Bioluminescence

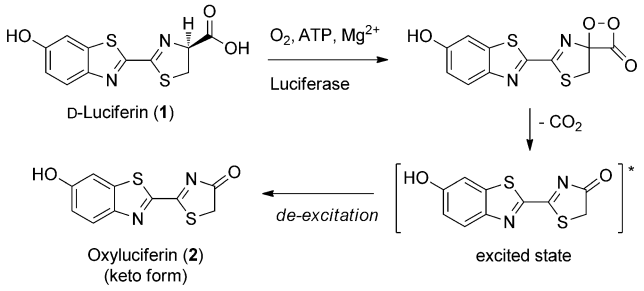
DOI: 10.1002/anie.201307972

Why is Firefly Oxyluciferin a Notoriously Labile Substance?**

Oleg V. Maltsev, Naba K. Nath, Panče Naumov, and Lukas Hintermann**

Abstract: *The chemistry of firefly bioluminescence is important for numerous applications in biochemistry and analytical chemistry. The emitter of this bioluminescent system, firefly oxyluciferin, is difficult to handle. The cause of its lability was clarified while its synthesis was reinvestigated. A side product was identified and characterized by NMR spectroscopy and X-ray crystallography. The reason for the lability of oxyluciferin is now ascribed to autodimerization of the coexisting enol and keto forms in a Mannich-type reaction.*

The fascinating phenomenon of firefly bioluminescence has been thoroughly investigated and finds application in the analysis of ATP and reactive oxygen species and, through suitable relay strategies, a multitude of bioanalytical imaging and monitoring techniques.[1] The basic chemistry of the process involves a reaction of firefly luciferin (**1**) with oxygen, ATP, and Mg^{2+} catalyzed by firefly luciferase to generate the emitter firefly oxyluciferin (**2**) in its excited state (Scheme 1).[2] The reaction product **2** is kinetically labile[3] under seemingly mild conditions,[4–7] and apparently highly sensitive to autoxidation.[6] Consequently, isolation of this material from natural sources is not feasible,[7–10] its chemical synthesis capricious,[11,12] and post-synthesis purification precluded.[6,12] Shimomura has stated that firefly oxyluciferin (**2**) "*[is] an extremely unstable compound; it has never been isolated in a completely pure form*".[2] Studies of the actual emitter **2** have remained scarce[10,12–14] and fundamental work has often been carried out on model compounds instead,[4,5,15] or in silico.[16]

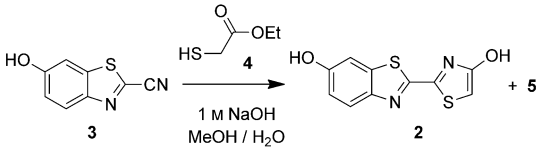


Scheme 1. Mechanism of firefly bioluminescence.

With recent interest in the color tuning mechanism of firefly bioluminescence[12,17,18] or potential applications of modified emission wavelength bioluminescence,[1b] a better understanding of the chemistry and spectroscopy of oxyluciferin (**2**) is required.[14] One of us has reported the first X-ray crystal structure of **2**,[12] and together with Sliwa et al., we have elucidated its triple acid/base and keto/enol equilibrium (pK_{a1}, pK_{a2}, pK_E) in aqueous buffers.[18] Here, we present a standardized chemical synthesis of **2** and clarify the reason for its surprising lability by identifying a key decomposition product and the mechanism of decomposition. We also detect for the first time the keto form of **2** by NMR spectroscopy in neutral solution.

Pure samples of oxyluciferin (**2**) were required for our recent spectroscopic studies.[18] By following the published synthesis from 2-cyano-6-hydroxybenzothiazole (**3**) and ethyl 2-mercaptoacetate (**4**) in aqueous/alcoholic base (Table 1),[6,11,12] we obtained a crude precipitate of oxyluciferin (**2**) containing a side product **5** in varying amounts (Table 1, entries 1, 2). In line with earlier reports, purification of **2** by either recrystallization or chromatography on SiO_2

Table 1: Optimization of the synthesis of oxyluciferin (**2**).[a]



Entry	NaOH [eqiuv]	T [°C][b]	t [min]	Yield [%][c]	**2**/**5**[d]
1[e]	0.5	RT	3	60	2.5:1
2[e]	0.5	3	3	44	>99:1
3	0.5	4	3	44	>99:1
4	0.25	4	3	60[f]	>99:1
5	1.0	4	3	77	>99:1
6	1.0	−1	3	54[f]	>99:1
7	1.0	4	6	79 (54)[g]	33:1
8	1.0	RT	30	93	0.72:1

[a] Unless otherwise specified, reactions were performed in air and quenched with HCl (aq) after the indicated reaction time. [b] Internal reaction temperature after mixing of reactants. [c] Yield of precipitated product calculated as **2**. [d] Molar ratio of **2**/**5** in the crude by ^{1}H NMR. [e] Reaction performed under argon. [f] Product contaminated with **3**. [g] Yield of pure **2** after column chromatography in brackets.

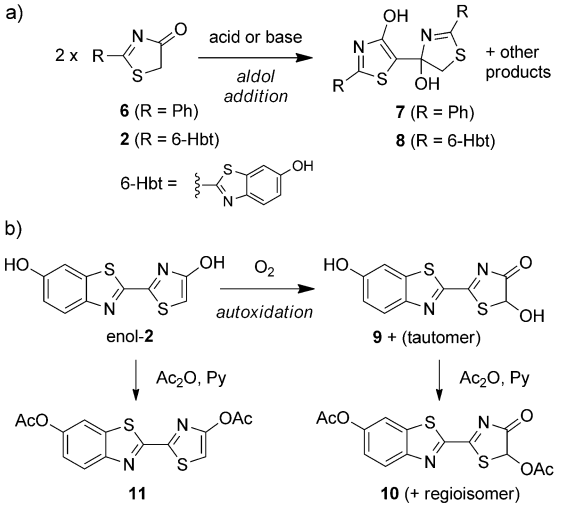
[*] Dr. O. V. Maltsev, Prof. Dr. L. Hintermann
Department Chemie, Technische Universität München
Lichtenbergstrasse 4, 85748 Garching b. München (Germany)
E-mail: lukas.hintermann@tum.de
Dr. N. K. Nath, Prof. Dr. P. Naumov
New York University Abu Dhabi
P.O. Box 129188 Abu Dhabi (United Arab Emirates)
E-mail: pance.naumov@nyu.edu

[**] Support by the Human Frontier Science Program (project RGY0081/2011, "Excited-State Structure of the Emitter and Color-Tuning Mechanism of the Firefly Bioluminescence") is acknowledged. We thank Profs. A. Wakamiya, Y. Murata, and N. Sato (Kyoto University) for making the single-crystal X-ray diffractometer available and for their help with the data collection of **14**. We thank Dr. S. C. Sahoo for collecting the single-crystal X-ray data for **13**.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201307972.

was precluded by facile decomposition.[6,12] To solve this problem, we wished to identify side product **5** and then try to suppress its formation in the synthesis.

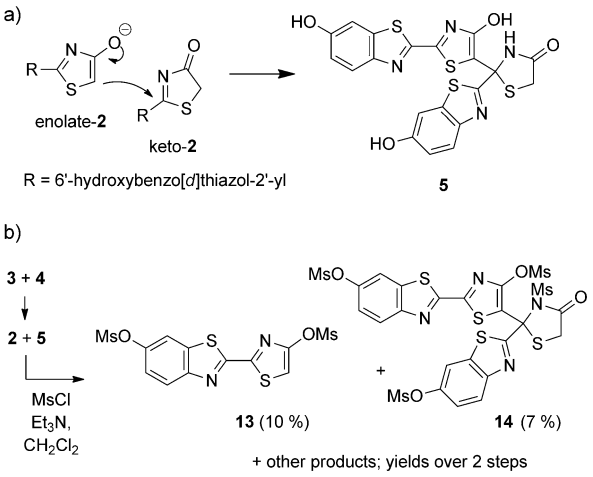
Two hypotheses have already been voiced to explain the lability of oxyluciferin (Scheme 2): White and co-workers proposed[4] that decomposition of **2** mirrors that of 2-phenyl-1,3-thiazol-4-one (**6**), the dimerization of which to **7** had been reported by Jensen et al.[19] By analogy, **5** could be aldol adduct **8**, or derived from it (Scheme 2a). Suzuki and Goto



Scheme 2. Proposed reasons for the lability of oxyluciferin (**2**): a) decomposition by aldol dimerization; b) by autoxidation.

reported that **2** readily autoxidizes to generate dioxyluciferin (**9**),[6,20] which was further characterized by conversion to a mixture of isomeric acetates **10** (Scheme 2b).[6,7,11]

The spectral data of our side product **5** were assessed against both hypotheses: LC–HRMS displayed a single molecular ion peak at m/z 500.9812 ($[M+H]^+$). This observation immediately excludes monomeric structures like **9** based on the Suzuki/Goto hypothesis. On the other hand, the ^{1}H NMR spectrum in [D$_6$]DMSO displayed two sets of benzothiazole signals, two doublets for diastereotopic hydrogen atoms within a CH$_2$ group ($\delta = 3.83$ and 3.90 ppm, $^2J_{H,H} = 15.3$ Hz), and four characteristic low-field singlets in the phenol/enol OH region between $\delta = 9.9$ and 11.8 ppm. These characteristics hinted at structure **8**, with a minor fault: the latter should have displayed three rather than four low-field singlets for OH groups! Acetylation of crude oxyluciferin (a mixture of **2**+**5**) gave a chromatographically separable mixture of oxyluciferin diacetate (**11**) and the new compound **12** (not shown), which must be a tetraacetate of **5**, based on mass (m/z 669.0233, $[M+H]^+$) and ^{1}H NMR data. Curiously, the ^{1}H NMR spectrum of tetraacetate **12** coincided with that reported for a mixture of dioxyluciferin diacetates **10**[6] obtained by acetylation of autoxidation product **9** (Scheme 2b).[20] So far, our data displayed partial agreement as well as discrepancy with both hypotheses on the nature of the decomposition products of oxyluciferin. However, we realized that there might be an alternative dimerization mode, consisting in the addition of an enolate of **2** to an imino group of another molecule of **2** in the keto tautomeric form (Scheme 3a): the resulting structure **5** accommodates the four observed low-field singlets (2 phenol-OH, 1 enol-OH, 1 amide-NH) and the diastereotopic CH$_2$ protons.



Scheme 3. a) Dimerization of oxyluciferin **2** to bisoxyluciferin **5**. b) Derivatization of crude oxyluciferin.

Structure **5** was in agreement with all 1D and 2D NMR data obtained, but unequivocal assignment was hampered by the ubiquity of quaternary carbon atoms and lack of suitable ^{1}H NMR signals for correlation. Complementary structural proof by X-ray crystallography was desirable. Derivatization of the mixture of **2** and **5** by mesylation gave oxyluciferin dimesylate **13** and bisoxyluciferin tetramesylate **14**,[21] both of which were amenable to X-ray crystal structure analysis, which corroborated the NMR findings (Figure 1).

The molecular structure of the oxyluciferin core is conserved in the crystal of **13**.[22] Apart of the mesylate

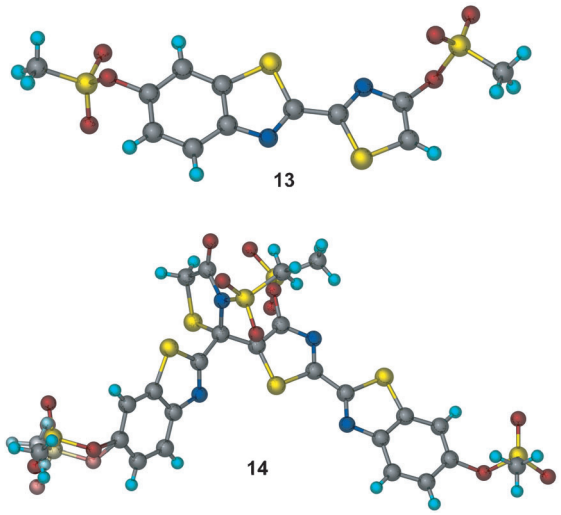


Figure 1. X-ray crystal structures of oxyluciferin dimesylate **13** and bisoxyluciferin tetramesylate **14**.[32] C gray, O red, H cyan, S yellow, N blue.

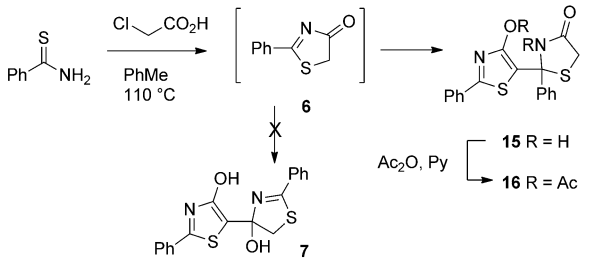
groups, the molecule is nearly planar. In the absence of strong hydrogen-bond donors, the molecules are connected by weak C–H···O/N hydrogen bonds and π···π interactions. In the nonsymmetric dimer **14**, two oxyluciferin moieties are joined between the *ipso* carbon atom (C-2) of the thiazole ring of one unit, and C-5 of another unit. The central thiazolidinone ring of **14** is planar and nearly perpendicular to the planes of both oxyluciferyl and the benzothiazolyl substituents at its quaternary C-2. The V-shaped molecules are packed by C–H···O hydrogen bonds and π···π interactions into centrosymmetric supramolecular dimers.

With the new compound firmly established as bisoxyluciferin (**5**),[23] the capricious nature of the synthesis of **2** becomes evident: the mildly basic reaction conditions in the condensation of **3** and **4** to oxyluciferin (**2**) also favor the dimerization of **2** to **5**! Initial synthesis attempts had suffered from insufficient cooling, and the reaction product was contaminated with **5** (Table 1, entry 1).[24] By standardizing the reaction conditions and strictly working in the cold, the precipitated **2** was almost pure, though obtained in lower yield (Table 1, entry 2). Contrary to earlier claims,[6,12] the presence or absence of oxygen had no influence on the outcome of the reaction (entries 2 versus 3); all ensuing syntheses were carried out in air. Coexistence of keto-**2** and enol-**2** or their conjugated bases causes the lability of oxyluciferin (**2**), particularly at high concentration!

Our recent species population analysis of **2** in phosphate buffers had shown that keto forms are significant below pH 8, but contribute very little above pH 9.[18] In line with this, the synthesis of **2** is improved when using more base, while retaining a stringent temperature and time protocol (Table 1, entry 5). At longer reaction times, the yield of **2** increases, and some **5** is coprecipitated (Table 1, entry 7). Interestingly, we find that oxyluciferin can be purified by chromatography on SiO_2, provided that a little acetic acid (1 vol%) is added to the eluent. The reaction conditions of entry 5 are suitable for obtaining pure samples of **2** (>99 mol%)[25] by direct precipitation, whereas the conditions of entry 7 optimize the yield of crude material, which can be chromatographically purified. Much bisoxyluciferin (**5**) is formed at prolonged reaction times without cooling (Table 1, entry 8).

Several assignments in the older literature must be modified: first, oxyluciferin is not sensitive to autoxidation, and the presumed autoxidation product "dioxyluciferin" **9** is in fact bisoxyluciferin **5**. This removes the paradox that oxyluciferin, which is generated by oxidation with O_2 and must be regenerated in vivo,[26] should be highly sensitive to oxygen. Second, the earlier described mixture of dioxyluciferin diacetates **10**[6] is now recognized as bisoxyluciferin tetraacetate **12**. Furthermore, the long known lability of 1,3-thiazol-4-ones (or the tautomeric 4-hydroxy-1,3-thiazoles)[27,28] is ascribed to a Mannich-type dimerization. This puts doubt on the structure of Holmberg's thiazolone, once thought to be thiazolone **6**,[29] but later reassigned to **7**[19] (Scheme 4): indeed, we find that Holmberg's thiazolone possesses structure **15**,[30] and its diacetate is **16**.[31]

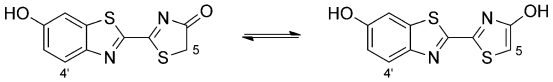
Having access to pure samples of **2** has allowed us to investigate another paradox in the chemistry of firefly bioluminescence: while the keto form is postulated as product



Scheme 4. Structure of the reaction product of thiobenzamide and chloroacetic acid (Holmberg's thiazolone): original proposal (**6**), first revision (**7**),[19] and our revision (**15**).

of the luminous reaction (Scheme 1),[2] NMR studies have so far exclusively detected the enol form of **2**.[6,10] We have now collected quantitative data of the keto–enol equilibrium in several solvents (Table 2).

Table 2: Time-dependent keto–enol equilibrium of oxyluciferin (**2**).[a]



Solvent	10 min	3 h	1 day	5.5 days
$[D_6]$DMSO	3.8:96.2	0.7:99.3	0.7:99.3	0.8:99.2
CD_3OD[b]	2.1:97.9	3.3:96.7	3.3:96.7	–[c]
$[D_6]$acetone	10.0:90.0	11.5:88.5	15.4:84.6	–[c]
CD_3CN	36.5:63.5	38.3:61.7	42.0:58.0	–[c]

[a] The keto/enol ratio was determined at ambient temperature by integration of the ^{1}H NMR signals for H-4′. Concentration 10–12 mM, or lower (saturated solution) in CD_3CN. [b] Deuterium incorporation at H-5: 53% (10 min), >99% (4 h). [c] Signals of the keto form were covered by those of decomposition products.

Tautomerization of **2** is slow in neutral, aprotic solvents, requiring hours to days to reach equilibrium. Dipolar aprotic solvents with weaker hydrogen-bond-acceptor character show higher levels of keto form, with over 40% abundance in acetonitrile. The enol form is found in single crystals, where it forms hydrogen-bonded dimers.[12] The relative importance of the keto form in methanol (ca. 3%) correlates with the lability of **2** in the aqueous-alcoholic medium of its synthesis. In terms of abundance, either tautomer must be considered a potential emitter in bioluminescence, where tautomerism in the excited state is conceivable.[14b]

In conclusion, we have reinvestigated the capricious synthesis of firefly oxyluciferin (**2**). In the process, we have clarified the cause of its lability, which is due to the presence of both keto and enol forms in protic solution that permits a Mannich-type dimerization. Earlier hypotheses suggesting aldol condensations or autoxidation of **2** can be rejected. Oxyluciferin is obtained in ≥99% purity when following a standardized synthesis protocol. These results will be of importance in future studies on the nature of the light-emitting species of the firefly bioluminescence.

Received: September 10, 2013
Published online: November 26, 2013

Keywords: bioorganic chemistry · heterocycles · luminescence · natural products · structure elucidation

[1] Recent reviews: a) T. Ozawa, H. Yoshimura, S. B. Kim, *Anal. Chem.* **2013**, *85*, 590; b) J. Li, L. Chen, L. Du, M. Li, *Chem. Soc. Rev.* **2013**, *42*, 662; c) M. Kelkar, A. De, *Curr. Opin. Pharmacol.* **2012**, *12*, 592; d) M. Keyaerts, V. Caveliers, T. Lahoutte, *Trends Mol. Med.* **2012**, *18*, 164; e) A. Roda, M. Guardigli, *Anal. Bioanal. Chem.* **2012**, *402*, 69.

[2] O. Shimomura, *Bioluminescence: Chemical Principles and Methods*, rev. ed., World Scientific, Singapore, **2012**.

[3] The term "labile" is used in the sense of "reactive", irrespective of thermodynamic stability. Oxyluciferin is a good example for a labile species (depending on conditions) that is not fundamentally unstable. See: IUPAC, *Compendium of Chemical Terminology (the "Gold Book")*, 2nd ed. (Eds.: A. D. McNaught, A. Wilkinson), Blackwell Scientific Publications, Oxford, **1997**.

[4] E. H. White, D. F. Roswell, *Photochem. Photobiol.* **1991**, *53*, 131.

[5] E. H. White, E. Rapaport, H. H. Seliger, T. A. Hopkins, *Bioorg. Chem.* **1971**, *1*, 92.

[6] a) N. Suzuki, M. Sato, K. Okada, T. Goto, *Tetrahedron* **1972**, *28*, 4065; b) N. Suzuki, M. Sato, K. Nishikawa, T. Goto, *Tetrahedron Lett.* **1969**, *10*, 4683.

[7] N. Suzuki, T. Goto, *Tetrahedron* **1972**, *28*, 4075.

[8] B. Bitler, W. D. McElroy, *Arch. Biochem. Biophys.* **1957**, *72*, 358.

[9] N. Suzuki, T. Goto, *Tetrahedron Lett.* **1971**, *12*, 2021.

[10] J. C. G. Esteves da Silva, J. M. C. S. Magalhaes, R. Fontes, *Tetrahedron Lett.* **2001**, *42*, 8173.

[11] N. Suzuki, T. Goto, *Agric. Biol. Chem.* **1972**, *36*, 2213.

[12] P. Naumov, Y. Ozawa, K. Ohkubo, S. Fukuzumi, *J. Am. Chem. Soc.* **2009**, *131*, 11590.

[13] O. A. Gandelman, L. Y. Brovko, N. N. Ugarova, A. Y. Chikishev, A. P. Shkurimov, *J. Photochem. Photobiol. B* **1993**, *19*, 187.

[14] a) K. Støchkel, C. N. Hansen, J. Houmøller, L. M. Nielsen, K. Anggara, M. Linares, P. Norman, F. Nogueira, O. V. Maltsev, L. Hintermann, S. B. Nielsen, P. Naumov, B. F. Milne, *J. Am. Chem. Soc.* **2013**, *135*, 6485; b) K. M. Solntsev, S. P. Laptenok, P. Naumov, *J. Am. Chem. Soc.* **2012**, *134*, 16452; c) Y. Erez, I. Presiado, R. Gepshtein, L. Pinto da Silva, J. C. G. Esteves da Silva, D. Huppert, *J. Phys. Chem. A* **2012**, *116*, 7452; d) P. Naumov, M. Kochunnoonny, *J. Am. Chem. Soc.* **2010**, *132*, 11566; e) C. Ribeiro, J. C. G. Esteves da Silva, *Photochem. Photobiol. Sci.* **2008**, *7*, 1085.

[15] a) T. Hirano, H. Nagai, T. Matsuhashi, Y. Hasumi, S. Iwano, K. Ito, S. Maki, H. Niwa, V. R. Viviani, *Photochem. Photobiol. Sci.* **2012**, *11*, 1281; b) T. N. Vlasova, O. V. Leontieva, N. N. Ugarova, *Biochemistry* **2006**, *71*, 555; c) O. V. Leontieva, T. N. Vlasova, N. N. Ugarova, *Biochemistry* **2006**, *71*, 51.

[16] Recent examples: a) I. Navizet, D. Roca-Sanjuan, L. Yue, Y. J. Liu, N. Ferre, R. Lindh, *Photochem. Photobiol.* **2013**, *89*, 319; b) V. R. Viviani, D. T. Amaral, D. R. Neves, A. Simoes, F. G. C. Arnoldi, *Biochemistry* **2013**, *52*, 19; c) L. Pinto da Silva, J. C. G. Esteves da Silva, *Int. J. Quantum Chem.* **2013**, *113*, 45; d) M. Anselmi, S. Marocchi, M. Aschi, A. Amadei, *Chem. Phys.* **2012**, *392*, 205; e) Z. S. Li, C. G. Min, A. M. Ren, L. Y. Zou, Z. Xu, *J. Photochem. Photobiol. A* **2012**, *243*, 7.

[17] a) L. W. Chung, S. Hayashi, M. Lundberg, T. Nakatsu, H. Kato, K. Morokuma, *J. Am. Chem. Soc.* **2008**, *130*, 12880; b) T. Nakatsu, S. Ichiyama, J. Hiratake, A. Saldanha, N. Kobashi, K. Sakata, H. Kato, *Nature* **2006**, *440*, 372; c) B. R. Branchini, M. H. Murtiashaw, R. A. Magyar, N. C. Porter, M. C. Ruggiero, J. G. Stroh, *J. Am. Chem. Soc.* **2002**, *124*, 2112.

[18] M. Rebarz, B.-M. Kukovec, O. V. Maltsev, C. Ruckebusch, L. Hintermann, P. Naumov, M. Sliwa, *Chem. Sci.* **2013**, *4*, 3803.

[19] a) K. A. Jensen, I. Crossland, *Acta Chem. Scand.* **1963**, *17*, 144; b) S. Gronowitz, B. Mathiasson, R. Dahlbom, B. Holmberg, K. A. Jensen, *Acta Chem. Scand.* **1965**, *19*, 1215.

[20] Note that compound **9** was considered to be a mixture of tautomers, and compound **10** a mixture of regioisomers, to explain the doubling of some signals in their ^{1}H NMR spectra.[6]

[21] Besides other products; see the Supporting Information.

[22] A more detailed discussion of the structures and crystallographic data are provided in the Supporting Information.

[23] The name bisoxyluciferin is proposed for dimer **5**. The term dioxyluciferin is ambiguous since it has been used for the hypothetic autoxidation product **9** of oxyluciferin, where the prefix "dioxy" implies that it is a product of twofold oxidation of luciferin.

[24] The original procedure asked for ice cooling, but the importance of accurate internal temperature control was not evident.[6,11]

[25] Purity is expressed in mol% and based on all components detected by ^{1}H NMR. A remaining impurity of ≤ 1 mol% of mercaptoester **4** was visible. Samples purified by chromatography retained traces of solvents even after prolonged drying in vacuum. By coevaporation with suitable solvents (such as MeOH) it was possible to remove all other volatile components, except that solvent, see the Supporting Information for details.

[26] a) K. Gomi, N. Kajiyama, *J. Biol. Chem.* **2001**, *276*, 36508; b) K. Okada, H. Iio, I. Kubota, T. Goto, *Tetrahedron Lett.* **1974**, *15*, 2771.

[27] G. C. Barrett, *Tetrahedron* **1980**, *36*, 2023.

[28] The lability is restricted to derivatives devoid of substituents at C-5, cf.: E. Täuscher, R. Beckert, J. Fabian, A. Assumpção, H. Görls, *Tetrahedron Lett.* **2011**, *52*, 2292.

[29] B. Holmberg, *Ark. Kemi Mineral. Geol.* **1945**, *20A*, No. 1.

[30] A compound with this structure has been reported, without noting its identity to Holmberg's thiazolone: I. E. Kopta, *Tetrahedron Lett.* **1988**, *29*, 3765.

[31] These reassignments rest upon extended spectral analyses including high-resolution mass spectra, ^{1}H, ^{13}C, and HMQC NMR; furthermore, the spectral data are fully consistent with those of **5** and **12**. See the Supporting Information for details.

[32] CCDC-952719 (**13**) and 952720 (**14**) contain the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

