The Structure of *Latia* Luciferin*

Osamu Shimomura and Frank H. Johnson

ABSTRACT: The substrate, luciferin, of the bioluminescence enzyme system of the freshwater limpet *Latia neritoides* has been purified by an improved method and the properties were studied in regard to the mass spectrum, nuclear magnetic resonance spectrum, ultraviolet and infrared absorption, and optical rotation. Properties of derivatives of luciferin, obtained by catalytic hydrogenation, ozonolysis, ammonolysis, $NaBH_4$ reduction of ammonolysis products, hydrolysis with NaOH to form luciferin aldehyde, and catalytic hydrogenation of luciferin aldehyde, have also been studied. On the basis of the results, the empirical formula of luciferin has been found to be $C_{15}H_{24}O_2$, with the proposed structure **4** of the text.

Isolation of the specific substrate and partial purification of the specific enzyme in the bioluminescent enzyme–substrate (luciferase–luciferin) system of the New Zealand freshwater limpet *Latia neritoides* were recently reported, together with some properties of the system and its components (Shimomura *et al.*, 1966). On the basis of the evidence in the present paper, the chemical structure of the luciferin molecule is now proposed.

Materials and Methods

Specimens collected during January–March, 1966–1967, were promptly frozen and preserved with Dry-Ice for air shipment and storage thereafter. Luciferase solutions for assays of luciferin activity were prepared by shaking 1 g (six to seven individuals) of frozen organisms in 100 ml of 0.005 M sodium phosphate buffer (pH 6.8) at room temperature and centrifuging to remove insoluble matter. The supernatant (5 ml) was added to a small volume (5–50 μl) of an ethanolic solution of luciferin at 25° and the total light emitted was measured in terms of "light units" (LU) with a photomultiplier light-integrating apparatus. Calibration with the luminol reaction according to the method of Lee *et al.* (1966) and with the luminescence reaction of the purified *Cypridina* system (Johnson *et al.*, 1962) gave results in close agreement, indicating that 1 LU = 1.0×10^{10} photons.

A Perkin-Elmer Model 202 was used for measuring ultraviolet absorption, a Perkin-Elmer Model 421 for infrared spectra, a Varian Model HA100 for nuclear magnetic resonance spectra, and a Cary Model 60 for optical rotation. Mass spectra of luciferin and its hydrogenation product (previously obtained) were measured at Nagoya University by means of the Hitachi Model RMU-6D, and all other mass spectra by the same model at the Morgan-Schaffer Corp. in Montreal, Canada. Samples for analysis were dissolved in *n*-hexane, which was evaporated before the measurements. Solvents used in this report (ethanol, ethyl acetate, *n*-hexane carbon tetrachloride, and acetic acid) were redistilled before use.

The method of purifying luciferin in this study was improved and simplified over that originally used. Initial processing was done with batches of 30 g of frozen specimens which were placed in a cold mortar and crushed with a pestle, first in 70–80 ml of cold ethanol (inactivating the luciferase), then repeating the extraction with 40–50 ml and finally, 30 ml. The combined extracts were filtered and added to 15 ml of butanol. This mixture was vacuum evaporated to about 15 ml to which an equal volume of ethanol was added before temporary storage with Dry-Ice while the remaining specimens were processed according to the same procedure. The combined crude extracts thus obtained were dark greenish. Some water was added and the mixture was extracted with *n*-hexane. The hexane layer was washed with water. The water layer was discarded and the hexane layer was poured onto a 4 × 30 cm column of silicic acid Mallinckrodt (100 mesh) prepared with *n*-hexane. The active material adsorbed on the column was eluted with 25% benzene in *n*-hexane. The eluate was evaporated until most of the benzene and hexane was removed and the active residue was redissolved in *n*-hexane. The solution was poured on a 2.5 × 20 cm silicic acid column and the active material was eluted with 8% ethyl acetate in *n*-hexane. The resulting luciferin solution had a light yellowish color. It was further purified by two repeated chromatographies on 1 × 15 cm columns of silicic acid Mallinckrodt (SilicAR CC-7, 200–325 mesh) eluting with 18% benzene in *n*-hexane each time. After evaporating most of the solvents, the product was stored in ethanolic solution at −25° until used. An activity of 3.8×10^5 LU was obtained from a portion of the solution which, after evaporation, left an

* From the Biology Department and the Program in Biochemical Sciences, Princeton University, Princeton, New Jersey 08540. *Received November 28, 1968.* This work was aided in part by National Science Foundation Grants GB-4086, GB-6836, and GF-143; Office of Naval Research Contract Nonr 4264(00); the Eugene Higgins Fund allocated to Princeton University; and facilities made available by the Whitehall and John A. Hartford Foundations. Publication in whole or in part by or for any purpose of the U. S. Government is permitted.

oily residue of approximately 500 μg. On this basis, the total activity of 2.8×10^7 LU in the pure luciferin obtained from about 2 kg of frozen specimens in this study represents a yield of about 37 mg of luciferin.

Results

Properties of Latia luciferin. The purified luciferin was colorless and under ultraviolet light it gave a slight, blue-purple fluorescence which was probably without significance. The ultraviolet absorption maximum was observed at 207 mμ (Figure 1) instead of the previously reported 212 mμ; wavelength calibration of the spectrophotometer verified the accuracy of the peak under the conditions involved in the present study. The infrared absorption spectrum was obtained with better resolution than previously by using a thinner cell and carbon tetrachloride as the solvent (Figure 2). The distinct peaks at 1742 and 1170 cm^{-1} indicate the presence of ester or lactone. The mass spectrum obtained by the direct introduction method (Figure 3A) indicates a molecular weight of 236, and the absence of halogens, sulfur, and nitrogen; absence of nitrogen was indicated also by negative results of spot tests by the method of Feigl and Amaral (1958). The isotope abundance M + 1 (16.63%) and M + 2 (1.73%) suggests that the empirical formula is most likely $C_{15}H_{24}O_2$ (M + 1, 16.67%; M + 2, 1.70%), although the possibility remains that it is $C_{14}H_{20}O_3$ (M + 1, 15.57%; M + 2, 1.73%). A metastable peak was found at *m/e* 153, indicating a fragmentation, 236 → 190 + 46, *i.e.*, splitting of formic acid or ethanol from the luciferin. Measurement by the indirect introduction method showed a definite increase of intensities at *m/e* 208, 193, 175, and 150, probably by decomposition.

The nuclear magnetic resonance spectrum is shown in Figure 4. From the integration curve (not shown), assuming that the peaks at 7.94 and 6.97 ppm represent one proton each, the total for the region of 2.2–0.8 ppm includes 23.5 protons, and for the peaks at 2.03, 1.88, 1.71, 1.59, 1.45, and 1.00 ppm the number of protons would be 3.7, 2.3, 3.0, 3.6, 3.3, and 5.9, respectively; considering the shape of the curve and the overlapping, these data are believed to indicate 4 H, 2 H, 3 H, 3 H, 4 H, and 6 H, respectively. The peak at 1.00 ppm (6 H), as well as the split peak in the 1375-cm^{-1} region of the infrared spectrum, indicate the presence of a *gem*-dimethyl group. No optical rotation was found in measurements on a 0.06% luciferin solution in ethanol over the range from 350 to 220 mμ.

Catalytic Hydrogenation of Luciferin. Approximately 5 mg of luciferin (4×10^6 LU) in 5 ml of ethanol was hydrogenated with 5 mg of PtO_2 catalyst at 5° for 24 hr. The catalyst was then removed by centrifuging. The product in the supernatant proved to be inactive in the bioluminescence reaction. Figure 1 shows that the product has an absorption maximum at 203 mμ with an optical density approximately half that of luciferin. The mass spectrum of the product showed the following peaks in *m/e*, with relative abundances in parentheses: 238 M^+ (5.4), 223 (21), 194 (4.1), 179 (8.1), 177 (4.9), 149 (9.8), 137 (7.4), 123 (100), and 121 (25). The spectrum below *m/e* 121 was almost identical with that of luciferin aldehyde (*vide infra*) except at *m/e* 31 (5.3). Evidently, one double bond of luciferin became saturated by the hydrogenation.

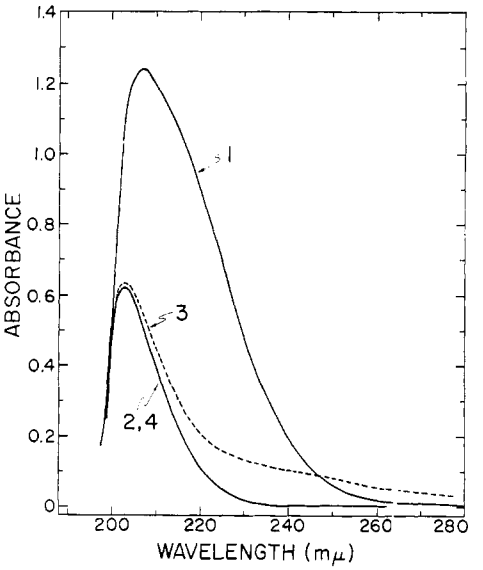


FIGURE 1: Ultraviolet absorption spectra of luciferin (1), hydrogenation product of luciferin (2), ammonolysis product of luciferin (3), and luciferin aldehyde (4), all in ethanol at the same molar concentration. Luciferin concentration of (1) was 16,000 LU/ml, giving an approximate value of 13,700 for the molecular extinction coefficient of the peak at 207 mμ.

Ammonolysis of Luciferin. Aqueous ammonia solution (2 ml of 30%) (freshly distilled) was added to 4 mg of luciferin in 5 ml of ethanol and left standing overnight at room temperature in a stoppered flask. The solvent and ammonia were removed by evaporation with an aspirator, leaving an oily product, which was not active in the bioluminescence reaction. The ultraviolet spectrum of the ammonolysis product (Figure 1) is similar to that of the hydrogenation product except for a bulge at about 240 mμ. The mass spectrum of the ammonolysis product was completely identical with that of luciferin aldehyde (hydrolysis product; *vide infra*) with M^+ at *m/e* 208, suggesting that these two products are the same, and moreover that luciferin is an ester but not a lactone, and also that the molecular weight of the alcohol portion of the ester is 208.

$NaBH_4$ Reduction of the Ammonolysis Product of Luciferin. An oily ammonolysis product prepared from approximately 5 mg of luciferin was dissolved in 5 ml of 95% ethanol and treated with 10 mg of $NaBH_4$. After 3 hr, the mixture was neutralized with CO_2, 10 ml of ethanol was added, and the inorganic salts which precipitated were removed by centrifugation. The supernatant was evaporated to about 0.5 ml, and the residue was taken up in 6 ml of *n*-hexane. The ultraviolet absorption spectrum of the product in ethanol was practically the same as that of the catalytic hydrogenation product of luciferin (Figure 1). The mass spectrum of the product showed the following distinctive peaks and relative abundances: 210 M^+ (22), 195 (25), 177 (21),

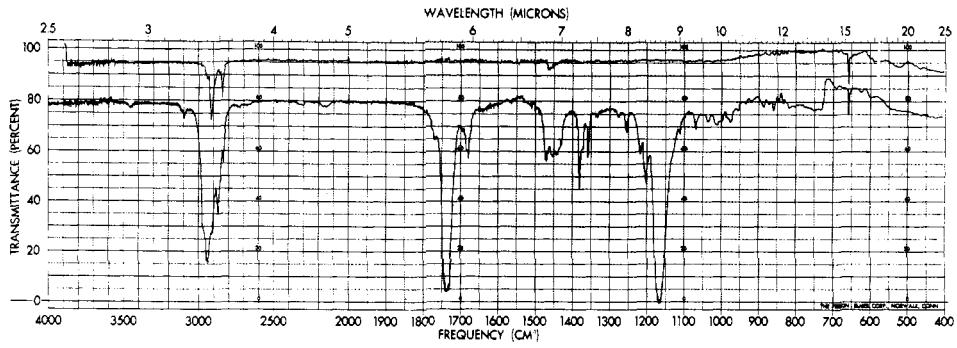


FIGURE 2: Infrared spectrum of luciferin, approximately 6 mg/0.1 ml, in carbon tetrachloride, using a 0.1-mm cell. The upper line is without luciferin.

149 (3.7), 137 (12), 123 (100), and the spectrum below *m/e* 123 was very similar to that of luciferin aldehyde. These results suggest the presence of only one carbonyl group in the ammonolysis product with a molecular weight of 208.

Hydrolysis of Luciferin. Approximately 5 mg of luciferin dissolved in 5 ml of ethanol was added to 2 ml of 0.1 N aqueous NaOH. After 15 min, the mixture was neutralized with CO_2, then shaken with 3 ml of water and 10 ml of *n*-hexane. The separated lower layer was concentrated to 5 ml *in vacuo* and used in testing for formate according to the method of Eegriwe (1937). The solution (1 ml) in a test tube was treated with 3 drops of 2 N HCl and some magnesium wire. After the formation of hydrogen gas ceased, 1 ml of 70% H_2SO_4 and 1–2 mg of chromotropic acid were dissolved in the mixture, then finally 1 ml of concentrated H_2SO_4 was placed at the bottom of the solution. The formation of a deep purple color at the boundary, in addition to the negative result obtained when the reduction step was omitted, clearly indicated the presence of formate in the test solution.

The upper (*n*-hexane) layer, which was not active in the bioluminescence reaction, was divided into two portions. A half-portion was first used for the measurement of ultraviolet and infrared spectra. The same sample after recovering was used for a mass spectrum, and also for testing for a double bond with tetranitromethane in carbon tetrachloride. The test gave a light yellow color, indicating the presence of a C=C bond. This C=C bond was further supported by a molecular ion peak at *m/e* 212, after catalytic hydrogenation of the remaining half-portion in acetic acid with platinum oxide and palladium black. No hydrogenation occurred when ethanol was used instead of acetic acid.

Based on the infrared spectrum and other properties, the hydrolysis product in the above *n*-hexane solution will be called luciferin aldehyde hereafter.

The ultraviolet spectrum of luciferin aldehyde was the same as that of the hydrogenation product, as shown in Figure 1. The infrared spectrum (Figure 5) indicates without doubt the presence of an aldehyde group by the absorption peaks at 1723 and 2702 cm^{-1}, and also suggests the absence of an ether group by the lack of absorption in the 1100–1200-cm^{-1} region. The mass spectrum with M^+ at *m/e* 208 shown in Figure 3B had a metastable ion peak at *m/e* 158.7, indicating

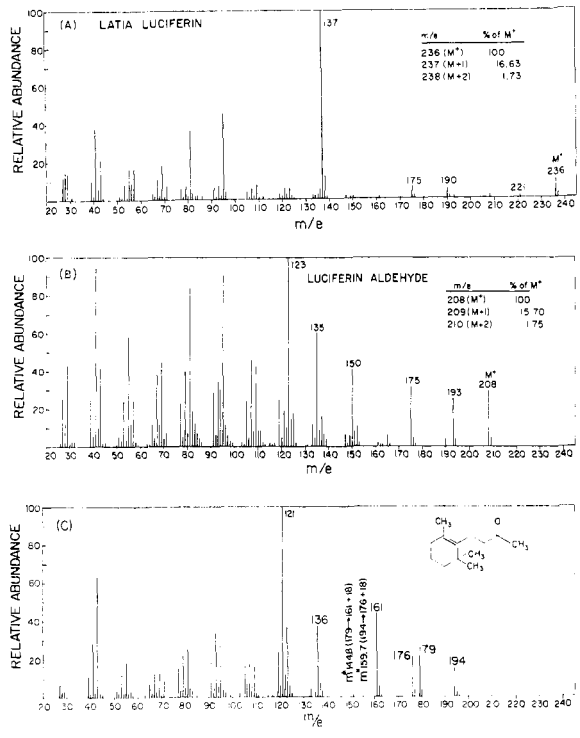


FIGURE 3: Mass spectra of luciferin (A), luciferin aldehyde (B), and dihydro-β-ionone (C). A was obtained by the direct sample introduction method, whereas B and C were by the indirect introduction method; 70 eV.

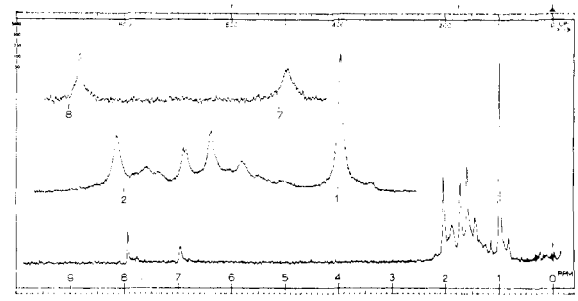


FIGURE 4: Nuclear magnetic resonance spectrum of luciferin in carbon tetrachloride in a microtube; 100 Mc. Sample concentration, approximately 5 mg/0.1 ml. Internal standard tetramethylsilane (0 ppm).

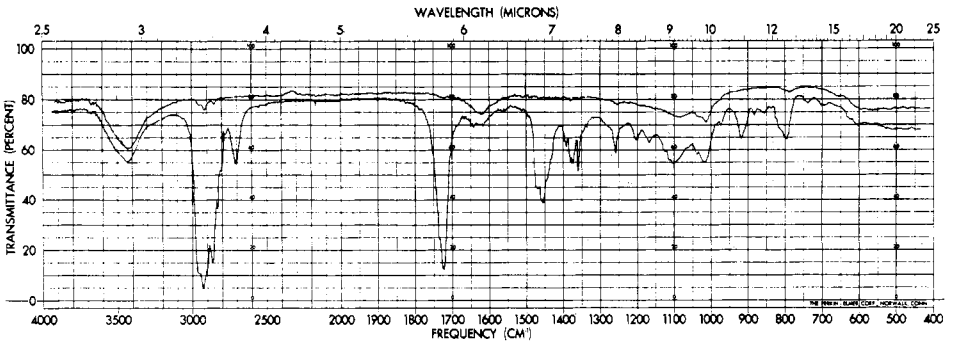


FIGURE 5: Infrared spectrum of luciferin aldehyde on a KBr disk. Approximately 2 mg of luciferin in a small amount of *n*-hexane was slowly added onto one side of the KBr disk, forming an oily film after evaporation of the solvent, before the measurement. No KBr disk was used for the reference beam. The upper line is without sample.

cleavage of water from the *m/e* 193 ion. The isotope abundance of M + 1 (15.70%) and M + 2 (1.75%) suggests the empirical formula of $C_{14}H_{24}O$ (M + 1 (15.55%), M + 2 (1.33%)) with a possibility of $C_{13}H_{20}O_2$ (M + 1 (14.45%), M + 2 (1.37%)).

Ozonolysis of Luciferin. Approximately 4 mg of luciferin was dissolved in a mixture of 5 ml of *n*-hexane and 5 ml of carbon tetrachloride, and a slow stream of ozone was bubbled through the solution for 15 min. After standing for 15 min, the solution was evaporated with an aspirator giving an oily residue, which was redissolved in 5 ml of ethanol. The ethanol solution was hydrogenated for 20 min using 5 mg of Adams' PtO_2 as the catalyst. The mass spectrum of the product showed that the sample contained several compounds which had been formed by different degrees of ozonization, hydrogenation, and decomposition. Relatively intense peaks above *m/e* 60 were as follows: 246, 244, 226, 212, 210, 208, 197, 194, 181, 179, 170, 167, 165, 163, 157, 149, 142, 127, 126, 123, 109, 99, 85 (strongest), 71, and 69.

Discussion

From the foregoing data, *Latia* luciferin is apparently a formyl ester. The carbonyl stretching absorption in the infrared spectrum at 1742 cm^{-1}, in contrast to 1720 cm^{-1} in the usual formyl ester, however, suggests the presence of a double bond between the α- and β-carbon atoms of the formyl group, and this is supported by the easy formation of an aldehyde after hydrolysis. The mass spectrum of luciferin aldehyde indicates the presence of a methyl group attached to the α-carbon atom of the aldehyde, by an intense peak at *m/e* 150 which is formed by the loss of $CH(CH_3)CHOH$ (*vide infra*). This methyl group is also noted in the nuclear magnetic resonance spectrum of luciferin at a peak at 1.71 ppm. Therefore, $CC(CH_3)=CHOCHO$ (**1**) is considered to be a part of luciferin.

Thus, the singlets of 7.94 and 6.97 ppm in the nuclear magnetic resonance spectrum of luciferin are most suitably assigned to the protons of C*H*O and C=C*H*, respectively; the latter of the two peaks shows a slight broadening caused by coupling through the double bond (model compounds: CH_3OCHO, δ = 8.03 ppm; $(CH_3)_2C=CHOCOCH_3$, δ = 6.79 ppm (Jackman and Wiley, 1960); $CH_3CH=CHCOOCH_3$, J_{HCH_3} = 1.82 (*trans*) and 1.67 cps (*cis*) (Fraser and McGreer, 1961). No information was obtainable for the geometrical configuration associated with the double bond in structure **1**.

The second double bond which could be hydrogenated in acetic acid must be fully substituted, because no more signal was found for the olefinic proton in the nuclear magnetic resonance spectrum. One of these substituents should be a methyl group fitting the nuclear magnetic resonance signal of 1.59 ppm. These considerations lead to the partial structure

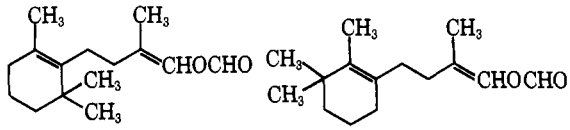
$$\begin{matrix} C \\ C \end{matrix}>C=C(CH_3)C \qquad (2)$$

A *gem*-dimethyl group, $>C(CH_3)_2$ (**3**), also should be present in the structure, as mentioned before. Since the peak at 2.03 ppm is too low in the field for methyl protons in the situation involved, no other methyl groups than those included in structures **1**–**3** are to be considered in the structure of luciferin according to the nmr spectrum.

On the other hand, luciferin aldehyde contains only one carbonyl group (C=O) and it does not have an ether group (COC), as already discussed. This leads conclusively to the empirical formula $C_{14}H_{24}O$ for luciferin aldehyde and $C_{15}H_{24}O_2$ for luciferin.

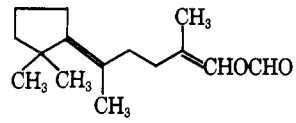
This empirical formula of luciferin shows that the luciferin must contain one cyclic structure as well as the partial structures of **1**–**3** and inevitably the double bond in **2** must be *endo*-cyclic or *exo*-cyclic. The ultraviolet spectrum shows that the two double bonds in **1** and **2** are not conjugated. All of the structures which satisfy the above conditions and have one or three carbon atoms between the double bonds were found not to be able to explain the nmr spectrum, with reference to the peaks at 2.03, 1.88, and 1.45 ppm. No structure which has four or more carbon atoms between the double bonds can satisfy the above conditions.

Structures **4**–**6** which have two carbon atoms between the double bonds satisfy the above conditions. In these structures, the nuclear magnetic resonance peak at 2.03 ppm (4 H) is readily explained by the protons of two methylene groups between the double bonds. Moreover, the peaks at 1.88 and 1.45 ppm are consistent with the cyclohexene ring of **4** and **5**, whereas they are in a



4

5



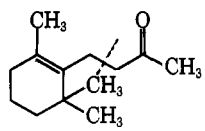
6

slightly higher field than expected for the cyclopentane ring of **6**.

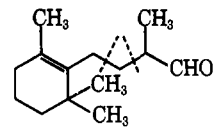
The mass spectrum of the ozonolysis product also favors **4** or **5** rather than **6**, by the presence of *m/e* 226 ($C_{13}H_{22}O_3$, triketone), as well as by the absence of m/*e* 112 ($C_7H_{12}O$, cyclopentanone) and *m/e* 114 ($C_6H_{10}O_2$, diketone). The presence of the peak at *m/e* 142 ($C(CH_3)_2COCH_2CH_2COCH_3$ + H) also favors structure **4** rather than **5**.

The mass spectrum of luciferin (Figure 3A) is well explained by structure **4**, *i.e.*, *m/e* 221 (M − CH_3), *m/e* 190 (M − HCOOH), *m/e* 175 (M − CH_3 and HCOOH), and the very intense base peak at *m/e* 137 (allylic cleavage in the middle of two double bonds).

A model compound, dihydro-β-ionone (**7**), would be predicted to cleave in between the C=O and C=C. Accordingly, *m/e* 136 in the mass spectrum of **7** (Figure 3C)[1] is considered to be the result of this cleavage, with hydrogen rearrangement, and with the charge remaining on the fragment including the skeleton. Thus, the base peak at *m/e* 121 is also explained by the further cleavage of CH_3 from the *m/e* 136 ion.



7



8

The aldehyde **8** derived by hydrolysis from **4** would be expected to show an allylic cleavage as well as a cleavage of the bond beta to C=O. The mass spectrum of the aldehyde of natural luciferin (Figure 3B) shows a weak peak at *m/e* 137 corresponding to the former cleavage, and a relatively intense peak at *m/e* 150, formed with hydrogen rearrangement, corresponding to the latter cleavage. In this latter cleavage, the charge remaining on the fragment which includes the skeleton is accounted for by a similar cleavage of the model compound **7**. Thus, *m/e* 135 is interpreted as (*m/e* 150 − CH_3).

The fragmentation patterns of Figure 3B,C below *m/e* 130 are closely similar except for higher intensities in Figure 3C for *m/e* 121 (explained above) and for *m/e* 43 (CH_3CO), thus suggesting that luciferin aldehyde (and consequently luciferin too) has the same skeleton as **7**.

Finally, the bioluminescence product of luciferin gave a mass spectrum which is judged essentially the same as that in Figure 3C, indicating that the structure of the product is structure **7** itself (O. Shimomura and F. H. Johnson, submitted for publication). On the basis of the reasons discussed above, we propose structure **4** as the structure of *Latia* luciferin.

[1] The data of Figure 3C are included here by courtesy of Professor R. K. Hill and Dr. M. Frachebold who synthesized the dihydro-β-ionone as a step toward the synthesis of luciferin.

Acknowledgments

The authors are pleased to acknowledge their indebtedness to a group of graduate students, led by Mr. James Watson, of the University of Auckland, New Zealand, for collecting and air shipping the specimens of *Latia*; to Mr. R. Hyde of the Program of Biochemical Sciences, Princeton University, for recording the infrared spectra; to Mr. Ed. Miller, Chemistry Department, Princeton University, for recording the nuclear magnetic resonance spectra; to Dr. T. Goto and Dr. A. Tatematsu, Nagoya University, Japan, for mass spectra obtained with previous preparations of luciferin; to Dr. O. J. Jacobus, Chemistry Department, Princeton University, for the data on optical rotation; to Dr. C. Gilvarg, Chemistry Department, Princeton University, for assistance with the ozonization of luciferin; and to Mrs. J. Sabadell and Miss Lynn Hum for technical assistance.

References

Eegriwe, E. (1937), *Z. Anal. Chem. 110*, 22.

Feigl, F., and Amaral, J. R. (1958), *Anal. Chem. 30*, 1148.

Fraser, R. R., and McGreer, D. E. (1961), *Can. J. Chem. 39*, 505.

Jackman, L. M., and Wiley, R. H. (1960), *J. Chem. Soc.*, 2881.

Johnson, F. H., Shimomura, O., Saiga, Y., Gershman, L., Reynolds, G. T., and Waters, J. R. (1962), *J. Cellular Comp. Physiol. 60*, 85.

Lee, J., Wesley, A. S., Ferguson, J. F., and Seliger, H. H. (1966), *in* Bioluminescence in Progress, Johnson, F. H., and Haneda, Y., Ed., Princeton, N. J., Princeton University, p 35.

Shimomura, O., Johnson, F. H., and Haneda, Y. (1966), *in* Bioluminescence in Progress, Johnson, F. H., and Haneda, Y., Ed., Princeton, N. J., Princeton University, p 391.