based on a pump pulse (shorter than the dynamics to be studied) that triggers photophysical processes and photochemical reactions, and a probe beam (continuous or pulsed, depending on the technique) that will analyze the dynamics of the systems after the photo-excitation. In general, the pump pulse should photoactivate a sufficient number of molecules/proteins while the probe pulse should analyze and not interfere in the dynamics, i.e. the energy per pulse is high for pump pulses while it is low for probe pulses[14]. Depending on the different properties that have to be investigated, the spectral domain of the probe beam differs. For example, a UV-vis probe gives information related to changes in electronic states, while IR gives information about vibrational changes. Recently, time-resolved crystallography which uses X-ray probe pulses, has allowed to obtain information on structural changes. In Figure 2.4, a general schematic representation of a stroboscopic time-resolved pump-probe experiment used in femtosecond time-resolved transient absorption can be seen, with the pump represented by the purple beam and white light continuum pulse as the probe beam. The probe beam is analyzing (spectrograph) the species excited by the pump beam (both beams are overlapping in the sample cuvette). As shown in Figure 2.3, the photo-dynamics of photoswitchable proteins will span several timescales from a few femtoseconds to several milliseconds where several excited and ground-state intermediates may exist transiently. Therefore, in order to study such a broad

timescale, different spectroscopic techniques with the appropriate temporal resolution are required. The resolution depends primarily on the temporal width of the pump pulse and secondarily on the electronics and width of the probe beam used to monitor the reaction[25]. The majority of electronic devices cannot measure signals faster than a few nanoseconds[26]. In the best cases, an excellent temporal resolution can be achieved until a few hundreds of picoseconds with gated streak cameras[27]. For cases where electronics can reach the desired temporal resolution (nanoseconds to milliseconds), the method developed for flash photolysis experiments can be used (probe beam is continuous, and an oscilloscope analyzes the time evolution of probe beam after excitation by the pump pulse). When electronics are not fast enough to transform the energy of a photon into an electronic signal that can be read, an alternative approach is needed. This can be summarized as a stroboscopic approach that takes pictures of the dynamics at certain delay after photo-excitation. This method is using a pulsed probe beam (time width equal or smaller than the desired time resolution) and a precise control of delay between pump and probe pulses (Figure 2.4). This is achieved using an optical delay line, the arrival time for pump and probe pulse are done by changing the optical path-length; 1 micrometer difference of optical path-length corresponds to 3 fs. Any detector can be used (photodiode, photomultiplier, camera) and the electronics do not limit the time resolution anymore since the electronic signal observed is coming from a pulsed probe beam (femtosecond, picosecond…) that is determining the time resolution of the picture. This technique afforded Prof. A. Zewail to get the Nobel Prize in 1999. He showed how to capture molecular spectra just in the transition state[28] with the use of powerful laser pulses of tens of femtoseconds which give birth to femtochemistry research field[14]. In each of the two approaches, the pump is generally either a nanosecond laser pulse for classical flash photolysis or a femtosecond pulse for the optical delay approach.

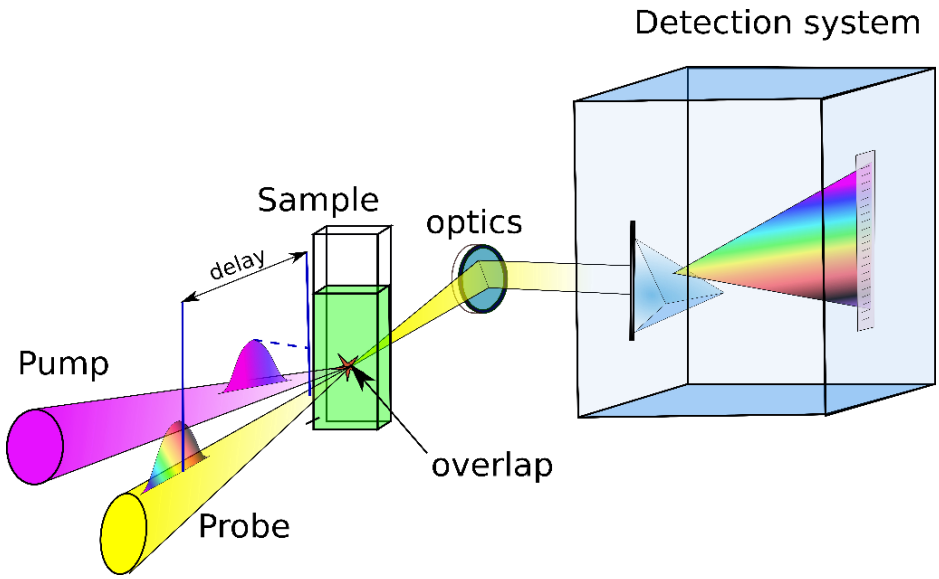


Figure 2. 4. Schematic representation of a pump-probe experiment given for femtosecond TRUV-vis. The probe pulse is a (fs) broad white-light continuum, and the detection system consists of some diffraction gratings and a detector in general, a spectrometer combined with a CCD camera. Figure made with Inkscape based on others[29].

Focusing on photo-active proteins, the study of structural dynamics was usually achieved using vibrational spectroscopies, i.e. Raman or IR, but recently the development of femtosecond X-ray pulse using a free electron laser (XFEL) allows to follow the dynamics of bR, PYP or myoglobin proteins using time-resolved crystallography with femtosecond time resolution. This powerful technique is however complex, and only the combination of time-resolved optical spectroscopy and crystallography affords to get a final photo-dynamical scheme. In the following section, we will give the details of the most common techniques that are used to study photoswitchable fluorescent proteins, i.e. time-resolved emission (section 2.3.1.a), time-resolved UV-Vis-NIR transient absorption (section 2.3.1.b), time-resolved IR (section 2.3.2) and time-resolved serial femtosecond X-ray crystallography (TR-SFX) (section 2.3.3). Finally, we will focus on the analysis of transient absorption experiments spectra, i.e. how to build a photo-dynamical scheme from their analysis and what are the limits (section 2.3.4). We divided time-resolved spectroscopy techniques on their energy domains and information obtained: UV-Vis time-resolved electronic spectroscopy, IR-Raman time-resolved vibrational spectroscopy and X-ray time-resolved crystallography (not spectroscopy).

2.3.1 Time-resolved electronic spectroscopy

2.3.1.a Time-resolved emission spectroscopy

In general, fluorescence lifetime for FP is about few nansoeconds[30]. Fluorescence is the most sensitive method (single protein lifetime can be measured) and is thus the best way to study excited state dynamics. Due to the spontaneity of the phenomena, there is no need to use probe beams to follow the fluorescent emission. The most common technique to measure fluorescence lifetimes is Time Correlated Single Photon Counting (TCSPC) technique. It measures the arrival time to the detector of a single photon emitted by fluorescence with respect to the time when the molecule is excited as a stopwatch would do it[25,31]. The detection of the arrival time of a single photon is done by the combination of a single photon counting card and detectors such as photomultipliers tubes or microchannel plates with intrinsically higher gain[25]. This technique is a statistical method that requires sufficient accumulated events (a few millions) to be able to fit the data and determine precisely a lifetime. Indeed, the resulting signal, which is counts versus time, is a histogram representing the statistical values of single photons arrival times to the detector where x represents the time. Each time a photon reaches the detector at a specific time, it increases the count value by one. Only one photon should be detected to be able to determine precisely the arrival time and thus no more than one single photon event per excitation[31]. With a high repetition rate laser excitation (MHz), millions of sequences can be measured in a few minutes. The precision of the clock for arrival time is about few picoseconds, and instrumental response function is about 30 ps for microchannel plate photomultiplier tube (MCP-PMT). Altogether using deconvolution, this technique affords to determine the lifetime of a few tens of picoseconds[25,31,32]. The practical details of the experiments are given in Appendix 1 (section A1.3).

The electronics inherent in TCSPC technique limit the measurement of lifetimes to a few tens of picoseconds. The principle of femtosecond fluorescence measurement is the same as in femtosecond transient absorption pump-probe spectroscopy. The problem of temporal resolution is overcome using stroboscopic technique, two femtoseconds optical pulses (reading and excitation) separated by a certain temporal delay will take a picture of the fluorescence response. The principle of femtosecond fluorescence is then to synchronize a reading "gate" femtosecond pulse with the fluorescence created by a femtosecond excitation

pulse and to be able to vary the delay between reading and excitation. Using pulses (excitation and reading) of a few tens of femtoseconds, it is in principle possible to catch the fluorescence for a time equal to the pulse width. By varying the delay, fluorescence kinetics are then reconstructed with femtosecond resolution. Two techniques have been developed for the reading gate pulses; both of them are based on a nonlinear phenomenon of wave mixing to temporally select the fluorescence signal. The main technique used is fluorescence upconversion: the optical gate is produced by mixing the frequencies of the fluorescence pulse and the gate pulse (generally the fundamental pulse of the Ti:Sa, 800 nm) in a non-linear second order crystal. A frequency sum pulse is created, which is proportional to the fluorescence intensity and the delay between the reading and excitation pulse[32].

2.3.1.b Time-resolved absorption spectroscopy

For time-resolved emission technique such as TCSPC, since the detection of a single photon can be achieved out of a non-fluorescent dark background, the signal to noise ratio is high. Time-resolved absorption is based on the measurement of absorption after excitation of the sample by the pump pulse. However, it is not possible to photo-convert 100% of proteins with the excitation of a single pulse. Indeed, only a few percent of the molecules are excited by the pump pulse (and it will be nonsense to use higher laser energies as these will induce non-linear processes such as photo-ionization). Therefore, the absorption spectra recorded after the pump pulse is composed of around 90 % of molecules which are not photo-activated. Consequently, the spectra before and after excitation are very similar. This difference is even smaller for longer decay times as photo-conversion yield does not exceed 50% of excited molecules (for example for bR). Therefore, to increase the clarity of the signals, the absorbance of the samples without excitation is subtracted. The spectrum without the excitation is measured at every time delay thus the difference of absorbance is calculated by measuring the absorbance with laser excitation (pump pulse) and by subtracting the absorbance spectrum without laser excitation measured immediately before as can be seen in equations 2.15.

$$\Delta Abs(\Delta t) = Abs_{\text{with pump}}(\Delta t) - Abs_{\text{without pump}}(\Delta t) \qquad \text{Equation 2.15}$$

As mentioned previously, there are two different techniques. For dynamics superior to few nanoseconds, time-resolved nanosecond transient absorption, known as flash photolysis experiments, is used. The pump pulse is a nanosecond pulse which is usually obtained from a Q-switch Nd:YAG laser (10 Hz, 5 ns, 1064 nm, 500 mJ). The visible pump is obtained using either nonlinear crystals (harmonics of Nd:YAG laser, 532, 355 and 266 nm) or an optical parametric oscillator (OPO) which covers all the UV-Vis domain[33]. For the probe beam, the most common light source is a Xenon lamp as it can cover from UV to NIR[33] regions (250-850 nm). The detection is done using a spectrometer coupled to a PMT (single wavelength measurements for different wavelengths to reconstruct the time-resolved different absorption spectra) or a gated intensified camera[27] (iCCD, measurement of difference time delay gates to reconstruct the time-resolved difference absorption spectra). For short time experiments (10 ns – 400 µs), the Xenon lamp can be pulsed to increase the signal to noise ratio. As mentioned above, these experiments consist of signal measurements with and without excitation with a second repetition rate (dynamics until a few tens of millisecond are measured), and the number of accumulations is usually low. Secondly, it should be stressed that the system should be back to its initial state before being re-excited by the pump laser[19]. This is usually not the case for RSFPs, therefore, the sample should be refreshed by stirring, flowing, or moving the sample cuvette. Classical characteristics for the experiments are an excitation pulse around 1-2 $mJ.cm^2$ and a protein solution having an absorbance about one at the excitation wavelength for an optical path of 1 cm. The number of accumulations is usually under 100.

For dynamics faster than few nanoseconds, pump and probe beams are femtosecond pulses. They are generally generated from self-mode-locking Ti: sapphire-based laser systems (details of our system are in Appendix 1) which were discovered in the early 1990s[34]. Such commercial laser systems can generate routinely 80-100 fs pulse widths. 30-40 fs and even 4-5 fs pulse widths are also available with specific designs[26]. Time and spectral widths are important parameters. Ultrashort pulses (< 50 femtoseconds) are used for coherent experiments because they have large spectral widths due to Heisenberg uncertainty principle, i.e. a 10 fs laser pulse at 800 nm has a spectral width of +/- 47 nm, while 100 fs pulses have narrow spectral width (+/- 4.7 nm). A compromise needs to be done between spectra and time widths[33]. The output from the Ti: sapphire-based lasers systems (oscillator + amplifier)

is generally peaking at 800 nm, the energy is about 1-5 mJ per pulse with a frequency about 1-5 kHz. This beam is the source of the pump and probe beams. The beam is divided in two, ~90% for the pump and ~10% for the probe beam. Each of the two beams should be converted into different wavelengths to excite photoactive systems: the pump, to a wavelength close to the maximum of the absorption spectrum of the sample and the probe into a white-light continuum beam to measure absorption spectra. Tuning the wavelength is usually done with an optical parametric amplifier (OPA, generation of a white-light continuum and amplification of one wavelength using a nonlinear crystal). The probe beam will be generated by focusing few micro-Joules of 800 nm into a 1-2 mm windows made of CaF_2, silica, or YAG. Indeed, high energy femtosecond pulses have a peak power that generates a continuum of wavelengths by nonlinear optical effects. For example, using a CaF_2 window a femtosecond white-light continuum from 350 to 800 nm is generated while using a YAG crystal, the continuum covers a spectral range between 450 nm to 2 micrometers.

Pump and probe arrival times between 0 to 2 nanoseconds are controlled by an optical delay line. The pump beam has an energy about 0.1-1 µJ (beam size 100-500 µm FWHM), and the sample has an absorbance about 0.1-0.5 for an optical path of about 0.1-1.0 mm. To avoid multiphoton ionization, a power density of few mJ/cm^2 is used. The pump-probe polarization configuration is usually set at the magic angle (54.7°) to avoid having a change of signal due to Brownian motion of the photoactive samples. The white light probe beam is then detected by a spectrometer coupled with a CCD camera or a photodiode array to calculate the difference absorption spectra. Similarly to flash photolysis experiments, the sample should be refreshed at kHz level, i.e. a flow cell is used to ensure it[35]. One time delay difference absorption spectrum needs about 1000 accumulations to get a signal below 1 milli-OD differential absorbance.

It should be noted that several laser centers offer measuring time-resolved absorption spectra from femtosecond to millisecond scale after a femtosecond excitation pulse. These systems are using a combination of different femtosecond laser systems that are electronically synchronized[36,37]. Such systems are known as time-resolved multi-probe spectroscopy (TRMPS).

2.3.2 Time-resolved vibrational spectroscopy (TRIR and TRSR)

TR-UVvis measurements can yield many hints regarding excited state dynamics and reaction pathways. However, the measured spectra mainly reveal information regarding the electronic transitions which are not specific to the protein cage and the interaction between the chromophore and the protein[38]. This makes it challenging to infer structural dynamics from such signals, except for simple organic molecules with few but clear structural changes[39].

On the contrary, the IR region of the spectra can give important vibrational information of the transient species. From these signals, structural information regarding the vibration motions that are important in the reaction coordinates can be deduced, indicating which bonds are broken or formed. Techniques that yield this type of information are IR and Raman spectroscopies[40]. With these techniques, small geometrical changes can easily be detected by vibrational frequency shifts. The time-resolved versions of these two techniques were initially limited to the nanosecond time scale (Step Scan and Time-resolved nanosecond resonance Raman). The development of femtosecond IR probe pulses, which can be obtained from nonlinear effects in a noncollinear optical parameter amplifier (NOPA)[26], were a great revolution in the 2000s[38], decreasing the time resolution to 200-300 femtoseconds.

TRIR spectroscopy

The measurement of time-resolved difference absorption spectra in the IR region is using a similar scheme as the UVvis TR (UV-Vis pump pulse generated by an OPA, flow cell, magic angle polarization). The differences are in the spectral width, i.e. IR probe beam has a spectral width of about 200 cm^{-1}, the detection part that is composed by a spectrometer and a detector which is composed of HgCdTe elements (usually 64 or 128 elements). One of the main drawbacks of TRIR spectroscopy is the high absorption of IR radiation by water molecules (solvent and moisture). The latter condition implies working with the thickness of a few tens of micrometer, exchanging H_2O to D_2O in water-based solvents and working under an inert atmosphere. Although TRIR measurements can yield crucial structural information, the molar absorption coefficient of IR transitions is lower than the UV-Vis one and higher concentrations of proteins (3 to 4 times) is needed to get at least 100 micro-absorbance signals. This can lead to some additional issues since high concentrated protein samples tend to aggregate and ultimately precipitate.

TRR spectroscopy

The second vibrational technique is time-resolved Raman spectroscopy. It was initially limited by a pulse width of the probe of hundreds of picoseconds. Notice that since the Raman peak shifts depend on the excitation wavelength, Raman experiments need narrow bandwidth laser (>10 cm^{-1} Fourier transform) to only have peak shifts originating from a single wavelength. As mentioned above, due to Heisenberg uncertainty principle, such a narrow pulse limits the temporal width to the picosecond range. This problem was overcome by time-resolved stimulated Raman (TRSR) experiments, which provide fundamental advantages compared to TRIR. This technique starts as the previous ones with a femtosecond pump pulse that triggers the photo-active protein[38]. The reaction is then probed at the desired time delay by two pulses that generate a stimulated Raman transition. These two pulses are a 2-3 picoseconds narrow-bandwidth Raman pulse and a 20 femtoseconds broadband continuum probe pulse[38]. This broadband pulse enables simultaneous probing of the vibrational features from around 700 to 2200 cm^{-1}. This technique allows recording vibrational structural information with a time resolution of 50 fs, which is similar to the vibrational periods of the probed bands[38]. The spectral resolution of TRSR is within 10 cm^{-1} [38]. One advantage of TRR compare TRIR is the absence of the water problem. Indeed, Raman transitions are linked to the polarizability of the chromophore. The bending modes of the water have a big influence in the dipolar moment (IR) but do not produce important changes on the polarizability, and as a consequence, H_2O bending mode Raman signal is not significantly important to obscure the sample signals. Thus, there is no need to use deuterated water samples. However, TRSR is not often used due to the complexity of the experiments (three beams) combined to the low signal intensity of Raman bands (long acquisition time) and unwanted fluorescence signal for fluorescent proteins (high background).

2.3.3 Time-resolved crystallography

None of the optical time-resolved techniques presented until now provide direct spatial or atomic information of the reaction intermediate structures. X-ray diffraction is a straightforward technique to obtain molecular structures of crystallized compounds. Some initial studies have determined the structures of the different photo-active forms of a protein, and the researchers have hypothesized the reaction pathway by comparing the initial and final

structures. Examples of such approaches can be found in the first structural studies of the RSFP dronpa[41]. X-ray diffraction has helped understanding many fundamental processes in biochemistry by determining the structure of proteins and other biomolecules[26]. Initial approaches to capture intermediate structures consisted of cryo-trapping. This technique consists of initializing the reaction by light and flash-cooling the crystal to try and capture an intermediate[42]. The first X-ray time-resolved crystallography experiments were developed several decades after the first flashphotolysis experiments[24]. High-intensity X-ray beam pulses are needed in order to observe small significant modifications in the observed diffracted patterns of the intermediates[26]. This high-intensity X-ray beams were available thanks to modern synchrotron facilities such as the Advanced Light Source (ALS) in Berkeley, USA or the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. By the end of the 1990s, two experiments on the photoactive yellow protein (PYP)[43] and myoglobine[44] were able to successfully capture intermediate structures in the millisecond and nanosecond timescales, respectively. Both studies used Laue crystallography and, and as pointed later by Gregory K. Farber[45], "synchrotron radiation has been the crucial ingredient in making the experiments work". Nearly a decade later, the studies on myoglobin were extended to the subnanosecond range also using synchrotron radiation[46]. This time resolution is the time-resolved limit that can be achieved at synchrotrons. As discussed above, the time resolution in pump-probe experiments is limited to the temporal length of the probe pulse, and 100-ps is the typical electron-pulse length at synchrotrons, preventing the investigation of ultrafast intermediates involved in photochemical processes. This problem was solved using femtosecond X-ray pulses generated by FEL.

X-ray free-electron lasers (XFEL) are so-called fourth generation light sources. The idea of XFEL came out at the beginning of the 1970s[47] and was patented in 1974[48]. Nevertheless, it took several decades for it to become a reality. The first XFEL available was the Linac Coherent Light Source (LCLS) at the Stanford Linear Accelerator Center (SLAC). The first publication by Emma et al[49] in 2010 reported that the LCLS delivered coherent soft and hard X-rays from 22 to 1.2 Å wavelengths, with a peak brightness of about ten orders of magnitude beyond conventional synchrotron sources. Furthermore, it was able to deliver a wide range of X-ray pulses durations from 5 to 500 fs. These outstanding beam characteristics made it possible to capture ultrafast intermediate structures and reaction dynamics at atomic level.

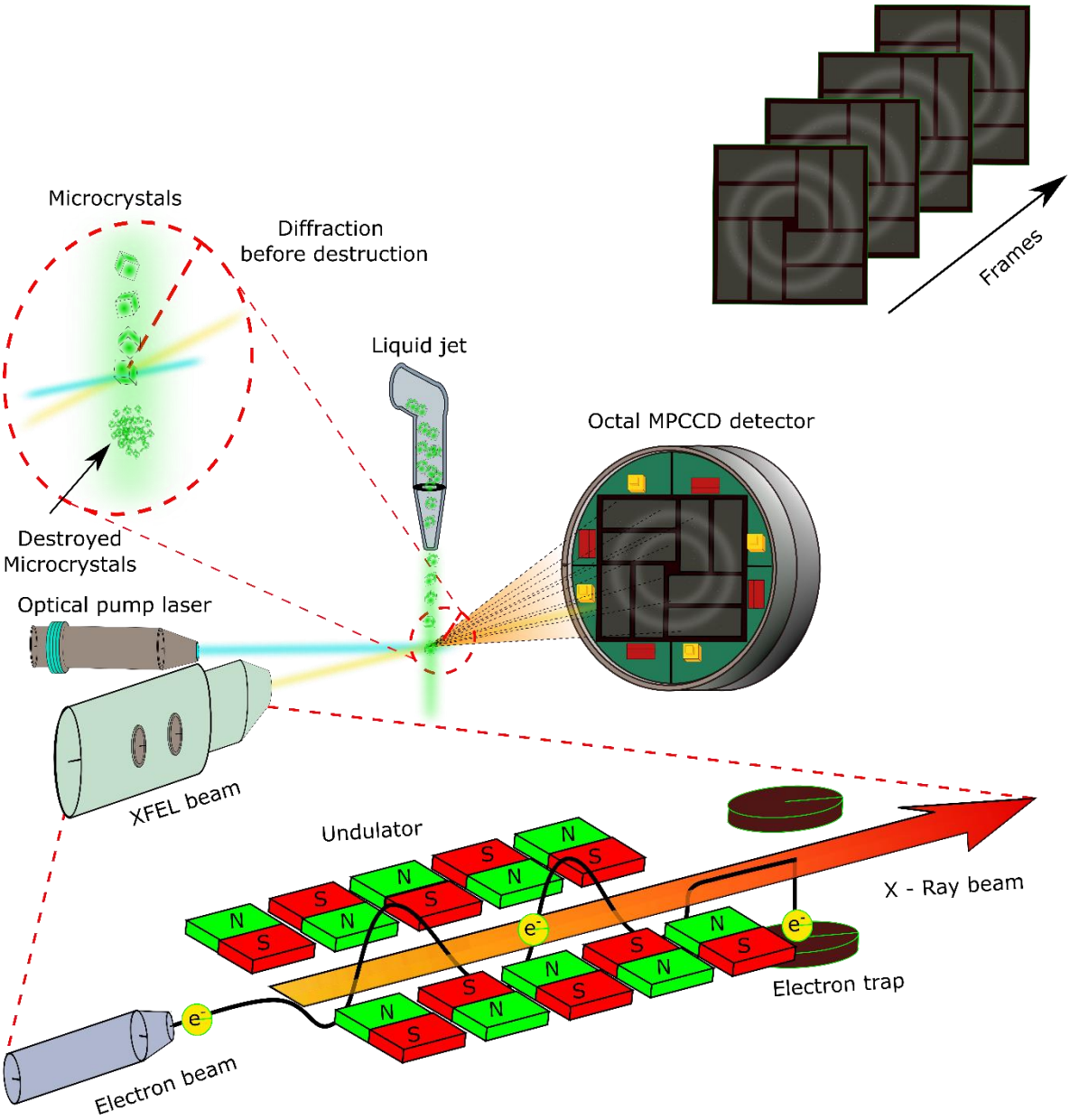


Figure 2.5. Time resolved serial femtosecond X-ray crystallography experiment setup. The representation includes the XFEL beam generation (undulator), the scheme of the pump-probe experiment, the MPCCD detector and the liquid jet sample injection. Figure made with Inkscape.

These incredibly brilliant X-ray beams are produced thanks to a very long linear accelerator (undulator)[50]. A bunch of electrons are injected into the undulator (Figure 2.5), where they oscillate and emit electromagnetic radiation. Since the electrons oscillate all together, their emitted waves are correlated. The E-field of these emitted waves are added, and since the

wave intensity is proportional to the square of the E-field, the emitted intensity exponentially increases until saturation generating such tremendous peak brilliance[50]. In traditional lasers systems, the path available for amplification is expanded by an external optical cavity. On the contrary, normal-incidence mirrors are extremely ineffective for X-ray wavelengths; thus, a very long undulator is needed[50]. In Figure 2.5, an undulator is illustrated. Such intense X-ray radiation destroys the protein, but it is possible to obtain structural information of microcrystals before the destruction of the protein. This is known as diffraction before destruction, and it was proposed in 2000 based on theoretical models[51]. The first X-ray structure resolved from data obtained with an XFEL proved experimentally this theory in 2011[52] (serial femtosecond crystallography; SFX). Finally, using an optical laser to photo-activate the entire microcrystals (pump beam) and the XFEL beam to probe structural changes, Aquila et al[53] performed in 2012 the first time-resolved experiments at an XFEL (TR-SFX).

There is no doubt that ultrafast sub-picosecond time-resolved spectroscopic experiments have provided unprecedented indirect information into the early events and chemical intermediates after light absorption for many light-sensitive systems. TR-SFX is relatively new, and there is still a lot to be investigated. While conceptually TR-SFX are similar to any other pump-probe technique, experimentally there are several differences between a molecule in crystal and in solution phase[54]. In Figure 2.5, a schematic representation of an TR-SFX experiment is presented. The microcrystals of the studied protein are injected with a liquid jet to ensure that only one single crystal is probed by the X-ray pulse. These crystals are excited by a femtosecond optical pump pulse which triggers the photoreaction. The microcrystals are then hit by the XFEL beam and produce a diffraction pattern. Technically, these experiments represent different challenges. First, the crystals are destroyed by the FEL pulse and then new microcrystals need to be presented to the beam for each XFEL pulse. Secondly, due to the intensity fluctuation of optical and FEL beams and the different orientation and forms of each new microcrystal, there is variability between two collected data frames. Therefore, a large number of images need to be collected to solve a protein structure by SFX. For TR-SFX, the number of frames is even higher because, similarly to transient absorption experiments, the signal is calculated as the difference between the data collected with and without pump laser. The high number of frames that needs to be collected,

together with the repetition rate of XFEL beam which is around in 100 Hz range (except the European XFEL in Hamburg, Germany, which has been recently inaugurated and has a MHZ repetition rate[55]) and the hit rate of an XFEL experiment, definitively scale the time of these experiments to a few hours for a single point and reduce the number of collected points after excitation to few delay times. This is in high contrast with any other kind of TR spectroscopy experiments which usually take less than a few hours to collect several hundreds of time delay points.

However, only TR-SFX experiments will provide the atomic structure of crucial intermediates that control the photo-activity of a protein. Such precise information is impossible to be provided entirely from any other technique. The exposed facts manifest the synergy between TR-SFX and optical TR experiments. TR spectroscopic experiments will give the main time constants and number of intermediates that need to be considered in the photodynamical scheme while TR-SFX will reveal their structure and interplay between chromophore and protein residues. Nevertheless, such synergy to get a precise photo-dynamical scheme is not trivial since proteins in solution and crystals can have different dynamics due to different environments (solvation environment, confinement effect). It needs to be pointed out that protein crystals are formed by approximately 50% (per volume) of solvent. Therefore, protein crystals are semi-rigid media where side chains and secondary structures can move, which translate in protein crystals still being biologically active[56]. Besides the relative unconstrained environment of protein crystals, it is evident that they still represent a more constrained media than a liquid solution. Therefore, the dynamics in such different media can definitively differ.

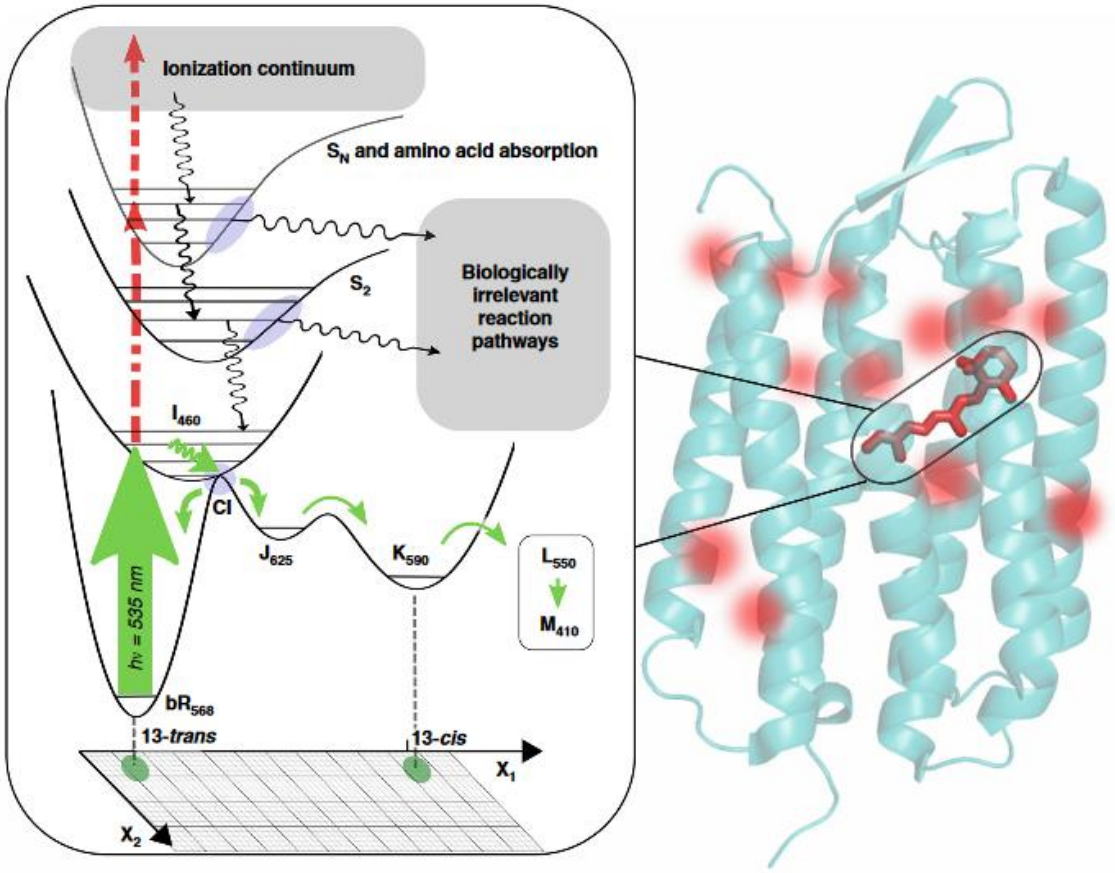


Figure 2.6. Photophysical and chemical pathways at high peak power fluences for bacteriorhodopsin reproduced from reference [57].

Moreover, in TR-SFX experiments, one tries to photo-convert as many of the proteins in a single microcrystal with a single optical pulse[58]. To achieve this photoconversion, knowing the chromophore absorption molar coefficient and the number of chromophores in a crystal, the energy needed is calculated. Up to date, all the sub-ps TR-SFX experiments have used tremendous excitation powers between 360 and 500 GW/cm^2 [35,54,59–62]. These high power densities correspond to far over more than one absorbed photon per chromophore, and as discussed previously, it can lead to multiphoton ionization processes or open non-biological pathways[58]. Even taking into account that a part of the incident light is scattered by the sample[60,63], which has been many times done to justify these high intensities, all performed TR-SFX experiment with sub-ps resolution are done until now are far from the excitation linear regime[35,54,59–61], as clearly pointed out by Miller et al.[57]. For example, the end of the linear regime for bacteriorhodopsin is around 30 GW/cm^2 (more than 100 times lower

energies than those in TR-SFX experiments), then electronic states higher than S_1 state are accessible via multiphoton absorption processes. Depending on the laser pulse width and associated peak power, some of the processes that occur are fully resonant coherent 2-photon, sequential resonant 2-photon, non-resonant 2-photon, to more complex n-photon transitions[57]. The relaxation pathways available from higher electronic states, which can explore many different potential energy surfaces, can lead to entirely different photoproducts and photodynamics, while only the dynamics from S_1 excited state will lead to the biologically relevant photoproduct[56]. The use of TR-SFX results needs to be validated by TR UV-Vis-IR spectroscopy for crystals and compared to the protein in solution. This will also be our strategy for rsEGPF2. Therefore, in the next section, we will give some details regarding the nature and analysis of the femtosecond difference absorption spectra to build a photo-dynamical scheme.

2.4 Data analysis of time-resolved UV-Vis and IR transient absorption spectra.

As given in the Equation 2.15, time-resolved absorption techniques consist in measuring and calculating the difference of absorption spectra at a specific delay time after excitation, i.e. the absorbance measured after laser excitation (pump pulse) minus the absorbance measured without laser excitation. Therefore the difference absorption spectra can be positive or negative related to the appearance of new absorbing intermediate species and photoproducts[64] and disappearance of the initial products. In TRUV-vis spectroscopy, negative and positive difference spectra are the result of different contributions: Ground state bleaching or depopulation bands (GSB) and stimulated emission bands (SE) which are negative signals, and induced absorption which is positive. These signals can be seen in Figure 2.6 for TRUV-vis and TRIR experiments. While induced absorption and GSB are common for TRUV-vis and TRIR, SE is specific for TRUV-vis measurements.

I. Ground state bleaching or depopulation bands (GSB): these negative bands are assigned to the decrease of the number of molecules in the ground state (few percent). The absorption spectra measured after the pump excitation compared to the one

measured without excitation will result in a negative absorption difference in the region where the system absorbs, i.e. the band (without other band contribution) is equal to the inverse of the steady-state absorbance of the sample. This signal is represented in green in Figure 2.6. It can be noted that for TR-UV-Vis spectra there is usually only one GSB band while for TR-IR several bands can be observed.

II. Stimulated emission (SE) bands: these negative bands correspond to the emission from excited molecules when they are dumped to the ground state by the probe. Compared to the absorbance without excitation, there is more light reaching the detector (that coming from the stimulated emission), which leads to a lower absorbance in this spectral region. Therefore, these negative bands are intrinsically located at lower energy (red-shifted in wavelength) with respect to GSB bands. Due to its origin, this type of signals are usually located in the fluorescence emission spectral region. However, as mentioned previously, SE and emission transition are different and scaled by frequency cube (Equation 2.5). This signal is represented in red in Figure 2.6a in red.

III. Excited-state absorption (ESA) band: this signal corresponds to new absorption bands formed by the molecules in the excited state (singlet state, triplet state, hot vibrational excited state, different intermediate states). This signal is represented in yellow in Figure 2.6.

IV. photoproduct (PA) bands: these bands are assigned to new species formed in either the ground or the excited state (and vibrationally excited or not). Technically there is no difference with ESA bands. The assignment of positive absorption bands to the existence of excited state or ground state intermediates is a continuous debate in photochemistry and will be developed in the next section for *cis-trans* isomerization[57].

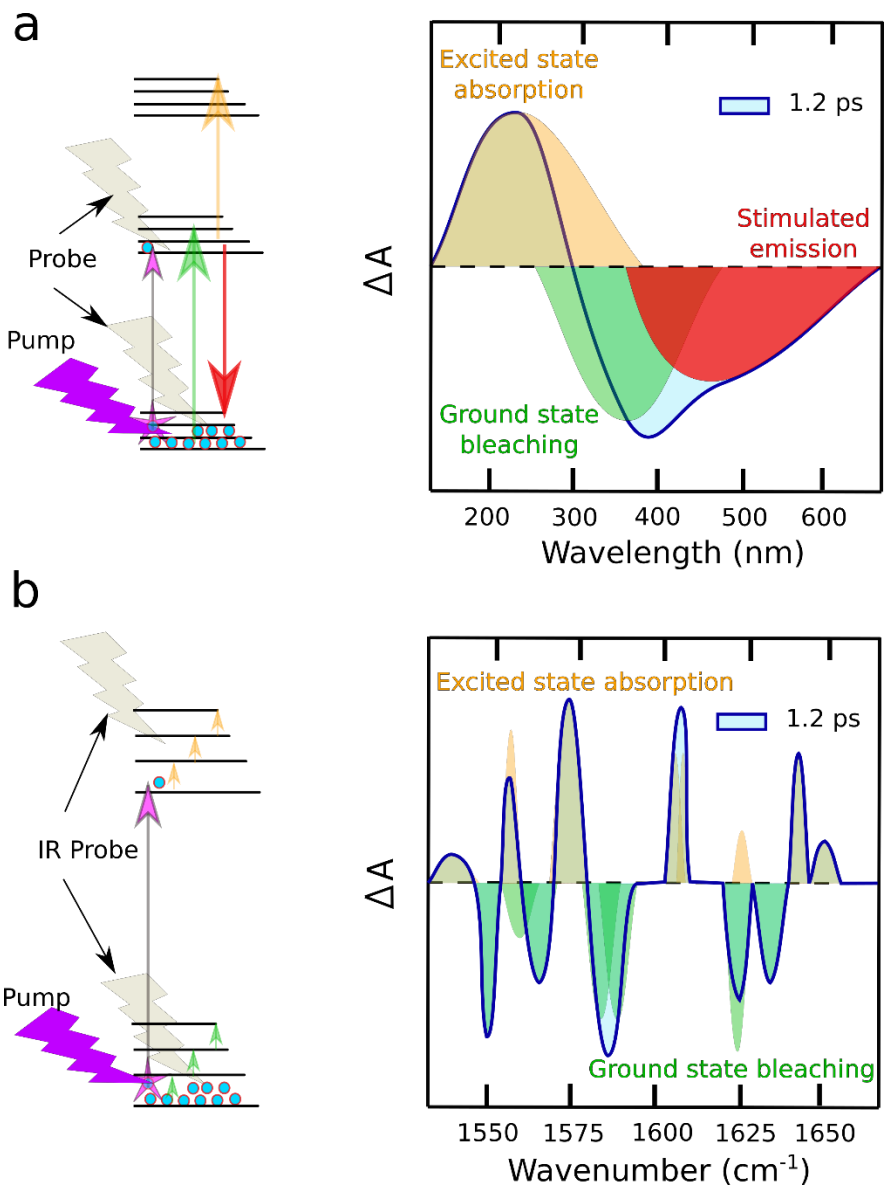


Figure 2.6. Simulated difference absorption spectra for a hypothetic pump-probe TR experiment recorded at 1.2 ps after pump excitation, together with the transitions represented in a simplified Jablonsky diagram. a) For a TRUV-vis experiment and b) for a TRIR experiment. Figure made with Inkscape.

Three remarks can be made regarding the characteristic of difference absorption spectra. (i) As mentioned above, there is no way to distinguish PA or ESA bands. The discussion can be done if an SE band which evolves within the same time constant is observed. Indeed, in this case, only an excited state can have a SE band. (ii) Difference absorption spectra are calculated by subtracting the absorbance without excitation; therefore, GSB bands are a part

of the spectra and cannot be dissociated from the rest of the signals. Relaxation of excited states to the initial state will be associated with the decrease of the GSB band. (iii) Since for an electronic state there are different vibrational states; all the bands are usually convoluted with cooling process such as (IVR and VR) and hot vibrational excited states (in excited or ground states). This increases the complexity of the transient spectra and its analysis.

TR spectroscopies often consist on detecting with a probe changes in absorption of the sample provoked by another laser pulse (the pump) over times with a stroboscopic or continuous detection. Thus, the resulting signal is a multivariate[65,66], with a first independent variable wavelength (UV-Vis) or wavenumber (IR), and a second independent variable being the time delay between pump and probe. In other words, a time-resolved data set is a collection of measurements done at different (distinct) times and wavelengths. The evolution of the absorption in time for the different wavelengths contains information such as reaction kinetics and molecular changes.

To unravel the chemical reactions and processes behind the multivariate data, model-based analysis of the data is mandatory[65–67]. The most straightforward approach to model a photochemical reaction is to assume that the concentration of transient species or reaction rate is the sum of independent species (Beer-Lambert law) and only depends on the concentration of these species (first-order reaction). The integration of a first order reaction can be described by exponential decay. According to this hypothesis, the data should be explained by a weighted sum of exponential functions, where every exponential time constant represents the decay of a transient species that follow a first-order reaction, and the pre-exponential value or weight, represents the concentration of this species (Equation 2.16). Fitting the kinetic traces recorded at one wavelength can yield an estimation of the number of transient species involved in the reaction and their decay constants. One of the most established approaches to analyze the different absorption spectra is to fit the time evolution for several wavelengths with the same sum of exponentials[65,66,68,69] (one wavelength can average with the adjacent ones to increase the signal to noise ratio). This approach has been demonstrated to be much robust and accurate than a separate or single wavelength analysis and is typically known as global fit analysis. Furthermore, from a global analysis of the data, not only the number of transient species and their decay time constants can be obtained but

also their decay associated spectral shapes. These are obtained from the weights of each exponentials associated to a species at each of the wavelengths, which constitute the so call Decay Associated Spectra (DAS).

$$\psi(t,\lambda) = \sum_0^n A_i * e^{(\frac{t-t_0}{\tau})} \qquad \text{Equation 2.16}$$

More complex analysis can be done if the data is fitted to a specific kinetic model with rate constants; this is known as target analysis or chemical reaction modelling[66]. Kinetic models are for example a cascade model in which the species evolve from one to the next one sequentially with 100% quantum yield efficiency, or it can be more complex kinetic models (parallel pathway) were the conversion yields between the different species and pathways can be fixed with known external parameters as photoconversion quantum yields. The result of such an analysis can give the so-called evolving associated difference spectra (EADS), which can yield a more comprehensive picture with rate constants of the different pathways of the photoreaction and not only time constants of species in DAS analysis.

Finally, more advanced analysis method can be applied. Due to the multilinearity characteristic of this type of data, multi-curve resolution alternating least squares method (MCR-ALS) has demonstrated to be efficient in decomposing the data in pure species spectra and their corresponding time evolution concentration profiles (soft modelling). From this concentration profiles, different models can be applied, e.g., lifetime decay can be obtained using an exponential fit. The rate constants for each species can also be obtained by applying a model and using hybrid hard-and soft-multivariate curve resolution methods[65,67,70].

For both simple and complex models, the goodness or quality of the fit should always be evaluated. In general, a correlation coefficient is used to examine “goodness of fit” between the model and experimental data. The most used one is the chi-squared χ^2. A high correlation coefficient does not implicitly imply that the model correctly describes the data, as the Anscombe's quartet points out[71]. One of the best ways to evaluate a model is to analyze the residual plot (spectra and time profiles). The residues of the fit should be similar to random noise and centered around zero ordinates with no trend with time evolution.

In practice, the selection of the number of species/components (number of exponential time constants in DAS) is not straightforward, as a higher number of components will always fit

better the data set, without necessarily describing better the photophysical model. The literature on similar systems can help in this step. Then a first visual analysis of the transient absorption spectra evolutions is already giving the main temporal and spectral evolutions. Finally, mathematical methods such as singular value decomposition (SVD) can also help. The SVD is a matrix factorization technique that decomposes the data (M) in three simpler matrices U, S and V[72] (Equation 2.17).

$$M = U \times S \times V^T \qquad \text{Equation 2.17}$$

Where U and V are the left and right singular vectors and S a diagonal matrix containing the singular values, the number of non-zero singular values is the rank or the minimum number of components that describes the data set M. Nevertheless, due to inevitably noisy data, the selection of non-zero singular values can be tricky. As can be seen, the selection of components is not trivial, and usually, several tests are done by adding or subtracting one component in the analysis, to find the minimum number of components that correctly describes the data. Once a model has been chosen, we try to describe the data with it (fitting), which consists in varying the model parameters iteratively until a set of parameters that describes the data is found. One of the problems with nonlinear minimization, particularly with increasing model complexity, is the possibility to arrive in a local minimum of the error function without finding the global minimum; in other words, there might be other parameters for the model that could describe better the data. Even more, in multivariate data, different sets of model parameters may result in exactly the same quality fit (in terms of lack of fit) thus the same χ^2. The best way to solve this issue is to repeat the process using a different set of initial model parameter guesses and determine whether an equivalent set of best-fit parameters is obtained.

Once the best set of the model parameters that describes the data correctly is chosen, the error should be estimated for each fitted parameter (time and rate/decay time constant, pre-exponential factors). The estimated error of a parameter is related to the minimum of the multidimensional χ^2 surface. The ambiguity in the estimation of the parameter should not be mixed up with the error of each parameter[73]. Error translates the uncertainty in the estimation of each parameter independently of the others. The estimation of the fitting error is usually given as a confidence interval (+/-) calculated by inverting the second-order derivative error

matrix where only the diagonal terms of are taken into consideration, (not interaction terms). This means that a single 1D trace of the multidimensional χ^2 surface is considered for each parameter. It can result that an ambiguous parameter may be estimated with a very low error. The error and the ambiguity in the parameter estimations are well-known problems in multi-exponential functions[74]. The best way to estimate the error is to fit the model to multiple data sets of the same sample. If all data sets are collected similarly, they should only differ in random variability such as noise. Finally, the variation in the individual model parameters fitted to all the recorded data sets will give the confidence interval or error of the parameter models. It is practically impossible to acquire a sufficiently number (1000) of experimental datasets of the same sample under the same conditions to have a large representative number of samples. One approach to overcome this problem is bootstrap, which is one method that allows generating new data by shuffling and substituting a part of the original points in the initial data set by synthetic data generation with the best parameters describing the model and including noise representative of the actual measurement noise[73]. This last approach was used in this thesis to judge the quality of fit for noisy data and has been explored in this thesis. More details can be found in Appendix 2.

2.5 Focus on two deactivation processes that occur for photo-active fluorescent proteins.

The first known deactivation process of fluorescent proteins is found for the famous avGFP. Excitation of the neutral form induces a fluorescence coming from anionic species through an excited-state proton transfer (ESPT). For most of RSFP proteins, such as rsEGFP2, the switch between states involves a *cis-trans* isomerization, a proton transfer and some protein rearrangements. Thus, here some definitions regarding basic concepts and literature examples of ESPT and *cis-trans* isomerization are given. More details on the photodynamics can be found in the next chapter.

2.5.1 Excited state proton transfer

The fluorescence of the avGFP neutral form is due to an excited state proton transfer (ESPT). Proton transfer reactions are widely found in a wide variety of different chemical systems. In a water-based planet and water-based organism, proton transfer reactions are one of the most basic chemical reactions which are yet not perfectly understood[75]. In the case of a photoreaction, a proton transfer can occur in both ground and excited states and includes formation and breaking of hydrogen bonds. It can be produced intramolecularly (hydrogen is coming from the same molecule) or with other molecules or solvent. Proton transfer reactions represent a research field by itself, and many aspects can be found in the book "Hydrogen-Transfer Reactions". Edited by J. T. Hynes et al.[75].

Proton transfer can occur in a single step or it can be characterized by a multi-step process through different metastable intermediates (both cases can be concerted). Finally, similar to electron transfer reactions, tunneling mechanism plays a vital role in proton transfer reactions. While electrons transfer reactions are well described by Marcus theory[76], there is no widely accepted theory for H-transfer up to date[77]. The kinetic isotopic effects (comparison of the rate constant for hydrogen (H) and deuterium (D)) constitute an essential tool to study proton transfer reactions although they do not reveal the existence of ESPT via tunneling[77]. With so many possibilities and types of proton transfer reactions, it is not easy to revise all mechanisms and examples in literature. Therefore, only some essential definitions regarding ESPT occurring in FPs will be given here.

Proton transfer reactions in the excited state can be divided in two types, intramolecular proton transfer i.e. ESIPT, or intermolecular one named ESPT. ESIPT is usually occurring with a shorter time constant (sub-100 fs can be found in the literature) while ESPT reaches few picoseconds. Some of the molecules exhibiting ESPT or ESIPT are GFP chromophores or analogues as the *orto*-HBDI[78,79] or the avGFP itself[80–82]. In 1986, Kasha[83] divided the ESPT reactions into four different types, only the two related with the avGFP and the HBDI will be described here. The first named "Intrinsic Intramolecular Proton Transfer" (ESIPT) is the one observed for *o*-HBDI analogues, while the one that exhibits by avGFP[80–82] is a "Proton-relay tautomerization", and is also observed for example in 7-hydroxyquinoline in methanol[84]. Besides their differences, both can be represented by similar schematic

dynamics. After excitation, an excited state A* converts to an unrelaxed excited tautomeric form I* by ESPT, which finally relaxes to an I state[83] (Figure 2.7).

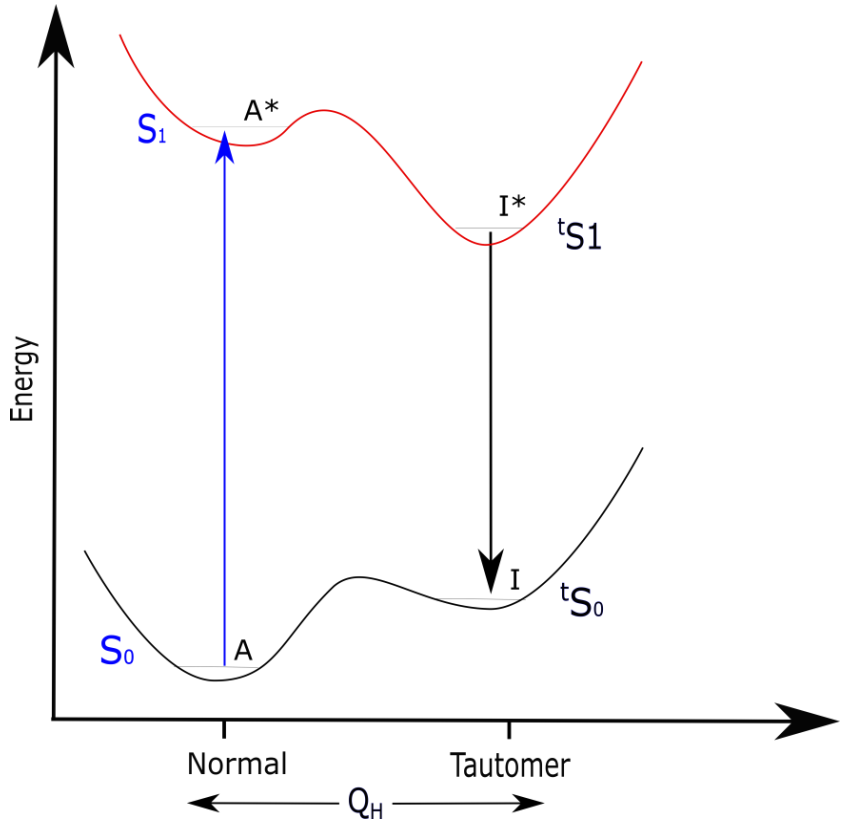


Figure 2.7. Double-well potential for proton-transfer in the excited state, with Q_H as the proton-transfer coordinate. Figure adapted with Inkscape from reference [83].

The characterization of ESPT is a weak fluorescence of the A initial stable form, which after excitation and proton transfer gives the tautomeric excited I* fluorescent form[85]. Due to the difference in energy for the transitions A→A* and I*→I, ESPT reactions display a significant Stokes shift for the fluorescence of A respect to the emission of I*, Figure 2.7. Therefore, one of the best ways to detect ESPT is via fluorescence upconversion, which will give the time constant of the ESPT. One of the parameters that controls the rate constant of the ESPT is slow motions of the molecular skeleton[82,86–88], which are known as coherently excited vibrations. In some cases, these modes are the relevant reaction coordinates, to the detriment of proton modes[85]. These coherently excited vibrations revealed as oscillatory signals contributing to the primary decay signal in time-resolved transient absorption spectroscopy if short duration excitation pulses of ~ 30 fs are used (coherent excitation)[88].

Applying the Fourier transform to these oscillatory signal components, the vibrational frequencies of the excited state modes can be deduced. The frequencies of the skeletal modes are compared to theoretically calculated ones with the help of quantum calculations, and can give excellent hints of the ESPT dynamics[85]. As mentioned above, they have been observed in organic molecules[86,87] as well as GFP proteins[82,88]. Finally, it needs to be mentioned that ESPT is also responsible for the behavior of some organic dyes, which after excitation becomes stronger acids or stronger bases in the electronic excited-state, these molecules are known as photoacids. The pK_a* of these molecules in the excited state can be estimated with the Föster cycle[89]; some GFP chromophore analogues have revealed such behaviour[90,91].

2.5.2 Cis-trans isomerization

In organic molecules, when two atoms are covalently bound by a single bond, they can freely rotate if nothing hinders this movement. Theoretically, all positions within a 360º dihedral rotation angle around a C-C bond can be adopted. In the case where the rotation around a particular angle is blocked, for example the presence of a double bond, only distinct positions of the substituents along the C-C bond may exist as stable rotamers called isomers. In the simplest case, in the ground state, there are two stable geometrical isomers accessible through the intrinsic isomerization barrier[92]. A simple double bond has two possible isomers; depending on the number n of double bonds, there can be 2^n isomers[92]. By convention, when the highest priority substituents (highest atomic number) of the double-bond are in the same side of the carbon chain, the isomer is in "cis" position, and *"trans*" when the substituents are in opposites sides[92].

Due to the steric hindrance of the substituents in the *cis* isomer, the *trans* conformer is generally the most stable one (Figure 2.8). The isomerization from *trans* to *cis* isomers can occur: (i). Thermally if the ΔG barrier can be overcome. (ii). Via a chemical reaction, generally via an intermediate for which the double bond is broken, which lowers the energy barrier. (iii). Via a photo-activated reaction, photoisomerization, where the energy barrier is overcome in the excited state. The identification and the description of the dynamics in the excited state and the existence or not of an intermediate is the main goal of ultrafast studies.

There are three mainly photoisomerization mechanisms proposed in the literature, One Bond Flip (OBF), Hula Twist (HT) and Bicyclic Pedal (BP)[92–95].

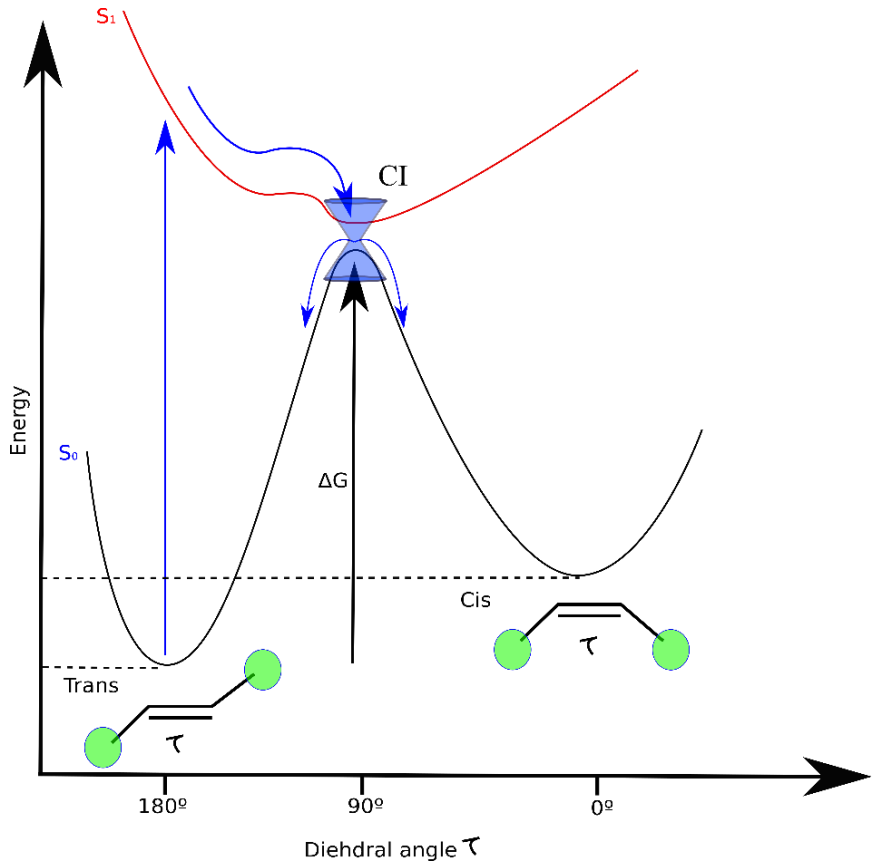


Figure 2.8. General schematic representation of a cis-trans isomerization of a double bond via an OBF mechanism. The blue arrows represent a hypothetical photoisomerization pathway via the S_1 state. Figure made with Inkscape.

The OBF mechanism was the firstly prevailing theory for *cis-trans* photoisomerization. After excitation, this mechanism implies a twisted excited S_1 state[92,95], with a dihedral angle near 90º (blue arrow, see the schematic representation in figure 2.8). The relaxation of the excited state via an adiabatic mechanism through a conical intersection (CI) to the ground state leads the chromophore to be in hot energy ground state from which it can relax to either one of the more stable isomers (cis or trans). For OBF mechanism, the final isomer has achieved a 180º turnover of one half of the molecule compared to the initial isomer (Figure 2.9). This has been represented by a blue arrow in Figure 2.9. This mechanism needs a high free volume and could not explain isomerization in constrained environments and structures of photo-

isomerized chromophores that show complete rotation. One of these cases is, for example, the retinal chromophores[96,97].

Therefore, new mechanisms were proposed. The BP mechanism was first proposed to explain the ultrafast isomerization of the batho-rhodopsin[98]. This mechanism involves at least two conjugated double bonds and rationalizes the isomerization of these two adjacent double bounds with only one photon. The mechanism was in agreement with the low-volume-demanding *cis-trans* isomerization of the 11-cis-retinyl chromophore inside the protein matrix but was not consistent with the final structures of bathorhodopsin (bR). The BP mechanism will not be further discussed since the HBDI chromophore isomerization cannot occur through this mechanism as it has only one double bond conjugated with two aromatic rings. Several years after the postulation of the BP isomerization mechanism, another volume conserving mechanism was suggested to be responsible for the *cis-trans* isomerization in bR. This new mechanism consists in a simultaneous rotation of the double bond τ together (blue arrow) with an additional 180º isomerization of one adjacent single bond Φ (red arrow), initially named *c*oncerted *t*wist at center *n*[99], was lately named hula twist (HT, Figure 2.9)[100].

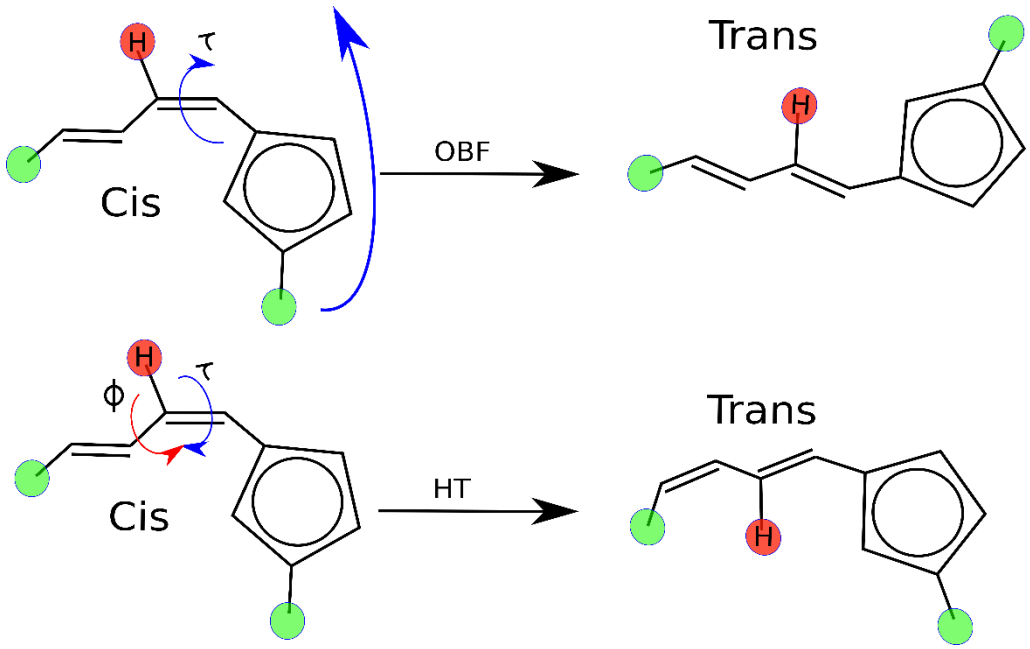


Figure 2.9. Schematic representation of a cis-trans photoisomerization of a double bond, via one bond flip (OBF) and hula twist (HT) mechanism. Figure made with Inkscape.

The general established idea is that, without constraint, the photoproduct after isomerization corresponds to the OBF one, while the HT volume conserving photoproduct is found in constraint environments[93–95,101]. The different products of an isomerization via OBF and HT pathways are represented in Figure 2.10. Since OBF and HT photoproducts are rotamers, which means that they can transform into each other via single bond rotation (SBR), there are several open questions: (1) if an OBF photoproduct can be formed via a HT or an OBF mechanism (2) the existence of twisted intermediates in the excited or ground state for both mechanisms. To understand these issues, the example of choice is the classical and well-studied *cis-trans* isomerization of stilbene and derivatives.

The stable form of these molecules is the *trans* isomer. In the excited state, it was hypothesized (theoretical calculations, ultrafast electronic and vibrational spectroscopy) that both isomers, *cis* and *trans*, have a twisted minimum isomer, for which the double bond is elongated, allowing a more free isomerization upon rotation (OBF). Starting from the *trans*-stilbene isomer, the photoisomerization characteristic time constant increases rapidly with the increase of solvent viscosity concomitantly with the increase of fluorescence quantum yield. Moreover, the photoisomerization is completely blocked and does not occur upon freezing. Such results are entirely in agreement with an OBF mechanism. On the contrary, starting from the *cis* isomer, upon sterical impediments and in frozen media, the *cis-trans* isomerization still occurs and become stereospecific. As discussed in literature, the behavior of *cis* conformers in frozen media can only be explained by a HT mechanism[70,71,79]. The complexity comes from the fact that the resulting photoproducts of *cis* stilbene isomerization from either a HT or OBF mechanisms are indistinguishable. Therefore, when studying the possible isomerization pathways of stilbene, it is essential to use substituents[95]. The isomerization in stilbenes have been extensively reviewed[102]. In the following discussion, the main results will be discussed.

First, two important points need to be commented for the difference between photoproducts resulting from a HT isomerization and those arising from an OBF one, the region-selectivity and the metastability[103]. The single bonds adjacent to the double bond can have a torsion angle, and thus two photoproducts can exist after an HT isomerization. On the contrary, there will be only one photoproduct after an OBF isomerization[95,103]. In Figure 2.10, the possible

products after photoisomerization starting from *cis*-stilbene isomer (via OBF and HT) are presented. As can be seen, the HT-1 photoproduct, due to the steric impediment, has a higher potential energy than the other two (OBF and HT-2) which are indistinguishable in solution or gas phase. The HT-1 photoproduct rapidly evolves to an OBF photoproduct with a simple rotation along the single carbon bond (black arrow). These two photoproducts are rotamers. It has been indeed demonstrated that HT photoproducts are metastable, and even at temperatures lower than 50 K, they can evolve to the OBF photoproduct[104]. For these aforementioned reasons, it is difficult to observe a HT photoproduct and discuss the nature of the isomerization mechanism.

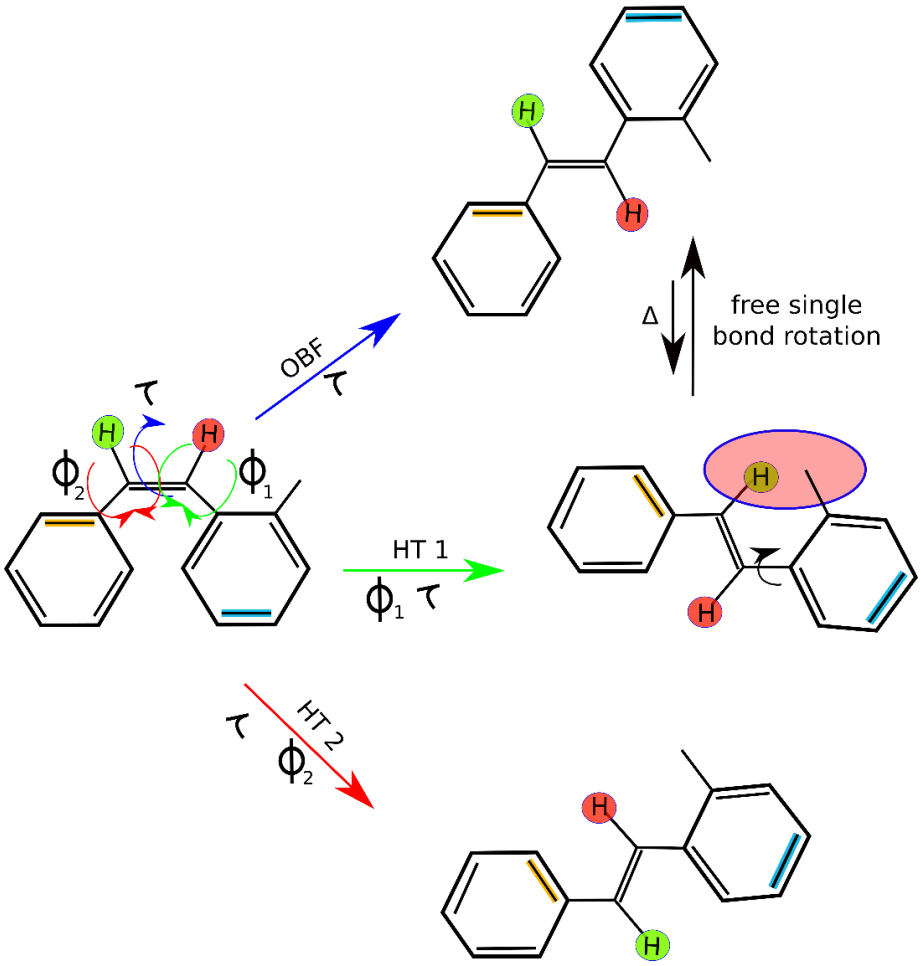


Figure 2.10. Cis-trans isomerization reactions of stilbene with a Hula-twist marker. Top: Conventional one-bond-flip (OBF) process showing the rotation of one-half of the molecule around τ. Middle and bottom: The Hula-Twist (HT-1 and HT-2) photoproducts around Φ and τ, showing the steric impediment for HT-1. Figure made with Inkscape.

The publication *"The case of medium-dependent dual mechanisms for photoisomerization: One-bond-flip and Hula-Twist"* in 2000 from Robert S. H. Liu and George S. Hammond[101], critically reviewed several results in different molecular systems and media, reinterpreting many previous findings initially attributed to OBF isomerization in favor of the HT mechanism. They refer to HT as a diabatic concerted mechanism higher in energy than OBF. Therefore, they postulated that isomerization via HT mechanism is the less probable isomerization process that only happens in restricted environment conditions, where the OBF is eliminated[101].

On the contrary, only a few years after, Oluvucci and Fuss[105] in an also extended publication, including quantum chemistry calculations, suggested that the general mechanism for all photoisomerization reactions, even in solution, should be HT type. They showed that a *cis-trans* isomerization through a CI requires a complex molecular distortion which, is impossible to be described in terms of just a single C=C double bond rotation (OBF), and corresponds to a concerted twisting around a double bond and an adjacent single bond. Furthermore, they concluded that the absence of the HT photoproduct is due to the fact that an isomerization through CI yields the HT photoproduct in a hot ground state, from where it can evolve upon relaxation to an OBF photoproduct via a single bond rotation (SBR), naming this mechanism "aborted hula twist". Altogether the authors concluded that the HT mechanism is the lowest-energy pathway for *cis–trans* isomerization, contrary to Hammond et al.[101]

There are several examples in the literature supporting the idea that HT mechanism is the general mechanism for *cis-trans* isomerization. In fact, it has been demonstrated that stilbene analogues, for which one of the two HT isomerization pathways is blocked by a carbon chain (single block stilbenes), still isomerizes with a similar time constant as free stilbene in around 135 fs. On the contrary, when the two HT isomerization pathways are blocked, avoiding the rotation of the two single bonds, the time constant for isomerization increases by about 5-6 times to around 800 fs[106,107]. Similar results can also be seen for cyanine dyes[108]. The general isomerization mechanism is still discussed up to date due to the difficulties in the characterization a HT intermidiate[109].

The same isomerization discussion for stilbene can also be found in photoactive proteins, where the nature of intermediate states, their geometry and their existence in excited or ground states are also a matter of debate. As commented above the HT photoproduct is assumed to be the only one observed in constrained media such as frozen solvent and inside protein matrixes. The HT mechanism was first postulated for isomerization of 11-*cis*-retinyl chromophore "sandwiched" within the rhodopsin protein matrix[99]. Several articles and reviews described the complete rhodopsin mechanism for the vision as a several steps process, as well as mechanisms observed in other retinal binding proteins such as bathorodopsin[93,96]. Recently, the structure of the primary intermediates have been elucidated by TR-SFX experiments[54,60].

Another example of *cis-trans* isomerization via HT is the pre-vitamin D, which can photo-transform to tachysterol under UV irradiation. UV-Vis absorption experiments done at 90 K on pre-vitamin by D. Fuss and coworkers, reported a volume conserving isomerization photoproduct; this photoproduct and the reviewed of several previous works lead to the authors to propose the HT isomerization mechanism in pre-vitamin D[110]. Nevertheless, several years after, new NMR and fluorescence experiments at even lower temperatures (77K) suggested a new mechanism via a one-bond twist (OBT) very similar to OBF mechanism[111]. More recently, a new extensive review of several works concluded that HT should be the main mechanism[112].

Finally, another example that shows the complexity of the spectroscopic experiments data and its interpretation in order to build a rationalized isomerization scheme is the photoactive yellow protein (PYP), which has also been extensively studied[92,93]. The PYP protein is a *para*-coumaric acid covalently bound to a protein matrix via a cysteine amino acid. The chromophore can reversibly be isomerized from *trans* to *cis* in the nanosecond time scale via several distinct intermediate conformations[93]. The PYP is one of the systems that Robert S. H. Liu and George S. Hammond reviewed, indicating that the isomerization results should be attributed to a HT[101] mechanism as the chromophore is inside a protein cage and the OBF isomerization would result in the same photoproduct as HT-2. New results on PYP based on quantum mechanic calculations show that isomerization still happens around a single bond when isomerization around the double bond is blocked[113]. To conclude, it is need to be

mentioned that the isomerization mechanism for RSFPs is still a matter of debate and will be the center part of this thesis.

2.6 References

(1) Weisskopf, V. F. How Light Interacts with Matter. *Scientific American* **1968**, *219*, 60–71.
(2) Condon, E. A Theory of Intensity Distribution in Band Systems. *Phys. Rev.* **1926**, *28*, 1182–1201.
(3) Franck, J.; Dymond, E. G. Elementary processes of photochemical reactions. *Transactions of the Faraday Society.* **1926**, *21*, 536.
(4) Condon, E. U. Nuclear Motions Associated with Electron Transitions in Diatomic Molecules. *Phys. Rev.* **1928**, *32*, 858–872.
(5) Born, M.; Oppenheimer, R. Zur Quantentheorie der Molekeln. *Ann. Phys.* **1927**, *389*, 457–484.
(6) Morse, P. M. Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels. *Physical Review* **1929**, *34*, 57–64.
(7) Jablonski, A. Efficiency of Anti-Stokes Fluorescence in Dyes. *Nature* **1933**, *131*, 839–840.
(8) Jabłoński, A. Über den Mechanismus der Photolumineszenz von Farbstoffphosphoren. *Z. Physik* **1935**, *94*, 38–46.
(9) Lewis, G. N.; Kasha, M. Phosphorescence and the Triplet State. *Journal of American Chemical Society* **1944**, *66*, 2100–2116.
(10) XXX. On the change of refrangibility of light. *Phil. Trans. R. Soc.* **1852**, *142*, 463–562.
(11) Franck, J.; Livingston, R. Remarks on the Fluorescence, Phosphorescence and Photochemistry of Dyestuffs. *The Journal of Chemical Physics* **1941**, *9*, 184–190.
(12) Hilborn, R. C. Einstein coefficients, cross sections, f values, dipole moments, and all that. *American Journal of Physics* **1982**, *50*, 982–986.
(13) Sorokin, P. P.; Lankard, J. R.; Hammond, E. C.; Moruzzi, V. L. Laser-pumped Stimulated Emission from Organic Dyes: Experimental Studies and Analytical Comparisons. *IBM J. Res. & Dev.* **1967**, *11*, 130–148.
(14) Zewail, A. H. Laser femtochemistry. *Science (New York, N.Y.)* **1988**, *242*, 1645–1653.
(15) Hell, S. W.; Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics letters* **1994**, *19*, 780–782.
(16) Klar, T. A.; Jakobs, S.; Dyba, M.; Egner, A.; Hell, S. W. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 8206–8210.
(17) 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. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 17565–17569.

(18) quantum yield, Φ. In *IUPAC Compendium of Chemical Terminology*; Nič, M., Jirát, J., Košata, B., Jenkins, A., McNaught, A., Eds.; IUPAC: Research Triagle Park, NC, 2009.
(19) Colletier, J.-P.; Sliwa, M.; Gallat, F.-X.; Sugahara, M.; Guillon, V.; Schirò, G.; Coquelle, N.; Woodhouse, J.; Roux, L.; Gotthard, G.*; et al.* Serial Femtosecond Crystallography and Ultrafast Absorption Spectroscopy of the Photoswitchable Fluorescent Protein IrisFP. *The journal of physical chemistry letters* **2016**, *7*, 882–887.
(20) Woodhouse, J.; Nass Kovacs, G.; Coquelle, N.; Uriarte, L. M.; Adam, V.; Barends, T. R. M.; Byrdin, M.; La Mora, E. de; Bruce Doak, R.; Feliks, M.*; et al.* Photoswitching mechanism of a fluorescent protein revealed by time-resolved crystallography and transient absorption spectroscopy. *Nature communications* **2020**, *11*, 741.
(21) Tokunaga, A.; Uriarte, L. M.; Mutoh, K.; Fron, E.; Hofkens, J.; Sliwa, M.; Abe, J. Photochromic Reaction by Red Light via Triplet Fusion Upconversion. *Journal of the American Chemical Society* **2019**, *141*, 17744–17753.
(22) Jacquet, M.; Uriarte, L. M.; Lafolet, F.; Boggio-Pasqua, M.; Sliwa, M.; Loiseau, F.; Saint-Aman, E.; Cobo, S.; Royal, G. All Visible Light Switch Based on the Dimethyldihydropyrene Photochromic Core. *The journal of physical chemistry letters* **2020**, *11*, 2682–2688.
(23) Eigen, M. Methods for investigation of ionic reactions in aqueous solutions with half-times as short as 10 –9 sec. Application to neutralization and hydrolysis reactions. *Discuss. Faraday Soc.* **1954**, *17*, 194–205.
(24) Norrish, R. G. W.; Porter, G. Chemical Reactions Produced by Very High Light Intensities. *Nature* **1949**, *164*, 658.
(25) Hirvonen, L. M.; Suhling, K. Wide-field TCSPC: methods and applications. *Meas. Sci. Technol.* **2017**, *28*, 12003.
(26) Reid, G. D.; Wynne, K. Ultrafast Laser Technology and Spectroscopy. In *Encyclopedia of Analytical Chemistry*; Meyers, R. A., Ed.; John Wiley & Sons, Ltd: Chichester, UK, 2006.
(27) Efremov, E. V.; Buijs, J. B.; Gooijer, C.; Ariese, F. Fluorescence rejection in resonance Raman spectroscopy using a picosecond-gated intensified charge-coupled device camera. *Applied spectroscopy* **2007**, *61*, 571–578.
(28) Dantus, M.; Rosker, M. J.; Zewail, A. H. Real-time femtosecond probing of ‘‘transition states’’ in chemical reactions. *The Journal of Chemical Physics* **1987**, *87*, 2395–2397.
(29) https://ellesgroup.ku.edu/ta_spectroscopy.
(30) Moeyaert, B.; Nguyen Bich, N.; Zitter, E. de; Rocha, S.; Clays, K.; Mizuno, H.; van Meervelt, L.; Hofkens, J.; Dedecker, P. Green-to-red photoconvertible Dronpa mutant for multimodal super-resolution fluorescence microscopy. *ACS nano* **2014**, *8*, 1664–1673.
(31) TCSPC - What is Time-Correlated Single Photon Counting? https://www.edinst.com/blog/what-is-tcspc/.
(32) Fron, E.; Sliwa, M.; Adam, V.; Michiels, J.; Rocha, S.; Dedecker, P.; Hofkens, J.; Mizuno, H. Excited state dynamics of the photoconvertible fluorescent protein Kaede revealed by ultrafast spectroscopy. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology* **2014**, *13*, 867–874.

(33) Ariese, F.; Roy, K.; Ravi Kumar, V.; Sudeeksha, H. C.; Kayal, S.; Umapathy, S. Time-Resolved Spectroscopy: Instrumentation and Applications. In *Encyclopedia of Analytical Chemistry*; Meyers, R. A., Ed.; John Wiley & Sons, Ltd: Chichester, UK, 2006; pp 1–55.
(34) Spence, D. E.; Kean, P. N.; Sibbett, W. 60-fsec pulse generation from a self-mode-locked Ti:sapphire laser. *Optics letters* **1991**, *16*, 42–44.
(35) Coquelle, N.; Sliwa, M.; Woodhouse, J.; Schirò, G.; Adam, V.; Aquila, A.; Barends, T. R. M.; Boutet, S.; Byrdin, M.; Carbajo, S.*; et al.* Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nature chemistry* **2018**, *10*, 31–37.
(36) Greetham, G. M.; Sole, D.; Clark, I. P.; Parker, A. W.; Pollard, M. R.; Towrie, M. Time-resolved multiple probe spectroscopy. *The Review of scientific instruments* **2012**, *83*, 103107.
(37) Konold, P. E.; van Stokkum, I. H. M.; Muzzopappa, F.; Wilson, A.; Groot, M.-L.; Kirilovsky, D.; Kennis, J. T. M. Photoactivation Mechanism, Timing of Protein Secondary Structure Dynamics and Carotenoid Translocation in the Orange Carotenoid Protein. *Journal of the American Chemical Society* **2019**, *141*, 520–530.
(38) Kukura, P.; McCamant, D. W.; Mathies, R. A. Femtosecond stimulated Raman spectroscopy. *Annual review of physical chemistry* **2007**, *58*, 461–488.
(39) Herek, J. L.; Materny, A.; Zewail, A. H. Femtosecond control of an elementary unimolecular reaction from the transition-state region. *Chemical Physics Letters* **1994**, *228*, 15–25.
(40) 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.*; et al.* Infrared spectroscopy reveals multi-step multi-timescale photoactivation in the photoconvertible protein archetype dronpa. *Nature chemistry* **2018**, *10*, 845–852.
(41) 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. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 13005–13009.
(42) Bourgeois, D.; Royant, A. Advances in kinetic protein crystallography. *Current opinion in structural biology* **2005**, *15*, 538–547.
(43) 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.*; et al.* Structure of a protein photocycle intermediate by millisecond time-resolved crystallography. *Science (New York, N.Y.)* **1997**, *275*, 1471–1475.
(44) Srajer, V.; Teng, T.; Ursby, T.; Pradervand, C.; Ren, Z.; Adachi, S.; Schildkamp, W.; Bourgeois, D.; Wulff, M.; Moffat, K. Photolysis of the carbon monoxide complex of myoglobin: nanosecond time-resolved crystallography. *Science (New York, N.Y.)* **1996**, *274*, 1726–1729.
(45) Farber, G. K. Laue crystallography: Lights! Camera! Action! *Current Biology* **1997**, *7*, R352-R354.
(46) Schotte, F.; Lim, M.; Jackson, T. A.; Smirnov, A. V.; Soman, J.; Olson, J. S.; Phillips, G. N.; Wulff, M.; Anfinrud, P. A. Watching a protein as it functions with 150-ps time-resolved x-ray crystallography. *Science (New York, N.Y.)* **2003**, *300*, 1944–1947.
(47) Madey, J. M. J. Stimulated Emission of Bremsstrahlung in a Periodic Magnetic Field. *Journal of Applied Physics* **1971**, *42*, 1906–1913.

(48) Madey, J. Stimulated emission of radiation in periodically deflected electron beam: US Patent 38 22 410 **1974**.
(49) Emma, P.; Akre, R.; Arthur, J.; Bionta, R.; Bostedt, C.; Bozek, J.; Brachmann, A.; Bucksbaum, P.; Coffee, R.; Decker, F.-J.*; et al.* First lasing and operation of an ångstrom-wavelength free-electron laser. *Nature Photon* **2010**, *4*, 641–647.
(50) Margaritondo, G.; Rebernik Ribic, P. A simplified description of X-ray free-electron lasers. *Journal of synchrotron radiation* **2011**, *18*, 101–108.
(51) Neutze, R.; Wouts, R.; van der Spoel, D.; Weckert, E.; Hajdu, J. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* **2000**, *406*, 752–757.
(52) 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.
(53) 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.*; et al.* Time-resolved protein nanocrystallography using an X-ray free-electron laser. *Optics express* **2012**, *20*, 2706–2716.
(54) Nass Kovacs, G.; Colletier, J.-P.; Grünbein, M. L.; Yang, Y.; Stensitzki, T.; Batyuk, A.; Carbajo, S.; Doak, R. B.; Ehrenberg, D.; Foucar, L.*; et al.* Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin. *Nature communications* **2019**, *10*, 3177.
(55) Decking, W.; Abeghyan, S.; Abramian, P.; Abramsky, A.; Aguirre, A.; Albrecht, C.; Alou, P.; Altarelli, M.; Altmann, P.; Amyan, K.*; et al.* A MHz-repetition-rate hard X-ray free-electron laser driven by a superconducting linear accelerator. *Nature Photon* **2020**, *14*, 391–397.
(56) Mozzarelli, A.; Rossi, G. L. Protein function in the crystal. *Annual review of biophysics and biomolecular structure* **1996**, *25*, 343–365.
(57) Miller, R. J. D.; Paré-Labrosse, O.; Sarracini, A.; Besaw, J. E. Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin in the multiphoton regime and biological relevance. *Nature communications* **2020**, *11*, 1240.
(58) Grünbein, M. L.; Stricker, M.; Nass Kovacs, G.; Kloos, M.; Doak, R. B.; Shoeman, R. L.; Reinstein, J.; Lecler, S.; Haacke, S.; Schlichting, I. Illumination guidelines for ultrafast pump-probe experiments by serial femtosecond crystallography. *Nature methods* **2020**, *17*, 681–684.
(59) Barends, T. R. M.; 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 (New York, N.Y.)* **2015**, *350*, 445–450.
(60) Nogly, P.; Weinert, T.; James, D.; Carbajo, S.; Ozerov, D.; Furrer, A.; Gashi, D.; Borin, V.; Skopintsev, P.; Jaeger, K.*; et al.* Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser. *Science (New York, N.Y.)* **2018**, *361*.
(61) Pande, K.; Hutchison, C. D. M.; Groenhof, G.; Aquila, A.; Robinson, J. S.; Tenboer, J.; Basu, S.; Boutet, S.; DePonte, D. P.; Liang, M.*; et al.* Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein. *Science (New York, N.Y.)* **2016**, *352*, 725–729.
(62) Weinert, T.; Skopintsev, P.; James, D.; Dworkowski, F.; Panepucci, E.; Kekilli, D.; Furrer, A.; Brünle, S.; Mous, S.; Ozerov, D.*; et al.* Proton uptake mechanism in bacteriorhodopsin captured by serial synchrotron crystallography. *Science (New York, N.Y.)* **2019**, *365*, 61–65.

(63) Arnlund, D.; Johansson, L. C.; Wickstrand, C.; Barty, A.; Williams, G. J.; Malmerberg, E.; Davidsson, J.; Milathianaki, D.; DePonte, D. P.; Shoeman, R. L.*; et al.* Visualizing a protein quake with time-resolved X-ray scattering at a free-electron laser. *Nature methods* **2014**, *11*, 923–926.
(64) Berera, R.; van Grondelle, R.; Kennis, J. T. M. Ultrafast transient absorption spectroscopy: principles and application to photosynthetic systems. *Photosynthesis research* **2009**, *101*, 105–118.
(65) Ruckebusch, C.; Sliwa, M.; Pernot, P.; Juan, A. de; Tauler, R. Comprehensive data analysis of femtosecond transient absorption spectra: A review. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* **2012**, *13*, 1–27.
(66) van Stokkum, I. H. M.; Larsen, D. S.; van Grondelle, R. Global and target analysis of time-resolved spectra. *Biochimica et biophysica acta* **2004**, *1657*, 82–104.
(67) Mouton, N.; Sliwa, M.; Buntinx, G.; Ruckebusch, C. Deconvolution of femtosecond time-resolved spectroscopy data in multivariate curve resolution. Application to the characterization of ultrafast photo-induced intramolecular proton transfer. *J. Chemometrics* **2010**, *24*, 424–433.
(68) Beechem, J. M.; Ameloot, M.; Brand, L. Global and Target Analysis of Complex Decay Phenomena. *Instrumentation Science & Technology* **1985**, *14*, 379–402.
(69) Meuwis, K.; Depuydt, G.; Boens, N.; C. De Schryver, F. Comparison of simultaneous biexponential and compartmental analyses of fluorescence decay surfaces of intermolecular two-state excited-state processes. *Chemical Physics Letters* **1995**, *246*, 641–648.
(70) Rooi, J. J. de; Devos, O.; Sliwa, M.; Ruckebusch, C.; Eilers, P. H. C. Mixture models for two-dimensional baseline correction, applied to artifact elimination in time-resolved spectroscopy. *Analytica chimica acta* **2013**, *771*, 7–13.
(71) Anscombe, F. J. Graphs in Statistical Analysis. *The American Statistician* **1973**, *27*, 17–21.
(72) Golub, G. H.; van Loan, C. F. *Matrix computations*, 3rd ed.; Johns Hopkins studies in the mathematical sciences; Johns Hopkins University Press: Baltimore, 1996.
(73) Costa, K. D.; Kleinstein, S. H.; Hershberg, U. Biomedical model fitting and error analysis. *Science signaling* **2011**, *4*, tr9.
(74) Istratov, A. A.; Vyvenko, O. F. Exponential analysis in physical phenomena. *Review of Scientific Instruments* **1999**, *70*, 1233–1257.
(75) *Hydrogen-transfer reactions*; Wiley-VCH-Verl.: Weinheim, 2007.
(76) Marcus, R. A. On the Analytical Mechanics of Chemical Reactions. Quantum Mechanics of Linear Collisions. *The Journal of Chemical Physics* **1966**, *45*, 4493–4499.
(77) Limbach, H.-H. Single and Multiple Hydrogen/Deuterium Transfer Reactions in Liquids and Solids. In *Hydrogen-Transfer Reactions*; Hynes, J. T., Klinman, J. P., Limbach, H.-H., Schowen, R. L., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2006; pp 135–221.
(78) Chen, K.-Y.; Cheng, Y.-M.; Lai, C.-H.; Hsu, C.-C.; Ho, M.-L.; Lee, G.-H.; Chou, P.-T. Ortho green fluorescence protein synthetic chromophore; excited-state intramolecular proton transfer via a seven-membered-ring hydrogen-bonding system. *Journal of the American Chemical Society* **2007**, *129*, 4534–4535.
(79) Hsieh, C.-C.; Chou, P.-T.; Shih, C.-W.; Chuang, W.-T.; Chung, M.-W.; Lee, J.; Joo, T. Comprehensive studies on an overall proton transfer cycle of the ortho-green fluorescent protein chromophore. *Journal of the American Chemical Society* **2011**, *133*, 2932–2943.

(80) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormö, M.; Remington, S. J. Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 2306–2311.
(81) Fang, C.; Frontiera, R. R.; Tran, R.; Mathies, R. A. Mapping GFP structure evolution during proton transfer with femtosecond Raman spectroscopy. *Nature* **2009**, *462*, 200–204.
(82) Fujisawa, T.; Kuramochi, H.; Hosoi, H.; Takeuchi, S.; Tahara, T. Role of Coherent Low-Frequency Motion in Excited-State Proton Transfer of Green Fluorescent Protein Studied by Time-Resolved Impulsive Stimulated Raman Spectroscopy. *Journal of the American Chemical Society* **2016**, *138*, 3942–3945.
(83) Kasha, M. Proton-transfer spectroscopy. Perturbation of the tautomerization potential. *Journal of the Chemical Society, Faraday Transactions 2* **1986**, *82*, 2379.
(84) Itoh, M.; Adachi, T.; Tokumura, K. Transient absorption and two-step laser excitation fluorescence spectra of the excited-state and ground-state proton transfer in 7-hydroxyquinoline. *Journal of American Chemical Society* **1983**, *105*, 4828–4829.
(85) Lochbrunner, S.; Schriever, C.; Riedle, E. Direct Observation of Nuclear Motion during Ultrafast Intramolecular Proton Transfer. In *Hydrogen-Transfer Reactions*; Hynes, J. T., Klinman, J. P., Limbach, H.-H., Schowen, R. L., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2006; pp 349–375.
(86) Lochbrunner, S.; Stock, K.; Riedle, E. Direct observation of the nuclear motion during ultrafast intramolecular proton transfer. *Journal of Molecular Structure* **2004**, *700*, 13–18.
(87) Lochbrunner, S.; Wurzer, A. J.; Riedle, E. Ultrafast excited-state proton transfer and subsequent coherent skeletal motion of 2-(2 ′ -hydroxyphenyl)benzothiazole. *The Journal of Chemical Physics* **2000**, *112*, 10699–10702.
(88) Litvinenko, K. L.; Meech, S. R. Observation of low frequency vibrational modes in a mutant of the green fluorescent protein. *Physical Chemistry Chemical Physics.* **2004**, *6*, 2012.
(89) Theodore. Förster. Die p H-Abhängigkeit der Fluoreszenz von Naphthalinderivaten. *eitschrift für Elektrochemie und angewandte physikalische Chemie,* **1950**, 531–535.
(90) Laptenok, S. P.; Conyard, J.; Page, P. C. B.; Chan, Y.; You, M.; Jaffrey, S. R.; Meech, S. R. Photoacid Behaviour in a Fluorinated Green Fluorescent Protein Chromophore: Ultrafast Formation of Anion and Zwitterion States.†. *Chemical science* **2016**, *7*, 5747–5752.
(91) Tolbert, L. M.; Solntsev, K. M. Design and Implementation of“Super” Photoacids. In *Hydrogen-Transfer Reactions*; Hynes, J. T., Klinman, J. P., Limbach, H.-H., Schowen, R. L., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2006; pp 417–439.
(92) Dugave, C., Ed. *Cis-trans isomerization in biochemistry*; Willey: Weinheim, 2006.
(93) Dugave, C.; Demange, L. Cis-trans isomerization of organic molecules and biomolecules: implications and applications. *Chemical reviews* **2003**, *103*, 2475–2532.
(94) Liu, R. S. Photoisomerization by hula-twist: a fundamental supramolecular photochemical reaction. *Accounts of chemical research* **2001**, *34*, 555–562.
(95) Liu, R. S. H.; Yang, L.-Y.; Zhao, Y.-P.; Kawanabe, A.; Kandori, H. Supramolecular Effects on Mechanisms of Photoisomerization: Hula Twist, Bicycle Pedal, and One-Bond-Flip. In

Supramolecular Photochemistry; Ramamurthy, V., Inoue, Y., Eds.; John Wiley & Sons, Inc: Hoboken, NJ, USA, 2011; pp 547–570.
(96) Mathies, R. A. Photons, femtoseconds and dipolar interactions: a molecular picture of the primary events in vision. *Novartis Foundation symposium* **1999**, *224*, 70-84; discussion 84-101.
(97) Hargrave, P. A.; McDowell, J. H.; Curtis, D. R.; Wang, J. K.; Juszczak, E.; Fong, S. L.; Rao, J. K.; Argos, P. The structure of bovine rhodopsin. *Biophysics of structure and mechanism* **1983**, *9*, 235–244.
(98) Warshel, A. Bicycle-pedal model for the first step in the vision process. *Nature* **1976**, *260*, 679–683.
(99) Liu, R. S.; Asato, A. E. The primary process of vision and the structure of bathorhodopsin: a mechanism for photoisomerization of polyenes. *Proceedings of the National Academy of Sciences of the United States of America* **1985**, *82*, 259–263.
(100) Ivn, K. J. Organic chemistry: The name game. A. Nickon and E. F. Silversmith, Pergamon Press, Oxford, 1987. *Polym. Int.* **1991**, *24*, 256.
(101) Liu, R. S.; Hammond, G. S. The case of medium-dependent dual mechanisms for photoisomerization: one-bond-flip and hula-twist. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 11153–11158.
(102) Ugo. Mazzucato/Fabio. Momicchioli. Rotational isomerism in trans-1,2-diarylethylenes.
(103) Fuß, W. Hula-twist cis–trans isomerization: The role of internal forces and the origin of regioselectivity. *Journal of Photochemistry and Photobiology A: Chemistry* **2012**, *237*, 53–63.
(104) Scoponi, M.; Gallinella, E.; Momicchioli, F. Rotational isomerism in trans-1,2-diarylethylenes. A Raman study of conformational equilibrium of styrylnaphthalenes. *Journal of the Chemical Society, Faraday Transactions 2* **1988**, *84*, 95.
(105) Sampedro Ruiz, D.; Cembran, A.; Garavelli, M.; Olivucci, M.; Fuß, W. Structure of the Conical Intersections Driving the cis–trans Photoisomerization of Conjugated Molecules¶. *Photochem Photobiol* **2002**, *76*, 622.
(106) Fuss, W.; Kosmidis, C.; Schmid, W. E.; Trushin, S. A. The photochemical cis-trans isomerization of free stilbene molecules follows a hula-twist pathway. *Angewandte Chemie (International ed. in English)* **2004**, *43*, 4178–4182.
(107) Imamoto, Y.; Kuroda, T.; Kataoka, M.; Shevyakov, S.; Krishnamoorthy, G.; Liu, R. S. H. Photoisomerization by hula twist: 2,2'-dimethylstilbene and a ring-fused analogue. *Angewandte Chemie (International ed. in English)* **2003**, *42*, 3630–3633.
(108) van der Meer, M.; Zhang, H.; Rettig, W.; Glasbeek, M. Femto- and picosecond fluorescence studies of solvation and non-radiative deactivation of ionic styryl dyes in liquid solution. *Chemical Physics Letters* **2000**, *320*, 673–680.
(109) Gerwien, A.; Schildhauer, M.; Thumser, S.; Mayer, P.; Dube, H. Direct evidence for hula twist and single-bond rotation photoproducts. *Nature communications* **2018**, *9*, 2510.
(110) Müller, A. M.; Lochbrunner, S.; Schmid, W. E.; Fuß, W. Low-Temperature Photochemistry of Previtamin D: A Hula-Twist Isomerization of a Triene. *Angewandte Chemie (International ed. in English)* **1998**, *37*, 505–507.
(111) Redwood, C.; Bayda, M.; Saltiel, J. Photoisomerization of Pre- and Provitamin D3 in EPA at 77 K: One-Bond-Twist, Not Hula-Twist. *The journal of physical chemistry letters* **2013**, *4*, 716–721.

(112) Fuß, W. Previtamin D: Z-E photoisomerization via a Hula-twist conical intersection. *Physical chemistry chemical physics : PCCP* **2019**, *21*, 6776–6789.
(113) Mustalahti, S.; Morozov, D.; Luk, H. L.; Pallerla, R. R.; Myllyperkiö, P.; Pettersson, M.; Pihko, P. M.; Groenhof, G. Photoactive Yellow Protein Chromophore Photoisomerizes around a Single Bond if the Double Bond Is Locked. *The journal of physical chemistry letters* **2020**, *11*, 2177–2181.

3 Photodynamics of photoswitchable fluorescent proteins (RSFPs): review and thesis's objectives

The discovery of the avGFP is the FPs starting point. Many of these proteins, including the rsEGFP2, share the same fluorescent chromophore: HBDI. Therefore, before the literature review of the RSFPs photodynamic, we will start by describing GFP fluorescent proteins and their general chemical-physical steady-state properties, including some RSFPs specific properties. Once the basic photophysical properties are established, in a second section, we will focus on the photodynamics of HBDI, avGFP and RSFPs in order to introduce the objectives and methodologies of the thesis.

3.1 Chemical and physical properties of GFP proteins: towards RSFPs

3.1.1 Introduction

The chromophore is the origin of fluorescence in GFP (Figure 3.1). The organic name given to this chromophore is 4-(p-hydroxybenzylidene) 5-imidazolinone (HBDI). Few years after the elucidation of avGFP structure, the chromophore was synthesized and thoroughly studied. The HBDI inside the avGFP can exist in both anionic and neutral forms with a pK_a value equal to 8.1[1]. The pK_a was assigned to the acid/basic equilibrium of the hydroxyl group. At low pH, the nitrogen of the imidazolinone group can also be protonated with an associated pK_a of 1.4[2]. At alkaline pH, the HBDI absorption spectrum has a maximum at 448 nm with a molar absorption coefficient of 44100 ($M^{-1}cm^{-1}$)[2], while this maximum is shifted to 384 nm at neutral or acidic pH. The HBDI is characterized by two aromatic rings linked together with a carbon double bond which forms a π-electron cloud where the electrons can be delocalized on the entire molecule and the existence of different mesomeric forms. The first allowed

dipole transition for the anionic form corresponds to the excitation of one electron from the HOMO to LUMO orbital (π –π* transition).

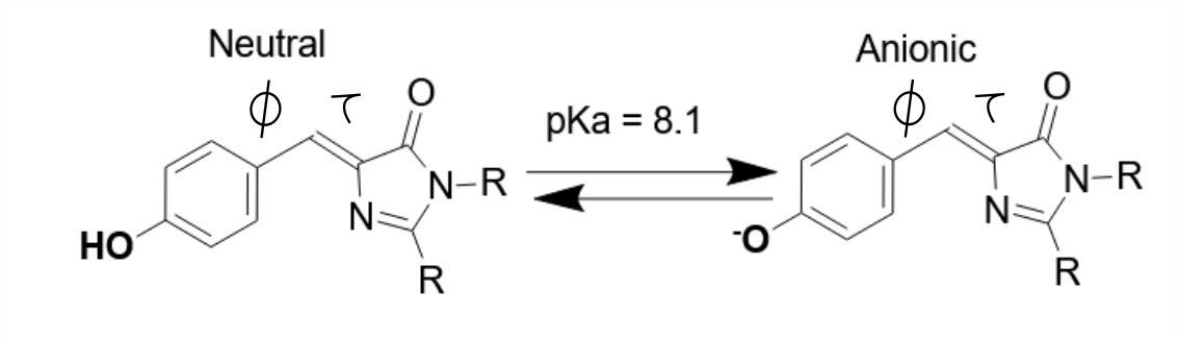


Figure 3.1. Neutral and anionic cis-forms of HBDI chromophore and its pK_a *value. Figure made with https://chemdrawdirect.perkinelmer.cloud.*

Depending on the two dihedral angles τ and ϕ that bridge the phenol and imidazolinone rings, the HBDI can adopt different possible geometries (*cis/trans* isomers). In the *cis* isomer, the phenol and imidazolinone rings are on the same side of the methylene double bond, while in the *trans* isomer they are in opposite sides. These angles are related to the efficiency of internal conversion and the existence of non-radiative pathways, which explain the absence of fluorescence of the chromophore in solution and unfolded proteins. These two results indicated that the non-constrained chromophore main excited-state deactivation is an internal conversion (IC) process. The absence of fluorescence is naturally in contrast with the high fluorescence quantum yield of avGFP, where the chromophore adopts *cis* neutral and anionic forms (Φ_F: 0.78 for the neutral form at 366 nm excitation, and 0.72 for anionic form at 470 nm excitation[3]). The chromophore rigidity inside the protein matrix is one the main reasons explaining these high fluorescence quantum yield values. Indeed, several studies have demonstrated that the fluorescence becomes again the main deactivation process for the chromophore in solution at 77 K[4,5] or when it is encapsulated in non-protein scaffolds[6] and non-native proteins such as human serum albumin[7].These results demonstrate that HBDI geometry and its interaction with the protein cage defines the photophysical properties of GFP proteins.

Three decades after the discovery of the avGFP by Osamu Shimomura in the sixties[8], the sequence of the protein consisting of 238 amino acid residues was elucidated by Prasher et

al. in 1992[9]. It was possible to isolate the GFP gene and consequently clone the protein[10,11]. In the following year, new studies showed that the HBDI chromophore is formed by the protein-peptide chain itself[12]. The chromophore of the avGFP protein is indeed formed by an autocatalytic cyclization reaction during protein maturation between the amino acid residues, Ser65, Gly67 and the oxidation of the α-β bond of the Tyr66[9,12] (see Figure 3.2). After the protein maturation and folding, the HBDI chromophore is localized in the center of a β-barrel covalently bound to the protein mainchain[13]. In 1996, the first X-ray crystallography structure was released, by Ormö et al.[13]. Nearly simultaneously and on the same year, another X-ray structure of its enhanced version EGFP was also elucidated by Yang et al.[14]. Ormö et al. determined that the protein fold consists of a 11-stranded β-barrel with a coaxial α-helix, with the chromophore formed in the central helix[13].This structure confers rigidity to the HBDI . The β barrel and the chromophore can be seen in Figure 3.2.

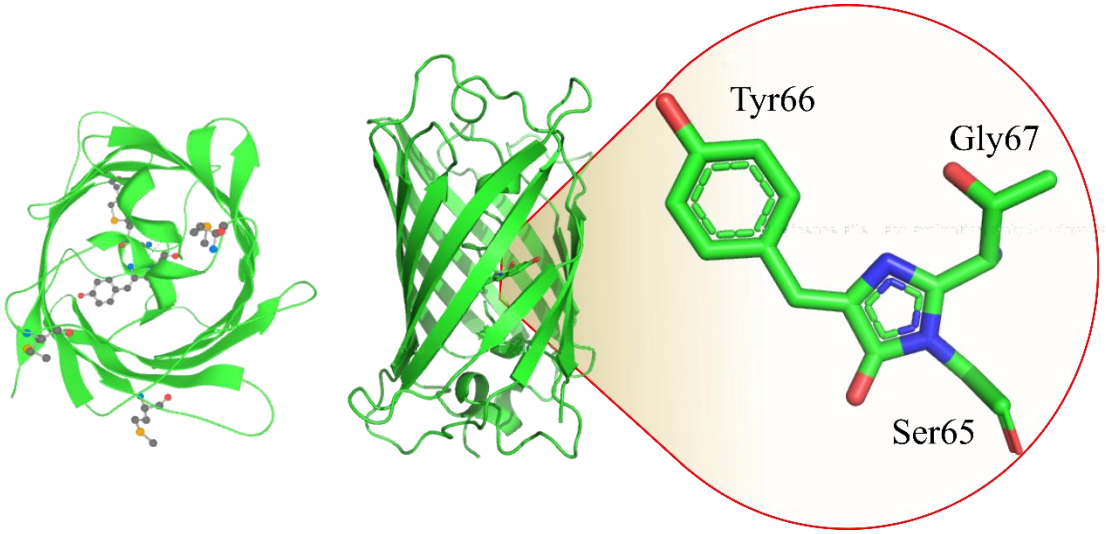


Figure 3.2. Aqua victoria Green Fluorescent Protein (avGFP) 3D structure, top and lateral perspective, represented together with the HBDI chromophore. The 3D representation, and chromophore has been done using pyMol from PDB generated by Ormö et al. 1996[13] (1EMA). The images have been combined using Inkscape.

The steady state absorption (blue) and emission spectrum after 400 nm excitation (green) at physiological pH for avGFP are shown in Figure 3.3. There are mainly two absorption maxima at 395 and 477 nm, which are logically assigned to neutral and anionic forms (also named A and B respectively) of HBDI respectively[15] (phenol/phenolate, Figure 3.1). At

physiological pH, there is an equilibrium between these two forms with predominance of neutral form in agreement with a pK_a of 8.1. After excitation with UV light, the fluorescence spectrum has its maximum at 510 nm while after 477 nm excitation the emission maximum is blue-shifted and peaks at 503 nm[10]. The 110 nm Stokes-shift after UV light excitation is explained by an excited-state proton transfer (ESPT).

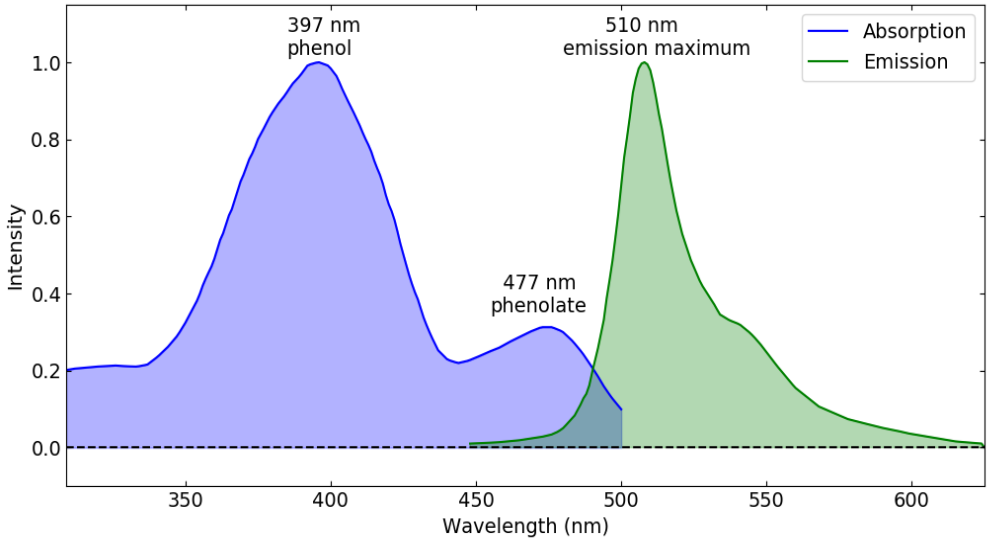


Figure 3.3. The absorption spectrum of avGFP at physiological pH (7.5) is represented in blue and the emission spectrum of avGFP after excitation with UV light is represented in green. The spectra have been reproduced from the fluorescent protein database were the avGFP associated code is 1XF1B[16] and plot using matplotlib python package.

A few number of mutations or even single point mutations can induce a variety of photophysical properties, including: i. red or blue shifts in the absorption or emission maxima, ii. changes in fluorescence stability, iii. changes in the fluorescent lifetime or the fluorescent quantum yield, and iv. the protein can even become photochromic with few muations[17]. Despite the enormous differences that can be found in optical properties of the different GFP proteins, their structure in form of a β-barrel is maintained across the GFP like proteins, with very few modifications among them[17,18]. This indicates that the protein cage and particularly the amino acids in the surroundings of the chromophore modified the photophysical properties of the HBDI chromophore. The protein should always be treated as a whole since the final photophysical properties do not only depend on the potential surfaces of the possible conformations of the chromophore, but also on the interactions with the

surrounding amino acids, and probably as well on the structural water molecules observed in the crystal structures which form complex hydrogen bond networks[18].

This vast range of variations among the GFP-like properties greatly extends their utility. The carboxyl and amino terminal sides at the ends of the GFP protein chain allow tagging other proteins by fusion at either ends of the protein of interest. The specificity to target different proteins makes the GFP-like proteins an excellent tool for live-cell imaging[19], e.g. different elements of a cell can be labelled with different proteins having different emission wavelengths[20]. GFP-like proteins can be classified depending on their photophysical properties and will be described in the next section.

3.1.2 Classification of fluorescent proteins

In 2002, only 22 fluorescent protein structures were listed in the protein data bank[21]. These 22 proteins were classified by Tsien into seven different types [15]: (For amino acid abbreviations refer to Appendix 3, Table A3.1)

- **Type I**: Proteins characterized by a visible absorption spectrum with two maxima corresponding to neutral and anionic forms. An example is avGFP.
- **Type II**: Proteins characterized by a visible absorption spectrum with one maximum corresponding to the phenolate anionic form due to mutation of the S65 by a Thr, Ala, or Gly. An example of this type is the EGFP[22].
- **Type III**: Proteins characterized by a visible absorption spectrum with one maximum corresponding to the phenol neutral form, attributed to the mutation of T203. An example is Sapphire[23].
- **Type IV**: Proteins having an anionic phenolate chromophore stacked by π-electron systems, due to a mutation of T203 to His, Trp, Phe, or Tyr. An example is the EYFP[13] and in general, are yellow fluorescent proteins.
- **Type V**: Proteins having a chromophore with an indole group. An example is the CFP[10]. These proteins have properties in between those of BFP[24] (Type VI) and EGFP[22] (Type II).

- **Type VI**: Proteins characterized by a chromophore with a phenyl and imidazole ring. Such proteins have an absorption spectrum with a maximum at 380 nm. An example is the BFP[24].
- **Type VII**: Proteins characterized by an excitation spectrum with a maximum below 300 nm.

Nowadays, there are more than 820 structures listed in the protein data bank (entering GFP as key research word) with proteins emitting fluorescence at every different colour of the visible spectrum[25] (Figure 3.5). The field has grown so fast that there is a separate fluorescent protein data base[26] where authors can upload new discovered or synthesized fluorescent proteins, with currently 773 different FP listed. With these many different proteins, the previously given classification is obsolete, and nowadays the most common classification is according to their light-induced changes, which define the applications exploiting them. The most basic application is their use as fluorescent tags. Proteins that after light-excitation fluoresce and have a photo-transformation are classified as photo-transformable FPs (PTFP). These can be further classified into three different types: photoactivatable, photoconvertible or photoswitchable (Figure 3.4).

- Photoactivatable (PAFP): Proteins that are able to go irreversibly from a dark non-fluorescent state to a bright state with light. Some examples of this type of proteins are PA-GFP[27] and PA-mCherry2[28].
- Photoconvertible (PCFP): PAFP Proteins having two distinct emissive states and are able to go irreversibly from one to another. For example, a green emitting state, that after irradiation with the correct wavelength goes irreversibly to a red emitting state. Examples of these proteins are Kaede[29], Dendra[30] and EosFP[31].
- Photoswitchable (RSFP): Reversible PAFP Proteins or photochromic FPs, i.e proteins that can be switched reversibly from a bright fluorescent (*On*) state and a dark (*Off*) non fluorescent state. Examples of these proteins are Dronpa[32], Dreiklang[33], Skyline[34] or rsEGFP2[35,36].

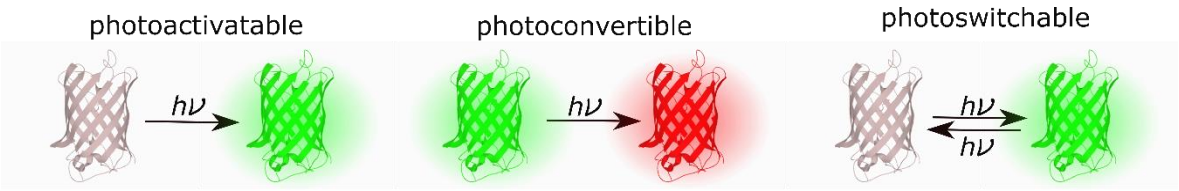


Figure 3.4. Light-induced phenomena in fluorescent proteins that allows classifying the phototransformable FPs. Figure made with Inkscape.

It is worth mentioning that a fluorescent protein can have simultaneously different types of phototransformations. For example, IrisFP[37] is photoconvertible from a green to a red fluorescent form, and each of these forms is photoswitchable[37]. In the FP database, a significant number of proteins correspond to those named "basic FP"[38] (not-PTFP). Nevertheless, during the last couple of years, the number of photoactivatable, photoconvertible and photoswitchable proteins has drastically increased, mainly due to their applications in super-resolution microscopies. In Figure 3.5, the maximum of fluorescence emission spectra for all FPs listed in the FP database are plotted against their absorption maximum; the colours represent the different type of proteins according to the classification mentioned above. Interestingly, PTFPs are grouped in specific clusters with similar absorption and fluorescence maxima. On the contrary, non-PTFP proteins cover the entire visible domain. PAFPs are quite rare, and the majority of them absorb and emit in the 590-630 nm region. RSFPs emit in a narrow region between 500 and 530 nm range, excluding the photoconverted red forms (i.e. red-IrisFP), and a few proteins emitting above 600 nm with low fluorescent quantum yield and brightness. PCFPs cover a broader emission region, but still only PSmOrange[39] and PSmOrange2[40] can be excited in the red. The development of new phototransformable proteins which can be excited above 600 nm is a challenge. [25].

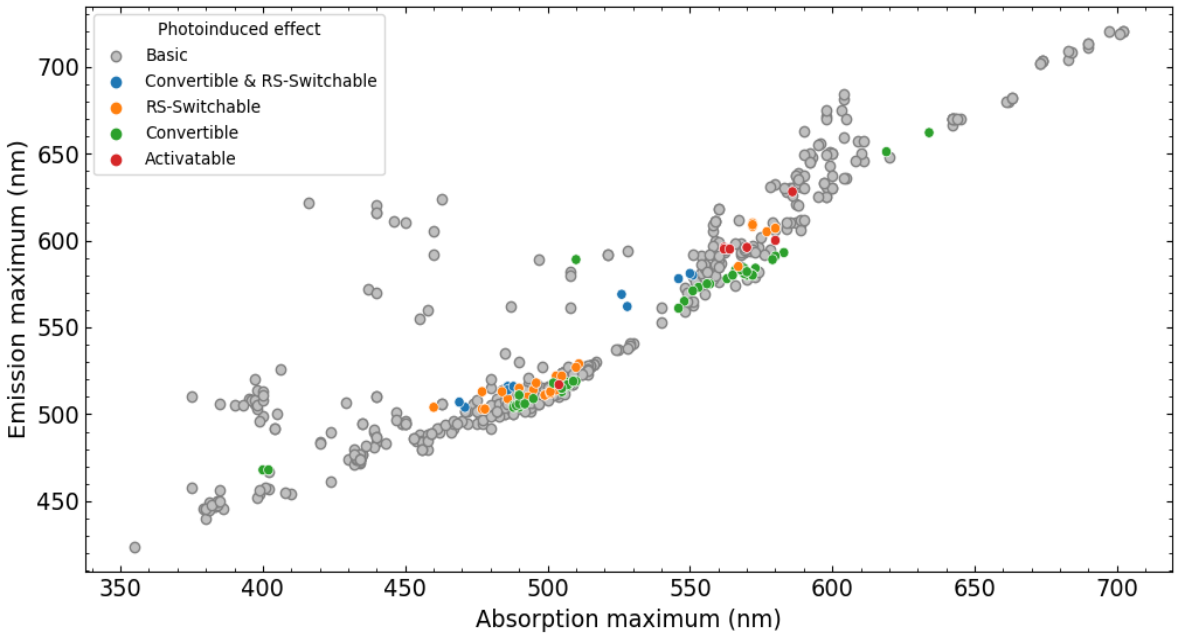


Figure 3.5. Fluorescence emission maximum of proteins listed in the FP database[26] represented according to their absorption maximum. The colours represent the different type of FPs: grey, blue, orange, and green for basic, RSFP, PCFP and PAFP respectively. Figure is plotted using matplotlib and seaborn python packages. (https://www.fpbase.org)

3.1.3 Correlations between structure and photophysical properties of FPs

The mechanism behind the different photo-transformations differs and their understanding for the design of new RSFPs is essential. For example, generally in PAFP, the photoactivation induces a decarboxylation of the glutamine E222 and a conversion from a neutral non-fluorescent state to an anionic fluorescent state[41]. For PCFPS, Kaede[29], Dendra[30] and EosFP[31], the mechanism of the photoconversion is a chemical modification which extends the π-system and breaks the protein backbone[25]. For RSFPs such as dronpa[32] or rsEGFP2[35,36], the photoswitching mechanism between the fluorescent and non-fluorescent states is based on a *cis-trans* isomerization and a protonation/de-protonation of the chromophore. Interestingly, for Dreiklang[33], the photoswitching mechanism is a photoinduced hydration/dehydration of the imidazolinone ring.

It is worth noting that there is not an established link between the FP structure and its photophysical properties (i.e. maxima of absorption and emission spectra, molar absorption coefficient, fluorescence quantum yield, brightness, fluorescence lifetime, switching yield, switching contrast, photostability, fatigue resistance, to name a few). As pointed out in an extensive study where the fluorescence of 51,715 proteins obtained by random mutagenesis of the avGFP was monitored[42], the prediction of properties for new FPs is challenging. All the proteins studied had between 1 and 15 mutations, with an average of 3.7 mutations per protein. They reported that 75% of the proteins with a single point mutation had a negative effect on fluorescence and ~10% of them had their fluorescence emission intensity (brightness) divided by 5 in comparison to the WT avGFP[42]. This study showed that the mutations affecting the most the fluorescence quantum yield are those in the amino acid residues oriented internally towards the chromophore. Furthermore, they generally observed that multiple mutations result in a more substantial negative effect than the sum of individual mutations[42].

Rationally designing new mutations is therefore essential for the development of new RSFPs with optimized photo-physical properties. In a recent publication[43], the group of Steven G. Boxer used an innovative technique to develop new FP mutants. This technique is based on FP split proteins[43–45]. Split proteins are two polypeptide chains having complementary parts of a complete entire protein sequence. These two peptides can then be fused together to form a complete functional protein[46–48]. A schematic representation of this technique applied to GFP type proteins is illustrated in Figure 3.6. The group of Boxer have studied the association of split proteins related to GFP for over a decade, discovering several unique light-dependent photodissociation and photoassociation of split FP fragments[49,50]. Using split proteins, they managed to precisely remove and substitute any amino acid in the avGFP protein sequence. They could incorporate non-canonical amino acids introducing electron-withdrawing and donating moieties directly to the chromophore and make several mutants targeting the different amino acids interacting directly with HBDI (R96, H148 T203, E222 and S65)[51] (see appendix 3 for amino acid table and chromophores environments). They were then able to study the influence that the different amino acid residues had on several photo-physical properties of avGFP[43] (Figure 3.7).

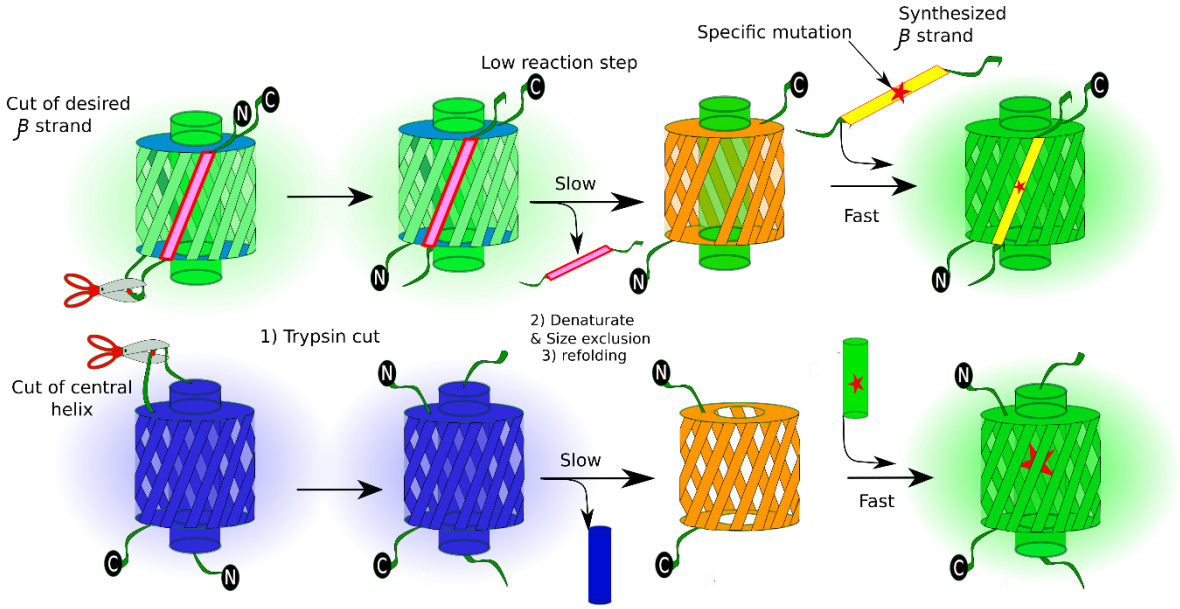


Figure 3.6. Semisynthetic route for creating new GFP mutants based on Split proteins. Proteins containing a trypsin cleavage inserted in a loop will be broken when the protein is digested with trypsin (step 1), this is followed by denaturation with guanidine hydrochloride and isolation of the larger piece of GFP by size exclusion. The dilution of these larger pieces into a solution containing the desire synthetic peptide yields the final modified protein. This can be applied to remove any specific β-strand (top) or the central chromophore (bottom). Figure made with Inkscape based on figure 2 from reference 45.

They further showed that electronic interactions of the amino acids with the chromophore control, among other properties, the emission maximum of fluorescence spectrum. Indeed, there is a correlation between the value of Stokes shift and absorption maximum for the fluorescence of a protein. The correlation follows the equations from the Marcus−Hush theory for mixed-valence compounds[52]. The HBDI chromophores can exist with different forms and is treated as a mixed-valence compound that can either have the charge on the oxygen of the phenol ring or the oxygen of the imidazolinone ring. The negative charge on either of the two rings is stabilized by electrostatic interactions with amino-acids of the protein [43]. For mutations on the Tyr66, the states depend on the capacity of the introduced chemical group to donate or withdraw electrons from HBDI[43]. They also studied spontaneous emission rate (*Kr* / fluorescence) relations with the molar absorption, transition dipole

moment, and absorption spectrum maxima. They reached to the conclusion that it is only possible to tune efficiently (several orders of magnitude) the fluorescence quantum yield of a chromophore by modifying the efficiency of competing nonradiative decay pathways, such as isomerization or excited state electron transfer[43].

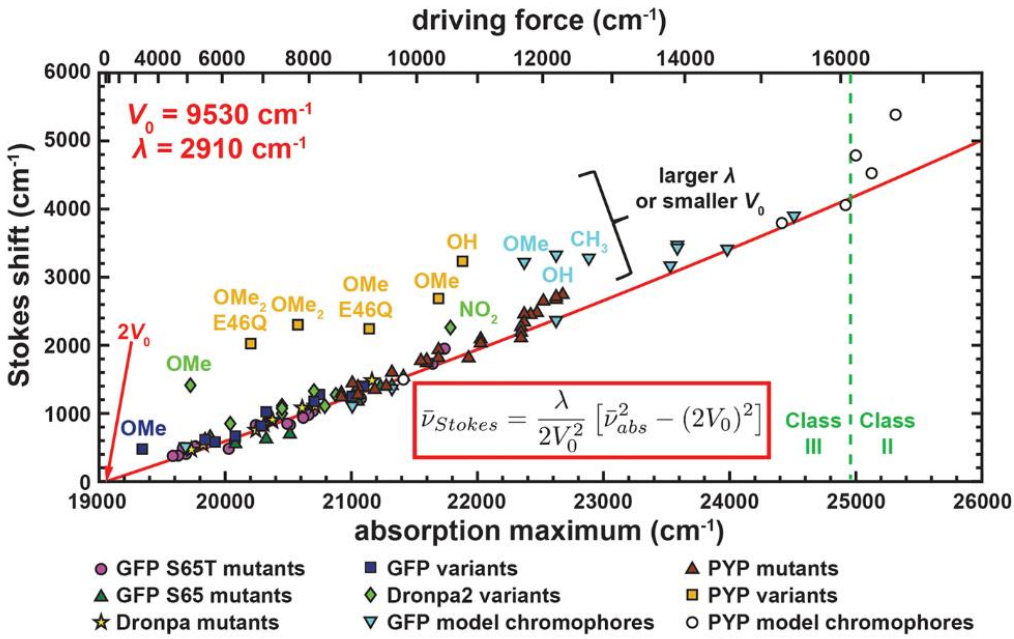


Figure 3.7. The Stokes shift plotted against the absorption maximum for all studied mutants in reference [43], variants, and model chromophores of GFP, Dronpa, and PYP at room temperature, The red line corresponds to the fit of data to the equation inserted in the figure. This Figure has been reproduced from reference [43].

3.1.4 Photophysical properties of RSFPs

As mentioned above, RSFPs are photochromic proteins that can be switched reversibly from a bright fluorescent (*On*) state and a dark state (*Off*). In addition to the previous classifications, RSFPs can be further classified according to their switching mode in three groups[53]:

- Negative RSFPs: Irradiation light induces switching from *On* to *Off* state and fluorescence. These types are the most common RSFP. Except for rsGamilus[54], in the

On fluorescent state the chromophore adopts a *cis* anionic form, while in the *Off* non-fluorescent state adopts *trans* neutral form.

- Positive RSFPs: Irradiation light induces switching from *Off* to *On* state and fluorescence. In general, the *Off* non-fluorescent state is formed by the *trans* anionic form of the chromophore, which, when irradiated, provokes an isomerization to the *cis* form, which is found in equilibrium between the protonated and deprotonated forms, the latest is the fluorescent form (*On*). The deprotonated *cis* form has an absorption band maximum close to the trans *Off* state. It is important to note that all positive RSFP have an Anthozoa origin and are derived from Dronpa[53].
- Decoupled RSFPs: These proteins have a decoupled fluorescence excitation from the irradiation switching light. There are only two reported proteins of this type, Dreiklang[33] and a mutant of it Spoon[55]. The switching between states is not based on a *cis-trans* isomerization but on a photoinduced hydration/dehydration of the imidazolinone ring. Indeed, in its *On* state, the chromophore is found in an equilibrium between the anionic and neutral forms. Excitation of the anionic form induces fluorescence, while irradiation of the neutral form induces the hydration to the *Off* state.

In addition to the parameters that characterized general FPs, for RSFPs, it is essential to characterize the different spectral forms, switching quantum yields between the fluorescent and the non-fluorescent state and, for super-resolution microscopy, the switching contrast and fatigue. There are currently 42 RSFP including the different forms of RSFP that can photoconvert from a green fluorescent state to a red fluorescent state. The majority of publications have focused on studying the photophysical properties that directly influence the resolution in RESOLFT, such as brightness (contrast, molar absorption coefficient, brightness and fluorescence quantum yield), fatigue resistance, switching speed and switching contrast[53]. The switching contrast is determined by the crosstalk between *On* and *Off* switching and the molecular brightness of *On* and *Off* states. Furthermore, fast thermal relaxation of the switched protein to the respective equilibrium state, as well as the population of intermediate states, can affect the reachable switching contrast. RSFPs useful for RESOLFT nanoscopy exhibit a contrast higher than 10 (a residual fluorescence below 10%), although smaller values of the residual fluorescence have been reported and are beneficial.

In figure 3.8, the reported brightness for all the RSFP in the fluorescent protein database[26] have been plotted versus the fluorescence maximum.

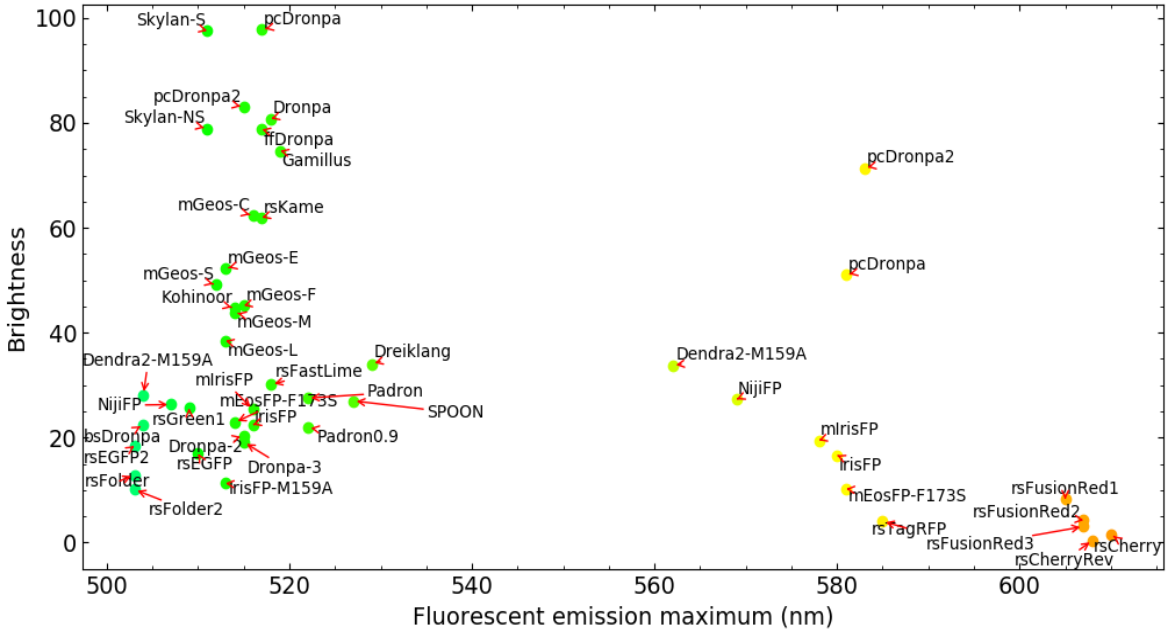


Figure 3.8. Brightness plotted against the fluorescence emission maximum for main RSFPs reported in the protein database[26]. The colour of the points corresponds to the wavelength value of the fluorescence emission maximum in RGB. Data plotted with matplotlib python package.

As commented in the introduction (Chapter 1), the switching quantum yields between the fluorescent and non-fluorescent states are parameters that can influence the acquisition times of images and the spectral resolution that can be achieved in RESOLFT[56]. However, the quantum yields of switching have not always been reported in the literature. These parameters are not easy to determine, and the reported values for a same protein vary between different publications (see Table 3.1). These differences can be attributed to the different methods used to meassure them; which are based on monitoring the changes invitro/invivo under irradiation of either, the fluorescence under a microscope, or the absorbance of the protein in solution. To our knowledge, out of the 42 RSFP reported in the fluorescent protein database[26], switching quantum yields have only been calculated for those proteins in Table 3.1. (a similar table with the molar absorption coefficients of most relevant RSFPs *On* state have been recently published by S. Jakobs and coworkers[53])

Table 3.1. Properties of RSFP with reported switching quantum yields.

Protein	Switch type	λ Absorption (nm) /λ Emission (nm)	ϕ_{fluo}	$\phi_{On\text{-}Off}$	$\phi_{Off\text{-}On}$	Ref
Dronpa	negative	503/517	0.85	0.032 e-2	0.37	[57]
			0.78	0.0160 e-2	0.165	[58]
			-	-	0.073	[59]
			0.93	0.043 e-2	0.21	[54]
Dronpa 2	negative	489/515	-	-	0.13	[59]
			0.28	4.7 e-2	-	[60]
ffDronpa	negative	507/519	0.75	0.007 e-2	0.172	[58]
pcDronpa (Green)	negative	505/517	0.85	0.056 e-4	0.092	[58]
rsGreen1	negative	486/509	0.42	0.42 e-2	0.14	[61]
rsGreenF	negative	486/509	0.39	0.87 e-2	0.18	[61]
rsEGFP	negative	491/510	0.42	0.55 e-2	0.17	[61]
rsEGFP2	negative	478/503	0.35	1.65 e-2	0.33	[62]
			0.34	0.089	0.12	[61]
			0.43	0.09	0.2	[54]
IrisFP (Green)	negative	488/516	0.43	1.4 e-2	0.5	[63]
Skyline-NS	negative		0.79	8.1 e-4	0.22	[54]
IrisFP (Red)	negative	551/580	0.47	0.2 e-2	0.047	[63]
mIrisFP (Green)	negative	486/516	0.54	0.69 e-2	0.36	[37]
mIrisFP (Red)	negative	546/578	0.59	0.31 e-2	0.28	[37]
Dendra2 M159A (Green)	negative	471/504	0.55	0.11 e-2	0.08	[64]
Dendra2 M159A (Red)	negative	528/562	0.75	0.32 e-2	0.01	[64]
NijiFP (Green)	negative	469/507	0.64	0.18 e-2	0.10	[64]
NijiFP (Red)	negative	526/569	0.65	0.10 e-2	0.10	[64]
mEosFP (Green)	negative	504/516	0.67	0.26 e-2	0.15	[64]
mEosFP (Red)	negative	569/581	0.66	0.0035 e-2	0.05	[64]
rsFolder	negative	478/503	0.25	2.1 e-2	0.44	[62]
rsFolder2	negative	478/503	0.23	1.98 e-2	0.28	[62]
Padron	positive	503/522	0.64	0.3 e-2	0.04	[65]
Kohinoor	positive	495/514	0.71	2.0 e-2	0.15	[65]
rsGamillus-S	negative	504/522	0.70	0.56 e-3	0.50	[54]
rsGamillus-F	negative	505/522	0.71	2.7 e-3	0.51	[54]

3.2 Photodynamics of FPs.

As discussed above, the photophysical properties of each of the different PTFP proteins and correlations with their structure are relatively complex and challenging. This difficulty is definitively associated with the excited state dynamics of the RSFP. Indeed, after excitation, RSFPs display a richer excited state dynamics than basic FPs[66,67] including transition between several fluorescents and non-fluorescent states (in excited and ground states). In the next section, we will first review the photodynamics of HBDI followed by the GFP, and we will end with RSFPs.

3.2.1 Photodynamics of HBDI

The photodynamics of the HBDI have been extensively reviewed[70,71]. As mentioned in chapter 2, the photodynamics of HBDI and its analogues can involve ESPT and *cis-trans* isomerization. The HBDI stable isomer is "*cis*" and can have different protonation states (phenol or imidazolinone ring). Early studies done using quantum mechanics simulations suggested the *cis-trans* isomerization or twisting of the exo-methylene double bond to be the main deactivation process for HBDI, at room temperature[68]. All the possible protonation states of the chromophore phenol group were studied with time-resolved spectroscopies. Some common conclusion can be drawn[5,69–72] and summarized in a fast bi-exponential decay, with a main characteristic time constant of few hundred of femtoseconds (~200-500 fs) followed by a longer decay in few picoseconds (~1-2 ps). The anionic chromophore excited state lifetime (~400 fs) is about two times longer than the neutral protonation state (~200 fs)[72]. The HBDI excited state dynamics showed almost no dependence with the solvent viscosity[5,70]. This result suggested that isomerization processes should happen via a pathway with conserving volume, and the "hula twist" (Chapter 2 section 2.5.2) of the methylene bridge connecting the phenol and imidazolinone rings was suggested. Similar results were predicted by computational simulations[73].

More recent studies regarding the *cis-tra*ns photoisomerization of *cis*-HBDI showed that, upon irradiation, the photostationary state between *cis* and *trans* chromophore isomers depends on the different protonation states and solvent (methanol/acetonitrile). Whilst for the neutral chromophore the concentration is around 40:60 of *trans* isomer in methanol, for the

anionic form the *trans* isomer only represents about 23 %[74]. The study of the excited state dynamics using ultrafast fluorescence up-conversion (resolution 50 fs) of these photo-stationary states (*cis*-*trans* mixture) showed that the anionic chromophore states always have the longest lifetimes[72,74]. Interestingly, identical decay times were found starting from 100% of *cis*-isomer as for *cis*-*trans* mixtures in both anionic and neutral forms[74]. The similarity between the lifetimes of the *cis* and *trans* states suggested that the excited states of the two isomers must also be similar[74]. Furthermore, NMR studies of the cationic, neutral and anionic forms of HBDI showed that the isomerization can also happen in the ground state[75]. Indeed, the isomerization energy barriers in the ground state were found to be very similar among the differently protonation states of the chromophore and relatively small compared to other similar compounds. Finally, this study showed that at room temperature, the ground state of the cationic chromophore is populated by *cis* and *trans* isomers. In contrast, only the *cis* isomer is present for the neutral and anionic chromophores forms[75].

To further describe the excited state dynamics of HBDI, and since there are a considerable large number of studies with different analogues[76–83], here, we will only focus on those tackling the most similar and straightforward ones. They are particularly interesting those that try to increase or decrease the planarity of HBDI[78,79]. For example, the above results contrast with studies done for a nonplanar chromophore generated through methylation of HBDI on the phenolic ring in meta position to the hydroxyl group. The non-planarity of the chromophore increases the deactivation by non-radiative pathways decreasing the excited state lifetime by a factor of two for the neutral forms (from 0.29 to 0.15 ps) and by a factor of five for the anionic forms (from 0.5 to 0.1 ps)[79]. On the contrary, blocking the isomerization pathways and increasing the planarity decreases the deactivation by non-radiative pathway and leads to an increase of the fluorescent quantum yield as for o-HBDI (see below)[78]. This observation confirms that the fluorescence in the avGFP is linked to the effect of the protein cage; which suppresses non-radiative relaxation by restricting large-amplitude motions and isomerization.

Some analogues that particularly draw the attention are those that have the -OH phenyl substituent in *meta* (*m*-HBDI)[80,81] and *orto*- (*o*-HBDI)[82,83] positions compared to the original HBDI where the OH is in *para*- position. The first significant difference for *m*-HBDI is that

the absorbance of the cationic form strongly overlaps with those of the neutral and anionic forms, and are practically indistinguishable among them. The photoconversion from *cis* to *trans*-isomers was about 35% in DMSO for *p*-HBDI[80] (similar values were reported for *p*-HBDI in methanol[75]). On the contrary, for *m*-HBDI, it was only about 7 %, together with an irreversible photo-degradation. The first excited state dynamics studies (using TCSPC) done in aprotic solvent, found longer lifetimes for *m*-HBDI (characteristic times of 15 and 90 ps in DMSO), compared to the *p*-HBDI with excited state lifetimes under the resolution limit of the TCSPC (~10 ps)[80]. This result correlated with the higher fluorescent quantum yields found for *m*-HBDI compared to *p*-HBDI. These first studies were expanded in protic solvents and not only to *m*-HBDI and *p*-HBDI but also to *m*-MeOBDI, *p*-MeOBDI, homologues that prohibit any kind of proton transfer from the phenyl to the solvent, and to *p*-HBDIMe+ which has an added methyl group to the nitrogen in the imidazolinone moiety[81]. The results for all *para* substituents were in agreement to those in previous studies, sub-picosecond excited state life time due to the existence of a near barrierless twisting chromophore in the excited state. On the contrary, they found a completely different behavior for meta conformers. For example, a sequential ESPT was observed for m-HBDI forms at neutral pH, involving first the adiabatically dissociation of the hydroxyl group in 0.7 ps (intermolecular proton transfer to the solvent) followed by the imidazolinone nitrogen protonation in 3.1 ps. For o-MeOBDI, the nitrogen protonation is the first step in 1.1 ps followed by *cis-trans* isomerization 27 ps. The essential result is that the overall photoinduced reactivity of the neutral anionic and cationic forms of m-HBDI in protic solvents is dominated by acid-base properties of the molecule[81]. These analogues are not the only ones exhibiting proton transfer to the solvent[76] (photo-acid).

The *o*-HBDI dynamics were published a few years later, and a schematic representation of its mechanism can be seen in Figure 3.9. The intramolecular hydrogen bond of the *cis* form stabilizes it thermodynamically and makes it the stable form over the *trans* conformer. The absorption spectrum of *o*-HBDI in aprotic solvents is mainly characterized by an absorption maximum around 385 nm, excitation at this wavelength induces a fluorescent emission peaking at 605 nm[82,83]. This large Stokes shift of about 120 nm is the consequence of an ultrafast excited-state intramolecular proton transfer (ESIPT). Fluorescence up-conversion experiments showed that growing of fluorescence signal was beyond the system resolution

(<150 fs)[82]. The experiments performed with deuterated phenol -OD demonstrated that the rate of ESIPT is insensitive to an H/D exchange (isotopic effect) and underlines the existence of an essentially barrierless potential energy surface along with the ESIPT reaction.

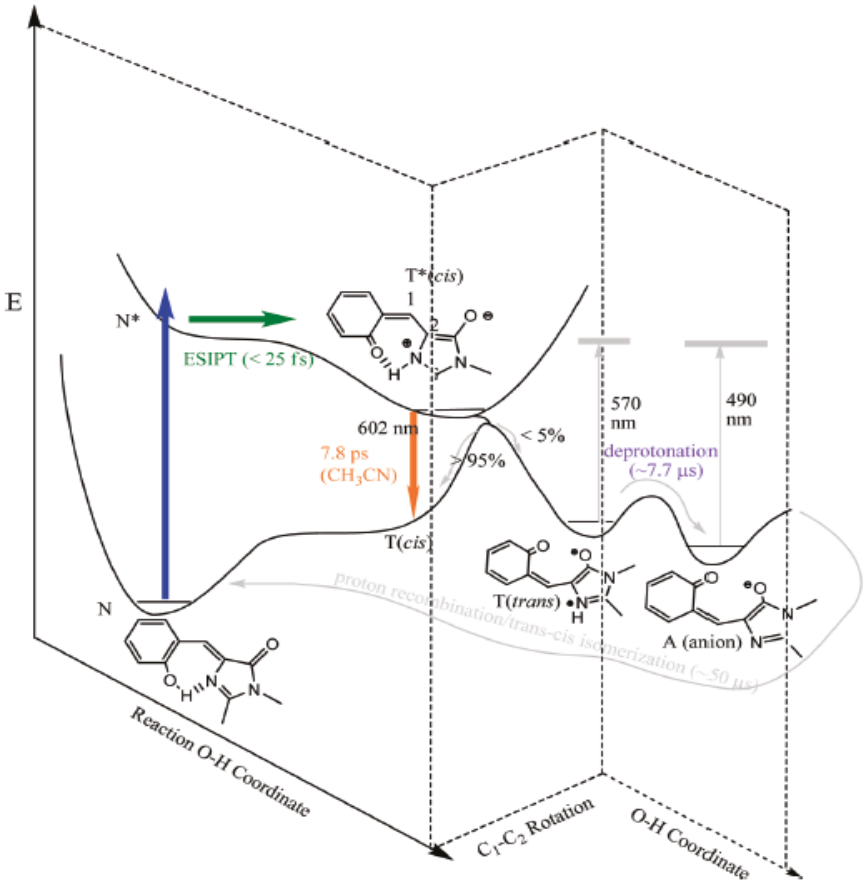


Figure 3.9. Proposed excited and ground-state dynamics of o-HBDI in CH3CN reproduced from Cheng-Chih Hsieh et al.[83]. The mechanism includes ESIPT and ground state, cis-trans isomerization, ground state deprotonation and a final reverse Proton Transfer.

These studies were further extended by an extensive work combining several ultrafast time-resolved (TR) spectroscopies including TR fluorescence up-conversion with a 40 fs resolution and TRUV/vis/mid-IR femtosecond-millisecond transient absorption[83]. TRUV-Vis femtosecond transient absorption experiments showed that after the ESIPT in 25 fs (characterized with fluorescent up-conversion) a new positive absorption band centered at 575 nm is growing in about 8 ps. This band was attributed to the photoproduct of *cis-trans* isomerization. The long-lived *trans* photoproduct had no evolution until the microsecond time range. Flash photolysis experiments were used to monitor the ground state dynamics of

this band, revealing a shift to 530 nm in the microsecond range (7.7 μs). This evolution was assigned to ground state deprotonation forming a *trans* anionic o-HBDI. All these results were further supported by similar time evolutions obtained by two-step laser-induced fluorescence (TSLIF) experiments and TRUV-pump/mid-IR transient absorption spectroscopy. Using the latter technique, the authors were able to identify a 6 ps lifetime species characterized by an exocyclic C=C double bond stretching at 1650 cm^{-1} and a C=O stretching of the imidazolidinone carbonyl group at 1720 cm^{-1}[183], clear features of *cis-trans* isomerization.

Table 3.2. Photodynamics of ortho, meta and para substituents for HBDI.

POSITION	DECAY TIMES	MECHANISM
***p*-HBDI**	<1 ps	Ultrafast cis-trans isomerization
***m*-HBDI**	0.7 and 3.1 ps	Phenol deprotonation + N imidazolinone protonation
***o*-HBDI**	25 fs, 7 ps, 7.7 μs	ESIPT + isomerization + GSPT

3.2.2 Photodynamics of avGFP

From the previous section, it is clear that the main deactivation pathways of the *p*-HBDI in solution is the *cis-trans* isomerization[68]. Furthermore, it has been mentioned above that the chromophore becomes fluorescent at very low temperatures (77 K)[4,5] and when it is encapsulated in non-protein scaffolds[6] or non-native proteins[7]. These results indicate that the protein cage β-barrel plays a major role in the fluorescence of the avGFP.

As shown in Figure 3.2, avGFP protein has two absorption maxima at 398 and 477 nm, the corresponding fluorescent emission spectra peak at 510 and 503 nm, respectively. The over 100 nm Stokes shift for the 510 nm fluorescent emission maximum respect to the 398 nm absorption maximum is explained by an ESPT. Boxer and coworkers[84] were the first to characterize it using fluorescence up-conversion. They showed that after excitation of the anionic form at 470 nm, there is an instantaneous rise of the fluorescence signal at 508 nm which lasts longer than 150 ps (longest delay meassured). While on the contrary, excitation

of the 398 nm maximum is characterized by an instantaneous fluorescence emission signal at 460 nm which decay in a few picoseconds and concomitantly the rise of the fluorescence emission at 508 nm occurs The significant isotopic effect observed (H/D) was attributed to an ultrafast ESPT in the ps range. Furthermore, they observed that the irradiation of the neutral form caused a small photoconversion to the ionic form attributed to an infrequent transition from I* to B*. Their studies, combined with the quantum calculations done by Weber et al.[68] allowed rationalizing the different species and time constants observed for the avGFP photodynamics (Figure 3.10). Thanks to the X-rays structures, the I* to B* transition was assigned to the rotation of the Tyrosine 203[85].

A few years later, using pump-dump-probe transient absorption (TA) experiments, J. Van Thor's group[86] demonstrated that the recovery of the neutral chromophore (A) from the tautomer form (I) occurs in a two-step process with associated times of 3 ps and 0.4 ns via one intermediate state both affected by isotopic effects, attributed to proton diffusion processes. This work was later extended by other groups where they studied the chromophore re-protonation reaction using pump-dump temperature dependent experiments in both light and heavy water to study isotopic effects. The researchers found that the thermal back reaction from I to A occurs via hopping of two protons in a synchronous but concerted manner and a tunnelling mechanism at the S205 oxygen, where the ionization-resistant S205 residue plays a major role acting as the limiting reaction step[87].

Finally, in 2015 the group of S.R. Meech reported the complete proton transfer cycle for avGFP using multiscale time resolve IR (TRMPS)[88]. They studied avGFP and two mutants T203V and S205V, which directly affect the hydrogen bond network[88]. They reported a similar time constant for avGFP and T203V. Excitation of the neutral form induces an ultrafast ESPT characterized by two time constants (13 ps and 100 ps), giving the photo-excited intermediate tautomer I*, which relaxes in 2.1 ns mainly by fluorescence to I. The recovery of the original A form in about 17 ns. Surprisingly, the ESPT transfer in S205V was reported to be much slower around ~730 ps attributed to an alternative proton wire previously suggested[89].

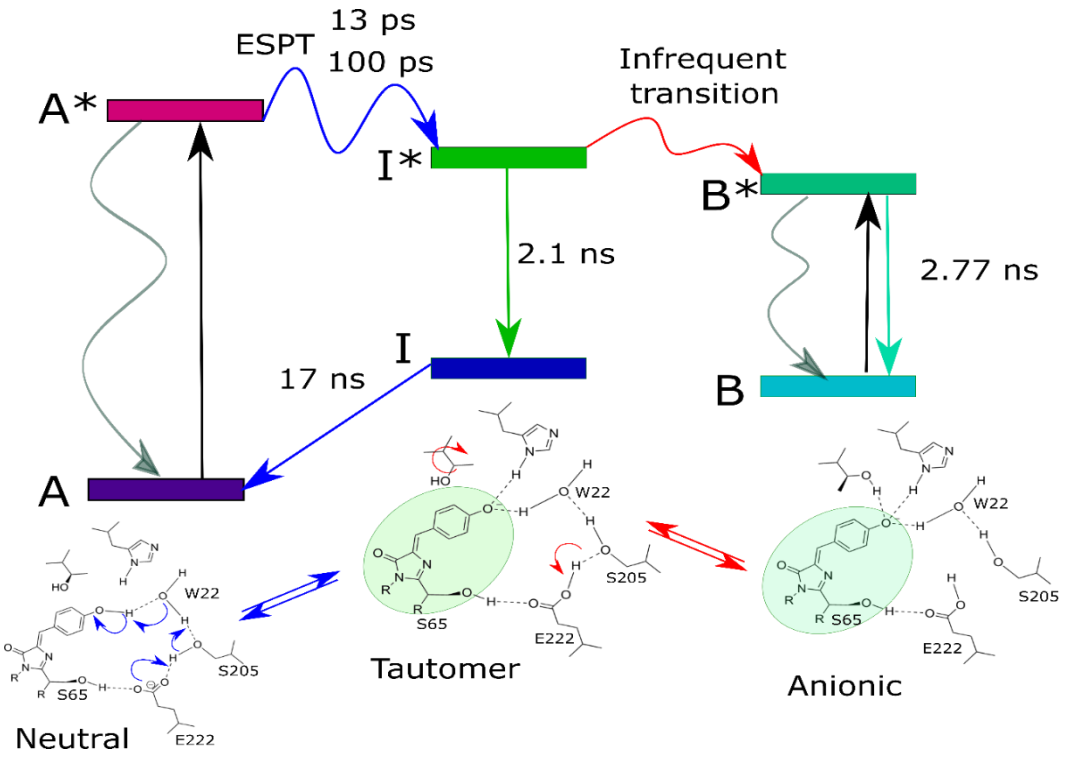


Figure 3.10. avGFP photodynamics scheme that combined results of Chattoraj et al. [84] and Weber et al. [68], together with the X-ray structures for the different states (neutral tautomer and anionic) from Brejc et al. [85]. The associated time constants have been taken from Laptenok et al[88] (experiment done D_2O buffer). Figure made with Inkscape.

The first step to elucidate the photodynamics of an FP is getting the time constant. Nevertheless, the most complex part is their assignment to a specific structure. A close look at Figure 3.10 shows the X-ray structures proposed by from Brejc et al. [85], suggesting that the ESPT forming the tautomer involves a water molecule, the phenol group of the chromophore and three protein-cage residues, S205, S65 (forming the chromophore), and E222. Therefore, the ESPT was rationalized by a proton relay (mechanism proposed by Kasha[90], Chapter 2 section 2.5.1) between the phenol and the E222. Furthermore, it can also be seen that hydrogen bonds between the chromophore and the protein cage seem to control the rigidity, i. e. the radiative decay yields the anionic (B*) form and the characteristic dynamics of avGFP[85]. TRIR[91] and femtosecond Raman spectroscopy[92] experiment have confirmed that the E222 is the proton acceptor[93] forming the tautomer. Furthermore, studies have shown that after excitation, the chromophore has some skeletal vibrations (280 fs characteristic time)[92,94] in the form of a "wagging" and, since protein cage prevents the

chromophore to isomerized, and the ESPT is triggered as lower energy barrier reaction[92,94]. Recent studies using TRUV-Vis and TRSR of a mutant (avGFP S65T/S205V) that lowers the ESPT time constant by a factor of 30, were able to determine that the first step is the proton transfer from the phenol to the water molecule, followed by a concerted transfer to E222[85,95]. Furthermore, it has been shown that the mutation of the glutamic acid E222 to glutamine renders the avGFP non-fluorescent[96].

So far, the excited state dynamics of the avGFP neutral absorption band (A) have been described. To understand the excited state dynamics of the avGFP anionic form (B), we will review that of the EGFP due to their similarities[22]. The EGFP is obtained by mutation of the Serine 65(S) by threonine (T), causing the predominant form to be the anionic chromophore (B) and not the neutral form (A). This is mainly because S65T favors the protonation of E222 and thus the deprotonation of HBDI[97]. In the avGFP after excitation of A the excited state tautomer I* is formed which decay via emission of fluorescent. The S65T in the EGFP distorts the hydrogen network provoking a higher fluorescence quantum yield of B form. However, at the same time, the deactivation of I* becomes mainly non-radiative. Interestingly, the fluorescence after excitation of the neutral A form is restored in the double mutants E222Q/H148D and S65T/H148D[96]. TRIR measurements have shown that in these two double mutants, an ESPT is occurring in about 100 ps from the chromophore to the aspartic acid. This demonstrates that multiple ESPT mechanisms through different amino acids are possible[96]. Single-molecule spectroscopy allows revealing silent and infrequent species such as dark state and triplet states that could be the origin of the avGFP blinking behavior. Initially, *cis-trans* isomerization was postulated by Weber et al.[68] as a possible reason for the momentaneous absence of fluorescence of the neutral avGFP form. Studies following the fluorescence under the microscope of EGFP mutant (T203Y) by Pellegrini and coworkers, demonstrated the existence of a dark state when this protein is excited at 365 nm, and assigned it to the existence of a *trans* state[98]. Furthermore, Pellegrini and coworkers have also shown that the presence of zwitterionic species can also explain blinking in EGFP protein[99]. These last studies are in contrast with the results from Gosh et al.[100], which attributed blinking of the EGFP to a heterogeneous ground state as a result of different amino acid side chains orientations. The authors were able to determine different ground state conformers using Fluorescence Lifetime Correlation Spectroscopy in EGFP[100]. They

recorded the fluorescent decay of the anionic form at 485 nm excitation, which needed two decay time constant of 2.1 and 3 ns to be correctly fitted. They used these time constants as filter functions to calculate fluorescent lifetime correlations identifying several ground state transitions between these two states and a dark state in the µs time scale. The three ground state conformers were assigned to the T203 rotation and two E222 conformations. These results are in agreement with new high resolutions X-ray structures which also reported different conformers for E222[101] in agreement with recent neutron diffraction studies which detected two different hydrogen bond networks for the two conformers of E222[97]. Interestingly, from this last study, it was also possible to infer possible alternative deprotonation paths. The crystal structures obtained by neutron diffraction from this publication are reproduced in Figure 3.11. Other works which are also in favor of the hypothesis of E222 being the origin for blinking in EGFP are new fluorescence lifetime measurement on EGFP with the E222 mutated to a histidine, showing a mono-exponential decay[102], and the fact that the mutation of E222 by glutamine renders the EGFP photo-switchable reversible[103].

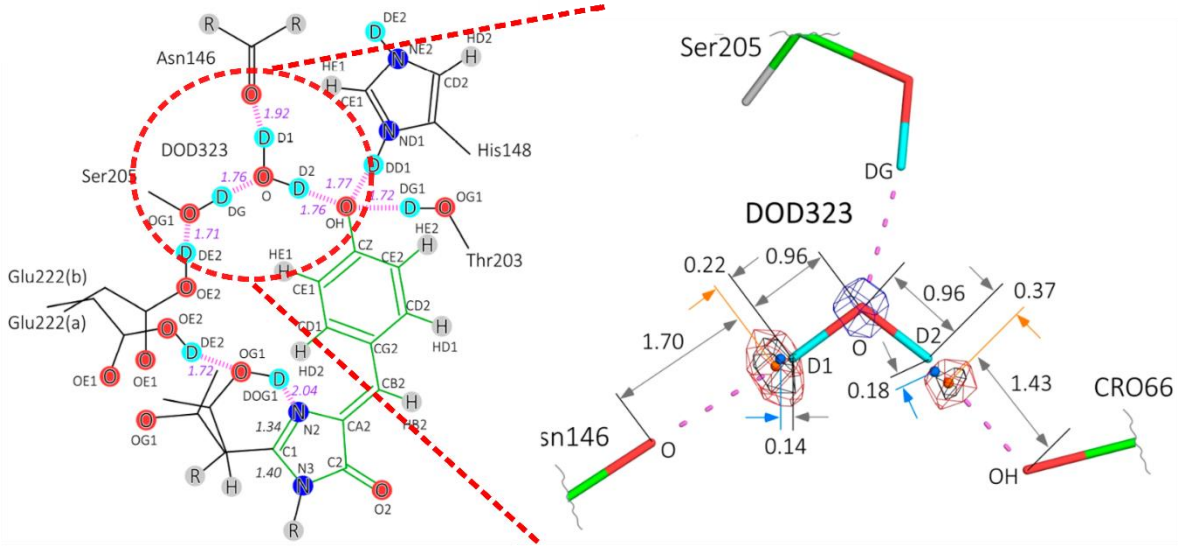


Figure 3.11. Bond model and neutron scattering length schematic drawing of the electron diffraction density maps. Including hydrogen, deuterium atoms and a magnified model for the water molecule close to the chromophore (named DOD323 in the figure). Reproduced from reference [97]. CRO acronym correspond to the HBDI chromophore.

We can try to summarize in a scheme all the different information from the ensemble time-resolved spectroscopies, single-molecule and X-rays results. Regardless of the origin of the multi-exponential decay and blinking of fluorescence in the GFP, the presence of different chromophore conformations in the ground state is clear[43,68,99,100]. From works reviewed here, it is possible to expand the photophysical scheme based on those proposed by Chattoraj et al.[84], which is based on the ESPT mechanism proposed by Kasha[90] (Chapter 2). The proposed photophysical scheme can be seen in Figure 3.12. In this figure, we can see a Kasha ESPT scheme (similar the one in Figure 2.13), adapted to the proposition of Chattoraj et al.[84] and Weber et al[68] for the neutral chromophore. Furthermore, the I1 and I2 states, revealed by pump-dump-probe experiments, have been added, displaying a multi-step ground process[86] characterized by the proton tunnelling effect at the Ser205[87]. In the middle part, we can see the anionic chromophore scheme treated as a mixed-valence compound,[43], which introduced at least two fluorescent states in agreement with others[100]. Finally, in the right part, the existence of dark states proposed by several authors[98–100] is represented. To conclude, the photodynamics of avGFP are complex, and this complexity will logically also be found in photo-transformable proteins and RSFPs.

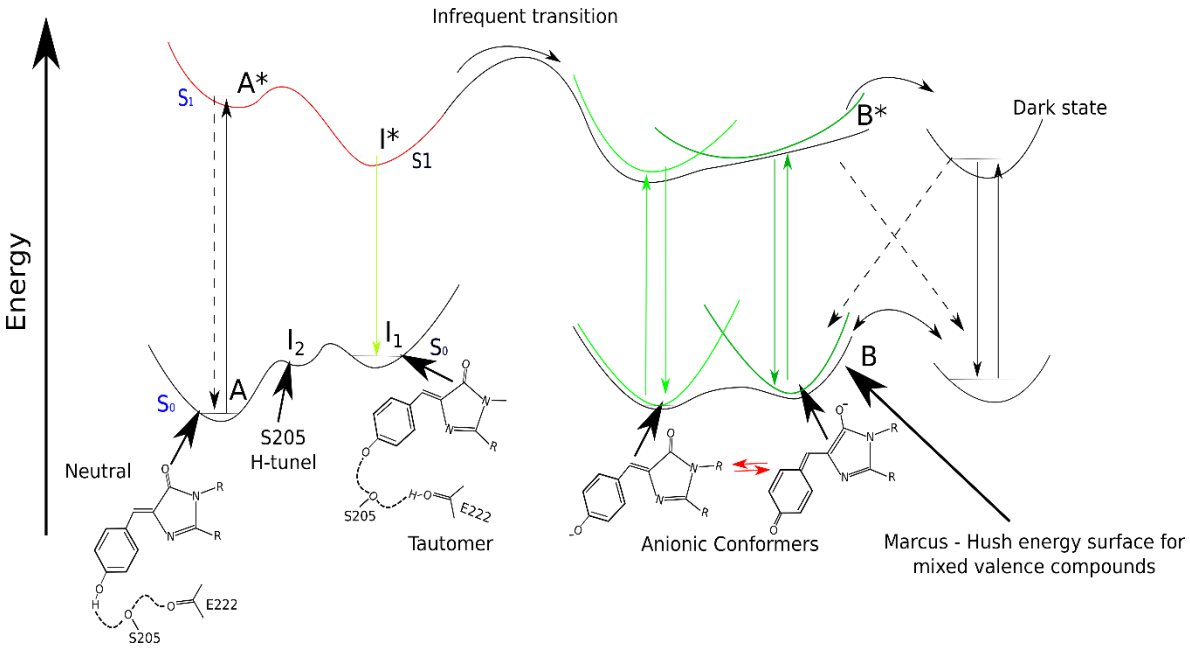


Figure 3.12. Suggested expanded photophysical scheme for GFP proteins, based on the one proposed by Chattoraj et al.[84] with the contributions of other recent works[43,86–88,98–100]. Figure made with Inkscape.

3.2.3 Photodynamics of RSFPs

As described above, there are currently 42 RSFP taking into account those that can be photoconverted from green to red. The majority of publications have focused on studying the photophysical properties of these proteins, such as fatigue brightness and fluorescence quantum yields. Nevertheless, the photodynamics of only very few of these proteins have been studied.

As shown in Figure 3.13, only 7 RSFP have been studied. Several studies focus on Dronpa[32,59,104,105] and Dronpa-2, the fast mutant M159T [32,59,106,107] that will be covered in the next section. Other RSFP studied are the positive photoswitcher Padron[108], Padron0.9[109] and the decoupled photoswitcher Dreiklang[110]. Regarding the multi-photochromic proteins, the photodynamics of IrisFP have been published. Nevertheless, this study only focused on the switching from *Off* to *On* form of the green part[111]. Finally, there are some studies on rsEGFP2[35,112]. From all of RSFP, the most studied are definitely Dronpa and Dronpa-2[32,59,104–107].

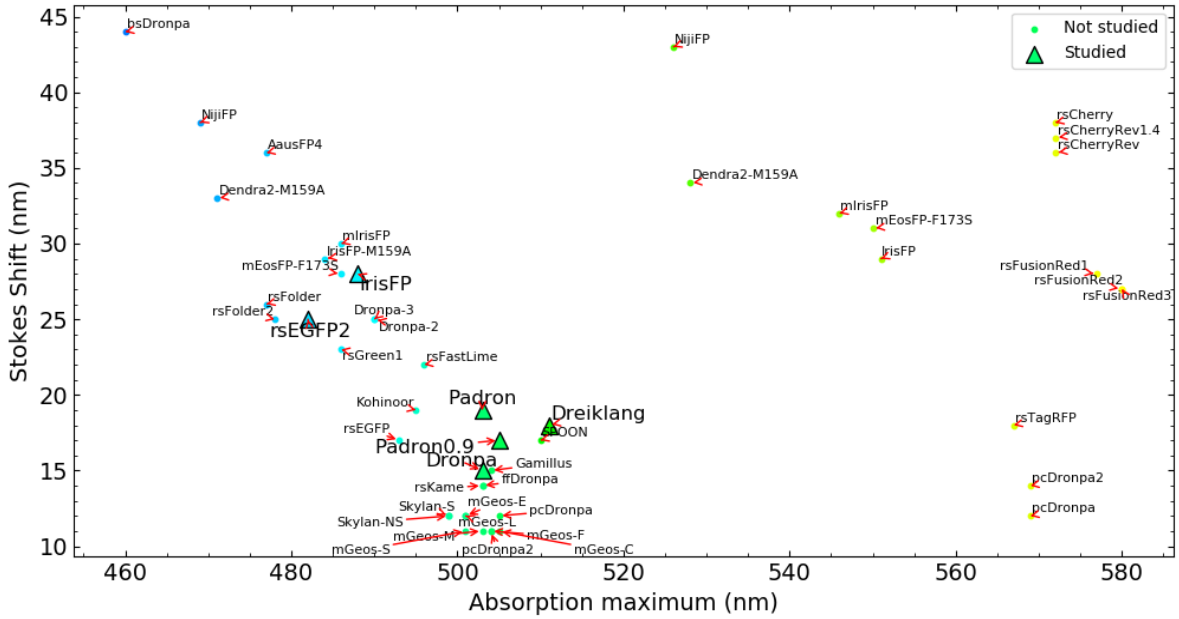


Figure 3.13. Stokes shift plotted against the absorption maximum for main RSFP proteins. The colour of the points corresponds to the absorption maximum. Proteins represented by a triangle correspond to those that photodynamics has been studied to our knowledge. Data plotted with matplotlib python package.

3.2.3.a Dronpa

Dronpa was the first RS-GFP-like protein discovered in 2004 by Miyawaki's group[57]. They named this protein dron (ninja term for vanishing) and pa (which stands for photoactivation). The switching properties of Dronpa have been widely studied. The structure of Dronpa is similar to the avGFP structure and GFP-like proteins. The protein cage is a 11 stranded β-barrel. The chromophore is also the HBDI which is formed by the tri-peptide Cys-62–Tyr-63–Gly-64 (CYG) in an α-helical segment inside this β-barrel[113]. Dronpa can be reversibly switched between a fluorescent *On*-state (resting state) that absorbs at 503[60] nm to a non-fluorescent *Off*-state, and reversibly switched back to the *On* state by irradiation of the *Off* state with UV-blue light (*Off* state absorption maximum is at 390 nm). The *On* state has a fluorescence quantum yield of 0.85[60] and a mono-exponential fluorescence decay with a characteristic time of 3.6 ns[114].

The first study by Andersen et al.[113] focused on the X-ray structures of *On* and *Off* states. In the *On* fluorescent state, the chromophore has a *cis*-anionic geometry, while on the *Off* non-fluorescent state, the chromophore adopts a *trans* neutral one. Besides the chromophore differences between *On* and *Off* states, there are also changes in the hydrogen bonding network and the protein cage, which mainly involve the amino acids Arg66, Ser142, Val157, and His193. From the X-ray structures, they concluded that the loss of fluorescence of the *Off* state is due to a change in the chromophore protonation state and two main structural differences between the states. The first is the planarity of the chromophore, which in the *On* state adopts an almost planar *cis* configuration, with a 15º dihedral angle between the phenol and imidazolinone rings, while in the *Off* state, the chromophore adopts a *trans* conformation with a 30º dihedral angle. The second is the freedom of movement coming from the fact that the chromophore in *On* state has 10 hydrogen bonds with the β-barrel cage whilst the *Off* form has only 8 hydrogen bonds resulting in a less bounded chromophore to the protein cage. This study showed that the switching mechanism between the two states in Dronpa requires several steps, which can be assigned to *cis–trans* isomerization, proton transfer, changes and rearrangement of the hydrogen bond network and amino acids of the protein cage[113]. Nevertheless, classical X-ray data do not handle any information regarding the time scale and order in which these processes may happen, and whether they take place in the ground

or excited states, and if all these rearrangements occur subsequently or in parallel to each other.

Several groups (Figure 3.14) have done ultrafast (Uv-vis)[59,104] or vibrational (IR)[32,105–107] transient absorption spectroscopies to address all these open questions for Dronpa. Since the photoswitching quantum yield of the *On* to *Off* reaction is several orders of magnitude lower than the *Off* to *On (*~1-4·10^{-4} vs ~0.1-0.4; see Table 3.1), the photoswitching studies for Dronpa are mainly focused on the later reaction. The study of *On* to *Off* reactions cannot be done directly, and only the fluorescence lifetime is giving some excited states information. Among all the *Off* to *On* photoswitching studies, it is clear that excited state species decay in the picosecond time range. Then the photoproducts in the ground state evolve to the final *On* state in the microsecond range. Besides the excited states lifetimes, which seems to be more or less clear among the literature, and the fact that there are processes in the ground state, the number and nature of species involved in the switching process and the order in which the structural changes happen, such as *cis-trans* isomerization, proton transfer or protein rearrangement, is still subject of debate.

The first time-resolved *Off* to *On* isomerization study of Dronpa for was done in 2007 by Fron et al.[104]. They used TRUV-Vis and found that after excitation of the *Off* state with a 395 nm 100 fs laser pulse, the excited state decay was characterized by a 4 ps time constant, which was affected by a 2 fold isotopic effect when the experiments were performed in heavy water. From these results, they concluded that the first step must be an excited-state proton transfer ESPT preceding a *cis-trans* isomerization in the ground state (see Figure 3.14). However, the study of Dronpa using IR femtosecond transient absorption led to different conclusions in 2013. Indeed, two other groups studied the photoswitching dynamics of Dronpa[105] and Dronpa-2[107], and none of the studies reported an ESPT. Dronpa-2 has a mutation of the Methionine 159 by a Thyrosine (Dronpa M159T), which increases the photoisomerization quantum yields (see table 3.1 reference 57 for an accurate comparison). Although both studies reported similar time constants for excited states decay, the assignment of the excited species nature was different. The first publication from Warren et al. [105] reported an excited state which decays in 9 ps to a long-lived photoproduct (more than 100 ps, limit of experiments). This long-lived GS species had no characteristic signal of the

phenolic group around 1500 cm^{-1} which would indicate chromophore deprotonation. In addition, some characteristic amino acid signals from the final cis anionic *On* state were missing. The authors then concluded that *cis-trans* isomerization occurs in 9 ps and is followed by millisecond ground-state proton transfer (GSPT)[105]. The authors also estimated that the *trans* to *cis* isomerization must occur from the *cis* anionic fluorescent state in 3.6 ns followed by a ground state proton transfer. Isomerization competes with emission deactivation, and this explains the low *On* to *Off* switching yield.

The second publication from Lukacs et al.[107], studied Dronpa-2 using also TRIR on a wider temporal range from 0 to1000 ps. The evolution of transient signals needed three-time constant to be fitted correctly. The first two (7 ps and 50 ps) were attributed to the recovery of the *trans* neutral isomer while the third one around (500 ps) was assigned to a long-lived metastable species in the ground state. Nevertheless, this long-lived species was lacking signal features of the final *On* state nor the characteristic signals that a photoproduct after either an ESPT (phenol deprotonation) or a *cis-trans* isomerization should have. These are characterized by a shift of the imidazolinone C=O mode or the formation of the phenolic group (1500 cm^{-1}) for ESPT and isomerization respectively. Thus, they concluded that in the first picosecond time constant, a protein rearrangement occurs forming a photoproduct, which further evolves to the cis conformer in longer time scales (>> ns). In summary, two similar studies agree with the fact that there is no ESPT, but they disagree in the order of processes. Whilst for Warren et al. [105] *cis-trans* isomerization is the first step in a picosecond range, for Lukacs et al.[107] it is a protein rearrangement. Such different conclusions rise from the same difficulties: i. none of the studies could monitor the evolution of the photoproduct between 1 nanosecond and 1 millisecond, and ii. assignment of IR band to *cis*/*trans* species is not trivial. Indeed, the assignment in both studies was based on the vibration of the Arg66 side change or the imidazolinone C=O mode, respectively. These two groups are hydrogen bonded in the *On* state. The Arg66 bands changed drastically during the *Off-On* photoswitching, making these bands a good indication of protein rearrangement. The changes caused by a *cis-trans* isomerization might not be apparent since the C=O vibrational bands shifts are also affected by the hydrogen bond with Arg66 and thus can be confused with those cause by the *cis-trans* chromophore conformations.

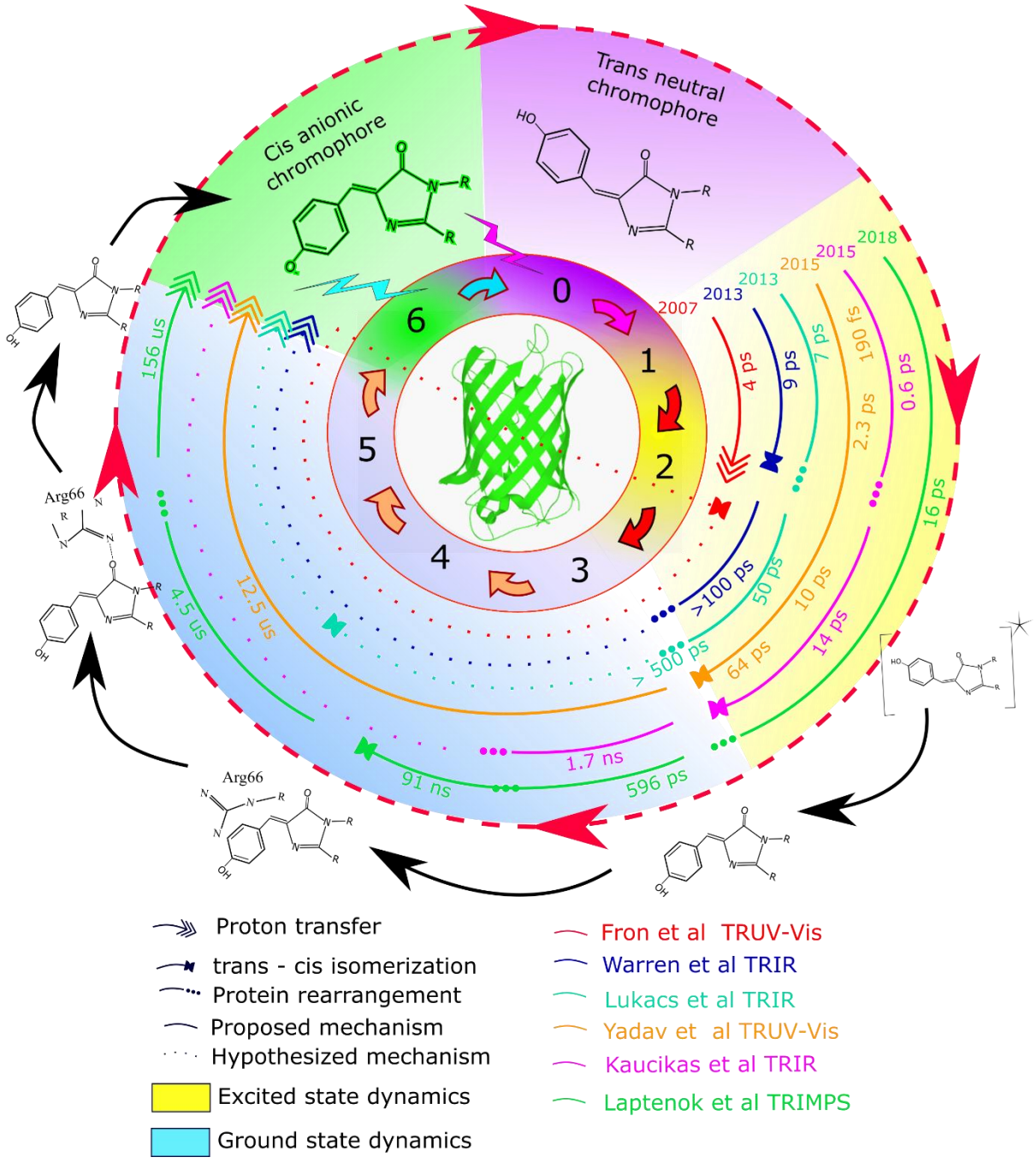


Figure 3.14. Different proposed photodynamics mechanism for the photoswitching of the trans neutral Off state to the cis anionic On state. The structures correspond to the mechanism proposed by Laptenok et al. [34]. Solid lines refer to proposed mechanisms from the recorded data. Dashed lines refer to hypothesized mechanisms in time scales longer than those recorded. Figure made with Inkscape.

Later on, in 2015, Kaucikas et al. [106], studied Dronpa-2, from 0 to 1800 ps using femtosecond TRIR and performed thorough quantum mechanics calculations to rationalize the different IR transient bands considering the different torsion dihedral angles for *cis* and *trans* isomer reported by Andersen et al.[113]. They concluded that the difficulty in getting a clear isomerization signature using the imidazolinone C=O mode in differential IR spectroscopy arises precisely from the *trans* distorted structure. This distortion implies that the discussion of the existence of *cis*/*trans* isomer is complex and cannot be only retrieved using imidazolinone C=O band shifts compared to quantum mechanics calculations done on the chromophore alone. All along their publications, they discuss the different TRIR transient signal attributions given by Lukacs et al.[107] and Kaucikas et al.[106] and compared their transient signals, which were mainly fitted by two time constants of 0.6 and 14 ps. According to their calculations, they concluded that the first step is the *trans-cis* isomerization, giving a long live species. This one will evolve via a ground state proton transfer to the final *On* state in time scales longer than 1.7 ns.

The first complete time-resolved spectroscopy experiments probing the entire photoswitching dynamic range of Dronpa and Dronpa 2 were done in 2015 by Yadav et al.[59] They used femtosecond TRUV-Vis spectroscopy to study excited states dynamics and nanosecond TRUV-Vis (flashphotolysis) to study the ground state species evolution. Four time constants were used to describe the fs-TRUV-Vis data 220 fs and 2.6, 15 and 160 ps for Dronpa-2 and 190 fs, 2.3, 10 and 64 ps for Dronpa. The first-time constant was attributed to a fast relaxation from FC state, and the following three time constants were assigned to the decay of different excited states. The existence of different decay times was attributed to an *Off* ground-state heterogeneity. They reported a small remaining spectral signature at 300 ps which remains unchanged until 1.5 ns (the longest time probed with fs-TRUV-vis). This photoproduct spectrum could be compared to the steady-state difference spectra associated with the *cis-trans* isomerization (*cis* neutral absorption spectrum minus *trans* neutral absorption spectrum), which confirmed the excited-state isomerization. Furthermore, their performed fs-TRUV-Vis anisotropy experiments and revealed a change in the transition dipole moment orientation of the photoproduct. These results strongly suggest that the *trans-cis* isomerization is the first step. In addition, they studied the evolution of the photoproduct and reported a deuterium isotopic effect (~3.5 longer evolutions) in the microsecond time

range. They concluded that *Off* to *On* photoswitching mechanism involves a *trans-cis* isomerization in the excited state. The photoproduct formed provokes a change in the HBDI pK_a, which triggers the microsecond deprotonation in the ground state. The authors considered that Dronpa *Off* to *On* switching was a two-step process consisting of a *trans* to *cis* isomerization followed by proton transfer reaction[59].

Finally, in 2018, Laptenok et al. [32], studied the whole switching dynamics for Dronpa and Dronpa-2 using a unique multiscale IR transient absorption technique[121]. This technique allows to study, after a femtosecond excitation, the difference IR absorption spectra from the femtosecond to the millisecond range. The outstanding set-up providing a high signal to noise ratio combined with site-specific ^{13}C isotope labelling of the chromophore enabled them to distinguish chromophore dynamics and the relaxation of the surrounding protein. They revealed a complex multi-step mechanism, with several intermediate states in the photocycle, with lifetimes from picoseconds to tens of microseconds for Dronpa and Dronpa-2 (Figure 3.14). They confirmed that *trans* to *cis* isomerization precedes the proton transfer but revealed a much more sophisticated mechanism than the previously characterized two-step process. First of all, they suggest that the isomerization starts with a first 16 ps step assigned to the decay of the excited state. The geometry of the chromophore after this step is still an open question; however, it is not a *cis-On* form (dihedral angle). This first step is followed by some protein rearrangement in the ground state in 596 ps, which trigger the formation of the *cis* neutral intermediate in 91 ns. This time constant is much slower than the one reported in previous publications for *cis-trans* isomerization. The next step is then the rearrangement of the Arg66 in 4.6 µs which precedes the ground state proton transfer occurring in 156 µs, yielding the final *cis* anionic *On* state[32]. Interestingly, they found that the same time constants obtained after fitting the data for Dronap-2 could be used to fit results for Dronpa (giving good residues). They suggested that Dronpa and Dronpa-2 might follow the same mechanism, but no firm conclusion could be drawn because when Dronpa data was fitted without constraints, it yielded decay times in a similar time range but not identical.

A comparison between the different proposed mechanisms and characteristic times for Dronpa and Dronpa 2 can be seen in Figure 3.14 and Table 3.2. As presented in Table 3.2,

the evolution of the time constants is very different for the same proteins, which manifests the complexity of the photodynamics of RSFP and the complexity of evaluating the TR data.

Table 3.2. Resume of photophysical studies on Dronpa and Dronpa-2

Journal first author	**Year/Ref**	**Technique**	**τSpecies in ES (τ ps)**		**Species in GS (τ, ps)**	
			Dronpa	Dronpa 2	Dronpa	Dronpa 2
JACS Fron, E	2007 [104]	TR UV-Vis	4***	NT*	NT*	NT*
Nature com Warren, M. M.	2013 [105]	TRIR	9	NT*	NT*	NT*
J. Phys Chem B Lukacs, A.	2013 [107]	TRIR	NT*	7 / 50	NT*	500 ps
J. Phys Chem B Yadav, D.	2015 [59]	TR UV-Vis	0.19 / 2.3 10 / 64	0.22 / 2.6 15 / 160	12.5 µs***	19.2 µs***
J. Phys Chem B Kaucikas, M	2015 [106]	TRIR	NT*	0.6 / 14	NT*	1.7 ns**
Nature chemistry Laptenok, S. P	2018 [32]	TRIR	16	16	596 ps; 91 ns; 4.5 µs; 156 µs	596 ps; 91 ns; 4.5 µs; 156 µs

* NT (Not Studied): correspond to works only center in the excited state or one of the proteins.
** Poorly determined data recorded until 1800 ps.

*** D_2O effect reported in the time constant.

Although we have discussed how the majority of studies had focused on the *Off* to *On* photodynamics, the group of S. G. Boxer, using the approach described above (Figure 3.6), has recently evaluated the contributions of steric and electrostatic interactions that different substituents have directly on the chromophore to quantitatively demonstrate how electrostatic effects bias the chromophore *On* to *Off* photoisomerization pathway in RSFP, using Dronpa-2 as reference[115]. Measuring the fluorescent decays of Dronpa-2 variants at different temperature (TCSPC), they showed that the primary dominant energetic feature governing

the competition between fluorescence and non-radiative decay is an excited-state barrier for chromophore bond rotations around the C=C-C bridge between the phenolate and the imidazolinone rings. If the rotation occurs around the C=C bond, a *cis-trans* isomerization may occur while an SBR around the C-C bond results in a phenolate-ring flip. With the use of electron withdrawing (i.e. halogen atoms or NO_2 groups) and donating moieties (methyl and methoxy groups) directly in the chromophore (mutating the Tyr66), they managed to study the influence on the two non-radiative paths (Figure 3.15; red and blue). For variants with electron-withdrawing substituents, the isomerization pathway is energetically downhill and thus preferred over the ring-flip pathway. On the contrary, electron-donating substituents have the opposite energetic effect and favour the phenolate ring flip pathway[115]. Since the distinctive charge redistributions in the methylene bridge generate different bonds rotations, the use of withdrawing or donating moieties can be used to finely control properties of interest, such as isomerization regioselectivity. A schematic representation of the ring-flip and chromophore isomerization reproduced from reference [115] can be seen in Figure 3.15.

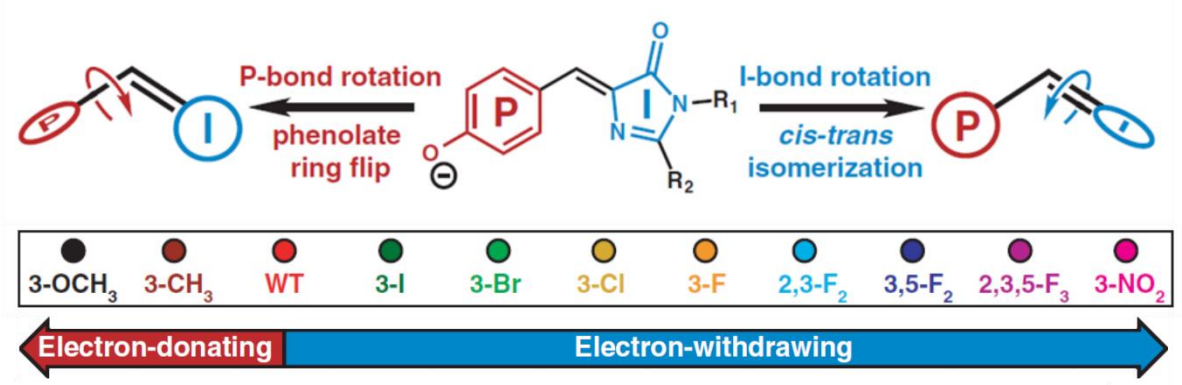


Figure 3.15. Schematic representation of the possible rotation that can occur about either the single or double bond, leading to a P-ring flip or cis-trans isomerization, respectively. R1 and R2 represent residues Gly64 and Cys62, respectively. Electron-donating and electron-withdrawing groups, introduced in the phenolate ring using non-canonical tyrosines in Dronpa2. Reproduced from reference 115.

3.2.3.b Other RSFPs

Meanwhile, several studies regarding other RSFP have been published. The first studied the photodynamics of Padron, a positive RSFPs[116]. The authors could characterize a 5.2 ps *cis-trans* isomerization step for *Off* to *On* switching, i.e. photoswitching from the *trans* anionic

Off state to the *cis* anionic *On* state. For the *On* to *Off* switching using 395 nm irradiation (starting from *cis* neutral form), they found two-time constants of 1 ps and 14.5 ps, attributed to ESPT and *cis-trans* isomerization. This first publication on Padron was followed by similar studies on one of its mutants - Padron0.9 by Walter et al.[109]. The authors, by selectively exciting at the proper wavelength and adjusting the pH to obtain the desired chromophore isomer, observed isotopic deuterium effects that allowed them to confirm the Padron suggested ps ESPT step for the *On* to *Off* photoswitching dynamics starting from the *cis* neutral form[108]. Furthermore, they could show a ps ground state dynamics between *cis*-neutral and *cis*-anionic forms, and correlated the dynamics with the previously published X-ray structures[117].

The publications regarding Padron and Padron0.9 were followed by Colletier et al.[111] who studied the *Off* to *On* photoisomerization for the IrisFP green form. This tetrameric protein was previously characterized by Adam et al.[63] who determined that in the fluorescent *On* state, the chromophore is a *cis*-anionic form, while in the O*ff* state the chromophore adopts a *trans* neutral form. After excitation of the *Off* state at 400 nm and 100 fs relaxation from FC, they reported two excited states decaying in 2 and 15 ps. The 2 ps excited state species was assigned to the relaxation of hot S_1 state. The second was assigned to an excited state species precursor of the isomerization. They also monitored the ground state dynamics from 10 ns to ms time scale. They found three-time components of 21 μs, 227 μs and 2.1 ms and a significant isotopic effect when experiments were done in D_2O (51 μs, 2.7 ms and 16 ms). These results were the first multistep ground state process to ever be reported in RSFP in the literature (before the publication of Laptenok et al.[32]). From these results, they suggested a first *cis-trans* excited-state isomerization in a picosecond range followed by a ground state deprotonation of the chromophore to adopt the final *cis* anionic conformation in few hundred of microseconds (Figure 3.16).

The last RSFPs that we need to mention in this chapter is Dreiklang. Lacombat et al.[110] studied the photodynamics of this protein using femtosecond TRUV-Vis spectroscopy in light and heavy water buffers. Dreiklang[33] together with Spoon[55] are the only two RSFPs that have a decoupled switching mode. Dreiklang was first reported by S. Hell and S. Jakobs coworkers[33]. Contrary to other RSFP, the *On* to *Off* switching is not based on a *cis-trans*

isomerization and a proton transfer[33,110]. From the X-ray structures, it was demonstrated that the switching mechanism is based on a reversible covalent addition of a water molecule to the C=N bond of the imidazolinone ring of the chromophore[33]. Lacombat et al. [110] showed that the *On* to *Off* switching dynamics after excitation of Dreiklang at 410 nm are characterized by 100 and 750 fs and 19, 108 ps excited states decay times (plus a long 1800 ps component). The first one was attributed to an ESPT exactly as the one happening in avGFP[84,92,94], i.e. they observed weak coherent oscillations [84,92,94]. Finally, they attributed the other three constants in the ps range to the decay of the tautomer I* to the ground state. The absence of absorption at 10 ns at 340 nm indicates that the final formation of the *Off* state is in the ground state and happen at longer delay times. The lacking of the ground state evolution did not allow further elucidation of the mechanism.

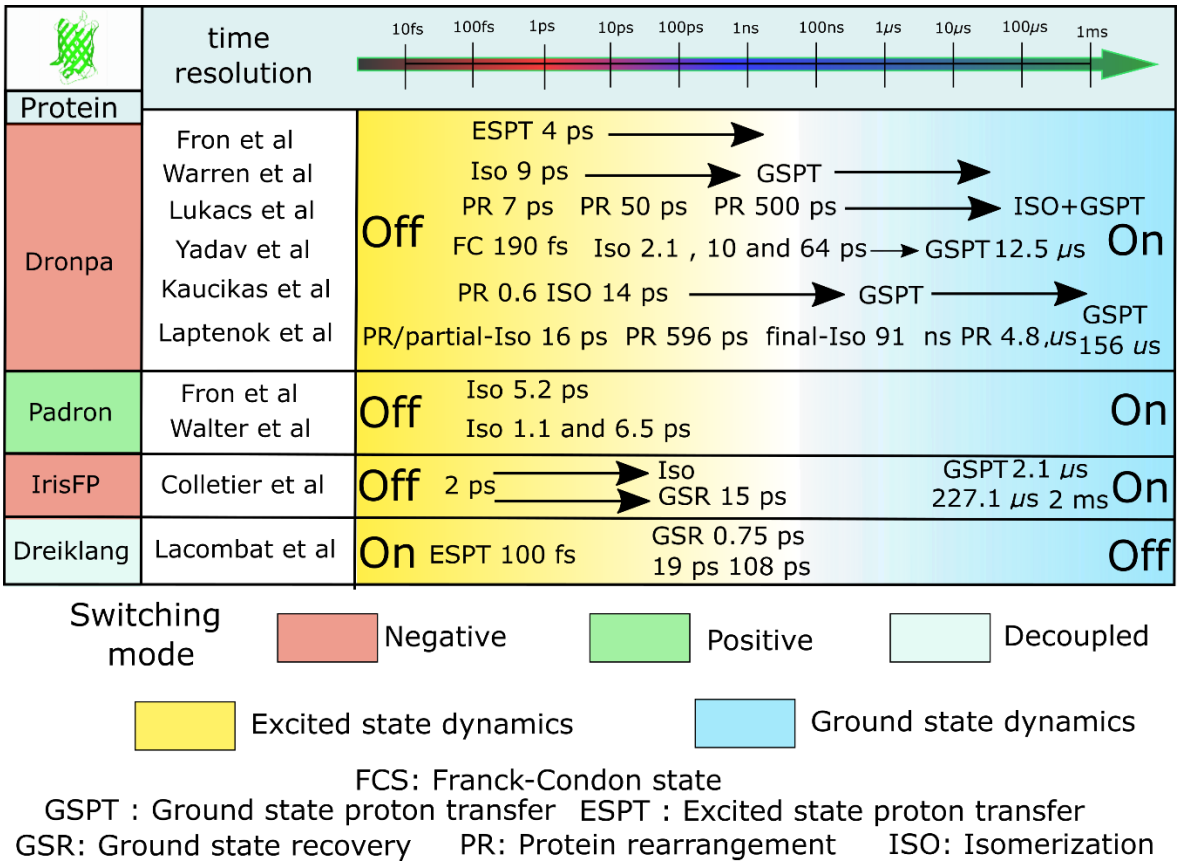


Figure 3.16. General overview of the proposed photodynamics for all the RSFP studied by ultrafast optical spectroscopy in the literature. Figure made with Inkscape.

The complete detailed mechanism between the different states in RSFP still represents an active research field. As we have seen, the different type of RSFP have different number of steps and mechanisms, mainly involving protonation states of the chromophore and *cis*/*trans* isomerization states[118]. Still the elucidation of the *On* to *Off* photoswitching mechanism for negative and positive RSFP needs to be studied in detail. Regarding decoupled RSFP, the studies of Spoon could complete the mechanism proposed for Dreiklang. In Figure 3.16, a summary of the above-discussed studies can be seen.

3.4 Photodynamics of rsEGFP2, objectives and methodologies of the thesis.

The rsEGFP2 protein is a negative RSFP developed by S. Hell and S. Jakobs group commonly used in RESOLFT microscopies. The resting state is the fluorescent *On* form: the chromophore adopts a *cis* configuration, and the phenol group is deprotonated (anionic *cis*-phenolate)[36,119]. The non-resting state of the protein is non-fluorescent, meaning that the chromophore has a *trans* configuration with the phenol group protonated (neutral *trans*-phenol). As pointed out previously, there is a high discrepancy on the photoswitching yields (Table 3.1). **Therefore, determining the precise photo-physical steady-state and photo-stationary properties of rsEGPF2 will be the first aim of the thesis.**

The switching dynamic between *On* and *Off* states involves at least a *cis-trans* isomerization, a proton transfer and some protein rearrangements[62]. Before I started my PhD thesis, the rsEGFP2 excited state dynamics were published by Coquelle et al.[35]. Using time-resolved serial femtosecond crystallography at an XFEL, combined with TRUV-vis spectroscopy, they showed the existence of a twisted chromophore configuration on the picosecond time-scale. Actually, after 400 nm excitation of the *trans* neutral *Off* state in solution, in 100 fs there is a relaxation from FC state, after this relaxation they reported the existence of two excited states decaying in 0.9 ps and 3.65 ps. The X-ray structure collected at 1 ps on protein microcrystals allowed assigning the 0.9 ps to a twisted chromophore, halfway in between the *trans* and *cis* isomers, as shown in Figure 3.17. This intermediate state is accommodated by a shift in the central α-helix in the excited state restricted by the close proximity to the V151

side chain. Mutation of the latter into an alanine show an increase of the *Off* to *On* photoswitching quantum yield. **The ground state evolutions, protein rearrangements and deprotonation mechanisms were not studied and represent one of the goals of my Ph.D.**

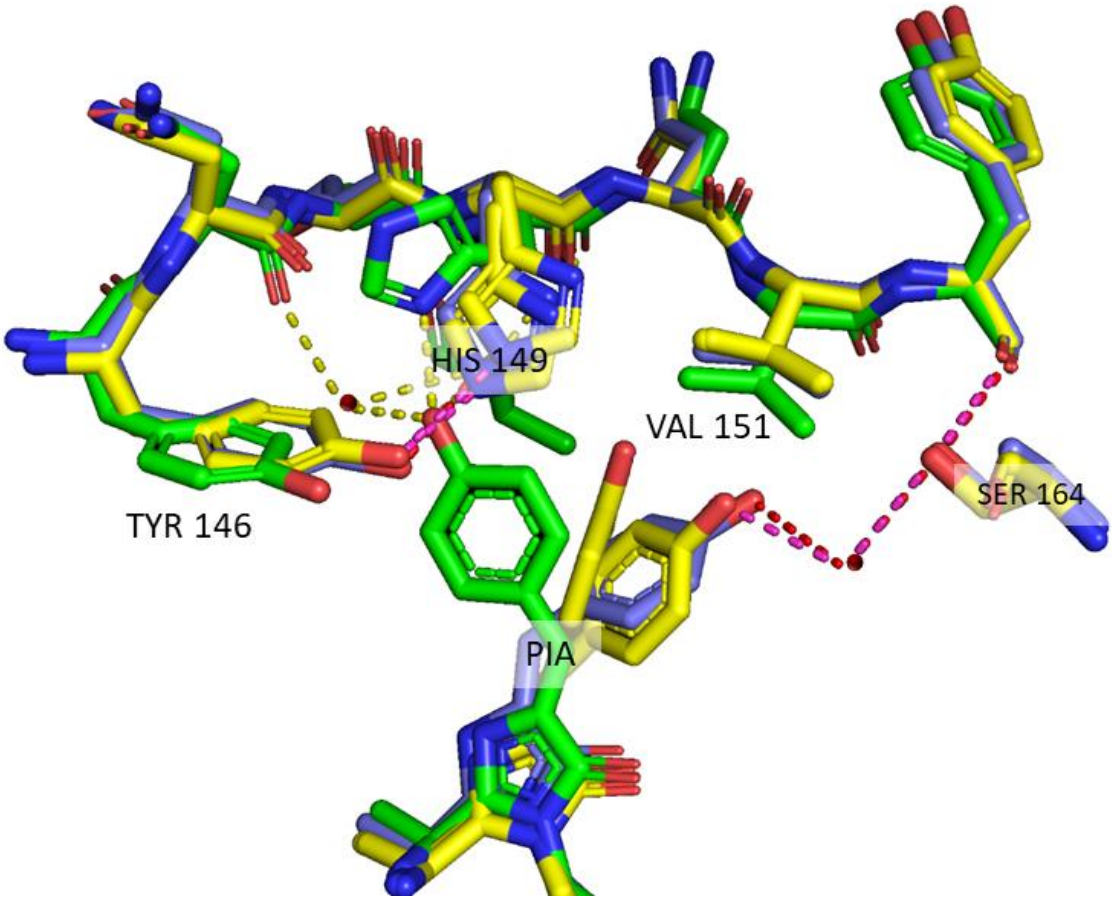


*Figure 3.17. rsEGFP2 chromophore structures in On and Off states in green and blue superposed with the 1 ps structured in yellow. Structures reproduced using PyMol from the PDB **5DTX**[62], **5DTY**[62] and **5O8B**[35] respectively.*

The analysis of transient UV-Visible data in Coquelle et al.[35] was based on the principle that the *trans* neutral *Off* state was populated by a single ground state conformer. This was based on the reported X-ray structures by El Khatib et al[62] obtained at low temperature in a synchrotron and the X-ray structure of the XFEL experiments[35]. In Figure 3.17, the *Off* form can be seen as the photoproduct derived from an OBF *On* to *Off* isomerization mechanism. However, some of our preliminary results, published in the thesis of Joyce Woodhouse[120], pointed out that depending on illumination conditions and crystal sizes, two *trans Off*

conformers could be found in the *trans* neutral *Off* state. These *Off* conformers are in agreement with one-bond-flip (OBF) isomerization photoproduct -corresponding to *trans1* in figure 3.19- or a hula-twist (HT) isomerization photoproduct -corresponding to *trans2* in figure 3.19-. The X-ray structures reproduced in Figure 3.18 correspond to Figure 3d in the page 101 of Joyce Woodhouse thesis[120]. Interestingly a (HT) photoproduct was previously observed by El Khatib et al[62] in rsFolder and rsFolder2 already in 2015.

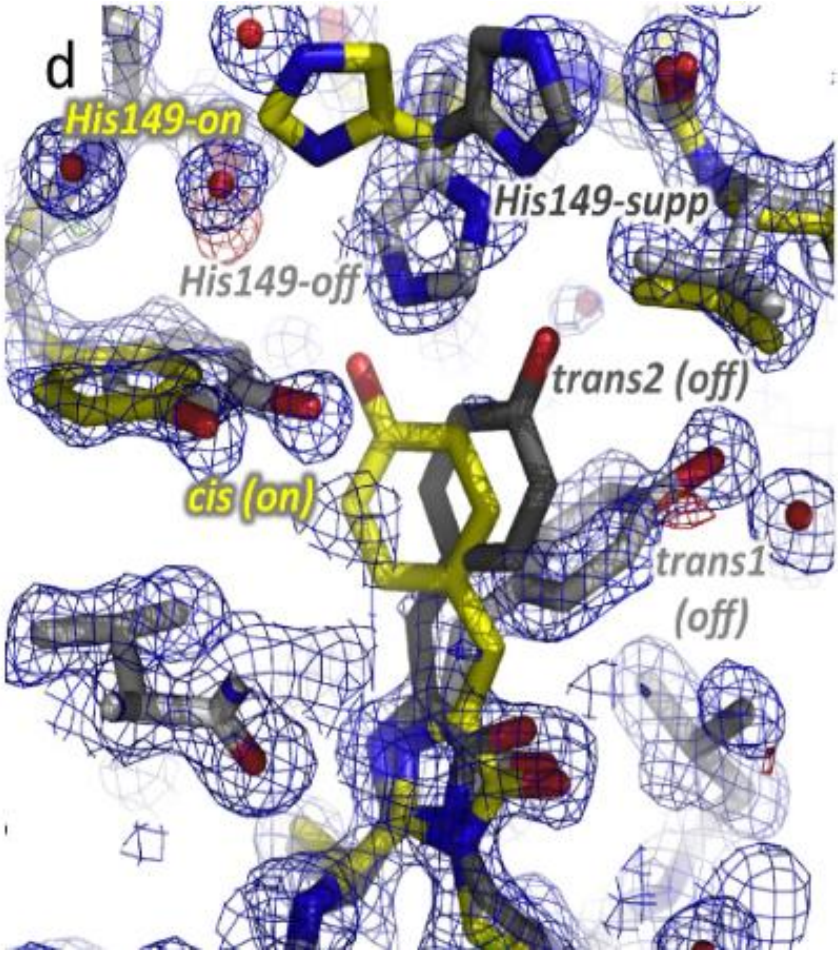


Figure 3.18. Chromophore (HBI) and its neighboring residues in rsEGFP2. The final laser-off model features triple conformations of His149 and the chromophore, i.e. His149-Off with trans1 (light grey), His149-On with cis (yellow) and the additional His149-superpose with trans2 (dark grey) conformations, at 70%, 10% and 20% occupancy, respectively. Reproduced from reference 120, page 101.

These results were reproduced by Jeffrey et al.[121] (group of S. G Boxer) in rsEGFP2 microcrystals with Cl-HBDI. They showed that after *On* to *Off* photoisomerization final *trans*

neutral photoproduct were similar to those reported in the PhD thesis of Joyce Woodhouse[120] and compatible with either OBF or hula-twist isomerizations[121]. As explained in chapter 2 (section 2.5.2), the OBF isomerization pathway is a mechanism that demands high volumes, while HT pathway can happen in sterically constraint environments, i.e. volume-conserving dynamics. A schematic representation of these two possible isomerization pathways is shown in Figure 3.19. Jeffrey et al.[121] showed that the control of the crystallization condition determines the protein cage volume, which in turn control the final isomerization photoproduct. In addition, using a chlorine derivative of HBDI (Figure 3.19, H is replaced by Cl), they could unequivocally identify the two different OBF and HT photoproducts. **The ground state *Off* conformers are entirely dependent on the environment, so far the HT photoproduct has only been observed in crystalline conditions. However, no proof has been given for the existence the rsEGFP2 *Off* ground-state heterogeneity in solution. Unravelling the *Off* and *On* ground-states heterogeneity of rsEGP2 in solution is another aim in this thesis.**

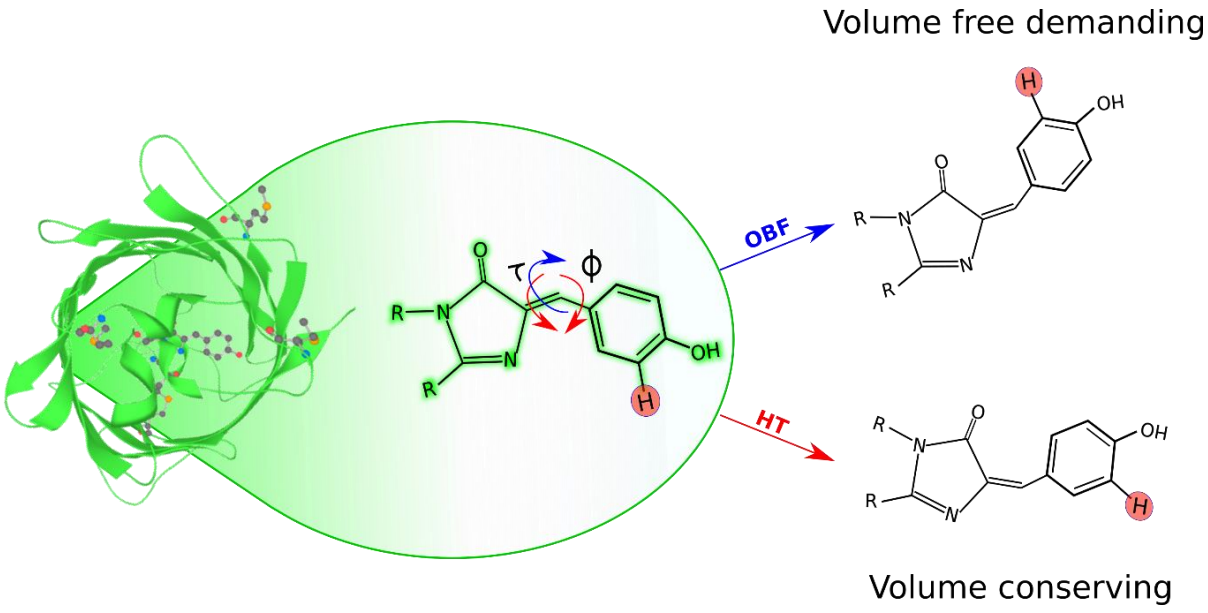


Figure 3.19. Schematic representation of rsEGFP2 chromophore cis-trans isomerization, via OBF involving a rotation around Φ, or via HT involving a rotation around Φ and τ. Figure made with Inkscape.

As mentioned previously, the results obtained by Coquelle et al.[35] allowed the development of a new RSFP where the Valine 151 was replaced by an Alanine and for which the *Off* to

On photoswitching quantum yield increases. Interestingly, the X-ray crystallographic structures in Figure 3.20, reveal that the mutation of the Valine 151 by either a leucine (V151L) or an alanine (V151A) generate variants that after photo-isomerization to the *Off* form only display one photoproduct, OBF (V151A) and HT (V151L). Not only the *trans* geometry is different, but also the hydrogen bond network and interaction with an amino acid of the protein cage. Remarkably, except for the introduced mutation in position 151**,** the *On*-state for all three variants is very similar. Furthermore, the phenol group of the chromophore in the *On* state has three possible hydrogen bonds, with a water molecule, with Thr204 and His149. The distance between the oxygen of the chromophore phenol group and His149 is 2.7 Å, 2.8 Å and 2.9 Å for WT, V151A and V151L respectively. After isomerization to the *Off* state, it is clear that there are not only rearrangement in the chromophore, but also in the protein cage; mainly involving the amino acids in the surroundings of the chromophore. A close look on Figure 3.17 reveals for the WT higher reorganizations in two amino acids, the Tyr146 and the His149. Interestingly, both amino acids are located in the surroundings of the phenol group, whilst no changes can be noticed in the surroundings of the chromophore imidazolinone carbonyl group. These differences are in contrast with Dronpa where one of the amino acids that display higher reorganizations is the Arg66 which is in direct contact with the imidazolinone carbonyl group in *On* state breaking this bond in the *Off* state[32,113] (Dronpa X-ray structures can be seen in Appendix 3). Furthermore, the chromophore in the *Off* structures of WT and V151A in Figure 3.20 adopts quite similar conformations. On the contrary, there are substantial differences between the *Off* states of the WT and V151L. The differences between *Off* conformations between WT and V151L are not only restricted to the chromophore configuration. In the *On* form, the three variants are hydrogen bond to the His149. This hydrogen bond between the chromophore and the His149 is broken and replaced by a new hydrogen-bond with a water molecule for WT and V151A variants (new distances of 2.8 and 2.9 Å respectively). At the same time, the His149 forms a new hydrogen bond with Tyr146, which adopts a new position. To the contrary, His149 in the V151L variant accompanied the conformational change of the chromophore to which it remained hydrogen-bonded -the distance between chromophore and His149 is 2.7 Å- and no conformational change is observed for Tyr146.

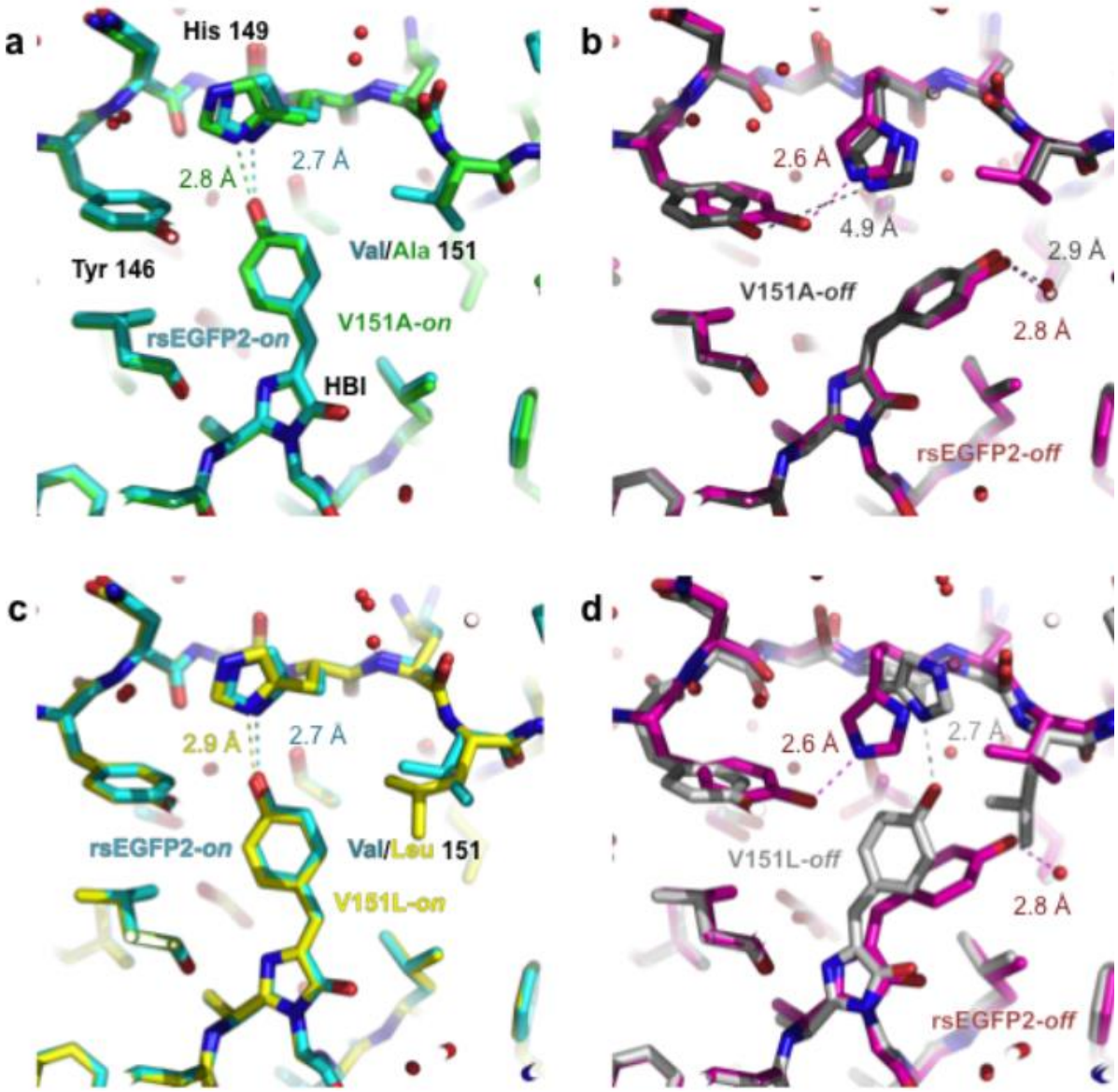


Figure 3.20. *Structural comparison between rsEGFP2 and its V151A and V151L variants. (a) Model of rsEGFP2 V151A in the On state (green) (b) Model of rsEGFP2 V151A in the Off state (dark grey) (c) Model of rsEGFP2 V151L in the On state (yellow) (d) Model of rsEGFP2 V151L in the Off state (light grey). The cis conformer was removed from rsEGFP2 V151L in the Off state for clarity. Light cyan (panels a and c) and magenta (panels b and d) correspond rsEGFP2 On and Off structures from reference [62]. Figure reproduced from Joyce Woodhouse thesis, page 129[120].*

These two mutants are ideal tools to clarify the *trans-cis* isomerization dynamics in rsEGFP2. Indeed, the details of double bond isomerization mechanisms within RSFP is still an open question, especially the existence (multiple twisted conformers) and the nature of excited and ground intermediate states[3]. **The main goal of this PhD thesis is to thoroughly characterize the photophysical properties (chapter 4) and the entire photodynamics (chapter 5) of V151A and V151L in comparison to the WT protein**. The same methodology used for the WT (Coquelle et al.[35]) will be used. **The combination of TR-SFX results (Grenoble and Heidelberg) and ultrafast optical spectroscopies (TRUV-Vis and TRIR) will allow to elucidate the photo-dynamical *Off* to *On* switching path for WT and the two variants.**

Besides, the thesis also coped with the fluorescent properties and *On* to *Off* dynamics of the WT variant (Chapter 6). The *On* to *Off* photodynamics of RSFP proteins is poorly studied mainly because the photoswitching quantum yield of the reaction is several orders of magnitude lower than the *Off* to *On*, (Table 3.1). The recent publication of S.G Boxer and coworkers showed that the study of fluorescence lifetimes can give indirect information of the isomerization[115]. The fluorescence quantum yield of rsEGFP2 is relatively low (0.34) in comparison to Dronpa (0.8), which implies a higher contribution of non-radiative excited state deactivation pathways, and thus, some information can be obtained probing non-fluorescent excited states by TR spectroscopies. **The combination of TR-SFX results (Grenoble and Heidelberg) and ultrafast optical spectroscopies (UV-Vis-IR) allows revealing the species that control the fluorescence and *On* to *Off* isomerization in rsEGPF2 (chapter 6). Finally, to validate our hypothesis, we will report preliminary results on the steady-state and photo-physical properties of 20 new rsEGFP2 variants (Chapter 7).**

3.3 References

(1) Ward, W. W.; Prentice, H. J.; Roth, A. F.; Cody, C. W.; Reeves, S. C. SPECTRAL PERTURBATIONS OF THE AEQUOREA GREEN-FLUORESCENT PROTEIN. *Photochem Photobiol* **1982**, *35*, 803–808.

(2) Ward, W. W. PROPERTIES OF THE COELENTERATE GREEN-FLUORESCENT PROTEINS. *Bioluminescence and Chemiluminescence*; Elsevier, 1981; pp 235–242.

(3) Morise, H.; Shimomura, O.; Johnson, F. H.; Winant, J. Intermolecular energy transfer in the bioluminescent system of Aequorea. *Biochemistry* **1974**, *13*, 2656–2662.

(4) Niwa, H.; Inouye, S.; Hirano, T.; Matsuno, T.; Kojima, S.; Kubota, M.; Ohashi, M.; Tsuji, F. I. Chemical nature of the light emitter of the Aequorea green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 13617–13622.

(5) Webber, N. M.; Litvinenko, K. L.; Meech, S. R. Radiationless Relaxation in a Synthetic Analogue of the Green Fluorescent Protein Chromophore. *J. Phys. Chem. B* **2001**, *105*, 8036–8039.

(6) Baldridge, A.; Samanta, S. R.; Jayaraj, N.; Ramamurthy, V.; Tolbert, L. M. Steric and electronic effects in capsule-confined green fluorescent protein chromophores. *Journal of the American Chemical Society* **2011**, *133*, 712–715.

(7) Baldridge, A.; Feng, S.; Chang, Y.-T.; Tolbert, L. M. Recapture of GFP chromophore fluorescence in a protein host. *ACS combinatorial science* **2011**, *13*, 214–217.

(8) Shimomura, O.; Johnson, F. H.; SAIGA, Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *Journal of cellular and comparative physiology* **1962**, *59*, 223–239.

(9) Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. Primary structure of the Aequorea victoria green-fluorescent protein. *Gene* **1992**, *111*, 229–233.

(10) Heim, R.; Prasher, D. C.; Tsien, R. Y. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **1994**, *91*, 12501–12504.

(11) Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science (New York, N.Y.)* **1994**, *263*, 802–805.

(12) Cody, C. W.; Prasher, D. C.; Westler, W. M.; Prendergast, F. G.; Ward, W. W. Chemical structure of the hexapeptide chromophore of the Aequorea green-fluorescent protein. *Biochemistry* **1993**, *32*, 1212–1218.

(13) Ormö, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. Crystal structure of the Aequorea victoria green fluorescent protein. *Science (New York, N.Y.)* **1996**, *273*, 1392–1395.

(14) Yang, F.; Moss, L. G.; Phillips, G. N. The molecular structure of green fluorescent protein. *Nature biotechnology* **1996**, *14*, 1246–1251.

(15) Tsien, R. Y. The green fluorescent protein. *Annual review of biochemistry* **1998**, *67*, 509–544.

(16) https://www.fpbase.org/protein/avgfp/.

(17) Hoesl, M. G.; Merkel, L.; Budisa, N. Synthetic Biology of Autofluorescent Proteins. In *Fluorescent Proteins I*; Jung, G., Ed.; Springer Series on Fluorescence; Springer Berlin Heidelberg: Berlin, Heidelberg, 2012; pp 99–130.

(18) Dedecker, P.; Schryver, F. C. de; Hofkens, J. Fluorescent proteins: shine on, you crazy diamond. *Journal of the American Chemical Society* **2013**, *135*, 2387–2402.

(19) Lippincott-Schwartz, J.; Patterson, G. H. Development and use of fluorescent protein markers in living cells. *Science (New York, N.Y.)* **2003**, *300*, 87–91.
(20) Ellenberg, J. Dual-colour imaging with GFP variants. *Trends in Cell Biology* **1999**, *9*, 52–56.
(21) Zimmer, M. Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. *Chemical reviews* **2002**, *102*, 759–781.
(22) Cormack, B. P.; Valdivia, R. H.; Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **1996**, *173*, 33–38.
(23) Cubitt, A. B.; Woollenweber, L. A.; Heim, R. Chapter 2: Understanding Structure—Function Relationships in the Aequorea victoria Green Fluorescent Protein. *Green Fluorescent Proteins*; Methods in Cell Biology; Elsevier, 1998; pp 19–30.
(24) Heim, R.; Tsien, R. Y. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology* **1996**, *6*, 178–182.
(25) Acharya, A.; Bogdanov, A. M.; Grigorenko, B. L.; Bravaya, K. B.; Nemukhin, A. V.; Lukyanov, K. A.; Krylov, A. I. Photoinduced Chemistry in Fluorescent Proteins: Curse or Blessing? *Chemical reviews* **2017**, *117*, 758–795.
(26) Lambert, T. J. FPbase: a community-editable fluorescent protein database. *Nature methods* **2019**, *16*, 277–278.
(27) Patterson, G. H.; Lippincott-Schwartz, J. A photoactivatable GFP for selective photolabeling of proteins and cells. *Science (New York, N.Y.)* **2002**, *297*, 1873–1877.
(28) Subach, F. V.; Patterson, G. H.; Manley, S.; Gillette, J. M.; Lippincott-Schwartz, J.; Verkhusha, V. V. Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nature methods* **2009**, *6*, 153–159.
(29) Ando, R.; Hama, H.; Yamamoto-Hino, M.; Mizuno, H.; Miyawaki, A. An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 12651–12656.
(30) Gurskaya, N. G.; Verkhusha, V. V.; Shcheglov, A. S.; Staroverov, D. B.; Chepurnykh, T. V.; Fradkov, A. F.; Lukyanov, S.; Lukyanov, K. A. Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. *Nature biotechnology* **2006**, *24*, 461–465.
(31) Wiedenmann, J.; Ivanchenko, S.; Oswald, F.; Schmitt, F.; Röcker, C.; Salih, A.; Spindler, K.-D.; Nienhaus, G. U. EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 15905–15910.
(32) 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.*; et al.* Infrared spectroscopy reveals multi-step multi-timescale photoactivation in the photoconvertible protein archetype dronpa. *Nature chemistry* **2018**, *10*, 845–852.
(33) Brakemann, T.; Stiel, A. C.; Weber, G.; Andresen, M.; Testa, I.; Grotjohann, T.; Leutenegger, M.; Plessmann, U.; Urlaub, H.; Eggeling, C.*; et al.* A reversibly photoswitchable GFP-like protein with fluorescence excitation decoupled from switching. *Nature biotechnology* **2011**, *29*, 942–947.

(34) Zhang, X.; Chen, X.; Zeng, Z.; Zhang, M.; Sun, Y.; Xi, P.; Peng, J.; Xu, P. Development of a reversibly switchable fluorescent protein for super-resolution optical fluctuation imaging (SOFI). *ACS nano* **2015**, *9*, 2659–2667.
(35) Coquelle, N.; Sliwa, M.; Woodhouse, J.; Schirò, G.; Adam, V.; Aquila, A.; Barends, T. R. M.; Boutet, S.; Byrdin, M.; Carbajo, S.*; et al.* Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nature chemistry* **2018**, *10*, 31–37.
(36) Grotjohann, T.; Testa, I.; Reuss, M.; Brakemann, T.; Eggeling, C.; Hell, S. W.; Jakobs, S. rsEGFP2 enables fast RESOLFT nanoscopy of living cells. *eLife* **2012**, *1*, e00248.
(37) Fuchs, J.; Böhme, S.; Oswald, F.; Hedde, P. N.; Krause, M.; Wiedenmann, J.; Nienhaus, G. U. A photoactivatable marker protein for pulse-chase imaging with superresolution. *Nature methods* **2010**, *7*, 627–630.
(38) https://www.fpbase.org.
(39) Subach, O. M.; Patterson, G. H.; Ting, L.-M.; Wang, Y.; Condeelis, J. S.; Verkhusha, V. V. A photoswitchable orange-to-far-red fluorescent protein, PSmOrange. *Nature methods* **2011**, *8*, 771–777.
(40) Subach, O. M.; Entenberg, D.; Condeelis, J. S.; Verkhusha, V. V. A FRET-facilitated photoswitching using an orange fluorescent protein with the fast photoconversion kinetics. *Journal of the American Chemical Society* **2012**, *134*, 14789–14799.
(41) Patterson, G. H.; Lippincott-Schwartz, J. A photoactivatable GFP for selective photolabeling of proteins and cells. *Science (New York, N.Y.)* **2002**, *297*, 1873–1877.
(42) Sarkisyan, K. S.; Bolotin, D. A.; Meer, M. V.; Usmanova, D. R.; Mishin, A. S.; Sharonov, G. V.; Ivankov, D. N.; Bozhanova, N. G.; Baranov, M. S.; Soylemez, O.*; et al.* Local fitness landscape of the green fluorescent protein. *Nature* **2016**, *533*, 397–401.
(43) Lin, C.-Y.; Romei, M. G.; Oltrogge, L. M.; Mathews, I. I.; Boxer, S. G. Unified Model for Photophysical and Electro-Optical Properties of Green Fluorescent Proteins. *Journal of the American Chemical Society* **2019**, *141*, 15250–15265.
(44) Kent, K. P.; Childs, W.; Boxer, S. G. Deconstructing green fluorescent protein. *Journal of American Chemical Society* **2008**, *130*, 9664–9665.
(45) Kent, K. P.; Oltrogge, L. M.; Boxer, S. G. Synthetic control of green fluorescent protein. *Journal of American Chemical Society* **2009**, *131*, 15988–15989.
(46) Carey, J.; Lindman, S.; Bauer, M.; Linse, S. Protein reconstitution and three-dimensional domain swapping: benefits and constraints of covalency. *Protein science : a publication of the Protein Society* **2007**, *16*, 2317–2333.
(47) Bauer, M. C.; Xue, W.-F.; Linse, S. Protein GB1 folding and assembly from structural elements. *International journal of molecular sciences* **2009**, *10*, 1552–1566.
(48) Lindman, S.; Hernandez-Garcia, A.; Szczepankiewicz, O.; Frohm, B.; Linse, S. In vivo protein stabilization based on fragment complementation and a split GFP system. *Proceedings of the National Academy of Sciences of the United States of America* **2010**, *107*, 19826–19831.
(49) Lin, C.-Y.; Both, J.; Do, K.; Boxer, S. G. Mechanism and bottlenecks in strand photodissociation of split green fluorescent proteins (GFPs). *Proceedings of the National Academy of Sciences of the United States of America* **2017**, *114*, E2146-E2155.

(50) Deng, A.; Boxer, S. G. Structural Insight into the Photochemistry of Split Green Fluorescent Proteins: A Unique Role for a His-Tag. *Journal of American Chemical Society* **2018**, *140*, 375–381.
(51) Do, K.; Boxer, S. G. GFP variants with alternative β-strands and their application as light-driven protease sensors: a tale of two tails. *Journal of American Chemical Society* **2013**, *135*, 10226–10229.
(52) Demadis, K. D.; Hartshorn, C. M.; Meyer, T. J. The localized-to-delocalized transition in mixed-valence chemistry. *Chemical reviews* **2001**, *101*, 2655–2686.
(53) Jensen, N. A.; Jansen, I.; Kamper, M.; Jakobs, S. Reversibly Switchable Fluorescent Proteins for RESOLFT Nanoscopy. In *Nanoscale Photonic Imaging*; Salditt, T., Egner, A., Luke, D. R., Eds.; Topics in Applied Physics; Springer International Publishing: Cham, 2020; pp 241–261.
(54) Shinoda, H.; Lu, K.; Nakashima, R.; Wazawa, T.; Noguchi, K.; Matsuda, T.; Nagai, T. Acid-Tolerant Reversibly Switchable Green Fluorescent Protein for Super-resolution Imaging under Acidic Conditions. *Cell chemical biology* **2019**, *26*, 1469-1479.e6.
(55) Arai, Y.; Takauchi, H.; Ogami, Y.; Fujiwara, S.; Nakano, M.; Matsuda, T.; Nagai, T. Spontaneously Blinking Fluorescent Protein for Simple Single Laser Super-Resolution Live Cell Imaging. *ACS chemical biology* **2018**, *13*, 1938–1943.
(56) Dedecker, P.; Hotta, J.; Flors, C.; Sliwa, M.; Uji-i, H.; Roeffaers, M. B. J.; Ando, R.; Mizuno, H.; Miyawaki, A.; Hofkens, J. Subdiffraction imaging through the selective donut-mode depletion of thermally stable photoswitchable fluorophores: numerical analysis and application to the fluorescent protein Dronpa. *Journal of the American Chemical Society* **2007**, *129*, 16132–16141.
(57) Ando, R.; Mizuno, H.; Miyawaki, A. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. *Science (New York, N.Y.)* **2004**, *306*, 1370–1373.
(58) Moeyaert, B.; Nguyen Bich, N.; Zitter, E. de; Rocha, S.; Clays, K.; Mizuno, H.; van Meervelt, L.; Hofkens, J.; Dedecker, P. Green-to-red photoconvertible Dronpa mutant for multimodal super-resolution fluorescence microscopy. *ACS nano* **2014**, *8*, 1664–1673.
(59) 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. *The journal of physical chemistry. B* **2015**, *119*, 2404–2414.
(60) Ando, R.; Flors, C.; Mizuno, H.; Hofkens, J.; Miyawaki, A. Highlighted generation of fluorescence signals using simultaneous two-color irradiation on Dronpa mutants. *Biophysical journal* **2007**, *92*, L97-9.
(61) Duwé, S.; Zitter, E. de; Gielen, V.; Moeyaert, B.; Vandenberg, W.; Grotjohann, T.; Clays, K.; Jakobs, S.; van Meervelt, L.; Dedecker, P. Expression-Enhanced Fluorescent Proteins Based on Enhanced Green Fluorescent Protein for Super-resolution Microscopy. *ACS nano* **2015**, *9*, 9528–9541.
(62) 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. *Scientific reports* **2016**, *6*, 18459.
(63) 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. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 18343–18348.

(64) Adam, V.; Moeyaert, B.; David, C. C.; Mizuno, H.; Lelimousin, M.; Dedecker, P.; Ando, R.; Miyawaki, A.; Michiels, J.; Engelborghs, Y.*; et al.* Rational design of photoconvertible and biphotochromic fluorescent proteins for advanced microscopy applications. *Chemistry & biology* **2011**, *18*, 1241–1251.
(65) Tiwari, D. K.; Arai, Y.; Yamanaka, M.; Matsuda, T.; Agetsuma, M.; Nakano, M.; Fujita, K.; Nagai, T. A fast- and positively photoswitchable fluorescent protein for ultralow-laser-power RESOLFT nanoscopy. *Nature methods* **2015**, *12*, 515–518.
(66) Adam, V.; Carpentier, P.; Violot, S.; Lelimousin, M.; Darnault, C.; Nienhaus, G. U.; Bourgeois, D. Structural basis of X-ray-induced transient photobleaching in a photoactivatable green fluorescent protein. *Journal of the American Chemical Society* **2009**, *131*, 18063–18065.
(67) Dickson, R. M.; Cubitt, A. B.; Tsien, R. Y.; Moerner, W. E. On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature* **1997**, *388*, 355–358.
(68) Weber, W.; Helms, V.; McCammon, J. A.; Langhoff, P. W. Shedding light on the dark and weakly fluorescent states of green fluorescent proteins. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 6177–6182.
(69) Larsen, D. S.; Vengris, M.; van Stokkum, I. H.; van der Horst, M. A.; Weerd, F. L. de; Hellingwerf, K. J.; van Grondelle, R. Photoisomerization and Photoionization of the Photoactive Yellow Protein Chromophore in Solution. *Biophysical journal* **2004**, *86*, 2538–2550.
(70) Litvinenko, K. L.; Webber, N. M.; Meech, S. R. Internal Conversion in the Chromophore of the Green Fluorescent Protein: Temperature Dependence and Isoviscosity Analysis. *J. Phys. Chem. A* **2003**, *107*, 2616–2623.
(71) Litvinenko, K. L.; Webber, N. M.; Meech, S. R. An ultrafast polarisation spectroscopy study of internal conversion and orientational relaxation of the chromophore of the green fluorescent protein. *Chemical Physics Letters* **2001**, *346*, 47–53.
(72) Mandal, D.; Tahara, T.; Meech, S. R. Excited-State Dynamics in the Green Fluorescent Protein Chromophore. *J. Phys. Chem. B* **2004**, *108*, 1102–1108.
(73) Ai, Y.-J.; Liao, R.-Z.; Fang, W.-H.; Luo, Y. Theoretical studies on the isomerization mechanism of the ortho-green fluorescent protein chromophore. *Physical chemistry chemical physics : PCCP* **2012**, *14*, 13409–13414.
(74) Addison, K.; Conyard, J.; Dixon, T.; Bulman Page, P. C.; Solntsev, K. M.; Meech, S. R. Ultrafast Studies of the Photophysics of Cis and Trans States of the Green Fluorescent Protein Chromophore. *The journal of physical chemistry letters* **2012**, *3*, 2298–2302.
(75) He, X.; Bell, A. F.; Tonge, P. J. Ground state isomerization of a model green fluorescent protein chromophore. *FEBS Letters* **2003**, *549*, 35–38.
(76) Chatterjee, T.; Lacombat, F.; Yadav, D.; Mandal, M.; Plaza, P.; Espagne, A.; Mandal, P. K. Ultrafast Dynamics of a Green Fluorescent Protein Chromophore Analogue: Competition between Excited-State Proton Transfer and Torsional Relaxation. *The journal of physical chemistry. B* **2016**, *120*, 9716–9722.
(77) Dong, J.; Abulwerdi, F.; Baldridge, A.; Kowalik, J.; Solntsev, K. M.; Tolbert, L. M. Isomerization in fluorescent protein chromophores involves addition/elimination. *Journal of the American Chemical Society* **2008**, *130*, 14096–14098.

(78) Hsu, Y.-H.; Chen, Y.-A.; Tseng, H.-W.; Zhang, Z.; Shen, J.-Y.; Chuang, W.-T.; Lin, T.-C.; Lee, C.-S.; Hung, W.-Y.; Hong, B.-C.*; et al.* Locked ortho- and para-core chromophores of green fluorescent protein; dramatic emission enhancement via structural constraint. *Journal of the American Chemical Society* **2014**, *136*, 11805–11812.
(79) Conyard, J.; Kondo, M.; Heisler, I. A.; Jones, G.; Baldridge, A.; Tolbert, L. M.; Solntsev, K. M.; Meech, S. R. Chemically modulating the photophysics of the GFP chromophore. *J. Phys. Chem. B* **2011**, *115*, 1571–1577.
(80) Dong, J.; Solntsev, K. M.; Poizat, O.; Tolbert, L. M. The meta-green fluorescent protein chromophore. *Journal of the American Chemical Society* **2007**, *129*, 10084–10085.
(81) Solntsev, K. M.; Poizat, O.; Dong, J.; Rehault, J.; Lou, Y.; Burda, C.; Tolbert, L. M. Meta and para effects in the ultrafast excited-state dynamics of the green fluorescent protein chromophores. *The journal of physical chemistry. B* **2008**, *112*, 2700–2711.
(82) Chen, K.-Y.; Cheng, Y.-M.; Lai, C.-H.; Hsu, C.-C.; Ho, M.-L.; Lee, G.-H.; Chou, P.-T. Ortho green fluorescence protein synthetic chromophore; excited-state intramolecular proton transfer via a seven-membered-ring hydrogen-bonding system. *Journal of the American Chemical Society* **2007**, *129*, 4534–4535.
(83) Hsieh, C.-C.; Chou, P.-T.; Shih, C.-W.; Chuang, W.-T.; Chung, M.-W.; Lee, J.; Joo, T. Comprehensive studies on an overall proton transfer cycle of the ortho-green fluorescent protein chromophore. *Journal of the American Chemical Society* **2011**, *133*, 2932–2943.
(84) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 8362–8367.
(85) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormö, M.; Remington, S. J. Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 2306–2311.
(86) Kennis, J. T. M.; Larsen, D. S.; van Stokkum, I. H. M.; Vengris, M.; van Thor, J. J.; van Grondelle, R. Uncovering the hidden ground state of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 17988–17993.
(87) Salna, B.; Benabbas, A.; Sage, J. T.; van Thor, J.; Champion, P. M. Wide-dynamic-range kinetic investigations of deep proton tunnelling in proteins. *Nature chemistry* **2016**, *8*, 874–880.
(88) Laptenok, S. P.; Lukacs, A.; Gil, A.; Brust, R.; Sazanovich, I. V.; Greetham, G. M.; Tonge, P. J.; Meech, S. R. Complete Proton Transfer Cycle in GFP and Its T203V and S205V Mutants. *Angewandte Chemie (International ed. in English)* **2015**, *54*, 9303–9307.
(89) Shu, X.; Leiderman, P.; Gepshtein, R.; Smith, N. R.; Kallio, K.; Huppert, D.; Remington, S. J. An alternative excited-state proton transfer pathway in green fluorescent protein variant S205V. *Protein science : a publication of the Protein Society* **2007**, *16*, 2703–2710.
(90) Kasha, M. Proton-transfer spectroscopy. Perturbation of the tautomerization potential. *Journal of the Chemical Society, Faraday Transactions 2* **1986**, *82*, 2379.
(91) Stoner-Ma, D.; Jaye, A. A.; Matousek, P.; Towrie, M.; Meech, S. R.; Tonge, P. J. Observation of excited-state proton transfer in green fluorescent protein using ultrafast vibrational spectroscopy. *Journal of American Chemical Society* **2005**, *127*, 2864–2865.

(92) Fang, C.; Frontiera, R. R.; Tran, R.; Mathies, R. A. Mapping GFP structure evolution during proton transfer with femtosecond Raman spectroscopy. *Nature* **2009**, *462*, 200–204.
(93) Meech, S. R. Excited state reactions in fluorescent proteins. *Chemical Society reviews* **2009**, *38*, 2922–2934.
(94) Fujisawa, T.; Kuramochi, H.; Hosoi, H.; Takeuchi, S.; Tahara, T. Role of Coherent Low-Frequency Motion in Excited-State Proton Transfer of Green Fluorescent Protein Studied by Time-Resolved Impulsive Stimulated Raman Spectroscopy. *Journal of the American Chemical Society* **2016**, *138*, 3942–3945.
(95) Tang, L.; Wang, Y.; Zhu, L.; Kallio, K.; Remington, S. J.; Fang, C. Photoinduced proton transfer inside an engineered green fluorescent protein: a stepwise-concerted-hybrid reaction. *Physical Chemistry Chemical Physics.* **2018**, *20*, 12517–12526.
(96) Stoner-Ma, D.; Jaye, A. A.; Ronayne, K. L.; Nappa, J.; Meech, S. R.; Tonge, P. J. An alternate proton acceptor for excited-state proton transfer in green fluorescent protein: rewiring GFP. *Journal of the American Chemical Society* **2008**, *130*, 1227–1235.
(97) Shibazaki, C.; Shimizu, R.; Kagotani, Y.; Ostermann, A.; Schrader, T. E.; Adachi, M. Direct Observation of the Protonation States in the Mutant Green Fluorescent Protein. *The journal of physical chemistry letters* **2020**, *11*, 492–496.
(98) Nifosì, R.; Ferrari, A.; Arcangeli, C.; Tozzini, V.; Pellegrini, V.; Beltram, F. Photoreversible Dark State in a Tristable Green Fluorescent Protein Variant. *J. Phys. Chem. B* **2003**, *107*, 1679–1684.
(99) Chirico, G.; Diaspro, A.; Cannone, F.; Collini, M.; Bologna, S.; Pellegrini, V.; Beltram, F. Selective fluorescence recovery after bleaching of single E2GFP proteins induced by two-photon excitation. *Chemphyschem : a European journal of chemical physics and physical chemistry* **2005**, *6*, 328–335.
(100) Ghosh, A.; Isbaner, S.; Veiga-Gutiérrez, M.; Gregor, I.; Enderlein, J.; Karedla, N. Quantifying Microsecond Transition Times Using Fluorescence Lifetime Correlation Spectroscopy. *The journal of physical chemistry letters* **2017**, *8*, 6022–6028.
(101) Arpino, J. A. J.; Rizkallah, P. J.; Jones, D. D. Crystal structure of enhanced green fluorescent protein to 1.35 Å resolution reveals alternative conformations for Glu222. *PloS one* **2012**, *7*, e47132.
(102) Kirk, W.; Allen, T.; Atanasova, E.; Wessels, W.; Yao, J.; Prendergast, F. Photophysics of EGFP (E222H) Mutant, with Comparisons to Model Chromophores: Excited State pK's, Progressions, Quenching and Exciton Interaction. *Journal of fluorescence* **2017**, *27*, 895–919.
(103) Bizzarri, R.; Serresi, M.; Cardarelli, F.; Abbruzzetti, S.; Campanini, B.; Viappiani, C.; Beltram, F. Single amino acid replacement makes Aequorea victoria fluorescent proteins reversibly photoswitchable. *Journal of the American Chemical Society* **2010**, *132*, 85–95.
(104) Fron, E.; Flors, C.; Schweitzer, G.; Habuchi, S.; Mizuno, H.; Ando, R.; Schryver, F. C. de; Miyawaki, A.; Hofkens, J. Ultrafast excited-state dynamics of the photoswitchable protein Dronpa. *Journal of the American Chemical Society* **2007**, *129*, 4870–4871.
(105) 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. *Nature communications* **2013**, *4*, 1461.

(106) 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. *The journal of physical chemistry. B* **2015**, *119*, 2350–2362.
(107) 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. *The journal of physical chemistry. B* **2013**, *117*, 11954–11959.
(108) Fron, E.; van der Auweraer, M.; Hofkens, J.; Dedecker, P. Excited state dynamics of photoswitchable fluorescent protein Padron. *The journal of physical chemistry. B* **2013**, *117*, 16422–16427.
(109) 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. *The journal of physical chemistry. B* **2015**, *119*, 5136–5144.
(110) Lacombat, F.; Plaza, P.; Plamont, M.-A.; Espagne, A. Photoinduced Chromophore Hydration in the Fluorescent Protein Dreiklang Is Triggered by Ultrafast Excited-State Proton Transfer Coupled to a Low-Frequency Vibration. *The journal of physical chemistry letters* **2017**, *8*, 1489–1495.
(111) Colletier, J.-P.; Sliwa, M.; Gallat, F.-X.; Sugahara, M.; Guillon, V.; Schirò, G.; Coquelle, N.; Woodhouse, J.; Roux, L.; Gotthard, G.*; et al.* Serial Femtosecond Crystallography and Ultrafast Absorption Spectroscopy of the Photoswitchable Fluorescent Protein IrisFP. *The journal of physical chemistry letters* **2016**, *7*, 882–887.
(112) Woodhouse, J.; Nass Kovacs, G.; Coquelle, N.; Uriarte, L. M.; Adam, V.; Barends, T. R. M.; Byrdin, M.; La Mora, E. de; Bruce Doak, R.; Feliks, M.*; et al.* Photoswitching mechanism of a fluorescent protein revealed by time-resolved crystallography and transient absorption spectroscopy. *Nature communications* **2020**, *11*, 741.
(113) 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. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 13005–13009.
(114) 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. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 9511–9516.
(115) Romei, M. G.; Lin, C.-Y.; Mathews, I. I.; Boxer, S. G. Electrostatic control of photoisomerization pathways in proteins. *Science (New York, N.Y.)* **2020**, *367*, 76–79.
(116) Andresen, M.; Stiel, A. C.; Fölling, J.; Wenzel, D.; Schönle, A.; Egner, A.; Eggeling, C.; Hell, S. W.; Jakobs, S. Photoswitchable fluorescent proteins enable monochromatic multilabel imaging and dual color fluorescence nanoscopy. *Nature biotechnology* **2008**, *26*, 1035–1040.
(117) Brakemann, T.; Weber, G.; Andresen, M.; Groenhof, G.; Stiel, A. C.; Trowitzsch, S.; Eggeling, C.; Grubmüller, H.; Hell, S. W.; Wahl, M. C.*; et al.* Molecular basis of the light-driven switching of the photochromic fluorescent protein Padron. *The Journal of biological chemistry* **2010**, *285*, 14603–14609.
(118) Bourgeois, D.; Adam, V. Reversible photoswitching in fluorescent proteins: a mechanistic view. *IUBMB life* **2012**, *64*, 482–491.

(119) Grotjohann, T.; Testa, I.; Leutenegger, M.; Bock, H.; Urban, N. T.; Lavoie-Cardinal, F.; Willig, K. I.; Eggeling, C.; Jakobs, S.; Hell, S. W. Diffraction-unlimited all-optical imaging and writing with a photochromic GFP. *Nature* **2011**, *478*, 204–208.
(120) Joyce Woodhouse. *Studying a reversibly switchable fluorescent protein by time-resolved crystallography using the X-ray free electron lasers. Biological Physics.*; Université Grenoble Alpes, 2018.
(121) Chang, J.; Romei, M. G.; Boxer, S. G. Structural Evidence of Photoisomerization Pathways in Fluorescent Proteins. *Journal of the American Chemical Society* **2019**, *141*, 15504–15508.

4 Photophysical properties of rsEGFP$_2$ WT, V151A and V151L.

rsEGFP2 is a negative RSFP. The resting state of the protein is the fluorescent *On* state; the chromophore adopts a *cis* configuration, and the phenol group is deprotonated (anionic *cis*-phenolate)[1,2]. The non-resting state of the protein is non-fluorescent, the chromophore displays mainly a *trans* configuration with the phenol group protonated (neutral *trans*-phenol)[1,2]. As pointed in the introduction, there is a high discrepancy in the switching yield determination, measured by absorbance in solution and fluorescence in a cuvette compared to proteins immobilized in a film. Since we will use time-resolved transient absorption, we decided to determine *On* to *Off* and *Off* to *On* switching yields monitoring, the light-induced changes of the proteins in solution under continues irradiation, with UV-vis absorption spectroscopy. The precise determination of the switching yield and the photo-stationary states obtained after a certain irradiation time is not trivial, i.e. the concentration of *On* and *Off* forms ($[on]^{\infty}$, $[off]^{\infty}$) at the thermodynamic equilibrium under specific irradiation needs to be determined. One more time, it should be stressed the difference between *On/Off* fluorescent states and *On*/O*ff* forms. A fluorescent state might be formed by several *cis/trans* anionic/neutral forms, and each form can be a mixture of different conformers. For example, a HT and OBF *trans Off* conformers compose the *Off* form, which in turn together with residual On form (cis anionic) compose the Off state. To get these parameters, we need to determine precisely molar absorption coefficient for *cis* anionic and *trans* neutral forms and the thermal back recovery from the *trans* neutral form to *cis* anionic (see equations 4.2, 4.3 and 4.4 and definitions in Chapter 2).

Furthermore, potentially in a photoactivation process, there are two possible intermediate forms that the chromophore could adopt a *cis*-phenol and a trans-phenolate. Although the *Off* to *On* photoisomerization first step is still a matter of debate[3–5], it is clear that the deprotonation is the last step of the photo-cycle[4–7]. Therefore, the formation of the *trans*-phenolate is very improbable as this would imply the proton transfer to occur in the excited state (ESPT) and previous to the isomerization. Thus here, we also determined the

spectroscopy properties of the *cis*-phenol form. For a correct determination of this form, the pK_a values of WT and variants need to be known.

Therefore, in this chapter, we determine the main photophysical properties for WT, V151A and V151L (Figure 4.1): (i) the molar absorption coefficient spectrum for *trans* neutral, *cis* neutral and *cis* anionic forms. (ii) The fluorescence spectrum, fluorescent quantum yields and brightness for the *cis* anionic form. (iii) The photoswitching quantum yields (*trans* neutral form to *cis* anionic form and vice versa). (iv) the thermal back recovery of *trans* neutral form. (v) the pK_a of *cis* neutral/*cis* anionic. (vi) absorption spectra for *trans* neutral and *cis* anionic form in crystal environment. (vii) The FTIR spectra for *cis* anionic and *trans* neutral forms (Shown in Appendix 1). The interpretation of different photophysical properties was made with the help of X-ray crystallographic structures (shown in Chapter 3, and obtained by our collaborators[8]).

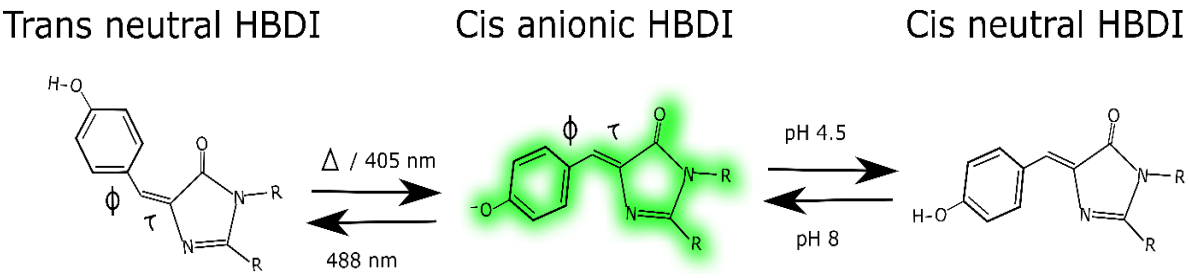


Figure 4.1 HBDI forms involved in the photodynamics of rsEGFP2: trans neutral (Off form), cis anionic (fluorescent, On form) and cis neutral form.

4.1. Molar absorption coefficient spectrum of trans neutral, cis neutral and cis anionic form

There are different ways to calculate the molar absorption coefficients of a FP. The main difficulty is to determine the protein concentration in solution. A general method is to consider the number of aromatic amino acids (phenylalanine, tryptophan and tyrosines) that the protein has and correlated with the absorption of the protein at 280 nm (mainly tryptophan contribution)**.** However, this method assumes that the chromophore does not absorb at this wavelength or have a minimal contribution. This is not the case for FP chromophore (HBDI) which contributes to the 280 nm absorbance band. The W. Ward method[9,10] for FP does not

suffer from this approximation. It is based on denaturing the protein and is one of the most established methods. The method consists in comparing the FPs denatured absorption spectrum with the pure chromophore spectrum (HBDI) in solution. **S**ince the molar absorption coefficient of the chromophore is known, one can calculate the protein concentration using Beer-Lambert equation (Chapter 2, Equation 2.1). The molar absorption coefficients were calculated by denaturation of the protein at basic pH. The details of the procedure used are given in the following section. The correct determination of these values has significant importance, as it is the starting point in the determination of several other photophysical properties.

4.1.1 Molar absorption coefficient spectrum of trans neutral and cis anionic form

Between 9-15 µL of protein were diluted into 3 mL buffer (50 mMol Hepes 50 mMol NaCl pH 8) and transferred to a spectroscopic cuvette (1 cm x 1 cm) to have an absorbance value between 0.3-0.5 at 480 nm. Protein solutions were irradiated under stirring with a homemade build LED (3W Ultraviolet UV LED (390-400nm) EPILED -already soldered to the *printed circuit board* (PCB)- and with an emission maximum at 395 nm; https://futureeden.co.uk/) to ensure a complete photo-transformation of RSFPs into the *cis*-anionic *On* form. We observed that the absence of irradiation at 395 nm would lower the value of *cis* anionic concentration, i.e. there was a mixture of *cis* anionic and *trans* neutral forms. The absorption spectrum of the pure *cis* anionic form (*On* state) was then measured. The absorption spectrum of *trans* neutral form was extracted from the photo-stationary absorption spectrum. This spectrum was obtained by irradiation of the protein solutions with a continuous LED light source at 505 nm (Thorlabs M505L3). To ensure the absence of photo-degradation, the solutions were switched back by 395 nm irradiation and the absorption spectrum compared to the initial one of the *On* form. Finally, 115 µL of NaOH 6M solution were added to denature the proteins. The protein concentration was then determined using the absorption coefficient of the anionic GFP denatured chromophore (44100 $cm^{-1}.M^{-1}$; Beer-Lambert equation)[10]. All the absorption spectra were collected with a Cary 3500 absorption spectrometer (spectral bandwidth of 0.4 nm and one point every 0.5 nm). Furthermore, for the *Off* state, two consecutive spectra were measured with a 5 minutes interval (the irradiation

LED at 505 nm is turn on within this time) to ensure that the photo-stationary absorption spectrum is collected (if two spectra are identical photo-stationary state has been reached).

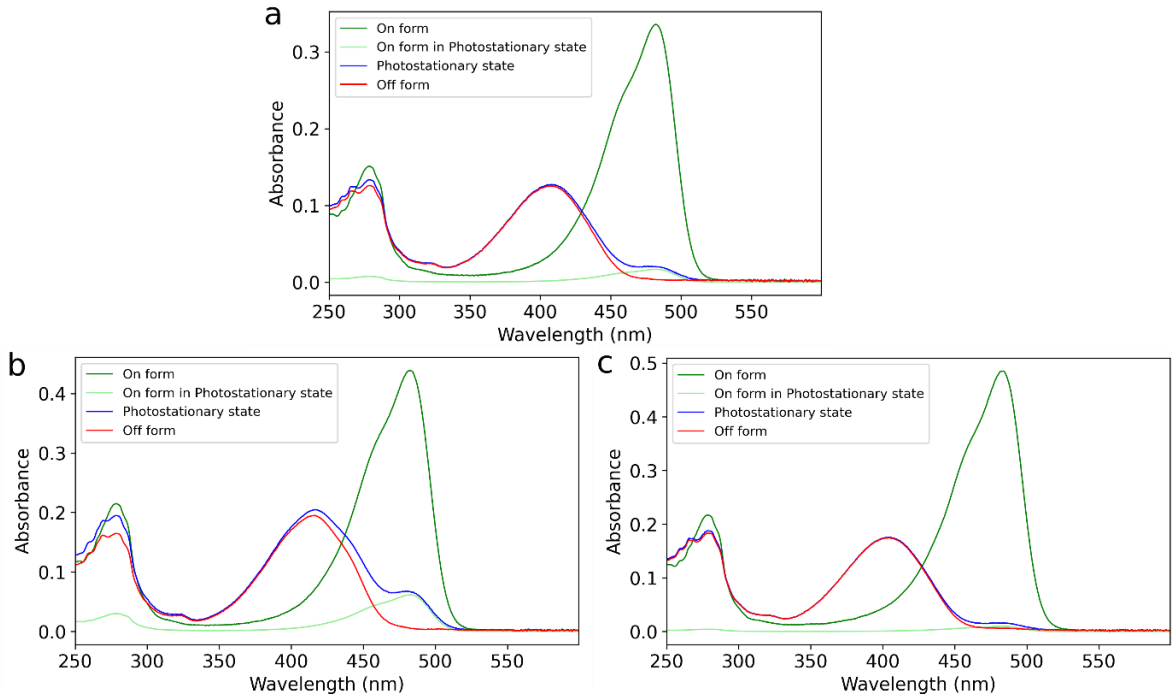


Figure 4.1 Absorption spectra for the photostationary Off state (blue), the On form (green), the Off form (red) and the On form contribution at the photo-stationary state (light green) for a) WT, b) V151L and c) V151A. Figure made with matplotlib python library.

The determination of the *trans* neutral spectrum is not trivial. The absorption spectrum at the photostationary state (*Off* state) is the sum of *cis* anionic and *trans* neutral forms absorption spectra because the chromophore is weakly switched back to the *On*-state by the *Off*-switching light. Therefore 100% conversion to the *trans* neutral form in the *Off* state can not be achieved. For WT and V151L, the *cis*-anionic concentration in the photo-stationary *Off* state is significant while for the V151A is nearly negligible. This indicates that at the photo-stationary state, the conversion to *trans* neutral form is much more efficient for V151A. The photostationary state spectra are represented in Figure 4.1a b and c for WT, V151L and V151A, respectively. The contribution of the residual *On* population (*cis* anionic form) can be calculated using the value of absorbance at 482 nm and subtracting it to the photo-stationary *Off* state spectra to get pure *trans* neutral *Off* form spectrum. These values were

5%, 14% and 2% of *cis* anionic form for WT, V151L and V151A at the photo-stationary state respectively (represented in light green in Figure 4.1).

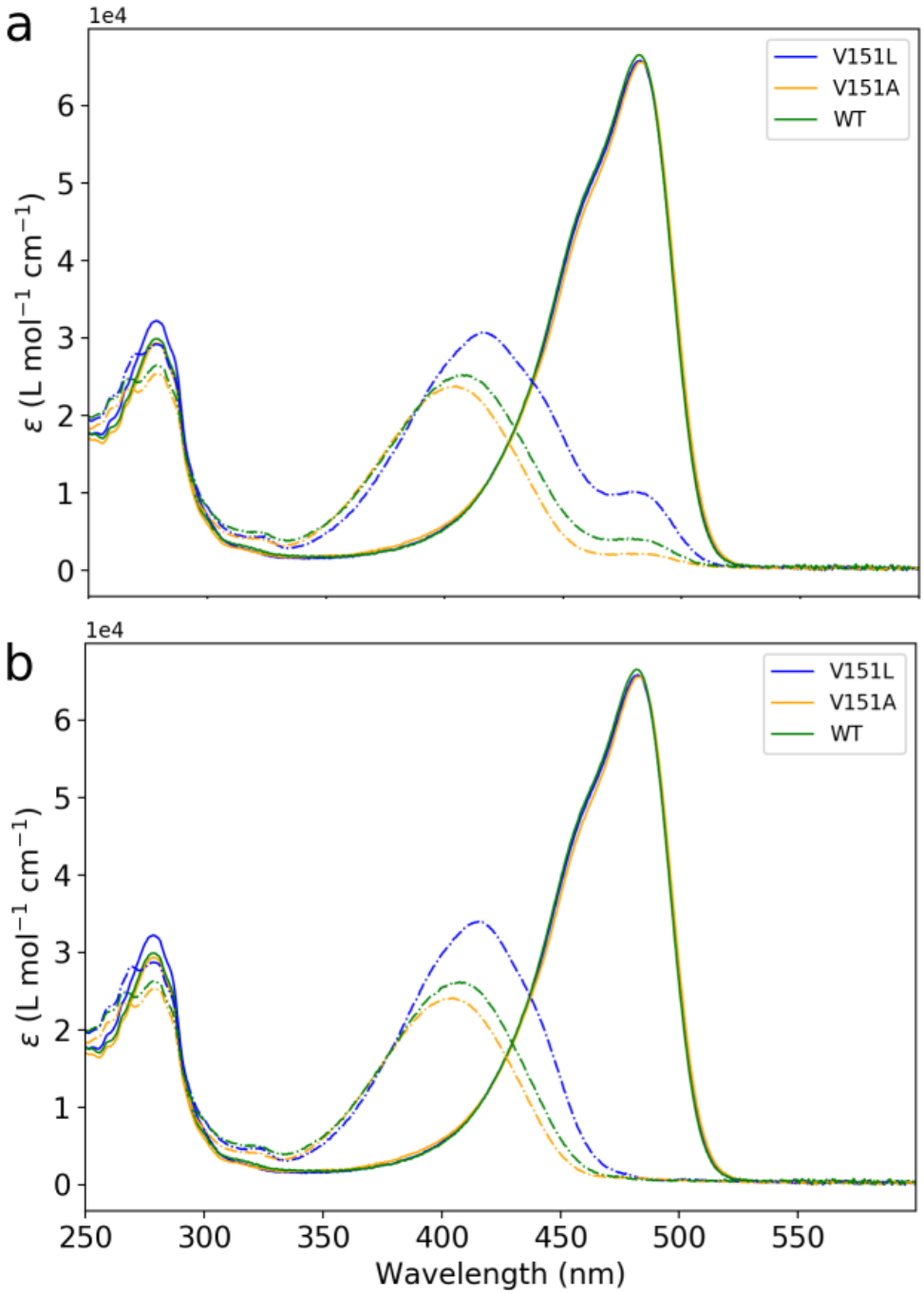


Figure 4.2. a) Molar absorption coefficient spectra for On (solid line) and Off state (photo-stationary state, dashed line) for WT (Green), V151L (blue) and V151A (orange). b) Molar absorption coefficient spectra for the cis anionic On form and trans neutral Off form. Figure made with matplotlib python library.

The molar absorption coefficient spectra for *On* form and *Off* state (a), and for the *cis* anionic and *trans* neutral form (b) at pH 8 are shown in Figure 4.2a and 4.2b. The maxima values for WT, V151A and V151L are summarized in Table 4.3 in the discussion section of the chapter. As expected from the X-ray structure (Chapter 3.4), the *cis*-anionic form for WT, V151L and V151A has the same absorption maximum at 482-483 nm and similar molar absorption coefficients (~66000 M^{-1} cm^{-1}; see Table 4.3 and Table 4S-1 at the end of the chapter) and assigned to π-π* transition**.** The *On* form absorption spectra are not symmetric; they all display a shoulder at 465 nm, revealing a vibronic structure. A similar result has been reported for PYP[11] and the avGFP at low temperature[12]. This vibronic spectral signature of the *On* form is assigned to a higher constrained chromophore environment compared to the *Off* form. The only difference between WT and the variants for the *O*n form is observed in the UV region. At 280 nm a hyperchromic effect is observed for V151L with an increase of 10% in amplitude, which shows the advantage of using the Ward method for a precise determination of the molar absorption coefficient. Contrary to the *cis* anionic spectra, the *trans* neutral forms of WT, V151A and V151L have different maxima and different molar coefficient values **(**WT and V151A display similar values of 26443 and 24640 M^{-1} cm^{-1}, while the V151L has nearly 50% higher values with 34231 M^{-1} cm^{-1}). The first absorption maximum is blue-shifted (hypsochrome effect) for V151A and red-shifted (bathochromic effect) for V151L in comparison to WT one. Indeed, the absorbance maximum for *trans* neutral form is at 416 nm, 408 nm and 405 nm for V151L, WT and V151A respectively. An important result is the existence of a hyperchromic effect (an increase of molar absorption coefficient) for V151L and hypochromic effect (a decrease of molar absorption coefficient) for V151A in comparison to WT. Therefore, the WT *trans* neutral spectrum can be rationalized by the sum of both *Off* forms, the OBF (V151A) and HT (V151L) forms, in line with the microcrystals results (Figure 3.18). For V151L variant, there is an extra shoulder at 440 nm that can be assigned to different hydrogen bond network compared to WT and V151A. From the X-ray structure in (Chapter 3.4), this shoulder, can be attributed to the H-bond with the His149 that leads to different electronic delocalization (contribution of a phenolate-type mesomeric form). Another interpretation of this shoulder is a more rigid chromophore for V151L. Theoretical quantum calculations to modelized the different effects will be needed to further interpret the results.

4.1.2 Molar absorption coefficient spectrum of *cis* neutral form

Before the determination of the *cis*-neutral molar absorption coefficient spectrum for each variant, the pK_a values for the *cis* form were determined. To measure the pK_a 14 buffer solutions containing 50 mMol glycine, 50 mMol phosphoric acid and 50 mMol citric acid were prepared. For each solution, the pH was adapted, adding NaOH to reach the corresponding pH and covering the range 4.0-10 with one point every 0.5 pH units. The chemicals, glycine, (BioUltra), for molecular biology, ≥99.0%, phosphoric acid, (BioReagent), suitable for insect cell culture, 85% and citric acid, (BioUltra), anhydrous, ≥99.5%, were purchase from sigma Aldrich.

For every pH point, a constant volume of concentrated proteins (between 3-6 µL depending on the protein variant) was diluted into 3 mL buffer and transferred to a cuvette (1 cm x 1 cm). Figure 4.3 shows the absorption spectra of WT and two variants recorded from pH 4 to pH 11 with 0.5 pH steps (Cary 3500 absorption spectrometer, a spectral bandwidth of 2 nm and one point every 2 nm). Proteins started to denature at pH higher than 11, and a different form appears (shift of the absorbance maximum) for pH values below 4.5, probably due to the protonation of the imadozolinone nitrogen[9].

To determine the pK_a, the Henderson–Hasselbalch equation (Equation 4.1) was used with [A^-= cis anionic form] and [HA = cis neutral form].

$$pH = pKa + log_{10}\left(\frac{[A^-]}{[HA]}\right) \qquad \text{Equation 4.1}$$

The values for [A^-] and [HA] can be calculated using Beer-Lambert law and the absorbance value at 480 nm for different pH. The assumption is that only the anionic form [A^-] absorbs at 480 nm and mass conservation (thus the sum of the molar fractions of [A^-] and [HA] is equal to 1). The spectra were normalized by the average value from the spectra at pH 8.5 to 9.5 at 480 nm. The data were then fitted using a homemade python 3.7 routine (Figure 4.3). The pK_a values obtained were 5.7, 5.8 and 6.3 for rsEGP2 WT, V151L and V151A, respectively.

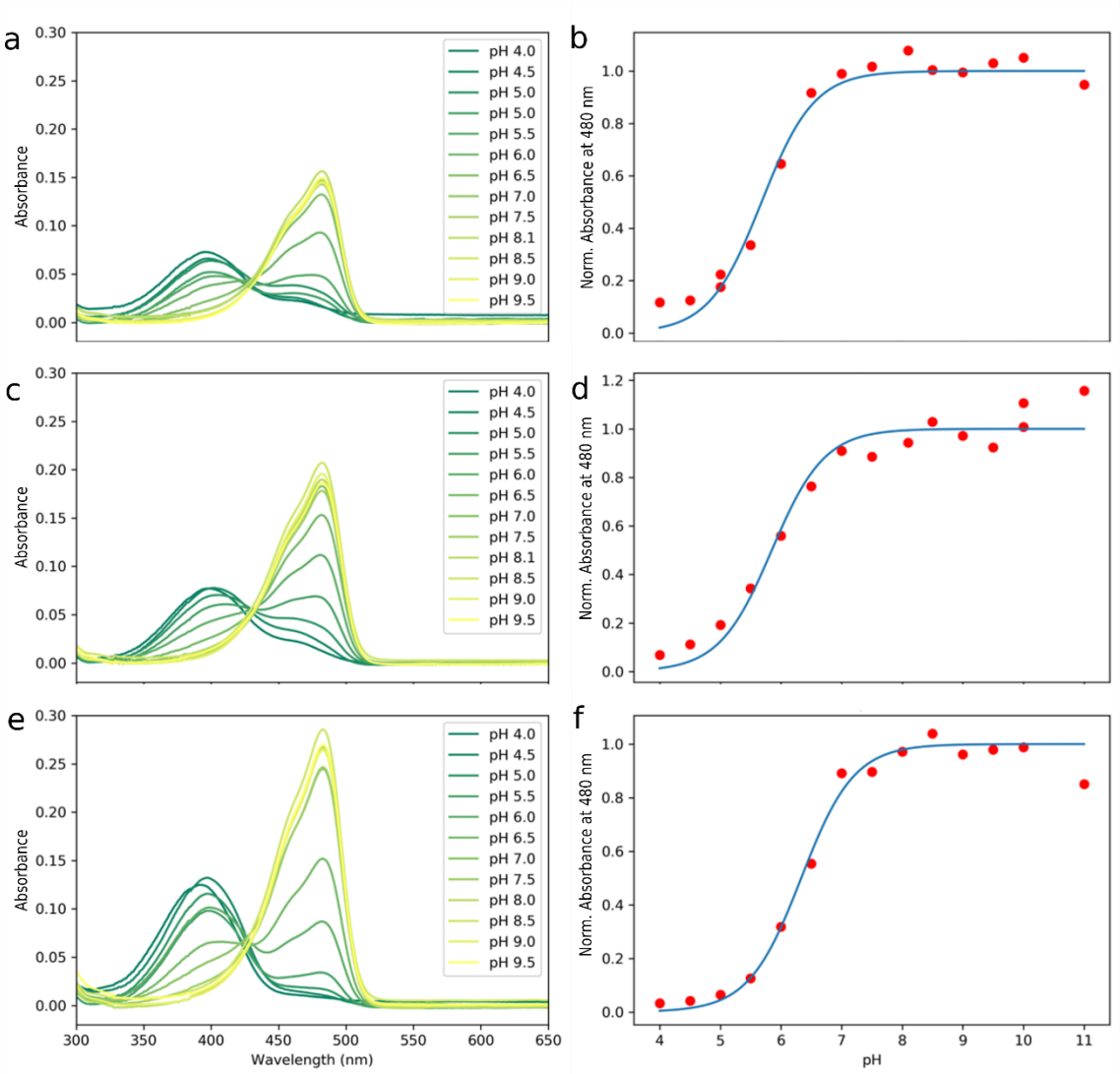


Figure 4.3. Absorption spectra at different pH for WT (a), V151L (b) and V151A (c). Absorbance at 480 nm for WT(b), V151L (d), V151A (f) and fit of Henderson–Hasselbalch equation in blue. The spectra were normalized by the average absorbance value of spectra at pH 8.5, 9 and 9.5 at 480 nm. Figure made with matplotlib python library.

Once the pK_a values have been determined, it is possible to calculate the *cis*-neutral molar absorption spectrum for WT and each variant. For this, the contribution of *cis* anionic form at pH 4.5 is calculated knowing the corresponding pK_a and using the Henderson–Hasselbalch (equation 4.1). Then this contribution is subtracted from the spectrum at pH 4.5 (at lower pH values the imidazolinone nitrogen is protonated). The example for V151L is given in Figure 4.4.

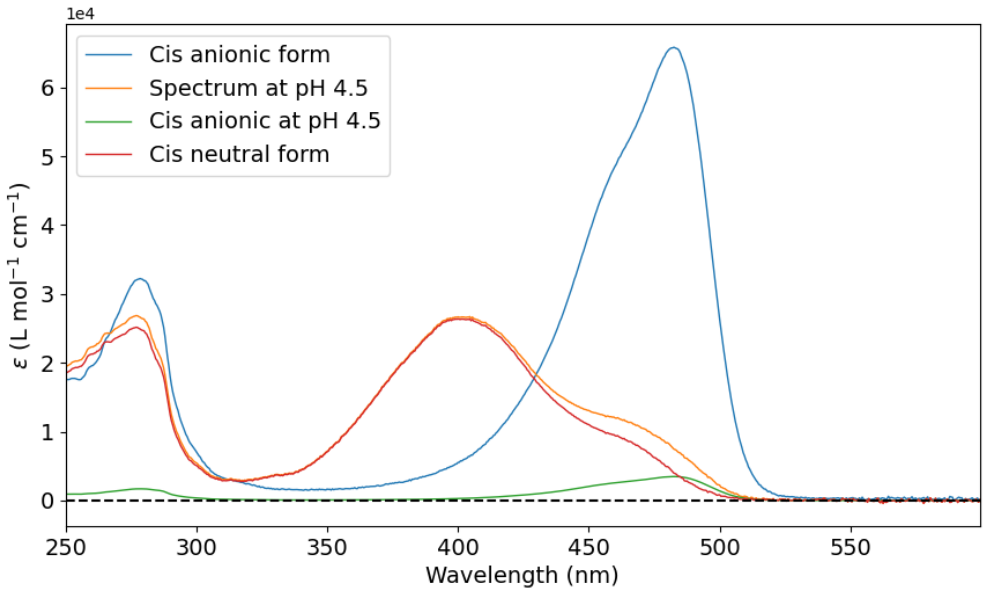


Figure 4.4. cis-neutral molar absorption coefficient spectrum for V151L. In orange the measured molar absorption coefficient spectrum at pH 4.5; in blue, the cis anionic molar absorption coefficient spectrum, in green the cis anionic contribution at pH 4.5 and in red, the cis neutral molar absorption coefficient spectrum after the subtraction of cis anionic contribution. Figure made with matplotlib python library.

All UV-Vis absorption molar coefficient spectra for the *cis* anionic, *cis* neutral and *trans* neutral forms determined for WT and two variants can be seen together in Figure 4.5. After the correction of the resting *On*-forms, the *cis* neutral form molar absorption coefficient spectra are characterized by one main band with maxima at 397, 401 and 395 nm for WT, V151L and V151A, and a shoulder at 468, 466 and 476 nm. The absorption maximum of the *cis* neutral form is blue-shifted in comparison with the absorption maximum of the *trans* neutral form. Contrary to what was previously stated for other proteins[7]; we think that this shoulders cannot be assigned to the *cis* anionic remaining *On* form because it was subtracted and the shoulder maximum is blue-shifted respect to the maximum of *On* form. This absorption shoulder is assigned to different mesomeric forms like in PYP[11]. For a precise comparison, theoretical calculation would be needed.

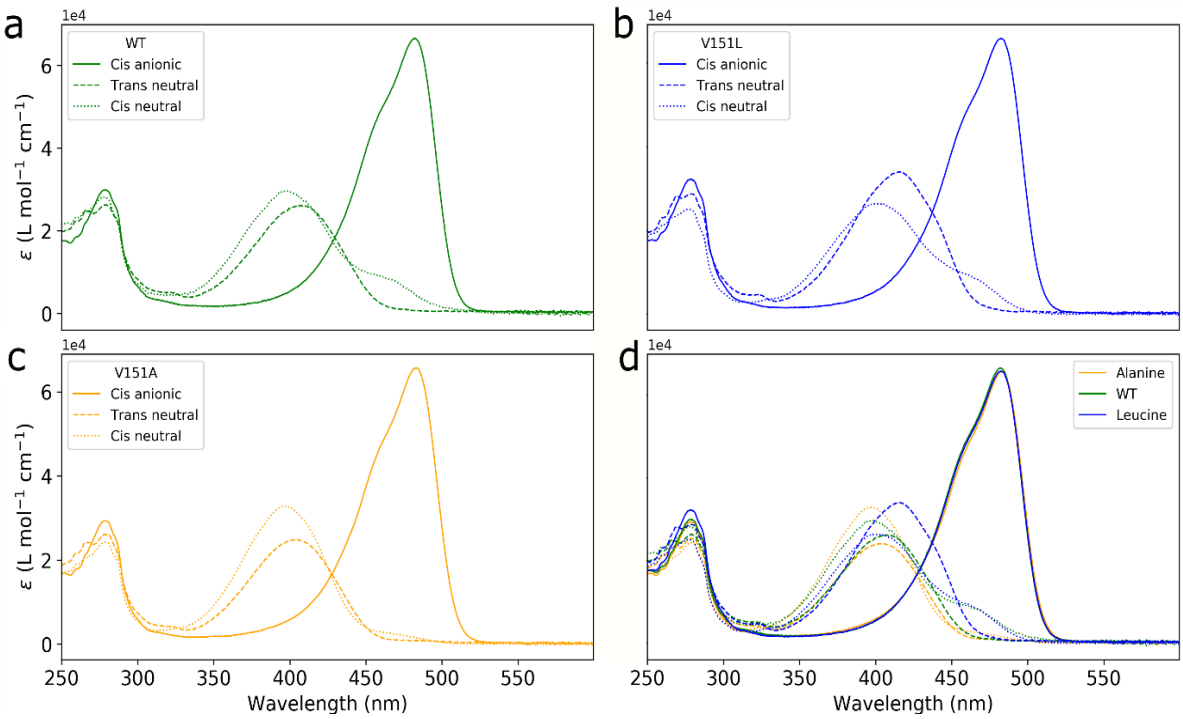


Figure 4.5. Trans neutral Off form (dashed line), cis anionic On form (solid line) and cis neutral form (dotted line) molar absorption coefficient spectra for WT (a), V151L (b) and V151A (c). All spectra are compared in (d). Figure made with matplotlib python library

4.1.3 cis anionic and trans neutral form absorption spectra in microcrystals.

The absorption spectra of the microcrystal colloidal solution were measured with a Nanodrop 2000c spectrometer (2 µL, baseline correction on the buffer, Thermo Scientific, nanodrop http://www.nanodrop.com), "before" and "after irradiation" (488-nm, 200 mW, Cobolt laser) using a homemade custom setup[13] to obtain the *On* and *Off* spectral state. The spectra with and without irradiation were recorded three times (refreshing the sample) and averaged. The recorded absorption spectra are considered as an overlapped of the microcrystal absorbance and a baseline artefact with a main scattering contribution (baseline shift). An example of the baseline and scattering subtraction is shown in Figure 4.6. The baseline and the scattering contribution were estimated, fitting the measured absorption spectra (red spectra in Figure 4.6a,b) by a sum of 3 Gaussian function plus an exponential function. The Gaussian function

reproduces the *On* and *Off* absorption spectra of rsEGFP2 proteins (maxima at 280 nm, 460 nm and 482 nm for on state; and 280 nm, 400 nm and 482 nm for off state) and the exponential simulates the scattering and baseline contribution. Note that a $1/\lambda^4$ correction has been tried, nevertheless, the comparison of the corrected spectra with those obtained in solution revealed an over estimation of the scattering contribution for the 250-300 nm region with a $1/\lambda^4$ polynom. The resulting estimation by an exponential function was subtracted to correct the spectra (Orange spectrum in Figure 4.20c,d) which was further smoothed with a Savitzky-Golay[14] filter using a 25 points window and a third-order polynomial. The same procedure was applied to all recorded *On* and *Off* spectra for WT and two variants.

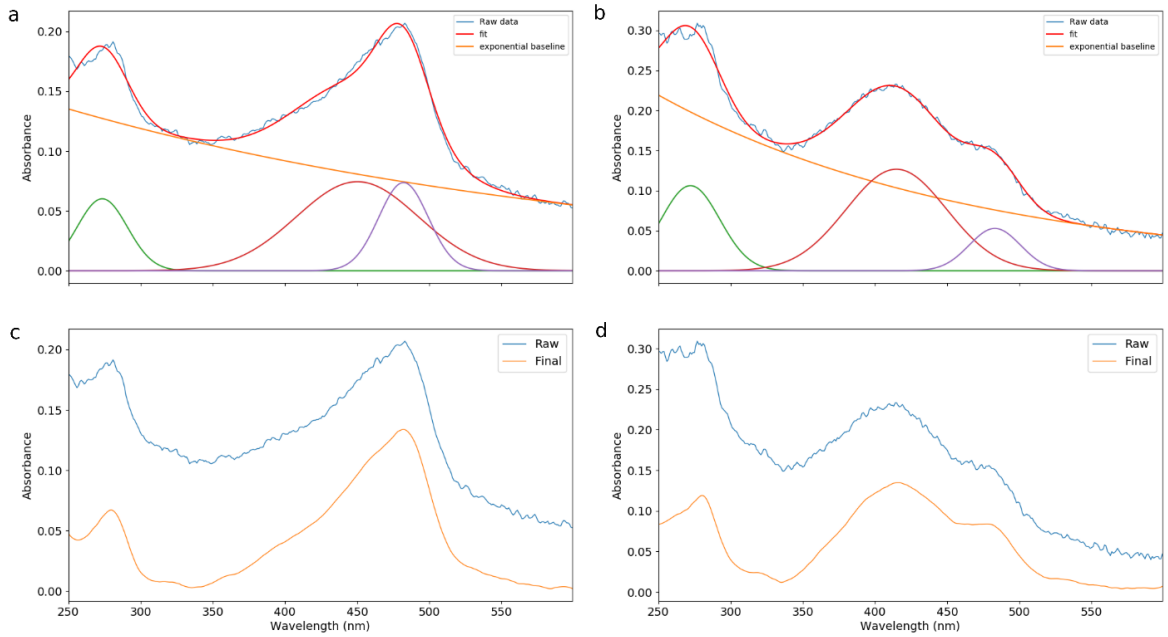


Figure 4.6. Top panels, correspond to the estimation of the baseline for V151A before irradiation spectrum by fitting it with a sum of 3 Gaussian functions to reproduce protein absorption and an exponential function (orange) to estimate the baseline artifact and the scattering contribution. a) V151A On form and b) V151A Off form. Bottom panel the raw data (blue) and final spectra (orange) can be seen.

After pre-processing, the "before" and "after" irradiation recorded spectra were normalized by the corresponding intensities at 280 nm for comparison. To calculate the percentage of *On* form that is switched to *Off* form after irradiation, we fitted the "after irradiation" spectrum (a mixture of *Off* and *On* forms) with two Gaussian functions to determine the

corresponding contributions of *On* and *Off* forms (Orange and Green lines in the Figure 4.7c,d). The absorbance of *On* contribution (thin green gaussian) at 482 nm was compared to the absorbance at 482 nm of the "without irradiation" spectrum (only *On* form) to determine the percentage of conversion. Using this procedure, we determined a conversion from *On* state to *Off* state of 85% for WT and V151A and 77% for V151L. The results are seen in Figure 4.7. The differences between crystal and solution absorption properties have been compared for the WT and the two variants. For this comparison, the different absorption bands have been fitted with a Gaussian to obtain the energy transition. The results can be seen in Table 4.1. A clear feature that can be extracted from the values in Table 4.1 and Figure 4.7 is that the maxima of the *trans* neutral form in the microcrystals are red-shifted respect to the maxima in solution. Moreover, the microcrystals absorption bands are wider. It also stands out that while we could see a shift in between the absorption maxima WT and V151A *trans* neutral spectra (*Off* form) in solution, for the microcrystals the maximum is found in both cases at 414 nm. Another result is that in both environments (crystal and solution) the V151L variant has a red-shifted absorption maximum. These results make reasonable to extrapolate the X-ray structure of *Off* state in crystal phase (particularly the geometry of HBDI) to the protein in solution.

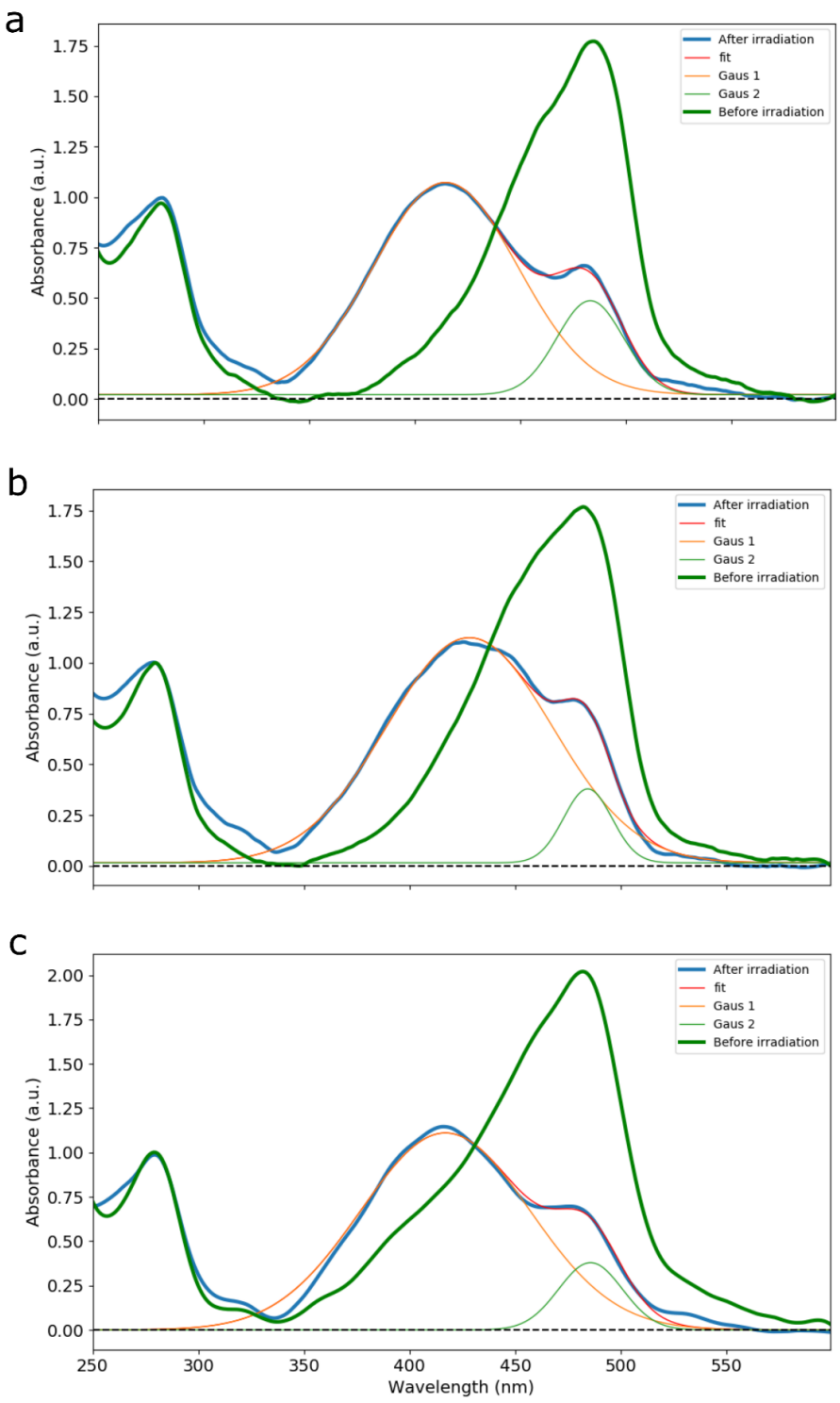


Figure 4.7. *UV-Vis absorption spectra for microcrystals "before" and "after" irradiation normalized at 280 nm (after subtraction of the baseline artifact and the scattering contribution) together with the estimation of On resting form (thin green gaussian) for WT (a), V151L (b) and V151A (c). Figure made with matplotlib python library.*

Table 4.1. *Off* state absorption maximum comparison between microcrystal and solution.

		solution	microcrystals
WT	**λ_{Max} / ε (M^{-1}, cm^{-1})**	408 / 26443	414
	Energy in cm^{-1}	4316	4762.6
V151L	**λ_{Max} / ε (M^{-1}, cm^{-1})**	416 / 34231	428
	Energy in cm^{-1}	4426.40	5234.6
V151A	**λ_{Max} / ε (M^{-1}, cm^{-1})**	405 / 24640	414
	Energy in cm^{-1}	4165.7	4864.0

4.2 Photoswitching quantum yields

Since the thermal back recovery values are needed to accurately determine the photoswitching quantum yields, these were previously determined. To perform these measurements, the protein concentration was adapted to have at 480 nm an absorbance of ~0.5 in 1cm at pH 8 (50 mMol Hepes, 50 mMol NaCl buffer). The proteins solutions were irradiated with a continuous LED light source 505 nm Thorlabs M505L3 (FWHM 37 nm) to switch the proteins to the photo stationary *Off* state. The absorbance at 482 nm and 408, 417 and 405 nm for WT, V151L and V151A respectively, was monitored for over 20 hours (140 h for V151A) (Cary 3500 absorption spectrometer, spectral bandwidth of 2 nm). One point per minute was collected, the exposition time per point was 0.1 s which ensures a minimum influence of the spectrometer light. During the whole experiment, the temperature was controlled with a Peltier cooling system and set to 23°C. The sample was continuously stirred to ensure homogenization. The recovery of the initial absorbance value before 505 irradiation prove the absence of photo-degradation. The experimental curves were globally fitted to a mono-exponential function yielding values of 1.81 h, 2.57 h and 40.3 h for WT, V151L and V151A respectively. The fitting was done using python 3.7 with the minimize function of lmFit package[15]. The results for the three proteins can be seen in Figure 4.8.

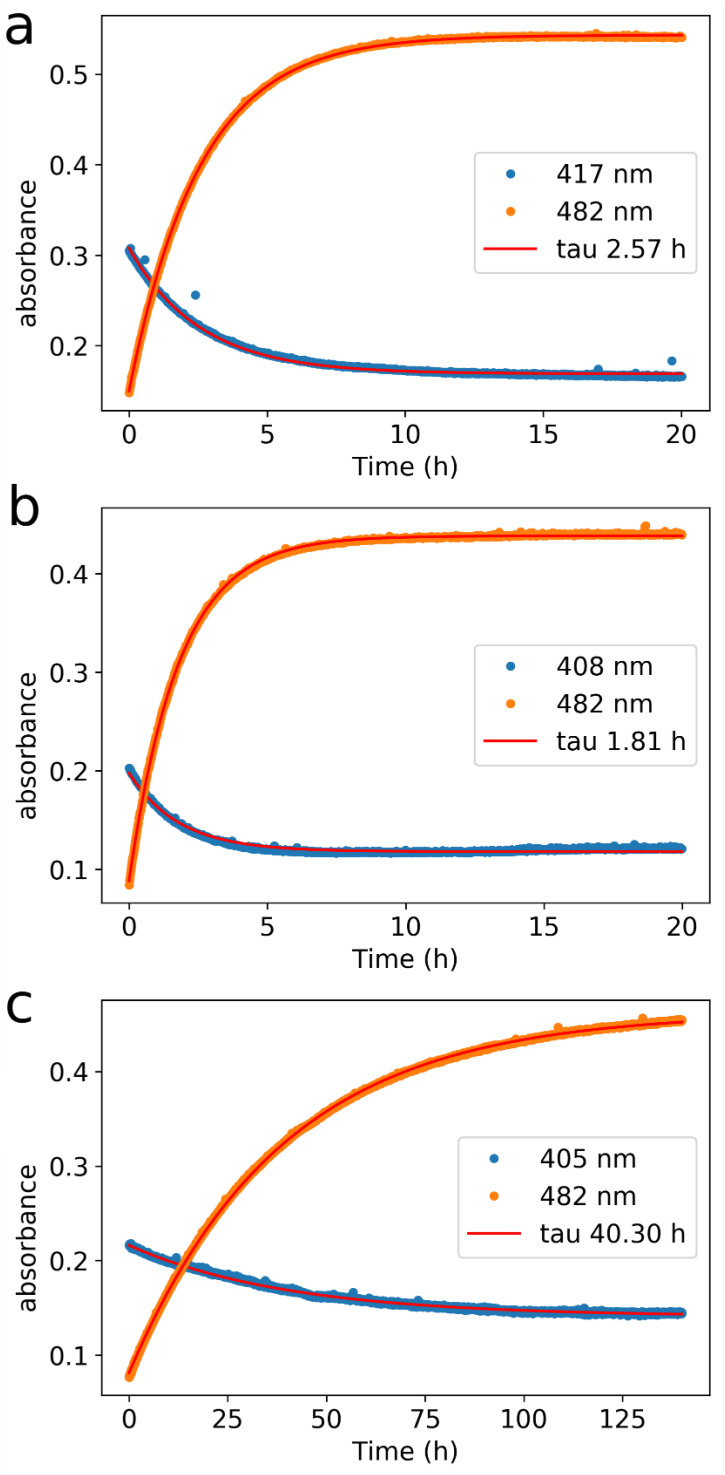


Figure 4.8. Evolution of absorbance starting from photo-stationary Off state after stopping the 505 nm light irradiation for (a) WT, (b) V151L, and (c) V151A. In blue the disappearance of the Off form, in orange the recovery of the On form and in red the fitted curves. Figure made with matplotlib python library.

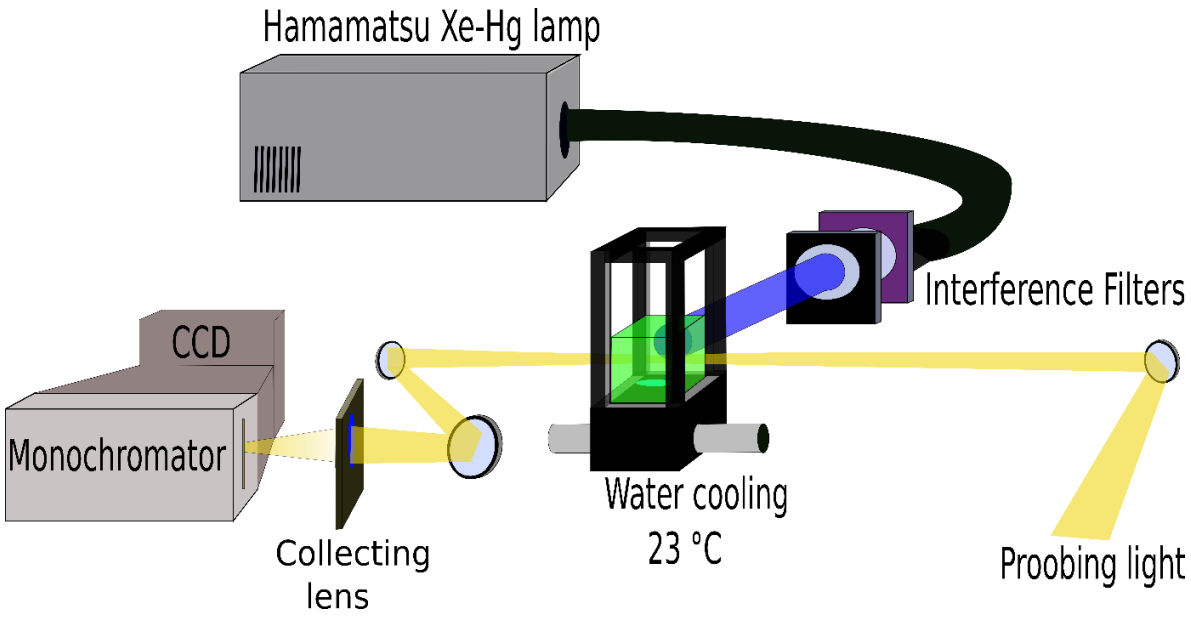


Figure 4.9. Schematic representation of the experimental setup used to monitor the changes in the absorption spectra under irradiation. Figure made with Inkscape.

Once the molar absorption coefficients and thermal back-reaction have been determined, it is possible to characterize the photoswitching quantum yields monitoring the light-induced UV-Vis absorption changes of the proteins in solution under continuous irradiation. For this purpose, between 6-9 µL of protein were diluted into 2 mL buffer (50 mMol Hepes 50 mMol NaCl pH 8) and transferred to a cuvette (1 cm x 1 cm) to have an absorption of ~0.5 over 1cm at 480 nm. The solutions were continuously irradiated with a Hamamatsu lamp (LC8 bulb Xe-Hg) at selected wavelengths using interference filters. For the *On* to *Off* form conversion, a filter Semrock FF01-485/20-25 and 2 metallic neutral filters Melles of optical densities of 1.0 and 0.5 were used. This yields an irradiation power of $4{,}19 \times 10^{-7}$ Einstein s^{-1} (Einstein unit is defined as energy in one mole of photons) centred at 485 nm with a 10 nm spectral width. For the *Off* to *On* form conversion, a filter Semrock FF01-406/15-25 and a metallic neutral filter Melles with an optical density of 3.0 which yields an irradiation power of $6{,}40 \times 10^{-8}$ Einstein s^{-1} centred at 405 nm with a 10 nm spectral width. The irradiation power was measured using a photodiode from Ophir (PD300-UV). During the irradiation, absorption spectra were continuously recorded using a white light beam (90° respect to the irradiation beam) and a spectrometer combined with a CCD (cooled with a Peltier -30 °C to

reduced thermal noise). The sample temperature was controlled with a water cooling system and set to 23°C. The sample was continuously stirred to ensure homogenization. The setup has been represented in Figure 4.9, and has been previously reported elsewhere[16].

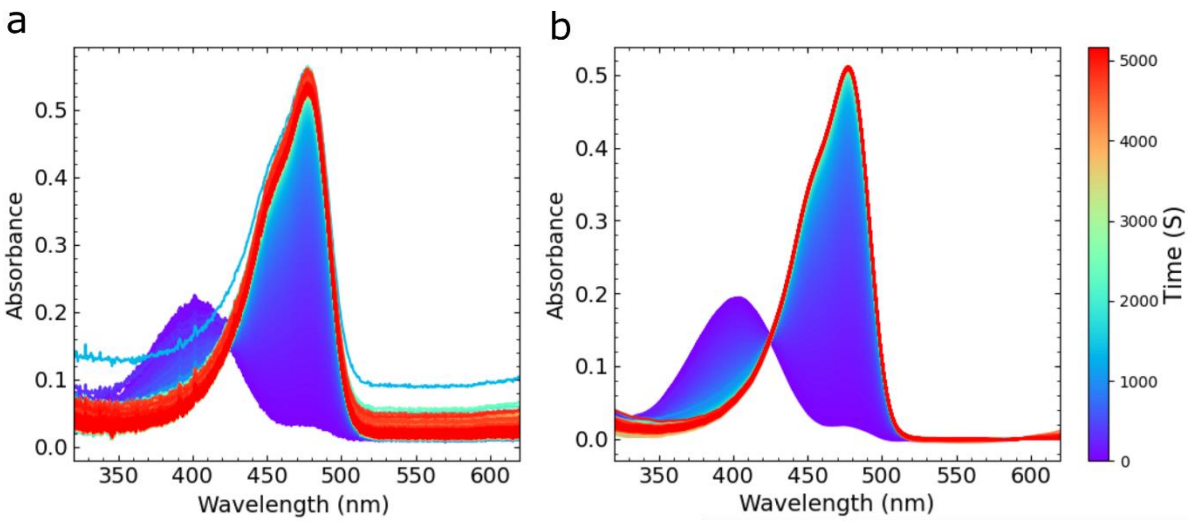


Figure 4.10. Example of data collected using the setup in Figure A1.18 for rsEGFP2 WT. a) raw data. b) data after pre-processing. The colour bar represents the recorded time in seconds for each of the spectra. Figure made with matplotlib python library.

The collection of spectra started 30 seconds before the irradiation lamp was turned on. The time between collected spectra was 0.68 s, this time includes the accumulation time and the readout of the camera. First, the *Off* to *On* photoswitching was measured, followed by the *On* to *Off* photoswitching. The measurements were performed sequentially, and the absence of degradation was checked. In general, for WT, V151A and V151L, between 45 and 55 minutes were enough to complete the *On* to *Off* and *Off* to *On* conversion. The lamp irradiation power was adapted to have enough time points to correctly fit the initial evolutions and reach the photostationary state in less than one hour. The raw data measurementfor WT can be seen in Figure 4.10a. These measurements are affected by some baseline drifts. Therefore, after a first correction step shifting the spectra to be equal to zero at 600 nm, if the difference in absorbance (at 480 nm) between two consecutive points was higher than 0.005, the spectra were filtered. Finally, after filtering, the spectra were smoothed using a Savitzky-Golay third-order polynomial with 25 points. The difference between the original spectra and pre-process ones can be seen in Figure 4.10a and b. In figure 4.11, absorption spectra

evolutions under continuous 405 nm irradiation (*Off* to *On switching)* can be seen for WT, V151L and V151A. Similar data were collected for *On to Off* photo-switching using 485 nm irradiation (Figure 4.12). As shown in these two figures, there is a clear isosbestic point around 426 nm for WT and 2 variants, which indicates a photo-transformation between two forms.

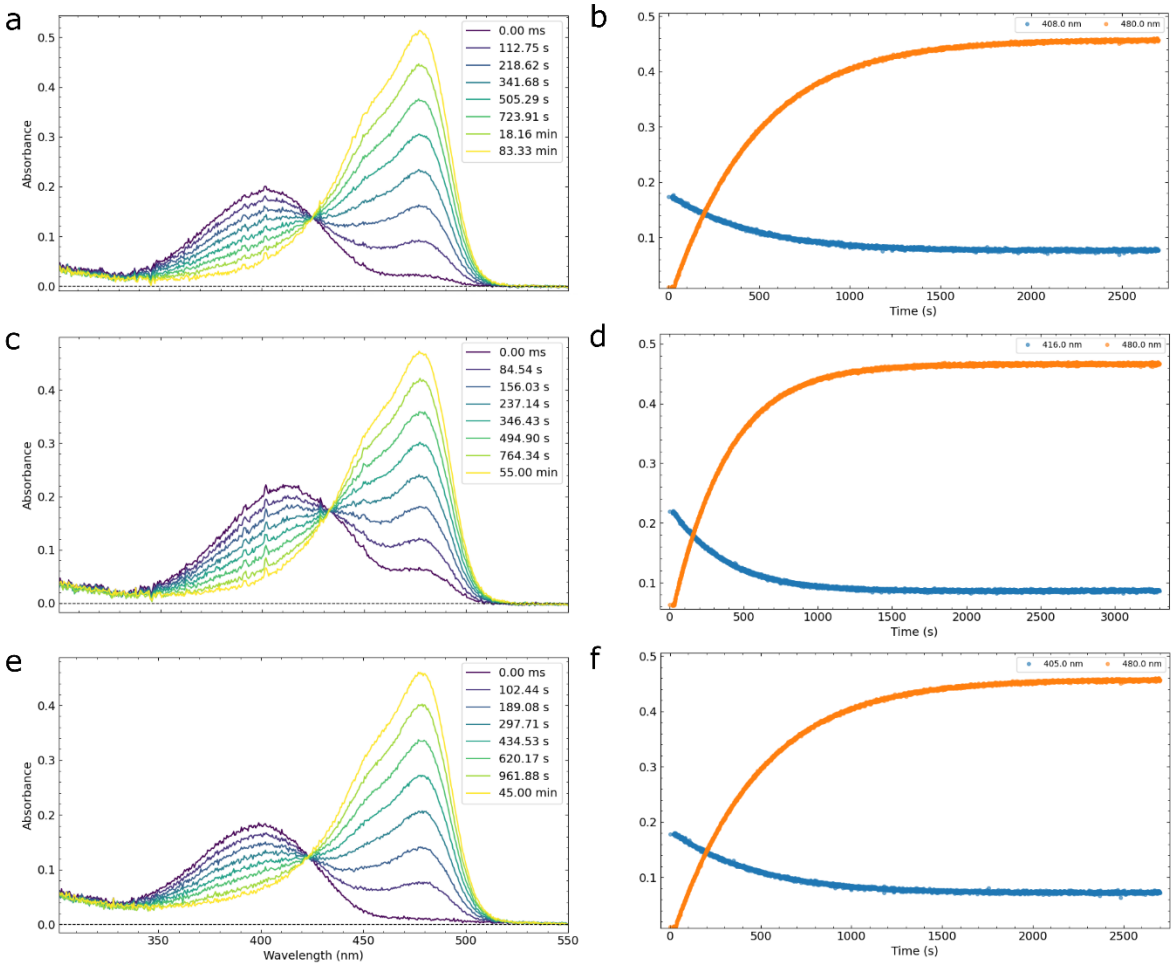


Figure 4.11. Time evolution of absorption spectra under continuous irradiation at 405 nm for: WT (a), V151L (b) and V151A (c). Time evolution of the absorbance at 480 nm (orange) and at the absorption maximum (blue) of Off form for WT (408 nm, d), V151L (416 nm, e) and V151A (405 nm, f). (data shown has been baseline corrected but not smooth). Figure made with UltraPyFit.

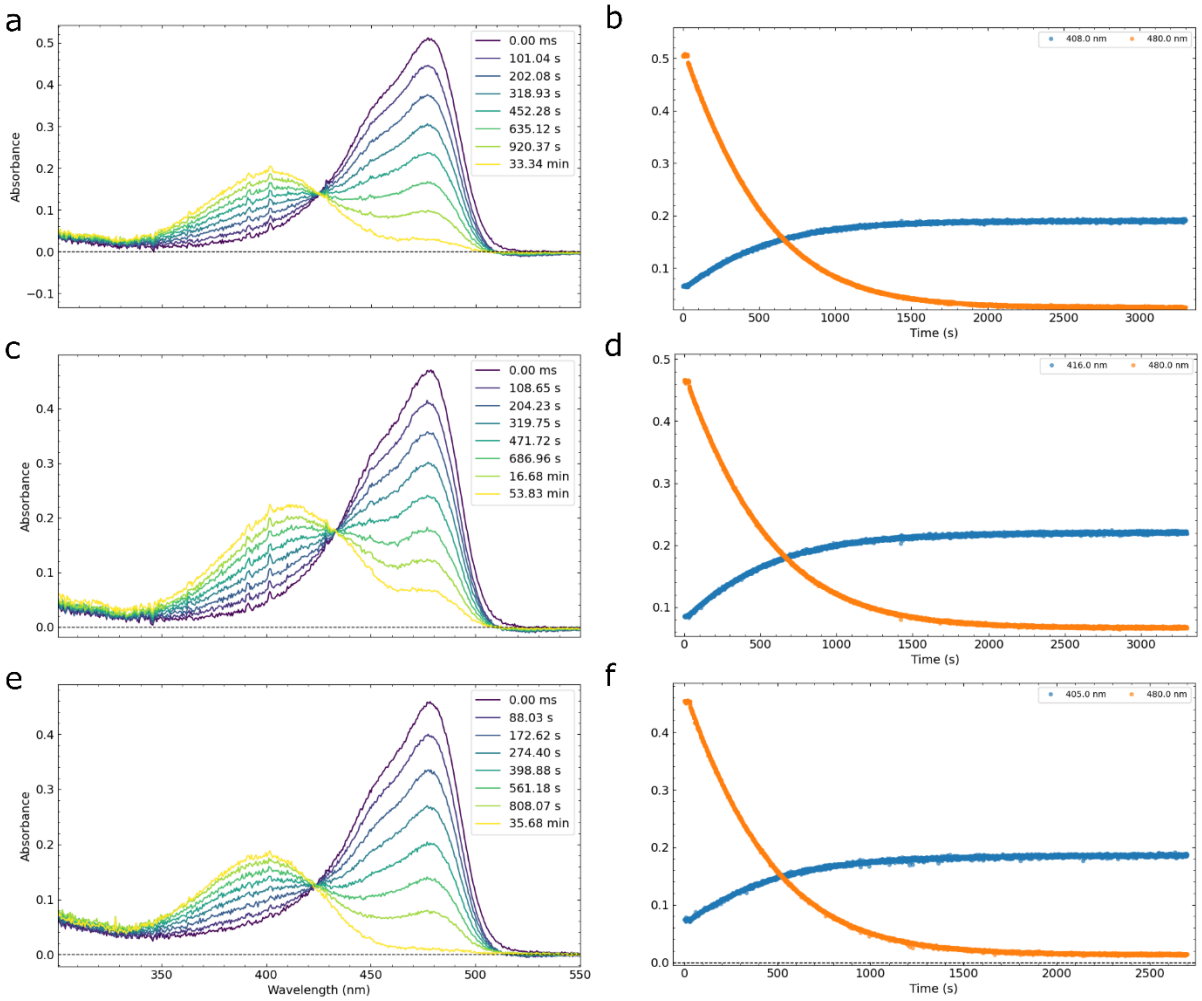


Figure 4.12 Time evolution of absorption spectra On to Off conversion under continuous irradiations at 485 nm. WT (a), V151L (b) and V151A (c). Time evolution of the absorbance at 480 nm (orange) and at the absorption maximum (blue) of Off form for WT (408 nm, d), V151L (416 nm, e) and V151A (405 nm, f). (data shown has been baseline corrected but not smooth). Figure made with UltraPyFit.

The formula to determine photoswitching quantum yields, $\Phi_{on\text{-}off}$ and $\Phi_{off\text{-}on}$, are given in the Equation 4.2:

$$\frac{dC_{off}}{dt} = -\frac{dC_{on}}{dt} = \phi_{on} I_{on}^{abs} - \phi_{off} I_{Coff}^{abs} - k_{off} C_{off} \qquad \text{Equation 4.2}$$

Where Φ_{On}, Φ_{Off} are the corresponding isomerization quantum yields. k_{off} is the thermal recovery rate constant from the *Off* to the *On* state. And finally, $I^{abs}_{On/Off}$ is the absorbed light at the irradiation wavelength of the different forms (equations 4.3 and 4.4).

$$I^{abs}_{On} = \frac{\varepsilon_{On} C_{On}}{\varepsilon_{Off} C_{Off} + \varepsilon_{On} C_{On}} I_0 [1 - 10^{-(\varepsilon_{Off} C_{Off} + \varepsilon_{On} C_{On})}] \quad \text{Equation 4.3}$$

$$I^{abs}_{Off} = \frac{\varepsilon_{Off} C_{Off}}{\varepsilon_{Off} C_{Off} + \varepsilon_{On} C_{On}} I_0 [1 - 10^{-(\varepsilon_{Off} C_{Off} + \varepsilon_{On} C_{On})}] \quad \text{Equation 4.4}$$

In equations 4.3 and 4.4, ε is the molar absorption coefficient of each species at the irradiation wavelength. I_0 represents the irradiation intensity of the Xe-Hg lamp that reaches the sample after the interference and neutral filters, and C the concentration of each species. As can be seen in equation 4.3 and 4.4 -which are linked to equation 4.2- to calculate the quantum yields, the profile concentrations of the *On* and *Off* forms are necessary. Therefore, these need to be calculated. Each spectrum was decomposed -using a least-square function coded in python 3.7- as a mixture of *Off* (trans neutral form) and *On* (cis anionic form) spectra multiplied by a certain weighting factor (Equation 4.5) and imposing a closure (mass conservation law). Thus, the profiles represent the relative molar fractions between *On* and *Off* forms (Figure 13b).

$$\text{Spectra}(t) = A(t) \times \text{Spectra\_On} + (1 - A(t)) \times \text{Spectra\_Off} \quad \text{Equation 4.5}$$

In equation 4.5, Spectra(t) is the recorded absorption spectrum at a particular time, Spectra_On is the *cis* anionic *On* form spectrum, and Spectra_Off is the *trans* neutral *Off* form spectrum. This last spectrum is obtained by subtracting to the final photo-stationary state the percentage of resting *On* form (see section 1). The absorption Spectrum in Figure 4.13a represents the two absorption spectra used to decompose each of the spectra in the series; the resulting profile concentrations can be seen in Figure 4.13b.

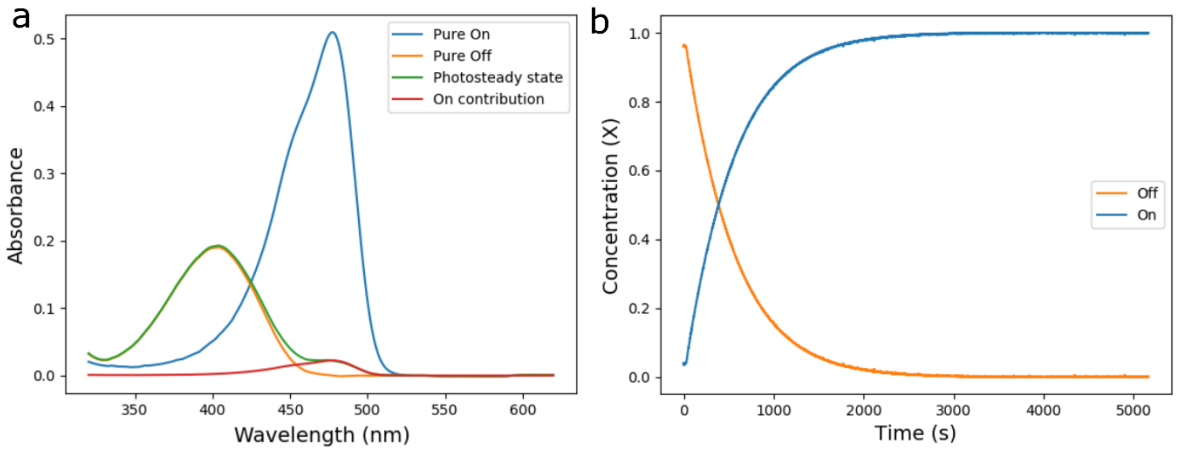


Figure 4.13. a) Absorption spectra used to decompose each of the spectra in figure 4.10 (WT). b) Profile obtained after the decomposition in molar fraction. For a) and b) orange curves represents "Off" form and blue one the "On" form. Figure made with matplotlib python library.

Knowing the molar absorption coefficient spectra of the *On* form, the total concentration of protein can be determined. With this value, the molar fractions profiles in Figure 4.12b were transformed into real concentrations profiles. Finally, to obtain the switching quantum yields, the same procedure explained above for the *Off* to *On* experiment was done for the *On* to *Off* experiment. The four concentration profiles (the two concentration profiles of the *On* to *Off* experiment plus the two profiles from the *On* to *Off* experiments) were globally fitted to the photo-kinetic equation 4.3 using Levenberg-Marquardt algorithm to minimize the χ^2 function, the results for the above-presented data can be seen in Figure 4.14.

The results for WT and two variants were $\Phi_{On\text{-}Off} = 0.0085$, $\Phi_{Off\text{-}On} = 0.11$ for WT, $\Phi_{On\text{-}Off} = 0.0075$, $\Phi_{Off\text{-}On} = 0.12$ for V151L and $\Phi_{On\text{-}Off} = 0.010$, $\Phi_{Off\text{-}On} = 0.14$ for V151A. The error value in this kind of experiments is estimated to be around 10%. The values for WT and V151A are different from those reported previously by Coquelle et al[3]. for *Off* to *On* switching which was 0.3 and 0.77 for WT and V151A, on the contrary, they agree with those reported by Duwé et al[17] for WT. It can be seen that V151A has both switching quantum yields slightly higher than the other two variants. On the contrary, the differences among WT, V151L and V151A are not as high as expected. However, logically, the less constrained chromophore (V151A) has the highest switching yield (*On* to *Off* and *Off* to *On*).

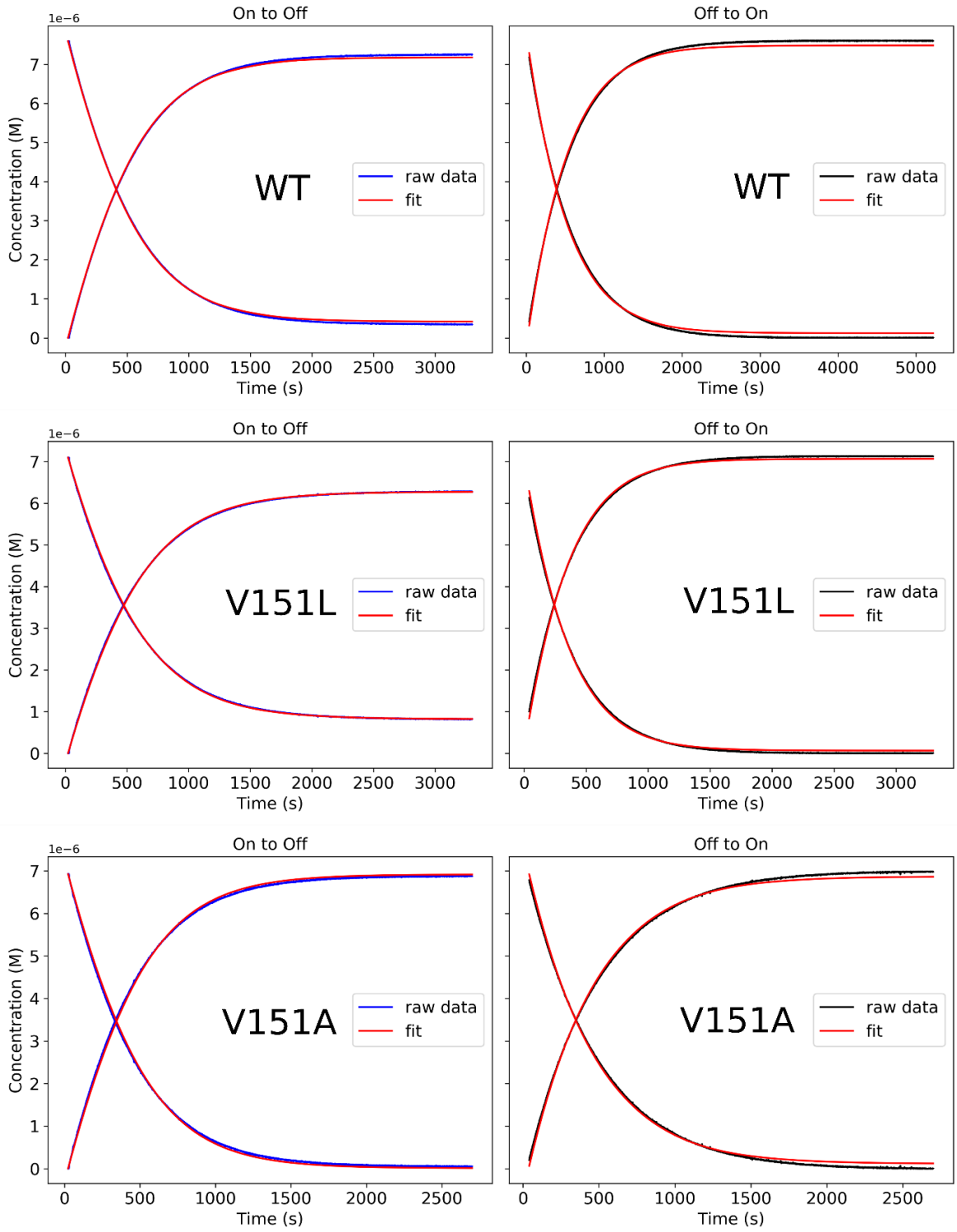


Figure 4.14. Results after globally fitting the profile concentrations of On to Off and Off to On experiments for rsEGFP2 variants. Top pannels for WT, middel pannels for V151L and bottom pannels for V151A. Figure made with matplotlib python library.

4.3 Fluorescence

Although some neutral forms have revealed to be fluorescent, mainly due to an ESPT as the avGFP[18–20], the *Off* neutral forms for WT and the two variants are not. Nevertheless, still, some signal can be detected since the completed photoisomerization of the proteins to the *trans* neutral *Off* form cannot be achieved. We first measured absorption and emission spectra (and excitation spectra for WT Figure 4.15b). For all the proteins, the emission spectra have a maximum at 502 nm, and a vibrational structure (shoulder) at 535 nm (with 470 nm excitation, red dashed vertical line Figure 4.15). Interestingly the absorption and excitation spectra for WT peak at different maxima revealing probably a ground state heterogeneity. The intensity ratio between the shoulder and the maxima is the same for WT, V151A and V151L. For the WT and the two variants, the emission maximum is red-shifted by about 20 nm respect to the absorption maximum (Stokes shift).

The fluorescence quantum yield has been measured following the methods described in reference [21], at two excitation wavelengths using two different references Cu-153 in EtOH with a fluorescent quantum yield of 0.53 and fluoresceine in 0,1 M NaOH with a fluorescence quantum yield of 0.89. The normalized fluorescence and absorbance spectra of these two compounds can be seen in Figure 4S-1 at the end of the chapter. To record *On* state fluorescence spectra, between 1-2 µL of protein were diluted into 3 mL buffer (50 mMol Hepes 50 mMol NaCl pH 8) and transferred to a cuvette (1 cm x 1 cm) to have at 480 nm an absorbance below 0.1 over 1cm. Then absorption spectra of the *On* and photo-stationary *Off* state were collected (Cary 1 absorption spectrometer spectral bandwidth of 2 nm and one point every 1 nm). The fluorescence spectra of the proteins and references were measured in a fluoromax-3 spectrofluorometer. The correction factors Sc/R were applied for the emission spectra. Sc is the fluorescent signal after application of the correction file "c" that takes in account the optical efficiency of the photomultiplier tube at each wavelength, and R is the correction for possible lamp intensity fluctuations. The quantum yields have been calculated using equation 4.6.

$$\Phi_S = \Phi_r \left(\frac{1-10^{\left(-Abs_{ref}\right)} \cdot Fluo_{sample} \cdot {n_{sample}}^2}{1-10^{\left(-Abs_{sample}\right)} \cdot Fluo_{ref} \cdot {n_{ref}}^2} \right) \qquad \text{Equation 4.6}$$

Φ_s and Φ_r are the fluorescence quantum yields of the sample and the reference, respectively. Abs_{ref} and Abs_{sample} the absorbance of proteins and the reference solutions at the excitation wavelength used. "n" is the different diffraction index for the solvents used, and $Fluo_{sample}$ and $Fluo_{ref}$ correspond to the number of emitted photons which was calculated integrating the fluorescent absorbance signal for samples and reference.

Table 4.2 Quantum yield results for rsEGFP2 WT, V151L and V151A

Sample	C153, exc. 450 nm	Fluoresceine, exc. 470 nm	Integrating sphere 470 nm	Average
WT	0.31	0.30	0.30	0.305
V151L	0.27	0.25	0.26	0.260
V151A	0.27	0.26	0.24	0.256

As shown in Table 4.2, the results between the different techniques are similar. For all cases the Φ_f of WT is higher than those of V151L and V151A, which are quite similar. The average values are 0.305, 0.260 and 0.256 for WT, V151L and V151A, respectively. With the molar absorption coefficient values and the Φ_f it is possible to calculate the brightness for each protein. The corresponding values for each rsEGFP2 variant can be seen in Table 4.3. Finally, the switching contrast calculated as the ratio between the areas under the fluorescence spectra before and after irradiation for the different variants are 20, 50 and 7.1 for WT, V151A and V151L respectively. These values are similar to the one determined using in absorbance and resting *On* state in the photo-stationary state: 16.7, 25 and 6.7 for WT V151A and V151L respectively. This means that the *Off* form (trans neutral form) is nearly non-fluorescent.

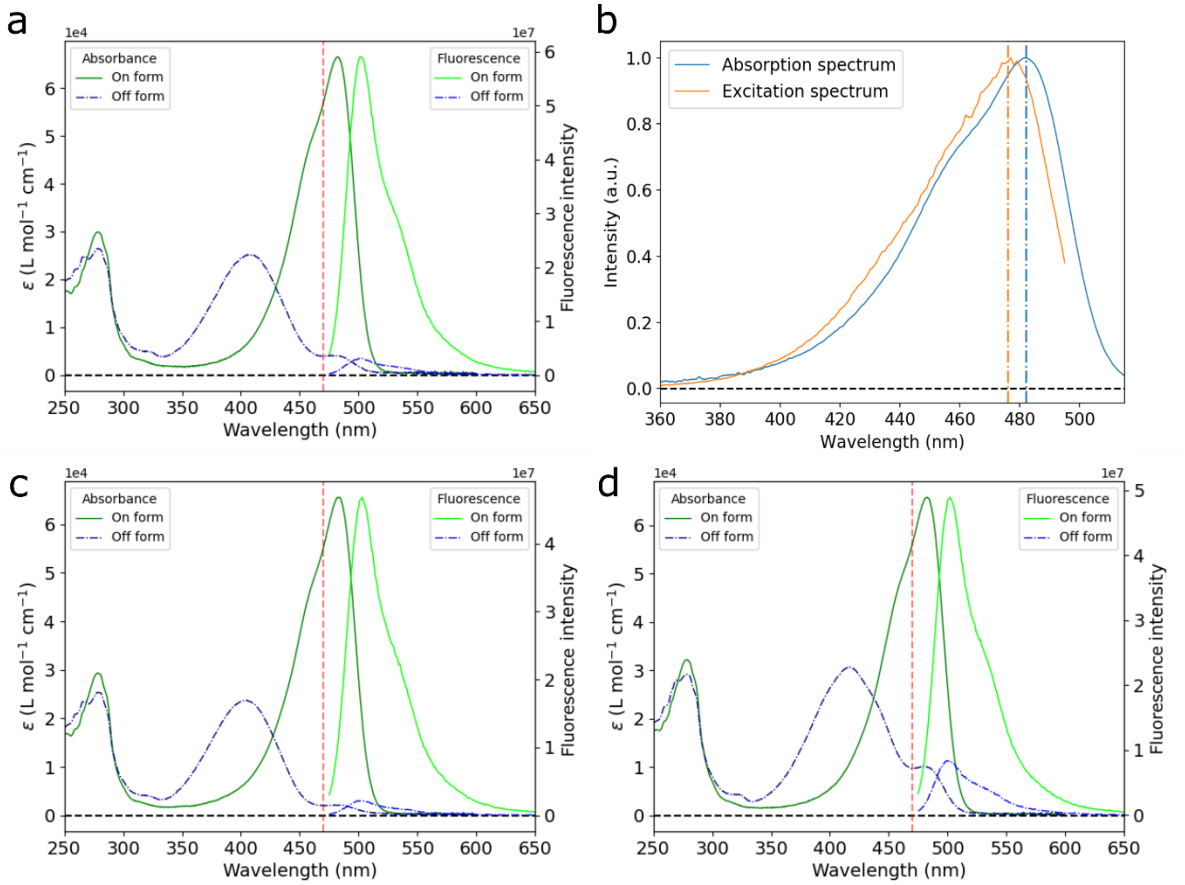


Figure 4.15. Fluorescence spectra of On form (light green solid line) and at the photo-stationary Off state (light blue dashed line) compared to the molar absorption coefficients of On (dark green solid line) and photo-stationary (dark blue dashed line) spectra. The red dashed vertical line at 470 nm is excitation wavelength. a) WT, c) V151A and d) V151L. b) correspond to a comparison between the absorption (maximum at 482 nm blue vertical line) and excitation spectra for WT (measured at 502 nm emission; maximum at 476 nm orange vertical line). Notice the emission and excitation spectra are normalized to their corresponding maxima. Figure made with matplotlib python library.

4.4 Discussion and conclusions

The main photophysical properties determined for the three protein samples are summarized in Table 4.3 In the resting state the chromophore adopts a *cis* anionic conformation (fluorescent *On state*). The *cis* anionic form of the three proteins displays the same absorption maximum at 482-483 nm and have similar molar absorption coefficient ~66000 M^{-1} cm^{-1}. The only difference observed is in the UV region at 280 nm (Figure 4.5). Upon irradiation of this *cis* anionic form with 490 nm visible light, the chromophore photo-converts into a *trans* neutral conformer (*Off* states). Contrary to *cis* anionic states, the *trans* neutral forms absorption spectra differ from each other, attributed to the different *Off* structures. Indeed, if a close comparison of the different chromophore in the *Off* states is made, it can be seen that they display different τ and Φ dihedral angles between the phenol and the imidazolinone moieties[6]. Furthermore, the protein environment of the chromophore and in particular, the H-bonding network are different. The His149 in the V151L variant accompanied the conformational change of the chromophore to which it remained hydrogen-bonded as commented just above. On the contrary, for WT and V151A, a new hydrogen bond is formed with a water structural molecule.

Furthermore, the photostationary spectra for the three proteins are composed of different percentage of *cis* anionic and *trans* neutral forms. Two factors influence these values, the thermal back recovery, which is 40.3 h, 1.81 h and 2.57 h for V151A, WT and V151L respectively and the absorption of the trans-neutral form at 480 nm. This lasts parameter is explained by the induced photochemical reaction to the original *cis* anionic, when the photoconverted *trans* neutral form is irradiated at 480 nm. A Higher absorbance at 480 nm for the *trans* neutral form will increase the *cis*-anionic contribution to the photostationary state. The V151A has both the longest thermal back recovery and the lowest *trans* neutral absorption at 480 nm which translate to a nearly 100% of conversion from *cis* anionic to *trans* neutral form. While WT and V151L variants have similar thermal back recoveries, the V151L *trans* neutral form absorption maximum is red-shifted respect WT and V151A which translates in a higher absorption coefficient at 480 nm. All together we can conclude that at the photostationary state we converted 95%, 86% and 98% of *On* form for WT, V151L and V151A (values for fluorescence). As the trans neutral state is considered to be non-

fluorescent, the switching contrast will be (Ion/Ioff) 20, 7.1 and 50, respectively. Therefore, V151L cannot be used in RESOLFT microscopy, as minimum values of 10 are needed[22,23].

Table 4.3 Photophysical properties of rsEGFP2 WT, V151L and V15A

Property	**WT**	**V151A**	**V151L**
Absorption maximum *On* (nm)	482	483	482
Absorption maximum *Off* (nm)	408	405	416
Molar absorption coefficient for On form at its maximum, pH 8 (M^{-1}, cm^{-1})	67210	65596	66149
Molar absorption coefficient for On form at 488 nm, pH 8 (M^{-1} cm^{-1})	61609	61926	61326
Molar absorption coefficient for Off form at its maximum, pH 8 (M^{-1} cm^{-1})	26443	24640	34231
Absorption maximum for Cis neutral (nm), pH 4.5	397	395	401
Molar absorption coefficient for Cis neutral form at its maximum, pH 4.5 (M^{-1} cm^{-1})	29221	32753	26231
Switching contrast at photostationary (fluorescence)	20	50	7.1
Switching contrast at photostationary (absorbance)	16.7	25.0	6.7
Fluorescence emission maximum	502	502	502
Fluorescence Quantum yield	0.34	0.27	0.28
On to *Off* switching quantum yield	0.0085	0.010	0.0078
Off to *On* switching quantum yield	0.11	0.14	0.12
Brightness at 488 nm	22851	17710	18521
Thermal recovery (h)	1.81	40.3	2.57
pK_a	5.7	6.3	5.8

Previous studies have shown that the rsEGFP2 WT upon irradiation with 490 nm visible light isomerized into a trans-neutral conformer adopting either HT type *Off* state as V151L or an OBF type as the V151A[8]. Interestingly our result shows that the WT *Off* state absorption maximum (408 nm) is in between those of the V151A (405 nm) and the V151L (416 nm). Therefore, the rsEGFP2 *Off*-state heterogeneity observed in crystals can also exist in solution. One could raise the objection that the differences between WT and V151A absorption spectra might not arise from the WT variant having an *Off* ground-state heterogeneity, but instead would originate from steric or electrostatic properties of the mutated amino acid at position 151. However, it is well known that HT and OBF conformers can rapidly exchange between them via a single bond rotation[24], as explained in Chapter 2. Indeed, the molar absorption coefficient spectrum of the *trans* neutral form for the WT variant can be perfectly reproduced by a weighted sum of the V151A (77.5%) and V151L (22.5%) spectra as can be seen in Figure 4.16. Notice that not only the main band at 408 nm but also the 280 nm maximum can be reproduced by the same sum of V151L and V151A spectra.

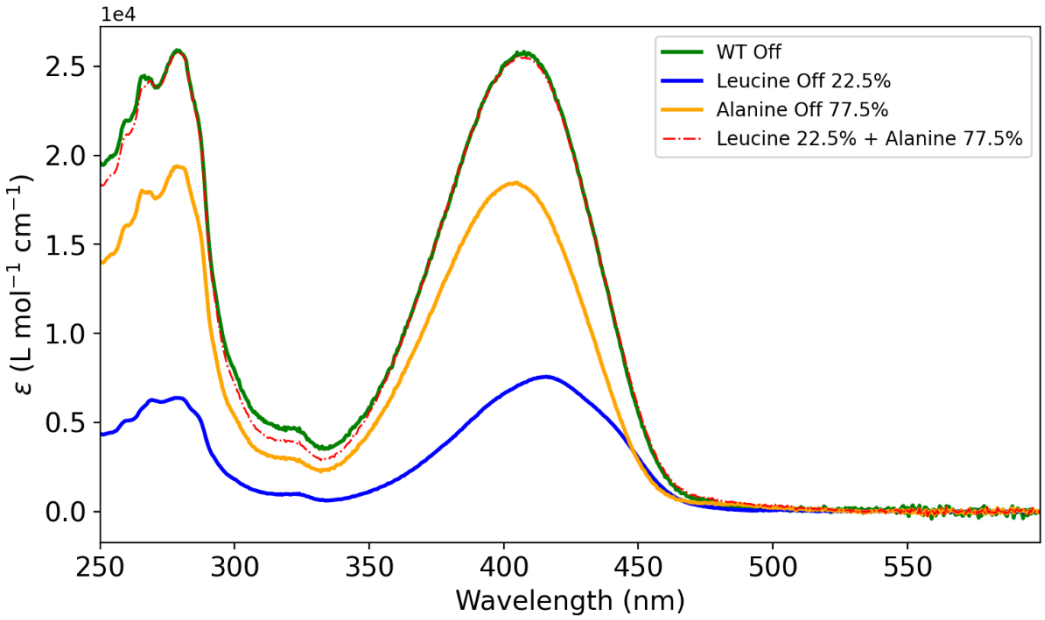


*Figure 4.16. rsEGP2 WT off trans-neutral UV-Vis absorption spectrum (green). 22.5 % of the V151L (HT Off conformer) in blue, 77.5 % V151A (OBF Off conformer) in orange, the sum of 0.225*V151L + 0.775*V151A gives the red dashed-dotted line which perfectly overlaps the WT trans-neutral Off molar absorption coefficient spectrum. Figure made with matplotlib python library.*

It is possible to assume that the observed differences in the absorbance of V151A and V151L variants in the solution may originate from steric or electrostatic properties of the amino acid at position 151. On the contrary, neither an alanine nor a leucine at position 151 are in direct interaction with the chromophore in neither of the *On* or *Off* states of the proteins and have no option to form hydrogen bonds with it. Therefore, we attribute **t**he red- and blue-shift of the absorption maxima in the V151L and V151A variants, to two possible reasons: i. the differences in τ and Φ dihedral angles of the chromophore in HT (V151L) and OBF (V151A) *Off* conformers. ii. differences in the local environment, that tunes the energy gap between the S_0 and S_1 electronic states, of the chromophore in V151L and V151A *Off* form. These differences are due to the different environments that the chromophore adopts in the two *trans* forms (HT vs OBF) rather than an electrostatic interaction with the amino acids at position 151. Shifts in the absorption band maximum position (both hypsochromic and bathochromic) accompanying mutation-induced changes in chromophore τ and Φ angles have previously been noted for green FPs with a *cis* chromophore[25] and for red FPs with a *trans* chromophore[26]. The bathochromic shift of the V151L spectrum appears to be consistent with a more extended electron delocalization in the near-planar chromophore of *HT* (dihedral angles φ = 9.5°, τ= -17.5°) as compared to in *OBF* (φ = -50.2°, τ= 9.2°) (unpublished results obtained from V151L and V151A *Off* state X-ray structures by Martin Weik and coworkers). Further studies based on high-level quantum chemistry calculations could provide a more detailed mechanistic view in the future.

From the values in Table 4.3, it stands out the relatively long thermal recovery, for V151A of over 40 h in comparison with those of WT and V151L which are 1.81 and 2.57 h respectively. In the same line is the pK_a value of the *On*-form which was determined to be 6.3 for V151A, 5.7 for WT and 5.8 for V151L**,** suggesting a possible relationship between these two properties. In the thermal back reaction, the trans-protonated state transforms into a *cis* anionic one. It is well known that the deprotonation in GFP-proteins occurs in the microsecond-millisecond time range constant, which is under the detection limit of our experiment, and thus instantaneous. Therefore mono-exponential thermal back evolution indicates that the *cis*-*trans* isomerization is the limiting reaction step. This further suggests that the isomerization provokes a change in the pK_a of the chromophore probably induced by

a different interaction of the surrounding amino acids, which then triggers the deprotonation reaction occurring nearly instantaneously.

The photoswitching quantum yields values, reported in Table 4.3 are in contrast to those previously reported by Coquelle et al.[3], the values are much smaller, especially regarding the *Off* to *On* photoswitching quantum yield. The values reported here were obtained following the protein absorbance changes in solution under continuous irradiation of the *On* and *Off* forms, respectively. Our reported values are in agreement with those reported for WT by Duwé et al. [17] using a similar approach. On the contrary, those reported by Coquelle et al. [3] were determined following the changes in fluorescence under a microscope for immobilized proteins in a film, which could explain the differences observed. Finally, logically a less constrained chromophore (V151A) has the highest switching yields (*On* to *Off* and *Off* to *On*)

4.5 Supporting material

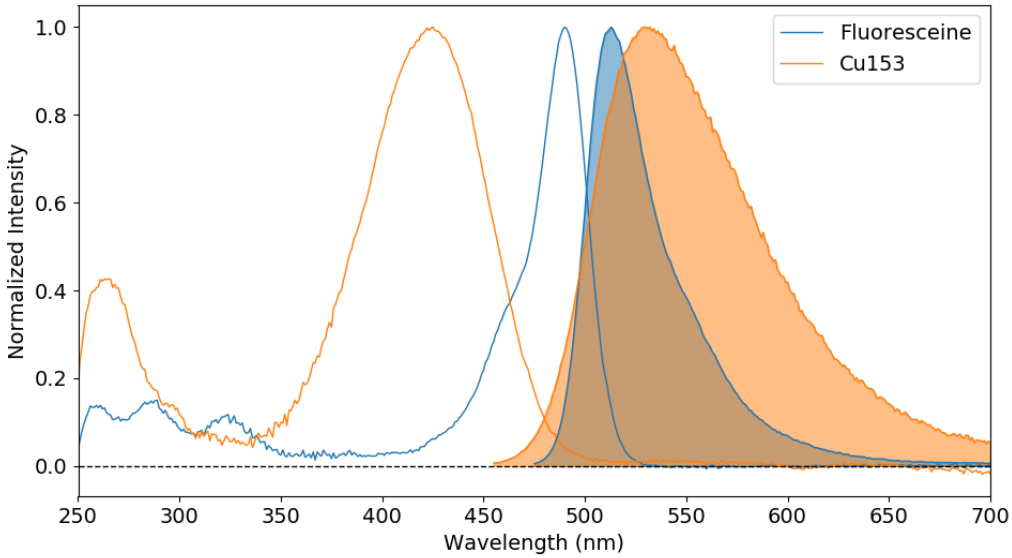


Figure 4S-1. Normalized absorbance and fluorescence spectra for fluoresceine 0.1M NaOH (blue) and Cu153 in EtOH (orange). Unfilled curves correspond to absorbance spectra and fill curves to fluorescence spectra. Figure made with matplotlib python library.

Table 4S-1. Molar absorption coefficient measurements.

	measurement	First	Second	Average
WT	Absorption *On* form max (nm)	482	482	482
	Epsilon *On* (L mol^{-1} cm^{-1})	67857.67	66561.63	67209.65
	Absorption *Off* state max (nm)	408	408	408
	Epsilon *Off* (L mol^{-1} cm^{-1})	26681.73	26203.4	26442.565
V151A	Absorption *On* form max (nm)	482.5	483	482.75
	Epsilon *On* (L mol^{-1} cm^{-1})	65552.02	65639.31	65595.665
	Absorption *Off* state max (nm)	405	405	405
	Epsilon *Off* at max (L mol^{-1} cm^{-1})	24624.65	24151.0	24387.825
V151L	Absorption *On* form max (nm)	482	482	482
	Epsilon *On* (L mol^{-1} cm^{-1})	66481.31	65817.59	66149.45
	Absorption *Off* form (nm)	416	416	416
	Epsilon *Off* (L mol^{-1} cm^{-1})	34051.88	34409.56	34230.72

4.6 References

(1) Grotjohann, T.; Testa, I.; Reuss, M.; Brakemann, T.; Eggeling, C.; Hell, S. W.; Jakobs, S. rsEGFP2 enables fast RESOLFT nanoscopy of living cells. *eLife* **2012**, *1*, e00248.

(2) 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. *Scientific reports* **2016**, *6*, 18459.

(3) Coquelle, N.; Sliwa, M.; Woodhouse, J.; Schirò, G.; Adam, V.; Aquila, A.; Barends, T. R. M.; Boutet, S.; Byrdin, M.; Carbajo, S.*; et al.* Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nature chemistry* **2018**, *10*, 31–37.

(4) 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.*; et al.* Infrared spectroscopy reveals multi-step multi-

timescale photoactivation in the photoconvertible protein archetype dronpa. *Nature chemistry* **2018**, *10*, 845–852.
(5) 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. *Nature communications* **2013**, *4*, 1461.
(6) Woodhouse, J.; Nass Kovacs, G.; Coquelle, N.; Uriarte, L. M.; Adam, V.; Barends, T. R. M.; Byrdin, M.; La Mora, E. de; Bruce Doak, R.; Feliks, M.*; et al.* Photoswitching mechanism of a fluorescent protein revealed by time-resolved crystallography and transient absorption spectroscopy. *Nature communications* **2020**, *11*, 741.
(7) 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. *The journal of physical chemistry. B* **2015**, *119*, 2404–2414.
(8) Joyce Woodhouse. *Studying a reversibly switchable fluorescent protein by time-resolved crystallography using the X-ray free electron lasers. Biological Physics.*; Université Grenoble Alpes, 2018.
(9) Ward, W. W. PROPERTIES OF THE COELENTERATE GREEN-FLUORESCENT PROTEINS. *Bioluminescence and Chemiluminescence*; Elsevier, 1981; pp 235–242.
(10) DeLuca, M. A.; McElroy, W. D. *Bioluminescence and Chemiluminescence: Basic Chemistry and Analytical Applications*; Elsevier Science: Saint Louis, 2014.
(11) Gromov, E. V.; Domratcheva, T. Four resonance structures elucidate double-bond isomerisation of a biological chromophore. *Physical chemistry chemical physics : PCCP* **2020**, *22*, 8535–8544.
(12) Salna, B.; Benabbas, A.; Sage, J. T.; van Thor, J.; Champion, P. M. Wide-dynamic-range kinetic investigations of deep proton tunnelling in proteins. *Nature chemistry* **2016**, *8*, 874–880.
(13) Schirò, G.; Woodhouse, J.; Weik, M.; Schlichting, I.; Shoeman, R. L. Simple and efficient system for photoconverting light-sensitive proteins in serial crystallography experiments. *J Appl Crystallogr* **2017**, *50*, 932–939.
(14) Savitzky, A.; Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627–1639.
(15) Matt Newville; Renee Otten; Andrew Nelson; Antonino Ingargiola; Till Stensitzki; Dan Allan; Austin Fox; Faustin Carter; Michał; Dima Pustakhod*; et al. lmfit/lmfit-py 0.9.14*; Zenodo, 2019.
(16) Métivier, R.; Badré, S.; Méallet-Renault, R.; Yu, P.; Pansu, R. B.; Nakatani, K. Fluorescence Photoswitching in Polymer Matrix: Mutual Influence between Photochromic and Fluorescent Molecules by Energy Transfer Processes. *J. Phys. Chem. C* **2009**, *113*, 11916–11926.
(17) Duwé, S.; Zitter, E. de; Gielen, V.; Moeyaert, B.; Vandenberg, W.; Grotjohann, T.; Clays, K.; Jakobs, S.; van Meervelt, L.; Dedecker, P. Expression-Enhanced Fluorescent Proteins Based on Enhanced Green Fluorescent Protein for Super-resolution Microscopy. *ACS nano* **2015**, *9*, 9528–9541.
(18) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 8362–8367.

(19) Fang, C.; Frontiera, R. R.; Tran, R.; Mathies, R. A. Mapping GFP structure evolution during proton transfer with femtosecond Raman spectroscopy. *Nature* **2009**, *462*, 200–204.
(20) Tang, L.; Wang, Y.; Zhu, L.; Kallio, K.; Remington, S. J.; Fang, C. Photoinduced proton transfer inside an engineered green fluorescent protein: a stepwise-concerted-hybrid reaction. *Physical Chemistry Chemical Physics.* **2018**, *20*, 12517–12526.
(21) Würth, C.; Grabolle, M.; Pauli, J.; Spieles, M.; Resch-Genger, U. Relative and absolute determination of fluorescence quantum yields of transparent samples. *Nature protocols* **2013**, *8*, 1535–1550.
(22) Jensen, N. A.; Jansen, I.; Kamper, M.; Jakobs, S. Reversibly Switchable Fluorescent Proteins for RESOLFT Nanoscopy. In *Nanoscale Photonic Imaging*; Salditt, T., Egner, A., Luke, D. R., Eds.; Topics in Applied Physics; Springer International Publishing: Cham, 2020; pp 241–261.
(23) Konen, T.; Grotjohann, T.; Jansen, I.; Jensen, N.; Hell, S. W.; Jakobs, S. *The Positive Switching RSFP Padron2 Enables Live-Cell RESOLFT Nanoscopy Without Sequential Irradiation Steps*, 2020.
(24) Gerwien, A.; Schildhauer, M.; Thumser, S.; Mayer, P.; Dube, H. Direct evidence for hula twist and single-bond rotation photoproducts. *Nature communications* **2018**, *9*, 2510.
(25) Shu, X.; Kallio, K.; Shi, X.; Abbyad, P.; Kanchanawong, P.; Childs, W.; Boxer, S. G.; Remington, S. J. Ultrafast excited-state dynamics in the green fluorescent protein variant S65T/H148D. 1. Mutagenesis and structural studies. *Biochemistry* **2007**, *46*, 12005–12013.
(26) Wilmann, P. G.; Petersen, J.; Pettikiriarachchi, A.; Buckle, A. M.; Smith, S. C.; Olsen, S.; Perugini, M. A.; Devenish, R. J.; Prescott, M.; Rossjohn, J. The 2.1A crystal structure of the far-red fluorescent protein HcRed: inherent conformational flexibility of the chromophore. *Journal of molecular biology* **2005**, *349*, 223–237.

5 *Off* state photodynamics of rsEGFP$_2$ WT, V151A and V151L

5.1 Introduction

After determining the photophysical properties for WT, V151A and V151L in solution (including the switching yield), we study in this chapter the *Off* to *On* photodynamics of these three proteins. As discussed in the introduction, after photoisomerization, the rsEGFP2 chromophore can have different *Off* state photoproduct isomers. These isomers are compatible either with an OBF or an HT isomerizations[1]. The OBF pathway is a large volume demanding mechanism and V151A fulfil this requirement. On the contrary, HT pathway is a volume-conserving pathway favoured in sterically constraint environments which is the case for V151L. An important result obtained in Chapter 4 is that WT *Off* state absorption spectrum can be interpreted as the sum of V151A (OBF) and V151L (HT) *Off* form spectra. Preliminary TR-SFX results obtained for WT microcrystals by our collaborators confirm that in certain conditions, it is possible to observe both *Off* photoproducts (Figure 3.18).

At the end of Chapter 3, it has been mentioned that the excited-state dynamics (from fs to 20 ps) of rsEGFP2 was published by Coquelle et al.[2]. The use of TR-SFX, combined with TRUV-Vis, demonstrated the existence of a twisted chromophore configuration at 1 ps, halfway between the *trans* and *cis* isomers attributed to an isomerization intermediate. The τ and Φ angles in the twisted intermediate are in agreement with a HT isomerization mechanism[1,2]. These results were further supported by QM/MM simulations which demonstrated that upon excitation, the protonated HBDI chromophore within the protein preferentially rotates around Φ to facilitate the isomerization as previously suggested for HBDI in solution[3]. Moreover, S. R. Meech and coworkers have recently demonstrated using TRIR that the *Off* to *On* isomerization in Dronpa and Dronpa 2 is significantly more complex than a two-step process involving a one-step *cis-trans* isomerization and a proton transfer[4]. They show the existence of several chromophore and protein motions involved all along the isomerization process which ends in 91 ns in the ground state (chapter 3). They used an innovative multiplex time-resolved infrared multiple probe spectroscopy (TRMPS)[4], that

allowed them to study in a single experiment the photodynamics from the femtosecond (fs) to the millisecond (ms) time range.

The *Off* to *On* photodynamics of V151L (HT *On* to *Off* photoproduct) and V151A (OBF *On* to *Off* photoproduct) from femtosecond to millisecond time scale are essential to understand the photodynamics of rsEGFP2 (WT) and will help to understand the general photoisomerization mechanism of RSFPs. The results in this chapter will be divided into two sections according to the different time ranges available in TRUV-Vis spectroscopy (explained in chapter 2): 0-2 ns (stroboscopic) and 2 ns-10 ms (oscilloscope). Each section will be sub-divided with the results of TRUV-Vis and TRIR probes used. Some figures were added at the end of the chapter to support our analysis (Supporting figures). The experimental details can be found in Chapter 2 (description of TR spectroscopies) and Appendix 1 (materials and methods). The structure at 10 ns obtained by TR-SFX (from Martin Weik's group in collaboration with Ilme Schlichting's group -see authors in section 5.3.3-) at 10 ns for WT, V151A and V151L are essential to link the optical TR spectroscopy results to the protein transient structures and build a final photo-dynamical scheme. Moreover, since the dynamic between solution and crystal can differ, we have also studied microcrystals colloidal solution using ultrafast optical spectroscopies to validate our methodology. Part of the results presented in this chapter are already published in Woodhouse et al.[5].

5.2 Dynamics between 0 and 2 ns

5.2.1 Transient absorption UV-Vis Spectroscopy

5.2.1.1 Excited-state dynamics, dynamics between 0-40 ps in H_2O buffer

Time-resolved difference absorption spectra between 0 and 40 ps for WT, V151A and V151L in solution (experimental condition are given in Appendix 1) can be seen in Figure 5.1 together with the main kinetic traces in Figure 5.2. Transient absorption spectra were recorded after excitation of the *Off* state with a 400 nm laser pulse (2.1 µJ, 1kHz, 500 µm $1/e^2$, 2.14 mJ/cm^2). Four different growing bands between -200 fs and 250 fs can be

identified for WT: two positive bands with maxima at 335 and 455nm, which are attributed to excited-state species absorption (ESA), and two negative bands. The first narrow negative band centered between 390 and 410 nm is attributed to the ground state bleaching (GSB, depopulation of the *trans* neutral form), and the second broadband negative band ranging from 490 to 700 nm is attributed to the stimulated emission (SE) of excited-state species. The SE band is characterized by a broad band (490-600 nm) with a minimum at 515 nm and a red-shifted tail (600-720 nm). Note that this latter tail keeps growing during 100 femtoseconds after the others. A small evolution of the two ESA bands is observed from 250 to 400 fs (see Figure 5.1d). Then in ~1 ps (from 400 fs to 1.7 ps), the ESA bands centered at 335 nm and 455 nm together with the GSB band and SE band centered at 515 nm decrease to around half of their intensities. Within the same time, the red tail part of the SE band completely decays. This evolution is followed by the decay of all transient absorption bands in ~10 ps to form at 40 ps a transient spectrum mainly characterized by a positive band at 375 nm and a small negative band between 465 and 580 nm with a maximum at 502 nm.

Similar experiments have been carried out for V151L and V151A variants. After excitation with a 400 nm laser pulse, the transient spectra which are growing until 200 fs exhibit similar features to WT ones. In both cases, after signal formation, at 200 fs, the transient spectrum shows mainly the same four bands, with two positive bands attributed to the ESA and two negative bands attributed to GSB and SE. However, the maxima, intensity and shape of these bands differ. Compared to WT, the maxima of ESA bands for V151A are blue-shifted at 330 nm and 445 nm while V151L has its ESA maxima red-shifted at 350 and 460 nm. Nevertheless, the main difference can be seen in the SE region. Firstly, while the SE band of V151L displays only one minimum centered at 515 nm with a shoulder at 550 nm and without any red tail contribution, the SE band of V151A is characterized by a broadband from 490 to 650 nm and centered at 530 nm. Notice that V151A has a weak positive extra ESA band with a maximum at 680 nm at 150 fs. The spectral evolutions among the two variants also differ. For V151L, there is a small growing of ESA band centered at 460 nm until 400 fs(also seen for WT). However, for V151A, in 400 fs, there is a new evolution of the SE band. Mainly there is a decrease and shift of the SE band maximum (from 530 to 550 nm), and simultaneously, the ESA band at 680 nm becomes negative, i.e. a new SE band. Notice the isosbestic point at 625 nm between these two bands (Figure 5.1.f). At 600 fs, the two SE band

contributions (550 nm and 680 nm) have almost the same intensity. This evolution is also accompanied by a red-shift of the second ESA band maximum from 445 to 450 nm. After this ultrafast spectral shift, we can observe a general decay of all bands in less than 2 ps. At 2.35 ps the value of ΔA for GSB and ESA bands decrease over 2/3 of the value at 600 fs. Interestingly, the SE band has almost completely decayed to zero. At 2.35 ps the SE band displays only a minimum at 510 nm with no signal over 650 nm. This ps evolution is followed by the continuous decay until 40 ps of the SE and GSB bands. At 40 ps the remaining transient spectrum has two positive maxima (375 and 460 nm, similar to WT), and no SE signal can be observed. Strikingly, for V151L the evolution of transient spectra after 400 fs are more simple. All the bands decay in about tens of picoseconds without shifts of the SE maximum always centered at 512 nm. At 40 ps, the transient spectrum has a positive band with a maximum at 375 nm, and two negative bands peaking at 445 nm and 500 nm. The remaining transient spectrum at 40 ps is a mixture of bands from the photoproduct (similar amplitude in agreement with the photo-switching yield values, 11% for WT, 12% for V151L, 14% for V151A), and the dynamics of the cis anionic *On* form present in photo-stationary *Off* state (5% for WT, 14% for V151L and 2% for V151A).

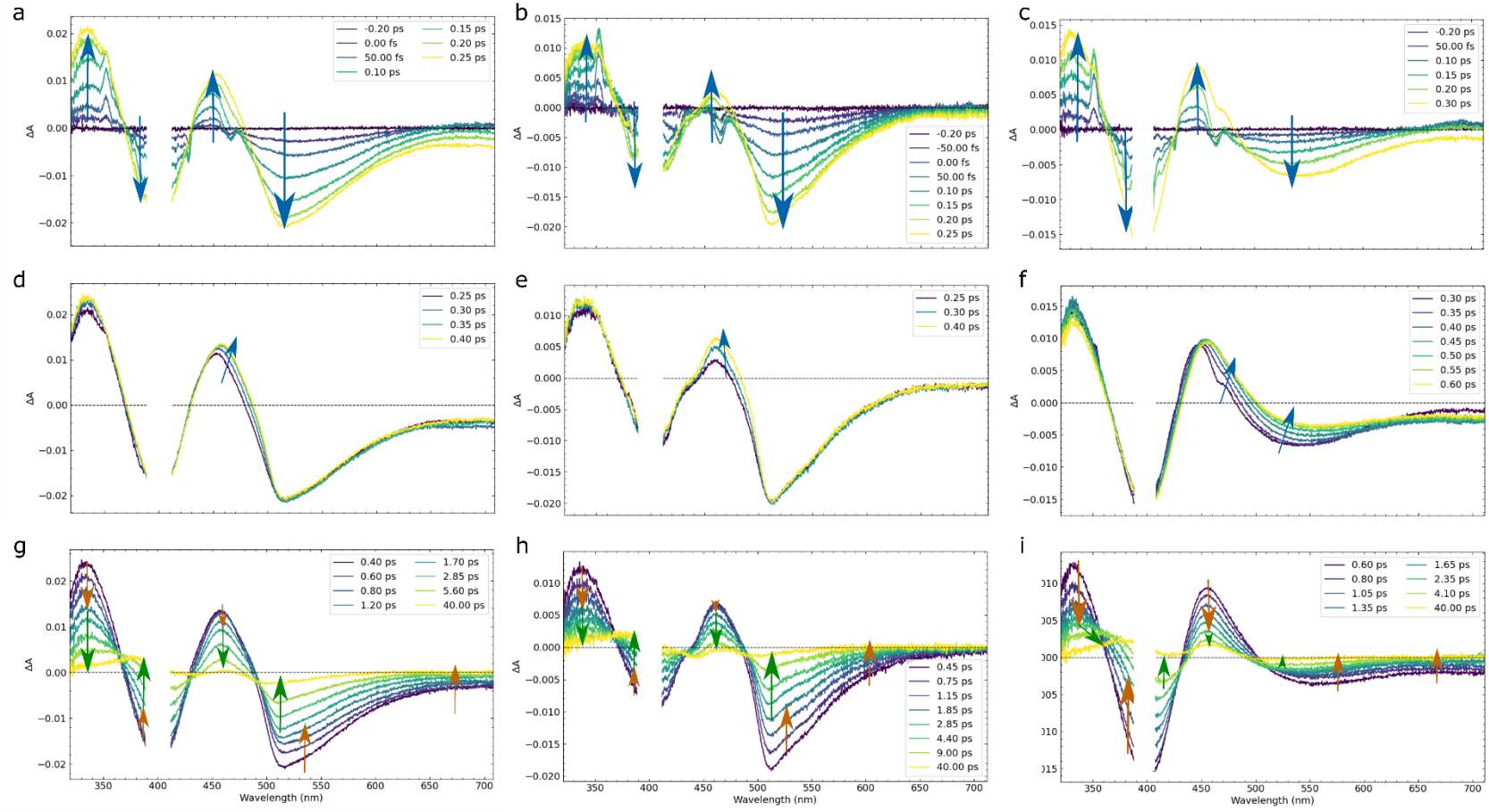


Figure 5.1. Femtosecond transient difference absorption spectra recorded at different time delays after a femtosecond laser excitation (400 nm) of the trans protonated Off state until 40 ps for WT (a,d g), V151L (b,e h) and V151A (c, f, i) in H2O solution (50 mM HEPES pH 8, 50 mM NaCl). The spectrum without laser excitation was subtracted to calculate the difference spectra. Upper panel correspond to the growing of the signal, middle panels to sub picosecond evolution and bottom panels to picosecond evolution. Figures made with Ultra Pyfit, arrows added with Inkscape.

The transient spectra evolutions until 40 ps for WT, V151A and V151L can be compared in Figure 5.1. The presence of SE bands indicates that the observed dynamics correspond to excited-state species. An interesting feature of the transient spectra in Figure 5.1 is the relative amplitude between the GSB and SE bands at 300 fs (Figure 5.1d,e,f). For V151L and WT the amplitude of SE band is higher than for the GSB whereas this is reverse for V151A. Secondly, the ratio between these bands is higher for V151L than for WT. Nevertheless, the most interesting difference is observed in the decay of the SE bands (after 600 fs time delay). For V151A the SE is completely red-shifted and decay without changes of shape and maximum (after 600 fs) and thus can be characterized by mainly one excited-state species with a fast decay (~90 %) and a second minor contribution decaying in longer time scales. For the V151L there is also short and a long component with similar contributions and shapes (Figure 5.2). Interestingly, for WT which also has two clear components; the shorter one has a red shifted contribution together with a minimum around 520 nm, thus, this band can be seen as a mixture of V151L and V151A short components (sub-ps) SE band. The global fit of the kinetic traces (convolved with the laser pulse 0.16 fs width and a constant τ = inf, see Appendix 2 for details) yielded three-time constants, for WT (0.17 ps, 0.89 ps and 4.33 ps), for V151L (0.13 ps, 0.70 and 4.20 ps) and for V151A (0.17 ps, 0.72 ps and 4.74 ps). The DAS obtained from the global fit are shown in Figure 5.2 together with the charcteristic kinetic traces for each of the transient signals. The residues of all fitted traces are shown in Figure 5S-1. The first time constant of ~170 fs is attributed to the build-up of electronically excited intermediates from the Franck–Condon state, and the other two time constants to their decay. The SE band of the corresponding DAS confirm that WT excited states dynamics can be seen as a mixture of those of V151L and V151A.

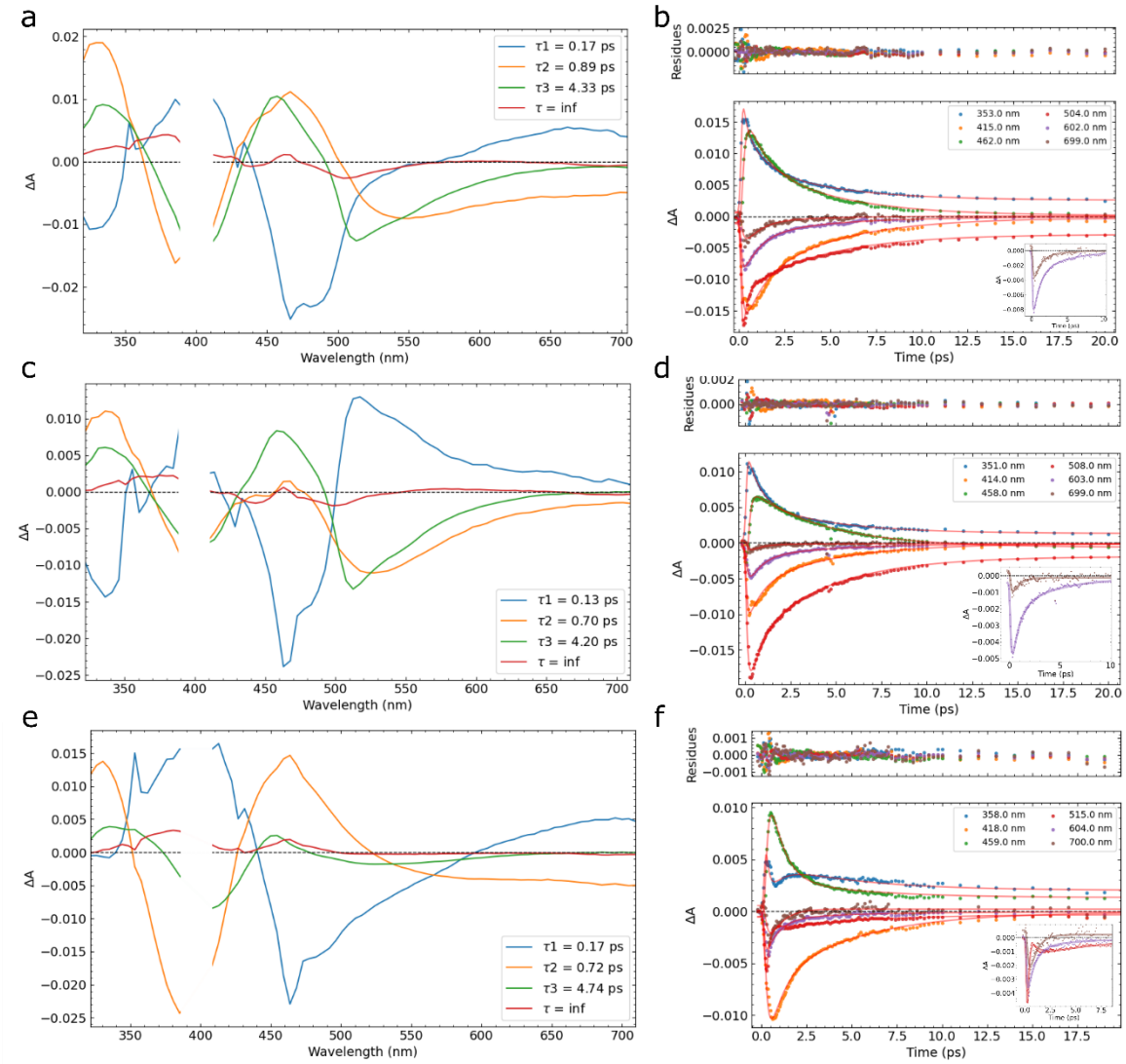


Figure 5.2. Left panels correspond to the decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis of kinetic traces (every 5 nm) using a weighted sum of three exponential functions convolved with a Gaussian function and a constant (inf). In the right panels, selected kinetic traces for wavelengths representing each of the transient absorption spectral main features (ESA, GSB and SE bands) together with the corresponding fits and residues. Panel a and b correspond to WT, c and d to V151L and e and f to V151A. Figure made with Ultra Pyfit.

5.2.1.2 Ground-state dynamics, dynamics between 40-2000 ps in H2O solution

The decay of the SE band at 40 ps indicates that excited-state species have relaxed to the ground state. Therefore, the spectral evolutions from 40 ps to 2000 ps are probing intermediate species in the electronic ground state.

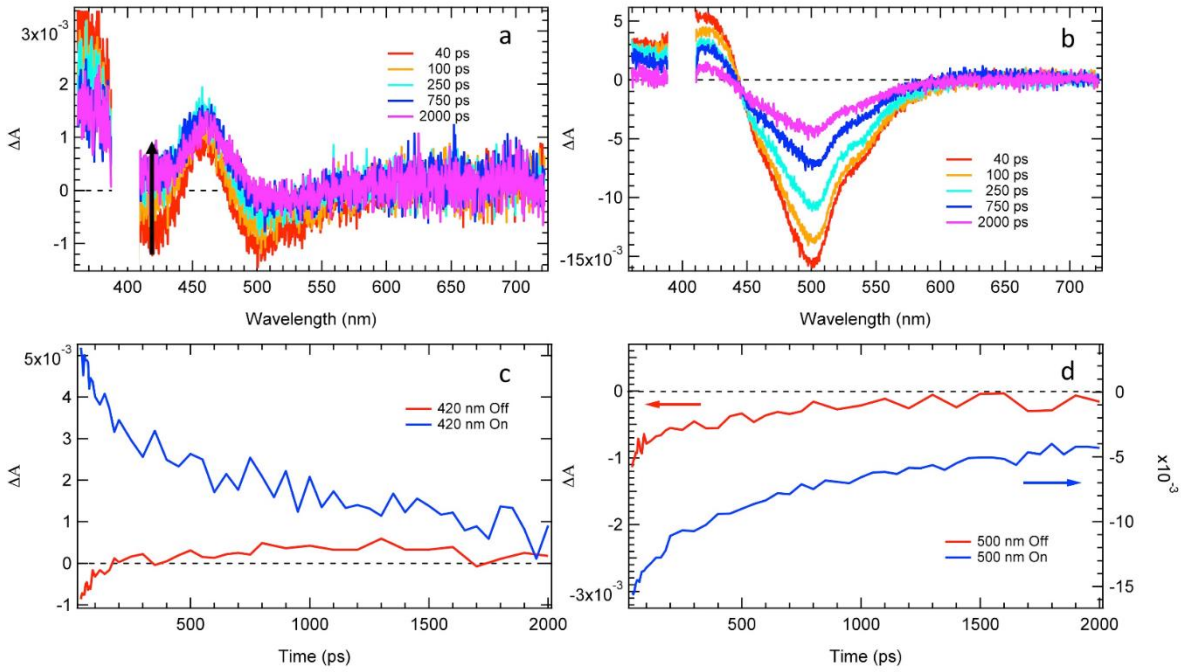


Figure 5.3. Time-resolved difference absorption spectra recorded between 40 ps and 2 ns after a 400 nm femtosecond excitation of the Off-state (a) and the cis-anionic on-state (b) of rsEGFP2 in H2O solution (50 mM HEPES pH 8, 50 mM NaCl). The black arrow (a) indicates the disappearance of the 420 nm band within 87 ps. Kinetic traces at 420 nm (c) and 500 nm (d) extracted from panels a) and b) are shown. Figure done with Igor pro.

The photoproduct ΔA value is in the limit of the instrument signal to noise ratio. To avoid long time experiments and photo-degradation of the protein, we did not measure at the same time short time delays (0-40 ps) and longtime delays (0-2 ns). For long time delays, we also measured transient spectra between 0 and 40 ps, but with a reduced number of points until 10 ps and on a new fresh sample. Therefore, the absolute value of ΔA is not exactly the same. It is worth noticing that the *Off* state is formed by contributions of *trans* neutral and *cis* anionic forms. Therefore, the evolutions for time delays longer than 40 ps are a mixture of the anionic excited state dynamics (negligible for V151A) and photoproducts of *trans* neutral

photodynamics. This explains the existence of a negative band at 500 nm assigned to SE of cis anionic *On* excited state. To confirm this attribution and solely identify the trans neutral photoproduct dynamics, we have followed *On* form spectral changes upon 400 nm femtosecond excitation (Figure 5S-2 and 5S-3). The *On* excited state dynamics will be discussed in Chapter 6, nevertheless, the essential result of these experiments is that the WT ESA, GSB and SE band of cis anionic form are decaying continuously until 2 ns (Figure 5.3b). Consequently, the growing of the positive band at 420 nm (Figure 5.3a and 5.3c) is solely due to photodynamics of the *trans* neutral photoproduct. The exponential fit of the kinetic trace at 420 nm with four components yielded an extra associated characteristic time constant of 87 ps (result published in Woodhouse et al.[5]). The new band at 420 nm seen for WT is even clearer for the V151A (Figure 5.4c) because the residual *On* form population in *Off* state is negligible. Therefore, all kinetic traces (0 to 2 ns) could be globally fitted with four components and increase the precision on time associated to the spectral evolution after 40 ps time delay. The results can be seen in (Figure 5S-4). The associated time constant for this extra evolution was 83 ps. For V151L the resting *On*-state in the *Off* state is estimated to be about 14 % (Chapter 4); therefore, the evolution of SE band around 512 nm is attributed to the residual *On*-state evolution Figure 5.4b. Besides this evolution, no additional changes from 40 ps until 2 ns can be seen (Figure 5.4b). This means that the *trans* neutral photoproduct formed after excited state relaxation does not evolve until 2 ns.

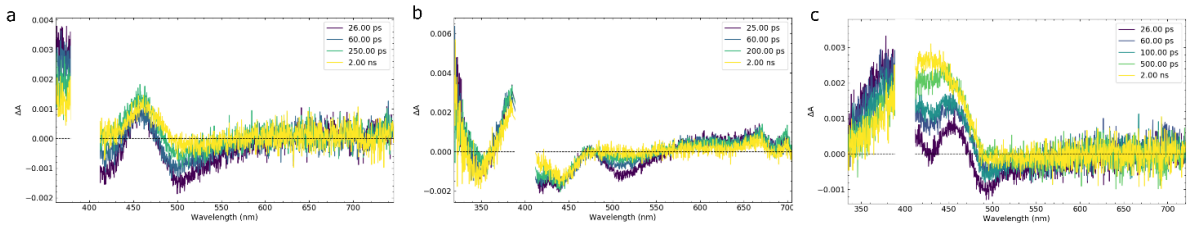


Figure 5.4. Femtosecond transient difference absorption spectra recorded from 20 ps to 2000 ps time delays after a femtosecond laser excitation (400 nm) of the Off state for a) WT, b) V151L and c) V151A. Figure made with Ultra Pyfit.

5.2.1.3 0-2 ns dynamics in D_2O solution

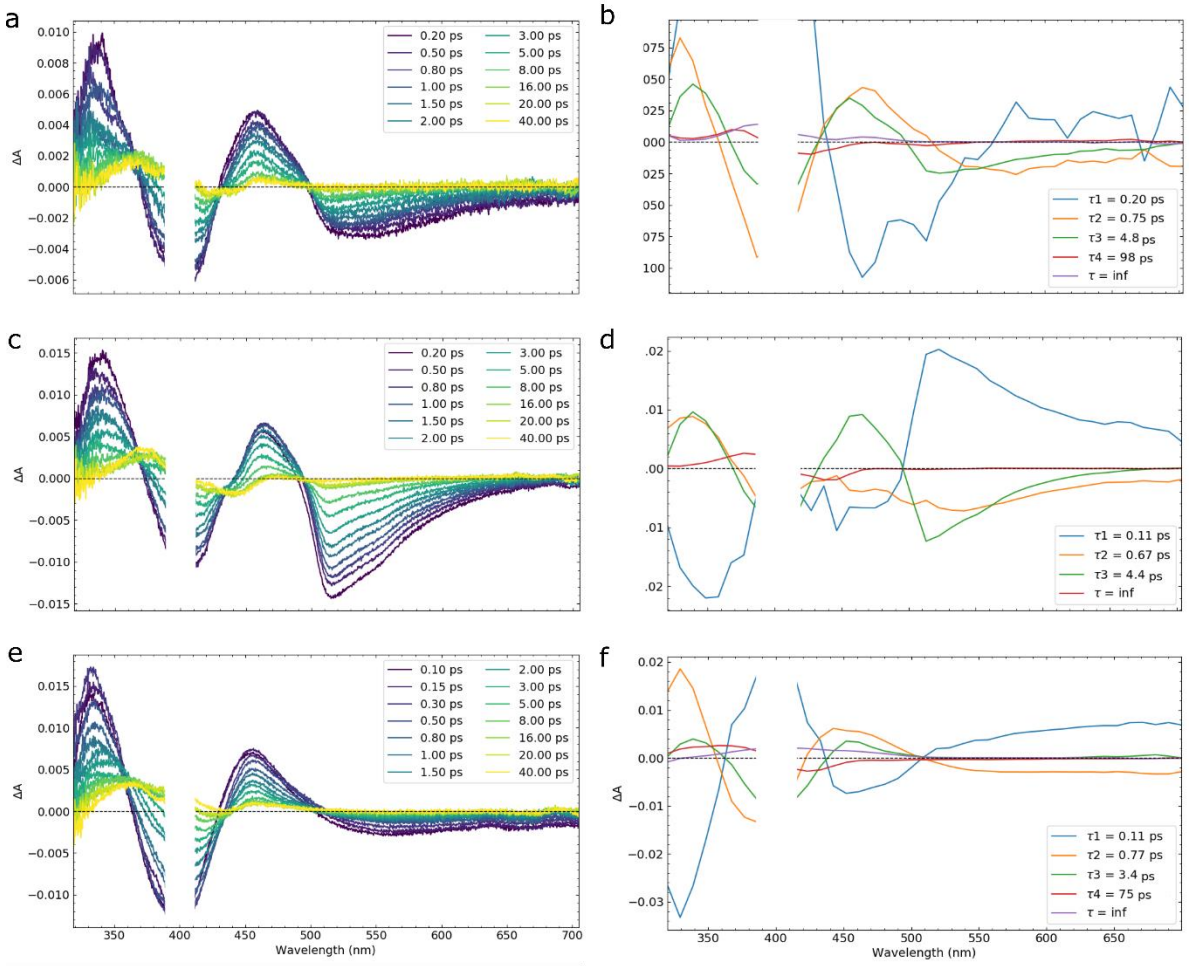


Figure 5.5. TRUV-Vis transient absorption experiments for WT, V151A and V151L in D_2O solution (50 mM HEPES pD 8, 50 mM NaCl). Left graphs (a, c, e) correspond to the time-resolved difference absorption spectra until 40 ps recorded after a femtosecond laser excitation (400 nm) of the Off state. Right graphs (b, d, f) correspond to the decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit (0 to 2 ns) of kinetic traces (every 10 nm) using a weighted sum of four exponential functions convolved with a Gaussian function and a constant (inf). ***WT*** *(top panels: a, b), V151L (middle panels: c, d) and V151A (bottom panels: e, f). Figure made with Ultra Pyfit.*

Similar experiments (Figure 5.5, only transient spectra until 40 ps is shown) to those reported above in H_2O have been done for WT, V151A and V151L in heavy water (50 mM HEPES pD 8, 50 mM NaCl in D_2O). Transient spectra and spectral evolutions correspond to those observed in H_2O. Three time constants (1 time constant for the growing of excited states and 2 time constants for their decay) are needed to fit the kinetics for V151L whereas a fourth time constant should be added to take into account the extra ground state evolution for WT and V151A. Global decay analysis of the data in heavy water yield for WT (0.2 ps, 0.75 ps, 4.8 ps and 98 ps), for V151L (0.11 ps, 0.67 and 4.4 ps) and for V151A (0.11 ps, 0.77 ps, 3.4 and 75 ps), the fits and residues can be seen in Figure 5S-1d, e and f. The comparison of time constants in H_2O and D_2O, reveals the absence of isotopic effects in the excited state dynamics.

5.2.1.4 0-2 ns dynamics for crystals

WT microcrystals were also studied using time-resolved UV-Vis transient absorption spectroscopy with the same setup. There are only two differences. Firstly, two LEDs emitting at 490 and 505 nm were used to keep the *Off* state during the experiment. Secondly, the optical path of the flow cell was increased from 250 µm to 500 µm. The number of time points collected was reduced to avoid photo-degradation. Since no significant changes can be observed after 60 ps, time points were collected only until 300 ps. The transient absorption spectra are shown in Figure 5.6. The results are comparable with the solution. The formation of excited states is within a few hundred femtoseconds. The transient spectrum at 200 fs displays 2 ESA positive bands and two negative bands attributed to GSB and SE (500 to 700 nm). In comparison to the solution, the ESA and SE maxima are red-shifted peaking at 475 nm and 550 nm. Kinetics traces (every 10 nm) were globally fitted from 0.2 ps to 60 ps with a bi-exponential function with associated decay times of 0.6 and 5.6 ps. The microcrystals red-shifted maxima indicate that the chromophore is feeling a different electrostatic field. Nevertheless, the DAS spectral features and time constants are similar, thus following the interpretation done for the protein in solution, the shorter decay time species (0.6 ps, red-shifted SE) is assigned to a twisted excited state intermediate during the *trans* to *cis* isomerization.

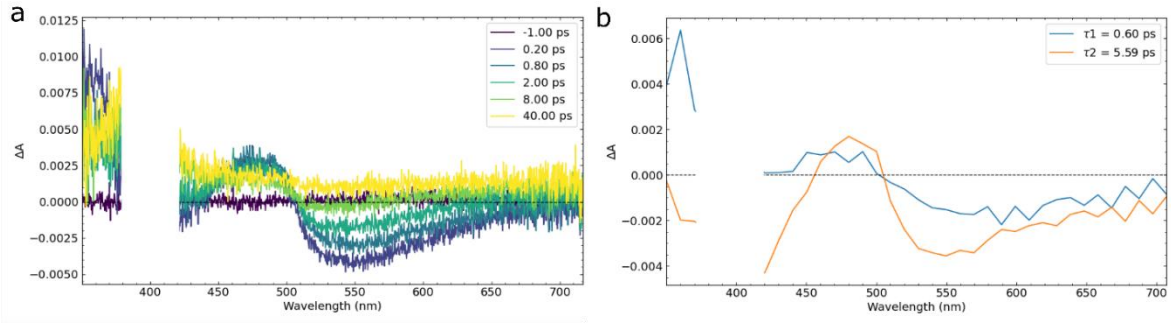


Figure 5.6. TRUV-Vis transient absorption experiments for rsEGFP2 microcrystals in 100 mM HEPES, pH 8, 2.5 M ammonium sulphate. a) time-resolved difference absorption spectra recorded after a femtosecond laser excitation (400 nm) of the trans protonated off state until to 2 ns. The spectrum without laser excitation was subtracted to calculate the difference spectra. b) decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit with a weighted sum of 2 exponential functions. Figure made with Ultra Pyfit.

The decay time associated with the isomerization (0.6 ps) is slightly shorter in crystals than in solution (0.88 ps) which can be the effect of crystal confinement. The second decay time seems to be slightly longer in microcrystals (5.99 vs 4.33 ps). However, the low S/N ratio of the microcrystals data makes difficult any precise conclusion. To come to an end, the excited state dynamics of rsEGFP2 microcrystal are dominated by two excited states, similar to those observed for the protein dynamics in solution, which validates the assignment of species done by Coquelle et al.[2] using TR-SFX.

5.2.2 Transient absorption Infrared spectroscopy

The photodynamics between 0-2 ns were also studied by TRIR spectroscopy[4] in heavy water. After irradiation of the *Off*-state with a (fs) laser pulse at 400 nm, some instantaneous signals are formed (0.3 ps is the shortest time delay that can be observed due to IR coherent artefact). As mentioned in Chapter 2, positive bands correspond to ESA and photoproduct intermediates and negative bands to GSB signals which correspond to the depopulation of the trans neutral *Off* state. The evolutions of transient spectra for different time delays can be seen in figure 5.7 for WT, V151A and V151L.

At 0.3 ps the WT spectrum displays a rich spectrum, the main bleaching bands are peaking at 1681, 1657, 1639, 1627, 1576, 1533, 1511 and 1488 cm^{-1} (*trans* neutral chromophore), those at 1681, 1657, 1639 and 1488 cm^{-1} are the ones with higher intensities and have similar amplitudes. Based on previous studies of Dronpa protein which has the same HBDI chromophore but different protein cage, the bleaching bands at 1681, 1639 and 1627 cm^{-1} are assigned respectively to C=O bond, delocalized C= C–N= C bonds and phenyl ring stretch vibration of the *trans* neutral chromophore[4]. The band at 1657 cm^{-1} is assigned to the C=C stretching vibration. Finally, the band at 1488 cm^{-1} is assigned to the residual *cis* anionic excited state dynamics. The main positive bands correspond to the signatures of excited state species and can be seen at 1666, 1649, 1611, 1594, 1560-1546, 1520 and 1500 cm^{-1}. Three time windows can be identified from the transient spectra evolutions: 0.3-8 ps, 8-50 ps and 50 ps-3 ns. From 0.3 to 8 ps some spectral evolutions can be seen between 1580 and 1700 cm^{-1}, primarily involving the decay of the negative GSB bands (1657 and 1680 cm^{-1}). At 8 ps the GSB bands have decreased by 2/3, this is assigned to the recovery of the *trans Off* form. New positive bands at 1691 cm^{-1} and 1587 cm^{-1} grow within the same time evolution. The 1691 cm^{-1} band is assigned to a cis photoproduct (see below). From 8 ps to 50 ps, the GSB bands continue to decay (for clarity the spectra at 0.3 and 50 ps are displayed in Figure 5S-5). Finally, from 50 ps to 3 ns the prominent evolution involves the decay of the GSB at 1488 cm^{-1}, which is accompanied by the decay of three positive bands between 1350 and 1470 cm^{-1}, and two positive bands peaking at 1627 and 1649 cm^{-1}. However, a comparison with *On* to *Off* transient spectra recorded after 480 nm excitation and DAS obtained by a global fit of all kinetic traces (Figure 5S-6), allows attributing the positive bands evolutions between 1350 - 1470 cm^{-1} and the bleach at 1488 cm^{-1} to the *On* residual form (cis anionic) ESA and GSB dynamics respectively.

Similar results can be seen for the variants V151L and V151A. Notice that the contribution of the *On* form dynamics after 50 ps is lower for V151A. Important differences can be seen in the transient bands assigned to the *Off* to *On* dyanmics for both variants. While at 0.3 ps the WT spectrum has two prominent GSB bands at 1639, 1627 cm^{-1}, only the first one is seen for V151A while the second is seen for V151L. Moreover, the excited state spectral evolutions are mainly finished at 8 ps for V151A. Importantly, it is possible to identify a new positive band at about 1691 cm^{-1}, assigned to a cis form photoproduct. A similar band for

V151L appears only after the 1681 cm^{-1} GSB recovery (see spectrum at 20 ps) and at 1685 cm^{-1}. The excited state decay for V151A is faster than for V151L. Importantly, after 20 ps for WT and V151A variants the most prominent positive band is at 1649 cm^{-1}, while on the contrary, this band has completely disappeared in V151L, which has as main positive band the 1587 cm^{-1} (also present in WT and V151A). The detailed analysis of photoproducts assigned to this positive bands is done in the discussion.

The *Off* to *On* kinetic traces were globally fitted taking into account the above-explained evolutions. To consider the resting *On* state contributions, a weighted sum of four exponential was used, and the longest time value was fixed to the *On* to *Off* photodynamics TCSPC value (see Table 7.1, the values are 2.8, 2.2 and 2.5 ns for WT V151L and V15A respectively). The associated fit time constant for WT were 1.51 ps, 6.26 ps and 180 ps. For V151L 1.46 ps, 5.88 ps and 90.9 ps. Finally, for V151A, 0.78 ps, 4.50 ps and 79.0 ps. The results can be seen in Figure 5.8 and the fits and residues in Figure 5S-7. For the three proteins the two first time constants can be assigned to the two excited species found in TRUV-Vis experiments (Table 5.1 next page). Time constants seem longer in TRIR and it is usually due to the presence of vibrational cooling convolved to excited state deactivation[6,7]. Importantly during the first species de-activation processes, we can see the formation of a positive band at 1649 cm^{-1}. Compared to the FTIR differential spectra (Figure A1.11 Appendix 1), allow us attributing this band to a *cis* type photoproduct. It is difficult to attribute the geometry and nature of excited species without theoretical quantum calculations of the spectra which are in progress. Nevertheless, the absence of evolutions in TRUV-Vis for V151L after 20 ps allows us attributing the longer components found for this variant to the *cis On* form. Similarly, the comparison of the DAS long component found for the *On* to *Off* photodynamics (Figure 5S-7) with those found here, allows assigning the last evolutions to the residual *On* form for V151A and WT. Contrary to TRUV-vis, **t**he first ground state evolutions found for V151A and WT variants (~100 ps) are difficult to observe with TRIR. All time constants retrieved from global fits done to the decay traces of TRUV-Vis in H_2O and D_2O and TRIR the are compared in Table 5.1.

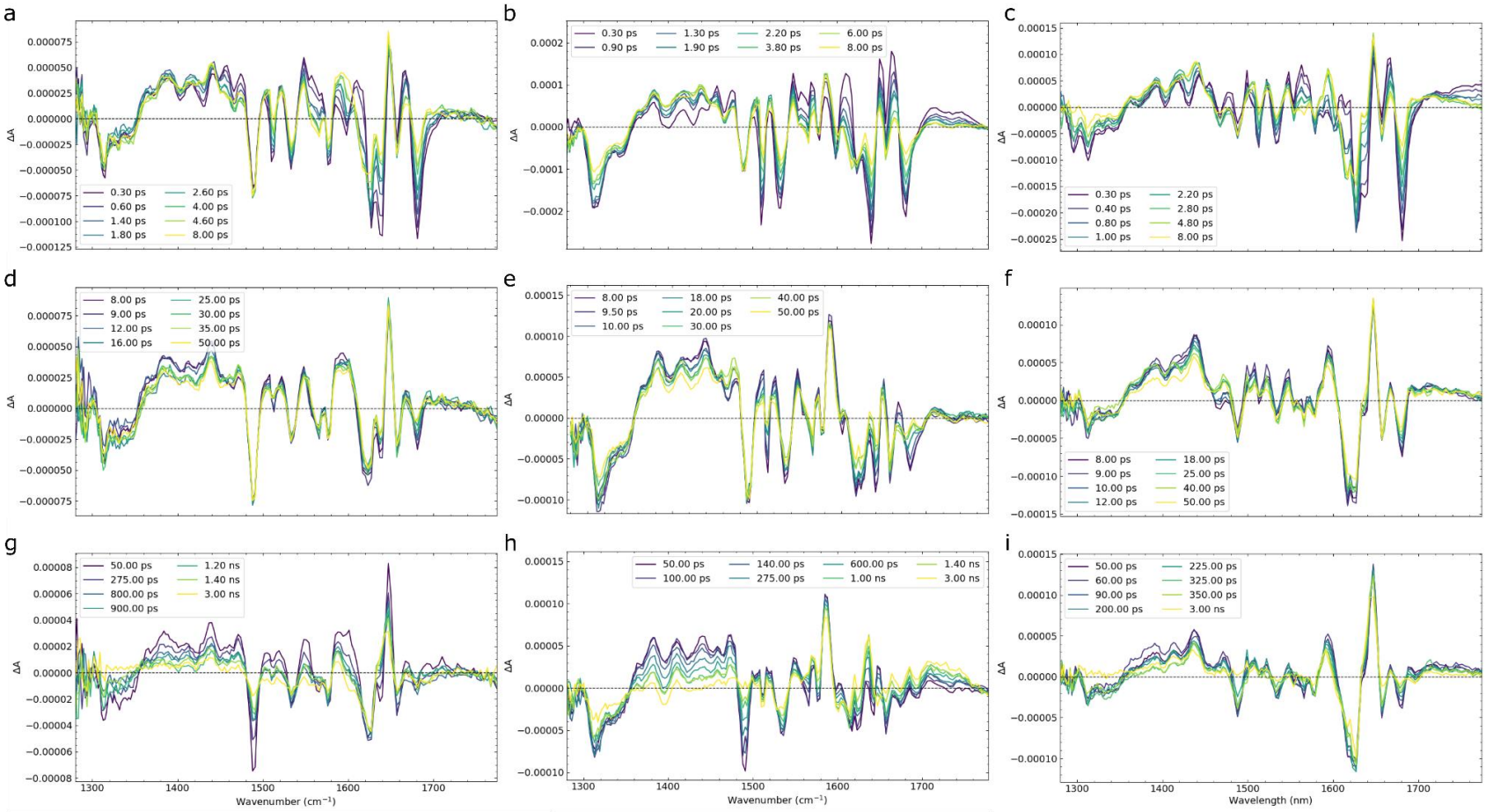


Figure 5.7. Femtosecond transient difference absorption IR spectra recorded at different time delays (until 3 ns) after a femtosecond laser excitation (400 nm) of the Off state (50 mM HEPES pD 8, 50 mM NaCl). Top, middle and bottom panels correspond to time delay between 0.3 and 8 ps, 8 and 50 ps, and 50 ps to 3 ns, respectively. WT variant panels are (a,d and g), V151L (b,e and h) and V151A (c, f and i). Figures made with Ultra PyFit.

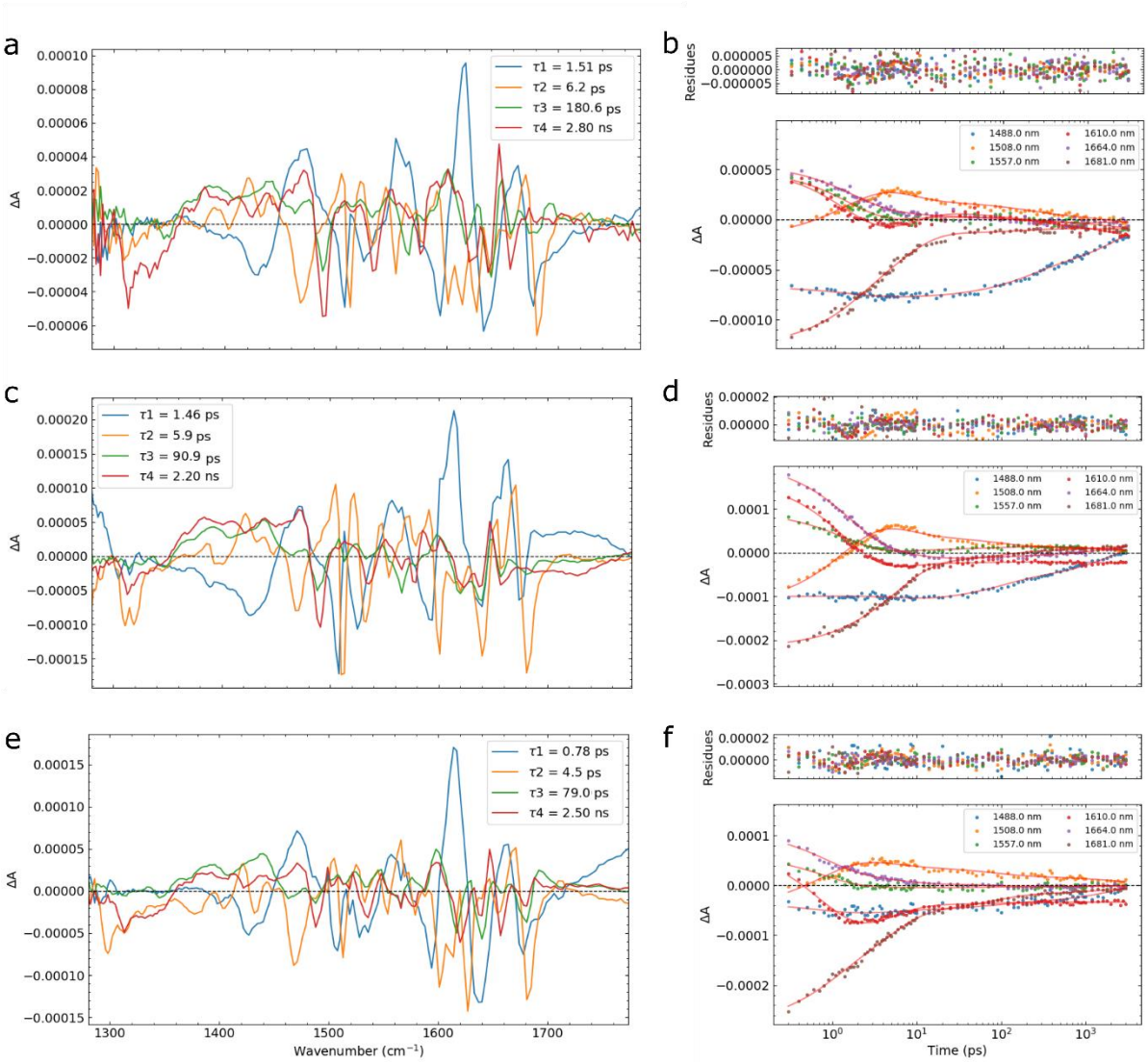


Figure 5.8. Left panels correspond to the decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis of all kinetic traces in data sets in Figure 5.7 with a weighted sum of four exponential functions. In the right panels, the selected traces representing the main transient bands together with their corresponding fits and residues. Notice that the time axis is set in a logarithmic scale. Panel a and b correspond to WT variant, c and d to V151L and e and f to V151A. Figure made with Ultra Pyfit.

Table 5.1 Time constants found after globally fitting the kinetic traces from TRUV-Vis (H_2O and D_2O) and TRIR (D_2O) data.

RSEGFP2	TIME CONSTANTS	TRUV-VIS (H_2O) (ps)	TRUV-VIS (D_2O)* (ps)	TRIR (D_2O)* (ps)
WT	τ1	0.17	0.20	NA
	τ2	0.89	0.75	1.51
	τ3	4.33	4.79	6.26
	τ4	87.0***	98.2	180.6
V151L	τ1	0.13	0.11	NA
	τ2	0.70	0.67	1.46
	τ3	4.20	4.44	5.88
	τ4	Not seen	Not seen	90.9**
V151A	τ1	0.17	0.11	NA
	τ2	0.72	0.77	0.78
	τ3	4.7	3.4	4.5
	τ4	83*	75	79

NA: Not applicable,
*The fit has been done including the last component to the *On* dynamics fixed (2.8, 2.5 and 2.77 ns for WT V151L and V15A respectively).
** Attributed unequivocally to *On* excited state dynamics.
*** Single wavelength fit[5]

The estimation of the errors of the fitted parameters have been calculated by inverting the Hessian matrix which represents the second order derivative matrix. From the square-root of the inverse of the diagonal, the standard errors for each variable parameter can be calculated independently and assuming that the components of the residual array are distributed around zero on the x-axis (residual plots) with a normal (Gaussian) distribution. This simple approach to calculate uncertainties ignores outliers, highly asymmetric uncertainties, or complex correlations between parameters. The values for the TRUV-Vis in H2O $\tau_1 = \pm 0.002$, $\tau_2 = \pm 0.010$ $\tau_3 = \pm 0.029$ and $\tau_4 = \pm 3.9$. For the TRUV-Vis in D_2O the $\tau_1 = \pm 0.03$, $\tau_2 = 0.032$ $t_3 =$

±0.23 and τ_4 = ±4.8. For TRIR in D_2O τ_2= ±0.044, t_3= ±0.32 τ_4 = ±5.7. Notice the lower values TRUV-Vis in H2O which are explained to the higher number of points and quality of the data in the 0-40 ps region.

5.3 Dynamics between 2 ns and 10 ms

5.3.1 Transient absorption UV-Vis Spectroscopy

5.3.2.3 Dynamics in H_2O and D_2O buffers

The spectral evolutions from 2 ns until 10 ns cannot be recorded since the excitation laser has a temporal width of ~8 ns. However, no evolutions can be seen between 10-100 ns in Figure 5S-8. Moreover, the comparison between the transient spectra at 2 ns (previous section) and 100 ns (410 nm excitation) shows the absence of spectral evolutions between 2 and 100 ns for V151A and V151L. The small difference for WT is assigned to different ratio of HT and OBF photoproduct in the *Off* state between experiments.

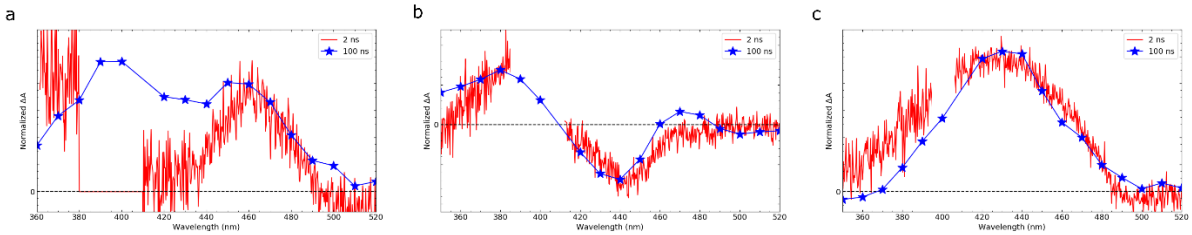


Figure 5.9 Normalized transient spectra at 2 ns (red) recorded after a femtosecond laser excitation (400 nm) compared to the spectra at 100 ns (blue) recorded after a 410 nm nanosecond excitation of the Off state (50 mM HEPES pH 8, 50 mM NaCl). a) WT, b) V151L and c) V151A.

Figure 5.10 shows spectral changes from 100 ns to 10 ms (20 ms for V151A) for the different RSFPs. The WT transient spectrum at 100 ns shows a broad positive band with a maximum at 400 nm (cis neutral photoproduct dark blue in Figure 5.10a). This band evolves within 5 µs to a spectrum with two positive maxima at 390 nm and 460 nm. Then within 50 µs, the first peak (390 nm) vanishes, while a negative band appears at 410 nm (GSB band). The second peak (460 nm) increases and shifts to 480 nm. Subsequently, and on a timescale from

100 µs up to 10 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 480 nm band is the characteristic absorption band of the *cis* anionic form[5] (Figure 4.2). The global fit of all kinetic traces identified three time constants, 5.43 µs, 34.40 µs and 827.71 µs (residues can be seen in Figure 5S-9). When similar experiments were carried out in D_2O solution (50 mM HEPES pD 8, 50 mM NaCl; 5S-10a and b), time evolution also required fitting with three exponential functions, yielding time constants of 5.14 µs, 78.48 µs and 2.05 ms. Thus, the first time constant is similar in H_2O and D_2O, but a significant isotope effect is observed for the two longer components (k_H/k_D= 2.45 and 2.47, respectively) and can thus be assigned to deprotonation steps (multi-step deprotonation[8]). Decay associated spectra (Figure 5.12a) show that the 5.47 µs time constant is mainly characterized by a growth of the positive band at 460 nm. The 34.5 µs time constant has some positive and negative contributions characteristic of the band shift observed, while the 827.7 µs time constant is mainly characterized by the respective decay and growth of the 390 nm and 480 nm bands.

Similar experiments were carried out for rsEGFP2 V151L and V151A variants (Figure. 5.10b,e and h, and Figure. 5.10c, f and i, respectively). Transient difference spectrum of V151L shows at 100 ns two bands, a positive one centered at 380 nm and a negative one centered at 440 nm (GSB band, purple in Figure 5.10b). This spectrum evolves in about 4 µs with the growing of a new positive band at 470 nm while the positive band at 380 nm decreases and the minimum shifts to 420 nm. This evolution is followed by the vanishing at 25 µs of the maximum at 390 nm, and the increase and shift of the negative band (from 420 to 410 nm) and the second positive one (from 470 to 480 nm). Finally, from 40 µs to 2 ms, the band at 480 nm keep increasing while the band at 410 nm shift to 400 nm and further decreases. Following the three different time windows (100 ns – 5 µs; 5 µs – 40 µs; 40 µs – 2 ms), the kinetic traces for all wavelengths were globally fitted with three exponentials, yielding time constants of 4.03 µs, 18.72 µs and 631.49 µs (Figure 5.12c and d). The analysis of the same experiments carried out in D_2O solution with the same sum of exponentials (50 mM HEPES pD 8, 50 mM NaCl; Figures 5S-10c, and d), yielded time constants of 4.59 µs, 33.05 µs and 1.4 ms. A significant isotope effect is observed for the two last time constants (k_H/k_D= 1.77 and 2.22, respectively), indicating as well that these last two evolutions can be assigned to deprotonation steps.

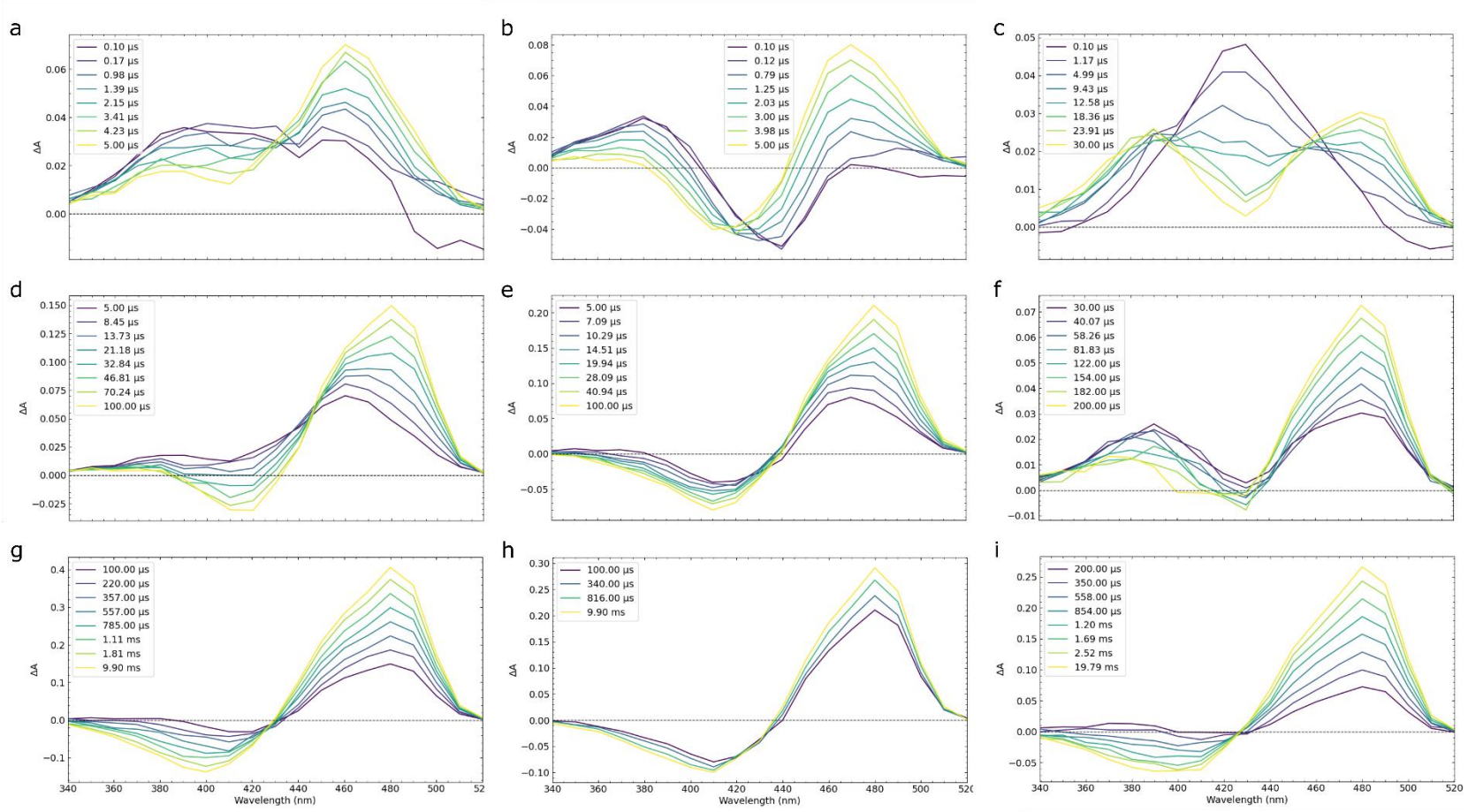


Figure 5.10.). Time-resolved difference absorption spectra recorded after a 410 nm nanosecond excitation of the Off state for three different time windows from 100 ns to 10 ms (50 mM HEPES pH 8, 50 mM NaCl). The spectrum without laser excitation was subtracted to calculate the difference spectra. WT variant panels (a,d and g), V151L (b,e and h) and V151A (c, f and i). Figures made with Ultra PyFit

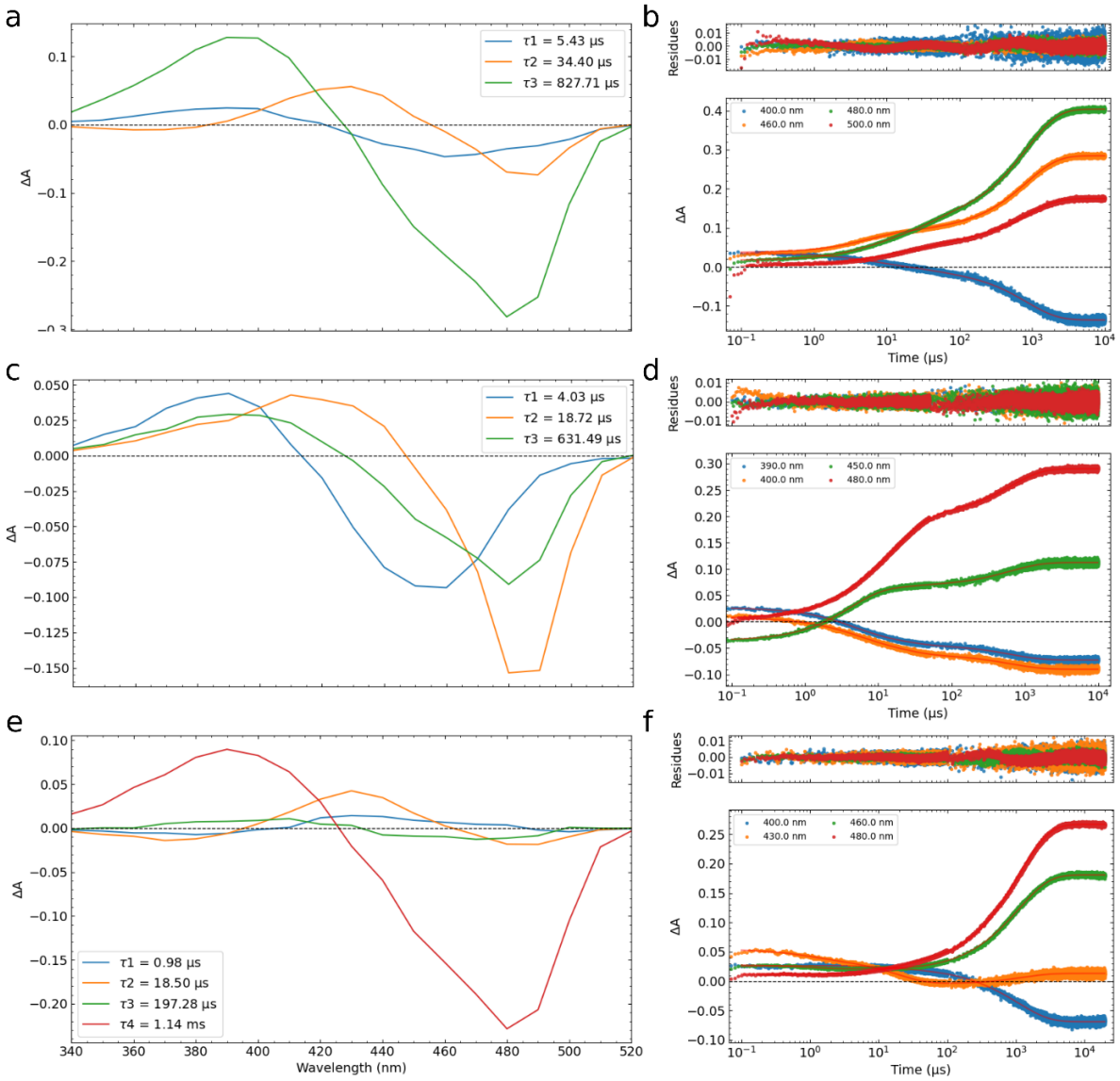


Figure 5.11. Left panels correspond to the decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis of the raw data in figure 5.10 with a weighted sum of exponential decay functions. In the right panels, the selected traces representing the main transient bands together with the corresponding fits and residues. Notice that the time axis has a logarithmic scale. Panel a and b correspond to WT variant, c and d to V151L and e and f to V151A. Figure made with Ultra Pyfit.

For V151A, the spectral evolutions slightly differ. Transient difference spectrum at 100 ns (figures 5.10c, and 5S-8c) shows a broad positive band with a maximum at 430 nm much more pronounced than for WT. This band decays and gives at 18.5 µs a transient spectrum (green) with two positive bands centered at 390 nm and 480 nm. Subsequently, in 260 µs, the 480 nm positive band continue increasing while the 390 nm band decreases and becomes negative with a minimum centered at 420 nm (Figure. 5.11f). Finally, in about 1150 µs the band at 420 nm continues to increase in amplitude and shift to 400 nm while the positive band at 480 further increases. Following the different evolutions, three time windows were identified (100 ns – 25 µs; 25 µs – 300 µs; 180 µs – 10 ms). Kinetic traces for all wavelengths were globally fitted with a weighted sum of four exponential functions, yielding time constants of 0.98 µs, 18.50 µs, 197.28 µs and 1.14 ms (figure 5.12e and f). When similar experiments were carried out in D_2O solution (50 mM HEPES pD 8, 50 mM NaCl; Figure 5S-10e, and f) the time constant were 1.76 µs, 20.91 µs, 251.16 µs and 3.95 ms. The isotope effects are assigned to deprotonation step which is clearly observed for the last step (k_H/k_D = 3.46). The first step shows also an isotope effect (1.8) indicating some protein rearrangement which involved hydrogen bond with HBDI. The 195 µs step characterized some band shift that can be interpreted by some non-single exponential process (diffusion, multi-step deprotonation).

5.3.2.3 Microcrystal dynamics

Transient absorption spectral evolutions and DAS are shown for WT in solution (50 mM HEPES pH 8, 50 mM NaCl) (Figure 5.12a), microcrystals suspensions (in 100 mM HEPES pH 8, 2.5 M ammonium sulphate) (Figure 5.12b) and WT in a partial crystalline sulphate buffer (50 mM HEPES pH 8, 50 mM NaCl, 1.25 ammonium sulphate) (Figure 5.12c). The measurement for crystal suspension can only be analysed from 600 ns (see Appendix 1 for methods). In general, it can be seen that transient microcrystals difference spectra are noisier; nevertheless, they are qualitatively similar to those measured in solution. i.e. have a positive band first growing at 460 nm and then at 480 nm are similar to those observed in solution. This result is important to validate the use of TR-SFX and linked the dynamics seen in solution to crystal structures.

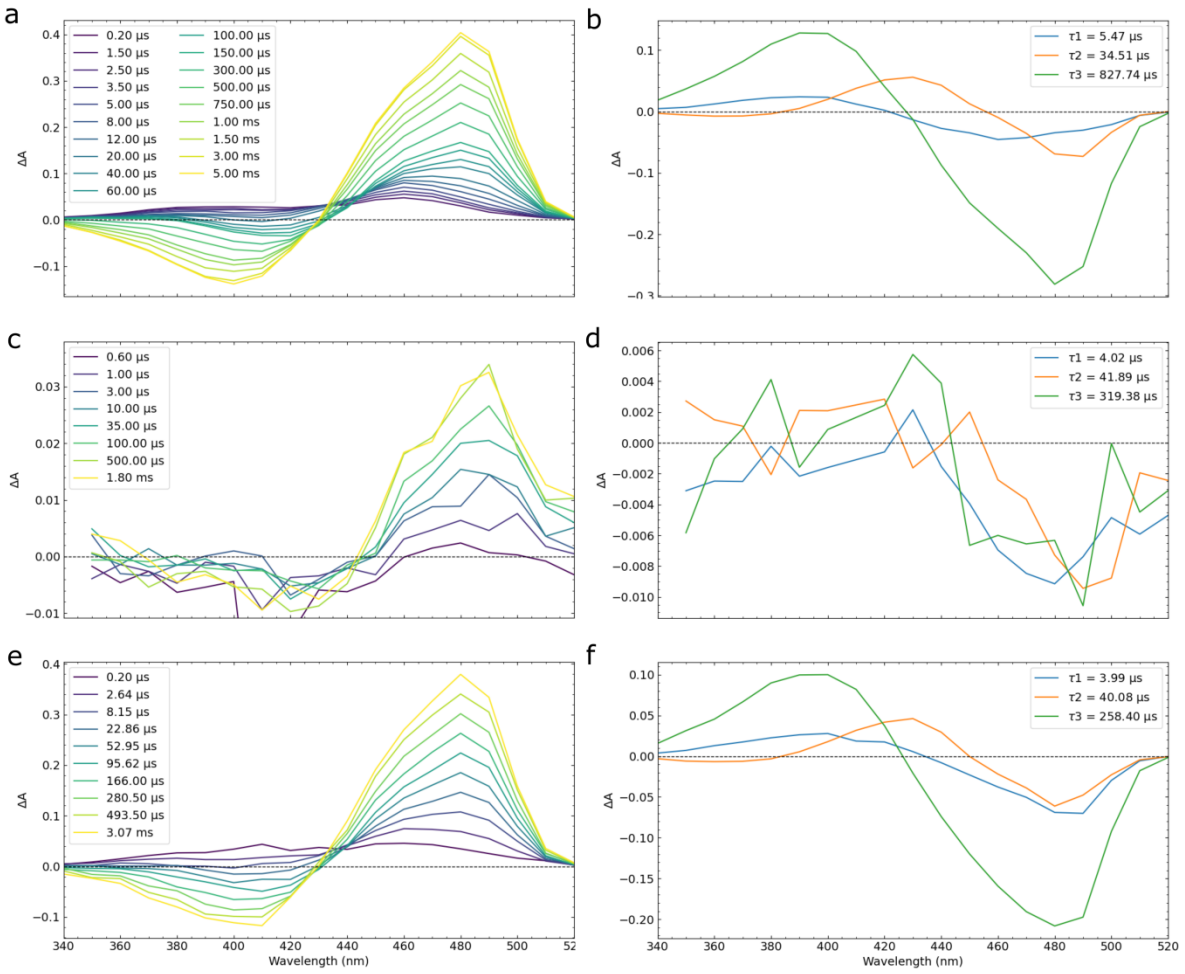


Figure 5. 12. Comparison of time resolved absorption results for WT in H_2O solution (50 mM HEPES pH 8, 50 mM NaCl, top panels), for a suspension of WT microcrystals (100 mM HEPES pH 8, 2.5 M ammonium sulphate, middle panels) and for protein in crystallization type buffer (50 mM HEPES pH 8, 50 mM NaCl, 1.25 M ammonium sulphate, bottom panels). Left graphs (a, c, d) correspond to the time-resolved difference absorption spectra recorded after a 410 nm nanosecond excitation of the trans protonated Off form in the time windows from 100 ns to 10 ms. The spectrum without laser excitation was subtracted to calculate the difference spectra. Right graphs (b, d, f) correspond to the decay associated spectra obtained from global fit of the raw data with a 3 exponential decay function.

The global fit of all microcrystals kinetic traces yielded time constants of 4.02 µs, 41.89 µs and 319.2 µs, *i.e.* significantly smaller than the rsEGFP2 in H_2O solution (50 mM HEPES pH 8, 50 mM NaCl). Interestingly, adding ammonium sulphate to the buffer (50 mM HEPES,

50 mM NaCl, pH 8, 1.25 M ammonium sulphate Figure 5. 12e and f) causes an effect to the dynamics, the time constants found where 3.99 µs, 40.08 µs and 258.4 µs. These values are much closer to the evolutions seen in microcrystals than those seen in solution. These results demonstrate that the solvent effect in the protein dynamics is more important than the confinement effect due to crystallization. The effect of ammonium sulphate in the solvent was further studied measuring kinetic traces at 480 nm for different buffer condition and in comparison, with D_2O. The results are shown in Figure 5.13a where the traces have been normalized to 1 (single shot measurements). All the traces were fitted with a 3 exponential function; the results can be seen in Table 5.2. The effect of ammonium sulphate with increasing concentrations from 0 to 1.5 M of ammonium sulphate was studied in D_2O solutions since they have more clear effects (Table 5.2). The result can be seen in the supporting Figure 5S-11 where it is clear that the acceleration of the deprotonation in microcrystals does not follow a linear dependence with ammonium sulphate concentration. On the contrary, small concentrations of ammonium sulphate induce a drastic decrease of the deprotonation time constant, with nearly no changes between 0.45 M to 1.5 M. To identify which ion, $NH_4^{+,}$ SO_4^{2-} or the ensemble of both, induced this effect similar experiments were repeated adding either Na_2SO_4 or NH_4Cl salts at equal NH_4^+ molar concentrations (Figure 5.13b). The results show that the deprotonation step is accelerated only when NH_4^+ is present. The fits can be seen in the supporting figures section Table 5S-1. The origin of the deprotonation acceleration with NH_4^+ is not clear. The variation of pH can be excluded since pH changes were not seen after adding NH_4^+ from 0 to 2.5 M to the buffer solution. The presence of tampon HEPES explains the absence of changes. The second hypothesis was that the positive charge of the NH_4^+ could stabilize the formation of a negative charge after the phenol deprotonation. On the contrary, the NH_4^+ and Na^+ are equally charged, and no effect could be seen when Na_2SO_4 was added. The reason for the deprotonation step been accelerated in the presence NH_4^+ is still elusive. Nevertheless, these experiments demonstrate that the differences between microcrystal and solution are caused by the presence of NH_4^+ and not by microcrystal structure.

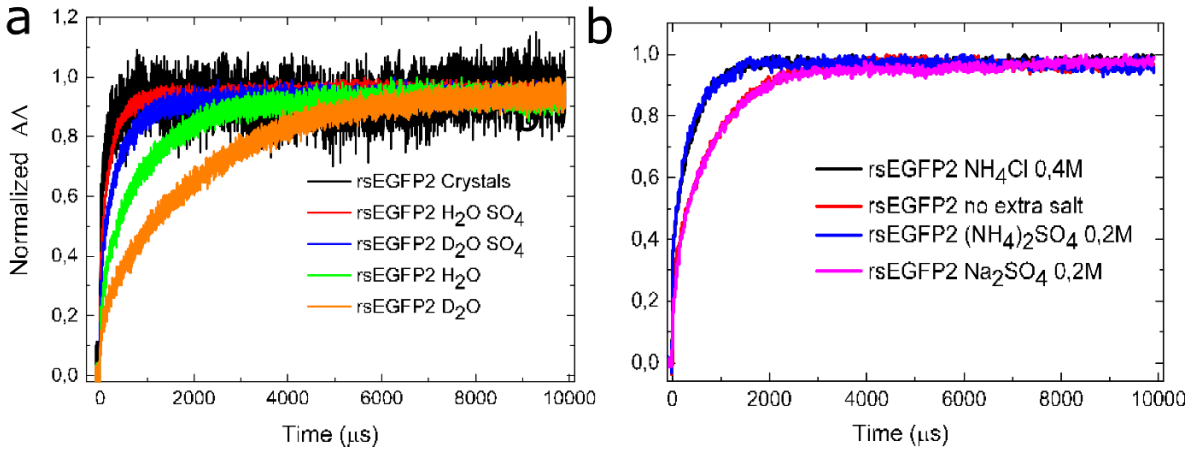


Figure 5.13. Normalized transient absorption kinetic traces at 480 nm, a) for a WT micro-crystals suspension (black), WT in H_2O and 1.25 ammonium sulphate (red), WTin D_2O and 1.25 ammonium sulphate (blue), WT in H_2O (green), rsEGFP2 in D_2O (orange). The base buffer for all solutions in panel a base buffer was 50 mM HEPES pH/D 8, 50 mM NaCl; for the micro-crystals suspension was 100 mM HEPES pH 8, 2.5 M ammonium sulphate. b) rsEGFP2 WT in 50 mM HEPES pH 8, 50 mM NaCl. With no added salt (red), with Na_2SO_4 0.2 M (pink), with NH_4Cl 0.4 M (black) and $(NH_4)_2SO_4$ 0.2 M (blue). Figure made with Origin Pro 8.

Table 5.2. Time constants retrieved using a weighted sum of three-exponential function to fit kinetic traces in Figure 5.13a

rsEGFP2 form	τ1 (**μs**)	τ2 (**μs**)	τ3 (**μs**)
Micro-crystals	3.1	56.7	338
H_2O $(NH_4)_2SO4$	3.7	41.2	310
D_2O $(NH_4)_2SO4$	6.7	86.3	504
H_2O	4.4	40.5	919
D_2O	5.5	73.8	2198

5.3.2 Transient absorption Infrared spectroscopy.

The TRIR evolutions in the ground state were followed using Time-resolved infrared multiple probe spectroscopy (TRMPS)[9]. The data have been corrected for baseline drift by fitting a 3 order polynomial to each spectra which was then substracted[4,9]. The TRMPS spectral evolutions in the ground state differ among the three proteins. Transient absorption spectra have been collected from 1 ps to 900 μs time range; the complete spectral evolutions can be seen in Figure 5.14. Notice that the initial spectral evolutions (top panels from figure 5.14) cover the same time range as the above-reported experiments using TRIR (0-2ns, Figure 5.7).

As discussed above, the *On* dynamics (Figure 5S-8 and Chapter 6) for all the variants can be described with two component of around ~120 ps and a longer component > 2 ns (obtained from TCSPC, see chapter 6). The DAS of these two components reveal that the contribution of the longer component to the signal is around 40%. Furthermore, at 2 ns the decay of signal attributed to 2 ns component is over 65 %. Therefore after 2 ns the *On* contribution (10-15% at photostationary state) is considered between 2 and 3 % of the signal and thus negligible. Since the shorter decay times are better resolved by TRIR data, the recorded TRMPS evolutions have been analysed up from 2 ns. After excitation WT and V151A variant display few bands and evolutions from 10 ns to 1 ms, while V151L display a richer spectrum with clear evolutions. At 100 ns the spectra of the WT and two variants have two main positive peaks at 1595 and 1650 cm^{-1} (1643 cm^{-1} for V151L) (Figure 5.15a) corresponding to the *cis* neutral photoproducts formed after excited-state isomerization. The main ground state bleaching is at 1634 cm^{-1} for WT and V151A, and 1605 cm^{-1} for V151L. Therefore at 100 ns the three variants display similar photoproduct signals.

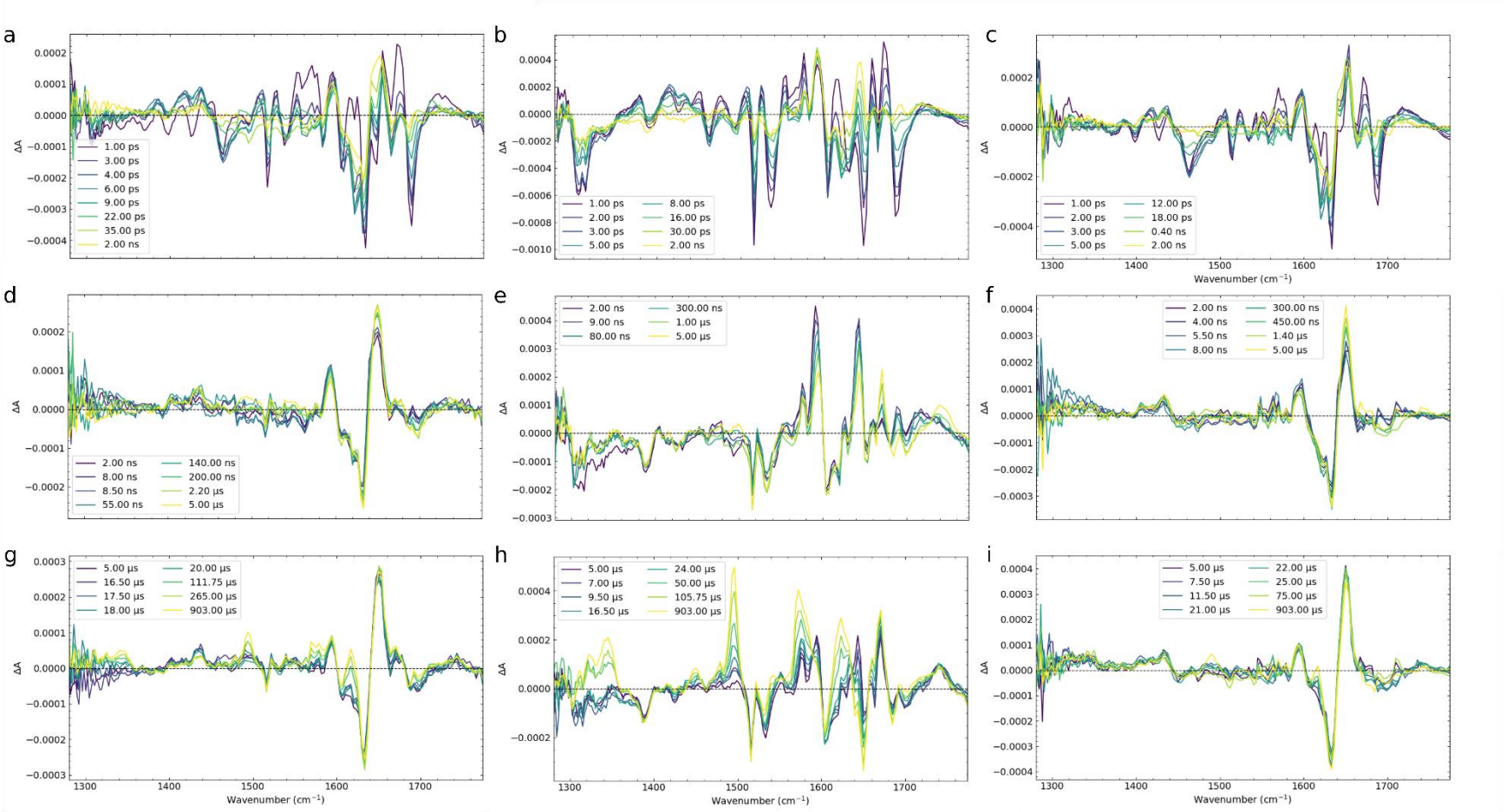


Figure 5.14. Selected transient IR spectra recorded at different time delays after a femtosecond laser excitation (400 nm) of the Off state from1 ps until 900 µs. The spectrum without laser excitation was subtracted to calculate the difference spectra. WT panels are (a,d and g), V151L (b, e and h) and V151A (c, f and i). Figure made with Ultra Pyfit.

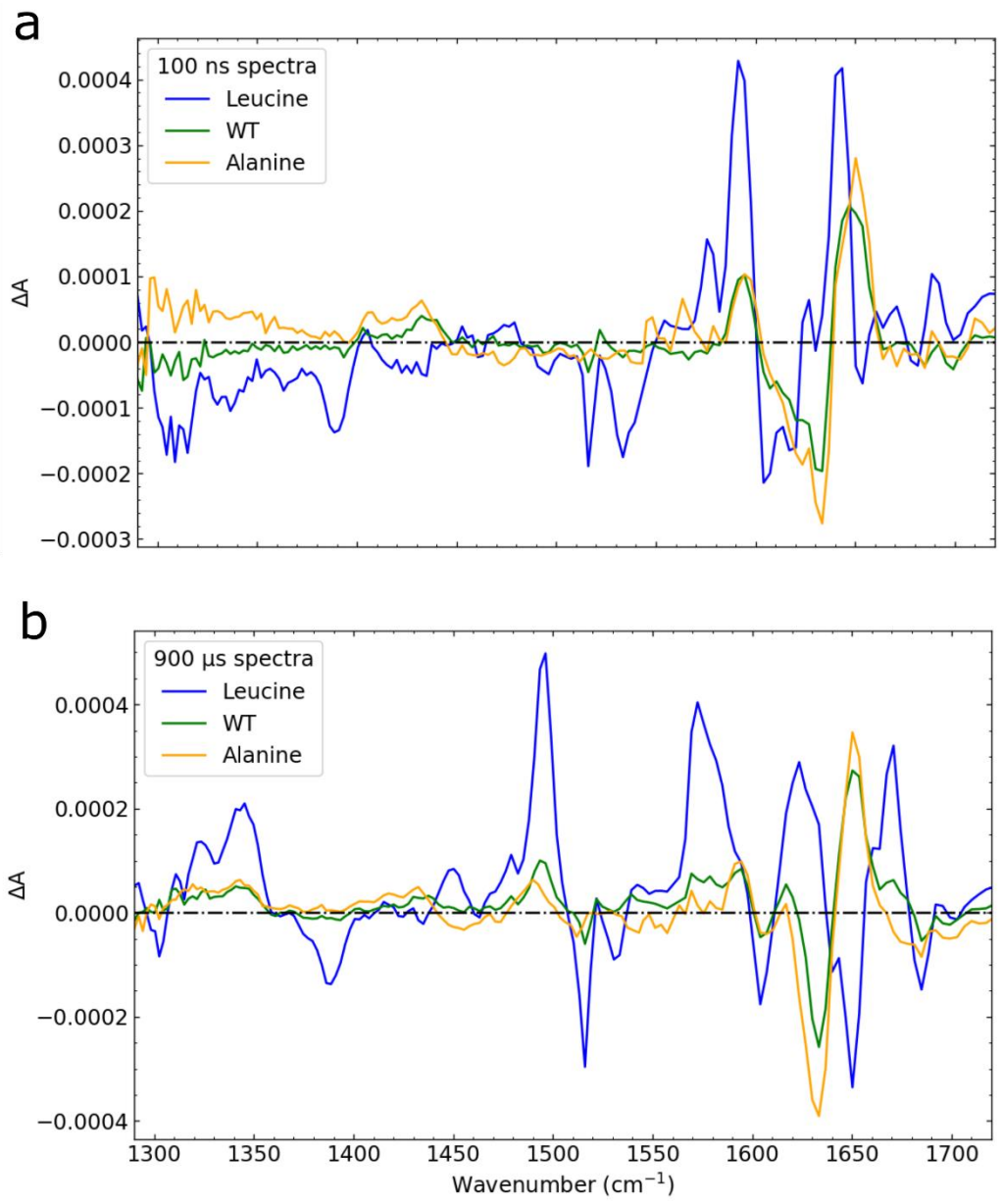


Figure 5.15. Difference IR spectra recorded after a femtosecond laser excitation (400 nm) of the Off form, for WT (green), V151L (blue) and V151A (yellow). The spectrum without laser excitation was subtracted to calculate the difference spectra. a) at 100 ns after excitation. b) at 900 µs after excitation. Figure made with matplotlib python library.

The comparison of transient spectra at 100 ns and 900 µs shows that there are nearly no major evolutions for WT and V151A, while for V151L the spectral changes are important (Figure 5.15 in blue) with clear disappearance and formation of bands. Indeed, for V151L the first initial positive bands peak at 1595 and 1643 cm^{-1}, both evolve and nearly disappear together

with the recovery of the red part of the bleaching band at 1605 cm^{-1}(transient spectra at 5 µs). Simultaneously, a new positive band is formed at 1670 cm^{-1}. These first evolutions are followed by the formation of four new positive signals in tens of µs (see transient spectra at 50 µs) at 1340, 1495, 1568, and 1615 cm^{-1}, together with the increase and shift of GSB band at 1650 cm^{-1} to 1644 cm^{-1} (also photoproduct evolution since the bleach of the *trans* neutral form is fixed after isomerization). The four new bands continue to grow until 900 µs. The bands peaking at 1495, 1523, 1568 and 1670 cm^{-1} can be seen as GSB band in the stationary FTIR spectrum of *Off* minus *On* differential spectrum (Appendix 1, Figure A1.17), and are thus assigned to the formation of *On cis* anionic form of V151L. Similarly, the GSB bands peaking at 1515, 1600, 1647 and 1682 cm^{-1} correspond perfectly to the positive FTIR band of the *Off trans* neutral form. All together proved the formation of the final *On* anionic form.

Contrary to V151L, WT and V151A spectral evolutions are less important in amplitude. However, the transient spectra at 900 µs is also characteristic by the disappearance of *cis* neutral photoproduct and the formation of *cis* anionic *On* form with the positive band peaking at 1490 cm^{-1}. This is an agreement with time-resolved UV-Vis transient absorption data which demonstrated that the deprotonation mainly occurs with time constants of 3.95 ms and 2.05 ms for V151A and WT respectively. Thus, only the initial evolutions can be inferred.

As previously all kinetic traces were globally fitted with a weighted sum of exponential. For all the data sets, the long associated time constant over the ms range found in TRUV-Vis D_2O data is fixed. The time constants (excluding the fixed one) for each mutant are: for WT 421 ns, 4.2 µs, and 98.1 µs; for V151L 377.12 ns, 11.1 µs and 57.8 µs; for V151A 530.7 ns and 26.7 µs. Interestingly a hundred ns time constant is needed to fit the data, whereas the shortest time constant found in TR UV-vis data is in the microsecond range. The nature of this extra species is not clear but has characteristic of a cis neutral state and have similar DAS with the first microsecond species. For V151A, this extra time constant could also be assigned to the 1.76 µs time constant found in the TRUV-Vis D_2O experiments. The assignment of their nature will be done in the discussion section after TR-SFX.

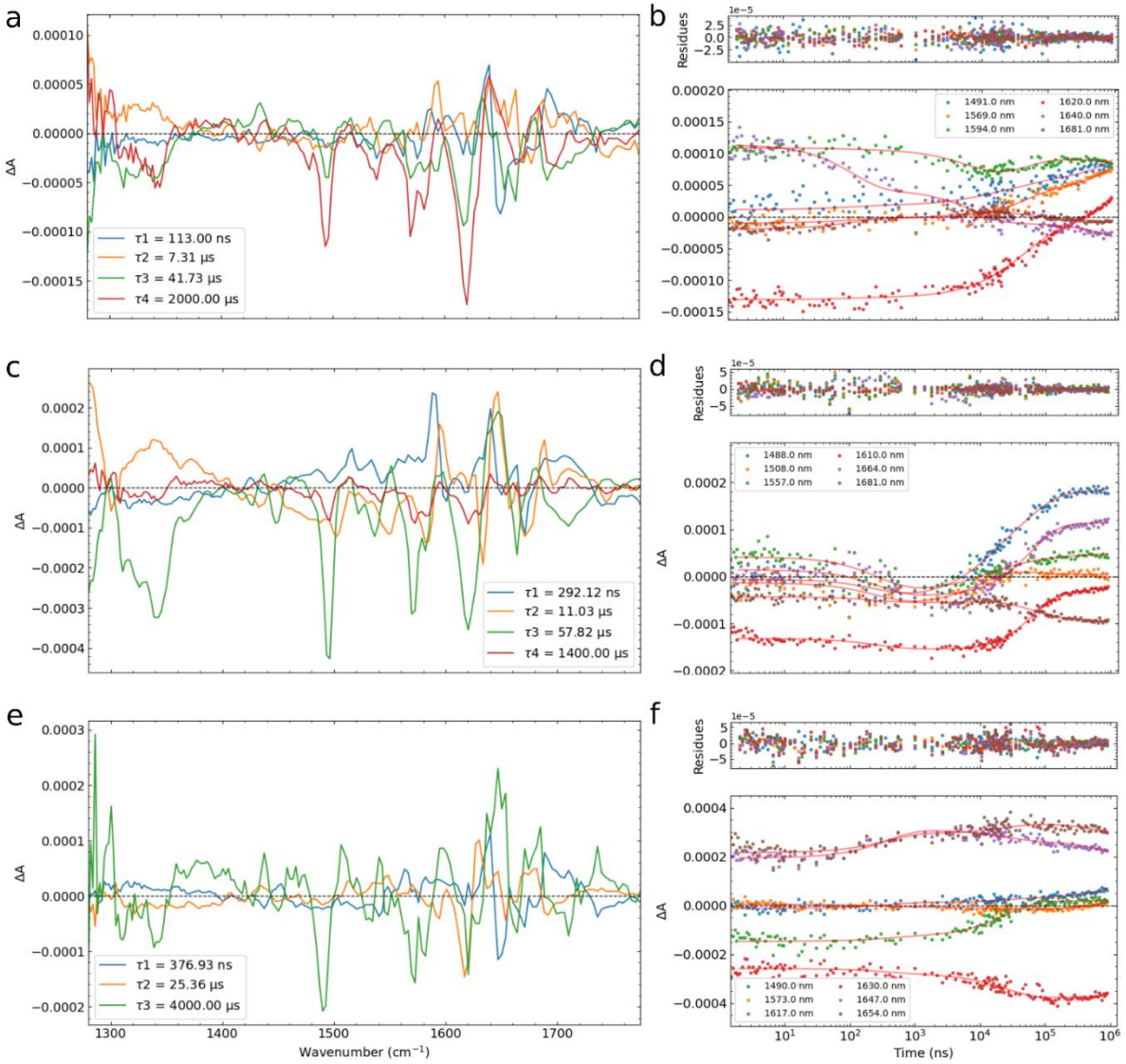


Figure 5.16. Left panels correspond to the decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis of all traces from 2 ns to 1 ms in the data sets displayed in Figure 5.14 with a weighted sum of four exponential function. In the right panels, the selected traces representing each of the transient absorption spectral main features (photoproducts absorption and GSB) together with the corresponding fits and residues can be seen. Notice that the time axis is set in a logarithmic scale. Panel a and b correspond to WT variant, c and d to V151L and e and f to V151A. Figure made with Ultra Pyfit.

5.3.3 TR-SFX structures at 10 ns

In this section, the TR-SFX results from the group of M. Weik in collaboration with I. Schlichting are reported. The WT structure at 10 ns after excitation is published and part of this PhD (figure 5.17)[5]. The time delay was chosen considering the TRUV-Vis results reported above, i.e. at 10 ns the structure should confirm the existence of cis neutral form and reveal the difference with *On* form helping in the assignation of the microsecond species. Two important results were reported:

- The *Off* state structure determined from TR-SFX data, features the chromophore in two *trans* conformations assigned to HT (trans2, 25% occupancy) and OBF (trans 1, 75% occupancy, Figure 5.17a), together with a remaining *On* form.
- The 10-ns structure, reveals that the side chain of His149 remains in an *Off*-like conformation while the chromophore has already isomerized to a *cis* form. Thus, the transition to the final *On* state conformation of His149 will occur on a timescale longer than 10 ns **(**Figure 5.17b**).**

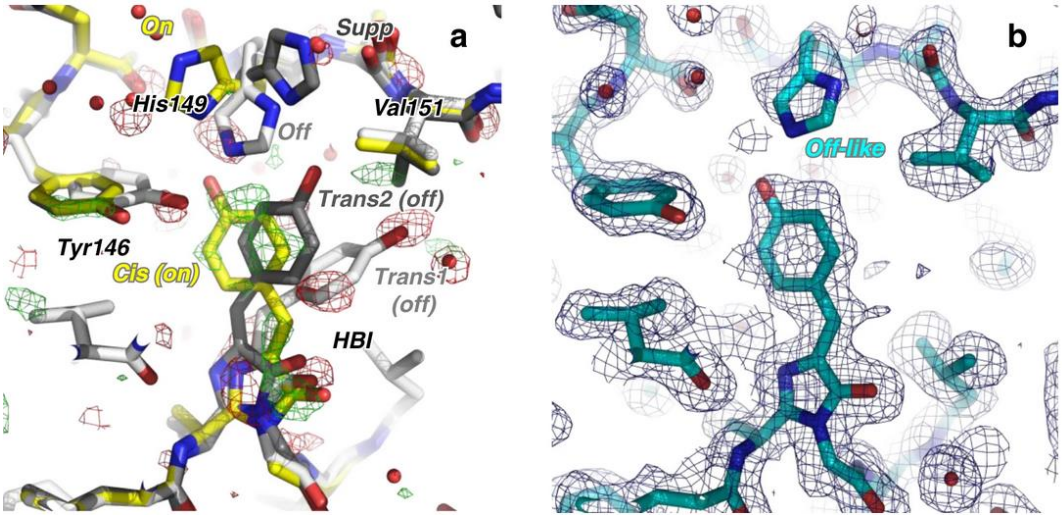


Figure 5.17. a) Q-weighted difference electron density map ($F_{obs}^{laser\_on\_\Delta 10ns}$–$F_{obs}^{laser\_off}$), determined from TR-SFX data with and without pump-laser activation, is contoured at +3σ (green) and −3σ (red) and overlaid onto the model determined from the laser_Off dataset. ***b*** *Model of the laser_on_Δ10 ns intermediate structure (cyan) determined by difference refinement at 1.85 Å resolution. $2F_{extrapolated}^{laser\_on\_\Delta 10ns}$–$F_{calc}$ (blue, 1σ) and $F_{extrapolated}^{laser\_on\_\Delta 10ns}$–$F_{calc}$ maps (green/red, ±3σ, respectively) are shown. Figure reproduced from reference [5].*

Time-resolved serial femto-second crystallography data using pump-probe[10] on V151A and V151L microcrystals are new data acquired on the BL2 – H3 experimental station of SACLA[11] (SACLA 2018A8026, 27- 29 July 2018). The samples were injected into the helium-filled Diverse Application Platform for hard X-ray Diffraction in SACLA (DAPHNIS)[12] using a Gas Dynamic Virtual Nozzle[13]. The X-ray beam (probe) was focused to 1.3 μm (H) × 1.4 μm (V) (FWHM) with nominal photon energy 7.6 keV, pulse length 10 fs at a repetition rate of 30 Hz. The energy pump laser was set to 2 μJ and 5 μJ for V151A and -V151L data collection, respectively. The pump laser parameters for both data collections were set as follow: 400 nm of wavelength with circular polarization, 180 μm (FWHM) of spot size and 6 ns of pulse duration at a repetition rate of 15 Hz. The pump-probe delay was at 10 ns. To switch the microcrystals from the resting *On* state to the *Off* state, there were pre-illuminated prior to the injection, using a 488 nm at 200 mW laser[14]. The efficiency of the pre-illumination was estimated at 80% and 75% for V151A and V151L microcrystals, respectively (chapter 4). There are two important results: at 10 ns the chromophore has isomerized to final *cis* form and similarly to the reported results for WT, the His149 has not occupied the final *On* like position.

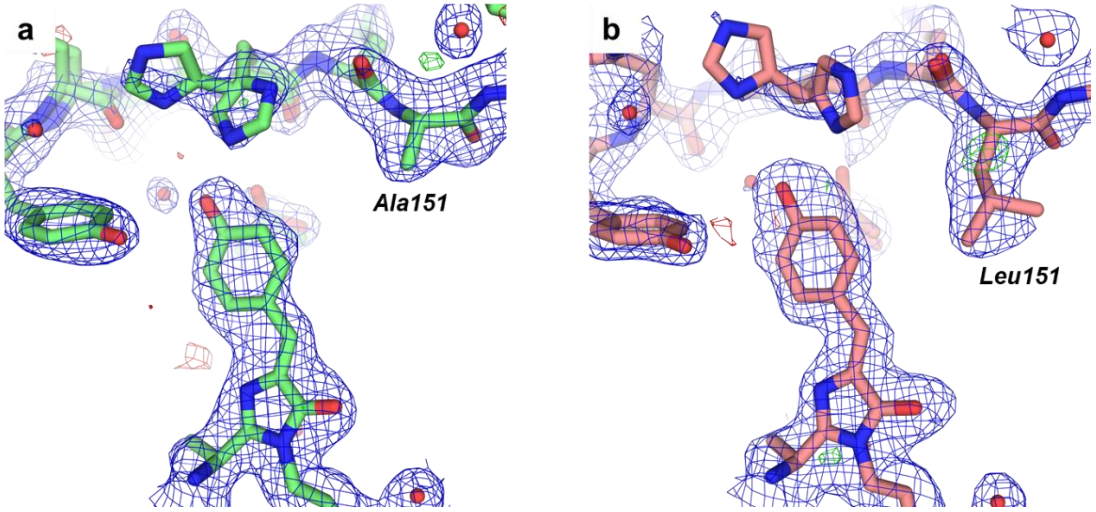


Figure 5.18: Models of the laser_on_Δ10 ns intermediate structures of rsEGP2-V15A ***(a)*** *and –V151L* ***(b)*** *in light green and in salmon, respectively. The models were determined by difference refinement at 1.95 and 2.1 Å resolution for rsEGP2-V15A and –V151L, respectively.* $2F_{extrapolated}^{laser\_on\_\Delta 10ns}–F_{calc}$ *(blue, 1σ) and* $F_{extrapolated}^{laser\_on\_\Delta 10ns}–F_{calc}$ *maps (green, +3σ and red,−3σ) are shown. Figure made by Kyprianos Hadjidemetriou unpublished.*

The refinement of the structure are still under progress by Kyprianos Hadjidemetriou (Doctoral Student in the group M. Weik). The members of the team that participate in the XFEL bem time SACLA 2018A8026, 27- 29 July 2018 are: Kyprianos. Hadjidemetriou, Elena Andreeva, Thomas Barends, Marco Cammarata, Jacques-Philippe Colletier, Nicolas Coquelle, Yasumasa Joti, R. Bruce Doak, Lutz Foucar, Alexander Gorel, Marie Gruenbein, Mario Hilpert, Marco Klooos, Gabriela Nass-Kovac, Shigeki Owada, Christopher Roome, Giorgio Schiro, Ilme Schilchting, Michel Sliwa, Miriam Stricker, Robert L. Shoeman, Kensuke Tono, Lucas. M. Uriarte, Joyce Woodhouse, Daehyun You and Martin Weik.

5.4 *Off* to *On* photodynamical scheme

As discussed in the introduction, after *On* to *Off* photoisomerization, the rsEGFP2 chromophore can have different *Off* states photoproducts isomers. These isomers are compatible with either an OBF or an HT isomerizations[1]. Preliminary TR-SFX results (previous section) from our collaborators confirm that in certain conditions they could observe both photoproduct in the *Off* state for WT in microcrystals. It has also been mentioned that the excited-state dynamics of rsEGFP2 were published by Coquelle et al., where they proposed an excited state isomerization[2]. On the contrary recent results in Dronpa and Dronpa2 by Laptenok et al.[4], showed, a *cis-trans* isomerization precursor in the excited state which ended in the ground state in 91 ns. Our transient spectroscopic data between 0 to 10 ms for WT, V151A and V151L are consistent with a trans-to-cis isomerization within the picosecond excited-state deactivation, followed by some protein rearrangements and ending in the ground-state proton transfer on a slower timescale, in line with previous studies on other RSFP[2,8,15,16]. Different results, indeed, confirm the absence of excited-state proton transfer (ESPT):

- No isotopic effect could be observed for the lifetime of excited states species in TRUV-Vis measurements performed in D_2O compared with H_2O buffers.
- By the time excited state has decayed to the ground state (time delay > 10 ps), the transient absorption spectra have no characteristic band of the phenolate form in the

UV-Vis (positive band around 480 nm Figure 4.2, TRUV-Vis result Figure 5.1) and the TRIR data (positive peak 1490 cm^{-1}, Figure 5.8 and Figure 5S-5).

- The difference of molar absorption coefficient between the spectra of *cis* neutral form minus *trans* neutral form are quite similar to UV-Vis transient spectra at 20 µs, 5 µs, and 50 µs for WT, V151L and V151A respectively (Figure 5.19). This indicates that at this points the only step to reach the final *cis* anionic form is the deprotonation of the chromophore.
- Isotopic effects (D_2O/H_2O) is found for UV-Vis kinetic traces in the µs-ms range, which proved that ground state proton transfer (GSPT) is the last step in the photo-switching.

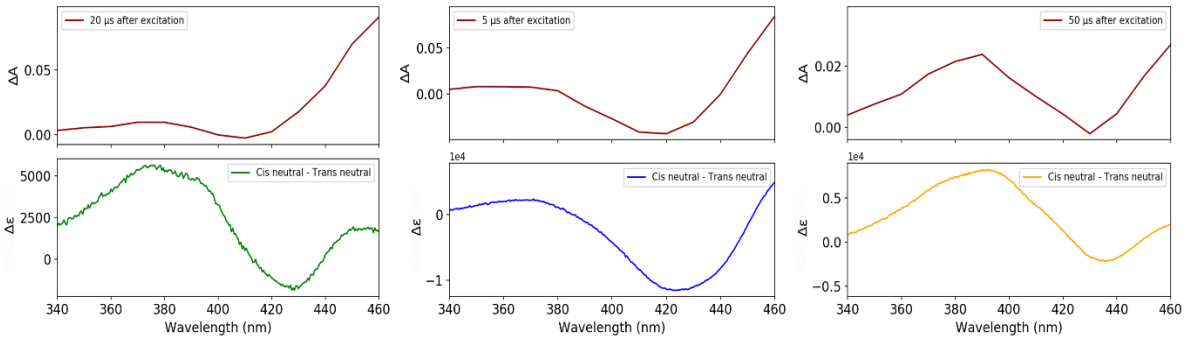


Figure 5.19. Top panel selected spectra after excitation of the trans Off-form that resemble the most to the difference molar absorption spectra of the On form at pH4.5 minus the trans Off spectrum at pH 8 (represented in Chapter 4 Figure 4.5). WT (a), V151L (b) and V151A (c). Figure made with Matplotlib python package.

Consequently, the *cis-trans* isomerization and protein rearrangements should have taken place at earlier times. TR-SFX experiments reported by Coquelle et al[2] attributed the isomerization to the deactivation of the excited state in the picosecond time scale. Nevertheless, several questions remained elusive. These are: i. the dynamics observed in crystal and in solution were not compared. ii. although they reported at 3 ps in TR-SFX the occupancy of a *cis* isomer, it was not clear if this *cis* isomer could be a twisted intermediate ground state or directly the cis *On* like form. iii. the nature of the isomerization mechanism starting from HT or OBF *Off* state. Moreover, the results obtained by Coquelle et al.[2] were based under the assumption that only one conformer exist in the Off ground state, which is not the case for WT. Contrary, V151A and V151L *Off* states are formed by either only OBF

or HT *Off* forms, respectively. The results obtained from the TR spectroscopy studies of these two variants compared to WT will help to clarify these open questions.

Dynamics between 0 and 40 ps.

Our TRUV-Vis and TRIR results reveal the formation in ~150 fs of two excited states. The first has a sub-picosecond lifetime and the second few picoseconds lifetime. The sub-picosecond excited state component DAS is (0.89 ps for WT, 0.70 ps for V151L and 0.72 ps for V151A; Figure 5.2) characterized by an SE band with a red tail (V151L) or a minimum at 680 nm (V151A, WT). Such a red-shifted SE band with large Stokes shift that grows with some band shift evolution from the FC state is assigned to a twisted excited state. The second excited state component has a DAS (4.20 ps for V151L, 4.33 ps for WT and 4.74 ps for V151A) characterized by a SE band with a minimum between 515-520 nm, and is assigned to a planar excited state that has a similar geometry to the ground state. The comparison of the ESA band at 450 nm and GSB at 390 nm (Figure 5.2) reveals that for V151L, the long-lived excited state is the major contribution, while for V151A it is the sub-picosecond one. Interestingly, for WT both excited states have almost the same contributions. The contributions calculated from the DAS in Figure 5.2 at 385 nm (GSB) can summarized in Table 5.2 for WT, V151L and V151A. Similarly, two excited states are observed in TR-IR experiments. At 1 ps the transient IR spectra for WT and V151A (Figure 5.7) display a positive peak at 1650 cm^{-1}. This band is present until 900 µs, by comparison with the *cis* anionic bands in the *Off* minus *On* FTIR spectrum (Figure A1.17 in Appendix 1), is assigned to a *cis* form. Moreover, after the decay of the excited state, the transient spectrum at 10 ps for WT, V151A and 20 ps for V151L is characterized by a positive band peaking around 1691 cm-1 (1685 cm^{-1} for V151L) near the main GSB band peaking at 1681 cm^{-1} (stretching C=O trans chromophore). The band at 1650 cm^{-1} has been attributed to the exocyclic C=C double bond stretching for cis o-HBDI[17] where a red shift of the C=O stretching of the imidazolidinone moiety is also observed. Moreover, the positive band at 1691 cm^{-1} has been attributed in Dronpa, and Dronpa2 also to C=O stretching of the chromophore in *cis* conformation and considered the primary *trans-cis* isomerization signature. The second time constant found in the TRIR experiments mainly is governed by the recovery of the GSB.

Table 5.2 Contributions of the different excited state species calculated using the variation of absorbance from the DAS spectra GSB at 385 nm (Figure 5.2).

Component	*WT τ(ps) / ΔA %*	*V151L τ(ps) / ΔA %*	*V151A τ(ps) / ΔA%*
τ1	0.89 / 65.7	0.70 / 57.0	0.72 / 87.4
τ2	4.33 / 34.3	4.20 / 43.0	4.74 / 12.6

Considering the red-shifted SE band of the sub-picosecond time constant and the formation of clear features of the *cis* isomer after the decay of the excited state we attribute the first excited state species to a twisted intermediate state that goes under isomerization (as previously reported by Coquelle et al[2]). Concomitantly with the decay of the short-excited state component an important recovery of the GSB (e.g. V151A) is observed, which is in agreement with the photo-isomerization yields ~10-14%. Therefore, the isomerization occurs through a conical intersection where part of the twisted excited state evolves to the original *trans* neutral conformer and the rest to a *cis* neutral chromophore. The second excited state is assigned to a planar excited species that goes back to the original trans neutral chromophore (100 %). The results are in agreement with "two-state two-coordinate" model previously proposed for bR[18] or the HBDI[19]. This model supposes the initial motion away from the Franck-Condon excited state is along a stretching coordinate, in agreement with the high number of stretching modes in TRIR, which provoke an energy redistribution between the different vibrational modes followed by a barrier-controlled evolution of the, at least partially equilibrated, twisted S_1 species toward CI for the subpicosecond time constant or towards a planar minimum in excited-state potential energy surface. These two different excited states (planar and twisted) were modelled for WT[2].

Altogether the isomerization dyanmics seems to be the same for V151A and V151L. However, the *Off* to *On* photoswitching mechanism for V151A and V151L starts from different trans neutral forms originated from an OBF and HT *On* to *Off* photoproducts. Therefore, one could expect that V151A will follow an OBF mechanism and V151L an HT mechanism. The differences between OBF and HT isomerization have been discussed in

Chapter 2. It has been reported that the isomerization time constant for an OBF is longer than for an HT[20,21-22]. This was shown for example, for stiff-stilbene, -a stilbene analogue that can only isomerize via an OBF mechanism- where the isomerization occurs in about 800 fs, which is 5 to 6 times longer compared to the isomerization of single-block and free stilbenes, both having similar characteristic isomerization times around 135 fs. For the latest, it has been shown that their isomerization occurs via HT[20,21]. Similar results can also be seen for cyanine dyes[22]. Therefore, if the isomerization of WT and V151A were occurring via an OBF mechanism, one would expect longer lifetime for the excited-state at the origin of the isomerization, compare to the V151L, which can only isomerize via a HT pathway due to the constrained environment. However as discussed above the time constants associated with the isomerization are nearly identical, 0.88, 0.70 and 0.72 ps for WT, V151L and V151A respectively, this result is in line with the three proteins sharing the same isomerization mechanism. Moreover, since we show here that the excited state dynamics are similar in crystal and solution (Figure 5.8), it is possible to validate the proposed trans to cis isomerization in the excited state by Coquelle et al.[2]. The authors reported a twisted excited chromophore identified at 1 ps (TR-SFX), halfway between the trans and cis isomers. The authors proposed the simultaneous rotation of the double bond and the adjacent single bond, thus suggested a HT isomerization pathway for the OBF photoproduct and confirmed by QM/MM calculations[2].

Dynamics between 40 ps and 10 ns.

From the aforementioned results, the photoproduct at 40 ps is attributed to a chromophore in *cis* form originated from an HT isomerization. For WT and V151A, a spectral evolution is observed in UV-Vis transient absorption spectra with 87 ps and 83 ps time constants. For V151L no spectral evolution is observed from 40 ps until 10 ns. The *Off* state X-ray structures of WT, V151A and V151L (Chapter 3 Figure 3.21) reveal that for WT and V151A *Off* forms, the hydrogen bond between the chromophore and the His-149 in the *On* state is broken and Tyr-146 relocates to establish a hydrogen bond with the His-149. The 10 ns X-ray structures for WT and V151A obtained in TR-SFX experiments (Figure 5.17 and 5.18) show that the Tyr-146 and the chromophore already occupy their respective final *On* state position.

Therefore, we attribute the first ground state evolution seen for WT (87 ps) and V151A (83 ps), which is absent for V151L, to the movement of the Tyr-146 which liberates a free space in the protein pocket cell allowing the relaxation of the chromophore via single bond rotation around Φ to the final *cis On* state position. The overall mechanism for WT and V151A is compatible with an aborted HT (Chapter 2). This mechanism was firstly described by Fuss and coworkers for several conjugated molecules[23] and consist of HT isomerization through a CI which further evolve via single bond rotation in the ground state forming the OBF photoproducts. Our results show that the chromophore isomerization is coupled to motions of the surrounding β-barrel, and explains the nearly identical isomerization times for WT, V151A and V151L.

Dynamics between 10 ns and 10 ms

After these evolutions in the ps range, we were unable to identify any other spectral evolution in the TRUV-Vis data until few µs. Contrary, the IR transient spectra (Figure 5.13 TRMPS) evolutions needed an extra-time constant of a few hundreds of ns to be correctly fitted. This was the case for the three variants with associated times of 113 ns, 292 ns and 376 ns for WT, V151L and V151A respectively. Although, a time constant was needed to correctly describe the data spectrally the evolutions seen in the IR are very small and mainly involve changes in the signal amplitude of the most prominent peaks. Except for V151A where a 0.96 µs component was observed (attributed to the 376 ns TRMPS component), the following evolutions are in the µs time scale where no isotopic effect was found, with characteristic times of 5.43 µs, 4.03 µs and 18.50 µs for WT V151L and V151A respectively. The 10 ns TR-SFX X-ray structure reveals that at 10 ns the His149 has not its final position. Moreover, since the movement of the His-149 is expected to occur before the deprotonation (discussed in Figure 5.19), the movement of the His149 could be attributed to either of these two decay times. Nevertheless, the absence of spectral changes in the UV-Vis data for the first evolution and the fact that the changes in the IR spectra only involve variations in the peaks intensity reveal rearrangements that provoke very small or no changes at all in the chromophore environment. Thus, since the His149 is hydrogen-bonded to the chromophore in the final *On* state, having a direct interaction with the HBDI, we attribute the last evolutions which have

a clear influence on the UV spectra to the movement of the His149. The decay component occurring in few hundreds of nanoseconds is within the time scale of the protein domain movements and thus is tentatively attributed to some relaxation of the β-barrel, especially to the β-strands 10, 7 and 8, which as shown by NMR studies, are those that suffer major distortion in between *On* and O*ff* states in rsFolder[24]. Moreover, the relaxation of this similar β- strands in Dronpa and some of its variants induce a broadening of the NMR signal occurring on a shorter time scale than NMR relaxation time[25] thus in agreement with relaxations on the ns time scale. Probably, the stabilization of the protein cage favours the movement of the His149 in the μs range.

Finally, the occupation of the final *On* like position for the chromophore and the surrounding amino acids triggers the chromophore deprotonation in the microsecond range with time constants of 36.1 and 825 μs for WT and, 197 μs and 1142 μs for V151A, where the main contribution is the longer one. For V151L two time constants of 18.7 μs and 631 μs were found with the main contribution being the shortest component. This results together with the V151L rich evolutions in TRIR with the appearance of characteristic *On* bands peaking at 1670 cm^{-1}, 1568, and 1615 cm^{-1}, already present for WT and V151A at earlier times, suggest that the His-149 accompanies the chromophore along the isomerization. For all the steps attributed to the chromophore deprotonations in the three proteins, substantial isotopic effects have been observed, supporting our assignations. Moreover, the presence of isosbestic points in the spectra evolutions in Figure 5.10, demonstrate that the ground state evolutions occur sequentially.

Comparison with Dronpa switching mechanism

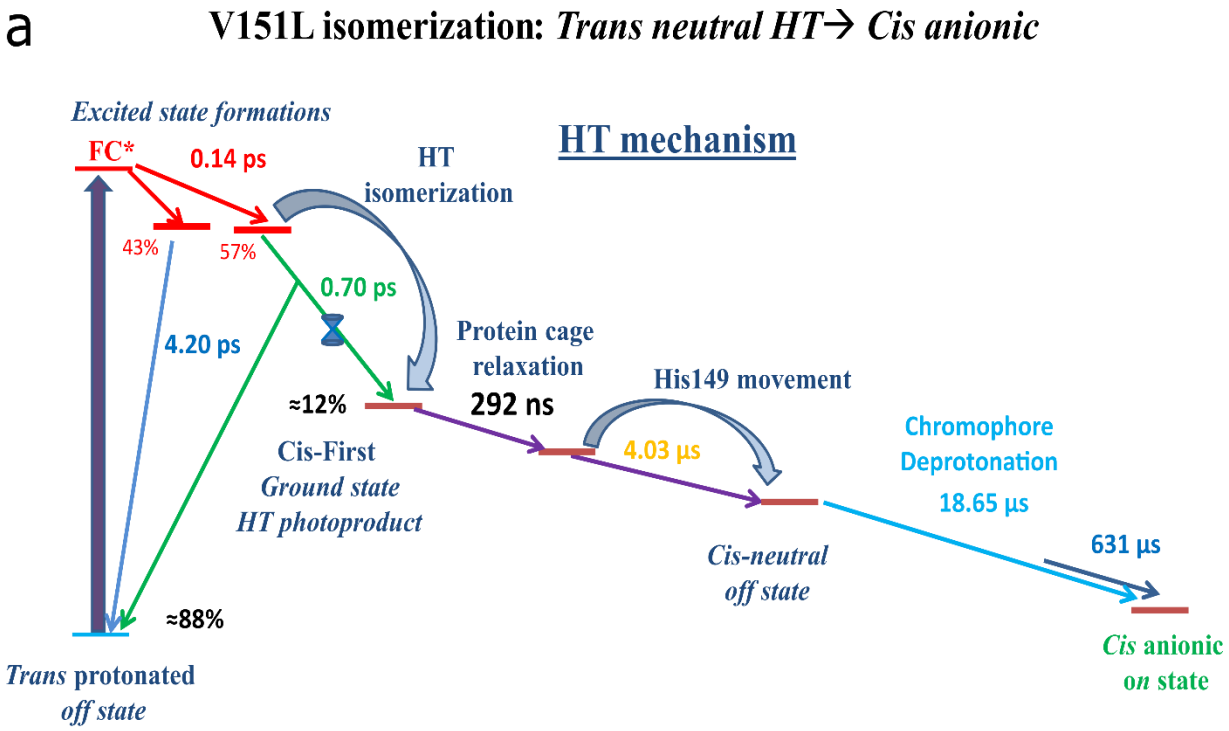
The mechanism observed for the three proteins differs from that observed in Dronpa. In previous studies, it has been demonstrated that the primary signal signature of *cis* to *trans* isomerization in Dronpa and Dronpa2 is a shift of the C=O chromophore stretching band, which is red-shifted to higher values for the *cis* compared to the *trans* isomer (Laptenok et al. [4]). For Dronpa2 this signature appeared in ns range (91 ns), while contrary, we can already see this shifts after few tens of ps for rsEGFP2 WT and variants. The photoisomerization

quantum yield is fixed by the excited state deactivation in 16 ps[4]. Therefore, since both proteins have the same chromophore, such a huge difference in time (ps vs ns) for the *cis* isomer signal between Dronpa and rsEGFP2 can only be attributed to the different protein cage rearrangements in the photoisomerization process of Dronpa compared to rsEGFP2. In Dronpa after the isomerization, the protein rearrangement provoked changes in the hydrogen bond networks of the carbonyl and phenyl groups of the chromophore. Contrary in rsEGFP2, the changes are exclusive to the phenyl group. Particularly important in Dronpa are the changes of Arg66, this amino acid has major rearrangements, and is found hydrogen-bonded to C=O carbonyl group of the chromophore imidazolinone in the *On* state but not in the *Off* state. The formation of a hydrogen bond between Arg66 and the chromophore will result in a shift of the C=O chromophore stretching band to higher values. This could explain the appearance of the shift in the C=O stretching in the nanosecond range, after the relaxation of the protein cage and the rearrangement of Arg66. It is worth mentioning that Laptenok et al.[4] founded a ground state protein rearrangement in around 596 ps. On the contrary, for rsEGFP2 only the hydrogen bonds of the phenyl group are different between *Off* and *On* states, but none of the hydrogen bonds of the carbonyl imidazolinone group is broken, assuring that any of the shifts in the C=O chromophore stretching band cannot be hindered by a hydrogen bond formation.

Global Off to On dynamics for V151A and V151L

The combination of TRUV-Vis with TRIR and TR-SFX experiments allows to propose a complete isomerization pathway for V151A and V151Lfrom either OBF or HT starting *Off* state conformers (Figure 5.18). Basically, both pathways start identically with an isomerization via HT. For the V151A (trans1) this is followed by Tyr146 movement to the final *On* position, which allows the rotation and movement of the chromophore to form the OBF photoproduct (aborted HT). Once the chromophore in both cases is in the final Cis neutral form as confirmed by the TR-SFX data at 10 ns, there is a relaxation of the β barrel favoring< the His149 to evolve to the final position which then triggers the final deprotonation of the chromophore.

a



b

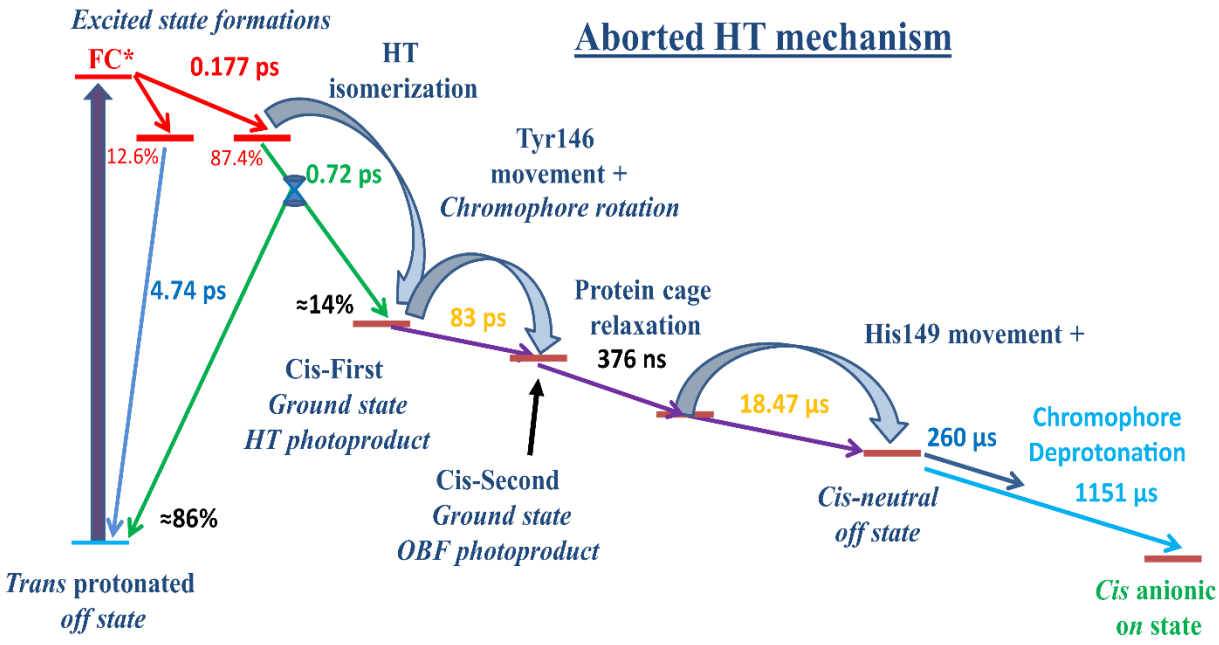


Figure 5.18. Main Off to On isomerization steps starting from: (a) hula twist and (b) One bond flip On to Off photoproducts for rsEGFP2 variants. Figure made with powerpoint.

5.5 Supporting Figures

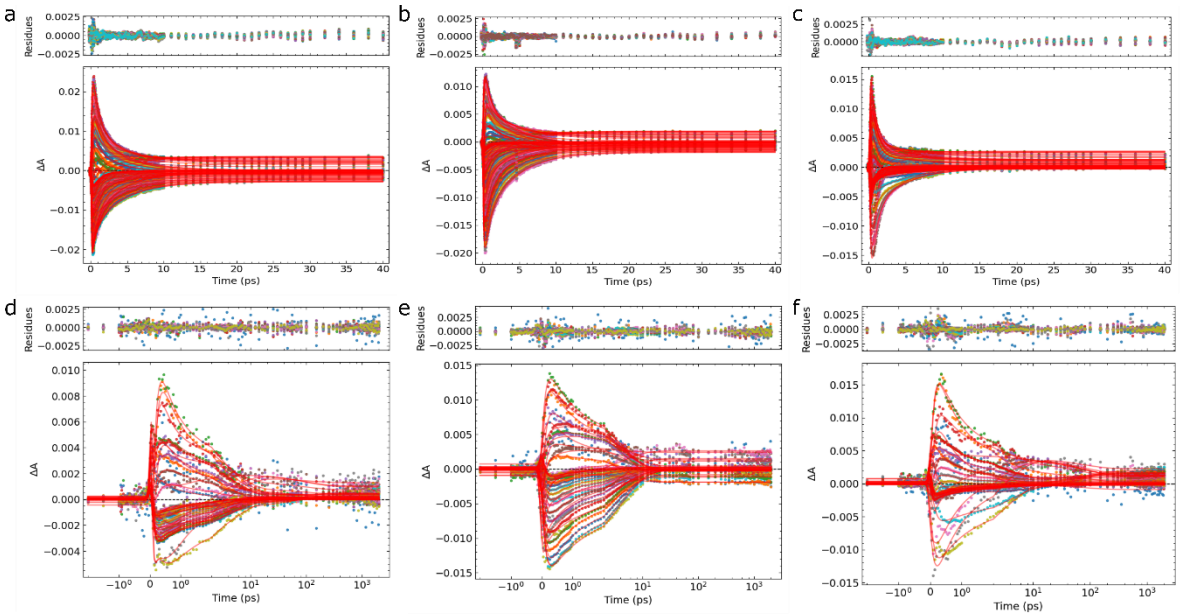


Figure 5S-1. Kinetic traces (every 10 nm) obtained from femtosecond TRUV-Vis experiment together with weighted sum of exponential function fit (convolved with a Gaussian function and a constant) with the corresponding residues. Top panels are the data sets measured in H_2O buffer (0-40 ps) and bottom panel are the data sets measured in D_2O buffer (0-2000 ps). Logarithmic scale is used from 2 ps. The panels a, d correspond to WT. b, e for V151L, and c, f for V151A. Figure made with Ultra Pyfit.

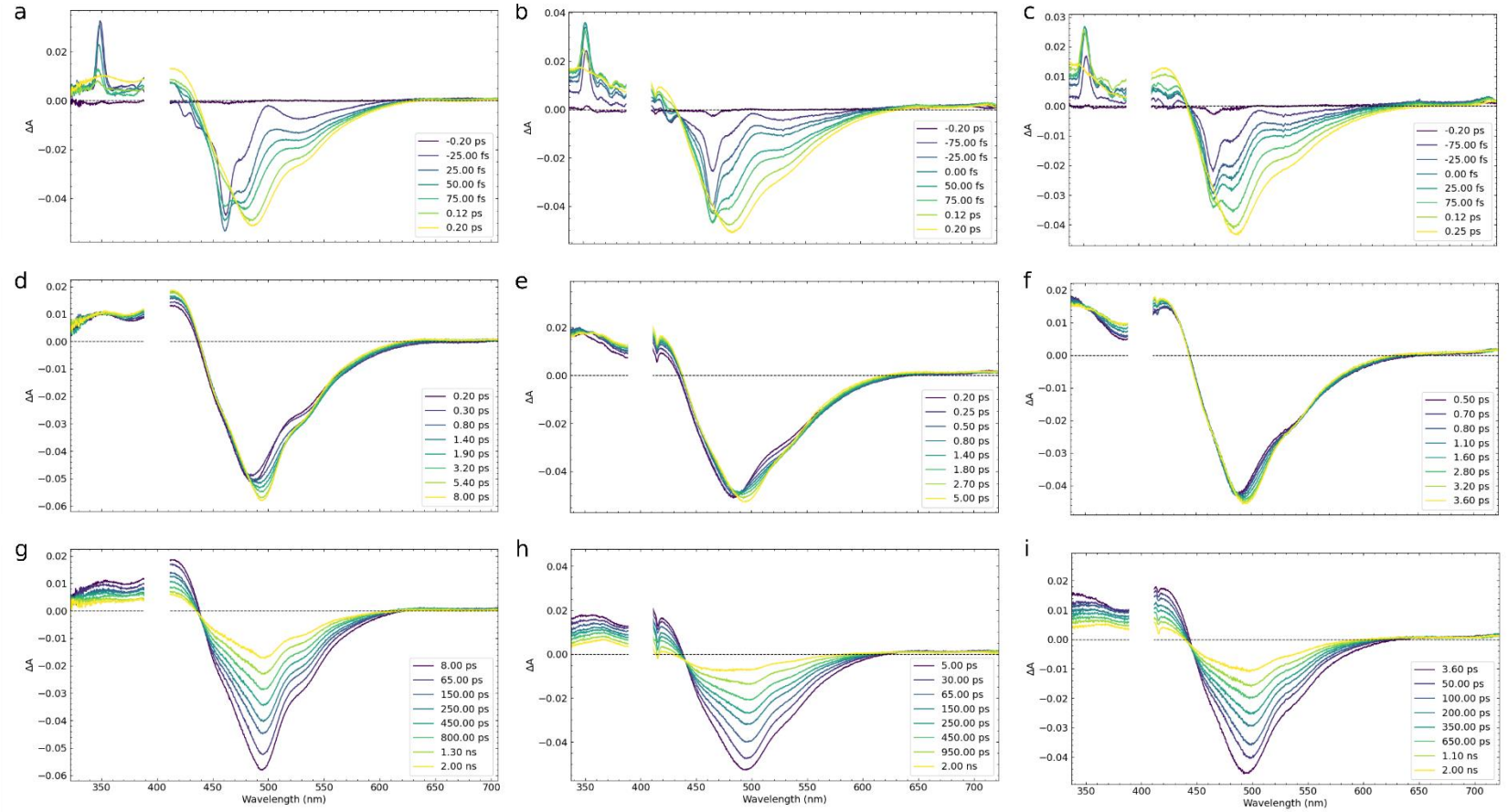


Figure 5S-2. Femtosecond transient difference absorption spectra recorded at different time delays after a femtosecond laser excitation (400 nm) of the cis anionic on state until to 2000 ps. The spectrum without laser excitation was subtracted to calculate the difference spectra. Top panels correspond to the signal formation. Middle panels correspond to a signal shift observed until 8 ps delay. Bottom panels signal decays until 2 ns. WT correspond to panels (a,d and g), V151L (b,e and h) and V151A (c, f and i). Figures made with Ultra PyFit.

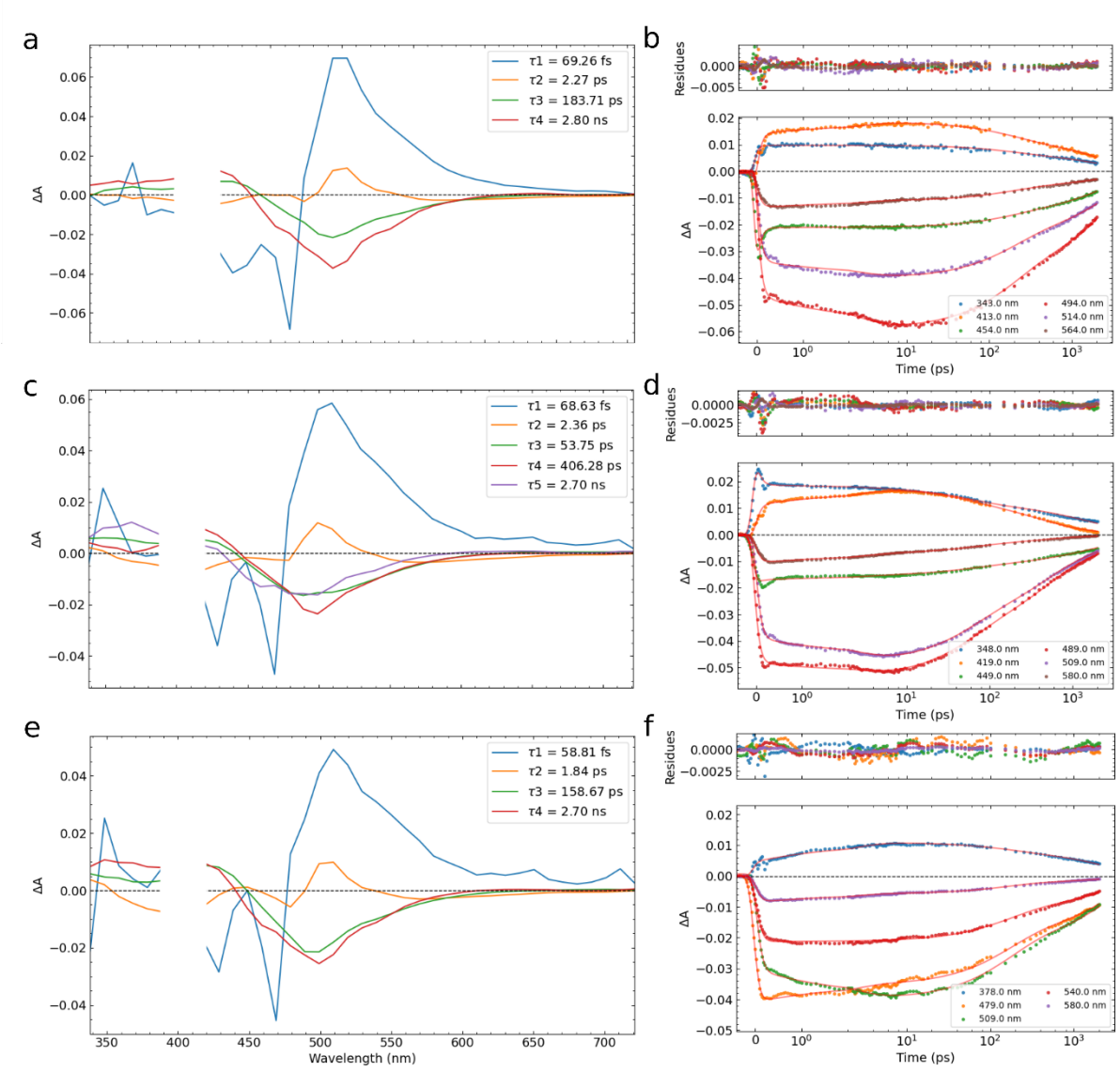


Figure 5S-3. Left panels correspond to the decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis of 1 decay traces every 5 nm form the data sets in Figure 5S-3 with weighted sum of exponential function convolved with a Gaussian function. In the right panels, the selected traces representing each of the transient absorption spectral main features (ESA, GSB and SE) together with the corresponding fits and residues can be seen. Panel a and b correspond to WT, c and d to V151A and e and f to V151A. Figure made with Ultra Pyfit.

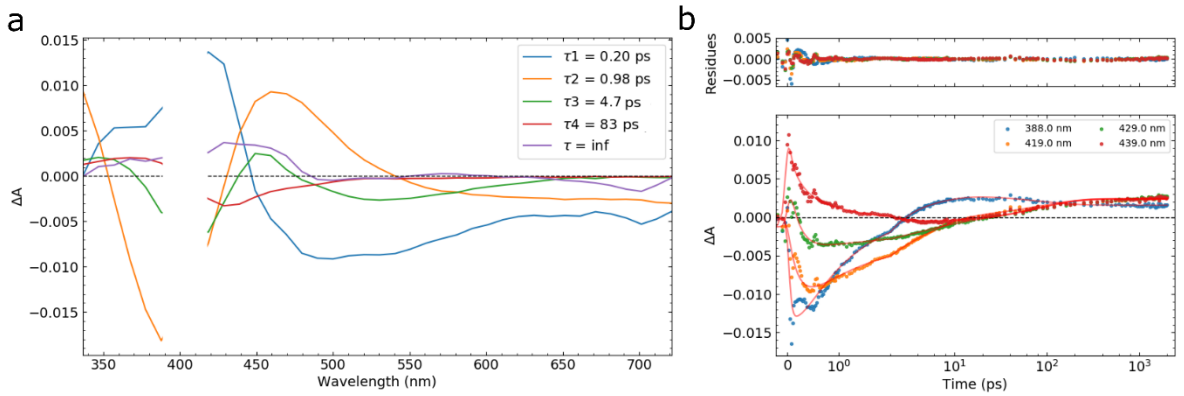


Figure 5S-4. V151A long spectral evolutions data set global analysis results. Left panels correspond to DAS reconstructed using the pre-exponential factors obtained from the global fit analysis of kinetic traces (every 10 nm) form the data sets in Figure 5.5c with a 4 exponential sum decay function convolved with a Gaussian function. In the right panels, the selected traces representing each of the transient absorption spectral main evolutions around 430 nm together with the corresponding fits and residues can be seen. Notice that the scale of the rigth panel is linear until 2 ps and logarithmic from 2 ps until 2 ns. Figure made with Ultra Pyfit.

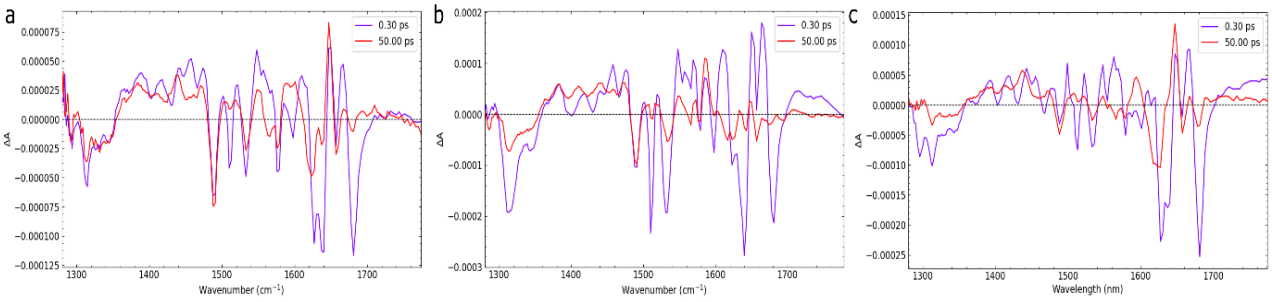


Figure 5S-5. Difference IR spectra recorded at 0.3 and 50 ps time delays time after a femtosecond laser excitation (400 nm) of the Off state. The spectrum without laser excitation was subtracted to calculate the difference spectra. WT (a), V151L (b) and V151A (c). Figure made with Matplotlib python package.

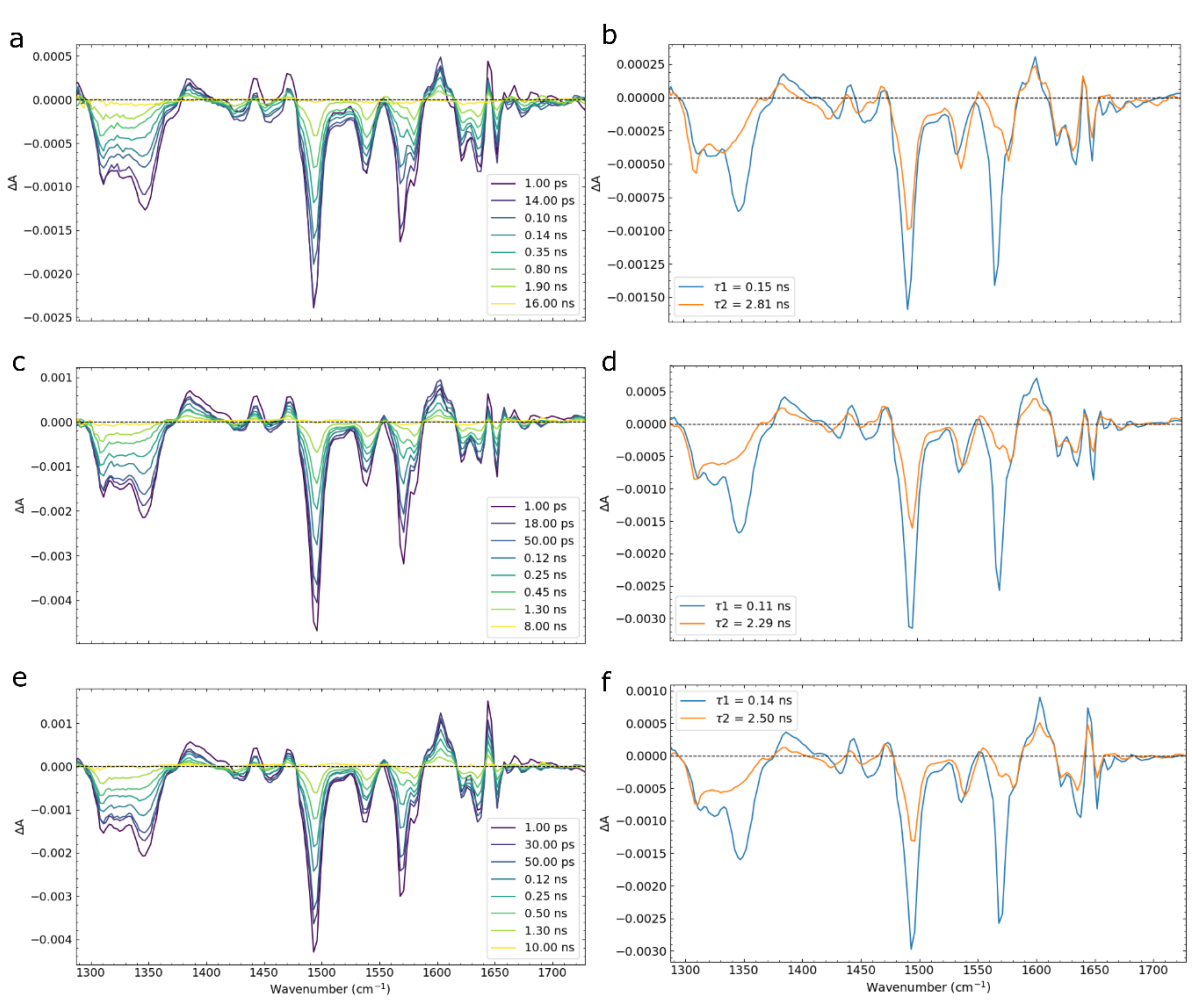


Figure 5S-6. rsEGFP2 variants in D_2O Transient absorption IR spectra recorded with TRMPS. Left graphs (a, c, d) correspond to the time-resolved difference absorption spectra recorded after a femtosecond laser excitation (482 nm) of the cis anionic On state until to 20 ns. The spectrum without laser excitation was subtracted to calculate the difference spectra. Right graphs (b, d, f) correspond to the decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis of 1 decay traces every 10 nm with a 2 exponential sum decay function convolved with a Gaussian function. WT (top panels: a, b), V151L (middle panels: c, d) and V151A (bottom panels: e, f). Figure made with Ultra Pyfit.

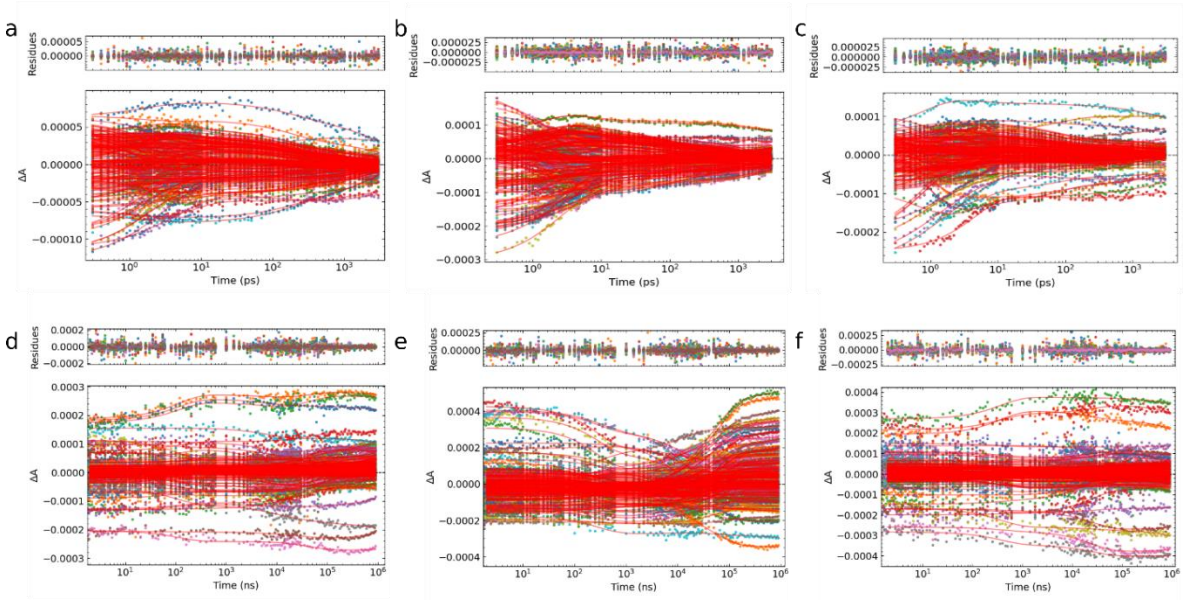


Figure 5S-7. Kinetic traces for IR experiments recorded in D_2O buffer together with the fit of weighted sum of exponential function and the corresponding residues. Top panels TRIR data from 0.3 until 3000 ps. Bottom panels TRMPS data from2 ns until 900 µs. The panels a, d correspond to WT. b, e for V151L. Finally, c, f for V151A. Figure made with Ultra Pyfit.

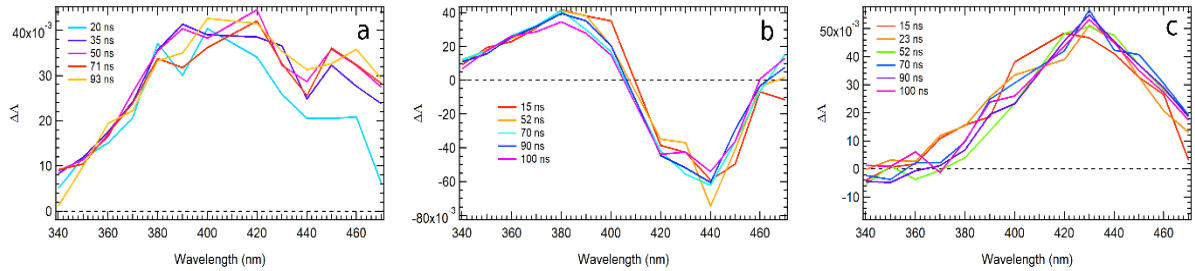


Figure 5S-8. Time-resolved UV-Vis difference absorption spectra recorded after a 410 nm nanosecond excitation of the trans protonated Off form between 10 ns and 100 ns, a b and c correspond to WT, V151L and V151A in H_2O solution (50 mM HEPES pH 8, 50 mM NaCl) respectively. Figure made with Igor pro.

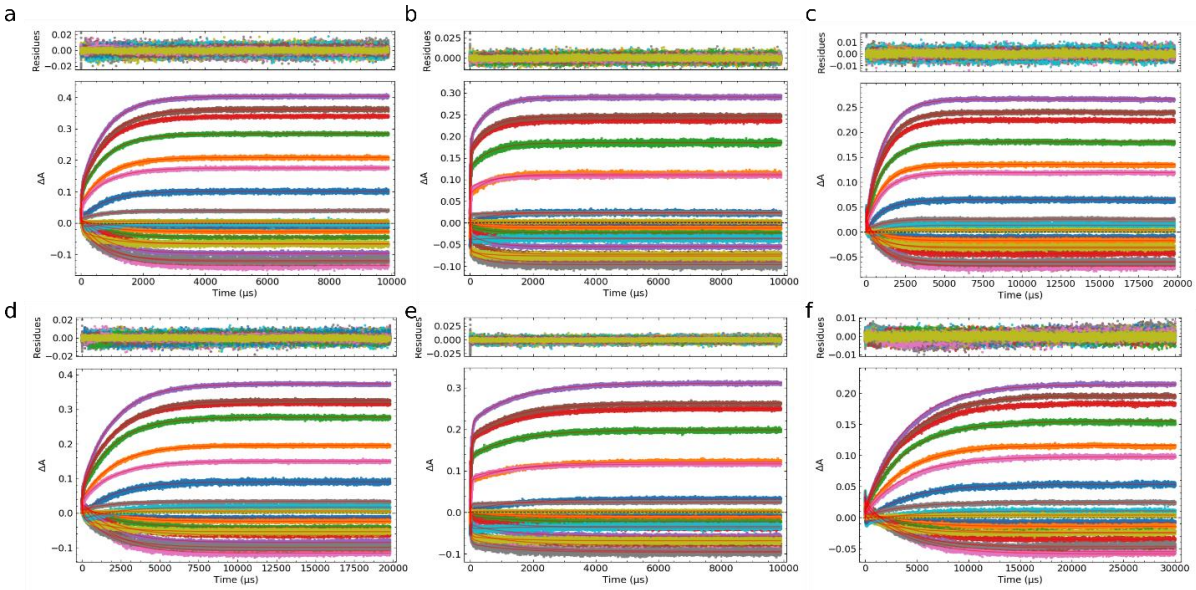


Figure 5S-9. Resulting decay traces from ns TRUV-Vis experiments between 10 ns until 30 ms together with the fit with a weighted sum of exponential function and the corresponding residues. Top panels are the data sets recorded in H_2O buffer. Bottom panels correspond to the data sets recorded in D_2O buffer. The panels a, d correspond to WT. b, e for V151L. Finally, c, f for V151A. Figure made with Ultra Pyfit.

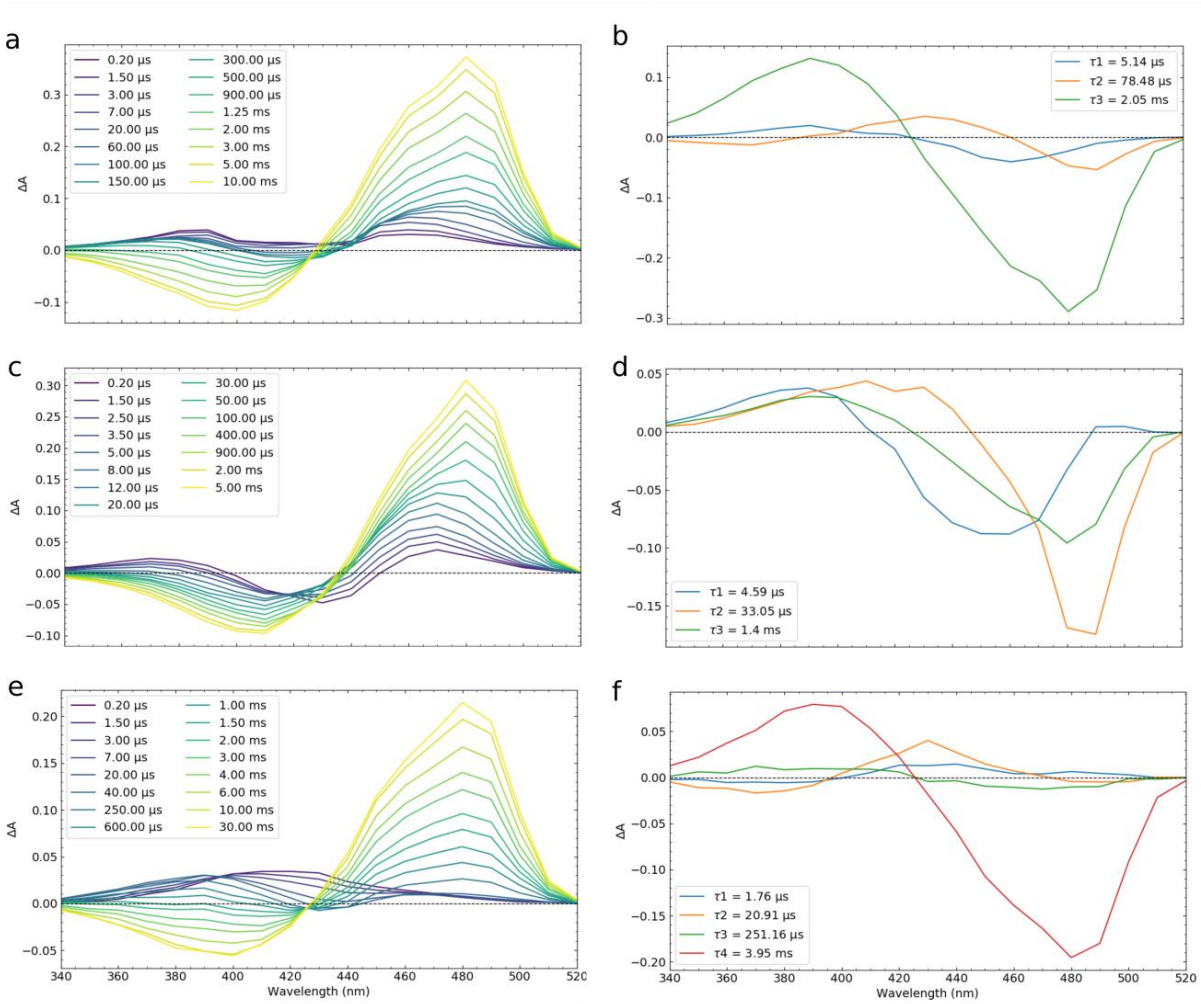


Figure 5S-10. rsEGFP2 variants ns TRUV-Vis in D_2O solution (50 mM HEPES pD 8, 50 mM NaCl). Left graphs (a, c, d) correspond to the time-resolved difference absorption spectra recorded after a 410 nm ns excitation of the Off state in the time windows from 100 ns to 10 ms. The spectrum without laser excitation was subtracted to calculate the difference spectra. Right graphs (b, d, f) correspond to the decay associated spectra obtained from global fitting the raw data with a 3 exponential decay function. WT (top panels: a, b), V151L (middle panels: c, d) and V151A (bottom panels: e, f). Figure made with Ultra Pyfit.

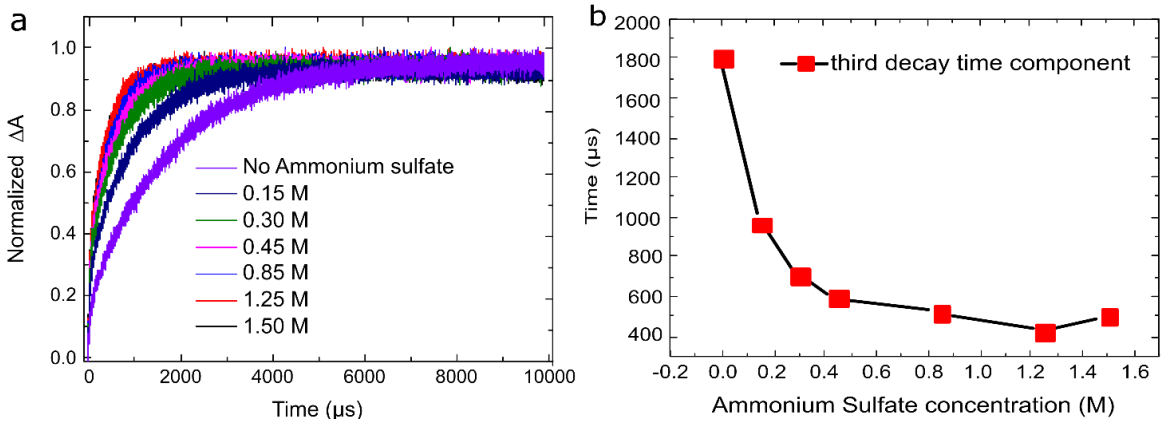


Figure 5S-11. a) Normalized recorded transient absorption time traces at 480 nm for rsEGFP2 in D_2O and different ammonium sulfate concentrations. b) Third decay time constants of a three exponential fit of traces in panel (a). Figure made with Origin Pro 8.

Table 5S-1. Time constants retrieved after a tri-exponential fit of the UV-Vis kinetic traces between 10 ns and 10 ms represented in Figure 5.13.

rsEGFP2 in:	**τ1 (μs)**	**τ2 (μs)**	**τ3 (μs)**
Buffer + NH_4Cl (0.4 M)	5.3	22.2	434
Buffer + $(NH_4)_2SO_4$ (0.2 M)	5.2	74.7	409
Buffer +Na_2SO_4 (0.2 M)	5.3	96.3	851
Buffer	5.2	33.2	820

* buffer contained 50 mM HEPES pH 8, 50 mM NaCl.

5.6 References

(1) Chang, J.; Romei, M. G.; Boxer, S. G. Structural Evidence of Photoisomerization Pathways in Fluorescent Proteins. *Journal of the American Chemical Society* **2019**, *141*, 15504–15508.

(2) Coquelle, N.; Sliwa, M.; Woodhouse, J.; Schirò, G.; Adam, V.; Aquila, A.; Barends, T. R. M.; Boutet, S.; Byrdin, M.; Carbajo, S.*; et al.* Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nature chemistry* **2018**, *10*, 31–37.

(3) Olsen, S.; Lamothe, K.; Martínez, T. J. Protonic gating of excited-state twisting and charge localization in GFP chromophores: a mechanistic hypothesis for reversible photoswitching. *Journal of the American Chemical Society* **2010**, *132*, 1192–1193.

(4) 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.*; et al.* Infrared spectroscopy reveals multi-step multi-timescale photoactivation in the photoconvertible protein archetype dronpa. *Nature chemistry* **2018**, *10*, 845–852.

(5) Woodhouse, J.; Nass Kovacs, G.; Coquelle, N.; Uriarte, L. M.; Adam, V.; Barends, T. R. M.; Byrdin, M.; La Mora, E. de; Bruce Doak, R.; Feliks, M.*; et al.* Photoswitching mechanism of a fluorescent protein revealed by time-resolved crystallography and transient absorption spectroscopy. *Nature communications* **2020**, *11*, 741.

(6) Debus, B.; Orio, M.; Rehault, J.; Burdzinski, G.; Ruckebusch, C.; Sliwa, M. Fusion of Ultraviolet-Visible and Infrared Transient Absorption Spectroscopy Data to Model Ultrafast Photoisomerization. *The journal of physical chemistry letters* **2017**, *8*, 3530–3535.

(7) Hamm, P.; Ohline, S. M.; Zinth, W. Vibrational cooling after ultrafast photoisomerization of azobenzene measured by femtosecond infrared spectroscopy. *The Journal of Chemical Physics* **1997**, *106*, 519–529.

(8) Colletier, J.-P.; Sliwa, M.; Gallat, F.-X.; Sugahara, M.; Guillon, V.; Schirò, G.; Coquelle, N.; Woodhouse, J.; Roux, L.; Gotthard, G.*; et al.* Serial Femtosecond Crystallography and Ultrafast Absorption Spectroscopy of the Photoswitchable Fluorescent Protein IrisFP. *The journal of physical chemistry letters* **2016**, *7*, 882–887.

(9) Greetham, G. M.; Sole, D.; Clark, I. P.; Parker, A. W.; Pollard, M. R.; Towrie, M. Time-resolved multiple probe spectroscopy. *The Review of scientific instruments* **2012**, *83*, 103107.

(10) Colletier, J.-P.; Schirò, G.; Weik, M. Time-Resolved Serial Femtosecond Crystallography, Towards Molecular Movies of Biomolecules in Action. In *X-ray Free Electron Lasers*; Boutet, S., Fromme, P., Hunter, M. S., Eds.; Springer International Publishing: Cham, 2018; pp 331–356.

(11) Yabashi, M.; Tanaka, H.; Ishikawa, T. Overview of the SACLA facility. *Journal of synchrotron radiation* **2015**, *22*, 477–484.

(12) Tono, K.; Nango, E.; Sugahara, M.; Song, C.; Park, J.; Tanaka, T.; Tanaka, R.; Joti, Y.; Kameshima, T.; Ono, S.*; 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* **2015**, *22*, 532–537.

(13) Weierstall, U.; Spence, J. C. H.; Doak, R. B. Injector for scattering measurements on fully solvated biospecies. *The Review of scientific instruments* **2012**, *83*, 35108.

(14) Schirò, G.; Woodhouse, J.; Weik, M.; Schlichting, I.; Shoeman, R. L. Simple and efficient system for photoconverting light-sensitive proteins in serial crystallography experiments. *J Appl Crystallogr* **2017**, *50*, 932–939.
(15) 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. *The journal of physical chemistry. B* **2015**, *119*, 2404–2414.
(16) 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. *Nature communications* **2013**, *4*, 1461.
(17) Hsieh, C.-C.; Chou, P.-T.; Shih, C.-W.; Chuang, W.-T.; Chung, M.-W.; Lee, J.; Joo, T. Comprehensive studies on an overall proton transfer cycle of the ortho-green fluorescent protein chromophore. *Journal of the American Chemical Society* **2011**, *133*, 2932–2943.
(18) González-Luque, R.; Garavelli, M.; Bernardi, F.; Merchán, M.; Robb, M. A.; Olivucci, M. Computational evidence in favor of a two-state, two-mode model of the retinal chromophore photoisomerization. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 9379–9384.
(19) Mandal, D.; Tahara, T.; Meech, S. R. Excited-State Dynamics in the Green Fluorescent Protein Chromophore. *J. Phys. Chem. B* **2004**, *108*, 1102–1108.
(20) Fuss, W.; Kosmidis, C.; Schmid, W. E.; Trushin, S. A. The photochemical cis-trans isomerization of free stilbene molecules follows a hula-twist pathway. *Angewandte Chemie (International ed. in English)* **2004**, *43*, 4178–4182.
(21) Imamoto, Y.; Kuroda, T.; Kataoka, M.; Shevyakov, S.; Krishnamoorthy, G.; Liu, R. S. H. Photoisomerization by hula twist: 2,2'-dimethylstilbene and a ring-fused analogue. *Angewandte Chemie (International ed. in English)* **2003**, *42*, 3630–3633.
(22) van der Meer, M.; Zhang, H.; Rettig, W.; Glasbeek, M. Femto- and picosecond fluorescence studies of solvation and non-radiative deactivation of ionic styryl dyes in liquid solution. *Chemical Physics Letters* **2000**, *320*, 673–680.
(23) Sampedro Ruiz, D.; Cembran, A.; Garavelli, M.; Olivucci, M.; Fuß, W. Structure of the Conical Intersections Driving the cis–trans Photoisomerization of Conjugated Molecules¶. *Photochem Photobiol* **2002**, *76*, 622.
(24) Christou, N.-E.; Ayala, I.; Giandoreggio-Barranco, K.; Byrdin, M.; Adam, V.; Bourgeois, D.; Brutscher, B. NMR Reveals Light-Induced Changes in the Dynamics of a Photoswitchable Fluorescent Protein. *Biophysical journal* **2019**, *117*, 2087–2100.
(25) Smyrnova, D.; Moeyaert, B.; Michielssens, S.; Hofkens, J.; Dedecker, P.; Ceulemans, A. Molecular Dynamic Indicators of the Photoswitching Properties of Green Fluorescent Proteins. *The journal of physical chemistry. B* **2015**, *119*, 12007–12016.

6 On state photodynamics for WT

6.1 Introduction

In the *On* state, the rsEGFP2 chromophore adopts a *cis* conformer where the phenol group is deprotonated. The *On* state fluorescent quantum yield (Φ_F) is 0.3, according to Grotjohann et al. 2011[1] and 2012[2]. *On* to *Off* switching quantum yield ($\Phi_{On \rightarrow Off}$) is around 1 % (Chapter 4) and over ten times smaller than the *On* to *Off* switching quantum yield ($\Phi_{Off \rightarrow On}$). As shown in Table 3.1 (Chapter 3), this is a common characteristic of RSFPs. These low values of photoswitching quantum yields make challenging to investigate the *On* to *Off* photoisomerization and usually limit the studies to the *On* state photodynamics.

While for fluorescent organic fluorophores the emission decay can be fitted by a single mono-exponential, for FPs a weighted sum of several exponential functions (multiexponential decay) is needed and usually accounts for the existence of several ground state conformers[3]. For example, the EGFP[4] fluorescence lifetime decay with two characteristics times of 2.0 ns and 3.1 ns, which are attributed to two possible positions of the amino acid E222 forming two different hydrogen bond networks[3]. For RSFPs, the average lifetime and fluorescence quantum yield are indirectly linked to the switching efficiency (chromophore isomerization). As indicated in chapter 2, the fluorescence quantum yield is the ratio between radiative rate (k_r) and the sum of all decay rates ($k_{nr} + k_r$) where the switching rate is included in the non-radiative rate constant (k_{nr}). Therefore the investigation of the *On* state emission lifetime, indirectly reveals essential information related to the RSFP switching dynamics[5]. Directly modifying the non-radiative decay pathways is the best way to optimize the fluorescent quantum yields[6]. The fluorescence lifetime of rsEGFP2 has just been reported with an average lifetime of 2 ns[7]. For Dronpa, the first single-photon counting (TCSPC) experiments reported a mono-exponential fluorescent lifetime of 3.6 ns[8]. On the contrary, the same studies using TRIR revealed a 40 ps minor contribution with no significant spectral difference with the longer 3.6 ns component[9] (370 ps and 1.6 ns for Dronpa2), this ps component was attributed to the solvation of the excited state following the direct electronic excitation. Similarly, Warren et al.[10] also reported a biexponential decay for Dronpa *On*

excited-state dynamics (16 ps and 2 ns). Although no experimental result could validate their hypothesis, they suggested an excited state isomerization followed by a proton transfer in the ground state after the fluorescent decay. The existence of two excited state decay times was also attributed to a heterogeneous ground state population[10]. Moreover, molecular dynamics simulations on Dronpa2[11] attributed the origin of the ground state structural heterogeneity to the chromophore phenolate moiety having one, two, or three hydrogen bonds with the protein cage, revealing that the single hydrogen bond subpopulation is responsible for chromophore isomerization (*Off*-switching). On the contrary, a higher number of hydrogen bonds inhibit the isomerization and promote fluorescence.

In this chapter, the photodynamics of rsEGFP2 *On* state is investigated. The studies are centred on identifying the number of ground-state conformers that control the *cis* anionic lifetime, the fluorescence quantum yield and indirectly the switching rate. Due to its sensitivity, TCSPC is one of the best techniques to study the *On* state dynamics; nevertheless, its time resolution is limited to 100 ps and to fluorescent species. Therefore, UV-Visible and IR transient absorption will be used to study sub-100 ps dynamics and non-fluorescent species complementing the TCSPC results. As commented above the *On* to *Off* photoswitching quantum yield is 1%; one of the highest *On* to *Off* switching yield among the negative RSFPs. The quantum yield is control by the excited state dynamics which yield a photoproduct in the ground state. The evolution of this photoproduct will be probed using ns UV-Vis transient absorption. Finally, as in Chapter 5, TR-SFX experiments will reveal essential structural information helping to identify the excited-state species that control the emission and photoswitching yield. Some figures supporting our analysis can be found at the end of the chapter (Supporting Information Figures). The experimental details can be found in Chapter 2 (description of TR spectroscopies) and Appendix 1 (materials and methods).

6.2 Photodynamics of rsEGFP2 WT in solution

6.2.1 Time-resolved emission using time-correlated single photon counting (TCSPC)

6.2.1.1 Emission decays of rsEGFP2 and comparison to other RSFPs

Absorption and emission spectra were reported in chapter 4 and are shown one more time in supporting Figure 6S-1 at the end of the chapter. The rsEGFP2 fluorescence decay (Figure 6.1) was collected at 510 nm after 485 nm excitation (50 mM HEPES pH 8, 50 mM NaCl buffer). The decay was fitted with a weighted sum of three exponential functions convolved with the instrumental response function (IRF, laser scattering measured with a ludox solution). The fit was done using Fluofit suite from PicoQuant[12] (the equations can be seen in Appendix 1). It is worth mentioning that the residuals. after fitting with a bi-exponential function are not well distributed around zero (Figure 6S-2), which is not the case for a three exponential fit, which yielded time constants of 2.72 ns (32%), 0.45 ns (18%) and 140 ps (50%) -values in parenthesis correspond to their relative contributions to the decay-. From these results stands the unique sub-nanosecond lifetime component with a 50% contribution to the excited state for an FP, which is assigned to a quasi non-fluorescent species (notice that its contribution to the steady-state fluorescence is only 7% in comparison to the 2.72 ns which represents 86%.). To verify the existence of this short component, we decided to measure a set of FP (EGFP) and RSFPs (IrisFP, rsfolder, rsfolder2) for comparison. The results can be seen in Figure 6.1. The fits results, time constants with their relative amplitude, can be seen in Table 6.1. The value for Dronpa found in literature[8] has also been added to the table for comparison.

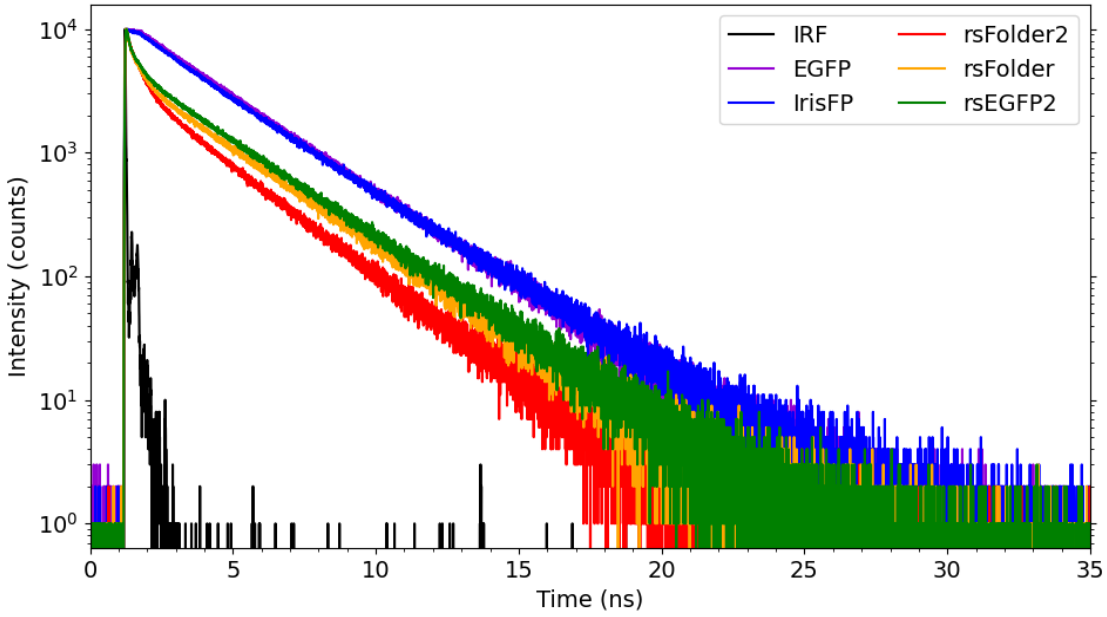


Figure 6.1. Fluorescence decays recorded at 510 nm (excitation 485 nm) for EGFP (Purple), IrisFP (Blue), rsFolder (Orange), rsFolder2 (Red) and rsEGFP2 (Green). The instrument response function (IRF) measured at 485 nm is plotted in black. Figure made with matplotlib python library.

Table 6.1. Fluorescence quantum yields and decay time constants with their relative amplitude (first value in percentage) obtained by TCSPC for different EGFP and different RSFPs at pH 8 (excitation at 485 nm, emission at 510 nm). The second value in percentage represents the contribution of each component to the steady state fluorescence.

	Dronpa*	**EGFP**	**IrisFP**	**rsEGFP2**	**rsFolder**	**rsFolder2**
τ1 (ns)	3.6 ns[8]	2.87	2.94	2.72	2.48	2.76
	100 %	87%/93/%	80%/90%	32%/86%	24%/73%	28%/78%
τ2 (ns)		1.39	1.22	0.45	0.54	0.46
		13%/7%	20%/10%	18%/7%	24%/15%	20%/12%
τ3 (ns)				0.14	0.18	0.15
				50%/7%	52%/11%	52%/10%
Qf*	0.68[8]	0.61[13]	0.47[14]	0.3[2]	0.25[15]	0.23[15]

*value from the literature

As expected for EGFP the fluorescence decay is fitted with two components of 1.39 (13%) and 2.87 (87%) ns. These results are in agreement with the results of Gosh et al.[4], who reported two decay times attributed to two fluorescent ground state conformers. IrisFP decay

is also bi-exponential, and the analysis yielded two decay time constants of 1.22 (20%) and 2.94 (80%) ns. Moreover, rsFolder and rsFolder2 confirm our results on rsEGFP2, i.e. a multi-exponential decay that needs three time constants, including a short decay time constant (< 200 ps) with an amplitude around 50% to the emission decay. This short lifetime species correlate with a lower value for the fluorescence quantum yield of rsEGFP2, rsFolder and rsFolder2 in comparison to EGFP, Dronpa or IrisFP.

6.2.1.2 Emission decays of rsEGFP2 in H_2O according to the excitation and emission wavelengths.

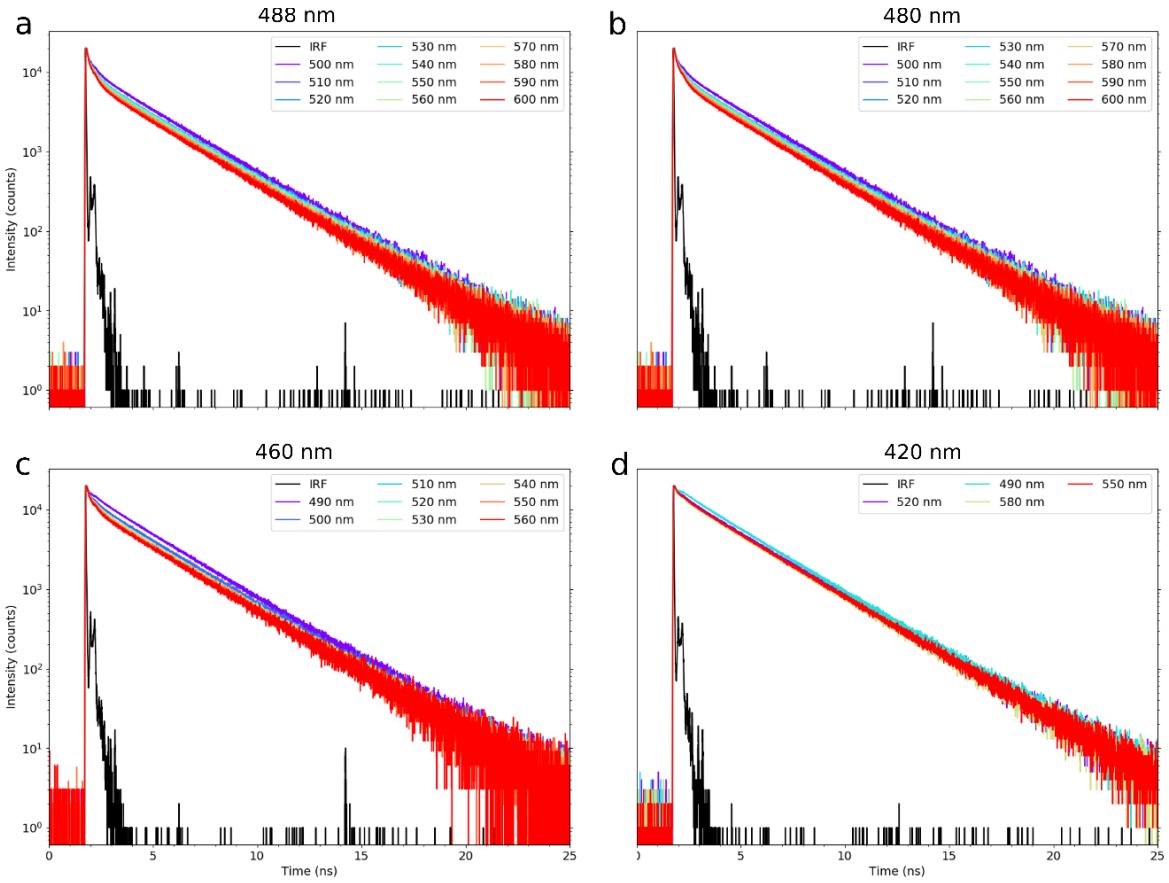


Figure 6.2. Fluorescence decays for rsEGPF2 at different excitation and emission wavelengths. a) 488 nm excitation and emission from 500 to 600 nm. b) 480 nm excitation and emission from 490 to 560 nm. c) 460 nm excitation and emission from 490 to 560 nm. d) 420 nm excitation and emission from 490 to 580 nm. Figure made with matplotlib python library.

To further investigate the excited state emission dynamics, fluorescence decays were recorded for rsEGFP2 WT according to the excitation and emission wavelengths, in 50 mM HEPES pH 8, 50 mM NaCl. The fluorescence decays dependence with the emission wavelength from 500 to 600 nm at four different excitation wavelengths (420 nm, 460 nm, 480 nm, 488 nm) are shown in Figure 6.2. Each data set has been globally fitted with a weighted sum of three exponential functions convolve with the IRF. Time constants are given in Table 6.2 (the given amplitudes in the table are those at 520 nm emission) and the amplitude for each emission wavelengths have been plotted in Figure 6.3.

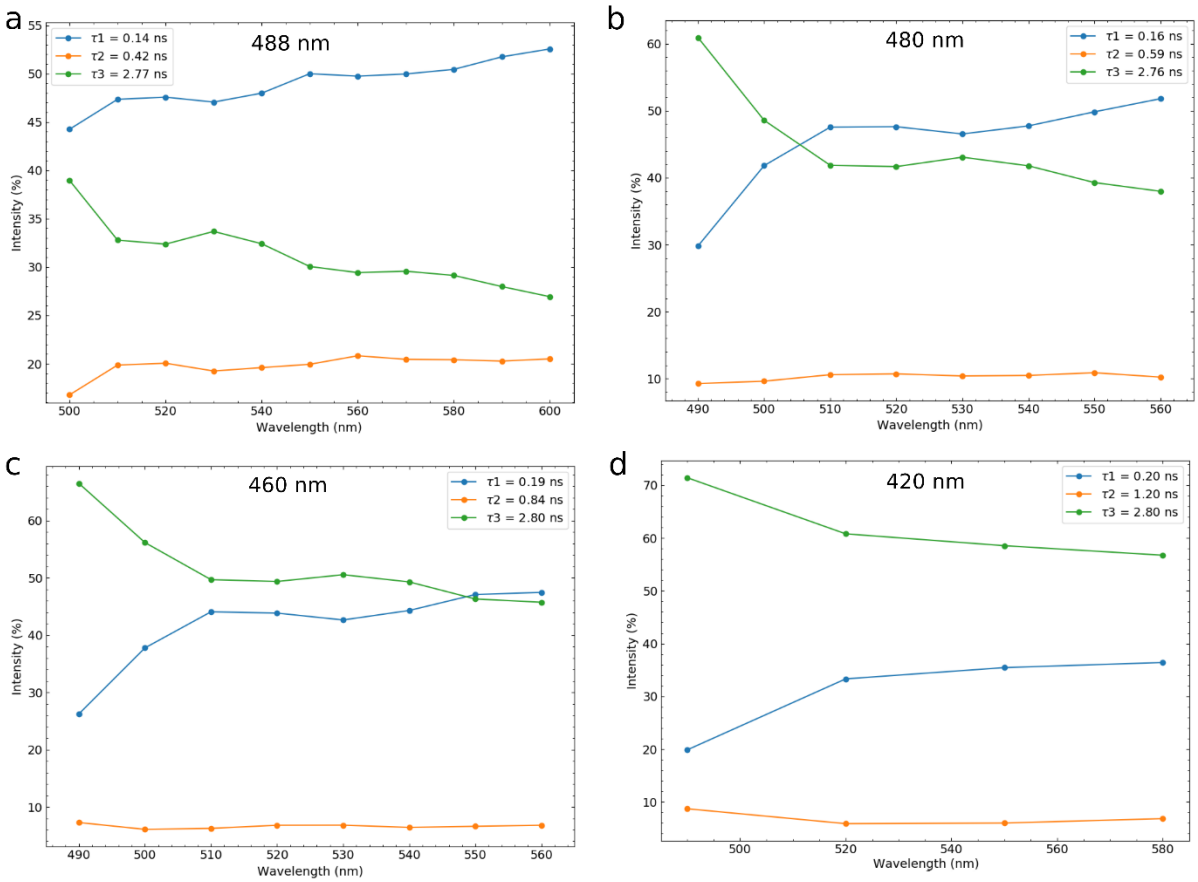


Figure 6.3. Relative (calculated in percentage %) pre-exponential amplitude for the three time constants obtained with a global analysis of recorded fluorescence decays in Figure 6.1. a) 488 nm. b)480 nm. c) 460 nm. d) 420 nm. Figure made with matplotlib python library.

Table 6.2. Global analysis results of rsEGFP2 fluorescence decays in recorded according to the excitation wavelength .

Excitation	**420 nm**	**460 nm**	**480 nm**	**488 nm**
τ1 (ns)	2,80	2,80	2,76	2.72
***A1(%)**	60.8	49.4	41.6	32.1
τ2 (ns)	1,20	0,84	0,59	0.45
***A2(%)**	5.9	6.8	10.7	17.6
τ3 (ns)	0,20	0,18	0,16	0.14
***A3(%)**	33.3	43.8	47.62	50.4

*Amplitudes correspond to the values at 520 nm

Several results can be highlighted from the values in Table 6.2 and Figure 6.3:

- the contribution of the short lifetime species (τ1; named non-fluorescent) increases at higher emission wavelengths (lower energies), no matter the excitation wavelength, while the contribution of the 2.8 ns species (fluorescent) decreases. On the contrary, the intermediate lifetime species roughly does not vary in amplitude within emission wavelengths.
- The intermediate lifetime species (τ2) is a minor species and has an amplitude which never exceeds 20%. Moreover, its contribution at higher excitation energy (lower wavelength) decreases in favour of the long component contribution.
- The lifetime of the intermediate species increases at higher excitation energy from 0.45 ns to 1.20 ns, (nearly a factor of 3).
- The fluorescence decay of rsEGFP2 is nearly bi-exponential at 420 nm excitation.

Altogether the long lifetime species (τ1) and the short lifetime species (τ3) are the main components and attributed to species existing in the ground state with slightly different

absorbance and emission spectra. The short component is mainly a non-fluorescent species and the long component fluorescent. On the contrary, the assignment of the intermediate component is not straightforward. Nevertheless, since the lifetime of the intermediate component increases and reach nanosecond values at high excitation energies, it is considered as a fluorescent species in equilibrium with the 2.8 ns.

6.2.1.3 Emission decays of rsEGFP2 in D2O and crystals buffer solution.

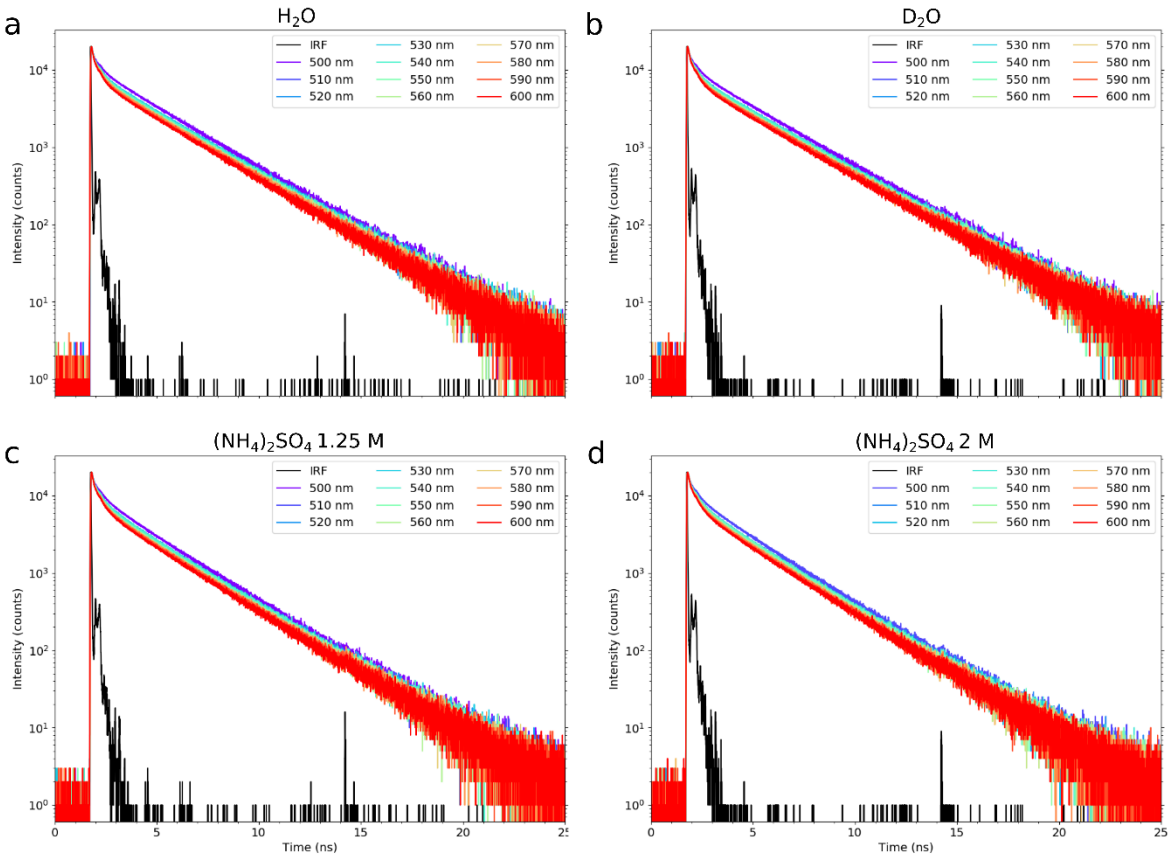


Figure 6.4. Fluorescence decays at different emission wavelengths for rsEGFP2 WT in different buffer solvents, all having 50 mMol NaCl and 50 mMol HEPES as a base. The excitation wavelength was 488 nm. a) in H_2O pH 8. b) in D_2O pD8. c) in H_2O and $(NH_4)_2SO_4$ 1.25 M pH 8. d) in H_2O and $(NH_4)_2SO_4$ 2 M pH 8. Figure made with matplotlib python library.

The fluorescence decays were also studied according to the solvent in D_2O solution and $(NH_4)_2SO_4$ at concentrations of 1.5, and 2.0 M. For all solutions the buffer also contained 50 mM HEPES pH/pD 8, 50 mM NaCl. The experiments in heavy water can help to identify the existence of an ESPT, like those in avGFP[16–18]. The $(NH_4)_2SO_4$ 2 M is a salt used to crystallize the protein. In the previous chapters, the flash-photolysis experiments were done in $(NH_4)_2SO_4$ 1.25 M since higher concentrations induced precipitation of the protein. Already at $(NH_4)_2SO_4$ 1.25 M, the protein dynamics in solution were highly affected and similar to the microcrystals dynamics. Due to the low concentrations of protein used in the TCSPC experiments, where this is adapted so that the solution has an absorption of around 0.07 over one centimetre at 480 nm for the *On* form, a 2 M $(NH_4)_2SO_4$ concentration did not induce sample precipitation. The recorded fluorescence intensity decays in the different buffers can be seen in Figure 6.4, and the global analysis fits results in Table 6.3. From the lifetime reported values, no evident dependence of the rsEGFP2 emission decays with the solvent can be inferred. All data sets have the same lifetime and amplitudes variations with the emission wavelength (Figure 6S-3). The exponential decay values in Table 6.3, reveal that the experiments in D_2O buffer have slightly longer decay lifetimes than those in H_2O, while the opposite effect is observed for decays with $(NH_4)_2SO_4$ present in the buffer. However, these variations are relatively small, and no clear trend can be deduced from the data. Therefore, we can conclude that fluorescence lifetime for rsEGFP2 is relatively independent of the studied buffers for at pH 8.

Table 6.3. Global analysis results of fluorescence intensity decays recorded in different buffers an at 488 nm excitation (Figure 6.4 and Figure 6S-3).

	H2O	**D2O**	**$(NH_4)_2SO_4$, 1.25M**	**$(NH_4)_2SO_4$ 2.0 M**
τ1	2.72	2.86	2.62	2.59
***A1(%)**	32	31	30	31
τ2	0.45	0.52	0.53	0.61
***A2(%)**	18	23	20	17
τ3	0.14	0.17	0.15	0.17
***A3(%)**	50	46	50	52

*Amplitudes correspond to the values at 520 nm

6.2.1.4 Emission decay for microcrystals

The fluorescence decays were recorded for a rsEGFP2 microcrystals suspension in 100 mMol HEPES and $(NH_4)_2SO_4$ 2M. The suspension was highly diluted to avoid that the scattering of the excitation laser damages the detector. Moreover, the fluorescence decay from a single crystal has been measured under a fluorescence lifetime microscope. The recorded decays are shown in Figure 6.5 and the fits in Table 6.4.

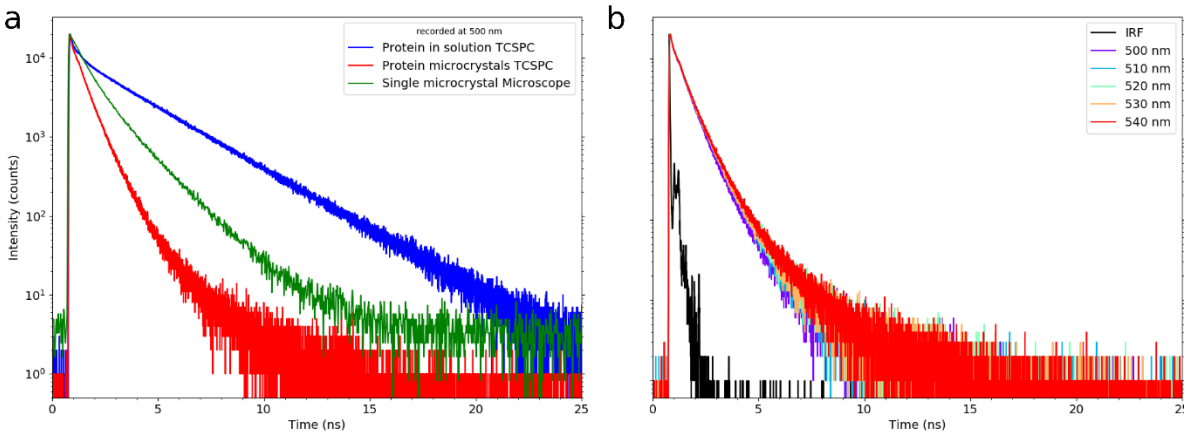


Figure 6.5. a) Fluorescence decays of rsEGFP2 at 500 nm (lifetime under microscope used a long pass filter at 500 nm). In red, colloidal solution of microcrystals, in green, single crystal, and in blue protein in solution (50 mMol NaCl and 50 mMol HEPES). The excitation wavelength was set to 480 for TCSPC experiments and 470 for microscope experiments. b) fluorescence decays for different emission wavelength for rsEGFP2 microcrystals with 488 nm excitation. Figure made matplotlib python package.

From the results in Figure 6.5a, it is clear that the fluorescence decays of rsEGFP2 in microcrystals and in solution differ. The long lifetime contribution (fluorescent species) decreases to a few percent in microcrystals while the shorter life time does not change (τ, %). Furthermore, it is also clear that the decay from a single crystal recorded in the microscope is different from the one recorded on a diluted microcrystals suspension. The differences are attributed to a lower time resolution (the short lifetime cannot be measured) and the use of a long-pass filter for the detection. The fit of the different traces yields three decay times constant; the results can be seen in Table 6.4, together with each component's pre-exponential percentage. Moreover, in Figure 6.5b we can see that for rsEGFP2 microcrystals suspension, there is no clear wavelength emission dependence of the fluorescence decay. This is in

contrast with the results obtained for the protein in solution, and in line with a rigid media. The main result is that the microcrystals fluorescent specie (F. Int) has a life time of 1.8 ns.

Table 6.4. Results of the fits of fluorescence decays in Figure 6.5a.

	Solution	**Microcrystals suspension**	***Single Microcrystal**
τ1 (ns)	2.76	1.80	1,67
F. Amp	41.6%	3.6%	35%
(F. Int)	85%	14.5%	
τ2 (ns)	0.59	0,58	0,57
F. Amp	10.7%	54.2%	65%
(F. Int)	8%	70.4%	
τ3 (ns)	0.14	0,16	NA
F. Amp	47.6%	42.2%	
(F. Int)	7%	15.1%	

*fit not deconvolved with the IRF

6.2.2 UV-Visible transient absorption

6.2.2.1 TRUV-Vis of rsEGFP2 in solution

The excited-state dynamics of the rsEGFP2 WT *On* state were also studied using transient absoprtion femtosecond UV-Vis spectroscopy. The results can be seen in Figure 6.6a and 6.6b. Transient absorption spectra were recorded after excitation of the *On* state with a 482 nm femtosecond laser pulse (0.9 μJ, 100 fs, 1 kHz). After signal formation, the spectrum at 0.35 ps, has three distinct transient features. (i) There is a positive band from 330 to 440 nm, which displays a maximum at 415 nm, and is attributed to excited-state absorption (ESA) band. (ii) There is also a negative band from 440 to 495 nm. This negative band is the ground state bleaching (GSB) band. (ii) The stimulated emission (SE) band from 495 to 650 nm, with a minimum at 495 nm close to the rsEGFP2 fluorescence steady-state spectrum

maximum (502 nm) The GSB and SE bands are overlapped. Similar to TCSPC experiments, the TRUV-Vis have been fitted with the same number of exponentials convolved with a Gaussian function (FWHM, 180 fs). The longest decay time was fixed to that found in the TCSPC experiments. The global fit yielded time constants of 113 ps, 586 ps and 2.76 ns (fixed). Thus, the decay times are similar to those found in TCSPC, with lifetimes of 160 ps, 590 ps and 2.76 ns. These results confirm the presence of three components participating in the excited state dynamics of rsEGFP2 *cis* anionic *On* state. DAS for the three species are relatively similar. In the TCSPC, the major contributions to the fluorescence decay are the 2.76 ns (~40%) and the short component 160 ps (~50%). The pre-exponential factors at 470 nm allow estimating the relative amplitude of each time constant to the recovery of the ground state (*On* state). These values are 50% for 113 ps component and 25% for the intermediate and long component (residues in Figure 6S-4).

The rsEGFP2 excited state dynamics after 400 nm fs laser pulse excitation have also been followed by TRUV-Vis spectroscopy Similar to the above-described experiments the signal is formed in about 350 fs. The spectrum at this time displays an ESA band from 330 to 440 nm, which has a clear maximum at 415 nm. Contrary to the experiments at 482 nm excitation, after signal formation, there is an extra evolution not seen previously from 0.35 to 7 ps, attributed to an excited state relaxation ($S_n \rightarrow S_1 + V_R$). Four exponentials were needed to globally fit the data (convolve with a Gaussian FWHM 120 fs) 69 fs, 2.3 ps, 184 ps and 2800 ps. The last time constant was fixed to the value founded in TCSPC for 420 nm excitation (Figure 6.2d). These results are in agreement with the TCSPC results, i.e. the long component is the main species (Pre-exponential factors amplitude at 490 nm). The results can be seen in the Figure 6S-5.

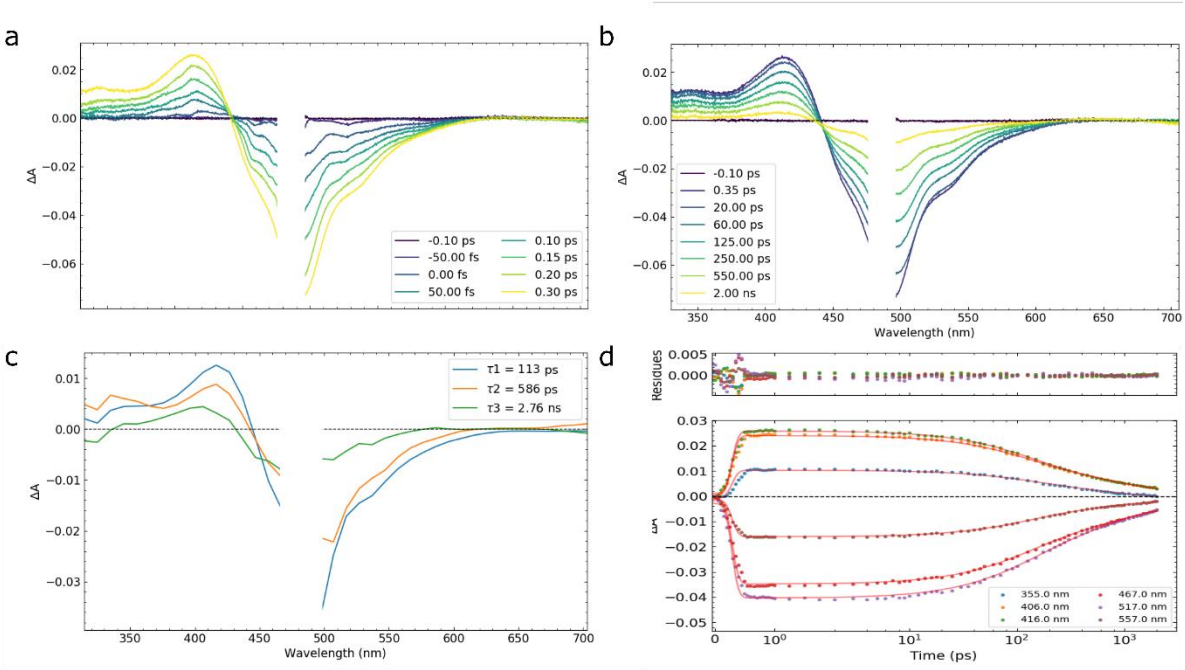


Figure 6.6. Transient absorption spectroscopy of rsEGFP2 (buffer: 50 mM HEPES pH 8, 50 mM NaCl). a) and b) Time-resolved difference absorption spectra recorded after a femtosecond laser excitation (482 nm) of the cis anionic On state until to 2 ns. The spectrum without laser excitation was subtracted to calculate the difference spectra. c) Decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis of 1 decay traces every 10 nm from panels (a) and (b) with a weighted sum of 3 exponential functions convolved with a Gaussian function (FWHM 180 fs). d) Selection of kinetic traces with the corresponding fits in red, the residues in the top panel correspond to the difference between the data and fits. Figure made with Ultra Pyfit.

6.2.2.2 TRUV-Vis of rsEGFP2 microcrystals

Microcrystals suspension were also studied using TRUV-Vis, the experimental result can be seen in Figure 6.7. Due to scattering of the excitation laser pulse, the spectra are cut between 450 and 500 nm. After signal formation in less than a few hundreds of femtoseconds, the spectrum at 0.30 ps displays a positive band ranging from 335 to 450 nm with a maximum at 435 nm. From 500 to 700 nm, we can see a negative band more pronounced until 650 nm. This negative band presents two close minima at 505 and 535 nm, with the first one being more intense at 0.30 ps. The evolutions of the transient signals are much faster for protein

microcrystals than for protein in solution (TRUV-Vis). First, the transient signals recover to half of their intensities in about 100 ps. Secondly, as can be seen in Figure 6.7b nearly all transient signals decay completely at 1.8 ns. Similar to previous TRUV-vis datasets, 1 trace every 10 nm was selected and all globally fitted. Since the amplitude of the 1.8 ns component is only 3.4 % in TCSPC, this component cannot be detected by TRUV-Vis absorption. Therefore, the data have been fitted with a weighted sum of 3 exponential functions convolved with a Gaussian pulse. The associated decay times are 0.16 ps, 93 ps and 497 ps, which are similar values to those found for the 3-exponential fit using TCSPC method (~0.16 ns and ~0.58 ns). The first decay time is associated with the signal formation. The DAS reveal that 93 and 497 ps have similar contributions ~50 % and the shortest component has an SE with a minimum red-shifted in comparison to the long component as in solution.

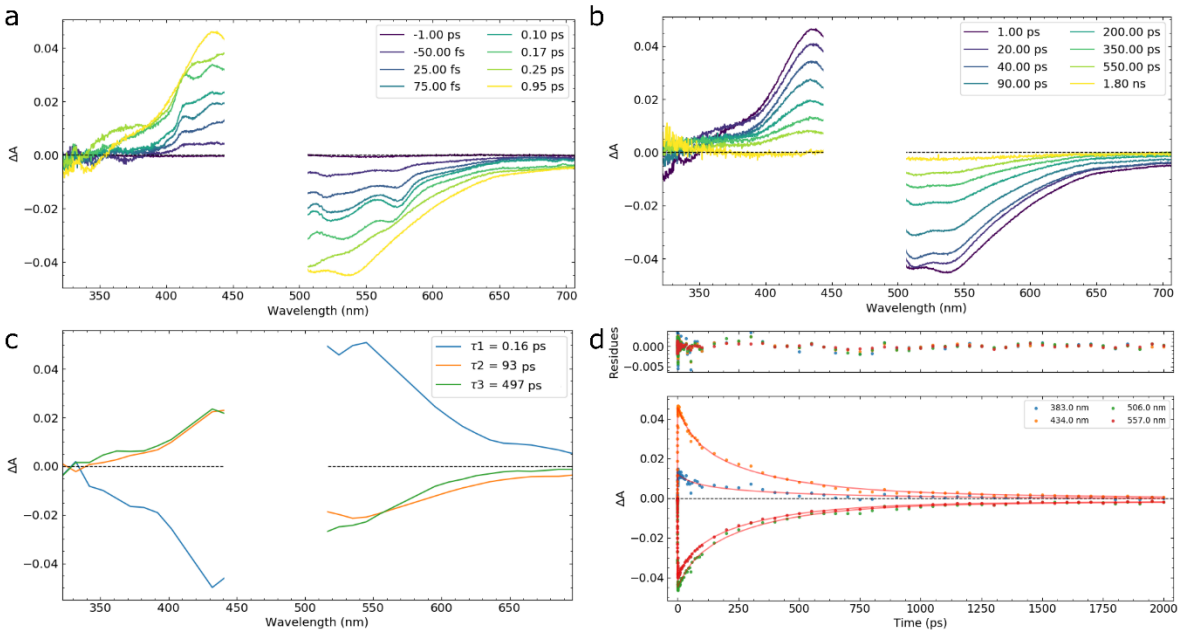


Figure 6.7. Transient absorption spectroscopy of rsEGFP2 microcrystals suspension (buffer: 50 mM HEPES pD 8, 2 M $(NH_4)_2SO_4$). a and b) Time-resolved difference absorption spectra recorded after a femtosecond laser excitation (482 nm) of the cis anionic On state until to 2 ns. c) decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis of 1 decay traces every 10 nm from panel (a and b) with a 3 exponential decay function convolved with a Gaussian function (FWHM 180 fs). d) Selected kinetic traces with the corresponding fits and residual plots Figure made with Ultra Pyfit.

The spectral signatures of rsEGFP2 microcrystals and the protein in solution differ as shown in Figure 6.8 where the spectra at 1 ps after laser excitation have been plotted. Two main differences stand out, the shifts of the ESA maximum between the microcrystals (435 nm) and the protein in solution (415 nm), and a difference in the SE band. While in solution the SE band has a minimum at 495 nm, for microcrystals, the SE band has two minima at 505 and 535 nm with similar amplitude. Finally, the related intensities of the SE and ESA band are also different. In solution, the SE band is much more prominent compare to the ESA band, while in the microcrystals the SE and ESA bands have comparable amplitudes. The spectral evolutions are also different between crystal and solution. The most important differences are related with the long decay component, which is much shorter (1800 ps in crystal vs 2760 ps in solution, considering TCSPC decay values) and have a much lower contribution in crystal (3.4%) than in solution phase (41.6%). Finally, considering the TCSPC values in Table 6.4, we can see that while in solution the excited state dynamics major contribution is the shortest decay time of 140 ps, in microcrystals the main contribution is the intermediate decay time of 580 ps. Similar conclusions can be obtained from the decay times retrieved from the TRUV-vis data.

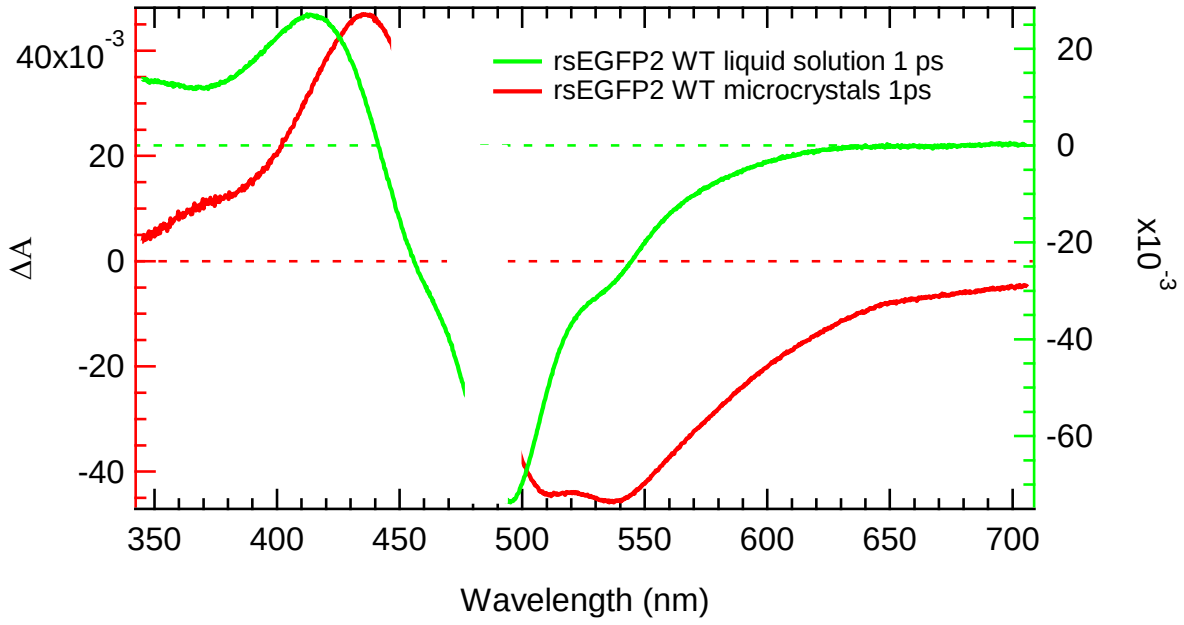


Figure 6.8. Time-resolved difference absorption spectrum recorded at 1ps after a femtosecond laser excitation (482 nm) of the cis anionic On state. rsEGFP2 in solution in green, and rsEGP2 microcrystals suspension in red. Figure made with Igor pro.

The above experiments are done at an excitation energy of 10 GW/cm^2. To validate the TR-SFX results obtained for *On* form, we have studied the photodynamic of rsEGFP2 microcrystals at higher excitation energies (26 GW/cm^2 and 52 GW/cm^2). The results are compared in Figure 6.9. The different data sets recorded at different excitation powers display similar shape and spectral evolution as those described above (10 GW/cm^2). For the highest excitation power (26 and 52 GW/cm^2) there is a new spectral evolution in the 0.5 – 5 ps time windows. This new evolution mainly involves the SE band, that keeps increasing in amplitude between 505-570 nm while no changes can be seen in the ESA band (350-450 nm). These evolutions can be seen in Figures 6.9e and 6.9f. The global fit of the traces was done excluding the minor long component (1.8 ns; residual plots can be seen in Figure 6S-6). The results can be seen in Figure 6.10, where the DAS and the kinetic traces with the corresponding fits and residues have been plotted. The DAS in Figures 6.10a, 6.10c and 6.10e, reveal that the relative contributions of the different components to the SE band do not significantly vary with the excitation energy. The extra time constant (about 2 ps) seen in the 26 and 52 GW/cm^2 datasets can be assigned to a hot excited state or a hot ground state. Extra energy is depleting S_1 state to the ground state (similar process than stimulated emission). Therefore, we can conclude that the *On* excited states dynamics is independent on excitation energy after 10 ps.

Table 6.5. Global fits results obtained after fitting the transient absorption spectra recorded after excitation of the WT microcrystals *On cis* anionic state with a 480 nm femtosecond laser pulse. The data is plotted in Figure 6.9.

	10 GW/cm^2	**26 GW/cm^2**	**52 GW/cm^2**
Signal formation (ps)	0.16	0.32	0.29
Extra component	-	2.29	2.15
τ1	93	102	104
τ2	497	636	528

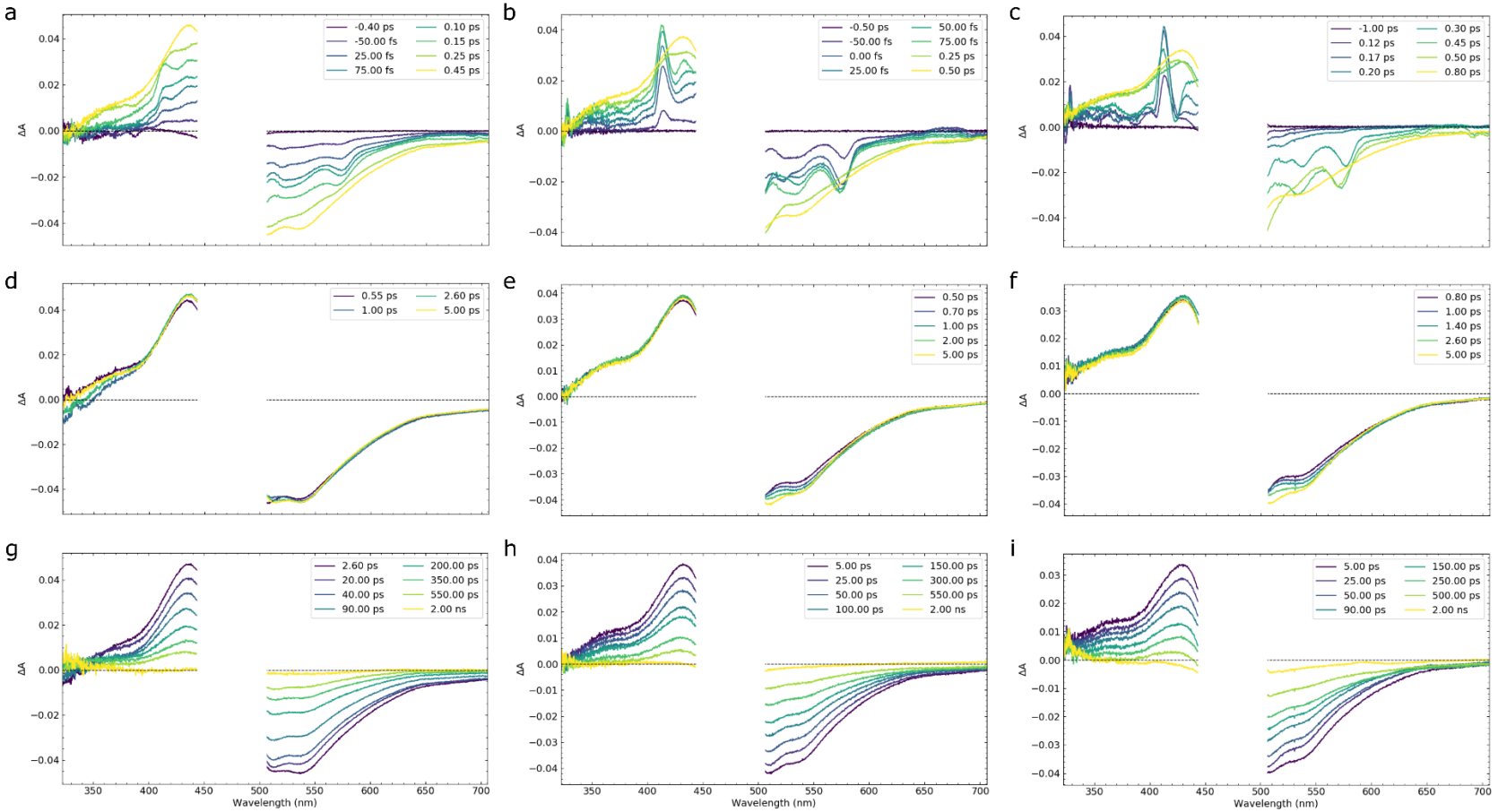


Figure 6.9. Femtosecond transient difference absorption spectra recorded at different excitation energies after excitation of the WT microcrystals On cis anionic state with a 480 nm femtosecond laser pulse. The spectrum without laser excitation was subtracted to calculate the difference spectra. Top panels correspond to the signal formation. Middle panels to signal shifts observed within 5 ps. Bottom panels to the signal evolutions until 2 ns. Excitation energies: 10 GW/cm^2 (a,d and g), 26 GW/cm^2 (b,e and h) and and 52 GW/cm^2(c, f and i). Figure made with Ultra Pyfit,

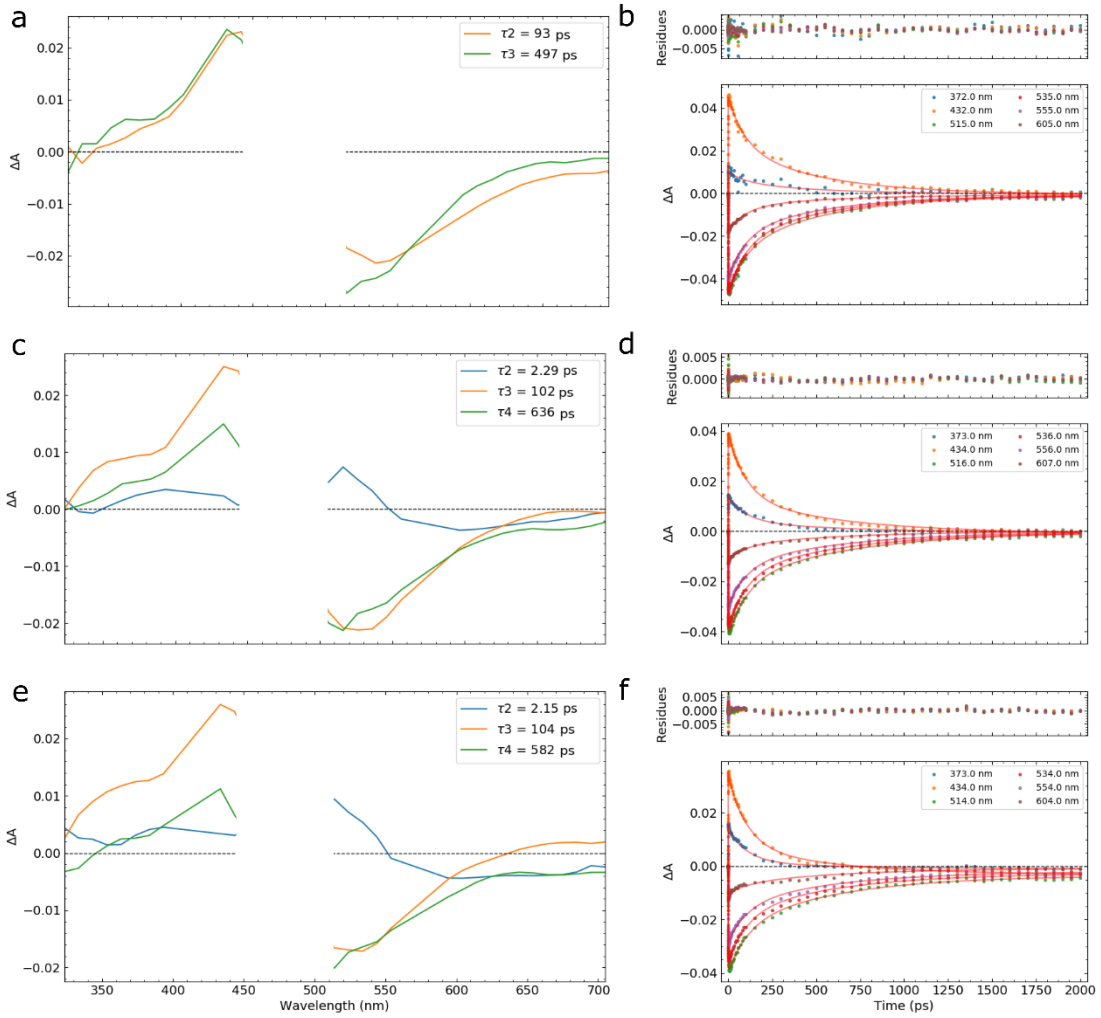


Figure 6.10. The left panels correspond to the decay associated spectra reconstructed using the pre-exponential factors obtained from globally fitting a selection of traces (1 decay traces every 10 nm) from the data set in Figure 6.9 with a weighted sum of exponential function convolved with a Gaussian. Note the shortest component attributed to the signal formation has been removed from all data sets for simplicity. In the right panels, the selected traces representing each of the transient absorption spectral evolution together with the corresponding fits and residues can be seen. Excitation energies: a), b) 10 GW/cm^2, c), d) 26 GW/cm^2, and e), f) 52 GW/cm^2. Figure made with Ultra Pyfit.

6.2.3 Transient absorption Infrared spectroscopy.

The excited-state dynamics of rsEGFP2 WT was studied in heavy water using time-resolved infrared multiple probe spectroscopy (TRMPS) spectroscopy. After irradiation of the *On* state with a 100 femtoseconds (fs) laser pulse at 480 nm, some prominent instantaneous signals are formed, the positive signals correspond to the absorption of *On* excited states (ESA) and the negative signals correspond to depopulation of the *On* form (GSB). The evolutions of the signals at different time delays can be seen in figure 6.11a. The main bleaching bands can be seen at 1310, 1346, 1492, 1538, 1568, 1580, 1621, 1637 and 1651 cm^{-1} with the band 1491 cm^{-1} having the highest amplitude. The most intense band is peaking at 1491 cm^{-1}, which according to Dronpa literature, corresponds to the C-O vibration of the phenolate *On* form[10]. The main positive bands correspond to the signatures of the first excited state species and can be seen at 1385, 1441, 1471, 1602 and 1644 cm^{-1}.

After the laser excitation and signal formation, the transient spectra from 1 ps to 20 ns have mainly 2 spectral evolutions. The signal remains constant for the first few ps and the recovery does not occur after few hundreds of ps. This first evolution is followed by the general decay of all transient signals in a few ns. The evolutions in the ground state of the protein were monitored until 900 μs, although no signal can be observed after 20 ns. Following the spectral evolutions, a singular value decomposition of the spectra yielded two main components. Global analysis of the time traces from 1 ps to 20 ns with a weighted sum of 2 exponential yielded time constants of 0.15 and 2.81 ns (Figure 6.11b). Using the amplitude of the 1491 cm^{-1} GSB, the contributions of the two excited state species has been estimated to be 61.4 % and 38.6 % for τ_1 (0.15 ns) and τ_2 (2.81 ns) respectively. The time constant and percentage found in TRMPS are very similar to those found in TCSPC in D_2O buffer. An important result is that the DAS for non-fluorescent (τ_1) state and fluorescent one (τ_2) are different in the 1300cm^{-1} – 1400 cm^{-1} region and at 1568 cm^{-1}. This means that the structure of these two excited states is different and justify the use of TR-SFX to reveal them.

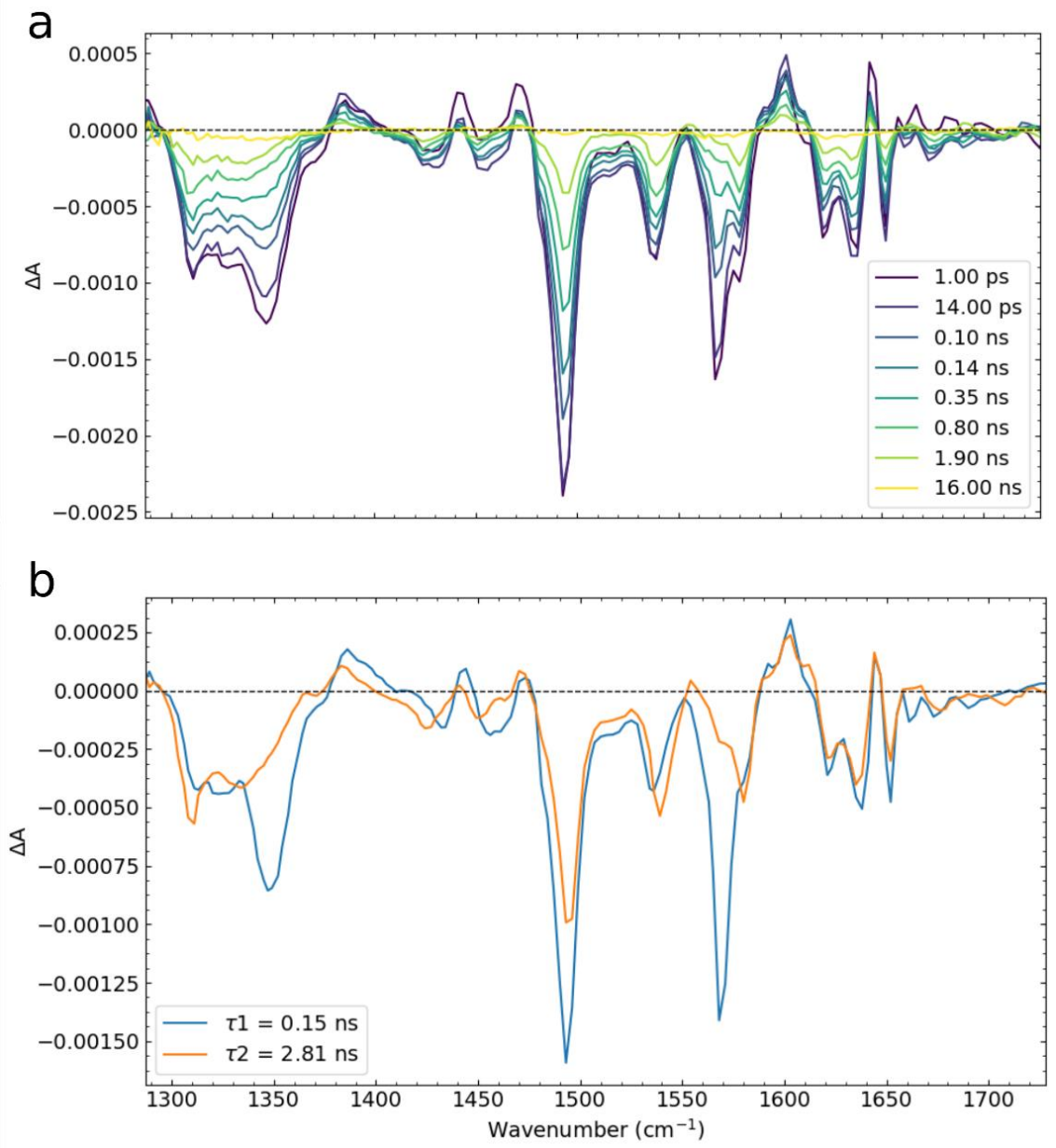


Figure 6.11. a) rsEGFP2 WT IR spectra recorded at different time delays after a 100 femtosecond laser excitation (480 nm) of the On state from 1 ps until 20 ns. The spectrum without laser excitation was subtracted to calculate the difference spectra. b) Decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis of traces from (a) with a weighted sum of 2 exponential functions. Figure made with Ultra Pyfit.

6.2.4 TR-SFX

Since in rsEGFP2, there is about 60% of a non-fluorescent species, it appears as a perfect target for TR-SFX studies. Structural changes were observed at time delay of 1, 10, 200 ps, 1000 ps and 20 ns after photoexcitation of rsEGFP2 in the *On* state. (LCLS LM47 23-27 June 2016: Visualization of excited states in light-switchable fluorescent proteins by femtosecond time-resolved SFX, PI Dr. M. Weik). Condition of the experiment are overall similar to those used in Coquelle et al. with the important difference that the femtosecond pumping laser was carried out at 488 nm, not 400 nm. Time-resolved data enabled the identification of three excited states structures, characterized by different lifetimes, in agreement with high excitation energy TRUV-Vis absorption data for microcrystals. These structural results are illustrated in Figure 6.12 (made by Jacques-Philippe Colletier). The figure displays the chromophore phenolate as well as the protein residues in its immediate vicinity, notably those to which it is H-bonded in the ground state, i.e. Thr 204, His 149, and a structural water. The first excited structure, observed 1 ps after excitation, is characterized by rupture of two H-bonds between the chromophore phenolate oxygen and its partners, viz. those to His149 and the structural water. This excited state is assigned to the 2 ps lifetime hot ground and excited state. The second excited state structure, observed 10 ps after-excitation, is characterized rupture of the last H-bond holding the chromophore (that to Thr204). With no H-bonds to the protein moiety, the two excited states characterized at 1 ps and 10 ps time delay feature a more flexible chromophore, which could be the reason for these being less fluorescent. The last excited-state species, observed 1 ns after excitation, reveals a situation wherein the chromophore phenolate is H-bonded to Thr204 and to the water molecule, but not to His149. In turn, the water is also H-bonded to Ser206, enabling restoration of the proton relay between the chromophore and Glu223. We assign this structure to the fluorescent excited state. Analysis are still under progress to rationalize the decrease of lifetime for the fluorescence species in crystals in comparison to solution.

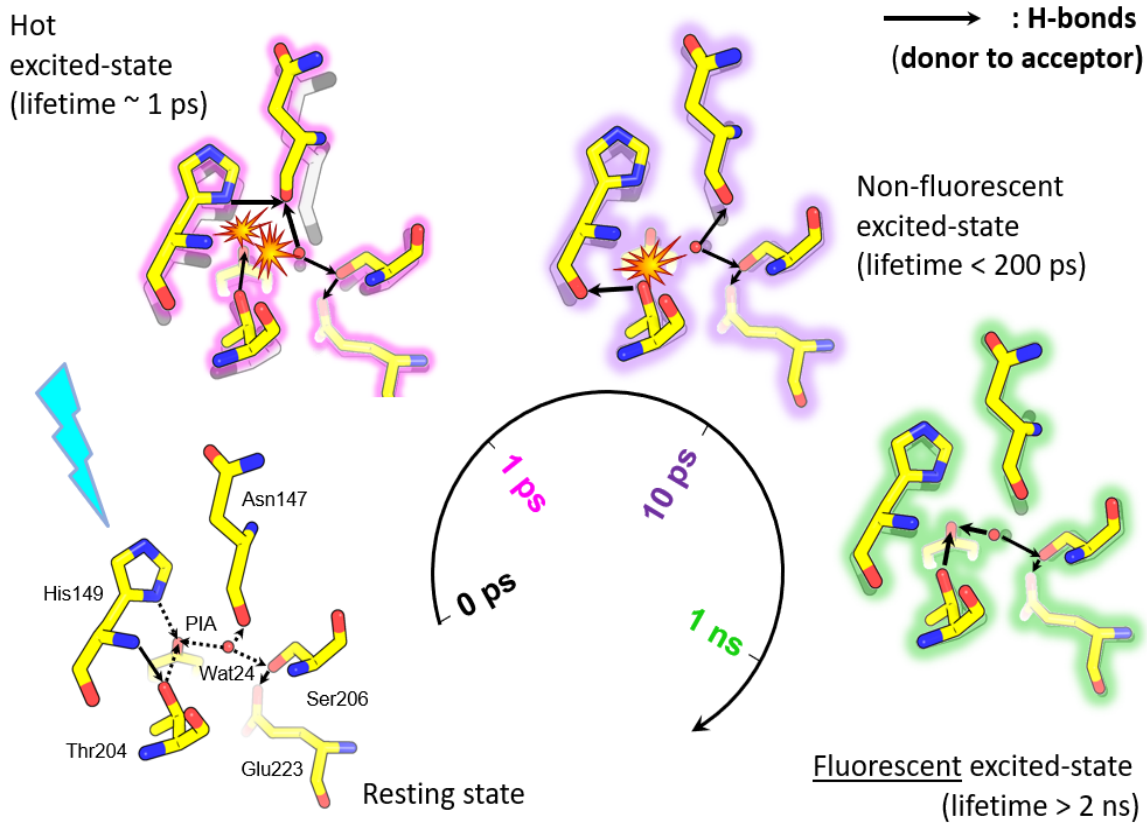


Figure 6.12. Schematic representation of the change in the structure obtained by TR-SFX experiments of hydrogen bonds (black arrow) with the chromophore (PIA) after a 488 femtosecond excitation at 1, 10 and 1000 ps time delay. In pale dark the structure of the resting On state (PIA = HBDI chromophore). Figure made by Dr Jacques-Philippe Colletier.

Unpublished results obtained at LCLC during LM47 beamtime, Andy Aquila, Thomas Barends, Sébastien Boutet, Marco Cammarata, Segio Carbajo, Jacques-Philippe Colletier, Nicolas Coquelle, R. Bruce Doak, Lutz Foucar, Mario Hilpert, Gabriela Kovacsova, Jason Koglin, Karol Nass, Christopher Roome, Giorgio Schiro, Robert L. Shoeman, Ilme Schlichting, Michel Sliwa, Martin Weik, Joyce Woodhouse.

6.2.5 QM/MM calculations

QM/MM simulations have been performed by Dr Matin Field to assign IR bands to specific vibrational modes and species. All the calculations have been carried out starting from the crystallographic *On* state conformation (PDB code *5O89*). In the *On* state structure, the phenolate has three hydrogen bonds. One of these bonds is with a water

molecule placed just in front of it as can be seen in Figure 6.12 or Figure 7.1. The QM/MM simulations have been done considering the three possible orientations and hydrogen bond networks that the structural water (Wat24 in Figure 6.12) molecule can have which were revealed by TR-SFX structural changes. Three names have been given according to the different orientations H-NC, H-NS and H-CS, where H are the water hydrogens, which can be oriented to either the chromophore (C), the serine-206 (S) or the aspargine-147 (N). The IR spectra have been calculated in the ground and excited states (S_0 and S_1), including the dynamics from neighbouring groups to simulate the protein cage. In Figure 6.13 The differential spectra between S_1 and S_0 states calculated for the three different orientations of the water molecule, have been plotted. In the same figure, the DAS spectra obtained from the global fit with two exponential functions performed on the TRMPS data (Figure 6. 11b) have been overlaid for comparison.

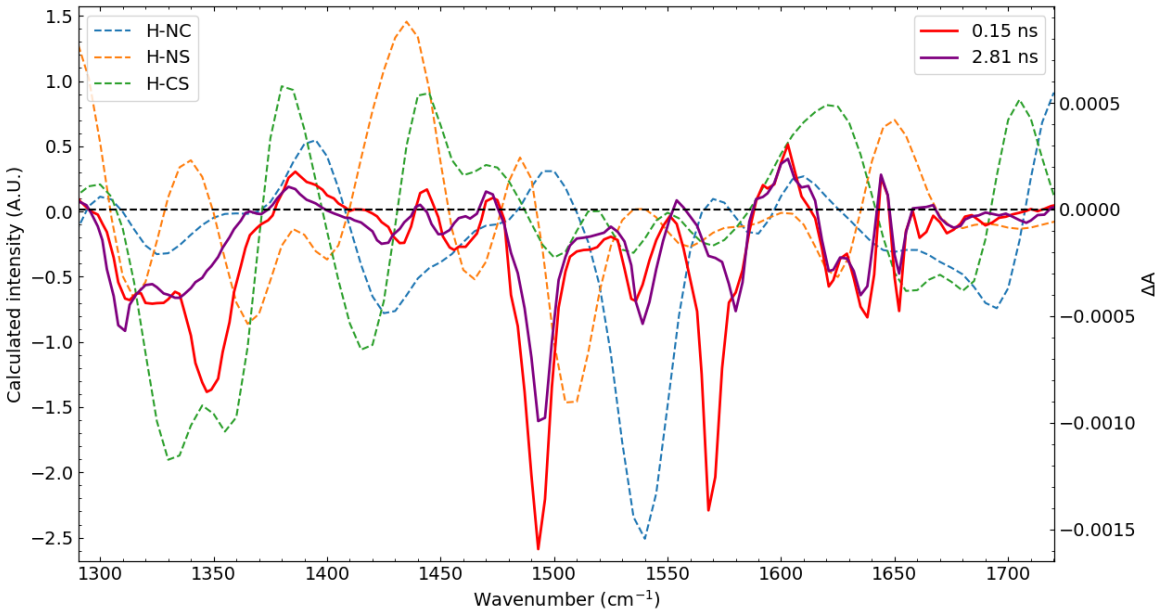


Figure 6.13. Differential spectra between S_1 and S_0 states calculated spectra for rsEGFP2 chromophore and surrounding amino acids considering three different orientations for the water molecule in front of the chromophore H-NC (dashed blue line), H-NS, (dashed orange line) and H-CS (dashed green line). The overlaid solid lines (red and purple) correspond to the decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis of traces from figure 6.7a with a weighted sum of two exponential function. Figure made with matplotlib python package.

The differential DAS spectra can be attributed to the sum of H-CS and H-NS calculated spectra. The H-CS conformer can be attributed to the longer lifetime species because it can model the prominent negative double peak between 1300-1375 cm^{-1} found in the 2.81

DAS. The non-fluorescent species (150 ps) can be attributed to the H-NS calculated conformer, mainly from the negative peak at 1490 cm^{-1} attributed to the C-O phenol stretching and together with the modelization that reproduce the double peak between 1300-1375 cm^{-1}.

6.2.6 TRUV-Vis flash-photolysis ground state dynamics

The ground state dynamics of rsEGFP2 WT in H_2O and D_2O were monitored from 100 ns to 200 μs using a ns transient absorption apparatus. After excitation of the *cis* anionic form with a 5 ns laser pulse at 482 nm, the transient difference absorbance spectrum at 100 ns shows a broad positive band with a maximum at 380 nm (dark blue in Figure 6.14a, the region from 460 to 540 nm is attributed to the residual fluorescence signal). In less than 1 μs, we can see the formation of a positive band at 510 nm while the GSB minimum becomes well defined at 480 nm. Within this time no evolution is observed in the 380 nm positive band. At 1 μs the spectrum displays three clear signatures, two positive bands with maxima at 380 and 510 nm and a negative band with a minimum at 480 nm. Following this first evolution, we can see the decrease of the band with a maximum at 510 nm while the GSB band decreases. Concomitantly the band with a maximum at 380 nm increases in amplitude and shifts to 410 nm. At 200 μs the differential spectrum has a negative maximum at 480 nm which is attributed to the disappearance of the *On cis* anionic form [16] and a maximum at 410 nm which corresponds to the absorbance of the *Off trans* neutral chromophore. Two time constants of 0.538 and 24.14 μs were identified after globally fitting the kinetic traces.

When similar experiments were carried out in D_2O solution (50 mM HEPES pD 8, 50 mM NaCl), the spectral evolutions are similar. The spectral evolutions also required a bi-exponential function for fitting, yielding time constants of 0.64 μs and 36.09 μs. The first time constant is similar in H_2O and D_2O, but a significant deuterium isotope effect is observed for the long time constant (k_H/k_D= 1.5). The latest evolution can thus be attributed to a ground state proton transfer (GSPT). The results for both H_2O and D_2O experiments can be seen in Figure 6.14. The fits and residues for both data sets are shown in Figure 6S-7 at the end of the chapter.

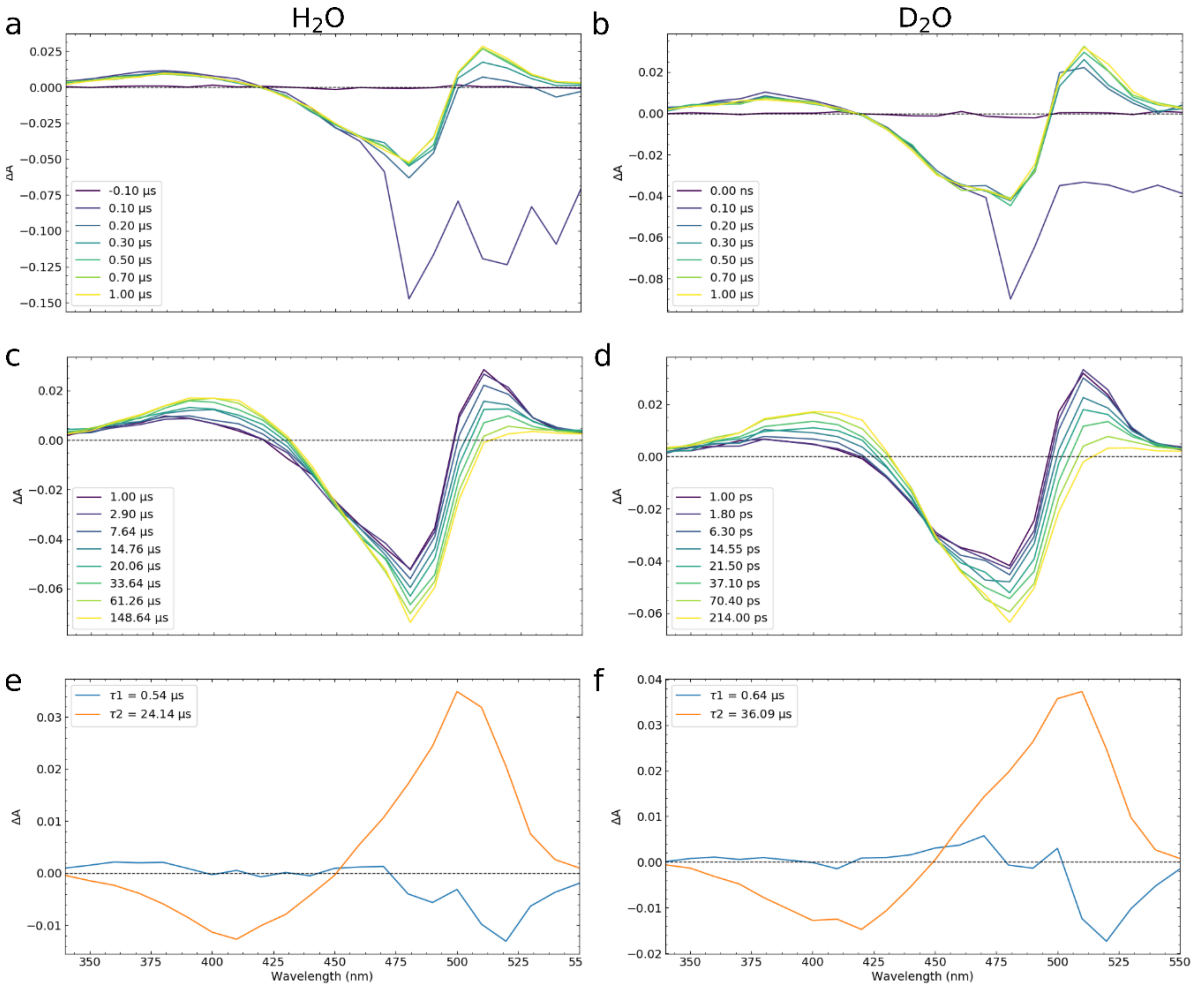


Figure 6.14. a), b), c) and d) Transient absorption spectroscopy of rsEGFP2 WT in H_2O and D_2O solution (50 mM HEPES pH/pD 8, 50 mM NaCl). Time-resolved difference absorption spectra recorded after a 482 nm nanosecond excitation of the cis anionic On state in the time windows from 100 ns to 200 µs. The spectrum without laser excitation was subtracted to calculate the difference spectra. e) and f) correspond to the decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis fitting of the raw data with a weighted sum of 2 exponential functions. Left panels H_2O results, right panel D_2O results. Figure made with Ultra Pyfit

6.3 Discussion and conclusions

The global fit of the fluorescence decays (510 nm) after 488 nm excitation of rsEGFP2 yielded 3 decay time constants, 0.14 ns, 0.45 ns and 2.72 ns, with relative amplitudes to the decay of 50%, 18% and 32 % respectively (Table 6.1). Therefore, three excited state species can be considered in the rsEGFP2 (*On* form) excited-state. The long and short are

the main components. Moreover, the ensemble of fluorescence decays reveals little variations in these two main components with the excitation and emission. The short component is attributed to a non-fluorescent state and the long component to a fluorescent state. The TRUV-Vis experiments using 480 nm excitation confirms the existence of three species with similar DAS contribution as those found in TCSPC. The amplitudes of the DAS spectra obtained from the TRUV-Vis global fit (Figure 6.6c) at 470 nm allows estimating the concentration of the three species, 50% for the non-fluorescent species (113 ps) and 25% for the intermediate and long species (586 ps and 2.72 ns). As no growing component is observed, the three species are formed within the excitation pulse and the most straightforward is to assign them to three species that exist in the ground state. The nature of the intermediate time constant (i.e fluorescent or non-fluorescent) is difficult to determined. The lifetime of this component has huge variations from 0.45 ns at 488 nm excitations to 1.20 ns at 420 nm. Moreover, its contribution to the excited state lifetime is overall small and decreases with the excitation energy, mainly, in favour of the long component (Figure 6.3). Finally, the fluorescence of microcrystals is mainly characterized by the short (42.2 %) and the intermediate (54.2 %) lifetime components; where the long component represents less than 3.6 % (Table 6.4).

Only two components were needed to fit the TRMPS data (Figure 6.11) with associated time constants of 0.15 and 2.81 ns which are identical to the shorter and longer components found in TCSPC. The DAS spectra reveal that, structurally, these two components are relatively similar but not identical. The bleaching signals at 1346 and 1568 cm^{-1}, are mainly observed for the short component. According to the literature[10], and compared to Dronpa, the band at 1568 cm^{-1} can be attributed to υ(C=N/ C=C) chromophore vibrations. Similar results have been reported for Dronpa2 where two decay times of 370 ps and 1.6 ns were found[9]**.** As demonstrated by Romei et al.[5], the primary non-radiative decay process lowering the chromophore's Φ_F are related with changes in the C=C-C bridge between phenolate and the imidazolinone rings. If these changes occur around C=C bond, a *cis-trans* isomerization may occur while an SBR around C-C bond results in a Phenolate-ringflip[5]. The vibrations of the C=C bound lowers the electron density of the double bond and thus favours the *cis-trans* isomerization of the chromophore. At the same time, the hydrogen bonds of the phenolate with the protein cage prevents it from ring-flipping. Interestingly, the TR-SFX results show two conformers H-CS conformer present at 1000 ps and H-NS present at short delays (10 ps).

The TRMPS data, reveals as well two decay times having very similar DAS but not identical. If these DAS are compared with the differential spectra obtained by the QM/MM simulations, we can attribute the longest decay time to the H-CS as revealed by TR-SFX (Figure 6.12). In such a conformation, the proton relay between the chromophore and the Glu223, that characterized the avGFP protein family, can be formed. Notice that the ensemble of measurements are in agreement. The complexity of the time-resolved IR spectra and the similarities between H-NC and H-NS calculated spectra make the structural assignment of the short component not straightforward. Nevertheless, from the ensemble TCSPC results, it is clear that the shorter decay component is a flexible excited state component that rapidly decay to the ground state. Therefore, this component should have a structure that reduces the rigidity of the chromophore. Such a situation is found in the structure with a reduced number of hydrogen bonds between the chromophore and the protein cage; in agreement with the 10 ps TR-SFX structure, where the structural water is oriented towards the Ser206 and the Asn147 (H-NS). In this structure, the hydrogen bond between the chromophore and the water molecule is broken and thus the proton relay. This assignment is in agreement with molecular dynamics simulations on Dronpa2[11]; which attributed the origin of the ground state structural heterogeneity to the chromophore phenolate moiety having several hydrogen bond networks. Moreover, the authors assigned the isomerization to a conformer with a reduced number of hydrogen bonds[11]. Therefore, since the short component has a reduced number of hydrogen bonds, we attribute the main deactivation process for S_1 to be the phenolate ring flip.

So far, we have not attributed the intermediate decay time ~500 ps. The assignment of this component is not straightforward since it is not seen in the TRMPS data. Its amplitude is about 10 % for the TCSPC data and supposes a minor contribution to the excited state decay for the protein in solution. On the contrary, this component becomes the predominant deactivation pathway in microcrystals. Comparatively, both microcrystal and solution, need three components to fit the TCSPC fluorescence decays. In solution, the retrieved decay times at 480 nm excitation are 0.16 ns (main component 47.6 %) 0.59 ns (10.7 %) and 2,76 ns (41.6 %), while in the microcrystals are 0.16 ns 0.58 ns and 1.8 ns where the main component is the 0.58 ns. Therefore, the only difference found is the long decay component, which is much shorter and its contribution much smaller in crystals than in solution (3.4% contribution in microcrystals vs 41.6% in solution). As a

result, this contribution could not be seen for the microcrystals in TRUV-Vis datasets. Similar to the protein in solution, the two components found in the microcrystals of 0.16 and 0.58 ns are attributed to two sub-populations already present in the ground state which are simultaneously excited, and they decay in parallel. In the previous chapter, we could assign the differences between the protein microcrystal and the protein in solution to the $(NH_4)_2SO_4$ salt, which is used to crystallize the protein and accelerates the chromophore deprotonation when it is present in the buffer. On the contrary, the fluorescent decays recorded in the presence of $(NH_4)_2SO_4$ are similar to the protein in solution. Therefore, the differences observed between the microcrystals and solution in the rsEGFP2 *On* state dynamics cannot be attributed to $(NH_4)_2SO_4$. Therefore, we attribute the absence of the long component in the microcrystals, to a modification of the energy landscapes induced by the crystalline structure, which mainly favours the presence of the intermediate component over the longer one in microcrystals. Probably, the crystalline structure force a distorted chromophore which induced a fast excited state decay in ~0.6 ns. Similar to the EGFP[4], we tentatively attribute the origin of the non-fluorescent components (~ 0.15 ns) to the rotation of the T204 which force the rotation of the water molecule causing different hydrogen bond networks between the protein cage and the chromophore H-NS (non-fluorescent) and H-CS (fluorescent).

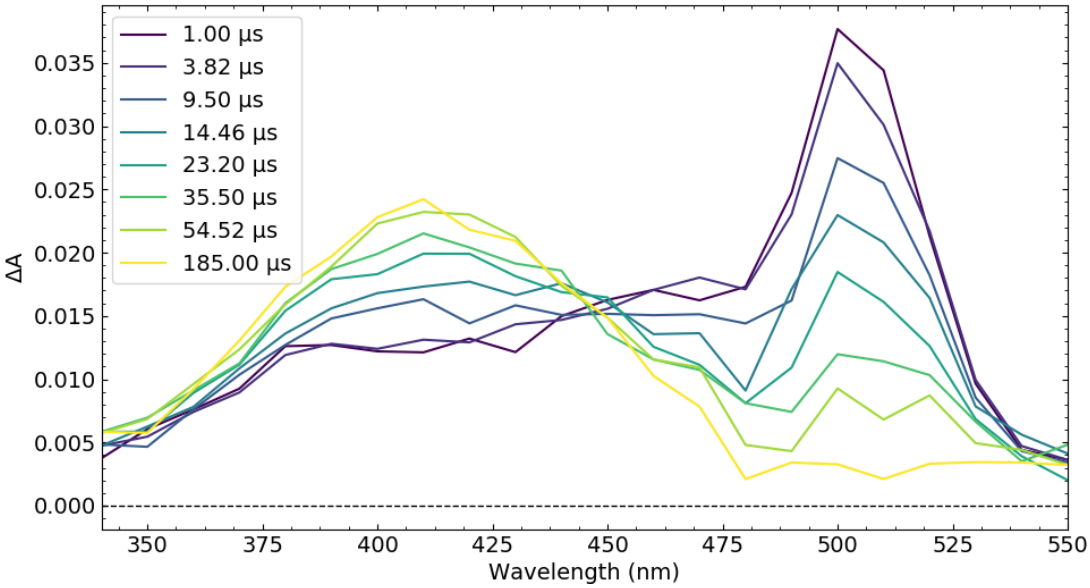


Figure 6.15. Transient absorption spectroscopy of rsEGFP2 WT in H_2O (50 mM HEPES pH 8, 50 mM NaCl), recorded after a 482 nm nanosecond excitation of the cis anionic on state in the time windows from 1 to 200 µs. The On spectrum was subtracted to calculate transient spectra. Figure made with Ultra Pyfit.

After the decay of the excited state, two decay evolutions were seen in solution. The first one is mainly an increase of signal at around 490-510 nm attributed to an artefact generated by the remaining fluorescence, therefore will not be further discussed. After the excited state isomerization through a conical intersection, the molecules mainly come back to the original *cis* anionic conformation as revealed by the low isomerization quantum yield around 1% (determined in chapter 4), but a small portion isomerized to the *trans* anionic form. The estimation of protein chromophores that have isomerized to the *Off* state in the flashphotolysis experiments can be calculated from the GSB signal (assuming the absorption of the *Off* form at 480 nm to be cero). The addition of the absorption *On* state spectrum to the transient spectra reveals the spectrum shape of the transient species (Figure 6.15). The results of this operation can be seen in Figure 6.15, where the spectrum at 1 μs reveals a positive band at 500 nm attributed to the *trans* anionic form. Interestingly, the *trans* anionic form has a 20 nm redshifted absorption maximum compare to the *cis* anionic form. The positive band at 500 nm disappears within a few ten's of microseconds and a positive band centered at 410 nm characteristic of the *trans* neutral form appears. The experiments performed in D_2O reveal an apparent isotopic effect (k_H/k_D= 1.5), which confirms the ground state proton transfer (GSPT). These results are to our knowledge the first ground state *On* to *Off* photoswitching protonation of an RSFP ever reported in the literature. Considering the ensemble results we proposed the following mechanism depicted in Figure 6.16.

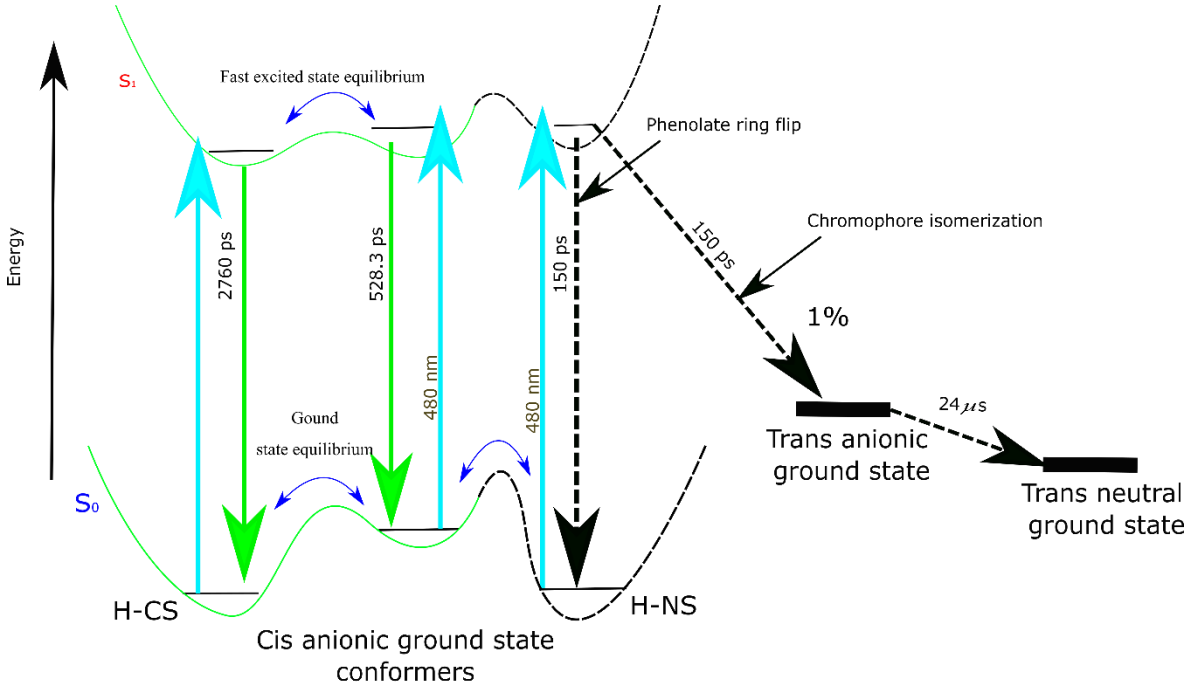


Figure 6.16. General scheme for the photodynamics of On cis anionic state for rsEGPF2 WT. Figure made with InkScape.

6.4 Supporting Figures

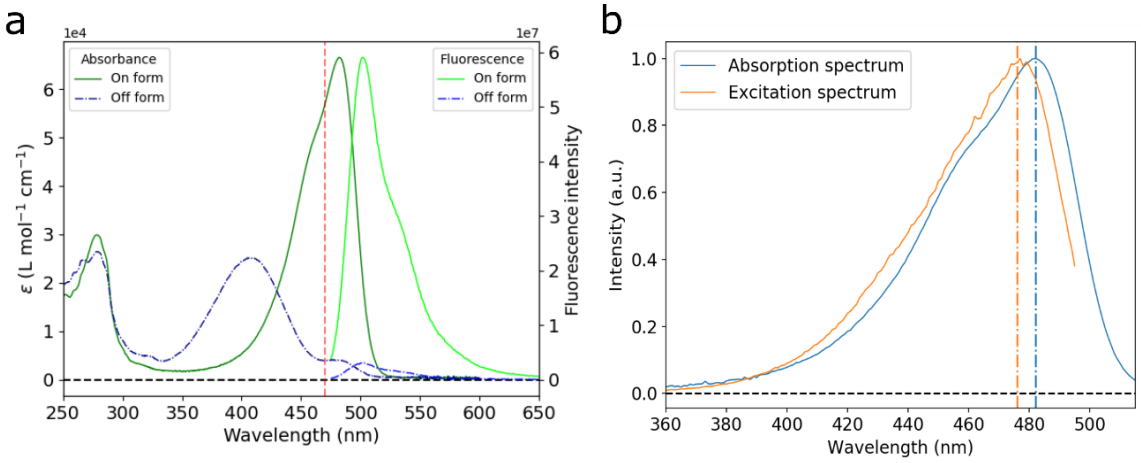


Figure 6S-1. a)rsEGFP2 Fluorescence spectra (light green solid line excited at 470 nm, represented by the vertical red line) compared to the absorption extinction coefficients of On and Off states. b) Comparison between rsEGFP2 absorption (maximum at 482 nm blue vertical line) and excitation spectra (emission at 502 nm; maximum at 476 nm orange vertical line). Figure made with matplotlib python library.

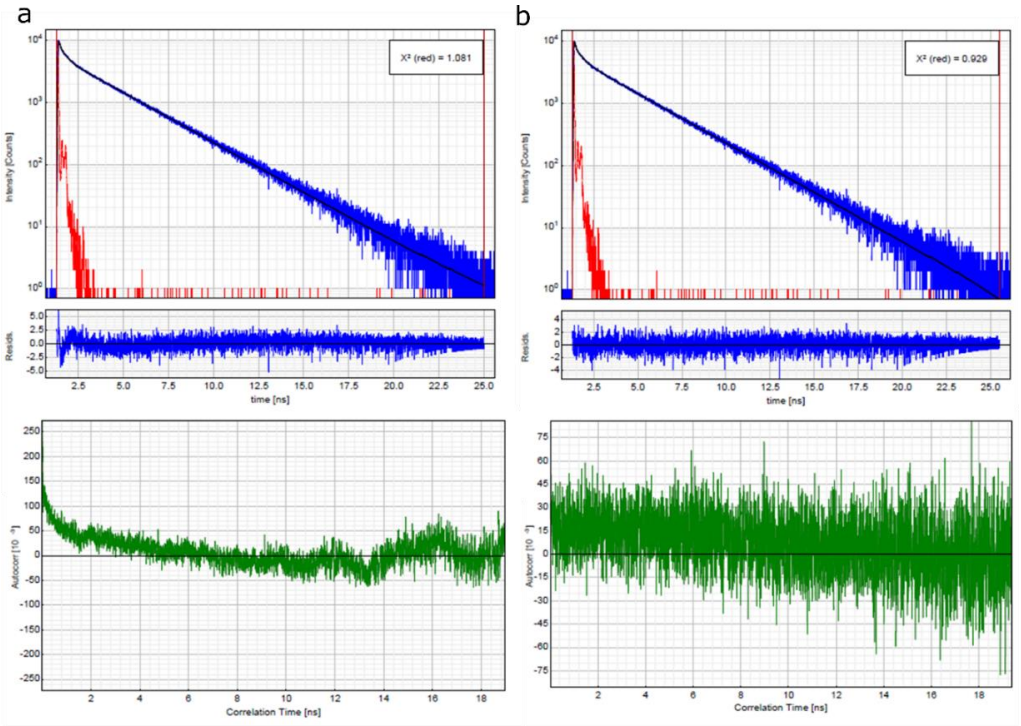


Figure 6S-2. Top panels: fluorescence decay of rsEGFP2 after 480 nm excitation in blue, measured IRF in red and the resulting fit in black. Reduced χ^2 show as an inset. Middle panels: residual plots. Bottom panels: autocorrelation plots. a) fit with a weighted sum of 2 exponential function convolved with the IRF. b) fit with a weighted sum of 3 exponential function convolved with IRF. Figure made with Fluofit.

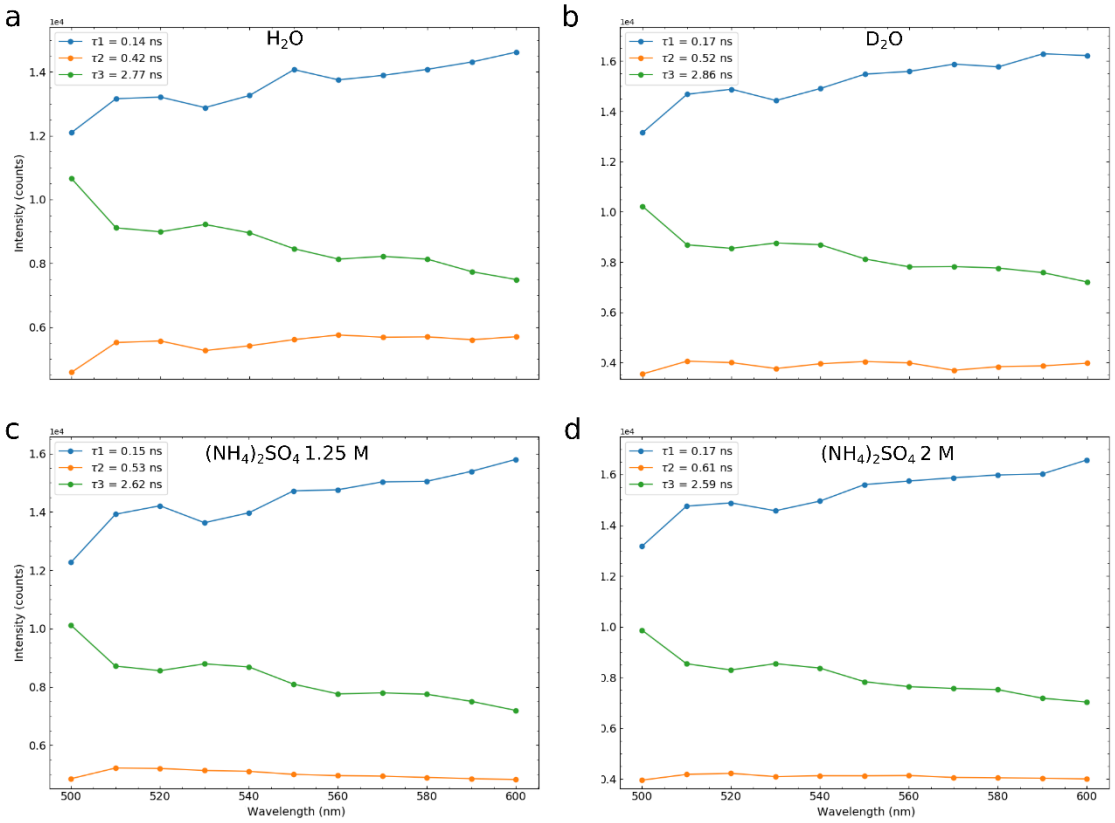


Figure 6S-3. Pre-exponential factor retrieved from the global analysis of the data sets shown in Figure 6.4. Figure made with matplotlib python package.

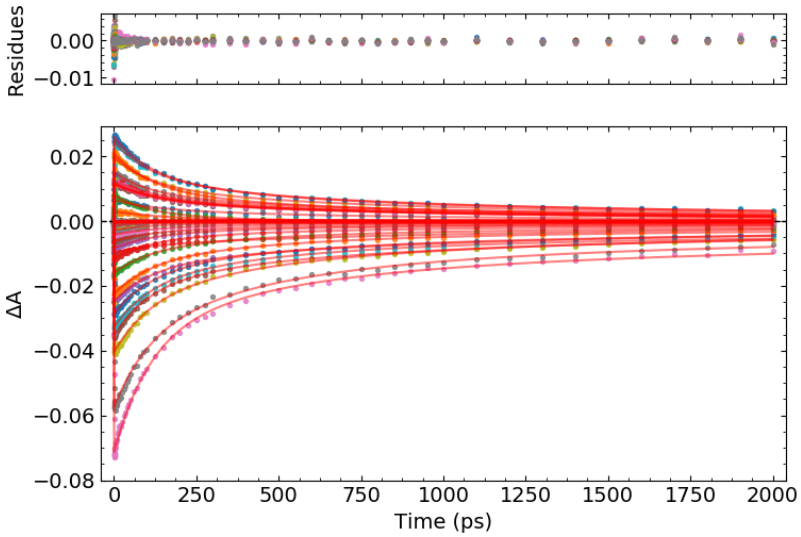


Figure 6S-4. bottom panel: Resulting fit (red) of sum of a three exponential decay convolved with a Gaussian function (FWHM 180 fs) to the kinetic traces (1 every 10 nm) recorded after a 480 nm excitation for rsEGFP2 On state. Top panel: the residues of the fit which are the differences between the fit and raw data. Figure made with Ultra PyFit.

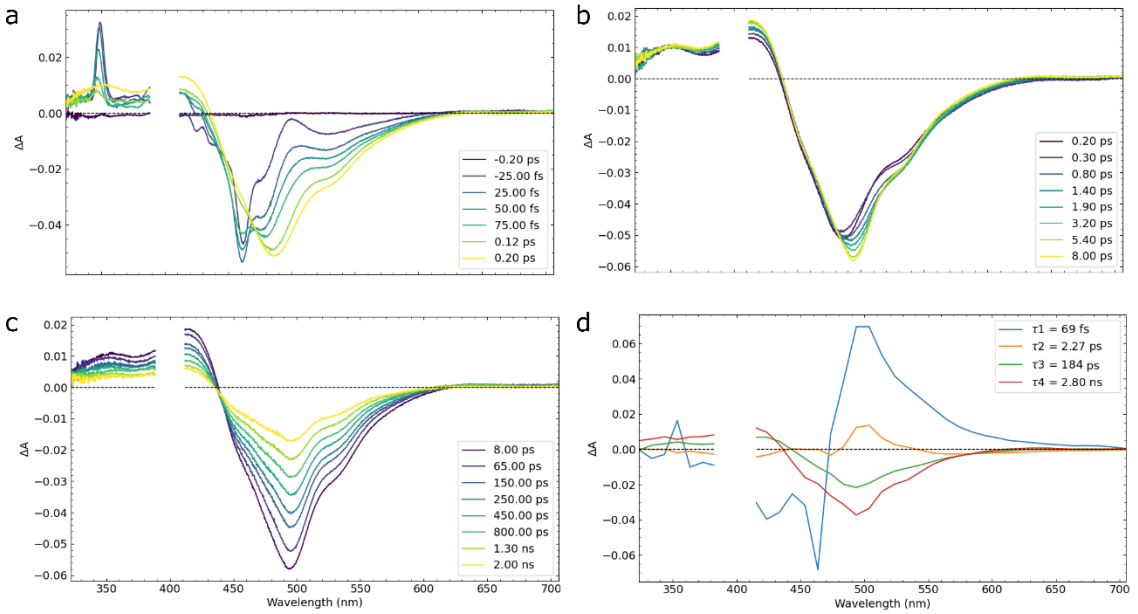


Figure 6S-5**.** *Transient absorption spectroscopy of rsEGFP2 (buffer: 50 mM HEPES pH 8, 50 mM NaCl). a) b) and c) Time-resolved difference absorption spectra recorded after a femtosecond laser excitation (400 nm) of the cis anionic On state until to 2 ns. d) decay associated spectra reconstructed using the pre-exponential factors obtained from the global fit analysis of 1 decay traces every 10 nm panel (a, b and c) with a 4 exponential sum decay function convolved with a Gaussian function. Figure made with Ultra Pyfit.*

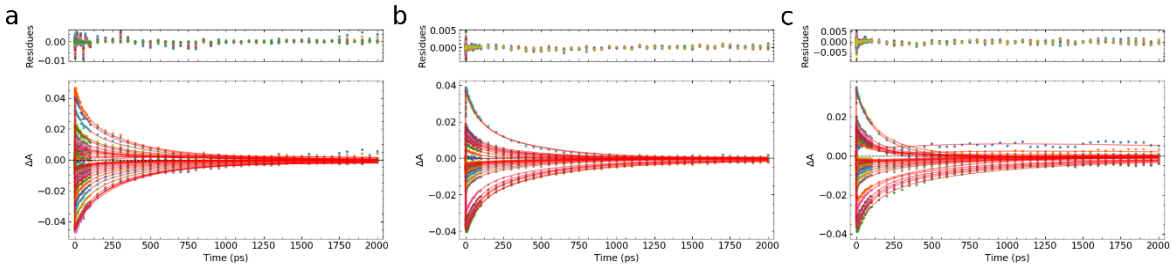


Figure 6S-6. bottom panel: Resulting sum of exponentials function fits in red and raw decay traces for the TRUV-Vis data sets recorded at different excitation energies. Top panel: the residues of the fit which are the differences between the fit and raw data. The corresponding graphs for each of the excitation energies: a) 10 GW/cm^2, b) 26 GW/cm^2, and c) 52 GW/cm^2. Notice that for the high energy excitation datasets, the traces that are distorted by huge stimulated Raman scattering peaks have not been fitted. Figure made with Ultra PyFit.

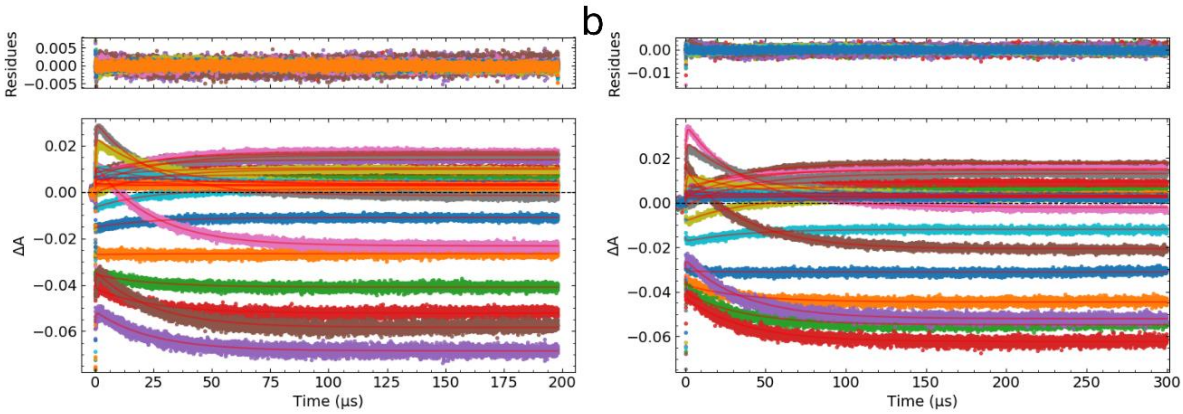


Figure 6S-7. bottom panel: Resulting sum of exponentials function fit in red and raw decay traces for the TRUV-Vis flashphotolysis experiments. Top panel: the residues of the fit which are the differences between the fit and raw data. The corresponding graphs for each of the excitation energies: a) H_2O buffer and b) D_2O buffer. *Figure made with Ultra PyFit.*

6.5 References

(1) Grotjohann, T.; Testa, I.; Leutenegger, M.; Bock, H.; Urban, N. T.; Lavoie-Cardinal, F.; Willig, K. I.; Eggeling, C.; Jakobs, S.; Hell, S. W. Diffraction-unlimited all-optical imaging and writing with a photochromic GFP. *Nature* **2011**, *478*, 204–208.
(2) Grotjohann, T.; Testa, I.; Reuss, M.; Brakemann, T.; Eggeling, C.; Hell, S. W.; Jakobs, S. rsEGFP2 enables fast RESOLFT nanoscopy of living cells. *eLife* **2012**, *1*, e00248.
(3) Shibazaki, C.; Shimizu, R.; Kagotani, Y.; Ostermann, A.; Schrader, T. E.; Adachi, M. Direct Observation of the Protonation States in the Mutant Green Fluorescent Protein. *The journal of physical chemistry letters* **2020**, *11*, 492–496.
(4) Ghosh, A.; Isbaner, S.; Veiga-Gutiérrez, M.; Gregor, I.; Enderlein, J.; Karedla, N. Quantifying Microsecond Transition Times Using Fluorescence Lifetime Correlation Spectroscopy. *The journal of physical chemistry letters* **2017**, *8*, 6022–6028.
(5) Romei, M. G.; Lin, C.-Y.; Mathews, I. I.; Boxer, S. G. Electrostatic control of photoisomerization pathways in proteins. *Science (New York, N.Y.)* **2020**, *367*, 76–79.
(6) Lin, C.-Y.; Romei, M. G.; Oltrogge, L. M.; Mathews, I. I.; Boxer, S. G. Unified Model for Photophysical and Electro-Optical Properties of Green Fluorescent Proteins. *Journal of the American Chemical Society* **2019**, *141*, 15250–15265.
(7) Ruhlandt, D.; Andresen, M.; Jensen, N.; Gregor, I.; Jakobs, S.; Enderlein, J.; Chizhik, A. I. Absolute quantum yield measurements of fluorescent proteins using a plasmonic nanocavity. *Communications biology* **2020**, *3*, 627.
(8) 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. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 9511–9516.

(9) 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.*; et al.* Infrared spectroscopy reveals multi-step multi-timescale photoactivation in the photoconvertible protein archetype dronpa. *Nature chemistry* **2018**, *10*, 845–852.
(10) 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. *Nature communications* **2013**, *4*, 1461.
(11) Morozov, D.; Groenhof, G. Hydrogen Bond Fluctuations Control Photochromism in a Reversibly Photo-Switchable Fluorescent Protein. *Angewandte Chemie (International ed. in English)* **2016**, *55*, 576–578.
(12) https://www.picoquant.com/products/category/software/fluofit-global-fluorescence-decay-data-analysis-software.
(13) Duwé, S.; Zitter, E. de; Gielen, V.; Moeyaert, B.; Vandenberg, W.; Grotjohann, T.; Clays, K.; Jakobs, S.; van Meervelt, L.; Dedecker, P. Expression-Enhanced Fluorescent Proteins Based on Enhanced Green Fluorescent Protein for Super-resolution Microscopy. *ACS nano* **2015**, *9*, 9528–9541.
(14) 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. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 18343–18348.
(15) 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. *Scientific reports* **2016**, *6*, 18459.
(16) Stoner-Ma, D.; Jaye, A. A.; Matousek, P.; Towrie, M.; Meech, S. R.; Tonge, P. J. Observation of excited-state proton transfer in green fluorescent protein using ultrafast vibrational spectroscopy. *Journal of American Chemical Society* **2005**, *127*, 2864–2865.
(17) Stoner-Ma, D.; Melief, E. H.; Nappa, J.; Ronayne, K. L.; Tonge, P. J.; Meech, S. R. Proton relay reaction in green fluorescent protein (GFP): Polarization-resolved ultrafast vibrational spectroscopy of isotopically edited GFP. *The journal of physical chemistry. B* **2006**, *110*, 22009–22018.
(18) Leiderman, P.; Gepshtein, R.; Tsimberov, I.; Huppert, D. Effect of temperature on excited-state proton tunneling in wt-green fluorescent protein. *The journal of physical chemistry. B* **2008**, *112*, 1232–1239.
(19) Istratov, A. A.; Vyvenko, O. F. Exponential analysis in physical phenomena. *Review of Scientific Instruments* **1999**, *70*, 1233–1257.

7 Preliminary results on rsEGFP2 variants: influence of the chromophore environment to the On-state photodynamics

7.1 Introduction

The previous chapter shows the importance of the chromophore surroundings to optimize the *On* state (*cis anionic* form) emission properties. The protein cage and chromophore surroundings of rsEGFP2 are very similar to those of the EGFP, as shown in Figure the 7.1, where the structure of the *On* state of the two proteins are overlaid. The rsEGFP2 structure corresponds to **5DTX** reported by El Khatib et al[1], and the EGFP corresponds to the **4EUL** reported by Arpino et al[2]. As shown in Figure 7.1, in the two structures, the phenyl group of the chromophore has three hydrogen bonds with the T204(203), H149(148) and a hydrogen with a structural water molecule, which in turn is hydrogen bounded to the S206(205) and N147(146) (values in parenthesis correspond to the EGFP). Furthermore, both proteins have the same proton relay to the E223(222). In the EGFP, the E223(222) adopts two different conformers, either bounded to the S206(205) or to the T66(65) which forms part of the chromophore. On the contrary, the mutation T66A in rsEGFP2 avoids the interaction between the E223 with the mutated alanine, and only one conformer for E223 is seen in in the *On* state rsEGFP2 X-ray structure (pdb: **5DTX**). The structures in figure 7.1 reveal the similarities between rsEGFP2 with its parent protein EGFP. Only four mutations differ between the two proteins, T66A, Q70L, V164S and A207K. If these four mutations are analysed, one by one, we can detect the main mutations that render EGFP reversible. As can be seen in Figure 7.1c, the mutation A207K is also shared by mEGFP[3] (monomeric-EGFP) which is not RSFP. In general, this mutation helps to generate monomeric proteins. The mutation T66A in rsEGFP2 is not present in rsEGFP. Consequently, there are two main mutations that can be identified, Q70L and V164S, which render the EGFP reversible. The rsEGFP2 chromophore in the *Off* form is near to L70 (spatially on top) and is hydrogen-bonded with a water molecule which is in turn hydrogen-bonded to the S164. Therefore, the Q70L mutation liberates space for the chromophore to adopt the *Off* conformation, which is stabilized by the

V164S mutation. Probably the carbonyl group of the Q70 prohibits the EGFP chromophore from adopting the *Off* from isomer. Interestingly, these two mutations are also shared by rsEGFP, rsGreen, rsFolder and rsFolder2, thus one can conclude that all RSFP derived from the avGFP share these two mutations. Finally, T66A is a very interesting mutation. As mentioned in chapter 6, in the EGFP, T66 directly interacts with the E223, which adopts two different conformations. On the contrary, as can be seen for rsEGFP2, the E223 has only one conformation. Thus, it is not possible to attribute the ground state conformers of rsEGFP2 to the different conformations of the E223, as it was attributed in EGFP to be the origin of the two fluorescent lifetimes by Gosh et al[4]. (Notice in chapter 6 it has been hypothesized that the ground state conformers attributed to different hydrogen bond networks are caused by the T204 rotation).

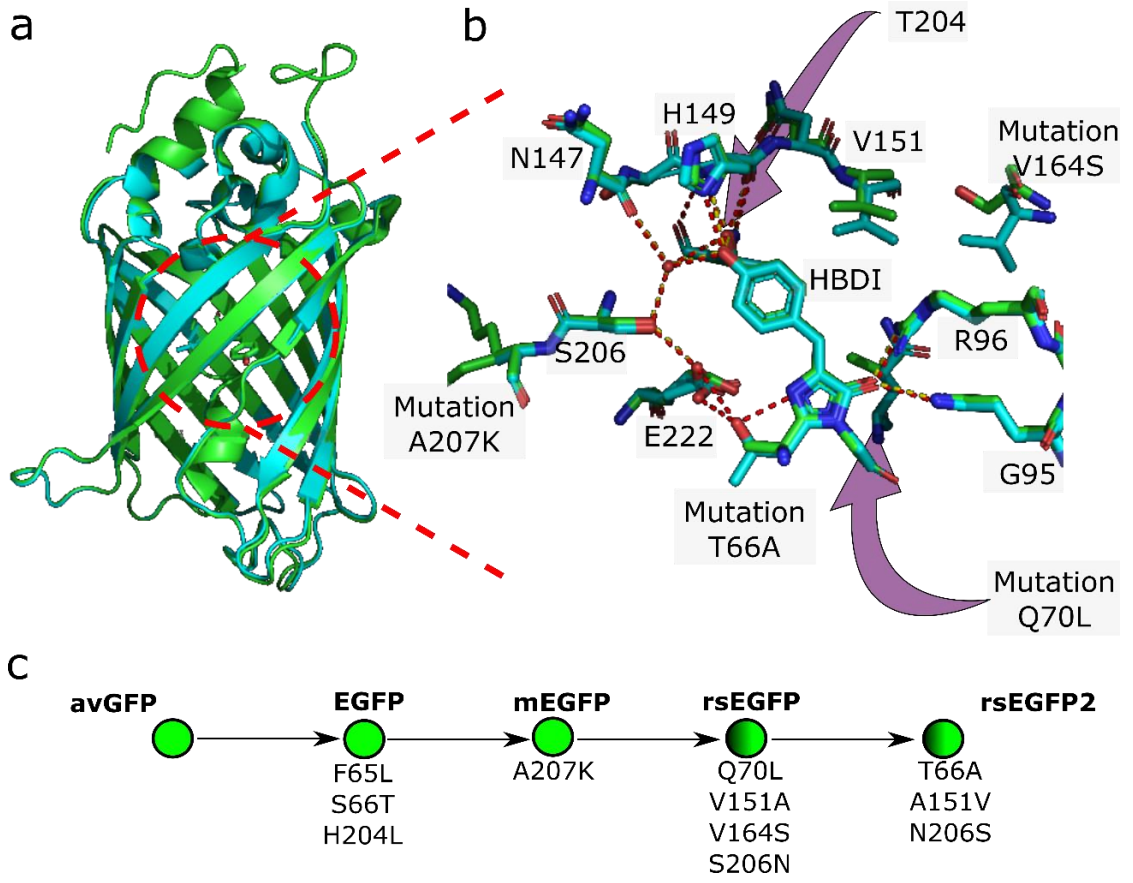


*Figure 7.1. a) X-ray structures of: rsEGP2 (green), PDB **5DTX**[1] and EGFP (blue), PDB **4EUL**[2]. b) Zoom on the chromophore (HBDI) region together with the main interactions of the protein cage with the chromophore (red interaction EGFP, yellow interactions rsEGFP2). c) Lineage of rsEGFP2 with the corresponding mutations from the direct parent protein. (Notice rsEGFP2 in the PDB **5DTX** has one extra aminoacid, thus T65 in EGFP is A66 in rsEGFP2 marked in the figure as T66A. Similarly, the numbers in panel c) are increased by 1). Figure made with PyMol[5] and InkScape.*

Recently, a new publication has come out targeting different mutations around the chromophore in rsGreen protein[6], where the authors showed that a single point mutation can introduce small biochemical changes in the surroundings of the chromophore which can trigger large changes in the spectroscopic properties of a fluorescent protein. They also found that the hydrogen bonding network in the surroundings of HBDI oxygen of the phenolate, and in particular by water molecules in direct connection to it, is the determinant key aspect of photoswitching efficiency. As shown in the previous chapter, fluorescence decay reveal important information regarding the emission properties and indirectly on switching. We showed that the fit of the fluorescent intensity decay (excitation at 480 nm and emission at 510 nm) of rsEGFP2 (WT) yielded 3 decay time constants: 0.15 ns (non-fluorescent, $\tau 3$), 0.47 ns (fluorescent, $\tau 2$) and 2.81ns (fluorescent, $\tau 1$), and the corresponding populations to the decay is 42.5% (A3), 15.5% (A2) and 42 % (A1) respectively. These populations are attributed to species already present in the ground state, which are simultaneously excited and they decay parallel one to the others. They are attribute to different hydrogen bond network around the chromophore and are in equilibrium in the ground and excited state. The non-fluorescent species is the more flexible species and more likely the precursor of isomerization. To study how the chromophore surrounding influence the *On* state photodynamics nineteen different variants were designed by our collaborators in Grenoble (JP. Colletier, M. Weik and coworkers). These nineteen different rsEGFP2 variants contain from a single point mutation to up to three, and all in the close proximity or directly hydrogen bonded to the HBDI chromophore. We report here the photophysical steady-state properties (absorption spectra, emission spectra, fluorescence quantum yield, fluorescence lifetime) for nineteen variants (including V151A and V151L) in comparison to WT rsEGFP2.

7.2 Photophysical properties of *On* state for different rsEGFP2 variants

As illustrated in Figure 7.2, the majority of the mutations are located in the surrounding of the HBDI phenolate group, addressing amino acids involved in the hydrogen bond network, i.e. either the T204, the H149 and the S206. One of the most interesting mutations is the S206N which restores the asparagine, the original amino acid for rsEGFP and rsGreen proteins. Moreover, the E223 was also mutated in order to check if the proton

relay that is responsible for the fluorescence in all avGFP derivated proteins. Finally, there are several double and triple mutations combining the above-mentioned amino acids. Figure 7.2 summarizes the different mutations and the values in parenthesis are the corresponding fluorescent quantum yield values. (measured at a pH for which only *On* form is present).

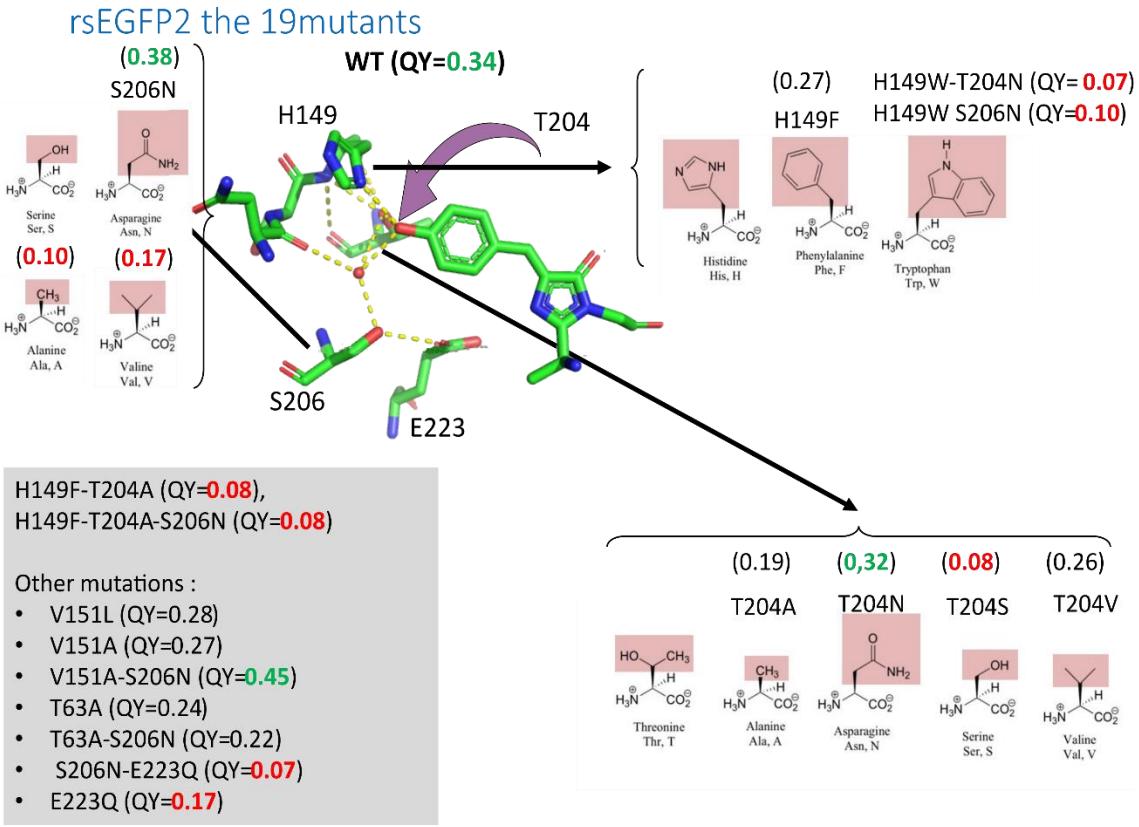


Figure 7.2. Scheme representing the main positions that have been mutated, the values in parenthesis indicate the fluorescence quantum yield. In the center the structure of the rsEGFP2 WT is represented in green. Figure adapted from an initial idea by Dr. Olivier Devos. X-Ray structures have been reproduced using Pymol from the PDB ***5DTX***[1]*. Figure made with power point.*

For all 19 purified mutants, emission decays were also measured using TCSPC at four different emission wavelengths from 490 to 550 nm every 20 nm, after 478 nm femtosecond laser pulse excitation. The measurements were done by Dr. Olivier Devos. The four traces have been globally fitted for each of the mutants with a weighted sum of three exponential functions (equation 7.1) convolved with the instrument response

function (IRF) measured in a Ludox solution. The decays were fitted to equation 7.1 with Fluofit suit from PicoQuant[7].

$$I(t) = \sum a_i \cdot e^{\frac{-t}{\sum \tau_i}} \qquad \text{Equation 7.1}$$

Where (a_i) are the pre-exponential factors and τ_i the corresponding lifetimes for each component. From the fits to Equation 7.1 several values have been extracted and are displayed in Table 7.1. These are: (i) the contributions of each species to the emission decay A_i; (ii) The contribution of each species to the fluorescence I_i; (iii) And the average lifetime in intensity. The contributions of each species to the excited state (A) can be calculated as a relation of pre-exponential factors with equation 7.2.

$$A_i = \frac{a_i}{\sum a_i} \qquad \text{Equation 7.2}$$

The contribution of each of the components to the fluorescence (I) was calculated with equation 7.3.

$$I_i(\%) = \frac{a_i \cdot \tau_i}{\sum a_i \cdot \tau_i} \qquad \text{Equation 7.3}$$

Finally, the average lifetime in intensity was calculated using equation 7.4.

$$< \tau >= \frac{\sum a_i \cdot \tau_i^2}{\sum a_i \cdot \tau_i} \qquad \text{Equation 7.4}$$

Where (t) are each of the decay times retrieved from the global fit, and (A) the contribution (percentage) to the fluorescence decay at 510 nm The results of the global fits for all the mutants can be seen in Table 7.1 Together with the excited state life times, the approximate pK_a is reported (pK_a values have been determined by Dr. Martin Byrdin from the IBS, Grenoble). According to the pK_a values in table 7.1, the emission decays and photo-physical properties (absorption and emission maxima, the Stokes shift and fluorescence quantum yield Φ_F) were measured at pH where cis anionic is the solely conformer (pH = 8 for pKa = 6; pH = 10 for pKa = 8 and pH = 11.5 for pKa > 10).

Pearson correlation matrix

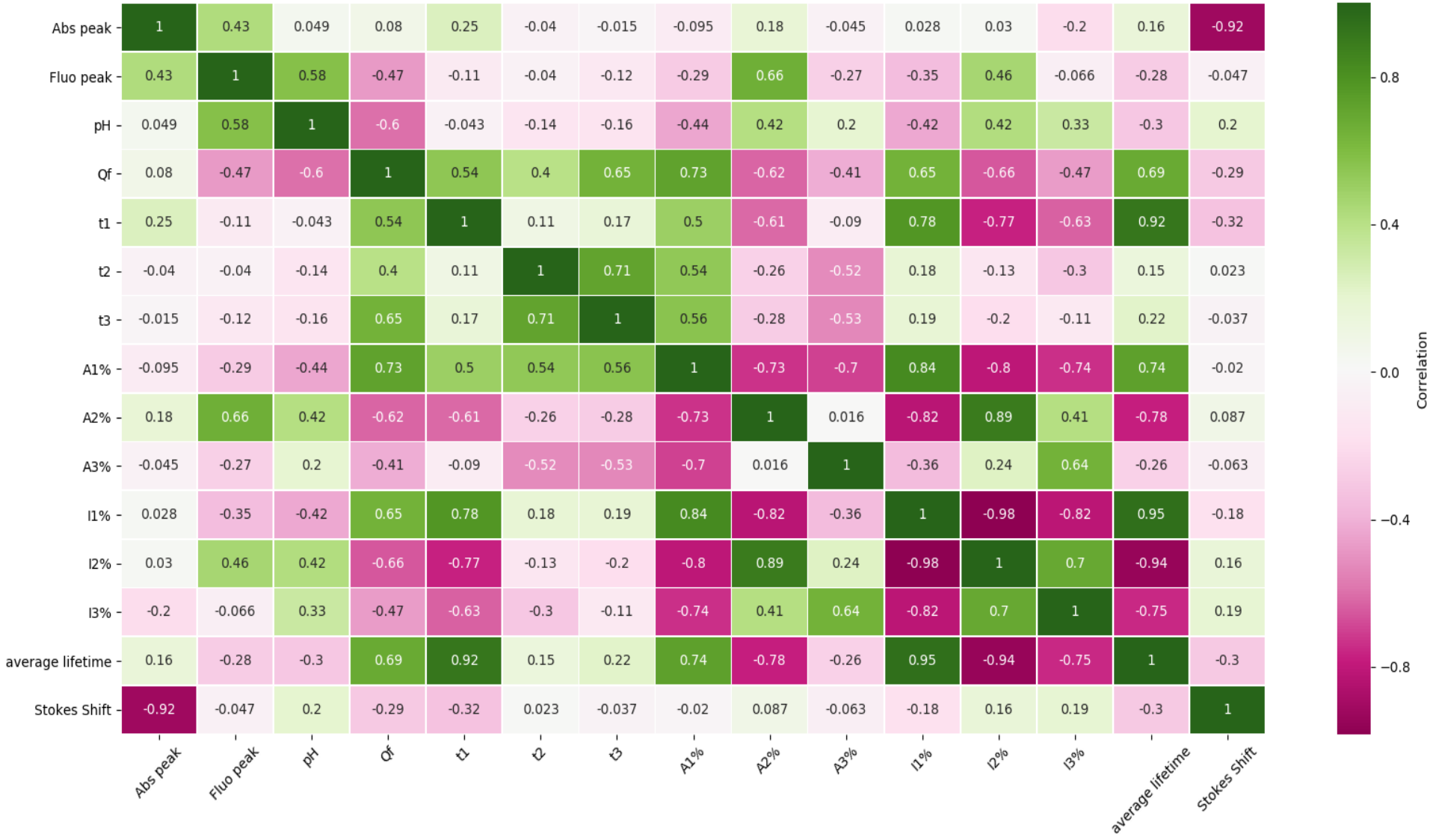
As commented by der Zitter et al[6], a correlation analysis can reveal important relations between parameters that might be difficult to directly infer otherwise. Therefore, the Pearson correlation coefficients have been calculated for the values reported in Table 7.1. The values have been put into a matrix and represented as a "heat map" in Table 7.2, where the dark green colors represent highly positive correlated variables (r= +1), white pale color represent non correlated variables (r= 0) and magenta highly negative correlated ones (r = -1).

The correlation matrix directly reveals expected positive correlation as, for example, between the fluorescent quantum yield (Qf in the correlation matrix; and in this chapter) with the average intensity lifetime or with the decay times. On the contrary, it also reveals very interesting correlations between parameters that are not so intuitively interconnected. For example, the fluorescence emission maximum (wavelength) has a negative correlation with the Qf and a positive correlation with A2 (contribution in amplitude to the emission decay of the second species). Another interesting correlation is between τ3 (lifetime of non-fluorescent species) and its pre-exponential factor A3. The increase of τ3 lifetime is correlated with a lower contribution of this component to the decay. Finally, it also needs to be mentioned that, while the lifetime of the fluorescent species (τ1) is not correlated with either of the other two decay times, the intermediate fluorescent species lifetime (τ2) and non-fluorescent species lifetime (τ3) are highly positively correlated. This correlation value is in line with the complexity of precisely determining two decay times for a three exponential fit when lifetime are similar[8].

Table 7.1. Photophysical properties for the 19 studied variants, WT rsEGFP2 and EGFP.

	λ Abs peak (nm)	λ Fluo peak (nm)	Stokes-shift (cm^{-1})	pK_a	Qf exc 470	Lifetime (exc 480 nm) Emission 510 nm			Contribution in amplitude (exc 480 nm; em 510 nm)			Contribution in intensity (exc 480 nm; em 510 nm)			average lifetime Intensity (ns)
						τ_1 (ns)	τ_2 (ns)	τ_3 (ns)	A1%	A2%	A3%	I1%	I2%	I3%	
EGFP	488	507	7679.4	6	0.60	2.87	1.39		86.9	13.1		93.2	6.8		2.77
V151A-S206N	486	503	6954.2	6	0.45	2.58	1.74	0.28	72.4	10.1	17.5	89.2	8.4	2.4	2.46
S206N	485	502	6982.4	6	0.38	2.36	1.1	0.26	70.3	13.7	16.0	89.6	8.1	2.3	2.21
rsEGFP2	482	502	8265.7	**6**	**0.34**	**2.81**	**0.47**	**0.15**	**41.9**	**15.6**	**42.5**	**89.7**	**5.6**	**4.7**	**2.55**
T204N	483	501	7438.5	6	0.32	2.81	0.45	0.14	41.5	15.1	43.4	90.1	5.3	4.6	2.56
V151L	483	501	7438.5	6	0.28	2.49	0.72	0.14	25.4	21.2	53.4	73.4	17.8	8.9	1.97
V151A	484	502	7408.4	6	0.27	2.77	0.33	0.11	33.0	23.5	43.4	87.9	7.4	4.7	2.47
H149F	497	507	3968.6	8	0.27	2.68	0.73	0.16	37.2	40.5	22.3	75	22.2	2.7	2.18
T204V	502	508	2352.8	8	0.26	2.97	0.5	0.18	37.4	24.2	38.4	85.4	9.3	5.2	2.59
T63A	482	503	8661.7	6	0.24	2.35	0.85	0.17	39.1	16.8	44.0	81	12.6	6.5	2.02
T63A-S206N	484	502	7408.4	6	0.22	1.72	0.75	0.15	27.1	37.6	35.3	58.1	35.3	6.7	1.28
T204A	501	502	397.6	8	0.19	2.75	0.56	0.1	31.5	9.1	59.4	88.7	5.2	6.1	2.47
H149W-S206N	475	508	13675.9	>10	0.10	2.23	0.6	0.12	53.8	22.7	23.5	88	9.9	2.1	2.02
T204S	484	504	8198.9	6	0.08	1.79	0.83	0.07	33.1	28.9	37.9	69.2	27.9	2.9	1.47
H149F-T204A	486	507	8522.7	>10	0.08	2.29	0.84	0.15	19.3	42.5	38.2	51.6	41.6	6.7	1.54
H149F-T204A-S206N	483	504	8626.6	>10	0.08	2.24	0.74	0.15	7.5	28.0	64.5	35.5	43.6	20.8	1.15
H149W-T204N	481	507	10661.5	>10	0.07	2.51	0.74	0.15	18.7	34.2	47.1	59.4	31.7	8.9	1.74
S206N-E223Q	486	504	7348.6	6	0.07	0.79	0.40	0.14	10.4	42.9	46.8	25.7	53.4	20.9	0.45
E223Q	478	503	10397.9	11.5	0.17	1.34	0.52	0.16	5.4	43.0	51.6	18.7	58.5	22.8	0.60
S206V	501	512	4288.3	10	0.17	2.15	0.76	0.14	11.8	52.3	35.9	36.2	56.7	7.1	1.22
S206A	492	510	7173.6	10	0.10	1.70	0.57	0.13	14.1	51.2	34.8	41.4	50.6	8.0	1.00

Table 7.2. Pearson correlation matrix of the spectroscopic properties in Table 7.1. Color code from green (r=1) over white (r= 0) to magenta (r = -1). (Representation made with Seaborn python package)



Principal component analysis

To further study the correlations between structural mutations of the rsEGFP2 WT protein and the photophysical properties in Table 7.1, a principal component analysis (PCA) has been done using Scikitlearn python package[8] (part of the code is shown at the end of the chapter). Such analysis transforms the real space into another space where the directions are orthogonal to the main direction, and in a such a way that the variance between samples points is minimized. The main direction or vector (PC1) explains the maximum variance of the sample. In that way, the PCA identifies linear combinations between properties. Before the PCA analysis, due to the huge differences in values between properties in Table 7.1, the different properties have been standardized, by subtracting the mean and dividing by the standard deviation. This means that the samples have been centered around 0 and have a standard deviation of 1. Therefore, properties with high values are not biasing the analysis. After applying the transformation of the data to the PCA space, we found that five principal components were required to represent 93.7% of the data variability, and eight for 99.3%. This highlights the complex relationships between the spectroscopic parameters. The distribution for the 20 proteins in the first two main principal components space can be seen in Figure 7.3a, where the form and color of the points represent the mutation. Figure 7.3b is the "biplot" representation of attributes contribution (protein properties) to each of the components. Finally, the circles in each of the figures represent the 95% confidence interval of the points which helps to identify mutations with clear distinct properties compared to the rest, and, in the case of the biplot figure, properties with high contributions to the principal component. The PC1-PC3 can be seen in Figure 7.3 and the PC1-PC4 at the end of the chapter (Supporting Figures 7S-1). Finally, the loadings of the spectroscopic parameters of the first four principal components are compared in Figure 7S-2.

How to read the PCA: the biplot figure

It is important to understand how to read these figures which are called biplot[9] figure (e.g. Figure 7.3). The biplot figure is a useful tool of data analysis and allows the visual evaluation of the structure of large data matrices. In a biplot, the samples (proteins) are displayed as points (Figure 7.3a) while the variables are displayed either as vectors, linear or axes (vector in this case; Figure 7.3b), in both cases the directions or axes of the figures are the PC.

Therefore, these figures transform the original matrix in a simple visualization in the PCA space. In these figures, the properties with the same direction are correlated with each other (e.g., Qf with A1 in PC1-PC2 Figure 7.3b). On the contrary, those in opposite directions are anti-correlated (e.g., the absorption maximum and the Stokes shift Figure 7.3b). Secondly, the correlations can be checked with table 7.2. Finally, the values for each individual variant in Table 7.1 can be compared with those of another variant with the used of their representations in the PCA space (e.g. Figure 7.3a). Samples in a similar region of the PCA space should have similar properties. Furthermore, it should be notice that even though the dashed circumference seems to be a circle, it is indeed an ellipse as the size of the axes are not identical. PC1 direction (x axes goes from -7 to 7) is indeed much higher than the PC2 x axes (from -4 to 4) which is in agreement with the variance ratio explained by each component.

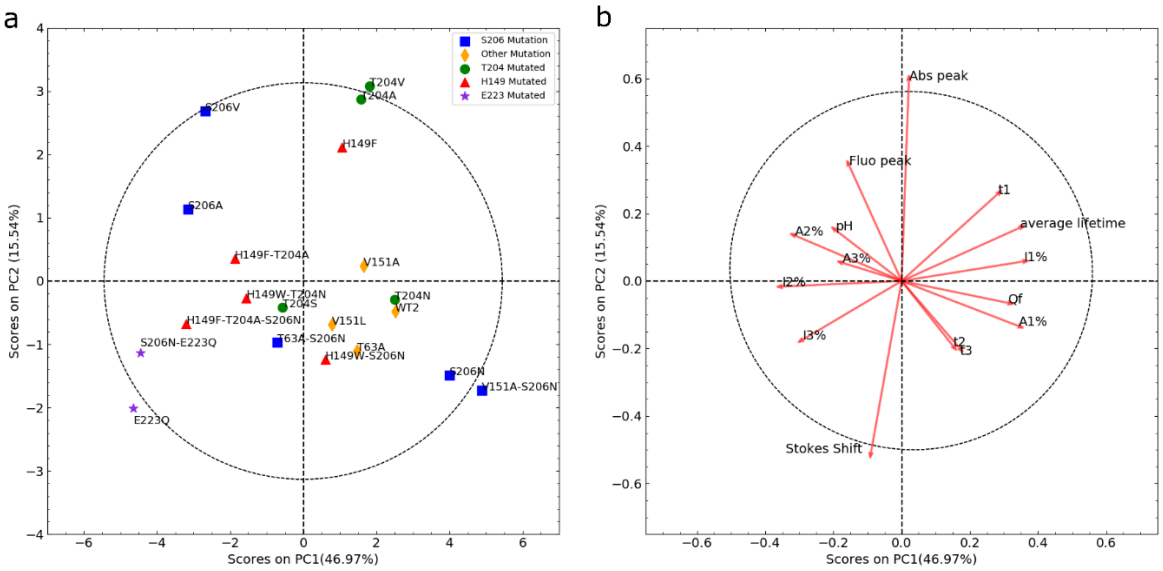


Figure 7.3. Principal component analysis. (a) Plot showing the scores for the 20 proteins "On fluorescent state" on PC1 and PC2. (b) Loadings of the spectroscopic parameters on the same principal components as in a), representing the so called biplot figure. Gray dashed circle represents the 95% confidence level of the points represented in each panel. Figure made with matplotlib python package.

PCA results

From these two representations in Figure 7.3, it is possible to attribute certain properties that force each of the 20 rsEGFP2 proteins into a certain area of the PCA space. Nevertheless, it is important to clarify that two variants with a mutation in a certain amino acid might lead to significant different positions in the PCA space, which is explained by the capabilities of the new amino acid to form Hydrogen bonds or electrostatic interactions with the chromophore. Finally, it also needs to be clarified that, although we will make some comparisons with the WT mutant, this one is not in the middle of the representation in the PC1 projection in Figure 7.3a.

Taking into consideration the above explanation and the biplot representation in Figure 7.3, some interesting correlations can be done. Firstly, as expected, we can see important anticorrelations between the absorption maximum and the Stokes shift, which is also revealed by the Pearson correlation matrix in Table 7.2. Furthermore, if we focus on the contribution of the different fluorescent species, we can see a strong correlation between A2 and A3 which in turn are anti-correlated with A1 and the fluorescent quantum yield (Qf). We can also found some interesting properties that have been defined by certain mutations. For example, the mutation of the T204 by amino acid with non-polar side chains, unable to form a hydrogen bond with the chromophore, will result in a red-shifted absorption maximum, and a nearly zero Stokes shift (Table 7.1). The same effect is found when the His149 is mutated to a amino acid with a non-polar side chain which cannot form a hydrogen bond with the chromophore (H149F variant). Similar to the conclusion drawn for T204 and H149, we can derive some important observations for the mutation of the S206. When it is mutated to an aspargine (N), which can form a hydrogen bond to the chromophore phenol group, the population of fluorescent species (A1) increases and thus the Qf. On the contrary, when this amino acid is mutated to an alanine or a valine which cannot form any bond to the chromophore, non-fluorescent species (A2 and A3) drastically increase in detriment of fluorescent species population (A1). This is followed with the corresponding decrease of fluroescence quantum yield.

The PC2 captures a slightly higher amount of information than the PC3. In the PC2 space, the absorption maximum and the Stokes shift have important weights (Figure 7S-2). On the contrary, PC3 captures nearly the same amount of variability as PC2 (15.54% and 14.87% for PC2 and PC3 respectively) and more importantly the weights of the different decay species A1, A2 and A3 and the decay times are high. Therefore, the same representation was done for PC1-PC3 in Figure 7.4. With the study of the first three PC we have a clear overview of the variant's properties as the first three PC together capture nearly 80% (77.4%) of the data variability in Table 7.1. Interestingly, we can see that A3 is anti-correlated with τ2 and τ3, which indicates that the smaller the contribution of A3 is, the higher the τ2 and τ3 decay times will be and vice-versa. The anti-correlation between A3 and τ3, indicates that, the higher the contribution of this short component (A3) to the excited state is, the faster is the decay (τ3). This is an important observation since τ3, and especially the contribution of this decay time to the excited state population (A3 values), are parameters that highly decrease the Qf.

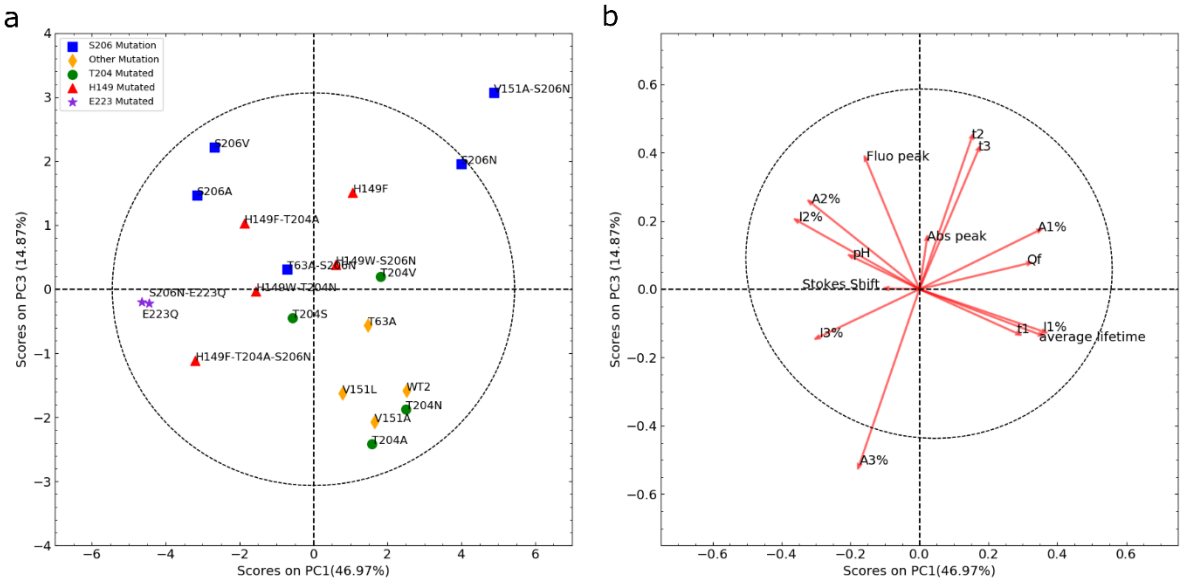


Figure 7.4. Principal component analysis. (a) Plot showing the scores for each RSFP "On fluorescent state" on PC1 and PC3. (b) Loadings of the spectroscopic parameters on the same principal components as in a, representing the so called biplot figure. Gray dashed circle represents the 95%confidence level of the points represented in each panel. Notice that even though the dashed circumference seems to be a circle, is indeed an ellipse as the size of the axes are not identical. Figure made with matplotlib python package.

Considering the representation of the RSFP proteins in the PC1-PC3 space, we can conclude that WT V151A, V151L and the proteins close to them have high contribution on the emission decay of non fluorescent species (A3) and a low contribution of A2, which results in a relatively high Qf. Fluorescent species contribution (A1) increases and becomes the major population in the emission decay. On the other hand, a mutation of the Ser206 by a amino acid with a non-polar side chain increases the intermediate lifetime (A2) contribution. Finally, through this representation, we can see that the E223Q mutants have a high contribution of A3 and I3, indicating that decays mainly via the non fluorescent species (high values of A3), in agreement with a low fluorescence quantum yield (Qf). A similar conclusion can be deduced for the triple mutant H149F-T204A-S206N. An interesting difference between the representations of the RSFP variants in the PC1-PC2 (Figure 7.3a) or the PC1-PC3 (Figure 7.4a) spaces is the T204A and T204V. We can see that, whilst they are close in the first space, they are separated in the second. This can mainly be explained by their similitudes in the absorption and Stokes shifts, but the differences in the non-fluorescent (A3) contributions. Furthermore, while the T204A has a high non fluorescent contribution, the T204V does not. Overall, we can see that mutating the S206 by an asparagine (N) increases the contribution of the fluoescent species (A1) which results in an increase of the Qf. On the contrary, we can see that the mutations of the T204 or the H149 by amino acids that cannot form an hydrogen bond to the chromophore makes the contribution of non fluorescent species (A3) and A2 (intermediate life time) to significantly increase, respectively. Therefore, we can see a decrease of the Qf for these mutants. More analysis will be done by adding X-ray structure of the different proteins (under progress)

7.2.1 Characterization of switching properties for selected variants

After identifying the mutations that allow us to control the different contributions to the excited state populations in rsEGFP2, a selection of these variants and several combinations with each other, were selected. *Off* state spectra and *On* to *Off* switching yield for these variants were measured (highlighted in green in Table 7.1; E223Q which is still currently under progress). For all the variants highlighted in green in Table 7.1, the *Off* state spectrum were recorded and the thermal back reaction measured. The results of these measurements

can be seen in Table 7.3 and the absorption and emission spectra in Figure 7.5. For some of them, the photoswitching quantum yield of *On* to *Off* could be determined.

The absorption spectra in Figure 7.5 were recorded in a Cary3500 spectrometer, except for H149F-T204A and H149F-T204A-S206N which, due to the fast thermal recovery, were recorded with a CCD camera. As shown in Table 7.3, before recording the spectra in Figure 7.5, the pH was adjusted according to the pK_a values reported in Table 7.1 to displace the proton equilibrium and ensure that the main form is the *cis* anionic (*On* form). After switching the proteins to the *Off* state, we could see the formation of a positive band around 400 nm for all the variants. Interestingly, a second band over 515 nm is formed for H149F-T204A, which became clearer for H149F-T204A-S206N. Moreover, if the H149F *Off* spectrum is analyzed in detail one can see a small increase of the absorption in the 515 nm region. For these variants the two bands formed (≈400 nm and ≈515 nm) are assigned to trans-neutral (blue maximum) and trans-anionic (red maximum) forms. The existence of a trans anionic form is due to the basic pH used for these measurements. This band is similar to the TRUV-Vis flash-photolysis spectra recorded at 1 µs excitation after correcting for the GSB shown in the previous chapter Figure 6.11, which has also a red-shifted maximum in respect to the *cis*-anionic form. Therefore, these results support the assignment done in Chapter 6 where the first ground state after excitation was assigned to a trans-anionic chromophore.

Interestingly, the mutation T204A showed a unique feature among the nineteen mutants studied. There is a difference in the emission maximum with the excitation wavelength. In fact, when the mutant is excited at 450 nm, the fluorescence maximum is at 501 nm, while when excited at 470 nm, the fluorescence maximum is red shifted to 507 nm. This result are attributed to different *cis* ground state conformers and reminds to the two emission maxima for A and B avGFP forms[10]. Finally, it is worth mentioning the low molar absorption coefficients values for the H149F-T204A-S206N and H149F-T204A mutants. Note that the absorbance of the *cis* anionic form is decreasing with increasing pH. In the future one should measure the dependence of the switching properties with the pH.

Table 7.3. *Off* state properties of rsEGFP2 variants highlighted in Table 7.1.

	pH	**Abs ON** λ_{max} /ε nm /(M^{-1}, cm^{-1})	**Abs Off** λ_{max} / ε nm /(M^{-1}, cm^{-1})	**Fluo λ_{max} Exc 450** nm	**Fluo λ_{max} Exc 470** nm	**Thermal recovery**	**Φs (%) On-Off**
V151A -S206N	8	486(55309)	405(24273)	502	503	38.2 h	0.60
S206N	8	485(66614)	405(26262)	501	502	3.31 h	0.39
rsEGFP2	8	483(67210)	408(26443)	501	502	1.81 h	0.40*
H149F	10	496	405	505	507	**3.7, 84.7 min**	-
T204A	10	501(95227) **	400 (20921) **	**501**	**507**	71 s	0.20
H149F-T204A	11.5	400/482 (4160/9707)**	396/515 (-/4298)**	506	506	130, 389 s	2.5
T204A-H149F-S206N	11.5	482(1690)**	513(787)**	506	504	183 s	38.1

*Note the value for WT is different than that reported in chapter 4, as this value is obtained from only fitting the *On* to *Off* data sets and that in chapter 4 from globally fitting the *On* to *Off* and *Off* to *On*. Therefore, results in chapter 4 is much more reliable.

** Molar absorption coefficient values are determined by comparison to WT using the absorbance at 280 mn, and not by denaturation.

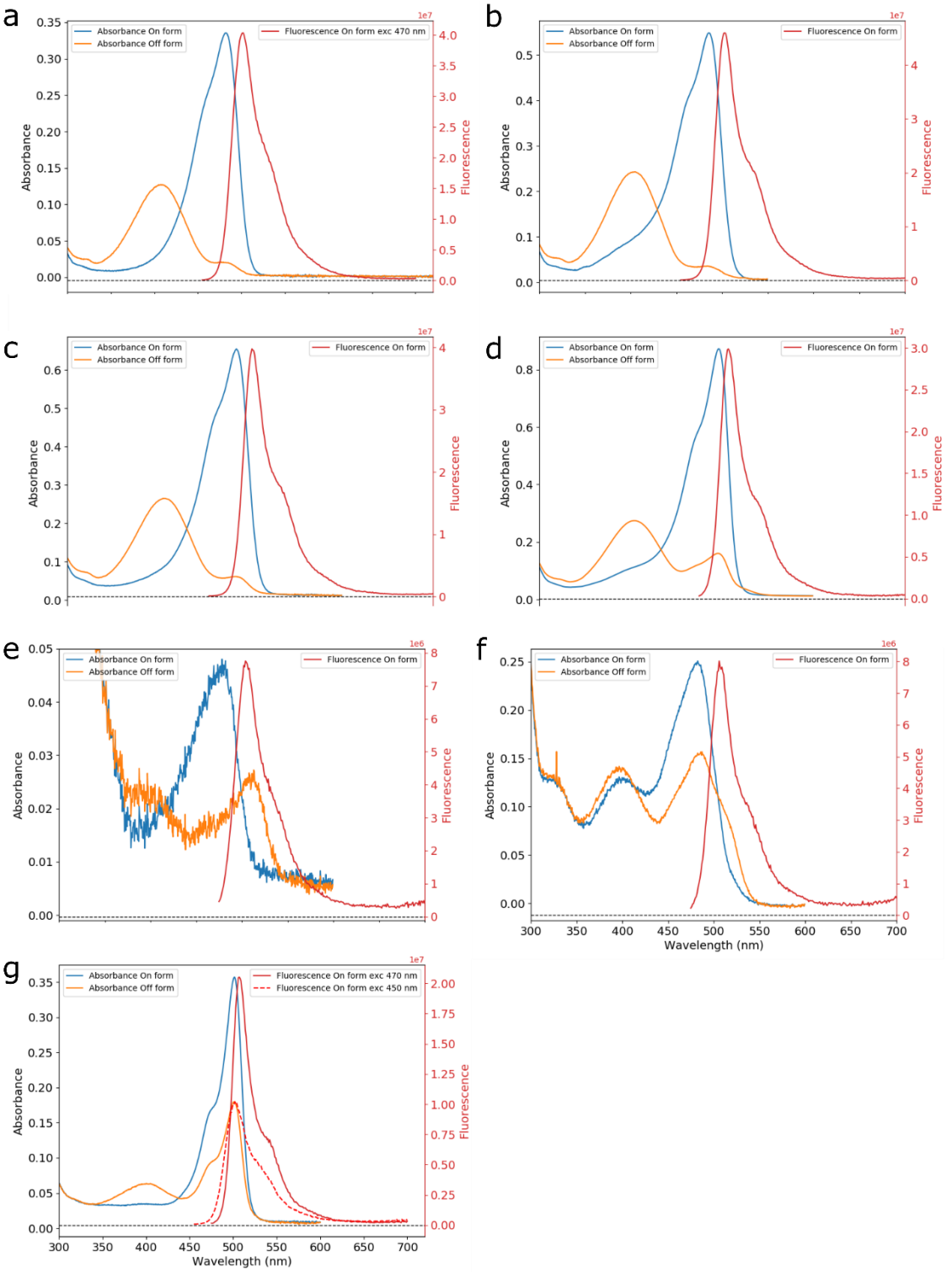


Figure 7.5. Absorption spectra of On (blue) and Off (orange) forms together with the fluorescence (red) of the different rsEGFP2 variants. (a) WT, (b) S206N-V151A, (c) S206N, (d) H149F, (e) H149F-T204A-S206N, (f)) H149F-T204A and (g) T204A. Figure made with matplotlib python package.

To determine the thermal recovery, starting from the photo-stationary *Off* state, we monitored the absorption maximum of the *On* and *Off* states for the different variants. The recorded traces for all the variants were then fitted with either a single or a double exponential function to recover the *On* state. (the traces and fits can be seen at the end of the chapter in Figure 7S-4). For H149F-T204A and H149F-T204A-S206N, we monitored the spectra between 320 and 620 nm to ensure a better study of the *trans* neutral and *trans* anionic forms of these two variants. In both cases, the recorded spectra were globally fitted with a single or double exponential function. The results for these two variants can be seen in Figure 7.6.

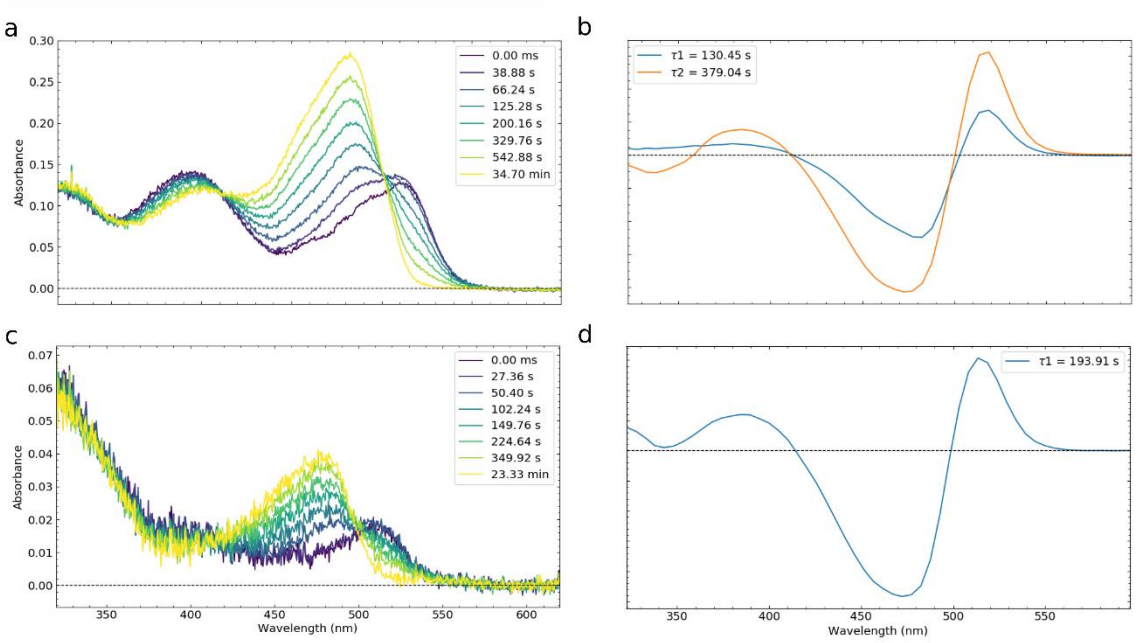


Figure 7.6.(a) and (c) Spectra recorded during the thermal back reaction of the rsEGFP2 variants H149F-T204A and H149F-T204A-S206N respectively. The proteins were initially irradiated to reach the photostationary Off state with a 490 nm LED (Thorlabs). (b) and (d) correspond to the decay associated spectra obtained from the global fit of the data in panels (a) and (c) respectively. The corresponding traces and fits can be seen at the end of the chapter in Figure 7S-3. Figure made with Ultra pyfit.

Remarkably, the thermal back recovery time constant showed a wide range of values among the variants. The mutation S206N (3.31 h) barely affects the thermal recovery as this variant has relatively similar time when compared to the WT (1.81 h), similarly the S206N-V151A (38.2 h) and the V151A (40.3 h) thermal recoveries are also comparable. On the contrary, all

the other variants exhibit much faster recovery times (several seconds) except for H149F which had two clear recovery times of 3.7 and 84.7 minutes. The fastest recovery time was found for T204A, being 71 s. Finally, at least two times were needed to correctly fit the traces of H149F-T204A, while only one was needed for H149F-T204A-S206N. Although in these mutants two bands at around 400 nm and at 515 nm are formed after irradiation, which can be attributed to the *trans* anionic and neutral forms respectively, it is clear that only one time constant is needed for H149F-T204A-S206N (Figure 7.6b,d). Moreover, and even though two times constants were needed to fit the H149F-T204A data, the spectral signature of these components in Figure 7.6a, reveals that the evolutions of the anionic and neutral trans forms occur concomitantly and in identical times, and thus the two time evolutions cannot be attributed to either the neutral or to the anionic form but to both of the conformers. After determining the thermal back reactions and the molar absorption coefficient for all the variants, it was possible to measure the *On* to *Off* photoswitching quantum yield. The procedure has been explained in chapter 4.5. For these mutants, the *On* to *Off* photoswitching experiment did not represent any problem. On the contrary, due to the fast recovery of the *Off* state, the *Off* to *On* measurements could not be accurately recorded for several mutants. Therefore, for all these mutants only the *On* to *Off* photoswitching quantum yield was calculated. The results are shown in Table 7.3, and the fitting results can be seen in Figure 7.7 (note that the H149F and E223Q variants have not been measured yet).

It is important to mention that the value for WT reported in Table 7.3 is different than the one reported in chapter 4, obtained from globally fitting the *On* to *Off* and *Off* to *On* together. This difference shows the complexity of these measurements, especially when determining values that are under 1% for switching yield, where the errors in the determination of thermal back reaction and the molar absorption coefficient are added to those intrinsic of the measurement and those of the fit. Nevertheless, we can see that whilst all the single point mutations have similar values to the one obtained for the WT, the H149F-T204A and H149F-T204A-S206N have increased switching quantum yields of over 10 and 100 times, respectively. These values are correlated with a high contribution of the non fluorescent species. Finally, it is worth mentioning that the extremely fast thermal recovery of the T204A variant causes the absence of formation of the *Off* form, as shown in Figure 7.7b

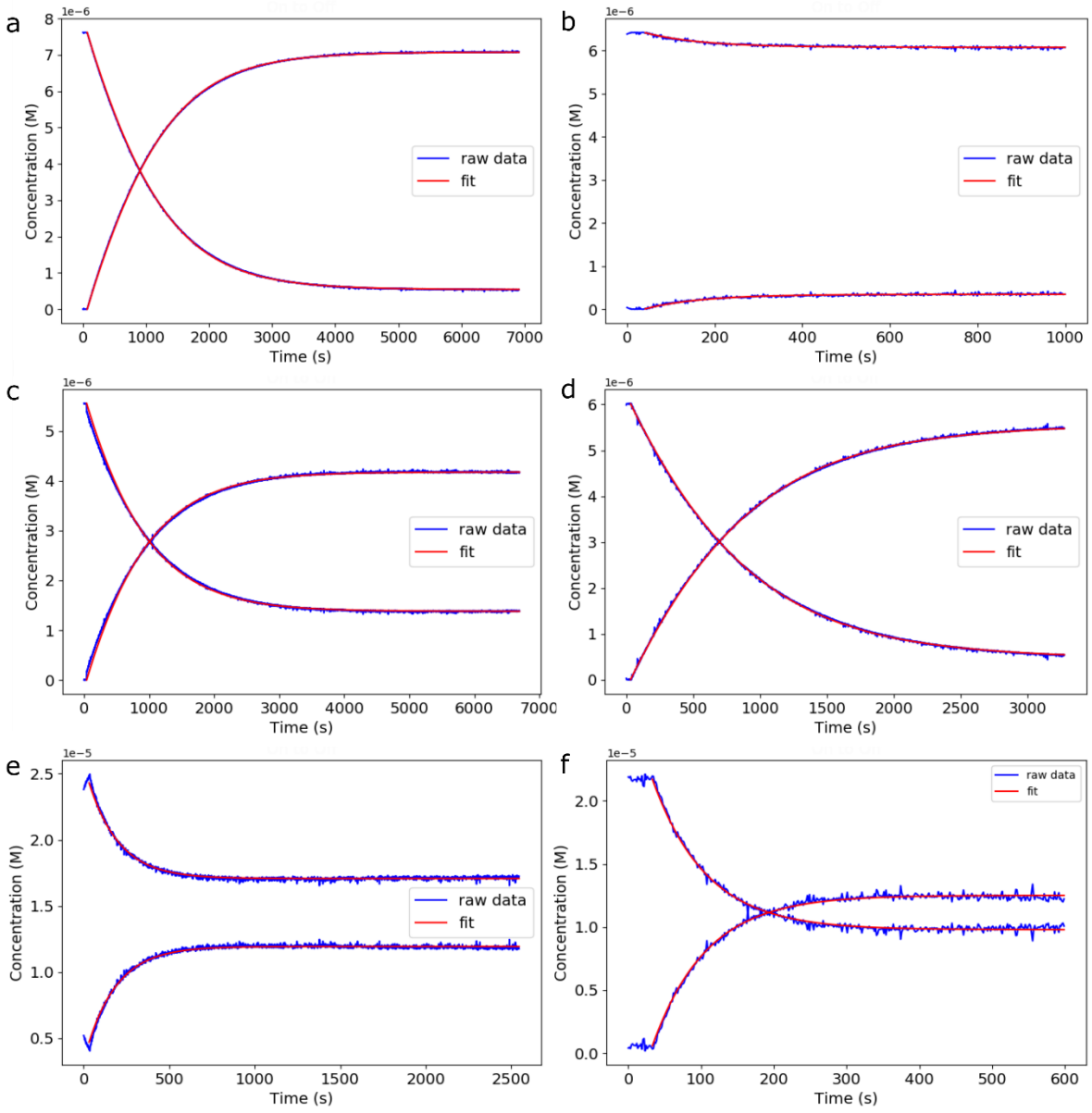


Figure 7.7. On to Off (blue) profile concentrations obtained from the continuous irradiation experiments together with thecorresponding fits (red) to equation 4.10 for the different rsEGFP2 variants. (a) WT, (b) T204A, (c) S206N, (d) S206N-V151A, (e) H149F-T204A and (f) H149F-T204A-S206N. Figure made with matplotlib python package.

7.3 Correlations between structure and properties

In this chapter, the *On* state of nineteen rsEGFP2 variants have been spectroscopically studied, and their fluorescence intensity decay recorded. In the previous chapter, the different exponential decays were attributed to at least two well-defined ground state conformers of the protein in the *cis* anionic *On* form. The heterogeneity is attributed to different hydrogen bond patterns of the phenolate with a structural water and residues of the protein cage. From the PCA analysis performed on the values in Table 7.1, several principal components were obtained. The representation of the different variable loadings in the PC1-PC2 space reveal the correlation between (Qf) and other variables. Particularly interesting is the correlation with A1 (fluorescent species) and its anti-correlation with A2 and A3 (non fluorescent species, Figure 7.2b) which manifest that the only way to increase the Qf is by decreasing competing non-radiative decay paths (i.e.: indirectly increase A1). The representation of the different mutants in the PCA space reveals that there are two main areas that exhibit different behaviours: either the S206-E223 region, and the H149 and T204 region (Figure 7.2). Nevertheless, similar variants with same amino acid being mutated give different photophysical properties and are in completely different areas in the PCA space. This points that the photo-physical properties of the final substituent and the new possible interactions with the chromophore are more important than the mutated position and thus difficult to predict.

The S206-E223 area: the proton relay

As already discussed in the introduction, the first area (S206-E223) represents one of the most studied amino acids in the avGFP protein family as they are involved in the proton relay. Already since the first X-ray avGFP structure[11], represented in Figure 3.2 and 3.10 (chapter 3), the origin of fluorescence in avGFP was attributed to the proton transfer of the phenolic proton via an ESPT[12] through a chain starting from a water molecule close to the phenyl ring, followed by the S206 and from there to the glutaminic acid (E223)[13,14]. The mutation of the E222 (E223 in rsEGFP2) by a glutamine (Q) renders the avGFP non-fluorescent[15]. Our results in the E223Q variant reveal that, although the fluorescence is not completely lost, it is drastically reduced. We can clearly see from the values in Table 7.1 that this decrease of fluorescence is provoked by an increase of the non fluorescent species (A3)

population with the corresponding reduction of the population of fluorescent species (A1). This decrease of Qf is further stressed by the decrease of the average excited state lifetime. The second rsEGFP2 variant with a mutated amino acid in this area is the S206N. Interestingly, for the two variants S206N and V151A-S206N, we can find the highest populations for the fluorescent species (A1) among all the twenty (nineteen variants + WT) studied proteins. Furthermore, we can also find the highest lifetime for intermediate species ($\tau 2$). This explain the highest Qf values for V151A-S206N followed by the S206N. Structurally, this can be correlated with a stabilization of the proton relay between the chromophore and the E223, leading to a more stable environment around the chromophore attributed to a stronger steric hindrance of the asparagine (N) side chain compared to the serine (S). It is worth noting that other RSFPs derived from avGFP as rsEGFP[16] or rsGreen[17] have an asparagine at 206 positions. Interestingly, when S206 is mutated to an alanine or a valine, we can see a clear decrease of the Qf, which is explained by an increase of the intermediate lifetime contribution. These two amino acids cannot form a hydrogen bond with the water molecule, and therefore the classical proton relay is broken. These two mutations reveal the importance of the position 206 as a key connector between the chromophore and the E223. In fact, studies in the avGFP have shown that the back proton transfer from I to A, occurs via a deep proton tunnelling where the S206 plays a major role[18]. Moreover, previous studies on similar mutants in the EGFP S206A[19] and S206V[20] have shown that an alternative proton transfer pathway is formed involving the chromophore hydroxyl, a bridging water molecule, T204 and E223. This alternative path has a slower proton transfer, explained by the long (~3.2 Å and presumably weak) hydrogen bond, between T204 (T203 in EGFP) and the water molecule, compared to the 2.7 Å normal hydrogen bond between the water molecule and S206 (S205 in EGFP)[19,20]. Regarding the double variant E223Q-S206N, one may expect a mixture effect between the increase of Qf caused by the S206N mutation, and a decrease provoked by the E223Q mutation. Interestingly, the E223Q-S206N variant has the lowest Qf of all nineteen variants studied. The influence of multiple mutations resulting in a more substantial negative effect than the sum of individual mutations have already been reported[21].

The H149-T204 area: Quinoidal form

The T204 mutation has also been previously studied and, indeed, corresponds to the Type-III (chapter 3) classification done by Tsien in 1998[10]. In general, this mutation causes a red-shift of the absorption maxima. In the avGFP family, the T204 is an important amino acid as in the crystal structure of monomeric wild-type GFP, T204 has two conformations: with the OH facing away from the HBDI phenol oxygen (85 %), and 15% with the OH rotated towards it[11]. This rotation was attributed to the photoconversion of neutral to anionic forms[12] (Figure 3.10 in chapter 3). Furthermore, in the EGFP, Gosh et al[4] revealed that the *On* state is formed by three ground state conformers, two fluorescence attributed to the different conformations of the E222, and one non-fluorescence attributed to the rotation of the T204. The variants with the T204 mutated, have the expected red-shifted absorption maximum when the T204 is the single mutation, and the mutated amino acid have a non-polar carbon side chain (T204A and T204V) but not when it is mutated to another amino acid with a polar side chain able to still form a hydrogen bond with the HBDI (T204N and T204S). An interesting feature of the T204A variant is that it has the major A3 (short decay) contribution to the excited state lifetime (Table 7.1 A3 value) with nearly 60% (except for the triple mutant H149F-T204A-S206N with 64.5 %). Furthermore, this variant also displays the fastest thermal back reaction (nearly in less than a minute). Jung et al[22] reported the T204V mutations in avGFP and attributed the absorption red-shift to the formation of the quinoidal form of the chromophore (Figure 7.8). As described in the introduction (Chapter 3, Figure 3.12), the HBDI chromophore can be treated as mixed valence compound that can either have the charge on the oxygen of the phenol group or on the oxygen of the imidazolinone moiety. Using the Marcus−Hush theory for mixed-valence compounds[23] for electron transfer, the Stokes shift has been perfectly correlated with the absorption maxima by S. G. Boxer and coworkers[24]. The Stokes shift in the proteins is related to a charge transfer in the excited state, which displaces the charge from the imidazolinone to the phenol group. Interestingly, in a quinoidal HBDI form, this displacement is already done[22]. The HBDI phenol group forms three hydrogen bonds with the protein cage in rsEGFP2, these are with the T204, the H149 and a water molecule. Our results with the different T204 variants demonstrate that out of these three hydrogen bonds, the HBDI-T204 interaction is the one stabilizing the benzoidal form (Figure 7.8) versus the quinoidal form.

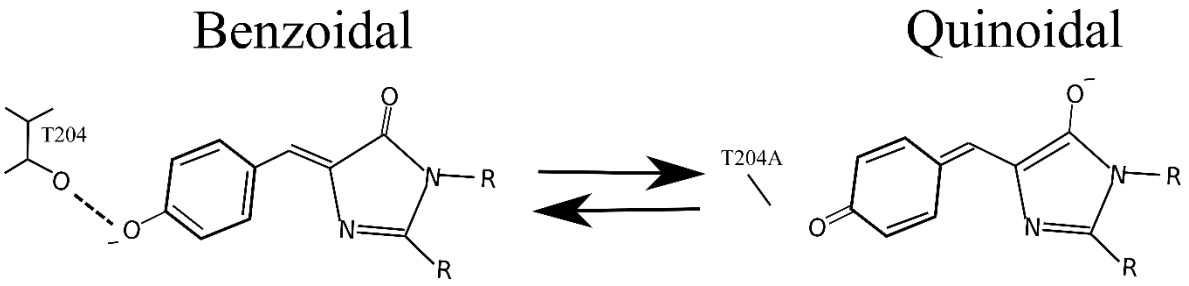


Figure 7.8 Benzoidal and quinoidal mesomers of the HBDI chromophore in its deprotonated form. Figure done with Inkscape.

Finally, the H149F variants also have important differences as the histidine interacts directly through a hydrogen bond (2.82 Å) with the HBDI phenolate, helping to stabilize the cis conformation. The introduction of a phenylalanine at this position breaks this interaction as the apolar aromatic chain cannot form an hydrogen bond with the chromophore. This variant also displays a red-shift in the absorption spectra, which can be attributed to the same stabilization of the quinoidal form as for T204A and T204V. In fact, when represented in the PC1-PC2 space, where PC-2 is mainly affected by the absorption maximum and Stokes shift values, the H149F variant is the closest variant to the T204A and T204V. Furthermore, the mutation of the histidine by a phenylalanine increases the intermediate lifetime population A2. It is worth noticing that this variant is the only one from the nineteen studied that has two clear separate thermal recovery times with two different time constant of 3.7 and 84.7 minutes. The reason behind these two different thermal recovery times is still elusive.

Multiple mutations:

Due to their spectroscopic properties of the double and triple mutation, H149F-T204A and H149F-T204A-S206N are quite different when compared to the single point mutation variants. What makes these two variants particularly interesting is the fact that two direct hydrogen bonds established in the WT are suppressed, which leads to an increase of the pK_a values for both variants. For the H149F-T204A variant, the absorption spectrum in Figure 7.5f and 7.6a has important contributions at between 390-440 nm. This reveals that the *On* ground state of this variant is formed by at least two different forms, where the major contribution is a *cis* anionic form (480 nm), and the band at 400 nm represents at least one neutral (probably *trans*) form. The high absorbance values between 420-440 nm may indicate

that more than one trans neutral conformer populates the *On* state. Contrary to the rest of variants after irradiation, we can see the rise of a band at around 515 nm which is attributed to the trans anionic form of the chromophore, together with a not so prominent raise of the band at around 400 nm corresponding to the *trans* neutral form. For this variant, two time constant were needed to fit the thermal recovery. From the associated spectra in Figure 7.6a,c it is clear that the evolutions of the anionic and neutral *trans* forms occur concomitantly and with identical times, and thus the two evolutions times cannot be attributed to either the neutral and anionic form, but rather to both states. On the contrary, for H149F-T204A-S206N, we can only see a clear band at 483 nm, suggesting that the ground state is only formed by a single cis anionic chromophore. Similar to the H149F-T204A, after irradiation we can see the formation of a band a 515 nm which is also attributed to the *trans* anionic form. Interestingly, these variants have very low Qf values. This is not a surprising result as it has been demonstrated that the origin of the HBDI chromophore fluorescence is the electrostatic interactions of the chromophore with the protein cage preventing it to going under ultrafast internal conversion (IC) via *cis-trans* isomerization after excitation, which is the case of the HBDI in soltuion[25–29]. Blocking the IC for example at very low temperatures 77 K[27,30] and encapsulated in non-protein scaffolds[31,32] makes the chromophore fluorescent again. Therefore, it is also not surprising that these two variants have the highest *On* to *Off* photo-switching quantum yield (IC process) of all the studied variants and low Qf (shown in Table 7.3.) as several interactions between the HBDI and the protein cage are broken.

7.3 Conclusion.

This chapter summarizes the main spectroscopical characterization of nineteen different rsEGFP2 variants, containing up to three single point mutations, and all concerning amino acids close or directly hydrogen bounded to the HBDI. For all the studied variants, the chromophore adopts in the resting state a *cis* anionic conformation which is the fluorescent form. The TCSPC results confirm that the ground state is composed of at least two different conformers and that at least three times constants are needed to fit the fluorescent decay recorded for all twenty variants. It is important to note that $\tau 2$ and $\tau 3$ are highly positively correlated and make their analysis within the 20 proteins difficult. The PCA analysis revealed that the Qf is correlated with the population of the fluorescent species A1 and anti-correlated with A2 and A3 (non fluorescent species). According to the PCA analysis, we have identified two important areas near to the HBDI chromophore: either the S206-E223 involving changes in the proton relay between the HBDI phenol and the E223, and the H149 and T204 area which causes changes in the ground state of the chromophore stabilizing the HBDI into its quinoidal form. From the PCA analysis we can conclude that the quinoidal form increases the population of the non-fluorescent species (A2 and A3) respectively decreasing the fluorescent quantum yield. On the contrary, the S206N mutation drastically increases the population of the fluorescent species which we have tentatively attributed to a stabilization of the chromophore cause by the steric hinderance of the asparagine. Therefore, as a main conclusion, we can say that from all the twenty studied proteins (nineteen variants + WT), the combination of the S206N mutation, together with the V151A mutation (S206N-V151A variant) is the most interesting one to test in cells and then explore its possibility as a new fluorescent probe in RESOLFT. Finally, it should be mentioned that the ensemble of our results reveals important differences between variants with new apolar amino acid side chains which cannot form any electrostatic interaction with the chromophore and those that can still form or even strengthen the interaction (e.g., comparison between S206N and S206A) demonstrating that the behaviour of the final substituent and the new possible interactions with the chromophore are more important than the mutated position.

7.4 Supporting Figures.

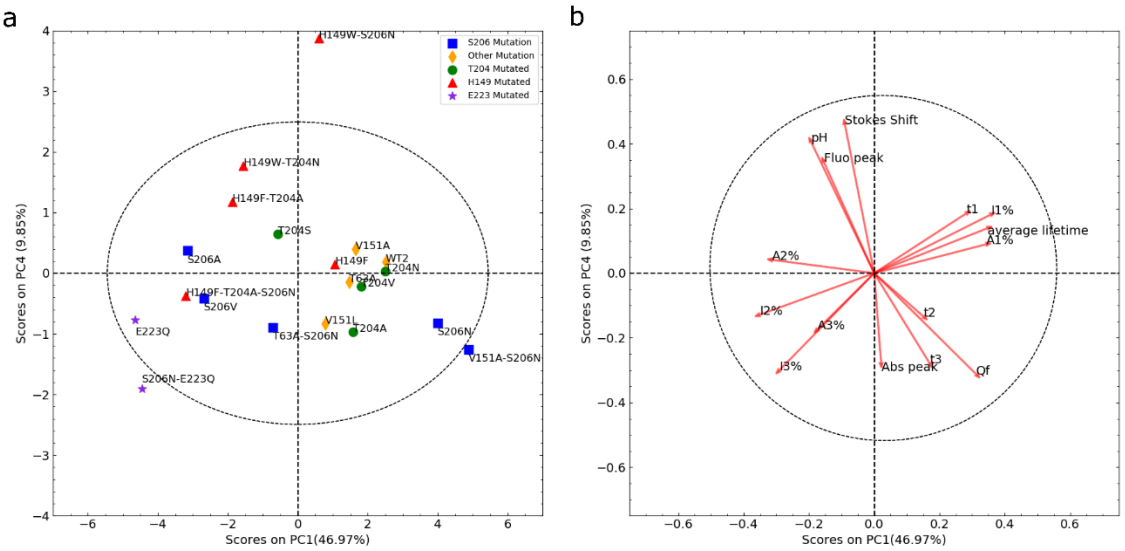


Figure 7S-1. Principal component analysis. (a) Plot showing the scores for each RSFP "On fluorescence state" on PC1 and PC4. (b) Loadings of the spectroscopic parameters on the same principal components as in (a), representing the so called biplot figure. Gray dashed circle represents the 95%confidence level of the points represented in each panel. Notice that even though the dashed circumference seems to be a circle, is indeed an ellipse as the size of the axes are not identical and PC1 direction (x axes) is indeed much bigger than the PC4 axes, as is logical from the ratio of variance explained by each component. Figure done with matplotlib python package.

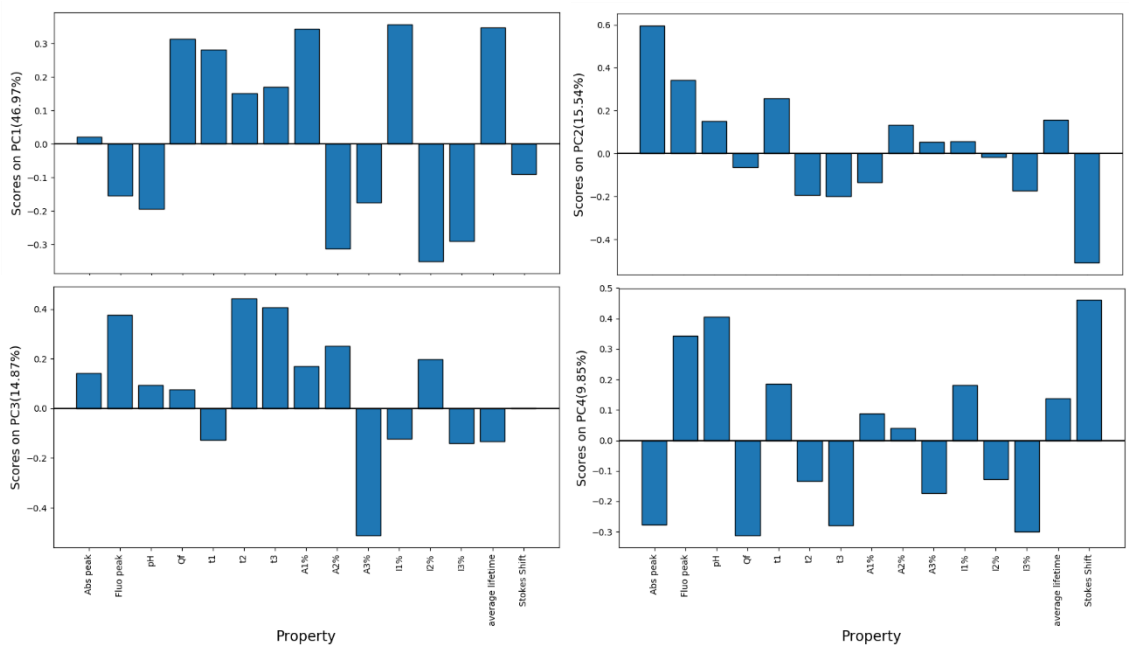


Figure 7S-2. Loadings of the spectroscopic parameters of the first four principal components. PC1 top-left, PC2 top-right, PC3, bottom-left and PC4 bottom-right. Notice that the first 4 PC explain nearly 90% of the Table 7.1 parameters variance. Figure done with matplotlib python package.

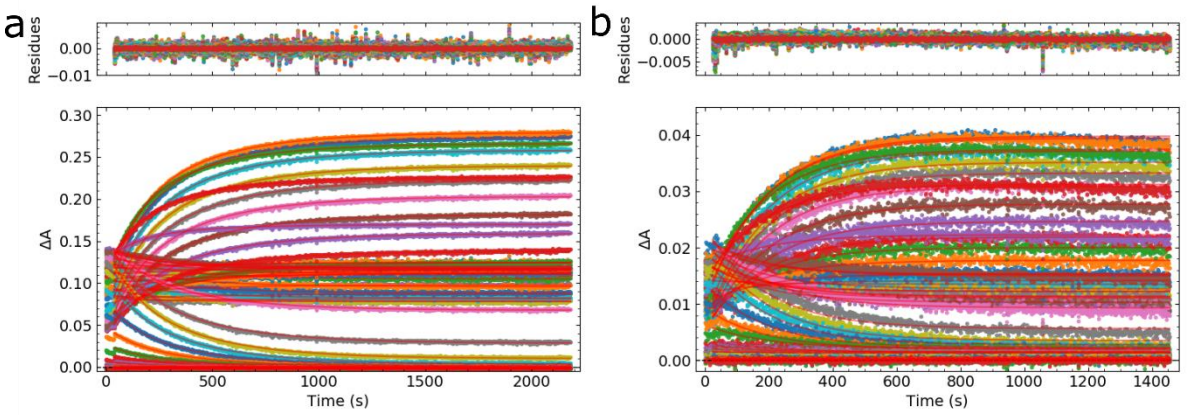


Figure 7S-3. bottom panel: Resulting fit of sum of exponential decay function in red and raw decay traces for the thermal recoveries sets recorded for H149F-T204A and (a) H149F-T204A-S206N (b). Top panel: the residues of the fit which are the differences between the fit and raw data. Figure done with Ultra Pyfit.

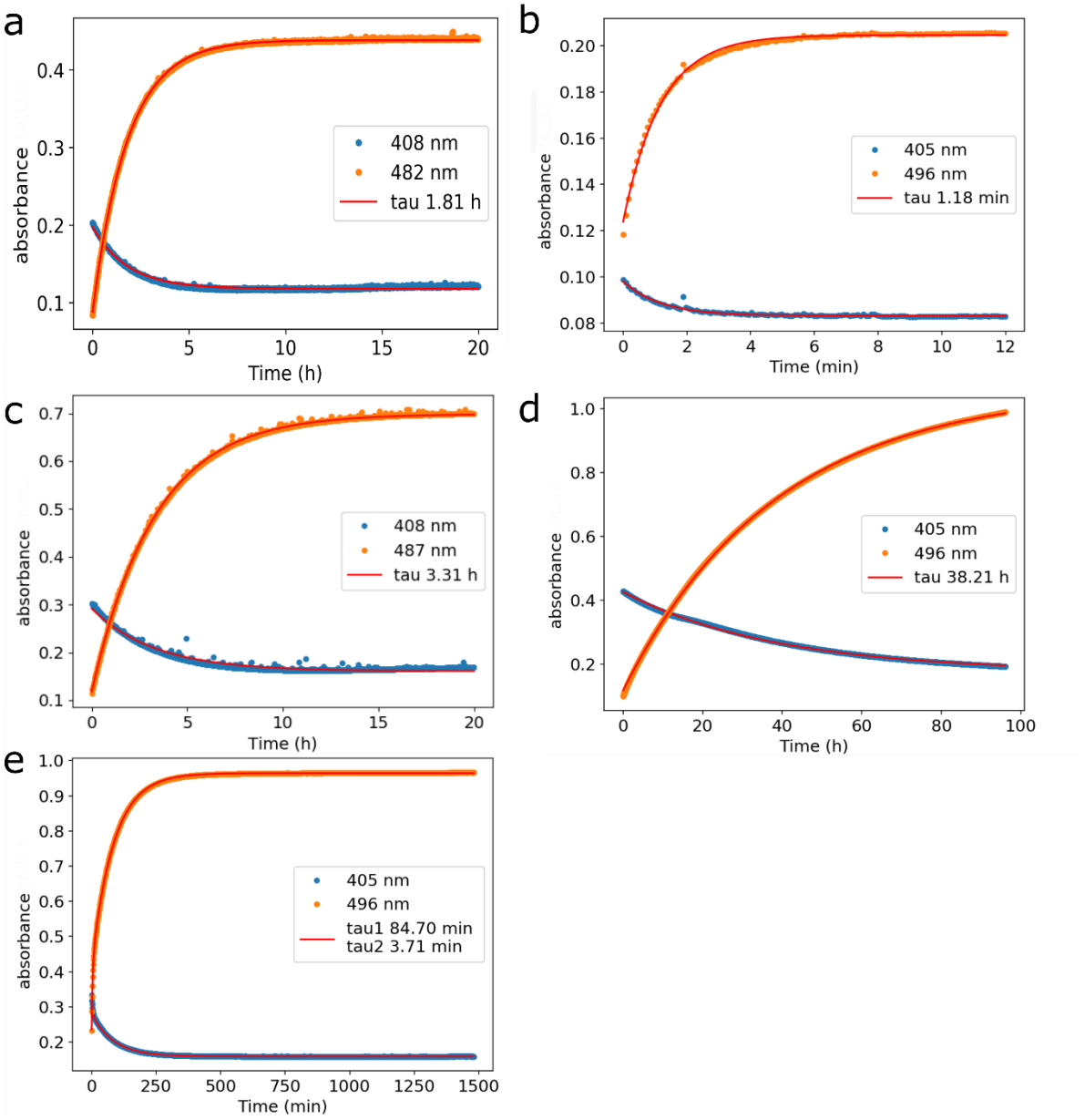


Figure 7S-4. Thermal recovery trace of Off (blue) and On (orange) absorbance maxima together with fitted traces (red) of the different rsEGFP2 variants. (a) WT, (b) T204A, (c) S206N, (d) S206N-V151A and (e) H149F. Figure done with matplotlib python package.

7.5 Python 3.7 simplified code for PCA analysis:

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.preprocessing import StandardScaler
from sklearn.decomposition import PCA
from matplotlib.patches import Ellipse

#import data
File = './Sumary preoperties On form3.txt'
Data = pd.read_csv(file, delimiter='\t', index_col=0).dropna(axis=1)

#Calculate Stokes shift in cm-1
data['Stokes Shift'] = abs(10**7/data['Fluo peak']-10**7/data['Abs
                                                        peak'])

data=data.drop(data.dtypes[data.dtypes=='object'].index, axis=1)

#instantiate Standar scaler
scaler = StandardScaler()

#instantiate PCA
pca = PCA()

#Standarize data
data2 = pd.DataFrame(scaler.fit_transform(data))

#Fit PCA
pca.fit(data2)

#Transform original data into PCA spacae
data_pca = pd.DataFrame(pca.transform(data2))

data_pca.index = data.index

#Explore results and variance explained
explained_var = pca.explained_variance_ratio_
data_pca.columns = [f"PC{i}" for i in range(1, data_pca.shape[1] + 1)]

print(pca.explained_variance_ratio_.cumsum())
```

7.6 References

(1) 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. *Scientific reports* **2016**, *6*, 18459.

(2) Arpino, J. A. J.; Rizkallah, P. J.; Jones, D. D. Crystal structure of enhanced green fluorescent protein to 1.35 Å resolution reveals alternative conformations for Glu222. *PloS one* **2012**, *7*, e47132.

(3) Zacharias, D. A.; Violin, J. D.; Newton, A. C.; Tsien, R. Y. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science (New York, N.Y.)* **2002**, *296*, 913–916.

(4) Ghosh, A.; Isbaner, S.; Veiga-Gutiérrez, M.; Gregor, I.; Enderlein, J.; Karedla, N. Quantifying Microsecond Transition Times Using Fluorescence Lifetime Correlation Spectroscopy. *The journal of physical chemistry letters* **2017**, *8*, 6022–6028.

(5) The PyMOL Molecular Graphics System. 1.5.0.4 edn (Schrödinger, LLC). https://pymol.org/2/support.html?#citing.

(6) Zitter, E. de; Hugelier, S.; Duwé, S.; Vandenberg, W.; Tebo, A. G.; van Meervelt, L.; Dedecker, P. *Integrated structure-function dataset reveals key mechanisms underlying photochromic fluorescent proteins*, 2020.

(7) https://www.picoquant.com/products/category/software/fluofit-global-fluorescence-decay-data-analysis-software.

(8) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.*; et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.

(9) GABRIEL, K. R. The biplot graphic display of matrices with application to principal component analysis. *Biometrika* **1971**, *58*, 453–467.

(10) Tsien, R. Y. The green fluorescent protein. *Annual review of biochemistry* **1998**, *67*, 509–544.

(11) Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormö, M.; Remington, S. J. Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 2306–2311.

(12) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 8362–8367.

(13) Fang, C.; Frontiera, R. R.; Tran, R.; Mathies, R. A. Mapping GFP structure evolution during proton transfer with femtosecond Raman spectroscopy. *Nature* **2009**, *462*, 200–204.

(14) Stoner-Ma, D.; Jaye, A. A.; Matousek, P.; Towrie, M.; Meech, S. R.; Tonge, P. J. Observation of excited-state proton transfer in green fluorescent protein using ultrafast vibrational spectroscopy. *Journal of American Chemical Society* **2005**, *127*, 2864–2865.

(15) Stoner-Ma, D.; Jaye, A. A.; Ronayne, K. L.; Nappa, J.; Meech, S. R.; Tonge, P. J. An alternate proton acceptor for excited-state proton transfer in green fluorescent protein: rewiring GFP. *Journal of the American Chemical Society* **2008**, *130*, 1227–1235.

(16) Grotjohann, T.; Testa, I.; Leutenegger, M.; Bock, H.; Urban, N. T.; Lavoie-Cardinal, F.; Willig, K. I.; Eggeling, C.; Jakobs, S.; Hell, S. W. Diffraction-unlimited all-optical imaging and writing with a photochromic GFP. *Nature* **2011**, *478*, 204–208.
(17) Duwé, S.; Zitter, E. de; Gielen, V.; Moeyaert, B.; Vandenberg, W.; Grotjohann, T.; Clays, K.; Jakobs, S.; van Meervelt, L.; Dedecker, P. Expression-Enhanced Fluorescent Proteins Based on Enhanced Green Fluorescent Protein for Super-resolution Microscopy. *ACS nano* **2015**, *9*, 9528–9541.
(18) Salna, B.; Benabbas, A.; Sage, J. T.; van Thor, J.; Champion, P. M. Wide-dynamic-range kinetic investigations of deep proton tunnelling in proteins. *Nature chemistry* **2016**, *8*, 874–880.
(19) Erez, Y.; Gepshtein, R.; Presiado, I.; Trujillo, K.; Kallio, K.; Remington, S. J.; Huppert, D. Structure and excited-state proton transfer in the GFP S205A mutant. *The journal of physical chemistry. B* **2011**, *115*, 11776–11785.
(20) Shu, X.; Leiderman, P.; Gepshtein, R.; Smith, N. R.; Kallio, K.; Huppert, D.; Remington, S. J. An alternative excited-state proton transfer pathway in green fluorescent protein variant S205V. *Protein science : a publication of the Protein Society* **2007**, *16*, 2703–2710.
(21) Sarkisyan, K. S.; Bolotin, D. A.; Meer, M. V.; Usmanova, D. R.; Mishin, A. S.; Sharonov, G. V.; Ivankov, D. N.; Bozhanova, N. G.; Baranov, M. S.; Soylemez, O.*; et al.* Local fitness landscape of the green fluorescent protein. *Nature* **2016**, *533*, 397–401.
(22) Jung, G.; Wiehler, J.; Zumbusch, A. The photophysics of green fluorescent protein: influence of the key amino acids at positions 65, 203, and 222. *Biophysical journal* **2005**, *88*, 1932–1947.
(23) Demadis, K. D.; Hartshorn, C. M.; Meyer, T. J. The localized-to-delocalized transition in mixed-valence chemistry. *Chemical reviews* **2001**, *101*, 2655–2686.
(24) Lin, C.-Y.; Romei, M. G.; Oltrogge, L. M.; Mathews, I. I.; Boxer, S. G. Unified Model for Photophysical and Electro-Optical Properties of Green Fluorescent Proteins. *Journal of the American Chemical Society* **2019**, *141*, 15250–15265.
(25) Larsen, D. S.; Vengris, M.; van Stokkum, I. H.; van der Horst, M. A.; Weerd, F. L. de; Hellingwerf, K. J.; van Grondelle, R. Photoisomerization and Photoionization of the Photoactive Yellow Protein Chromophore in Solution. *Biophysical journal* **2004**, *86*, 2538–2550.
(26) Litvinenko, K. L.; Webber, N. M.; Meech, S. R. Internal Conversion in the Chromophore of the Green Fluorescent Protein: Temperature Dependence and Isoviscosity Analysis. *J. Phys. Chem. A* **2003**, *107*, 2616–2623.
(27) Webber, N. M.; Litvinenko, K. L.; Meech, S. R. Radiationless Relaxation in a Synthetic Analogue of the Green Fluorescent Protein Chromophore. *J. Phys. Chem. B* **2001**, *105*, 8036–8039.
(28) Litvinenko, K. L.; Webber, N. M.; Meech, S. R. An ultrafast polarisation spectroscopy study of internal conversion and orientational relaxation of the chromophore of the green fluorescent protein. *Chemical Physics Letters* **2001**, *346*, 47–53.
(29) Mandal, D.; Tahara, T.; Meech, S. R. Excited-State Dynamics in the Green Fluorescent Protein Chromophore. *J. Phys. Chem. B* **2004**, *108*, 1102–1108.
(30) Niwa, H.; Inouye, S.; Hirano, T.; Matsuno, T.; Kojima, S.; Kubota, M.; Ohashi, M.; Tsuji, F. I. Chemical nature of the light emitter of the Aequorea green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 13617–13622.

(31) Baldridge, A.; Samanta, S. R.; Jayaraj, N.; Ramamurthy, V.; Tolbert, L. M. Steric and electronic effects in capsule-confined green fluorescent protein chromophores. *Journal of the American Chemical Society* **2011**, *133*, 712–715.
(32) Baldridge, A.; Feng, S.; Chang, Y.-T.; Tolbert, L. M. Recapture of GFP chromophore fluorescence in a protein host. *ACS combinatorial science* **2011**, *13*, 214–217.

8 Conclusions and perspectives

In this work, we employed electronic and vibrational TR absorption spectroscopy from the femtosecond to the minute time scales to study the photodynamics of wild-type rsEGPF2, and two variants V151A and V151L, which are negative RSFPs having as origin the avGFP. The time-resolved transient absorption spectroscopy experiments have been combined with time-resolved crystallography results obtained by our collaborators (Martin Weik group IBS). These two approaches have allowed inferring the photoswitching mechanism of rsEGFP2 and its variants. We will discuss our results regarding the different objectives raised at the beginning of this PhD (chapter 3) and give also some perspectives. **The first objective of the PhD was to determine the steady-state photophysical properties of rsEGPF2, V151A and V151L**. The most relevant result is the *Off*-to-*On* photo-switching quantum yields of 11, 12 and 14% that were determined for WT, V151L and V151A, respectively (Chapter 4). Such small differences are reflected in similar isomerization mechanism. **Rationalizing the *Off* to *On* isomerization mechanism was the main query of the PhD**. Based on the new results described here, we can now propose detailed isomerization pathways for both HT (*trans* 2) and OBF (*trans* 1) *Off trans* neutral forms (Figure 8.1). The excitation of the HT and OBF *Off* forms triggers an excited-state isomerization to a *cis* chromophore within a few picoseconds via a twisted excited state[1]. The nearly identical characteristic time constants for the isomerization steps of V151A and V151L demonstrates that both have similar isomerization mechanism and this being the HT as the V151L cannot isomerized via an OBF mechanism. After excited-state de-activation, the mechanisms differ. For V151A there is an evolution in the ground-state in around 83 ps, which from the X-ray structures, is associated to the movement of the Tyr146 which breaks its hydrogen bond with the His149 and relocates to the *On cis* final position, allowing thus the rotation and movement of chromophore to the final *On cis* form. This step is in agreement with an OBF isomerization product formed via an aborted HT mechanism. The resulting photoproduct (*cis* protonated chromophore) remains stable until 100 ns. The next evolution seen by TRMPS is attributed to the protein cage relaxation in few hundreds of ns for both V151A and V151L, due to the absence of spectral changes in the UV-vis data. This is followed by the movement of the His-149 to its final position which triggers a two-steps deprotonation in the µs-ms range.

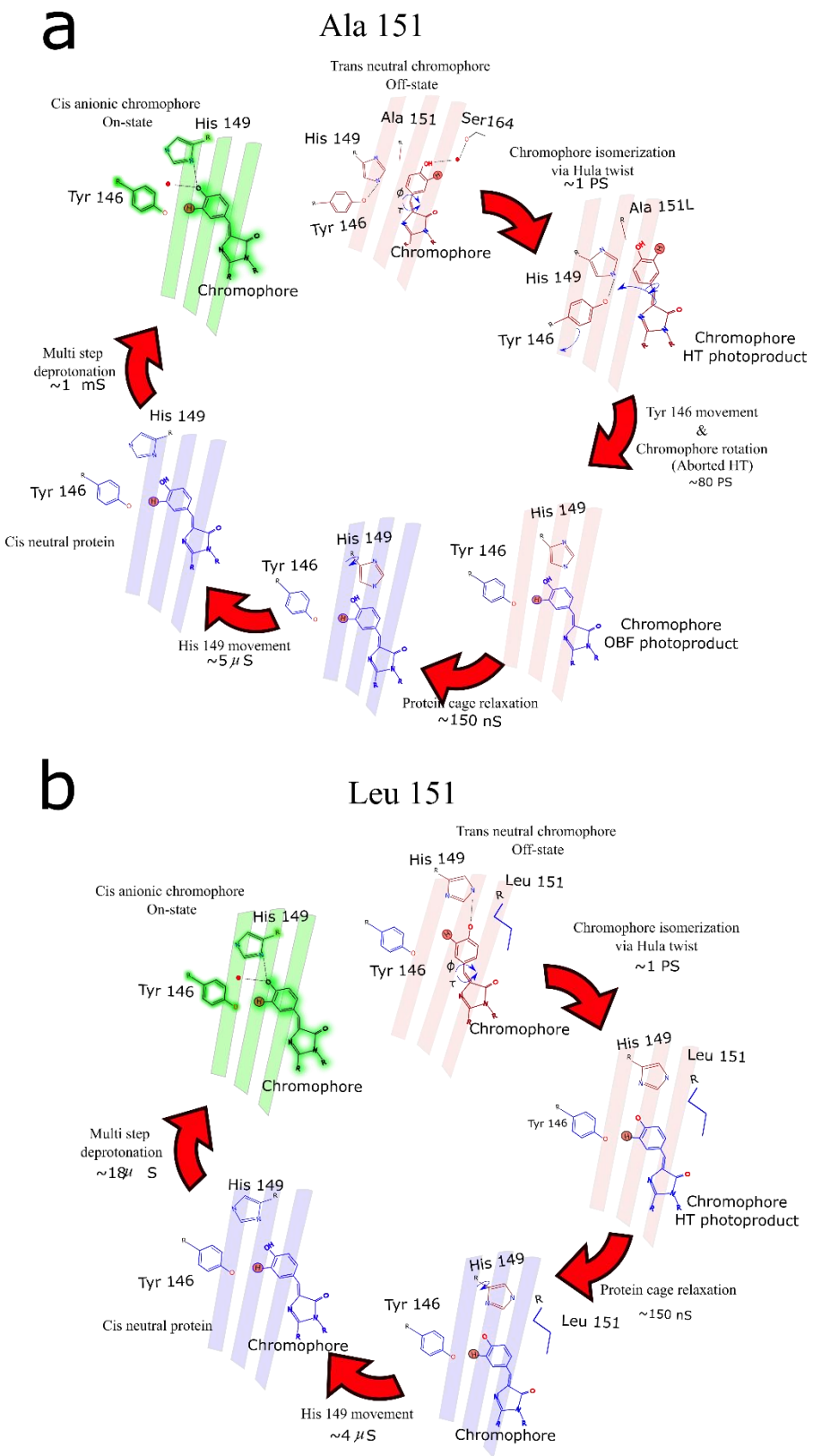


Figure 8.1. Schematic representation of the different Off to On photoswitching mechanisms for V151A (a) and V151L (b) rsEGFP2 variants. Figure made with Inkscape.

Moreover, our results could demonstrate that the rsEGFP2 WT transient absorption spectra can be seen as a mixture between the excited state dynamics of V151L and V151A. The same conclusion can be done from the steady state absoption spectra. **Thus we can conclude that the *Off* ground-state heterogeneity seen in rsEGP2 microcrystals seems to be maintained in solution. This answers the ground state heterogeneity question opened in Chapter 3.**

Finally, in chapter 5, the *Off* to *On* photodynamics of rsEGFP2 in solution and microcrystals have been compared. Interestingly a red-shift of the bands corresponding to the ESA and SE signal can be seen for microcrystals spectra compared to those in solution. These results indicate that the crystal phase modified the potential energy surface of the protein. On the contrary, no difference has been found in the *Off* neutral excited state dynamics of rsEGFP2 WT in crystal and solution; both data sets were fitted with two-time constants 0.6 ps (0.88 ps solution) and 5.59 ps (4.74 ps solutions). Surprisingly, clear differences have been found in the ground state dynamics, especially regarding the deprotonation steps, which have been found to be much faster in crystal than in solution. Nevertheless, these could be explained by the crystallization agent used $(NH_4)_2SO_4$. Further analysis, has identified the NH_4^+ as the responsible cation for the accelerated deprotonation step.

As discussed in chapter 2, the search of the existence of intermediate excited state during the cis-trans isomerization has been an important research topic throughout the last 50 years. Protein crystals are formed by about 50% of water molecules where side chains and secondary structures can move, which translates into protein crystals still been biologically active[2]. On the contrary crystallized organic molecules are rigid and not photoactive. Therefore, RSFPs such as V151A and V151L may be the only way to find and investigate, using TR-SFX, the different structures of excited state intermediates for HT and OBF primary isomerization steps. Thus TR-SFX experiment targeting the subpicosecond intermediates should be a future target.

To investigate species that control the fluorescence and *cis-trans* isomerization in rsEGPF2 was another goal of this project. In chapter 6, the photodynamics of rsEGFP2 *On* state has been presented. From the results presented, we confirm the presence of two main

ground state conformers for the protein *On cis* anionic state. These two ground state species have two different lifetimes, a short one of about 200 ps, which is also observed in rsFolder and rsFolder 2 (Figure 6.1). The second component is the long live fluorescent excited state, which takes of account of ~90% of the fluorescence. The TRMPS data and the comparison of the two main components with TR-SFX experiment and quantum mechanics calculations allow the attribution of these two ground-state conformers to different orientations of the water molecule in front of the chromophore (HBDI) phenol group. Since no difference has been found in the excited state dynamics of microcrystals, the TR-SFX structures at 10 ps and 1 ns after excitation can be used to get the structure of non fluorescent and fluorescent species. The structure at 1 ps (non fluorescent species) shows that the water molecule is hydrogen bonded to the Asn147 and the Ser206. While at 1000 ps (fluorescent species) the water molecule is hydrogen bonded to the HBDI phenolate and the Ser206. Therefore, the chromophore-Glu223 connexion via a proton relay, same as in avGFP, is broken in the first structure while established in the second. **Overall the chapters 5 and 6 highlight the combination of TR-SFX results and ultrafast optical spectroscopies (UV-Vis-IR) to investigate photoactive biological systems.**

Ultimately, the rsEGFP2 microcrystals and protein in solution have been compared. In both cases the excited state dynamics is composed of mainly two components, a short one <200 ps non-fluorescent and a long one with different decay times in crystal and solution. The differences of the long components in solution and crystals have been tentatively attributed to the chromophore environment in the microcrystals. On the contrary, in solution, the chromophore has enough space to adopt a planar excited state conformation, allowing it to decay with either 500 ps or via a long-lived excited state 2.8 ns. The origin of these two fluorescent components might be the same as those of the EGFP which is the rotation of T204[3]. Finally, after the decay to the ground state, the ns transient absorption spectroscopy experiments with relatively high S/N ratio compared to the TRMPS data, and performed in H_2O and D_2O, allows us to study the ground state switching dynamics. From the spectra in Figure 6.15, we show that after isomerization, at 1 µs, the transient spectra has a maximum of absorption at 500 nm, which is unequivocally attributed to the *trans* anionic chromophore. The significant isotopic effect (k_H/k_D= 1.5) reveals that this transient specie gets protonate in about 24.14 µs forming the final *Off trans* neutral form. This is the first-ever observed GSPT

from the *On* to *Off* state in an RSFP. Pump-dump experiment could confirm that the short lived excited state is the precursor of the final *trans* neutral photoproduct.

Moreover, the thesis also deal with the photo-physical properties of nineteen other variants of rsEGFP2 (Chapter 7, Figure 7.2). From this study, we can confirm the existence of different ground states with different fluorescence lifetimes with one order of magnitude of difference (100 ps vs 2 ns). From all the nineteen studied variants the most interesting one to test its expression in cells and then explore its possibility as a new fluorescent probe in RESOLFT is the rsEGFP2 S206N-V151A variant. The study is still ongoing. Overall, the outcomes of these studies will contribute to a better understanding of the photo-physics of RSFPs but more particularly into negative RSFP. Finally, they should open newer perspectives towards the design of optimized RSFPs for advanced bio-imaging application.

This thesis has study rsEGFP2, which is a negative RSFP. Nevertheless, although these proteins represent the major group among the RSFP proteins, and practically all RESOLFT implementations are based on them, they have several practical disadvantages compared to positive RSFP when applied in microscopy. The main one is that when excited they emit fluorescence and get deactivated at the same time. Therefore, during the readout in RESOLFT, they will turn *Off*, and consequently, if the number of collected photons is low, they need to be switched back to their *On* fluorescent state, and the RESOLFT excitation-deexcitation sequence repeated[7]. Positive RSFP have the advantage that fluorescence and *On* switching are trigger by the same excitation wavelength, during readout, they get switched to the *On* state. Therefore, they overcome the problem and simplify the RESOLFT sequence of pulses. Currently, to my knowledge, all positive RSFPs are derived from Dronpa, and only the photodynamics of Padron has been partially studied. Therefore the study of new positive RSFP as Padron2[7] should be a priority to understand the sequence of events that control photodynamics and switching events of the positive RSFPs which can open new paths for the development of new positive RSFPs with other origins than Dronpa.

8.1References

(1) Coquelle, N.; Sliwa, M.; Woodhouse, J.; Schirò, G.; Adam, V.; Aquila, A.; Barends, T. R. M.; Boutet, S.; Byrdin, M.; Carbajo, S.*; et al.* Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nature chemistry* **2018**, *10*, 31–37.

(2) Mozzarelli, A.; Rossi, G. L. Protein function in the crystal. *Annual review of biophysics and biomolecular structure* **1996**, *25*, 343–365.

(3) Ghosh, A.; Isbaner, S.; Veiga-Gutiérrez, M.; Gregor, I.; Enderlein, J.; Karedla, N. Quantifying Microsecond Transition Times Using Fluorescence Lifetime Correlation Spectroscopy. *The journal of physical chemistry letters* **2017**, *8*, 6022–6028.

(4) Arpino, J. A. J.; Rizkallah, P. J.; Jones, D. D. Crystal structure of enhanced green fluorescent protein to 1.35 Å resolution reveals alternative conformations for Glu222. *PloS one* **2012**, *7*, e47132.

(5) Shibazaki, C.; Shimizu, R.; Kagotani, Y.; Ostermann, A.; Schrader, T. E.; Adachi, M. Direct Observation of the Protonation States in the Mutant Green Fluorescent Protein. *The journal of physical chemistry letters* **2020**, *11*, 492–496.

(6) 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. *Scientific reports* **2016**, *6*, 18459.

(7) Konen, T.; Grotjohann, T.; Jansen, I.; Jensen, N.; Hell, S. W.; Jakobs, S. *The Positive Switching RSFP Padron2 Enables Live-Cell RESOLFT Nanoscopy Without Sequential Irradiation Steps*, Biorxiv, **2020**.

Appendix 1. Materials and methods

This appendix contains relevant information for the majority of experiments and setups used to acquire the data described in this thesis. Here, detailed descriptions of the different setups can be found as well as different procedures and preprocessing used in the analysis of the data acquired. The appendix is organized as follow:

- Short description of sample preparation.
- Description of time-resolved setups in LASIRE (France).
- Short description of time-resolved experiments done in Japan at Miyasaka's laboratory.
- Short description of time-resolved experiments done in England in the ultrafast laser laboratory.
- Description of time steady-state FTIR measurements.

A.1.0 Sample preparation

rsEGFP2 samples were expressed and purified by researchers in the IBS (Ninon Zala and Kyprianos Hadjidemetriou) as described previously[1–3]. Briefly, rsEGFP2 proteins fused to an N-terminal polyhistidine tag were expressed in *E. coli* BL21 (DE3). After cell lysis, the fluorescent proteins were purified by Ni-NTA affinity chromatography followed by size exclusion chromatography using a HiLoad 16/600 Superdex 75 column (GE Healthcare, Freiburg, Germany). The purified proteins were concentrated by ultrafiltration and equilibrated in buffer solutions (50 mM NaCl, 50 mM HEPES pH 8). Deuterated samples were prepared by exchanging the storage buffer with the same deuterated one (50 mMol Hepes, 50 mMol NaCl and pD 8), using 10k-Da concentrators. The process was done three times to ensure 99.9% of D2O[3]. I personally only participated one week in the purification of V151L sample and deuteration of small samples.

A1.1 Pump-probe fs-ns transient absorption spectroscopy

A1.1a System description

Time-resolved transient absorption spectroscopy in the fs-ns range system uses fs laser pulses as pump and probe beams. The system in LASIR[4,5] consists of fs Ti: sapphire laser pumped by an Nd: YVO4 continuous laser with 4.5 mW power (532 nm). This Ti: sapphire laser produces pulses of 180 nJ energy, 100 fs width centered at 800 nm, with a repetition rate of 76 Mhz. These Laser pulses are stretched, further amplified and finally recompressed in a regenerative Ti: sapphire amplifier pumped with an Nd: YFL laser. The delivered laser pulses at the output of the amplifier have around 100 fs pulse width, 0.9 mJ energy and are centered at 800 nm. These fundamental 800 nm pulses are divided into two beams. Around 10 % of the pulse energy is focused in a 2 mm CaF_2 rotating plate to generate the white light continuum probe beam. The remaining energy pulse is used to generate the pump pulse by focusing the 800 nm pulse into a BBO crystal to obtain the second harmonic 400 nm for the *Off* to *On* experiments and into an OPA (Palitra) to obtain 488 nm for the *On* to *Off* experiments. The polarization between the pump and probe pulse was always set at the magic angle (54.7°) to avoid anisotropy effects.

TA spectra were recorded in the 320-710 nm spectral range at different delays after the 400 nm pump excitation. The delay between the probe and pump pulses was obtained with an optical delay line. Each time delay spectrum is then averaged a certain number of frames (typically 1200). Each frame corresponds to 4 pump excitations pulses. Laser intensity fluctuations were corrected using a reference beam. Both beams, reference and signal, are recorded using multichannel spectrograph equipped with a CCD camera (Princeton Instruments). A $CuSO_4$ solution was placed in front of the spectrograph to eliminate the remaining 800 nm laser pulse. The Rayleigh scattered light from the pump was suppressed using a Notch filter at the respective pump excitation wavelength (Semrock). Every experimental setup was controlled using a home-made LabVIEW interface program. The setup is illustrated in Figure A1.1.

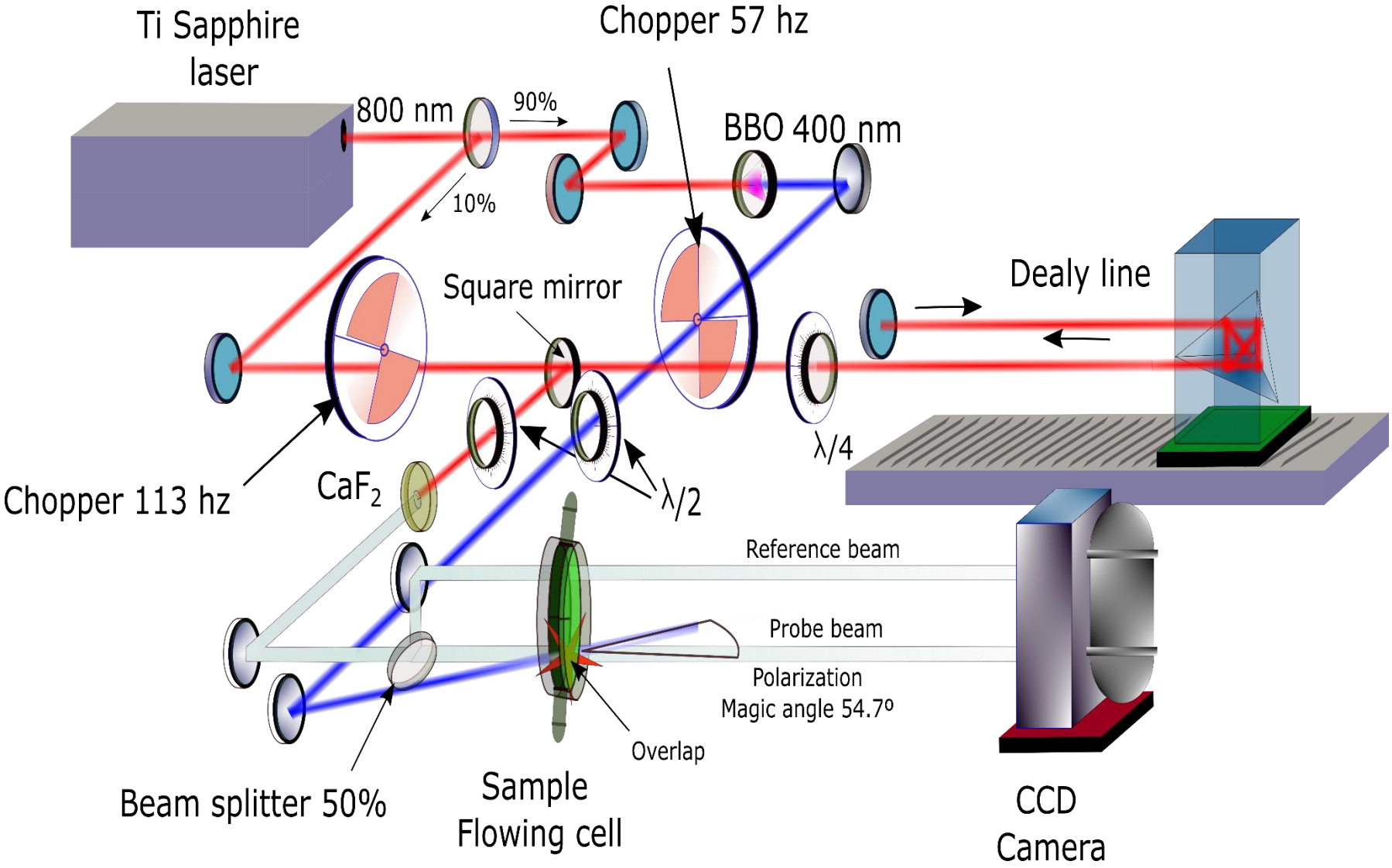


Figure A1.1. Pump-probe transient absorption spectroscopy setup scheme. Figure made with Inkscape.

The samples were placed inside a flow cell with 1 mm CaF_2 windows and an internal thickness of 250 µm to refresh the probe volume (500 x 500 x 250 µm) between consecutive pump pulse. Due to the conversion of the samples to long lived photoproducts by the pump pulse, samples were irradiated in an external reservoir with a CW LED (Thorlabs), at 490 nm and 255 mW to keep the *Off* state concentration constant or 400 nm 100 mW home-made assembled LED to keep the *On* state concentration constant during the experiments. The sample volume for the total flow cell system is 2.75 mL; thus, in general, 3.5 mL volumes has been used. A schematic representation of the flow cell has been done in Figure A1.2.

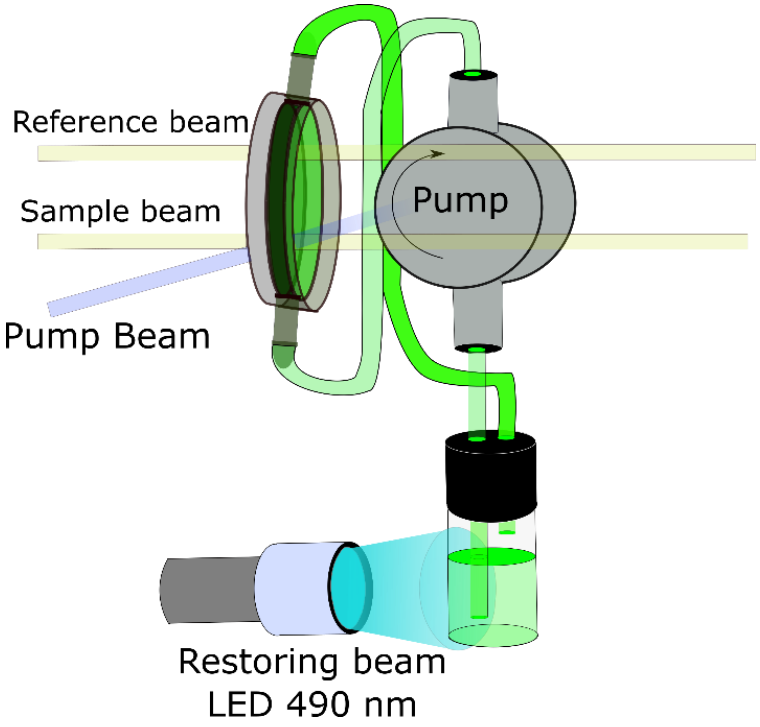


Figure A1.2. Flow cell used in TR-UVvis experiments to restore the photo-stationary Off state for Off to On experiment by illuminating an external reservoir. For On to Off experiments the restoring beam used was a home-made build 400 nm LED. Figure made with Inkscape.

A1.1b System characterization

The correct alignment of the system is verified with a compound that generates a long live species that do not decay all along the measured times. The chosen compound was Tris (bipyridine) ruthenium(II) chloride, as it is a compound that can be easily dissolved in water $[Ru\ (bpy)_3]^{2+}$. This compound can absorb UV and visible light and is generally used as a reference in transient absorption spectroscopy[6]. In aqueous solution, the compound has an absorption band at 452 ± 3 nm with an extinction coefficient of 14,600 $M^{-1}cm^{-1}$. Excitation of this band provokes an almost instantaneous (sub-ps) metal to ligand charge transfer (MLCT) to a triplet state. As explained in the chapter "light and matter interaction", triplet excited states are long live sates. The lifetime of this triplet MLCT state in water is 650 ns. Therefore, this compound is perfect for verifying the alignment of the setup from a few fs to around 2-3 ns delay times that can be achieved with an optical delay line, as there is no decay of the signal. Therefore, $[Ru\ (bpy)_3]^{2+}$ spectra from 0-2 ns were recorded before the experiments.

An example of an alignment experiment can be seen in Figure A1.3. For measurement that displayed fluctuations in the variation of absorption between 10 ps and 2ns, two main origins were identified: either a poor overlap between the pump and probe beams, or loss of continuum probe light at longer delay times. This is explained by the fact that at long delay, focused on the CaF2 plate can change due to misalignment of the optical delay line.

In figure A1.3a for the spectrum at 0 ps (excitation 400 nm), two sharp peaks can be seen: one negative at around 460 nm and one positive around 350 nm. These two peaks correspond to the stimulated Raman amplification signal of the solvent (Stokes and anti-Stokes signal). From these signals, it is possible to retrieve the time resolution of the setup as they give the cross-correlation time between the pump and the probe beams. This measurement should be performed directly on the sample data when possible. In case these peaks signals are not visible in the data, a solvent measurement can give a reasonable estimation of it.

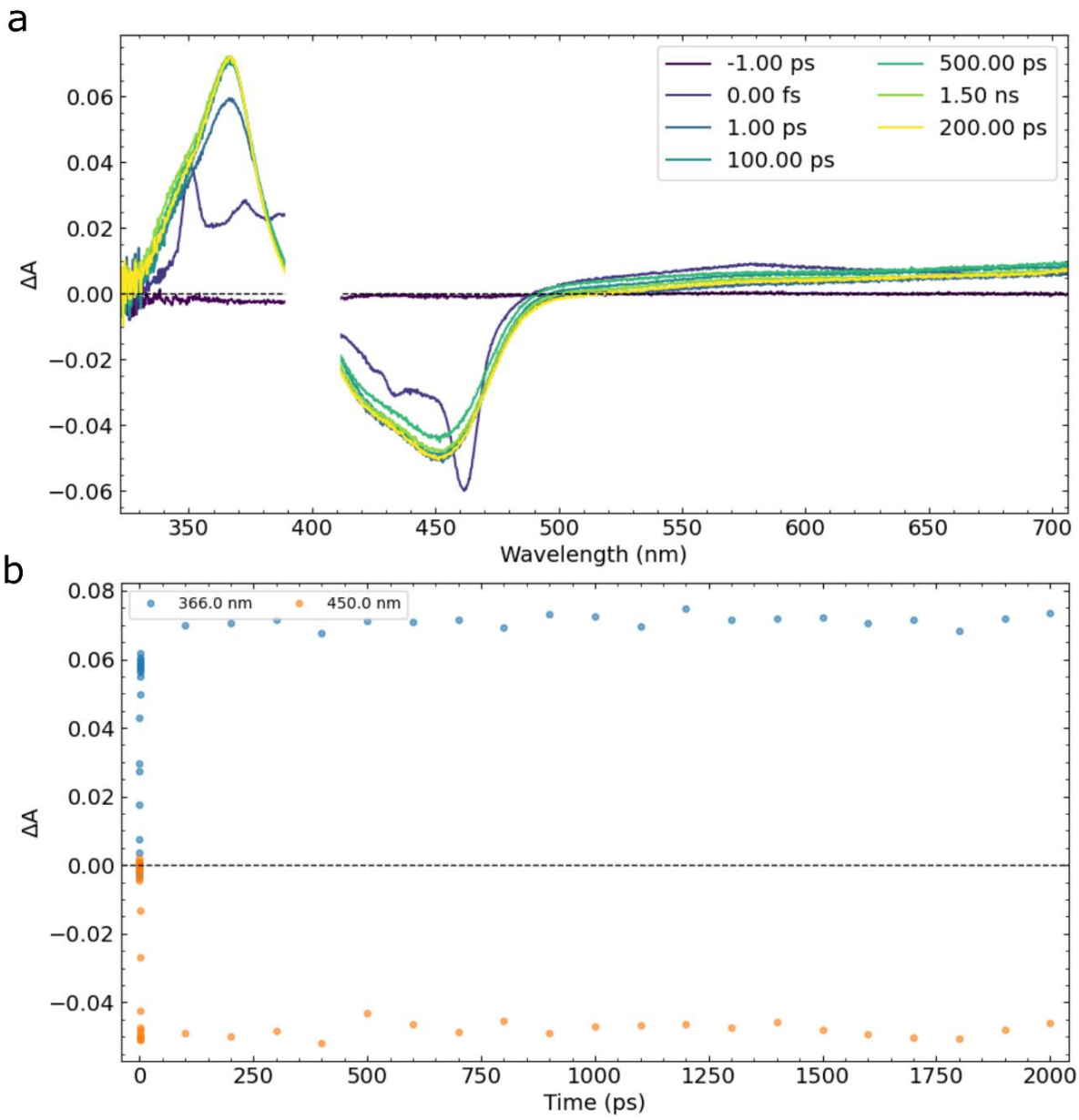


Figure A1.3. $[Ru(bpy)_3]^{2+}$ transient absorption spectra after GVD correction (excitation 400 nm fs). a) Selected spectra from -1 ps to2 ns. b) Time traces at 366 and 450 nm. Figures made with UltraPyFit

The cross-correlation between the pump and the probe beam can be considered as the FWHM of the stimulated Raman amplification signal of the solvent peaks. To retrieve this value, the stimulated Raman peaks can be fitted to Gaussian function (Equations A1.1) since the laser pulses have a Gaussian profile. The FWHM is thus estimated from the parameters of the fit.

$$f(x, A, \mu, \sigma) = \frac{A}{\sigma\sqrt{2\pi}} e^{\left[-(x-\mu)^2/2\sigma^2\right]} \qquad \text{Equation A1.1}$$

Where the width of the function (σ) is related to the FWHM as follow: FWHM $= 2\sqrt{\ln 2}\,\sigma$. In Figure A1.4 the Raman peak at 352 nm from Figure A1.3 has been plotted against time. After fitting this signal with a Gaussian function, the estimated FWHM was 190 fs.

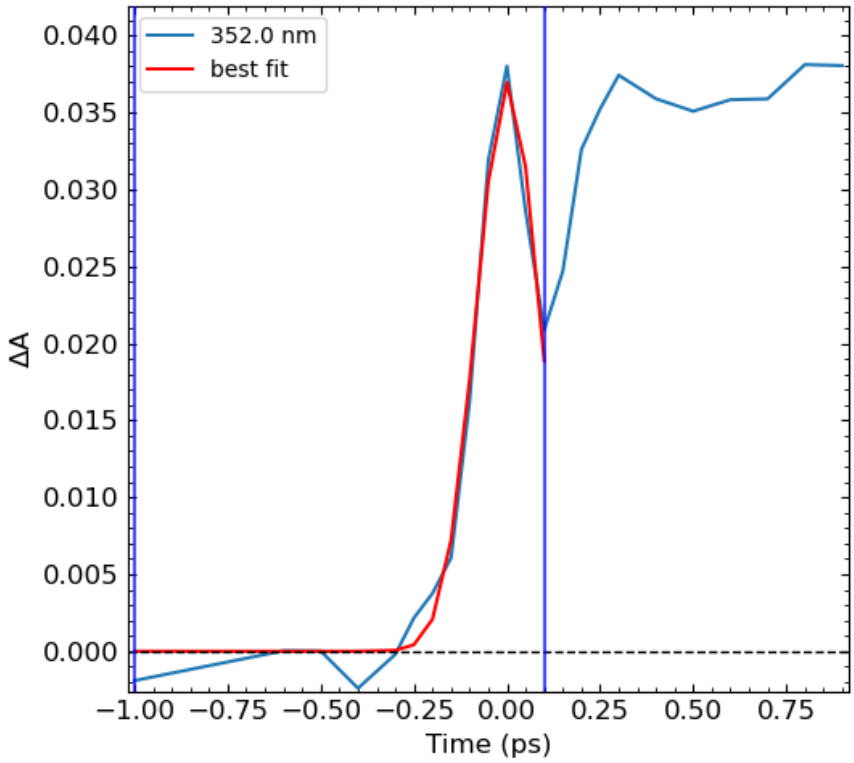


Figure A1.4. Early time evolution of kinetic trace at 352 nm for [Ru (bpy)$_3$]$^{2+}$ excited by 400 nm femtosecond pulse (in blue). In red, a fit to the data point between the two vertical blue lines with a Gaussian function. The estimated FWHM is 0.19 ps. Figure made with UltraPyFit.

A1.1c Data acquisition

The signal at a chosen delay consists of a variation of absorbance between data collected when the pump is blocked and data when it is unblocked. Moreover, to correct the laser intensity fluctuation, not only the light from the sample (the point where pump and probe overlap, see Figure A1.1) is recorded, but also a laser reference beam. Finally, the background signals for each beam are also taken into account, with the particularity that also the pump beam can produce emission. All these datasets are stored in separate buffers and compared to each other when enough frames have been acquired. It is important to note that, due to the readout speed of the camera, each frame corresponds to four pump laser shot and not to a single one. The calculation of delta absorbance are shown in equations A1.2 and A1.3.

$$\Delta Abs(\Delta t) = Abs_{after\ pump}(\Delta t) - Abs_{before\ pump}(\Delta t) \qquad \text{Equation A.12}$$

$$\Delta Abs(\Delta t, \lambda) = -log\left(\frac{I_{sample}^{pump}(\lambda,\Delta t) - I_{sample-bkg}^{pump}(\lambda)}{I_{ref}^{pump}(\lambda,\Delta t) - I_{ref-bkg}^{pump}(\lambda)}\right) + log\left(\frac{I_{sample}^{No\ pump}(\lambda,\Delta t) - I_{sample-bkg}^{No\ pump}(\lambda)}{I_{ref}^{No\ pump}(\lambda,\Delta t) - I_{ref-bkg}^{No\ pump}(\lambda)}\right)$$

Equation A1.3

Each term corresponds to reference and sample intensities signals and background signal:

$I_{sample}^{pump}(\lambda, \Delta t)$: Intensity of the sample beam with the pump excitation.

$I_{ref}^{pump}(\lambda, \Delta t)$: Intensity of the reference beam with pump excitation.

$I_{sample-bkg}^{pump}(\lambda)$: Background signal with the pump excitation.

$I_{ref-bkg}^{pump}(\lambda)$: Background signal with pump excitation.

$I_{sample}^{No\ pump}(\lambda, \Delta t)$: Intensity of the sample beam when the pump is blocked.

$I_{ref}^{No\ pump}(\lambda, \Delta t)$: Intensity of the reference beam when the pump is blocked.

$I_{sample-bkg}^{No\ pump}(\lambda)$: Background signal when the pump is blocked.

$I_{ref-bkg}^{No\ pump}(\lambda)$: Background signal when the pump is blocked.

The signals are recorded in a dual-channel CCD camera (1340 x 200 x 2). One channel detects the probe beam and the other the reference beam, and each channel records the

corresponding backgrounds. The dark background counts are recorded with pump and probe shutters closed. To take into account possible fluorescent emission and dispersion photons that this pulse might generate, a second background signal is recorded only with the pump pulse shutter open. These background signals are measured before and after the experiments. The rest of the signals are recorded at every delay time, where the pump pulse is blocked two times more than the probe pulse by the chopper (pump 57 Hz; probe 113 Hz), to take measurements without and with laser excitation. A home-made Labview software automatically controls the measurement acquisition. Once all the frames for the selected delay have been recorded, the optical delay is moved to collect the next time delay frames.

A1.1d GVD correction.

The phase velocity of a wave is related to the speed of propagation in a medium[7]. The speed is directly related to the diffractive index of the medium which can be expressed as a function of the wavelength or frequency of the wave. This phenomenon can cause different dispersion of light, which can be desired or non-desired. An example of a significant consequence of this effect is the change in the angle of refraction of different colors. This effect can be used to construct spectrometers. On the contrary, it also provokes the chromatic aberration in lenses.

For the case of femtosecond pulses, the propagation of wave packets of the pulse is essential. Due to the Heisenberg incertitude principle[8], these very short pulses have very large spectral widths (different wavelengths) as explained in chapter 2 “interaction of light and matter”. Therefore, these pulses have different frequency components which, when travelling through a dispersive media, have different propagation speeds. The ensemble effect produced is known as group velocity dispersion (GVD). This effect is even more pronounced for white light continuum pulses used as probe beams as they have a wider range of frequencies. As a consequence, the longer wavelength components of the white light continuum travel faster than the shorter wavelength when the pulse travels through a medium with positive group-velocity dispersion. In general, this translates into a pulse becoming *positively chirped*. When travelling through a negative group-velocity dispersion media, the opposite occurs. The

dispersion media in the systems: i. lenses, ii. the filters, ii. the CaF2 plates to generate a white light continuum, iv. the window cells and v. the sample itself. All are positive group-velocity dispersive media and have a direct ensemble effect on the recorded data. In figure A1.5a, representing the collected data point from -1 to 2 ps from an *Off* to *On* experiment for V151L variant, it can be seen how for red part of the data set, the wavelength has travelled faster than the blue part of the spectrum, and thus at very short time delays, the recorded spectrum might have a raising signal in the blue part of the spectrum whilst for the other parts there is still no signal. The GVD introduces a nonlinear effect in the data which, in order to fit a global model, can introduce errors in the estimated parameters. Therefore, this effect needs to be corrected.

There are several ways to correct GVD effects on the data. One popular way is to fit a polynomial function of at least order 3 to estimate the chirp. Nevertheless, since the GVD depends on the diffraction index of the media, this dispersion can be simulated. The equation 4 shows that the GVD can be expressed as a function of the diffraction index of a dispersive media and the wavelength. Therefore, by calculating the GVD for each wavelength and taking into account the thickness of the dispersive media, it is possible to model the experimental data displacement introduced indirectly in the dataset for further corrections.

$$GVD(\lambda) = \frac{\lambda}{C} \cdot \frac{d^2 n}{d\lambda^2} \qquad \text{Equation A1.4}$$

In the above equation, "C" is the speed of light and "n" is the refraction index of the dispersive media. The variation of "n" according to the wavelength can be calculated from the Sellmeier equation[9] (Equation A1.5).

$$n^2(\lambda) = 1 + \frac{B_1\lambda^2}{\lambda^2 - C_1} + \frac{B_2\lambda^2}{\lambda^2 - C_2} + \frac{B_3\lambda^2}{\lambda^2 - C_3} \qquad \text{Equation A1.5}$$

Where B1-B3 and C1-C3 are constant values experimentally determined for each dispersive media. In the thesis, the data sets have been corrected considering three different dispersive media: CaF_2, SiO_2 and BK_7. The corresponding B and C constants are stated in Table A1.1 and values were obtained from references [10,11] and [12].

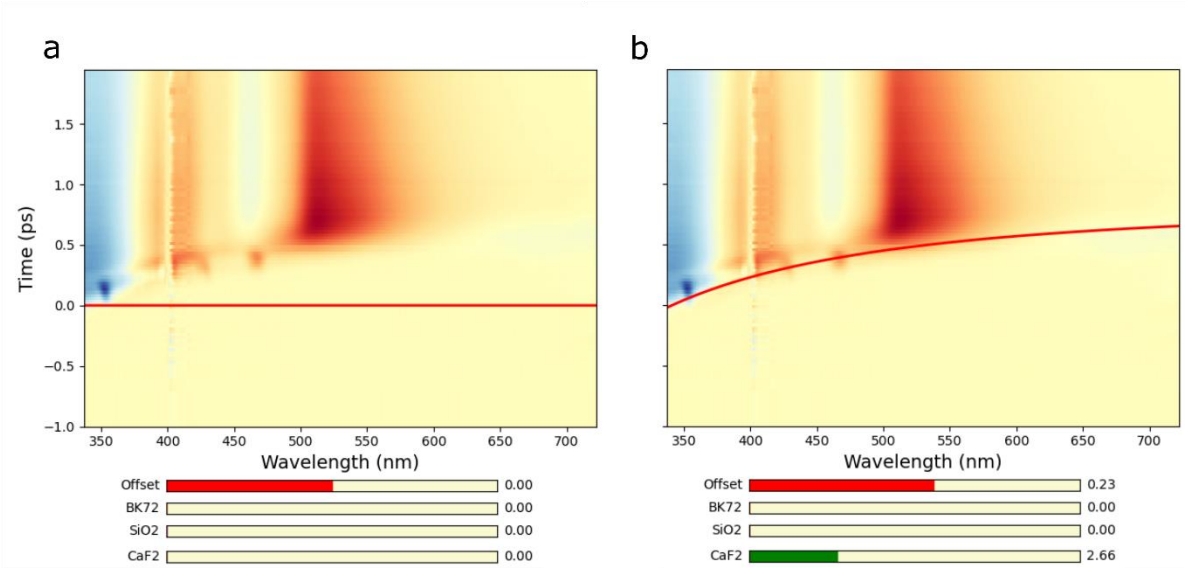


Figure A1.5. Representation of the collected data point from -1 to 2 ps versus wavelength for V151L Off to On experiment. In a) the red line represents the theoretical time zero that all wavelengths should have. In b) the red line represents the calculated t0 for each wavelength obtained after modelized the GVD using 2.66 mm of CaF2 and adjusting the offset to 23 fs. Figures made with UltraPyFit.

In figure A1.5a, the data set before correction is shown, whilst figure A1.5b, shows that the estimation of the GVD can be correctly done using only 2.66 mm CaF_2 thickness and the equations above. The used value is in agreement with the experimental setup as a 2 mm CaF2 plate (see figure A1.1) has been used to generate the white light continuum and the flowing cell has 1mm CaF2 windows.

Once the GVD calculated, the data needs to be corrected. Here, we have always corrected it using the points from the dataset itself as proposed by Nakayama et al[13]. The method used consists of subtracting the estimated dispersion curve (red line in Figure A1.5) for every independent wavelength. Therefore, every wavelength will be shifted differently, and thus data will be corrected. For every time point in a single wavelength, the estimated delay produced by the GVD was subtracted, yielding a new shifted time point. The estimations of the intensity of this new shifted time point were done as the pondered average of the two closest real-time points from the old-time grid to this new time point (See Equation A1.6 for pondered average and Figure A1.6 for the implemented algorithm). This results in a single

corrected time point. The same procedure was applied for the correction of every point in each recorded wavelength. The result of the final correction are shown in Figure A1.6.

$$\Delta A(t_{new}) = W \cdot \Delta A\left(t_{old}^{high}\right) + (1 - W) \cdot \Delta A\left(t_{old}^{low}\right) \qquad \text{Equation A1.6}$$

Where t_{old} represents time points in the old-time grid, and W the weighting factor, which is calculated from equation A1.7.

$$W = abs\left(\frac{t_{old}^{high} - t_{new}}{t_{old}^{high} - t_{old}^{low}}\right) \qquad \text{Equation A1.7}$$

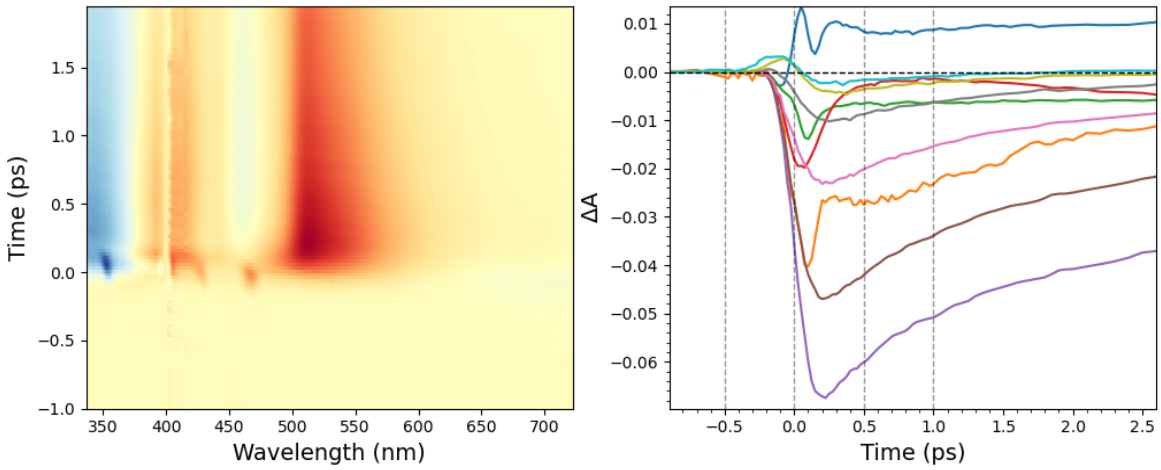


Figure A1.6. Left panel, representation of the collected data point from -1 to 2 ps versus wavelength for V151L Off to On experiment after the GVD correction. The right panel corresponds to the selected time traces from -1 to 2.5 ps from the data in the right panel, where it can be seen the rise of the simultaneous rise of the signal for all traces. Figures made with UltraPyFit.

Table A1.1 B and C values for CaF_2 SiO_2 and BK7 used to calculate the diffraction index

	B1	**C1**	**B2**	**C2**	**B3**	**C3**
BK7	1.0396121	$6.00069 \cdot 10^{-3}$	0.2317923	$2.001791 \cdot 10^{-2}$	1.0104694	103.5606
SiO_2	0.696160	$4.67915 \cdot 10^{-3}$	0.4079426	$1.351206 \cdot 10^{-2}$	0.8974794	97.93400
CaF_2	0.5675888	0.0502636	0.4710914	0.1003909	38.484723	34.64904

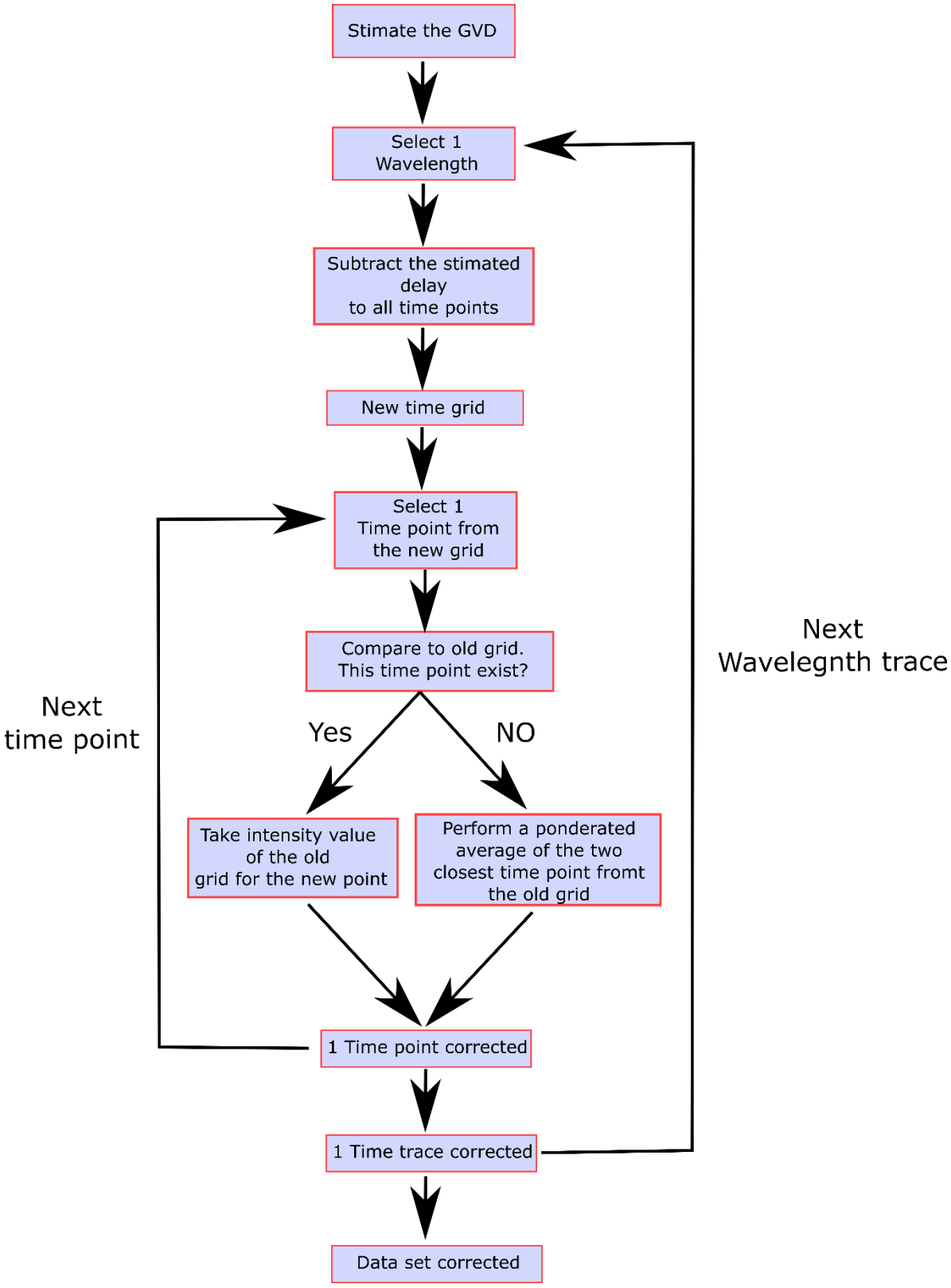


Figure A1.7 Schematic representation of the implemented algorithm to correct the GVD. Figure done with Inkscape.

In this thesis, a routine to correct the GVD as the one presented above has been developed in python 3.7. Finally, a similar routine to calculate the GVD as a polynomial function was also developed. In both cases, once the GVD dispersion is calculated, the correction algorithm presented just above is applied to correct the data (Figure A1.7.). The algorithm and the two different methods to calculate the GVD have been integrated into the Graphical user interface (GUI) named UltraPyFit, or as methods from the class GlobalFit python class. The codes and GUI are available on-demand, and more details about them are found in the next appendix, specially dedicated to the details of the developed GUI.

A1.2 Pump-probe ns-ms transient absorption spectroscopy

The nano-millisecond transient absorption experiments were performed by using a laser flash photolysis apparatus[14]. Excitation pulses were obtained from a 10 Hz Nd: YAG laser (Continuum Surelite II) coupled to an OPO (Continuum Panther EX OPO) and SH05 shutters (Thorlabs). The probe light was obtained from a pulsed Xe lamp (XBO 150W/CR OFR, OSRAM). The transmitted light was dispersed by a monochromator (Horiba Jobin–Yvon, iHR320) and analyzed with a photomultiplier (R1477-06, Hamamatsu) coupled to a digital oscilloscope (LeCroy 454, 500 MHz). Synchronization of the excitation pulses and the acquisition time was secured with PCI-6602 8 channel counter/timer (National Instruments card). The experiment was controlled by a home-made software written in the LabView. The nanosecond pump laser pulses at 410 nm (8 ns, 1.6 mJ) were used and focused onto a 1 cm × 1 mm cell containing 200 µl of the protein solution (absorbance of about 1 at 410 nm in a 1 cm path-length). Before pump-laser excitation, rsEGFP2 was switched from the *On* to the *Off* state by 2 min irradiation with light from a continuous wave modulated laser diode (Cobolt 06-MLD 488) at a power of 200 mW and with a nominal wavelength of 488 nm. After pump-laser excitation, the time traces were recorded from 340 to 520 nm every 10 nm to reconstruct the transient absorption spectra. To increase the signal to noise ratio and avoid laser intensity fluctuations, traces were recorded 8 times which were then averaged. Between each single excitation shot, rsEGFP2 in solution was converted to its *Off* state by visible-light irradiation using LED with nominal wavelength 490 nm, 255 mW over 30 seconds. To

prevent the sample degradation, this was renewed every 56 laser shots, and absorption spectra were recorded before and after the laser exposure as a control measurement. The described setup is illustrated in Figure A1.8. Similar to the above described fs-ns pump-probe setup, the ns-ms setup was also optimized using $[Ru(bpy)_3]^{2+}$. To do so, the signal at 360 nm, which corresponds to the absorption of the excited triplet state, was maximized. After maximizing the signal, and without moving the cuvette, the $[Ru(bpy)_3]^{2+}$ sample was replaced by the protein solutions after correctly cleaning with water and buffer solutions. It is worth noticing that the probe beam travels through 1cm of the sample while the pump beam excites the sample in a 90° configuration over 1 mm. This 90° configuration ensures the complete excitation of the probed sample, and over 10 times higher signal to noise ratio compared to a collinear excitation since 1 cm of sample is probed (See figures A1.8 and A1.11.

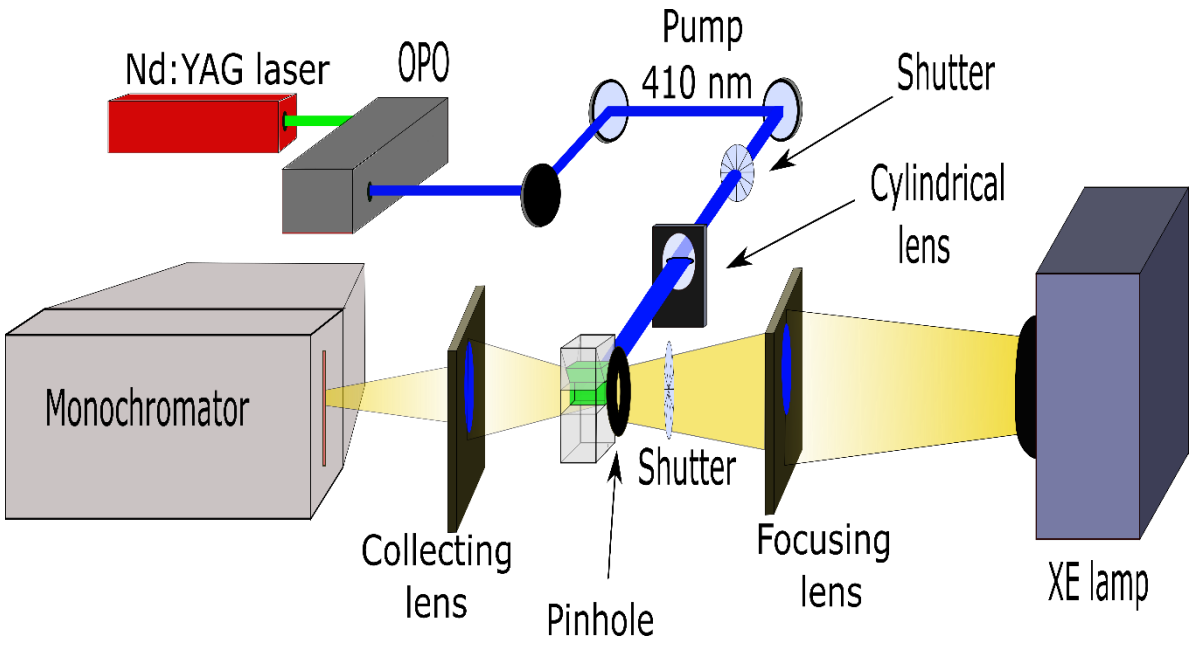


Figure A1.8. Schematic design of the flashphotolysis used to measure the ground state evolution from a few ns up to 10 or 20 ms. Figure made with Inkscape.

The experiment was repeated three times to obtain different time spectral windows of 10000 points each, which increases the number of points in the time domain of the recorded traces and ensures the collection of enough points in each interval. The selected windows were, from 0-5 µs and 0-100 µs for all studied samples plus a long time window. This last window was from 0-10000 µs for H2O solutions, 0-20000 µs for the D2O solution and 0-5000 µs for

solutions with ammonium sulfate. A 20MHz filter was used for all spectral windows to filtrate electronically the signal. For the time window 0-5 µs, the slit of the spectrometer was minimized to 1 nm to eliminate the laser dispersion and decrease the fluorescence contribution for traces recorded at wavelengths longer than 470 nm. For the remaining time windows, the slit was set to 1.5 nm, and the signal was amplified with an external resistance of 100 Ohm for 0-100 µs, and 560 Ohm for longer time domains. $[Ru(bpy)_3]^{2+}$ traces were measured to characterize the electronic modification of the signal that the different resistance introduces. The results are shown in Figure A1.9. Finally, for the two shortest time windows, the Xe lamp used flash pulses for probing, whilst no flash was used for the longer time windows.

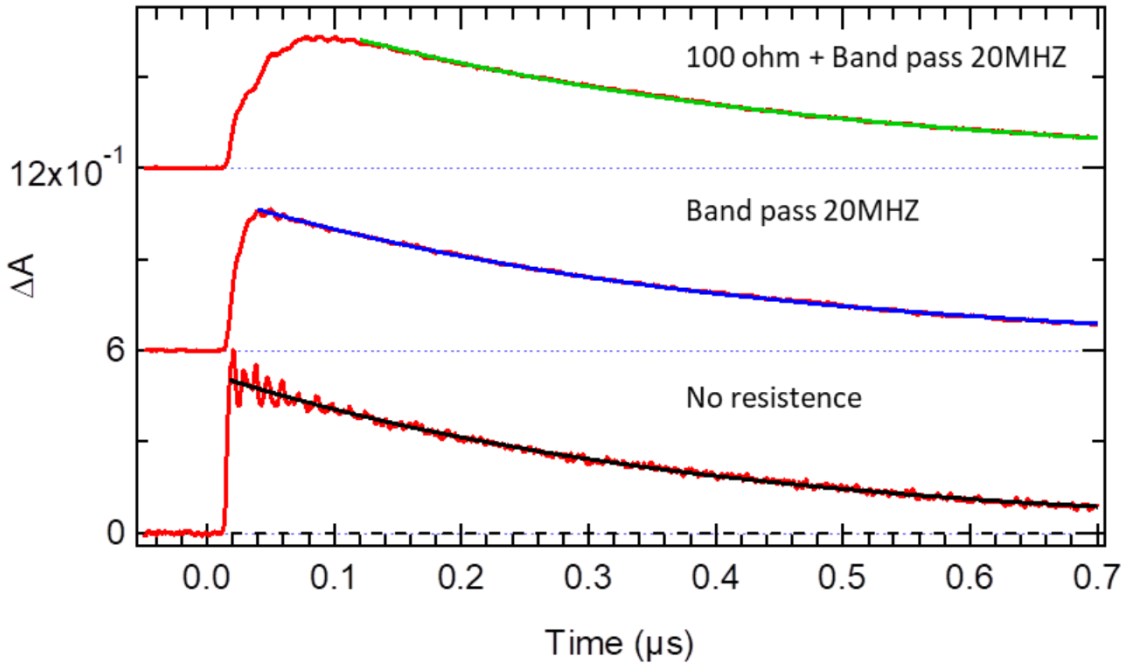


Figure A1.9. Variation of the instrument responses according to the resistance and bandpass used to amplify the signal and reduced the noise measured in $[Ru(bpy)_3]^{2+}$ water solution at 360 nm. Figure made with Origin pro 8.

The time evolutions measured are relatively long (ms range) and much longer than those found in dronpa (µs range). Thus, the question regarding the last evolutions observed in such long times may be raised. For example, diffusion of molecules may occur due to Brownian movement in such time scales which can cause diffusing inside the probed beam of proteins

that were not pumped, and vice-versa. Therefore, to verify this effect, a time trace evolution at 600 nm of an organic photochromic dye (named here PBOX) was recorded with the flash photolysis apparatus configuration. PBOX is an organic molecule with a much lower molecular mass compared to the rsEGFP2 protein and a long-live state in the second range, which is formed instantaneously after 410 nm laser excitation. The results are shown in figure A1.9. As can be appreciated, the signal recorded is a flat line. Therefore, one can conclude that the time evolutions seen at such long delays are originated from the protein.

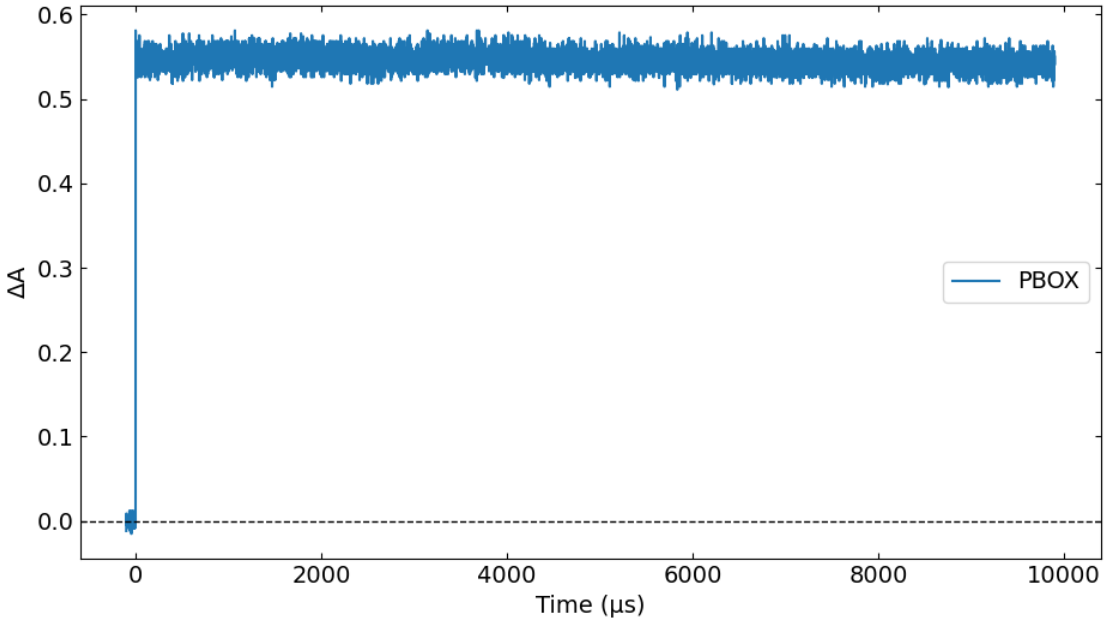


Figure A1.10. Variation of absorbance over time recorded in the flasphotlisys system for PBOX at 600 nm using 560 Ohm resistance and 20 MHz bandpass. The result ensures that the variations of the signal at long time delays measured for the proteins are intrinsic to them. Figure made with Matplotlib python package.

For spectroscopy experiments on microcrystal suspensions (rsEGFP2 microcrystals), the setup was modified (collinear geometry between the pump and probe beams) to minimize light diffusion according to the literature. The pump laser (410 nm, 5 mJ) was focused using a spherical lens onto the cell containing the colloidal microcrystal solution. The probe light source was not modified from conventional experiments. The focal spot size of both pump and probe beams was 2 mm. To maximize the signal to noise ratio a 560 Ohm resistance was used, and thus, data can only be analysed up from 600 ns which was characterized as that for

100 Ohm resistance (Figure A1.9). 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. A close comparison between the perpendicular and collinear configurations between the pump and probe beams can be seen in Figure A1.11. Note that the microcrystal measurements have a lower S/N ratio than the measurements in solution. This is explained by three main characteristics of the microcrystals suspension measurements: firstly, there are single-shot measurements which reduce the S/N ratio considerably. Secondly, the cuvette optical path used to measure the microcrystals is 10 times shorter than the one used to measure the protein in solution. And finally, the diffusion of the probing light by the microcrystals.

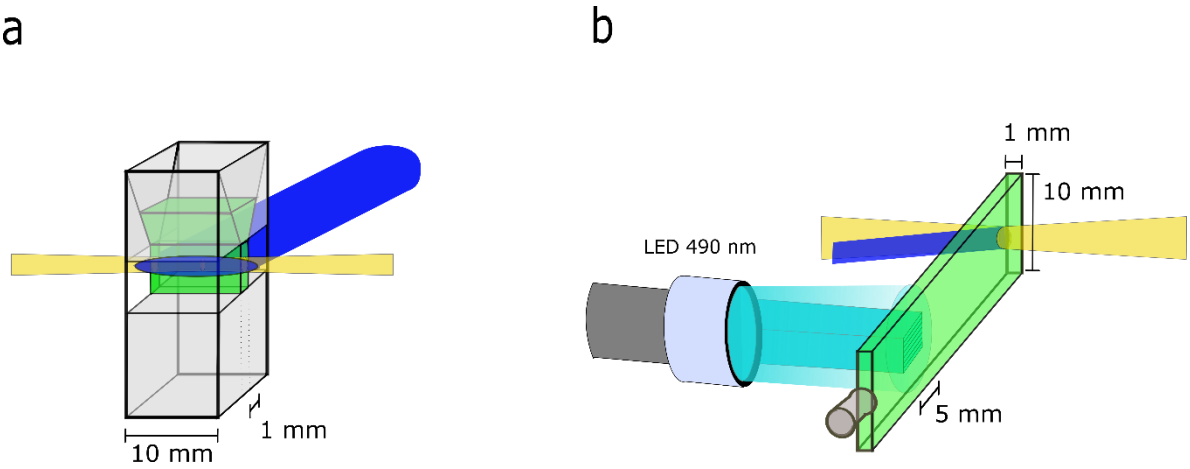


Figure A1.11 a) perpendicular pump-probe configuration used for solution samples. b) collinear pump-probe configuration used for microcrystals suspension samples. Figure made with Inkscape.

A1.3 TCSPC

The setup used has already been described earlier[15]. Briefly, the excitation source for the time-correlated single-photon counting is originated in a Ti: sapphire laser (Coherent Chameleon Ultra II) with a repetition rate of 80 MHz and tunable output wavelength range of 700 1200 nm. The laser can deliver 200 fs width pulses. The output of the laser is coupled with a pulse picker (4 MHz pulse repetition rate). Finally, there is also a second harmonic

generator (SHG/THG, APE) that allow the production of excitation wavelengths between 350-600nm. The fluorescence emission is recorded using an FT200 PicoQuant spectrometer. The emission of fluorescence is further filtrated with a Czerny-Turner type monochromator equipped with 0.5 mm slits that render a 4 nm spectral resolution. Finally, the photons are detected with a cooled microchannel plate (Hamamatsu R3809U) photomultiplier tube.

The time correlation of photons is done with a PicoHarp 300 TCSPC system coupled with the detector and a photodiode that triggers the signal. The time resolution of the setup is determined by the instrument response function (IRF). This signal was measured using a Ludox solution to disperse the excitation beam for each excitation wavelength used. The IRF varies between the different excitations used. Nevertheless, in general, the full width at half maximum of the IRF was around to 32 ps. A scheme of the set up is shown in Figure A1.11. The polarization between the pump pulse and collected photons was set at the magic angle (54.7°) to avoid anisotropy effects.

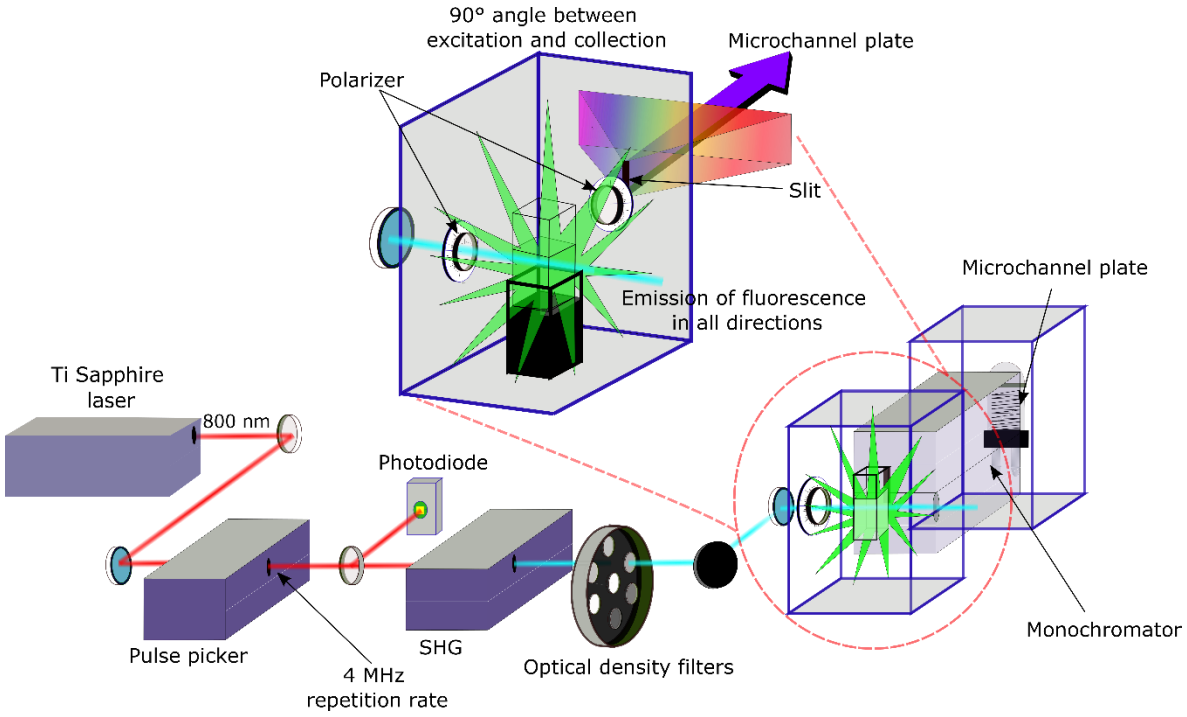


Figure A1.12. Scheme for the time-correlated single-photon counting setup. Figure made with Inkscape.

A1.4 Description of setups not present in LASIR

A1.4a Set-up in Miyasaka laboratory - Japan

The laser system used for transient absorption spectroscopy has already been described elsewhere[16]. A femtosecond laser pulse from a Ti: Sapphire oscillator (Tsunami, Spectra-Physics, 802 nm, 100 fs, 820 mW, 80 MHz) was injected into a regenerative amplifier at 1-kHz repetition rate (Spitfire, Spectra-Physics). Similar to the setup described above, the output pulses were divided into two portions. The majority of the 802-nm pulse was led to the other OPA (TOPAS-Prime, Light-Conversion) and converted into 1180 nm. This near-infrared pulse was focused into a 2-mm CaF_2 to generate a femtosecond white-light continuum which was used as a probing pulse. The rest of the 802-nm fundamental pulse was focused into a BBO crystal pulse to obtain the second harmonic 400 nm for the *Off* to *On* experiments and into OPA system (TOPAS-Prime, Light-Conversion) to obtain the 488 nm for laser pulses in the *On* to *Off* experiments.

The white light was further divided into two different beams - sample and reference. This last one is used to correct fluctuations of the probe pulse. Both white light pulses were monitored by two multichannel photodiode arrays (PMA-10, Hamamatsu) equipped with polychromators. The polarization of the excitation pulse was set at the magic angle with respect to that of the probe pulse.

The sample solution was set in a rotation cell with an optical length of 1 mm, and absorbance of the solution at excitation wavelength is around 0.5. Samples were irradiated in with a CW LED, at 490 nm and 255 mW to regenerate the *Off* state (Thorlabs) to keep the photostationary state.

A1.4b Setup used in the ultrafast laser facility at RAL - England

The setup used, named Ultra, was published in 2012 by Greetham et al. [17] and has 10 kHz multiplexed infrared broadband probe pulses. The excitation is at 1 kHz with sub-100 fs. The setup allows to recorde data points using these two pump-probe beams in time delays from 100 fs to 1 ms range. The setup used to generate the delays between the two pulse was based on different methods: the first delay times were obtained using an optical delay to collect fs–ns data, and times longer than those that cannot be obtained from a delay line were electronically obtained with a delay generator up to 100 μs (ns- μs range). Finally, a multiple probe approach (flash photolysis-like) was used to collect μs–ms data and beyond. Therefore, with this approach that combines several methods to obtain all the time ranges from fs-ms, it is possible to measure time scales across 9 orders of magnitude on a single instrument. This makes it an extremely powerful setup as it allows comparing spectral evolutions accurately and in a single experiment.

The excitation energy pulses were < 1 μJ energy in a 100-μ diameter spot. Lower intensities were checked to yield identical spectra and kinetics on the nanosecond timescale. The IR probe beam is split before the sample to provide probe and reference spectra, which can be normalized to reduce shot-to-shot probe intensity fluctuations. The spectrometer and detector systems are extensively described in detail in reference[18].

The samples were placed inside a flow cell with 1 mm CaF_2 windows and an internal thickness of 50 μm. To avoid spatial overlap in between consecutive laser shot in the same sample spot the cell was also moved continuously in the plane (raster). Samples were irradiated in an external reservoir with a CW LED at 505 nm (500 mW) to regenerate the *Off* state (Thorlabs) or 405 nm to regenerate the *On* state (Thorlabs). The minimum sample volume for the total system is 650 μL.

A1.5 FTIR

The *On* to *Off* photo-switching reaction was also studied using FTIR spectra in the region of 1300-1800 cm^{-1}. The experiments were done on deuterated samples to avoid the overlap of the water bending modes with the protein signals. The protein samples were deuterated using 10k-Da concentrators. The process was done three times to ensure 99.9% of D2O, and concentrated to approximately 15mg/mL.

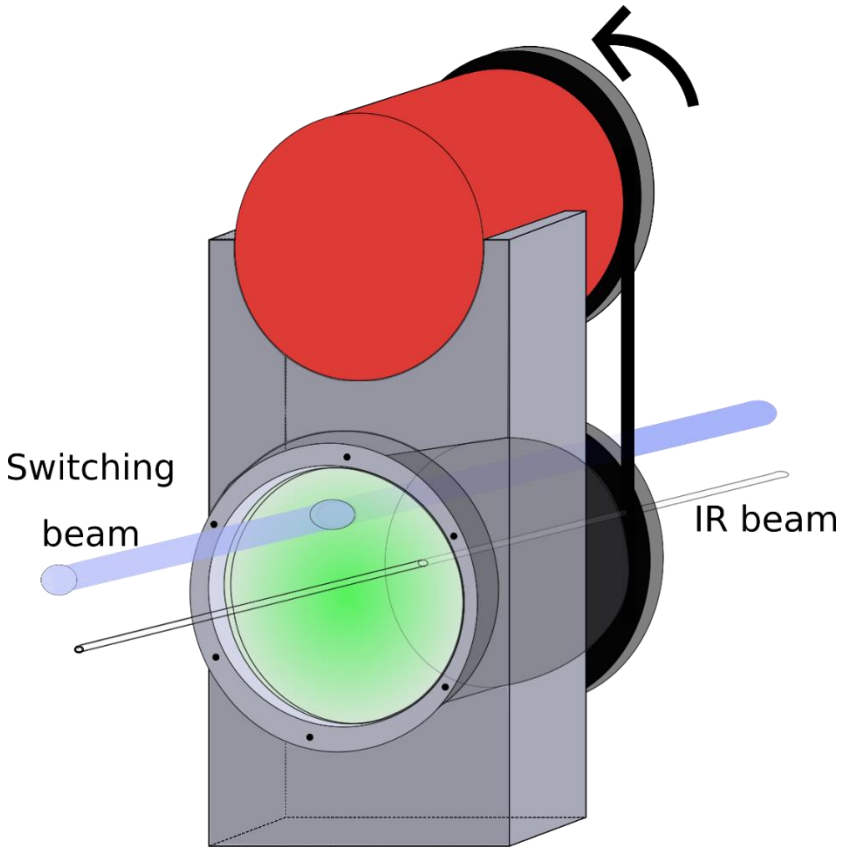


Figure A1.13 Rotating cell designed in LASIRE and used in IR On to Off photoswitching experiments.

The samples were introduced in a rotating cell with CaF_2 1 mm windows and a thickness between windows of 50 µm (see figure A1.13). IR spectra were taken with an FTIR spectrometer Vertex 70 with a detector MCT D316 cooled with liquid nitrogen, with a spectral resolution of 4 cm-1. The time to stabilize vapor in the air (moisture) before the measure was 12 minutes. The spectra reference was air. The range studied goes from 7000

to 550 cm^{-1}. Each spectrum was accumulated 128 times, and the time between two consecutive spectra was 28 seconds. A specific box was designed with polystyrene, and a transparent plastic window on top that entirely occupies the spectrometer space allowing to introduce the irradiation system. As irradiation source a fibered Hamamatsu Hg lamp with visible bandpass and an inference filter at 480 nm was used. The power of the lamp was set to 10%, to acquire enough spectra until reaching complete photoconversion of the samples.

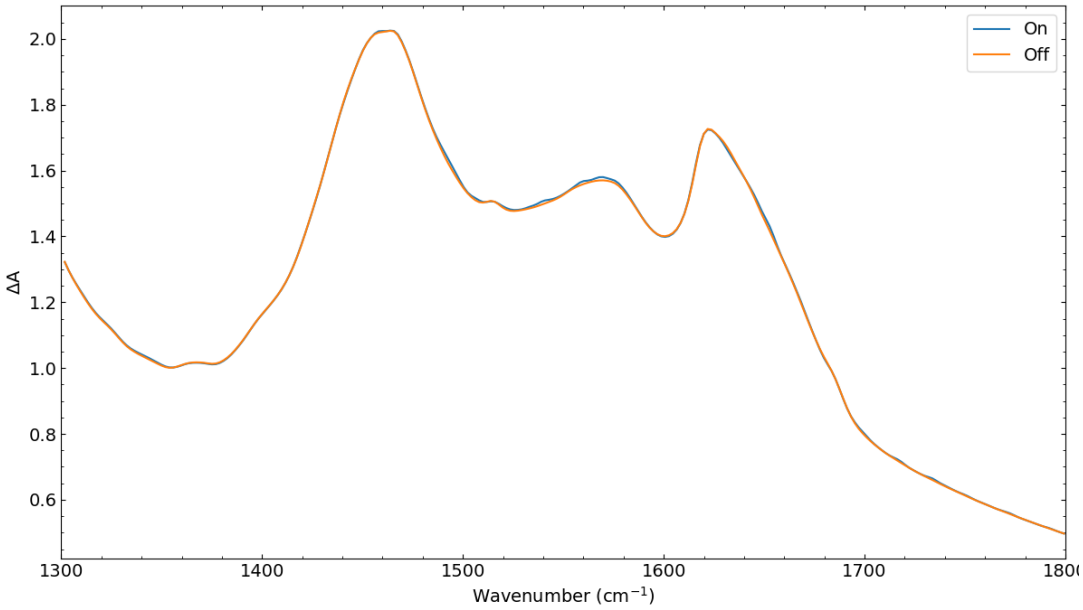


Figure *A1.14*. FTIR raw spectra obtained for WT variant between 1300-1800 cm^{-1}, On form in blue and *Off* form in orange. Figure made with matplotlib python library.

The FTIR spectra in the region of 1300-1800 cm^{-1} are mainly composed by two prominent bands corresponding to the strong absorption of amide 1 and 2 protein vibrations, which are found in every amino acid bond, as shown in Figure A1.14.. Therefore, *On* and *Off* spectral forms are very similar since the IR region probed the entire proteins and not only the chromophore. As a result, very little can be extracted from the raw data. Therefore, the initial spectrum was subtracted to all of the consecutive spectra to eliminate signals that have identical contributions in the *On* and *Off* forms and obtain more obvious variations of the light-induced conformational changes. These variations of absorbance are minimal and are strongly affected by the baseline drifts. To correct this effect, each spectrum mean value was subtracted to itself. The results of this correction for WT variant are shown in Figure A1.15.

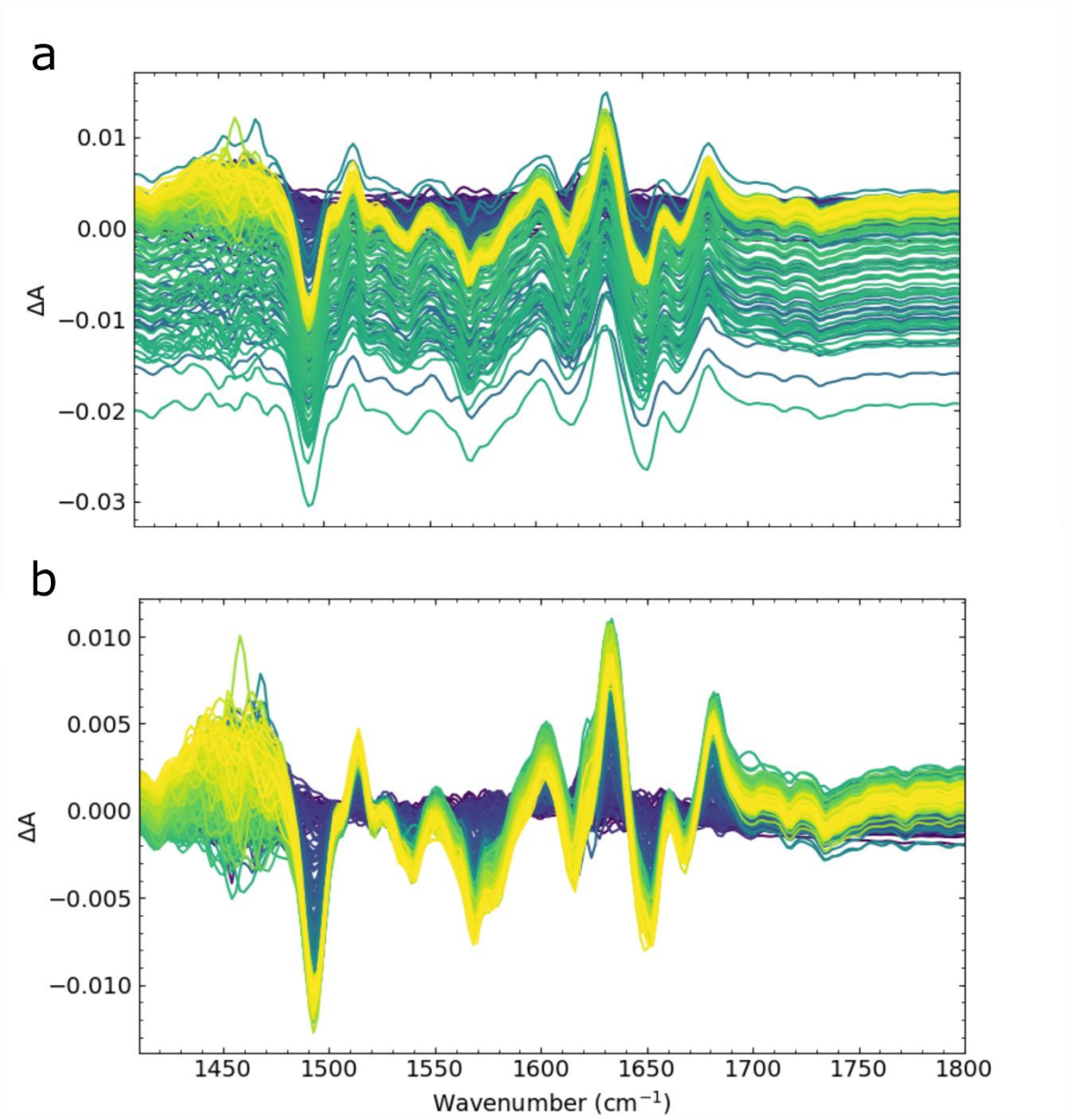


Figure A1.15. FTIR differential spectra obtained for WT variant between 1400-1800 cm^{-1}, a) correspond to uncorrected spectra, b) after baseline correction. Figure made with matplotlib python library.

A selection of the differential spectra measured for WT variant are shown in Figure A1.16. Similarly, to femtosecond transient absorption spectroscopy, negative bands correspond to the depopulation of the *On*-state form, while the positive bands correspond to the growing of the *Off* form.

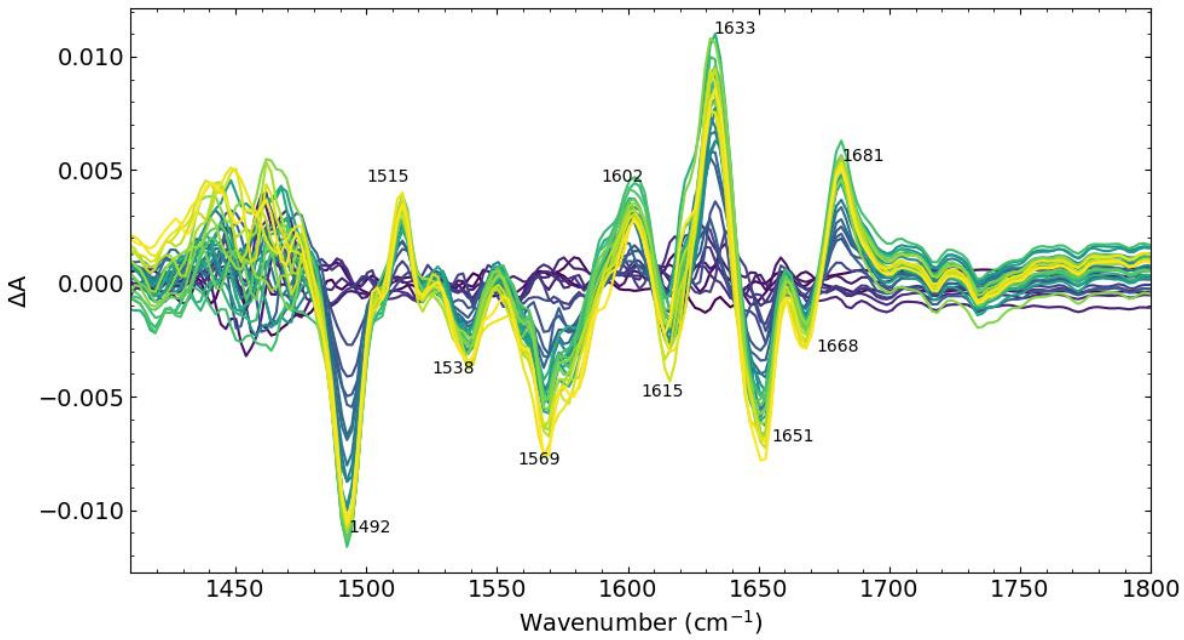


Figure A1.16. On to Off FTIR differential spectra for WT in D2O buffer under continuous irradiations at 480 nm. Figure made with matplotlib python library.

Several clear bands arise after irradiation of the protein. In comparison to Dronpa which has the same chromophore, the most intense band corresponds to the 1492 cm^{-1} peak which can be assigned to the C-O vibration of the phenolate for the *On*-form. The essential bands for the *On*-form are those at 1538, 1569, 1615, 1651 and 1668 cm^{-1}. The main *Off*-form characteristic band can be seen at 1681 cm^{-1}, and according to the literature[12–15], corresponds to the C=O stretching of the chromophore imidazolinone group. The other *Off* form main bands can be seen at 1514, 1602 and 1633 cm^{-1}. The peak at 1514 cm^{-1} corresponds to the phenol symmetric stretching.

Similar experiments were carried out for V15L and V151A. The differential spectra for the three proteins of *Off* forms minus On form at the photo-stationary state are displayed in Figure 4 A1.17. The spectra were corrected from baseline drift and normalized with the intensity of the highest peak (absolute value) which, for all three variants, corresponds to the C-O vibration of the phenolate *On* form.

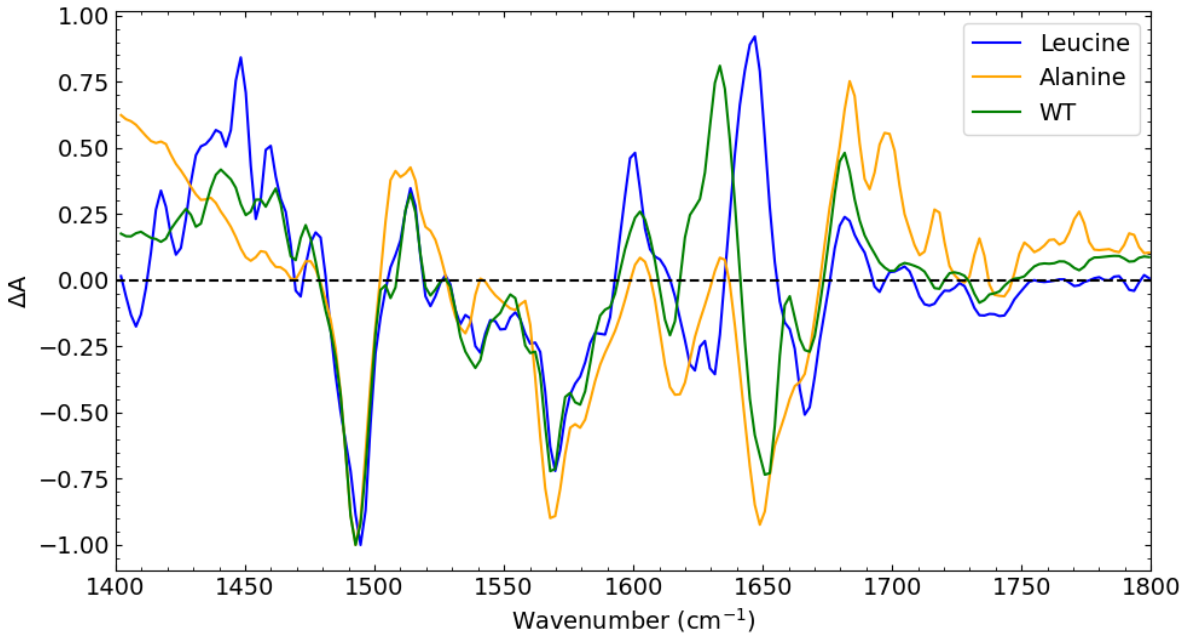


Figure A1.17, On minus Off FTIR differential spectra for WT (green), V151L (blue) and V151A (orange). Negative peaks correspond to On form, and positive peaks correspond to Off form. Figure made with matplotlib python library.

There are clear differences between the positive peaks of the differential spectra which correspond to the *Off* form for the three variants. Interestingly, and besides the similitudes found between the *On* forms X-ray structures and the UV-vis absorption spectra for the three variants, the FTIR negative peaks of the spectra also present several differences amongst them. Due to the differential character of the spectra represented in Figure A1.17, these differences could be originated by real structural differences between the variants *On* forms. However, most likely, they could be generated by differences between the *Off* spectra provoked as the *Off* forms have substantial structural differences. From the differences between the negative peaks, it is worth mentioning the red-shift of the C-O stretching band from 1492 to 1495 cm^{-1} for V151L compared to the other two variants which indicate a weaker C-O bond. It is also important the absence of the band at 1650 cm^{-1} for V151L. Nevertheless, this might be due to the presence of an *Off*-form intense positive peak at 1648 cm-1 present for V151L but not for WT and V151A. Finally, V151L has two negative bands at 1623 and 1631 cm^{-1} which are not visible for WT and V151A.

The negative peaks corresponding to the *Off* form have a clear difference, which is the positive band at 1633 cm^{-1} for WT. This band has a much lower intensity in V151A and is is absent in V151L, which has a unique positive band red-shifted at 1648 cm^{-}. These differences between V151L when compared to WT and V151A highlight the differences between the *Off* states *of the variants.* All the values for the different peaks for the three rsEGFP2 variants are resumed in Table A1.2.

Table A1.2 FTIR main peaks for differential spectra in Figure A1.17.

Positive peaks cm^{-1}			**Negative peaks cm^{-1}**		
WT	**V151L**	**V151A**	**WT**	**V151L**	**V151A**
		1507	1492	1495	1492
1514	1514	1514	1538	1542	1535
1602	1599	1602	1569	1570	1569
1633	1645	1634	1615		1617
1681	1682	1684		1623	
		1698		1631	
			1651		1649
			1668	1667	

Importantly, from the TRIR differential spectra, the WT variant displays two negative peaks at 1650 cm^{-1} and 1668 cm^{-1}, whilst the first peak is absent for V151L, only displaying the 1668 cm^{-1} peak. The V151A variant has the opposite feature displaying the peak at 1650 cm^{-1} and the absence of that at 1668 cm^{-1}. It is worth noticing that since the amount of protein of the TRIR spectra is unknown, the same decomposition done for the UV-Vis absorption spectra cannot be performed.

A1.6 References

(1) Coquelle, N.; Sliwa, M.; Woodhouse, J.; Schirò, G.; Adam, V.; Aquila, A.; Barends, T. R. M.; Boutet, S.; Byrdin, M.; Carbajo, S.*; et al.* Chromophore twisting in the excited state of a photoswitchable fluorescent protein captured by time-resolved serial femtosecond crystallography. *Nature chemistry* **2018**, *10*, 31–37.

(2) 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. *Scientific reports* **2016**, *6*, 18459.

(3) Woodhouse, J.; Nass Kovacs, G.; Coquelle, N.; Uriarte, L. M.; Adam, V.; Barends, T. R. M.; Byrdin, M.; La Mora, E. de; Bruce Doak, R.; Feliks, M.*; et al.* Photoswitching mechanism of a fluorescent protein revealed by time-resolved crystallography and transient absorption spectroscopy. *Nature communications* **2020**, *11*, 741.

(4) Tokunaga, A.; Uriarte, L. M.; Mutoh, K.; Fron, E.; Hofkens, J.; Sliwa, M.; Abe, J. Photochromic Reaction by Red Light via Triplet Fusion Upconversion. *Journal of the American Chemical Society* **2019**, *141*, 17744–17753.

(5) Jacquet, M.; Uriarte, L. M.; Lafolet, F.; Boggio-Pasqua, M.; Sliwa, M.; Loiseau, F.; Saint-Aman, E.; Cobo, S.; Royal, G. All Visible Light Switch Based on the Dimethyldihydropyrene Photochromic Core. *The journal of physical chemistry letters* **2020**, *11*, 2682–2688.

(6) Müller, P.; Brettel, K. Ru(bpy)(3)(2+) as a reference in transient absorption spectroscopy: differential absorption coefficients for formation of the long-lived (3)MLCT excited state. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology* **2012**, *11*, 632–636.

(7) Brillouin, L. *Wave propagation and group velocity*, 7. pr; Pure and applied physics 8; Acad. Pr: San Diego, 1988.

(8) Heisenberg, W. ber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Z. Physik* **1927**, *43*, 172–198.

(9) Sellmeier. Zur Erklärung der abnormen Farbenfolge im Spectrum einiger Substanzen. *Ann. Phys. Chem.* **1871**, *219*, 272–282.

(10) Ghosh, G. Sellmeier coefficients and dispersion of thermo-optic coefficients for some optical glasses. *Appl. Opt.* **1997**, *36*, 1540–1546.

(11) Malitson, I. H. Interspecimen Comparison of the Refractive Index of Fused Silica*,†. *J. Opt. Soc. Am.* **1965**, *55*, 1205.

(12) Malitson, I. H. A Redetermination of Some Optical Properties of Calcium Fluoride. *Appl. Opt.* **1963**, *2*, 1103.

(13) Nakayama, T.; Amijima, Y.; Ibuki, K.; Hamanoue, K. Construction of a subpicosecond double-beam laser photolysis system utilizing a femtosecond Ti:sapphire oscillator and three Ti:sapphire amplifiers (a regenerative amplifier and two double passed linear amplifiers), and measurements of the transient absorption spectra by a pump-probe method. *Review of Scientific Instruments* **1997**, *68*, 4364–4371.

(14) Stoll, T.; Gennari, M.; Serrano, I.; Fortage, J.; Chauvin, J.; Odobel, F.; Rebarz, M.; Poizat, O.; Sliwa, M.; Deronzier, A.*; et al.* Rh(III)(dmbpy)2Cl2+ as a highly efficient catalyst for visible-light-

driven hydrogen production in pure water: comparison with other rhodium catalysts. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2013**, *19*, 782–792.
(15) Ghose, A.; Rebarz, M.; Maltsev, O. V.; Hintermann, L.; Ruckebusch, C.; Fron, E.; Hofkens, J.; Mély, Y.; Naumov, P.; Sliwa, M.*; et al.* Emission properties of oxyluciferin and its derivatives in water: revealing the nature of the emissive species in firefly bioluminescence. *The journal of physical chemistry. B* **2015**, *119*, 2638–2649.
(16) Nagasaka, T.; Sotome, H.; Morikawa, S.; Uriarte, L. M.; Sliwa, M.; Kawai, T.; Miyasaka, H. Restriction of the conrotatory motion in photo-induced 6π electrocyclic reaction: formation of the excited state of the closed-ring isomer in the cyclization. *RSC Adv.* **2020**, *10*, 20038–20045.
(17) Greetham, G. M.; Sole, D.; Clark, I. P.; Parker, A. W.; Pollard, M. R.; Towrie, M. Time-resolved multiple probe spectroscopy. *The Review of scientific instruments* **2012**, *83*, 103107.
(18) Chung, H. S.; Khalil, M.; Smith, A. W.; Tokmakoff, A. Transient two-dimensional IR spectrometer for probing nanosecond temperature-jump kinetics. *The Review of scientific instruments* **2007**, *78*, 63101.

Appendix 2. Ultra PyFit A Python 3.7 self-made Software

This appendix presents the most relevant information and features of Ultra-PyFit. This software application is a self-made graphical user interface (GUI) done using PyQT, which mainly interacts with the GlobalFit python class used in the thesis to fit the data, and both have been fully developed during the course of the PhD. The main target of this class is to accelerate every process between the collection of raw data from the lab and the final output results using the most common classical analysis for time-resolved spectroscopy data, but also more advanced techniques. The software includes data correction processes, fitting, validation and analysis of the error and model used to fit the data and, finally, the results are presented in a clear and understandable manner with the use of full functionalities of matplotlib and Seaborn python packages offering nicely formatted figures to the user.

The origin of such a tool came from the necessity of a python package which allowed data-fitting including from simple approaches (weighted sum of exponentials) to a wide range of target models to multi-way spectroscopy data. Apart from the obsolete Pytra package[1] for python 2.7, there was no python 3.5-3.8 available packages (January 2021). It is worth noting that such tools for fitting a wide range of models to multi-way spectroscopy data with different approaches already exist in other languages such as R[2] or MatLab[3–5]. Nevertheless, in none of them, a proper analysis of the fitting error is done and nor the most simple approach of fitting the data with a weighted sum of exponentials is possible. One of the most significant improvements of this python Class is the implementation of error analysis and model validation for fitting a wide range of models to multi-way spectroscopy data.

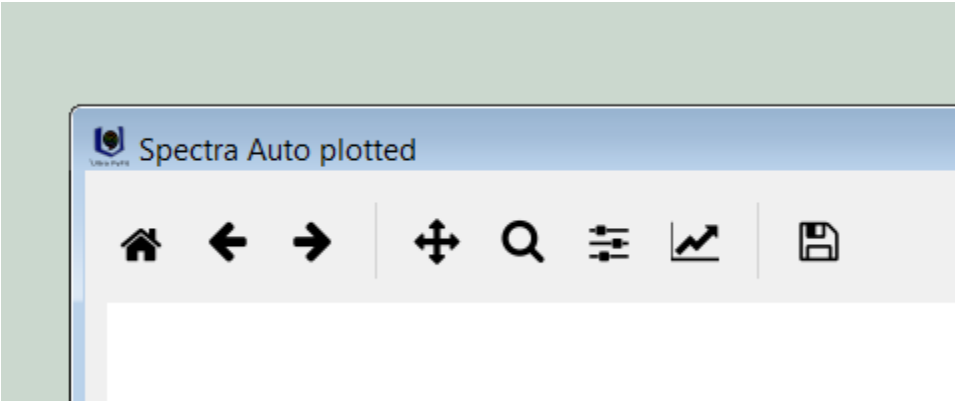
Finally, regarding the development of the GUI, it is clear that there are also already developed Matlab Toolboxes such as Ultrafast toolbox[3] and Optimus[4], or Glotaran[6] a Java-based GUI developed for Timp-R package[2]. However, still today, they all require Matlab or either R environments to be installed in the computer to work. On the contrary, Ultra PyFit has been compiled with pyinstaller package, and an installator has been created with Inno set-Up.

Therefore, there is no need to have Python installed on the computer. Moreover, the huge advantages for formatting figures and presenting results of Matplotlib[7] and seaborn[8] python packages together with the many widgets available in PyQt, creates a user-friendly and accessible GUI, for correcting data, designing the models, fitting and validating the models, and exporting the results. Ultimately, the development of the GUI came along with the necessity of transferring and making available the developed features that can be done with the GlobalFit Class, for users and lab members with zero or little programming and Python knowledge, for whom knowing the python syntax and the parameters of the GlobalFit functions and their many options may represent a challenge. This project was developed with the contributions of Stanislaw Niziński, especially regarding creation of a model based interface to easily generate models for a target fit (section A2.5b). Finally, a β- version of the program is available where the fitting algorithms have been verified. Although there is still work in progress, specially making the GUI more user friendly, writing documentation and development python test.

A2.1 Introduction of Ultra Pyfit.

Figure A2.1 Screenshot of the application window with windows OS with the main components: a) subwindow area (light-green). b) action menu (light-grey) with 4 sub-menus and fitting button.

Ultra PyFit is based on a main window consisting on two parts: an action menu in the left side of the window (light-grey) and a middle area (light-green) Figure A2.1. The first menu has 4 sub-menus, which are “File menu”, “Preprocess”, “Datasets” and “Plot menu”. File menu contains actions related to loading data, loading a previous project, saving the current project, displaying a report of actions and exporting data and results. The pre-process menu deals with all available pre-process actions, and is presented later on (section A2.4**)**. The dataset menu is used to recover and undo any pre-process action. Finally, the plotting menu deals with presenting the results of the current activated data. Besides these four submenus, there are two buttons: "Exp settings" which is used to introduce the units of the current data sets (time unit and wavelength unit), and the Fitting button, which has two options - Exponential Fit and Target Fit. The second area is used for displaying an infinite number of sub-windows which will be activated subsequently for the user to interact with and guide him/her through the current process. For all displayed figures, there is an action bar designed for interacting with them, with eight buttons this is the standard matplotlib navigation toolbar Figure A2.2. The first five buttons can be used to zoom inside the figures and come back to the original perspective view. The sixth button is for interacting with the figure and white space, and the seventh for interacting with the actual elements of the figures like legends, plotted lines and axis labels. Finally, the last button is designed for exporting the figure in several different formats.



FigureA2.2 Matplotlib navigation toolbar integrated inside ultra PyFit figures.

The application works on a main single thread (“a thread of execution is the smallest sequence of programmed instructions that can be managed independently by a scheduler”, Wikipedia), which mainly handles the GUI processes and fast actions that do not require

much time to be finished. This means that the execution is driven in response to user interactions such as clicking on a button which creates an *event. This event*, when executed, produces an expected output. The events are handled sequentially and are therefore pushed onto and taken of an *event queue* to be processed, meaning a new event is not executed until the precedent event has finished. Therefore, running a process that can take several minutes, such as fitting a model to the data, will force the new events to be stacked in the *events queue* which can make the OS interpret to believe that the program is being blocked and not responding. To avoid such a situation, several "working threads" created from QThread class are instantiated when the application is started so that certain actions inside the application such as long calculation processes are transferred to these working threads to be handled. Finally, as commented above, the program has been compiled. Therefore, Ultra Pyfit is a complete application based on PyQt that doesn't require any other installed programs or interpreters, while still benefitting from the main functionalities that can be obtained from lmFit and Scipy Python packages for data analysis, and matplotlib and seaborn for Interactive data exploration via a specific and straightforward designed interface with specifically designed data editor. The main characteristics of the application are:

- Importing data in various known data file formats, with an interactive menu to specify any specific options such as decimal use column separator, and organized file, which can read the main data files such as *.dat, *.txt and *.csv among others. (figure A2.3).

- Several pre-processing options to correct the data sets from physical generated effects, such as the GVD (Appendix1 A1.1d), or correct initial signals by baseline subtraction, or eliminating data errors such as outliers, or saturated areas.

- Interactive data exploration via an exploring data window, where it is possible to directly interact with the data set to directly select spectra and traces to visualize separately, via a 3D plot of the data and single trace fitting for a fast initial estimation of parameters, and full singular value decomposition (SVD) exploration.

- Clear separation of fitting procedures to either an exponential fit more widely used where no model is imposed, or a target fit where any kind of designed photochemical model can be easily designed via the custom-designed model creator window, which allows any model to be saved.

- Pre-fittings options which allow verification of the initial model or initial estimations of the parameters.

- Interactive results inspection via specific designed windows where several fits can be compared to each other.

- Analysis of errors and model validation via Bootstrap and F-test which, although they are lengthy processes, can give a good estimation of the quality of the fit.

- Tracking of the actions performed, which can help towards reproducibility and verifying analysis of actions by any external user or by the user itself months or years later.

- Interactive figures, thanks to a self-made matplotlib cursor which allows direct interactions with the data. Finally, nicely formatted figures that can be directly saved into several types of possible formats thanks to Matplotlib library.

A2.2 Importing data and GlobalFit Class instantiation.

Importing data into Ultra PyFit it is straightforward via the sub-menu File-menu where the first option is Load data, after right-click a window where the computer explorer can be opened will pop out. After selection of the desired file, in the main menu several options will become available for specifically indicating the file specifications, such as header wavelength and time Column/row and the separator and a window where the file is ready for the user to inspect it. After indicating the parameters, they can be saved as default for a faster initialization next time. Once the data is loaded, the file will be read using the functions from

pandas package read_csv() and a homemade function to identify possible strings and transform then into data for the time and wavelength vectors. After the process, a wavelength vector, a time vector and a data matrix will be generated which will be then used to instantiate the GlobalFit class named as "Experiment". This is all automatically done by the program by clicking on the "Load by default" or "Load data" buttons.

Once instantiated, it is important to know that from the data set it is possible to select traces from the pre-process sub-menu or fitting buttons. For all the functions related to the traces, such as plotting traces, fitting them or exploring the singular value decomposition, the functions will only work on the subset of selected traces. On the contrary, for the remaining functions, like for example plotting spectra, the complete data set will be taken into account. Finally, a new selection of traces can be done at any moment.

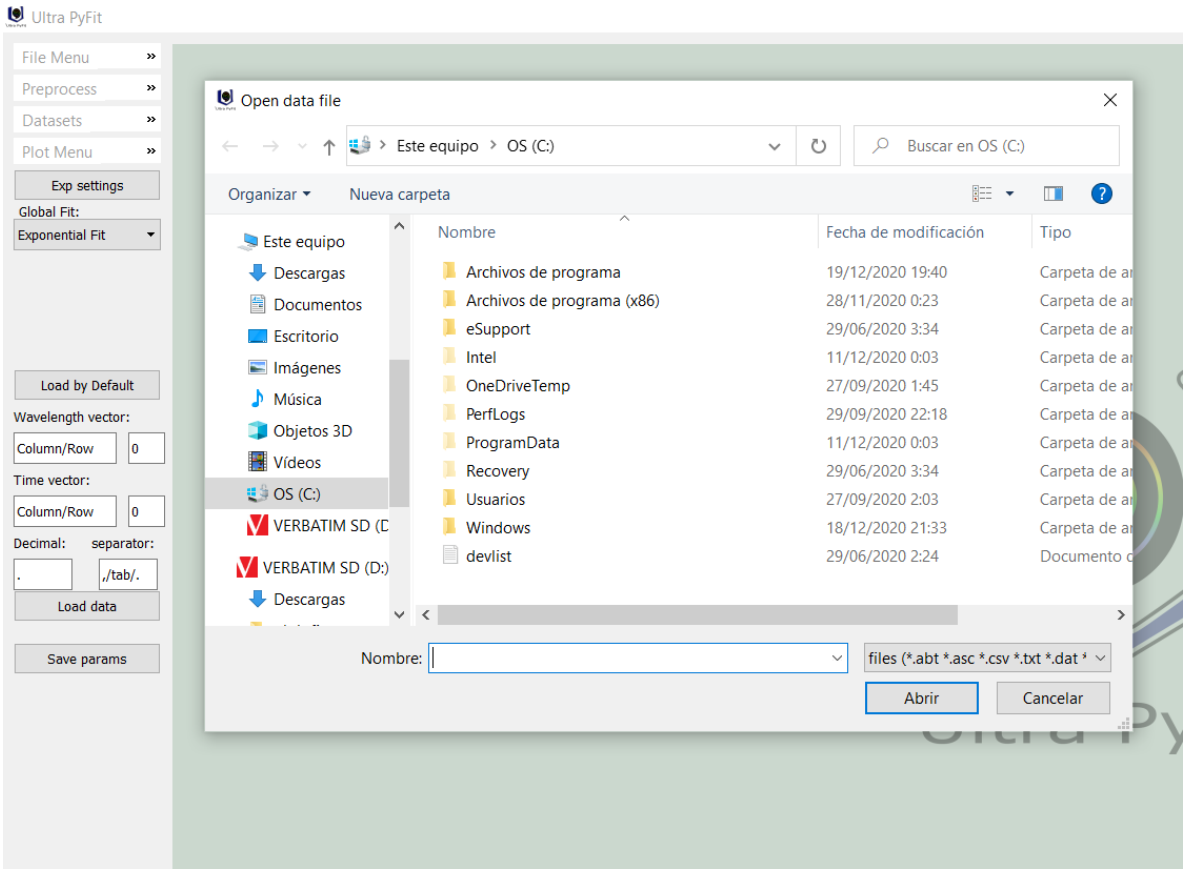


Figure A2.3 Interface for loading data into Ultra-PyFit and instantiate the GlobalFit python Class.

A2.3 Exploring the data

After loading the data, the first thing to do is to explore the spectral and time dimensions. To do so, in the plotting menu, there is the option of exploring the data. Clicking on the button will pop out a window where the figures will be displayed (see Figure A2.4), the first one top left is a 2D map representation of the data where warm colours represent positive differences of signal, and cold colours represent negative values. The right figure represents the spectra at the time where the slider is and the left bottom the corresponding traces. These two last figures are interactive. From the spectra, any trace can be selected, and vice-versa, and the selection of traces and spectra can be further plotted. Finally, the traces can be selected for fitting.

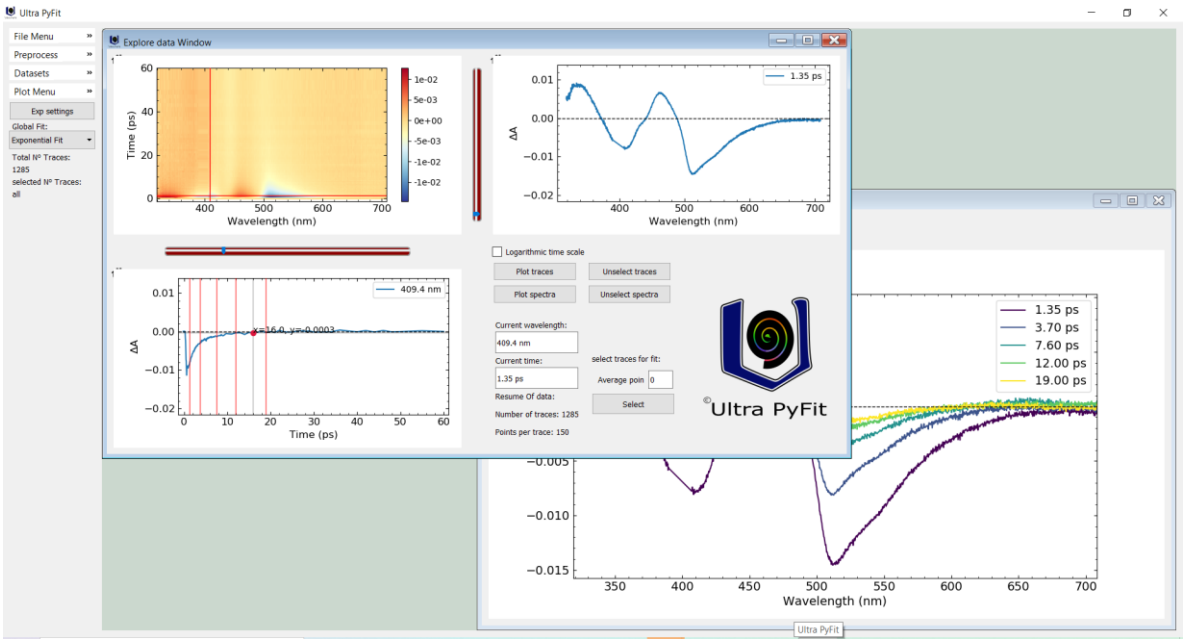


Figure A2.4. Exploring the data window, the window under the explore data panel corresponds to the spectra selected directly on the time trace (vertical red lines). Note, this is the only window where the figures cannot be saved; this is done on purpose as the rest of plotting functions yield nice formatted figures.

To further explore the data, we propose three fast plotting options. The first one is to plot the data in 3D, which will pop out a 3D map of the data surface which can be turned to interact and visualized from different directions. The second is to plot the spectra automatically; this function plots eight spectra equally spaced at the maximum absolute value of the dataset. At

the same time, new buttons in the left side main menu will be available after plotting the spectra automatically to have a personalized plotting. The possibilities are to select the time region for plotting, select the number of points to average, the legend properties, the number of spectra to plot and the wavelength value at which they should be equally spaced. Finally, as a fast plotting option, it is also possible to plot ten separated traces from the dataset. In Figure A2.4, a 3D plot and the result of plotting the spectra automatically together with the menu can be seen. Moreover, all colours maps available in Matplotlib are available in the program for a personal customization of the figures, and can be selected in the plot sub-menu in the plot spectra option "change colour". By default, the colour map is “Viridis”. Finally, it is worth mentioning that by adjusting the size of the figure window, we directly affect the size of the exported and saved figures.

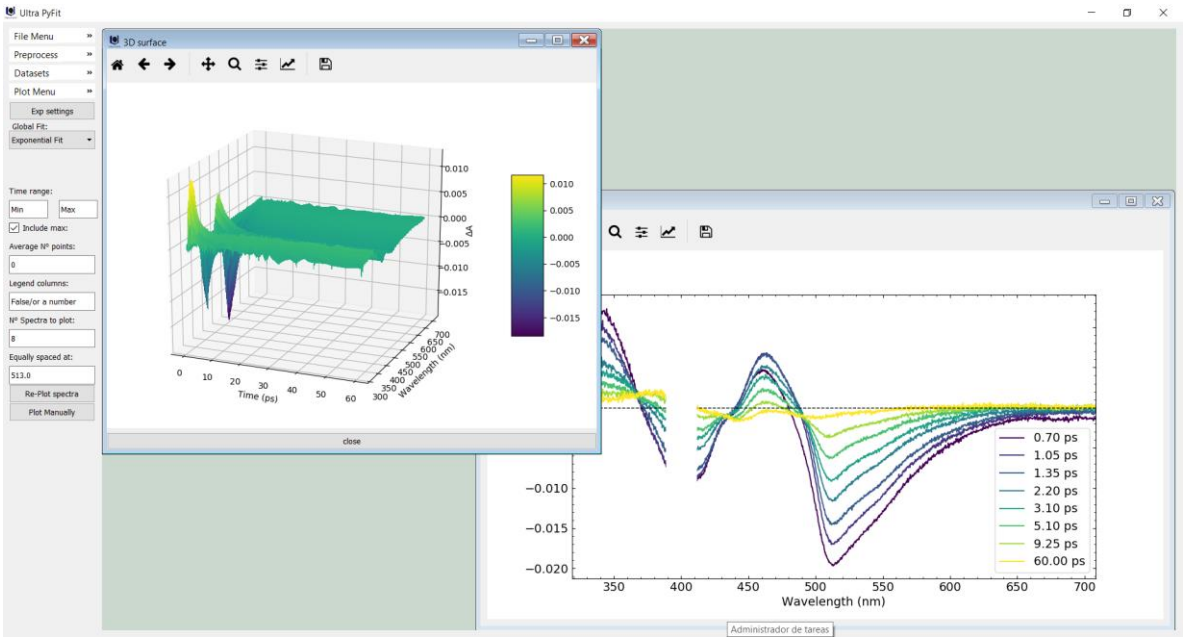


Figure A2.5. 3D plot, the automatically spectra plotted windows and the menu to manually customize auto-plotting.

After spectrally exploring the data, the singular value decomposition (SVD) of the dataset can help to know the number of spectrally and temporally independent components. The SVD is a matrix factorization technique[9]. See next section.

The complete calculation of all the singular vectors is a cost calculation for a matrix with a high number of rows or columns. Therefore GlobalFit uses scipy.sparse.linalg.svds() function which computes a N number of singular values/vectors (truncated SVD) for a sparse

matrix, where the order of the singular values is not guaranteed although is generally correct for the main vectors. This function uses Arpack[10], which is a Fortran package which provides routines for quickly finding a few eigenvalues/eigenvectors of large sparse matrices. In experimental datasets, the noise causes the most significant singular vectors to be clearly defined and might stand out from the rest. Nevertheless, in some cases, it is difficult to differentiate the number of significant SV. In Ultra PyFit 15, SV are calculated by default. For the exploration of the SVD space in the plotting menu, there is the option of plotting only the singular values or plot the singular values together with the right and left singular vectors. As is explained further along, there is an option to fit the left singular vectors. In figure A2.5, the SVD window can be seen with the singular values and left and right singular vectors.

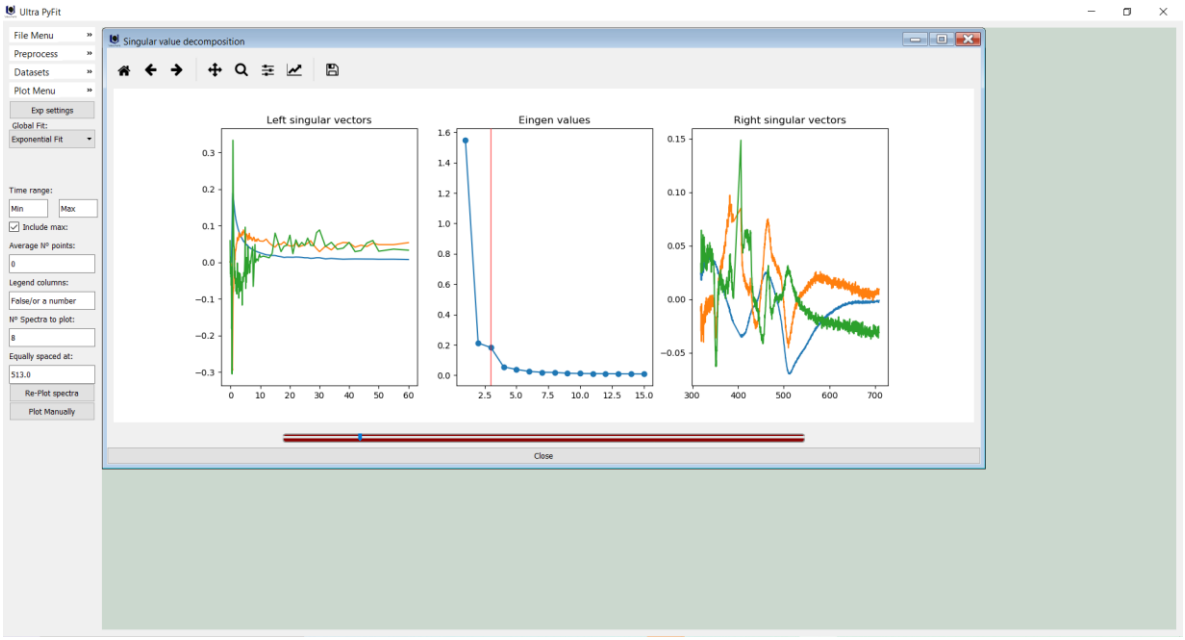


Figure A2.5. Ultra Pyfit window displaying the complete SVD, including the singular values and left and right singular vectors

A2.4 Pre-processing options

There are several basic pre-processing features available in UltrPyFit which will be briefly described in this section:

Calibrations of the wavelength vector. This can be done with four different options using a specific self-made Calibration python class: 1) introducing the first and final points of the

calibration curve, assuming the relation between the pixel and wavelengths is linear. 2) Copying a vector of wavelengths. 3) Fitting a third-order polynomial giving at least four points or more for the fit. 4) Interactively from the spectra of a reference sample taken with the spectrometer and the known spectrum of this reference. For each of the possibilities, either a window will pop-out or new options will be made available in the menu bar. In Figure A2.6, the window for calibrating with the spectra of a reference can be seen, with some polyester spectra as an example.

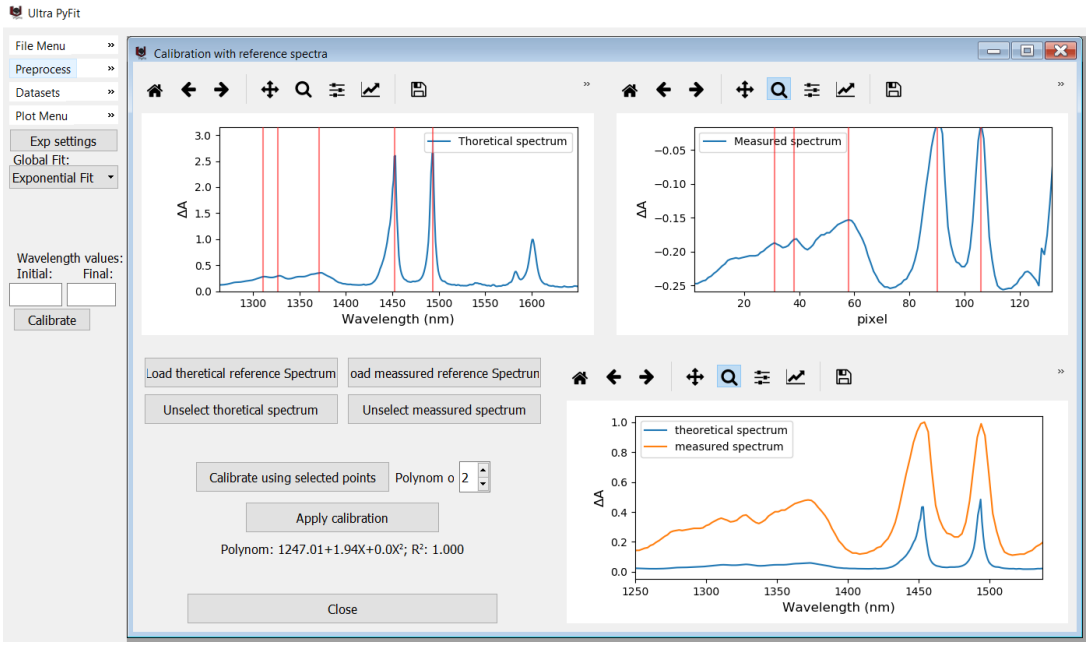


Figure A2.6 Sub-window to calibrate the dataset with a reference spectra. The figures after calibration can be independently saved. The selection of calibration points can be easily done with the cursor that directly pops into the window, with a left-click to select a new point, and a right-click to de-select the last point.

Baseline subtraction: two options are proposed, to either subtract a specific selected spectra, or a selection of spectra within a range. This will generate a new corrected data set, although it is always possible to go back to the original dataset.

Group velocity dispersion correction: For the correction of the GVD[11], we propose to either calculate it fitting a polynom or calculating it using the sellmeier equation[12], as has been explained in the Appendix 1 (A1.1d section). Once calculated, a window will pop out

to verify the correction. Examples of the correction can be seen in Figure A1.5 and Figure A1.6 in the Appendix 1, and the algorithm to correct the data set can be seen in Figure A1.7.

IRF fitting: for fitting instrument response function, we propose to fit the stimulated Raman amplification signal of the solvent. This can be done directly from the data set if the peaks are visible, or alternatively, we propose to load a solvent measurement where to fit it. The solvent measurement will be loaded with the same parameters as the loaded file, as we assume they are measured in the same apparatus. For details on the fitting, see the Appendix 1 (section A1.1b).

Cutting or selecting areas: in the program, it is possible to select or cut areas of the data for further analysis. A classical area that will be cut in TRUV-Vis data is the laser excitation area.

Removing single points: Via an interactive figure, it is also possible to eliminate single points in the data that can, for example, come from a dead pixel.

Averaging points in time dimension: This option is proposed as in many time-resolved experiments, such as flash photolysis or TRMPS, the number of recorded in longer delay times is very high. Thus by averaging points in areas where the dynamics between points do not vary, we can improve the S/N ratio. For averaging the time points, we proposed to either average the points in between constant increments or to increase the steps for the points to be average.

Smooth and derivation: As seen in Chapter 8, an alternative to correct the baseline fluctuations is to derivate the data. This is done with fitting a Stavisky-Golay polyonomy and derivate, the result, thus a zero order derivation will result in only smoothing the data.

A2.5 Fitting options in ultra PyFit.

GlobalFit class uses lmfit python package[13], which is a free software using an Open Source license, to fit the data. The interface can be used to manually build a number of models and instantiating parameter to these models with the use of lmFit package. This library provides a high-level interface to non-linear optimization and curve-fitting problems for Python. It is

based on Scipy optimize function[14] and extends many of the optimization methods. Although lmFit has several different optimizer algorithms for minimizing the error surface, GlobalFit only uses the popular Levenberg-Marquardt method from scipy. optimize.leastsq. SciPy is the most standard python library for optimization (it also has many other utilities such as integration, interpolation, eigenvalue problems, algebraic equations, to name a few). It is built on top of the Numeric array data structure from Numpy[15], and it has become one of the most used libraries in Python, with over 600 unique code contributors, thousands of dependent packages and over 100,000 dependent repositories and millions of downloads per year[14].

A2.5a Fitting with a weighted sum of exponential

Physical explanation: To unravel the chemical reactions and processes behind the multiway data, a model-based analysis of the data is mandatory[5,16–18]. The most simple approach to model a photochemical reaction is to assume that the concentration of transient species or reaction rate follows a first-order reaction. According to this hypothesis, the data should be explained by a weighted sum of exponential functions, where every exponential time represents the decay of a transient species that follow a first-order reaction, and the pre-exponential value or weight is the concentration of this species. Logically, fitting one of the kinetic traces records at one wavelength can yield an estimation of the number of transient species involved in the reaction and their decay constants. One of the most established approaches to analyze a certain number of selected traces (multi wavelengths) with the same sum of exponentials[5,16,19,20]. This approach, typically known as global analysis, has been demonstrated to be much more robust and accurate than separate or single wavelength analysis. Furthermore, from a global analysis of the data, not only the decay time constants and number of transient species can be obtained, but also their spectral shapes from the weights (pre-exponential factor) of the sum of exponential at each of the wavelengths, which constitute the so call Decay Associated Spectra (DAS).

$$\psi(t) = \sum_0^n A_i * e^{(\frac{x-x_0}{t})}$$ Equation A2.1

$$\psi(t) = \frac{1}{2} e^{\left(\frac{x-x_0}{t} + \frac{1}{2}\left(\frac{\sigma}{t}\right)^2\right)} * \left(1 + \text{erf}\left(\frac{(x-x_0) - \frac{\sigma^2}{t}}{2\sqrt{\sigma^2}}\right)\right)$$ Equation A2.2

Ultra PyFit allows for single and global fitting with a sum of weighted exponential function (Equation A2.1), which can be either convolved with a Gaussian IRF (Equation A2.1) or not. It also allows fitting a selection of left singular vectors, with different output windows which allow always to understand what procedure we are applying.

Single trace fitting: after clicking the Exponential fit button, there is available a new submenu for selecting traces to be fitted. The first option is "Select one trace and fit" this will generate an auto-plotted spectra figure from where a single trace the average points, and the number of exponentials can be selected. Just by clicking on the "select and fit", a first fit will be performed with all conditions unrestricted and an initial estimation of the parameters done automatically by the program. This fit is nearly automatically completed, and the results will be displayed in a new window Figure A2.7. This window consists of two main parts; the first one is a figure where the selected trace is scattered in blue points with the fit result in red and the residual plot on top of its trace with the fitting result. The second part is the fit-report with all the details of the fit and buttons to change the fitting conditions and re-Fit with the new selected conditions. It is worth mentioning that the fitting regions can also be selected by directly clicking on the left-figure. If the fit is not successfully achieved, the parameters can be changed or fixed to known values manually to improve the fit in the open window. Finally, once the fit is completed and we have obtained logical parameters of the single trace fit, it is possible to export them and initialized them as initial guess for a global fit. An example of the single fit window can be seen in Figure A2.7. A single trace fitting is an excellent way to have an initial estimation of the decay times of the dataset. It is always advisable to fit independently several of the most representative traces to have a global idea of the dataset before starting a global fit where several traces are fitted together. Finally, the single trace fit can be stored in memory and can always be retrieved after form the previous fit button, from where it can be compared to any other single fit or integral band fit saved.

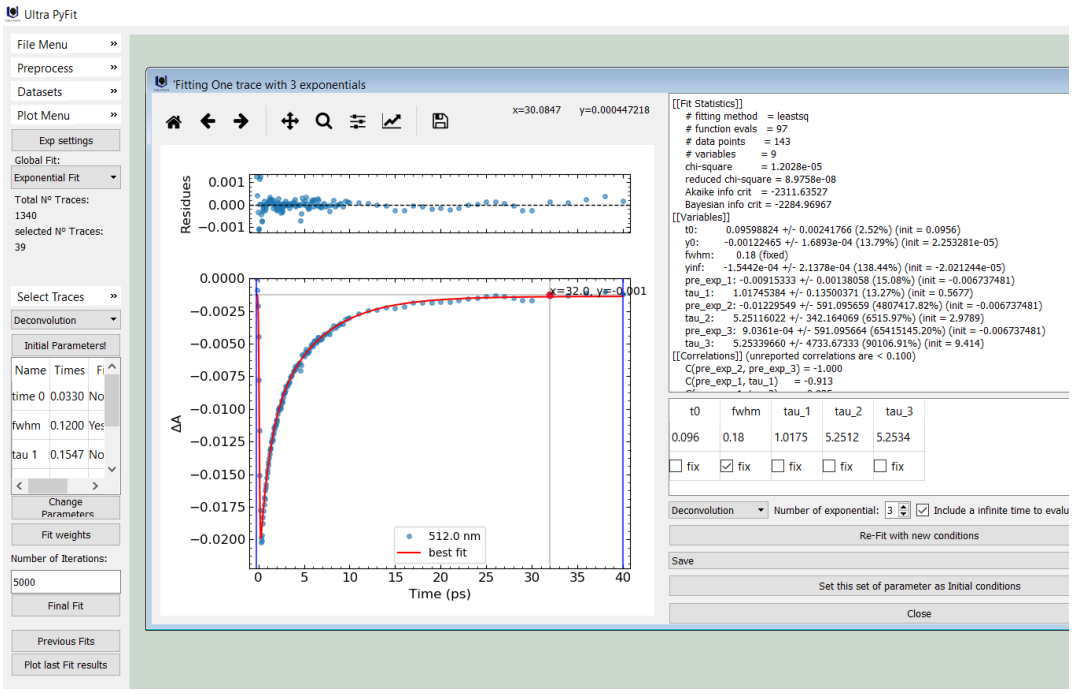


Figure A2.7. Example of window interface after single-trace fitting that is automatically opened after selection of the trace to fit. Notice all parameters are unfixed.

Integral band fit: Similarly to the single trace fit, it is possible to perform an integral band fit. Such an analysis allows for example to remove features and time constants that are the result of a simultaneous I negative and positive contribution in the spectra. Some processes causing such features are vibrational cooling or solvation phenomena. The Interface for the integral band fitting is identical to the single trace fitting except that in the first selection graph two traces need to be selected to define the integration area and therefore no average can be selected.

Global trace fitting: As explained above**,** a global fit consists on fitting several traces simultaneously, where the decay times and IRF factors (and t0 if the GVD has been corrected) are globally estimated from the ensemble of traces, while the pre-exponential factors are independently fitted for every trace. To perform a global fit, the first thing to do is to initialize the parameter's value. This can be done in the exponential fit section from the single trace fitting interface (see section above), or via the initial parameters button. This button displays a window where an infinite number of exponentials can be chosen and the option to fix their decay times or not. Once initialized, they will be visible in the parameter table together with all other global parameters. Finally, as expected, a global fit generates a

large surface where the absolute error between the data and the weighted sum of exponentials is minimized. Thus, such analysis cannot be performed to the ensemble of the data set. Nevertheless, this is not a problem as only fitting a few number of traces globally can give a good estimation of the decay times. Ultra PyFit can handle up to 75 traces without problems and allows to fit generally up to 180 traces (Depending on the number of parameters), although it is not recommended to fit more than 100 as this increases the fitting time but not the estimation of the times. Therefore, before doing a global fit we recommend selecting traces using the function "select a series of traces" from the select traces sub-menu. This function allows to select traces equally spaced according to an indicated wavelength distance, and it also allows to average the selected traces with the just direct adjacent traces which increase the S/N ratio of the exported final trace and can help to improve the quality of the fit. Finally, it is also possible to exclude regions where traces should not be selected, for example, traces affected by the Raman scattering signal from the solvent or laser excitation, which can affect the estimation of decay times.

Once the parameters are initialized and the traces selected, the program will automatically generate the initial estimations for the fit. It is possible to perform a pre-fit, where the selected traces will be fitted independently with the decay times fix; thus only the pre-exponential factors are optimized, and the results can be seen by the plot fit results button. This allows having a hint of how good the initial estimations are. If they are not good enough, they can be changed with the change parameters button.

By clicking on the “Final fit” button, the global fit will start. The parameters to be optimized will be changed to either convergence, or the maximum iterations are reached (by default 5000). The fitting menu will transform into a progress bar. It is possible to perform other actions while fitting, although we recommend waiting as it generally takes less than a minute. Once finished, a report window will appear, and the fitting menu will become available again. The report window contains the main information, details of the fit and six extra buttons to further explore the fit.

- “Verify Fit”, which displays at a time a single trace and its corresponding fit from the global fit.

- “Plot DAS” which plot the decay associated spectra (DAS), which corresponds to the pre-exponential factors of the sum of exponentials fitted.
- "Plot Fit", displays the fits plot, which includes all traces with the corresponding fits and residual plot. From this figure, it is possible to select independent traces to have a clearer view of the most interesting traces independently. If less than ten traces are selected, the wavelength will be displayed in a legend.
- "Extended fit report" displays an extended report with details of the fit and all of the parameters as well as correlations between them and the standard error estimation for each parameter.
- “Calculate confidence interval” opens a window to verify the quality of the model and confidence intervals, using an F-test or via a bootstrap of the residues. Check the specific section for more details.
- “Export to Target fit", this button exports the retrieved times and creates populations or species that have each of the associated times, to facilitate the creation of a model to initialized a target fit. Check the next section for more details.

The window fit associated to a global fit can be seen together with the DAS plot and the fit plot in Figure A2.8.

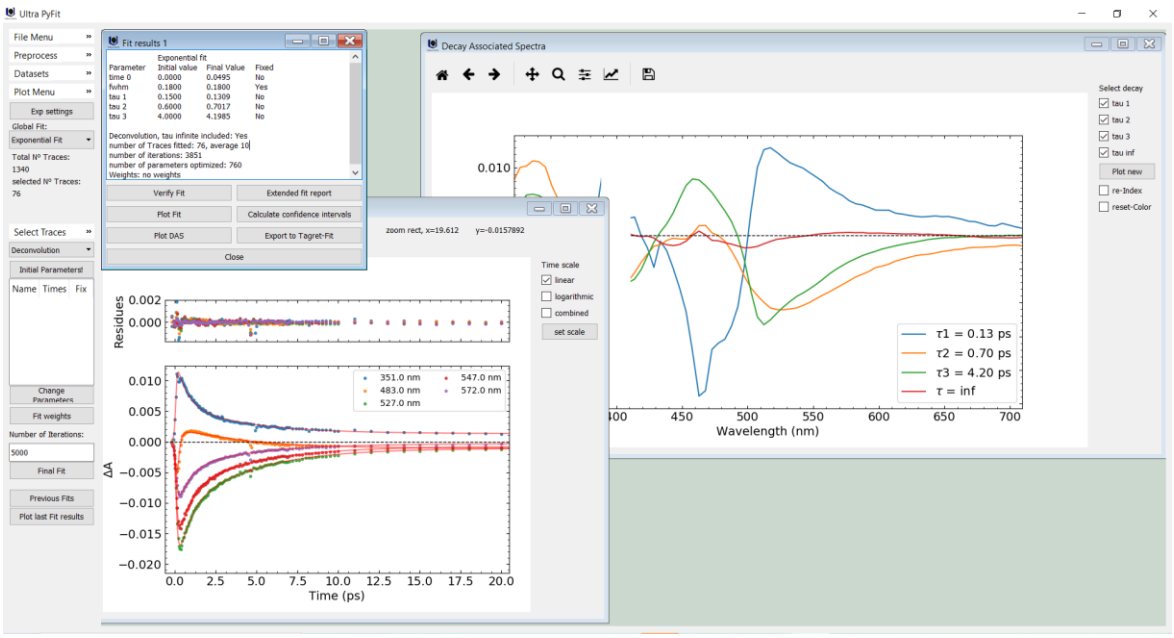


Figure A2.8. Fitting report window with details of the fit together with the DAS spectra and the general plot-fit.

Global fitting to left singular vectors:

As previously explained, one of the best options to estimate the number of parameters in a sample is an SVD decomposition. Indeed, an SVD decomposition consists on taking the original data matrix (M) and decomposing in three simpler matrices U, S and V (Equation A2.3).

$$M = U \times S \times V^T \qquad \text{Equation A2.3}$$

Considering that M is *m* by *n*, U is an order *m* matrix, V is an *n* order matrix, and S is an *m* by *n* diagonal matrix; $^{\mathbf{T}}$ indicates matrix transposition. Where the non-nulls diagonal elements of S matrix are the singular values of M matrix, and thus if S has a number (*k*) of non-null diagonal elements M has rank *k*. Considering the Figure A2.5 middle panel where the singular values of the data are plotted and taking into account that the SV are in a descendent order, we can clearly observe three non-null elements. We could also consider the four values to be a non-null element. Nevertheless, due to inevitable experimental noisy data, the four elements may be already representing the noise. After an SVD decomposition, the non-null left singular vectors contain all the temporal evolution of the data. Therefore, they can be globally fitted to obtain the time constants in a dataset with the greatest certainty. On the other hand, when the resulting pre-exponential factors which are in a global fit associated to the spectra of the transient species, in this case, do not necessarily have a physical meaning. On the contrary, the time constants retrieved from the fit can be used as constant values to independently fit all the traces separately. The pre-exponential factors obtained from this last fit do have the same physical meaning as the DAS.

Ultra Pyfit allows to fit the left singular vectors and can be selected directly from select traces submenu. Once fitted automatically, every single trace is fitted with the times found from the global fit of the left singular vectors (Scheme A2.1).

Importantly the fit window has two slight variations compared to the traces fit. The first important thing is that the fit plot is that corresponding to the left-SV fit. On the contrary, the verify fit, which can be alternated with the verify spectra button and the DAS button, corresponds to the separated traces fitted done with the time constant retrieved from singular vectors fit fixed. The results of a fit to left SV can be seen in Figure A2.9. The data set used

is the same as the one globally fitted previously (Figure A2.8). It is important to note that the fit-figure correspond to the three selected left SV globally fitted while the DAS figure corresponds to the pre-exponential factors of the single fits done to each of the traces fixing the time values to those found in the global fit of the left SV.

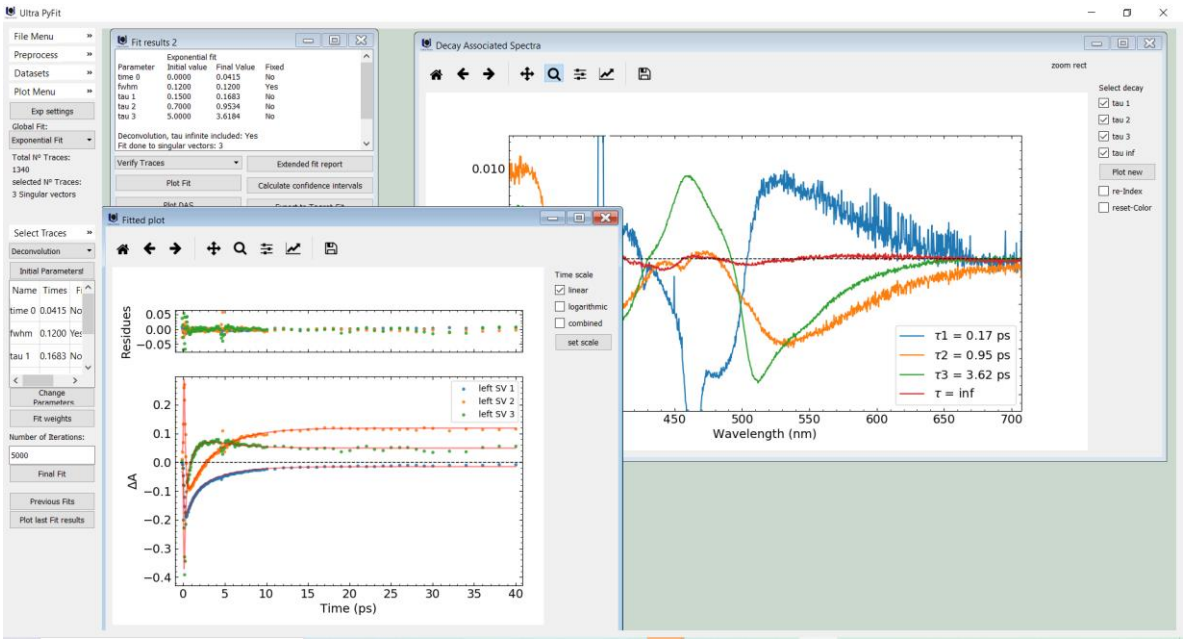


Figure A2.8. Fitting report window with details of the fit together and the general plot-fit and the DAS spectra obtained by fitting each of the traces fixing the time values to those found in the global fit of the left SV.

A2.5b Fitting data to a model.

In case the experimentalist has some knowledge about possible kinetic models that may occur in the system, it is possible to try to describe the data in a more complex way than the assumption done above where we estimate that all decay components have a parallel evolution independent from each other. For example, it is possible to have a model where the components can evolve one from the other (sequential model) or more complex models. This kind of approach is also known as target fit. An important assumption done in the possible kinetic models that can t with Ultra PyFit is that all components follow first-order kinetics in which their life-time is independent of the rest of components in the sample. Therefore, the

concentrations of the different components are described by linear differential equations. The objective of a target fit is to describe the concentrations of the components over time. Even when a certain model perfectly describes the data, it doesn't necessarily mean that the data follows this scheme as other models may also describe perfectly the same data, in such a case we can consider the system is structurally unidentifiable.

The kinetics and evolutions of the initial components into the next ones are described by their life-time and the kinetic rates. In other words, a component may have several kinetic rates but can only have one decay component. In this case, the splitting factor between the two kinetic rates can be given by an external parameter which can be the photo-reaction quantum yield. The ensemble of rate constants forms the so-called "the transfer K matrix". The diagonal elements of this matrix contain the total decay rates of each component and are related to the decay time constant of its Equation A2.4. The non-diagonal elements correspond to the rate constant that indicates transformation between components. Thus, if all non-diagonal components are cero, the model is equivalent to a parallel decay.

$$Tau_i = 1/k_{ii}$$ Equation A2.3

To build a particular model, all steps described above should be done to obtain the final k matrix, together with the initial concentrations of all the components. This is not a straightforward step and can be very tricky to build a K matrix that fulfilled all the above-described steps. Therefore, in the software, a specific builder model window has been created to build and design models in a graphical way from where the program directly builds the K matrix and from there transferred to the global fit class to initialize the fitting parameters. An example of such a window can be seen in Figure A2.9 where a sequential model has been designed. The window has been designed in a way that elements in it can be dragged and moved to design a model, and each of the components (boxes) and rate constants (arrows) can be modified by double-clicking. In Figure A2.9, the specifications for the second component are visible. As we have just seen, modifying a decay time of a component will directly affect the rate constant, and thus these are automatically actualized. Equally, if a rate constant is set, the decay time is automatically set. Once the model is set, the program directly verifies the consistency of it and allows to set the parameters or raises an error message with

the inconsistency. We recommend to set the decay times (tau) of the components and, if needed, the splitting factors and let the program set the rate constants values (k).

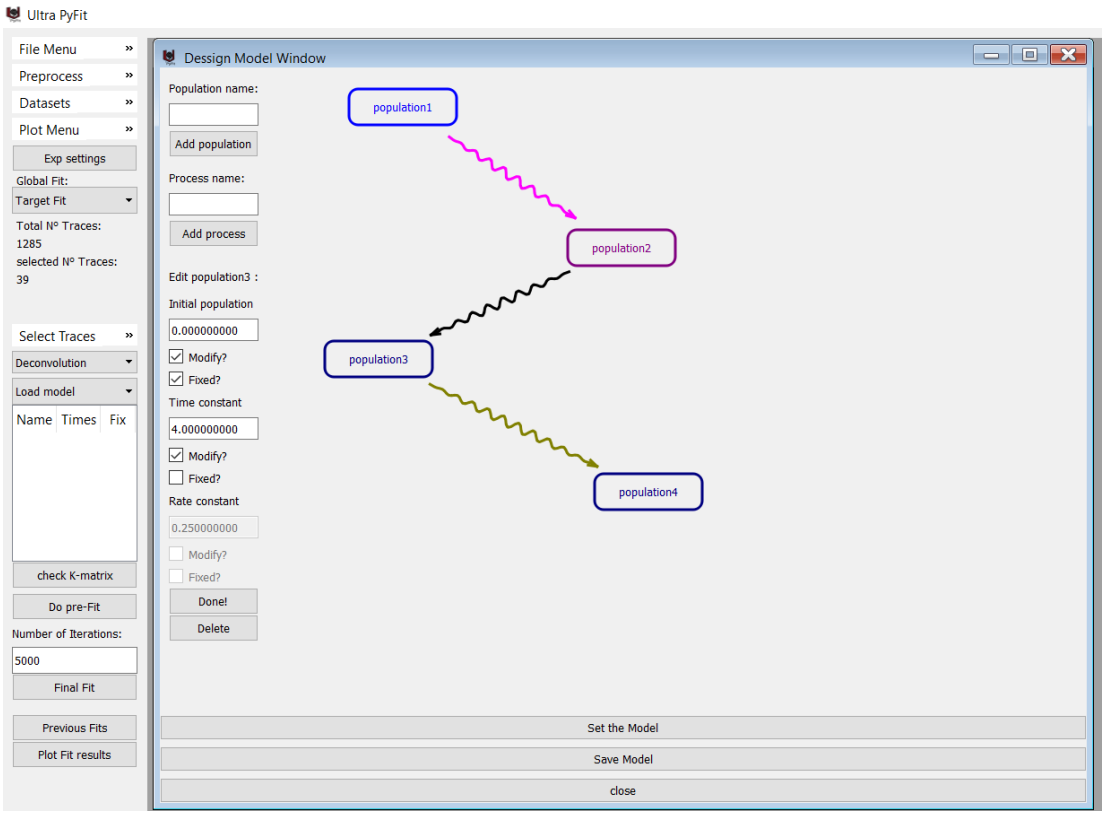


Figure A2.9. Ultra PyFit builder model window example where a sequential model has been designed. In the left side of the window, the menu for specifically indicating the characteristic of population 2 component can be seen.

Once finished, the model can be saved for future uses in other data sets. Loading an already built model is done directly by selecting the load model option in the model combo-box. Once the model is set, the K-matrix can be checked, and ultimately, the parameters can be modified, although the program will not verify this modification. Therefore, it is recommended to modify the existing model directly from the model window rather than from the K matrix all parameters window.

As mentioned above, the data can be described by a number of components, where a spectrum at a certain delay time corresponds to the sum of all components and their concentrations. The evolution of the concentrations of the different components is described by linear differential equations. The objective is to describe the concentrations of the component over

time. The total concentration is the sum of all components concentration where each of them is described by Equation A2.4**.**

$$\frac{dC_i}{dt} = K_{ii}C_i(t) + \sum K_{ij}C_j(t) \qquad \text{Equation A2.4.}$$

Where Kii represent the rate constant (diagonal element of K matrix), and Kij represent possible inputs of other components in the sample that evolve to this component (non-diagonal K matrix elements). For the evaluation of the exponential of a non-diagonal Kmatrix, the program decomposes the K matrix in igenvector and eigenvalues using the scipy.linalg.eig() function. With the eigen vector matrix and the initial concentrations, the relative concentrations are solved by the linear equation set (eigenvectors*coeficients = Intial_Concentrations) to obtain the relative proportion of each component at a certain delay time. To solve this linear equation, the program uses `scipy.linalg.solve`() function. Once these coefficients are obtained, the concentration of each component at a certain delay can be obtained from an exponential decay and the corresponding eigenvalue (Equation A2.5)

$$C_i(t) = e^{-k_{ii}t} \qquad \text{Equation A2.5}$$

Finally, the data, as said at the beginning, is the linear combinations of each concentration and their spectral shape Equation A2.6. The difference between the data and the model is minimized to obtain the parameters that best fit the model.

$$\psi(t) = \sum C_i(t) * Spec_i \qquad \text{Equation A2.6}$$

In Ultra PyFit, either the left singular vectors or an ensemble of traces can be globally fitted with any custom-designed model. Once the fit is finished, similarly to the exponential global fit, a window fit will pop out. This window is identical to that of the exponential global fit, except for the fact that, instead of the DAS, there is the option to plot the evolving associated spectra (EAS), which correspond to the *Spectra* of each component (*spec* in equation A2.6) and their concentrations profile. Finally, instead of the Export to target fit button, there is the plot model button which retrieves the positions of the designed element that the user has designed in the builder model window and generates a matplotlib figure that can be saved. An example of the final fit window can be seen in Figure A2.10.

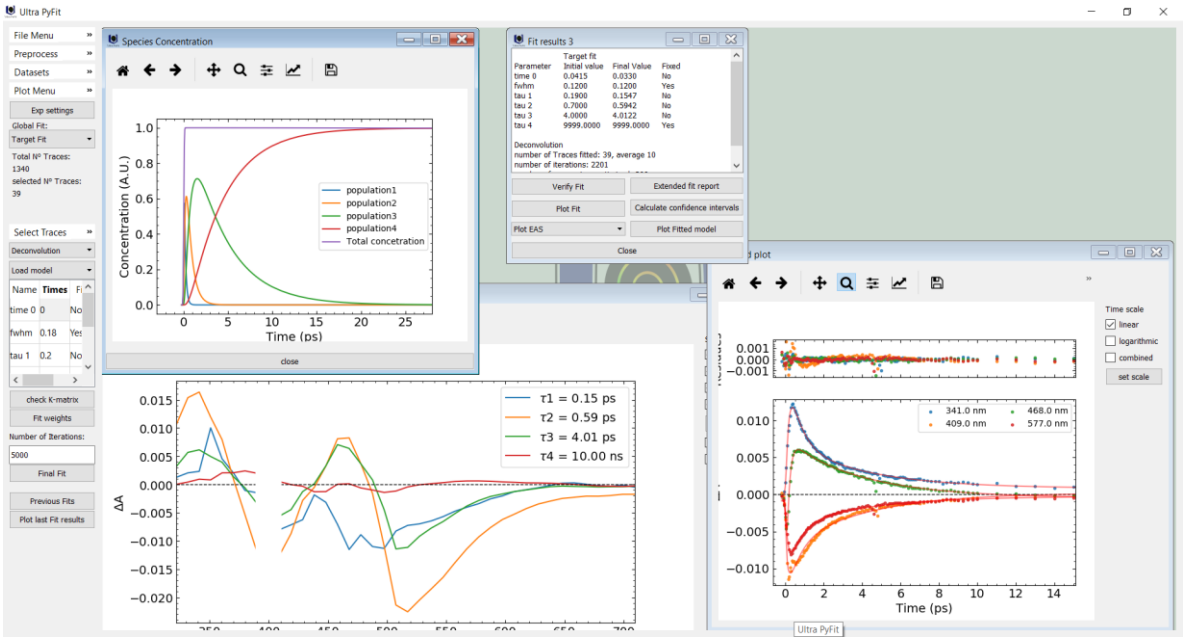


Figure A2.10. Fitting report window with details of the fit, together with windows containing the EAS spectra, and the a selection of traces with the corresponding fit and residues (sequential). The results were obtained in the same data set used previously.

A2.6 Analysis of the error in Ultra PyFit.

The model ambiguity is a known issue in multi-exponential fitting, except for mono-exponential models (the most know ambiguity is the slow-fast ambiguity of the parameters in two exponential fittings)[21]. Moreover, one of the problems with non-linear minimization, particularly with increasing model complexity, is the possibility of reaching a local minimum of the error function without finding the global minimum, in other words, there might be other parameters for the model that could describe better the data. The best way to determine this id to repeat the process using a different set of initial model parameter guesses and determine whether an equivalent set of best-fit parameters is obtained. The fact that different sets of model parameters may results in precisely the same quality fit (in terms of lack of fit). This is inherent to the method itself. In general, a correlation coefficient (in our case $\chi 2$) is used to examine "goodness of fit" between the model and data[18]. However, a high correlation coefficient does not implicitly imply that the data are correctly described by the model as can be seen in the Anscombe's quarter[22] in chapter 2. One of the best ways to evaluate a model

is the residual plot. A good model should be able to describe the different trend of the data. Therefore the residues of the fit should be similar to random noise and uniformly spaced along the abscissa and around zero ordinates with no trend[18,22].

Once it is clear that the model correctly describes the data, the next step is to evaluate the uncertainty on each of the estimated parameters that best fit the data. The simple approach to estimating uncertainties and correlations is by inverting the second derivative matrix assuming that the components of the residual array are distributed around 0 with a normal Gaussian distribution. The error translates the uncertainty in the estimation of each parameter independently of the others (only the diagonal terms of the second-order derivative error matrix are taken into consideration no interaction term). This means that a 1D trace of the multidimensional qui-squared surface is considered for each parameter and we assume that the residues are well distributed around 0 with a normal (Gaussian distribution). It results that an ambiguous parameter may be estimated with a very low error. The estimated error of a parameter is related to the minimum of the multidimensional qui-squared surface. This simple approach to assessing uncertainties ignores outliers, highly asymmetric uncertainties, or complex correlations between Parameters. (e.g. the error fit of the estimated *Off* to *On* quantum yields are around 10^{-4} which not necessarily implies that the real parameter have this error of the same order or that the values is correctly estimated)

Therefore, for determining the error in the estimated model parameter values, more advance methods are needed. The best way to do this would be to fit the model to multiple data sets of the same sample[18]. If all data sets are collected similarly, they should only differ on random variability such as random noise. Finally, the variation in the individual model parameters will give the confidence interval of each of the parameters. In order to do this, it is necessary to acquire a sufficiently large number of experimental datasets of the same sample under same conditions to have a large representative number of replicates between 500-1000 samples to be able to estimate the uncertainties of the parameters accurately; as the reader might expect, this, in practice, is impossible. Therefore, the best approaches is the simulation of synthetic data sets using known parameter values, but also including noise representative of the actual measurement noise[18]. This approach of estimating the error in the parameter values is known as inverse modelling. The simulated data sets can be generating using the

inverse model itself or by using more complex forward models that represent the experimental system. In either case, it is crucial to include noise representative of the actual measurement noise. There are two different main approaches to do this: Monte Carlo simulations, from where the noise can be simulated following a known distribution (i.e. Normal Gaussian distribution for random noise) and selecting for each point of the data set values extracted from this distribution.

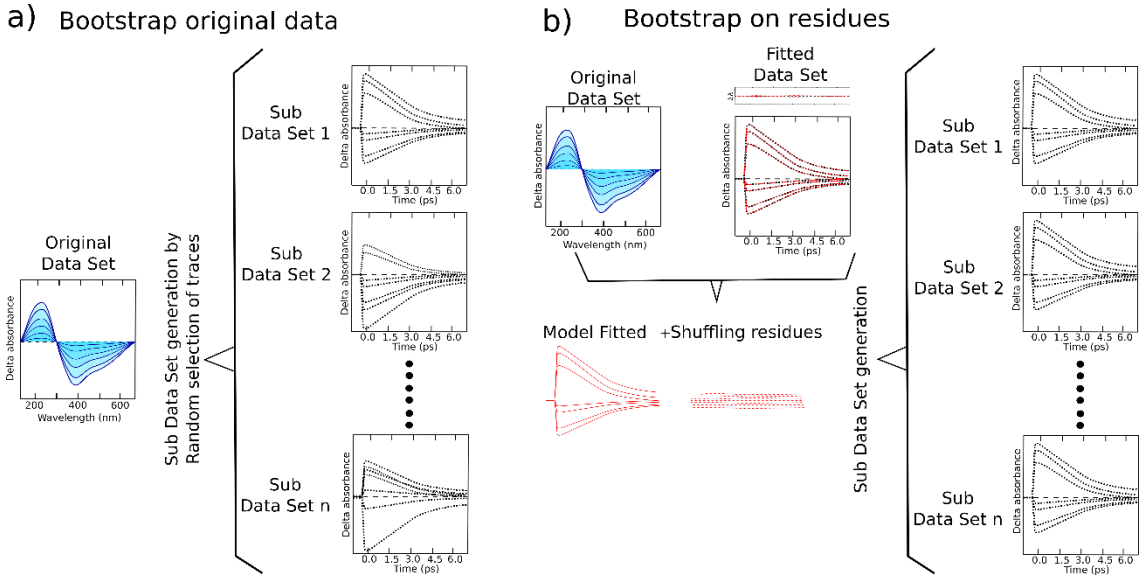


Figure A2.11. Two bootstrap approaches for generating a number n of data sets as replicates for determining the error in the estimated fitted model parameters. a) Bootstrap on the data. b) Bootstrap on the residues. Figure done with Inkscape.

The second option is the bootstrap technique which is based on generating data by shuffling and substituting data points from the original data set. In the case of a global fit where only a subset of traces has been fitted, two types of bootstrap can be imagined. The first is directly to the original data set, as generally from the dataset, not all traces are fitted but only a sub-set of them. Therefore, no assumptions are made at all to generate the synthetic data sets as they are indeed directly taken from the original complete set of data. The second option is to assume that the fitted model correctly describes the data. In such a case, a synthetic data set can be generated by "inverse model" (using the retrieved parameters) and shuffling and substituting the residues to obtain actual noise representative of the measurement. This approach is known as residues bootstrap and we have previously applied it in time-resolved

spectroscopy data (nano second flashphotolysis data presented in chapter 5)[23]. The two approaches are represented in Figure A2.11.

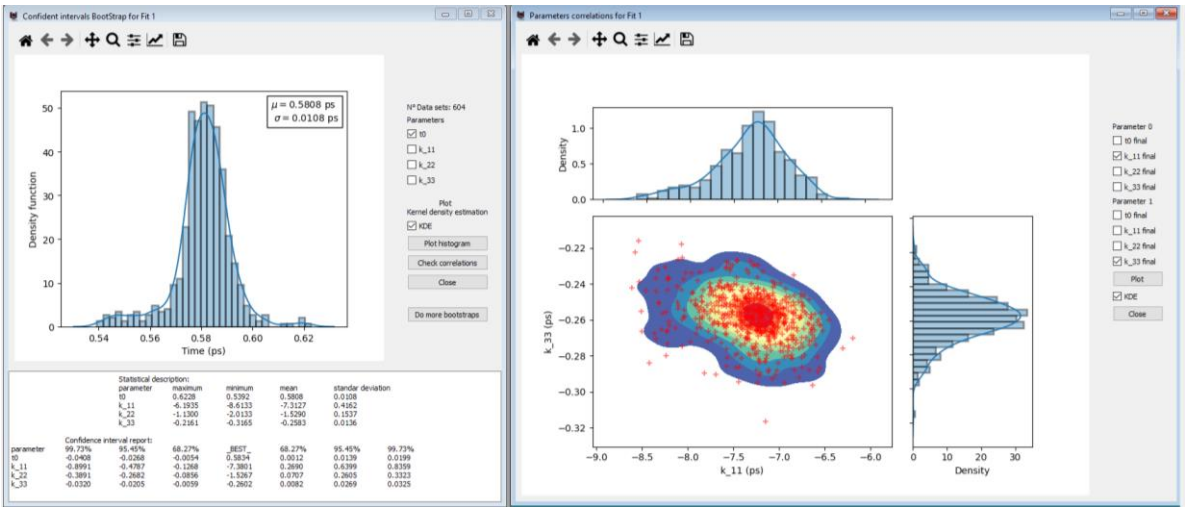


Figure A2.12. Fitting report window with details of confidence intervals studied using a residues bootstrap, where the histogram for t0 and the correlation plot for K_11 and K_33 can be seen.

Although both of the options have been coded in the GlobalFit python class, only the residues bootstrap are so far available in the Ultra PyFit GUI. After performing a global fit via an exponential or a target fit, the button confidence interval calculations offers the option to directly generate the desired number of data sets, after selecting the percentage of residue traces that will be shuffled. The results of a bootstrap can be seen once the calculations are finished. The result will be displayed as histograms for all parameters that have been globally fitted with a statistical description of the data including minimum, maximum, mean, standard deviation and an estimation of the 99.8 95.5% and 68% of the confidence intervals in both directions for all globally fitted parameters. Finally, the correlations between parameters can be studied via a triple plot figure where the middle figure the scatter plot of the two variables that want to be studied are represented together with a kernel density estimation, and in each of the sides the histograms of each variable can be seen. The bootstrap analysis results can be seen in Figure A2.12 for the fitted performed in section A2.5b and 600 replicates.

Alternatively, the confidence intervals calculation can be done performing an F-test. This is the best option offered by lmfit python package[13] calculating confidence intervals, from

where the code has been adapted to our specific case. The F-test is particularly useful in cases where the fit has been done to the left singular vector and thus is not possible to perform a bootstrap to the data or the residues, as there is no option to select other singular vectors and then shuffling, which is necessary to obtain actual noise representative of the measurement to generate new replicates by the bootstrap of the residues.

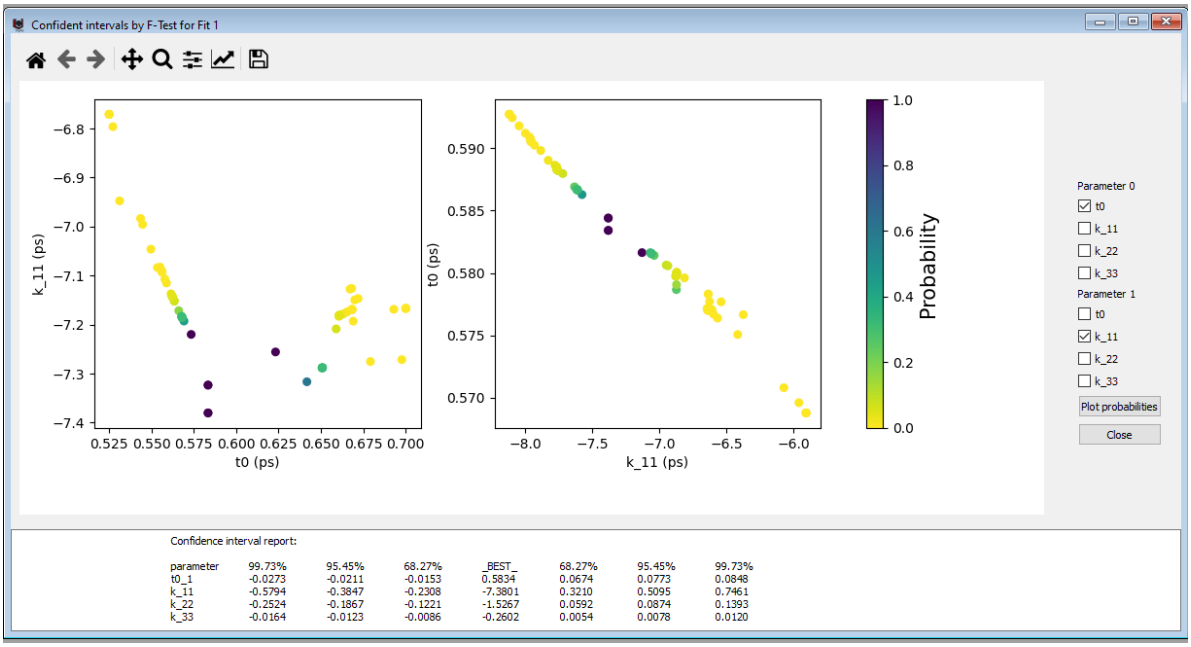


Figure A2.13. Fitting report window with details of confidence intervals studied F-test, where values probabilities for t0 and K_11 can be seen.

For these cases, after finding a set of best-fit parameters of the fitted model, we have obtained the lowest chi^2 possible. The F-test is used to compare fitted models, which is (best fit found), with an alternate model, where one of the parameters is fixed to a specific value (Equation A2.7[13]). This alternative fit will always be worse and thus is up in the error surface (greater chi^2), which directly translates on an increased values F for comparing those two models. This is repeated iteratively until the confidence level desired for the parameter is reached. In other words, the value is changed until the difference between $\chi 20$ and $\chi 2f$ can't be explained by the loss of a degree of freedom within a certain confidence. This method calculates for each variable parameter the probabilities for the corresponding cumulative variables. This can be used to show the dependence between two parameters. An example is

shown in Figure A2.13, were also an estimation of the 99.8, 95.5% and 68% of the confidence intervals in both directions for all globally fitted parameters can be seen.

$$F(P_{fix}, N-P) = \left(\frac{\chi^2 f}{\chi^2 0} - 1\right)\frac{N-P}{P_{fix}} \qquad \text{Equation A2.7}$$

Both methods are alternatives ways of estimating a more realistic value of the uncertainties and correlations of the parameters that best fit the data than those obtained by inversion of the second derivative matrix and assuming that the components of the residual array are distributed around 0 with a normal Gaussian distribution.

A2.7 Keeping track of actions.

At all moments, the analysis done can be saved, with the use of pickle python package, it is possible to generate a python object of the GlobalFit instance. This can be reloaded again in another computer with the program installed or later on for further analysis. Not only every fit that has been done can be re-plotted and checked effortlessly by clicking on the previous fit button, but also every action and steps done to reach to the fitted model can be seen. Indeed, the program keeps track of all preprocessing actions and details of these actions, for example, the number of spectra subtracted as baseline, and how the GVD was corrected. Finally, it also keeps track of the data file that was loaded initially. All of this information can be fast accessed by only clicking on display report button, where the user can see all the details of the preprocessing actions, all the results, the different fits performed, and a sequence of actions that have been done in a clear, self-explanatory page. An example of the report of the experiment is shown in Figure A2.14.

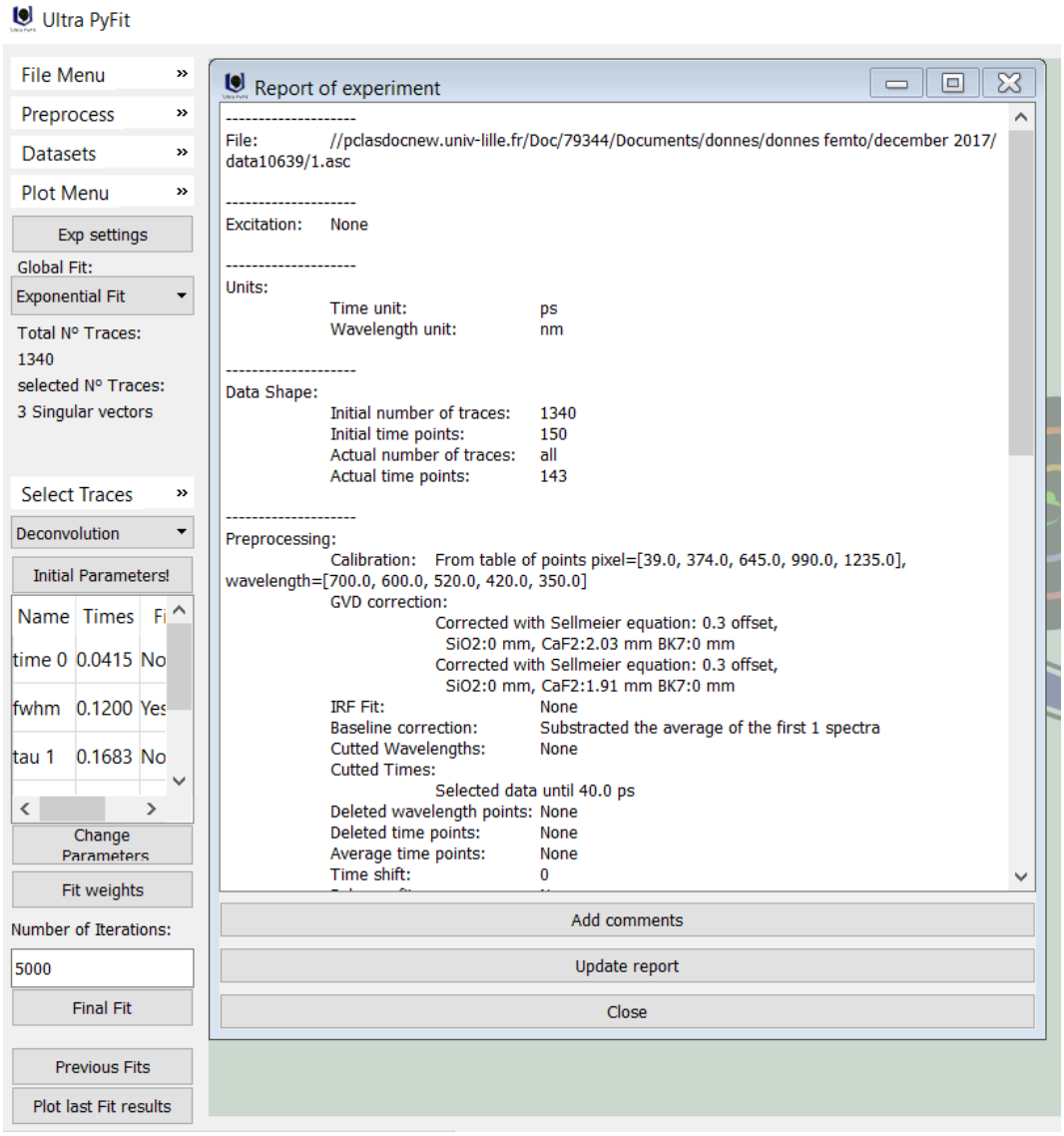


Figure A2.14. Report experiment page where all details of preprocessing results and sequence of actions are clear and summarized in a single page.

At all moments, the analysis done can be saved, with the use of pickle python package, it is possible to generate a python object of the GlobalFit instance. This can be reloaded again in another computer with the program installed or later on for further analysis. Not only every fit that has been done can be re-plotted and checked effortlessly by clicking on the previous fit button, but also every action and steps done to reach to the fitted model can be seen. Indeed, the program keeps track of all preprocessing actions and details of these actions, for example, the number of spectra subtracted as baseline, and how the GVD was corrected. Finally, it also keeps track of the data file that was loaded initially. All of this information

can be fast accessed by only clicking on display report button, where the user can see all the details of the preprocessing actions, all the results, the different fits performed, and a sequence of actions that have been done in a clear, self-explanatory page. An example of the report of the experiment is shown in Figure A2.14.

A2.7 References

(1) Martin, J.; Wang, X.; Simpson, C.; Siekmann, I. PyTrA: ultra-fast transient absorption data analysis software. *IJNT* **2014**, *11*, 601.

(2) Mullen, K. M.; van Stokkum, I. H. M. TIMP : An R Package for Modeling Multi-Way Spectroscopic Measurements. *J. Stat. Soft.* **2007**, *18*.

(3) van Wilderen, L. J. G. W.; Lincoln, C. N.; van Thor, J. J. Modelling multi-pulse population dynamics from ultrafast spectroscopy. *PloS one* **2011**, *6*, e17373.

(4) Slavov, C.; Hartmann, H.; Wachtveitl, J. Implementation and evaluation of data analysis strategies for time-resolved optical spectroscopy. *Analytical chemistry* **2015**, *87*, 2328–2336.

(5) Ruckebusch, C.; Sliwa, M.; Pernot, P.; Juan, A. de; Tauler, R. Comprehensive data analysis of femtosecond transient absorption spectra: A review. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* **2012**, *13*, 1–27.

(6) Snellenburg, J. J.; Laptenok, S. P.; Seger, R.; Mullen, K. M.; van Stokkum, I. H. M. Glotaran : A Java -Based Graphical User Interface for the R Package TIMP. *J. Stat. Soft.* **2012**, *49*.

(7) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95.

(8) Waskom, M.; Botvinnik, O.; O'Kane, D.; Hobson, P.; Lukauskas, S.; Gemperline, D. C.; Augspurger, T.; Halchenko, Y.; Cole, J. B.; Warmenhoven, J.*; et al. Mwaskom/Seaborn: V0.8.1 (September 2017)*; Zenodo, 2017.

(9) Golub, G. H.; van Loan, C. F. *Matrix computations*, 3rd ed.; Johns Hopkins studies in the mathematical sciences; Johns Hopkins University Press: Baltimore, 1996.

(10) Lehoucq, R. B.; Sorensen, D. C.; Yang, C. *ARPACK users' guide: Solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*; Software, environments, tools; SIAM: Philadelphia, 1998.

(11) Brillouin, L. *Wave propagation and group velocity*, 7. pr; Pure and applied physics 8; Acad. Pr: San Diego, 1988.

(12) Sellmeier. Zur Erklärung der abnormen Farbenfolge im Spectrum einiger Substanzen. *Ann. Phys. Chem.* **1871**, *219*, 272–282.

(13) Newville, M.; Stensitzki, T.; Allen, D. B.; Ingargiola, A. *LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python*; Zenodo, 2014.

(14) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.*; et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods* **2020**, *17*, 261–272.

(15) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.*; et al.* Array programming with NumPy. *Nature* **2020**, *585*, 357–362.
(16) van Stokkum, I. H. M.; Larsen, D. S.; van Grondelle, R. Global and target analysis of time-resolved spectra. *Biochimica et biophysica acta* **2004**, *1657*, 82–104.
(17) Mouton, N.; Sliwa, M.; Buntinx, G.; Ruckebusch, C. Deconvolution of femtosecond time-resolved spectroscopy data in multivariate curve resolution. Application to the characterization of ultrafast photo-induced intramolecular proton transfer. *J. Chemometrics* **2010**, *24*, 424–433.
(18) Costa, K. D.; Kleinstein, S. H.; Hershberg, U. Biomedical model fitting and error analysis. *Science signaling* **2011**, *4*, tr9.
(19) Beechem, J. M.; Ameloot, M.; Brand, L. Global and Target Analysis of Complex Decay Phenomena. *Instrumentation Science & Technology* **1985**, *14*, 379–402.
(20) Meuwis, K.; Depuydt, G.; Boens, N.; C. De Schryver, F. Comparison of simultaneous biexponential and compartmental analyses of fluorescence decay surfaces of intermolecular two-state excited-state processes. *Chemical Physics Letters* **1995**, *246*, 641–648.
(21) Istratov, A. A.; Vyvenko, O. F. Exponential analysis in physical phenomena. *Review of Scientific Instruments* **1999**, *70*, 1233–1257.
(22) Anscombe, F. J. Graphs in Statistical Analysis. *The American Statistician* **1973**, *27*, 17–21.
(23) Woodhouse, J.; Nass Kovacs, G.; Coquelle, N.; Uriarte, L. M.; Adam, V.; Barends, T. R. M.; Byrdin, M.; La Mora, E. de; Bruce Doak, R.; Feliks, M.*; et al.* Photoswitching mechanism of a fluorescent protein revealed by time-resolved crystallography and transient absorption spectroscopy. *Nature communications* **2020**, *11*, 741.

Appendix 3 X-ray structures of rsEGFP2 and Dronpa

X-ray structures of rsEGFP2.

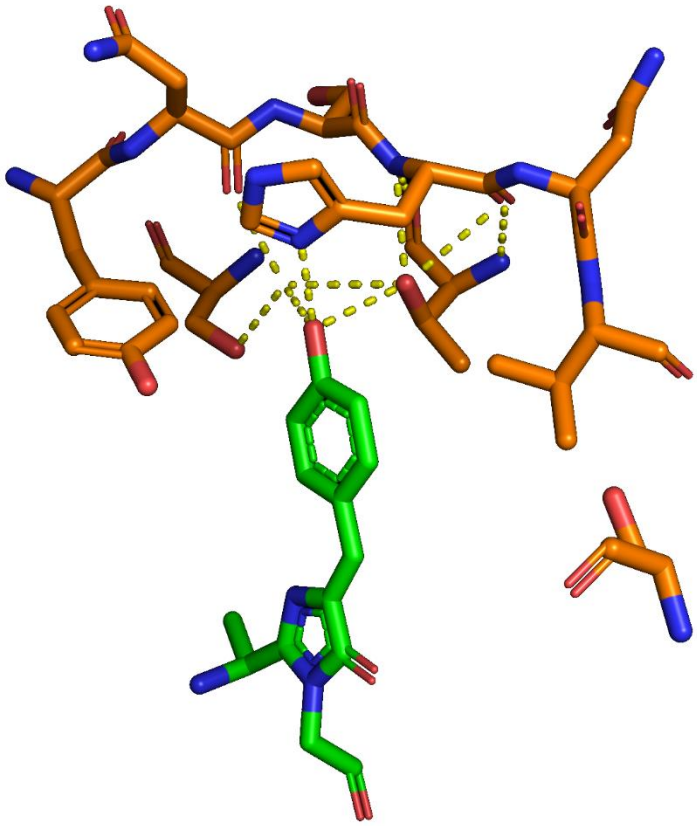


*Figure A3.1. X-ray structure of rsEGFP2 cis anionic On state. The figure shows the main hydrogen bond between the protein cage and the HBDI phenol group. Reproduced from PDB **5DTY**[1] using pymol [2].*

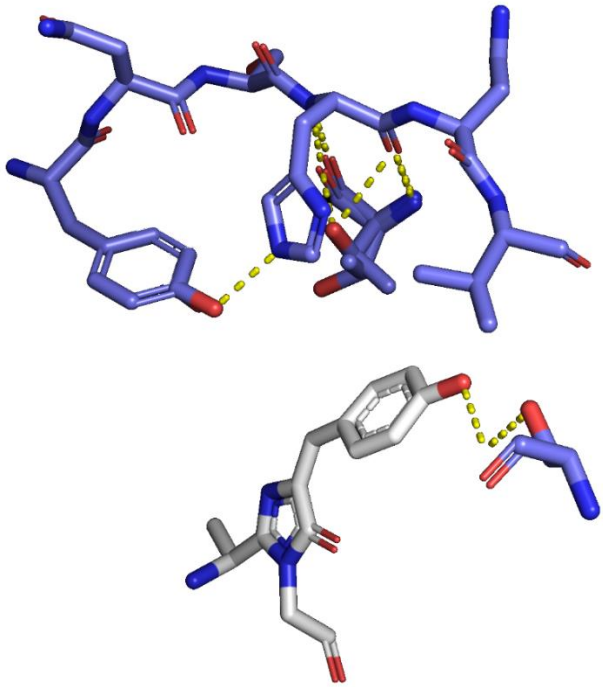


Figure A3.2. X-ray structure of rsEGFP2 trans neutral Off state. The figure shows the main hydrogen bond between the protein cage and the HBDI phenol group. Reproduced from PDB ***5DTX***[1] *using pymol* [2].

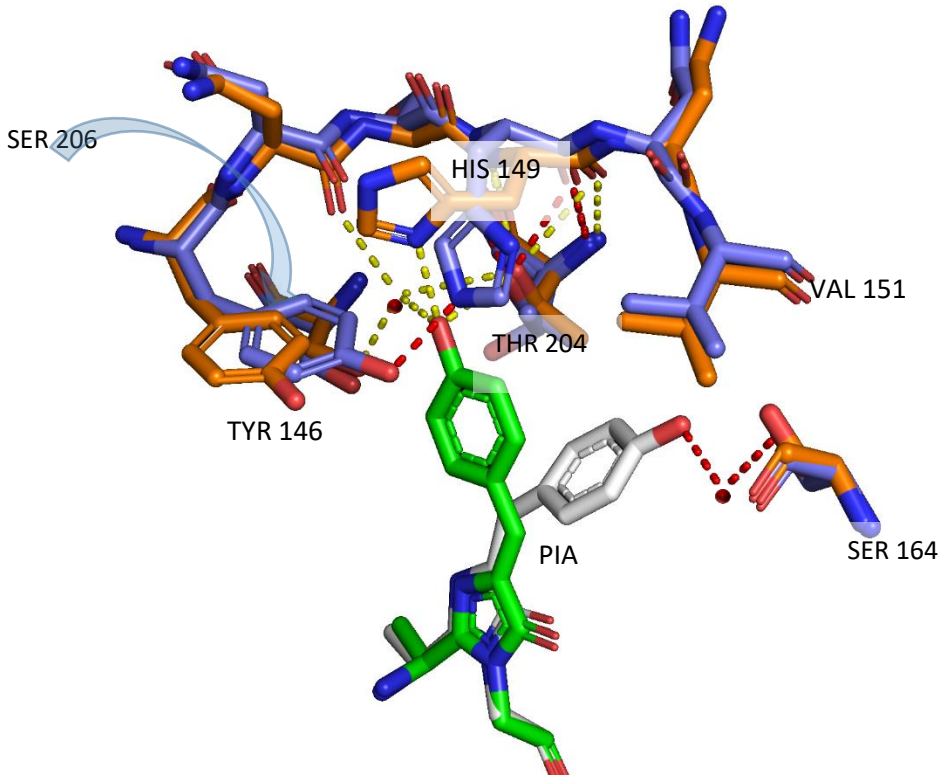


Figure A3.3. Overlaid of figures A3.1 and A3.2. the differences in between On and Off states can be observed.

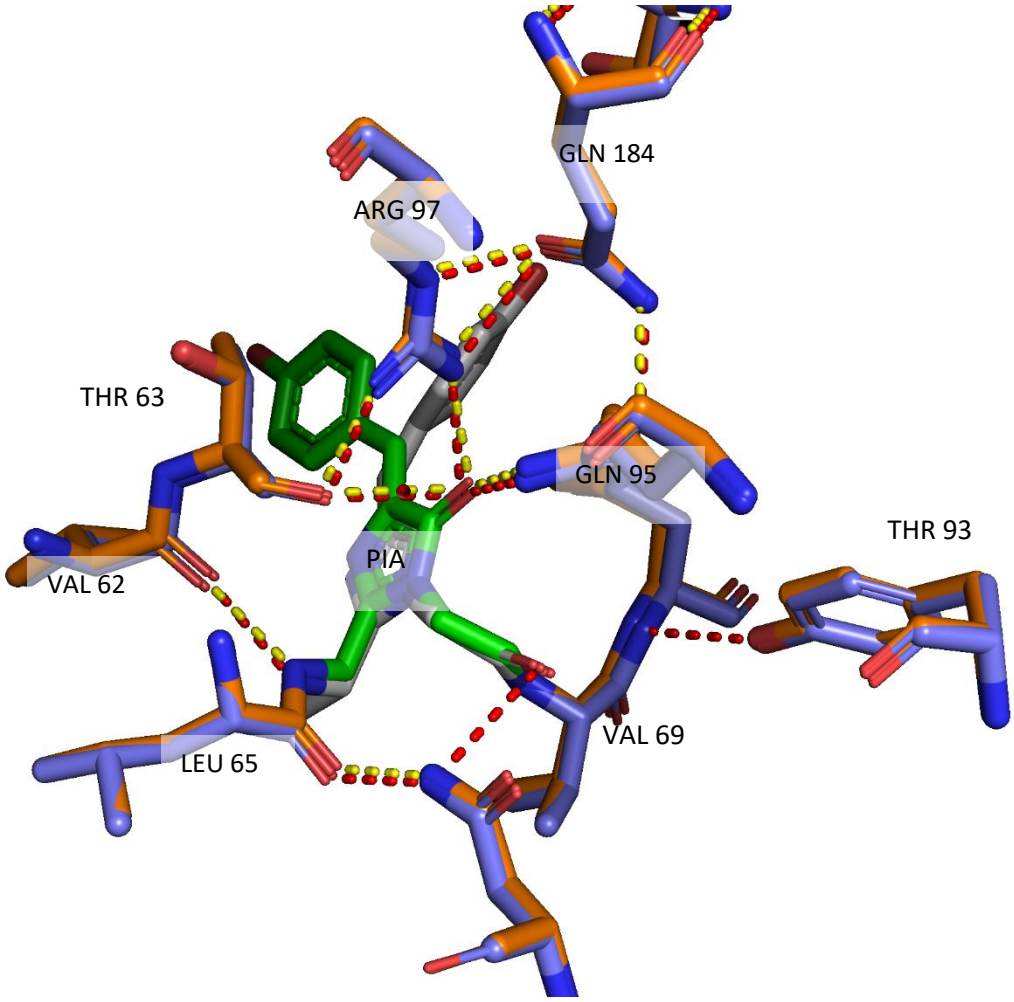


Figure A3.4. X-ray structure of rsEGFP2 trans neutral Off state in blue and cis anionic On state in orange. The figure shows the main hydrogen bond between the protein cage and the HBDI imidazolinone group. No difference can be observed. Reproduced from PDB ***5DTY***[1] *and* ***5DTX***[1] *using pymol* [2].

X-ray structures of Dronpa.

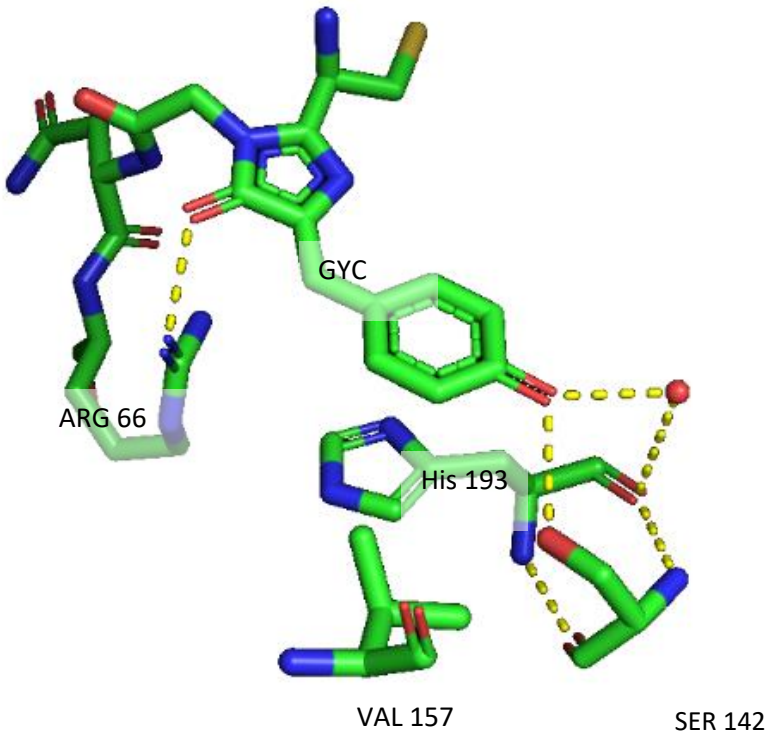


Figure A3.5. X-ray structure of Dronpa cis anionic On state. The figure shows the main hydrogen bond between the protein cage and the HBDI imidazolinone and phenol groups. Reproduced from PDB 2POX[3] using pymol [2].

μ

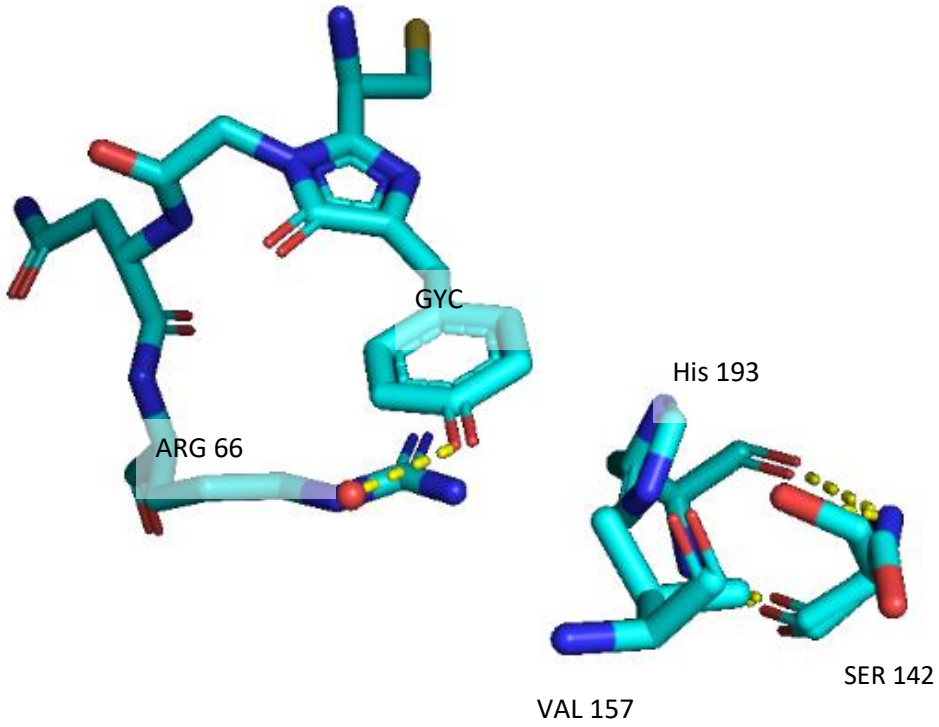
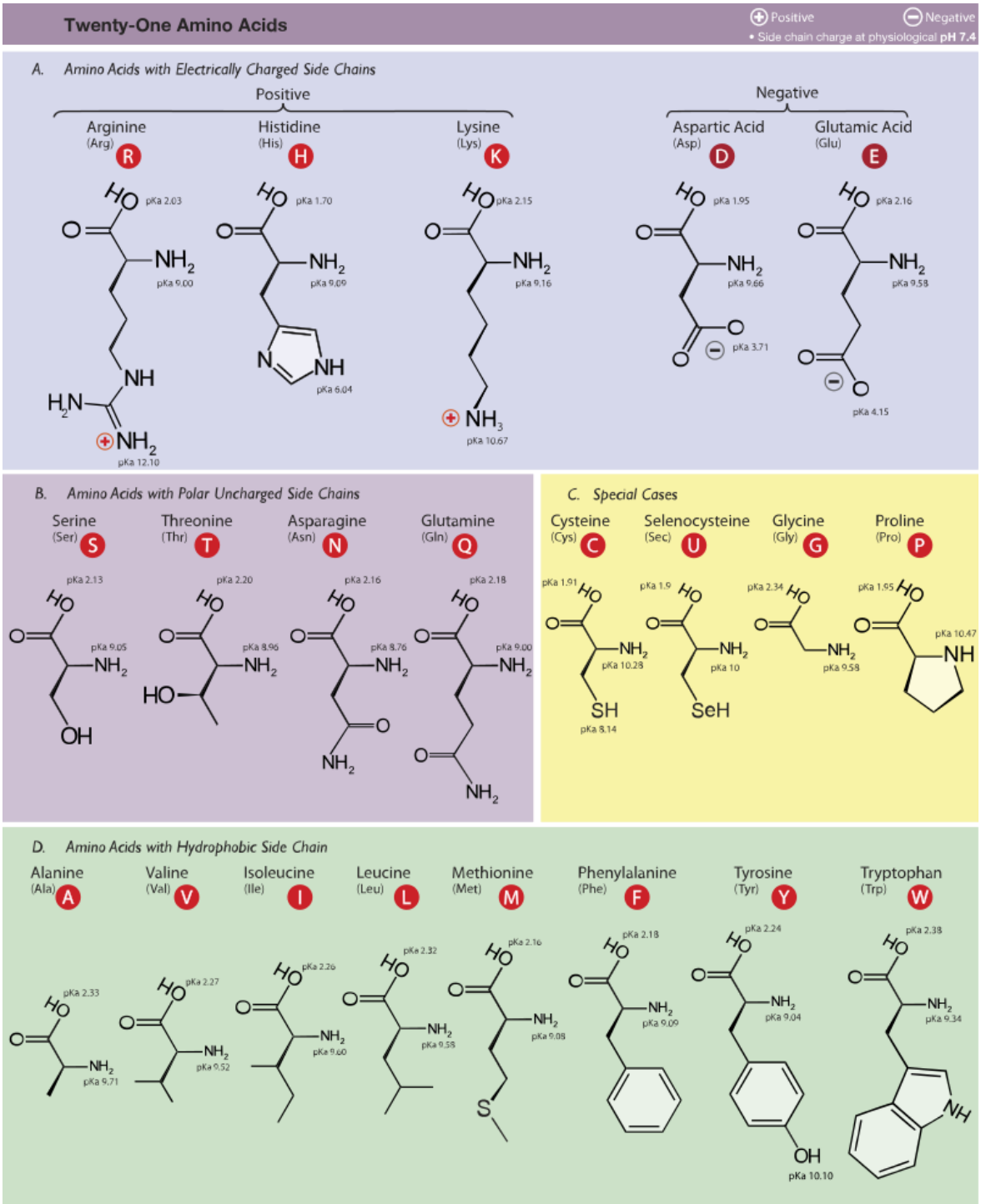


Figure A3.6. X-ray structure of Dronpa trans neutral Off state. The figure shows the main hydrogen bond between the protein cage and the HBDI imidazolinone and phenol groups. Reproduced from PDB 2IOV[3] using pymol [2].

Table of amino acids

Table A3.1 Table representing the twenty-one main amino acids. Reproduced from https://en.wikipedia.org/wiki/**Amino**_acid.



References

(1) 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. *Scientific reports* **2016**, *6*, 18459.

(2) The PyMOL Molecular Graphics System. 1.5.0.4 edn (Schrödinger, LLC). https://pymol.org/2/support.html?#citing.

(3) 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. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 13005–13009.

Appendix 4. List of scientific contributions

A4.1 Publications

A4.1.1-Publications PhD main research topic

1. J. Woodhouse, G. Nass-kovacs#, N. Coquelle#, **L. M. Uriarte#**, V. Adam#, T. R. M. Barends, M. Byrdin, E. de la Mora, R. B. Doak, M. Feliks, M. Field, F. Fieschi, V. Guillon, S. Jakobs, Y. Joti, P. Macheboeuf, K. Motomura, K. Nass, S. Owada, C. M. Roome, C. Ruckebusch, G. Schirò, R. Shoeman, M. Thepaut, T. Togashi, K. Tono, M. Yabashi, M. Cammarata, L. Foucar, D. Bourgeois, M. Sliwa, J-P. Colletier, I. Schlichting, M. Weik. *Photoswitching mechanism of a fluorescent protein revealed by time-resolved serial femtosecond crystallography and transient absorption spectroscopy*. **Nat Commun 2020**, 11**,** 741 #these authors contributed equally to this work.

Under preparation:

2. **Lucas M. Uriarte,** Raffaele Vitale, Stanislaw Nizinski, Andras Lukacs, Steve Meech, Michel Sliwa, Cyril Ruckebusch. *Undirected baseline correction on in time-resolved IR spectroscopy: example with the photo-dynamics off rsegfp2and transient.*
3. **Lucas M. Uriarte,** Kyprianos Hadjidemetriou, Stanislaw Nizinski, Olivier Devos, Cyril Ruckebusch, Jacques-Philippe Colletier, Martin Weik and Michel Sliwa. *Trans to cis isomerization photo-dynamics in rsEGFP2 starting from OBF or HT conformer.*
4. Kyprianos Hadjidemetriou, N. Coquelle#, **L. M. Uriarte#**, V. Adam#, T. R. M. Barends, M. Byrdin, E. de la Mora, R. B. Doak, M. Feliks, M. Field, F. Fieschi, V. Guillon, S. Jakobs, Y. Joti, P. Macheboeuf, K. Motomura, K. Nass, S. Owada, C. M. Roome, C. Ruckebusch, G. Schirò, R. Shoeman, M. Thepaut, T. Togashi, K. Tono, M. Yabashi, M. Cammarata, L. Foucar, D. Bourgeois, M. Sliwa, J-P. Colletier, I. Schlichting, M. Weik *Rational control of off-state heterogeneity in a photoswitchable fluorescent protein provides switching contrast enhancement.*

A4.1.2-Publications PhD not main research topic

1. T. Nagasaka; H. Sotome, S. Morikawa, L. M. Uriarte, M. Sliwa, T. Kawai; H. Miyasaka. Restriction of the conrotatory motion in photo-induced 6π electrocyclic reaction: formation of the excited state of the closed-ring isomer in the cyclization RSC Adv., 2020,10, 20038-20045
2. M. Jacquet, **L. M. Uriarte**, F. Lafolet, M. Boggio-Pasqua, M. Sliwa, F. Loiseau, E. Saint-Aman, S. Cobo, G Royal The Journal of Physical Chemistry Letters **2020** 11 (7), 2682-2688
3. 3- A. Tokunaga; L. M. Uriarte; K. Mutoh; E Fron; J. Hofkens; M.Sliwa J. Abe Photochromic Reaction by Red Light via Triplet Fusion Upconversion. Journal of American Chemical Society 2019, 141, 44.

Under preparation:

4. Stanislaw Nizinski, Adjéle Wilson, Lucas Uriarte, Cyril Ruckebusch, Elena Andreeva, Ilme Schilichting, Jacques-Philippe Colletier, Diana Kirilovsky*, Gotard Burdzinski*, Michel Sliwa*, *Balance between picosecond lifetime intermediate states controls the primary photo-activation quantum yield of the Orange Carotenoid Protein from Synechocystis.*

A4.2 Contribution to congress being presenter

A4.2.1 Oral presentations:

1. **Lucas M. Uriarte**, Kyprianos Hadjidemetriou, Olivier Devos, Tadeo Moreno-Chicano, Jacques-Philippe Colletier, Martin Weik, and Michel Sliwa. Influence of the environment on the photo-dynamics of a photo-switchable fluorescent protein: solution vs micro-crystals, SolvATE2020: **Scientific Meeting of the GdR SolvATE 25-26 Nov 2020, (Online).**

International conference

2. **Lucas M. Uriarte**, Kyprianos Hadjidemetriou, Olivier Devos, Cyril Ruckebusch, Jacques-Philippe Colletier, Martin Weik and Michel Sliwa. Trans to cis isomerization photo-dynamics in rsEGFP2 starting from OBF or HT conformer. **5th Nanosynergetics Workshop, 13th Nov. 2020 (Online)**.
(**Best talk award).**

International conference

3. **L. M. Uriarte**, O. Devos, K. Hadjidemetriou, C. Ruckebusch, S. Meech, D. Bourgeois, J.P. Colletier, M. Weik and M. Sliwa. *First study of the photodynamics of an hydrozoan fluorescent photo-switchable protein: existence of different switching mechanism.* **9th International Symposium On Photochromism (Paris, France, 2019)**.

International conference

4. **L. M. Uriarte**, O. Devos, K. Hadjidemetriou, C. Ruckebusch, S. Meech, D. Bourgeois, J.P. Colletier, M. Weik and M. Sliwa. *First study of the photodynamics of an hydrozoan fluorescent photo-switchable protein: existence of different switching mechanism.* **17th Congress of the International Union of Photobiology & 18th Congress of the European Society for Photobiology (Barcelona, Spain, 2019)**.

5. **L. M. Uriarte**, O. Devos, C. Ruckebusch, S. Meech, D. Bourgeois, M. Weik and M. Sliwa. *First study of the photodynamics of an hydrozoan fluorescent photo-switchable protein: existence of different switching mechanism.* **Journées Annuelles de la SP2P (Lille, France, 2019).**

International conference

6. **L. M. Uriarte**, Cyril ruckebusch, Michel sliwa. *photoswitching dynamics of the reversible photoswitchable fluorescence protein rsEGFP2: crystal and solution*. **Workshop for Young Researchers on Photo-active materials with cooperative and Synergetic Responses. (Lille, France, 2018).**

A4.2.2 Posters presentations:

International conference

1. **L. M. Uriarte**, O. Devos, K. Hadjidemetriou, C. Ruckebusch, S. Meech, D. Bourgeois, J.P. Colletier, M. Weik and M. Sliwa. *Study of the photodynamics of a hydrozoan photo-switchable fluorescent protein: existence of different switching mechanisms*. **Let there be..... Light: A symposium on the occasion of the 80th birthday of Prof. Frans De Schryver (Leuven, Belgium 2019).**
(**best poster award)**

International conference

2. **L. M. Uriarte**, O. Devos, K. Hadjidemetriou, C. Ruckebusch, S. Meech, D. Bourgeois, J.P. Colletier, M. Weik and M. Sliwa Study of the photodynamics of a hydrozoan photo-switchable fluorescent protein: existence of different switching mechanisms. **9th International Symposium On Photochromism (Paris, France, 2019).**

3. **L. M. Uriarte,** R. Vitale, C. Ruckbusch, M. sliwa. *Photoswitching dynamics of the reversible photoswitchable fluorescence protein rsegfp2: crystal and solution.* **3ème Réunion plénière du GDR Ultrafast Phenomena. (Paris, France, 2018)**

International formation

4. **L. M. Uriarte,** R. Vitale, C. Ruckbusch, M. sliwa. *Photoswitching dynamics of the reversible photoswitchable fluorescence protein rsegfp2: crystal and solution.* **Ecole de physique des houches. Fluorescence markers for advance mycroscopy winter 2018. (Les Houches, France, 2018)**

Abstract

Among all applications of fluorescence proteins (FP), the most important one is undoubtedly their use as markers in fluorescence microscopy for *in-vivo* studies (Chemistry Nobel Prize 2008). In the last decade, a new type of FPs, reversible photoswitchable fluorescent proteins (RSFP), has found growing applications in super-resolution fluorescence microscopy (Chemistry Nobel Prize 2014). RSFPs can be toggled back and forward between a fluorescent state and a non-fluorescent state. The image resolution and the image acquisition speed parameters are linked to their photodynamics. However, the switching mechanism that controls these parameters is still a matter of debate. In this thesis, time resolved spectroscopy and crystallography permitted to infer the photo-switching mechanism of different RSFPs. These results will open crucial perspectives towards the design of new RSFPs, capable of successfully facing the most recent challenges of modern advanced fluorescent microscopy.

Résumé

Parmi toutes les applications des protéines fluorescentes, la plus répandue est sans aucun doute leur utilisation comme marqueurs en microscopie de fluorescence (Prix Nobel de Chimie 2008). Ces dernières années ont vu l'émergence d'un nouveau type de protéine fluorescente, les protéines photo-commutables réversibles qui sont maintenant couramment utilisées dans les microscopies de fluorescence super-résolues (Prix Nobel de chimie 2014). Ces protéines sont caractérisées par une commutation réversible entre un état fluorescent et un état non fluorescent. Les paramètres de résolution et de vitesse d'acquisition des images sont liés à leur photo-dynamique de commutation. Cependant, son mécanisme est encore sujet à question. Dans cette thèse, la spectroscopie et cristallographie résolue dans le temps nous a permis de déterminer le mécanisme de commutation pour différentes protéines. Ces résultats devraient ouvrir des perspectives dans la conception de nouveaux marqueurs pour la bio-imagerie.