PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

Hidden Photoinduced Reactivity of the Blue Fluorescent Protein mKalama1

Russell B. Vegh,[‡a,b,c] Dmitry A. Bloch,[‡d] Andreas S. Bommarius,[a,b,c] Michael Verkhovsky,[d] Sergei Pletnev,[e,f] Hideo Iwaï,[d] Anastasia V. Bochenkova,*[g,h] and Kyril M. Solntsev*[a]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
DOI: 10.1039/b000000x

Understanding the photoinduced dynamics of Fluorescent Proteins is essential for their applications in bioimaging. Despite numerous studies on the ultrafast dynamics, the delayed response of these proteins, which often results in population of kinetically trapped dark states of various origins, is largely unexplored. Here, by using transient absorption spectroscopy spanning the time scale from picoseconds to seconds, we reveal a hidden reactivity of the bright blue-light emitting protein mKalama1 previously thought to be inert. This protein shows no excited-state proton transfer during its nanosecond excited-state lifetime; however, its tyrosine-based chromophore undergoes deprotonation coupled to non-radiative electronic relaxation. Such deprotonation causes distinct optical absorption changes in the broad UV-to-NIR spectral range (*ca.* 300 - 800 nm); the disappearance of the transient absorption signal has a complex nature and spans the whole microsecond-to-second time scale. The mechanisms underlying the relaxation kinetics are disclosed based on the X-ray structural analysis of mKalama1 and the high-level electronic structure calculations of proposed intermediates in the photocycle. We conclude that the non-radiative excited-state decay includes two major branches: internal conversion coupled to intraprotein proton transfer, where a conserved residue E222 serves as the proton acceptor; and ionization induced by two consecutive resonant absorption events, followed by deprotonation of the chromophore radical cation to bulk solvent through a novel water-mediated proton-wire pathway. Our findings open up new perspectives on the dynamics of fluorescent proteins as tracked by its optical transient absorption in the time domain extending up to seconds.

Introduction

The Green Fluorescent Protein (GFP), widely known due to its ubiquitous use as an *in vitro* and *in vivo* fluorescent marker, is an 11-stranded β-barrel with an α-helix running through the center. The chromophore (Cro) is midway of the α-helix and is formed by three amino acids S65, Y66, and G67 through an autocatalytic post-translational cyclization, only requiring molecular oxygen.[1] Wild-type GFP (wtGFP) has two main absorption maxima at 398 nm (band A) and 478 nm (band B) corresponding to the Cro and its conjugate base, respectively. Excitation of band A simultaneously yields a weak 460 nm and a strong 508 nm emission. The strong fluorescence at 508 nm occurs through an excited-state proton transfer (ESPT) mechanism, where the Cro becomes deprotonated at the Y66 phenol group in the excited state, resulting in a ~50 nm red shift of the emission band.[2,3] It has been established that, in wtGFP, the proton is translocated via a major, well-defined Cro-Water22-S205-E222 proton-wire pathway.[4,5] The existence of a less efficient, alternative pathway that involves a nearby T203 residue has later been found.[6] Finally, even more extended proton wires that connect the Cro to the bulk are also suggested.[7] According to this scenario, the protein may function as a light-driven proton pump, expelling the proton to bulk solvent via a switchable T203 exit pathway upon excitation, while acquiring the proton back in the ground state from E222, which is, in turn, protonated from the outside environment via an E5 entry pathway located near the N-terminus of GFP.

Ultrafast spectroscopy techniques have been used to track proton transfer events along the major pathway in the GFP photocycle, occurring on ps and sub-ns time scales in the excited[2,5,8] and ground[9] states, respectively. Multiple intermediates have been revealed in both states,[9,10] also located and characterized by quantum chemistry calculations.[11]

Mutations of wtGFP have resulted in a plethora of fluorescent proteins with various photophysical and photochemical properties, thus tailoring them to new applications. In particular, blue fluorescent proteins (BFPs) are essential in multilabel/multicolor imaging schemes.[12,13] BFPs reported in the literature are usually based on histidine or phenylalanine chromophores, thus enforcing neutrality of the Cro.[14,15] The BFP variants with an original tyrosine-derived chromophore have been discovered through the directed evolution studies of enhanced

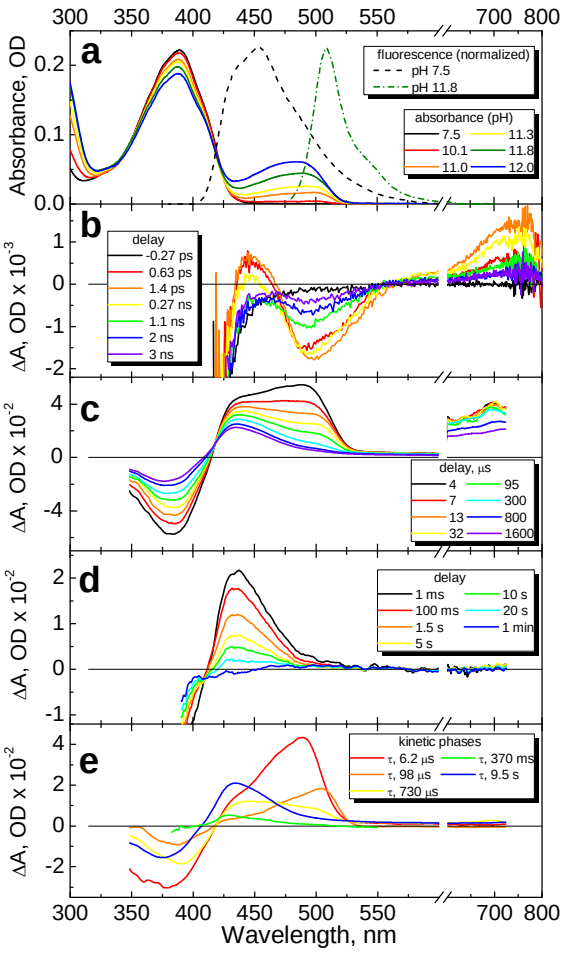

Fig. 1 Optical properties of mKalama1. (a) Absorption and emission spectra of mKalama1 at various pH. Buffer: CCPBS (20 mM CAPS, 20 mM CABS, 20 mM potassium phosphate, 200 mM NaCl); pH adjusted with KOH. (b) Transient absorption of mKalama1 at pH 7.5 after excitation with 100 fs 355 nm laser pulse. Note that the 445 nm band heavily overlaps with the strong negative stimulated emission signal, thus significantly lowering the observed amplitude of this band. (c) Transient absorption of mKalama1 at pH 7.5 after excitation with 4 ns 355 nm laser pulse. The 600-750 nm part is magnified by a factor of 10 for better visualization. (d) Same as (c), but monitored in the 1 ms - 1 min time window. (e) Results of the best multi-exponential decay fit of the data from graph (c). Buffer: PBS (pH 7.5). In (b-d), the delay after the laser flash instance is indicated for each curve.

GFP (EGFP) by introducing mutations in close proximity to the Cro that are capable of better stabilizing the neutral form in the ground state and blocking ESPT.[12] A series of 24 mutations relative to EGFP results in the monomeric, blue-light emitting protein mKalama1[12] with intrinsically high brightness. mKalama1 ($\lambda_{ex/em}$= 385/456 nm, fluorescent quantum yield φ = 0.45)[12], as well as a double mutant GFP T203V/S205V ($\lambda_{ex/em}$= 390/459 nm, φ = 0.29)[6] are the variants of wtGFP, in which both the main (via S205) and the alternative (via T203) proton wires are blocked. Therefore, the bright single-band fluorescence of these proteins belongs to the Cro neutral form in the most stable, planar *Z* configuration, with the imidazole heterocyclic N atom in the *cis* position towards the phenol ring, Z^N.[4] It is worth noting, that a double mutant GFP T203V/S205A still exhibits ESPT, though with a much slower rate and small quantum efficiency, resulting in two emission bands, as observed recently.[16] The X-ray structure of GFP T203V/S205A does not reveal a viable pathway for the proton transfer; however, the molecular dynamics studies show that ESPT in this mutant may occur through formation of transient proton-wire pathways formed by water molecules infrequently entering the protein structure from the bulk.[16]

Numerous studies have been conducted to investigate the dynamics in fluorescent proteins (FPs); however, measurements in the microsecond to second time domain are rarely reported and mostly involve the Fluorescence Correlation Spectroscopy (FCS),[17-21] single-wavelength phosphorescence,[22,23] transient absorption[24,25] and photobleaching[26]. Processes with relaxation times in this range have been ascribed to protonation and/or conformational dynamics of the GFP Cro.[18] However, no detailed spectral information about the nature of transients in the μs-to-second time domain is available and the mechanisms still remain elusive. Recently, we have performed first broadband transient absorption measurements on a series of red FPs in the extended time domain to identify long-lived dark intermediates responsible for their phototoxicity.[27] Along with a negative impact of the non-radiative excited-state decay, transient absorption of dark states can efficiently be used to optically deplete these kinetically trapped intermediates and hence to modulate steady-state emission of bright states in FPs, most notably in BFPs.[28] This idea has successfully been employed in optically modulated fluorescence bioimaging, which enables to visualize FPs in high autofluorescent background.[29]

In this work, using a combination of time-resolved spectroscopy, X-ray crystallography, and electronic structure theory, we study the photochemical and photophysical properties of the BFP mKalama1. Although optimized for brightness and photostability, this protein surprisingly shows a hidden reactivity. We unveil the broadband transient absorption of mKalama1 on the time scale ranging from picoseconds to seconds, and give a full theoretical account of spectroscopically identified intermediates in the photocycle. We show that the Cro can undergo a photoinduced delayed intraprotein deprotonation coupled to internal conversion (IC), even in the absence of viable proton-wire pathways connecting the hydroxyl group of the Cro and E222. Our experimental and theoretical results also support the existence of an alternative, novel proton transfer pathway[16] that leads from the Cro to the bulk solvent through the well-defined water chain. In mKalama1, this pathway plays a negligible role in ESPT; however, it appears to be very efficient in proton transfer from the radical cation Cro formed upon consecutive two-photon ionization.

Materials and methods

Materials and methods are described in detail in ESI. It includes the details of (a) protein expression and purification, (b) crystallization, data collection, structure solution and refinement, (c) transient absorption and (d) fluorescence spectroscopy measurements, (e) H/D kinetic isotope effect measurements, (f)

aerobic and anaerobic measurements, (g) protein structure simulations, and (h) multi-state multi-reference perturbation theory XMCQDPT2[30]/EFP[31] multi-scale excited-state calculations within the Firefly quantum chemistry package[32].

Results

Transient fluorescence and transient absorption spectroscopy

The absorption and emission spectra of mKalama1 at neutral pH have almost structureless peaks with maxima at 390 and 455 nm, respectively, which are close to the earlier reported values (Fig. 1a).[12] Upon excitation with a 372 nm diode, the fluorescence at 460 nm is found to decay biexponentially, with lifetimes of 0.8 ns and 1.9 ns and an amplitude ratio of 0.25 (ESI, Fig. S1). To check whether there is a contribution from the anionic emission to the long-wavelength part of the broad spectrum of mKalama1 at neutral pH, we have also measured the fluorescence decay at 510 nm upon excitation at 372 nm, observing the same lifetimes. If mKalama1 exhibited ESPT, one would expect to see a rise time component in the fluorescence kinetics at this wavelength, as well as an increase in the decay time, since the lifetime of the anion is longer (3.9 ns, as determined at pH 11.25, ESI, Fig. S2) than that of the neutral (0.8 ns and 1.9 ns). The absence of both a rise time component and an increase of the decay time[33] in the fluorescence kinetics of mKalama1 at 510 nm indicates that there is no anion emission and hence no ESPT.

The subpicosecond transient absorption (TA) spectra, acquired at pH 7.5, are shown in Fig. 1b. The transient decays measured at different wavelengths have the same lifetime (τ = 1.6 ns) for all spectral features, which is close to the fluorescence lifetimes in mKalama1 (ESI, Fig. S1). Therefore, we assign the time evolution of these relatively simple spectral bands at 740 nm and 445 nm to the decay of the singlet excited-state population of the neutral Cro ($^*Z^N$). Remarkably, the fs-ns transient absorption spectra of mKalama1 are very close to that of *p*-HOBDI[34] (*p*-hydroxybenzylidene-imidazolinone, the synthetic Cro of wtGFP and mKalama1), but the S_1 lifetime of the latter is three orders of magnitude shorter due to fast photoisomerization-induced deactivation.[35]

Increasing pH results in a slight decrease of the major absorption band of the ground-state Cro at 390 nm concomitant with the appearance of a green band ($\lambda_{max} \cong$ 480-505 nm) belonging to the Cro anion, presumably also in the *Z*-configuration (Z^A) (Fig. 1a). The majority of the high pH-induced changes are fully reversible (ESI, Fig. S3), suggesting a lack of significant denaturation of the protein at high pH. Upon increasing pH, a blue shift of the anionic peak is also observed. Interestingly, the apparent pK_a value of the Cro in the protein is unusually high (>12), much higher than those of any of the HOBDI derivatives. The excitation and emission spectra of mKalama1 acquired at pH 11.8 demonstrate that the anionic form is also fluorescent with the excitation and emission maxima at 490 and 509 nm, respectively (Fig. 1a and ESI, Fig. S4).

These data show no deviations from the expected behavior of mKalama1 as a brightly fluorescent, inert BFP. However, the microsecond-to-second TA measurements surprisingly reveal a hidden photoinduced reactivity of this protein. The transient signal in the 1 µs – 2 ms time window consists of the ground-state bleach at 385 nm and rise in color at the 430 nm and 490 nm bands (Fig. 1c). Also, a very weak but detectable broad band is observed at around 700 nm (Fig. 1c, inset and ESI, Fig. S5). Upon time evolution, the intensity of the 430 nm peak decreases insignificantly, while the 490 nm band exhibits a slight bathochromic shift and disappears. The remaining peak at 430 nm is monitored separately in the 1 ms - 1 min time window (Fig. 1d), where it disappears at the end of this time interval without any change of its spectral shape. The best multiexponential fit of the experimental time-wavelength-absorption surfaces gives three major spectral components with submillisecond time constants τ_1 = 6±2 µs, τ_2 = 98±10 µs, and τ_3 = 730±100 µs; a minor component with τ = 370 ms, and a very long-lived component with τ > 10 s (Fig. 1e). All submillisecond components have an isosbestic point at 418±1 nm; the two faster components have sharp peaks at 489 nm (6 µs) and 505 nm (98 µs), respectively, both contributing largely to the ~490 nm band of the photolysis (light-*minus*-dark) difference spectrum. The selected kinetic traces and residuals are presented in Fig. 2. The data are also fit using the algorithm for a sequential reaction model,[19,20,36,37] implying *n* consecutive irreversible steps. The sequential reaction fit yields similar spectral states as the global multiexponential fit, when the slower components are subtracted from the original data matrix (*n* = 3, ESI, Fig. S6). The TA kinetics also reveals a pH-sensitivity of the second component (ESI, Fig. S7).

The τ = 370 ms and τ > 10 s components have almost identical spectra with an isosbestic point at 407 nm and a bright maximum at 430 nm. To rule out diffusion in the disappearance of the long-lived component (τ > 10 s), the sample was illuminated with 30 sequential flashes upon stirring, and the spectra were recorded at 0, 16, 40, 85, and 170 min after the 30 flashes (ESI, Fig. S8-S10). This allowed us to produce enough of the long-lived component, which was seen in the TA spectra after stirring, and, therefore,

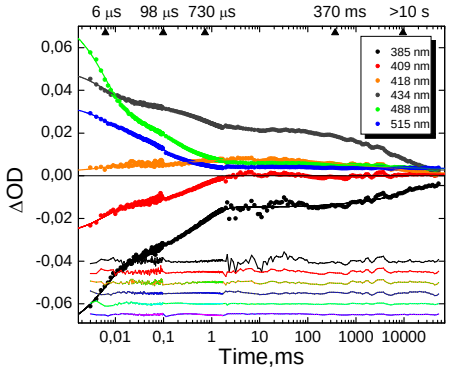

Fig. 2 Absorbance decays (dots) with kinetic fits (solid lines) at various wavelengths. The residuals are displayed at the bottom of the graph with different vertical offsets. The time constants of the exponential decay components are shown at the top. The kinetic transients presented are the slices of the experimental time-wavelength-absorption matrices at certain wavelengths, as indicated. The matrices are combined from the ANDOR and Basler data, collected separately. The conditions are those as for Fig. 1c-e. For further details, see Materials and Methods (section c) in ESI.

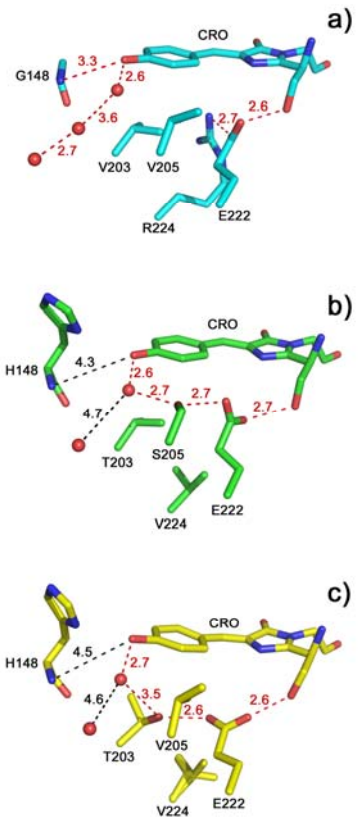

Fig. 3 Proton transfer pathways in (a) mKalama1 (PDB ID, 4ORN, this work), (b) wtGFP (PDB ID, 1GFL)[38], and (c) GFP S205V (PDB ID, 2QLE)[6]. The red dashed lines show the H-bonds; the black dashed lines indicate the distance between the respective conservative groups, where such H-bonds are not formed. The distances between heavy atoms are shown in Å.

diffusion did not play a role. There was a slow disappearance of the long-lived product with a synchronous increase in the ground state absorption (ESI, Fig. S9 and S10), indicating that the long-lived component was partially and slowly reversible.

X-ray structural analysis of the initial dark-adapted state

To explore the structural basis for the observed photoinduced processes in mKalama1, we have determined the crystal structure of this protein at 1.71 Å resolution (PDB ID: 4ORN; for details, see ESI, Table S1). The superposition of mKalama1 and wtGFP by C_α atoms results in a root-mean-square (RMS) deviation of 0.6 Å, indicating close similarity of the two structures. As in many other fluorescent proteins described to date, the Cro of mKalama1 adopts a Z-configuration in the dark state. mKalama1 differs from wtGFP by 26 amino acids; the following seven: L64F, Y145M, H148G, V163A, T203V, S205V, and V224R are worth noting since they are close to the Cro. Four of them (Y145M, H148G, T203V, and S205V) make the protein cavity containing the Cro *p*-hydroxyphenyl ring entirely hydrophobic. Residues V205 and V203 terminate all hydrogen bond connections between the hydroxyl group of the Cro and E222 in mKalama1, similarly to the S205V/T203V double mutant of wtGFP described earlier.[6] The H148G mutation makes the β-strand fragment V147-G148-N149 move closer to the hydroxyl oxygen of the Cro by 1.3 Å, forming a novel hydrogen bond between the backbone N_{G148} and the phenolic oxygen atom of the Cro (Fig. 3). Thus, the area occupied by the H148 side chain in wtGFP now becomes open to the solvent in mKalama1. Despite the hydrophobic nature of the residues forming the Cro phenol ring cavity, the latter contains structural water molecules close to the protein exterior (Fig. 3a). The hydroxyl oxygen of the Cro thus forms a strong hydrogen bond with the nearest conserved water molecule.

Similarly to wtGFP, the environment of the imidazolone ring and of the side chain of the Cro-forming residue S65 is mostly hydrophilic in mKalama1. In both proteins, the imidazolone ring is surrounded by a cluster of hydrophilic residues creating an extensive hydrogen bonding network (ESI, Fig. S11). The presence of R224 and C70 (oxidized to sulfinic acid in mKalama1) increases the hydrophilicity of the imidazolone environment as compared to wtGFP, in which the residue at position 224 is valine and C70 is not oxidized. The side chain of R224 protrudes towards the methylidene group of the Cro and forms a strong hydrogen bond (ion pair) with E222 and a water-mediated hydrogen bond with the hydrophilic cluster, consisting of the Q69, Q94, R96, Q183, and N185 side chains. The latter residues are interconnected by numerous hydrogen bonds creating a network, which propagates further, through several water molecules, to the residues T62, T63, T108, N121, and H181. E222 also forms a strong hydrogen bond with the hydroxyl group of the Cro-forming residue S65 (Fig. 3a).

Therefore, the Cro binding pocket in the dark-adapted state of mKalama1 differs from that of wtGFP[38] by three major structural features: (i) the absence of H-bonded proton transfer pathways between the phenolic oxygen of the Cro and the side chain of E222; (ii) the presence of a novel, proton transfer pathway between the phenolic oxygen of the Cro and the bulk; (iii) the formation of an ion pair between the side chains of E222 and R224.

Electronic structure calculations: photocycle intermediates

To identify the nature of the ground-state intermediates in the photocycle of mKalama1, we have carried out *ab initio* quantum mechanical/molecular mechanical (QM/MM) calculations of equilibrium structures and vertical excitation energies of the initial and transient states in the protein. The initial, dark-adapted state of mKalama1 has been obtained "from-scratch" by introducing the corresponding mutations to the well-known structure of GFP S205V[6] and performing molecular dynamics (MD) simulations (ESI, Fig. S12). Fig. 4 illustrates a close similarity between the structures of the dark-adapted state of mKalama1, obtained theoretically and through the X-ray analysis. The superposition by C_α atoms results in a RMS deviation of less than 1 Å. Such good agreement also comes from the fact that the X-ray structures of wtGFP, GFP S205V, and mKalama1 are very close to each other. The "from-scratch" modeling well reproduces all new features of the Cro binding pocket in mKalama1. It fully supports a novel type of the water-mediated pathway that connects Cro Y66 with the solvent and also reveals a reduced distance between the backbone of G148 and Cro (3.8 Å) compared to those of wtGFP (4.3 Å) and GFP S205V (4.5 Å). Further calculations however show that this distance is adjustable and can be increased to accommodate additional water molecules, as in the structures of transients. In the dark-adapted, final MD

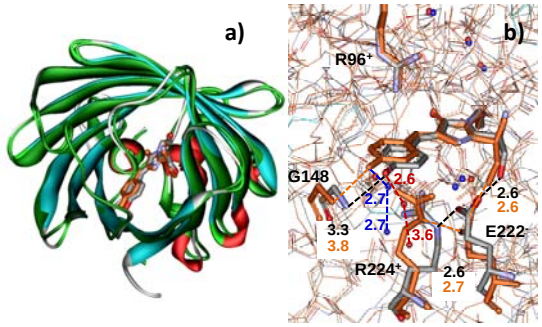

Fig. 4 Superposition of the final MD (green ribbons) and X-ray (chain A, cyan ribbons) dark-adapted structures of mKalama1: (a) the top view and (b) the Cro binding pocket. The Cro and the nearby amino acid residues are shown in the stick representation. The simulated structure is shown in orange. Water molecules are shown in red (experiment) and blue (theory). All distances are in Å. Note that the different conformations of the non-polar part of the side chain of R224 do not affect vertical excitation energies. The calculations also reveal a conformational flexibility of the ion pair E222-R224. See also ESI, Fig. S13.

structure, this distance is also enlarged compared to that of the X-ray structure (3.3 Å) to adjust the lengths of H-bonds in the new water-mediated pathway between the Cro and the bulk (Fig. 4b). The distance between the two nearest to the Cro water molecules is significantly reduced from 3.6 Å (X-ray) to 2.7 Å (final MD).

The anionic form of the Cro is a primary candidate in search for the ground-state intermediates in the photocycle. Since there is a well-defined water chain between Cro Y66 and the bulk, we first assume that the proton is released to the solvent. We will however show below that this transient does not result from a direct photoinduced proton transfer, but it is rather formed on the way back to the dark-adapted state along a radical branch. The resulting QM/MM//MD optimized structures of the Cro-binding pockets in the initial neutral and stabilized anionic states in mKalama1 are shown in Fig. 5. In both cases, the Cro adopts a planar conformation. The deprotonation of the Cro induces certain changes in the nearby environment. In particular, additional water molecules enter the binding pocket to stabilize the deprotonated Cro through formation of a hydrogen-bonded network with the nearby ion pair E222-R224.

The obtained anionic structure exhibits a blue-shifted absorption compared to that of wtGFP (band B), as it follows from the XMCQDPT2/EFP calculations of the corresponding vertical excitation energy, which is 440 nm. We note that the same effect is observed for β-naphthol[39] and 5-cyano-2-naphthol[40] in solvents of differing proton-donating ability. The accommodation of the additional water molecules results in a disruption of the preexisting proton transfer pathway that leads from the bulk solvent to the Cro, therefore, increasing a lifetime of this blue-shifted anionic intermediate upon its external reprotonation.

Other possible intermediates in the photocycle, which can transiently be formed after internal conversion as a result of a photoinduced shift in acid-base equilibrium inside the protein, includes the anionic and zwitterionic forms of both *Z*- and *E*-isomers and E222 as a primary proton acceptor. Various local H-bonded environments and multiple conformations of the ion pair

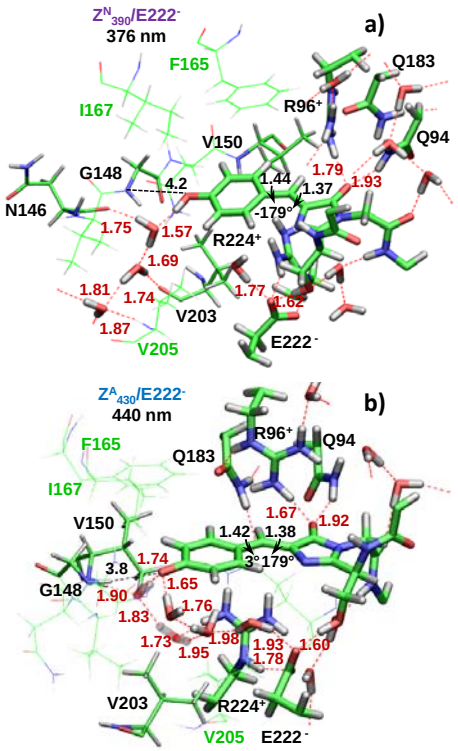

Fig. 5 The QM/MM ground-state equilibrium structures of mKalama1 in (a) the dark-adapted state with the neutral Cro, Z^{N}, and (b) the blue-shifted, stabilized transient state with the anionic Cro, Z^{A}_{430}. Shown are the QM parts (sticks) and the nearby MM residues (lines and bold lines). The hydrogen bonds are depicted in red. The N_{G148}–O_{Cro} distances, the bond lengths and the dihedral angles in the Cro methylidene bridge are shown in black. All distances are in Å. The structures are labeled according to the charge state of the Cro and to the assigned experimental absorption maxima. Shown also are the calculated vertical excitation energies. See also Table S2 for the assignment of the Cro structures.

E222-R224 have been considered. The structures of the Cro radical forms have also been analyzed. The final assignment of the transient absorption components are made based on the forms summarized in Fig. 6 (see also ESI, Table S2).

Singlet states

The transients characterized by short lifetimes of microseconds to sub-milliseconds should involve only local rearrangements in the protein. Therefore, the intraprotein proton transfer to a nearby residue, that serves as a proton acceptor, seems to be the most probable process. The calculated structure of the dark-adapted state, however, shows that all direct proton transfer pathways from the Cro to amino acid residues are disrupted inside the protein. The X-ray analysis supports this by revealing an entirely hydrophobic protein environment close to the Y66 phenol ring. This is also fully consistent with the results of the TA and fluorescence studies that indicate the absence of ESPT in mKalama1. At the same time, IC induced by twisting in the methylidene group of the Cro can offer alternative pathways for the proton transfer. Remarkably, E222, the primary proton acceptor during ESPT in wtGFP, remains to be the most reasonable partner in mKalama1: upon twisting, the distance

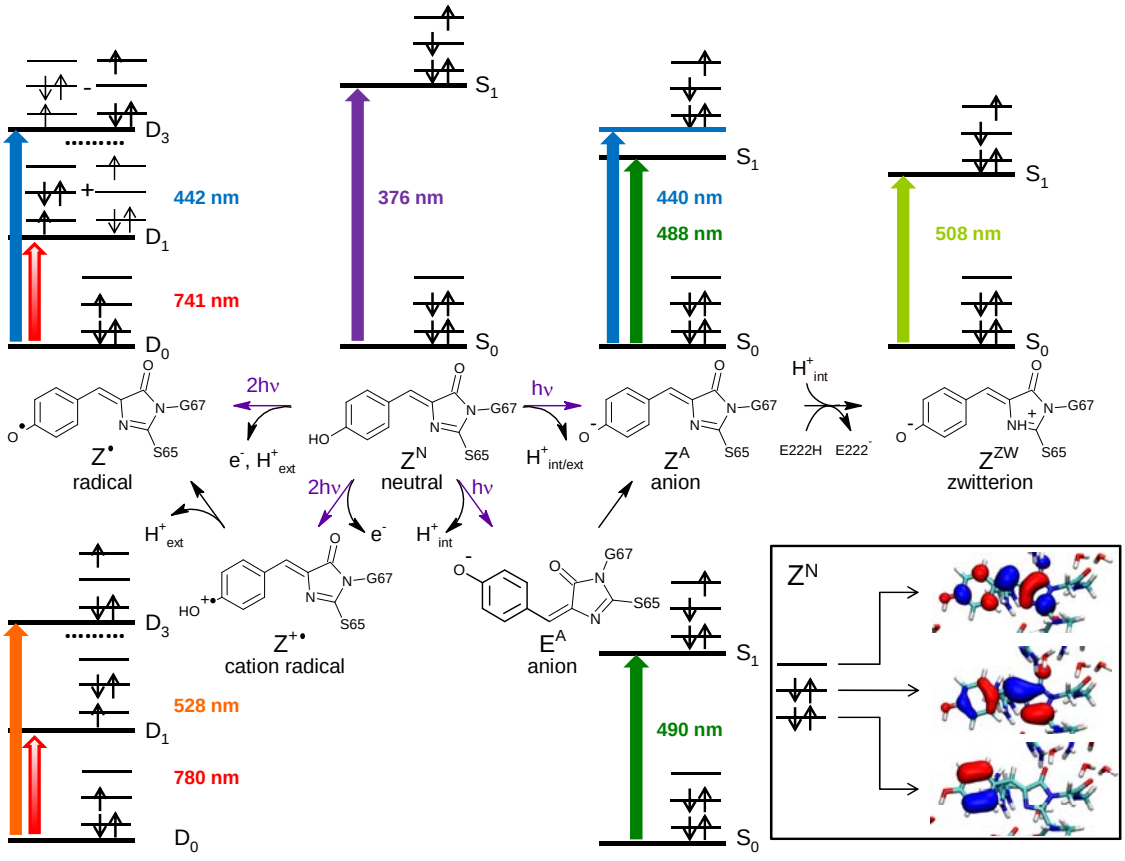

Fig. 6 Dark-adapted and transient states of mKalama1 containing the Cro in various protonation and oxidation forms. Leading electronic configurations for the singlet (S) and doublet (D) ground and excited states are given together with the calculated vertical excitation energies for the bright transitions. The inset shows the typical structure of the natural orbitals involved in electronic transitions for all the forms. Note the mixture of different electronic configurations in the excited states of the neutral radical, which is also true for the radical cation but to a lesser extent.

between the Cro phenolic OH and COO^-_{E222} can vary and the H^+ transfer pathway can transiently be formed. IC coupled to deprotonation results in either *Z*- or *E*-anion. The latter subsequently undergoes a thermal backward $E \rightarrow Z$ isomerization, which, in turn, may or may not be coupled with a simultaneous reprotonation that regenerates the neutral dark-adapted ground state.

The total QM/MM energies of the ground-state equilibrium structures of mKalama1 containing the Cro *Z*-anion and the protonated E222 residue are similar to those of the protein with the neutral Cro and the deprotonated E222, when additional water molecules stabilize the Cro binding pocket, Z^A/E222H ↔ Z^N_{wat}/E222$^-$ (ESI, Fig. S14). Due to the hydrophobic protein environment of Y66, these water molecules are confined in the cavity between the phenol ring and the ion pair E222-R224, forming H-bonded cyclic networks between OH_{Y66} and OH_{S65} through R224 and further to E222, which is, in turn, connected either to the N atom of the Cro heterocyclic ring or back to Y66. Upon the protonation of E222, the ion pair breaks up and the deprotonated Cro becomes stabilized by R224 through these water-mediated hydrogen bonds (Fig. 7a). All the structures, which exhibit a chain of water molecules that serves as a proton wire for direct transferring the proton back to Y66 (ESI, Fig. S14b), lie somewhat higher in energy (by *ca.* 10 kcal/mol) than those which have a disrupted water chain (ESI, Fig. S14a). Remarkably, these structures can be referred to those of the anionic ground-state intermediates in the photocycle of wtGFP – I and B,[4] with the intact and disrupted H-bonded pathway between E222 and Y66, respectively.

The time-resolved measurements reveal a complex nature of the relaxation kinetics with several sub-ms timescale components in the spectral region of 485-505 nm. The shortest lifetime component of 6 μs refers to the transient state absorbing at 489 nm. It is reasonable that this shortest-lived ground-state intermediate is formed in the *E* anionic form, which undergoes a fast backward isomerization and simultaneous reprotonation. Unlike typical *cis-trans* interconversion time scales of 10-100 ms[17,28], the μs lifetime of this dark state in mKalama1 indicates that the Cro *E*-anion is energetically highly disfavored. The calculations indeed reveal that the optimized structure of the *E*-isomer of the Cro anion is distorted from a planarity (Fig. 7b). The anion adopts a pronounced quinoide-like electronic structure, with the two bridge bonds adjacent to the *p*-hydroxyphenyl (Ph-C) and imidazolone (Im-C) rings being equal to 1.39 and 1.41 Å, respectively. The *E*-anion is pre-twisted by 17° about the Im-C bond, thus facilitating a rather fast backward $E \rightarrow Z$ isomerization in the ground state. The QM/MM energy difference of *ca.* 20 kcal/mol between the anionic forms of the *E*- and *Z*-isomers

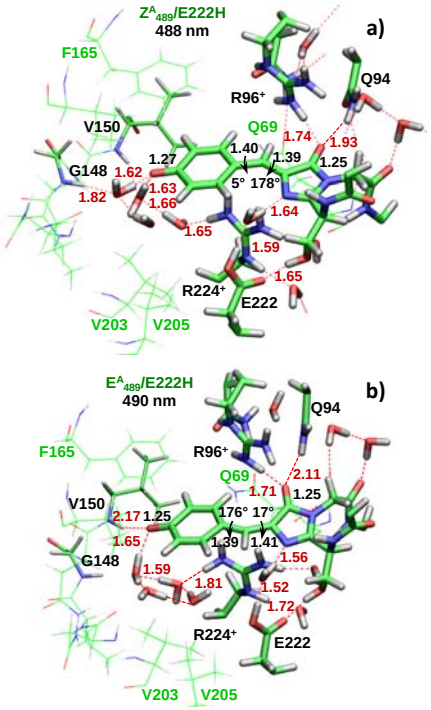

Fig. 7 The QM/MM ground-state equilibrium structures of the transient states in mKalama1 formed after internal conversion coupled to the proton transfer to E222: (a) *Z*-anion, Z^{A}_{489}, and (b) *E*-anion, E^{A}_{489}. Shown are the QM parts (sticks) and the nearby MM residues (lines). Two residues of the β-strand fragment, G148 and V150, involved in a hydrogen-bonding stabilization of the Cro anion, are displayed as the bold lines. The hydrogen bonds are depicted in red. Shown also are the calculated vertical excitation energies. All distances are in Å.

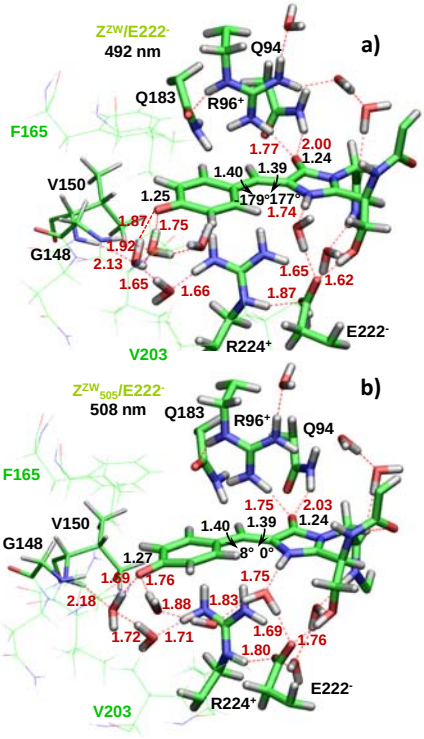

Fig. 8 The QM/MM ground-state equilibrium structures of the zwitterionic transient states in mKalama1 with different local water-mediated H-bonded networks: (a) the lowest-energy structure with a disrupted proton back-transfer pathway, and (b) the higher-energy structure that precedes regeneration of the neutral form. Shown are the QM parts (sticks) and the nearby MM residues (lines). Two residues of the β-strand fragment, G148 and V150, involved in a hydrogen-bonding stabilization of the Cro zwitterion, are displayed as the bold lines. The hydrogen bonds are depicted in red. Shown also are the calculated vertical excitation energies. All distances are in Å.

provides clear evidence that the *Z*-isomer is the most stable one, and the *E*-anion dark state of mKalama1 cannot be populated efficiently. This is fully consistent with the recent experimental studies on a series of BFPs that reveal no significant steady-state population of the Cro *E*-anion state in mKalama1 on a time scale of tens of milliseconds.[28] The calculated vertical excitation energies of the Cro *Z*- and *E*-anion inside the protein, containing the protonated E222 residue, are 488 nm and 490 nm, respectively (Fig. 6). This is consistent with the observed transient absorption signal at *ca.* 489 nm. Both anionic isomers exhibit very close absorption maxima and are, therefore, spectroscopically indistinguishable.

In the transient state Z^{A}_{489} (Fig. 7a), the proton transferred to E222 may subsequently be transferred forward to the Cro imidazolone ring, thus forming a Cro zwitterion (Fig. 8a). Indeed, the calculations show that the anionic and zwitterionic *Z*-forms have very similar ground-state energies and they exist in a nearly perfect equilibrium, Z^{A}/E222H ↔ Z^{ZW}/E222^{-} (ESI, Fig. S15). The stabilization of such zwitterion is gained through the formation of a quadrupole with the ion pair E222-R224 located in close vicinity to the Cro and regenerated upon deprotonation of E222. The zwitterionic forms exhibit a slightly red-shifted absorption: the calculated vertical excitation energies vary from 492 to 508 nm, depending on the particular structure of the local environment (Fig. 8). Such red-shift is consistent with the TA measurements and the solution measurements of the model synthetic GFP Cro,[41] indicating that the zwitterion state can indeed be populated during the proton transfer from E222 back to Z^{A}. Furthermore, the intermediate structure, which exhibits a water-mediated proton back-transfer pathway (Fig. 8b), lies only 6 kcal/mol higher in energy than the lowest-energy zwitterionic structure, thus rendering the overall step-wise zwitterionic pathway likely upon regenerating the dark-adapted state. Therefore, our results give rise to at least three intermediates responsible for the transient absorption signal in the spectral range of 489-505 nm: the *E*-anion (E^{A}), the *Z*-anion (Z^{A}), and the *Z*-zwitterion (Z^{ZW}).

Doublet states

The measured average excited-state singlet lifetime of few nanoseconds may also allow the neutral Cro in the first excited state to sequentially absorb a second 355 nm (3.5 eV) photon during a 4 ns laser-pulse duration. In order to check whether this process is energetically allowed, we have estimated an adiabatic electron binding energy of the dark-adapted state of mKalama1 based on the calculated energy difference of the ground electronic states of the protein, containing the neutral and the radical cation forms of the Cro. It is found that the electron is bound with an energy of 8.6-8.9 eV, depending on the local environment of the Cro. Upon photoionization, a Cro radical cation and a solvated

electron are formed. Free electrons in an aqueous solution are known to be bound with an energy of 3.3 eV.[42] Therefore, sequential absorption of two photons with a total energy of 7 eV can potentially ionize the neutral Cro inside the protein. However, the Cro-binding pocket contains three charged residues, R96, R224 and E222, thus making an overall charge of the cavity equal to +1. Therefore, the Cro radical cation should undergo a barrierless deprotonation due to Coulomb repulsion in the Cro-binding cavity. The proton can be expelled from the protein to the bulk through the well-defined water chain bonded to the planar Cro (Fig. 3a and 5a). The optimized structure of this transient state indeed reveals that the $Z^{+\bullet}$ Cro is only partially protonated, with the O-H bond length of 1.13 Å and the H-bond distance to the nearest water molecule of 1.29 Å (ESI, Fig. S16).

The calculated spectra of the Cro radical forms exhibit three low-lying transitions. The lowest lying transition has an energy of less than 1.77 eV (700 nm) and is characterized by a low oscillator strength, which is *ca.* 10 times lower than that of the third bright transition located at 442 nm ($Z^{\bullet}$) and 528 nm ($Z^{+\bullet}$) (Fig. 6). The second transition at 522 and 578 nm for $Z^{\bullet}$ and $Z^{+\bullet}$, respectively, is optically dark, with an oscillator strength less than 0.005. The detectable band in the red part of the spectrum at *ca.* 700 nm and the transient band at 430 nm found experimentally can then be assigned to the neutral radical, whereas the radical cation does not live long enough to be spectroscopically detectable. At *ca.* 700 nm, the radical TA also overlaps with the broad absorption band of solvated electrons that extends from 550 up to 800 nm.[43]

Discussion

In this study, we report the multiexponential TA decay upon excitation of the BFP mKalama1 observed in the microsecond-to-second time domain. This is the first account on the relaxation kinetics in BFPs tracked by broadband TA spectroscopy. The relaxation kinetics in the ground electronic state reveal specific pH, D/H exchange, and [O_2] sensitivity, and is interpreted as the combination of two parallel processes: (1) the microsecond to sub-millisecond ground-state intraprotein reprotonation of the Cro anion coupled to its reisomerization; (2) recombination of the Cro radical followed by the external reprotonation of the Cro anion on the time scale extending to seconds. We first discuss the high-resolution X-ray structure of the dark-adapted state of mKalama1 and possible proton transfer pathways inside the protein. The photocycle of mKalama1 is then presented based on the theoretical account of the observed transient absorption, where the QM/MM//MD calculations provide an important insight into the structures of possible transients and their spectroscopy.

Proton transfer pathways

The obtained X-ray structure (Fig. 3a) shows that the essentially hydrophobic protein environment of the Y66 cavity does not allow a proton transfer from the Cro to E222 or to any other possible proton acceptors, so that both main (Cro-water-S205-E222) (Fig. 3b) and alternative (Cro-water-T203-E222) (Fig. 3c) ESPT pathways found in wtGFP and GFP S205V are non-functional in mKalama1. In wtGFP, there is also a minor probability for the proton to be released to the bulk through the switchable T203 exit pathway near H148, replenished in the ground state from E222 and the entry pathway near E5.[7] A similar entry pathway exists in mKalama1, too; however, it is interrupted at E222 due to the lack of the hydrophilic residues at positions 203 and 205 disabling proton transfer to the Y66 hydroxyl. Moreover, R224 interferes with this pathway by forming a hydrogen bond with S72, thus preventing a direct pathway from S72 to E222 through a water chain (ESI, Fig. S11). mKalama1 also lacks the switchable T203 exit pathway, since this position is mutated to valine.

Recently, an entirely water-mediated ESPT pathway transiently formed in GFP T203V/S205A, which exhibits a dual emission band, has been suggested.[16] Although the X-ray structure of this double mutant does not reveal any viable proton-wire pathways, a formation of the H-bonded water chain through molecules infrequently entering the protein structure from the bulk has been observed during MD simulations. The analysis of the X-ray structure of mKalama1, for the first time, allows us to experimentally identify the chain of water molecules that links the protein surface with the Y66 hydroxyl group (Fig. 3a). Note that the similar chain of water molecules is present in wtGFP as well (Fig. 3b); however, in wtGFP, the distance between the water molecule H-bonded to Y66 and the next water molecule in the chain exceeds 4.5 Å, disabling proton transfer, whereas in mKalama1 this distance is 3.6 Å. Furthermore, in mKalama1 these water molecules are indirectly connected through the backbone carbonyls of N146 (2.8 and 3.3 Å) and of V203 (2.8 Å).

Optimized for brightness, the blue-light emitting mKalama1 with 26 mutations relative to wtGFP appears to be more exposed to the solvent than wtGFP and its single and double mutants, so that the water-mediated H-bonded pathway is well resolved in the X-ray structure of this protein. However, neither the GFP-related BFP mKalama1, nor any other GFPs studied so far, reveal this pathway to be functional in ESPT, except for GFP T203V/S205A.[16] But even in the latter case, ESPT occurs with a remarkably small quantum efficiency. Therefore, even though we might always expect these transient or permanent water-mediated pathways to exist in the GFPs, their efficiency in transferring proton to the bulk from the neutral Cro upon excitation is low. Indeed, the presence of this water-mediated channel in mKalama1 does not enable ESPT, as confirmed by the present transient fluorescence spectroscopy measurements.

The fact that no ESPT is observed in the present study can well be rationalized by considering the proton transfer to the bulk as a step-wise process. The first proton acceptor is the water molecule directly bound to the Y66 hydroxyl group (Fig. 3a and 5a). Whether or not this water molecule may efficiently mediate the proton transfer depends on its pK_a compared to that of the Y66 phenol in the excited state of the Cro. The single water molecule appears to be a poor proton acceptor, taking into account that the charged residue $R224^+$, which is only 7.2 Å apart, disfavors a formation of H_3O^+ close to the Cro anion, so that the neutral Cro prevails. As opposed to mKalama1, the substitution V224R is not present in GFP T203V/S205A, and the inefficient step-wise ESPT is enabled through a transiently formed water chain.[16]

At the same time, the water-mediated proton wire should indeed facilitate the deprotonation of the ionized planar Cro in its radical cation form, which is highly disfavored in the positively charged

Cro binding pocket of mKalama1. The optimized equilibrium structure of $Z^{+\bullet}$ supports this by revealing a partial formation of the H_3O^+ ion bound to the Cro neutral radical (ESI, Fig. S16b). Therefore, in mKalama1, the planar Z^N Cro can undergo deprotonation only following its ionization. However, the present transient absorption studies provide clear evidence for the formation of ground-state anionic intermediates that are likely to be formed during the non-radiative deexcitation, induced by twisting in the Cro methylidene bridge, as discussed below.

Excited-state decay channels

To interpret the relaxation kinetics, we first analyze the rapid, ns-timescale electronic relaxation out of the first excited state. The observed multi-exponential fluorescence decay of mKalama1 (ESI, Fig. S1) can be caused either by a ground-state heterogeneity or by a Cro flexibility leading to photoisomerization-induced deactivation; the latter has been shown for both GFP[44] and the free Cro in all protonation states.[34,45] Our group[46-48] and others[49] have demonstrated almost the complete restoration of fluorescence in "locked" compounds. The correlation between the phenyl-imidazolone dihedral angle in several FP chromophores and their fluorescent lifetimes has also been revealed.[50] One cannot, however, ignore other deactivation pathways such as the torsional deformation of the phenyl single bond.[51-52] For the isolated Cro anion, we have previously shown that there are indeed two distinct types of the so-called conical intersections, where the crossing between the ground and excited states is induced by twisting in the Cro methylidene group in both cases.[53]

The various protein environments alter the branching in the FPs deexcitation pathways. In wtGFP, IC is largely suppressed due to high barriers that prevent Cro from twisting. The barriers are imposed by the extended H-bonded network involving the Y66 phenol ring. In mKalama1, the hydrophobic protein cavity of Y66 facilitates the intramolecular rotation. The protein can accommodate such twisting by adjusting the two dihedral angles in the methylidene bridge. Moreover, IC proceeds through a twisted intramolecular charge transfer pathway accompanied by distinct changes in the Cro electronic structure,[53] thus changing a pK_a of its phenol ring. Therefore, the photoinduced twisting influences both the Cro acidity and a distance between the proton donor and acceptor inside the protein, enabling a proton transfer otherwise locked for the planar Cro in mKalama1. As in wtGFP, the E222 residue located close to the Cro remains to be the proton acceptor, as it follows from the X-ray crystal structure analysis and is further supported by the QM/MM calculations.

One-photon IC-induced branches

In mKalama1, the photoinduced deprotonation proceeds faster than 1 µs. A huge fluorescence intervention prevents us from the experimental tracking of TA on the sub-µs time scale, and the exact sequence of isomerization/deprotonation events (parallel/coupled/consecutive) is yet to be studied. Based on the striking similarity between the high pH-induced, dark-adapted state spectra of mKalama1 (Fig. 1a) and the spectra of the 6 µs and 98 µs TA components at pH 7.5 (Fig. 1e), we propose that the sub-ms kinetic phases reflect the "delayed" reprotonation of the ground-state deprotonated Cro. Since the formation of Z^A at pH 7.5 is thermodynamically unfavorable and the $^*Z^A$ form has

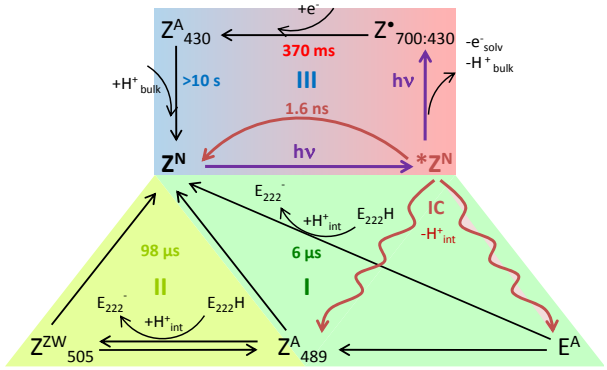

Fig. 9 Three branches in the photocycle of mKalama1. Shown are the characteristic decay times and the wavelengths of the spectroscopically detected intermediates.

not been detected in the emission spectra, we conclude that the anionic species detected in the TA spectra at *ca.* 489 nm at early times (t < 1 µs) are the *E*- and *Z*-anions (E^A_{489} and Z^A_{489}), which are formed following IC coupled to the intraprotein proton transfer (Fig. 9, branch I). This is fully consistent with the calculated vertical excitation energies of these anionic forms inside the protein containing the protonated E222: 488 nm and 490 nm for Z^A_{489} and E^A_{489}, respectively (Fig. 7). At a later stage, the bathochromic shift to 505 nm can be referred to the TA of the zwitterionic form, Z^{ZW}_{505} (Fig. 8), efficiently populated when a transient water-mediated proton back-transfer pathway has not been formed yet (Fig. 9, branch II).

Supporting our hypothesis for the "delayed" reprotonation, the kinetics of both first (6 µs) and second (98 µs) components exhibit a pronounced 3-fold H/D isotope effect (ESI, Fig. S17) in the pH region where both are nearly pH-independent (ESI, Fig. S7). The latter excludes a thermodynamic isotope effect and points out to a true kinetic effect; this also excludes the possibility that some of these phases are rate-limited by Cro conformational changes, without proton rearrangements. The 98 µs component is pH-sensitive above pH ~8.5, while the 6 µs component remains mostly unchanged until pH 11 (ESI, Fig. S7). Therefore, the faster component represents the reprotonation of the Cro by an internal (not in rapid equilibrium with the bulk) proton donor coupled to the simultaneous *E*→*Z* backward isomerization, whereas the 98 µs component refers to the regeneration of the dark-adapted state from the planar *Z*-forms of the Cro through a transient formation of water-mediated proton wires inside the Cro binding pocket.

The remarkably short µs lifetime of the ground-state *E*-anion is fully consistent with its distorted quinoide-like structure, which is pre-twisted about the Im-C bond (Fig. 7b), thus facilitating the fast backward isomerization of the Cro. The longer time scale reprotonation of the *Z*-anion, which is formed either right after IC or following the ground-state *E*→*Z* isomerization without simultaneous proton transfer, can occur through at least two alternative pathways. The direct proton back-transfer pathway from E222 to the Cro *Z*-anion includes the formation of a low-barrier hydrogen bond structure, which lies *ca.* 10 kcal/mol above the structure with a disrupted chain of water molecules, Z^A_{489} (ESI, Fig. S14). The second step-wise pathway, along which the Cro zwitterionic state is populated prior to the dark-adapted state

is regenerated, occurs through the transient states that are much closer in energy than those along the first route. The largest energy gap between the two structures with a disrupted and transiently formed water chain is *ca.* 6 kcal/mol (ESI, Fig. S15). Therefore, the experimental spectral shift in the TA absorption that agrees well with the difference in the calculated vertical excitation energies of the anionic and zwitterionic Cro forms indicate that both forms are populated at different stages in the photocycle of mKalama1. We also note that reprotonation of the ground-state anion is a continuous phenomenon, and the two distinct kinetic phases might not be enough for the full description; the two phases should, therefore, be regarded as a "first approximation" to the relaxation kinetics.

Further corroboration for a coupled deprotonation(reprotonation)/isomerization process comes from the study of other GFP variants. The characteristic interconversion time scales observed in mKalama1 match those of the GFP variant GFPmut2, with three mutations S65A, V68L, and S72A.[54] The photodynamics of GFPmut2 and its photoswitchable variant with the E222Q replacement (GFPmut2Q) have extensively been studied using FCS and single-wavelength TA.[17-18,25,55] The FCS analysis of GFPmut2 reveals two components with similar lifetimes to those found in mKalama1: a pH-dependent component with τ=10-100 μs and a laser power-dependent component with τ=100-500 μs.[18] The former is attributed to protonation-deprotonation of the GFP Cro, whereas the latter refers to an excited-state process unrelated to the proton transfer. The sub-ms protonation dynamics has also been investigated in GFP S65T, using pH-jump experiments.[56] It has been found that the proton transfer inside GFP is coupled to the protein dynamics.[57]

At the same time, a class of optically modulatable BFPs that originate from a double mutant GFP T203V/S205V exhibits the photoinduced ground-state $E^A \rightarrow Z^N$ relaxation dynamics on a longer time scale of 10-100 ms,[28] relative to that observed in mKalama1. These proteins absorbing at 390 nm are specifically optimized to enable efficient steady-state population of anionic dark states with a shifted transient absorption to *ca.* 500 nm, with ms-lifetimes of their kinetically trapped *E*-anionic states. The mutations H148K and V150A are shown to be of particular interest in this respect.[28] A singly mutated protein mKalama1 G148K exhibits the relaxation kinetics on a time scale of 10 ms, which is remarkably longer than that found in both mKalama1 (τ = 6-98 μs) and GFP T203V/S205V (τ ~ 240 μs, see ESI, Fig. S18). The analysis of the calculated structure of the transient *E*-anion state found in this work well explains such different relaxation timescales of the anionic transients. While the backbone of both G148 and K148 contributes to only a weak stabilization of the *Z*-anion (Fig. 7a), the positively charged flexible side chain of K148 is capable of stabilizing *E*-anion through formation of a strong H-bond (Fig. 7b). Also, upon the V150A point mutation, a less bulky side chain at the position 150 should be favorable for the *E*-anion formation. Therefore, all the data available strongly support our assignment.

Two-photon radical branch

The decay of the 430 nm TA band (Fig. 1d) contains two components, with τ = 370 ms and τ > 10 s (Fig. 1e). To study the possibility of multiphoton excitation we have measured the power dependence of the transient absorption of mKalama1 by

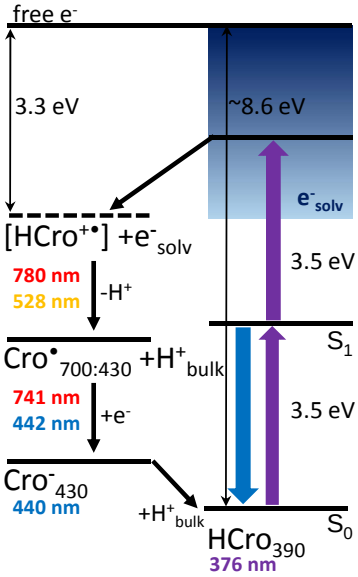

Fig. 10 Schematic representation of consecutive two-photon absorption events, which result in the protein photoionization and further the deprotonation. The relaxation kinetics along this branch includes the radical recombination followed by the re-protonation on the time scale of seconds.

excitation at different flash energies (ESI, Fig. S19). The kinetic and spectral analysis of the TA demonstrates that the 430 nm transient has no appreciable rise within the observation time window ($t \geq 1$ μs) and appears synchronously with the Z^A/E^A anions within the experimental error. Since the spectral shape of the 430 nm transient does not change over six orders of magnitude in time, and its lifetime is much longer than the characteristic time of the $Z^A/E^A \rightarrow Z^N$ evolution, we exclude this intermediate from branches I and II in the photocycle of mKalama1 (Fig. 9). The absorption maximum of this transient matches the calculated vertical excitation energy of the neutral radical Cro ($Z^•$), with the brightest transition at 442 nm (Fig. 6). It is also worth noting, that the Cro radical also exhibits a weak absorption at around 700 nm. The saturation curves in the power dependence of the TA show a remarkably identical behavior for these spectral components at the time delays of 1 μs - 100 μs (ESI, Fig. S20). This is in contrast to the behavior of other spectral components, for example, at 630 nm. It indicates a common origin of the transient spectral components observed at 430 and 700 nm, thus supporting our assignment.

Since the Cro in mKalama1 is initially in the neutral protonated state and does not undergo ESPT upon excitation, the mechanism of the radical formation needs special attention. In particular, we discuss two aspects: (i) the protonation state and (ii) the nature of the photoionization process. We note, that the yield of both components at 430 and 700 nm exhibits an n=1 hyperbolic power dependence typical for a one-photon absorption process (ESI, Fig. S20). However, we might expect to observe a pseudolinear power dependence, if two photons are absorbed sequentially through the first excited state[58], in the range of flash energies used in the present work.[43] Therefore, we consider the two-photon ionization of mKalama1 with a total energy of 7 eV to be likely, also supported by the QM/MM calculations (Fig. 10). The

estimated ionization energy of mKalama1 (8.6-8.9 eV) is consistent with the smaller vertical detachment energy of the Cro in the anionic state inside GFP S65T (7.1 eV), previously calculated at the MRMP2/EFP level of theory.[59] The protein ionization in solution results in formation of a Cro radical and a solvated electron, where the latter is bound with a considerable energy of 3.3 eV in the bulk of water molecules,[42] thus resulting in the overall energy of the ionized protein and the solvated electron to be 1.4-1.7 eV lower than the ionization energy of two 355 nm photons. The broad, weak TA band of solvated electrons, spanning beyond 500 nm,[60,61] has indeed been observed in the present experiments (Fig. 1c and ESI, Fig. S6). Solvated electrons have previously been registered in photoactive yellow protein (PYP), containing a *p*-coumaric acid chromophore, upon two-photon ionization,[43,58,61] as well as for free chromophores of PYP[43,58] and GFP[34,62] in solution.

The lifetime of the solvated electrons is similar to that of the third relaxation component ($\tau \cong$ 730 μs) and is also oxygen-sensitive (ESI, Fig. S21 and S22). This is not surprising since they can react with oxygen to form superoxide.[63-65] Additionally, long-lived radicals have been detected in various fluorescent proteins upon photolysis using EPR spectroscopy.[22] The optical absorption extinction coefficient of solvated electrons is difficult to estimate, therefore, we are unable to determine what fraction of the protein enters a photoinduced pathway producing solvated electrons; however, since the yield of other products is high, the production of solvated electrons cannot be very efficient in mKalama1.

The Cro radical cation has the bright transition, which is significantly red-shifted to 528 nm compared to that of the neutral radical (Fig. 6). The TA has not been observed in this spectral window unlike the one of the neutral radical. Therefore, the multiphoton ionization also induces a proton transfer (ESI, Fig. S16). A presumably very low pK_a of the radical cation inside the positively charged binding cavity of the protein makes this process irreversible, occurring on an ultrafast time scale.

It is worth noting that the radical formation can be responsible for blinking phenomena observed in many FPs.[66] Indeed, we have also registered the formation of radical species $Z^{\bullet}$ in wtGFP, characterized by only slightly red shifted TA (ESI, Fig. S23). At the same time, multiphoton ionization through sequential resonant absorption of two photons requires relatively powerful laser pulses, such as those of a nanosecond NdYag laser used in the present work, and this process is less relevant to most common microscopy applications of FPs. The BFP proteins, including mKalama1 mutants, have successfully been used in two-laser imaging experiments where dark anionic species are optically depleted.[28] The role of radical species in these experiments at a very low laser power is negligible. Precautions should, however, apply when using focused, pulsed femtosecond lasers with high instantaneous light intensities, which drastically increase a probability of high-order mixed resonant and non-resonant absorption events, as for example in multiphoton microscopy[67].

The transient absorption at 430 nm is also consistent with the spectral window for anionic tyrosine-based chromophores in teal FPs ($\lambda_{ex/em}$= 453/488 nm),[68] blue-shifted relative to GFPs. This shift is caused by electrostatic interactions of the Cro anion with a nearby positively charged residue H197. The ion pair R224-E222 of mKalama1 can be regarded as the analogue of H197-E215 in TFPs. Therefore, upon transient formation of the Cro anion in the presence of this ion pair, a blue-shifted TA band might also be expected in the case of mKalama1. Since the spectral components at *ca.* 700 and 430 nm, both assigned to the radical formed at the earlier stages of the photocycle, have essentially different lifetimes, we conclude that a radical recombination proceeds faster than re-protonation of the Cro from the bulk. This gives rise to the formation of a stabilized blue-shifted anion, Z^{A}_{430} (Fig. 9, branch III). The structure of this long-lived intermediate is shown in Fig. 5b. The calculated vertical excitation energy of 440 nm supports the assignment.

To summarize, the complete photocycle of mKalama1 shown in Fig. 9 spans the time frame of up to seconds. Our studies reveal diversity of the photochemistry in FPs and disclose hidden proton transfer events, which have previously been unexplored.

Conclusions

We have for the first time performed broadband transient absorption measurements on the microsecond-to-second time scale to identify dark transient states in the BFP mKalama1. The complete photocycle of this protein spanning fifteen orders of time magnitude is proposed based on the results obtained through time-resolved spectroscopy, X-ray crystallography and high-level *ab initio* QM/MM calculations. The transient absorption data analyzed in a pH range from 7 to 12 indicate that mKalama1 undergoes a delayed photoinduced proton transfer. The three-fold H/D kinetic isotope effect exhibited by 6 μs and 98 μs exponential phases in the relaxation kinetics supports this. This isotope effect along with the hydrophobic nature of Y66 in the chromophore-binding protein cavity suggests that upon excitation the neutral chromophore should undergo internal conversion induced by isomerization/twisting to deprotonate.

The X-ray structural analysis reveals no viable proton transfer pathways inside the protein, except for a novel water-mediated channel that links the chromophore directly to the bulk. A lack of excited-state proton transfer shows that this channel is non-functional in mKalama1, primarily due to the substitution V224R. At the same time, upon photoinduced twisting or isomerization of the chromophore, the proton transfer events include E222 as the primary proton acceptor inside the protein.

We provide a full theoretical account for all observed transient ground-state absorption features that have been assigned to the anionic, zwitterionic, as well as radical forms of the chromophore in either *Z* or *E* configurations. Based on these results, we suggest that the non-radiative excited-state decay includes two parallel branches: (i) internal conversion coupled to the intraprotein proton transfer, and (ii) radical formation upon consecutive absorption of two photons, which causes the neutral chromophore to ionize. A radical cation, energetically disfavored, undergoes almost barrierless deprotonation to the bulk through a novel, water-mediated proton wire well resolved in the X-ray structure. At the same time, resonant two-photon ionization requires relatively powerful laser pulses, and the radical branch is less relevant to most common microscopy applications of FPs.

Upon regenerating the dark-adapted state of mKalama1, the intraprotein proton transfer events occur in several phases, all within the sub-ms time frame. We conclude that the thermal re-

isomerization coupled with simultaneous re-protonation, $E^A \rightarrow Z^N$, occurs on a remarkably fast μs-time scale; whereas the $Z^A \rightarrow Z^N$ conversion proceeds within a time frame of *ca.* 100 μs – 1 ms and includes formation of transient water-mediated proton wires inside the chromophore-containing protein cavity. Along the two-photon branch, the relaxation through external re-protonation of Z^A, following radical recombination $Z^{\bullet}+e^- \rightarrow Z^A$, occurs much slower on a time scale of milliseconds to seconds.

Prior to this work, no full spectral information about the transients appearing at the microsecond-to-second time scale has been available, and the mechanistic details of photoinduced processes in BFPs have remained elusive. This delayed reactivity, previously unexplored, will not only bring us to the understanding of their photoinduced dynamics, but also opens up wide perspectives for long-wavelength optical modulation of steady-state fluorescence in BFPs, the keystone to modern super-resolution microscopy.[28-29,69]

Acknowledgements

This work was supported by the U.S. National Science Foundation (CHE-1213047, K.M.S.), the U.S. National Institutes of Health (R21EB009976-01, A.S.B. and HHSN26120080001E, S.P.), the Intramural Research Program of the NIH, U.S. National Cancer Institute, Center for Cancer Research (S.P.), the Academy of Finland, the Sigrid Jusélius Foundation, Biocenter Finland, and Biocentrum Helsinki (D.A.B., M.V., and H.I.), as well as by the Russian Foundation for Basic Research (RFBR No. 14-03-00887, A.V.B.) and a Marie Curie FP7 CIG Grant (A.V.B.). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This work was granted access to the HPC resources of the Leibniz and RZG Supercomputing Centers (Garching, Germany) made available within the FP7 DECI/PRACE-2IP initiative, as well as to the Lomonosov Supercomputing Center at Moscow State University. Crystallization was performed at the Crystallization Core Facility, Institute of Biotechnology, Helsinki University. Seija Mäki and Jesper Ømig are greatly acknowledged for their excellent assistance in the crystallization, as well as Dr. Konstantin Kogan for the data collection, and the staff at the DIAMOND synchrotron beamline for their support. We thank Laren Tolbert, Robert Dickson and Konstantin Lukyanov for stimulating discussions. We are grateful to the members of Joseph Perry's group for the help with fs TA measurements. We thank Mårten Wikström for hosting two research visits of RBV to Helsinki, providing access to their equipment, and for his interest to this study. We would like to dedicate this work to the memory of Dmitry Bloch and Michael Verkhovsky.

Notes and references

[a] *School of Chemistry and Biochemistry and* [b]*Petit Institute of Bioengineering and Bioscience,* [c]*School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, United States. E-mail: solntsev@gatech.edu (K.M.S.)*
[d]*Research Program in Structural Biology and Biophysics, Institute of Biotechnology, University of Helsinki, Helsinki 00014, Finland*
[e]*Synchrotron Radiation Research Section, Macromolecular Crystallography Laboratory, National Cancer Institute, Argonne, IL 60439, United States*
[f]*Basic Research Program, Leidos Biomedical Research Inc, 9700 South Cass Avenue, Argonne, IL 60439, United States*
[g]*Department of Physics and Astronomy, Aarhus University, Aarhus, Denmark. E-mail: bochenkova@phys.au.dk (A.V.B.)*
[h]*Department of Chemistry, Moscow State University, Moscow, Russia*

† Electronic Supplementary Information (ESI) available: Materials and Methods, details for TCSPC measurements, transient absorption kinetic measurements and their analysis, results of the X-ray structural analysis and of the QM/MM//MD calculations. See DOI: 10.1039/b000000x/

‡ These authors contributed equally to the work.

1 R. Y. Tsien, *Annu. Rev. Biochem.*, 1998, **67**, 509.
2 M. Chattoraj, B. A. King, G. U. Bublitz and S. G. Boxer, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 8362.
3 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.
4 K. Brejc, T. K. Sixma, P. A. Kitts, S. R. Kain, R. Y. Tsien, M. Ormö and S. J. Remington, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 2306.
5 D. Stoner-Ma, A. A. Jaye, P. Matousek, M. Towrie, S. R. Meech and P. J. Tonge, *J. Am. Chem. Soc.*, 2008, **127**, 2864.
6 X. Shu, P. Leiderman, R. Gepshtein, N. R. Smith, K. Kallio, D. Huppert and S. J. Remington, *Protein Sci.*, 2007, **16**, 2703.
7 N. Agmon, *Biophys. J.*, 2005, **88**, 2452.
8 C. Fang, R. R. Frontiera, R. Tran and R. A. Mathies, *Nature*, 2009, **462**, 200.
9 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. USA*, 2004, **101**, 17988.
10 M. Di Donato, L. J. G. W. van Wilderen, I. H. M. van Stokkum, T. C. Stuart, J. T. M. Kennis, K. J. Hellingwerf, R. van Grondelle and M. L. Groot, *Phys. Chem. Chem. Phys.*, 2011, **13**, 16295.
11 B. L. Grigorenko, A. V. Nemukhin, I. V. Polyakov, D. I. Morozov and A. I. Krylov, *J. Am. Chem. Soc.*, 2013, **135**, 11541.
12 H.-W. Ai, N. C. Shaner, Z. Cheng, R. Y. Tsien and R. E. Campbell, *Biochemistry*, 2007, **46**, 5904.
13 R. N. Day and F. J. Schaufele, *Biomed. Opt.*, 2008, **13**, 031202.
14 M. A. Mena, T. P. Treynor, S. L. Mayo and P. S. Daugherty, *Nat. Biotechnol.*, 2006, **24**, 1569.
15 G.-J. Kremers, J. Goedhart, D. J. van den Heuvel, H. C. Gerritsen and T. W. J. Gadella, Jr., *Biochemistry*, 2007, **46**, 3775.
16 V. Wineman-Fisher, R. Simkovitch, S. Shomer, R. Gepshtein, D. Huppert, M. Saif, K. Kallio, S. J. Remington and Y. Miller, *Phys. Chem. Chem. Phys.*, 2014, **16**, 11196.
17 V. Quercioli, C. Bosisio, S. C. Daglio, F. Rocca, L. D'Alfonso, M. Collini, G.Baldini, G. Chirico, S. Bettati, S. Raboni and B. Campanini, *J. Phys. Chem. B*, 2010, **114**, 4664.
18 C. Bosisio, V. Quercioli, M. Collini, L. D'Alfonso, G. Baldini, S. Bettati, B. Campanini, S. Raboni and G. Chirico, *J. Phys. Chem. B*, 2008, **112**, 8806.
19 P. Schwille, S. Kummer, A. A. Heikal, W. E. Moerner and W. W. Webb, *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 151.
20 U. Haupts, S. Maiti, P. Schwille and W. W. Webb, *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 13573.
21 J. Widengren, Ü. Mets and R. Rigler, *Chem. Phys.*, 1999, **250**, 171.
22 R. B. Vegh, K. M. Solntsev, M. K. Kuimova, S. Cho, Y. Liang, B. L. W. Loo, L. M. Tolbert and A. S. Bommarius, *Chem. Commun.*, 2011, **47**, 4887.
23 A. Jimenez-Banzo, S. Nonell, J. Hofkens and C. Flors, *Biophys. J.*, 2008, **94**, 168.
24 R. Bizzarri, R. Nifosi, S. Abbruzzetti, W. Rocchia, S. Guidi, D. Arosio, G. Garau, B. Campanini, E. Grandi, F. Ricci, C. Viappiani and F. Beltram, *Biochemistry*, 2007, **46**, 5494.
25 S. Abbruzzetti, E. Grandi, C. Viappiani, S. Bologna, B. Campanini, S. Raboni, S. Bettati and A. Mozzarelli, *J. Am. Chem. Soc.*, 2005, **127**, 626.
26 K. M. Dean, J. L. Lubbeck, J. K. Binder, L. R. Schwall, R. Jimenez and A. E. Palmer, *Biophys. J.*, 2011, **101**, 961.

27 R. B. Vegh, K. B. Bravaya, D. A. Bloch, A. S. Bommarius, L. M. Tolbert, M. Verkhovsky, A. I. Krylov and K. M. Solntsev, *J. Phys. Chem. B*, 2014, **118**, 4527.

28 A. E. Jablonski, R. B. Vegh, J.-C. Hsiang, B. Bommarius, Y.-C. Chen, K. M. Solntsev, A. S. Bommarius, L. M. Tolbert and R. M. Dickson, *J. Am. Chem. Soc.*, 2013, **135**, 16410.

29 J.-C. Hsiang, A. E. Jablonski and R. M. Dickson, *Acc. Chem. Res.*, 2014, **47**, 1545.

30 A. A. Granovsky, *J. Chem. Phys.*, 2011, **134**, 214113.

31 P. Day, J. Jensen, M. Gordon, S. Webb, W. Stevens, M. Krauss, D. Garmer, H. Basch and D. J. Cohen, *Chem. Phys.*, 1996, **105**, 1968.

32 A. A. Granovsky, Firefly, version 8.0, http://classic.chem.msu.su/gran/firefly/index.html.

33 G. Jung, J. Wiehler and A. Zumbusch, *Biophys. J.*, 2005, **88**, 1932.

34 K. M. Solntsev, O. Poizat, J. Dong, J. Rehault, Y. B. Lou, C. Burda and L. M. Tolbert, *J. Phys. Chem. B*, 2008, **112**, 2700.

35 L. M. Tolbert, A. Baldridge, J. Kowalik and K. M. Solntsev, *Acc. Chem. Res.*, 2012, **45**, 171.

36 D. A. Bloch, A. Jasaitis and M. I. Verkhovsky, *Biophys. J.*, 2009, **96**, 4733.

37 A. Rakowski, *Z. Phys. Chem. Stochiom.Verwandtschafts*, 1906, **57**, 321.

38 F. Yang, L.G. Moss and G.N. Phillips Jr., *Nature Biotechnol.*, 1996, **14**, 1246.

39 K. M. Solntsev, D. Huppert and N. Agmon, *J. Phys. Chem. A*, 1998, **102**, 9599.

40 K. M. Solntsev, D. Huppert and N. Agmon, *J. Phys. Chem. A*, 1999, **103**, 6984.

41 J. Dong, K. M. Solntsev and L. M. Tolbert, *J. Am. Chem. Soc.*, 2006, **128**, 12038-12039.

42 K. R. Siefermann, Y. Liu, E. Lugovoy, O. Link, M. Faubel, U. Buck, B. Winter and B. Abel, *Nat. Chem.*, 2010, **2**, 274.

43 J. Zhu, L. Paparelli, M. Hospes, J. Arents, J. T. M. Kennis, I. H. M. van Stokkum, K. J. Hellingwerf and M. Louise Groot, *J. Phys. Chem. B*, 2013, **117**, 11042.

44 W. Weber, V. Helms, J. A. McCammon and P. W. Langhoff, *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 6177.

45 R. Gepshtein, D. Huppert and N. Agmon, *J. Phys. Chem. B*, 2006, **110**, 4434.

46 A. Baldridge, K. M. Solntsev, C. Song, T. Tanioka, J. Kowalik, K. Hardcastle, L. M. Tolbert, *Chem. Commun.*, 2010, **46**, 5686.

47 M. S. Baranov, K. A. Lukyanov, A. O. Borissova, J. Shamir, D. Kosenkov, L. V. Slipchenko, L. M. Tolbert, I. V. Yampolsky and K. M. Solntsev, *J. Am. Chem. Soc.*, 2012, **134**, 6025.

48 M. S. Baranov, K. M. Solntsev, N. S. Baleeva, A. S. Mishin, S. A. Lukyanov, K. A. Lukyanov and I. V. Yampolsky, *Chem. Eur. J.*, 2014, **20**, 13234.

49 L. Wu and K. Burgess, *J. Am. Chem. Soc.*, 2008, **130**, 4089.

50 C. M. Megley, L. A. Dickson, S. L. Maddalo, G. J. Chandler and M. Zimmer, *J. Phys. Chem. B*, 2008, **113**, 302.

51 I. V. Polyakov, B. L. Grigorenko, E. M. Epifanovsky, A. I. Krylov and A. V. Nemukhin, *J. Chem. Theor. Comput.*, 2010, **6**, 2377.

52 M. E. Martin, F. Negri and M. Olivucci, *J. Am. Chem. Soc.*, 2004, **126**, 5452.

53 A. V. Bochenkova and L. H. Andersen, *Farad. Discuss.*, 2013, **163**, 297.

54 B. P. Cormack, R. H. Valdivia and S. Falkow, *Gene*, 1996, **173**, 33.

55 S. C. Daglio, N. Banterle, L. D'Alfonso, M. Collini, G. Chirico, *J. Phys. Chem. B*, 2011, **115**, 10311.

56 R. Mallik, J. B. Udgaonkar and G. Krishnamoorthy, *Proc. Ind. Acad. Sci. Chem. Sci.*, 2003, **115**, 307.

57 A. M. Saxena, J. B. Udgaonkar and G. Krishnamoorthy, *Protein Sci.*, 2005, **14**, 1787.

58 D. S. Larsen, M. Vengris, I. H. M. van Stokkum, M. A. van der Horst, F. L. de Weerd, K. J. Hellingwerf and R. van Grondelle, *Biophys. J.*, 2004, **86**, 2538.

59 A. V. Bochenkova, B. Klærke, D. B. Rahbek, J. Rajput, Y. Toker and L. H. Andersen, *Angew. Chem. Int. Ed.*, 2014, **53**, 9797.

60 F.-Y. Jou and G. R. Freeman, *J. Phys. Chem.*, 1979, **83**, 2383.

61 E. C. Carroll, S.-H. Song, M. Kumauchi, I. H. M. van Stokkum, A. Jailaubekov, W. D. Hoff and D. S. Larsen, *J. Phys. Chem. Lett.*, 2010, **1**, 2793.

62 M. Vengris, I. H. M. van Stokkum, X. He, A. F. Bell, P. J. Tonge, R. van Grondelle and D. S. Larsen, *J. Phys. Chem. A*, 2004, **108**, 4587.

63 J. F. Wishart and P. Neta, *J. Phys. Chem. B*, 2003, **107**, 7261.

64 W. M. Draper and D. G. Crosby, *J. Agric. Food. Chem.*, 1983, **31**, 734.

65 B. H. Milosavljevic and O. I. Micic, *J. Phys. Chem.*, 1978, **82**, 1359.

66 R. M. Dickson, A. B. Cubitt, R. Y. Tsien and W. E. Moerner, *Nature*, 1997, **388**, 355.

67 M. Drobizhev, C. Stoltzfus, I. Topol, J. Collins, G. Wicks, A. Mikhaylov, L. Barnett, T.. E. Hughes and A. Rebane, *J. Phys. Chem. B*, 2014, **118**, 9167.

68 J. N. Henderson, H. W. Ai, R. E. Campbell and S. J. Remington, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 6672.

69 B. Huang, M. Bates and X. W. Zhuang, *Annu. Rev. Biochem.*, 2009, **78**, 993.