

A time-dependent DFT/molecular dynamics study of the proton-wire responsible for the red fluorescence in the LSSmKate2 protein

Carlos Randino · Marc Nadal-Ferret ·
Ricard Gelabert · Miquel Moreno ·
José M. Lluch

Received: 16 October 2012 / Accepted: 28 December 2012 / Published online: 8 January 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Fluorescent proteins (FP) have become a major topic in the recent biochemical research due to their applications as *in vivo* markers in biological systems. In particular, Red fluorescent proteins (RFP) present some advantages since they require less harmful radiations to be excited and show less light-scattering. In this paper, we are focusing on the LSSmKate2 protein, a RFP that, together with LSSmKate1 and mKeima, is well known for the outstanding difference between absorption and emission wavelengths, which is usually referred as Large Stokes Shift (LSS). It is commonly accepted that an excited state proton transfer accounts for the fluorescence observed in the three proteins. In this work, a molecular dynamics simulation of the LSSmKate2 protein has been carried out, and from different snapshots, a series of excited states have been calculated and analyzed. Our molecular dynamics simulation has proved the availability of the two-link proton-wire suggested by Piatkevich et al. and has furnished a new one-link relay, more prone to take place. The statistical treatment of the excited states can reproduce the electronic absorption spectrum in a reasonable way, and the analysis of the involved orbitals confirms that one

absorption wavelength maximum corresponds to an acidification of the chromophore, regardless of the hydrogen-bonded acceptor residue. All this work constitutes an important step in what should be a thorough and complete study of the photochemistry of the LSSproteins.

Keywords Red fluorescent proteins · LSSmKate2 protein · Electronic absorption spectrum · Time-dependent density functional theory · Molecular dynamics simulation

1 Introduction

Since the extraction, purification and characterization of Green Fluorescent Protein (GFP) by Shimomura from the jellyfish *Aequorea victoria* in the early 1960s [1, 2], a wide range of natural and artificial fluorescent proteins (FP) have appeared [3], contributing to add new improvements for these *in vivo* markers in biological systems [4–7]. Many theoretical studies have appeared in recent years due to the interest of fluorescent proteins [8–21]. Among these, the red fluorescent proteins (RFPs) are particularly interesting due to their better biochemical features, suitable to minimize some drawbacks caused by the use of energetic and dangerous wavelengths in living organisms. Moreover, the RFPs reduce the autofluorescence on the surrounding cells and show deeper light penetration and lower light-scattering [22].

In 2010, two new red fluorescent proteins, LSSmKate1 and LSSmKate2, were developed by Piatkevich et al. [23, 24] from the far-red fluorescent protein mKate. LSSmKate1 and LSSmKate2 differ from each other in the amino acid sequence at position 160 (Glu160 for LSSmKate1 and Asp160 for LSSmKate2) and in the conformation of the chromophore (*cis* for LSSmKate1 and *trans* for LSSmKate2). Both proteins

Published as part of the special collection of articles derived from the 8th Congress on Electronic Structure: Principles and Applications (ESPA 2012).

C. Randino · M. Nadal-Ferret · R. Gelabert (✉) · M. Moreno ·
J. M. Lluch
Departament de Química, Universitat Autònoma de Barcelona,
08193 Bellaterra, Barcelona, Spain
e-mail: Ricard.Gelabert@uab.cat

J. M. Lluch
Institut de Biotecnologia i de Biomedicina,
Universitat Autònoma de Barcelona,
08193 Bellaterra, Barcelona, Spain

share the special property of showing a big gap between absorption and emission wavelengths, commonly named in the field of biology large Stokes shift (LSS). This feature allows the possibility of obtaining a multicolor image useful to select and distinguish different parts of biological systems. LSSmKate1 and LSSmKate2 present absorption maxima at 463/460 nm and emission maxima at 624/605 nm, respectively.

Because of the similarities that LSSmKate1 and LSSmKate2 share with GFP, the suggested photochemical process could follow a similar scheme. For both proteins, the mechanism involves two species: the protonated and deprotonated chromophore. The protonated chromophore absorbs light and starts a proton transfer reaction. Based on their X-ray structures, a one-link excited state proton transfer (ESPT) for LSSmKate1 and a two-link ESPT for LSSmKate2 are proposed. LSSmKate1 would support a direct proton migration from the hydroxyl group of the chromophore to the carboxylic group of Glu160. For LSSmKate2 (Fig. 1), the first proton moves from the hydroxyl group of the chromophore to the hydroxyl group of Ser158 and the second one from Ser158 to the carboxylic group of Asp160. Once the different proton transfers occur, the fluorescence in the deprotonated chromophores takes place. Our group's research interest is focused on the multiple proton transfer reactions [11–13, 25], and so in this paper, we will center on LSSmKate2 which has been proposed to operate via a two-step proton-wire.

Due to the size and the wide mobility of this kind of biochemical systems, the study becomes highly complex. Focusing only on the available crystallographic structure forces us to work with a simple and static model of the whole protein and therefore, restricting us to one small and scarce representation, representative only of the solid state of the crystal structure. However, the protein as encountered naturally is not in the solid phase: it is surrounded by

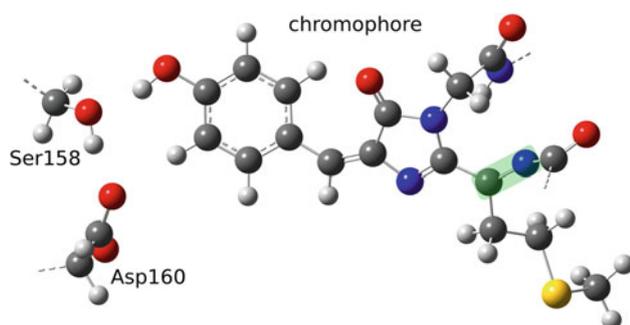


Fig. 1 Representation of the chromophore and the acceptor amino acids involved in the ESPT in LSSmKate2, Ser158 and Asp160. The C and N atoms, highlighted in green, have had to be reparameterized, as their likes are not present in the standard CHARMM force field. Color code: dark gray, carbon; white, hydrogen; red, oxygen; blue, nitrogen; yellow, sulfur

solvent (water) and ions and at some finite temperature. In these conditions, the protein is a very dynamic entity, flexible and that jostles around due to thermal motion. To understand the protein's properties in these conditions, it is necessary to sample a number of configurations of it, obtained through for instance, molecular dynamics. To achieve this aim, we will carry out a molecular dynamics simulation using classical mechanics. So the main purpose of this work is to provide this variety of geometries from a theoretical point of view, including a suitable environment. With this information, we will simulate the electronic absorption spectrum for LSSmKate2, and we will analyze along the molecular dynamics the thermal fluctuations of the three residues directly involved in the two proton transfers responsible for the fluorescence.

2 Methods

2.1 Parameterization and classical dynamics simulation

A 40-ns molecular (classical) dynamics (MD) simulation has been run with a dual aim: on the one hand, to check the dynamical stability of the crystallographic structure [24] supplied at the Protein Data Bank, and on the other hand, to furnish an ensemble of structures from which further excited state electronic calculations have been done, thus providing a tool to introduce thermal effects in the spectrum, as explained in the following subsection.

To properly run the MD simulation, the crystallographic structure has had to undergo some refinements. First of all, as hydrogen atoms are not detected in the X-ray structure, they have been included by using the PDB2PQR [26, 27] server at a neutral pH. As for the chromophore, which is not a standard amino acid, it has been protonated as Piatkevich et al. [24] suggest. Secondly, crystallographic structures are obtained at very low temperatures and do not contain bulk solvent molecules. Physiological conditions imply the inclusion of the solvent effects and require heating and equilibrating the system to physiological temperatures before running the MD simulation. All these processes (including the MD simulation) have been performed using the CHARMM22 [28–30] force field by means of the CHARMM-35b1 [31, 32] software. In all, 34952 atoms have been included in the molecular dynamics simulation.

Nevertheless, before tackling these very steps, another delicate issue has had to be dealt with, namely, the fact that the chromophore is not a standard residue, new parameters are therefore needed. Since the LSSmKate2 chromophore is similar to the GFP one, [33] most of the necessary parameters have been taken from the parameterization done for GFP's by Thiel and co-workers, [34] consistent

with the CHARMM22 force field. Compared to GFP's, the LSSmKate2 chromophore has a methionine instead of a serine, so for these very atoms, the CHARMM22 force field parameters have been taken. However, there are two atoms that differ from any other amino acid: the nitrogen and carbon atoms that form a double bond (Fig. 1, highlighted in green). Parameters for all stretchings, bendings, torsions and improper torsions, partial charges and Van Der Waals interactions involving these two atoms have had to be found. For the sake of consistency, we have tried to follow Thiel's method [34] as thoroughly as possible, although we have calculated each parameter separately from the rest (except for the nonbonding and partial charge), as explained elsewhere [35].

The system has been solvated with a cubic box of water molecules of 69 Å edge, and periodic boundary conditions have been set to account for the long-range electrostatic interactions of the protein. Two rough energy minimizations (each one with 200 steps of Adopted Basis Newton–Raphson method) have been performed to avoid bad contacts, firstly involving just the water molecules, and then the whole system. Next, the system must be heated and equilibrated. The heating and equilibration have been carried out using the CHARMM-implemented Verlet integrator algorithm with steps of 1 fs. The temperature has been increased 25 K every 5 ps. This way the system is softly heated from 0 to 300 K. The periodic boundary conditions imply the construction of a cubic lattice, in which the protein moves freely and the potential energy is calculated with a 20 Å cutoff. The motion of the backbone is constrained by harmonic forces. When the system has reached 300 K, the constrained dynamics simulation is restarted with the same integration step for 160 ps, without further increase of temperature and gradually releasing all the harmonic constraints applied on the system. Afterward, 360 ps more has been run without any harmonic constraints. Once the system is fully equilibrated, the MD simulation is run for 40 ns production time.

2.2 Theoretical simulation of the electronic absorption spectrum

Excited state electronic structure calculations have been performed on 160 snapshots from the ensemble of structures obtained from the molecular dynamics simulation (one every 0.25 ns). The calculated absorption wavelengths constitute a fairly good approximation to the experimental absorption spectrum as it takes into account the thermal effects.

The electronic calculations have been done as follows: each structure resulting from the MD simulation has been divided in two regions, namely, a quantum mechanics (QM) region and a molecular mechanics (MM) region.

The QM region includes the chromophore in its entirety with the carbonyl group of the previous amino acid (Phe60) and the amino group of the next amino acid (Ser66). The side chain of Ser158 and a deprotonated Glu160, the two responsible amino acids participating in the ESPT, are also included (in total, 53 QM atoms + 4 link atoms). All the QM atoms of the model are represented as spheres in Fig. 1. The rest of the protein has been treated as MM region to introduce polarization effects on the QM part. The polarization of the protein due to the chromophore is not taken into account, but we think that this effect will not be very significant in the spectrum, see for instance the work by Murugan et al. [36] The QM region has been calculated quantum-mechanically using the Gaussian 09 [37] software, and the MM atoms have been treated as point charges according to the CHARMM-22 force field. We have used the Chemshell [38] package to prepare the inputs. Link atoms have been used to treat the QM/MM boundary.

The quantum mechanical calculations have been done with the density functional theory (DFT) for the ground state, and the time-dependent density functional theory (TDDFT) for the excited states, all of them with the CAM-B3LYP [39] functional and the 6–31 + G(d,p) basis set. TDDFT is considered a proper method to tackle the analysis of an ESPT, as it has been stated in published works about other chromophores [40, 41]. The alleged involvement of charge-transfer excited states in LSSmKate2 [42, 43] fully justifies the election of the hybrid Coulomb-attenuated CAM-B3LYP, specifically designed to avoid the spurious effects resulting from the inability of pure DFT methods to properly describe long-range exchange effects [39, 44]. As a matter of fact, we have recently shown that CAM-B3LYP provides good results for LSSmKate2, specially compared with conventional hybrid functional such as B3LYP [43].

15 excited states have been calculated for each of the 160 frames. Then, all the different excited states with oscillator strength (f) greater than 0.1 have been sorted by their frequencies, and a wavelength relative histogram has been done with a bin size of 10 nm. This histogram represents an average oscillator strength for each 10 nm wavelength window, weighted by the number of snapshots which have an excited state accessible ($f \geq 0.1$) through excitation at the wavelength range from the ground state. The sum of the counted oscillator strengths around a given wavelength value gives an approximate idea of how likely it is to detect absorption around it regardless of the nature of the excited state involved. It is not a thorough reproduction of the spectrum, but it is a feasible and plausible way to simulate it. A similar procedure has already been used by Thiel and co-workers [19] and other groups [20, 45] to characterize the absorption spectrum of different fluorescent proteins. Other approaches have also been

applied to obtain the optical properties of varied chemical systems [36, 46–49].

3 Results

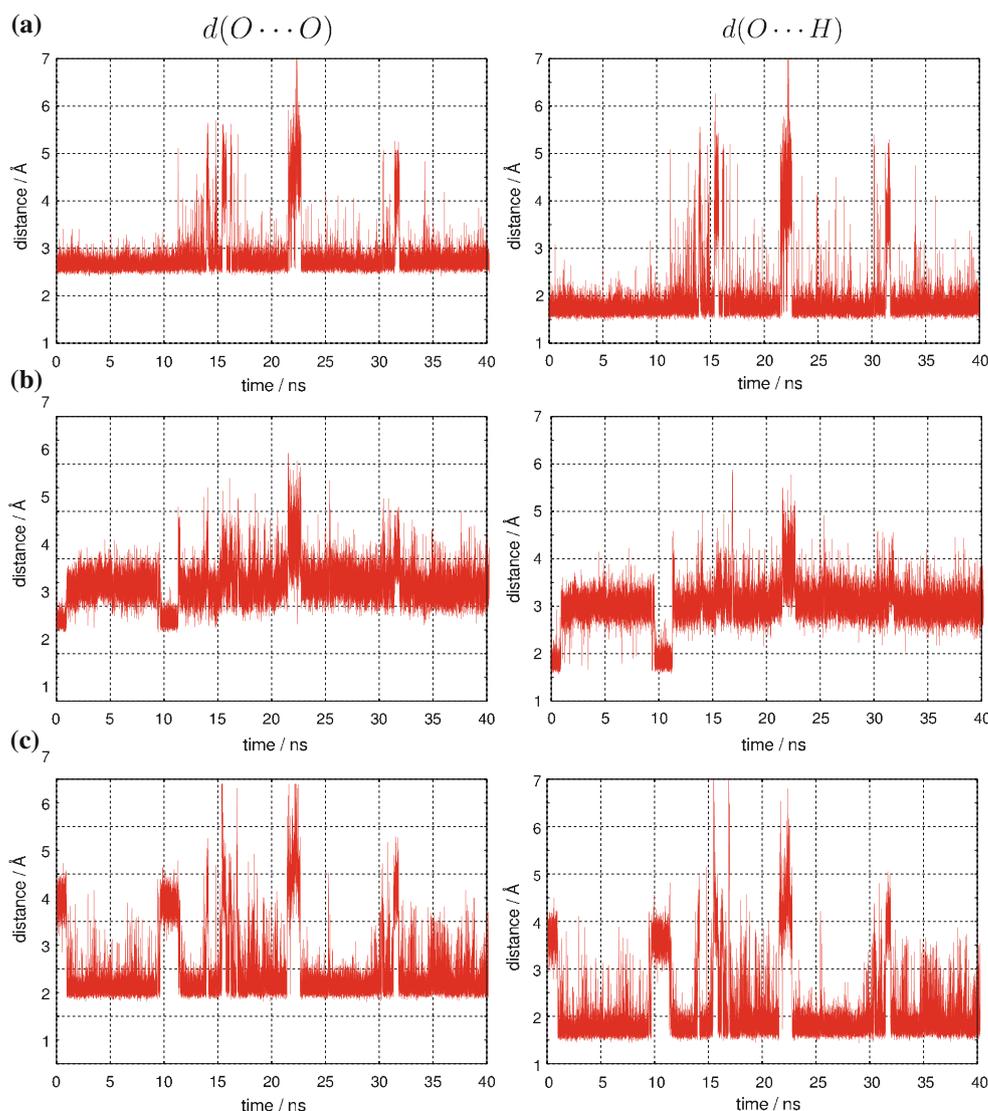
3.1 Classical dynamics simulation

The 40-ns MD simulation has yielded interesting results to figure out possible paths available for the excited state proton transfer. To begin with, it is seen that the protein is stable along the dynamics simulation, that is to say, there are no relevant conformational changes. As for the chromophore and the residues that are proposed to host the proton transfer, namely, Ser158 and Asp160, several hydrogen bonds are observed. By analyzing the distances between the oxygen atoms of Ser158 and Asp160 (Fig. 2a) and the distances between the hydrogen atom of Ser158

and the oxygen atom of Asp160 (Fig. 2a), several features can be noticed.

First of all, leaving apart some pronounced and random fluctuations that appear occasionally, the range of the oscillations is small (about 0.5 Å, slightly more for the hydrogen–oxygen distance). Besides, the distance between the oxygen atoms oscillates between 2.5 and 3 Å, and the distance between the oxygen and the hydrogen atoms oscillates between 1.5 and 2 Å. This constant difference of 1 Å coincides with the average hydrogen chemical bond distance, which implies that the three atoms are almost collinear and the hydrogen bond is well formed and is maintained along the MD simulation. Following the same criteria, we can also identify in Fig. 2b, c when the chromophore is hydrogen bonded either to Ser158 or Asp160. On the contrary, the corresponding hydrogen bond does not exist when the range of oscillation of the oxygen–oxygen distance is about 1 Å, while the oxygen–oxygen distance is

Fig. 2 Distances between the residues involving the hydrogen bonds in the putative proton-wire along the MD simulation, taking into account the oxygen–oxygen distances (*left*) or the oxygen–hydrogen distances (*right*) for Ser158–Asp160 **a**, Cro–Ser158 **b** and Cro–Asp160 **c** hydrogen bonds



beyond 3 Å (Ser158) or 4 Å (Asp160). Then, except for short temporary hydrogen bond disruptions, the chromophore is always hydrogen bonded to either to Ser158 or Asp160, but never to both simultaneously.

Considering all that, we can identify two kinds of hydrogen bond networks between the other residues that constitute two qualitatively different situations. Firstly, one in which, as it appears in the crystallographic structure, the chromophore is hydrogen bonded to Ser158 and Ser158 is hydrogen bonded to Asp160, thus enabling a putative two-link proton relay. This situation, as can be observed in Fig. 2a, b, is not the most probable in terms of residence time as it only occurs between 0 and 1.5 ns, and between 9.5 and 11.5 ns approximately. Secondly, another situation can be observed, in which the chromophore is hydrogen bonded to Asp160, which is also hydrogen bonded to Ser158 (Fig. 2a, c). This is the most likely situation in terms of residence time, as it happens between 1.5 and 9.5 ns, and from 11.5 ns to the end of the simulation. Nonetheless, from about 14–17 ns, about 22–23 ns and about 31–32 ns, all the distances shown in Fig. 2 increase. This corresponds to temporary disruptions of the hydrogen bonds, and in these periods, the above-mentioned residues are hydrogen bonded to other residues (mainly water molecules), but they always return to the latter situation, in which the chromophore is hydrogen bonded to Asp160 despite some spurious peaks. This latter configuration would enable a single link proton relay, in which the proton could be transferred directly from the chromophore to Asp160, analogously to what is suggested to occur for LSSmKate1, [24] where the phenoxy moiety proton in the chromophore is thought to be directly transferred to Glu160 (already in the X-ray structure).

As it has been checked by the analyses of the excited state electronic structure calculations presented in the following subsection, there are no appreciable differences between those two different configurations in terms of the nature of excited states found for each of them. Furthermore, both configurations would eventually lead to the same chemical species: an anionic chromophore and an aspartic residue.

3.2 Simulated electronic absorption spectrum

The simulated electronic absorption spectrum of LSSmKate2 (Fig. 3) built as we have explained in the methods section takes into account several configurations of the protein. Among them, both proton-relay configurations (that is to say, the Cro-Ser158-Asp160 and Cro-Asp160 ones, described in the former subsections) have been considered for the simulated absorption spectrum. To figure out whether both proton-relay configurations present different absorption frequencies or not, the absorption frequencies

corresponding to the 15 excited states calculated for each of the 160 snapshots (including only those transitions with oscillator strength equal to or greater than 0.1) have been represented versus $d_{\text{Cro-Ser}} - d_{\text{Cro-Asp}}$, where $d_{\text{Cro-Ser}}$ is the oxygen–oxygen distance between the chromophore and Ser158 and $d_{\text{Cro-Asp}}$ is the oxygen–oxygen distance between the chromophore and Asp160. When $d_{\text{Cro-Ser}} - d_{\text{Cro-Asp}}$ is positive, the chromophore is closer to Asp160 than to Ser158 (typical of the Cro-Asp160 proton-relay configuration). Conversely, when the value is negative, the chromophore is closer to Ser158 than to Asp160 (which would be typical of the Cro-Ser-Asp proton relay). The results are shown in Fig. 4. Most points are concentrated on the positive abscissa axis, this being a consequence of the predominance of the Cro-Asp proton-relay configuration along the whole dynamics. Notwithstanding this, no obvious correlation can be observed between the abscissa value

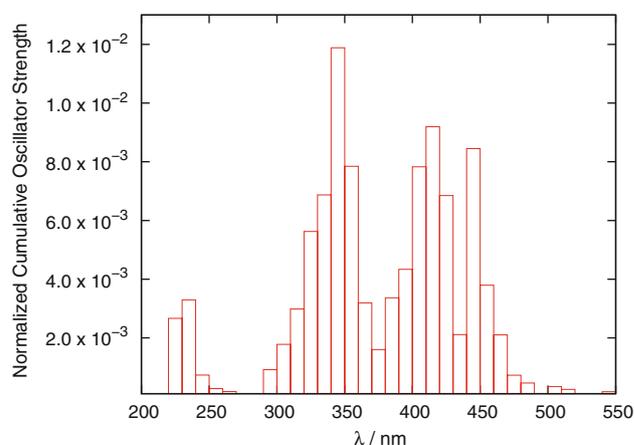


Fig. 3 Simulated electronic absorption spectrum of LSSmKate2. The red columns represent the sum of oscillator strength for each set of 10 nm (properly scaled)

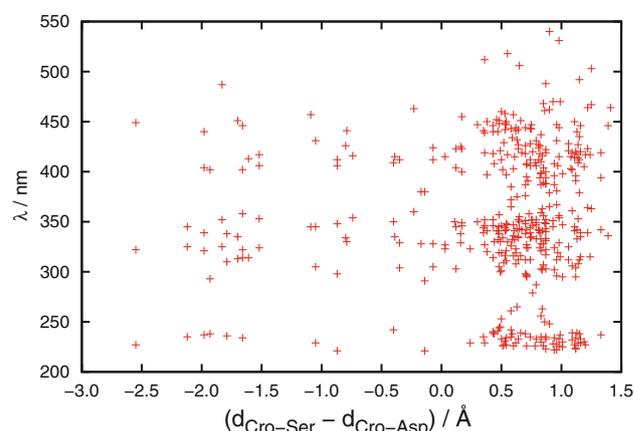


Fig. 4 Representation of the wavelengths obtained while calculating the spectrum versus the difference of O–O distances between Cro-Ser158 and Cro-Asp160

and the absorption wavelength, so both configurations appear equally spread along the spectrum.

The simulated absorption spectrum of LSSmKate2 (shown in Fig. 3) clearly presents two main bands in agreement with the available experimental data. The more energetic band has a maximum at about 345 nm and the second one shows two peaks at 415 and 445 nm. The latter absorption band, the one experimentally analyzed and known to produce red fluorescence at 605 nm, is well reproduced when compared with the experimental absorption maximum (460 nm). Nevertheless, the higher energy band between 300 and 375 nm is red-shifted by about 65 nm with respect to its corresponding experimental absorption band (graphically estimated at 280 nm [24]). Shifts of similar magnitude have been found in related studies [36]. Both bands present narrower bandwidths than the corresponding experimental values.

The spectrum depicted in Fig. 3 has been obtained from data consisting of 160 snapshots each of them contributing up to 15 excited states. From a statistical point of view, a possible concern could be the degree of convergence of this spectrum. To assess this point, we have proceeded as follows: we have computed different spectra based on data with increasingly large length of the simulation in steps of 10 ns and computed the spectra using exactly the same procedure and convoluted with a Gaussian function of $\sigma = 10$ nm to attenuate statistical noise. The results are shown in Fig. 5, and it can be seen that the overall shape of the spectra (number, intensity, width and position of the bands) is already stable from the beginning. As for the precise positions of the peaks, they undergo some small shifts and could change slightly if more samples were included.

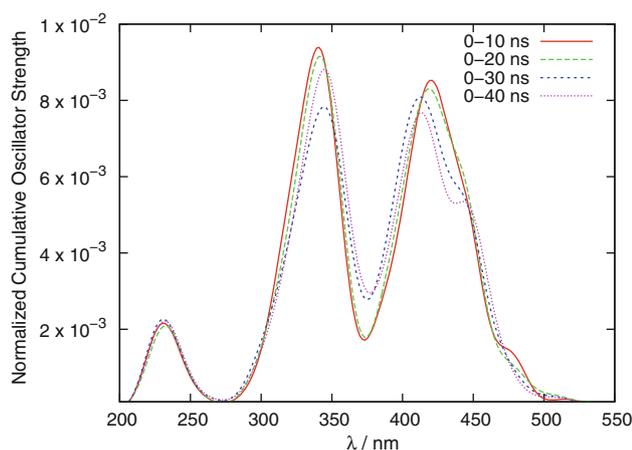


Fig. 5 Convergence of spectrum along the simulation. Each of the spectra has been computed on an increasingly longer part of the simulation. A convoluting Gaussian function of $\sigma = 10$ nm has been used to smooth the statistical noise

The use of molecular dynamics simulations has allowed us to obtain complete absorption bands that are attributable to thermal fluctuations (configurational changes) on the chromophore and its environment, adding contributions from different excited states reproducing the whole spectrum.

Analyzing the description of the electronic states accessed upon photoexcitation in terms of the orbital description, we can observe three $\pi\pi^*$ molecular orbitals that participate in the $\pi\pi^*$ excitations leading to the UV–vis spectrum shown in Fig. 3: HOMO, LUMO and LUMO + 1. All of them are located over the chromophore, preserving the approximate molecular plane as a node. In Fig. 6, we have depicted these three molecular orbitals for Cro-Asp160 (taken at 4 ns of the MD simulation) and Cro-Ser158-Asp160 (taken at 10.5 ns of the MD simulation) hydrogen-bonded structures. It is clear that these molecular orbitals are identical for the two structures, so justifying that both are compatible with the experimental electronic absorption spectrum.

The right band (low energy) of the spectrum could be assigned to a HOMO \rightarrow LUMO transition. As can be seen in Fig. 6, the HOMO \rightarrow LUMO excitation represents an electron shift from the phenoxy moiety to the pyrimidine moiety of the chromophore. This represents effectively a “charge transfer” of intramolecular type and justifies the methodological choice made. Based on the nature of the molecular orbitals involved in this excited state, the excitation results in a depletion of charge of the phenoxy moiety

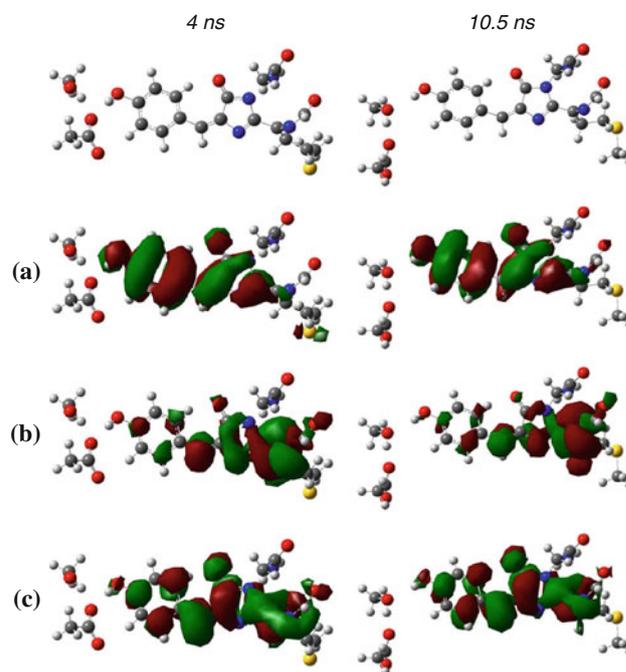


Fig. 6 Orbital representations of two selected frames at 4 and 10.5 ns of the MD simulation. The HOMO (a), LUMO (b) and LUMO + 1 (c) involved in the $\pi\pi^*$ transitions are shown

which likely results in an increase of the acidity of the chromophore. Thus, this charge shift provides support for the ESPT hypothesis as a basis of the mechanism of generation of the fluorescent species. On the other hand, the higher energy band corresponds mostly to a HOMO \rightarrow LUMO + 1 excitation. In this case, LUMO + 1 is an orbital evenly spread over the full chromophore, and the main difference with HOMO lies in the increased number of nodal planes perpendicular to the molecular plane on the phenoxy moiety of the chromophore (from HOMO with 2 to LUMO + 1 with 3).

4 Conclusions

In the present work, we have run a MD simulation of the LSSmKate2 protein. From this simulation, we have been able to analyze the evolution of the plausible structures prone to host the proton transfer reaction that accounts for the fluorescence. Besides, from a series of snapshots, the electronic excited states have been calculated, furnishing a reasonable simulation of the experimental electronic absorption spectrum and allowing the analysis of the molecular orbitals involved in the absorption bands.

As for the molecular dynamics, the 40-ns simulation has proved the general structure of the protein to be dynamically stable, as already stated. However, apart from corroborating the availability of the two-link proton-wire configuration reported in the crystallographic structure, [24] a new single link proton-wire involving the chromophore and Asp160 has emerged as an even more likely configuration at least in terms of residence time. This new scenario would imply an analogy with the LSSmKate1, in which a one-link proton relay is suggested [24] to allow for the formation of the anionic chromophore, the species that would be responsible for the fluorescence. This series of different paths can provide interesting insights about the actual mechanism of proton transfer.

Focusing now on the methodology, the use of molecular dynamics simulations together with the calculation of excited states via TDDFT and the functional CAM-B3LYP has allowed us to simulate a full absorption spectrum that reasonably reproduces the available experimental data.

Regardless of the fact that two different configurations allow the formation of the anionic chromophore along, respectively a one-link or two-link proton transfer, the π molecular orbitals involved in the electronic excitation are not appreciably modified, which justifies that both configurations are compatible with the experimental electronic absorption spectrum. The two bands present a $\pi\pi^*$ orbital transition. A transition between the HOMO to the LUMO accounts for the absorption band responsible for the red fluorescence. The migration of the electronic density toward the pyrimidine ring facilitates the shift of the proton

of the chromophore in the excited state, making a more favorable reaction. Conversely, the UV-band can be assigned to the HOMO \rightarrow LUMO + 1 transition, implying a very minor displacement of the electronic density over the chromophore.

Acknowledgments This work was supported by the “Ministerio de Economía y Competitividad” through project CTQ2011-24292 and by the “Generalitat de Catalunya” through project 2009SGR409. Use of computational facilities at the “Centre de Serveis Científics i Acadèmics de Catalunya (CESCA)” is gratefully acknowledged. C.R. acknowledges the “Ministerio de Economía y Competitividad” for a fellowship within the FPU program. M.N.-F. thanks the “Secretaria d’Universitats i Recerca (SUR)” of the “Departament d’Economia i Coneixement (DEC)” of the “Generalitat de Catalunya” and the European Social Fund (ESF) for a fellowship within the FI-DGR program.

References

1. Prasher DC, Eckenrode VK, Ward WW, Prendergast FG, Cormier MJ (1992) Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene* 111:229–233
2. Shimomura O, Johnson FH, Saiga Y (1962) Extraction, purification and properties of Aequorin, a bioluminescent protein from luminous hydromedusa. *Aequorea J Cell Compar Phys* 59(3): 223–239
3. Matz MV, Fradkov AF, Labas YA, Savitsky AP, Zaraisky AG, Markelov ML, Lukyanov SA (1999) Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat Biotechnol* 17(10): 969–973
4. Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA (2010) Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol Rev* 90(3):1103–1163
5. Davidson MW, Campbell RE (2009) Engineered fluorescent proteins: innovations and applications. *Nat Methods* 6(10):713–717
6. Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC (1994) Green fluorescent protein as a marker for gene-expression. *Science* 263(5148):802–805
7. Day RN, Davidson MW (2009) The fluorescent protein palette: tools for cellular imaging. *Chem Soc Rev* 38(10):2887–2921
8. Jonasson G, Teuler J-M, Vallverdu G, Mérola F, Ridard J, Lévy B, Demachy I (2011) Excited state dynamics of the green fluorescent protein on the nanosecond time scale. *J Chem Theor Comput* 7(6):1990–1997
9. Laurent AD, Assfeld X (2010) Effect of the enhanced cyan fluorescent protein framework on the UV/visible absorption spectra of some chromophores. *Interdiscip Sci Comput Life Sci* 2(1):38–47
10. Laurent AD, Mironov VA, Chapagain PP, Nemukhin AV, Krylov AI (2012) Exploring structural and optical properties of fluorescent proteins by squeezing: modeling high-pressure effects on the mStrawberry and mCherry red fluorescent proteins. *J Phys Chem B* 116(41):12426–12440
11. Vendrell O, Gelabert R, Moreno M, Lluch JM (2006) Potential energy landscape of the photoinduced multiple proton-transfer process in the green fluorescent protein: classical molecular dynamics and multiconfigurational electronic structure calculations. *J Am Chem Soc* 128:3564–3574
12. Vendrell O, Gelabert R, Moreno M, Lluch JM (2008) A potential energy function for heterogeneous proton-wires. Ground end photoactive states of the proton-wire in the green fluorescent protein. *J Chem Theor Comput* 4:1138–1150

13. Vendrell O, Gelabert R, Moreno M, Lluch JM (2008) Operation of the proton wire in green fluorescent protein. A quantum dynamics simulation. *J Phys Chem B* 112(17):5500–5511
14. Bravaya KB, Bochenkova AV, Granovskii AA, Nemukhin AV (2008) Modeling of the structure and electronic spectra of green fluorescent protein chromophore. *Russ J Phys Chem B* 2(5): 671–675
15. Li X, Chung LW, Mizuno H, Miyawaki A, Morokuma K (2010) Competitive mechanistic pathways for green-to-red photoconversion in the fluorescent protein kaede: a computational study. *J Phys Chem B* 114(49):16666–16675
16. Filippi C, Buda F, Guidoni L, Sinicropi A (2012) Bathochromic shift in green fluorescent protein: a puzzle for QM/MM approaches. *J Chem Theor Comput* 8(1):112–124
17. Bravaya KB, Grigorenko BL, Nemukhin AV, Krylov AI (2012) Quantum chemistry behind bioimaging: insights from Ab Initio studies of fluorescent proteins and their chromophores. *Acc Chem Res* 45(2):265–275
18. Zhang H, Sun Q, Wang S, Olsen SL, Smith SC (2011) Hydrogen bonding and transfer in the excited state. Wiley, New York, p 815
19. Sanchez-Garcia E, Doerr M, Thiel W (2010) QM/MM study of the absorption spectra of DsRed.M1 chromophores. *J Comput Chem* 31(8):1603–1612
20. Imhof P (2012) Computational study of absorption spectra of the photoconvertible fluorescent protein EosFP in different protonation states. *J Chem Theor Comput* 8(11):4828–4836
21. Jackson S, Sanders J (2009) Green fluorescent protein issue. *Chem Soc Rev* 38(10):2821
22. Piatkevich KD, Efremenko EN, Verkhusha VV, Varfolomeev DD (2010) Red fluorescent proteins and their properties. *Russ Chem Rev* 79(3):243–258
23. Piatkevich KD, Hulit J, Subach OM, Wu B, Abdulla A, Segall JE, Verkhusha VV (2010) Monomeric red fluorescent proteins with a large Stokes shift. *Proc Natl Acad Sci USA* 107(12):5369–5374
24. Piatkevich KD, Malashkevich VN, Almo SC, Verkhusha VV (2010) Engineering ESPT pathways based on structural analysis of LSSmKate red fluorescent proteins with large Stokes shift. *J Am Chem Soc* 132(31):10762–10770
25. Ortiz-Sánchez JM, Gelabert R, Moreno M, Lluch JM, Anglada JM, Bofill JM (2010) Bipyridyl derivatives as photomemory devices: a comparative electronic-structure study. *Chem-Eur J* 16(22):6693–6703
26. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucl Acids Res* 35:W522–W525
27. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucl Acids Res* 32:W665–W667
28. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorcikiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102(18):3586–3616
29. MacKerell AD, Feig M, Brooks CL (2004) Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* 126(3):698–699
30. MacKerell Jr AD, Brooks B, Brooks III CB, Nilsson L, Roux B, Won Y, Karplus M (1998) CHARMM: The energy function and its parametrization with an overview of the program. *Encyclopedia of computational chemistry*, P. v.R.Schleyer, N.L Allinger, T. Clark, J. Gasteiger, P.A. Kollman, H.F. Schaefer III, P.R. Schreiner, Chichester vol. 1
31. Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caffisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoseck M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30(10):1545–1614
32. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) Charmm—a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4(2):187–217
33. Zimmer M (2002) Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. *Chem Rev* 102(3):759–781
34. Reuter N, Lin H, Thiel W (2002) Green fluorescent proteins: empirical force field for the neutral and deprotonated forms of the chromophore. Molecular dynamics simulation of the wild type and S65T mutant. *J Phys Chem B* 106(24):6310–6321
35. Nadal-Ferret M, Gelabert R, Moreno M, Lluch JM (2012). Submitted
36. Murugan NA, Kongsted J, Rinkevicius Z, Agren H (2012) Color modeling of protein optical probes. *Phys Chem Chem Phys* 14(3):1107–1112
37. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery J, J. A., Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam NJ, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2011) Gaussian 09, Revision C.1. Gaussian, Inc, Wallingford CT, 2009
38. Sherwood P, de Vries AH, Guest MF, Schreckenbach G, Catlow CRA, French SA, Sokol AA, Bromley ST, Thiel W, Turner AJ, Billeter S, Terstegen F, Thiel S, Kendrick J, Rogers SC, Casci J, Watson M, King F, Karlson E, Sjøvoll M, Fahmi A, Schafer A, Lennartz C (2003) QUASI: a general purpose implementation of the QM/MM approach and its application to problems in catalysis. *J Mol Struct-Theochem* 632:1–28
39. Handy NC (2004) The molecular physics lecture 2004: (i) density functional theory, (ii) quantum monte carlo. *Mol Phys* 102 (23–24):2399–2409
40. Aquino AJA, Lischka H, Hattig C (2005) Excited-state intramolecular proton transfer: a survey of TDDFT and RI-CC2 excited-state potential energy surfaces. *J Phys Chem A* 109: 3201–3208
41. Sobolewski AL, Domcke W (1999) Ab initio potential-energy functions for excited state intramolecular proton transfer: a comparative study of o-hydroxybenzaldehyde, salicylic acid and 7-hydroxy-1-indanone. *Phys Chem Chem Phys* 1(13):3065–3072
42. Zhang MY, Wang JY, Lin CS, Cheng WD (2011) First-principles study of one- and two-photon absorption of the H-bonding complexes from monomeric red fluorescent proteins with large stokes shifts. *J Phys Chem B* 115(36):10750–10757

43. Randino C, Moreno M, Gelabert R, Lluch JM (2012) Peek at the potential energy surfaces of the LSSmKate1 and LSSmKate2 proteins. *J Phys Chem B* 116(49):14302–14310
44. Tawada Y, Tsuneda T, Yanagisawa S, Yanai T, Hirao K (2004) A long-range-corrected time-dependent density functional theory. *J Chem Phys* 120(18):8425–8433
45. Isborn CM, Götz AW, Clark MA, Walker RC, Martínez TJ (2012) Electronic absorption spectra from MM and ab Initio QM/MM molecular dynamics: environmental effects on the absorption spectrum of photoactive yellow protein. *J Chem Theor Comput* 8(12):5092–5106
46. Fantacci S, De Angelis F, Sgamellotti A, Marrone A, Re N (2005) Photophysical properties of $[\text{Ru}(\text{phen})_2(\text{dppz})]^{2+}$ Intercalated into DNA: an integrated Car–Parrinello and TDDFT study. *J Am Chem Soc* 127(41):14144–14145
47. De Angelis F, Fantacci S, Sgamellotti A (2007) An integrated computational tool for the study of the optical properties of nanoscale devices: application to solar cells and molecular wires. *Theor Chem Acc* 117(5–6):1093–1104
48. Eriksen JJ, Olsen JMH, Aidas K, Agren H, Mikkelsen KV, Kongsted J (2011) Computational protocols for prediction of solute NMR relative chemical shifts. A case study of L-Tryptophan in aqueous solution. *J Comput Chem* 32(13):2853–2864
49. Murugan NA, Aidas K, Kongsted J, Rinkevicius Z, Agren H (2012) NMR Spin–Spin coupling constants in polymethine dyes as polarity indicators. *Chem-Eur J* 18(37):11677–11684