Protonation States of Molecular Groups in the Chromophore-Binding Site Modulate Properties of the Reversibly Switchable Fluorescent Protein rsEGFP2

*Bella L. Grigorenko [a,b], Tatiana Domratcheva [a,c], Igor V. Polyakov [a,b], Alexander V. Nemukhin [a,b] ***

[a] Department of Chemistry, Lomonosov Moscow State University, Moscow, 119991, Russia

[b] Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, 119334, Russia

[c] Max Planck Institute for Medical Research, Department of Biomolecular Mechanisms, Jahnstrasse 29, 69120 Heidelberg, Germany

* Corresponding author:

Prof. Alexander V. Nemukhin, Chemistry Department, Lomonosov Moscow State University, Leninskie Gory 1/3, Moscow, 119991, Russian Federation

Phone: +7-495-939-10-96

E-mail: anemukhin@yahoo.com; anem@lcc.chem.msu.ru

† Electronic supplementary information (ESI) available.

KEYWORDS: fluorescent proteins, rsEGFP2, photoswitching, absorption spectra, proton wires, hydrogen bonds, quantum chemistry, QM/MM.

Abstract

The role of protonation states of the chromophore and its neighboring amino acid side chains of the reversibly switching fluorescent protein rsEGFP2 upon photoswitching is characterized by molecular modeling methods. Numerous conformations of the chromophore-binding site in computationally derived model systems are obtained using the quantum chemistry and QM/MM approaches. Excitation energies are computed using the extended multiconfigurational quasidegenerate perturbation theory (XMCQDPT2). The obtained structures and absorption spectra allow us to provide interpretation of the observed structural and spectral properties of rsEGFP2 in the active ON- and inactive OFF-states. To identify intermediates along the routes of chromophore transformations between the ON- and OFF-states, molecular dynamics trajectories with the QM/MM potentials are examined. The results demonstrate that in addition to the dominating anionic and neutral forms of the chromophore, the cationic and zwitterionic forms may participate in the photoswitching of rsEGFP2. Conformations and protonation forms of the Glu223 and His149 side chains in the chromophore-binding site play an essential role in stabilizing specific protonation forms of the chromophore.

Introduction

Reversibly photoswitchable fluorescent proteins (rsFP) are important markers used in advanced bioimaging techniques applicable to living-cell studies.[1–4] They can be reversibly photoconverted between two optical states termed ON (active, bright) and OFF (inactive, dark) depending on their spectral properties. Despite wide applications in biotechnology, the mechanism of action of rsFPs is still under active investigation. In this work, we focus on an engineered member of the green fluorescent protein (GFP) family, rsEGFP2.[5–9] Photoswitching in this protein is attributed to the cis-trans isomerization of its chromophore (Chro). The ON→OFF transition is initiated by the 488 nm irradiation; it proceeds with a low quantum yield (QY) of 4%.[7] The back transfer OFF→ON takes place either after the 400 nm irradiation with the QY an order of magnitude greater (40%) or by the thermal activation. [7]

Fig. 1 shows the chromophore (Chro) of rsEGFP2 matured from the tripeptide Ala65-Tyr66-Gly67. According to the experimental results,[6,7] Chro occurs in the cis conformation (left panel) in the fluorescent active ON-state, and in the trans conformation (right panel) in the dark OFF-state. The Chro resides in the chromophore-binding site that includes the side chains of Arg97, His149, Thr204, Ser206, Glu223 and water molecules.
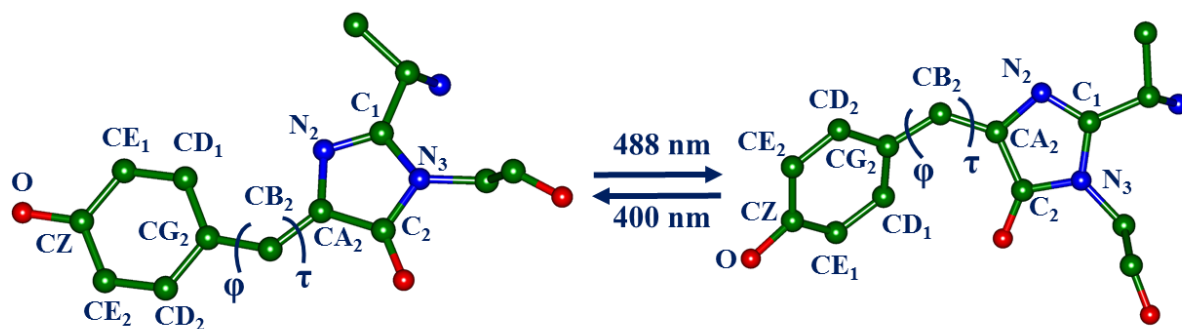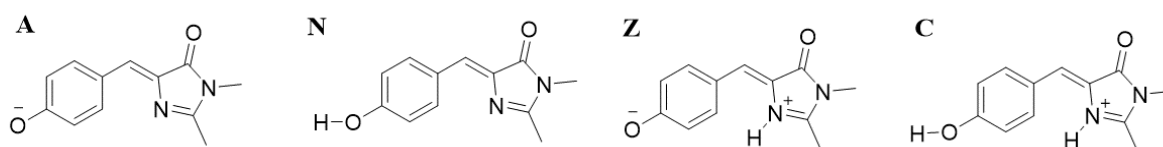

Figure 1. The rsEGFP2 chromophore in the ON- and OFF-states. Here and in other figures, carbon atoms are colored green, oxygen – red, nitrogen – blue.

Several structures [6–8] of rsEGFP2 have been deposited to the Protein Data Bank. They were determined by conventional cryocrystallography as well as by serial femtosecond X-ray crystallography (SFX) at room temperature. These studies provided not only the high-resolution structure of the ON-state, but also resolved the structure of the photoconverted OFF state obtained by irradiation of the crystallized protein in the ON-state. Moreover, the SFX studies allowed mapping the structural pathway of the photochemical OFF → ON switching.[7,8]

The structural data coupled with the results of spectral studies were used to propose a mechanism of chemical transformations of Chro inside the chromophore-binding site. [5–8] This proposal assumes changes in the protonation forms of Chro and of the nearby molecular groups. In particular, they favor the occurrence of Chro in the cis-anionic form in the ON-state and in the trans-neutral form in the OFF-state. From the very first studies of the GFP-like chromophores, the anionic (**A**) and neutral (**N**) forms correspond to the unprotonated and protonated oxygen of the phenolic ring. In other considered forms, [10] the cationic (**C**) and zwitterionic (**Z**) forms, an additional proton is attached to the imidazolinone ring nitrogen as compared to the neutral and anionic forms, respectively, as shown in Scheme 1.

Scheme 1. From left to right – anionic (**A**), neutral (**N**), zwitterionic (**Z**) and cationic (**C**) forms of the GFP-type chromophores shown here in the cis-conformation.

It is well known [11] that the anionic and neutral forms of the GFP-like chromophores demonstrate distinct absorption bands around 500 nm and 400 nm, respectively. In the case of rsEGFP2, a notable band shift from 478 nm to 408 nm observed [5] in the ON-state upon decreasing pH from 9 to 5 demonstrates an equilibrium involving various protonation forms. Similar spectral changes are observed in the ON → OFF photoswitching, indicating a change in the Chro pKa upon cis-trans isomerization in the protein. [5] The Chro protonation state alteration as well as of the surrounding Chro molecular groups was also put forward by studies of the OFF → ON transformations of rsEGFP2. [7,8]

When discussing protonation in rsEGFP2, we note the very recent paper [12] that reports results of the NMR studies for a related GFP-like reversibly switchable protein rsFolder with the same composition of the chromophore and only few variations in amino acid residues neighboring to the chromophore. [6] The obtained data show that the relative population of the chromophore ON- and OFF-states vary in a complex pH-dependent manner. The protonated and deprotonated forms of Chro and nearby residues play an important role in hydrogen-bond dynamics, altering the relative stability of the cis and trans configurations of Chro. Protonation-dependent dynamics determines the rsFolder contrast, [12] an important parameter in applications of the reversibly switchable fluorescent probe.[5]

Therefore, the distinct properties of the cis and trans forms of Chro revealed in recent experimental studies of rsEGFP2 and rsFolder [5–8,12] are linked to a specific location of protons at various molecular groups in the chromophore-binding sites. This knowledge is indispensable to understand mechanistic details of photoswitching. Although quantum-based molecular modeling may provide an essential contribution to elucidate this issue, scarce computational studies are devoted to this problem. Ref. [7] includes a fragment describing QM/MM simulations of rsEGFP2 focused on the OFF → ON pathway after irradiation of the OFF-state. The author focused on the excited state dynamics of the twisted trans-neutral Chro, i.e. the geometry configuration revealed in the SFX study, at a modest theoretical level.

In this work, we relate the observed structural data and pH dependent spectral changes in rsEGFP2 to specific protonated forms of Chro in cis and trans configurations and protonation and conformation changes of the critical residues Glu223 and His149 forming hydrogen bonds with Chro. We apply full-protein models treated at the quantum mechanics/molecular mechanics (QM/MM) level and cluster models treated at the quantum chemistry level to simulate ground state properties and excitation energies of rsEGFP2. The applied high-level approaches allowed us to identify a number of conformations in the chromophore-binding site, which may contribute to photophysical properties of rsEGFP2. Following the results of simulations, consistent with the available experimental data, we put forward a hypothesis highlighting the role of protonation forms in reversible photoswitching. In particular, we propose that the cationic cis form of Chro is involved in the photoswitching of the ON-state leading to the OFF-state with the neutral trans form of Chro. The decay of the photoexcited OFF-state with the trans Chro proceeds over a set of structures containing the neutral cis Chro.

Models and Methods

Two types of model systems were constructed in this work, both prompted by the available crystal structures of rsEGFP2. At initial stages of the project, large molecular clusters composed of about 225 atoms were created following the PDB ID 5DTX (containing cis Chro),[6] PDB ID 5O89 (containing cis Chro),[7] and PDB ID 6T39 (containing cis and trans Chro)[8] templates. These clusters, which simulate properties of the sole chromophore-binding site, include the extended Chro-containing fragment Thr63-Leu71, surrounding groups from the residues Arg97, His149-Val151, Thr204-Ser206, Glu223 and water molecules. Protons were variably assigned to Chro, His149 and Glu223, thus mimicking different protonation forms in these species at high and low pH in the related

experiments. Correspondingly, total charges of the clusters were set as -1, 0 and +1. Equilibrium geometry parameters were optimized at the density functional theory (DFT) level using the PBE0 functional [13] and the cc-pVDZ basis set. Calculations of transition energies between the lowest singlet states S_0 and S_1 as well as the corresponding transition dipole moments were performed with the extended multiconfigurational quasidegenerate perturbation theory in the second order (XMCQDPT2).[14] This method, which is based on the complete active space self-consistent field (CASSCF) wavefunctions, gained a solid reputation in accurate calculations of the excitation energies of biological chromophores.[15–17] In this work, the minimal active space featuring 2 electrons over 2 orbitals, CASSCF(2,2), was applied followed by the state-averaging over 3 first electronic states. The Firefly package [18] was used in all cluster model calculations. The use of cluster approaches in modeling properties of photoactive proteins is well documented in the literature, see, e.g. Ref. [19] Since the geometry optimization of atomic coordinates in clusters is carried out with a reference to the parent crystal structures by constraining the coordinates of the terminal heavy atoms, the obtained model systems describe the active sites corresponding to specific crystallization conditions.

Navigated by the obtained results of the clusters models, we carried out simulations of rsEGFP2 properties using QM/MM calculations by considering the entire protein system. We denote this series of models as the protein models. In this approach, the only source of atomic coordinates was the crystal structure of the ON-state PDB ID 5DTX (containing cis Chro).[6] Hydrogen atoms were added using molecular mechanics tools; the side chains of Arg and Lys were assumed as positively charged, the side chains of Glu and Asp as negatively charged. The model protein molecule was fully surrounded by explicit water molecules. All structures were optimized in QM/MM calculations using the DFT(PBE0/6-31G*) approach in QM and the AMBER force field in MM. A large fraction of the chromophore-binding site was assigned to the QM-part. It included 162 or 163 atoms from Chro, the residues Arg97, Glu223, Thr204, Gln205, Ser 206, His 149, Asn150, Val151, Gln184, Ser164, and 6 water molecules. First, the series of structures with the cis Chro was created as described below. Structures with the trans Chro were prepared manually when starting from those with the cis Chro and guided by the PDB 5DTY crystal structure. These calculations were performed by using the NWChem program.[20] Calculations of transition energies and transition dipole moments were performed with the XMCQDPT2 method on the CASSCF(16,12) wavefunctions using state-averaging over 10-15 routes; the same number of states was taken into account in the PT2 part. The Firefly package [18] was applied for molecular clusters corresponding to the QM-subsystem taken at the QM/MM optimized geometry configurations.

Molecular dynamics (MD) simulations with QM/MM potentials with the atom-centered basis functions and hybrid functionals in DFT are now available and prove their efficiency in modeling chemical reactions in proteins, e.g. Refs. [21,22] In this work, we perform calculations at the QM(PBE0/6-31G**)/MM(CHARMM) MD level using NAMD [23] and TeraChem [24] packages interfaced according to Ref. [25] The QM part comprised 81 atoms including Chro and the side chains of Arg97, His149 and Glu223.

Results and Discussion

We illustrate in Fig. 2 the key molecular groups in our model systems taking one of the obtained ON-state structures of rsEGFP2 as an example. The left panel presents a view of the solvated protein showing explicitly the chromophore and nearby residues Arg97, His149, Ser206 and Glu223. We emphasize that His149 is located close to the protein barrel surface, and its side chain is solvent accessible. The right panel in Fig. 2 shows the groups, which are in the focus of the present simulations. Here, the chromophore is in the anionic cis form and Glu223 is protonated.
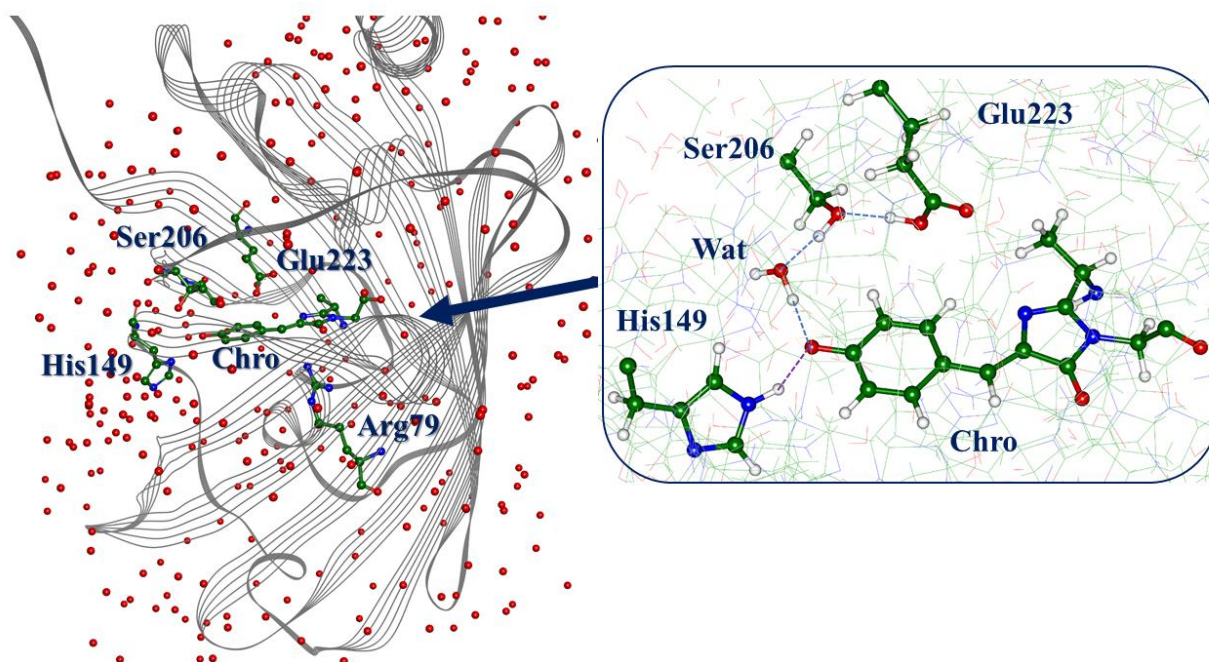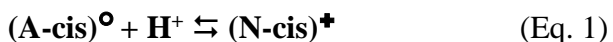

Figure 2. The protein model (left) and a fragment of the chromophore-binding site (right).

Different conformations and protonation forms of Chro, His149 and Glu223 will be characterized below. The Arg97 side will appear only in some of the subsequent figures. Although

evidence has been presented that hydrogen bonding between arginine at this position in GFP-like proteins and the chromophore carbonyl group may be important for photoisomerization, e.g. in Dronpa,[26] we do not find that this residue plays an important role in the stabilization of various protonation forms of rsEGFP2.

In this paper, we adopt the following notation of model systems created computationally. Letters **A, N, C, Z** refer to the conventional anionic, neutral, cationic and zwitterionic forms of the chromophore. Next, either **cis**, or **trans** configuration of Chro is designated. We describe here two series of model systems, which differ by an amount of protons in the area shown in the right panel in Fig. 2. According to proposal formulated in the experimental works,[5–8] the anionic and neutral forms of Chro are the immediate candidates for consideration of rsEGFP2 photoswitching. Correspondingly, we consider systems of the type shown in the right panel in Fig. 2 with the anionic Chro and systems with an added proton to the active site. We use the symbols "$^{\circ}$" and "$^{+}$" (plus a proton) to distinct systems of the first and second type. For instance, the right panel in Fig. 2 illustrates the **(A-cis)$^{\circ}$** system, whereas the notation **(N-cis)$^{+}$** refers to the system with an added proton, and it describes the neutral Chro. If necessary, the Roman numerals distinguish structures within the same type, which are ordered by their relative energies. In these terms, the symbol **(A-cis)$^{\circ}$-I** refers to the lowest-energy structure with the cis-anionic chromophore. According to the previous studies, [5–8] this structure dominates in the ON-state of rsEGFP2. The symbol **(N-trans)$^{+}$-I** refers to the lowest-energy structure with the neutral Chro in the trans-form, which presumably dominates in the OFF-state. In the following, the Roman numerals may be omitted from the symbolic names, if there is no ambiguity.

When studying a closely related protein rsFolder, Cristou et al.[12] concluded that solvent protons can easily migrate to the chromophore-containing pocket and protonate the hydroxybenzylidene ring of the chromophore "on a submilliseconds time scale". In line with this finding, which presumably holds for rsEGFP2 as well, the following equilibrium is addressed by our models of the ON-state

$$\textbf{(A-cis)}^{\circ} + \textbf{H}^+ \leftrightarrows \textbf{(N-cis)}^{+} \qquad \text{(Eq. 1)}$$

At neutral pH, the **(A-cis)$^{\circ}$** structure dominates, whereas at low pH, the **(N-cis)$^{\circ}$** structure is a predominant species in the ON-state. As shown in the left part in Fig. 2, the side chain of His149 most likely serves as a "gate" for migration of protons. Importance of proton wires connecting the

chromophore-containing area and the protein exterior in photochemistry of GFP-like proteins is documented in the literature. [27,28]

Structures and spectra of systems corresponding to the ON-state with the cis-Chro.

As described in Models and Methods, molecular model systems were constructed following motifs of the crystal structures of the ON-state of rsEGFP2 (PDB IDs 5DTX, 5O89) obtained at pH near 8, which implies the anionic state of the cis Chro in the series marked by the "$^{\circ}$" sign. The latter refers to the rsEGFP2 properties at neutral and high pH. Extensive search of minimum energy points on the ground state potential energy surface resulted in series of model structures. The right side panels (b-d) in Fig. 3 show chromophore-binding sites of three selected systems. The left panel (a) shows a superposition of the key molecular groups taken from the crystal structure PDB 5DTX and from two low-lying computationally derived structures with the cis-anionic Chro. Table 1 collects the relative energies, excitation energies and corresponding wavelengths and transition dipole moments of these systems. Supplementary Materials contain more data on these and other computationally derived structures.
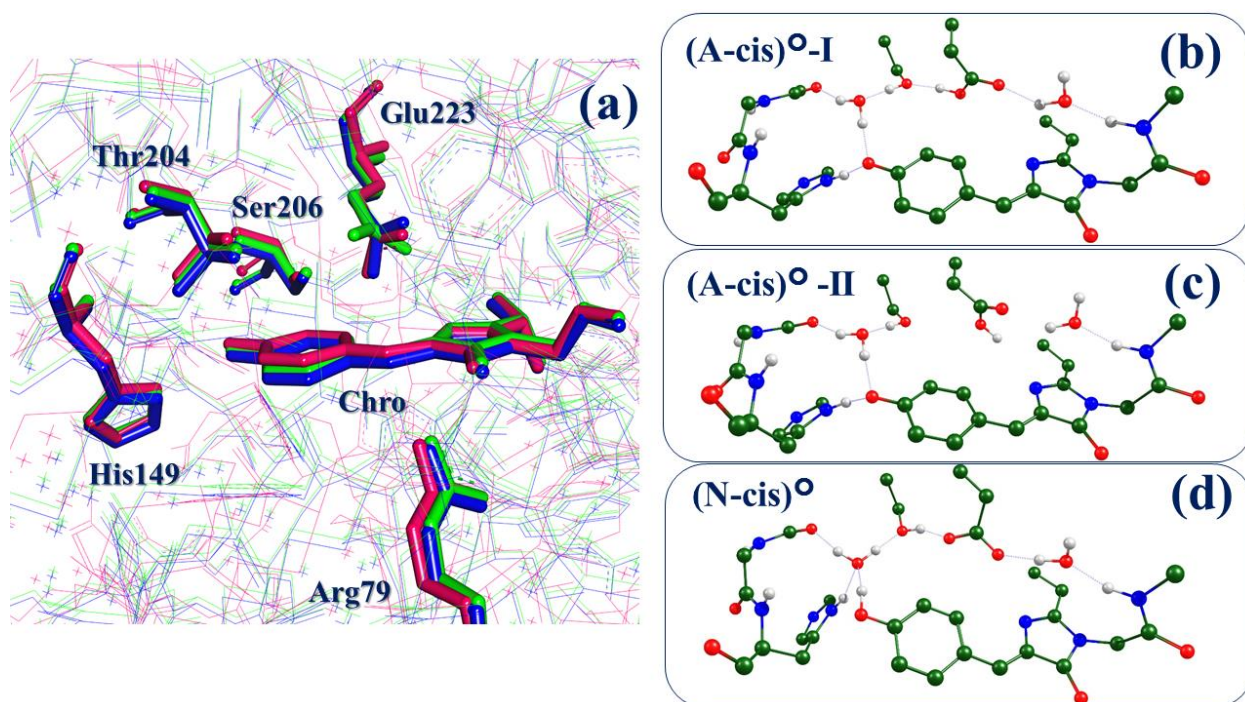

Figure 3. Selected molecular systems with the cis-Chro of the "$^{\circ}$"-series. Panel (a): superposition over the heavy atoms in the protein model systems **(A-cis)$^{\circ}$-I** (green sticks), **(A-cis)$^{\circ}$-II** (blue sticks)

9

and the crystal structure PDB ID 5DTX (red sticks). Panels (b, c, d) illustrate an arrangement of the hydrogen bonding patterns in the chromophore-bonding site in the systems **(A-cis)°-I**, **(A-cis)°-II**, **(N-cis)°**.

Table 1. Properties of selected model systems simulating conformations in the ON state of rsEGFP2 at neutral and high pH.

Property\System	(A-cis)°-I		(A-cis)°-II		(N-cis)°	
	Protein model	Cluster model	Protein model	Cluster model	Protein model	Cluster model
Relative energy, kcal/mol	0	0	0.2	2.2	7.1	6.7
Excitation energy, eV	2.52	2.49	2.65	2.38	3.04	2.94
λ_{max} , nm	491	498	468	521	408	422
Transition dipole moment, a.e.	3.8	3.3	3.4	3.4	1.2	2.8

We note that the lowest energy model system **(A-cis)°-I** computed with both models (the protein and cluster models) perfectly reproduces the structural and spectral experimental data.[5–7] The anionic phenolate is stabilized by the hydrogen bonds with His149, water and Thr204. We point out that discrepancies between computational and experimental distances between heavy atoms in the important hydrogen-bonded pairs, e.g. Chro-His149 and Ser206-Glu223, do not exceed 0.3 Å. The computed excitation energies (491 nm (2.52 eV) and 498 nm (2.49 eV)) are in a perfect agreement with the observed absorption band maximum at 478 nm (2.59 eV).[5,6] Structures with the anionic chromophore are the candidates to be excited by 488 nm light and to emit at about 500 nm in the related experiments.

We also note that simulations predict several structures, the energies of which are very close (within 3 kcal/mol) to the level of the major conformation **(A-cis)°-I**. The **(A-cis)°-II** structure shown in Fig. 3c is nearly isoenergetic to **(A-cis)°-I** according to our protein model. These structures differ by conformations of the Glu223 or Thr204 side chains (cf. panels (b) and (c) in Fig. 3). The critical residue Glu223 is protonated donating a hydrogen bond either to Ser206 or to the imidazolinone ring of Chro. Rotation of both side chains Glu223 or Thr204 induces slight changes in the excitation energies of the anionic Chro.

Simulations with the protein model predict occurrence of the low-lying structure **(Z-cis)°-I** with the chromophore in the zwitterionic form. This structure is formed by a rearrangement of the Glu223 side chain concomitant with proton transfer. Its energy is only 1.5 kcal/mol above the level of the major conformation **(A-cis)°-I**, and its excitation energy 2.32 eV corresponds to the red-shifted absorption band maximum at 534 nm and the corresponding transition dipole moment is 3.9 a.e. In previous simulations of GFP-like proteins, Schäfer et al.[29,30] proposed that photophysical properties of the chromoprotein asFP595 from the GFP family could be described by consideration of the zwitterion as the dominant chromophore form in the ground electronic state. This conclusion was obtained in calculations using the ZINDO approach; however, it was not confirmed in the subsequent *ab initio* type modeling of electronic transitions in asFP595, but structures with the zwitterionic chromophore were obtained in the excited state.[28] Here, we obtain a low lying local minima on the ground state potential energy surface with the zwitterionic Chro for rsEGFP2.

The lowest-energy structure with the cis-anionic Chro shown in Fig. 3b is interesting in the sense that it illustrates features of the well-known proton transport wire in GFP,[11] which extends here over the phenolic ring of Chro, water molecule Wat, Ser206 and Glu223. In the wild-type GFP with the Chro precursor Ser65-Tyr66-Gly67, the structure with the cis-neutral Chro is slightly lower in energy (about 1 kcal/mol according to the calculations in Ref. [31]) than that with the cis-anionic Chro obtained by concerted proton transfer from the phenolic oxygen to Glu222 via Ser205. In the case of rsEGFP2, a structure **(N-cis)°** with the neutral Chro prepared by proton transfer starting from **(A-cis)°-I** is illustrated in Fig. 3c and characterized in Table 1. This structure is about 7 kcal/mol higher in energy than **(A-cis)°-I.** The estimated energy difference disfavors appearance of **(N-cis)°** in the ON-state of rsEGFP2. This is consistent with the observation that the Ser65Ala replacement on the way from GFP to rsEGFP2 disrupts the hydrogen-bond pattern near Glu223. This is also consistent with the pKa value of the cis-form of Chro in rsEGFP2 [5] and demonstrates stabilization of the anionic cis Chro by the protein, including interactions with the residues forming the proton wire.

The computed excitation energy of **(N-cis)°** (3.04 eV or 2.94 eV (Table 1)) corresponding to the wavelengths 408 nm or 423 nm match the observed [5] band maximum near 400 nm in the ON-state of rsEGFP2 at pH ~ 5. However, we think that this experimental band building up at low pH values, which apparently refers to the neutral Chro, should be assigned to the model system **(N-cis)⁺**, i.e. to the system with a proton added to the chromophore-binding site according to the above-mentioned equilibrium of Eq. 1. The pKa estimate (~ 6 units) of the chromophore in the ON-state, [5] assumes that about 10% of the protonated state is expected to be present at neutral pH. The neutral form of Chro

can appear in the ON-state of rsEGFP2 either by protonation with participation of His149 or by protonation of the proton transfer wire. Both ways were explored in the present simulations. We prepared the protein and cluster models corresponding to the **(N-cis)$^+$** species, optimized their geometry parameters and evaluated the absorption band maxima. The most important results are illustrated in Fig. 4 and listed in Table 2.
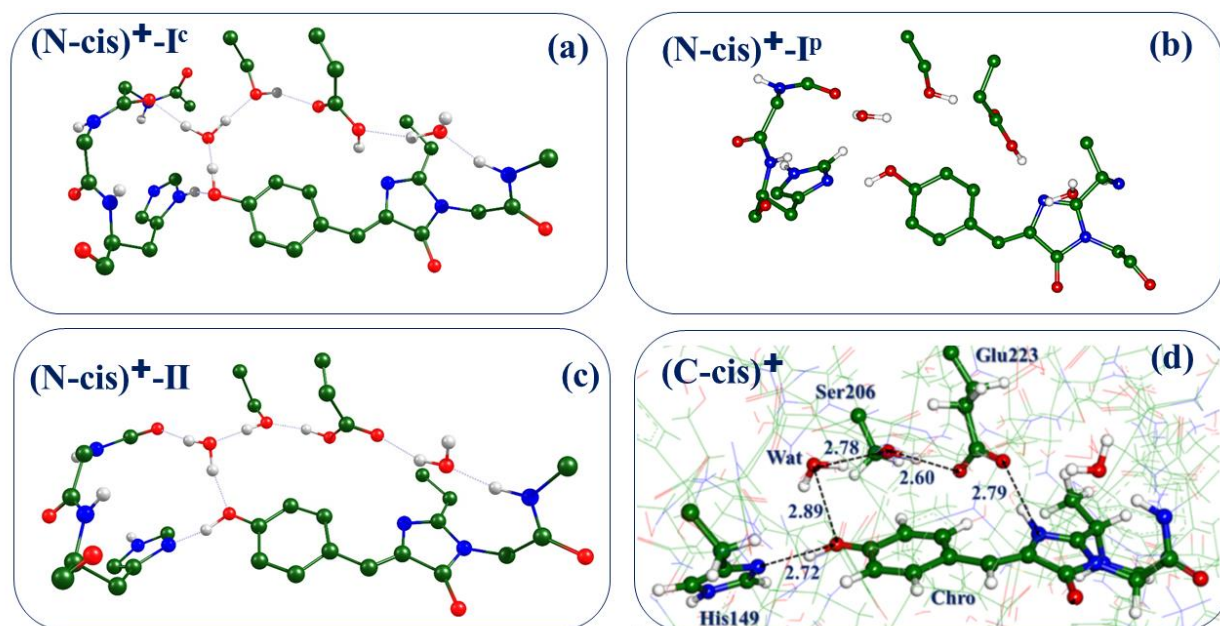

Figure 4. Selected molecular models with the cis-Chro with an extra proton in the system. Panels (a - c) illustrate an arrangement of the key groups in the chromophore-binding site in the **(N-cis)$^+$** model systems. Panel (d) shows the system with the cationic Chro. Distances between heavy atoms are given in Å.

Table 2. Properties of selected model systems simulating conformations in the ON state of rsEGFP2 at low pH.

Property\System	(N-cis)$^+$-I^p	(N-cis)$^+$-I^c	(N-cis)$^+$-II		(C-cis)$^+$	(A-cis)$^+$
Relative energy, kcal/mol	Protein model	Cluster model	Protein model	Cluster model	Protein model	Cluster model
	0	0	2.0	1.2	1.2	7.5
Excitation energy, eV	2.95	2.98	2.93	3.09	2.70	2.92
λ_{max} , nm	420	416	423	401	460	424
Transition dipole moment, a.e.	1.6	2.9	1.2	2.8	3.7	2.9

The side chain of His149 is located fairly close to the protein surface (see Fig. 2); therefore, it is reasonable to assume that His149 may serve as a proton shuttle accepting a proton from the solvent at the initially deprotonated imidazole nitrogen and donating a proton initially bound to another imidazole nitrogen to the phenolic oxygen of Chro. The model cluster with the anionic chromophore **(A-cis)⁺**, which contains the cationic protonated His149 side chain (Table 2), can be located, but it is 7.5 kcal/mol higher in energy than the model with the protonated phenolate Chro and neutral His149.

In this series of calculations, we obtained slightly different conformations of the lowest energy system **(N-cis)⁺-I** in the protein model and in the cluster model. Correspondingly, we add superscripts "**p**" or "**c**" to the system notations. The most important fragments of the systems **(N-cis)⁺-Iᵖ** and **(N-cis)⁺-Iᶜ** are presented in the panels (a) and (b) in Fig. 4. These structures show different hydrogen bonding patterns in the His149-Chro-Wat fragment. Despite this dissimilarity, both types of models agree that formation of a hydrogen bond between Glu223 and the imidazolinone ring of Chro is energetically favorable, as indicated by comparing energies of the **(N-cis)⁺-I** and **(N-cis)⁺-II** models (see Fig. 4 and Table 2). These structures differ by conformations of the Glu223 and Ser206 side chains. The computed excitation energies of **(N-cis)⁺-I** and **(N-cis)⁺-II** in all models are fairly close (Table 2) and match the experimental absorption band maximum at 400 nm (3.10 eV) observed in rsEGFP2 at low pH. [5]

Importantly, the protein model **(C-cis)⁺** (Fig. 4d) predicts that a structure with the cationic Chro stabilized by hydrogen-bond patterns is only 1.2 kcal/mol higher in energy than **(N-cis)⁺-I**. The excitation energy of **(C-cis)⁺** is 2.70 eV (460 nm), what makes this conformation a likely candidate for consideration in protein photophysics for the 488 nm photoexcitation. Involvement of the cationic GFP-like chromophore in molecular processes in solution and proteins has been proposed in earlier works, see, e.g. Refs. [32,33] Structures with the cationic Chro were identified previously for the kindling fluorescent protein (KFP),[28] however, at the excited state potential energy surface, whereas the **(C-cis)⁺** minimum for rsEGFP2 is located on the ground state potential energy surface.

Structures and spectra of systems modeling the OFF-state with the trans-Chro

Following the crystallography studies, two structural variants are proposed in the OFF-state of rsEGFP2; one with the highly twisted Chro turned away from the His149 side chain (called *Trans1* in Ref. [8], and a nearly planar Chro hydrogen-bonded to His149 (called *Trans2*). [8] In the present

simulations, several conformations of chromophore-binding sites with the trans Chro were considered after geometry optimization of the corresponding model systems. In Supplementary Materials we collect cluster models with various compositions of the chromophore-containing pocket. In the main text, we compare *Trans1* and *Trans2* containing active sites with the same number of protons as in the series **(A-cis)⁺**, **(N-cis)⁺** and **(C-cis)⁺** in the ON-state simulations (see Fig. 4 and Table 2). Fig. 5 and Table 3 present the results of the most important systems with the trans Chro.

Fig. 5a illustrates the protein model system **(N-trans)⁺-II** with the twisted Chro corresponding to the *Trans1* variant. The obtained values of the conventional torsional angles at the bridge between the phenolic and the imidazolinone rings φ and τ (see Fig. 2) are -31° and -5°, respectively, which are consistent with the crystal structures in the OFF-state as well as with the results (-46 ± 11° and 2 ± 8°) of QM/MM-based simulations described in Ref. [7]. As shown in Table 3 the system with the twisted Chro **(N-trans)⁺-II** lies slightly higher in energy than that with the planar Chro, **(N-trans)⁺-I**, illustrated in Fig. 5b. Superposition of the protein model **(N-trans)⁺-I** with one of the conformations in the crystal structure PDB 5O8A depicted in Fig. 5e provides support to the present simulation results.
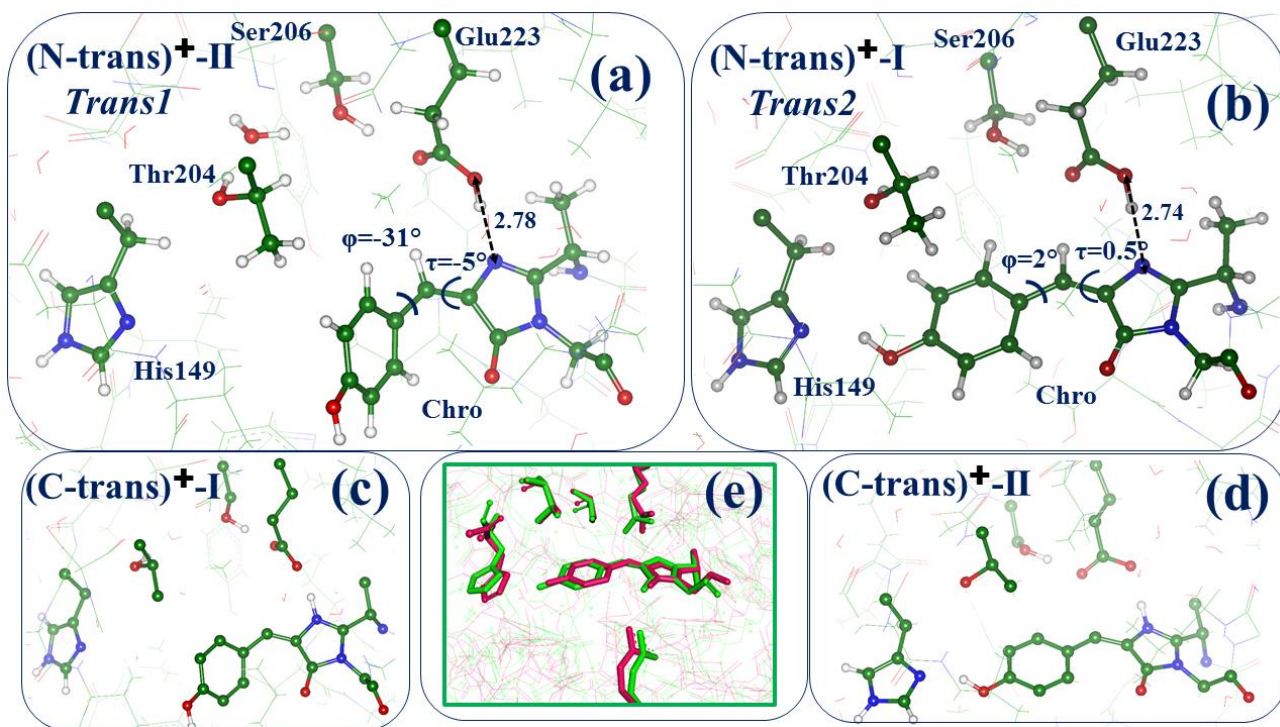

Figure 5. Selected structures with the trans-Chro. We point out designation of the structures *Trans1* and *Trans2* as in Ref. [8] along with the notations used in this work. In panel (e) we show superposition

of the **(N-trans)$^+$-I** model system, which correlates with the *Trans2* variant, and the crystal structure PDB ID 5O8A with the planar Chro.

Table 3. Properties of selected model systems modeling the OFF-state of rsEGFP2.

Property\System	**(N-trans)$^+$-I** (*Trans2*) Planar Chro		**(N-trans)$^+$-II** (*Trans1*) Twisted Chro		**(C-trans)$^+$-I** Twisted Chro	**(C-trans)$^+$-II** Planar Chro
	Protein model	Cluster model	Protein model	Cluster model	Protein model	Protein model
Relative energy, kcal/mol	0	---	1.2	---	1.0	1.3
Excitation energy, eV	2.95	2.91	2.93	3.14	2.86	2.72
λ_{max} , nm	420	426	423	395	434	456
Transition dipole moment, a.e.	1.2	2.8	1.2	2.4	1.9	2.4

The excitation energies and transition dipole moments computed for the **(N-trans)$^+$-I** and **(N-trans)$^+$-II** model systems in both our models correspond to bands with the wavelengths 395 ÷ 426 nm, well consistent with the observed absorption band maximum at 408 nm.[5]

In Figs. 5c, 5d and in Table 3 we show the computed parameters of model systems with a cationic Chro in the twisted conformation **(C-trans)$^+$-I** and in the planar conformation **(C-trans)$^+$-II**. Remarkably, the protein models place the systems with the cationic Chro very close to the energy level of the lowest energy isomer. Occurrence of systems with the cationic Chro is not unexpected in rsEGFP2; the short (Glu223)O-N(Chro) distances, 2.78 and 2.74 Å (Figs 5a, 5b) suggest easy migration of a proton along this hydrogen bond. The excitation energy in these C-trans models is somewhat reduced as compared to the N-trans models (Table 3).

Theoretical reconstruction of the absorption spectral bands in the ON-state of rsEGFP2

The model systems described above provide an insight in the observed features of the absorption spectral bands at different pH. The central part in Fig. 6 reproduces contours of the experimental bands recorded in the ON-state; [5] the panels in the upper and lower rows in Fig. 6

illustrate the structures of model systems, which according to our simulations, contribute to the observed bands.
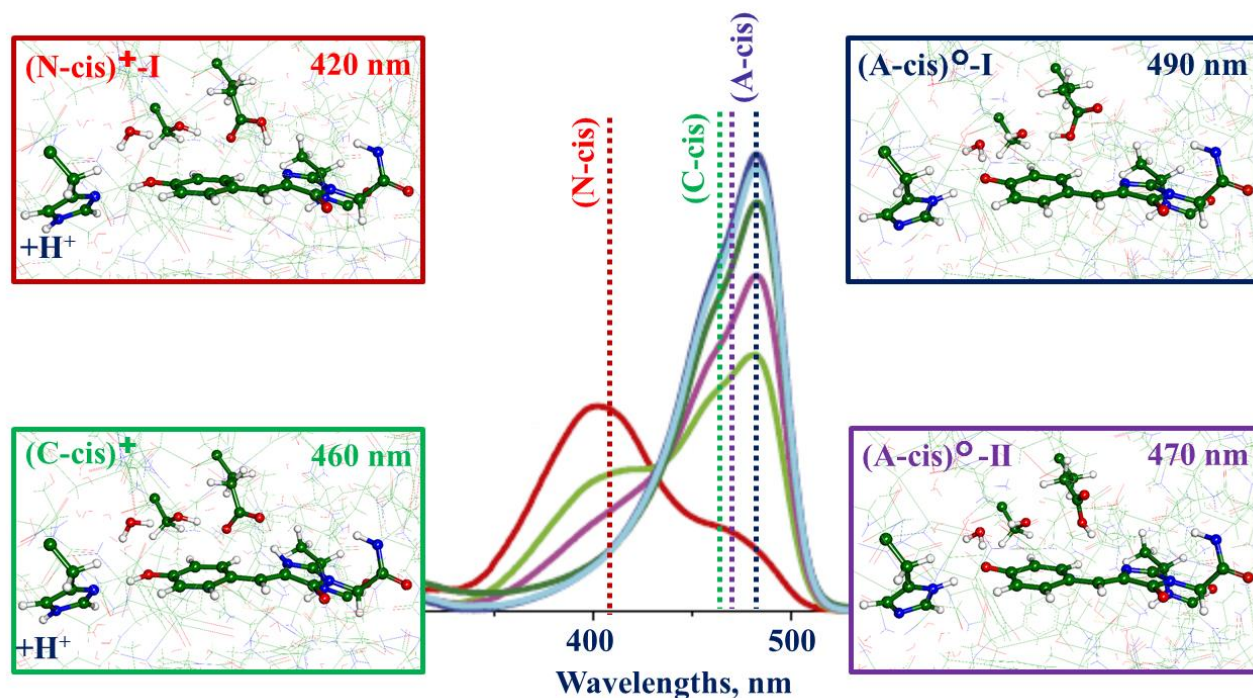

Figure 6. Reconstructing the pH dependence of the absorption bands in the ON-state of rsEGFP2 by the results of the present simulations. The dotted vertical lines in the central panel, which are colored dark blue, magenta, light green, and red, respectively, mark positions of the computed absorption band maxima of the structures shown in the framed panels. The color of every frame and of the characters inside the frame is the same as of the corresponding vertical line. Experimental data digitized and replotted from Ref. [5] show the shapes of the absorption bands at pH 9 (dark blue), pH 8 (light blue), pH 7.5 (dark green), pH 6.5 (magenta), pH 6 (light green), pH 5 (red).

The shapes of absorption bands corresponding to high pH, which apparently refer to the protein conformations with the cis-anionic Chro, show a pronounced shoulder between 450 and 500 nm shifted to lower wavelengths from the major absorption peak. This is typical for GFP-like proteins, and possible explanations connect this shoulder with vibrational satellites or with contributions from different protein conformations (see Refs. [11,34,35]). The major peak and, to a lesser extent, the shoulder decrease with lowering pH. According to the present simulations, we propose that the major peak is due to the **(A-cis)°-I** conformation (upper right panel in Fig. 6), whereas the shoulder is due to the **(A-cis)°-II** conformation (lower right panel in Fig. 6). The major peak at low pH (the red line in the experimental band) corresponds to the structure with the cis-neutral Chro **(N-cis)⁺-I** (upper left panel

16

in Fig. 6). The corresponding band also includes a shoulder, this time shifted to higher wavelengths. We propose to relate this feature to the presence of **(C-cis)⁺** species with the band maximum at 460 nm and the relative energy of 1.2 kcal/mol above the **(N-cis)⁺-I** system. Only a small fraction of proteins may adopt the conformation of the **(C-cis)⁺** model at high and neutral pH; however, its weight at low pH may be high enough.

QM/MM MD simulations of the ground state trans-cis transformation of Chro

The model systems described above also provide an insight in the observed features of the OFF → ON transformations. [8] According to the proposal formulated in Ref. [8], the reaction route initiated by the excited-state structure with the twisted Chro in the trans-neutral form, crosses the conical intersection point and proceeds on the ground state energy surface over possible intermediates containing cis-neutral Chro and cis-anionic Chro. These transformations that depend on the stability of hydrogen-bonding interactions in the chromophore-binding site are addressed in our QM/MM MD simulations.

As seen in Figs. 2-5, the hydrogen-bonded proton wires connect molecular groups in the chromophore-containing pocket. They include the conventional GFP-type proton transport wire Chro-Wat- Ser206-Glu223 as well as the extended wire with inclusion of His 149 (see Fig. 2). It is instructive to simulate dynamics of model systems upon OFF → ON transformations to clarify proton redistribution around the chomophore. The recently developed QM/MM MD approaches, which allow using the hybrid functionals in DFT and the atom-centered basis functions, provide suitable modeling tools. [25] The proton shuttle along the conventional GFP-type proton transport wire has been extensively studied and simulated; see, e.g. one of the recent papers.[36] Here, we executed and analyzed ground-state QM/MM MD trajectories starting from the model system **(N-cis)°** (see Fig. 3). In these simulations, trajectories arrive to lower-energy systems with the cis-anionic Chro during first 200 fs. More interesting results were obtained in QM/MM MD simulations of the extended proton wire. Specifically, we started from the protein conformation corresponding to the structure PDB ID 5O8B (conformer D) [7] with the highly twisted chromophore. In our terms, the starting point corresponds to the **(N-trans)⁺-II** system. A part of computed trajectories describe the return of the system back to the OFF-state with the trans-neutral Chro, whereas some trajectories finally relaxed to the system with the cis chromophore, i.e. to the **(N-cis)⁺-I** conformation. The panels in Fig. 7 illustrate major features of such a pathway along the 40 ps trajectory.
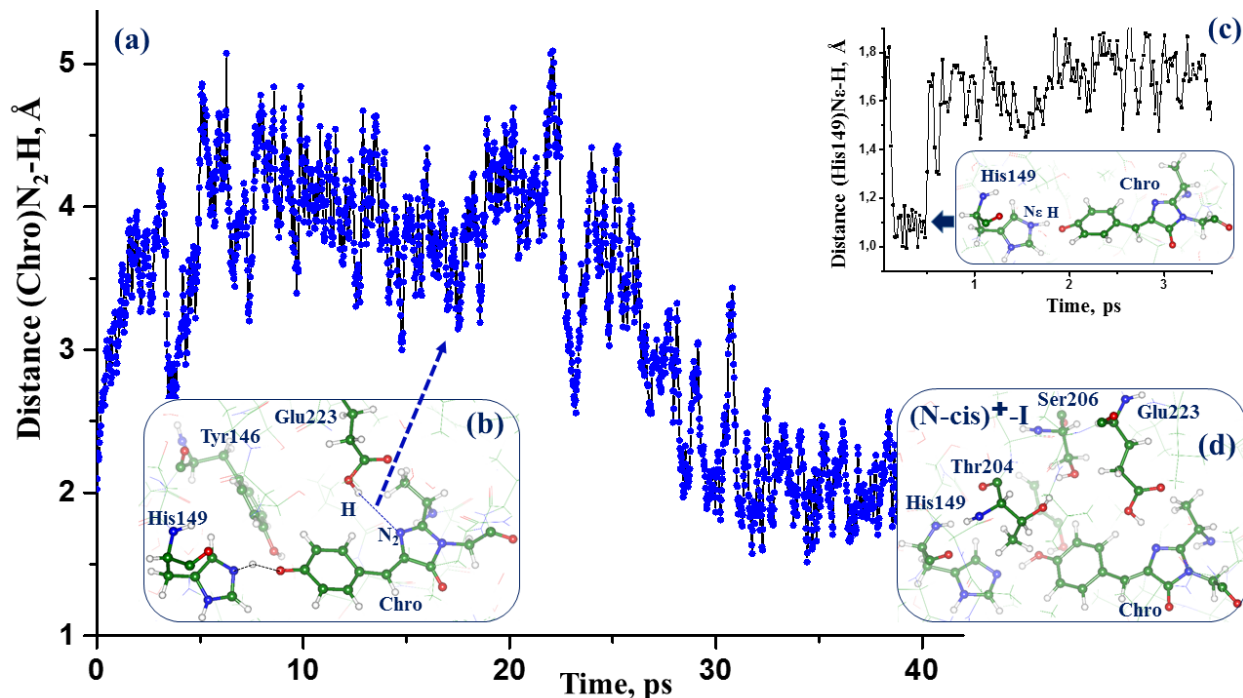
Figure 7. Features along the QM/MM MD trajectories relevant to the ground-state OFF→ON transformations. Panel (a) shows the evolution of the distance between the N_2 atom of Chro and the hydrogen atom attached to the Oε atom of Glu223; panel (b) illustrates a related fragment of the model system. Panel (c) shows the evolution of the distance between the Nε atom of His149 and the hydrogen atom shared between His149 and Chro. Panel (d) illustrates the destination point, which corresponds to the **(N-cis)$^+$-I** system.

According to these simulations, the transformation of the twisted trans form of Chro to the cis form occurs at the very first time steps; already after 200 fs we observe a slightly twisted cis-Chro, which is getting planar at 500 fs. At this time scale, we see proton migration between phenolic oxygen of Chro and His149 (Fig. 7c). The graph in Fig. 7a shows that the break of the hydrogen bond between Glu223 and Chro takes place during first 25 ps; the side chain of Glu223 drifts far from the chromophore (up to 4 Å). The side chains of Thr204 and Tyr146 closely interact with Chro; rotations of the corresponding molecular groups of Glu223, Thr204, Tyr146 take place. However, after 40 ps, the system arrives to a stable conformation corresponding to the **(N-cis)$^+$-I** model system (Fig. 7d). Further stabilization, at a considerably longer time scale that is not accessible with the QM/MM MD method, but predicted by our protein and cluster model geometry optimization, would require

migration of proton from the His149 side chain to the bulk leaving the system in the **(A-cis)°-I** conformation as presumed in Refs.[7,8]

Discussion of the switching pathways

As demonstrated above, several close in energy protein conformations contribute to the observed structural and spectral properties of rsEGFP2 in the ON-state. The **(A-cis)°-I** and **(A-cis)°-II** model systems are important, featuring rearrangements in conformations of the groups constituting proton transport wires. These rearrangements are in line with the fast proton exchange in the chromophore-binding site as suggested by the NMR characterization of the related protein rsFolder.[12] In particular, the **(A-cis)°-II** model is a precursor of the isomers with the neutral and cationic Chro **(N-cis)⁺** and **(C-cis)⁺**, which may form upon protonation of Chro, most likely, with the help of the His149 side chain. These protonated forms constitute a minor fraction of species at neutral pH. Remarkably, the excitation energy in **(C-cis)⁺** corresponds to the absorption band maximum at 460 nm close to that of the anionic form at 478 nm. The corresponding energy value (2.70 eV) is close to the experimental excitation energy 2.54 eV related to the 488 nm wavelength [5] required for photoswitching. Hence, the minor fraction corresponding to the **(C-cis)⁺** system is likely to be excited along with the major fraction in the **(A-cis)°** conformation.

According to our previous studies,[37] the GFP-like chromophores in the anionic form are stable against cis-trans isomerization in the photoexcited state, because the pathway from the corresponding conical intersection point rather leads to systems with the initial cis form. In the case of rsEGFP2, the photoswitching is unlikely applicable to the most populated systems with the anionic Chro, then the corresponding QY would be higher that the observed 4%.[5] In contrast, chromophores with the neutral Chro form readily undergo cis-trans photoinduced isomerization. [37] However, in the case of rsEGFP2 this cannot be accomplished by the 488 mn (2.54 eV) irradiation, since the excitation energy of the neutral form is substantially larger and amounts to 3.1 eV (400 nm). We propose that the low QY of the photoswitching is consistent with a hypothesis that a small, but sufficient fraction of the protein absorbing at longer wavelength undergoes photoisomerization. Such fraction may correspond to the **(C-cis)⁺** conformation of the Chro-containing pocket undergoing photoconversion to the trans-configuration of Chro either in the cationic or neutral form. We suggest that the structure **(C-cis)⁺** is a likely candidate to account for the ON → OFF photoswitching in rsEGFP2. Irradiation by 488 nm initiates excitation and isomerization of **(C-cis)⁺** eventually leading to the **(N-trans)⁺-I**. Participation of the protonated **(C-cis)⁺** form of Chro in photoconversion also suggests that the isomerization yield

may increase by lowering pH. To support this proposal, we note the results of the recent studies of the related protein rsFolder; the graphs shown in Fig. 6 in Ref. [12] demonstrate that the efficiency of cis-trans isomerization upon 488 nm illumination increases with lower pH.

Our QM/MM and QM/MM MD simulations are consistent with the proposal that on the descent from the excited OFF-state, the so-called "cis-protonated states" [7] may appear. However, according to the energies of our optimized models, we expect a mixture of **(N-cis)⁺/(A-cis)°** species in the recovered ON-state rather than the occurrence of "cis anionic ON-state" as assumed in Ref. [7] A predominant fraction at neutral and high pH corresponds to **(A-cis)°-I**, but a small fraction of **(C-cis)⁺** is ready for the next photocycle.

Conclusions

We report the following novel aspects on the structure and dynamics of an important reversibly switching fluorescent protein rsEGFP2. Quantum-based simulations predict a variety of the chromophore-binding structures with fairly close energies, containing Chro in the cis and trans forms. These structures differ in conformations and protonation states of the chromophore and indicate substantial flexibility of the critical amino acid side chains His149, Ser206 and Glu223. The computed structures and excitation spectra of the series of model systems are consistent with the observed absorption spectra of the protein including the pH-dependence in the ON-state. We propose that protonation of the imidazolinone ring of Chro by Glu223 facilitates cis-trans isomerization involved in the ON→OFF photoswitching. According to the results of simulations, the cationic cis form of Chro may be involved in photoexcitation of the ON-state finally leading to the OFF-state with the neutral trans Chro. The hydrogen bonds of Chro with His149 and Glu223 are breaking and reforming in the course of OFF→ON photoswitching along the decay of systems with the protonated trans Chro to those with the neutral cis Chro. The obtained group of structures with the cis-Chro includes, in particular, species with the cis-anionic Chro and cis-cationic Chro, which are ready for the next photocycle.

Associated content

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/XXX.

Coordinates of the structures obtained in geometry optimization of molecular systems within the protein and cluster models (PDF).

References

(1)   Zhou, X. X.; Lin, M. Z. Photoswitchable Fluorescent Proteins: Ten Years of Colorful Chemistry and Exciting Applications. *Curr. Opin. Chem. Biol.* **2013**, *17* (4), 682–690. https://doi.org/10.1016/j.cbpa.2013.05.031.

(2)   Nienhaus, K.; Ulrich Nienhaus, G. Fluorescent Proteins for Live-Cell Imaging with Super-Resolution. *Chem. Soc. Rev.* **2014**, *43* (4), 1088–1106. https://doi.org/10.1039/C3CS60171D.

(3)   Bourgeois, D.; Adam, V. Reversible Photoswitching in Fluorescent Proteins: A Mechanistic View. *IUBMB Life* **2012**, *64* (6), 482–491. https://doi.org/10.1002/iub.1023.

(4)   Adam, V.; Berardozzi, R.; Byrdin, M.; Bourgeois, D. Phototransformable Fluorescent Proteins: Future Challenges. *Curr. Opin. Chem. Biol.* **2014**, *20*, 92–102. https://doi.org/10.1016/j.cbpa.2014.05.016.

(5)   Grotjohann, T.; Testa, I.; Reuss, M.; Brakemann, T.; Eggeling, C.; Hell, S. W.; Jakobs, S. RsEGFP2 Enables Fast RESOLFT Nanoscopy of Living Cells. *Elife* **2012**, *1*. https://doi.org/10.7554/eLife.00248.

(6)   El Khatib, M.; Martins, A.; Bourgeois, D.; Colletier, J.-P.; Adam, V. Rational Design of Ultrastable and Reversibly Photoswitchable Fluorescent Proteins for Super-Resolution Imaging of the Bacterial Periplasm. *Sci. Rep.* **2016**, *6* (1), 18459. https://doi.org/10.1038/srep18459.

(7)   Coquelle, N.; Sliwa, M.; Woodhouse, J.; Schirò, G.; Adam, V.; Aquila, A.; Barends, T. R. M.; Boutet, S.; Byrdin, M.; Carbajo, S.; De la Mora, E.; Doak, R. B.; Feliks, M.; Fieschi, F.; Foucar, L.; Guillon, V.; Hilpert, M.; Hunter, M. S.; Jakobs, S.; Koglin, J. E.; Kovacsova, G.; Lane, T. J.; Lévy, B.; Liang, M.; Nass, K.; Ridard, J.; Robinson, J. S.; Roome, C. M.; Ruckebusch, C.; Seaberg, M.; Thepaut, M.; Cammarata, M.; Demachy, I.; Field, M.;

Shoeman, R. L.; Bourgeois, D.; Colletier, J.-P.; Schlichting, I.; Weik, M. Chromophore Twisting in the Excited State of a Photoswitchable Fluorescent Protein Captured by Time-Resolved Serial Femtosecond Crystallography. *Nat. Chem.* **2018**, *10* (1), 31–37. https://doi.org/10.1038/nchem.2853.

(8)   Woodhouse, J.; Nass Kovacs, G.; Coquelle, N.; Uriarte, L. M.; Adam, V.; Barends, T. R. M.; Byrdin, M.; de la Mora, E.; Bruce Doak, R.; Feliks, M.; Field, M.; Fieschi, F.; Guillon, V.; Jakobs, S.; Joti, Y.; Macheboeuf, P.; Motomura, K.; Nass, K.; Owada, S.; Roome, C. M.; Ruckebusch, C.; Schirò, G.; Shoeman, R. L.; Thepaut, M.; Togashi, T.; Tono, K.; Yabashi, M.; Cammarata, M.; Foucar, L.; Bourgeois, D.; Sliwa, M.; Colletier, J.-P.; Schlichting, I.; Weik, M. Photoswitching Mechanism of a Fluorescent Protein Revealed by Time-Resolved Crystallography and Transient Absorption Spectroscopy. *Nat. Commun.* **2020**, *11* (1), 741. https://doi.org/10.1038/s41467-020-14537-0.

(9)   Chang, J.; Romei, M. G.; Boxer, S. G. Structural Evidence of Photoisomerization Pathways in Fluorescent Proteins. *J. Am. Chem. Soc.* **2019**, *141* (39), 15504–15508. https://doi.org/10.1021/jacs.9b08356.

(10)   Weber, W.; Helms, V.; McCammon, J. A.; Langhoff, P. W. Shedding Light on the Dark and Weakly Fluorescent States of Green Fluorescent Proteins. *Proc. Natl. Acad. Sci.* **1999**, *96* (11), 6177–6182. https://doi.org/10.1073/pnas.96.11.6177.

(11)   Zimmer, M. Green Fluorescent Protein (GFP): Applications, Structure, and Related Photophysical Behavior. *Chem. Rev.* **2002**, *102* (3), 759–782. https://doi.org/10.1021/cr010142r.

(12)   Christou, N. E.; Giandoreggio-Barranco, K.; Ayala, I.; Glushonkov, O.; Adam, V.; Bourgeois, D.; Brutscher, B. Disentangling Chromophore States in a Reversibly Switchable Green Fluorescent Protein: Mechanistic Insights from NMR Spectroscopy. *J. Am. Chem. Soc.* **2021**, *143* (19), 7521–7530. https://doi.org/10.1021/jacs.1c02442.

(13)   Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110* (13), 6158–6170. https://doi.org/10.1063/1.478522.

(14)   Granovsky, A. A. Extended Multi-Configuration Quasi-Degenerate Perturbation Theory: The New Approach to Multi-State Multi-Reference Perturbation Theory. *J. Chem. Phys.* **2011**, *134* (21), 214113. https://doi.org/10.1063/1.3596699.

(15)   Gozem, S.; Huntress, M.; Schapiro, I.; Lindh, R.; Granovsky, A. A.; Angeli, C.; Olivucci, M.

Dynamic Electron Correlation Effects on the Ground State Potential Energy Surface of a Retinal Chromophore Model. *J. Chem. Theory Comput.* **2012**, *8* (11), 4069–4080. https://doi.org/10.1021/ct3003139.

(16) Gozem, S.; Melaccio, F.; Lindh, R.; Krylov, A. I.; Granovsky, A. A.; Angeli, C.; Olivucci, M. Mapping the Excited State Potential Energy Surface of a Retinal Chromophore Model with Multireference and Equation-of-Motion Coupled-Cluster Methods. *J. Chem. Theory Comput.* **2013**, *9* (10), 4495–4506. https://doi.org/10.1021/ct400460h.

(17) Gozem, S.; Luk, H. L.; Schapiro, I.; Olivucci, M. Theory and Simulation of the Ultrafast Double-Bond Isomerization of Biological Chromophores. *Chem. Rev.* **2017**, *117* (22), 13502–13565. https://doi.org/10.1021/acs.chemrev.7b00177.

(18) Granovsky, A. A. Firefly version 8, www http://classic.chem.msu.su/gran/firefly/index.htm (last accessed July 20, 2021).

(19) Mroginski, M.; Adam, S.; Amoyal, G. S.; Barnoy, A.; Bondar, A.; Borin, V. A.; Church, J. R.; Domratcheva, T.; Ensing, B.; Fanelli, F.; Ferré, N.; Filiba, O.; Pedraza-González, L.; González, R.; González-Espinoza, C. E.; Kar, R. K.; Kemmler, L.; Kim, S. S.; Kongsted, J.; Krylov, A. I.; Lahav, Y.; Lazaratos, M.; NasserEddin, Q.; Navizet, I.; Nemukhin, A.; Olivucci, M.; Olsen, J. M. H.; Pérez de Alba Ortíz, A.; Pieri, E.; Rao, A. G.; Rhee, Y. M.; Ricardi, N.; Sen, S.; Solov'yov, I. A.; De Vico, L.; Wesolowski, T. A.; Wiebeler, C.; Yang, X.; Schapiro, I. Frontiers in Multiscale Modeling of Photoreceptor Proteins. *Photochem. Photobiol.* **2021**, *97* (2), 243–269. https://doi.org/10.1111/php.13372.

(20) Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. A. NWChem: A Comprehensive and Scalable Open-Source Solution for Large Scale Molecular Simulations. *Comput. Phys. Commun.* **2010**, *181* (9), 1477–1489. https://doi.org/10.1016/j.cpc.2010.04.018.

(21) Khrenova, M. G.; Tsirelson, V. G.; Nemukhin, A. V. Dynamical Properties of Enzyme–Substrate Complexes Disclose Substrate Specificity of the SARS-CoV-2 Main Protease as Characterized by the Electron Density Descriptors. *Phys. Chem. Chem. Phys.* **2020**, *22* (34), 19069–19079. https://doi.org/10.1039/D0CP03560B.

(22) Khrenova, M. G.; Grigorenko, B. L.; Nemukhin, A. V. Molecular Modeling Reveals the Mechanism of Ran-RanGAP-Catalyzed Guanosine Triphosphate Hydrolysis without an Arginine Finger. *ACS Catal.* **2021**, 8985–8998. https://doi.org/10.1021/acscatal.1c00582.

(23)   Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* **2020**, *153* (4), 044130. https://doi.org/10.1063/5.0014475.

(24)   Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Isborn, C. M.; Kokkila-Schumacher, S. I. L.; Li, X.; Liu, F.; Luehr, N.; Snyder, J. W.; Song, C.; Titov, A. V.; Ufimtsev, I. S.; Wang, L.; Martínez, T. J. TeraChem : A Graphical Processing Unit-accelerated Electronic Structure Package for Large-scale Ab Initio Molecular Dynamics. *WIREs Comput. Mol. Sci.* **2021**, *11* (2). https://doi.org/10.1002/wcms.1494.

(25)   Melo, M. C. R.; Bernardi, R. C.; Rudack, T.; Scheurer, M.; Riplinger, C.; Phillips, J. C.; Maia, J. D. C.; Rocha, G. B.; Ribeiro, J. V; Stone, J. E.; Neese, F.; Schulten, K.; Luthey-Schulten, Z. NAMD Goes Quantum: An Integrative Suite for Hybrid Simulations. *Nat. Methods* **2018**, *15* (5), 351–354. https://doi.org/10.1038/nmeth.4638.

(26)   Warren, M. M.; Kaucikas, M.; Fitzpatrick, A.; Champion, P.; Timothy Sage, J.; van Thor, J. J. Ground-State Proton Transfer in the Photoswitching Reactions of the Fluorescent Protein Dronpa. *Nat. Commun.* **2013**, *4* (1), 1461. https://doi.org/10.1038/ncomms2460.

(27)   Shinobu, A.; Palm, G. J.; Schierbeek, A. J.; Agmon, N. Visualizing Proton Antenna in a High-Resolution Green Fluorescent Protein Structure. *J. Am. Chem. Soc.* **2010**, *132* (32), 11093–11102. https://doi.org/10.1021/ja1010652.

(28)   Grigorenko, B. L.; Polyakov, I. V.; Savitsky, A. P.; Nemukhin, A. V. Unusual Emitting States of the Kindling Fluorescent Protein: Appearance of the Cationic Chromophore in the GFP Family. *J. Phys. Chem. B* **2013**, *117* (24), 7228–7234. https://doi.org/10.1021/jp402149q.

(29)   Schäfer, L. V.; Groenhof, G.; Klingen, A. R.; Ullmann, G. M.; Boggio-Pasqua, M.; Robb, M. A.; Grubmüller, H. Photoswitching of the Fluorescent Protein AsFP595: Mechanism, Proton Pathways, and Absorption Spectra. *Angew. Chemie Int. Ed.* **2007**, *46* (4), 530–536. https://doi.org/10.1002/anie.200602315.

(30)   Schäfer, L. V.; Groenhof, G.; Boggio-Pasqua, M.; Robb, M. A.; Grubmüller, H. Chromophore Protonation State Controls Photoswitching of the Fluoroprotein AsFP595. *PLoS Comput. Biol.* **2008**, *4* (3), e1000034. https://doi.org/10.1371/journal.pcbi.1000034.

(31)   Grigorenko, B. L.; Nemukhin, A. V.; Polyakov, I. V.; Morozov, D. I.; Krylov, A. I. First-

Principles Characterization of the Energy Landscape and Optical Spectra of Green Fluorescent Protein along the A→I→B Proton Transfer Route. *J. Am. Chem. Soc.* **2013**, *135* (31), 11541–11549. https://doi.org/10.1021/ja402472y.

(32) He, X.; Bell, A. F.; Tonge, P. J. Ground State Isomerization of a Model Green Fluorescent Protein Chromophore. *FEBS Lett.* **2003**, *549* (1–3), 35–38. https://doi.org/10.1016/S0014-5793(03)00761-0.

(33) Sen, T.; Ma, Y.; Polyakov, I. V.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. Interplay between Locally Excited and Charge Transfer States Governs the Photoswitching Mechanism in the Fluorescent Protein Dreiklang. *J. Phys. Chem. B* **2021**, *125* (3), 757–770. https://doi.org/10.1021/acs.jpcb.0c09221.

(34) Suto, K.; Masuda, H.; Takenaka, Y.; Tsuji, F. I.; Mizuno, H. Structural Basis for Red-Shifted Emission of a GFP-like Protein from the Marine Copepod Chiridius Poppei. *Genes to Cells* **2009**, *14* (6), 727–737. https://doi.org/10.1111/j.1365-2443.2009.01305.x.

(35) Davari, M. D.; Ferrer, F. J. A.; Morozov, D.; Santoro, F.; Groenhof, G. The Lineshape of the Electronic Spectrum of the Green Fluorescent Protein Chromophore, Part I: Gas Phase. *ChemPhysChem* **2014**, *15* (15), 3236–3245. https://doi.org/10.1002/cphc.201402355.

(36) Donati, G.; Petrone, A.; Caruso, P.; Rega, N. The Mechanism of a Green Fluorescent Protein Proton Shuttle Unveiled in the Time-Resolved Frequency Domain by Excited State Ab Initio Dynamics. *Chem. Sci.* **2018**, *9* (5), 1126–1135. https://doi.org/10.1039/C7SC02803B.

(37) Polyakov, I. V.; Grigorenko, B. L.; Epifanovsky, E. M.; Krylov, A. I.; Nemukhin, A. V. Potential Energy Landscape of the Electronic States of the GFP Chromophore in Different Protonation Forms: Electronic Transition Energies and Conical Intersections. *J. Chem. Theory Comput.* **2010**, *6* (8), 2377–2387. https://doi.org/10.1021/ct100227k.

TOC graphics