## Cloning and expression of cDNA for the luciferase from the marine ostracod Vargula hilgendorfii

(bioluminescence/oxygenase/luciferin/aequorin/gene expression)

ERIC M. THOMPSON\*, SHIGEKAZU NAGATA\*, AND FREDERICK I. TSUJI\*†

\*Osaka Bioscience Institute, 6-2-4 Furuedai, Suita, Osaka 565, Japan; and <sup>†</sup>Marine Biology Research Division, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92093

Communicated by Martin D. Kamen, June 7, 1989 (received for review May 12, 1989)

The marine ostracod Vargula hilgendorfii ABSTRACT ejects luciferin and luciferase into seawater to produce a bright luminous cloud. The light is due to the oxidation of luciferin, an imidazopyrazine compound, by molecular oxygen, catalyzed by luciferase. The mechanism of the reaction has been studied extensively and the 60 kcal/mol required for the blue emission have been shown to be derived from the oxidation of luciferin via a dioxetanone intermediate, in which the excited state oxyluciferin bound to luciferase is the emitter. However, only limited information is available regarding the properties of the enzyme. This paper reports the cloning and sequence analysis of the cDNA for Vargula luciferase and the expression of the cDNA in a mammalian cell system. The primary structure, deduced from the nucleotide sequence, consists of 555 amino acid residues in a single polypeptide chain with a molecular weight of 62,171. Two regions of the enzyme show significant amino acid sequence homology with an N-terminal segment of the photoprotein aequorin. The Vargula luciferase gene, which contains a signal sequence for secretion, should be well suited as a reporter in studies of gene expression.

The marine ostracod crustacean Vargula hilgendorfii (1), formerly Cypridina hilgendorfii (2), has played a central role in contributing to an understanding of the chemical and physical bases of bioluminescence. V. hilgendorfii is indigenous to the south coastal waters of Japan, where it is commonly referred to as "umi botaru" or the "sea firefly." In the past the animal occurred in great abundance, but their numbers have been reduced significantly in recent years. The earliest studies on the bioluminescence of V. hilgendorfii were carried out by Harvey (3), and subsequent studies by various investigators resulted in numerous publications (4).

V. hilgendorfii is a small animal ( $\approx$ 3 mm long) with nocturnal habits. It lives in sand at the bottom of shallow waters and becomes an active feeder at night. When disturbed it ejects a copious secretion of luciferin (substrate) and luciferase into sea water, producing a bright luminous cloud. The light results from an enzyme-substrate reaction, catalyzed by luciferase, in which luciferin, an imidazopyrazine compound, is oxidized by molecular oxygen (1, 5). The other products of the reaction are oxyluciferin and carbon dioxide (6, 7). The excited-state oxyluciferin bound to luciferase is the emitter in the reaction (8). Following the determination of the imidazopyrazine structure of Vargula luciferin (5), it was found that imidazopyrazine compounds are widely used as substrates in the bioluminescence reactions of marine organisms. One such reaction which has been studied in detail is that of the jellyfish Aequorea victoria, from which the cDNA for the calcium-binding photoprotein aequorin has recently been cloned (9, 10).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Interestingly, V. hilgendorfii luciferin, or a compound almost identical to it, is used as a substrate by luminescent fish. These include the North American batrachoidid fish Porichthys notatus, which possesses more than 700 dermal light organs (11); apogonid fishes of the Far East, in which the light organs are connected to the digestive tract by a duct (12); and the Japanese pempherid fish Parapriacanthus ransonneti, in which the light organ communicates directly with the pyloric caeca (13). The luciferins and luciferases of all of these animals give reciprocal light-emitting cross-reactions with Vargula luciferin and luciferase. Since luminescent Vargula is found in the same waters as these fishes, questions have been raised as to the origin of their luciferins and luciferases (14).

The present study reports the cloning and nucleotide sequence analysis of the cDNA for *Vargula* luciferase and the expression of this cDNA in a mammalian cell system.<sup>‡</sup> Because of the use of imidazopyrazine compounds by taxonomically diverse marine organisms and since the primary structure of aequorin is the only one presently known, it is of interest to compare the structures of *Vargula* luciferase and aequorin. The study of the structures of these two proteins and others should lead to a better understanding of how bioluminescent reactions take place and how the luminescent systems of various marine organisms are related evolutionarily.

## **MATERIALS AND METHODS**

Purification and Sequencing of Vargula Luciferase. Vargula luciferase, partially purified as described (15), was purified to homogeneity, first by using a tryptamine affinity column (Pierce) equilibrated in 2.0 M NaCl/0.07 M Tris·HCl, pH 7.2, and eluting stepwise with 30% (vol/vol) ethylene glycol/0.17 M NaCl/0.07 M Tris HCl, pH 7.2, and second, after concentration by ultrafiltration, by using a p-aminobenzamidine affinity column (Pierce) under the same chromatographic conditions (F.I.T., unpublished work). Aliquots (120  $\mu$ g) of the purified protein were subjected to endopeptidase digestions by trypsin (Boehringer Mannheim), lysyl endopeptidase (Wako Pure Chemical, Osaka, Japan) or arginyl endopeptidase (Takara Shuzo, Kyoto, Japan). Peptide fragments were separated by reverse-phase chromatography on a C<sub>18</sub>protein-peptide HPLC column (Vydac). Undigested luciferase and purified peptides were sequenced by Edman degradation using a gas-phase protein sequenator (Applied Biosystems; model 477A).

**mRNA Preparation and cDNA Library Construction.** V. hilgendorfii were collected (15) at Chiba, Japan, and frozen immediately in liquid nitrogen. The ostracods (wet weight, 5 g) were ground to a fine powder in liquid nitrogen with an ultraturrax homogenizer (Janke & Kunkel, Staufen, F.R.G.). Total cellular RNA was extracted from the powder by the

<sup>&</sup>lt;sup>‡</sup>The sequence reported in this paper has been deposited in the GenBank data base (accession no. M25666).



FIG. 1. NaDodSO<sub>4</sub>/polyacrylamide gradient gel electrophoresis of purified Vargula luciferase. Samples were analyzed on a 10–15% PhastGel (Pharmacia) gradient and proteins were visualized by using a PhastGel silver staining kit (Pharmacia). Lane 1, Pharmacia low molecular weight markers ( $M_r$  given  $\times 10^{-3}$ ); lane 2, 50 ng of affinity chromatography purified Vargula luciferase.

guanidine thiocyanate/cesium chloride method (16) and  $poly(A)^+$  RNA was purified by oligo(dT)-cellulose column chromatography (17). The cDNA library was constructed as described (18) except that the double-stranded *Eco*RI digested cDNA was size-selected by two purifications on low-melting agarose (Bio-Rad).

Screening and Analysis of cDNA Clones. Peptide sequences with minimum codon ambiguity were used to synthesize oligonucleotide probes by the phosphoramidite method (19), using a DNA synthesizer (Applied Biosystems; Model 381A). The oligonucleotide probes were 5'-end-labeled (specific activity,  $5-6 \times 10^6$  cpm/pmol) with T4 polynucleotide kinase (Takara Shuzo) and  $[\gamma^{-32}P]ATP$  (222 TBq/mmol; New England Nuclear) and used to screen  $1 \times 10^6$  recombinant phages by plaque hybridization (20). Hybridization conditions were as described (21) except that the hybridization temperature was lowered to 28°C and the filters were washed at 37°C with 1× SSC (0.15 M NaCl/15 mM sodium citrate, pH 7.0). EcoRI fragments of positive clones were subcloned in pUC8 and nucleotide sequence analyses were carried out by the dideoxynucleotide chain termination technique (22, 23)using a 7-deaza sequencing kit (Takara Shuzo) and  $[\alpha$ -<sup>32</sup>P]dCTP (222 TBq/mmol) (New England Nuclear).

**Expression of Vargula Luciferase.** HindIII and Bgl II linkers (Takara Shuzo) were ligated to Vargula luciferase cDNA at the 5' and 3' ends, respectively. The plasmid pRSVL (obtained from S. Subramani, Univ. of California, San Diego) (24) was digested with Sma I and ligated with Bgl II linker. The linearized plasmid and the cDNA were digested with

HindIII and Bgl II and the HindIII-Bgl II fragment containing Vargula luciferase cDNA was inserted in place of the firefly luciferase cDNA. The expression vector was designated pRSVVL.

Monkey COS cells (25)  $(7 \times 10^6)$  were seeded on 10-cm dishes in 10 ml of Dulbecco's modified Eagle's medium (Nissui Seiyaku, Tokyo) containing 10% fetal calf serum (HyClone). The cells were transfected with 10  $\mu$ g of plasmid DNA containing the cDNA insert by using the calcium phosphate method (26, 27). After 48 hr of incubation, the medium was collected and the cells were harvested. Cell extracts were prepared by repeated cycles of freezing and thawing followed by centrifugation (24). To assay for Vargula luciferase activity, aliquots of the medium and of the cell extract were diluted with 200 mM Tris HCl, pH 7.6 (total volume of 1.5 ml) in a 20-ml scintillation vial and placed in a Mitchell-Hastings photometer (28). Vargula luciferin, prepared as described (15), was dissolved at 50 nM in 200 mM sodium phosphate buffer, pH 6.8, and 1.5 ml of this solution was injected into the scintillation vial. Conversion of light intensity to quanta per second was made by calibrating the photometer with a  $[^{14}C]$ -hexadecane light standard (29).

## RESULTS

**Peptide Sequences and Probe Design.** When analyzed by polyacrylamide gel electrophoresis in the presence of NaDodSO<sub>4</sub>, the purified Vargula luciferase showed a single band of  $M_r$  68,000 (Fig. 1). Edman degradation of the intact protein did not yield an assignable amino acid, suggesting that the N terminus is blocked. The luciferase was then digested with endopeptidases, and several peptides were subjected to Edman degradation.

A portion of one of the peptide fragments had the sequence Thr-Met-Glu-Asn-Leu-Asp-Gly-Gln-Lys, which was used to design the complementary oligodeoxynucleotide probe 5'-(T/C)TT(T/C)TGICC(A/G)TCIAGGTT(T/C)TCCATIGT-3'. Deoxyinosine (30) was included at three positions wherethere was high codon degeneracy. A second complementaryprobe with A substituted for G at position 15 was alsosynthesized, and a mixture of these two probes was used toscreen the Vargula cDNA library. Using the plaque hybridization technique (20), we isolated eight positive clones fromthe library; two of these, designated VL16 and VL18, withinsert lengths of 1.2 and 1.5 kilobases (kb), were selected forfurther analysis.

Nucleotide Sequence of Vargula Luciferase. The restriction map of Vargula luciferase cDNA was constructed by using clones VL16 and VL18 as shown in Fig. 2. Clones VL16 and VL18 each had a single internal *Eco*RI site and had a



FIG. 2. Restriction map of Vargula luciferase cDNA and strategy for the construction of a full-length clone. The restriction map (a) was generated from nucleotide sequence analysis of clones VL16 (b) and VL18 (c). After digestion with EcoRI, the 5' fragment of VL16 was ligated to the 3' fragment of VL18 to form the full-length clone (d). The full-length clone was inserted into the expression vector (described in the text) via a 5' HindIII linker and a 3' BamHI linker. The shaded portion of the restriction map (a) indicates the portion of the nucleotide sequence believed to code for the putative signal sequence. bp, Base pairs.

10 50 70 90 GANACCGACACCACC ATG AAG ATA ATA ATT CTG TCT GTT ATA TTG GCC TAC TGT GTC ACC GAC AAC TGT CAA GAT GCA TGT CCT GTA GAA Met Lys Ile Ile Leu Ser Val Ile Leu Ala Tyr Cys Val Thr Asp Asn Cys Gin Asp Ala Cys Pro Val Glu 110 GCG GAA CCG CCA TCA AGT ACA CCA ACA GTT CCA ACT TCT TGT GAA GCT AAA GAA GGA GAA TGT ATA GAT ACC AGA TGC GCA ACA TGT AAA Ala Glu Pro Pro Ser Ser Thr Pro Thr Val Pro Thr Ser Cys Glu Ala Lys Glu Gly Glu Cys Ile Asp Thr Arg Cys Ala Thr Cys Lys 30 40 190 210 230 250 270 CGA GAT ATA CTA TCA GAT GGA CTG TGT GAA AAT AAA CCA GGG AAG ACA TGC TGT AGA ATG TGC CAG TAT GTG ATT GAA TGC AGA GTA GAA Arg Asp lle Leu Ser Asp Gly Leu Cys Glu Asn Lys Pro Gly Lys Thr Cys Cys Arg Met Cys Gln Tyr Val Ile Glu Cys Arg Val Glu 60 70 290 310 330 350 GCA GCT GGT TAT TTT AGA ACG TTT TAC GGC AAA AGA TTT AAT TTT CAG GAA CCT GGT AAA TAT GTG CTG GCT AGG GGA ACC AAG GGT GGC Ala Ala Gly Tyr Phe Arg Thr Phe Tyr Gly Lys Arg Phe Asn Phe Gln Glu Pro Gly Lys Tyr Val Leu Ala Arg Gly Thr Lys Gly Gly 100 100 110 370 390 410 430 450 GAT TGG TCT GTA ACC CTC ACC ATG GAG AAT CTA GAT GGA CAG AAG GGA GCT GTG CTG ACT AAG ACA ACA CTG GAG GTT GCA GGA GAC GTA Asp Trp Ser Val Thr Leu Thr Met Glu Asn Leu Asp Gly Gln Lys Gly Ala Val Leu Thr Lys Thr Thr Leu Glu Val Ala Gly Asp Val 470 490 510 530 ATA GAC ATT ACT CAA GCT ACT GCA GAT CCT ATC ACA GTT AAC GGA GGA GCT GAC CCA GTT ATC GCT AAC CCG TTC ACA ATT GGT GAG GTG Ile Asp Ile Thr Gln Ala Thr Ala Asp Pro Ile Thr Val Asn Gly Gly Ala Asp Pro Val Ile Ala Asn Pro Phe Thr Ile Gly Glu Val 150 150 160 170 550 570 590 610 630 630 ACC ATT GCT GTT GAA ATA CCG GGC TTC AAAT ATC ACA GTC ATC GAA TTC TTT AAA CTA ATC GTG ATT GAT ATT CTG GGA GGA AGA TCT Thr Ile Ala Val Val Glu Ile Pro Gly Phe Asn Ile Thr Val Ile Glu Phe Phe Lys Leu Ile Val Ile Asp Ile Leu Gly Gly Arg Ser 650 670 690 710 GTG AGA ATT GCT CCA GAC ACA GAC AAA GGA CTG ATA TCT GGT ATC TGT GGT AAT CTG GAG ATG AAT GAC GCT GAT GAC TTT ACT ACA Val Arg Ile Ala Pro Asp Thr Ala Asn Lys Gly Leu Ile Ser Gly Ile Cys Gly Asn Leu Glu Met Asn Asp Ala Asp Asp Phe Thr Thr 210 220 230 730 750 770 790 810 GAT GCA GAT CAG CTG GCG ATC CAA CCC AAC ATA AAC AAA GAG TTC GAC GGC TGC CCA TTC TAT GGC AAT CCT TCT GAT ATC GAA TAC TGC Asp Ala Asp Gln Leu Ala Ile Gln Pro Asn Ile Asn Lys Glu Phe Asp Gly Cys Pro Phe Tyr Gly Asn Pro Ser Asp Ile Glu Tyr Cys 240 250 250 830 850 870 890 AAA GGT CTG ATG GAG CCA TAC AGA GCT GTA TGT CGT AAC AAT ATC AAC TTC TAC TAT TAC ACT CTA TCC TGT GCC TTC GCT TAC TGT ATG Lys Gly Leu Met Glu Pro Tyr Arg Ala Val Cys Arg Asn Asn Ile Asn Phe Tyr Tyr Tyr Thr Leu Ser Cys Ala Phe Ala Tyr Cys Met 280 910 930 950 970 990 GGA GGA GAA AGA GGT AAA CAC GTC CTT TTC GAC TAT GTT GAG ACA TGC GCT GCG CCG GAA ACG AGA GGA ACG TGT GTT TTA TCA GGA Gly Glu Glu Arg Ala Lys His Val Leu Phe Asp Tyr Val Glu Thr Cys Ala Ala Pro Glu Thr Arg Gly Thr Cys Val Leu Ser Gly 300 320 1010 1030 1050 1070 CAT ACT TTC TAT GAC ACA TTC GAC AAA GCA AGA TAT CAA TTC CAG GGC CCA TGC AAG GAG ATT CTG ATG GCC GCA GAC TGT TAC TGG AAC His Thr Phe Tyr Asp Thr Phe Asp Lys Ala Arg Tyr Gln Phe Gln Gly Pro Cys Lys Glu Ile Leu Met Ala Ala Asp Cys Tyr Trp Asn 330 340 1090 1110 1130 1150 1170 ACA TGG GAT GTA AAG GTT TCA CAT AGA GAC GTC GAA TCA TAC ACT GAG GTA GAG AAA GTA ACA ATC AGG AAA CAG TCA ACT GTA GTA GAT Thr Trp Asp Val Lys Val Ser His Arg Asp Val Glu Ser Tyr Thr Glu Val Glu Lys Val Thr 11e Arg Lys Gln Ser Thr Val Val Asp 370 1190 CTC ATT GTG GAT GGC AAG CAG GTC AAG GTT GGA GGA GTG GAT GTA TCT ATC CGG TAC AGC TCT GAG AAC ACT TCC ATA TAC TGG CAG GAT Leu Ile Val Asp Gly Lys Gln Val Lys Val Gly Gly Val Asp Val Ser Ile Pro Tyr Ser Ser Glu Asn Thr Ser Ile Tyr Trp Glg Asp 400 410 1270 1290 GGA GAC ATC CTG ACG ACG GCC ATC CTA CCT GAA GCT CTT GTC GTT AAG TTC AAC TTT AAG CAG CTC CTT GTA GTT CAT ATC AGA GAT CCA Gly Asp Ile Leu Thr Thr Ala Ile Leu Pro Glu Ala Leu Val Val Lys 20 410 410 1310 1310 1330 1330 1330 1330 1350 TTC GAT GGA AAG ACA TGC GGC ATA TGT GGT AAC TAT AAT CAA GAT TCA ACT GAT GAT GAT TTC TTT GAC GCA GAA GGA GGA GCA TGC GCT CTA ACC Phe Asp Gly Lys Thr Cys Gly Ile Cys Gly Asn Tyr Asn Gln Asp Ser Thr Asp Asp Phe Phe Asp Ala Glu Gly Ala Cys Ala Leu Thr 450
460
460
460
460 
 1450
 1470
 1490
 1510
 1530

 CCC AAC CCC CCA GGA TGT ACA GAG GAA CAG AAA CCA GAA GCT GAG CGA CTT TGC AAT AAT CTC TTT GAT TCT ATC GAC GAG AAA TGT
 Pro Asn Pro Gly Cys Thr Glu Glu Gln Lys Pro Glu Ala Glu Arg Leu Cys Asn Asn Leu Phe Asp Ser Ser Ile Asp Glu Lys Cys
 400
 1550 1570 1590 1610 AAT GTC TGC TAC AAG CCT GAC CGG ATT GCC CGA TGT ATG TAC GAG TAT TGC CTG AGG GGA CAA CAA GGA TTT TGT GAC CAT GCT TGG GAG Asn Val Cys Tyr Lys Pro Asp Arg Ile Ala Arg Cys Met Tyr Glu Tyr Cys Leu Arg Gly Gln Gln Gly Phe Cys Asp His Ala Trp Glu 510 520 530 1630 1650 1670 1690 1710 TTC AAG AAA GAA TGC TAC ATA AAA CAT GGA GAC ACT CTA GAA GTA CCA CCT GAA TGT CAA TAA ACGTACAAAGATACAGAAGCTAAGGCTACTACAGC Phe Lys Glu Cys Tyr IIe Lys His Gly Asp Thr Leu Glu Val Pro Pro Glu Cys Gln End 540 550 

FIG. 3. Nucleotide sequence and predicted amino acid sequence of *Vargula* luciferase cDNA. Numbers above each line refer to nucleotide position and numbers below each line refer to amino acid position. Horizontal arrows indicate regions where the amino acid sequences were identical with those obtained by endopeptidase digestions. Consensus sequences for N-glycosylation are boxed, and the AATAAA polyadenylylation signal is underlined.

sequence overlap of 830 nucleotides. Nucleotide sequence analysis demonstrated that clone VL16 encodes the Nterminal portion of luciferase and clone VL18, the C-terminal portion. The complete nucleotide sequence of the *Vargula* luciferase cDNA is shown in Fig. 3 together with the deduced amino acid sequence.

The sequence contains an open reading frame of 1665 nucleotides coding for a protein of 555 amino acids with a

calculated molecular weight of 62,171. The open reading frame contains the amino acid sequence used to construct the oligonucleotide probe. In addition, the amino acid sequences of seven other peptides determined by Edman degradation were in complete agreement with those deduced from the nucleotide sequence. The translation initiation codon was assigned to the first ATG codon at nucleotide position 16. The nucleotide sequence around the putative initiation codon is in



FIG. 4. Assay of active Vargula luciferase synthesized and secreted by COS cells. The cells, transfected with the expression plasmid pRSVVL, were incubated for 48 hr and the culture medium was assayed for luciferase activity in the presence of Vargula luciferin. Light emission was monitored with a Mitchell-Hastings (28) photometer attached to a Chromatocorder 12 (System Instruments). Shown above is the assay of the culture medium in which the transfected cells were incubated. A scintillation vial containing 100  $\mu$ l of the culture medium was placed in the photometer and the shutter was opened. At time zero, 50 nM luciferin was injected into the vial and at 1.5 min the shutter was closed.

good agreement with the consensus sequence CC(A/G)CCAUGG, characteristic of many eukaryotic mRNAs (31). The N-terminal amino acid sequence also has many of the features of a signal sequence (32), in accord with the biological role of luciferase as a secretory protein.

Consistent with the positive staining of luciferase by the periodic acid-Schiff reaction, there are two potential Nglycosylation sites (Asn-Xaa-Ser/Thr), at amino acid residues 186 and 408. At residue 258 the sequence Asn-Pro-Ser occurs, but this sequence is generally not efficiently glycosylated (33). N-glycosidase F was found to reduce the molecular weight of luciferase by 2000-3000, while no reduction in molecular weight was detected in the digestions specific for O-glycosylation (data not shown). These results indicate that Vargula luciferase is N-glycosylated and that carbohydrate moieties may account for the difference in the calculated molecular weight of the polypeptide deduced from the nucleotide sequence and that measured for the native protein by gel electrophoresis (Fig. 1) and by gel filtration and sedimentation equilibrium analysis (34). The other notable feature of Vargula luciferase is the very cysteine-rich region present in the N-terminal portion, where nine cysteine residues can be found between amino acid residues 39 and 82.

Expression of Vargula Luciferase in COS Cells. To establish that the cloned cDNA actually encodes Vargula luciferase, a mammalian cell system was used to express the cDNA. As shown in Fig. 2, a full-length cDNA was constructed from clones VL16 and VL18, and it was placed under the promoter of the Rous sarcoma virus long terminal repeat. The hybrid gene was introduced into COS cells and the transfected cells were incubated for 48 hr. As shown in Fig. 4, luciferase activity was clearly detected in the culture medium of transfected COS cells with some luciferase activity also present in the cell extract. Luminescence from the media and cell extracts of COS cells which were either mock (no plasmid DNA) transfected or transfected with pSV0CAT (35) or pRSVL (24) was more than two orders of magnitude less intense than that shown in Fig. 4. These results confirm that the reconstructed cDNA is the full-length cDNA for Vargula luciferase and that it also encodes the signal for protein secretion. The luminescence emitted from the culture medium of COS cells was directly proportional to the volume of medium assayed, and a clear signal above background could be detected from as little as 5  $\mu$ l of the 10 ml of culture medium obtained from COS cells transfected with the Vargula luciferase expression vector.

## DISCUSSION

The present results are of interest not only because they contribute to a further understanding of the bioluminescent reaction in *Vargula* but also because they provide a more general understanding of the bioluminescent reactions of a large number of marine organisms in which structurally analogous luciferins and similar reaction mechanisms are used. Among these organisms, a direct comparison can now be made with the bioluminescent reaction of the jellyfish *Aequorea victoria* (36).

Aequorea possesses, in the margin of its umbrella, a calcium-binding protein, aequorin, which emits light in the presence of calcium ions. Aequorin is a complex of apoaequorin (apoprotein), coelenterazine, and molecular oxygen. When aequorin binds calcium ions, a conformational change takes place, converting the protein into an oxygenase. Coelenterazine is then oxidized via an intramolecular reaction, as shown in Fig. 5. The emitter in the reaction is the excitedstate coelenteramide bound to apoaequorin (37). The cDNA for apoaequorin has been cloned and expressed, and subsequent analyses have shown apoaequorin to be composed of 189 amino acid residues ( $M_r = 21,400$ ) with three EF-hand structures characteristic of Ca<sup>2+</sup>-binding sites (9, 10, 38, 39).

Despite the close similarities in substrate structure and mechanisms of the bioluminescent reaction in Vargula and Aequorea, the enzymes and substrates of the two systems show only slight light-emitting cross-reactions. However, the two proteins are seen to show limited but significant amino acid sequence homologies in two regions of Vargula luciferase (residues 97-154 and 353-411) with one in apoae-



FIG. 5. Bioluminescence reactions of Vargula and Aequorea. In literature describing the Aequorea reaction, luciferin and oxyluciferin are also referred to as coelenterazine and coelenteramide, respectively.

a b	120 KRFNFQEPGKYVLARGTKGGDWSVTLTME ::.:: KKLATDELEKYAKNEPTLIRIWGDAL-FD 90 110	140 NLDGQKGAVLTKTTLEVAGDVIDITQATAD .::
a b	360 YWNTWDVKVSHRDVESYTEVEKVTIRKQS 	400 T V V D L I V D G K Q V K V G G V D V S I P Y S S E N T S I 

FIG. 6. Sequence homology between Vargula luciferase (a) and jellyfish (Aequorea) aequorin (b). Double dots indicate identical residues and single dots indicate amino acids with homologous scores in the Dayhoff mutation data matrix (40). Numbers indicate the position of each residue with respect to the N terminus of each protein.

quorin (residues 82-144) (Fig. 6). Since the entire apoaequorin molecule consists of 189 amino acid residues, about one third the number in Vargula luciferase, it is conceivable that one or both of the homologous regions in Vargula luciferase may be involved in light emission. The larger size of Vargula luciferase may also be due to its hydrolase activity (8). Possibly, it optimizes the luminescent reaction in seawater. in contrast to the aequorin reaction, which occurs intracellularly.

Since the N terminus of Vargula luciferase is blocked, it was not possible to determine which amino acid residue constitutes the N terminus of the mature polypeptide. However, considering the fact that the region in Vargula luciferase (residues 97-154) homologous to acquorin (residues 87-144) is shifted by 10 amino acid residues, it is possible that, as an approximation, the leader peptide for secretion is the first 11 amino acid residues and that the N terminus of luciferase is tyrosine (residue 12; Fig. 3). Should such a peptide be cleaved, it is likely to be at position 11, since alanine and the adjacent amino acid residues possess a pattern compatible with a signal sequence cleavage site (32).

Vargula luciferase does not appear to contain any clear EF-hand structures characteristic of calcium-binding sites, as present in aequorin. Consequently the previously reported inhibition of Vargula luciferase by EDTA (41) may be attributed to a different action from the inhibition observed with aequorin. Examination of the primary structure of Vargula luciferase reveals the presence of a relatively high content of glutamic and aspartic residues, as in aequorin. There is also a high content of cysteine residues. Since no free sulfhydryl group was detected in Vargula luciferase (34), the cysteine residues presumably exist in intramolecular disulfide bridges.

The lux operon of the bacterium Vibrio fischeri (42) and the luciferase gene of the firefly Photinus pyralis (24, 43) have been used as indicators of promoter activity and in other applications (44). The simplicity and specificity of the Vargula reaction, involving only the substrate luciferin and molecular oxygen, the high sensitivity with which light can be measured, and the fact that Vargula luciferase is secreted into the culture medium, suggest that the Vargula luciferase gene may find wide biomedical applications.

We thank Dr. Y. Haneda and Mr. M. Hayashi for assistance in the collection of Vargula and Dr. M. Nishizawa of Osaka Bioscience Institute for helpful discussions. This work was supported in part by research grants (DMB 85-17578 and DMB 88-44459) from the National Science Foundation (to F.I.T.), a Grant-in-Aid (62480131) from the Ministry of Education, Science and Culture of Japan (to S.N.), and a postdoctoral fellowship from the Science and Technology Agency of Japan (to E.M.T.).

- Harvey, E. N. (1952) Bioluminescence (Academic, New York), pp. 1. 297-331.
- 2. Kornicker, L. S. & Baker, J. H. (1977) Proc. Biol. Soc. Wash. 90, 218-231.
- Harvey, E. N. (1917) Am J. Physiol. 42, 318-341. 3.
- 4. Johnson, F. H. & Shimomura, O. (1978) Methods Enzymol. 57, 331-364.

- 5. Kishi, Y., Goto, T., Hirata, Y., Shimomura, O. & Johnson, F. H. (1966) in Bioluminescence in Progress, eds. Johnson, F. H. & Haneda, Y. (Princeton Univ. Press, Princeton, NJ), pp. 89-113.
- Stone, H. (1968) Biochem. Biophys. Res. Commun. 31, 386-391.
- Shimomura, O. & Johnson, F. H. (1971) Biochem. Biophys. Res. Commun. 44, 340-346.
- 8. Shimomura, O., Johnson, F. H. & Masugi, T. (1969) Science 164, 1299-1300.
- 9. Inouye, S., Noguchi, M., Sakaki, Y., Takagi, Y., Miyata, T., Iwanaga, S., Miyata, T. & Tsuji, F. I. (1985) Proc. Natl. Acad. Sci. USA 82, 3154-3158.
- 10. Prasher, D., McCann, R. O. & Cormier, M. J. (1985) Biochem. Biophys. Res. Commun. 126, 1259-1268.
- Tsuji, F. I., Haneda, Y., Lynch, R. V., III, & Sugiyama, N. (1971) 11. Comp. Biochem. Physiol. 40A, 163-179. Sie, E. H.-C., McElroy, W. D., Johnson, F. H. & Haneda, Y. (1961)
- 12. Arch. Biochem. Biophys. 93, 286–291.
- Johnson, F. H., Sugiyama, N., Shimomura, O., Saiga, Y. & Haneda, Y. 13. (1961) Proc. Natl. Acad. Sci. USA 47, 486-489. Thompson, E. M., Toya, Y., Nafpaktitis, B. G., Goto, T. & Tsuji, F. I.
- 14. (1988) J. Exp. Biol. 137, 39-51. Tsuji, F. I. (1978) Methods Enzymol. 57, 364-372.
- 15
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) Biochemistry 18, 5294-5299. 16.
- 17. Aviv, H. & Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408-1412. 18. Morishita, K., Kubota, N., Asano, S., Kaziro, Y. & Nagata, S. (1987) J.
- Biol. Chem. 262, 3844-3851. 19. Caruthers, M. H. (1987) in Synthesis and Applications of DNA and RNA,
- ed. Narang, S. A. (Academic, New York), pp. 47-94. Benton, W. D. & Davis, R. W. (1977) Science 196, 180-182. 20.
- Wahl, G. M., Stern, M. & Stark, G. R. (1979) Proc. Natl. Acad. Sci. USA 21. 76, 3683-3687
- 22. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- 23 Chen, E. Y. & Seeburg, P. H. (1985) DNA 4, 165-170.
- deWet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R. & Subramani, 24. S. (1987) Mol. Cell. Biol. 7, 725-737
- 25. Gluzman, Y. (1981) Cell 23, 175-182.
- 26. Graham, F. L. & van der Eb, A. J. (1973) Virology 52, 456-467.
- 27. Wigler, M., Pellicer, A., Silverstein, S. & Axel, R. (1978) Cell 14, 725-731.
- 28. Mitchell, G. W. & Hastings, J. W. (1971) Anal. Biochem. 39, 243-250.
- Hastings, J. W. & Weber, G. (1963) J. Opt. Soc. Am. 53, 1410-1415. 29.
- 30. Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y. & Matsubara, K. (1985) J. Biol. Chem. 260, 2605-2608.
- Kozak, M. (1978) Cell 15, 1109-1123. 31.
- von Heijne, G. (1983) Eur. J. Biochem. 133, 17-21. 32.
- 33. Marshall, R. D. (1972) Annu. Rev. Biochem. 41, 673-702.
- Tsuji, F. I., Lynch, R.V., III, & Stevens, C. L. (1974) Biochemistry 13, 34. 5204-5209
- 35. Gorman, C. M., Moffat, L. F. & Howard, B. H. (1982) Mol. Cell. Biol. 2, 1044-1051.
- Johnson, F. H. & Shimomura, O. (1978) Methods Enzymol. 57, 271-291. 36.
- Shimomura, O. & Johnson, F. H. (1973) *Tetrahedron Lett.*, 2963–2966. Prasher, D. C., McCann, R. O., Longiaru, M. & Cormier, M. J. (1987) 37.
- 38. Biochemistry 26, 1326-1332.
- 39. Kurose, K., Inouye, S., Sakaki, Y. & Tsuji, F. I. (1989) Proc. Natl. Acad. Sci. USA 86, 80–84.
- Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. (1978) in Atlas of 40. Protein Sequence and Structure (National Biomedical Research Foundation, Washington, DC), Vol. 5, pp. 345-352.
- Lynch, R. V., III, Tsuji, F. I. & Donald, D. H. (1972) Biochem. Biophys. Res. Commun. 46, 1544-1550. 41.
- 42 Engebrecht, J., Simon, M. & Silverman, M. (1985) Science 227, 1345-1347.
- 43 Ow, D. W., Wood, K. V., DeLuca, M., deWet, J. R., Helinski, D. R. & Howell, S. H. (1986) Science 234, 856-859.
- 44 Hauber, R., Miska, W., Schleinkofer, L. & Geiger, R. (1988) J. Clin. Chem. Clin. Biochem. 26, 147-148.