JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

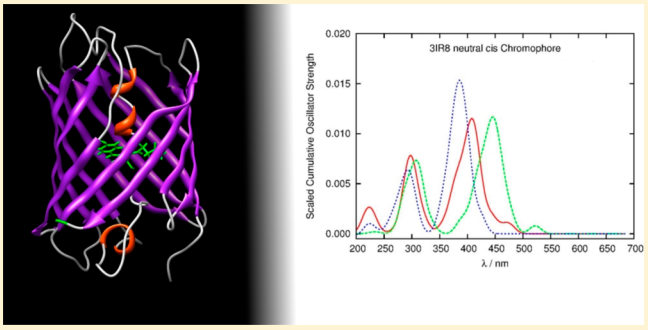
How Does the Environment Affect the Absorption Spectrum of the Fluorescent Protein mKeima?

Marc Nadal-Ferret,[†] Ricard Gelabert,[†] Miquel Moreno,*[,†] and José M. Lluch[†,‡]

[†]Departament de Química and [‡]Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

Supporting Information

ABSTRACT: The absorption spectrum of a fluorescent protein is determined by its chromophore, but the residues that surround it also have a remarkable role, leading to noticeable spectral shifts. We have theoretically analyzed the monomeric protein Keima (mKeima), a red fluorescent protein most remarkable for an outstanding difference between the absorption and emission frequencies, and potentially suited for multicolor imaging applications. In the present work, we have performed excited state electronic calculations on the chromophore with an increasing number of atoms surrounding it, and we have compared these results with the excited states calculations on an ensemble of structures obtained from a molecular dynamics simulation of the complete protein. The importance of the inclusion of the effects of the whole protein in the electronic calculations has been proved, and it is concluded that only with the consideration of the thermal effects can the absorption spectra of the protein be properly characterized.



1. INTRODUCTION

Fluorescent proteins (FP) are nowadays a major topic in biochemical and biomedical research since they allow for the study of dynamic processes in live cells.[1−3] In particular, *Aequorea victoria* Green Fluorescent Protein (GFP, see the Nobel lectures by Shimomura,[4] Chalfie,[5] and Tsien,[6] and references therein) has been the most outstanding for its high quantum yield, its stability, and the autocatalytic formation of its chromophore.[7] Since the discovery of GFP, a huge amount of fluorescent proteins, whose fluorescence covers the full range of the visible electromagnetic spectrum, have been developed.[1−3,8−11] This palette of fluorescent proteins furnishes a powerful tool for multicolor imaging, namely, an approach to the study of the above-mentioned processes in live cells. Therefore, it is crucial to have a deep understanding of the photochemical processes involved in these proteins to be able to track their photophysical and photochemical properties to their structural characteristics. In this sense, a large amount of work has addressed the optical properties of the fluorescent proteins.[12−67]

There is solid agreement now that the way in which a chromophore has been matured determines the fluorescent properties of a protein,[68] even if formed by similar amino acids. For instance, the common tripeptide Met−Tyr−Gly is able to form chromophores spanning a 175-nm range for the maxima of emission spectra.[69] Although the nature of the chromophore and its maturation are crucial to account for the fluorescence wavelength, it has been reported that local environmental features such as the position of charged amino acid residues, hydrogen bonds, or hydrophobic interactions can yield shifts of up to 40 nm in absorption and emission maxima.[69]

Among the spread of fluorescent proteins, let us focus on the family of Red Fluorescent Proteins (RFP), in particular, the subgroup represented by drFP5833[11] (commercial name DsRed). In this kind of chromophore, the π-system of the GFP chromophore is extended by an additional N-acylimine moiety; however, the mechanism of the chromophore maturation is not fully understood yet.[68] Recently developed proteins like mKeima,[1] LSSmKate1,[2] and LSSmKate2[2] (440−463 nm absorption, 605−624 nm emission), or proteins like mNeptune,[8] eqFP670[70], and TagRFP657[71] (592−611 nm absorption, 646−670 nm emission), all of them containing a DsRed-like chromophore, have exhibited novel photochemical properties by means of rationally modifying the chromophore environment.[8] Such a phenomenon suggests that the absorption and fluorescence spectra are affected not only by the chromophore but also by the surrounding residues and the solvent.

Considering therefore the fact that the absorption properties of the chromophore are related to the environment, we should bear in mind that a single protein at a given temperature can show different wavelengths, as it displays a collection of configurations in which charges, hydrogen bonding networks, or long-range interactions are constantly changing. In other words, the experimental data available concerning radiation

Received: November 15, 2012
Published: February 11, 2013

absorption consist of, in general, the absorption spectrum, which describes how much radiation is absorbed at a given wavelength. It is usually a band encompassing a wide range of wavelengths, more or less symmetric and showing at least a maximum. The band can be thought of as consisting of individual absorption signals (specific spectral lines) originated in different radiation−molecule interaction events, each one corresponding to a particular structure of the system, and what is registered is the envelope.

In this work, we have considered it to be of the utmost interest to determine to what extent the protein environment can modulate the specific spectral lines (each one corresponding to a concrete structure of the system associated to its chromophore). Particularly, we have focused our studies on monomeric Keima (mKeima), a red fluorescent variant of a protein from the *Montipora* sp. stony coral.[1] Its backbone structure consists of 11 β sheets displaying a barrel-like shape with one α helix going through it (Figure 2). Its chromophore is a DsRed-like chromophore, formed by a tripeptide that involves glutamine, tyrosine, and glycine. mKeima absorbs and emits light maximally at 440 and 620 nm, respectively. This remarkable difference between the absorption and emission wavelengths has been referred to as the Large Stokes Shift (LSS),[1,8] although, strictly speaking, the absorbing and emitting species are not the same. The LSS makes mKeima a suitable target for dual-color single-laser fluorescence cross-correlation spectroscopy.[1] Besides its tantalizing usefulness, this protein is particularly suited to theoretical studies since several crystallographic structures have been recently published,[72,73] showing significant differences between them.

Blanchoin and co-workers[72] have observed two absorption maxima at the wavelengths of 440 and 590 nm and one emission maximum at 620 nm. It has been assumed[72] that the 440 nm absorption maximum corresponds to the neutral chromophore, whereas the anionic one yielded the 590 nm maximum. From their work, three crystallographic structures are available in the Protein Data Bank (PDB): one obtained at pH 3.8 (ID 2WHS), which presents a *cis* chromophore; one obtained at pH 8.0 (ID 2WHU), which shows a *trans* chromophore; and a third one presenting both chromophores at a time, obtained at pH 5.6. On the other hand, Remington and co-workers[73] do not mention the 590 nm absorption maximum, and the crystallographic structure supplied has been obtained at pH 7.0 (ID 3IR8) possessing a *cis* chromophore. They do not find evidence of a *trans* chromophore, though its presence with less than a 10% proportion is not excluded. In Figure 1, both configurations of the DsRed-like chromophore are displayed. In both isomers, an excited state proton transfer (ESPT) is thought to be responsible for the observed fluorescence, but the proton wire responsible for the ESPT process is foreseen to be different from that originally found in wtGFP involving the motion of three protons.[74] From the crystallographic data, a proton wire of two protons is predicted for the *cis* chromophores whereas just one proton is expected to transfer in the *trans* isomer. But there is agreement neither with the isomerization of the chromophore nor with its protonation state, for Blanchoin and co-workers[72] suggest there is a counterintuitive reverse relation with the pH of the medium. These controversies make mKeima particularly suitable to tackle a theoretical study of the modulation of the spectrum mediated by the environment that could shed some light on the particular issues of its isomerization and protonation state. It is a matter of fact that the different X-ray structures were found

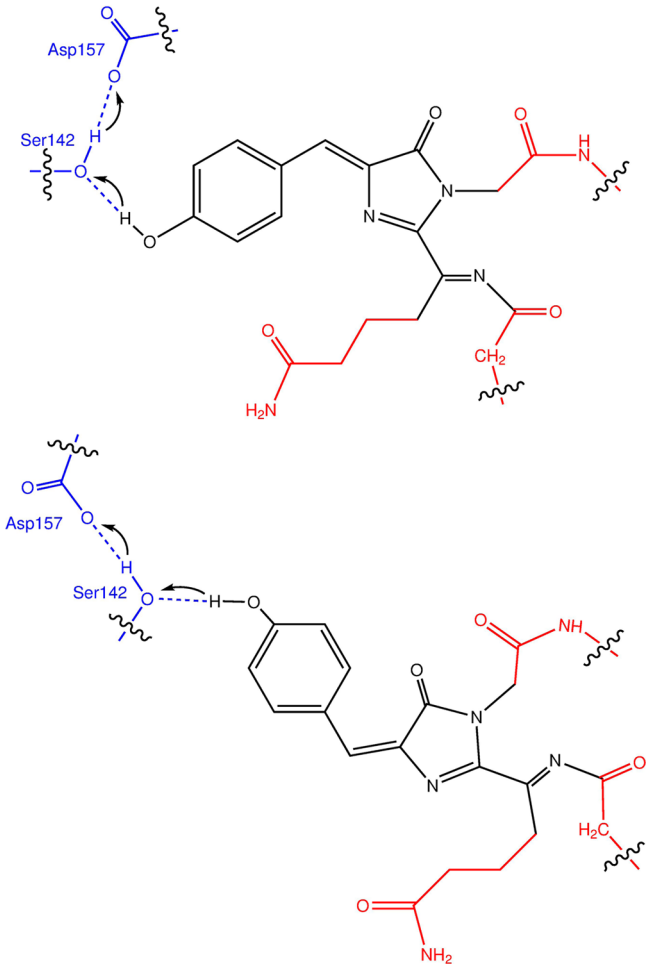


Figure 1. The protonated form of the mKeima *cis* (up) and *trans* (down) chromophore and atoms of its immediate surroundings. Arrows indicate the proton transfer(s) that lead(s) to the deprotonated (anionic) chromophore structure, as some authors have suggested.[72,73] Wave lines indicate covalent bonds that have been cut and saturated with hydrogen atoms in some reduced models of the protein considered here (see text for the details). In black: atoms considered in the *optimized chromophore* model. In red: atoms that have been added for the *only chromophore* (unoptimized) model. In blue: atoms of the Ser142 and Asp157 residues additionally considered in the *optimized wire* model.

following different procedures, but there is not any more available information regarding the mKeima structure and its plausible reaction pathways nowadays.

So as to deal with the mentioned topic, several theoretical procedures have been considered here. To begin with, the development of a series of increasingly sophisticated static models has been undertaken. At an early stage, Quantum Mechanics (QM) calculations for excited states (ES) have been performed on the chromophore alone with the geometry it has in the resolved crystallographic structures. For the QM calculations, Time Dependent Density Functional Theory (TDDFT) has been used. To have some idea of the relevance of surrounding steric effects, these very same calculations have been performed on a geometrically optimized chromophore. Then, to calibrate the spectrum modulation due to the environment, those residues that can plausibly host a proton transfer and stabilize the chromophore have been included in the quantum mechanical calculation, namely Ser142 and

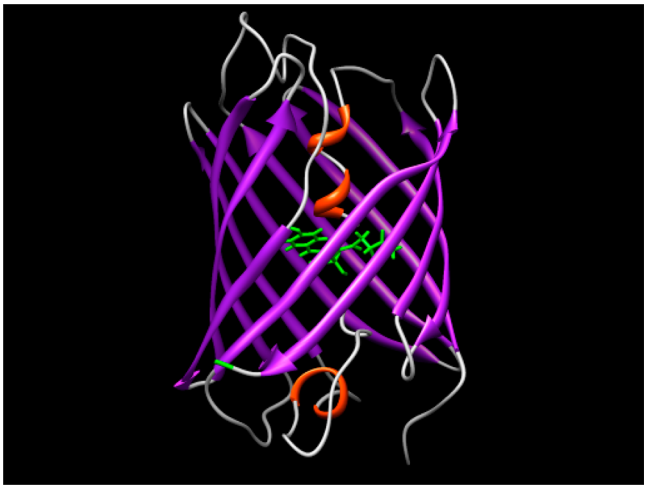


Figure 2. Backbone structure of the mKeima protein, with the chromophore in its center.

Asp157 and some crystallographic water molecules. Again, here we have analyzed the effect of optimizing this small model of the whole protein. The next step in this series of "static" calculations has included all the residues and the crystallographic water molecules within 3.5 Å from the chromophore. Finally, these calculations have been compared with quantum mechanic calculations involving the proton wire and taking account of the rest of the whole protein by means of point charges. These calculations have been performed for a molecular mechanics (MM) optimized structure and for a hybrid QM/MM optimized one.

From the results of these static models, it is concluded that the effect of the environment is highly remarkable, but not enough to satisfactorily account for the absorption spectra. At this point, an extensive molecular dynamics simulation has been carried out, with a double scope: to ensure to what extent the crystallographic structures represent also stable dynamical structures during a molecular dynamics simulation and to provide an ensemble of structures from which the excited states' properties could be calculated, thus allowing for thermal fluctuations that yield a largely improved description of the observed physical and chemical properties of mKeima. We remark that the study presented here has not considered any bond cleavage at all. In particular, the hydrogen atoms involved in the proton wire have remained definitely bound to the corresponding donor atom throughout.

2. METHODS

2.1. Excited State Electronic Calculations. By means of Gaussian 09,[75] several excited-state electronic calculations have been done. All these calculations have been performed for each of the three mentioned crystallographic structures, and for each protonation state of the chromophore. For the ground state, DFT calculations have been used, with the CAM-B3LYP[76] functional and the 6-31+G(d,p)[77−90] basis set. A total of 15 excited electronic states have been evaluated using TDDFT, with the same functional and basis. TDDFT has been established as a suitable and proper way to study and analyze the ESPT like those from the fluorescent proteins.[91−93] Recent works[91,94] have disclosed the involvement of charge-transfer excited states in a closely related RFP. This is why we have decided to work with the hybrid Coulomb-attenuated CAM-B3LYP functional, specifically designed to avoid such spurious effects.

The mKeima chromophore (see Figure 1) comes from the common tripeptide sequence Gln−Tyr−Gly. At an earliest stage, our calculations involved only the chromophore directly taken from the corresponding crystallographic structure with some atoms belonging to the contiguous residues (atoms shown in black and red in Figure 1). Cuts in the sequence that bonds the chromophore to the rest of the protein (indicated with wave lines in Figure 1) were saturated with hydrogen atoms. Then the chromophore was geometrically optimized. In this case, to prevent some bends of the long tails of the chromophore that cannot take place in the full protein pocket, some atoms (depicted in red within Figure 1) were eliminated, taking care that the π-system of the chromophore was fully preserved. Again, the cleaved bonds were saturated with hydrogen atoms.

Next, in order to consider the influence of the surrounding atoms and residues, a larger model (referred by us as *wire*) was considered. This model was built up by adding to the chromophore the residues Ser142 and Asp157 (putative proton donors or acceptors) and two or three (depending on the case) crystallographic water molecules prone to H-bonding them. First of all, the crystallographic geometry was frozen, and later on, to assess the influence of steric effects, a reduced wire model was constructed by removing the crystallographic water molecules, the red atoms in Figure 1, and some atoms from the residues Ser142 and Asp157. The final molecular system, called *optimized wire* (atoms depicted in black and blue in Figure 1), was then geometrically optimized. Finally, we considered a larger model (which has been named *shell*) by adding most residues containing at least one atom within 3.5 Å of the chromophore in the crystallographic structure. In all, residues Ser142, Asp157, Arg91, Glu211, Met159, and Arg193 were included for all *shell* models. Then, for 3IR8 and 2WHS structures (*cis* chromophore), residues Pro59, Leu61, and Ser65 were included, whereas for the 2WHU structure (*trans* chromophore), Glu144 and Phe173 were included. The 3IR8 *shell* contains four crystallographic water molecules. The 2WHS *shell* contains three crystallographic water molecules, and the 2WHU *shell* contains two crystallographic water molecules. All these considered models are explicitly depicted in Figures S2−S11 in the Supporting Information.

Our results show that optimizing the chromophore blue-shifts the spectra, and including more residues around it (*wire* or *shell* models) red-shifts it. A combination of both (*optimized wire*) furnishes reasonable results, but we cannot exclude that this might be due to error compensation. For this reason, the whole protein has been taken into account. To this purpose, the crystallographic structures have been solvated and partly optimized by means of the CHARMM22 force field[95−97] just to eliminate bad contacts. These procedures will be explained in the following subsection (Parameterization and Classical Dynamics). From the resulting structures, the ChemShell[98] interface has been used to calculate the excited states with Gaussian 09 (as before) on the *wire* structure (with some of the Ser142 and Asp157 atoms not included) and to include the rest of the protein as point charges. The charge corresponding to each atom has been determined by the CHARMM22 force field. This model will be referred to as *Whole System*.

The next step in the exploration of the effects of the medium on the spectra has consisted of repeating the latter calculations but on a QM/MM optimized structure (*ChemShell-optimized*

Whole System from now on). The initial structures in this optimization are the same as the *Whole System* ones, and the QM region is again the *wire* structure (again, with some of the Ser142 and Asp157 atoms not included). The QM energy has been calculated with B3LYP/6-31+G(d,p) using the Gaussian 09 software. The MM energy has been calculated with the CHARMM22 force field using DL-POLY,[99] a ChemShell device for MM calculations. QM and MM regions are connected by link atoms. The QM/MM interaction energy has been calculated by ChemShell with the *charge-shift* scheme. That is to say, QM/MM electrostatic interactions are handled by the QM code with QM polarization, charge close to the link atoms is shifted away, and point dipoles are added to compensate. The optimization has been done using HDLCopt[100]—another ChemShell device—with an active region including all residues at 20 Å from the chromophore. The optimization has been performed with the default HDLCopt parameters: in particular, the maximum gradient component is set toler = 0.00045 au. From these QM/MM optimized structures, we have used the same procedure as with the *Whole System* model to calculate 15 excited electronic states.

At this point, as the wavelength of the resulting excited states is not always in agreement with the experimental data, it is reasonable to think that the range of wavelengths experimentally observed could be as well a result of the fluctuations in the positions of the chromophore. Not necessarily does a minimum provide more realistic wavelengths if the minimum of potential energy finds itself in a region of configurational space seldom visited by the system in the conditions of temperature of the experiment. For this reason, classical dynamics simulations have been run for all the structures. That way, not only is it possible to check the stability of the supplied crystallographic structures but also to somehow reproduce the absorption spectra. In the Results section, we present a simple, yet reasonable way to simulate the absorption spectra taking into account the full dynamics of the protein.

2.2. Parameterization and Classical Dynamics. To run the classical dynamics simulations (for the three X-ray solved structures, 2HWS, 2WHU, and 3IR8, each one with protonated and anionic chromophore, that is to say, six simulations), some steps have had to be done as the PDB structure is obtained at very low temperatures and may be significantly different at physiological temperature. Moreover, crystallographic structures do not include hydrogen atoms, so they have been placed by means of the PDB2PQR server.[101,102] The protonation of the chromophore, if needed, is quite straightforward. The only tricky decision has been excluding a hydrogen atom from the peptide bond between the chromophore and the leucine 61, thus assuming that there is a double bond between the nitrogen (named N in Figure S1) and what would be the α carbon (CA1 in Figure S1) of the glutamine (both part of the chromophore). But these two atoms are coplanar with the two other atoms linked to the carbon, so a double bond can be expected, and some published works[8,103] support this decision. For all the following processes, the CHARMM22[96] force field and the CHARMM-35b1[104,105] software have been used.

Another delicate issue to deal with is that the chromophore is not a standard residue, so new parameters are needed. Since the mKeima chromophore is very similar to the GFP one, most of the necessary parameters have been taken from the parametrization done for GFP by Thiel and co-workers.[106] Compared to GFP, the mKeima chromophore has a glutamine instead of a serine, so for these very atoms, the CHARMM22 force field parameters for a glutamine have been taken. However, there are two atoms that differ from any other amino acid: the above-mentioned nitrogen and carbon atoms that form a double bond (Figures 1 and S1). All bonds, angles, dihedrals and improper angles, partial charges, and nonbonding parameters involving these two atoms have had to be found.

For the sake of consistency, we have tried to follow Thiel's method[106] as thoroughly as possible, although we have calculated each parameter separately from the rest (except for the nonbonding and partial charges). Besides, Thiel and co-workers have found the GFP chromophore parameters starting from already existing structures in amino acids, but no amino acid in the CHARMM topology file has this kind of double bond. Given this lack of structures to compare with, we have taken the whole mKeima *cis* chromophore for our calculations. Nevertheless, we have always followed the same level of calculation they used, so that our set of obtained parameters is consistent with theirs and with the CHARMM22 force field. For bonds, angles, and improper angles, energy profiles have been scanned around plausible minima with B3LYP/6-31G(d). This is the level of calculation at which Thiel and co-workers have adjusted these parameters for the GFP chromophore. In this way, a consistent set of parameters can be obtained. All these potentials are treated harmonically by the CHARMM force field, so the force constants and equilibrium angles and distances have been found by adjusting the profile to a parabola. Dihedral angles have not been calculated as the steric effects do not allow for the exploration of the profile, and anyway the values for dihedral parameters tend to be much lower than the other sort of parameters, so we have directly taken the parameters calculated by Thiel or, otherwise, the already existing ones in the CHARMM22 force field.

Regarding the nonbonding van der Waals interaction parameters and partial charges, they have been obtained with the same algorithm. Two energy profiles for a water molecule H-bonded to one atom (either the involved carbon or nitrogen atom) have been calculated (sometimes removing some neighboring atom to avoid unwanted contacts) with an ab initio RHF/6-31G(d) calculation (this is the level of calculation at which Thiel and co-workers have obtained the GFP nonbonding parameters). Then, the CHARMM22 energy has been calculated with suitable initial parameters for partial charges and van der Waals parameters. The squared differences between the electronic RHF and CHARMM22 energies for each point are summed, also over the two profiles so that both profiles have been consistently adjusted at a time. This sum is a function of the parameters to be found, so through minimization, it is possible to obtain the desired values of them. Nonetheless, CHARMM topology files force the total charge of a group to be an integer. To avoid program failure, slight modifications upon some atom charges have been done. All the new parameters that have been used are presented in the Supporting Information. Note that, for the anionic chromophore, as the main changes are expected for the tyrosine moiety, we have not deemed it necessary to fit the parameters again.

To take into account the effects of the solvation, stochastic boundary conditions have been used. A CHARMM automatic algorithm has been used to set different solvation water molecules all over a sphere centered at the origin (which lies very close to the chromophore) within a 30 Å radius. These water molecules first, and then the whole protein, have been roughly optimized (500 steps of adopted basis Newton–

Table 1. Calculated Absorption Wavelengths (nm) According to the Model and Structure Used[a]

structure[b]		only chromophore	optimized chromophore	wire	optimized wire	shell	whole system	ChemShell optimized whole system
3IR8 (*cis*)	N	551 (0.67)	369 (0.77)	561 (0.63)	430 (0.99)	557 (0.72)	422 (0.97)	391 (1.09)
	A	575 (0.89)	486 (1.07)	593 (0.92)	472 (1.11)	588 (0.75)	492 (1.17)	448 (1.22)
2WHS (*cis*)	N	519 (0.59)	369 (0.77)	568 (0.56)	430 (0.99)	489 (0.69)	457 (0.95)	426 (1.13)
	A	561 (0.94)	486 (1.07)	571 (0.97)	472 (1.11)	517 (0.82)	504 (1.19)	460 (1.20)
2WHU (*trans*)	N	494 (0.66)	363 (0.83)	505 (0.83)	420 (1.30)	490 (0.46)	444 (1.25)	427 (1.28)
	A	560 (1.00)	458 (1.43)	550 (1.03)	439 (1.42)	530 (0.67)	491 (1.37)	453 (1.40)
QM atoms		46−50	35−36	80−83	49	203-213	59−62	59−62
MM atoms		0	0	0	0	0	11685−12795	11685−12795

[a]In parentheses, the oscillator strength is given. Letters N and A refer to Neutral or Anionic chromophore, respectively. The differences in the number of atoms in comparable models (for instance, columns 2 and 3) are explained in subsection 2.1. [b]The structure 3IR8 has been taken from ref 73, whereas the structures 2WHS and 2WHU have been taken from ref 72.

Table 2. Distances (in Å) between the Donor and Acceptor Oxygen Atoms of the Cited Residues[a]

structure		crystallographic structure	chromophore protonation state	optimized wire	ChemShell optimized whole system	dynamics average
3IR8[73] (*cis*)	Cro-Ser	2.66	neutral	2.49	2.61	2.81 ± 0.14
			anionic	2.56	3.85	2.86 ± 0.15
	Ser-Asp	2.67−2.74	neutral	2.49	2.79	2.91 ± 0.31
			anionic	2.54	4.35	2.91 ± 0.27
	Cro-Arg91	2.83−2.86	neutral		2.72	2.86 ± 0.15
			anionic		2.70	2.80 ± 0.12
	Cro-Glu211	2.78−2.84	neutral		3.14	3.24 ± 0.23
			anionic		3.25	4.19 ± 0.36
2WHS[72] (*cis*)	Cro-Ser	2.70−2.77	neutral	2.49	2.57	2.89 ± 0.19
			anionic	2.56	2.65	2.95 ± 0.34
	Ser-Asp	2.93−3.22	neutral	2.49	2.80	3.14 ± 0.43
			anionic	2.54	3.41	4.08 ± 0.62
	Cro-Arg91	2.80−2.94	neutral		2.77	2.84 ± 0.13
			anionic		2.74	2.82 ± 0.13
	Cro-Glu211	2.90−3.0	neutral		3.43	3.54 ± 0.28
			anionic		3.39	4.12 ± 0.41
2WHU[72] (*trans*)	Cro-Asp	2.92−3.16	neutral	2.45	2.58	2.70 ± 0.17
			anionic	2.44	2.52	2.73 ± 0.10
	Asp-Ser	2.56−3.61	neutral	2.72	2.84	2.94 ± 0.41
			anionic	2.83	2.93	4.27 ± 0.41
	Cro-Arg91	2.78−3.35	neutral		2.78	2.94 ± 0.18
			anionic		2.76	2.91 ± 0.22
	Cro-Glu211	2.69−3.20	neutral		3.29	3.38 ± 0.26
			anionic		3.53	3.99 ± 0.38

[a]The uncertainty range in the last column corresponds to the standard deviation.

Raphson method) so as to avoid bad contacts. The structures resulting from these optimizations have been used for *Whole System* excited states electronic calculations, as explained in the previous subsection.

Prior to running the molecular dynamics simulation, the minimized structures have been heated and equilibrated. In all these simulations, the system is divided into three regions. The chromophore is centered at the origin. There is a Newtonian zone (where all the residues containing at least one atom at 20 Å or less from the origin are included) in which Newton equations govern the dynamics, then a Langevin zone (where all residues not belonging to the Newtonian zone containing at least one atom at 24 Å or less from the origin are included) in which Langevin (dissipative) dynamics governs the equations. Finally, there is the reservoir region (that contains the rest of the system) in which atoms cannot displace at all. This separation is justified by the fact that no great conformational changes are expected, as in terms of structure mKeima is very similar to GFP, which has a fairly rigid structure.[13,107] The reduction of degrees of freedom expected to be not so relevant has made calculations much more feasible. The equations of motion (Newton's or Langevin's, depending on the region) for the heating, equilibration, and dynamics have been integrated with a leapfrog algorithm, as implemented in the CHARMM program.

The heating has been done as follows: the minimized structure is heated up to 100 K. Random velocities are originally assigned following a Gaussian distribution. The integration step is 1 fs. Every 5 fs, the system rises 5 K up to 100 K by the rescaling of velocities. After 50 ps, the coordinates are saved. Starting from those, again random velocities are assigned, and the same procedure is repeated for 50 ps to heat the system up to 200 K. The same thing is done till the system reaches 300 K. Afterward, the system is equilibrated for 150 ps, following the described procedure, but this time the velocities are not updated and the temperature remains constant. Once

the system is fully equilibrated, a 20 ns dynamics simulation is run. This procedure has been carried out six times: each model (2WHS, 2WHU, 3IR8) has been simulated with the chromophore protonated and ionized. This corresponds to a total simulation for the production phases of 120 ns.

3. RESULTS

3.1. Excited-Stated Electronic Calculations on Static mKeima Structures. As it has been previously mentioned, the experimental spectrum of mKeima exhibits an absorption maximum at 440 nm associated with the protein with the protonated (neutral) chromophore, though some authors[2,72] report a 590 nm absorption maximum as well, thought to correspond to the protein with the deprotonated (anionic) chromophore. In Table 1, the results of the different excited-state electronic calculations are displayed.

It can be seen (Table 1) that when the chromophore alone is geometrically optimized, it exhibits absorption wavelengths noticeably shorter than the ones obtained with the geometry directly taken from the X-ray structure. This is so because the optimization of the ground state stabilizes it more than the other states. On the other hand, by adding more atoms (*wire* and *shell* models), a variety of results is obtained. In the *wire* model, the wavelengths tend to be longer than in the crystallographic structure. However, in the even bigger *shell* model, the wavelengths are either longer or shorter depending on the crystallographic structure. For this reason, we cannot extract conclusions about which the most suitable model is. The *wire* and *shell* models describe fairly well the 590 nm maximum associated with the anionic chromophores but yield poor results for the protonated ones. Conversely, the *optimized wire* models furnish good results for the protonated chromophore but bad for the anionic one, so these small models fail to properly describe the whole absorption spectra of mKeima. A similar result has been obtained for the closely related LSS-Kate models,[91] so we conclude that these reduced models, quite habitual in theoretical studies of photochemistry, are not suited to deal with large systems such as a FP. The different models considered here are depicted in the Supporting Information.

For this reason, to thoroughly assess the effects of the whole environment, the *Whole System* model has been considered. That model provides results that are found to be more reliable and significantly closer to the 440 nm maximum, specifically for the protonated chromophore. A further improvement of the results could be expected if these structures were optimized. As explained in the previous subsection, this was done at the QM/MM level with the ChemShell program, yielding the *ChemShell optimized Whole System* results. Nonetheless, the wavelengths of the resulting excited states describe almost systematically worse the experimental absorption spectra (Table 1), particularly for the anionic *cis* chromophore models, in which the minimum structures yield partially disrupted proton wires (see Table 2).

The data in Table 1 enable a very interesting kind of analysis that is in essence only possible from the theoretical point of view: it shows that there are two effects that, originating in the environment of a given chromophore, contribute to the final absorption wavelength of a particular structure. First, the environment could play a mere geometrical effect by constraining the chromophore's structure to be different than *in vacuo*. This can be seen when comparing any two columns referring to the same kind of environment but differing solely on whether optimization has been done, like for instance columns 2 and 3 (where environment is not present), columns 4 and 5 (the *wire* model), and so on. This usually blue-shifts the absorption, because the optimization is done in the ground state and, usually, makes the excitation energy larger. Second, the effect can be of polarization type, where the environment can affect the charge distribution of the chromophore and, in this way, alter the energy difference between the ground and photoactive state. This is seen comparing columns where the chromophore has the same geometry but differs by the presence of an environment like, for instance, columns 2, 4, 6, and 7. In actuality, both effects can, and do, act together.

Although the *Whole System* describes very well the 440 nm absorption maximum for the protonated chromophore samples, the anionic chromophore maximum excitation seems to shift too much toward the blue, when it is suggested[72] that this could account for the 590 nm absorption maximum. One could expect that adding more atoms furnishes fairly better and—most importantly—more reliable results. But optimizing the structure does not necessarily make it more realistic, as it is hard to know whether the minimum reached is the most representative one in view of the very large number of different minima that can be found. More importantly, the protein samples from which experimental spectra are recorded are usually in solution. A minimization would produce, if anything, a structure close to what a low temperature sample would render, likely a solid or crystalline sample. That is why classical molecular dynamics simulations are needed to provide a spread of structures from which the same *Whole System* calculations can be done. In the next subsection, the classical dynamics simulations are analyzed, and following it, the results of the excited states electronic calculations on the different frames along the dynamics are presented.

3.2. Classical Molecular Dynamics Simulations on the Whole Protein. Molecular dynamics simulations have been run for the six analyzed cases (three X-ray initial structures, either with neutral or with anionic chromophore, as explained in the previous subsection). A first noticeable result of the molecular dynamics simulations is that they prove that all the considered structures are fairly stable. In this sense, it is of the utmost relevance to pay attention to the distances between the residues that can host the proton transfer wire, as no covalent bonds exist to bind them together. For protonated *cis* configurations, the proton is expected to be transferred from the chromophore to the Ser142 and from the Ser142 to the Asp157 (see Figure 3), whereas for protonated *trans* configurations the proton would be transferred directly from the chromophore to the Asp157, being in this case the Ser142, H-bonded to the Asp157, a mere spectator of the proton transfer as depicted in Figure 1.

In Table 2, distances from the oxygen atoms implied in the proton wire are shown for the crystallographic structures and their corresponding geometrical optimizations. We have considered the structures obtained by Blanchoin and co-workers[72] and Remington and co-workers.[73] Although the protein is monomeric, the crystallographic structures contained two[73] or four[72] monomers, so the maximum and minimum distance values are shown for each case. Also, it must be considered that the Gaussian *wire* optimization has been the same for both *cis* chromophore models, so the distances are the same. Some of the distances displayed in Table 2 that have been obtained with the ChemShell optimization are far greater than the corresponding ones from the X-ray structure, from the *wire* optimization, or from the molecular dynamics simulation.

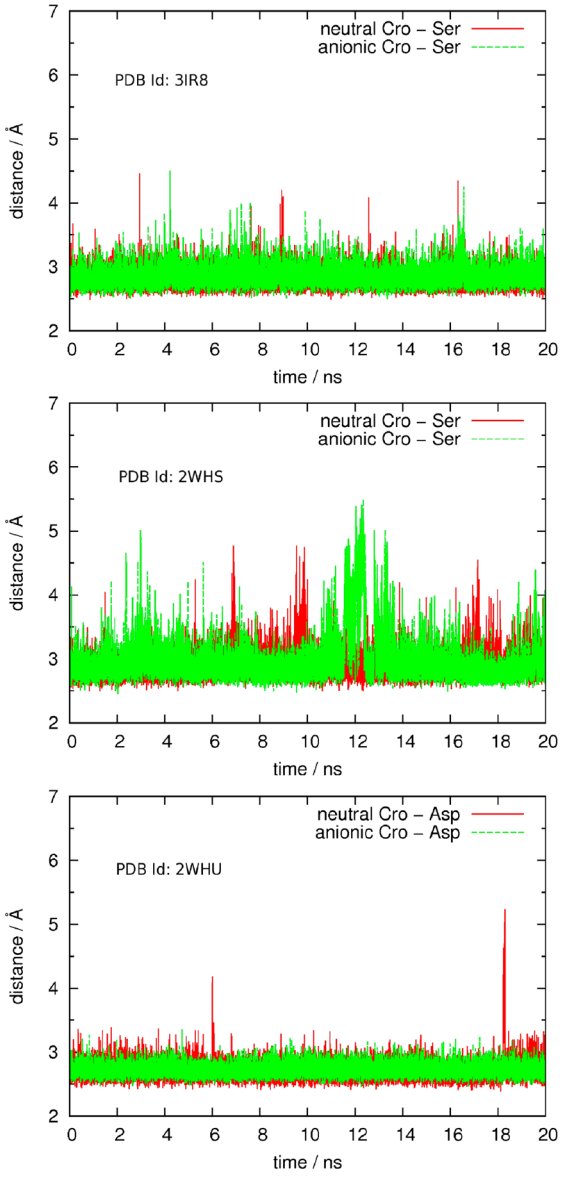


Figure 3. Distances between the oxygen atoms of the chromophore and the corresponding residue throughout the molecular dynamics simulations.

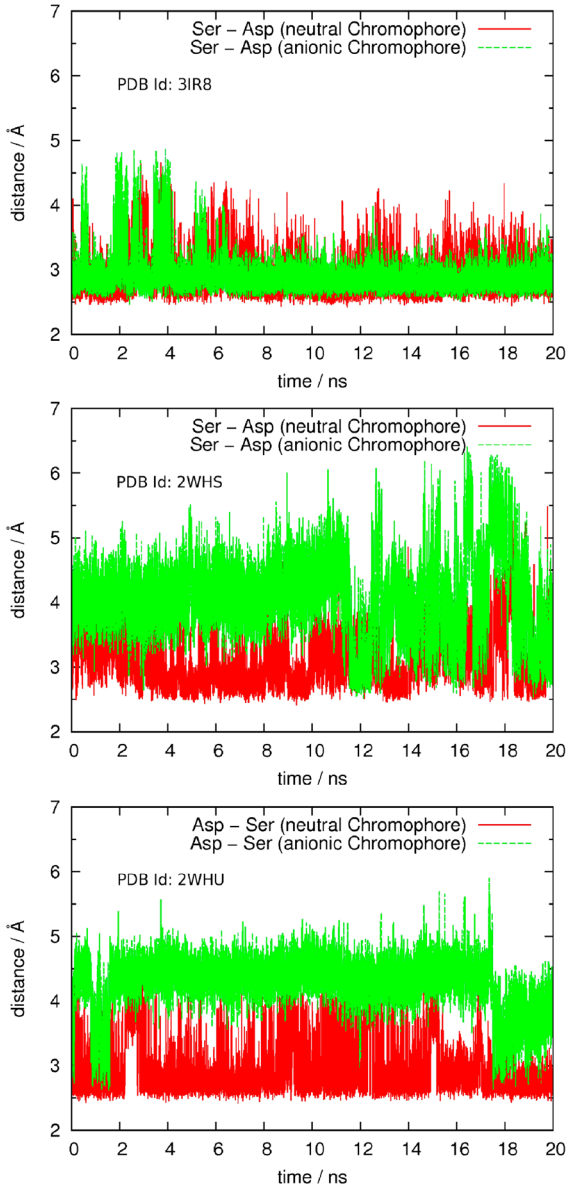


Figure 4. Distances between the oxygen atoms of the serine and the aspartic residues throughout the molecular dynamics simulations.

This is so because along the optimization a water molecule has ended up H-bonding the chromophore, thus pushing away the Ser142 and somehow disrupting the proton wire. The last column contains the average of the values all over the dynamics simulation, with the corresponding standard deviation.

The results of the 20-ns-long molecular dynamics simulations are presented in Figures 3 and 4. It is clearly seen that the first link of the proton wire is well formed and stable over time for all the structures as the O−O distances betweeen the chromophore and the corresponding acceptor residue oscillate most of the time between 2.5 and 3 Å (Figure 3). The same can be said for the distances between the oxygen atoms of the serine and the aspartate residues (Figure 4) as far as the neutral chromophore is considered. However, for the structures originated from Blanchoin and co-workers, this distance noticeably increases for the anionic chromophores.

The last column of Table 2 gives the average of these distances for each case so that we can compare them with the X-ray data and the values previously obtained through energy minimizations. Qualitatively, the averaged values along the dynamics are in good accordance with X-ray data, though the former values are slightly larger. This is to be expected in part given the fact that the molecular dynamics were devised to simulate the protein under physiological conditions where the averaged distances between different residues are expected to be larger than the values measured when the structure is frozen. At any rate, it is worth noting that the X-ray values usually fall within the statistical range of variation of these distances along the whole simulation, measured by the corresponding standard deviation, a value which is also given in Table 2.

As for the comparison with the previous optimizations, it is clearly seen in Table 2 that optimizing just the fragment of the protein implied in the proton wire provides O−O distances clearly shorter than X-ray values (and so also much shorter than the MD-averaged values). This result clearly emphasizes the importance of taking into consideration the protein environment when analyzing the structure of the active center (in this

dx.doi.org/10.1021/ct301003t | J. Chem. Theory Comput. 2013, 9, 1731−1742

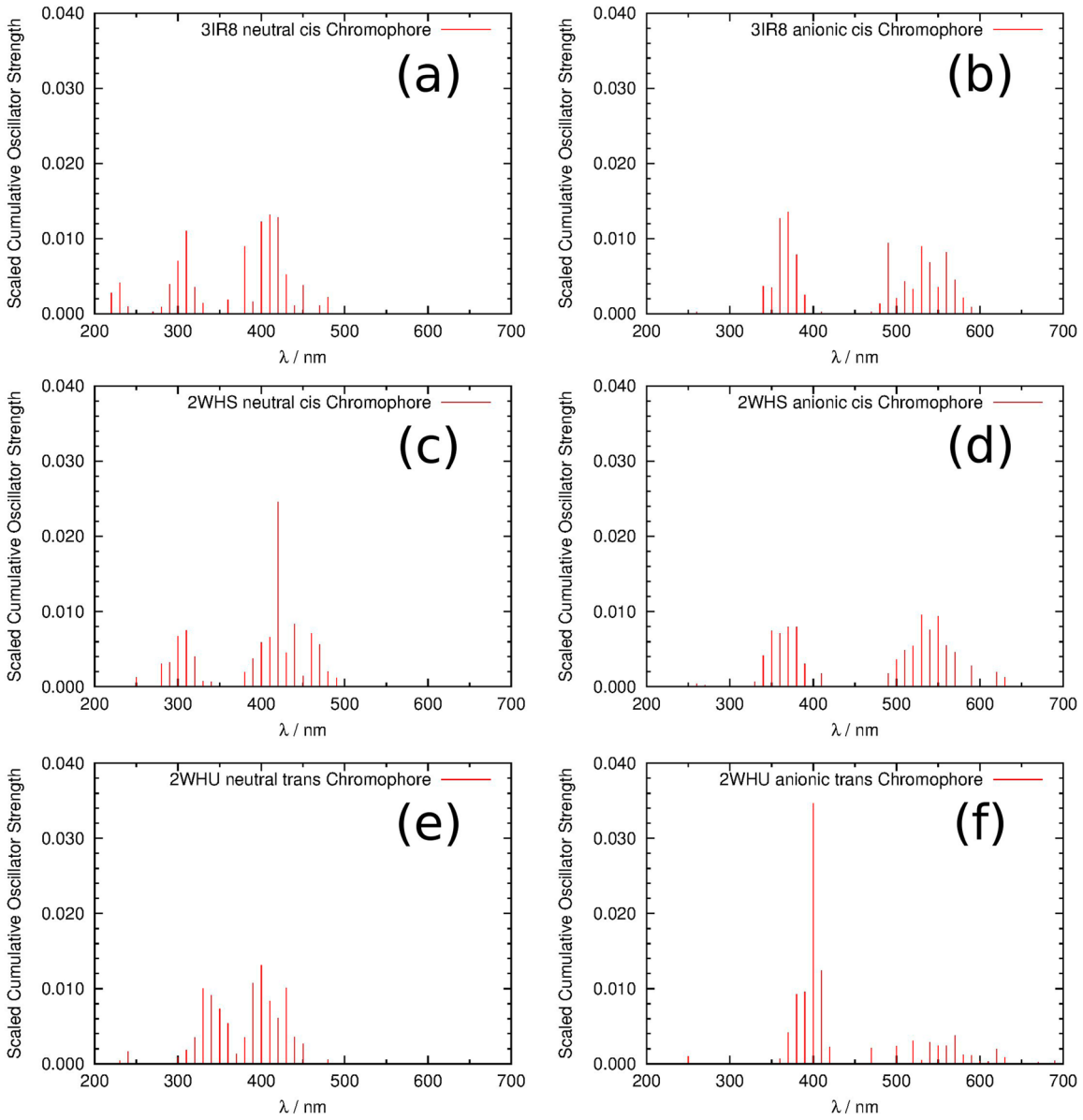


Figure 5. Representation of the absorption spectra calculated throughout the classical dynamics simulations (see text for the details). The graphics correspond to the models 3IR8[73] (*a* and *b*), 2WHS[72] (*c* and *d*), and 2WHU[72] (*e* and *f*), the chromophore being neutral (left) or anionic (right).

case the proton wire). It is tempting to verify whether more accurate results should be obtained when optimizing the whole system. Results in Table 2 show that, as is the case when comparing the values of the absorption wavelengths, they lead to quite disparate results. This confirms that optimizing a large molecular system might provide a geometry that is not necessarily representative of the actual structure of the molecule, as its dynamics can span a large region of the potential energy surface in a relatively short period of time.

A qualitative analysis of Figure 2 in reference 72 shows that the position of the chromophore and its planarity could also be stabilized because of the charged amino acids Arg91 and Glu211, which would have a nonbonding interaction with, respectively, one oxygen and one nitrogen atom situated in one of the coplanar rings of the chromophore. It can be seen in Table 2 that, when the chromophore is neutral, both the optimization and the molecular dynamics lead to distances that account for a nonbonding stabilizing interaction. When the chromophore is anionic, the Cro−Arg91 distance remains small. As for the Cro−Glu211 distance, it becomes larger than found in the crystal structure after optimizing or during the molecular dynamics, but not enough to make the nonbonding stabilizing interaction irrelevant. In all cases, we have considered the shortest interaction distance. Results in Table 2 suggest that these two residues, especially Arg91, play an important role in the dynamics of the chromophore and, consequently, in modulating its photochemical properties.

3.3. Theoretical Simulation of the Absorption Spectra.

Each of the 20-ns-long simulations has been used to extract a snapshot every 0.5 ns (40 snapshots per simulation, 240 snapshots total). The snapshots derived from each molecular dynamics run form an ensemble of structures that is expected to be more representative of the state of the protein in solution under physiological conditions than the optimized minimum would be.

That way, we have repeated the *Whole System* calculations on the considered snapshots of our dynamics simulations. Then, all the different excited states of a single dynamics simulation with

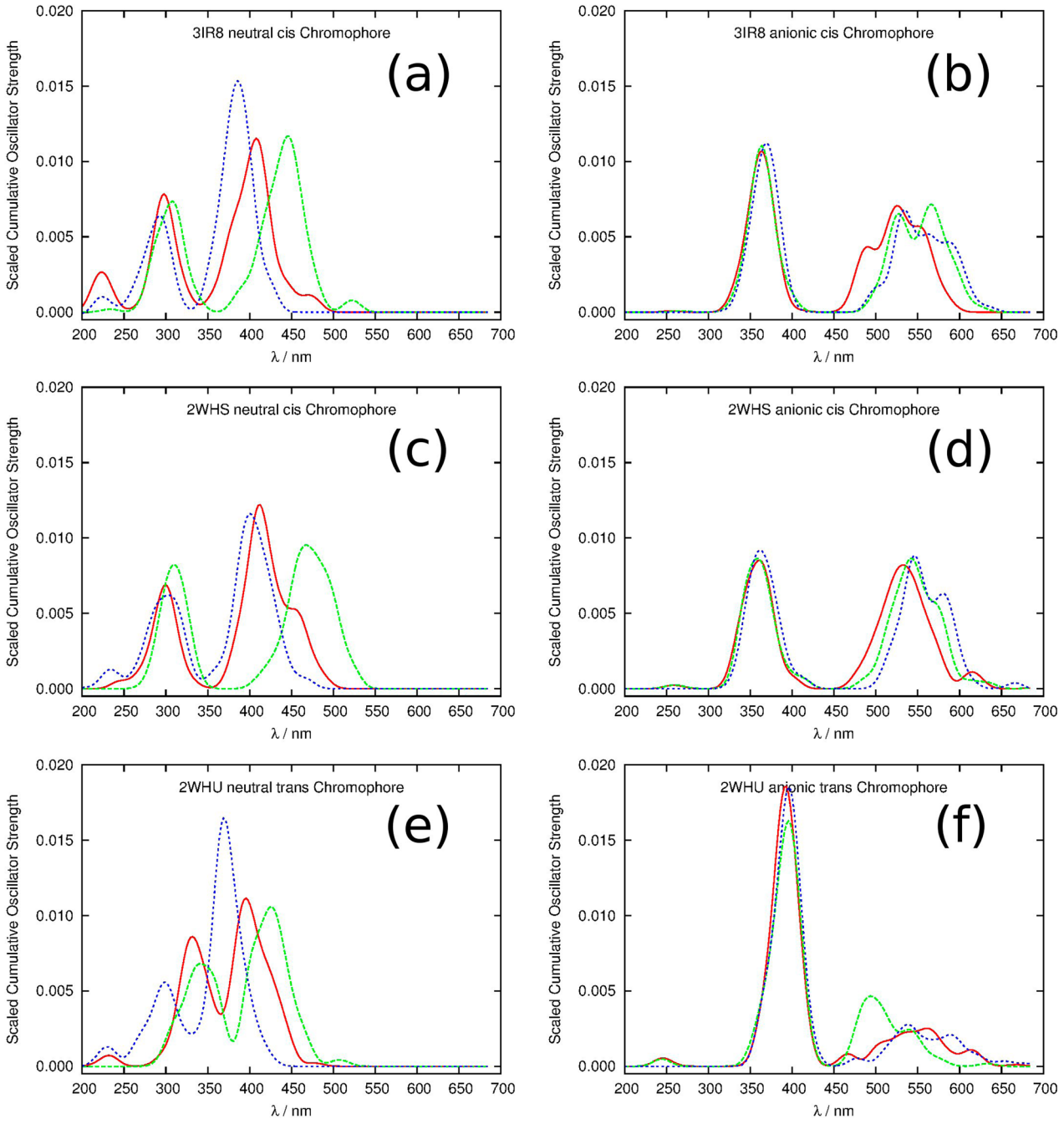


Figure 6. Representation of the absorption spectra considering the *Whole System* (red), the *Wire* atoms (green), and the *Only Chromophore* (blue) models. These correspond to the PDB structures 3IR8 with a protonated (a) and anionic (b) chromophore, 2WHS with a protonated (c) and anionic (d) chromophore, and 2WHU with a protonated (e) and anionic (f) chromophore. In all cases, a convoluting Gaussian function of σ = 10 nm has been used to smooth the statistical noise. Note that the red curves in spectra a–f above correspond to the convolution of data shown in Figure 5.

oscillator strength (f) greater than 0.1 are sorted by their frequencies and a wavelength histogram is done with a bin size of 10 nm (Figure 5) and normalizing the integral. These histograms represent 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 spectra, but it is a feasible and plausible way to simulate it. A similar procedure has already been used by Thiel and co-workers,[108] by Imhof,[109] by Martínez and co-workers,[110] and by us[111] to characterize the absorption spectra of the chromophore of other FPs.

Apart from the fact that excited states with higher energies have been found, all protonated chromophore (both *cis* and *trans*) models exhibit an absorption band with a maximum between 400 and 450 nm, which matches pretty well with the 440 nm absorption maximum experimentally detected. As for the anionic chromophore models, both *cis* and *trans* show a

dx.doi.org/10.1021/ct301003t | J. Chem. Theory Comput. 2013, 9, 1731–1742

band between 500 and 600 nm. This band is slightly blue-shifted with respect to the detected absorption maximum at 590 nm, yet it is not necessarily a bad description, for it must be considered that the calculations involve energies, which do not depend linearly of wavelengths, so the same energy difference errors, for lower energies, yield greater errors in wavelengths.

In all the calculated absorption spectra shown in Figure 5, there appears a higher energy (lower wavelength) band between 300 and 400 nm. This band comes from transitions to excited states of higher energy that have a non-negligible oscillator strength ($f > 0.1$, as explained above). For the neutral chromophore cases (spectra on the left of Figure 5), this band is clearly less intense than the "main" absorption band centered at wavelengths slightly above 400 nm. On the contrary, for the anionic chromophores (spectra on the right of Figure 5), this higher energy band competes on equal grounds with the experimentally measured band at wavelengths above 500 nm and, in the case of the *trans* chromophore (Figure 5f), is clearly the most intense band. This band cannot be compared with the experiments, as it lies outside the published spectral range in previous works.[2,72]

We have deemed it interesting to analyze separately the effect on the spectra caused by thermal agitation (given the multiplicity of molecular configurations) and the electrostatic effect of the environment. We have tackled this by taking the different geometries corresponding to the *only chromophore* and *wire* models from the ensemble of configurations with which the spectra have been simulated. With these geometries, the same TDDFT calculations have been done to each snapshot, and similar simulated spectra have been done. For the sake of clarity, the results are presented in Figure 6 as convoluted spectra with a convoluting Gaussian function of $\sigma = 10$ nm. A similar idea to assess the electrostatic effect of the environment in a protein has already been reported by Lopez and co-workers.[52]

Comparing Table 1 with Figure 6, we can easily notice the importance of taking an ensemble of structures instead of only one. For example, for the structure 3IR8, the neutral *only chromophore* presents an absorption wavelength at 551 nm (Table 1). However, Figure 6a shows that the blue line (which corresponds to the isolated chromophore) shows no signal at this wavelength. The *optimized chromophore* absorbs at 369 nm (Table 1), and again Figure 6a shows some signal at this wavelength, but it is not the maximum. A similar analysis for the *wire* model (green line in Figure 6a) could be done with similar conclusions. For the anionic *only chromophore*, it happens the other way around. The crystallographic structure absorbs at 575 nm (Table 1), which in Figure 6b corresponds to a populated region (although not the maximum), but the optimized chromophore absorbs at 486 nm (Table 1), which corresponds to a completely flat area of the blue line in Figure 6b. Again a completely analogous analysis could be done for the *wire* model.

It can be seen that introducing electrostatic effects affects the position and width of the band. However, the separation between thermal and electrostatic effects is not thorough because all the structures have been generated with molecular dynamics performed using a force field that includes all the electrostatic interactions from the environment with the chromophore, instead of making a simulation where these interactions had been taken out: these simulations would likely tear apart these models.

4. CONCLUSIONS

In this paper, we have studied the light absorption properties of mKeima, a red fluorescent protein of particular interest: not only for its potential biomedical applications but because several crystallographic structures have been recently published, showing significant differences between them. To this purpose, different TDDFT calculations have been performed.

Taking different static models, ranging from a few atoms to the whole protein, and considering even QM/MM optimizations of the structure, it has been impossible to reproduce simultaneously the transitions corresponding to the neutral and anionic chromophores (440 and 590 nm, respectively). One plausible explanation of this is that the energy minimum is not necessarily populated in terms of residence time. Besides, single point calculations cannot reproduce the absorption band. To avoid these problems, we have done an extensive sampling of the regions that are populated at 300 K, so that thermal effects are introduced by combining classical dynamics simulations and TDDFT electronic calculations. This way, we have been able to characterize both bands, thus relying on the nature of the absorption signals associated with the protonation of the structures. We have seen that the protein environment that surrounds the chromophore (both the identity of the residues and the thermal motion) has a noticeable effect on the absorption spectra (position and width of the bands).

Our classical molecular dynamics simulations on mKeima have proved, as far as our results indicate, that the several existing crystallographic structures are stable in solution at room temperature, and the distances between the residues involved in the proton transfer have remained stable at the common hydrogen bond distances.

Starting from different crystallographic structures of the protein, we have asserted that all protonated chromophore (both *cis* and *trans*) models exhibit an absorption band with a maximum between 400 and 450 nm, which matches pretty well with the 440 nm absorption maximum experimentally detected. As for the anionic chromophore models, both *cis* and *trans* show a band between 500 and 600 nm.

■ ASSOCIATED CONTENT

Supporting Information

Values of the fitted classical parameters for the CHARMM22 force field specific to the mKeima chromophore and some of the structures corresponding to the calculations in Table 1. This material is available free of charge via the Internet at http://pubs.acs.org

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Miquel.Moreno@uab.cat.

Notes

The authors declare no competing financial interest.

■ 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. 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) Kogure, T.; Karasawa, S.; Araki, T.; Saito, K.; Kinjo, M.; Miyawaki, A. *Nat. Biotechnol.* **2006**, *24*, 577−581.
(2) Verkhusha, V. V.; Piatkevich, K. D.; Hulit, J.; Subach, O. M.; Wu, B.; Abdulla, A.; Segall, J. E. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 5369−5374.
(3) Verkhusha, V. V.; Subach, F. V. *Chem. Rev.* **2012**, *112*, 4308−4327.
(4) Shimomura, O. *Angew. Chem., Int. Ed.* **2009**, *48*, 5590−5602.
(5) Chalfie, M. *Angew. Chem., Int. Ed.* **2009**, *48*, 5603−5611.
(6) Tsien, R. Y. *Angew. Chem., Int. Ed.* **2009**, *48*, 5612−5626.
(7) Zimmer, M. *Chem. Rev.* **2002**, *102*, 759−781.
(8) Piatkevich, K. D.; Malashkevich, V. N.; Almo, S. C.; Verkhusha, V. V. *J. Am. Chem. Soc.* **2010**, *132*, 10762−10770.
(9) Hsiao, Y. W.; Thiel, W. *J. Phys. Chem. B* **2011**, *115*, 2097−2106.
(10) Shcherbakova, D. M.; Hink, M. A.; Joosen, L.; Gadella, T. W. J.; Verkhusha, V. V. *J. Am. Chem. Soc.* **2012**, *134*, 7913−7923.
(11) Verkhusha, V. V.; Lukyanov, K. A. *Nat. Biotechnol.* **2004**, *22*, 289−296.
(12) Camilloni, C.; Provasi, D.; Tiana, G.; Broglia, R. A. *J. Phys. Chem. B* **2007**, *111*, 10807−10812.
(13) Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Am. Chem. Soc.* **2006**, *128*, 3564−3574.
(14) Grigorenko, B.; Savitsky, A.; Topol, I.; Burt, S.; Nemukhin, A. *Chem. Phys. Lett.* **2006**, *424*, 184−188.
(15) Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Chem. Theory Comput.* **2008**, *4*, 1138−1150.
(16) Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Phys. Chem. B* **2008**, *112*, 5500−5511.
(17) Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Phys. Chem. B* **2008**, *112*, 13443−13452.
(18) Schafer, L. V.; Groenhof, G.; Klingen, A. R.; Ullmann, G. M.; Boggio-Pasqua, M.; Robb, M. A.; Grubmuller, H. *Angew. Chem., Int. Ed.* **2007**, *46*, 530−536.
(19) Kamiya, M.; Saito, S.; Ohmine, I. *J. Phys. Chem. B* **2007**, *111*, 2948−2956.
(20) Taguchi, N.; Mochizuki, Y.; Nakano, T.; Amari, S.; Fukuzawa, K.; Ishikawa, T.; Sakurai, M.; Tanaka, S. *J. Phys. Chem. B* **2009**, *113*, 1153−1161.
(21) Gonzalez, E. M.; Guidoni, L.; Molteni, C. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4556−4563.
(22) Lelimousin, M.; Adam, V.; Nienhaus, G. U.; Bourgeois, D.; Field, M. J. *J. Am. Chem. Soc.* **2009**, *131*, 16814−16823.
(23) Filippi, C.; Zicchedda, M.; Buda, F. *J. Chem. Theory Comput.* **2009**, *5*, 2074−2087.
(24) Voityuk, A. A.; Michel; Beyerle, M. E.; Rosch, N. *Chem. Phys.* **1998**, *231*, 13−25.
(25) Weber, W.; Helms, V.; McCammon, J. A.; Langhoff, P. W. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 6177−6182.
(26) Laino, T.; Nifosi, R.; Tozzini, V. *Chem. Phys.* **2004**, *298*, 17−28.
(27) Helms, V.; Winstead, C.; Langhoff, P. W. *J. Mol. Struct.: THEOCHEM* **2000**, *506*, 179−189.
(28) El Yazal, J.; Prendergast, F. G.; Shaw, D. E.; Pang, Y. P. *J. Am. Chem. Soc.* **2000**, *122*, 11411−11415.
(29) Marques, M. A. L.; Lopez, X.; Varsano, D.; Castro, A.; Rubio, A. *Phys. Rev. Lett.* **2003**, 90.
(30) Martin, M. E.; Negri, F.; Olivucci, M. *J. Am. Chem. Soc.* **2004**, *126*, 5452−5464.
(31) Sinicropi, A.; Andruniow, T.; Ferre, N.; Basosi, R.; Olivucci, M. *J. Am. Chem. Soc.* **2005**, *127*, 11534−11535.
(32) Das, A. K.; Hasegawa, J. Y.; Miyahara, T.; Ehara, M.; Nakatsuji, H. *J. Comput. Chem.* **2003**, *24*, 1421−1431.
(33) Toniolo, A.; Olsen, S.; Manohar, L.; Martinez, T. J. *Faraday Discuss.* **2004**, *127*, 149−163.
(34) Xie, D. Q.; Zeng, J. *J. Comput. Chem.* **2005**, *26*, 1487−1496.
(35) Nemukhin, A. V.; Topol, I. A.; Burt, S. K. *J. Chem. Theory Comput.* **2006**, *2*, 292−299.
(36) Wan, S. B.; Liu, S. S.; Zhao, G. J.; Chen, M. D.; Han, K.; Sun, M. T. *Biophys. Chem.* **2007**, *129*, 218−223.
(37) Hasegawa, J. Y.; Fujimoto, K.; Swerts, B.; Miyahara, T.; Nakatsuji, H. *J. Comput. Chem.* **2007**, *28*, 2443−2452.
(38) Bravaya, K. B.; Bochenkova, A. V.; Granovskii, A. A.; Nemukhin, A. V. *Russ. J. Phys. Chem. B* **2008**, *2*, 671−675.
(39) Li, X.; Chung, L. W.; Mizuno, H.; Miyawaki, A.; Morokuma, K. *J. Phys. Chem. Lett.* **2010**, *1*, 3328−3333.
(40) Matsuura, A.; Hayashi, T.; Sato, H.; Takahashi, A.; Sakurai, M. *Chem. Phys. Lett.* **2010**, *484*, 324−329.
(41) Li, X.; Chung, L. W.; Mizuno, H.; Miyawaki, A.; Morokuma, K. *J. Phys. Chem. B* **2010**, *114*, 1114−1126.
(42) Li, X.; Chung, L. W.; Mizuno, H.; Miyawaki, A.; Morokuma, K. *J. Phys. Chem. B* **2010**, *114*, 16666−16675.
(43) Bravaya, K. B.; Khrenova, M. G.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. *J. Phys. Chem. B* **2011**, *115*, 8296−8303.
(44) Hasegawa, J. Y.; Fujimoto, K. J.; Nakatsuji, H. *ChemPhysChem* **2011**, *12*, 3106−3115.
(45) Voityuk, A. A.; Michel-Beyerle, M. E.; Rosch, N. *Chem. Phys.* **1998**, *231*, 13−25.
(46) Voityuk, A. A.; Michel; Beyerle, M. E.; Rosch, N. *Chem. Phys. Lett.* **1998**, *296*, 269−276.
(47) Voityuk, A. A.; Kummer, A. D.; Michel; Beyerle, M. E.; Rosch, N. *Chem. Phys.* **2001**, *269*, 83−91.
(48) Nifosi, R.; Amat, P.; Tozzini, V. *J. Comput. Chem.* **2007**, *28*, 2366−2377.
(49) Epifanovsky, E.; Polyakov, I.; Grigorenko, B.; Nemukhin, A.; Krylov, A. I. *J. Chem. Phys.* **2010**, 132.
(50) Toniolo, A.; Granucci, G.; Martinez, T. J. *J. Phys. Chem. A* **2003**, *107*, 3822−3830.
(51) Demachy, I.; Ridard, J.; Laguitton; Pasquier, H.; Durnerin, E.; Vallverdu, G.; Archirel, P.; Levy, B. *J. Phys. Chem. B* **2005**, *109*, 24121−24133.
(52) Lopez, X.; Marques, M. A. L.; Castro, A.; Rubio, A. *J. Am. Chem. Soc.* **2005**, *127*, 12329−12337.
(53) Mochizuki, Y.; Nakano, T.; Amari, S.; Ishikawa, T.; Tanaka, K.; Sakurai, M.; Tanaka, S. *Chem. Phys. Lett.* **2007**, *433*, 360−367.
(54) Hasegawa, J.; Ise, T.; Fujimoto, K. J.; Kikuchi, A.; Fukumura, E.; Miyawaki, A.; Shiro, Y. *J. Phys. Chem. B* **2010**, *114*, 2971−2979.
(55) Filippi, C.; Buda, F.; Guidoni, L.; Sinicropi, A. *J. Chem. Theory Comput.* **2012**, *8*, 112−124.
(56) Bravaya, K. B.; Grigorenko, B. L.; Nemukhin, A. V.; Krylov, A. I. *Acc. Chem. Res.* **2012**, *45*, 265−275.
(57) Ma, Y. Y.; Sun, Q.; Li, Z.; Yu, J. G.; Smith, S. C. *J. Phys. Chem. B* **2012**, *116*, 1426−1436.
(58) Grigorenko, B. L.; Nemukhin, A. V.; Morozov, D. I.; Polyakov, I. V.; Bravaya, K. B.; Krylov, A. I. *J. Chem. Theory Comput.* **2012**, *8*, 1912−1920.
(59) Topol, I.; Collins, J.; Polyakov, I.; Grigorenko, B.; Nemukhin, A. *Biophys. Chem.* **2009**, *145*, 1−6.
(60) Polyakov, I. V.; Grigorenko, B. L.; Epifanovsky, E. M.; Krylov, A. I.; Nemukhin, A. V. *J. Chem. Theory Comput.* **2010**, *6*, 2377−2387.
(61) Topol, I.; Collins, J.; Savitsky, A.; Nemukhin, A. *Biophys. Chem.* **2011**, *158*, 91−95.
(62) Topol, I.; Collins, J.; Nemukhin, A. *Biophys. Chem.* **2010**, *149*, 78−82.
(63) Sanchez-Garcia, E.; Doerr, M.; Hsiao, Y. W.; Thiel, W. *J. Phys. Chem. B* **2009**, *113*, 16622−16631.
(64) Nemukhin, A. V.; Grigorenko, B. L.; Savitsky, A. P. *Acta Nat.* **2009**, *1*, 41−52.
(65) Wang, S. F.; Smith, S. C. *Chem. Phys.* **2006**, *326*, 204−209.
(66) Zhang, H.; Sun, Q.; Wang, S.; Olsen, S.; Smith, S. C. In *Hydrogen Bonding and Transfer in the Excited State*; John Wiley & Sons: New York, 2011; p 815.
(67) Jackson, S.; Sanders, J., Eds. 2009 Green Fluorescent Protein Issue, *Chem. Soc. Rev.* **2009**

(68) Bravaya, K. B.; Subach, O. M.; Korovina, N.; Verkhusha, V. V.; Krylov, A. I. *J. Am. Chem. Soc.* **2012**, *134*, 2807−2814.
(69) Shaner, N. C.; Patterson, G. H.; Davidson, M. W. *J. Cell Sci.* **2007**, *120*, 4247−4260.
(70) Shcherbo, D.; Shemiakina, I. I.; Ryabova, A. V.; Luker, K. E.; Schmidt, B. T.; Souslova, E. A.; Gorodnicheva, T. V.; Strukova, L.; Shidlovskiy, K. M.; Britanova, O. V.; Zaraisky, A. G.; Lukyanov, K. A.; Loschenov, V. B.; Luker, G. D.; Chudakov, D. M. *Nat. Methods* **2010**, *7*, 827−U1520.
(71) Morozova, K. S.; Piatkevich, K. D.; Gould, T. J.; Zhang, J. H.; Bewersdorf, J.; Verkhusha, V. V. *Biophys. J.* **2010**, *99*, L13−L15.
(72) Violot, S.; Carpentier, P.; Bourgeois, D.; Blanchoin, L. *J. Am. Chem. Soc.* **2009**, *131*, 10356−10357.
(73) Henderson, J. N.; Osborn, M. F.; Koon, N.; Gepshtein, R.; Huppert, D.; Remington, S. J. *J. Am. Chem. Soc.* **2009**, *131*, 13212−13213.
(74) Lill, M. A.; Helms, V. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 2778−2781.
(75) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Milliam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Kakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.
(76) Handy, N. C. *Mol. Phys.* **2004**, *102*, 2399−2409.
(77) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724−728.
(78) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257−2261.
(79) Harihara, Pc; Pople, J. A. *Mol. Phys.* **1974**, *27*, 209−214.
(80) Gordon, M. S. *Chem. Phys. Lett.* **1980**, *76*, 163−168.
(81) Harihara, Pc; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213−222.
(82) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; Defrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654−3665.
(83) Binning, R. C.; Curtiss, L. A. *J. Comput. Chem.* **1990**, *11*, 1206−1216.
(84) Blaudeau, J. P.; McGrath, M. P.; Curtiss, L. A.; Radom, L. *J. Chem. Phys.* **1997**, *107*, 5016−5021.
(85) Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. *J. Chem. Phys.* **1998**, *109*, 1223−1229.
(86) Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A. *J. Comput. Chem.* **2001**, *22*, 976−984.
(87) Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Allaham, M. A.; Shirley, W. A.; Mantzaris, J. *J. Chem. Phys.* **1988**, *89*, 2193−2218.
(88) Petersson, G. A.; Allaham, M. A. *J. Chem. Phys.* **1991**, *94*, 6081−6090.
(89) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265−3269.
(90) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. *J. Comput. Chem.* **1983**, *4*, 294−301.
(91) Randino, C.; Moreno, M.; Gelabert, R.; Lluch, J. M. *J. Phys. Chem. B* **2012**, *116*, 14302−14310.
(92) Sobolewski, A. L.; Domcke, W. *Phys. Chem. Chem. Phys.* **1999**, *1*, 3065−3072.
(93) Aquino, A. J. A.; Lischka, H.; Hattig, C. *J. Phys. Chem. A* **2005**, *109*, 3201−3208.
(94) Zhang, M. Y.; Wang, J. Y.; Lin, C. S.; Cheng, W. D. *J. Phys. Chem. B* **2011**, *115*, 10750−10757.
(95) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph; McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz; Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(96) MacKerell, A. D., Jr.; Brooks, C., III; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Ed.; John Wiley & Sons: Chichester, U. K., 1998; pp 271−277.
(97) MacKerell, A. D.; Feig, M.; Brooks, C. L. *J. Am. Chem. Soc.* **2004**, *126*, 698−699.
(98) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjovoll, M.; Fahmi, A.; Schafer, A.; Lennartz, C. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 1−28.
(99) Smith, W.; Forester, T. R. *J. Mol. Graphics* **1996**, *14*, 136−141.
(100) Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177−2186.
(101) Baker, N. A.; Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A. *Nucleic Acids Res.* **2004**, *32*, W665−W667.
(102) Baker, N. A.; Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; Klebe, G. *Nucleic Acids Res.* **2007**, *35*, W522−W525.
(103) Olsen, S.; Smith, S. C. *J. Am. Chem. Soc.* **2007**, *129*, 2054−2065.
(104) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(105) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(106) Reuter, N.; Lin, H.; Thiel, W. *J. Phys. Chem. B* **2002**, *106*, 6310−6321.
(107) Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 8362−8367.
(108) Sanchez; Garcia, E.; Doerr, M.; Thiel, W. *J. Comput. Chem.* **2010**, *31*, 1603−1612.
(109) Imhof, P. *J. Chem. Theory Comput.* **2012**, *8*, 4828−4836.
(110) Isborn, C. M.; Götz, A. W.; Clark, M. A.; Walker, R. C.; Martínez, T. J. *J. Chem. Theory Comput.* **2012**, *8*, 5092−5106.
(111) Randino, C.; Nadal-Ferret, M.; Gelabert, R.; Moreno, M.; Lluch, J. M. *Theor. Chem. Acc.* **2013**, *132*, 1327.