## Isolation and expression of a cDNA encoding *Renilla reniformis* luciferase

(bioluminescence/Renilla luciferase/green fluorescent protein/gene expression)

W. WALTER LORENZ\*, RICHARD O. MCCANN\*, MATHEW LONGIARU<sup>†</sup>, AND MILTON J. CORMIER<sup>\*‡</sup>

\*Department of Biochemistry, University of Georgia, Athens, GA 30602; and <sup>†</sup>Hoffman-La Roche, Nutley, NJ 07110

Communicated by W. D. McElroy, December 26, 1990 (received for review, October 12, 1990)

ABSTRACT Renilla reniformis is an anthozoan coelenterate capable of exhibiting bioluminescence. Bioluminescence in Renilla results from the oxidation of coelenterate luciferin (coelenterazine) by luciferase [Renilla-luciferin:oxygen 2-oxidoreductase (decarboxylating), EC 1.13.12.5]. In vivo, the excited state luciferin-luciferase complex undergoes the process of nonradiative energy transfer to an accessory protein, green fluorescent protein, which results in green bioluminescence. In vitro, Renilla luciferase emits blue light in the absence of any green fluorescent protein. A *Renilla* cDNA library has been constructed in  $\lambda gt11$ and screened by plaque hybridization with two oligonucleotide probes. We report here the isolation and characterization of a luciferase cDNA and its gene product. The recombinant luciferase expressed in Escherichia coli is identical to native luciferase as determined by SDS/PAGE, immunoblot analysis, and bioluminescence emission characteristics.

Renilla reniformis (class Anthozoa) is a bioluminescent soft coral found in shallow coastal waters of North America, which displays blue-green bioluminescence upon mechanical stimulation (1, 2). The components involved in *Renilla* bioluminescence have been described in detail (3). *Renilla* luciferase [*Renilla*-luciferin:oxygen 2-oxidoreductase (decarboxylating), EC 1.13.12.5] catalyzes the oxidative decarboxylation of coelenterazine in the presence of dissolved oxygen to yield oxyluciferin, CO<sub>2</sub>, and blue light ( $\lambda_{max} = 480$  nm) (4). This reaction has a bioluminescence quantum yield of  $\approx 7\%$ . The stoichiometry of this reaction and the detailed mechanism leading to excited-state formation have been described (4, 5).

The color of *in vitro*-catalyzed bioluminescence changes from blue to green upon addition of submicromolar amounts of an energy-transfer acceptor green fluorescent protein (GFP), which has been purified from *Renilla* and characterized (6). This green fluorescence ( $\lambda_{max} = 509$  nm) is identical to the *in vivo* emission in *Renilla*. The energy-transfer process is nonradiative; an increase in both the quantum yield (6) and calculated lifetimes has been determined for this process (7). Luciferase and GFP form a specific 1:1 rapid equilibrium complex in solution (8).

The elucidation of mechanisms involved in nonradiative energy transfer processes as well as determination of detailed structural information on both luciferase and GFP have been hindered by a lack of material. To overcome this, we have cloned, sequenced, and expressed in *Escherichia coli* a cDNA encoding *Renilla* luciferase.<sup>§</sup>

## MATERIALS AND METHODS

Amino Acid Sequence Determination of *Renilla* Luciferase. Native *Renilla* luciferase was isolated as described (4). Purified luciferase was digested with *Staphlococcal* protease V-8 (9). The resulting peptides were purified by HPLC and subjected to  $NH_2$ -terminal Edman sequencing as described (10). Based on these peptide sequences two 17-base oligonucleotide probes were synthesized with an Applied Biosystems DNA synthesizer at the Molecular Genetics Instrumentation Facility at the University of Georgia.

Construction of a cDNA Library in Agt11. Live R. reniformis were collected at the University of Georgia Marine Institute located at Sapelo Island. The animals were frozen immediately in liquid  $N_2$  and stored at  $-80^{\circ}$ C. Frozen tissue was ground to a fine powder in liquid  $N_2$  with mortar and pestle. Total RNA was isolated from the frozen powder by the guanidine thiocyanate method (11), and  $poly(A)^+$  RNA was isolated by oligo(dT)-cellulose chromatography (12). cDNA was synthesized by the method of Gubler and Hoffman (13). Phosphorylated EcoRI linkers (Collaborative Research) were ligated to the cDNAs, which were then digested with EcoRI. Separation of cDNA from free linkers after EcoRI digestion as well as size selection of cDNAs were accomplished by electrophoresis in low-melting-temperature agarose (NuSieve, FMC) (14). cDNAs were ligated into the *Eco*RI site of  $\lambda$ gt11 (15). The library was amplified in Y1088 cells (16) by a plate method (17).

Isolation and DNA Sequence Determination of a Luciferase cDNA. Oligonucleotide probes were 5' end-labeled with T4 polynucleotide kinase (Bethesda Research Laboratories) and  $[\gamma^{32}P]ATP$  (3000 Ci/mmol; 1 Ci = 37 GBq; ICN) to specific activities  $\geq 1 \times 10^8$  cpm/ $\mu$ g (18). A total of  $6 \times 10^5$  recombinant plaque-forming units were screened by plaque hybridization (19). Phage DNA was isolated as described (20). A luciferase cDNA, isolated from the clone  $\lambda$ RLuc-6, was subcloned into the M13 sequencing vectors mp18 and mp19, and sequencing templates were prepared (21). The DNA sequence of both strands was determined by the dideoxynucleotide chaintermination technique by using a Sequenase kit (United States Biochemical) and  $[\alpha^{-35}S]dATP$  (400 Ci/mmol; Amersham) (22). The M13 universal primer and a  $\lambda$ gt11 sequencing primer (Amersham) were used to prime the sequencing reactions.

**Expression of Recombinant Luciferase (r-luciferase).** Positive clones were converted to lysogens in *E. coli* ¥1089 cells (16). Lysogens were grown at permissive temperatures and induced with 1 mM isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG). Crude cell extracts were prepared and assayed for luciferase activity as described below. The plasmid pTZR-Luc-1 was constructed by ligation of a 2.2-kilobase-pair (kbp) *EcoRI/Sst* I  $\lambda$ RLuc-6 restriction fragment into the plasmid pTZ18R (Pharmacia), which contains the *lacZ'* gene. *E. coli* TG-1 cells (23) were transformed with pTZRLuc-1 (24). Single colonies were isolated and grown at 37°C in LB

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.

Abbreviations: GFP, green fluorescent protein; IPTG, isopropyl  $\beta$ -D-thiogalactopyranoside; r-luciferase, recombinant luciferase; ORF, open reading frame.

<sup>&</sup>lt;sup>‡</sup>To whom reprint requests should be addressed.

<sup>&</sup>lt;sup>§</sup>The sequence reported in this paper has been deposited in the GenBank data base (accession no. M63501).

medium containing ampicillin (100  $\mu$ g/ml) to an OD<sub>600</sub> = 0.6-0.8 unit and induced with 1 mM IPTG for 4 hr. The cells were centrifuged at 10,000 × g and frozen solid at -20°C. The pellets were thawed and resuspended 1:5 in 10 mM EDTA, pH 8, and lysozyme at 4 mg/ml (Sigma). After 20-min incubation at 25°C, the cells were placed on ice for 1 hr and then sonicated for 30 sec with a Branson cell disrupter. The cell lysate was clarified by centrifugation at 30,000 × g. The clarified lysate was used in subsequent bioluminescence assays and emission studies.

Assay for *Renilla* Luciferase Activity and Determination of Emission Spectra. Bioluminescence assays (4) were done with a Turner model TD-20e luminometer, and peak light intensities were determined. Bioluminescence intensity was converted to quanta per second by calibrating the instrument relative to a radioactive <sup>63</sup>Ni light standard that emits in the 460- to 480-nm region (25). Corrected emission spectra were collected on an on-line computerized fluorimeter (26). A 100- $\mu$ l sample of a clarified pTZRLuc-1 cell extract was added to 1 ml of luciferase assay buffer (4) or to 1 ml of "energy-transfer buffer" containing 1 × 10<sup>-6</sup> M GFP (8). An excess of coelenterazine (0.47 mM) dissolved in MeOH was added to maintain a strong emission signal.

Genomic Southern Blot Analyses. A 790-bp EcoRI/BamHIcDNA restriction fragment was labeled to specific activities  $\geq 1 \times 10^9$  cpm/µg with both [ $\alpha$ -<sup>32</sup>P]dATP and dCTP (4000 Ci/mmol, ICN) by the random hexamer-priming method (27). Genomic DNA was isolated from *Renilla* by a guanidine thiocyanate method developed for coelenterate DNA isolation (D. Prasher, personal communication). DNA samples were digested with the appropriate enzymes and resolved in a 0.8% agarose gel, followed by transfer to nitrocellulose filters (Schleicher & Schuell) (28). Aqueous hybridizations and washes were done at high stringencies as described for a homologous probe (17).

Electrophoretic Analysis of Protein. Protein samples were analyzed on 12.5% SDS/PAGE gels that were fixed and stained with Coomassie blue as described (29). Immunoblots were done as described (30). Proteins were transferred to nitrocellulose (Schleicher & Schuell) and incubated in a 1:50 dilution of rabbit anti-native luciferase antibody. Detection of the secondary antibody (horseradish peroxidase-conjugated goat anti-rabbit IgG) signal was determined according to the vendor's instructions (Bio-Rad).

**Computer-Facilitated DNA and Amino Acid Sequence Analyses.** The DNA sequence was compiled and manipulated using MicroGenie sequence software (Beckman).

## RESULTS

Synthesis of Luciferase Oligonucleotides. Seven luciferase peptides (V8-1–V8-7) were purified by HPLC, and their amino acid sequences were determined. Two of the peptides contained regions of relatively low codon degeneracy. Amino acid sequence from these regions was used to synthesize two mixed-sequence 17-base oligonucleotide hybridization probes with the following sequences; RLP-1 {GAR-AAYAAY-TTY-TTY-GT} and RLP-2 {AAR-AAR-TTY-CCN-AAY AT}, which are 32- and 64-fold degenerate, respectively.

Nucleotide and Deduced Amino Acid Sequence Analyses of *Renilla* Luciferase. Six clones were isolated from the *Renilla* cDNA library. One clone,  $\lambda$ RLuc-6, hybridized to both oligonucleotide probes. The cDNA insert could not be isolated after *Eco*RI digestion, as one of the linker sites was lost during cloning. A double digest with *Eco*RI and *Sst* I produced a 2.2-kbp fragment that contained a 1.2-kbp cDNA with 1 kbp of  $\lambda$ gt11 DNA at the 3' end. This fragment was subcloned into the M13 vectors mp18 and mp19. DNA sequencing provided the locations of six base restriction sites (Fig. 1), which were used to generate specific sequencing subclones. The entire 1.2-kbp luciferase cDNA was sequenced on both strands.

The cDNA, excluding the EcoRI linkers, is 1196 nucleotides long and encodes an open reading frame (ORF) of 314 amino acids (Fig. 2). Although an ATG in-frame codon is found at the 5' end of the cDNA, the intrinsic mRNA may contain additional 5' coding nucleotides. If the first ATG codon in the ORF is designated as the initiation codon, the predicted 311 amino acid sequence is essentially identical in size (34 kDa) and composition to native *Renilla* luciferase (4).

Comparison of the deduced amino acid sequence with the native peptides reveals that  $\lambda$ RLuc-6 encodes a luciferase cDNA (Fig. 2). One discrepancy lies at amino acid residue 222, which is leucine in the peptide sequence but tryptophan in the deduced sequence. Sequencing autoradiograms from this region of the clone have been examined carefully and found free of any irregularities. The protein sequence also contains a consensus N-linked glycosylation site (Asn-Xaa-Ser) beginning at residue 92.

Genomic Southern Analysis. A Renilla genomic Southern blot was probed with a 790-base-pair (bp) EcoRI/BamHI luciferase cDNA restriction fragment (Fig. 3). The BamHI digest, lane A, contains two hybridizing bands as does the Sma I digest, lane B. The Bgl II digest, lane C, contains four bands. If luciferase is encoded by a single gene containing no introns, a single band would be expected in the BamHI and Sma I digests, as these two sites are not spanned by the hybridization probe. Similarly, two bands would be expected in the Bgl II digest. That the BamHI and Sma I digests contain two hybridizing bands shows either that there is more than one luciferase gene or that the luciferase gene(s) has introns containing BamHI and Sma I sites. The four bands seen in the Bgl II could be explained by two very large introns containing Bgl II sites. When genomic DNA was digested with restriction enzymes having no sites within the cDNA sequence, there were always at least two or more hybridizing bands (data not shown). These results suggest that luciferase is encoded for by more than one gene, which may or may not contain introns.

Luciferase Expression in E. coli. The  $\lambda$ RLuc-6 lysogen is capable of low-level r-luciferase expression as determined by light emission from clarified, crude extracts (5 × 10<sup>10</sup> hv·sec<sup>-1</sup>·ml<sup>-1</sup>). When these cells are induced with 1 mM IPTG, light emission decreases by 2-fold; this happens because the cDNA is reversely oriented with respect to the  $\lambda$ gt11 *lacZ* promoter. Presumably, when IPTG is absent, the luciferase gene is transcribed from a promoter in the right end of  $\lambda$ gt11, as reported (31).

The 2.2-kbp *Eco*RI/*Sst* I fragment was subcloned into the plasmid pTZ18R, which uses the *lacZ* promoter. The ORF of





1	AGC	TTA Leu	AAG Lvs	* <u>ATG</u> Met	ACT Thr	TCG Ser	* AAA Lvs	GTT Val	TAT Tvr	* GAT Asp	CCA Pro	GAA Glu	C <b>AA</b> Gln	* AGG Arq	AAA Lys	CGG Arg	* ATG Met	ATA Ile	ACT Thr	* GGT Gly	60 20
-			-1-	*			*		-	*				*	-	-	*			*	
61 21	CCG Pro	C <b>A</b> G Gln	TGG Trp	TGG Tro	GCC Ala	AGA Arg	TGT Cvs	AAA Lvs	C <b>AA</b> Gln	ATG Met	AAT Asn	GTT Val	CTT Leu	GAT Asp	TCA Ser	TTT Phe	ATT Ile	AAT Asn	TAT Tyr	TAT Tyr	120 40
		•=		*		3	*	-1-		*				*			*			*	
121	GAT	TCA	GAA	AAA	CAT	GCA	GAA	AAT	GCT	GTT Val	ATT	TTT Phe	TTA Leu	CAT His	GGT Glv	AAC Asn	GCG Ala	GCC Ala	TCT Ser	TCT Ser	180 60
41	нэр	Ser	Gru	±		AIG	•			•				*	,		*			*	
181	TAT	TTA	TGG	ĊGA	CAT	GTT	GTG	CCA	CAT	ATT	GAG	CCA	GTA	GCG	CGG	TGT	ATT	ATA	CCA	GAT	240
61	Tyr	Leu	Trp	Arg	His	Val	vai	Pro	HIS	11e	GIU	PIO	vai	ALA	ALG	Cys	116	116	FIO	тэр	00
241	CTT	ATT	GGT	* ATG	GGC	ААА	* TCA	GGC	ААА	TCT	GGT	AAT	GGT	* ТСТ	TAT	AGG	TTA	CTT	GAT	CAT	300
81	Leu	Ile	Gly	Met	Gly	Lys	Ser	Gly	Lys	Ser	Gly	Asn	Gly	Ser	Tyr	Arg	Leu	Leu	Asp	His	100
301	тас	***	ТАТ	* СТТ	АСТ	GCA	* TGG	ттт	GAA	* CTT	CTT	ААТ	TTA	* CCA	AAG	AAG	* ATC	ATT	TTT	* GTC	360
101	Tyr	Lys	Tyr	Leu	Thr	Ala	Trp	Phe	Glu	Leu	Leu	Asn	Leu	Pro	Lys	Lys	Ile	Ile	Phe	Val	120
			~~~	*		~~~	*		~~~	*	~	<b>m</b> > m	200	* 	<b>CN</b> C	<b>C N T</b>	*	CAT	MG	*	420
361 121	GGC Gly	CAT His	GAT Asp	Trp	GGT Gly	Ala	тст Суз	Leu	Ala	Phe	His	Tyr	Ser	Tyr	Glu	His	Gln	Asp	Lys	Ile	140
				*			*			*				*			*			*	
421 141	AAA Lvs	GCA Ala	ATA Ile	GTT Val	CAC His	GCT Ala	GAA Glu	AGT Ser	GTA Val	GTA Val	GAT Asp	GTG Val	ATT Ile	GAA Glu	TCA Ser	TGG Trp	GAT Asp	GAA Glu	TGG Trp	CCT Pro	480 160
	-			*			*			*				*			*			*	
481	GAT	ATT	GAA	GAA	GAT	ATT	GCG	TTG	ATC	AAA	TCT	GAA	GAA	GGA	GAA	AAA	ATG	GTT	TTG	GAG Glu	540 180
101	Азр	TTe	GIU	GIU	Азр	TTe		Deu	116	- 193	Jei										l
541	ልልጥ	AAC	TTC	* TTC	GTG	GAA	ACC	ATG	TTG	CCA	TCA	ААА	ATC	ATG	AGA	AAG	TTA	GAA	CCA	GAA	600
181	Asn	Asn	Phe	Phe	Val	Glu	Thr	Met	Leu	Pro	Ser	Lys	Ile	Met	Arg	Lys	Leu	Glu	Pro	Glu	200
	~ • •		~~~	*	<b></b>	0.000	*		mmc	*	<b>CA</b> C		CCT	*	CTT	COT	*	600	202	* TTA	660
201	GAA Glu	Phe	Ala	Ala	TAT	Leu	Glu	Pro	Phe	Lys	GIu	Lys	Gly	Glu	Val	Arg	Arg	Pro	Thr	Leu	220
		+		*			*			*				*			*			*	
661 221	TCA Ser	TGG	CCT	CGT	GAA	ATC Ile	CCG Pro	TTA Leu	GTA Val	AAA Lys	GGT	GGT Gly	AAA Lys	CCT Pro	GAC Asp	GTT Val	GTA Val	C <b>AA</b> Gln	ATT Ile	GTT Val	720 240
	L	-		*			*			*				*			*			*	
721	AGG	AAT	TAT	AAT	GCT	TAT	CTA	CGT	GCA	AGT	GAT	GAT	TTA	CCA	AAA	ATG	; TTT	ATT	GAA	TCG	780
241	Ατα					m		A		C ~ ~	200		1 011	Dro		Mot	Dha		11-11		1200
		ASI	Tyr	Asn	Ala	Tyr	Leu	Arg	Ala	Ser	Asp	Asp	Leu	Pro	Lys	Met	. Phe	ITe	GIU		4
781	GAT	CCA	GGA	Asn * TTC	Ala TTT	Tyr TCC	Leu * AAT	Arg GCT	Ala ATT	Ser * GTT	Asp GAA	GGC	GCC	Pro * AAG	Lys	Met	Phe *	AAT		GAA	840
781 261	GAT Asp	CCA	GGA GIY	Asn * TTC Phe	Ala TTT Phe	Tyr TCC Ser	Leu * AAT Asn	Arg GCT Ala	Ala ATT Ile	Ser * GTT Val	Asp GAA Glu	GGC GIY	GCC GCC	AAG Lys	AAC Lys	TTI TTI	* CCT	AAT Asn	ACT	GAA Glu	840 280
781 261 841	GAT Asp	CCA	GGA GIY	Asn * TTC Phe	Ala TTT Phe	Tyr TCC Ser	Leu * AAT Asn *	Arg GCT Ala	Ala ATT Ile	Ser GTT Val	Asp GAA Glu	GGC GGC GIY GAA	GCC GCC Ala GAI	AAG	AAC	TTI TTI Phe	CCT CCT Pro	AAT Asn ATG	ACT Thr	GAA Glu *	840 280 900
781 261 841 281	GAT Asp TTT Phe	CCA Pro GTC Val	GGA GGA GIy AAA Lys	Asn * TTC Phe * GTA Val	Ala TTT Phe AAA Lys	Tyr TCC Ser GGT Gly	AAT Asn * CTT Leu	Arg GCT Ala CAT His	Ala ATT Ile TTT Phe	Ser * GTT Val * TCG Ser	Asp GAA Glu CAA Glr	GGC GIY GAA GLU	GCC Ala GAT Asp	* AAG Lys * GCA	AAG Lys Lys CCI Pro	TTI Phe GAI Asp	* CCT * CCT * Pro	AAT Asn ATG Met	ACT Thr GGA Gly	GAA Glu * AAA Lys	840 280 900 300
781 261 841 281	GAT Asp TTT Phe	CCA Pro GTC Val	GGA GIY AAA Lys	Asn * TTC Phe * GTA Val	Ala TTT Phe AAA Lys	Tyr TCC Ser GGT Gly	Leu * AAT Asn * CTT Leu	GCT Ala CAT His	Ala ATT Ile TTT Phe	Ser * GTT Val * TCG Ser *	Asp GAA Glu CAA Glr	GGC GIY GAA GIU	GCC Ala GAT Asp	AAG AAG Lys GCA Ala	AAG	GAT GAT	CCT CCT Pro GAA Glu	AAT Asn ATG Met	ACT Thr GGA Gly	GAA Glu * AAA Lys	840 280 900 300
781 261 841 281 901 301	GAT Asp TTT Phe TAT Tyr	CCA Pro GTC Val ATC Ile	GGA GGA GIY AAA Lys AAA Lys	Asn * TTC Phe * GTA Val * TCC Ser	Ala TTT Phe AAA Lys TTC Phe	Tyr TCC Ser GGT Gly GTT Val	AAT Asn * CTT Leu GAG	Arg GCT Ala CAT His CGA	Ala ATT Ile TTT Phe GTT	Ser * GTT Val * TCG Ser * CTC	Asp GAA Glu CAA Glr AAP	GGC GIY GAA Glu AA1 AA1	GAT GAT GAT GAT GAT	A CAP	A CCI CCI Pro	GAT GAT	CCT CCT Pro F GAA Glu * CTTT	AAT Asn ATG Met	ACT Thr GGA Gly	GAA Glu AAA Lys *	900 300 960 314
781 261 841 281 901 301	GAT Asp TTT Phe TAT Tyr	GTC GTC Val	GGA GIY AAA Lys AAA Lys	Asn * TTC Phe & GTA Val * TCC Ser	Ala TTT Phe AAA Lys TTC Phe	Tyr TCC Ser GGT Gly GTT Val	AAT Asn CTT Leu GAG GIU	Arg GCT Ala CAT His CGA	Ala ATT Ile TTT Phe GTT Val	Ser * GTT Val * TCG Ser * CTC	Asp GAA Glu CAA Glr AAP	GGC GIY GAA GAA Glu AAT AAT	GCC Ala GAT Asp GAA	AAG AAG Lys GCA Ala CAP	AAG AAG Lys ACCI Pro ATAA	GAT GAT ASP	CCT CCT Pro GAA Glu * CTTT	AAT Asn ATG Met	ACT Thr GGA Gly	GAA GIU * AAA Lys * AT	840 280 900 300 960 314
781 261 841 281 901 301 963	GAT Asp TTT Phe TAT Tyr	CCA Pro GTC Val ATC Ile	GGA GGY AAA Lys AAA Lys Lys Lys	Asn * TTC Phe * GTA Val * TCC Ser CCC	Ala TTT Phe AAA Lys TTC Phe	Tyr TCC Ser GJy GJy GTT Val	AAT Asn CTT Leu GAG GIU	Arg GCT Ala CAT His CGA Arg	Ala ATT Ile TTT Phe GTT Val	Ser * GTT Val * TCG Ser * CTC Leu	Asp GAA Glu CAA Glr AAP	GGC GIY GAA GIU AAT AAT	GAT GAT GAT GAT GAT	A CAP	A CC1 A CC1 A CC1 A Pro	TTT TTTAC	CTTT CTTGA	AAT Asn ATG Met GGTI ATAI	ACT Thr GGA Gly TTTT	GAA GIU * AAA Lys * AT *	840 280 900 300 960 314 1032
781 261 841 281 901 301 963 1033	GAT Asp TTT Phe TAT Tyr TT GG	GTC Val ATC Ile ACAT	GGA GIY AAA Lys AAA Lys Lys TTTTT	Asn * TTC Phe * GTA Val * TCG Ser CCC ATF	Ala TTT Phe AAA Lys TTC Phe CGGGI	Tyr TCC Ser GGT Gly GTT Val * TTA	AAT AAT ASN CTT Leu GAG GIU ATAA	Arg GCT Ala CAT His CGA Arg ATATA	Ala ATT Ile TTT Phe GTI Val	Ser * GTT Val * TCG Ser * CTC CTC	Asp GAA Glu CAM Glr AAP	GGC GIY GAA GAA GIU AAT AAT	GAT GAT GAT GAT GAT GAT GAT CTCTC	A CAP CAP CAP CAP CAP CAP CAP CAP	A CCI CCI CCI TAZ	GAT GAT GAT ATTAC	CTTT CTGA	AAT Asn ATG Met GGTI ATAT	ACT Thr GGA Gly TTTT	GAA GIU * AAA Lys * TAT * CA	840 280 900 300 960 314 1032 1102
781 261 841 281 901 301 963 1033 1103	GAT Asp TTT Phe TAT Tyr TT GG CT	CCA Pro GTC Val ATC Ile ACAT GAAG	GGA GGA GIy : AAA : Lys : AAA : Lys : AAA : Lys : AAA : Lys : AAA : Lys	Asn * TTC Phe * GTA Val * CCC Ser CCC ATA	Ala TTT Phe AAA Lys TTC Phe CGGGT	Tyr TCC Ser GgT Gly Val * TTTA TTTG	AAT AAT ASN CTTT Leu GIU ATAA ATTA	Arg GCT Ala CAT His CAT His CAT TATA ATTT	Ala ATT Ile TTT Phe GTT Val	Ser * GTT Val * TCG Ser * * CTCC Lev CGTCA	Asp GAA Glu ; CAA ; Glr ; AAA ; Lys	Asp GGC Gly GAA Glu AAT * C AA T AC	GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT	* AAG * AAG * CAA * CAA * CAA GIT * TTTTA	A TAP	TTAC	CCT Pro Pro GAA Glu CTTT * CTGA TTGG ATTT	AAT Asn ATG Met GGTT ATAT AATA GTTF	ACT Thr GGA Gly TTTT	AAA Glu * AAA Lys * AT * CA	840 280 900 300 960 314 1032 1102 1172

FIG. 2. Nucleotide sequence and translated amino acid sequence of the *Renilla* luciferase cDNA. Putative and known translation control elements, as well as oligonucleotide hybridization sites, are underlined. Positions of native luciferase peptide sequences are boxed and, except at one residue (+), are identical to the deduced amino acid sequence obtained from the luciferase cDNA. Some of the native luciferase peptide sequences overlap at glutamic acid residues.

the cDNA is not in frame with the *lacZ'* gene ORF of pTZ18R. Supernatants were prepared from IPTG-induced pTZRLuc-1 cells, as described, and the level of luciferase expression was measured by the standard luciferase assay. A high level of r-luciferase activity,  $2 \times 10^{15}$  hv·sec<sup>-1</sup>·ml<sup>-1</sup>, is detected in clarified crude extracts of pTZRLuc-1 cells. This level of activity is 7-fold greater than in uninduced pTZR-Luc-1 cells.

Luc-1 crude extracts (Fig. 4). Crude extracts of IPTG-induced pTZRLuc-1 cells were analyzed by immunoblotting (Fig. 5). The protein band that reacts with antiluciferase antibody, lane B, corresponds to the same band seen in the Coomassie-stained gel (Fig. 4). Native luciferase was used as a positive control, lane A. No signal is detectable in the crude extract of pTZ18R cells, lane C. A duplicate filter incubated with preimmune serum showed no detectable signal.

A prominent protein band ( $M_r = 34,000$ ) migrating to the position of native luciferase is seen after SDS/PAGE of pTZR-

**Bioluminescence Emission Spectra.** The r-luciferasecatalyzed bioluminescence emission spectrum (Fig. 6a) is



FIG. 3. Genomic Southern blot of *R*. reniformis DNA hybridized to a 790-bp EcoRI/BamHI luciferase cDNA fragment. Genomic DNA had been digested with the following restriction enzymes: BamHI (lane A); Sma I (lane B); and BglII (lane C). Samples were resolved in a 0.8% agarose gel and transferred to nitrocellulose. Each lane contains 20  $\mu$ g of digested genomic DNA. Molecular size markers are in kbp.

very similar to that seen with native luciferase (32). The r-luciferase emission spectrum has a  $\lambda_{max} = 480$  nm and a slight shoulder at 400 nm, which correspond to emission from the excited-state oxyluciferin monoanion and neutral species, respectively. Disproportionation between these species is sensitive to environmental factors (7); thus, this spectrum indicates the strong similarity of the active-site environment between r-luciferase and the native enzyme. Although an increase in quantum yield has yet to be determined, r-luciferase can clearly transfer energy in the presence of *Renilla* GFP (Fig. 6b). The emission spectrum dramatically shifted



FIG. 4. SDS/PAGE analysis of total protein from IPTG-induced *E. coli* cells transformed with either pTZRLuc-1 or pTZ18R. Tenmilliliter cultures were grown to an OD<sub>600</sub> = 0.8 and induced with 1 mM IPTG for 4 hr. One-milliliter of cell culture, OD<sub>600</sub> = 5.0, was pelleted and resuspended in 0.5 ml of SDS sample buffer. Samples were boiled for 5 min, and 20  $\mu$ l was loaded per lane: native luciferase (10  $\mu$ g) (lane A); pTZRLuc-1 cells (lane B); and pTZ18R cells (lane C). Molecular weight ( $M_r \times 10^{-3}$ ) standard positions are indicated. Arrow shows position of native luciferase (L'ase).



14

FIG. 5. Immunoblot analysis of total protein. Sample preparation and electrophoresis were the same as in Fig. 4. Native luciferase  $(2 \mu g)$  (lane A);  $10 \mu l$  of pTZRLuc-1 cell extract (lane B); and  $10 \mu l$  of pTZ18R cell extract (lane C). Molecular weight ( $M_r \times 10^{-3}$ ) standard positions are indicated.

from the broad emission band generated by r-luciferase to the narrow, structured emission band ( $\lambda_{max} = 509$ ) seen when GFP is present. The emission spectrum generated with r-luciferase and GFP is very similar to the same spectrum generated with native luciferase and GFP (33).



FIG. 6. Bioluminescence emission spectra generated with crude r-luciferase and r-luciferase plus GFP. Crude pTZRLuc-1 cell extracts were prepared, as described, and 100  $\mu$ l,  $\approx 1 \times 10^{-6}$  M r-luciferase, as determined by peak light emission, was used to generate each spectrum. (a) Emission spectrum of crude r-luciferase. (b) Spectrum that results when  $1 \times 10^{-6}$  M Renilla GFP is added to crude r-luciferase in energy-transfer buffer.

## DISCUSSION

This work describes the isolation of a 1.2-kbp *R. reniformis* luciferase cDNA capable of directing the expression of r-luciferase. The cDNA contains an ORF encoding a 314-amino acid sequence in which all of the native luciferase peptide sequences obtained from V8-protease digestion are found. Rescreening the cDNA library with a 790-bp luciferase cDNA fragment as a hybridization probe has failed to produce other clones that contain the 5' noncoding region; therefore, whether the luciferase cDNA is full-length is not known. This uncertainty can be resolved by sequencing genomic clones corresponding to the 5' end of the luciferase gene.

The genomic Southern hybridization data indicates that *Renilla* luciferase is probably encoded by more than one gene, which may or may not contain intervening sequences. Further characterization of luciferase genomic clones will be required before the genetic organization of the luciferase gene(s) can be defined.

A putative initiation codon located at triplet position 4 of the ORF may be the translation initiation site for Renilla luciferase; the 311-amino acid sequence is essentially identical to native luciferase with respect to its composition and predicted molecular weight. Irrespective of whether this cDNA is full length, the luciferase that it encodes is expressed in pTZRLuc-1 cells and is catalytically active. The expression data demonstrate that r-luciferase is the same size as native luciferase on SDS/PAGE gels and is reactive with polyclonal rabbit antibodies raised against native Renilla luciferase. Expression of r-luciferase from the plasmid pTZR-Luc-1 is "leaky" because activity can be detected from uninduced cell cultures. The luciferase cDNA ORF is not in frame with the short lacZ' ORF contained in this construct. Any translation product initiating at the  $\beta$ -galactosidase sequence of pTZRLuc-1 would be terminated at a stop codon immediately adjacent to the putative initiation codon in the luciferase cDNA. Thus, the r-luciferase seen in SDS/PAGE gels does not contain any  $\beta$ -galactosidase sequence. We propose that expression of r-luciferase by pTZRLuc-1 is due to a translation coupling mechanism (34).

r-luciferase displays two very important characteristics of native luciferase: the ability to catalyze coelenterazine oxidation with the concomitant emission of blue light ( $\lambda_{max} = 480$ nm) and the ability to transfer energy to Renilla GFP with the production of green light ( $\lambda_{max} = 509$  nm). The two emission bands at 400 nm and 480 nm in the r-luciferase spectrum verify the strong similarity between the native and recombinant proteins and suggest that the integrity of the luciferase active site has been maintained. Furthermore, that energy transfer occurs in the presence of GFP shows that the luciferase domain(s) required for the interaction between luciferase and GFP is present in r-luciferase. Once pure r-luciferase is available, energy transfer quantum yield measurements will offer a more quantitative determination of the efficiency of the nonradiative energy-transfer process. Finally, the data demonstrate that N-linked glycosylation is not required for luciferase activity because E. coli do not perform this modification (35). r-luciferase in E. coli crude extracts behaves like the native enzyme by every criterion examined thus far.

We thank Yu-Ching Pan and Jeffery Hulmes (Roche Diagnostic Systems) for the partial amino acid sequence data on native luciferase; John Wunderlich for the synthesis of oligonucleotide probes; Dennis O'Kane and John Wampler for help in obtaining bioluminescence emission spectral data; Leon Dure, Claiborne Glover, and Douglas Prasher for their critical comments; and Judy Gray for technical assistance. This is contribution number 659 from the University of Georgia Marine Institute, Sapelo Island, GA. This work was funded, in part, by a grant from ELA Technologies Inc.

- 1. Harvey, E. N. (1952) in *Bioluminescence*, ed. Harvey, E. N. (Academic, New York), pp. 148-194.
- Morin, J. G. (1974) in *Coelenterate Biology: Reviews and* Perspectives, eds. Muscatine, L. & Lenhoff, H. M. (Academic, New York), pp. 397-438.
- Cormier, M. J. (1978) in *Bioluminescence in Action*, ed. Herring, P. J. (Academic, New York), pp. 75-108.
- 4. Matthews, J. C., Hori, K. & Cormier, M. J. (1977) Biochemistry 16, 85-91.
- 5. Hart, R. C., Stempel, K. E., Boyer, P. D. & Cormier, M. J. (1978) Biochem. Biophys. Res. Commun. 81, 980-986.
- Ward, W. W. & Cormier, M. J. (1979) J. Biol. Chem. 254, 781-788.
- 7. Hart, R. C., Matthews, J. C., Hori, K. & Cormier, M. J. (1979) Biochemistry 18, 2204-2210.
- Ward, W. W. & Cormier, M. J. (1978) Photochem. Photobiol. 27, 389-396.
- Houmard, J. & Drapeau, G. R. (1972) Proc. Natl. Acad. Sci. USA 69, 3506-3509.
- Charbonneau, H., Walsh, K. A., McCann, R. O., Prendergast, F. G., Cormier, M. J. & Vanaman, T. C. (1985) *Biochemistry* 24, 6762–6771.
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* 18, 5294–5299.
- 12. Aviv, H. & Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408-1412.
- 13. Gubler, U. & Hoffman, B. J. (1983) Gene 25, 263-269.
- 14. Weislander, L. (1979) Anal. Biochem. 98, 305-309.
- 15. Young, R. A. & Davis, R. W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194-1198.
- Ausubel, F. M., Brert, R., Kingston, R. E., Morre, D. D., Seidmon, J. G., Smith, J. A. & Struhl, K. (1987) Current Protocols in Molecular Biology (Wiley, New York).
- 17. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
- Chaconas, G. & van de Sande, J. H. (1980) Methods Enzymol. 68, 75-88.
- Wood, W. I., Gitschier, J., Lasky, L. A. & Lawn, R. M. (1985) Proc. Natl. Acad. Sci. USA 82, 1585-1588.
- 20. Grossberger, D. (1987) Nucleic Acids Res. 15, 6737.
- 21. Yanisch-Perron, C., Vieria, J. & Messing, J. (1985) Gene 33, 103-119.
- 22. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- 23. Gay, N. J. & Walker, J. E. (1985) Biochem. J. 225, 707-712.
- Kushner, S. R. (1978) in Genetic Engineering, eds. Boyer, H. W. & Nicosia, S. (Elsevier Biomed., Amsterdam), pp. 17-23.
- 25. O'Kane, D. J. & Lee, J. (1990) Photochem. Photobiol. 52, 723-734.
- Mulkerrin, M. G. & Wampler, J. E. (1982) Anal. Chem. 54, 1778–1782.
- 27. Feinberg, A. P. & Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
- 28. Southern, E. M. (1975) J. Mol. Biol. 98, 503-517.
- 29. Laemmli, U. K. (1970) Nature (London) 277, 680-685.
- 30. Burnett, W. H. (1981) Anal. Biochem. 112, 195-203.
- 31. Chirala, S. S. (1986) Nucleic Acids Res. 14, 5935.
- Matthews, J. C., Hori, K. & Cormier, M. J. (1977) Biochemistry 16, 5217-5220.
- 33. Ward, W. W. & Cormier, M. J. (1978) Methods Enzymol. 57, 257-267.
- Stormo, G. D. (1986) in Maximizing Gene Expression, eds. Rezinikoff, W. & Gold, L. (Butterworths, Boston), pp. 195-224.
- 35. Kornfeld, R. & Kornfeld, S. (1985) Annu. Rev. Biochem. 54, 631-664.