Tetrahedron Letters No.29, pp. 3427-3436, 1966. Pergamon Press Ltd.
Printed in Great Britain.

CYPRIDINA BIOLUMINESCENCE I*
STRUCTURE OF CYPRIDINA LUCIFERIN

Yoshito Kishi, Toshio Goto, Yoshimasa Hirata
Chemical Institute, Faculty of Science, Nagoya University
Chikusa, Nagoya, Japan

Osamu Shimomura**
Water Research Laboratory, Faculty of Science, Nagoya University
Chikusa, Nagoya, Japan

Frank H. Johnson
Department of Biology, Princeton University
Princeton, New Jersey, U. S. A.

(Received 4 May 1966)

BIOLUMINESCENCE of Cypridina hilgendorfii (Japanese name:umi-hotaru) has been studied extensively since 1917, when Harvey (1) observed the luciferin-luciferase reaction with extracts of this animal. The luminescent system of Cypridina is one of the simplest among luminous animals and plants so far investigated; only luciferin, luciferase and oxygen being required for the light production in aqueous solutions. Thus, it has been considered to be most suitable for the study of bioluminescence mechanisms. General structures of the luciferin have been proposed by many workers (2), but since the materials used were impure, the evidence was not substantial.

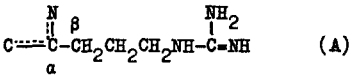
* A paper under the same title is scheduled for publication in F. H. Johnson and Y. Haneda "Bioluminescence in Progress", Princeton University Press, Princeton, N. J. (1966).

** Present address: Department of Biology, Princeton University, Princeton, New Jersey, U. S. A.

We obtained crystalline Cypridina luciferin in 1957 (3), and suggested the presence of an indole nucleus (3) and a mono-substituted guanidine group (4) in the molecule. In this communication we wish to report the complete structure of Cypridina luciferin.

Treatment of Cypridina luciferin (I) with Cypridina luciferase in the presence of oxygen gives, with emission of light, Cypridina oxyluciferin (II) [formerly oxyluciferin A (3)] and Cypridina etioluciferin (III) [formerly oxyluciferin B (3)]. The same products were obtained also with ammonia, instead of luciferase, but in this case no luminescence was observed. II is the initial product of the reaction, and treatment of II with dil. HCl affords III, which is rather stable to oxygen alkalis, and acids compared with I or II, and III is thus suitable for chemical investigation.

Structure of Cypridina Etioluciferin (III). Etioluciferin, m.p. (in sealed tube) (HCl salt) 235-237° dec.; (HBr) 226-229° dec.; (HNO_3) 222-224° dec., exhibits UV: $\lambda_{max.}^{MeOH-0.1N\ HCl}$ mμ (ε) 223 (26,000), 306 (20,000), 410 (5,000); $\lambda_{max.}^{MeOH-0.1N\ NaOH}$ 227 (24,000), 273 (18,000), 365 (7,600); IR: $\nu_{max.}^{KBr}$ cm^{-1} 3500-3000, 1650(s), 1620, 1530, 1430, 750; and is a diacidic base having pKa's 2.9 and above 11 (50% EtOH). Elemental analyses of its salts fit the formula $C_{16}H_{19}N_7 \cdot 2HX$ (X: Cl, Br or NO_3), which is supported by high-resolution mass spectrometry of etioluciferamine (IV) (see below) (5). III has a mono-substituted guanidine group (positive Sakaguchi reaction and pKa' above 11) and hydrolysis of III with barium hydroxide yields etioluciferamine (IV), $C_{15}H_{17}N_5$, the molecular formula of which is established by the high-resolution mass spectrum (M^+: obs. 267.145; calc. 267.148) (5). IV exhibits the UV spectrum identical with that of III, and shows pKa's 2.4 and 9.6 in 50% ethanol; the latter pKa' is shifted to 6.45 by addition of formalin (6), suggesting the presence of an amino group. Thus, the reaction III $\rightarrow$ IV can be represented as $C_{15}H_{15}N_4\text{-}NH\text{-}C(=NH)NH_2 \rightarrow C_{15}H_{15}N_4\text{-}NH_2$, and evidently the guanidine group in III takes no part in the chromophore. Since the acid hydrolysis of luciferin (I) affords arginine (7) (Table 1), the guanidine group must be part of the arginine moiety. The NMR spectrum of III (Fig. 1A) shows the presence of three methylene groups as in the form:



(A)

That the signal for the β-protons is triplet and about 1 ppm lower field than that of arginine suggests the absence of a hydrogen atom at the α-carbon atom and, instead, the presence of a double bond (8).

Catalytic hydrogenation of III with PtO_2 in ethanol gives hydro-etioluciferin (V), an amorphous powder, which shows a positive Sakaguchi but negative Ehrlich reaction. It exhibits a typical UV spectrum of the indole nucleus and slightly bathochromic shifts (2-5 mμ) are observed in its maxima when the medium is changed from acidic to alkaline (Fig. 2). Examination of UV spectra of model compounds listed below

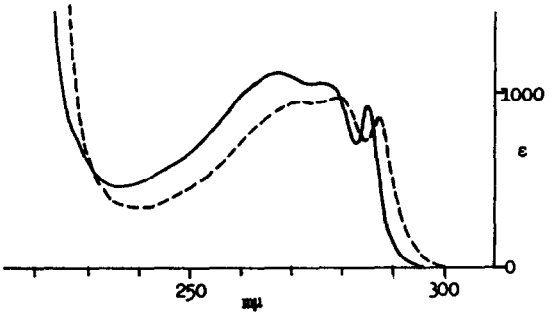
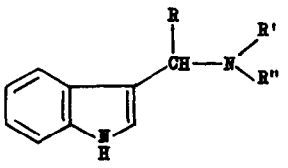


Fig. 2
UV spectra of hydro-etioluciferin (V)
——— in MeOH-0.1N HCl
----- in MeOH-0.1N NaOH

indicates that these shifts are associated with the β-(aminomethyl)-indole structure. Tryptamine and tryptophane as well as indole itself do not show such shifts.



R	R'	R"
H	CH_3	CH_3 (gramine)
H	H	CH_3
H	H	$CH_2CH_2NH_2$
H	H	CH_2CH_2OH
CH_3	CH_3	H

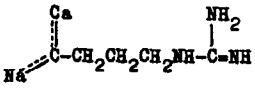
The β-(aminomethyl)indole structure in V is further supported by the following evidence. The NMR spectrum of III in DMSO (Fig. 1A) shows, besides of four aromatic ring protons, an indole-NH signal (disappears by addition of D_2O) at 11.7 ppm, and an indole α-H signal at 8.0 ppm; a spin-coupling is observed between these signals as is con-

firmed by a double resonance experiment. Then, V must have a β-substituted indole nucleus without a substituent at the α-position. In general, indole derivatives, such as tryptamine, bearing no substituent at the α-position show a positive Ehrlich reaction, but no reaction is actually observed with the model compounds listed above, as well as hydroetioluciferin (V). On heating or heating with alkali, V affords indole, possibly by a reverse Mannich reaction.

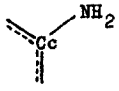
Since the UV spectra of III and IV are different from that of indole, the indole moiety in III and IV must conjugate with the chromophore in the side chain (partial structure B).

III is fairly resistant to acid treatments, but it reacts smoothly at 0° with nitrous acid to give an orange color when β-naphthol is added. When a solution of III is warmed at 50° after treatment with nitrous acid, a new compound (VI), which shows a UV spectrum different from that of III, is obtained. These results indicate the presence in III of a primary aromatic amino group (partial structure C), which takes part in the chromophore of III.

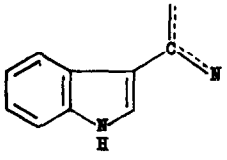
Although III is stable to acids, VI can be hydrolyzed with 6N HCl to give glycine. Luciferin (I) itself also gives glycine (ca. 1 mole) as well as arginine (ca. 0.5 mole) by heating with 6N HCl (Table 1). These results indicate that the nitrogen atom of the glycine moiety is not common with that of the arginine moiety or of the aromatic amino group. In the NMR spectrum of III, there remains at 8.3 ppm only one singlet signal corresponding to one proton, which has not been assigned. This signal is best assigned to the proton of the glycine moiety, which must exist as the partial structure D. From the above considerations, the following partial structures can be written for etioluciferin (III).



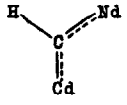
(A)



(C)

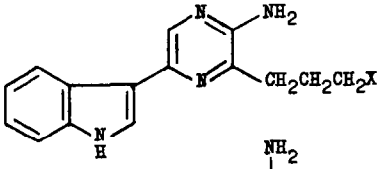


(B)



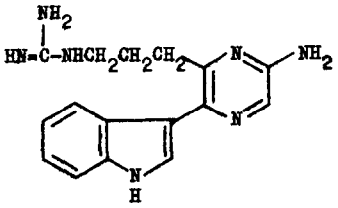
(D)

In the partial structures, A to D, one nitrogen atom and two carbon atoms must be overlapped since the actual molecular formula of III is $C_{16}H_{19}N_7$. Since the glycine part (D) cannot be overlapped with the arginine part (A) (Ca ≠ Cd, Na ≠ Nd), there remain only two possible formulas, IIIa (Cb = Cd, Ca = Cc, Na = Nb) and IIIb (Cc = Cd, Ca = Cb, Na = Nb).

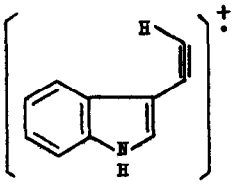


(IIIa) X = $-NH-\overset{NH_2}{\overset{|}{C}}=NH_2$

(IVa) X = $-NH_2$



(IIIb)



m/e 141

The structure IIIb may be excluded from the high-resolution mass spectrum of etioluciferamine (IV) (5). The spectrum shows a peak at m/e 141 ($C_{10}H_7N$) as the largest N_1-fragment. If IIIb were correct, it would be strongly expected that the N_1-fragment peaks containing more than ten carbon atoms would appear instead of the m/e 141 peak (structure shown above). Furthermore, from the biogenetic point of view the structure IIIa is more reasonable than IIIb, since IIIa is composed of tryptamine and arginine moieties. Here the structures IIIa and, hence, IVa are assigned to etioluciferin and etioluciferamine. The lower pKa' of III (2.9) and IV (2.4) in 50% ethanol are attributed to the aminopyrazine nucleus. 2-Aminopyrazine itself has a pKa' of 3.07 in water (9). Unequivocal evidence for this structure has been obtained by a total synthesis of III (10).

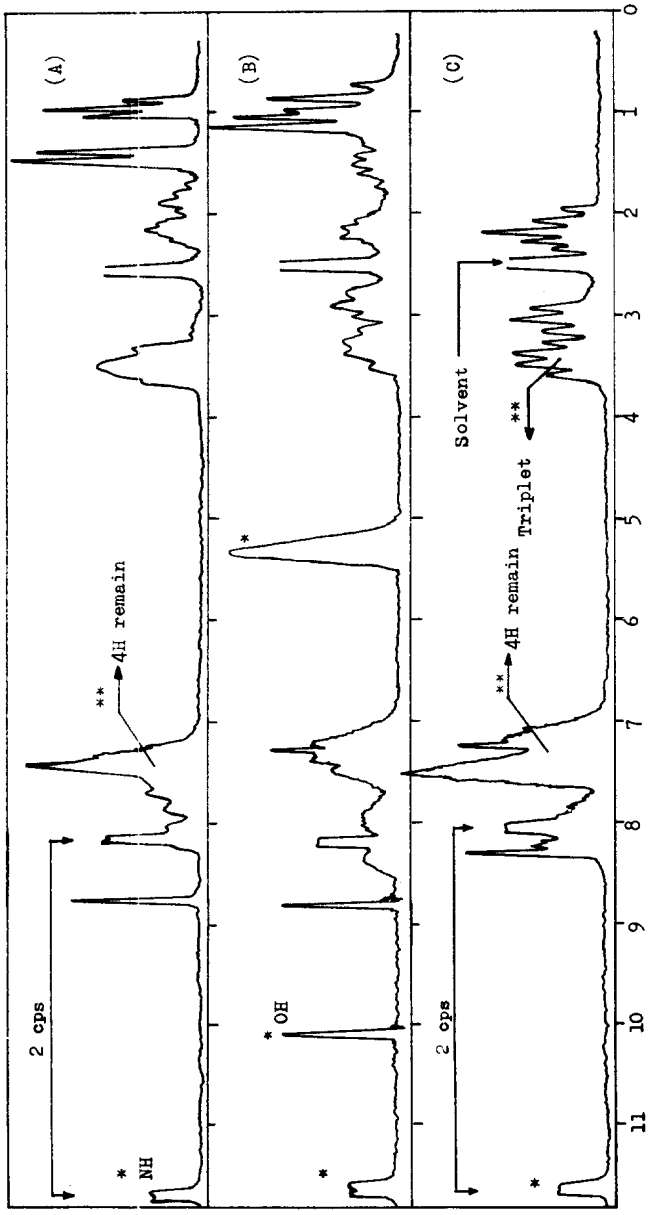


Fig. 1. NMR spectra of luciferin (I), oxyluciferin (II) and etioluciferin (III) at 60 Mc in $(CD_3)_2SO$; p.p.m. from internal tetramethylsilane (0 p.p.m.).

* Signals that disappear by the addition of D_2O.

** Signals that change in shape by the addition of D_2O.

Structure of Cypridina Oxyluciferin (II). Treatment of the oxyluciferin (II) with hydrochloric acid affords α-keto-β-methylvaleric acid (VII) as well as etioluciferin (III); the former being identified as its 2,4-dinitrophenylhydrazone.

II hydrochloride, m.p. 240-245° dec. (in sealed tube), exhibits UV: $\lambda_{max.}^{MeOH}$ mμ (ε) 220 (28,000), 271 (14,000), 302 (13,500), 347 (15,400); IR: $\nu_{max.}^{KBr}$ cm^{-1} 3500-3050, 2970-2800, 1650 (s), 1600, 1540, 1450, 1160, 1120, 750; and elemental analysis fits the molecular formula $C_{22}H_{27}O_2N_7 \cdot 2HCl$ (Found: C, 53.64; H, 5.77; O, 6.59; N, 19.83%), which indicates that II is constructed from III + VII - H_2O.

Table 1 shows amounts of amino acids obtained by hydrolysis of luciferin (I) and hydroluciferin (VIII) obtained by hydrogenation of I.

Table 1

Amino acids obtained from luciferin and hydroluciferin (7, 11)

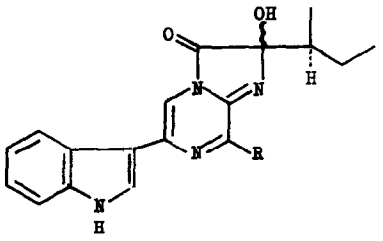
	Luciferin (I)		Hydroluciferin (VIII)	
	vacuum	O_2	vacuum	O_2
Glycine	0.96mole	0.93	0.03	0.12
Arginine*	0.62	0.38	0.03	0.41
Isoleucine**	0.16	0.16	0.98	0.96
Ammonia	0.27	0.33	0.40	0.41

*Combined amounts of arginine, γ-guanidinobutyric acid and proline.
**Combined amounts of isoleucine and alloisoleucine; both acids being produced in nearly the same amounts.

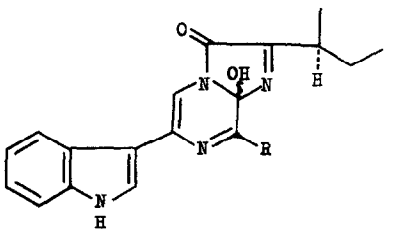
From this result the possibility that the nitrogen atom of isoleucine moiety might be overlapped with that of arginine or glycine moiety can be excluded. As the keto acid VII must be derived from the isoleucine moiety of luciferin, the ketonic carbon atom of VII must be linked with the primary amino group in etioluciferin (III).

The NMR spectrum of oxyluciferin (II) (Fig. 1B) reveals that II contains isoleucine, indole, glycine and arginine moieties. A sharp singlet at 10.1 ppm, which disappears on the addition of D_2O, is not included in the above moieties and may be assigned to a hydroxyl group. From the above considerations, the structure IIa is proposed for oxy-

luciferin (II). An alternative formula IIb is less probable since acid treatment of II gives III almost quantitatively (UV and paper chromatography). (Absolute configuration: see below).



(IIa)

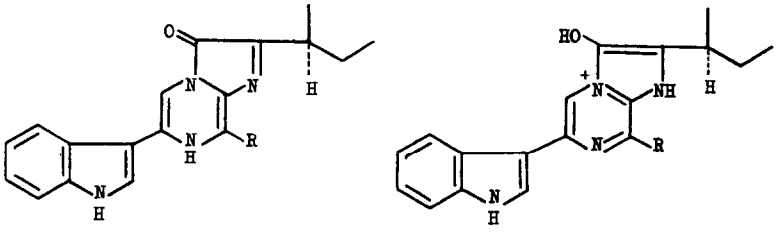


(IIb)

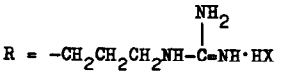
R = $-CH_2CH_2CH_2NH-\overset{NH_2}{\overset{|}{C}}=NH$

Structure of Cypridina Luciferin (I). The NMR spectrum of luciferin (I) hydrobromide (Fig. 1C) shows that I has the indole, glycine, arginine, and isoleucine moieties like oxyluciferin (II), but no hydroxyl signal is observed at around 10 ppm. Spin-decoupling experiments indicate the presence of the indole N-H and the α-H. Since the conversion of I to II involves oxidation corresponding to -2H [consumption of oxygen (12), redox titration (13), etc.], the molecular formula of I must be $C_{22}H_{29}O_2N_7$ or its equivalent in oxidation stage. Elemental analysis data of I hydrobromide, m.p. 252-254° dec. (in sealed tube); $\nu_{max.}^{KBr}$ cm^{-1} 3500-3200, 3010, 1685-1645 (s), 1620, 1590, 1545, 745; pKa' (in H_2O) below 2, 8.35, and above 11; (in 50% EtOH) below 2, 8.3 and above 11; (Found: C, 46.20; H, 5.46; C, 3.78; N, 17.20; Br, 27.57%) fits the formula $C_{22}H_{27}ON_7 \cdot 2HBr$. It forms almost colorless crystals, but its solution in methanol gives an orange-red color [$\lambda_{max.}^{MeOH-H_2O}$ mμ (ε) 218 (27,500), 270 (17,000), 310sh (11, 500), 435(9,000)]. When anhydrous hydrogen bromide is introduced into the solution of the hydrobromide in methanol, the color of the solution fades and the UV spectrum of the solution [$\lambda_{max.}^{MeOH-HBr}$ 221 (29,000), 243sh (12,000), 270-280sh (14,000-15,000), 307 (20,000), 320sh (6,000)] becomes very similar to that of etioluciferin (III). Thus, the chromophore of luciferin in

weakly acidic solution (orange-red) is different from that in strongly acidic, anhydrous solutions (nearly colorless). Two forms of luciferin are best formulated as Ia (orange-red) and Ib (colorless).



(Ia) (Ib)



The α-carbon atom of the isoleucine moiety in luciferin is not asymmetric. This is supported by the results that the hydrolyzates of luciferin and hydroluciferin always contain nearly the same amounts of isoleucine and alloisoleucine (Table 1). The absolute configuration at C_β of the isoleucine moiety was determined as follows: The mixture of isoleucine and alloisoleucine (nearly the same amounts) produced by hydrolysis of hydroluciferin (VIII) was treated with L- and D-amino acid oxidases and then analyzed the composition of the products by means of amino acid analyzer; only isoleucine was decomposed when L-amino acid oxidase from Crotalus Adamanteus venom (14) was employed, whereas D-amino acid oxidase from Hog kidney (14) consumed only alloisoleucine. This result indicates that the mixture contained L-isoleucine and D-alloisoleucine. Then, the configuration at C_β must be that shown in Ia.

Thus, luciferin is composed of tryptamine, arginine and L-isoleucine moieties, and the extended dihydropyrazine structure is responsible for the oxidation during the luminescence.

The authors wish to thank Mr. Samon Miyamoto and Mr. Hazime Miyamoto, Setoda, Hiroshima, Japan, for collecting Cypridina hilgendorfii. One of us (O. S.) wishes to express his thanks to the Japan Society for the Promotion of Science and the Rockefeller Foundation

for support of this work. One of us (F. H. J.) is grateful to the J. S. P. S., N. S. F., and O. N. R. for support of this work.

REFERENCES

1. E. N. Harvey, "Bioluminescence", Academic Press, New York (1952). p. 299.
2. F. I. Tsuji, A. M. Chase and E. N. Harvey in F. H. Johnson, "The Luminescence of Biological Systems", Am. Assoc. for the Advancement of Science, Washington, D. C. (1955). p. 127.
3. O. Shimomura, T. Goto and Y. Hirata, Bull. Chem. Soc. Japan 30, 929 (1957).
4. Y. Hirata, O. Shimomura and S. Eguchi, Tetrahedron Letters No. 5, p. 4 (1959); O. Shimomura, J. Chem. Soc. Japan 81, 179 (1960).
5. Y. Kishi, T. Goto, S. Eguchi, Y. Hirata, E. Watanabe and T. Aoyama, Tetrahedron Letters
6. T. Goto, Y. Hirata, S. Hosoya and N. Komatsu, Bull. Chem. Soc. Japan 30, 729 (1957).
7. S. Eguchi, J. Chem. Soc. Japan 84, 86 (1963).
8. Another possibility that two nitrogen atoms are linked to the α-carbon atom can obviously be excluded in the later stage.
9. A. Albert, J. Chem. Soc. 2240 (1948).
10. Y. Kishi, T. Goto, S. Inoue, S. Sugiura and H. Kishimoto, Tetrahedron Letters
11. Interpretation of this result is given in reference 5.
12. F. H. Johnson, O. Shimomura, Y. Saiga, L. C. Gershman, G. T. Reynolds and J. R. Waters, J. Cell. Comp. Physiol. 60, 85 (1962).
13. Data to be published in near future.
14. Purchased from Sigma Chemical Co.; J. P. Greenstein and M. Winitz, "Chemistry of the Amino Acids", John Wiley and Sons, Inc., New York (1961). p. 185, 1782, 2064.