A novel-type luciferin from Siberian luminous earthworm *Fridericia heliota*: structure elucidation by spectral studies and total synthesis

Valentin N. Petushkov[a,b], Maxim A. Dubinnyi[c], Aleksandra S. Tsarkova[c], Natalja S. Rodionova[a,b], Mikhail S. Baranov[c], Vadim S. Kublitski[c], Osamu Shimomura[a,d] and Ilia V. Yampolsky[c,*]

[a] *Laboratory of Bioluminescent Biotechnologies, Institute of Fundamental Biology and Biotechnology, Siberian Federal University, pr. Svobodnyi, 79, Krasnoyarsk 660041, Russia*

[b] *Laboratory of Photobiology, Institute of Biophysics, Siberian Branch of the Russian Academy of Sciences, Akademgorodok, Krasnoyarsk 660036, Russia*

[c] *Institute of Bioorganic Chemistry, Russian Academy of Sciences, Miklukho-Maklaya 16/10, Moscow 117997, Russia*

[d] *Marine Biological Laboratory Woods Hole, MA 02543, U.S.A.*

ABSTRACT

We report structure elucidation and synthesis of the luciferin from the recently discovered luminous earthworm *Fridericia heliota*. This luciferin represents a key component of a novel ATP-dependent bioluminescence system. The UV, fluorescence, NMR and HRMS spectral studies were performed on 5 mkg of the isolated substance, and gave four isomeric structures, conforming with spectral data. These isomers were chemically synthesized and one of them was

found to produce light in the reaction with a protein extract from *Fridericia*. The novel luciferin was found to have an unusual deeply modified peptidic nature, implying an unprecedented mechanism of action.

Keywords:

Bioluminescence

Fridericia heliota

Luciferin

Total synthesis

Natural products

---

* Corresponding author. Tel./fax: +7 499 724 81 22.

E-mail address: ivyamp@ibch.ru (Ilia V. Yampolsky).

Bioluminescence is a fascinating phenomenon of visible light emission by live organisms. Hundreds of species of bioluminescent animals, fungi, protists and bacteria are known, possessing estimated 30 different chemical mechanisms, underlying generation of "cold light".[1] In all known cases, the energy required for light production is generated by oxidation of a small organic molecule - luciferin, facilitated by a specific enzyme - luciferase. Presently, structures of only seven natural luciferins are known (figure S1).[1] The latest structural characterization of luciferin from dinoflagellates dates back 25 years.[2]

Recently, a novel bioluminescent species has been discovered in Siberia.[3] *Fridericia heliota* is a small (~15 mm in length, 0.5 mm in diameter and ~2 mg in weight), white-yellowish oligochaete worm which inhabits forest soils and emits blue light (the luminescence maximum at 478 nm) when mechanically stimulated. The luminescence of *F. heliota* is located in the region of epidermal cells (figure 1B).

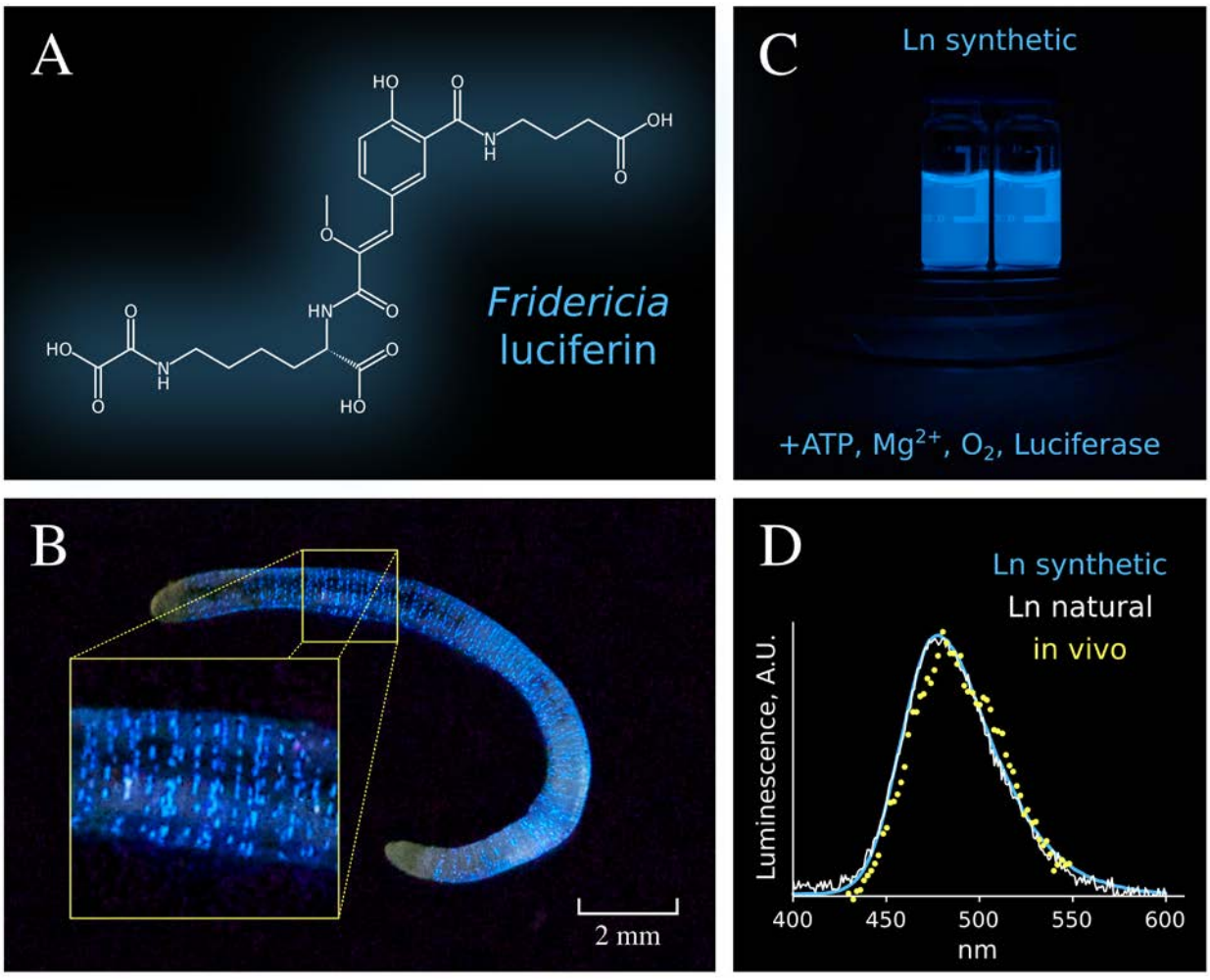


Figure 1. (A) Structure of *Fridericia* luciferin. (B) Bioluminescence of *Fridericia heliota*. The photograph is courtesy of Alexander Semenov (White Sea Biological Station, Biology Department of Lomonosov Moscow State University). (C) Luminescence of synthetic Fridericia

luciferin. (D) Comparison of *in vivo* bioluminescence spectra of worms, with *in vitro* bioluminescence spectra of natural and synthetic samples of luciferin.

There was a general concept about the common nature of luminescence of earthworms, which was based on the results of comparative studies of physiology and biochemistry of 12 species belonging to 6 genera (*Diplocardia, Diplotrema, Fletcherodrilus, Octochaetus, Pontodrilus,* and *Spenceriella*).[4] All these bioluminescent oligochaetes secrete luminescent liquid containing coelomic cells, in the granules of which the luminescence is localized. Bioluminescence of oligochaetes is characterized by a common feature: the involvement of hydrogen peroxide. Luciferin of *Diplocardia longa*, *N*-isovaleryl-3-amino-propanal, serves as a substrate for luciferases of all bioluminescent earthworms. In addition, *D. longa* luciferase is active in cross reactions with luciferins of other worms.[5]

The bioluminescence of potworms (Fam. Enchytraeidae) is known since 1838, but in lesser detail than that of megadriles, and is confined to the sole genera *Henlea* and *Fridericia*.[6]

We demonstrated the light production mechanism of *F. heliota* to be unique, since cross-reactions of luciferase or luciferin from *Fridericia* with luciferins and luciferases from other organisms, were all negative. We found five components to be essential for *F. heliota* luminescence: *Fridericia* luciferase, *Fridericia* luciferin, ATP, Mg^{2+} and atmospheric oxygen.[7,8]

Isolation and structural characterization of *Fridericia* luciferin was seriously hindered by the scarcity of the earthworms biomass (manual harvesting gave about 30 g/year), and the low content of luciferin (~0.1 μg per gram of wet biomass).[9] In the course of our extensive efforts aimed at purification of *Fridericia* luciferin, we isolated CompX - a component of *Fridericia*

biomass, similar to luciferin by chromatographic and UV spectral behavior.[10] Spectral studies supported by total synthesis revealed CompX to be an unusual derivative of tyrosine (figure 2).

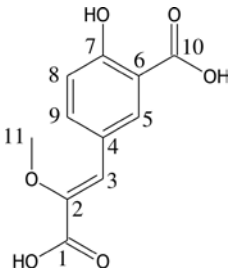


Figure 2. CompX, a structural analog of *Fridericia* luciferin, found in *Fridericia* biomass.

Total amount of luciferin, obtained form 90 g of biomass was 0.005 mg, which allowed us to obtain only the ^{1}H, COSY and partial ^{13}C-HSQC NMR spectra (figures S4-S6). These data revealed the following three fragments of luciferin structure: substituted CompX and residues of lysine and of gamma-aminobutyric acid (GABA). The scarcity of luciferin did not allow us to obtain 1D ^{13}C and HMBC spectra, which would reveal the connectivity of these three fragments and the presence of non-hydrogenated carbon atoms. Obviously, the most plausible way for these fragments to form a stable compound, while leaving all their C-H bonds intact (the requirement, imposed by available NMR data) is the formation of peptide bonds between some of their four carboxylic and three amino groups. Therefore, we performed the ^{1}H NMR titration of luciferin in the pH range 3.1 ÷ 7.5 in order to discriminate between the free carboxylic groups, and those forming peptide bonds with the amino groups of lysine and GABA residues. The pH-dependence of chemical shifts of protons, adjacent to titratable carboxyls, indicated that the carboxylic groups of lysine and GABA are free, whereas the two carboxyls of the CompX moiety are probably involved in peptide bonds (figure S7).

The HRMS spectra of luciferin showed a molecular ion with m/z=524.1851, for which the closest molecular formula is $C_{23}H_{30}N_3O_{11}^+$. The difference between this formula and a sum (CompX + Lysine + GABA - $2H_2O$) is C_2O_3. In this calculation, the loss of two water molecules is assumed to result from the formation of two peptide bonds between the three specified fragments. Also, the MS spectra showed two abundant peaks corresponding to the loss of CO_2 and of CO_2 and CO, but not CO alone (figure S8). All these data taken together suggest a monosubstituted oxalic acid residue to be a missing structural fragment of luciferin.

Four isomeric structures **1**-**4** (figure 3) were consistent with NMR and mass-spectroscopic data summarized above. These isomers differ only by the order of peptide bonds connecting the four residues, identified as the building blocks of *Fridericia* luciferin: CompX, lysine, GABA and oxalate. Derivatives of L-lysine and racemization-preventing conditions were used throughout all syntheses.

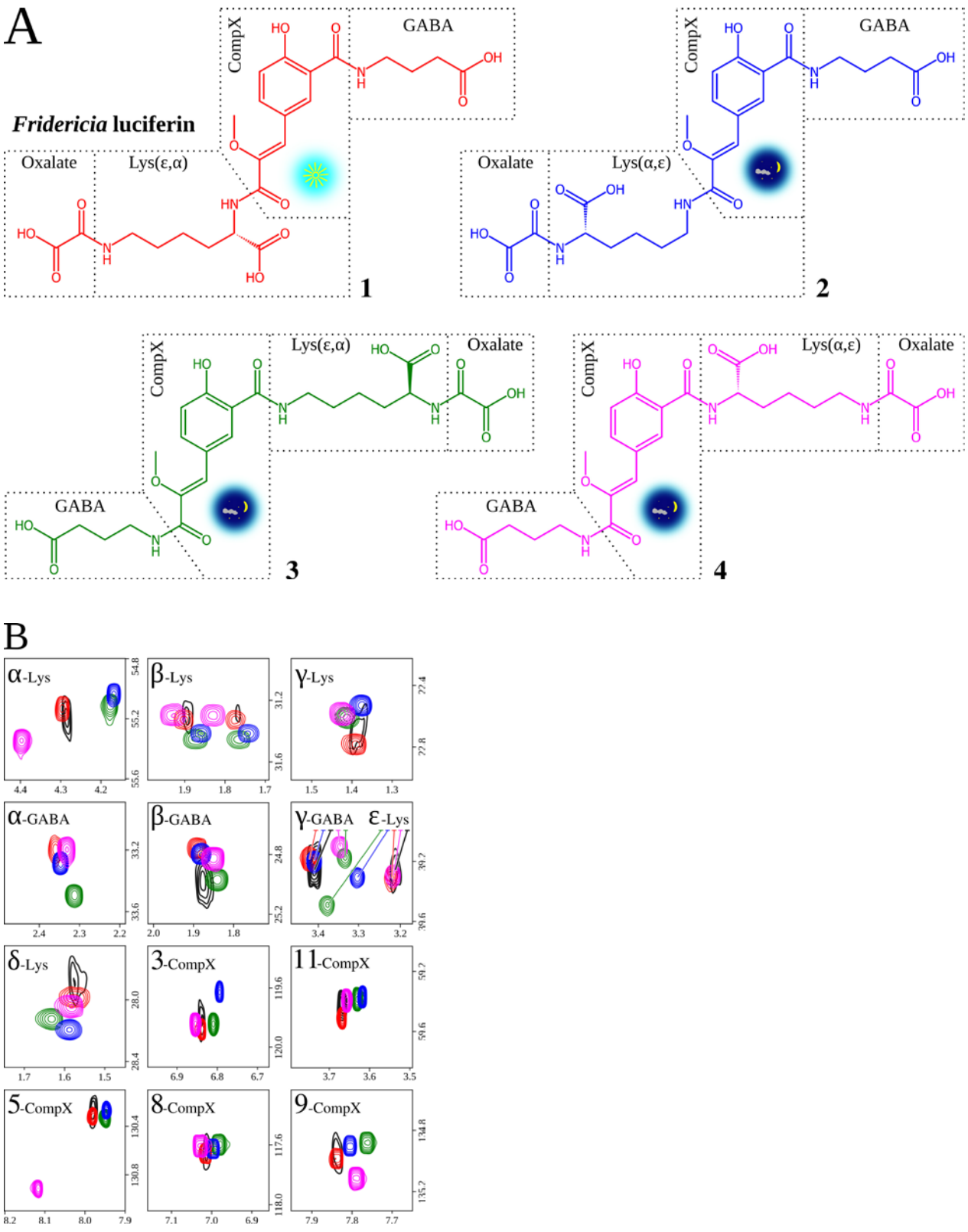


Figure 3. **(A)** Structures of synthetic isomeric peptides **1-4**. Only **1** produced light when mixed with *Fridericia* luciferase. **(B)** Comparison of selected fragments of ^{13}C-HSQC NMR spectra of

Fridericia luciferin and compounds **1**-**4** (D_2O, 30°C, pH 5.0). Colors of the peaks correspond with (A), peaks of natural luciferin are shown in black.

We synthesized all the isomeric peptides **1**-**4** and compared their NMR spectra with those of the natural sample (table S1). The synthesis started from CompX monomethyl ester[10] and utilized common peptide chemistry methods (schemes S1-S4). All synthetic compounds possessed similar NMR spectra in D_2O at pH 5.0. However, only **1** showed ^{1}H and ^{13}C chemical shifts completely matching those of the natural luciferin (figure 3B, table S1).

We explored the ability of synthetic compounds **1**-**4** to produce light upon addition to the crude *Fridericia* luciferase in the presence of ATP and $MgSO_4$. Only synthetic compound **1** showed luminescence in these conditions, with the luminescence spectrum (figure 1 C, D) and intensity-concentration dependence identical to those of the natural luciferin (figure S3).

Thus, we report structure of a novel luciferin which represents a key component of a novel ATP-dependent bioluminescence system from a Siberian earthworm *Fridericia heliota*. This luciferin probably represents a fundamentally novel chemical mechanism of luminescence, to be evaluated in the nearest future. Presumably, in the course of luminescent reaction an oxalate moiety is oxidized, while a fluorescent CompX moiety serves as a light emitter, in a way similar to the chemistry underlying the "glow stick" chemiluminescence toys. The role of CompX moiety as the light emitter is supported by close similarity of luciferin fluorescence emission spectrum with bioluminescence spectrum of *Fridericia heliota* (λ_{max} 466 and 480 nm, respectively).[10] Another interesting question arising is the biosynthetic route, leading to *Fridericia* luciferin. In a recent research by Clardy et al., a carbonylated analog of tyrosine, similar to CompX fragment, has been

identified as a result of genome screening for the members of ATP-grasp enzyme family.[11] Considering the nonribosomal peptidic nature of the novel luciferin, its biosynthesis might probably utilize the enzymes of this family. A set of unusual peptides containing CompX moiety is found in *F. heliota* biomass. One of such peptides (tyrosine-CompX-lysine derivative designated AsLn2) has been described in our recent paper.[10b] The structures of other members of this family and their role in luciferin metabolism are presently under exploration in our group. Our further efforts will be focused on structural characterization of luciferin biosynthetic precursors and oxyluciferin, evaluation of the role of ATP and on sequencing and cloning of *Fridericia* luciferase.

[1] O. Shimomura, *Bioluminescence: Chemical Principles and Methods*; World Scientific Publishing: Singapore, **2006**.

[2] H. Nakamura, Y. Kishi, O. Shimomura, D. Morse, J. W. Hastings, *J. Am. Chem. Soc.* **1989**, *111*, 7607-7611.

[3] E. Rota, N. T. Zalesskaja, N. S. Rodionova, V. N. Petushkov, *J. Zool.* **2003**, *260*, 291-299.

[4] J. E. Wampler, B. G. M. Jamieson, *Comp. Biochem. Physiol.***1980**, *66B*, 43-50.

[5] H. Ohtsuka, N. G. Rudie, J. E. Wampler, *Biochemistry* **1976**, *15*, 1001-1004.

[6] E. Rota, *Lights on the ground: A historical survey of light production in the Oligochaeta. In: "Bioluminescence in Focus– A Collection of Illuminating Essays"* Ed by VB Meyer-Rochow, Research Signpost, Kerala, India, **2009**.

[7] V. N. Petushkov, N. S. Rodionova, V. S. Bondar, *Dokl. Biochem. Biophys.* **2003**, *391*, 204.

[8] N. S. Rodionova, V. S. Bondar, V. N. Petushkov, *Dokl. Biochem. Biophys.* **2003**, *392*, 253.

[9] V. N. Petushkov, N. S. Rodionova, *J. Photochem. Photobiol. B* **2007**, *87*, 130-136.

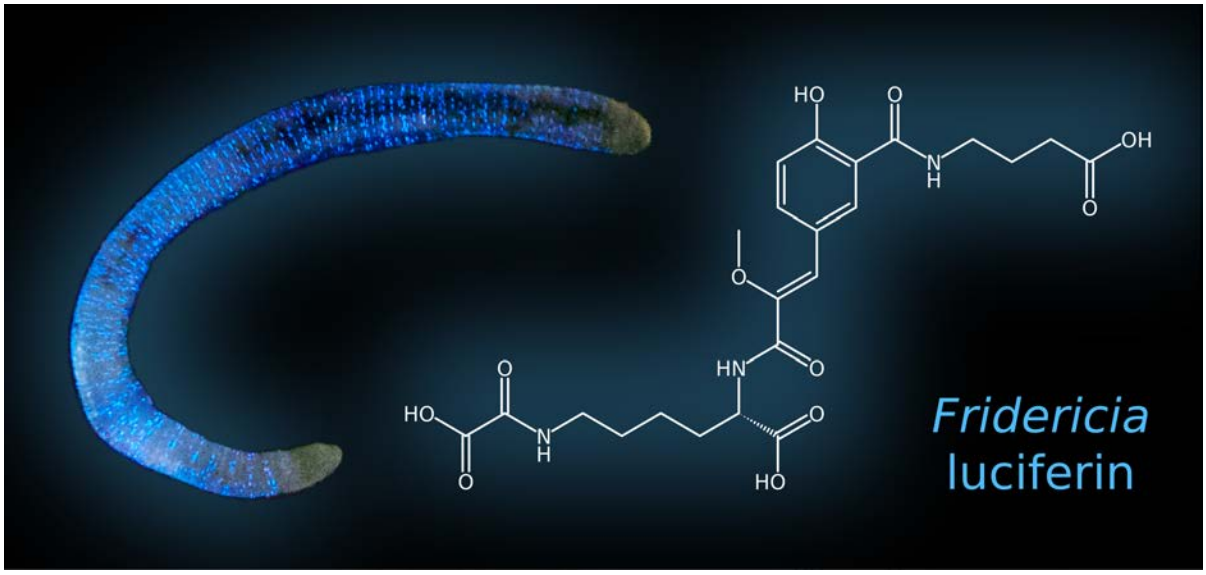
[10] a) V. N. Petushkov, A. S. Tsarkova, M. A. Dubinnyi, N. S. Rodionova, S. M. Marques, J. C. G. Esteves da Silva, O. Shimomura, I. V. Yampolsky, *Tetrahedron Lett.* **2014**, *55*, 460-462; b) V. N. Petushkov, M. A. Dubinnyi, N. S. Rodionova, K. D. Nadezhdin, S. M. Marques, J. C.G. Esteves da Silva, O. Shimomura, I. V. Yampolsky, *Tetrahedron Lett.* **2014**, 55, 463-465; c) S. M. Marques, V. N. Petushkov, N. S. Rodionova, J. C. G. Esteves da Silva, *J. Photochem. Photobiol. B* **2011**, *102*, 218–223.

[11] L. C. Blasiak, J. Clardy, *J. Am. Chem. Soc*. **2010**, *132*, 926-927.

Acknowledgements

The authors thank Dr. Alexander O. Chizhov for recording mass spectra and Dr. K. S. Mineev for NMR analysis of synthetic intermediates. We acknowledge support from the Program of the Government of the Russian Federation “Measures to attract leading scientists to Russian educational institutions” (grant no. 11. G34.31.0058), the programs MCB RAS, President of the Russian Federation “Leading science school” (grant 3951.2012.4) and the Russian Foundation for Basic Research (grant 14-03-01015). B.M.S. was supported by a stipend from the Program of the President of the Russian Federation.

Table of Contents



Luciferin from a novel ATP-dependent bioluminescent system is a peptide, formed by the residues of oxalic acid, lysine, modified tyrosine and gamma-aminobutyric acid.