Structural Information on the Trans-to-Cis Isomerisation Mechanism of the Photoswitchable Fluorescent Proteins rsEGFP2 Revealed by Multiscale Infrared Transient Absorption.

Journal:	*The Journal of Physical Chemistry Letters*
Manuscript ID	jz-2021-02920w.R1
Manuscript Type:	Letter
Date Submitted by the Author:	07-Nov-2021
Complete List of Authors:	Uriarte, Lucas Martinez ; Univ. Lille, LASIRE Vitale, Raffaele; Université de Lille, Laboratoire Avancé de Spectroscopie pour les Interactions, la Réactivité et l'Environnement Nizinski, Stanislaw; Uniwersytet im Adama Mickiewicza w Poznaniu, Faculty of Physics Hadjidemetriou, Kyprianos; Univ. Grenoble Alpes, CEA, CNRS Zala, Ninon; Université Grenoble Alpes, Institut de Biologie Lukacs, Andras; University of Pecs, Department of Biophysics Greetham, Gregory; Rutherford Appleton Laboratory, Central Laser Facility Sazanovich, Igor; Central Laser Facility, Research Complex at Harwell Weik, Martin; Institut de Biologie Structurale, Laboratoire de Biophysique Moleculaire Ruckebusch, Cyril; Universite Lille 1, LASIR Meech, Stephen; University of East Anglia, Chemical Sciences and Pharmacy Sliwa, Michel; Univ. Lille, LASIRE

Structural Information on the Trans-to-Cis Isomerisation Mechanism of the Photoswitchable Fluorescent Proteins rsEGFP2 Revealed by Multiscale Infrared Transient Absorption.

Lucas M. Uriarte[a,], Raffaele Vitale[a], Stanisław Niziński[a,b], Kyprianos Hadjidemetriou[c], Ninon Zala[c], Andras Lukacs[d], Gregory M. Greetham[e], Igor V. Sazanovich[e], Martin Weik[c], Cyril Ruckebusch[a], Stephen R. Meech[f], Michel Sliwa[a,*].*

AUTHOR ADDRESS

[a] Univ. Lille, CNRS, UMR 8516, LASIRE, LAboratoire de Spectroscopie pour les Interactions, la Réactivité et l'Environnement, Lille 59000, France

[b] Quantum Electronics Laboratory, Faculty of Physics, Adam Mickiewicz University, Uniwersytetu Poznańskiego 2, Poznan 61-614, Poland

[c] Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, 38000 Grenoble, France.

[d] Department of Biophysics, Medical School, University of Pecs, Szigeti ut 12, 7624 Pecs, Hungary

[e] Central Laser Facility, Research Complex at Harwell, Harwell Science and Innovation Campus, Didcot, Oxon OX11 0QX, UK

[f] School of Chemistry, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, U.K.

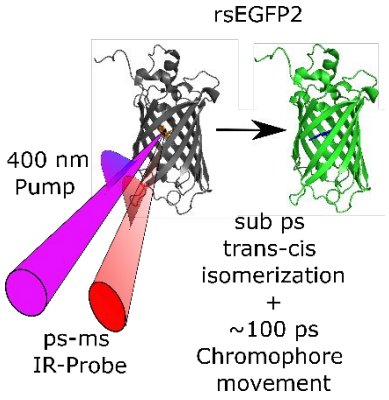
AUTHOR INFORMATION Corresponding Author

lucasmuriarte@gmail.com; michel.sliwa@univ-lille.fr

ABSTRACT

The rsEGFP2 is a reversible photoswitchable fluorescent protein (RSFP) used in super-resolved optical microscopies, which can be switched between a fluorescent On state and a non-fluorescent Off state. Previous studies have shown that the Off-to-On photoactivation extends over the femtosecond to the millisecond range and involves two picosecond-lifetime excited states and four ground state intermediates, reflecting a *trans*-to-*cis* excited state isomerization, a millisecond deprotonation and protein structural reorganizations. Femtosecond to millisecond time resolved multiple probe infrared spectroscopy (TRMPS-IR) is ideal to reveal the ~~structure~~ structural aspects of intermediate species in solution. Here we apply TRMPS-IR to rsEGFP2 and implement a ~~along with the application of~~ Savitzky-Golay derivative analysis, to correct ~~for~~ baseline drifts. The results reveal that a sub-picosecond twisted excited state precursor controls the *trans*-to-*cis* isomerization and a previously unreported step assigned to structural relaxation of the β-barrel occurs before the final millisecond deprotonation of the chromophore.

TOC

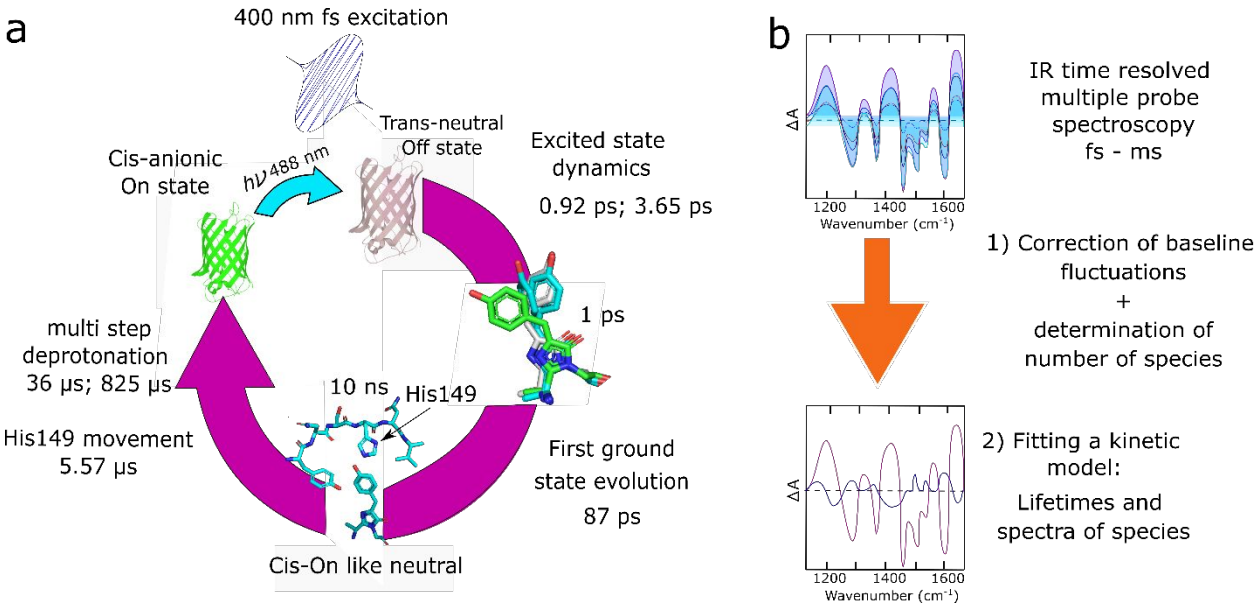


The structure determination of intermediates involved in the early stages (femtosecond to picosecond time scales) of protein photoactivation mechanisms has recently become possible with the advent of X-ray free-electron lasers (XFEL), coupled to a femtosecond UV-Vis pump source (time-resolved serial femtosecond crystallography, TR-SFX)[1–4]. Vibrational time resolved (TR) spectroscopies such as Raman or IR can disclose relevant structural information that complements the TR-SFX results. Importantly, they can uncover chromophore-protein interactions (not accessible by TRUV-Vis which is mainly sensitive to chromophore electronic transitions) and have been used to study complex photoactive biological systems such as fluorescent proteins[5–8] and other photoactive bio-systems[9–12]. Moreover, when used in a multiscale approach (time-resolved multiple probe spectroscopy, TRMPS) the dynamics from a few hundreds of femtoseconds (fs) up to the milliseconds (ms) time range can be measured[13].

The photoswitching mechanism of reversible photoswitchable fluorescent proteins (RSFPs) involves in general a *trans*-to-*cis* isomerization, a proton transfer and protein structural reorganizations[1,6,7,14]. Recently, the Off-to-On photodynamics of isotopically labelled Dronpa2, were studied using TRMPS-IR[6]. The formation of the first ground state photoproduct after a few picoseconds (ps) is followed by a protein reorganization in 596 ps and the formation of the final *cis* protonated chromophore (On like) state ~~*cis* (On like sate) protonated chromophore~~ in 91 nanoseconds (ns) ~~was reported~~. A second protein reorganization in 4.8 microseconds (µs) takes place before the final ~~microsecond (µs)~~ deprotonation[6]. On the contrary, recent studies from the fs-ps[1] and ps-ms[14] time scales by TR-SFX and TRUV-Vis spectroscopy on a similar RSFP[1], the rsEGFP2[15], proposed a different mechanism (Scheme 1a and Scheme S1 in SI). This protein has the same chromophore as Dronpa (p-hydroxybenzylidene-2,3-dimethylimidazolinone, HBDI) but a different protein cage (Fig. S5). The fs-ps study by TRUV-Vis revealed two excited state decays on the ps time scale involved in the *trans*-neutral (Off state) to *cis*-anionic (On state) photoswitching (Scheme S1, and Table S2), and the TR-SFX captured a twisted and a planar chromophore conformation at 1 ps and the formation of the *cis* On like chromophore at 3 ps. The TRUV-Vis ps-ms study[14] revealed the existence of four ground-state intermediates in solution with lifetimes of 87 ps, 5.57 µs, 36.1 µs and 825 µs, where the last two were assigned to deprotonation steps (D_2O isotopic effect). The X-ray structure determined at 10 ns featured a *cis*-On like chromophore geometry, where the main difference with the On state is found in the His149 side

chain position[14]. Therefore, the movement of the His149 side chain to the final On state was tentatively assigned to the 5.57 µs time constant[14].

Although the general Off-to-On photoswitching mechanism of rsEGFP2 is established, quantum chemical calculations predicted an almost zero transition dipole moment for the twisted model found in TR-SFX, which cannot correspond to either of the excited-state species identified by TRUV-Vis spectroscopy[1]. Thus, the precise geometry of the chromophore excited state species in solution and several of those in the ground state is still elusive. Here we use TRMPS-IR to gain structural information of the excited state dynamics and the different ground state intermediates in the solution phase switching of rsEGFP2.



Scheme 1. a) Current rsEGFP2 Off-to-On photoswitching mechanisms based on TRUV-vis and TR-SFX results[1,14]. The first chromophore structures (1 ps) represented in the panel correspond to the Off state in light grey (reproduced from PDB 5DTY), the twisted and planar chromophore structures determined at 1 ps in cyan (reproduced from PDB 5O8B) and the On state in green (reproduced from PDB 5O89). The cyan protein chromophore with part of the protein cage (10 ns) corresponds to the structure determined at 10 ns (reproduced from PDB 6T3A). b) General TRMPS-IR data analysis scheme.

The intrinsically low molar absorption coefficients in the IR spectral region and modest quantum yield of product formation makes TRMPS-IR signals of RSFPs typically very weak compared to those in TRUV-Vis. Consequently, despite advances in experimental methods[16–19], TRMPS-IR transient spectra of RSFPs are affected by intensity fluctuations[13,20]. These fluctuations translate

into baseline drifts and offsets, making preprocessing and signal correction key steps before the analysis of TRMPS-IR data, which typically involve a multi-exponential global-fit of the time traces[21–24]. The standard procedure to correct the baseline drift (Scheme 1b) is to fit a polynomial baseline function to each individual spectrum and subtract it from the raw data[17]. However, polynomial baseline correction methods commonly rely on parameters that are specified by the user (order of the polynomial, location of points for fitting). A major drawback of modelling the baseline shifts as a polynomial is, in some cases, the inability to select appropriate points that are both sufficiently well spread throughout spectral range and genuinely offset-free for the whole set of spectra. A common approach is to choose isosbestic points in the data, which are generally assumed to be crossing points (overlap between ground state depopulation and induced absorption) of modes that have shifted[17]. Altogether, these aspects can hinder the determination of the number of species and time constants involved in the photodynamics. Here we applied Savitzky-Golay derivative filter[25], a method used in different fields[26–29], to pre-process the TRMPS-IR data. The advantage of the Savitzky-Golay derivative approach rests on its non-arbitrary choice of support points and on the local nature of the derivative filter exploited to correct local drifts.

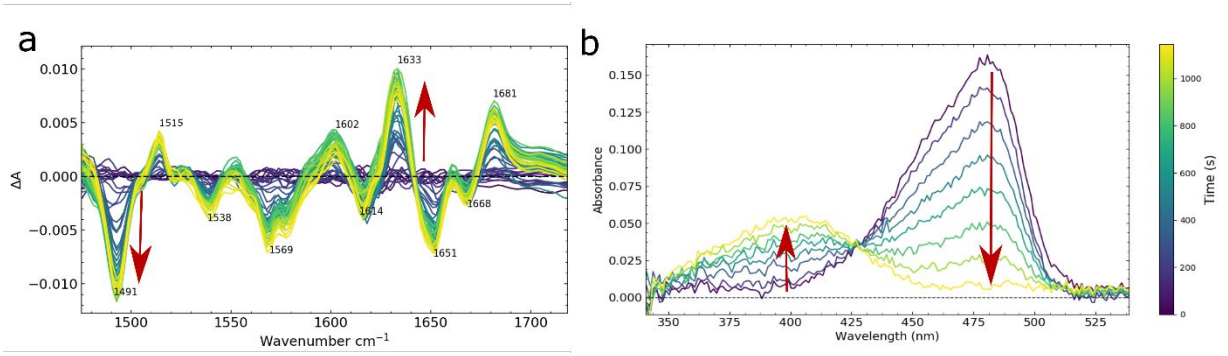


Fig. 1. rsEGFP2 On-to-Off photoswitching reaction (480 nm irradiation) at pD 8(50 mM NaCl, 50 mM HEPES) followed by a) difference FTIR spectra in the 1475-1720 cm^{-1} region (obtained by subtracting the non-irradiated On state spectrum) and b) UV-Vis steady state absorption spectroscopy between 340 to 540 nm.

To characterize the On and Off forms, first the On-to-Off kinetics of rsEGFP2 were measured by FTIR (Fig. 1a) and UV-Vis spectroscopy (Fig. 1b) under steady-state 480 nm irradiation in deuterated water at pD 8. UV-Vis absorbance spectra (Fig.1b) show an isosbestic point between the absorption bands of the chromophore in the Off state (*trans*-neutral conformation; 408 nm), and On state (*cis*-anionic conformation; 482 nm). Assuming that at pD 8 the thermodynamically stable state is the On state with the *cis*-anionic chromophore fully occupied, the photostationary

Off state is calculated to be a mixture of 90% *trans* protonated form and 10 % *cis*-anionic form. The switching quantum yield was determined to be 1% and 18% for On-to-Off and Off-to-On, respectively (Figure S2, SI), and the thermal back recovery time was fitted with a single exponential decay of 1.81 hours (Figure S3). These values are in agreement with those reported by Duwé et al.[30]. Similar to UV-Vis experiments, the FTIR difference absorbance spectra in Fig 1a (the resting On state spectrum is subtracted) are characterized by two species, the negative bands corresponding to the depopulation of the *cis*-anionic On state and the positive bands corresponding to the *trans*-neutral Off state formation. The assignment of the different vibrational bands can be made by comparison with Dronpa, which has the same chromophore[6–8,31] and HBDI literature[32]. The main On state bands are at those at 1491, 1538, 1569, 1614, 1651 and 1668 cm^{-1}. The 1491 cm^{-1} and 1569 cm^{-1} bands are characteristics of phenolate vibrations. The 1668 cm^{-1} is assigned to the imidazolinone C=O stretching mode[32], and the 1538 cm^{-1} to the C=N/C–C stretching mode[33]. The negative bands that cannot be assigned to the HBDI, correspond to infrared active modes of protein residues which couple differently to the *cis*-anionic and *trans*-neutral chromophore; these are the 1614 cm^{-1} and 1651 cm^{-1} bands. The main *trans*-neutral chromophore vibrations are observed at 1681, 1633, 1602 and 1515 cm^{-1}. The 1681 cm^{-1} was assigned to the C=O stretching of the chromophore imidazolinone group[6–8,31]; while the band at 1633 cm^{-1} can be assigned to the delocalized C=C–N=C bond (which may have an underlying protein contribution and phenyl ring stretch vibration)[8]. The 1602 cm^{-1} band is assigned to phenyl vibration[6], and the 1515 cm^{-1} one corresponds to the phenol group[7].

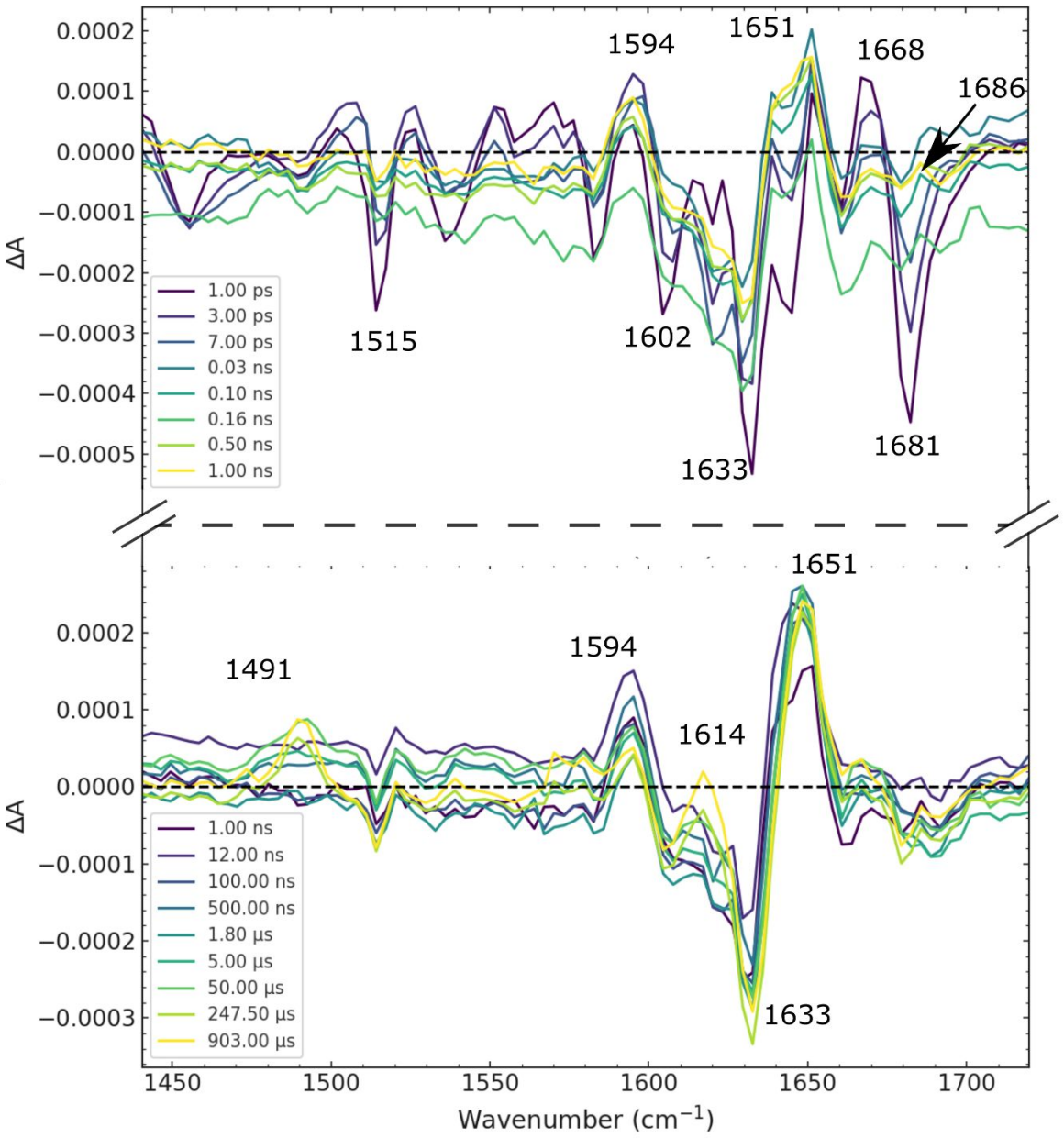


Fig 2. Transient IR absorption spectra of rsEGFP2 after 400 nm fs laser excitation of the trans-neutral chromophore (Off state*) at pD 8 (50 mM NaCl, 50 mM HEPES, pD 8); recorded by TRMPS-IR from 1 ps up to the ms range.*

The TRMPS-IR spectra recorded after a 400 nm laser excitation of the rsEGFP2 Off state (1 ps - 900 µs), are displayed in Figure 2. The corresponding TRUV-Vis data were published up to 40 ps[1] and are here extended until 2 ns in SI (Fig. S1). At 1 ps the rsEGFP2 TRMPS-IR spectrum is characterized by several negative bands (ground state bleaching, GSB), assigned to the depopulation of the *trans*-neutral Off form. Note that several of the negative bands in Fig. 2

correspond to positive bands in the FTIR spectra in Fig. 1a, while there are some new bands in Fig 2 which reflect prompt perturbations of the protein vibrational spectrum.

The ultrafast Off-to-On dynamics studied by TRUV-Vis fs spectroscopy (Fig. S1) shows the formation of the transient signal in about 300 fs. At this time several bands can be seen, 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 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 large band (490-600 nm) with a minimum at 515 nm together with a redshifted tail (600-720 nm). The growing of the later one is occurring with a certain delay in comparison to the one center at 515 nm. A fast initial evolution around 1 ps can be seen in the ESA bands centered at 335 nm and 455 nm together with the GSB band and SE band centered at 515 nm, these bands decrease to around half of their intensities (Fig. S1a). Importantly, within the same time, the red tail of the SE band decays completely. This evolution is followed by the decay of all the transient absorption bands in ~10 ps to form at 40 ps a transient spectrum characterized by a positive band at 375 nm and a small negative band between 465 and 580 nm with a maximum at 502 nm. The absence of SE at 40 ps after this time indicates that processes in the ground state are being probe. These evolutions are followed by the formation of a new positive band at 390 nm Fig. S1b. The global analysis yielded three time constants, an ultrafast component of 0.2 ps assigned to decay of the Franck Condon (FC) excited state followed by two excited state decays with lifetimes of 0.7 ± 0.1 and 4.8 ±0.3 ps (Fig. S1d), in agreement with previous results[1,14] (Table S2).

Similar evolutions of species can be observed in TRMPS-IR. The relaxation of the first excited state is characterized in TRMPS-IR spectra by the recovery of the GSB bands (~30% decrease of the 1681 cm^{-1} band at 3 ps) and a decrease and growth of positive bands at 1668 cm^{-1} and 1651 cm^{-1}, respectively. This initial evolution is followed by an incomplete recovery of the GSB bands during the relaxation of the second excited state. At 30 ps, both excited states have relaxed to the electronic ground state, and the formation of a band at 1686 cm^{-1} superimposed to a broad bleach band can be observed. The total recovery of the GSB bands is around ~80%, in agreement with the determined Off-to-On switching quantum yield of 18%. The growth of the band at 1686 cm^{-1},

slightly blue shifted with respect to the *trans* C=O mode, is assigned to the formation of a *cis*-neutral photoproduct based on observations in Dronpa and the HBDI associated DFT calculations[6] (discussed below). The assignment of the other chromophore modes at lower wavenumbers is complex due to overlapping protein modes (e.g., at 1651 cm^{-1}). From 30 ps to 1 ns further spectral evolutions are difficult to resolve due to baseline fluctuations (see 160, 500 and 1000 ps at 1600 cm^{-1}). The rsEGFP2 transient spectrum recorded at 1 ns (Fig. 2) displays two main positive bands at 1594 cm^{-1} and 1651 cm^{-1} and a ground state bleach at 1633 cm^{-1}. Strong baseline fluctuations hinder the observation of the photoproduct spectral evolution from 1 to 900 µs. Within the first 500 ns there is an increase in the amplitude of the positive band at 1651 cm^{-1} and the GSB at 1633 cm^{-1}. Finally, the formation of two positive bands at 1491 and 1614 cm^{-1}, characteristic of the *cis*-anionic On state, can be seen in a few hundred µs (see transient spectrum at 900 µs and FTIR spectra in Fig. 1a).

As discussed above, baseline fluctuations need to be corrected when performing a global decay analysis to obtain intermediate species decay associated difference spectra (DADS) and their lifetimes. Here we apply a Savitzky-Golay derivative filter, which to the best of our knowledge, has not been applied to TRIR or TRMPS-IR data so far. The validation of the Savitzky-Golay derivative filter to correct baseline drifts was carried out on simulated data sets as described in the supporting information (Fig. S6-S8) through a comparison with standard polynomial correction methods (see Fig. S9-S13 and Table S3). Considering the simulations, after the baseline corrections, the most relevant feature of the Savitzky-Golay derivative filter compared with the polynomial correction is a much accurate determination by a singular value decomposition (SVD) of the underlying number of components involved in the dynamics (see the outcomes on the simulated data Fig. S9). The analysis of TRMPS-IR data using the Savitzky-Golay derivative filter method is shown in Fig. 3 (data from 1 ns to 900 µs). The filter removes the baseline drift, revealing the spectral evolutions in the derivative spectra (Fig. 3b). The associated time constants and (derivative) DADS obtained from the global analysis of the preprocessed derivative data can be seen in Fig. 3c (bottom right panel). The final baseline drifts corrected DADS are obtained *via* a simple cumulative integration, to yield the DADS (Fig 3d).

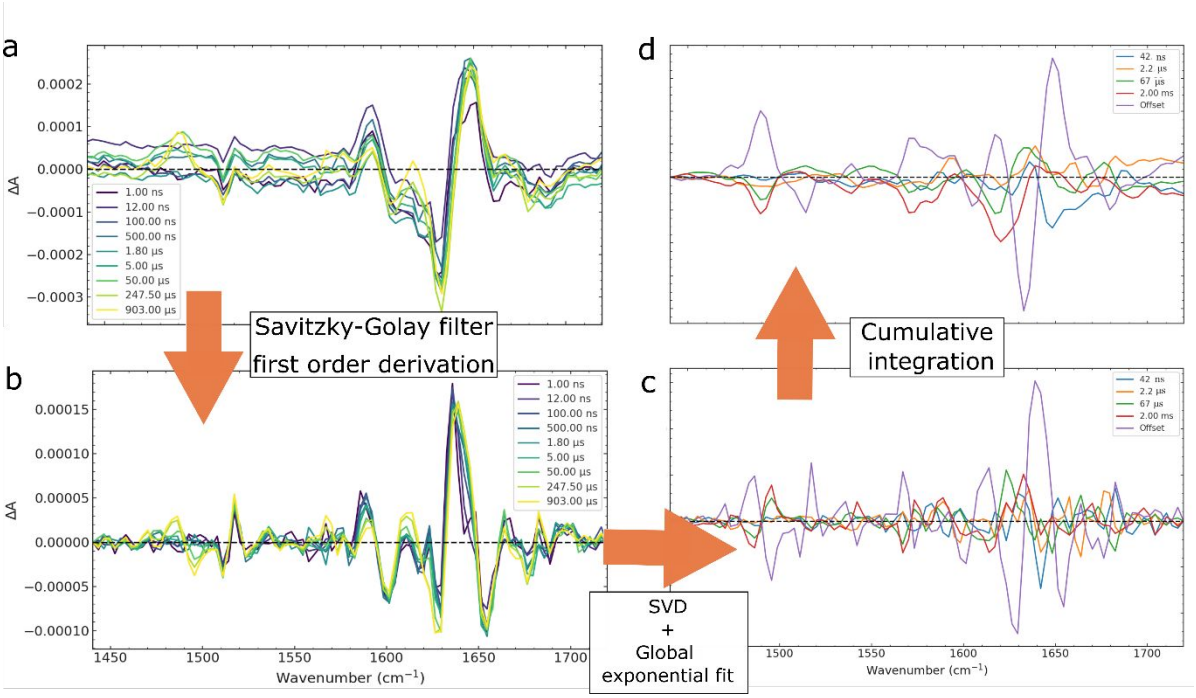


Fig. 3. Methodology used to apply Savitzky-Golay derivative filter to the TRMPS-IR data and obtain DADS spectra and associated time constants. a) IR transient absorption spectra of rsEGFP2 after 400 nm fs laser excitation of the Off form. b) Savitzky-Golay (5 point spectral window; third order polynomial interpolation) first order derivative of the time-resolved spectra. c) Determination of the number of components by SVD followed by a global decay analysis to obtain the derivative DADS and time constants. d) Cumulative integrated DADS.

After the application of the Savitzky-Golay-filter, the global fit of the data has been performed in two steps, from 1 ps to 2 ns and then from 1 ns to 900 µs; this prevents the high amplitudes of the initial evolution biasing the analysis of the long-time ones which exhibit lower signal amplitudes. The SVD of the entire rsEGFP2 TRMPS-IR derivative data set highlighted the presence of seven underlying components (three from 1 ps to 2 ns and four from 1 ns to 900 µs). The results are displayed in Fig. 4. Considering the higher S/N of the TRUV-Vis data and the number of time points collected between 0 and 1 ps (first time point recorded in TRMPS-IR is 1 ps), we fixed the first time constant to 0.7 ps and ~~leave~~left the other two~~rest~~ free for the TRMPS-IR data[34], obtaining a value of 5.5 ± 0.1 ps and 117± 7 ps. The DADS of the first component (0.7 ps) is characterized by a positive contribution of the excited state at 1620 cm^{-1} and a negative one at 1686 cm^{-1} (a negative DADS for a positive signal in the transient spectra implies the formation of this signal, Fig. 4a blue curve), consistent with the appearance of the positive band at 1686 cm^{-1} in the original data (discussed above, Figure 2, 30 ps). The 1620 cm^{-1} band can be assigned to the chromophore

C=C vibration by comparison with the attributions done by Warren et al.[7] in Dronpa, although there may be contributions from protein modes according to Laptenok et al.[6]. The 1686 cm^{-1} band is assigned to the C=O stretching of a *cis* chromophore and is located near the C=O stretching band of the *trans* chromophore peaking at 1681 cm^{-1} (positive band in Fig. 1). This redshift of the C=O stretching peak between a *cis* and a *trans* chromophore has been observed and considered the primary *trans*-to-*cis* isomerization signature for Dronpa and Dronpa2 and the o-HBDI analogue[35]. Concomitantly with the decay of the 0.7 ps excited state, a partial recovery of the GSB is observed. The partial recovery of the GSB signals indicates that the ultrafast isomerization occurs through a conical intersection where part of the chromophore evolves back to the original *trans*-neutral conformer and the remainder forward to a *cis*-neutral chromophore.

Contrary to the 0.7 ps component, the 5.5 ps (4.8 ps in TRUV—Vis) DADS (Fig. 4a orange curve) is mainly characterized by GSB recovery and the excited state absorption (ESA) decay. The existence of an SE band for the 4.8 ps DADS indicate the presence of an ESA. The presence of a positive band that is red shifted in comparison to the maximum of the GSB band can also be assigned to the presence of a hot ground state. The origin of this hot ground state is attributed to the relaxation of part of the excited state species that, after the decay through a conical intersection in 0.7 ps, relax back to the original *trans* form. Such vibrational relaxation could explain the longer time component found for the second ESA by TRMPS-IR, 5.5 ps vs 4.8 ps, since IR signals are more affected by vibrational cooling processes. Within the 5.5 ps component in TRMPS-IR, essentially, the recovery of the main characteristic bands of the *trans*-neutral chromophore peaking at 1515 cm^{-1}, 1602 and 1682 cm^{-1} is observed. The first two bands correspond to the phenyl vibrations and the last one to the C=O stretching of the imidazolinone group. Therefore, the 5.5 ps component is attributed to a conformation that does not lead to isomerization, while 0.7 ps excited state corresponds to the *cis* photo-product precursor. Since the DADS of the TRUV-Vis data short component (0.7 ps, Fig. S1f) has a redshifted stimulated emission, it is assigned to a twisted chromophore in the excited state, while the 5.5 ps is assigned to a planar chromophore in the excited state, similarly to fluorophores that show twisted intramolecular charge transfer, the emission from the TICT is red shifted in comparison to the one from the localized excited state which present no change of geometry[36]. Moreover, twisted HBDI chromophores have much shorter lifetimes than planar HBDI[37,38]. The twisted geometry of the chromophore in the excited state suggested by UV-Vis and IR probably differs from the one captured by TR-SFX at 1 ps

considering that quantum mechanical calculations predicted a zero transition dipole moment for the latter[1]. On the contrary the same calculations predicted that the formation of the twisted chromophore in the excited state is completed in less than 0.5 ps via a Hula-Twist mechanism. For the S_1 state they yielded a major twisted chromophore conformation (τ=-73°, Φ = 11°), and an additional minor planar conformation (τ =-30°, Φ = -11°)[1]. The two excited state species (0.7 ps; 5.5 ps) may arise from a bifurcation of the FC state in 200 fs, populating two distinct excited states. Alternatively, two distinct chromophore ground state conformations may lead to two excited states with different decay times; note that the presence of two different co-existing *trans* forms was recently reported for crystalline rsEGFP2[14,39,40]. After the decay of the 5.5 ps component the GSB bands recover to about 80 %. Since no extra recovery of the main GSB band can be observed at longer delay times, we consider that there is no thermal back reaction from the following intermediates photoproduct back to the Off state.

After the decay to the ground state, a species with a time constant of 117 ± 7 ps was observed (Fig. 4a green curve) in agreement with the 98 ± 16 ps time constant retrieved from the UV-Vis data (Fig. S1). The corresponding species is characterized by negative and positive bands at 1680 cm^{-1} and 1691 cm^{-1} assigned to a shift of the C=O stretching. Similarly, negative signals at 1650 cm^{-1} and 1620 cm^{-1} assigned to a protein residues and phenyl ring stretches are also observed. The shift of the C=O stretching after the decay of the excited state suggests the relaxation of the *cis*-neutral chromophore to its final *cis* On like neutral position (TR-SFX structure captured at 10 ns, Scheme 1) and not solely to a protein rearrangement as previously reported[14]. ~~This is much faster than the larger-scale ground state chromophore reorganization in 91 ns reported for Dronpa2~~. The appearance of the *cis* isomer (1691 cm^{-1} band) for rsEGFP2 in a few ps following an excited state isomerization reaction is different to observations on the related Dronpa2 protein[6], where the cis isomer (1702 cm^{-1} band) appears in 91 ns, after a protein ground state reorganization of the picosecond formed metastable intermediate. In drompa2 the assignment of 1702 cm^{-1} band to *cis* isomer was confirmed by ^{13}C isotope labelling[6]. In rsEGFP2 the isomerization precedes protein structure change. The difference is ascribed to significant difference in the structure between the two proteins (see X-ray structures of Dronpa and rsEGFP2 in Fig. S5), which include both On and Off structures of the chromophore ground state and the H-bonding environment of the C=O himidazolinone group with the surrounding residues.

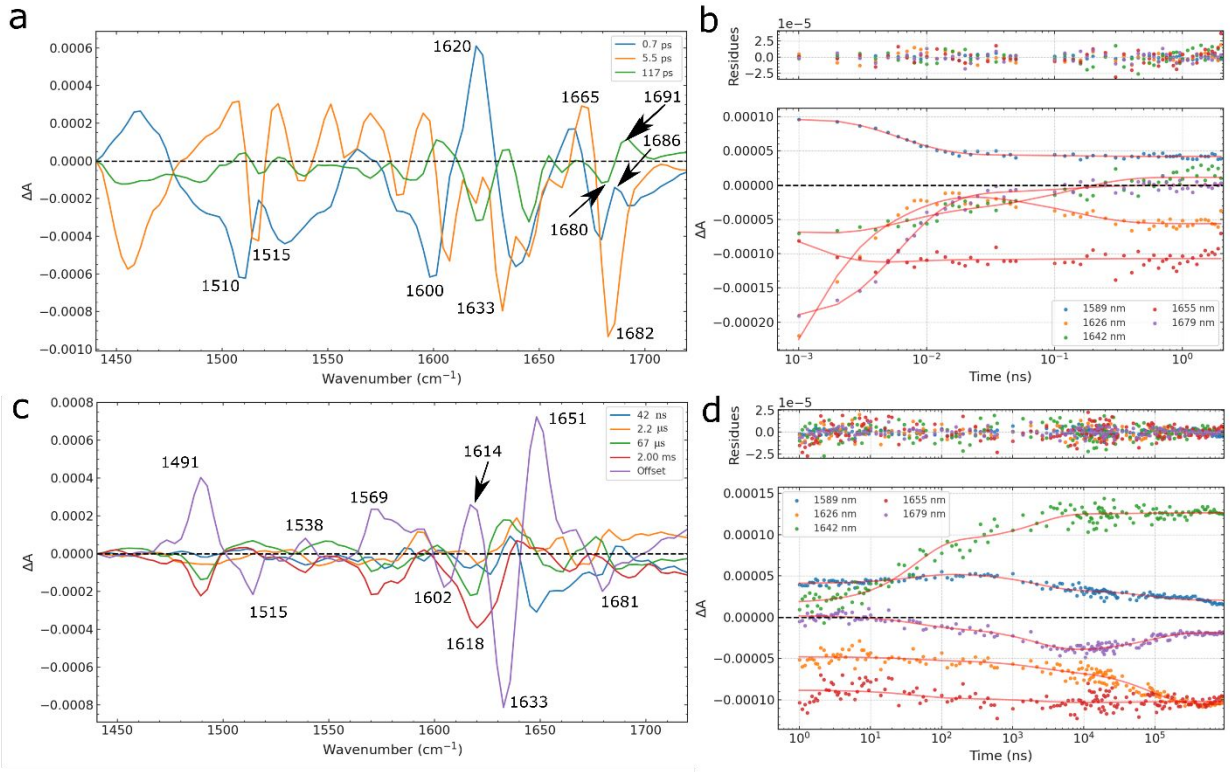


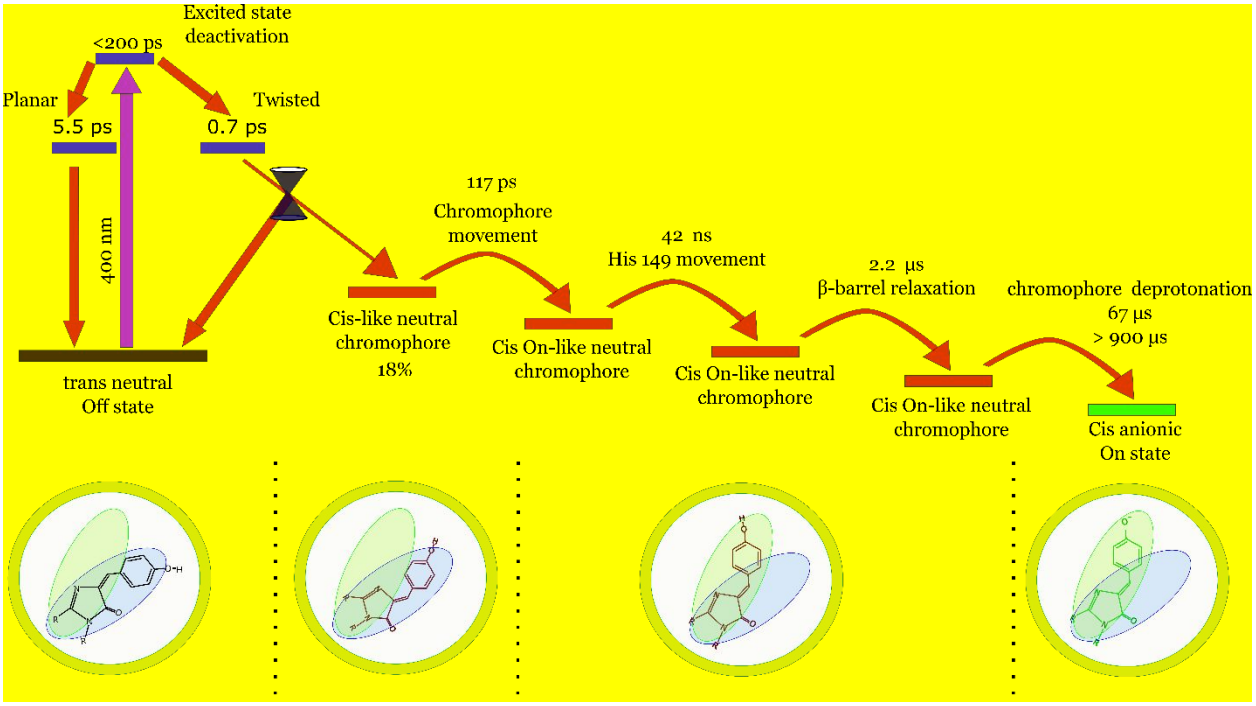
Fig. 4. TRMPS-IR data global analysis results**.** Left panels correspond to the cumulative integrated DADS reconstructed using the pre-exponential factors obtained from the global fit analysis of the preprocessed traces with a weighted sum of four exponential functions (Note the DADS are already reintegrated). a) Data from 1 ps to 2ns. b) Data from 1 ns to 900 µs. In the right panels, a selection of traces represented together with the corresponding fits and residuals. c) Data from 1 ps to 2ns. d) Data from 1 ns to 900 µs. (Note that the time axes are set on a logarithmic scale).

Four time constants were needed to fit the signal evolutions after 1 ns (Fig. 4c,d). Note that the 2 ms time constant reported elsewhere (in D_2O)[14] cannot be obtained precisely as the longest time delay probed by TRMPS-IR is 900 µs, and thus ~~this value~~the 2 ms was used as a constant in the fit. If it is omitted, the quality of the fit is unaffected, but the final state (offset) then fails to reproduce the FTIR data, as required at long times (Fig. S4b). The other three time constants found are 42 ± 2 ns, 2.2 ± 0.1 µs and 67 ± 3 µs. In addition to the 2 ms time constant, our previous UV-Vis study[14] reported two time constants of 5.16 µs and 88.4 µs (no evolution in the ns range was reported). The 10 ns intermediate state structure obtained by TR-SFX features the His149 in an Off-like position[14]. Thus the first UV-Vis time (5.16 µs) constant was attributed to the structural evolution of His149 prior to the multi-step chromophore deprotonation (88.4 µs, 2 ms; time

constants both show significant deuterium isotopic effects)[14]. The 67 µs and 2 ms species IR DADS (Fig. 4c, the two DADS are represented solely in Figure S4a for clarity) revealed the formation of the CO^- phenolate stretching band at 1491 cm^{-1} and, therefore, agrees with the D_2O isotope effect and multi-step deprotonation observed by Woodhouse et al.[14]. The movement of the His149 side chain was expected to occur before the deprotonation steps[14] and, thus, can be attributed either to the 42 ns or to the 2.2 µs constants found in TRMPS-IR. The δ nitrogen atom (ND1) of His149 ~~side chain~~ is hydrogen-bonded to the chromophore phenolate in the On state. The 42 ns evolution has important contributions in the 1651 cm^{-1} positive band (assigned to the amide mode of residues in interaction with the chromophore) and in the 1633 cm^{-1} GSB representative of the *trans* form (FTIR Fig. 1a) and thus agrees with the movement of an amino acid side chain in interaction with the chromophore. The fact that the 2.2 µs DADS shows only small variations in the band intensities suggests rearrangements that provoke small or no changes in the chromophore environment. Therefore, we tentatively attribute the 2.2 µs evolution to the β-barrel relaxation, involving particularly the 7th, 8th and 10th, β-strands, which, as shown by rsFolder NMR studies[41] or Dronpa[42], are the strands undergoing major distortions in the photoswitching process. Moreover, NMR studies on GFP have shown that the β-barrel is rigid on the ps-ns, while significant flexibility on the µs-ns time scales has been observed in the β-strands 3, 7, 8, and 10 and in the α-helix carrying the chromophore[43]. This time scales agrees with the time constant of 2.2 µs being attributed to the β-barrel relaxation.

In summary, we hypothesize that once the *cis*-neutral chromophore (117 ps) and the surrounding amino acids (42 ns) adopt their final position, a general relaxation of the protein occurs (2.2 µs), that provoke changes in the pKa which in-turn triggers the chromophore deprotonation in the µs-ms range (Scheme 2). The offset obtained from the fit (Fig 4c) is in agreement with the FTIR spectra and confirms both the deprotonation step occurring in the ms range and the formation of the final *cis*-anionic On state[14]. The two deprotonation process can be due to either a multi-step deprotonation via Thr204 and His149 to the solvent outside the protein as suggested for GFP[44] which is triggered by a change in the pKa once all protein rearrangements have taken place, or to an average signal of several protonation/deprotonation events between the chromophore and the His149 in a sub-ms time scale, as suggested by recent studies on rsFolder[45], that finally connects the water molecule that is hydrogen bond to the chromophore with the solvent via His149 and a chain of three water molecules[14]. Altogether, the UV-Vis, TR-SXF and TRMPS-IR results allow

us to suggest the following modified Off-to-On photo-dynamical scheme for rsEGFP2 in solution (Scheme 2).



Scheme 2. Proposed Off-to-On photoswitching mechanism for rsEGFP2.

The proposed scheme could be obtained thanks to the study of the rsEGFP2 photodynamics by TRMPS-IR combined with the Savitzky-Golay derivative filter (only one *trans* Off state is considered)[1]. We revealed a previously unreported decay time in the ns range, which has been tentatively attributed to the movement of the His149 side chain, and the few µs evolution reassigned to the relaxation of the β-barrel (Scheme 2). More importantly, the appearance of a *cis*-like chromophore band in few ps is ~~assigned to~~evidence for a sub-ps twisted excited state precursor that controls the *trans*-to-*cis* isomerization followed by a hundred ps relaxation in the ground state to the protonated *cis*-On like chromophore. On-going theoretical calculations of IR vibrations that take into account the protein residues will allow us to assign precisely the different bands observed to unveil the different chromophore-protein interactions. Future studies on rsEGFP2 variants mutating the His149 could verify the hypothesis that the nanosecond time constant is assigned to histidine movement.

ACKNOWLEDGMENTS

LMU thanks the Lille University for the PhD fellowship and Jean-Julien Dubois for technical support in time-resolved spectroscopy experiments. MENESR – Univ. Grenoble Alpes PhD fellowship to KH. The Chevreul Institute (FR 2638), the Ministère de l'Enseignement Supérieur et de la Recherche, the Région Hauts de France and FEDER are acknowledged for financial support. ANR grant to M.W. and M.S.l. (BioXFEL). S.R.M. acknowledges EPSRC for financial support (EP/N033647/1 and EP/M001997/1). A.L. acknowledges funding from EFOP-3.6.2-16-2017-00005. The authors acknowledge UKRI (STFC) for access to the Central Laser Facility.

ASSOCIATED CONTENT

Supporting Information. Materials and methods, TRUV-Vis of rsEGFP2 in D_2O, determination of photoswitching quantum yields in D_2O, thermal back reaction, simulated data sets and application of the Savitzky-Golay derivate filter, comparison of the Savitzky-Golay derivate filter with the polynomial correction.

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. *Nat. Chem.* **2018**, *10*, 31–37.

(2) 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. *Nat. Commun.* **2019**, *10*, 3177.

(3) 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.

(4) 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.

(5) 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. *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 9303–9307.
(6) 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. *Nat. Chem.* **2018**, *10*, 845–852.
(7) Warren, M. M.; Kaucikas, M.; Fitzpatrick, A.; Champion, P.; Sage, J. T.; van Thor, J. J. Ground-state proton transfer in the photoswitching reactions of the fluorescent protein Dronpa. *Nat. Commum.* **2013**, *4*, 1461.
(8) Lukacs, A.; Haigney, A.; Brust, R.; Addison, K.; Towrie, M.; Greetham, G. M.; Jones, G. A.; Miyawaki, A.; Tonge, P. J.; Meech, S. R. Protein photochromism observed by ultrafast vibrational spectroscopy. *J. Phys. Chem. B.* **2013**, *117*, 11954–11959.
(9) Iuliano, J. N.; Hall, C. R.; Green, D.; Jones, G. A.; Lukacs, A.; Illarionov, B.; Bacher, A.; Fischer, M.; French, J. B.; Tonge, P. J.*; et al.* Excited State Vibrations of Isotopically Labeled FMN Free and Bound to a Light-Oxygen-Voltage (LOV) Protein. *J. Phys. Chem. B* **2020**, *124*, 7152–7165.
(10) Pirisi, K.; Nag, L.; Fekete, Z.; Iuliano, J. N.; Tolentino Collado, J.; Clark, I. P.; Pécsi, I.; Sournia, P.; Liebl, U.; Greetham, G. M.*; et al.* Identification of the vibrational marker of tyrosine cation radical using ultrafast transient infrared spectroscopy of flavoprotein systems. *Photochem. Photobiol. Sci.* **2021**, *20*, 369-378.
(11) Kübel, J.; Chenchiliyan, M.; Ooi, S. A.; Gustavsson, E.; Isaksson, L.; Kuznetsova, V.; Ihalainen, J. A.; Westenhoff, S.; Maj, M. Transient IR spectroscopy identifies key interactions and unravels new intermediates in the photocycle of a bacterial phytochrome. *Phys. Chem. Chem. Phys.* **2020**, *22*, 9195–9203.
(12) Fang, C.; Tang, L. Mapping Structural Dynamics of Proteins with Femtosecond Stimulated Raman Spectroscopy. *Annu. Rev. Phys. Chem.* **2020**, *71*, 239–265.
(13) Greetham, G. M.; Sole, D.; Clark, I. P.; Parker, A. W.; Pollard, M. R.; Towrie, M. Time-resolved multiple probe spectroscopy. *Rev. Sci. Instrum.* **2012**, *83*, 103107.
(14) 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. *Nat. Commun.* **2020**, *11*, 741.
(15) 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.
(16) Towrie, M.; Grills, D. C.; Dyer, J.; Weinstein, J. A.; Matousek, P.; Barton, R.; Bailey, P. D.; Subramaniam, N.; Kwok, W. M.; Ma, C.*; et al.* Development of a broadband picosecond infrared spectrometer and its incorporation into an existing ultrafast time-resolved resonance Raman, UV/visible, and fluorescence spectroscopic apparatus. *Appl. Spectrosc.* **2003**, *57*, 367–380.
(17) Greetham, G. M.; Burgos, P.; Cao, Q.; Clark, I. P.; Codd, P. S.; Farrow, R. C.; George, M. W.; Kogimtzis, M.; Matousek, P.; Parker, A. W.*; et al.* ULTRA: A Unique Instrument for Time-Resolved Spectroscopy. *Appl. Spectrosc.* **2010**, *64*, 1311–1319.

(18) Matousek, P.; Towrie, M.; Stanley, A.; Parker, A. W. Efficient Rejection of Fluorescence from Raman Spectra Using Picosecond Kerr Gating. *Appl. Spectrosc.* **1999**, *53*, 1485–1489.
(19) Ghosh, A.; Smits, M.; Bredenbeck, J.; Dijkhuizen, N.; Bonn, M. Femtosecond time-resolved and two-dimensional vibrational sum frequency spectroscopic instrumentation to study structural dynamics at interfaces. *Rev. Sci. Instrum.* **2008**, *79*, 93907.
(20) Feng, Y.; Vinogradov, I.; Ge, N.-H. Optimized noise reduction scheme for heterodyne spectroscopy using array detectors. *Opt. Express.* **2019**, *27*, 20323–20346.
(21) Beechem, J. M.; Ameloot, M.; Brand, L. Global and Target Analysis of Complex Decay Phenomena. *Instrum. Sci. Technol.* **1985**, *14*, 379–402.
(22) 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. *Chem. Phys. Lett.* **1995**, *246*, 641–648.
(23) van Stokkum, I. H. M.; Larsen, D. S.; van Grondelle, R. Global and target analysis of time-resolved spectra. *Biochim. Biophys. Acta.* **2004**, *1657*, 82–104.
(24) Ruckebusch, C.; Sliwa, M.; Pernot, P.; Juan, A. de; Tauler, R. Comprehensive data analysis of femtosecond transient absorption spectra: A review. *J. Photochem. Photobiol. C* **2012**, *13*, 1–27.
(25) Savitzky, A.; Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627–1639.
(26) Zimmermann, B.; Kohler, A. Optimizing Savitzky-Golay parameters for improving spectral resolution and quantification in infrared spectroscopy. *Appl. Spectrosc.* **2013**, *67*, 892–902.
(27) Cobas, J. C.; Bernstein, M. A.; Martín-Pastor, M.; Tahoces, P. G. A new general-purpose fully automatic baseline-correction procedure for 1D and 2D NMR data. *J. Magn. Reson.* **2006**, *183*, 145–151.
(28) Rinnan, Å.; van den Berg, F.; Engelsen, S. B. Review of the most common pre-processing techniques for near-infrared spectra. *Trends. Analyt. Chem.* **2009**, *28*, 1201–1222.
(29) Vitale, R.; Bevilacqua, M.; Bucci, R.; Magrì, A. D.; Magrì, A. L.; Marini, F. A rapid and non-invasive method for authenticating the origin of pistachio samples by NIR spectroscopy and chemometrics. *Chemom. Intell. Lab. Syst.* **2013**, *121*, 90–99.
(30) 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.
(31) Kaucikas, M.; Tros, M.; van Thor, J. J. Photoisomerization and proton transfer in the forward and reverse photoswitching of the fast-switching M159T mutant of the Dronpa fluorescent protein. *J. Phys. Chem. B.* **2015**, *119*, 2350–2362.
(32) He, X.; Bell, A. F.; Tonge, P. J. Isotopic Labeling and Normal-Mode Analysis of a Model Green Fluorescent Protein Chromophore. *J. Phys. Chem. B* **2002**, *106*, 6056–6066.
(33) van Thor, J. J.; Ronayne, K. L.; Towrie, M.; Sage, J. T. Balance between ultrafast parallel reactions in the green fluorescent protein has a structural origin. *Biophys. J.* **2008**, *95*, 1902–1912.

(34) 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. *J. Phys. Chem. Lett.* **2017**, *8*, 3530–3535.
(35) 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. *J. Phys. Chem. B* **2016**, *120*, 9716–9722.
(36) Kumpulainen, T.; Lang, B.; Rosspeintner, A.; Vauthey, E. Ultrafast Elementary Photochemical Processes of Organic Molecules in Liquid Solution. *Chem. Rev.* **2017**, *117*, 10826–10939.
(37) Conyard, J.; Heisler, I. A.; Chan, Y.; Bulman Page, P. C.; Meech, S. R.; Blancafort, L. A new twist in the photophysics of the GFP chromophore: a volume-conserving molecular torsion couple. *Chem. Sci.* **2018**, *9*, 1803–1812.
(38) 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.
(39) Chang, J.; Romei, M. G.; Boxer, S. G. Structural Evidence of Photoisomerization Pathways in Fluorescent Proteins. *J. Am. Chem. Soc.* **2019**, *141*, 15504–15508.
(40) Adam, V.; Hadjidemetriou, K.; Jensen, N.; Shoeman, R. L.; Woodhouse, J.; Aquila, A.; Banneville, A.-S.; Barends, T. R. M.; Bezchastnov, V.; Boutet, S.*; et al.* Rational control of structural off-state heterogeneity in a photoswitchable fluorescent protein provides switching contrast enhancement; *bioRxiv*, **2021**. DOI: 10.1101/2021.11.05.462999.
(41) 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. *Biophys. J.* **2019**, *117*, 2087–2100.
(42) Mizuno, H.; Mal, T. K.; Wälchli, M.; Kikuchi, A.; Fukano, T.; Ando, R.; Jeyakanthan, J.; Taka, J.; Shiro, Y.; Ikura, M.*; et al.* Light-dependent regulation of structural flexibility in a photochromic fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9227–9232.
(43) Seifert, M. H. J.; Georgescu, J.; Ksiazek, D.; Smialowski, P.; Rehm, T.; Steipe, B.; Holak, T. A. Backbone dynamics of green fluorescent protein and the effect of histidine 148 substitution. *Biochemistry* **2003**, *42*, 2500–2512.
(44) Agmon, N. Proton pathways in green fluorescence protein. *Biophys. J.* **2005**, *88*, 2452–2461.
(45) Christou, N. E.; Giandoreggio-Barranco, K.; Ayala, I.; Glushonkov, O.; Adam, V.; Bourgeois, D.; Brutscher, B. Disentangling Chromophore States in a Reversibly Switchable Green Fluorescent Protein: Mechanistic Insights from NMR Spectroscopy. *J. Am. Chem. Soc.* **2021**, *143*, 7521–7530.