Mendeleev Communications

Mendeleev Commun., 2027, **37**, in press

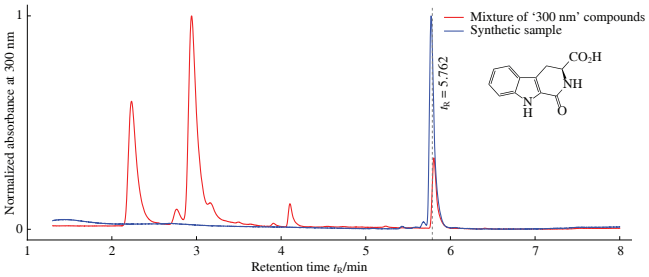
Identification of one more metabolite from the Siberian bioluminescent earthworm *Henlea* sp.

Anastasia N. Khokhlova,*[a,b] Renata I. Zagitova,[a,b] Valentin N. Petushkov,[c] Natalia S. Rodionova,[c] Matvey V. Vavilov,[a,b] Olga A. Belozerova,[a] Maxim A. Dubinnyi[a] and Ilia V. Yampolsky[a,d]

[a] *M. M. Shemyakin–Yu. A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 117997 Moscow, Russian Federation. E-mail: anastasia.n.khokhlova@gmail.com*
[b] *National Research University Higher School of Economics (HSE University), 101000 Moscow, Russian Federation*
[c] *Institute of Biophysics, Siberian Branch of the Russian Academy of Sciences, 660036 Krasnoyarsk, Russian Federation*
[d] *N. I. Pirogov Russian National Research Medical University, 117997 Moscow, Russian Federation*

DOI: 10.71267/mencom.8105

Tryptophan-2′-carboxylate derivative possessing pyrido-[3,4-*b*]indole core was discovered in the unique bioluminescent system of the *Henlea* earthworms, and its structure was confirmed by the independent synthesis. The synthetic compound exhibits the characteristic 300 nm absorption and can be proposed as a candidate for oxyluciferin or a related oxidized luciferin product.



Keywords: bioluminescence, metabolite, *Henlea* earthworm, tryptophan, pyrido[3,4-*b*]indole, natural products, mass spectrometry.

The study of the bioluminescent (BL) system of *Henlea* earthworms began in 2002 with the discovery of two new luminescent species, *H. petushkovi* and *H. rodionovae*, in the Krasnoyarsk and Baikal regions of Siberia, Russia.[1] The species were found in small quantities in Siberian taiga together with *Fridericia heliota* earthworms and in much larger amounts in private gardens.[2] The BL system turned out to be unique, consisting of an approximately 80 kDa luciferase,[1] a low-molecular-weight luciferin (hypothesized to be a tryptophan derivative), calcium ions,[3] oxygen, and a deazaflavin activator.[4] The fluorescence emission spectrum of the deazaflavin activator (ActH) and its structural analog ActS matches the spectrum of the BL reaction (maximum at 464 nm). Moreover, the deazaflavin activator, if present, shifts the BL emission maximum *in vitro* from violet (approximately 430 nm) to blue (464 nm) by a proposed natural BRET mechanism[5–7] and increases the intensity of the BL signal by a factor of 33.[4]

Earlier, we identified several low-molecular-weight compounds associated with luciferin: tryptophan-2′-carboxylate (T2C) and its derivatives.[5,8] All T2C compounds exhibit an absorption wavelength maximum at 300 nm (Figure 1); it differs from that of tryptophan (280 nm) due to the presence of a carboxy group at the position 2 of the indole ring. This property allows T2C compounds to be readily identified in the mixture of natural products in the worm extract. The structures of T2C and its *N*-acetyl derivative were unambiguously established by independent synthesis.[5] Our knowledge about the *Henlea* BL system is now consistent with the hypothesis that the actual yet unknown oxyluciferin should be among the set of at least eleven T2C derivatives in the *Henlea* extract.[5] In the present work, we performed the identification of the lowest-mass T2C compound by chemical synthesis and comparison of its chromatographic profile and HRMS spectrum with those of the natural one.

In our previous study,[5] we reported that chromatographic fractionation of worm extracts yielded fractions containing a chromophore with an absorption maximum at 300 nm (see Figure 1). The ESI-HRMS analysis of the fraction with t_R = 5.76 min revealed a molecular ion at $[M+H]^+$ 231.0762, consistent with the molecular formula $C_{12}H_{10}N_2O_3$ (calculated $[M+H]^+$ 231.0764),[9] which corresponds to the mass of tryptophan-2′-carboxylate without H_2O. The MS^2 analysis revealed a key fragment ion at *m*/*z* 185.0709, corresponding to the loss of the carboxy group (Figure 2). Based on its retention time in the hydrophobic region of the chromatogram, we hypothesized that the water elimination could occur intramolecularly between the amino group of tryptophan and the carboxy group at the position 2 of the indole ring to give compound **1** comprising pyrido[3,4-*b*]indole core [see Figure 1(*c*)]. However, the quantity of the purified natural substance was insufficient for NMR analysis. We hypothesized that its structure did not contain additional stereocenters or other sources of structural ambiguity.

Synthesis of compound **1** from 2-vinylaniline has been reported previously.[10] In this work, we employed a different approach (Scheme 1). Initial compound **2** was prepared according to a previously reported procedure[5] from commercially available N_α-Boc-L-tryptophan using the Pictet–Spengler reaction with formaldehyde.[11] After protection of the amino group,[12] the resulting compound **3** was subjected to oxidation with selenium dioxide[13] to yield aldehyde **4**, which was further

© 2027 Mendeleev Communications

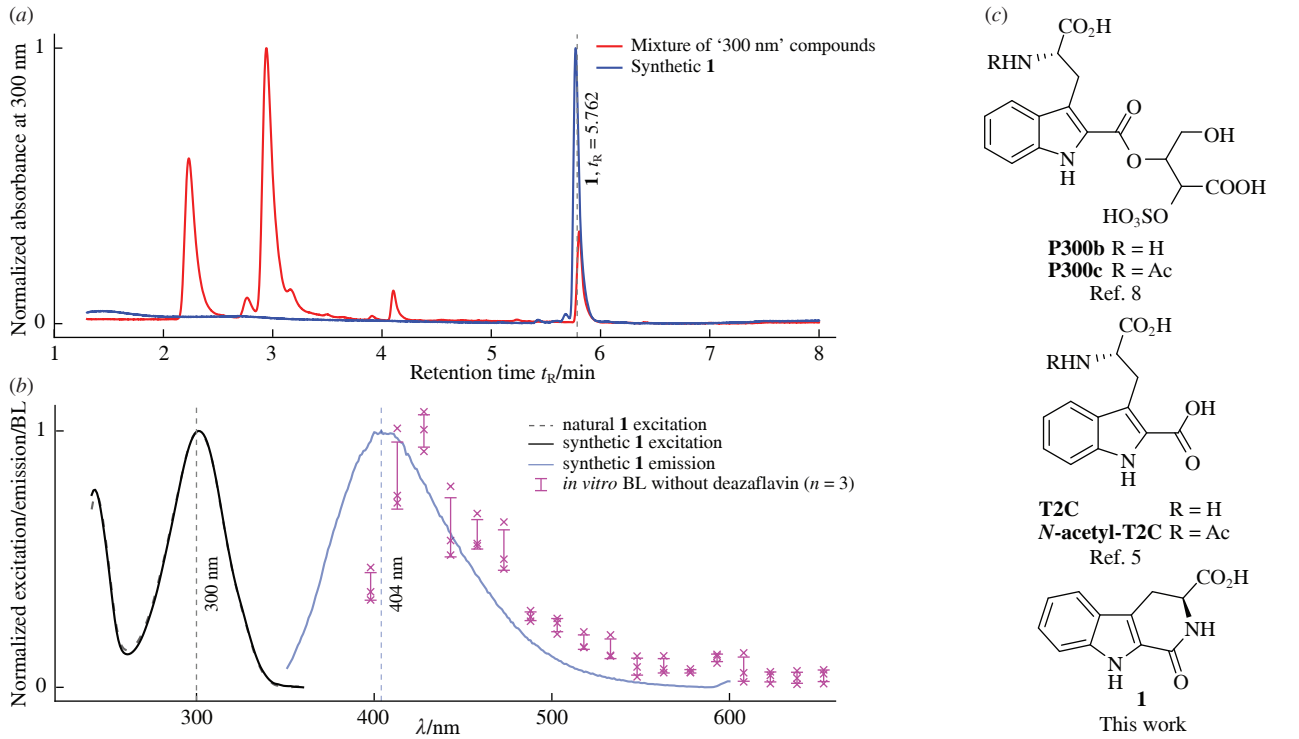


Figure 1 Characterisation of metabolite **1** by chromatography, absorption and fluorescence spectroscopy. (*a*) HPLC with monitoring at 300 nm. Synthetic compound **1** matches the most hydrophobic peak found in the *Henlea* extract. (*b*) Excitation spectra of natural and synthetic **1**, emission (fluorescence) spectrum of synthetic **1** (0.1% formic acid, pH 2.88) and bioluminescence (BL) spectrum of *Henlea* measured without deazaflavin chromophore (16 mM MOPS pH 7 with 1 mmol of $CaCl_2$, data collected with 15 nm intervals, $n = 3$ experiments are shown[5]). (*c*) All known T2C compounds. The structures of **P300b**,**c** were determined with precision up to the position of the sulfo group and the threonic acid linkage site to T2C.

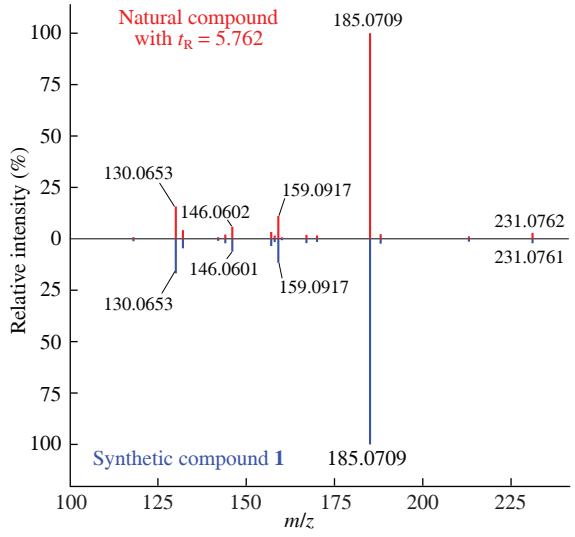
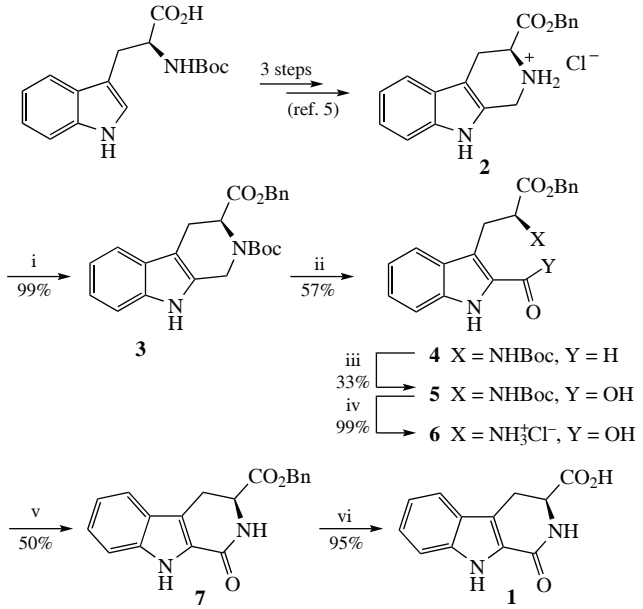


Figure 2 Comparison between tandem mass spectra of the natural and synthetic compounds in positive-ion mode. The parent ion of compound **1** is at *m*/*z* 231.0760.



Scheme 1 *Reagents and conditions*: i, Boc_2O, K_2CO_3, THF/H_2O; ii, SeO_2, dioxane, reflux; iii, $KMnO_4$, H_2O/acetone; iv, HCl/dioxane; v, TBTU, DIPEA, CH_2Cl_2; vi, H_2, Pd/C, MeOH.

oxidized with potassium permanganate in aqueous acetone[14] to afford the corresponding carboxylic acid **5**. For the cyclization, the Boc protecting group was removed with hydrogen chloride in dioxane[15] to obtain intermediate **6**, followed by intramolecular cyclization using *O*-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate (TBTU) and diisopropylethylamine (DIPEA) to give product **7**.[16] Final hydrogenolysis of the benzyl ester group[17] afforded the target compound **1** in an overall yield of 9% over six steps from reactant **2**.

The synthesized compound was characterized by HPLC and HRMS. Its retention time coincided with that of the corresponding component from the worm extract (see Figure 1). The MS^2 fragmentation pattern of synthetic **1** (blue) was identical to that of the natural compound (red) (see Figure 2). These data collectively confirm that we have elucidated the structure of a novel T2C natural product from *Henlea*. We would postulate that the synthetic sample **1** has (*S*)-configuration since it was obtained from (*S*)-tryptophan whereas the synthetic methods employ relatively mild conditions, so the racemizarion if occurs should proceed in small extent. The compound from the earthworm is highly likely (*S*)-configured being the natural one.

As anticipated, synthetic compound **1** also exhibits an absorption wavelength of 300 nm, even though the carboxy

group is replaced by a secondary amide. The emission spectrum was measured at an excitation wavelength of 300 nm and was shown to match the bioluminescence spectrum obtained without the deazaflavin activator, which corresponds to approximately 430 nm [see Figure 1(*b*)]. This suggests that compound **1** may be the oxyluciferin or a related oxidized luciferin product in the *Henlea* bioluminescent system. The minor discrepancy observed between the fluorescence and BL spectra can be attributed to the different microenvironments (free *vs*. enzyme-bound). Final confirmation of this hypothesis can only be achieved after the synthesis of the luciferin. Such work is currently in progress in our group.

To conclude, the synthesis and comparative analysis reliably established the structure of a newly discovered natural T2C derivative from *Henlea* earthworms. These results may help to elucidate the molecular mechanism of the bioluminescent system in future studies.

This work was supported by the grant of The Russian Science Foundation, Russia, grant no. 24-14-00421, https://rscf.ru/project/24-14-00421/.

Online Supplementary Materials

Supplementary data associated with this article can be found in the online version at doi: 10.71267/mencom.8105.

References

1 V. N. Petushkov, N. S. Rodionova, K. V. Purtov and V. S. Bondar', *Dokl. Biol. Sci.*, 2002, **385**, 310; https://doi.org/10.1023/A:1019936112111.
2 E. Rota, S. Martinsson and C. Erséus, *Org. Divers. Evol.*, 2018, **18**, 291; https://doi.org/10.1007/s13127-018-0374-6.
3 N. S. Rodionova, V. S. Bondar and V. N. Petushkov, *Dokl. Biochem. Biophys.*, 2002, **386**, 260; https://doi.org/10.1023/A:1020759527927.
4 V. N. Petushkov, M. V. Vavilov, I. A. Ivanov, R. H. Ziganshin, N. S. Rodionova, I. V. Yampolsky, A. S. Tsarkova and M. A. Dubinnyi, *Org. Biomol. Chem.*, 2023, **21**, 415; https://doi.org/10.1039/D2OB01946A.
5 V. N. Petushkov, M. V. Vavilov, A. N. Khokhlova, R. I. Zagitova, O. A. Belozerova, A. S. Shcheglov, S. I. Kovalchuk, A. S. Tsarkova, N. S. Rodionova, I. V. Yampolsky and M. A. Dubinnyi, *Biochem. Biophys. Res. Commun.*, 2024, **708**, 149787; https://doi.org/10.1016/j.bbrc.2024.149787.
6 J. G. Morin and J. W. Hastings, *J. Cell. Physiol.*, 1971, **77**, 313; https://doi.org/10.1002/jcp.1040770305.
7 S. Pi and Y.-J. Liu, *Luminescence*, 2026, **41**, e70454; https://doi.org/10.1002/bio.70454.
8 M. A. Dubinnyi, I. A. Ivanov, N. S. Rodionova, S. I. Kovalchuk, Z. M. Kaskova and V. N. Petushkov, *ChemistrySelect*, 2020, **5**, 13155; https://doi.org/10.1002/slct.202003075.
9 C. Seger, S. Sturm and H. Stuppner, *Nat. Prod. Rep.*, 2013, **30**, 970; https://doi.org/10.1039/C3NP70015A.
10 T. Shen, Y. Zhang, Y.-F. Liang and N. Jiao, *J. Am. Chem. Soc.*, 2016, **138**, 13147; https://doi.org/10.1021/jacs.6b08094.
11 Y. Saiga, I. Iijima, A. Ishida, T. Miyagishima, K. Homma, T. Oh-ishi, M. Matsumoto and Y. Matsuoka, *Chem. Pharm. Bull.*, 1987, **35**, 3284; https://doi.org/10.1248/cpb.35.3284.
12 A. Y. S. Balazs, R. J. Carbajo, N. L. Davies, Y. Dong, A. W. Hird, J. W. Johannes, M. L. Lamb, W. McCoull, P. Raubo, G. R. Robb, M. J. Packer and E. Chiarparin, *J. Med. Chem.*, 2019, **62**, 9418; https://doi.org/10.1021/acs.jmedchem.9b00716.
13 F. Gatta and D. Misiti, *J. Heterocycl. Chem.*, 1987, **24**, 1183; https://doi.org/10.1002/jhet.5570240449.
14 J. I. Urgel, M. Di Giovannantonio, K. Eimre, T. G. Lohr, J. Liu, S. Mishra, Q. Sun, A. Kinikar, R. Widmer, S. Stolz, M. Bommert, R. Berger, P. Ruffieux, C. A. Pignedoli, K. Müllen, X. Feng and R. Fasel, *Angew. Chem., Int. Ed.*, 2020, **59**, 13281; https://doi.org/10.1002/anie.202001939.
15 M. Meppen, B. Pacini, R. Bazzo, U. Koch, J. F. Leone, K. A. Koeplinger, M. Rowley, S. Altamura, A. Di Marco, F. Fiore, C. Giuliano, O. Gonzalez-Paz, R. Laufer, V. Pucci, F. Narjes and C. Gardelli, *Eur. J. Med. Chem.*, 2009, **44**, 3765; https://doi.org/10.1016/j.ejmech.2009.04.043.
16 F. Albericio, R. Chinchilla, D. Dodsworth and C. Najera, *Org. Prep. Proced. Int.*, 2001, **33**, 203; https://doi.org/10.1080/00304940109356592.
17 T. Furukawa, T. Akutagawa, H. Funatani, T. Uchida, Y. Hotta, M. Niwa and Y. Takaya, *Bioorg. Med. Chem.*, 2012, **20**, 2002; https://doi.org/10.1016/j.bmc.2012.01.050.

Received: 2nd June 2026; Com. 26/8105

Contents entry

Identification of one more metabolite from the Siberian bioluminescent earthworm *Henlea* sp.

Anastasia N. Khokhlova, Renata I. Zagitova, Valentin N. Petushkov, Natalia S. Rodionova, Matvey V. Vavilov, Olga A. Belozerova, Maxim A. Dubinnyi, Ilia V. Yampolsky

