J|A|C|S
ARTICLES

Published on Web 07/23/2009

Structure and Spectroscopy of Oxyluciferin, the Light Emitter of the Firefly Bioluminescence

Panče Naumov,* Yutaka Ozawa, Kei Ohkubo, and Shunichi Fukuzumi

Department of Material and Life Science, and Frontier Research Base for Global Young Researchers, Graduate School of Engineering, Osaka University, 2-1 Yamada-oka, Suita, Osaka, Japan

Received May 27, 2009; E-mail: npance@wakate.frc.eng.osaka-u.ac.jp

Abstract: The crystal structures of the pure, unsubstituted firefly emitter oxyluciferin ($OxyLH_2$) and its 5-methyl analogue ($MOxyLH_2$) were determined for the first time to reveal that both molecules exist as pure *trans*-enol forms, enol-$OxyLH_2$ and enol-$MOxyLH_2$, assembled as head-to-tail hydrogen-bonded dimers. Their steady-state absorption and emission spectra (in solution and in the solid state) and nanosecond time-resolved fluorescence decays (in solution) were recorded and assigned to the six possible *trans* chemical forms of the emitter and its anions. The spectra of the pure emitter were compared to its bioluminescence and fluorescence spectra when it is complexed with luciferase from the Japanese firefly (*Luciola cruciata*) and interpreted in terms of the intermolecular interactions based on the structure of the emitter in the luciferase active site. The wavelengths of the emission spectral maxima of the six chemical forms of $OxyLH_2$ are generally in good agreement with the theoretically predicted energies of the S_0-S_1 transitions and range from the blue to the red regions, while the respective absorption maxima range from the ultraviolet to the green regions. It was confirmed that both neutral forms, phenol-enol and phenol-keto, are blue emitters, whereas the phenolate-enol form is yellow-green emitter. The phenol-enolate form, which probably only exists as a mixture with other species, and the phenolate-enolate dianion are yellow or orange emitters with close position of their emission bands. The phenolate-keto form always emits in the red region. The concentration ratio of the different chemical species in solutions of $OxyLH_2$ is determined by several factors which affect the intricate triple chemical equilibrium, most notably the pH, solvent polarity, hydrogen bonding, presence of additional ions, and $\pi-\pi$ stacking. Due to the stabilization of the enol group of the 4-hydroxythiazole ring by hydrogen bonding to the proximate adenosine monophosphate, which according to the density functional calculations is similar to that due to the dimerization of two enol molecules observed in the crystal, the phenolate ion of the enol tautomer, which is the predominant ground-state species within the narrow pH interval 7.44−8.14 in buffered aqueous solutions, is the most probable emitter of the yellow-green bioluminescence common for most wild-type luciferases. This conclusion is supported by the bioluminescence/fluorescence spectra and the NMR data, as well the crystal structures of $OxyLH_2$ and $MOxyLH_2$, where the conjugated acid (phenol) of the emitter exists as pure enol tautomer.

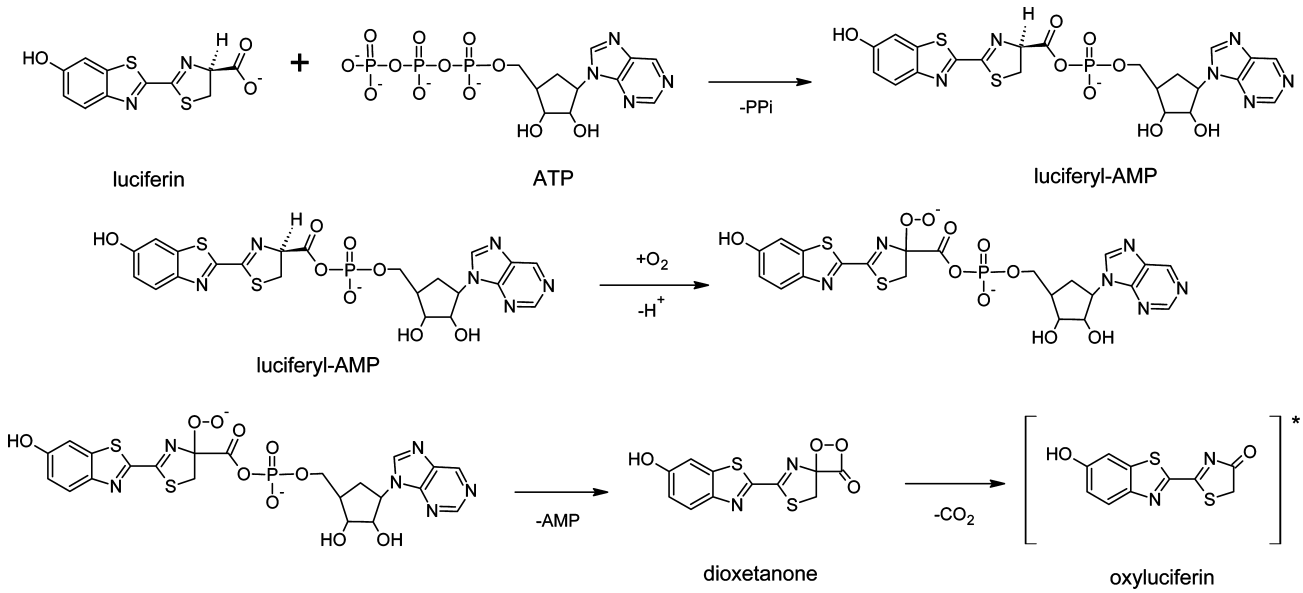
Introduction

Bioluminescence, the phenomenon of biological generation of visible light by deexcitation of chemically produced excited states, exhibited by a number of organisms including certain species of bacteria, insects, jellyfish, mushrooms, worms, and squid,[1,2] has been attracting the attention of researchers for decades because of its appealing visual effect. Recently, the interest in bioluminescence and chemiluminescence was revived due to the increasing applications in biochemistry which include rapid analysis of bacterial contamination,[3] multiple gene regulation and expression,[4] and *in vivo* bioluminescence imaging.[5] A critical limitation in utilization of the chemiluminescent reactions for analytical applications is their low quantum yield, which for most purposefully designed chemiluminescent reactions is usually less than 3−5% and rarely surmounts 10%. Relative to the chemiluminescence, the bioluminescence reactions are appreciably more efficient in terms of the overall yield.[6] In practice, the highest conversion of the chemical energy to light appears in the bioluminescence of fireflies, catalyzed by the synthase luciferase, with the recently redetermined value for the quantum yield of 41.0 ± 7.4%.[7] The primary structures of about 20 different beetle luciferases determined thus far[8] are consistent with the same suggested four-step reaction mechanism

(1) Johnson, F. H.; Shimomura, O. *Photophysiology* **1972**, *7*, 275.
(2) McCapra, F. *Acc. Chem. Res.* **1976**, *9*, 201.
(3) Griffiths, M. W. *Food Technol.* **1996**, *50*, 62.
(4) (a) Kricka, L. *Anal. Chem.* **1995**, *67*, 499R. (b) Price, R. L.; Squirrell, D. J.; Murphy, M. J. *J. Clin. Ligand Assay* **1998**, *21*, 349. (c) Kricka, L. *Methods Enzymol.* **2000**, *305*, 333. (d) Ohkuma, H.; Abe, K.; Kosaka, Y.; Maeda, M. *Luminescence* **2000**, *15*, 21.
(5) (a) Greer, L. F.; Szalay, A. A. *Luminescence* **2002**, *17*, 43. (b) Shinde, R.; Perkins, J.; Contag, C. H. *Biochemistry* **2006**, *45*, 11103. (c) Contag, C. H.; Bachmann, M. H. *Annu. Rev. Biomed. Eng.* **2002**, *4*, 235.
(6) The overall quantum yield of the bioluminescence based on the amount of the substrate, ϕ has been defined in the past as product of three variables, $\phi = \phi_C\phi_e\phi_{fl}$, where ϕ_C, ϕ_e and ϕ_{fl} are the chemical yield of the reaction, the proportion of the product molecules which are produced as excited, and the fluorescence yield, respectively.

10.1021/ja904309q CCC: $40.75

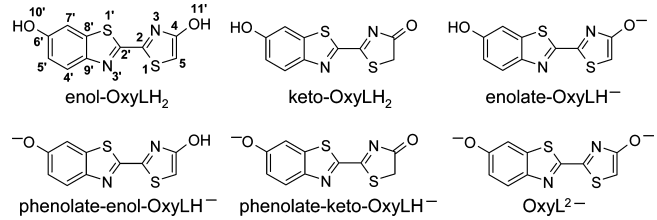
Chart 1. Suggested Reaction Mechanism of the Luciferase-Catalyzed Firefly Bioluminescence[a]



[a] The visible photon is emitted from the excited product oxyluciferin, presented here in its neutral keto form.

outlined in Chart 1.[2,9,10] The heterocyclic acid D(−)-luciferin (LH_2) appears as substrate in all cases, which by chirality-specific[11] adenylation with ATP is converted at the luciferase active site to luciferyl-adenosine monophosphate (luciferyl-AMP). This intermediate is subsequently oxidized with molecular oxygen and after removal of AMP it gives substituted dioxetanone intermediate.[12−16] The dioxetanone loses CO_2, generating the emitter molecule oxyluciferin ($OxyLH_2$) in its first single excited state ($OxyLH_2$*), which subsequently decays by emission of a visible photon.[17] The reactant LH_2 is partially enzymatically regenerated from the product $OxyLH_2$.[18,19]

Scheme 1. Possible *trans*−C2−C2′ Chemical Forms of the Firefly Emitter Oxyluciferin, $OxyLH_2$



The structure of the pure reactant, D(−)-LH_2, was determined about 30 years ago,[20−22] but due to marked instability and associated impediments with purification and crystallization,[23] very little is known about the structure of the real light-emitting molecule, $OxyLH_2$, and its crystal structure has not been determined yet. If only the more stable *trans* conformations around the C2−C2′ bridge are considered, by means of triple chemical equilibrium (deprotonation of the two hydroxyl groups and keto−enol tautomerism of the 4-hydroxythiazole ring) $OxyLH_2$ can exist as six chemical forms, as shown in Scheme 1. Except for several absorption/emission maxima, no detailed systematic study has been reported of the solvent/pH effects on the absorption and emission spectra of $OxyLH_2$. In fact, the exact chemical form in which $OxyLH_2$ exists at various conditions during the chemiluminescence and bioluminescence reactions, as well as the structures, absorption and emission spectra of its forms outlined in Scheme 1 have long remained a subject of speculations. The ultrafast dynamics of the first singlet excited state (S1) of the emitter and its ions as free species or complexed with luciferase have not been investigated in detail either, although the spectroscopy of the reactant LH_2 was reported.[24−26] Actually, it was argued[27] that due to rapid

(7) The earlier reported value of 88 ± 25%: (a) Seliger, H. H.; McElroy, W. D. *Arch. Biochem. Biophys.* **1960**, *88*, 136. (b) Selinger, H. H.; McElroy, W. D. *Biochem. Biophys. Res. Commun.* **1959**, *1*, 21. was recently re-determined and corrected to 41.0 ± 7.4%. (c) Ando, Y.; Niwa, K.; Yamada, N.; Enomoto, T.; Irie, T.; Kubota, H.; Ohmiya, Y.; Akiyama, H. *Nat. Photonics* **2007**, *2*, 44.
(8) Ugarova, N. N.; Brovko, L. Y. *Luminescence* **2002**, *17*, 321.
(9) White, E. H.; Rapaport, E.; Seliger, H. H.; Hopkins, T. A. *Bioorg. Chem.* **1971**, *1*, 92.
(10) Suzuki, N.; Goto, T. *Agr. Biol. Chem.* **1972**, *36*, 2213.
(11) Nakamura, M.; Maki, S.; Amano, Y.; Ohkita, Y.; Niwa, K.; Hirano, T.; Ohmiya, Y.; Niwa, H. *Biochem. Biophys. Res. Commun.* **2005**, *331*, 471.
(12) Rauhut, M. M. *Acc. Chem. Res.* **1969**, *2*, 80.
(13) McCapra, F. *J. Chem. Soc. Chem. Commun.* **1977**, 946.
(14) McCapra, F.; Burford, A. *J. Chem. Soc. Chem. Commun.* **1977**, 874.
(15) McCapra, F. *Chem. Commun.* **1968**, 155.
(16) Plant, P. J.; White, E. H.; McElroy, W. D. *Biochem. Biophys. Res. Commun.* **1968**, *31*, 98.
(17) In the earlier literature: (a) Koo, J.-Y.; Schmidt, S. P.; Schuster, G. B. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 30. (b) Koo, J.-y.; Schuster, G. B. *J. Am. Chem. Soc.* **1977**, *99*, 6107. a "chemically initiated electron exchange luminescence (CIEEL)" mechanism has been invoked to explain the emission of light from the dioxetanone, according to which intramolecular electron transfer occurs from the polycyclic heterocyclic portion of the molecule to the dioxetanone part, whereby an intermediate diradical anion is creation and decomposed. This mechanism, however, has been discredited in the more recent reports based on quantum yield considerations. We thank the anonymous reviewer for bringing this to our attention.
(18) Okada, K.; Iio, H.; Kubota, I.; Goto, T. *Tetrahedron Lett.* **1974**, *32*, 2771.
(19) Gomi, K.; Kajiyama, N. *J. Biol. Chem.* **2001**, *276*, 36508.
(20) Dennis, D.; Stanford, R. H. *Acta Crystallogr.* **1973**, *B29*, 1053.
(21) Blank, G. E.; Pletcher, J.; Sax, M. *Acta Crystallogr.* **1974**, *B30*, 2525.
(22) Blank, G. E.; Pletcher, J.; Sax, M. *Biochem. Biophys. Res. Commun.* **1971**, *42*, 583.
(23) Suzuki, N.; Sato, M.; Okada, K.; Goto, T. *Tetrahedron* **1972**, *28*, 4065.
(24) Wada, N.; Shibata, R. *J. Phys. Soc. Jpn.* **1997**, *66*, 3312.

condensation reactions[28] the real luciferase emitter $OxyLH_2$ has probably never been isolated in pure form. Although the structure of the main decay products was tentatively assigned based on physicochemical data,[23] there is no precise information on the actual rates of decomposition and its effects on the UV−visible spectra of the emitter. Except for one report of nanosecond fluorescence measurements in water and ethanol,[29] where the blue, green and red emission of the emitter were assigned to the phenol, phenolate and keto-phenolate forms, respectively, the earlier spectroscopic results pertain mainly to the more stable 5-methyl ($MOxyLH_2$) and 5,5-dimethyl (DMOxyLH) analogues,[23,27,30,31] and they usually refer to single solvent and/or concentration. While this paper was in preparation, a thorough spectroscopic study on the 5,5-dimethyl analogue DMOxyLH was reported by Hirano et al.,[32] which provides a good basis to assign the spectra of the real, unsubstituted emitter, $OxyLH_2$. It should be noted, however, that both 5-methyl derivatives, $MOxyLH_2$ and DMOxyLH, bind more weakly to the luciferase relative to the unsubstituted emitter,[31] and therefore the methyl groups impose significant steric hindrance for binding of the emitter to the active site of the real protein system. Moreover, the 4-hydroxythiazole oxygen atom of DMOxyLH is fixed by the double substitution to its keto form, so that this model molecule can not account for the effects in the emission spectra that may be due to the keto−enol tautomerization. Therefore, although by being more stable than $OxyLH_2$, $MOxyLH_2$, and DMOxyLH are very convenient and useful models for the firefly bioluminescence, these molecules are structurally different from the real emitter and they can provide only indirect information on the bioluminescence reaction.

A second issue that has been thoroughly debated in the related literature is the origin of the difference in color of firefly bioluminescence. Namely, although the emitter molecule is identical in all cases, the *in vivo* emitted color depends on the species[33] and ranges from green to red:[34] firefly (Lampyridae) luciferases produce yellow-green light, click beetles (Elateridae) produce green to orange light, while railroad worms (Phengodidae) emit green to red light.[8] The color change can also be induced artificially, by introducing point mutations proximate to the active site in the luciferase.[35−39] Although several mechanisms have been suggested to explain the color difference, no consensus has been reached yet of as to the exact origin of the phenomenon. The structure of uncomplexed luciferase from the North-American firefly *Photinus pyralis*, determined at 2.0 Å resolution[40] and combined with the results on the solution-state reaction kinetics of luciferase mutants, provided indirect evidence for setting working models[37] for the binding of the substrate (LH_2) to the enzyme, which indicated that polarity and rigidity of the emitter binding site may be responsible for the yellow-green emission. The recent structure determination by Nakatsu et al.[41] and comparison of the wild-type (WT) luciferase from the Japanese firefly *Luciola cruciata* complexed with $OxyLH_2$ and AMP as reaction products, with WT and red mutant *L. cruciata* luciferases complexed with a stable model molecule for the reaction intermediate luciferyl-AMP indicated that changes in the interaction of $OxyLH_2$* with the protein environment, such as those caused by conformational change of the hydrophobic chain of the proximate isoleucine residue, affect the emitted color. These results have clearly unraveled the exact, three-dimensional structure of the binding site, as well as the position and orientation of the ligand $OxyLH_2$ in the active pocket. In combination with the spectra of the emitter in model solutions and as complex with the enzyme, they now provide good prospect for more quantitative consideration of the environment effects on the bioluminescence color. However, the conditions under which each of the chemical forms in Scheme 1 can exist, as well as their spectroscopic properties, have yet to be clarified.

We report herein the first crystal structure determination of the pure firefly emitter, $OxyLH_2$ and its 5-methyl substituted derivative, $MOxyLH_2$. To our surprise, and contrary to our and other authors' quantum mechanical theoretical calculations (see below) which have regularly favored the respective keto form keto-$OxyLH_2$ as the more stable tautomer, the structure determination clearly demonstrated that in the solid state both compounds exist as pure enol forms, enol-$OxyLH_2$ and enol-$MOxyLH_2$. Prompted by this apparent inconsistency between the experimental and theoretical results, and aiming to downsize to the atomic level the structural factors determinant to the structure of $OxyLH_2$ and the emitted color, we studied in detail the steady-state absorption/emission spectra and nanosecond time-resolved fluorescence decays of oxyluciferin. In particular, we studied the effects on the electronic and structural properties of $OxyLH_2$ of solvent, pH, methyl substitution, single/double deprotonation, presence of base and oxygen in much greater detail than it has been done before, and we compared the results to the laser-induced fluorescence and chemically produced bioluminescence of the emitter in the scaffold of the *L. cruciata* luciferase. By working with the unsubstituted emitter, we intended to unravel the spectroscopic manifestation of the structural changes that accompany the triple chemical equilibrium of the emitter, without the interferences from the methyl substitution. From the new spectroscopic and structural data, by considering the environment effects on pure $OxyLH_2$ based

(25) Wada, N.; Mitsuta, K.; Kohno, M.; Suzuki, N. *J. Phys. Soc. Jpn.* **1989**, *58*, 3501.
(26) Brovko, L. Y.; Cherednikova, E. Y.; Chikishev, A. Y.; Chudinova, E. A. *Moscow Univ. Trans. Phys. Astron. Ser.* **1998**, *3*, 26.
(27) White, E. H.; Roswell, D. F. *Photochem. Photobiol.* **1991**, *53*, 131.
(28) Jensen, K. A.; Crossland, I. *Acta Chem. Scand.* **1963**, *17*, 144.
(29) Gandelman, O. A.; Brovko, L. Y.; Ugarova, N. N.; Chikishev, A. Y.; Shkurimov, A. P. *J. Photochem. Photobiol.* **1993**, *B19*, 187.
(30) Leont'eva, O. V.; Vlasova, T. N.; Ugarova, N. N. *Biochemistry (Moscow)* **2006**, *71*, 51.
(31) Vlasova, T. N.; Leontieva, O. V.; Ugarova, N. N. *Biochemistry (Moscow)* **2006**, *71*, 555.
(32) Hirano, T.; Hasumi, Y.; Ohtsuka, K.; Maki, S.; Niwa, H.; Yamaji, M.; Hashizume, D. *J. Am. Chem. Soc.* **2009**, *131*, 2385.
(33) (a) Lloyd, J. E. In *Bioluminescence in Action*; Herring, P. J., Ed.; Academic Press: New York, 1978; pp 241−272. (b) Hastings, J. W. In *Cell Physiology Source Book*; Sperelakis, N., Ed.; Academic Press: New York, 1995; pp 665−681.
(34) (a) Wood, K. V.; Lam, Y. A.; Seliger, H. H.; McElroy, W. D. *Science* **1989**, *244*, 700. (b) Wood, K. V. *Photochem. Photobiol.* **1995**, *62*, 662.
(35) Branchini, B. R.; Magyar, R. A.; Murtiashaw, M. H.; Andreson, S. M.; Helgerson, L. C.; Zimmer, M. *Biochemistry* **1999**, *38*, 13223.
(36) (a) 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. (b) Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Magyar, R. A.; Gonzalez, S. A.; Ruggiero, M. C.; Stroh, J. G. *Biochemistry* **2004**, *43*, 7255.
(37) Branchini, B. R.; Magyar, R. A.; Murtiashaw, M. H.; Portier, N. C. *Biochemistry* **2001**, *40*, 2410.
(38) Tafreshi, N.Kh.; Hosseinkhani, S.; Sadeghizadeh, M.; Sadeghi, M.; Ranjbar, B.; Naderi-Manesh, H. *J. Biol. Chem.* **2007**, *282*, 8641.
(39) Ugarova, N. N.; Maloshenok, L. G.; Uporov, I. V.; Koksharov, M. I. *Biochemistry (Moscow)* **2005**, *70*, 1262.
(40) Conti, E.; Franks, N. P.; Bricks, P. *Structure* **1996**, *4*, 287.
(41) Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A.; Kobashi, N.; Sakata, K.; Kato, H. *Nature* **2006**, *440*, 372.

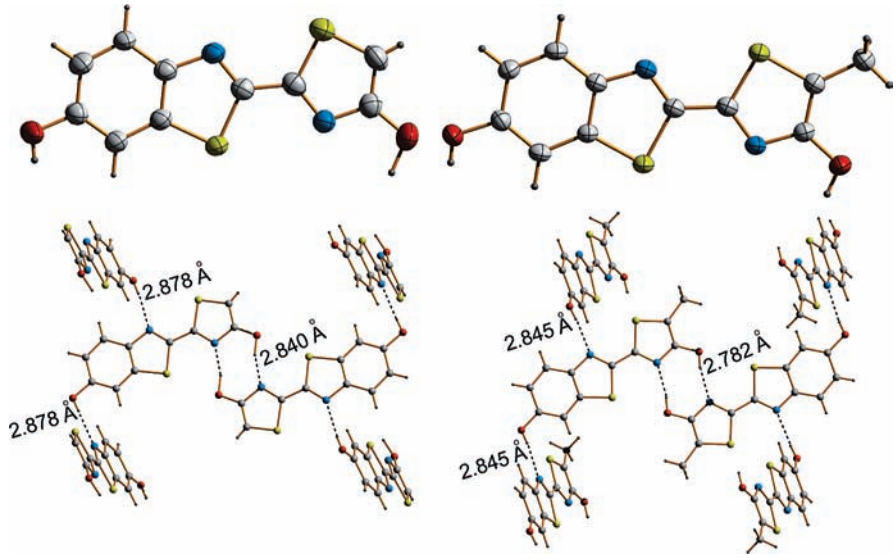


Figure 1. Molecular (top, ORTEP) and packing (bottom, stick-and-plot) structures of oxyluciferin ($OxyLH_2$, left) and 5-methyloxyluciferin ($MOxyLH_2$, right) in the crystals. The values in the plots on the bottom correspond to the N···O distances. Selected intramolecular distances (the values are listed in Ångströms; the labels refer to the top left diagram in Scheme 1. The complete list of the structure parameters can be retrieved from the structure data in CIF format deposited as Supporting Information): **$OxyLH_2$**, S1′−C2′, 1.732(3), S1′−C8′, 1.734(3), S1−C5, 1.688(5), S1−C2, 1.705(3), O10′−C6′, 1.358(4), N3′−C2′, 1.291(4), N3′−C9′, 1.408(4), O11′−C4, 1.339(4), N3−C2, 1.328(4), N3−C4, 1.361(4); **$MOxyLH_2$**, S1−C2, 1.707(2), S1−C5, 1.719(2), S1′−C8′, 1.727(2), S1′−C2′, 1.740(2), O10′−C6′, 1.361(2), N3−C2, 1.314(2), N3−C4, 1.361(3), O11′−C4, 1.341(2), N3′−C2′, 1.302(2), N3′−C9′, 1.391(2).

on the structure of its complex with the enzyme, we obtained new insights on the mechanism of color variation of firefly bioluminescence.

Results and Discussion

Crystal Structures and Structural Comparison with the Complexed Emitter. Except for the structure of the reactant D(−)-LH_2[20−22] and the recently reported DMOxyLH,[32] no other crystal structures of molecules including the 2-(benzothiazol-2′-yl)thiazole moiety have been determined thus far. $OxyLH_2$ and $MOxyLH_2$ were synthesized by condensation reactions[23] and obtained in pure form (in the case of $OxyLH_2$, for the first time) by recrystallization by slow evaporation over $CaCl_2$ under nitrogen. Due to the pronounced instability of $OxyLH_2$ with prolonged storage in solution, repeated attempts at its purification including evaporation, diffusion and seeding methods, proved an extremely difficult task, always returning a mixture of an amorphous-like yellow phase as major component in addition to red microcrystalline powder of $OxyLH_2$. After repeated trials at various conditions, by evaporating isopropanol solutions of the as-synthesized product under nitrogen, sufficient amount of $OxyLH_2$ was obtained in form of bundled, red needle-like single crystals, which were fairly stable in air at room temperature. The physicochemical characterization (see the Experimental Section) confirmed that the compound was obtained as solventless product[42] of 100% purity. Similar colorless prismatic single crystals of $MOxyLH_2$ were prepared by slow evaporation of its solution in mixture of methanol (MeOH) and dichloromethane under nitrogen. Single crystals of both compounds were used for X-ray diffraction analysis[43] (complete crystallographic data are deposited as Table S1 in the Supporting Information).

The structure determination of $OxyLH_2$ and $MOxyLH_2$ at room temperature unraveled that in the crystals both molecules exist as pure *trans*-enol forms, enol-$OxyLH_2$ and enol-$MOxyLH_2$, assembled as head-to-tail hydrogen-bonded dimers (Figure 1). No residual electron density was observed that could indicate tautomeric equilibrium in the solid state. The thiazole nitrogen atom N3 of each molecule is hydrogen bonded to the 4-hydroxyl group O11′H of its centrosymmetric counterpart, with d(O11′−H···N3) = 2.840(4) Å (∠(O11′−H···N3) = 162(5)°) and 2.782(2) Å (169(3)°) for $OxyLH_2$ and $MOxyLH_2$, respectively. Furthermore, the benzothiazole nitrogen N3′ and the C6′-hydroxyl group of the benzothiazole moiety are each involved in additional intermolecular hydrogen bonds to the hydroxyl group and the benzothiazole nitrogen of two neighbor molecules, d(N3′···O10′) = 2.878(4) Å (∠(N3′···H−O10′) = 174(5)°) and 2.845(2) Å (174(3)°) for $OxyLH_2$ and $MOxyLH_2$, respectively. The existence of the enol forms in the crystals is contradictory to the theoretical calculations,[44−48] which practically without exception predict the keto form as the more stable form of the two neutral tautomers. Along with the other theoretical predictions, our single-point calculations on the optimized structures (B3LYP/6-31G(d)) within the density

(42) $OxyLH_2$ and $MOxyLH_2$ were prepared according to a literature method (ref 23). Unlike the original procedure, which reported $OxyLH_2$ as hemihydrate, after recrystallization we obtained the compound as solventless product, which was confirmed by the excellent agreement of the analytical data. We presume that the earlier reported formula $OxyLH_2 \cdot 0.5H_2O$ is a result of partial decomposition of the product to dioxyluciferin, an approximate 1:1 mixture of which with $OxyLH_2$ would appear in the elemental analysis results as 0.5 water molecules per formula unit.

(43) Basic crystallographic data: **$OxyLH_2$**, $C_{10}H_6N_2O_2S_2$, monoclinic, $P2_1/n$, a = 4.8806(4), b = 18.8242(14), c = 11.1298(8) Å, β = 99.260(5)°, V = 1009.21(13) Å^3, Z = 4, Θ_{max} = 26.4°, GoF = 1.037, $R_1[I > 2\sigma(I)]$ = 0.0524, wR_2 = 0.1239. **$MOxyLH_2$**, $C_{11}H_8N_2O_2S_2$, monoclinic, $P2_1/n$, a = 6.3027(4), b = 14.8046(10), c = 11.6515(7) Å, β = 91.419(4)°, V = 1086.86(12) Å^3, Z = 4, Θ_{max} = 27.5°, GoF = 1.069, $R_1[I > 2\sigma(I)]$ = 0.0337, wR_2 = 0.0875.

(44) Nakatani, N.; Hasegawa, J.; Nakatsuji, H. *J. Am. Chem. Soc.* **2007**, *129*, 8756.
(45) Dahlke, E. E.; Cramer, C. J. *J. Phys. Org. Chem.* **2003**, *16*, 336.
(46) Li, Z.; Ren, A.; Guo, J.; Yang, T.; Goddard, J. D.; Feng, J. *J. Phys. Chem.* **2008**, *112*, 9796.
(47) Ren, A.; Guo, J.; Feng, J.; Zou, L.; Li, Z.; Goddard, J. D. *Chin. J. Chem.* **2008**, *26*, 55.
(48) Yang, T.; Goddard, J. D. *J. Phys. Chem.* **2007**, *111*, 4489.

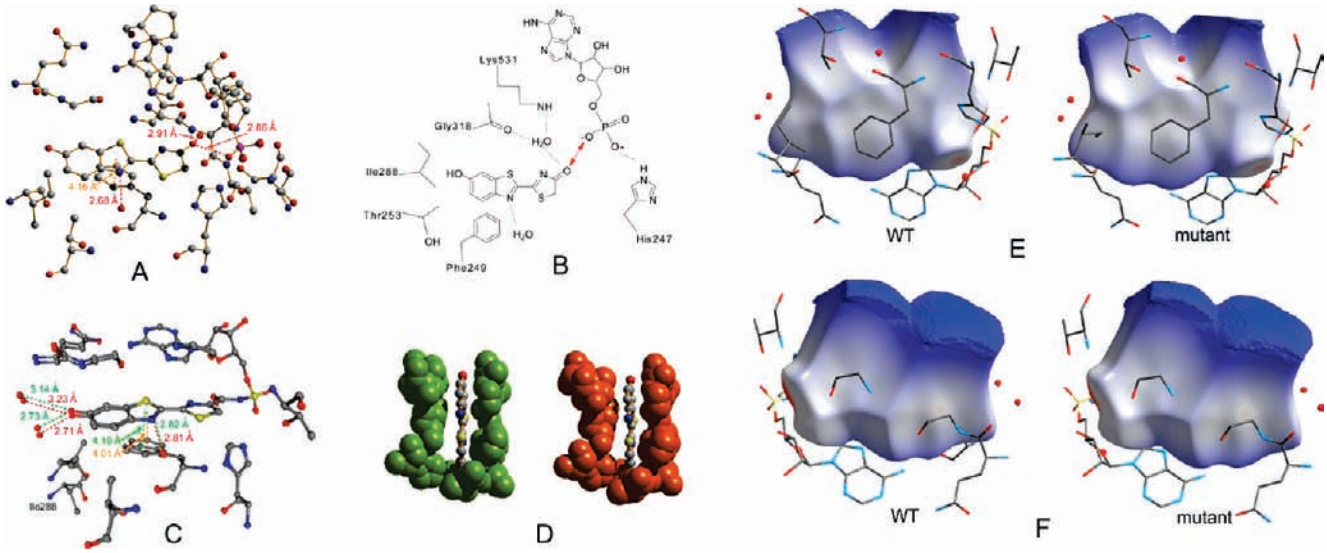


Figure 2. Structure of the emitter oxyluciferin ($OxyLH_2$) and a model compound (DLSA) for the intermediate state of the bioluminescence reaction in luciferase from *L. cruciata* according to ref[41] (the coordinates were retrieved from the Protein Data Bank (PDB), with the following codes: 2D1R, 2D1S and 2D1T; The hydrogen atoms are omitted for clarity.) (A) Three-dimensional structure of the nearest surrounding of $OxyLH_2$, assigned to the neutral keto form keto-$OxyLH_2$. The hydrogen bonds and the possible $\pi-\pi$ stacking interaction are shown in red and orange colors, respectively. (B) Schematic representation of the environment shown in panel A. Note the short distance between the oxygen of $OxyLH_2$ and the oxygen atom of AMP (double-headed arrow), which could stabilize the enol group of the 4-hydroxythiazole ring. (C) Overlapped structures of DLSA in the green (green-colored bonds) and red (red-colored bonds) bioluminescence. The possible $\pi-\pi$ stacking interaction is shown in light green and orange color. (D) Space-filling plot of the immediate surrounding of the model DLSA in green (left) and red (right) emitting luciferase showing the possible mechanism of color tuning according to Nakatsu et al.[41] (E) Top and (F) bottom views of the Hirshfeld surfaces[67] (plotted as d_{norm}) of the ligand portion of DLSA in the structures of the WT and mutant luciferase from *L. cruciata* complexes with the intermediate.

functional theory (DFT) indicate that for isolated molecules, the resonance stabilizes the keto form (keto-$OxyLH_2$) better than the enol form (enol-$OxyLH_2$), the energy difference with the respective enol forms ΔE[(enol-$OxyLH_2$)-(keto-$OxyLH_2$)] being 6.57 kcal mol^{-1} ($OxyLH_2$) and 3.99 kcal mol^{-1} ($MOxyLH_2$). The stabilization of the enol form in the crystals can be explained by considering the energy gain provided by the dimerization: the calculated stabilization energy per monomer unit of a model dimer $(OxyLH_2)_2$ which has been constrained to planarity at the central eight-member ring, ΔE[(enol-$OxyLH_2$)-(keto-$OxyLH_2$)] = −6.83/2 kcal mol^{-1}, confirms that the enol form is stabilized in the dimer as a result of the hydrogen bonding between the two enol-$OxyLH_2$ (or enol-$MOxyLH_2$) molecules, which is not possible with the keto-$OxyLH_2$ form.

A second unexpected feature of the structure of $OxyLH_2$ is the slight bending of the molecule (∼6°) at C2′ perpendicular to the bridge between the two heterocyclic rings, which turns the dimers away from coplanarity. The molecular distortion of enol-$MOxyLH_2$ is about half of that of enol-$OxyLH_2$, and is midway between bending and twisting. Similarly to these structures, the molecule of DMOxyLH, which is constrained to its keto form, is also bent, although the phenyl-methyl ether analogue of DMOxyLH is planar.[32] The bending of the molecule of $OxyLH_2$ is not immediately expected from the extended electron conjugation and the theoretically optimized structure, and it might have role in determination of the absorption and emission colors (note that while the crystals of $MOxyLH_2$ are pale yellow, nearly colorless, the crystals of $OxyLH_2$ are red, due to the absorption shoulder at ∼546 nm, Figure S1, Supporting Information).

Detailed comparison of the crystal structure of enol-$OxyLH_2$ in the crystal (Figure 1) with keto-$OxyLH_2$ embedded in the scaffold of *L. cruciata*[41] (panels A and B in Figure 2) reveals important analogy in their hydrogen bonding potential. First, the intradimer N···O distance in the crystal of enol-$OxyLH_2$ (2.840 Å) is very close to that between keto-$OxyLH_2$ in the luciferase and the phosphate group of the product AMP (2.86 Å), indicating that intermolecular interaction similar to the interdimer hydrogen bonds observed in the crystal of pure oxyluciferin could stabilize the enol functionality in the luciferase pocket, especially if the negative charge on the phosphate ion of AMP is considered. Indeed, constraining the distance between $OxyLH_2$ and oxygen atom of a model phosphate ion to the experimental value (2.86 Å) results in theoretical stabilization energy ΔE[(enol-$OxyLH_2$) − (keto-$OxyLH_2$)] = −3.24 kcal mol^{-1} (B3LYP/6-31G(d)) of enol-$OxyLH_2$ over keto-$OxyLH_2$. This stabilization is very similar with the hydrogen bonding stabilization per monomer unit in the $(enol\text{-}OxyLH_2)_2$ dimer, −6.83/2 kcal mol^{-1}. Second, in both cases the benzothiazole atom N3′ participates in hydrogen bonding—to hydroxyl group in the crystal of pure enol-$OxyLH_2$ and to water molecule in the enzyme. Third, the phenol group of pure enol-$OxyLH_2$ is hydrogen bonded to nitrogen atom (2.878 Å, Figure 1). In the structure of the complex of *L. cruciata* luciferase with keto-$OxyLH_2$ (Figure 2B), the hydroxyl group appears free (unbound). However, in the structure of the intermediate with both WT and mutant luciferases (Figure 2C),[41] the phenol group is bifurcated to two water molecules, assigned to density peaks at 2.73 and 3.14 Å from the oxygen atom. This may be an indication that during the bioluminescence reaction these positions are occupied by one or two dynamic water molecules, which can bind reversibly to the phenol group. Finally, in the protein, the thiazole half of the benzothiazole ring approaches, in a displaced face-to-face arrangement, phenyl ring from Phe249 at 4.16 Å (Figure 2A) with which it could $\pi-\pi$ interact at some step during the bioluminescence reaction. Similarly, in the crystal of the pure compound, the adjacent dimers are offset to each other, so that there are two weak $\pi-\pi$ contacts, at 3.972 Å (thiazole-benzothiazole) and 3.942 Å (benzothiazole-benzothiazole; not shown in Figure 1). These

four important similarities, which include $\pi-\pi$ stacking in addition to hydrogen bonding, favor the structure of pure enol-$OxyLH_2$ as viable small-molecule model for some of the relevant interactions of the ground-state emitter in the enzyme scaffold. It should be noted, however, that similar to what was observed for other emitters,[49] the stronger intermolecular interactions and tighter packing in the enol-$OxyLH_2$ crystal results in complete quenching of the fluorescent emission in the solid state (Figure S1 in the Supporting Information), and thus direct comparison of the emission spectra is not possible.

Steady-State UV−Visible Spectra of the Chemical Forms of Oxyluciferin. Scattered, partial and occasionally contrasting results exist in the literature on the UV−visible spectra of pure $OxyLH_2$. The major spectral data of oxyluciferin-related compounds published before 1991 were compiled by White and Roswell, and consist of band maxima measured at different concentration/solvent combinations.[27] In the meantime, additional results, including high-level theoretical spectroscopic simulations, have become available, and the spectra of DMOxyLH were published recently,[32] which provide a basis for in-depth spectral analysis.

Generally, the spectral assignment of $OxyLH_2$ is burdened by the triple chemical equilibrium (two ionizable hydroxyl groups and keto−enol tautomerism of the 4-hydroxythiazole moiety), resulting in the *trans* chemical forms in Scheme 1 and their *cis* counterparts. It is further complicated by the possible coexistence of several forms within a narrow p*K* range and ambiguities due to overlaps caused by additional spectral shifts from the specific and nonspecific interactions with the environment. Except for the neutral enol form (enol-$OxyLH_2$) which exists as nearly pure form in aprotic solvents (e.g., dimethyl sulfoxide, DMSO),[27,50] at physiologically relevant pH in aqueous solutions $OxyLH_2$ probably exists as mixture of different molecular species. Additional difficulty to the spectral interpretation of $OxyLH_2$ is imposed by the chemical instability of the emitter, which becomes more pronounced in basic media. To arrive at firmer evidence about the chemical nature and the spectra of the emitter in solution, by working under anaerobic conditions and strictly within the time-window of full chemical stability of the analyte, we succeeded in recording reproducible spectra of $OxyLH_2$ without major difficulties. Actually, our measurements showed that in absence of base, $OxyLH_2$ is sufficiently stable in argon-purged solvents even under aerobic conditions. Although the stability was decreased in presence of base, and the spectra changed on a time scale of minutes (see below), this was still sufficiently long period to record reproducible steady-state and time-resolved spectra of deprotonated $OxyLH_2$. As it has been already noted,[27] $MOxyLH_2$ was more stable relative to $OxyLH_2$ under all conditions.

The steady-state absorption and emission spectra of $OxyLH_2$ and $MOxyLH_2$ in DMSO with and without added *t*-BuOK are shown in Figure 3 (panels A and B; *t*-BuOK was selected as nonhydroxylic base in order to exclude formation of additional hydrogen bonds between the nitrogen atoms of the emitter with hydroxyl groups, and also to assess the effect of such interactions by comparison with Bu_4NOH). Along with the previous observations,[27,50] it was confirmed by ^1H NMR spectroscopy that in absence of base in DMSO-d_6, in DMSO both $OxyLH_2$ ($\lambda_{max}^{abs} = 377$ and $\lambda_{max}^{em} = 448$ nm) and $MOxyLH_2$ ($\lambda_{max}^{abs} = 386$ and $\lambda_{max}^{em} = 456$ nm) exist as practically pure neutral enol forms enol-$OxyLH_2$ and enol-$MOxyLH_2$. The 5-methyl substitution merely red-shifts the absorption and emission maxima of 9 and 8 nm, respectively. Both neutral enol molecules are blue emitters, and exhibit Stokes shifts of 70 and 71 nm, respectively. The values of the Stokes shifts indicate that in DMSO both molecules remain protonated in the first excited state, so that the respective blue emitter is assigned to enol-$OxyLH_2$*. The deprotonation has identical effect on the spectra of $OxyLH_2$ and $MOxyLH_2$ (Figure 3, panels A and B). Addition of *t*-BuOK results in strong red-shifts (484 and ∼575 nm for $OxyLH_2$, 504 nm and ∼590 nm for $MOxyLH_2$) whereupon a mixture of species is obtained, as evidenced by the shoulders accompanying the absorption bands and the complex emission profiles. Comparison of the emission from $OxyLH_2$ in presence of base excited at 500 nm and at 600 nm unraveled coexistence of *three* emitters, which are excited with different efficiency at the two wavelengths, and emit at ∼570 nm, 629 nm and ∼690 nm. A large excess of base, which produces the pure dianion $OxyL^{2-}$, results in small red shift of the absorption and *blue* shift of the emission maximum to 590 nm, showing that the S_1 dianion $(OxyL^{2-})$* emits at higher energy than the respective monoanion(s) $(OxyLH^-)$* and can be related to the overlapped band at 570 nm. The other two bands (629 nm and ∼690 nm) are assigned to the monoanions. As it will be detailed below, this assignment is supported by the analogous blue shift observed in the spectra recorded from aqueous solutions. The Stokes shifts of the dianions, 97 nm for $OxyL^{2-}$ and 88 nm for $MOxyL^{2-}$, are larger than the values of the respective neutral molecules, indicating larger charge separation in $(OxyL^{2-})$* and $(MOxyL^{2-})$*.

Solvent Effects and Assignment of the UV−Visible Spectra. The solvent effects on the spectrum of $OxyLH_2$ with and without added Bu_4NOH are shown in Figure 4, and the relevant spectral data are listed in Table 1. The absorption spectra of pure $OxyLH_2$ in $CHCl_3$, CH_2Cl_2, DMSO, acetone, MeOH and acetonitrile (ACN) are dominated by a single strong band which is weakly sensitive to the solvent, the respective maxima ranging between 367 nm (acetone) and 377 nm (DMSO) (Figure 4A). Relative to the organic solvents, the absorption maximum in the aqueous solution is strongly red-shifted ($\lambda_{max}^{abs} = 409$ nm) due to the singly deprotonated species $OxyLH^-$. An additional, weak band from the small amount of dianion $OxyL^{2-}$ appears at 525 nm. The spectra of $OxyLH_2$ in MeOH and in ACN exhibit long-wavelength tails >440 nm, apparently due to presence of additional species.

The solvent effects on $OxyLH_2$ (in absence of base) are better discernible in the emission spectra (Figure 4B), due to more pronounced oxyluciferin-solvent interactions in the S_1 state relative to the S_0 state. The emission bands can be divided into four groups according to their relative wavelengths: (acetone, DMSO) < (MeOH, ACN) < (CH_2Cl_2, $CHCl_3$) < water. The strong deviation from linear correlation in the Lippert-Mataga analysis[51,52] of the solvent effects on the spectrum of pure $OxyLH_2$ (Figure 5) confirms that specific interactions and,

(49) (a) Yoshida, K.; Uwada, K.; Kumaoka, H.; Bu, L.; Watanabe, S. *Chem. Lett.* **2001**, *30*, 808. (b) Dong., J.; Solntsev, K. M.; Tolbert, L. *J. Am. Chem. Soc.* **2009**, *131*, 662.

(50) Esteves da Silva, J. C. G.; Magalhaes, J. M. C. S.; Fontes, R. *Tetrahedron Lett.* **2001**, *42*, 8173.

(51) The solvent parameters for calculation of Δf were taken from: (a) Reichardt, C. *Solvents and solvent effects in organic chemistry*, 3rd ed.; Wiley-VCH: Weinheim, 2003. (b) Lippert, E. Z. *Electrochem.* **1957**, *61*, 962. (c) Mataga, N.; Kaifu, Y.; Koizumi, M. *Bull. Chem. Soc. Jpn.* **1956**, *29*, 465.

(52) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; Springer: Tokyo, 2006.

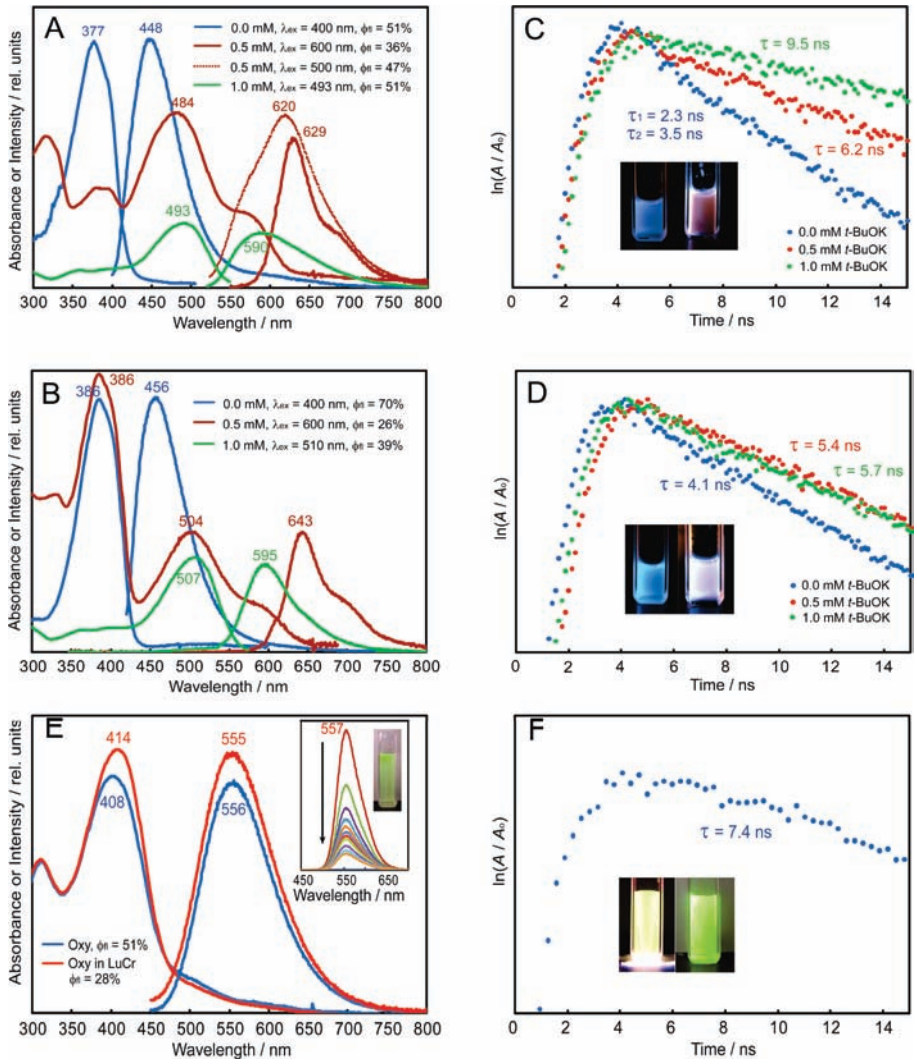


Figure 3. (A−D) Effects of deprotonation and fluorescence decay lifetimes on the absorption and emission spectra of solutions of oxyluciferin ($OxyLH_2$; A, C) and 5-methyloxyluciferin ($MOxyLH_2$; B, D) in DMSO. The conditions (c(*t*-BuOK), excitation wavelengths λ_{ex} and fluorescence yields ϕ_{fl}) are given in the legends; $c(OxyLH_2) = c(MOxyLH_2) = 5 \times 10^{-5}$ M. The insets in panels C, D and F show the fluorescence, induced by continuous-wave UV excitation, of deoxygenated DMSO and aqueous solutions before and after addition of *t*-BuOK. (E and F) Effects of complexation of $OxyLH_2$ and AMP with luciferase from *L. cruciata* on the absorption and emission spectra and the fluorescence decay ("Oxy" in panel E denotes oxyluciferin). Conditions: aqueous solution, c(AMP) = 1.6 mM, $c(MgSO_4)$ = 5 mM, c(luciferase) = 1 mg·mL^{-1}, $c(OxyLH_2)$ in the final solution was half of the saturated stock solution, λ_{ex} = 400 nm. (Inset in panel E) Time-resolved bioluminescence spectra of $OxyLH_2$ produced *in situ* in luciferase from *L. cruciata*. Conditions: aqueous solution, c(ATP) = 1.6 mM, $c(MgSO_4)$ = 5 mM, c(luciferase) = 1 mg·mL^{-1}, c(D(−)-LH_2) = 0.3 mM (the uppermost and lowest spectra correspond to time points of 0 and 20 min from the reaction start, respectively). In panels A, B and E, the values of λ_{max} (in nanometers) are shown above the respective bands.

particularly, hydrogen bonding, play crucial role in the structure of $OxyLH_2$*. Whereas in DMSO and in aprotic solvents (CH_2Cl_2, $CHCl_3$, acetone) $OxyLH_2$ exists as enol-$OxyLH_2$ (e.g., ^{1}H NMR spectra in DMSO-d_6: δ_{OH}(phenol) = 10.1 ppm, δ_{OH}(enol) = 11.0 ppm), upon excitation (λ_{ex} = 400 nm) in MeOH and ACN solutions it exhibits secondary, strongly red-shifted emission bands, at $\lambda_{max}^{em} \approx 552$ and 540 nm. The complex ^{1}H NMR spectrum in ACN-d_3 (Figure S2, Supporting Information) evidenced existence of more than one species. In MeOH-d_4 (Figure S3, Supporting Information), none of the hydroxyl protons could be observed due to fast proton exchange with the solvent. The secondary emission bands in these solvents are of much higher intensity than the respective secondary absorption bands. This is strong indication for occurrence of excited-state proton transfer from $(OxyLH_2)$*, so that the main emission $S_1 \rightarrow S_0$ originates from the ionic species, although their concentration before the excitation (S_0) was relatively smaller. In fact, it is reasonable to assume that the polar solvents ACN and MeOH form hydrogen bonded complexes $OxyLH_2 \cdots$ACN and $OxyLH_2 \cdots$MeOH with the emitter in the ground state, so that after the excited-state proton transfer has occurred, the actual emitting species are the respective ion pairs, $(OxyLH^- \cdots ACNH^+)$* and $(OxyLH^- \cdots MeOH_2^+)$*, although formation of larger hydrogen bonded complexes is possible as well. It should be noted that the secondary emission bands in these solvents are at close positions to the emission of $OxyLH_2$ in phosphate-buffered aqueous solution at pH 7.8, λ_{max}^{em} = 553 nm, and in excess of *t*-BuOK, λ_{max}^{em} = 540 nm. As it was shown by combined spectrophotometric-titration measurement (see below), at pH 7.8 $OxyLH_2$ exists mainly as the phenolate-enol-$OxyLH^-$ form (Scheme 1). This ion is present as a secondary species in MeOH due to partial dissociation of the phenolic hydroxyl group. In ACN, the secondary species can be assigned either to the phenolate-enol-$OxyLH^-$ ion or to the $OxyL^{2-}$ ion,

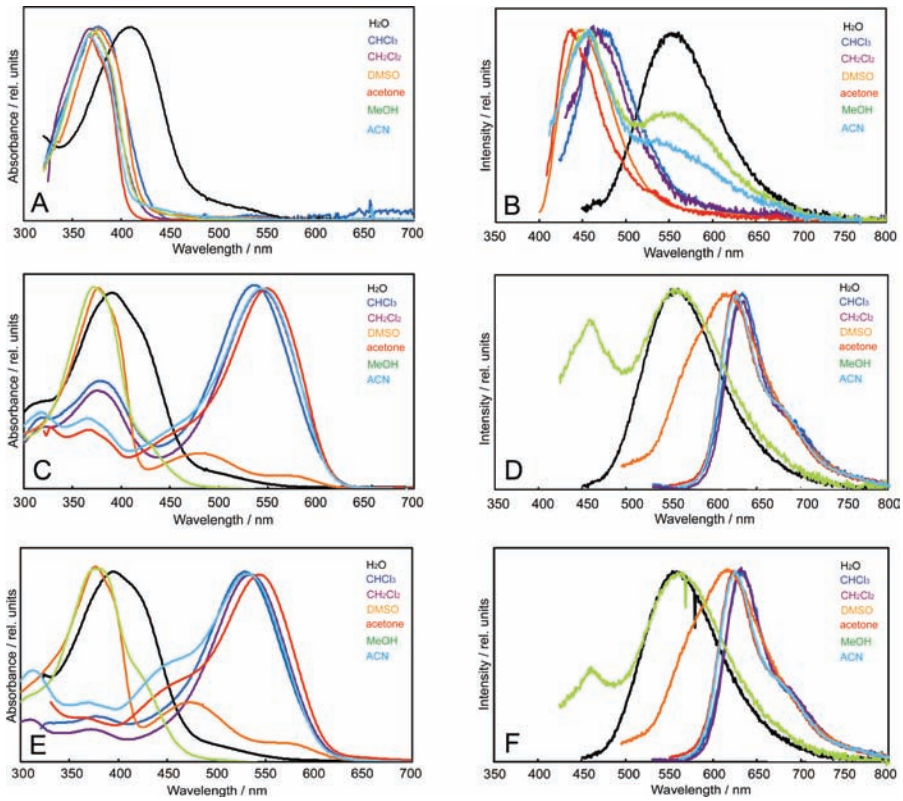


Figure 4. Solvent effects on the absorption (left-side panels) and emission (right-side panels) spectra of oxyluciferin (A, B) and its deprotonated species (C−F) obtained by treatment with Bu_4NOH in 1:1 (C, D) and 1:2 (E, F) ratio.

Table 1. Effects of Solvent and Deprotonation on the Absorption and Emission Spectra, and Assignments of the Chemical Forms of Oxyluciferin ($OxyLH_2$) in Various Solvents

	maximum wavelength[a]/nm (fluorescence yield/%)				
	without base		with base[b]		assignment[c]
solvent	absorption	emission	absorption	emission	
H_2O	**409**, 525sh	**553** (47.0)	**396**, 423sh, 500sh[d]	**559** (42.3)	phenolate-enol-$OxyLH^-$; $OxyL^{2-}$
MeOH	**372**	**457**, 552b (34.5)	**383**, 430sh	**463, 563** (40.3)	enol-$OxyLH_2$; enolate-$OxyLH^-$; phenolate-enol-$OxyLH^-$; $OxyL^{2-}$
ACN	**369**	**457**, 540sh (30.5)	370w, 450sh, **535**	**626**, 690sh (15.9)	phenolate-keto-$OxyLH^-$; phenolate-enol-$OxyLH^-$
acetone	**367**	**436** (27.7)	367w, 453sh, **546**	**625**, 690sh (20.9)	phenolate-keto-$OxyLH^-$; phenolate-enol-$OxyLH^-$
DMSO	**377**	**448** (51.1)	**377**, 476, 570b	574sh, **617** (44.4)	phenolate-keto-$OxyLH^-$; phenolate-enol-$OxyLH^-$
CH_2Cl_2	**369**	**465** (23.1)	375w, **535**	**633**, 690sh (22.2)	phenolate-keto-$OxyLH^-$; phenolate-enol-$OxyLH^-$
$CHCl_3$	**376**	**471** (30.8)	381w, **531**	**626**, 690sh (18.9)	phenolate-keto-$OxyLH^-$; phenolate-enol-$OxyLH^-$

[a] The main bands are given by bold numbers. W, weak band; b, broad band; sh, shoulder. [b] Two molar equivalents of Bu_4NOH. [c] The labels refer to Scheme 1. [d] When *t*-BuOK is used for deprotonation, $\lambda_{max}^{abs} = 377$ and $\lambda_{max}^{em} = 445$ nm.

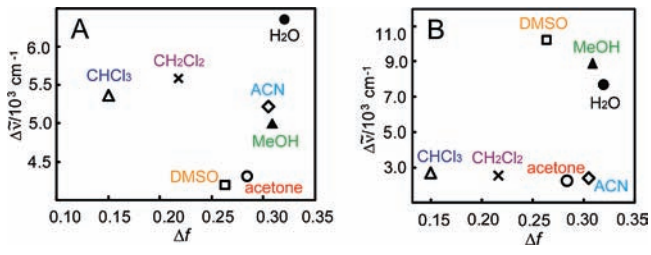


Figure 5. Lippert-Mataga plots of the solvent effects on the spectra of oxyluciferin (left) and its dianion (right). The solvent constants employed to calculate Δf were taken from ref 51a. The band maxima and fluorescence yields are listed in Table 1.

of which the latter can be readily produced by adding excess base to the aqueous solution.

Due to the limited solubility of *t*-BuOK in some organic solvents, to examine the solution effects on the spectra of deprotonated $OxyLH_2$, a stoichiometric ratio of appropriately diluted standard solution of Bu_4NOH was added to solutions of $OxyLH_2$ under anaerobic conditions. In aprotic solvents, this base can also account for the interactions of $OxyLH_2$ with the hydroxyl groups from the surrounding water molecules in the predominantly hydrophobic luciferase scaffold.[41] In the presence of Bu_4NOH, the fluorescence spectra of partially deprotonated $OxyLH_2$ are resolved better relative to the neutral species and *four* distinct regions can be distinguished based on the emission maxima (Figure 4C−F) and will be discussed in detail.

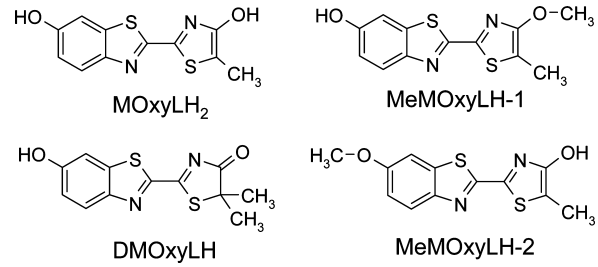
Region I (430−490 nm, *blue* emission) contains the higher-energy band in MeOH. Inspection of this band shows two components, the main peak at 463 nm and a shoulder at ∼442 nm. These two components correspond to the forms enol-$OxyLH_2$ and keto-$OxyLH_2$, both of which are *blue* emitters, with predicted maxima in DMSO at 439 and 421 nm,[44] respectively (note that the value of the observed split between the enol and keto forms, 21 nm, is also very close to the theoretical value, 18 nm). Moreover, the emission maxima in

region I are close to the emission of enol-$OxyLH_2$* in neutral DMSO solution (448 nm), for which the NMR data confirmed that both hydroxyl groups (enol and phenol) are in protonated form, and in acetone (436 nm), where $OxyLH_2$ exists mainly as the keto form, keto-$OxyLH_2$*. The assignment to keto-$OxyLH_2$ is further supported by comparison of its emission spectrum with the absorption spectrum of DMOxyLH, which is constrained to the neutral keto form keto-DMOxyLH by the substitution (in DMSO, $\lambda_{max}^{abs} = 391$ nm,[9,32] in MeOH, $\lambda_{max}^{abs} = 388$ nm[32,35]), by accounting for the expected Stokes shifts (actually, according to ref,[27] the emission of DMOxyLH in DMSO is weak). In accordance with this assignment, if the amount of base is increased, the emission intensity of region I decreases relative to that of region II (Figure 4D and F).

In a previous study,[29] in addition to enol-$OxyLH_2$ and keto-$OxyLH_2$, the enolate of the phenol form, enolate-$OxyLH^-$, was also assigned as blue emitter. Due to the stronger acidity of the phenol relative to the enol group, the accumulation in sufficient amount to detect clearly this species in solution was very difficult. However, for enolate-$OxyLH^-$ the theoretical calculations[44] indicate a strong red shift of the emission outside of the blue region (calculated in DMSO: 580 nm, in the presence of K^+: 564 nm). Therefore, it does not seem likely that enolate-$OxyLH^-$ is a blue emitter, but rather yellow or perhaps yellow-orange emitter. According to these considerations, only the neutral species of $OxyLH_2$, that is, enol-$OxyLH_2$ and keto-$OxyLH_2$, emit in the blue spectral region.

The emission in region II (~530−590 nm, *yellow-green* emission) appears as major emission in protic solvents (MeOH and H_2O), and as minor component in DMSO. As it will be elaborated in detail by the aid of the titration results below, the emission in this region is due to the monoanion phenolate-enol-$OxyLH^-$ and the dianion $OxyL^{2-}$. Due to the close values of the two pK_as of $OxyLH_2$, the assignment of the two species seems difficult, because the relative order of their emission energies is expected to depend strongly on the actual experimental conditions, including the nature of the solvent and added base, and could be exchanged by solvent/cation effects. Furthermore, the emission bands in region III are broader relative to region I, and the associated band overlap adds up to the difficulties in assignment. It has been found that in the real system, for example, the WT luciferase from the firefly *Luciola mingrelica*,[39] at least three forms coexist in wide range of pH values between 6 and 10. Accounting for the above difficulties, the relative order of the bands of phenolate-enol-$OxyLH^-$ and $OxyL^{2-}$ can be assessed experimentally: increasing the amount of base, and thus the equilibrium concentration of $OxyL^{2-}$, results in red shift of the absorption maximum and blue shift of the emission spectrum, indicating that $OxyL^{2-}$ absorbs at lower energies but emits at higher energy relative to phenolate-enol-$OxyLH^-$. At high excess of base (*t*-BuOK) in DMSO, where the ^{1}H NMR spectra revealed complete deprotonation, $OxyLH_2$ absorbs at 493 nm, and emits at 590 nm with a tail in the red region (Figure 3; orange-red emitter, predicted in DMSO: 600 nm[44]), and thus it exists as pure $OxyL^{2-}$. Similarly, the dianion of $MOxyLH_2$, $MOxyL^{2-}$, absorbs at 507 nm and emits at 595 nm. Due to the partial proton transfer and exchange with the solvents, the enol-phenolate form phenolate-enol-$OxyLH^-$ is observed in H_2O and MeOH even in absence of base, and emits at 553 and 552 nm (broad band), respectively. Addition of small amount of Bu_4NOH shifts the equilibrium enol-$OxyLH_2$/keto-$OxyLH_2$ toward the phenolate-enol-$OxyLH^-$ ion, which becomes the main component in H_2O (559 nm) and in MeOH (563 nm), but it is also created in DMSO (~574 nm). The coexistence in neutral MeOH solution of the neutral forms enol-$OxyLH_2$ and keto-$OxyLH_2$, and the ionic form phenolate-enol-$OxyLH^-$ is clearly a result of the rapid proton exchange processes between $OxyLH_2$ and the solvent, which appears as absence of the hydroxyl proton peaks in the ^{1}H NMR spectrum. It should be noted that the possibility that the monoanions phenolate-enol-$OxyLH^-$ and enolate-$OxyLH^-$ are green-yellow emitters in addition to the dianion $OxyL^{2-}$ has been already pointed out by White and Roswell.[27]

Scheme 2. Chemical Structures of the 5-Methyl Substituted Oxyluciferin Analogues $MOxyLH_2$ and DMOxyLH, and Two Methyl Ethers of $MOxyLH_2$ (MeMOxyLH-1 and MeMOxyLH-2) which are Used in the Discussion for Structural and Spectroscopic Comparison



The assignment of the spectrum of phenolate-enol-$OxyLH^-$ as the yellow-green emitter is supported by spectral comparison with the enol-methylated derivative MeMOxyLH-1 in Scheme 2. Spectral comparison of $OxyLH_2$ with MeMOxyLH-1 and MeMOxyLH-2, where the deprotonation of the respective enol and phenol groups is blocked by methylation, provides useful insight into the complicated chemical equilibrium of $OxyLH_2$ in solution, and can aid greatly the assignment of its electronic spectra. The solution absorption and emission spectra of both MeMOxyLH-1 and MeMOxyLH-2 in MeOH, in nonbasic and basic media were reported, as well as the respective spectra of MeMOxyLH-1 in water.[27] MeMOxyLH-1 is a convenient model of the phenol−enol form enol-$OxyLH_2$ because the thiazole hydroxyl group is constrained to the enol form by methylation. This assumption is justified by the close similarity between the band positions of MeMOxyLH-1 in MeOH ($\lambda_{max}^{abs} = 377$ and $\lambda_{max}^{em} = 450$ nm) and $OxyLH_2$ in DMSO ($\lambda_{max}^{abs} = 377$ and $\lambda_{max}^{em} = 448$ nm), where $OxyLH_2$ exists as pure enol-$OxyLH_2$ form. Upon addition of weak or strong base, the spectrum of MeMOxyLH-1 in MeOH exhibits strong red-shifts of 18 nm (377 nm → 395 nm) and 90 nm (450 nm → 540 nm) due to deprotonation of its only hydroxyl (phenol) group. In H_2O, the respective shifts for MeMOxyLH-1 are 30 nm (375 nm → 405 nm), and 5 and 80 nm (455 nm → 460/535 nm). If these shifts are applied to $OxyLH_2$, assuming similar magnitude of the effects due to annulation of the methyl substitution, the predicted maxima for phenolate-enol-$OxyLH^-$ are 390 nm (λ_{max}^{abs}) and 547/642 nm (λ_{max}^{em}) in MeOH, and 439 nm (λ_{max}^{abs}) and 558/633 nm (λ_{max}^{em}) in H_2O. The higher-energy emission maxima are in good correspondence for the region of emission of phenolate-enol-$OxyLH^-$ (~552−563 nm in MeOH, and 553−559 nm in water). Similarly, the compound MeMOxyLH-2 has only its hydroxyl functionality on the hydroxythiazole ring as ionizable group and can serve as useful model for the enol−enolate equilibrium. The nearly identical band maxima in MeOH ($\lambda_{max}^{abs} = 377$ and $\lambda_{max}^{em} = 445$ nm) with MeMOxyLH-1 ascertain that the compound is not ionized in neutral solution.

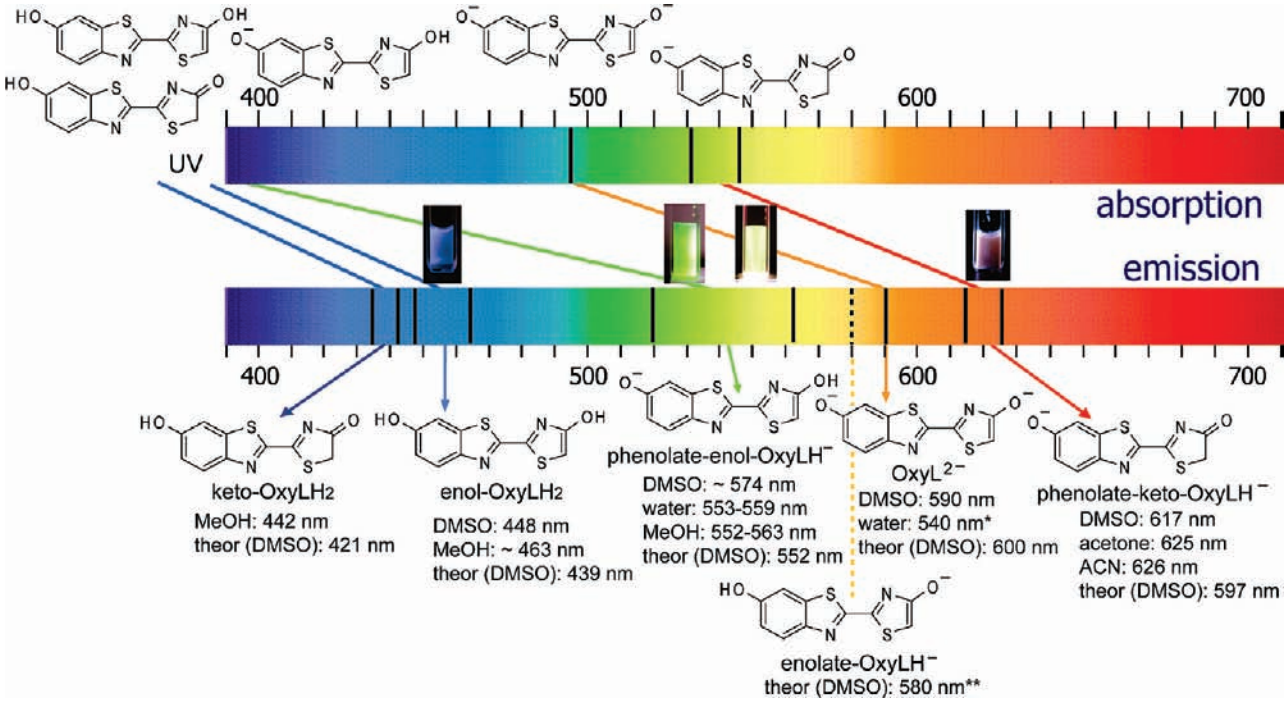


Figure 6. Assignment of the absorption and emission spectra of the firefly emitter oxyluciferin, based on experimental and theoretical spectroscopic data (*Unstable species; **Difficult to be determined experimentally because of the small concentration in mixture with other species). The calculated values of the emission maxima shown for comparison were taken from a paper by Nakatsuji et al.[44] The figures in the middle show the color change of oxyluciferin (from left to right) in DMSO without added base, in water with and without base, and in DMSO in presence of base.

Region III (620−626 nm, *red* emission), contains the main emission maxima of $OxyLH_2$ in basic ACN, CH_2Cl_2, $CHCl_3$, DMSO and acetone solutions. Note that when the molecule is excited at 600 nm, the emission spectrum of $OxyLH_2$ in DMSO (Figure 3), becomes identical to those in ACN, CH_2Cl_2, $CHCl_3$ and acetone (panel D in Figure 4), so that the spectra in all these solvents are similar, and in the same time, different from those recorded from MeOH and H_2O. This emission can be assigned to the keto-monoanion form phenolate-keto-$OxyLH^-$ (Scheme 1) based on the emission spectra of the $DMOxyL^-$ ion, $\lambda_{em}{}^{max} = 632-633$ nm in DMSO,[27] which exists as keto form ($\lambda_{abs}{}^{max} = 489$ nm in MeOH,[35] $\lambda_{abs}{}^{max} = 578-580$ nm in DMSO),[27] and accounting for the accumulative red shifts due to the double methyl substitution. The spectra of the phenolate-keto-$OxyLH^-$ form are the most red-shifted among the oxyluciferin forms in Scheme 1, and therefore this species appears as red emitter.

Based on the above discussion, the approximate spectral regions covered by the absorption and emission of the chemical forms of $OxyLH_2$ can be summarized as shown in Figure 6. The spectra of various species are generally in good agreement with the theoretical emission energies predicted by Nakatsuji et al.[44] Although the exact values of the emission maxima will be normally affected by the solvent polarity, specific interactions including hydrogen bonding and presence of cations, it becomes clear that the neutral forms of the firefly emitter emit exclusively in the blue region, and the phenolate of the enol form is the yellow-green emitter. The phenol-enolate form, which probably does not exist in pure form in solution and the dianion of oxyluciferin are yellow or orange emitters with close, overlapping emission bands, while the keto form of the phenolate monoanion is red emitter.

As it can be concluded from the Lippert-Mataga plots in Figure 5, the polarity of the protein environment is expected to exert significant effect on the spectral Stokes shifts of the emitter in the luciferase active site. If in the first approximation the effect of the environment on the bioluminescence is modeled by the fluorescence of the emitters in solutions, and if the special case of solution of the neutral emitter in water (where, as it was explained above, complete ionization occurs) is excluded, then Figure 5 shows that the solvent polarity has opposite effect on the Stokes shifts of the neutral emitter molecule and its ion. Namely, while *less polar* environment (CH_2Cl_2, $CHCl_3$) generally causes *larger* Stokes shift to the emission of a neutral molecule than the more polar one (DMSO, MeOH), just the opposite is true for the anion. This means that the deprotonation of oxyluciferin causes switching of the effect of the environment polarity on its emission, this effect possibly being combined with the effect on the spectra of the keto−enol tautomerism. Notably, the magnitudes of the Stokes shifts are nearly double in the case of the ion relative to the neutral molecule. This result indicates that in the protein matrix, the emission of the ionized oxyluciferin molecule will be much more susceptible to the presence of polar components (e.g., water molecules) than the neutral emitter. Therefore, it can be expected that the presence (or not) of dynamic water molecule(s) close to the phenolate group, which may come close to the emitter during some stage of the bioluminescence reaction, will result in significant red shifts of the emission.

Protonation and Keto−Enol Equilibria of Oxyluciferin and Its Derivatives. According to the earlier theoretical calculations,[46] the prevalence of the enol or keto form of the 4-hydroxythiazole depends critically on its substituents. If only the isolated 4-hydroxythiazole ring is considered, the enol form is preferred in the gas phase. Substitution at C2 with the benzothiazolyl moiety stabilizes the keto form over the enol form. Further single and double methyl substitution at C5, such as in the structures of $MOxyLH_2$ and DMOxyLH, stabilize the

enol form, but in all cases mixtures of isomers are expected, with the solvation effects favoring the keto tautomer. Actually, all previous high-level theoretical studies of the unsubstituted oxyluciferin[44−48] as well as our DFT calculations (see above) showed that from the two *trans*-planar neutral forms, the keto form, keto-$OxyLH_2$, is more stable than the enol form, enol-$OxyLH_2$. As it was mentioned above, according to the DFT single point energy calculations, the neutral keto forms of $OxyLH_2$ and $MOxyLH_2$ *in vacuo* are stabilized for ΔE[(enol-$OxyLH_2$)-(keto-$OxyLH_2$)] = 6.57 and 3.99 kcal mol^{-1} relative to the respective enol forms. From the high-level, detailed theoretical CASSCF and CASPT2 results of Goddard and Yang[48] we find that similar or even larger stabilization is expected for the respective phenolate ions in S_0: for example, if the CASSCF electronic energies are considered, within the configuration spaces CASSCF(6,6), CASSCF(8,8) and CASSCF(10,10), the ground-state phenolate-keto form is more stable than the phenolate-enol form for ΔE[(phenolate-enol-$OxyLH^-$)-(phenolate-keto-$OxyLH^-$)] = 16.57, 21.03, and 5.28 kcal mol^{-1}, respectively.

The main problem with the determination of pK_a values of $OxyLH_2$—the close values of the two constants—is amplified with the concomitant keto−enol tautomeric equilibrium and the enhanced chemical instability in basic media. Mainly as a result of the narrow pH interval within which different forms of $OxyLH_2$ coexist and the instability problems in basic media, the usual titrations experiments are not informative and there are no commonly accepted values for the two pK_as of $OxyLH_2$. For the phenol−phenolate equilibrium of $OxyLH_2$, the experimental values of 7.4,[53] 8.5[29] and 8.60[54] were reported. For $MOxyLH_2$, the value of pK_a of the phenol−phenolate equilibrium was estimated as 7.8, but it could not be determined precisely due to instability.[30] The pK_a of the phenol−phenolate reaction of DMOxyLH was reported as 7.8 from the UV spectra.[30] For the enol−enolate equilibrium, the pK_a value was reported as 6.85 for $OxyLH_2$[54] and 7.6 for $MOxyLH_2$.[27] In the first exited state, the estimated pK*of $OxyLH_2$, $MOxyLH_2$ and DMOxyLH are −0.5,[29] 0.7 and −3.91,[30] respectively.

To avoid the obstacles with the instability of the emitter in solution, which increase with the time at high pH, we performed indirect spectrophotometric titration of $OxyLH_2$ in water: the dependence of its absorption spectrum on the pH in the physiologically important region between 6.0 and pH 9.0 was recorded from a series of 30 phosphate-buffered aqueous solutions (the exact pH value was confirmed by direct measurement). Fresh aqueous solution of $OxyLH_2$ was added in identical concentration to each of the buffered solutions, and the absorption spectrum was recorded immediately after the mixing and before any significant decomposition had occurred. The plots of the maximum absorption wavelength, λ_{max}^{abs}, and the absorbance at two characteristic wavelengths, 370 nm (λ_{370}^{abs}) and 450 nm (λ_{450}^{abs}) in Figure 7 show the combined effects of deprotonation (λ_{max}^{abs}) and keto−enol tautomerization (λ_{370}^{abs} and λ_{450}^{abs}).

As shown in the inset in Figure 7, increasing pH results in decreased absorption at 370 nm and increased absorption at 415 nm. The large number of very closely lying inflections on the titration curves, each inflection point being characterized by a small number of measurement points, as well as imperfections in the measurement conditions for each individual solution due to both the titration method and the sample instability in some cases, prevented strictly quantitative analysis of the titration results. Under such circumstances, the inflection points were calculated simply as average of the two adjacent points which were considered closest to the transition, and based on the slopes on both sides from the transition. The three titration curves of λ_{max}^{abs}, λ_{370}^{abs} and λ_{415}^{abs} show five transition points, at pH 6.54, 7.04, 7.44, 8.14, and 8.65. The most prominent changes in all three curves, and especially of λ_{max}^{abs}, occur at pH 7.44 and 8.14, which apparently correspond to the two pK_a values of $OxyLH_2$. On the basis of the predicted value for the first pK_a in water,[45] which show that at the first pK_a the phenolate species exists in 20-fold excess over the enolate, the lower pK_a value (7.44) is prescribed to the dissociation of the phenol group and the higher value (8.14) to that of the enol group. This is supported by the assignment UV−visible spectral maxima for the phenolate and the enolate species in water, discussed above. At pH 7.04, where the absorption at λ_{415}^{abs} changes without change in the trend of λ_{max}^{abs}, the keto−enol tautomeric transformation of the neutral emitter occurs. The higher and lower pH values than 7.04 correspond to the neutral enol and keto forms, because at higher pH (7.04−7.44) the enol group of the 4-hydroxythiazole moiety, and thus the enol-$OxyLH_2$ form, is stabilized by formation of $-OH\cdots OH^-$ hydrogen bonds with the small excess of base. The neutral keto form, keto-$OxyLH_2$, is stabilized in weakly acidic medium. The point pH 6.54 is not accompanied by a significant change in λ_{max}^{abs} and probably corresponds to either formation of weakly hydrogen complex $C{=}O\cdots H_3O^+$ between keto-$OxyLH_2$ and the hydroxonium ion, or to the first protonation point ($OxyLH_3^+$). At more basic media, above pH 8.65, the structure of the emitter is affected by chemical isomerization (decomposition) reactions.

The titration results show that in aqueous solution, the green emitter phenolate-enol-$HOxy^-$ exists in a very narrow pH interval of about 0.7 units (7.44−8.14), most probably in equilibrium with smaller amounts of other species. It becomes clear that subtle changes in pH will bring to drastic changes in the ratio of the species and will shift the equilibrium toward

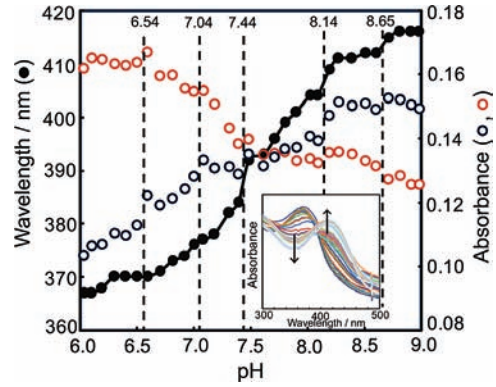


Figure 7. pH-dependence of the maximum absorption wavelength, λ_{max}^{abs} (left scale, black filled circles), λ_{370}^{abs} (right scale, red empty circles) and λ_{415}^{abs} (right scale, blue empty circles) of oxyluciferin in deoxygenated water. The change of the lowest energy absorption bands is shown in the inset. To prevent decomposition, each individual spectrum was recorded under anaerobic conditions, from a fresh solution of oxyluciferin treated with phosphate-buffered aqueous solution. The major species within the pH ranges denoted with vertical bars are assigned to: $OxyLH_3^+$ or $OxyLH_2\cdots H_3O^+$ (<6.54), keto-$OxyLH_2$ (6.54−7.04), enol-$OxyLH_2$ (7.04−7.44), phenolate-enol-$OxyLH^-$ (7.44−8.14), $OxyL^{2-}$ (8.14−8.65), byproduct (>8.65).

(53) Gandel'man, O. A.; Brovko, L. Yu.; Ugarova, N. N.; Shchegolev, A. A. *Biochemistry (Moscow)* **1990**, *55*, 785.
(54) Goto, T.; Kubota, I.; Suzuki, N.; Kishi, Y.; Inoue, S. In *Chemiluminescence and Bioluminescence*; Cormier, M. J., Hercules, D. M., Lee, J., Eds.; Plenum Press: New York, 1973; pp 325−335.

the dianion $OxyL^{2-}$at higher pH (8.14−8.65), or toward the neutral enol form enol-$OxyLH_2$ at lower pH (7.04−7.44). At still lower pH values (<7.04) enol-$OxyLH_2$ isomerizes to the keto form, keto-$OxyLH_2$, which exists as dominant species in a narrow interval (6.54−7.04) before the molecule is protonated at lower pH values. More importantly, the pH interval where the form phenolate-enol-$OxyLH^-$ dominates the equilibrium (7.44−8.14) includes the value pH 7.8, where the luciferase exhibits maximum activity. Therefore, at the optimum pH, in buffered aqueous solution $OxyLH_2$ exists mainly as the enol-phenolate form phenolate-enol-$OxyLH^-$, which emits yellow-green light (Figure 6). Our conclusions support the previous assumptions of the existence of the ionic phenolate form in the luciferase based on the pH-dependence of the bioluminescence spectra of $OxyLH_2$, $MOxyLH_2$ and DMOxyLH in WT and mutant luciferases from *L. mingrelica*.[31,39]

Bioluminescence Spectra. In the structure of the complex of $OxyLH_2$ and AMP with *L. cruciata* luciferase, the product $OxyLH_2$, created *in situ* by reaction of D(−)-LH_2 in solution and crystallized, was reported in its neutral form keto-$OxyLH_2$, with the oxygen on the thiazole ring as keto group.[41] However, elucidation of the precise ligand distances in macromolecular structures is oftentimes precluded by limitations related to the resolution of the diffraction data. Moreover, it should be noted that both the absorption and emission spectra of firefly luciferase are pH-dependent.[39] As the protein crystal structures, as well as the chemical state of the ligand and its surrounding depend on the pH of the crystallization solution,[55] in the case of ligands with multiple and close pK_a values around neutral pH, minute differences in the ionic strength or polarity of the crystallization solution or of the protein environment may trap the ligand in certain chemical form which does not necessarily correspond to the form in which it exists after the deexcitation in the real system. Finally, according to the spectral assignments, the neutral form keto-$OxyLH_2$, as well as its neutral enol tautomer enol-$OxyLH_2$, is blue emitter under all solvent conditions where it remains neutral. This is in marked contrast with the natural bioluminescence of the firefly luciferases, which do not exhibit blue emission at any pH value. Therefore, the excited neutral keto form (keto-$OxyLH_2$)* is not the real emitting state of the firefly emitter.

To unravel the chemical state of ($OxyLH_2$)* in the luciferase active pocket, we prepared its complex with luciferase from *L. cruciata* in two ways: (a) by complexation of AMP and $OxyLH_2$ with excess luciferase (where the high binding constant and the excess enzyme ensured complete inclusion) and subsequent laser-excitation of the emitter in the active site, and (b) with *in situ* generation of the excited emitter by activation of a precomplexed substrate D(−)-LH_2 with the enzyme by treatment with ATP (Figure 3E). Upon complexation of pure $OxyLH_2$ with the enzyme, which at pH 7.8 exists mainly as the phenolate-enol-$OxyLH^-$ form, a red shift of the single absorption maximum from 408 to 414 nm was observed, which confirmed that complete inclusion of the ligand had occurred. At the same time, the emission maximum remained unaffected (555 and 556 nm before and after complexation, respectively), indicating that the chemical form of the emitter molecule in S_1, (phenolate-enol-$OxyLH^-$)* did not change upon complexation. Accordingly, the bioluminescence spectrum of *in situ* produced excited emitter is also consistent with the observation that in phosphate-buffered aqueous solution at pH 7.8, corresponding to the maximum of firefly luciferase activity, the bioluminescence exhibits yellow-green emission $\lambda_{max}^{em} = 557$ nm (see the inset in Figure 3E). The maximum emission wavelength of the *in situ* produced emitter coincides with the fluorescence of laser-induced emission of the monoanion when it is uncomplexed in solution, $\lambda_{max}^{em} = 556$ nm, and when it is complexed with luciferase, $\lambda_{max}^{em} = 555$ nm, clearly indicating that in all cases, the emission appears from the excited (phenolate-enol-$OxyLH^-$)* form. In the two luciferase structures determined thus far, that of uncomplexed *P. pyralis* luciferase[40] and complexed *L. cruicata* luciferase,[41] the active pocket was found to be quite hydrophobic (note, however, that in the latter case a water molecule is situated close to the hydroxyl group of the emitter in the intermediate state, which could change significantly the hydrophobicity of the active site and affect the emission; Figure 2C). If the solvent effects on the absorption maxima of the pure neutral emitter are considered (Table 1), the decreased polarity of the enzyme active site is additionally supported by the red shift observed for the absorption maximum upon complexation (408 nm → 414 nm). The assignment of the enol-phenolate form to the excited emitter state is consistent with the yellow-green emission in aprotic solvents (e.g., DMSO), which corresponds to phenolate-enol-$OxyLH^-$, λ_{em}(DMSO) ≈ 574 nm, if blue-shifts of about 20 nm are accounted for that are normally expected to occur upon the complexation of the emitter.[31]

In the case of *P. pyralis*, *in vitro* bioluminescence in acidic medium (optimum at pH of about 6.8) is red, but it turns to yellow-green in weakly basic media (pH of about 8).[9] Similarly, the bioluminescence maximum of luciferase from *L. mingrelica* depends on the protonation and keto−enol equilibria, which are in turn affected by the pH.[39] Three of the phenolate forms of $OxyLH_2$ were assigned to explain the observed spectral change: in the WT enzyme, the species assigned to the keto form ($\lambda_{max}^{em} = 618$ nm) decreases to 20% with pH and remains constant at pH 7.0. The emission of this species and the assignment corresponds to the phenolate-keto-$OxyLH^-$ form in our UV−visible spectra. The species previously assigned as the enol form phenolate-enol-$OxyLH^-$ ($\lambda_{max}^{em} = 587$ nm) reaches a maximum of about 30% at pH 7.0 and decreases at more basic conditions. The content of the third species, assigned to the enolate ion $OxyL^{2-}$ ($\lambda_{max}^{em} = 556$ nm), increases with pH and reaches 60% at pH above 7.0. The relative order of band maxima in the fluorescence spectra, explained above, indicates that the assignment of the latter two species should be probably exchanged.

To obtain additional information on the chemical identity of $OxyLH_2$ in the luciferase scaffold, nanosecond fluorescence decay lifetimes ($\lambda_{ex} = 400$ nm) were recorded of the emitter complexed with *L. cruciata* luciferase and compared to those of various forms of $OxyLH_2$ obtained by treatment with base (*t*-BuOK) in DMSO. In the case of the pure emitter, two species were observed, with close lifetimes $\tau_1 = 2.3$ ns and $\tau_1 = 3.5$ ns. At least one of these species can be ascribed to (enol-$OxyLH_2$)*, which was determined by ^{1}H NMR. The other species may be due to the (keto-$OxyLH_2$)*. In the presence of a small amount of base, where the excited solution is enriched with (phenolate-enol-$HOxy^-$)*, the lifetime of the emitter obeys single exponential decay law with $\tau = 6.2$ ns. Large excess of base resulted in a single decay with longer lifetime, $\tau = 9.5$ ns, corresponding to the fluorescence decay of the excited dianion ($OxyL^{2-}$)*. Therefore, the fluorescence lifetime of the emitter

(55) Elsliger, M.-A.; Wachter, R. M.; Hanson, G. T.; Kallio, K.; Remington, J. S. *Biochemistry* **1999**, *38*, 5296.

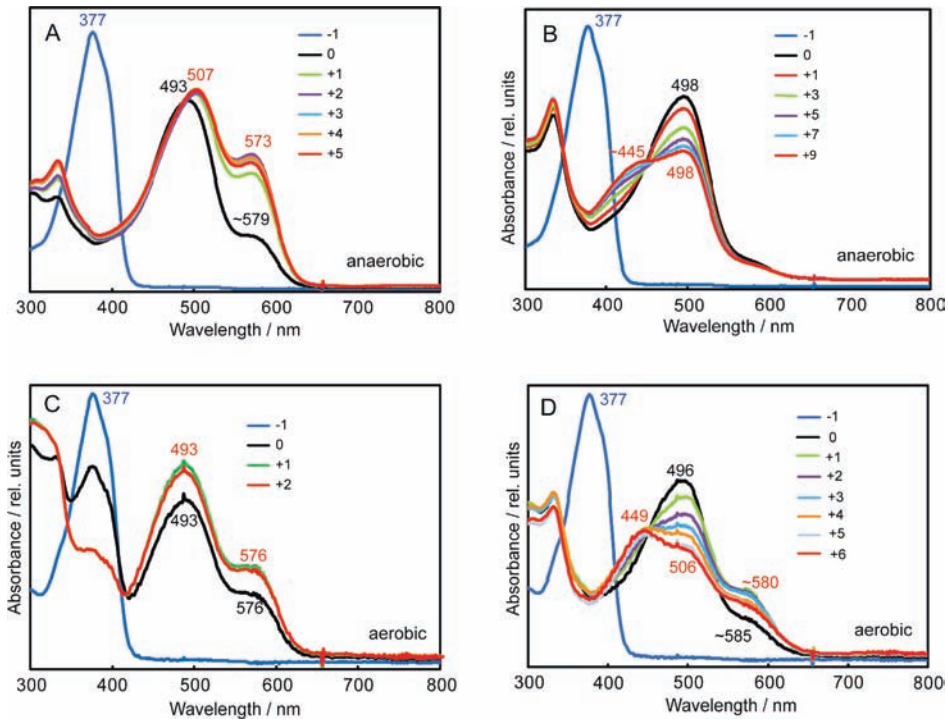


Figure 8. Temporal profiles of the absorption spectra of oxyluciferin and its deprotonated forms in DMSO under anaerobic (A, B) and aerobic (C, D) conditions. The spectra in panels A and C correspond to small amount of base (*t*-BuOK) and those in panels B and D correspond to large excess of base. The numbers in the legends are the respective times (in minutes) from mixing of the reactants ("−1" stands for solution before the addition of base). The numbers represent the band maximum wavelengths in nanometers.

in model DMSO solutions increases by increasing the degree of deprotonation. The fluorescence decay of $OxyLH_2$ in luciferase was fitted to a single exponential function, which gave lifetime $\tau = 7.4$ ns (Figure 3F). This value is close to the lifetime of $OxyLH_2$ in presence of small amount of base, that is (phenolate-enol-$OxyLH^-$)*, in support of the above discussion of the existence of the emitter mainly as the phenolate anion of the enol form in the luciferase active site. This conclusion is also in agreement with the absorption maxima of $OxyLH_2$ in water, where, as discussed above, the yellow-green emission is due to deprotonation of the phenol group.

Effects of Decay of Oxyluciferin on the UV−Visible Spectra. In the absence of air and base, $OxyLH_2$ is stable in all solutions. As shown in Figure 8, when base is added in DMSO, the spectra show temporal changes both in aerobic and anaerobic conditions. In all cases, two bands were observed immediately after the addition of base, a stronger band at 493 nm, which shifts to 496 − 498 nm higher wavelength at higher base concentrations, and weaker band at 576−579 nm, which appears as ∼585 nm at higher base concentrations. Analogous effects were observed for $MOxyLH_2$ (Figure S4, Supporting Information).

Implications for the Mechanism of Color Tuning. Four main basic mechanisms have been suggested to explain the color difference of firefly bioluminescence. According to the early "*keto-enol mechanism"* suggested by White et al.,[9,56] based on analogy to the chemiluminescence of active esters in solution, the difference in emitted color is due to switching between the enol and keto forms of the thiazole hydroxyl group, which emit yellow-green and red lights, respectively. According to the alternative "*twisted intramolecular charge-transfer (TICT) mechanism*", suggested by McCapra et al.,[57] the color difference is associated with the rotational freedom of the keto form in its first excited state: conformations close to the planar form were suggested as green emitters, while significant distortion around the bridging bond C2−C2′ was presumed as origin of red emission. The TICT mechanism was discredited recently by the high-level theoretical calculations of Goddard and Yang,[48] and Nakatsuji et al.,[44] both showing that it is not likely that the twisted conformation represents a minimum in the first excited state. Two additional alternative explanations were also given that are related to the effects of the environment of the emitter in the luciferase. According to the "*microenvironment mechanism*",[8,29] based on comparison of the bioluminescence spectra between WT and mutant luciferases and their dependence on the pH, the shifts in the spectral maximum in *L. mingrelica* luciferase are either due to proton transfer between the phenolate group of $OxyLH_2$ and the positively charged arginine residue (Arg218), or to the keto−enol−enolate equilibrium of the hydroxythiazole ring directed by interaction with His247 and Thr345.[39] The recent report[36] of green bioluminescence of *in situ* produced DMOxyLH, which is constrained to the keto form, showed that, depending on the conditions, the keto form alone can emit both red *and* yellow-green light. This observation was later employed by Hirano et al.[32] to develop a more explicit model of the color

(56) White, E. H.; Rapaport, E.; Hopkins, T. A.; Selkinger, H. H. *J. Am. Chem. Soc.* **1969**, *91*, 2178.

(57) McCapra, F.; Gilfoyle, D. J.; Young, D. W.; Church, N. J.; Spencer, P. In *Bioluminescence and Chemiluminescence. Fundamentals and Applied Aspects*; Campbell, A. K., Kricka, L. J., Stanley, P. E., Eds.; John Wiley & Sons: Chichester, UK, 1994; p 387.

(58) It should be noted, however, that unlike the real intermediate (luciferyl adenylate), in which the oxyluciferin residue is flexibly bound to the rest of the molecule, the model dehydroluciferyl ester is fixed to planarity by the hybridization at the atom C4, and it also includes sulfonylamido instead of phosphate group. These differences might result in different interactions with the environment due to the decreased flexibility of the model intermediate.

change based on the interaction with the environment. This result supports the TICT mechanism, but as carefully stated by Branchini et al., who were the first to observe green emission from the keto form, it can not exclude the keto−enol mechanism: "While our results also demonstrate that the keto-enol tautomerization mechanism of White is not required to explain red and green firefly bioluminescence, they do not preclude the occurrence of this phenomenon".[36a] On the basis of these results, a "*resonance-structure mechanism*" was suggested, according to which the charge resonance in the phenolate of the keto form alone is responsible for the color change.

In principle, the quest for the origin of the different bioluminescence colors includes two different points: (a) the chemical identity of the emitter, that is, the form in which it exists in the protein, and (b) the factors which affect the deexcitation energy of the emitter. The neutral forms of oxyluciferin are blue emitters under all conditions where they remain neutral in S_0 as well as in S_1. Because blue emission has not been observed in the real system yet, the emitter in the excited state is ionized. The titration results showed larger acidity of the phenol hydroxyl group, and thus, the excited emitter exists as a phenolate ion. This does not exclude the possibility that the ion is created in the excited state, after a proton transfer has occurred from the neutral form; however, the comparison of the fluorescence of the emitter in buffered aqueous solutions and its bioluminescence in the protein matrix confirmed that the emitting species exists as phenolate ion even in the ground state. Now, a question occurs as to whether the emitting phenolate ion exists as its enol or keto form? Although the data obtained on molecules similar to $OxyLH_2$ are very useful, it seems difficult to answer this question solely from the spectra-structural data obtained from model compounds other than the real emitter. Namely, the single or double methyl substitution in compounds such as $MOxyLH_2$ or DMOxyLH, where additional red shifts relative to the unsubstituted emitter appear as a result of the substitution, and the methyl-hindered docking of these analogues to the active site known from fluorescence titration experiments,[31] partially preclude direct assessment of the spectral effects. Our fluorescence and bioluminescence spectra of the unsubstituted emitter in various solvents, the pH dependence of the spectra, the crystal structures of the emitter and the model theoretical calculations are all consistent with the yellow-green emitter in the luciferase existing as the *phenolate ion of the enol form*, phenolate-enol-$OxyLH^-$. Moreover, the red shift observed by complexation shows that when included in the active pocket, the ground-state phenolate ion is enclosed in a less polar environment than the aqueous solution. This implies that there are no water molecules proximate to the phenolate group of the emitter, which would otherwise greatly affect the environment polarity, or that the effect from the eventual water molecule(s) to the overall polarity is overweighed by interaction with hydrophobic protein residues, including alkyl residues (Ile) and phenyl rings (Phe). On the other hand, the insensitivity of the emission energy to complex formation with the protein shows that the change of polarity by complexation does not affect the excited state of the emitter. The excited state of the phenolate ion decays before significant reorganization of the protein matrix has occurred, that is, the protein dynamics are slower relative to the deexcitation. A second proton transfer from the hydroxythiazole ring is not likely to occur; thus the emitter appears as a monoanion, a conclusion which is in agreement with the above discussion and is also supported by the comparison of the fluorescence lifetimes.

The reasons behind the color tuning phenomenon of the firefly luciferase can be considered at two levels, as being caused by *intra*molecular or *inter*molecular factors. Because the previously considered intramolecular factors, mainly the TICT mechanism, were discredited on theoretical grounds,[44,48] we will consider here only the intermolecular factors. The crystal structure of *L. cruciata* luciferase complexed with $OxyLH_2$ and reaction intermediate determined by Nakatsu et al.[41] (Figure 2) provides much more detailed insight into the nature of the binding site of the emitter than it has been available before. Based on careful analysis of the environment in these structures, it is possible to address some of the viable reasons for the differences in the emitted color which are not accessible from the solution-state studies. It has been already suggested[41] that the color change can be caused by reversible closing/opening of the pocket around the emitter by conformational change of the proximate isoleucine residue (Ile288) which is induced by point mutation (Figure 2D). Indeed, as it can be also inspected from the Hirshfeld surfaces of the ligand portion of the model intermediate shown in Figure 2E, the (attached) ligand is placed in nearly identical environments in the green- and red-emitting luciferases, except for the structure around Ile288. Moreover, from the blue regions of the surfaces in Figure 9F it is evident that one side of the active pocket remains open and provides relatively large freedom to the nearly planar ligand to shift to much less restricted space. The structure of the unbound emitter in the same pocket (Figure 2A), however, shows that in the WT (green-emitting) enzyme, the emitter does not slip far from the product AMP. There are several possible factors that are responsible of the ligand remaining close to the AMP in the active pocket, and each of these could change during the color tuning. The most apparent reason is the hydrogen bond of the thiazole oxygen atom to a water molecule (2.91 Å), which further bonds to the protein scaffold (Gly318 and Lys531) through two additional hydrogen bonds (Figure 2A and B). Moreover, the thiazole oxygen is close to the negative phosphate residue (2.86 Å) and therefore, as explained above, it can exist in the enol form fixed by a strong hydrogen bond to the phosphate. Shifting the ligand away from the AMP and the water molecule would weaken or destroy these hydrogen bonds, which would destabilize the enol form and the emitter could switch to its keto form. The corresponding change of hydrogen bonding network will inevitably result in redistribution of the negative charge within the ion and thus, it will affect the emitted color. Additional detail which has not been pointed out in the previous studies, but becomes apparent from Figure 2C, is that the phenyl ring from the residue Phe249 approaches the planar molecule laterally and could $\pi-\pi$ interact with its π system. Being attached by a flexible bond, this phenyl group could reversibly approach or retreat from the ligand during the bioluminescence reaction. Changes of the $\pi-\pi$ stacking distance as small as 0.02−0.10 Å are expected to result in significant shifts in the emission maxima by modulation of the deexcitation pathways. In turn, the orange emitter $OxyL^{2-}$, the yellow-orange emitter enolate-$OxyLH^-$ or perhaps even the yellow-green emitter phenolate-enol-$OxyLH^-$ could shift their emission to the orange or red regions. Moreover, a dynamic water molecule that comes close to the phenolate group of the emitter during the process may change the overall situation radically.

The preceding discussion generally implies that the emitter is maintained close to the site of its creation by collective intermolecular interactions, which include hydrogen bonding (with water molecules, amino acid residues and/or AMP),

Coulomb interactions (with the phosphate of the AMP), and $\pi-\pi$ interactions (with a phenyl ring). Then, which of these factors is the most important for the color tuning? Which effect will prevail will be ultimately determined by the relative positions of the ligand in the yellow-green- and red-emitting states, and therefore on the conformational change of the whole protein network. At the present, the structure of the free (unattached) emitter only in the WT green-emitting protein is known. However, because the protein framework acts as a cooperative ensemble, it seems more sensible to suggest that it will affect the structure of the emitter in a multiple ways, with the specific and nonspecific interactions participating simultaneously.

Conclusions

As summarized in Figure 6, by the first detailed study of the absorption and emission spectra, and the crystal structure of oxyluciferin, the firefly bioluminescence emitter, we have shown that the triple chemical equilibrium of the unsubstituted emitter can provide a range of emission energies, spanning wavelengths from the blue to the red region of the visible spectrum, without the need of constraining its molecular structure to the keto form. Actually, together with the steady-state and time-resolved spectral analysis, the structure of pure oxyluciferin and its 5-methyl analogue indicate that the enol group of the hydroxythiazole ring can be notably stabilized in the protein scaffold, and that the phenolate ion of the enol form exhibits closest properties to the green-yellow emission observed in the real, WT (and most common) yellow-green-emitting system. Our detailed study on the dependence of the spectrum on the pH showed that within a relatively narrow pH region, including conditions that are close to the physiological conditions, a complex mixture of several species with distinct spectral properties can exist, so that minute changes in the pH causes drastic shifts in the equilibrium and the emitted wavelength. In addition to the pH, the interplay among the three equilibria can be affected by the polarity of the environment, presence of ionic species in close proximity to the emitter, intramolecular charge redistribution caused by directional intermolecular interactions, and $\pi-\pi$ stacking. Considering the possible interplay of multiple factors, based on model solutions of oxyluciferin or its analogues, it seems futile at the present to give clear and definite preference to any of these factors as the primary controlling factor, because there does not seem to be a direct way to apply the conclusions approximated on model systems in solution to the real biological system *in vivo*. In the direct application of the conclusions that are based on fluorescence (or chemiluminescence reactions) to explain the color tuning of the bioluminescence, the limitation of the environment effects to any single factor would represent an oversimplification, especially considering that even presence of a single water molecule in protein scaffold could result in drastic change of the properties of the environment. Moreover, unlike other fluorescent proteins, most notably the green fluorescent protein, where the emitter is attached to the protein framework, in the case of the firefly luciferase the emitter has larger freedom for movement inside the active pocket, and thus the *dynamic aspects* of the emitter-environment interaction need to be considered as well, if an ultimate evidence of the color tuning mechanism is to be sought for. Instead, it seems viable to accept that several factors can act collectively in the protein scaffold during the tuning of the emission color, until direct evidence is obtained about the structural changes that occur upon the color change.

Experimental Section

Materials. Commercial D(−)-LH_2 (95% by HPLC; AnaSpec), 5-adenylic acid (AMP; TCI), disodium adenosine-5′-triphosphate (ATP; Chameleon), 2-cyano-6-hydroxybenzothiazole (98.2%; Sequoia), ethyl thioglycolate (TCI) and ethyl 2-mercaptopropionate (TCI), as well as spectroscopic grade solvents (DMSO, MeOH, ACN, CH_2Cl_2 and $CHCl_3$; Nacalai Tesque) were obtained as commercial products and used without prior purification. Recombinant luciferase from *L. cruciata* expressed in *Escherichia coli* (activity: 6 milliunits/mg) was purchased from Wako. The phosphate buffer was prepared by mixing, in corresponding ratios, 100 mM NaH_2PO_4 and 100 mM Na_2HPO_4, and purged with nitrogen prior to use (the very high or low pH values were adjusted with aqueous NaOH or HCl).

Synthesis of Oxyluciferin ($OxyLH_2$). The general procedure for synthesis of $OxyLH_2$ of Suzuki et al.[23] was adopted. Because we found that small changes in the experimental conditions, especially in the last stage, greatly affect the yield and purity of the product, the synthetic procedure is described here in detail. All solvents were purged with nitrogen prior to the use. Through the first funnel of a four-neck reactor, a fresh mixture of 1 mol·L^{-1} NaOH in nitrogen-purged water (1.2 mL), 20 mL nitrogen-purged MeOH and 750 μL ethyl thioglycolate was added. After thermal equilibration, a solution of 2-cyano-6-hydroxybenzothiazole was added through the second funnel, which had been prepared by dissolving 500 mg 2-cyano-6-hydroxybenzothiazole in previously nitrogen-purged 60% MeOH in water (110 mL). The mixture was stirred gently for 2 min, and then a solution of 0.43 mL HCl in 50 mL nitrogen-bubbled water was gradually added through the third funnel. The orange-red microcrystalline precipitate which formed immediately was filtered, dried under stream of nitrogen, and stored sealed under nitrogen at 273 K (an additional amount of the product could be obtained by adding excess acid, but it is usually contaminated with a yellow side-product). Extensive crystallization experiments including slow evaporation, diffusion (gel, gas−liquid and liquid−liquid) and seeding, either in air or in inert atmosphere, were performed in order to crystallize the crude product, which always returned red material of poor crystallinity, contaminated with a single yellow phase with amorphous appearance. Column chromatography (silica gel, elution with increasing amount of MeOH in toluene) provided more of the major yellow phase in addition to $OxyLH_2$. Bundles of red needles of pure $OxyLH_2$ were obtained by slow evaporation at room temperature from MeOH, THF and $CHCl_3$, and could be easily separated mechanically from the yellow side product, which crystallized first. Small amount of regular red needle crystals suitable for XRD measurement were obtained by slow evaporation of isopropanol solution of the crude product over $CaCl_2$ under nitrogen at room temperature. Anal. Calcd. for $C_{10}H_6N_2O_2S_2$: C, 47.98; H, 2.42; N, 11.19. Found: C, 48.00; H, 2.40; N, 11.00. EIMS, *m/z*: M^+ calcd for $C_{10}H_6N_2O_2S_2$, 250.30; found, 250. IR (KBr, cm^{-1}): 3292, 3110, 2936, 1560, 1458, 1442, 1213, 1172, 1004, 934, 884, 842, 805. ^{1}H NMR (600 MHz, DMSO-d_6, δ): 6.51 (s, 1H, C5H), 7.02 (q, 1H, C5′H), 7.43 (d, 1H, C4′H), 7.86 (d, 1H, C7′H), 10.06 (s, 1H, $O_{phenol}H$), 11.01 (s, 1H, O_{enol}).

Synthesis of 5-Methyloxyluciferin ($MOxyLH_2$). $MOxyLH_2$ was prepared at room temperature, under similar conditions (dark, nitrogen) to that described above for $OxyLH_2$.[23] A well-mixed solution of 2-cyano-6-hydroxybenzothiazole (300 mg) in 39 mL MeOH and 33 mL water were sequentially added from the first septum-closed funnel to the reaction flask, followed by addition of 450 μL ethyl 2-mercaptopropionate. The mixture was stirred, and then aqueous solution (25 mL) of 1 g NaOH was added, and the stirring was continued during 2.5 h. The solution was acidified with diluted aqueous solution of HCl, affording large amount of yellow microcrystalline precipitate, which was filtered, dried with nitrogen

above calcium chloride and stored under nitrogen at 270 K. Long prismatic pale yellow crystals of $MOxyLH_2$ suitable for X-ray diffraction were obtained by slow evaporation of MeOH solution of the compound over $CaCl_2$ under static atmosphere of nitrogen. Anal. Calcd. for $C_{11}H_8N_2O_2S_2$: C, 49.98; H, 3.05; N, 10.60. Found: C, 49.66; H, 3.09; N, 10.50. EIMS, *m/z*: M^+ calcd for $C_{11}H_8N_2O_2S_2$, 264.32; found, 264. IR (KBr, cm^{-1}): 3140, 1608, 1578, 1503, 1456, 1407, 1324, 1284, 1231, 1167, 1122, 948, 891, 837. ^1H NMR (600 MHz, DMSO-d_6, δ): 2.26 (s, 3H, CH_3), 7.00 (t, 1H, C5′H), 7.40 (d, 1H, C4′H), 7.84 (d, 1H, C7′H), 10.02 (s, 1H, $O_{phenol}H$), 10.75 (s, 1H, O_{enol}).

Characterization. The elemental analyses were performed with standard methods and the mass spectra (electron ionization method) were recorded at Osaka University analytical center. The FTIR spectra were recorded from KBr pellets with Nicolet iS10 instrument (Thermo Scientific). The ^1H NMR spectra were recorded with Varian UNITY INOVA600 instrument operating at 600 MHz. To extract the proton, *t*-BuOK in DMSO or standard solution of Bu_4NOH were added to the solutions.

Steady-State Spectra. The absorption UV−visible spectra were recorded with Hewlett-Packard 8453 diode-array spectrophotometer. The fluorescence spectra were recorded on Shimadzu spectrofluorophotometer (RF-5300PC). The absorption and emission spectra of pure $OxyLH_2$ in water were recorded from saturated aqueous solution in nitrogen-purged phosphate buffer. The experimental conditions for spectral measurement in organic solvents are detailed in the Results and Discussion section. All solvents (acetone, ACN, CH_2Cl_2, CH_3Cl, MeOH, H_2O, DMSO) were deoxygenated by bubbling with high-purity argon or nitrogen. To study the solvent effects on the absorption and emission spectra of $OxyLH_2$ and its ions, stock solutions of Bu_4NOH (10 mM) and $OxyLH_2$ (10 mM) were prepared in DMSO. The stock solution of $OxyLH_2$ (10 μL) was diluted in the respective solvent (2 mL) to prepare working solutions. Single or double volumes of the base solution were added for deprotonation. In another set of experiments, stock solutions of $OxyLH_2$ or $MOxyLH_2$ were prepared in DMSO (2 mL, $5\cdot10^{-5}$ M) and used to record the absorption, fluorescence spectra and fluorescence decay curves. For deprotonation, a solution of *t*-BuOK (100 μL, 10 mM) in DMSO was added. The absorption and fluorescence spectra were also measured in CH_2Cl_2, $CHCl_3$, MeOH, ACN, acetone, water, and phosphate-buffered aqueous solution at pH 7.8. In case of CH_2Cl_2, $CHCl_3$, ACN and water, saturated solutions of $OxyLH_2$ and $MOxyLH_2$ were used. In MeOH and acetone, the concentrations were adjusted to absorbance of 0.1−0.2. For recording the pH-dependent absorption spectra of $OxyLH_2$, a series of nitrogen-purged aqueous solutions of phosphate buffer (pH 6−9) were prepared and the pH was recorded. Saturated solution of $OxyLH_2$ in water (1 mL) was added to the phosphate buffer (1 mL) and the absorption spectra were recorded immediately.

Time-Resolved Spectra and Fluorescence Lifetimes and Fluorescence Yields. The fluorescence decay profiles were recorded with Photon Technology GL-3300 instrument (nitrogen laser), coupled with a Photon Technology International GL-302 (pigment laser), and a nitrogen laser-pumped dye laser system equipped with four-channel digital delay/pulse generator (Standard Research System Inc. DG535) and motor driver (Photon Technology International MD-5020). The excitation wavelengths were selected between 385 nm when using the POPOP dye (Wako) and 400−440 nm when using 1,4-bis(4-methyl-5-phenyl-2-oxazolyl)benzene (Dojindo). All spectra were recorded under nitrogen, with solutions prepared from nitrogen-bubbled spectral grade solvents. Fluorescence quantum yields of dye thin films were determined by the photoluminescence (PL) method using a Hamamatsu DynaSpect C 9920−02 absolute PL quantum yield measurement system, which is made up of an excitation light source (150 W xenon lamp), a monochromator, an integrating sphere and a multichannel spectrometer.

Bioluminescence Spectra. To record the spectra and fluorescence lifetimes of $OxyLH_2$ and AMP complexed with luciferase (products), saturated solution of $OxyLH_2$ was prepared in aqueous phosphate buffer containing 3.2 mM AMP and 10 mM $MgSO_4$, to which an equal volume of freshly prepared solution of the protein was added to 1 mg mL^{-1}. The excitation wavelength was 400 nm. For the emission spectra of $OxyLH_2$ produced *in situ* in luciferase, initially, a solution containing 100 mM sodium phosphate buffer (pH 7.8), 0.6 mM luciferin, 3.2 mM ATP and 10 mM $MgSO_4$ in nitrogen-purged water was prepared and used to record the background spectra. An equal volume of freshly prepared solution of *L. cruciata* luciferase in aqueous buffer was injected, and the emission spectrum was quickly recorded. The change of spectrum was monitored at certain time intervals after the mixing of the reactants.

X-Ray Diffraction. The single crystal X-ray diffraction data were collected at room temperature in ω-scan mode with APEX2 diffractometer (Bruker AXS),[59] using Mo Kα X-rays obtained from a rotating anode source, a confocal multilayer X-ray mirror as monochromator and a CCD as area detector. The integrated and scaled data[59] were empirically corrected for absorption effects with SADABS.[60] The structures were solved by using direct methods[61] and refined[62] on F_0^2 with SHELXL. All non-hydrogen atoms were treated anisotropically and all aromatic hydrogen atoms were included as riding bodies.

Theoretical Calculations. The molecular geometries were optimized[63] within the Density Functional Theory[64,65] and the stationary points were checked for minima by subsequent harmonic vibrational analyses. All calculations were performed with the Gaussian03 program suite.[66]

Acknowledgment. This work was supported by a Global COE program, "the Global Education and Research Center for Bio-Environmental Chemistry" from the Ministry of Education, Culture, Sports, Science and Technology, Japan and KOSEF/MEST through WCU project (R31-2008-000-10010-0).

Supporting Information Available: Solid-state absorption and emission spectra, ^1H NMR spectra, UV−visible spectra of $MOxyLH_2$, crystallographic data, complete ref 66, and structures in CIF format. This material is available free of charge via the Internet at http://pubs.acs.org.

JA904309Q

(59) *APEX2 (ver. 2.1−4) and SAINT (ver. 7.34A)*; Bruker AXS Inc.: Madison, WI, 2007.
(60) Sheldrick, G. M. *SADABS*; University of Göttingen: Göttingen, Germany, 1996.
(61) (a) Sheldrick, G. M. *Acta Crystallogr.* **2008**, *A64*, 112. (b) Altomare, A.; Cascarano, G.; Giacovazzo, C.; Guagliardi, A.; Burla, M. C.; Polidori, G.; Camalli, M. *J. Appl. Crystallogr.* **1994**, *27*, 435.
(62) Sheldrick, G. M. E. *SHELXL-97*; University of Göttingen: Göttingen, Germany, 1997.
(63) Schlegel, H. B. *J. Comput. Chem.* **1982**, *3*, 214.
(64) Lapinski, L.; Rostkowska, H.; Nowak, M. J.; Kwiatkowski, J. S.; Leszczynski, J. *Vib. Spectrosc.* **1996**, *13*, 23.
(65) Hess, B. A.; Schaad, L. J.; Ćarsky, P.; Zahradník, R. *Chem. Rev.* **1986**, *86*, 709.
(66) Frisch, M. J.; et al. *Gaussian03*; Gaussian, Inc.: Wallingford, CT, 2004 (the complete reference is provided as Supporting Information).
(67) Wolff, S. K.; Grimwood, D. J.; McKinnon, J. J.; Jayatilaka, D.; Spackman, M. A. *CrystalExplorer*; University of Western Australia: Perth, Australia, 2007.