Photoinduced proton transfer inside an engineered green fluorescent protein: A stepwise-concerted-hybrid reaction

Journal Name

ARTICLE

Photoinduced proton transfer inside an engineered green fluorescent protein: A stepwise-concerted-hybrid reaction†

Longteng Tang,[a] Yanli Wang,[a] Liangdong Zhu,[a] Karen Kallio,[b] S. James Remington,[b] and Chong Fang*[a]

Photoactivated proton transfer (PT) wire is responsible for the glow of green fluorescent protein (GFP), which is crucial for bioimaging and biomedicine. In this work, a new GFP-S65T/S205V double mutant is developed from wild-type GFP in which the PT wire is significantly modified. We implement femtosecond transient absorption (fs-TA) and femtosecond stimulated Raman spectroscopy (FSRS) to delineate the PT process in action. The excited state proton transfer proceeds on the ~110 ps timescale, which infers that the distance of one key link (water to T203) in the PT wire of GFP-S205V is shortened by the extra S65T mutation. The rise of an imidazolinone ring deformation mode at ~871 cm^{-1} in FSRS further suggests that this PT reaction is in a concerted manner. A ~4 ps component prior to large-scale proton dissociation through the PT wire is also retrieved, indicative of some small-scale proton motions and heavy-atom rearrangement in the vicinity of the chromophore. Our work provides deep insights into the novel hybrid PT mechanism in engineered GFP and demonstrates the power of tunable FSRS methodology in tracking ultrafast photoreactions with the desirable structural specificity in physiological environments.

Introduction

Photoinduced proton transfer (PT) is a ubiquitous and important phenomenon found in many chemical and biological systems.[1-6] One of the most famous examples is the wild-type green fluorescent protein (wtGFP), whose discovery and development have revolutionized molecular and cellular biology and many related fields.[7-11] In the electronic ground state (S_0), its absorption spectrum has a major peak at ~395 nm due to the neutral form (A) of the chromophore, and a minor peak around 475 nm from the anionic form (B).[12] Excitation at 400 nm leads to green fluorescence at 508 nm. Such a large Stokes shift (~5300 cm^{-1}) is attributed to excited state proton transfer (ESPT) through the well-established chromophore (CRO)-W22-S205-E222 hydrogen (H)-bonding network (Fig. 1a).[3,13-15] The bright green fluorescence is emitted from the deprotonated form (I*) in an unrelaxed protein environment.

Previous studies showed that mutating the residue S205, which is an integrated part of the ESPT chain in wtGFP (Fig. 1a), to valine (S205V) does not block proton transfer, but slow the

process down by ca. 30-fold.[16-18] Examining the crystal structure of the GFP-S205V mutant reveals that an adjacent residue T203 undergoes rotation and bridges the PT wire. The long (~3.5 Å) hydrogen bond between T203 and a water molecule is likely the rate-determining step in the ESPT process. Further mutation of T203V abolishes ESPT and achieves a blue fluorescent protein with a Tyr-based chromophore.[16] Another interesting mutation is replacing Ser65 with Thr (S65T) on the chromophore, close to the PT terminal site, which can suppress the A form and increase the deprotonated B state with brighter emission. Later research showed that this mutation leads to an alternative conformation of the final proton acceptor E222, thus significantly impacts the H-bonding network near the TYG chromophore.[19,20]

In the current work, we investigate the wtGFP variant with both S205V and S65T mutations (see Fig. 1b for illustration). Two complementary ultrafast spectroscopic techniques, femtosecond transient absorption (fs-TA) and femtosecond stimulated Raman spectroscopy (FSRS), are implemented to unravel the PT mechanism. TA provides a wealth of information in the excited electronic domain[21] while FSRS tracks molecular structural evolution on the electronic ground and excited state with simultaneously high temporal and spectral resolutions.[4,22-25] Both techniques have been successfully used to study GFP and the GFP-based biosensors,[23,26-29] particularly for tracking the I* rise in addition to A* decay. The ESPT mechanism inside our GFP double mutant is revealed to be a hybrid of concerted and

a. Oregon State university, Department of Chemistry, 263 Linus Pauling Science Centre (lab), 153 Gilbert Hall (office), Corvallis, OR 97331, USA. *E-mail: Chong.Fang@oregonstate.edu; Fax: +1 541 737 2062; Tel: +1 541 737 6704; Web: http://fanglab.oregonstate.edu/

b. Institute of Molecular Biology and Department of Physics, University of Oregon, Eugene, Oregon 97403, USA.

† Electronic Supplementary Information (ESI) available: Fig. S1–S2 with additional discussions and references. See DOI: 10.1039/x0xx00000x
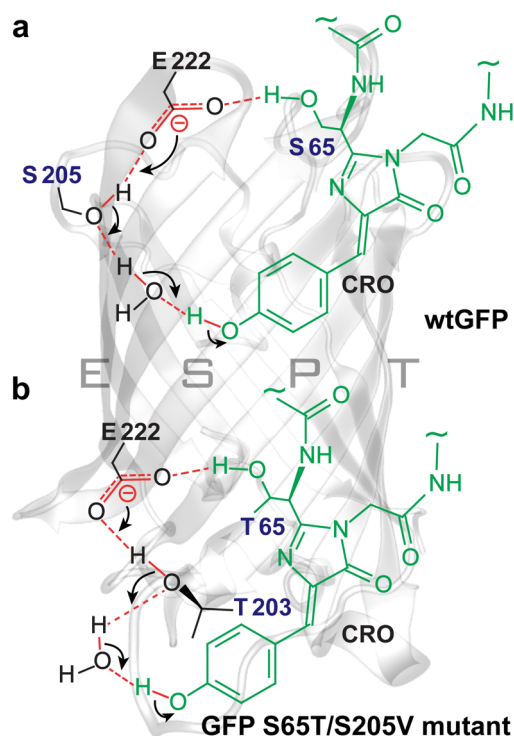
Fig. 1 Scheme of the excited state proton transfer (ESPT) wire in (a) wtGFP and (b) the S65T/S205V mutant. The overall β-barrel structure of a common GFP is shown as the background. The arrows indicate the transfer direction of electrons.

Fig. 2 Steady-state electronic spectra of the GFP-S65T/S205V mutant in aqueous buffer solution (pH=7.9). The normalized absorbance spectrum is shown in solid blue. The normalized fluorescence spectra under 400 and 505 nm excitation are in solid and dashed green traces, respectively. A small shoulder on the red side is noted.

stepwise nature with a notable separation of characteristic atomic motions on different timescales. The correlation of the ESPT rate with the chromophore local environment especially along the H-bonding chain (PT wire) also sheds light on the important roles of protein residues as well as bridging water molecules in determining the fluorescence properties.

Results and discussion

We note that GFP and its many derivatives have found wide applications in bioimaging and biomedicine, which make it almost routine to visualize and track cellular processes. At this stage, what role can the mechanistic understanding play in practical applications? In general, the use of GFP-based probes and sensors has benefited from trial-and-error approaches from a large library of candidates by researchers and engineers. But a rational design strategy which promises more hypothesis-driven discoveries with higher efficiency requires fundamental knowledge about the underlying photochemistry, which is the focus of this work. Besides achieving improved properties to light up the interior of a cell or a hidden state of a biological process, these colorful biomolecules are excellent model systems to study and elucidate the foundational chemistry and biophysics principles with chemical bond resolution on intrinsic molecular timescales, as detailed as the transfer of a proton. The uncovered structure-function relationships can be applied to other biomolecular systems in physiological environments.
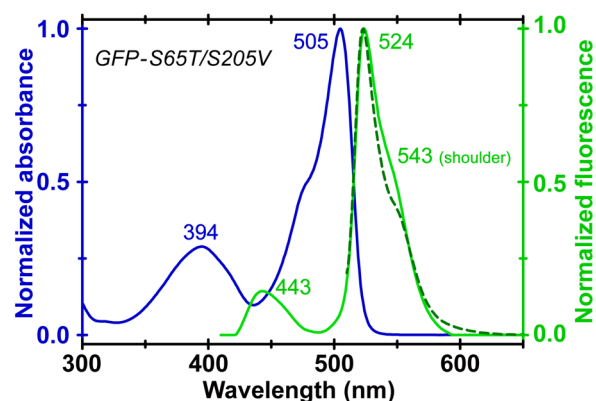
Steady-state electronic spectra of GFP-S65T/S205V mutant

Unlike wtGFP and its S205V mutant, which are largely in the protonated A state at ground state, the GFP-S65T/S205V mutant has a dominant deprotonated B form absorbing at 505 nm and a minor A species absorbing at 394 nm (Fig. 2) which is a result of the S65T mutation, resembling the enhanced GFP (EGFP).[10,14,30] Exciting the B form by 505 nm light leads to green fluorescence centered at 523 nm.[10,14] Upon 400 nm excitation of the A form, the mutant emits a major fluorescence peak around 524 nm, confirming that the A species undergoes ESPT after electronic excitation to the S_1 state. The protein also emits in the blue at ~443 nm, which comes from the protonated form in S_1 (A*).[16,23] The notable Stokes shift (394 to 443 nm) indicates that the A* potential energy surface (PES) has a steep slope after photoexcitation, indicative of large electronic redistribution with some nuclear rearrangement even without the ESPT step. Notably, a small shoulder appears at ~543 nm in fluorescence spectrum, which could be due to vibronic coupling to a mode at ~670 cm^{-1}. An adjacent mode with imidazolinone motions is shown by ground-state DFT calculations of the chromophore (see Computational Methods below) and observed later in our tunable FSRS with the vibrational dynamics information.

Transient electronic spectra imply chromophore dynamics

In the fs-TA spectra of this GFP double mutant following 400 nm photoexcitation (Fig. 3a), two stimulated emission (SE) bands within ca. 460–600 nm and two excited state absorption (ESA) bands within ca. 430–460 nm and 630–730 nm are observed inside our detection window. We note that for the two main SE bands, particularly the one below 500 nm, contributions from ground state bleaching (GSB) are likely present as shown by the steady-state absorption profile in Fig. 2. However, the SE contribution is dominant because the 400 nm excitation wavelength is centred at A not B form. Both the ESA bands reach the intensity maximum upon 400 nm
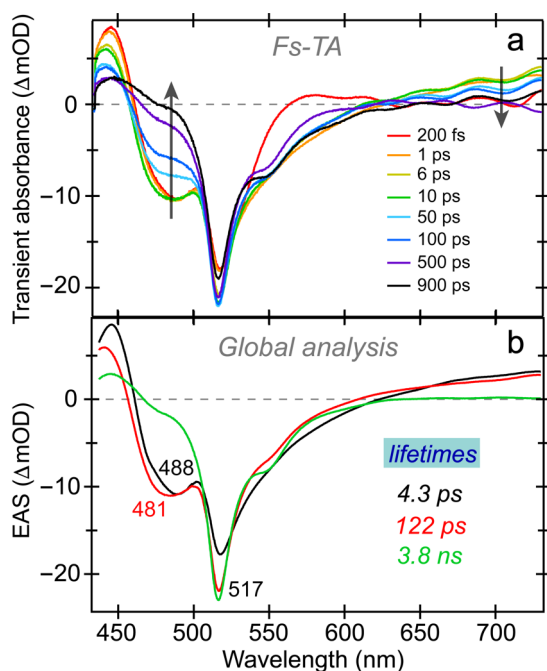
Fig. 3 Fs-TA analysis implies ESPT reaction mechanism inside the GFP-S65T/S205V mutant. (a) Fs-TA spectra at time delays from 200 fs to 900 ps following 400 nm photoexcitation. The arrows depict time evolution. (b) Evolution-associated spectra (EAS) of the GFP double mutant using a sequential kinetic scheme in global analysis. The lifetimes of three EAS components are 4.3 ps (black), 122 ps (red), and 3.8 ns (green).

photoexcitation and show decay with time, indicating their A* origin. However, the SE band intensity rises on the hundreds of ps timescale before an ensuing decrease on the ns timescale, showing its I* nature.

To unveil the hidden states and ESPT dynamics in the electronic domain, we perform global analysis using Glotaran[31,32] with a sequential model and uncover three components (Fig. 3b, see the ESI† for more details), each one termed as an evolution-associated spectrum (EAS). The earliest EAS has a major SE peak at 517 nm with a minor SE peak at 488 nm. It transforms to the second EAS with a time constant of 4.3 ps. During this time, the position of the dominant SE peak remains largely intact with a notable increase of magnitude, but the shoulder peak blueshifts to 481 nm while retaining its intensity. Since this minor SE peak position does not correspond to any of the peaks or shoulders in Fig. 2, it could be assigned to an intermediate A* species prior to the main ESPT step because the I* species after ESPT would lead to the SE peak at 517 nm. In other words, the blueshift of an intermediate A* SE peak from 488 to 481 nm on the few ps timescale suggests that the chromophore undergoes some structural changes to climb a small potential energy barrier and reach a more suitable geometry for ESPT, which could slightly blueshift the SE band of the largely protonated chromophore. The second EAS decays with a time constant of 122 ps, with the disappearance of the aforementioned 481 nm A* SE peak and the appearance of a new SE shoulder at ~543 nm in the third EAS, confirming the

ESPT progress as A* converts to I*. The third EAS decays with a time constant of 3.8 ns, attributed to fluorescence emission along with other nonradiative pathways. When comparing the normalized EAS with the fluorescence spectrum (see Fig. S1 in the ESI†), it is clear that (1) the third EAS spectrum resembles the I* fluorescence feature with a major peak and a shoulder at longer wavelength, and (2) the difference between main SE band at 517 nm and fluorescence peak at 524 nm is due to the transient nature of fs-TA probing the unrelaxed I* species.

FSRS with 800 nm Raman pump monitoring A* decay

Fig. 4a presents the time-resolved contour plot of the FSRS spectra of this GFP double mutant. The actinic pump is set at 400 nm to specifically excite the protonated chromophore (A species) and track the ESPT process. The Raman pump is chosen at 800 nm which is likely within the broad A* ESA band above 650 nm (Fig. 3) and has been shown to be a suitable wavelength to enhance the A* Raman modes of GFP and its derivatives.[23,33] Several characteristic modes of the GFP chromophore (1137, 1173, 1259 cm^{-1}) are observed which gradually decay away on the hundreds of ps timescale, confirming their A* nature. However, no concomitant I*
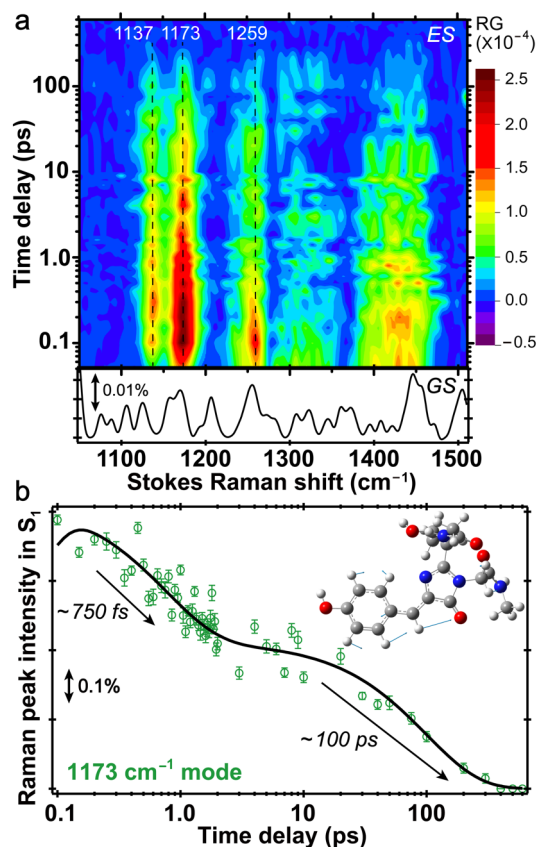
Fig. 4 Time-resolved FSRS with an 800 nm Raman pump elucidates structural evolution of A* species. (a) Semilogarithmic contour plot of Stokes FSRS of the GFP-S65T/S205V mutant after 400 nm photoexcitation. Transient A* Raman modes are labeled. Ground state (GS) spectrum is plotted below. The Raman gain (RG) is shown by the colorbar. (b) Semilogarithmic plot of the A* 1173 cm^{-1} mode intensity decay. The associated atomic motion from TD-DFT calculation is presented in the upper right inset.

Fig. 5 Ground state FSRS of GFP-S65T/S205V mutant with 550 nm (red) and 800 nm (black) Raman pump. The double-arrowed line denotes the stimulated Raman gain magnitude of 0.01%. The spectrum with 550 nm Raman pump is scaled by a factor of 0.044 for normalization to the peak at ~840 cm^{-1} using 800 nm Raman pump. The spectral changes are apparent.

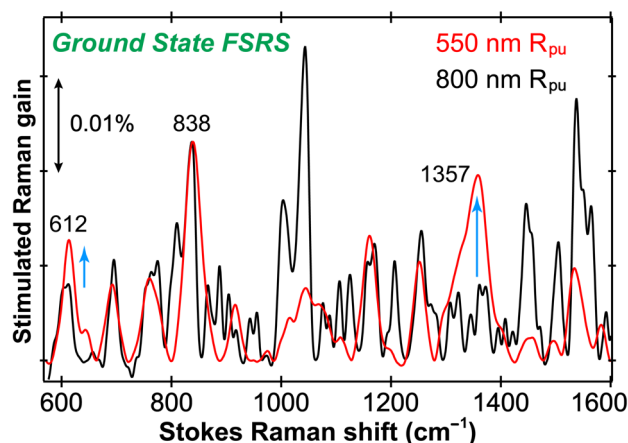modes are clearly observed in the FSRS spectra with an 800 nm Raman pump, which seems to be a common phenomenon for the Thr-Tyr-Gly (TYG) chromophore-based GFP and Ca^{2+} biosensors at this resonance condition.[33,34]

Each A* peak is fitted with a Gaussian profile to obtain their frequency and intensity information. Notably, the least-squares curve fitting with integrated intensity plot effectively filters out noise and better exposes the Raman modes.[35] We plotted the intensity dynamics of the strongest A* mode at 1173 cm^{-1} in Fig. 4b. This mode is a phenol ring-H and bridge C−H rocking motion based on our time-dependent density-functional theory (TD-DFT) calculation in Gaussian 09 software,[36] and the mode has been proven to be highly sensitive to ESPT in former studies.[33,34] The least-squares biexponential fit yields a faster decay time constant of 750 fs, which is commonly observed in GFP and the GFP-based Ca^{2+}-biosensors and could be attributed to the initial wavepackets moving out of the Franck-Condon (FC) region,[4,23,33] accompanied by some ultrafast small-scale nuclear motions likely involving several adjacent internal water molecules. The second time constant of ~100 ps is similar to the ~120 ps value (Fig. 3b) that we retrieved from fs-TA, indicative of a common ESPT origin.

Tunable FSRS with 550 nm Raman pump tracking I* rise

To further confirm the occurrence of ESPT and reveal the rise dynamics of the product I* modes, the wavelength-tunable FSRS with 550 nm Raman pump is implemented. The selection of Raman pump is based on the fs-TA spectra which exhibit a strong SE band of I* near 520 nm (Fig. 3), hence a 550 nm Raman pump achieves pre- or quasi-resonance with this SE band and enhances the Raman signals.[34,35] Notably, a Raman pump at 520 nm (right on resonance) should be avoided in this case since it could induce tilted baselines and lead to dispersive lineshapes due to interactions with lower electronic
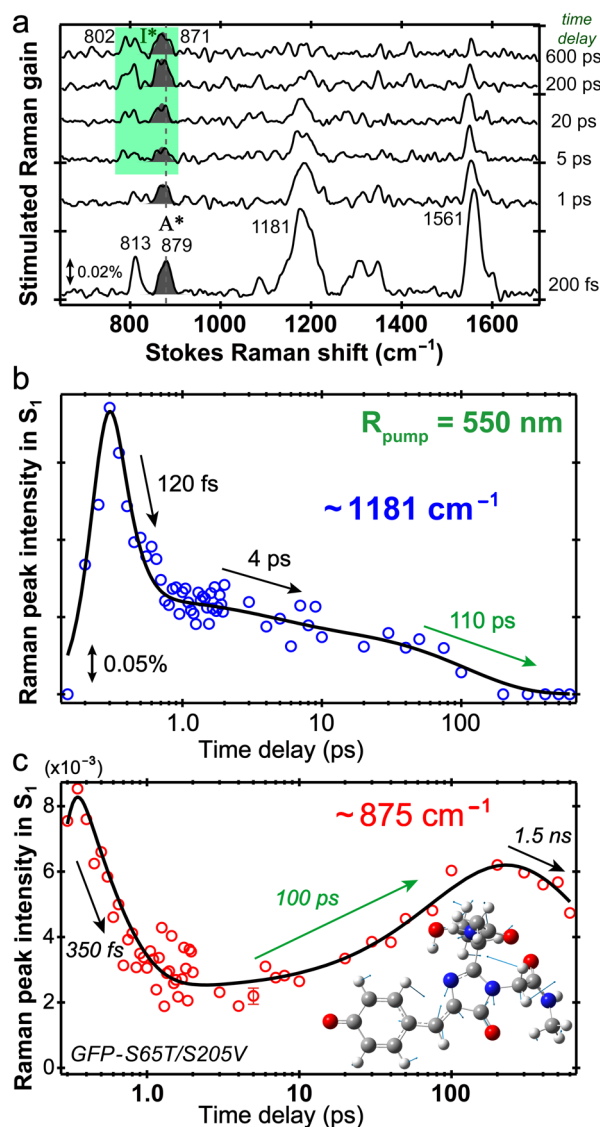

Fig. 6 Tunable FSRS with 550 nm Raman pump tracking ESPT in action. (a) Excited-state FSRS spectra at selected time delay points. The dashed line highlights the A* peak frequency which redshifts to I*. (b) and (c): Intensity dynamics of the I* mode at ~1181 and 875 cm^{-1}, respectively, with the least-squares fit time constants denoted by the arrows. In (c), the transient Raman mode intensity rise dynamics are fit after a ~5 ps dwell. The green color highlights the main ESPT process. A typical error bar of data points is shown at 5 ps.

states (e.g., the interference between vibrational coherences on the electronic excited state and ground state during the FSRS signal generation).[37]

The ground state (GS) FSRS spectrum displays notable changes as the Raman pump wavelength is tuned toward the visible range (see Fig. 5). With a 550 nm Raman pump, the GS spectrum becomes much more enhanced when compared to that with an 800 nm Raman pump. This is because the 550 nm Raman pump is much closer to the B form absorption peak of this GFP mutant at 505 nm (Fig. 2). Most Raman peaks show similar frequencies and intensity ratios but with key differences, especially the rise of the 1357 cm^{-1} mode which consists of phenolic C−O stretch with bridge C−H rocking and

some C–N stretch in the deprotonated chromophore by DFT calculation in the ground state. The ~612 cm^{-1} mode corresponds to in-plane phenol ring deformation with increased polarizability in the chromophore with partial deprotonated character.[34] These observations indicate that different Raman pump wavelengths can specifically enhance a certain species of the protein.

The excited-state Raman spectra of the GFP double mutant from tunable FSRS with 550 nm Raman pump are stack-plotted at representative time delays after 400 nm photoexcitation (Fig. 6a). It is clear that strong A* modes (e.g., 1181 and 1561 cm^{-1} modes) quickly rise and reach their maximal intensity near time zero, then gradually decay. With the decrease of A* modes, the ~802 and 875 cm^{-1} modes gradually rise on the hundreds of ps timescale with a slight redshift (Fig. 6a), indicating their identity as the product I* modes. The intensity decay of the A* mode at 1181 cm^{-1} is plotted in Fig. 6b, and according to our TD-DFT calculations of a geometrically optimized neutral TYG chromophore,[33,34] this mode mainly consists of phenol ring-H rocking and bridge-H rocking motions. Notably, this marker band frequency is blueshifted by 8 cm^{-1} from the 1173 cm^{-1} mode in FSRS with 800 nm Raman pump, which implies that the 550 nm Raman enhances a slightly different A* species which could arise from a subpopulation of A with a tighter H-bonding configuration in S_0 (thus being partially deprotonated).[34] The mode intensity decay can be fit by three exponentials with time constants (amplitude weight percentages) of ~120 fs (78%), 4 ps (9%) and 110 ps (13%) with a common time zero, indicative of three relaxation stages for the chromophore to navigate the PES starting from A* to I*. The 120 fs component can be attributed to the initial FC dynamics along the photoinduced reaction coordinate, and a similar time constant was observed for the TYG chromophore in a GFP-based calcium biosensor involving transient, small-scale proton motions toward a charge-transfer state.[33] The retrieved 4 ps may involve further proton motions (e.g., exchange with surroundings) and some heavy-atom motions close to the chromophore before reaching the main stage for ESPT. This time constant was also observed in another GFP-based Ca^{2+} biosensor with a TYG chromophore from the tunable FSRS data.[34] The 110 ps time constant could be associated with the proton hopping through the dynamic H-bonding network to E222. The retrieved second and third time constants in tunable FSRS match global analysis results from the fs-TA, which corroborates the species observed in tunable FSRS with 550 nm Raman pump being the major player undergoing ESPT reaction in this GFP double mutant.

The intensity of the ~875 cm^{-1} mode of I*, an imidazolinone ring deformation motion along with sidechain motions near the E222 site (see Fig. 6c inset from TD-DFT calculation of a deprotonated chromophore), tracks the photoproduct rise after ESPT. We note that the observation of this marker band benefits from the pre-resonance enhancement of I* species by the 550 nm Raman pump. This mode is initially at 879 cm^{-1} in the largely protonated A* state, and redshifts to 871 cm^{-1} in the deprotonated I* form. Because the corresponding mode is at ~840 cm^{-1} in S_0 (Fig. 5), this photoinduced blueshift (i.e.,

840→879 cm^{-1}) seems to be a common feature among the GFP chromophores. For example, the SYG chromophore inside a GFP-derived Ca^{2+} biosensor displays a similar magnitude of blueshift of its 855 cm^{-1} mode in S_0 to 885 cm^{-1} mode in S_1 of the protonated chromophore species after 400 nm electronic excitation.[23] However, in that study only the 800 nm Raman pump was used so we could not clearly observe the I* mode appearance in the pertinent spectral region or any frequency shift on the hundreds of ps timescale. The observed transient mode frequency redshift and intensity rise of TYG chromophore in the GFP double mutant (i.e., 879→871 cm^{-1}) is thus a consequence of gradual accumulation of further deprotonated state I* from the A* species already active at early time, effectively enhanced by the 550 nm Raman pump (see Fig. 6a).

The initial ~350 fs decay time constant in Fig. 6c is attributed to the A* mode at 879 cm^{-1}, which is in between the above-mentioned ~120 fs initial decay of the 1181 cm^{-1} A* mode with 550 nm Raman pump (Fig. 6b) and ~750 fs initial decay of the 1173 cm^{-1} A* mode with 800 nm Raman pump (Fig. 4b). This key comparison manifests the intrinsic structural inhomogeneity and mode specificity in tracking the excited state PES, and the necessity to dissect the sub-ps process in greater details with the tunable FSRS methodology. For instance, it is reasonable that a more delocalized skeletal motion at 879 cm^{-1} (Fig. 6c) exhibits a lengthened initial decay than a more localized high-frequency mode at 1181 cm^{-1} (Fig. 6b) when the pertinent A* subpopulation with deprotonated character is accessed by the 550 nm Raman pump (but not by the 800 nm Raman pump). This spectral analysis is supported by previous structural findings that the S65T mutation in GFP pocket significantly impacts the local H-bond network near the chromophore and the proton acceptor E222 exhibits at least two conformations which could help define the chromophore charge state at the opposite end.[20] If we use a minimal model to differentiate subpopulations based on the charge state at the chromophore phenolic end, then the A* subpopulation we track in Fig. 6a and 6b mainly involves the chromophore with some deprotonated character, which has a slightly redshifted absorption band (i.e., underneath the broad A absorption band near 394 nm in Fig. 2, closer to the Raman pump at 550 nm) in comparison to the more protonated chromophore.[28,34]

Furthermore, since there is no significant intensity for the I* mode before ~5 ps, we fit the transient Raman intensity rise after a short dwell of 5 ps. We uncover a ~100 ps time constant (Fig. 6c, green arrow) that matches the 100–110 ps decay time constant of an A* marker band (Figs. 4b and 6b). Therefore, the main ESPT reaction proceeds after a few-ps ultrafast preparation stage for key structural rearrangement (i.e., some heavy-atom motions) prior to transferring the chromophore phenolic proton via an optimized H-bonding chain.[4,33,38] Notably, the absence/presence of the 4 ps component in the I*/A* mode intensity rise/decay (Fig. 6c and 6b) suggests a more complex ESPT mechanism than a purely concerted, synchronous proton transfer from one end to the other side of the chromophore via an H-bonding chain (see below).

We note that a global fitting approach as employed for the fs-TA spectra (see Fig. 3) would be ideal as an additional method to perform the kinetic analysis of FSRS data.[39] However, the intrinsic difference between fs-TA and FSRS makes global analysis less useful or impactful for FSRS because its transient peak intensity not only depends on the electronic population (like in fs-TA), but also depends on the resonance conditions and Raman polarizabilities.[35,40] This is why the mode-dependent kinetic analysis typically adopted in FSRS with DFT-calculation-aided assignment has been the field-proven effective way in revealing structural dynamics on the multidimensional potential energy surface.[4,24] In other words, the independent assessment of transient vibrational mode dynamics essentially provides a powerful level ground for us to retrieve time constants for direct comparison, correlating similarities and interpreting differences, so we can gain deeper mechanistic insights into various photoactivated processes.

Spectroscopic studies infer key structural modification

It is well established that ESPT in wtGFP goes through a CRO-W22-S205-E222 proton wire with E222 as the terminal proton acceptor.[14,17,41,42] The corresponding O···O distances are 2.7 Å, 2.6 Å, 2.6 Å between each pair from the beginning to the end of the proton wire based on its crystal structure (PDB ID: 2WUR) as shown in Fig. S2a.[43] The S205V mutation leads to an alternative ESPT pathway as the CRO-Water-T203-E222 chain shows 2.7 Å, 3.5 Å, 2.6 Å O···O distances (PDB ID: 2QLE) with only one value significantly different from the wtGFP case.[16] The ESPT time constant is lengthened from ~10 ps in wtGFP to ~300 ps in the GFP-S205V mutant. This ESPT rate decrease can be explained by a large O···O distance (3.5 Å) between a bridging water molecule and the repositioned T203 residue in the GFP-S205V mutant, which does not directly involve either end of the chromophore. This result infers that the ESPT reaction in the GFP chromophore pocket is a concerted process and the "weakest" link with the longest H-bonding distance is the rate-determining step. The structural changes of these surrounding residues are typical of the S205V and S205A mutations due to steric effects and the change of hydrophobicity, which also explains the different acidity of the chromophore in the GFP-S205V mutant (pK_a>11), GFP-S65T mutant (pK_a≈6), and GFP-S65T/S205V mutant (pK_a≈7.1) in comparison to wtGFP (pK_a≈8.3).[16] Interestingly, such a structural rearrangement is also reflected by the consistent redshift of the B absorption peak position from wtGFP (475 nm), GFP-S65T (490 nm) or GFP-S205V (490 nm),[3,14,16] to GFP-S65T/S205V (505 nm, see Fig. 2).

In a related GFP-S205A mutant, the O···O distance from water to T203 is measured to be 3.2 Å based on crystal structure and the ESPT time is around 150 ps.[44] This comparison further suggests that the pertinent ESPT reaction proceeds concertedly from the chromophore phenolic side to the imidazolinone side: with a shorter O···O distance for the rate-determining step at the "middle" region of the PT wire, the overall ESPT reaction becomes accelerated with still one rate constant. In our current study with the GFP double mutant, adding an extra S65T mutation to the S205V mutant speeds up the ESPT process to ~100 ps. This result indicates that the S65T mutation may alter the position of E222 and also indirectly affect T203 and therefore shorten the O···O distance between T203 and the bridging water molecule to below 3.2 Å. Whilst a protein crystal structure is currently unavailable, the careful utilization of a spectroscopic approach to retrieve time constants may provide an alternative method to compare with literature and obtain key structural constraints with functional importance.

A concerted and stepwise hybrid ESPT reaction

Using tunable FSRS with 550 nm Raman pump, we reveal a prominent I* mode at ~871 cm^{-1} which mainly consists of the deformation of the chromophore imidazolinone ring and some sidechain motions near E222, with a characteristic decay-rise-decay pattern on the fs to hundreds of ps timescale (see Fig. 6c). Notably, a related mode at ~670 cm^{-1} can be retrieved from the vibronic progression in the fluorescence spectrum (Fig. 2), suggesting its similar origin as an imidazolinone mode of a deprotonated chromophore which is also corroborated by TD-DFT and DFT calculations in S_1 and S_0, respectively. Except for this mode, other motions which involve the phenolic ring (i.e., the starting point of the ESPT chain) are not strong in the S_1 spectra (Fig. 6a) during the main ESPT step with the ~100 ps time constant. This rather unexpected result suggests that the polarizability of the imidazolinone ring is specifically enhanced which could arise from large electron redistribution and electrostatic changes during the ESPT process. One plausible explanation is that the carboxylate of E222 receives a proton on the 100 ps timescale which affects the imidazolinone ring side, starting after a short dwell of ~5 ps as shown in Fig. 6a and 6c. This key observation further supports that the chromophore proton transfer is overall concerted in GFP, which impacts the imidazolinone ring electron and nuclear motions from very early time out of the A* FC region after actinic electronic excitation.

Besides the long time constant, we retrieve a 4 ps process from global analysis of the fs-TA spectra (Fig. 3b) as well as the intensity decay of the 1181 cm^{-1} mode in tunable FSRS (Fig. 6b), but it is not seen from the 1173 cm^{-1} mode dynamics in the conventional FSRS (Fig. 4b). This intriguing result could be rationalized by the 800 nm Raman pump in a conventional FSRS setup[33,38] focusing on the more protonated chromophore species, which is not the major contributor in this GFP double mutant because it has a large population of deprotonated species at pH=7.9 (Fig. 2) due to its aforementioned pK_a≈7.1. The 550 nm Raman pump, however, is located much closer to the deprotonated chromophore absorption peak and has proven effective in enhancing the loosely protonated or partially deprotonated A* subpopulation which contributes to the dominant transient modes after 400 nm photoexcitation. This analysis substantiates the importance of considering the intrinsic heterogeneity of the system, even within the same neutral chromophore population with a broad absorption band as seen in Fig. 2, that UV light irradiation can excite at

least two A subpopulations of the protein chromophore with distinguishable protonation states.[34]

In wtGFP, a time-resolved fluorescence study showed that the A* to I* transition is biphasic with 2.2 and 8.1 ps time constants.[3] It remains unclear about what the 2.2 ps component does during ESPT reaction, except that it likely involves the initial structural motions that facilitate efficient ESPT.[4] If the proton transfer is purely concerted, shouldn't there be only one time constant? Both of these time constants are increased upon deuteration but with a different kinetic isotope effect (KIE) of 3.7 for the first process and 5.7 for the second one.[3] The difference in KIE indicates that they could be due to two distinct processes. To assign the origin of the initial process, we consider the following experimental evidences. First, 2.2 ps is close to the 4 ps time constant in this GFP double mutant. Second, the O···O distance between CRO and W22 is 2.7 Å in wtGFP which remains unchanged in the GFP-S205V mutant (see Fig. S2, it is safe to assume that this distance still holds in the GFP-S205V/S65T mutant since the extra S65T mutation is on the far end of the H-bonding chain), so the PT process that occurs between CRO and an adjacent water molecule should have similar time constants. Third, the I* 871 cm^{-1} mode intensity in tunable FSRS changes only slightly on the few ps timescale, showing an insignificant accumulation of I* species before a dwell of ~5 ps (see Fig. 6c).

Notably, these discussion points suggest that the 2.2 ps and 4 ps components could be attributed to some small-scale proton motions (confined between CRO and internal water at the phenolic end) in wtGFP and the GFP-S205V/S65T mutant, respectively. A similar time constant is observed in the ESPT process of a four-ring photoacid pyranine in water, which has been attributed to the contact ion pair formation prior to significant proton separation.[35,45] The relatively smaller KIE (i.e., 3.7<5.7) also implies that more heavy-atom motions occur during the first process which could well be the aforementioned ultrafast preparation stage as an integral part of ESPT, witnessing a substantial reconfiguration of the chromophore and its surrounding protein environment. The second time constant is associated with the large-scale concerted proton transfer through a well-established PT wire on the ~10 Å length scale, which coincides with a KIE of ~6 that is much closer to the symmetric limit of ~7 with the theoretical vibrational zero-point energy difference between adiabatic states without significant fluctuations of the pertinent reactant and product PESs.[18,38,46] This temporal component is increased by >30-fold when the middle link (O···O distance) of the PT chain is lengthened to 3.5 Å in the GFP-S205V mutant (Fig. S2b) which exhibits an overall KIE of ~5.[16] Jointly, these results reveal that although the large-scale proton transfer proceeds in a concerted manner in our GFP-S65T/S205V double mutant impacting both ends of the chromophore on the ~100 ps timescale, there is a previously unknown stage involving small-scale proton motions and some heavy-atom motions largely confined to the phenol-ring side on the ~4 ps timescale. The likely occurrence of an intermediate state between the initial ultrafast relaxation of the excited state chromophore-surrounding PES and the long-scale PT process

represents a significant new piece of information, which sheds more light on the physical chemistry processes governing the fluorescence tuning in GFP derivatives and other analogous autofluorescent proteins.[47,48]

Furthermore, the absence of this ~4 ps intermediate component in the 879 cm^{-1} A* mode after photoexcitation (see Fig. 6a and 6c) indicates that the initial "stepwise" phase of ESPT reaction is indeed largely concentrated at the phenolic end of the TYG chromophore in the GFP-S65T/S205V double mutant. Meanwhile as we stated above, the absence of this ~4 ps component in the 1173 cm^{-1} A* mode using 800 nm Raman pump (Fig. 4b) indicates the specificity of Raman pump in enhancing various subpopulations of A*, which could be more protonated at the chromophore phenolic end than the A* subpopulation enhanced by the 550 nm Raman pump (i.e., in the visible spectral region). Put another way, the initial ground state subpopulations being probed by the Raman pump with different wavelengths could be different as shown in Fig. 5; but after actinic photoexcitation, we expect that the small-scale coherent proton motions on the ~120 fs timescale leads to more deprotonated chromophores in all the A* subpopulations.[27,33] The 550 nm Raman pump thus provides higher sensitivity to the subsequent vibrational dynamics of A* decay due to the increased electric polarizability over a conjugated chromophore ring system, corroborated by the resonance Raman enhancement effect (see relevant SE band positions in Fig. 3).

This added resolving power of tunable FSRS allows us to capture the characteristic ESPT reaction phase when the chromophore proton further moves along the H-bonding chain in the immediate vicinity of the phenolic end surrounded by internal water molecules and protein pocket residues. Such a mechanistic understanding of initial structural events prior to the main ESPT process with clearly distinguishable time constants (ca. 4 and 100 ps) supports a stepwise-concerted-hybrid ESPT model along the PT wire inside a protein.

Our main findings point to the crucial importance of a site-specific alteration with a holistic view of the proton wire in enhancing ESPT in biomolecules, ideally with an improved fluorescence quantum yield.[16,28] The entire chromophore structure needs to be considered in the H-bonding network beyond the immediate vicinity of the proton dissociation site to predict the functional motions leading to the main ESPT event, with sufficient time resolution to resolve the sequence of distinct molecular events. This stepwise-concerted-hybrid ESPT mechanism has now been shown for chromophore populations with slightly different protonation states, which suggests its general importance for a multi-link, dynamic, and efficient PT wire in biomolecules. Inherently, the protein adopts such an ESPT pathway to avoid the high-energy intermediates and effectively transfer a proton with the energy provided by photoexcitation.[25,35,42] Based on this knowledge, potential new strategies targeting an expanded group of atomic sites from the chromophore ring system to protein matrix in controlling ESPT efficiency and rates can be further developed.

Experimental and computational methods

Protein sample preparation and steady-state spectroscopy

The site-specific S65T/S205V double mutation was introduced into wtGFP using the QuikChange method (Stratagene), and the protein expression and subsequent purification followed previously established procedures.[16] The final protein aqueous buffer solution with 50 mM HEPES (4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid) has a pH of 7.9. The steady-state electronic absorption and emission spectrum is collected by the Thermo Scientific Evolution 201 UV-Vis and Hitachi F-2500 fluorescence spectrophotometer, respectively.

Femtosecond transient absorption and stimulated Raman methods

Our experimental fs-TA and FSRS setups have been reported.[33,35,49,50] In brief, a ~35 fs, 4 mJ, 800 nm fundamental pulse train with 1 kHz repetition rate is produced from a regenerative laser amplifier (Legend Elite-USP-1K-HE, Coherent) seeded by a fs mode-locked Ti:sapphire oscillator (Mantis-5, Coherent). The ~400 μW, 400 nm actinic pump is generated by doubling the frequency of a portion of the fundamental pulse in a 0.3-mm-thick type-I β-barium borate (BBO) crystal. For fs-TA measurements, we generated the broadband white light probe (ca. 420–750 nm) in a 2-mm-thick quartz cell filled with deionized water.[37] In conventional FSRS, a ~3 ps, ~4 mW, 800 nm Raman pump is achieved by a home-built spectral filter.[38,51] A ~35 fs, ~100 nJ/pulse supercontinuum white light (ca. 840–960 nm) acts as the Raman probe which is generated in a 2-mm-thick Z-cut sapphire crystal plate, followed by fused silica prism-pair compression. In tunable FSRS, the 550 nm Raman pump is produced by a home-built three-component system:[50] (1) a ~2 ps, 400 nm pump by a home-built second harmonic bandwidth compressor (SHBC); (2) a tunable ps seed pulse by a fs noncollinear optical parametric amplifier (NOPA) with a grating-slit-based spectral filter; and (3) amplification of the ps seed with the ps pump through a two-stage ps NOPA.

For the exited-state FSRS data collection, the three beams are focused on the protein sample solution (OD of ~0.4/mm at 400 nm), which is continuously stirred by a home-made magnetic stir bar in a 1-mm-pathlength quartz cell to minimize thermal effects induced by the pulsed laser irradiation. Notably, the OD of this sample is lower than a normal OD=1/mm sample, which is one of the reasons why the signal is not very strong in excited state FSRS (see Figs. 4 and 6).[23,35] The actinic pump is controlled by a motorized translational stage and the Raman pump and probe pair arrives at the sample spot simultaneously. After the sample, only the probe and the collinear FSRS scattering signal are focused into a Princeton Instruments SP2356 spectrograph (600 grooves/mm and 1000 nm blaze for conventional FSRS, whereas 1200 grooves/mm and 500 nm blaze for tunable FSRS) and imaged onto a front-illuminated 1340×100 CCD camera array (PIXIS 100F, Princeton Instruments) synchronized with the 1 kHz laser repetition rate. At least ten sets of excited-state FSRS data in the Stokes region (i.e., the Raman probe is redder than the Raman pump) are collected at each time delay point between −5 and 650 ps. During each set, eight ground-state FSRS spectra are also measured periodically to monitor the sample condition and laser stability, providing an *in situ* standard to be subtracted from the excited-state signal so the time-resolved difference spectra can be reliably obtained (see Figs. 4a and 6a).

DFT and TD-DFT calculations

To retrieve the vibrational normal modes of the chromophore in the electronic ground state (S_0) and first singlet excited state (S_1), quantum chemical calculations of the neutral and anionic TYG chromophore are performed in Gaussian 09.[36] For Raman mode frequencies, we perform the density-functional theory (DFT) calculation of an optimized protonated or deprotonated TYG chromophore capped with methyl groups using RB3LYP 6-31G+(d,p) basis sets *in vacuo*. To obtain vibrational frequencies of a geometrically optimized, protonated or deprotonated TYG chromophore in S_1, a time-dependent DFT (TD-DFT) calculation with the same basis sets are carried out. The calculated normal mode frequencies are all scaled with a factor[34] of 0.97 and 0.96 to compare with experimental values in S_0 and S_1, respectively. Due to deviations from realistic environments with a partially deprotonated chromophore and extensive H-bonding network, the frequency exact match is not expected.[23,34] For instance, the two representative excited-state vibrational motions depicted in Figs. 4b and 6c are respectively calculated at ca. 1210 and 880 cm^{-1} after frequency scaling of the TD-DFT results.

Conclusions

By combining fs-TA and wavelength-tunable FSRS methods, the main photoinduced proton transfer time constant inside a GFP-S65T/S205V double mutant is retrieved to be ~110 ps. Spectroscopic results reveal that the PT wire is significantly altered from wtGFP and slightly modified in comparison with the GFP-S205V mutant. The O···O distance between a bridging water molecule and T203 is expected to be shorter than 3.2 Å in this GFP double mutant. The observation of the 871 cm^{-1} marker band of the emerging I* species in excited-state FSRS is achieved by a strategically tuned picosecond Raman pump at 550 nm with high selectivity. Based on a proposed correlation between the ESPT time constant and the H-bonding distance in the proton wire, the real-time tracking of this imidazolinone ring mode with one dominant ~100 ps rise time further proves that the proton transfer proceeds in a concerted manner from the chromophore to E222 (hence the promptly enhanced A* and I* modes after 400 nm photoexcitation, and the notable growth of the 871 cm^{-1} I* mode in the excited-state FSRS data).

Furthermore, a ~4 ps process is uncovered from global analysis of the fs-TA spectra and the decay dynamics of the A* 1181 cm^{-1} mode in tunable FSRS prior to main ESPT event (i.e., stepwise reaction on the few-ps timescale). This initial process

most likely involves small-scale proton motions between the chromophore phenolic end and an adjacent water molecule along with some surrounding residue structural motions, which set up the stage for the subsequent concerted proton transfer (on the ~100 ps timescale) and efficiently release the sudden influx of photoexcitation energy to guide the protein toward its fluorescent state. These observations enrich our understanding of the ESPT mechanism in close correlation with the H-bonding wire modification in the GFP pocket, which can aid the rational design and development of novel fluorescent proteins and other fluorescence-based bioimaging probes in the future.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

Notes and references

1   M. Eigen, *Angew. Chem. Int. Ed.*, 1964, **3**, 1-72.
2   A. Douhal, S. K. Kim and A. H. Zewail, *Nature*, 1995, **378**, 260-263.
3   M. Chattoraj, B. A. King, G. U. Bublitz and S. G. Boxer, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 8362-8367.
4   C. Fang, R. R. Frontiera, R. Tran and R. A. Mathies, *Nature*, 2009, **462**, 200-204.
5   J. N. Henderson, M. F. Osborn, N. Koon, R. Gepshtein, D. Huppert and S. J. Remington, *J. Am. Chem. Soc.*, 2009, **131**, 13212-13213.
6   Y. Zhang, K. de La Harpe, A. A. Beckstead, R. Improta and B. Kohler, *J. Am. Chem. Soc.*, 2015, **137**, 7059-7062.
7   O. Shimomura, F. H. Johnson and Y. Saiga, *J. Cell. Comp. Physiol.*, 1962, **59**, 223-239.
8   M. Chalfie, Y. Tu, G. Euskirchen, W. Ward and D. Prasher, *Science*, 1994, **263**, 802-805.
9   M. Ormö, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien and S. J. Remington, *Science*, 1996, **273**, 1392-1395.
10   R. Y. Tsien, *Annu. Rev. Biochem.*, 1998, **67**, 509-544.
11   J. K. M. Sanders and S. E. Jackson, *Chem. Soc. Rev.*, 2009, **38**, 2821-2822.
12   R. Heim, D. C. Prasher and R. Y. Tsien, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 12501-12504.
13   H. Lossau, A. Kummer, R. Heinecke, F. Pöllinger-Dammer, C. Kompa, G. Bieser, T. Jonsson, C. M. Silva, M. M. Yang, D. C. Youvan and M. E. Michel-Beyerle, *Chem. Phys.*, 1996, **213**, 1-16.
14   K. Brejc, T. K. Sixma, P. A. Kitts, S. R. Kain, R. Y. Tsien, M. Ormö and S. J. Remington, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 2306-2311.
15   G. J. Palm, A. Zdanov, G. A. Gaitanaris, R. Stauber, G. N. Pavlakis and A. Wlodawer, *Nat. Struct. Mol. Biol.*, 1997, **4**, 361-365.
16   X. Shu, P. Leiderman, R. Gepshtein, N. R. Smith, K. Kallio, D. Huppert and S. J. Remington, *Protein Sci.*, 2007, **16**, 2703-2710.
17   R. Simkovitch, A. Huppert, D. Huppert, S. J. Remington and Y. Miller, *J. Phys. Chem. B*, 2013, **117**, 11921-11931.
18   S. P. Laptenok, A. Lukacs, A. Gil, R. Brust, I. V. Sazanovich, G. M. Greetham, P. J. Tonge and S. R. Meech, *Angew. Chem. Int. Ed.*, 2015, **54**, 9303-9307.
19   A. Royant and M. Noirclerc-Savoye, *J. Struct. Biol.*, 2011, **174**, 385-390.

20   J. A. J. Arpino, P. J. Rizkallah and D. D. Jones, *PLOS ONE*, 2012, **7**, e47132.
21   R. Berera, R. van Grondelle and J. T. M. Kennis, *Photosynth. Res.*, 2009, **101**, 105-118.
22   P. Kukura, D. W. McCamant and R. A. Mathies, *Annu. Rev. Phys. Chem.*, 2007, **58**, 461-488.
23   B. G. Oscar, W. Liu, Y. Zhao, L. Tang, Y. Wang, R. E. Campbell and C. Fang, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 10191-10196.
24   D. R. Dietze and R. A. Mathies, *ChemPhysChem*, 2016, **17**, 1224-1251.
25   D. P. Hoffman and R. A. Mathies, *Acc. Chem. Res.*, 2016, **49**, 616-625.
26   J. T. M. Kennis, D. S. Larsen, I. H. M. van Stokkum, M. Vengris, J. J. van Thor and R. van Grondelle, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 17988-17993.
27   Y. Wang, L. Tang, W. Liu, Y. Zhao, B. G. Oscar, R. E. Campbell and C. Fang, *J. Phys. Chem. B*, 2015, **119**, 2204-2218.
28   S. R. Tachibana, L. Tang, Y. Wang, L. Zhu, W. Liu and C. Fang, *Phys. Chem. Chem. Phys.*, 2017, **19**, 7138-7146.
29   L. Tang, Y. Wang, W. Liu, Y. Zhao, R. E. Campbell and C. Fang, *J. Phys. Chem. B*, 2017, **121**, 3016-3023.
30   R. Heim, A. B. Cubitt and R. Y. Tsien, *Nature*, 1995, **373**, 663-664.
31   J. J. Snellenburg, S. Laptenok, R. Seger, K. M. Mullen and I. H. M. van Stokkum, *J. Stat. Softw.*, 2012, **49**, 1-22.
32   C. Chen, W. Liu, M. S. Baranov, N. S. Baleeva, I. V. Yampolsky, L. Zhu, Y. Wang, A. Shamir, K. M. Solntsev and C. Fang, *J. Phys. Chem. Lett.*, 2017, **8**, 5921–5928.
33   L. Tang, W. Liu, Y. Wang, Y. Zhao, B. G. Oscar, R. E. Campbell and C. Fang, *Chem. Eur. J.*, 2015, **21**, 6481-6490.
34   L. Tang, W. Liu, Y. Wang, L. Zhu, F. Han and C. Fang, *J. Phys. Chem. Lett.*, 2016, **7**, 1225-1230.
35   W. Liu, Y. Wang, L. Tang, B. G. Oscar, L. Zhu and C. Fang, *Chem. Sci.*, 2016, **7**, 5484-5494.
36   M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision B.1*, Gaussian, Inc., Wallingford, CT, 2009.
37   B. G. Oscar, C. Chen, W. Liu, L. Zhu and C. Fang, *J. Phys. Chem. A*, 2017, **121**, 5428-5441.
38   F. Han, W. Liu and C. Fang, *Chem. Phys.*, 2013, **422**, 204-219.
39   C. R. Hall, J. Conyard, I. A. Heisler, G. Jones, J. Frost, W. R. Browne, B. L. Feringa and S. R. Meech, *J. Am. Chem. Soc.*, 2017, **139**, 7408-7414.
40   W. Liu, L. Tang, B. G. Oscar, Y. Wang, C. Chen and C. Fang, *J. Phys. Chem. Lett.*, 2017, **8**, 997–1003.
41   D. Stoner-Ma, A. A. Jaye, P. Matousek, M. Towrie, S. R. Meech and P. J. Tonge, *J. Am. Chem. Soc.*, 2005, **127**, 2864-2865.
42   S. R. Meech, *Chem. Soc. Rev.*, 2009, **38**, 2922-2934.
43   A. Shinobu, G. J. Palm, A. J. Schierbeek and N. Agmon, *J. Am. Chem. Soc.*, 2010, **132**, 11093-11102.
44   Y. Erez, R. Gepshtein, I. Presiado, K. Trujillo, K. Kallio, S. J. Remington and D. Huppert, *J. Phys. Chem. B*, 2011, **115**, 11776-11785.
45   R. Gepshtein, P. Leiderman, L. Genosar and D. Huppert, *J. Phys. Chem. A*, 2005, **109**, 9674-9684.
46   N. Agmon, *Chem. Phys. Lett.*, 1995, **244**, 456-462.
47   M. Z. Lin, M. R. McKeown, H.-L. Ng, T. A. Aguilera, N. C. Shaner, R. E. Campbell, S. R. Adams, L. A. Gross, W. Ma, T. Alber and R. Y. Tsien, *Chem. Biol.*, 2009, **16**, 1169-1179.
48   I. M. Dobbie, N. F. Lowndes and K. F. Sullivan, Autofluorescent Proteins, in *Methods in Cell Biology*, ed. K. F. Sullivan, Academic Press, 2008, vol. 85, pp. 1-22.

49   W. Wang, W. Liu, I.-Y. Chang, L. A. Wills, L. N. Zakharov, S. W. Boettcher, P. H.-Y. Cheong, C. Fang and D. A. Keszler, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 18397-18401.
50   L. Zhu, W. Liu and C. Fang, *Appl. Phys. Lett.*, 2014, **105**, 041106.
51   D. W. McCamant, P. Kukura, S. Yoon and R. A. Mathies, *Rev. Sci. Instrum.*, 2004, **75**, 4971-4980.

Please do not adjust margins