J|A|C|S ARTICLES

Published on Web 12/16/2009

Color-Tuning Mechanism of Firefly Investigated by Multi-Configurational Perturbation Method

Isabelle Navizet,[‡,||] Ya-Jun Liu,*[,‡] Nicolas Ferré,[⊥] Hong-Yan Xiao,[§] Wei-Hai Fang,*[,‡] and Roland Lindh*[,#]

College of Chemistry, Beijing Normal University, Beijing 100875, China, Université Paris-Est, Laboratoire Modélisation et Simulation Multi Echelle, MSME FRE3160 CNRS, 5 bd Descartes, 77454 Marne-la-Vallee, France, Universités d'Aix-Marseille I, II et III-CNRS UMR 6264: Laboratoire Chimie Provence, Equipe: Chimie Théorique, Faculté de St-Jérome, Case 521, 13397 Marseille Cedex 20, France, Key Laboratory of Photochemical Conversion and Optoelectronic Materials, Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, Beijing 100190, China, and Department of Theoretical Chemistry, Chemical Centre, Lund University, P.O. Box 124, S-22100 Lund, Sweden

Received September 22, 2009; E-mail: yajun.liu@bnu.edu.cn (Y.-J.L.); fangwh@bnu.edu.cn (W.-H.F.); roland.lindh@teokem.lu.se (R.L.)

Abstract: This is the first report on a multiconfigurational reference second-order perturbation theory−molecular mechanics study of the color modulation of the observed bioluminescence of the oxyluciferin−luciferase complex of the Japanese genji-botaru firefly using structures according to recent X-ray data. Our theoretical results do not support the experimentally deduced conclusion that the color modulation of the emitted light primarily depends on the size of the compact luciferase protein cavity embedding the excited oxyluciferin molecule. Rather, we find, in agreement with recent experimental observations, that the wavelength of the emitted light depends on the polarity of the microenvironment at the phenol/phenolate terminal of the benzothiazole fragment in oxyluciferin.

Introduction

Fireflies communicate with each other through flashes of visible light. The chemical origin of the color modulation in bioluminescence is still not well understood. For example, the same luciferin molecule, within various mutated forms of luciferase, can emit light at slightly different wavelengths, ranging from red to yellow to green.[1−5] To explain the variation in the color of the bioluminescence, four factors have been discussed[6] and five hypotheses proposed for firefly luminescence.[7] The generally accepted mechanism of firefly bioluminescence is shown in Figure 1. This luciferase-catalyzed reaction proceeds via the initial formation of an enzyme-bound luciferyl adenylate intermediate (luciferyl-AMP, see Figure 1). Nakatsu et al.[8] reported several X-ray crystal structures of the Japanese genji-botaru firefly (*Luciola cruciata*) luciferase. They replaced luciferyl-AMP by the stable analogue 5′-*O*-[*N*-(dehydroluciferyl)sulfamoyl] adenosine (DLSA, see Figure 2) and captured three key 'snapshots' of the reaction: (1) luciferase bound to the reactants (ATP-Mg^{2+}), (2) luciferase bound to the DLSA, and (3) luciferase bound to the products oxyluciferin-AMP ($OxyLH_2$-AMP). The structures of luciferase bound to the reactants and the products (1 and 3) are quite similar: both possess an active site that the authors qualified as "open" since the substrate is less packed by the environment surrounding the protein as compared to that for the luciferase−DLSA complex (2). When the protein is bound to DLSA, SER286 forms hydrogen bonds with TYR257 and ASN231 via a water molecule. This conformational change is also accompanied by a 6° rotation of the side chain of PHE249 toward DLSA and rotation of the side chain of ILE288 by 131° (see Figure 3a in

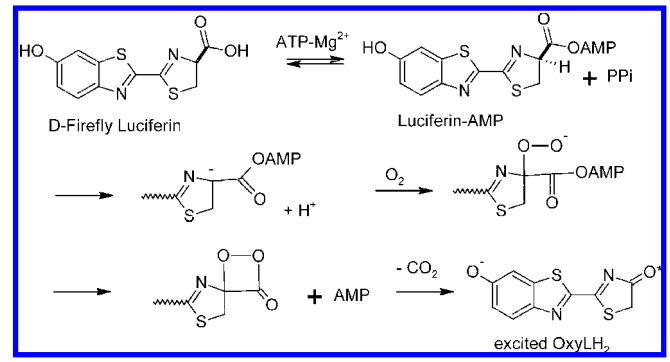

Figure 1. Generally accepted mechanism of firefly bioluminescence.

‡ Beijing Normal University.
|| Université Paris-Est.
§ Chinese Academy of Sciences.
⊥ Universités d'Aix-Marseille I, II et III-CNR.
Lund University.

(1) Hastings, J. W. *Gene* **1996**, *173*, 5–11.
(2) Kajiyama, N.; Nakano, E. *Protein Eng.* **1991**, *4*, 691–693.
(3) Mamaev, S. V.; Laikhter, A. L.; Arslan, T.; Hecht, S. M. *J. Am. Chem. Soc.* **1996**, 118.
(4) Ugarova, N. N.; Maloshenok, L. G.; Uporov, I. V.; Koksharov, M. I. *Biochemistry (Moscow)* **2005**, *70*, 1262–1267.
(5) Wood, K. V. *Photochem. Photobiol.* **1995**, *62*, 662–673.
(6) Coto, P. B.; Strambi, A.; Ferré, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17154–17159.
(7) Liu, Y.-J.; De Vico, L.; Lindh, R. *J. Photochem. Photobiol. A* **2008**, *194*, 261–267.
(8) Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A.; Kobashi, N.; Sakata, K.; Kato, H. *Nature* **2006**, *440*, 372–376.

10.1021/ja908051h

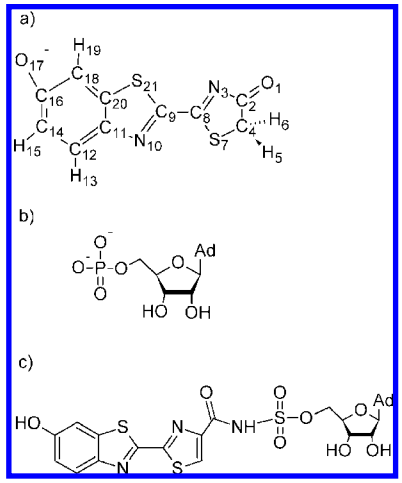

Figure 2. Structures of (a) **keto-1**, (b) AMP, and (c) DLSA.

ref 8). The result is a "closed" form of the active site, creating a structure in which the benzothiazole ring of DLSA is tightly sandwiched in a hydrophobic pocket including ILE288. Nakatsu et al. have proposed that controlling the hydrophobic microenvironment of the first excited single state (S_1) of the $OxyLH_2$ molecule is a key to understanding the bioluminescence color variations. To confirm this, they have linked the observed color modulation to the size of the protein pocket surrounding $OxyLH_2$. The authors have suggested that the "open" form allows some energy loss in the $OxyLH_2$ excited state, leading to the emission of low-energy red light, while the "closed" form, with its much more rigid microenvironment, minimizes this energy loss, thereby emitting a yellow-green light.

The proposed mechanism of controlling the amount of energy loss based on the size of the luciferase protein cavity contrasts with the conclusion of two recent experimental reports.[9,10] Hirano et al. studied the color modulation of the light emission as a function of various base/solvent combinations of an $OxyLH_2$ analog.[9] It was suggested that the color tuning might not be regulated by the rigidity of the active site but rather by the changes of polarity near the phenolate oxygen of the keto form of $OxyLH_2$. An investigation by Naumov et al. shed some additional light on this topic.[10] It was concluded that the enol form of $OxyLH_2$ can play a role in the emission and that several factors (pH, polarity of the microenvironment, presence of ionic species in the cavity, etc.) can act collectively in the protein scaffold during the tuning of the emission color.

Some previous theoretical studies dealing with the emission spectrum of the $OxyLH_2$ have been reported.[7,9,11−15] First, most of them were done *in vacuo*,[7,13,15] thus ignoring the steric and electrostatic contributions arising from the protein. Second, some of them were done at the density functional level of theory,[9,11,13,14] using standard functionals which are known to be unable to treat correctly the charge-transfer states.[16,17] Some theoretical studies have already been performed using the X-ray structures of firefly luciferase from *Photinus pyralis* determined with the absence of substrates[18] and *Luciola cruciata* as discussed above.[8] The two proposed models for the active site in *P. pyralis*, introduced by the Branchini and Ugarova groups,[19,20] were later confirmed by their similarities to the more current *L. cruciata* data. Two quantum-mechanics/molecular mechanics (QM/MM) studies have recently been reported.[15,21] In both cases, single-reference methods were employed for the QM subsection, and the red-to-green shift was not well reproduced. Furthermore, the microenvironmental contributions were not analyzed in detail with respect to the H-bonding network around $OxyLH_2$.

To improve upon the deficiencies of past theoretical studies, verify the conclusion in ref 8, solve the experimental discrepancy, and investigate the influence the microenvironment exerts on the multicolor emission of the firefly, we performed QM/MM calculations and classical molecular dynamics (MD) simulations based on the "open" and "closed" X-ray structures in ref 8 The present study utilizes geometry optimizations with the complete active space self-consistent field (CASSCF)[22] method. Subsequent calculations of the $S_1 \rightarrow S_0$ emission energy were established at the multiconfiguration reference second-order perturbation theory (CASPT2)[23,24] level of approximation. This combined approach has been demonstrated in the past to be reliable in studies of other photoactive proteins.[25] To the best of our knowledge, such multiconfigurational studies of the S_1 excited state of oxyluciferin in the luciferase protein are reported here for the first time. Considering the multiconfigurational nature of the conjugated system in oxyluciferin and the charge-transfer nature of the first excited state, this level of theory is mandatory.[15] The report is supplemented by Supporting Information providing additional information and details.

Computational Details

QM/MM Calculations. The calculations were performed using the QM/MM coupling scheme implemented on the basis of a modification of MOLCAS[26] and TINKER.[27,28] The electrostatic

(9) Hirano, T.; Hasumi, Y.; Ohtsuka, K.; Maki, S.; Niwa, H.; Yamaji, M.; Hashizume, D. *J. Am. Chem. Soc.* **2009**, *131*, 2385–2396.
(10) Naumov, P.; Ozawa, Y.; Ohkubo, K.; Fukuzumi, S. *J. Am. Chem. Soc.* **2009**, *131*, 11590–11605.
(11) Li, Z.-w.; Ren, A.-m.; Guo, J.-F.; Yang, T.; Goddard, J. D.; Feng, J.-K. *J. Phys. Chem. A* **2008**, *112*, 9796–9800.
(12) Nakatani, N.; Hasegawa, J.-Y.; Nakatsuji, H. *Chem. Phys. Lett.* **2009**, *469*, 191–194.
(13) Orlova, G.; Goddard, J. D.; Brovko, L. Y. *J. Am. Chem. Soc.* **2003**, *125*, 6962–6971.
(14) Ren, A.; Guo, J.; Feng, J.; Zou, L.; Li, Z.; Goddard, J. D. *Chin. J. Chem.* **2008**, *26*, 55–64.
(15) Yang, T.; Goddard, J. D. *J. Phys. Chem. A* **2007**, *111*, 4489–4497.
(16) Kobayashi, R.; Amos, R. D. *Chem. Phys. Lett.* **2006**, *420*, 106–109.
(17) Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. O.; Serrano-Andrés, L. *Mol. Phys.* **1999**, *97*, 859–868.
(18) Conti, E.; Franks, N. P.; Brick, P. *Structure* **1996**, *4*, 287–298.
(19) Branchini, B. R.; Magyar, R. A.; Murtiashaw, M. H.; Anderson, S. M.; Zimmer, M. *Biochemistry* **1998**, *37*, 15311–15319.
(20) Sandalova, T. P.; Ugarova, N. N. *Biochemistry (Moscow)* **1999**, *64*, 962–967.
(21) Tagami, A.; Ishibashi, N.; Kato, D.-I.; Taguchi, N.; Mochizuki, Y.; Watanabe, H.; Ito, M.; Tanaka, S. *Chem. Phys. Lett.* **2009**, *472*, 118–123.
(22) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem. Phys.* **1980**, *48*, 157–173.
(23) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483–5488.
(24) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218–1226.
(25) Ferré, N.; Olivucci, M. *J. Am. Chem. Soc.* **2003**, *125*, 6868–6869.
(26) Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. *Comput. Mater. Sci.* **2003**, *28*, 222–239.
(27) Aquilante, F.; Vico, L. D.; Ferré, N.; Ghigo, G.; Malmqvist, P.-Å.; Neogrády, P.; Pedersen, T. B.; Pitoňák, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. *J. Comput. Chem.* **2010**, *31*, 224–247.
(28) Ponder, J. W. *TINKER, Software Tools for Molecular Design*, version 4.2; Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine: St. Louis, MO, 2004. The most updated version for the TINKER program can be obtained from the Internet at http://dasher.wustl.edu/tinker.

potential fitted (ESPF) method[29] was used to compute the interaction between the charge distribution of the QM section and the external electrostatic potential created by the MM point charges. The microiterations technique was used to converge the MM subsystem geometry every QM minimization step. In all of the computational models, **keto-1** was selected as the QM region, while the remaining atoms were included in the MM region. The geometries were optimized for the S_1 state at the state-averaged CASSCF (SA-CASSCF) level of theory, as our study focused on the calculation of T_e values, i.e. the energy of the vertical emission transition from the S_1 state minimum to the S_0 state. Since no analytical second derivatives are currently available in the MOLCAS/TINKER QM/MM implementation and since a numerical procedure would be too expensive, the exact nature of the optimized QM/MM stationary points has not been verified. A 16-in-14 active space was selected, including all π/π^* orbitals except the one centered on the thiazolone sulfur atom. The 6-31G* basis set[30] was employed for the QM region and AMBER Parm99 force field,[31] as described below for the MM region. The transition energies have been computed at the CASPT2 level of theory using 2-roots SA-CASSCF wave functions.

Force Field Parameters for the S_1 State of OxyLH_2 and AMP. The AMBER Parm99 force field was used to model the residues of the protein. In accordance with the procedure by Meagher et al. in developing the parameters for ADP,[32] some missing parameters for AMP and OxyLH_2 were elaborated (see Supporting Information for details).

Models. The "open" form with AMP and OxyLH_2 (PDB code: 2D1R.pdb) and the "closed" form with DLSA (PDB code: 2D1S.pdb), both of them from ref 8, were used as starting structures. In the "closed" form the DLSA was replaced by keto-form-OxyLH_2 anion (**keto-1**) + AMP (see Figure 2). Although the latest experiments[10] suggest that OxyLH_2 could adopt forms besides **keto-1**, it is demonstrably the most probable template for firefly luminescence, according to both experimental[9,33] and theoretical[7,13,34] investigations.

The hydrogen atoms, the missing atoms, and some TIP3P water molecules[35] were added by the LEAP module of the AMBER9 suite of programs[36] to ensure a neutral rectangular system for a total of about 63 000 atoms. Building constraints were relaxed by using molecular mechanics minimization in AMBER9, resulting in two starting structures: Model-0-open/closed (see details in Supporting Information). It should be noted that the notion of one single cavity is somewhat misleading in the case of the luciferin−luciferase complex. There are actually two cavities, one surrounding the benzothiazole terminal and another around the thiazolone terminal of the oxyluciferin molecule. This will have some impact on the analysis of the hydrogen-bonding networks and their influence on color modulation. In the following section, three sets of structures denoted Model-x (x = 1,2,3) that are used in the subsequent calculations will be described.

Model-1. This set of structures includes **keto-1**, water molecules within 10 Å of **keto-1**, AMP, and luciferase residues, all extracted from the Model-0 structures. A QM/MM geometry optimization was performed with **keto-1** in its S_1 state included in the QM section. AMP, six side chains (Arg220, Phe249, Thr253, Ile288, Ser349, and Ala350), and water molecules within 5 Å of **keto-1** were also allowed to relax in the MM subsystem, while the other MM water and protein atoms were kept frozen (see Figure 3). The resulting molecular structures are denoted Model-1-open/closed.

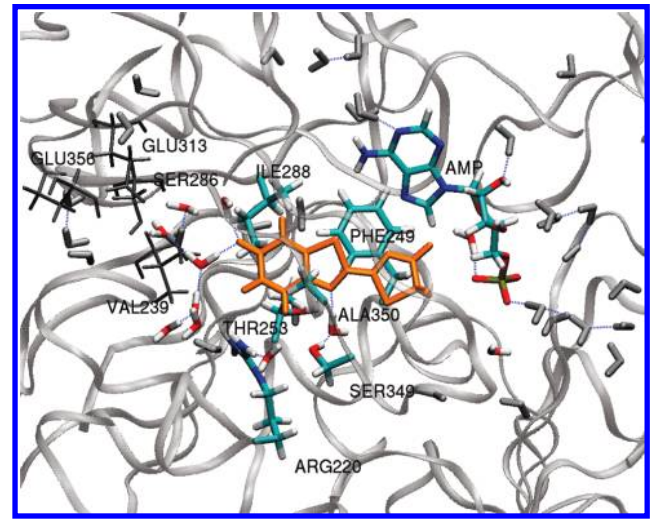

Figure 3. QM/MM computational model for Model-1-closed. The QM/MM optimization was performed with **keto-1** (in orange) in its S_1 state in the QM section. AMP and the six side chains (Arg220, Phe249, Thr253, Ile288, Ser349, and Ala350) and selected water molecules within 5 Å of **keto-1** (O in red, C in cyan, H in white, and P in tan) were allowed to move. The other atoms were frozen. The residues at the entrance of the cavity are drawn in licorice. H-bonds in the water network are colored blue. Model-1-open, Model-2-open/closed, and Model-3-open/closed have the same QM and MM sections, their only difference being the number of water molecules.

Model-2. The X-ray structures of the "open" and "closed" forms do not share the same number of water molecules inside the protein cavity. Therefore, the number of water molecules located at less than 10 Å of **keto-1** are not equal in Model-1-open and Model-1-closed. Two new structures have been built by adding eight missing water molecules to the X-ray "open" form in the positions found in X-ray "closed" form (corresponding to the phenol side of the cavity) and adding one missing water molecule near AMP to the X-ray "closed" form (see Figure S2, Supporting Information). After the MM minimization, the structures of **keto-1**, the water molecules within 10 Å of **keto-1**, AMP, and all protein residues were extracted. Using the same QM/MM partition as described for Model-1, the geometry of the S_1 state of **keto-1** was then optimized; the resulting molecular structures are denoted Model-2-open/closed.

Model-3. Starting with Model-0-open/closed, 2 ns of MD simulation were produced at a constant temperature (T = 300 K) and pressure (P = 1 atm), with a time step of 2 fs, default cutoff radius of 8 Å for nonbonding interactions, and periodic boundary conditions using the particle mesh Ewald method[37] and the SHAKE algorithm.[38] The minimum-energy structure that was found between 500 ps and 1.5 ns of the trajectories was extracted. During this period, the artifacts brought by the construction phase were removed, while the criteria defined in ref 8 (in order to discriminate between the "open" and the "closed" forms) were still satisfied. After the MM minimization, the QM/MM partition defined above was again employed, and the same geometry optimizations were performed. The corresponding structures are denoted Model-3-open/closed.

(29) Ferré, N.; Angyan, J. G. *Chem. Phys. Lett.* **2002**, *356*, 331–339.
(30) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213–222.
(31) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
(32) Meagher, K. L.; Redman, L. T.; Carlson, H. A. *J. Comput. Chem.* **2003**, *24*, 1016–1025.
(33) Branchini, B. R.; Murtiashaw, M. H.; Magyar, R. A.; Portier, N. C.; Ruggiero, M. C.; Stroh, J. G. *J. Am. Chem. Soc.* **2002**, *124*, 2112–2113.
(34) Nakatani, N.; Hasegawa, J.-y.; Nakatsuji, H. *J. Am. Chem. Soc* **2007**, *129*, 8756–8765.
(35) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(36) Case, D. A.; et al. *AMBER*, 9; University of California: San Francisco, CA, 2006.
(37) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
(38) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.

Table 1. Comparison of the CASPT2 Emission Energies (in eV) of **keto-1** Calculated in the Protein and *in Vacuo*[a]

computational model	in protein	*in vacuo*
Model-1-open/closed	2.05/2.13	2.03/2.02
Model-2-open/closed	2.15/2.01	2.03/2.00
Model-3-open/closed	2.25/2.22	2.05/2.04
exptl	2.05/2.21	

[a] All calculations have been done on the QM/MM CASSCF optimized S_1 geometries.

More details on the protocols employed for building the structures can be found in the Supporting Information.

Results and Discussion

Distortion of keto-1. All the structures show emission energies T_e in the range of the red and yellow-green visible light (see Table 1). The optimized structure of **keto-1** in the S_1 state in the Models-1,2,3-open/closed structures shows a slight distortion (bending and twisting) of oxyluciferin in the angle between the two heterocyclic rings ranging from 6° to 13°, in contrast to the planar structure obtained *in vacuo*.[7,13−15] This distortion is apparently a consequence of the asymmetry of the cavity in which the oxyluciferin moiety lies. It barely shows a small variation between the different "open" or "closed" structures, while other parameters, such as the bond lengths, are consistent with the optimized parameters *in vacuo* (see Table 2), in agreement with the recent experimental result showing a bending of 6° in the crystallographic structure of $OxyLH_2$.[10]

Model-1. For the Models-1,2,3 structures, we have verified that the cavities remain "closed" or "open", as described by Nakatsu and co-workers[8] (see Table S5, Supporting Information). According to Table 1, the calculated T_e values for the "open" and "closed" structures of Model-1 are in agreement with the experimental values.[8] The Model-1-open T_e is 0.08 eV red-shifted compared to the Model-1-closed T_e. To analyze the origin of this shift, the same QM/MM optimized geometries of **keto-1** in Model-1-open/closed were used to calculate the corresponding emission energies *in vacuo*. Both structures *in vacuo* showed close T_e values (2.03 and 2.02 eV) and differ from the corresponding values in the protein (2.05 and 2.13 eV) (see Table 1). Hence, the protein cavity not only perturbs the planar structure of the **keto-1**, something which does not directly change the emission energy,[7] but also affects the emission indirectly by a mechanism whose nature is still not clear.

Nakatsu et al.[8] have suggested that the smaller cavity size found in the "closed" form results in a constraint by the hydrophobic wall on the $OxyLH_2$ S_1 state, thereby minimizing the energy loss (due to thermal relaxation or a reorganization of specific interactions) before the emission. According to their hypothesis, the **keto-1** structures coming from Model-1-open and Model-1-closed should have different emission energies *in vacuo*, with the former being smaller. This suggestion is inconsistent with our present calculations. In fact, the two **keto-1** geometries are rather similar (see Table 2). Therefore, the criterion of the rigidity of the environment allowing less energy loss from the excited state does not explain the color modulation. Rather, the theoretical findings support the polarity of the microenvironment as the reason behind the color modulation, which would accord with experimental reports.[9]

In a detailed analysis of the environment close to **keto-1**, one notes that the H-bond networks of water molecules in Model-1-open/closed are significantly different (see Figure 4a and b). First, in the "open" form the thiazolone-ring O_1 builds an H-bond with a water molecule connected to the GLY318 residue, while in the "closed" form the water molecule nearest to O_1 is able to build two H-bonds with AMP and THR529. Second, the benzothiazole O_{17} has one H-bond with a water molecule in "open" form while it builds two H-bonds in the "closed" form. Moreover, residues like TYR229 (in Model-1-open) or HIS223 (in Model-1-closed) also take part in the H-bond network around O_{17}, leading to completely different microenvironments of **keto-1** in the "closed" or "open" structures. Is this an artifact of the structure-construction procedure employed? Is it due to the different number of crystallographic water molecules, depending on the substrate, or to the nature of the "open"/"closed" cavity? To investigate the first possibility, we analyzed structures with the same number of water molecules in the cavity (i.e., Model-2 structures).

Model-2. The Model-2 structures share the same cavity characteristics as Model-1; however, **keto-1** is now in a different H-bond network as compared to Model-1 (Figures 4c and S3 and S4, Supporting Information). The number of water molecules in the cavity on the side of benzothiazole is increased in Model-2-open as compared to Model-1-open (see Figures 4a and 4c). Now the water H-bond network in the cavity of Model-2-open also involves the residues ARG220 and ALA224. This shows that the H-bond network in the cavity depends on the number of water molecules used during the construction of the Model. In Model-1, we used the crystallographic water molecules found in the crystal structures, whose numbers depend on the substrate (DLSA or $OxyLH_2$) inside the protein. In Model-2, constructed with same number of water molecules in "open" and "closed" forms, the H-bond network depends on the orientation of the residues inside the cavity.

One notes that the T_e value calculated is now larger in Model-2-open than in Model-2-closed, which contrasts with the results of the Model-1 study (see Table 1). However, the corresponding **keto-1** QM/MM optimized geometries, Models-1,2-open/closed, are nearly identical and show almost the same computed *in vacuo* T_e values. This demonstrates that the H-bond network, the number and position of water molecules, and **keto-1** inside the cavity (see Figure S5, Supporting Information) are the major factors influencing the emission color.

The foregoing finding can be rationalized by the following arguments:

First, the S_1 to S_0 transition leads to an internal negative charge transfer from the thiazolone ring near the AMP to the benzothiazole ring (phenolate) (see graphic HOMO and LUMO in Figure 5). The electron transfer is approximately 0.2e for all structures (Table S7, Supporting Information), agreeing with the fact that a twisted **keto-1** leads to a significant electron transfer, as compared to that with the planar form[11]—even if the electron transfer is not as large as the 0.7e found in the fully twisted **keto-1** structure reported by Li et al.[11] Hence, any factor favoring the stabilization of the charge on the benzothiazole moiety will induce a stabilization of the ground state relative to the excited state, i.e., blue shifting of the transition energy. Second, different water (and residue) arrangements lead to different external electrostatic potentials on the QM **keto-1** moiety, particularly the benzothiazole fragment. If the external potential on the benzothiazole part stabilizes a charge there, then the transition energy should increase. In fact, in the transition from Model-1-open to Model-2-open, the presence of more water molecules in the cavity near the phenolate oxygen leads to a larger external potential felt by the phenolate oxygen atom

Table 2. CASSCF Optimized Geometries of **keto-1** S_1 State in the Different Models

	Model-1-open	Model-2-open	Model-3-open	Model-1-closed	Model-2-closed	Model-3-closed	*in vacuo*[a]
β^b	9.0	8.8	6.0	7.7	9.3	13.3	0.03
$C_{16}-O_{17}{}^c$	1.25	1.25	1.25	1.25	1.25	1.26	1.25
$C_2-O_1{}^c$	1.21	1.22	1.22	1.21	1.22	1.22	1.21
$C_9-C_8{}^c$	1.43	1.43	1.42	1.43	1.43	1.43	1.43
$N_{10}-C_9{}^c$	1.31	1.31	1.31	1.31	1.31	1.31	1.30
$N_3-C_8{}^c$	1.31	1.32	1.32	1.31	1.32	1.32	1.31
ΔE^d	2.03	2.03	2.05	2.02	2.00	2.04	

[a] CASSCF(18,15)/ANO-RCC-VDZP geometry from ref 7. [b] Angle between the (S_{21},C_9,N_{10}) plane and the (S_7,C_8,N_3) plane in degrees. [c] Bond length in Å. [d] Emission energy (in eV) in the gas phase employing the same structure optimized in the protein.

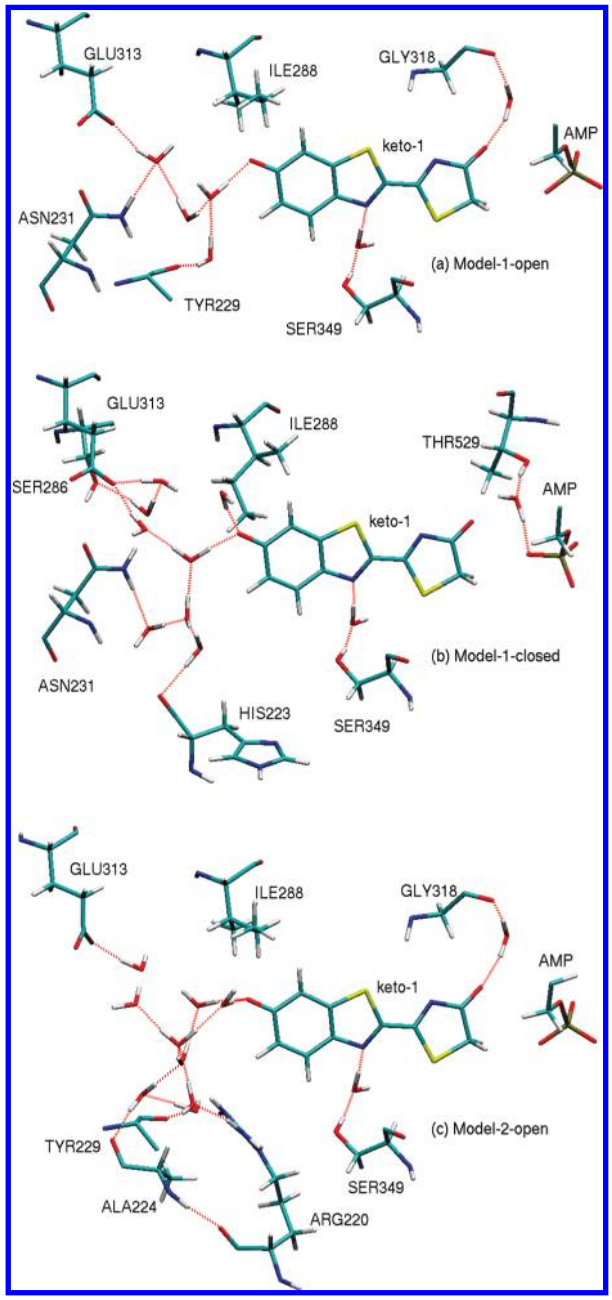

Figure 4. H-bond network of the water molecules near **keto-1** in Model-1-open (a), Model-1-closed (b), and Model-2-open (c). The residues or part of the residues represented are involved in the H-bond network. Additionally, ILE288 and the phosphate terminal of AMP are included. Atoms are color coded as O red, H white, C cyan, N blue, S yellow, and P tan.

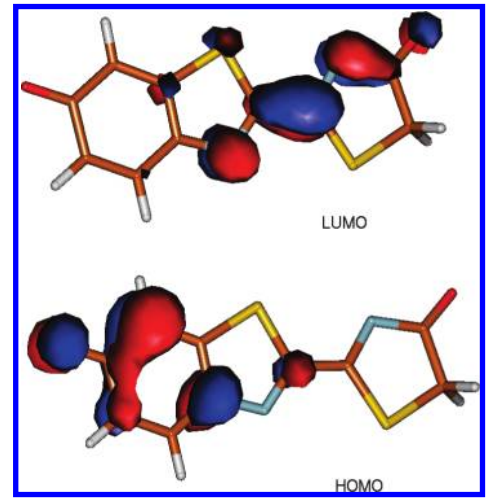

Figure 5. Calculated CASSCF orbitals LUMO and HOMO.

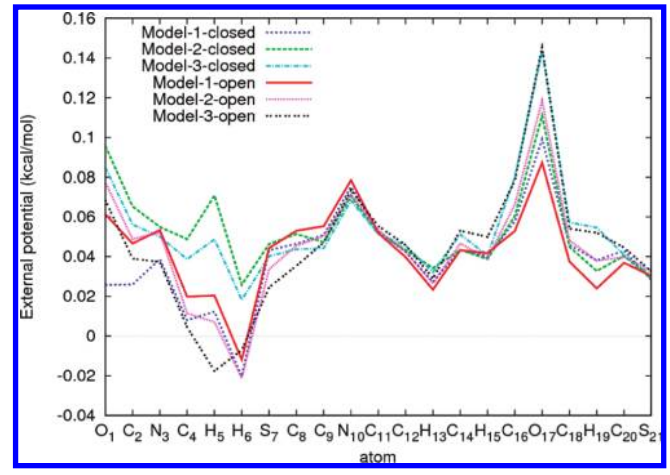

Figure 6. External potential applied to the QM **keto-1** structure due to the surrounding MM protein and water molecules (atom numbering of Figure 2).

(see Figure 6), thereby blue-shifting the transition energy by 0.1 eV. Note that the external potential acting on the thiazolone ring of $OxyLH_2$ can also influence the transition energy, but to a lesser extent. If the external potential created on the thiazolone oxygen by the surroundings is increased, then the excited state is more stabilized than the ground state, inducing a lower-energy transition. In Model-1-closed the phosphate group is oriented toward the **keto-1** thiazolone O_1, while it is not in Model-2-closed (see Figure S4, Supporting Information). The resulting external potential does not change much in the other parts of the molecule but is considerably stronger on the atoms of the thiazolone ring in Model-2-closed than in Model-1-closed (see Figure 6). The resulting transition energy in Model-2-closed is lowered by 0.12 eV, as compared to that of Model-1-closed. On the basis of the above calculations, one can conclude that

the emission energy can also be modulated by the position of AMP and the corresponding H-bonds in the cavity around the thiazolone fragment.

Model-3. Between the CO_2 elimination and the light emission, a process which takes more than one nanosecond,[9] the protein containing oxyluciferin in its excited state has sufficient time to partially relax, something not taken into account in Models-1,2. Hence we carried on MD simulations according to the protocol presented above. After 500 ps, the "open" and "closed" criteria remained valid. However, the Model-3-open/closed structures exhibited nearly the same emission energies (see Table 1), in disagreement with the "cavity size" hypothesis.[8] Again, for both structures, the emission energies computed *in vacuo,* using the corresponding QM/MM optimized geometries in protein, are nearly the same. Moreover, all the T_e values *in vacuo* are nearly the same for the six different structures (see Table 1), once again showing that cavity size has little or no influence on the **keto-1** geometry. The external potential is the same near the phenolate oxygen atom and not significantly different on the thiazolone-ring oxygen atom (see Figure 6). The difference in the cavity sizes of the "open" or "closed" structures is minimal, but it should be noted that, after equilibration, both cavities become more open to the external solvent reservoir (see Table S6, Supporting Information). Water molecules can enter and leave the cavities along the trajectories until three water molecules are finally H-bonded to O_{17} in both structures (see Figure S6, Supporting Information).

Additional Tests. To further confirm the above interpretation of our theoretical results, especially the influence of the water H-bond network bordering O_{17} and the orientation of AMP toward **keto-1**, some additional calculations were performed. The results of the different calculations are summarized in Table S4 of Supporting Information. Some tests concern the technical choice of the basis set (6-31G* or the larger ANO-L-VTZP[39] basis set) and the size of the active space (16-electrons-in-14-orbitals compared to 18-in-15). Another test shows that when the external potential contribution arising from the water molecules nearest to O_{17} is switched off in the Model-1-closed structure, a red-shift of 0.1 eV is observed, confirming the importance of the water molecules near the benzothiazole ring.

We also tried to characterize the influence of the protonation state of AMP. The calculation of emissions with protonated models (denoted AMPH) gave results somewhat similar to the one obtained using AMP in the model, without giving any information about the preferred protonation state of AMP (see Table S4, Supporting Information).

We also performed some tests with neutral keto-form $OxyLH_2$ (**keto**). The calculated T_e value of **keto** is 3.11 eV in the "closed" form, which is not within the visible light spectrum. Moreover, the oscillator strength of the **keto** in the "closed" form S_1 to S_0 transition (about 0.1) is much lower than those of the **keto-1** (about 0.9). Such a low oscillator strength is not compatible with an intense luminescence. As discussed by Hirano et al.,[9] the anion form is the most probable one for the light-emitter structure. The range of fluorescence emission maxima (2.32−2.95 eV) for the neutral-$OxyLH_2$ analogues in different solvents has only a small overlap with the bioluminescence emission maxima (1.94−2.34 eV). The same authors suggest that the transition might come from the protonation of one backbone residue of the protein.[9] In all the structures we have investigated, there is always a water molecule playing the role of the protonated residue near the phenolate oxygen, as described in the Hirano study.[9]

A New Reading of the Mutation Experiments. Our results show that the number of water molecules inside the cavity around the benzothiazole terminal has an impact on the color tuning. It may explain why the experimental mutations of residues located at the entrance to the benzothiazole cavity or involved in the water H-bond network of this cavity affect the color emission of the system. This appears to be the case for the singly mutated mutants of *Luciola cruciata* S286N and V239I (see Figure 3).[2] Likewise, the native red bioluminescent *Phrixothrix hirtus* luciferase features the presence of ARG353, corresponding to a missing residue near GLU356 (*L. cruciata* numbering).[40] This residue is also situated at the entrance to the water-filled benzothiazole cavity. It has similarly been shown that mutations in the loop between residues 223 to 235 (also located at the cavity door) play a major role in bioluminescence color and pH-sensitivity of luciferase.[41] All these experiments can be interpreted in the light of our new study.

In summary, these results and our calculations clearly indicate that the **keto-1** anion is one of the main oxyluciferin emitters. Moreover, microenvironmental modifications may destabilize or stabilize the charge localized on the benzothiazole moiety. These modifications can be induced by mutating some residues during experimental studies or by controlling the number of water molecules inside the cavity while performing *in silico* investigations. In both cases it is not the size of the cavity and its "open/closed" nature, but the H-bond network in the cavity involving water molecules and protein residues that mainly affects the charge redistribution in the oxyluciferin emitter during the S_1 to S_0 transition. This process is experimentally observed as the color modulation.

Conclusion

On the basis of the crystalline structures[8] of firefly *L. cruciata* luciferase, the mechanism of the firefly's multicolor bioluminescence was theoretically studied using a QM(CASPT2)/MM method for the first time. Our results demonstrate that (a) the phenomenon of multicolor bioluminescence is mainly related to the polarization of the close surroundings of $OxyLH_2$ (in support of two recent experimental findings),[9,10] and (b) the experimentally deduced luciferase cavity size model[8] does not impact the relative structures of the oxyluciferin substrate sufficiently to account for the observed color modulation. Indeed, our study points out the importance of the H-bond networks formed between the water molecules, **keto-1**, and the residues involved in the protein cavity. It may be noted that a similar dependence of the wavelength of the fluorescence on the hydrogen bond network connected to the fluorophore has been documented at the same level of theory as for that of the green fluorescence protein.[42] The mutations of the residues involved in the H-bond network can lead to different polarizations acting on the benzothiazole moiety and change the color of the emitted light accordingly. This study has been done on the **keto-1** form of $OxyLH_2$, which is broadly considered to be the species responsible for light emission inside the wild firefly. Other conformers suggested by recent experiments,[10] such as the anionic enol-form of $OxyLH_2$ and mutant proteins, are currently under study.

(39) Widmark, P.-O.; Persson, B. J.; Roos, B. O. *Theor. Chim. Acta* **1990**, *77*, 291–306.
(40) Alipour, B. S.; Hosseinkhani, S.; Ardestani, S. K.; Moradi, A. *Photochem. Photobiol. Sci.* **2009**, *8*, 847–855.
(41) Viviani, V. R.; Neto, A. J. S.; Arnoldi, F. G. C.; Barbosa, J. A. R. G.; Ohmiya, Y. *Photochem. Photobiol.* **2008**, *84*, 138–144.
(42) Sinicropi, A.; Andruniow, T.; Ferré, N.; Basosi, R.; Olivucci, M. *J. Am. Chem. Soc.* **2005**, *127*, 11534–11535.

Acknowledgment. This study was supported by the National Nature Science Foundation of China (Grant Nos. 20873010, 20673102 and 20720102038) and the Major State Basic Research Development Programs (Grant Nos. 2004CB719903 and 2007CB815206). The Swedish Research Council (VR) has also provided assistance. Some of the computations have been performed at the "Centre Régional de Compétences en Modélisation Moléculaire" in Marseilles, France. We thank Teddy Primack for assistance in preparation of the manuscript in English.

Supporting Information Available: Complete ref 36; absolute energies (in Hartrees) of all structures calculated; force field parameters; computational details; details on the molecular dynamic; superposition of different models; geometric data on the different models; evolution of the structures during the dynamic; and corresponding Tables S1−S8 and Figures S1−S8. This material is available free of charge via the Internet at http://pubs.acs.org.

JA908051H