OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1 Jan 98	Final report 15 April 1994-14April 1997

4. TITLE AND SUBTITLE
Molecular Biology and Biochemical Mechanisms of Marine Microorganisms

5. FUNDING NUMBERS
N00014-94-1-0575

6. AUTHOR(S)
J. Woodland Hastings

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Harvard University
Department of Molecular and Cellular Biology
16 Divinity Avenue, Cambridge, MA 02138

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217-5000

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

19990326 037

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Dinoflagellate genes of interest, including those associated with bioluminescence, were isolated from a *Gonyaulax polyedra* cDNA library by PCR amplification, screening by homologous probes or antibodies. Sequences were compared to those in gene and protein data bases to gain insight into the evolutionary status of *G. polyedra*. Genomic sequences were used to look for the existence of introns, the +1 site of the mRNAs, and to identify promoter sequences. Cloned sequences were expressed in *E. coli* as fusion proteins and activities determined by standard assays. The luciferase gene was discovered to possess three contiguous and homologous repeated sequences, each having luciferase activity as an individual peptide expressed in *E. coli*.

14. SUBJECT TERMS
Bioluminescence, dinoflagellates, peridinin-chlorophyll protein
Repeated domains

15. NUMBER OF PAGES
3

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	UL

FINAL
~~ANNUAL PROGRESS~~ REPORT

GRANT #: N00014-94-1-0575 R&T CODE: 4101-001

PRINCIPAL INVESTIGATOR: Dr. J. Woodland Hastings

INSTITUTION: Harvard Univ., Dept Molecular & Cellular Biology

GRANT TITLE: Molecular Biology and Biochemical Mechanisms of Marine Microorganisms

REPORTING PERIOD: 1 June 1996- 31 May 1997

AWARD PERIOD: 15 April 1994- 14 April 1997

OBJECTIVE: Structure, organization and expression of genes of *Gonyaulax polyedra*, a bioluminescent marine dinoflagellate.

APPROACH: Dinoflagellate genes of interest are isolated from a cDNA library by PCR amplification, screening by homologous probes or antibodies. Sequences are compared to those in gene and protein data bases to gain insight into the evolutionary status of *G. polyedra*. Genomic sequences are used to look for the existence of introns, the +1 site of the mRNAs, and to identify promoter sequences. Cloned sequences are expressed in *E. coli*,as glutathione-S-transferase fusion proteins and activities determined by standard assys.

ACCOMPLISHMENTS (last 12 months): The luciferase (*lcf*) gene of *Gonyaulax polyedra* has been cloned and sequenced at both cDNA and genomic levels. Several novel features have emerged. The gene contains three homologous and contiguous repeated sequences, which extend to the stop codon at the 3′ end. The rest of the gene at the 5′ end is homologous with a similarly located sequence in the gene coding for the luciferin binding protein (LBP). Each of the three repeat units, cloned and expressed separately in *E. coli* as glutathione-S-transferase fusion proteins, has luciferase activity. Sequence identity between repeats is greater in their central regions (>90%), as might be expected if they code for active sites. But synonymous substitutions are also significantly less frequent in this region, for which there is no known precedent and no evident explanation. The genomic *lcf* sequence, including its flanking regions, were determined. Transcription initiation site was identified using RNase protection assays and primer extension. Sequence analysis shows that, like the *lbp* gene, *lcf* does not contain introns. Genomic Southern blots, inverse PCR, and sequencing revealed that the *lcf* gene is organized as tandem repeats in the genome. The spacer region between the *lcf* genes, which probably contains the promoter elements necessary for transcription initiation, has no TATA box or other known promoter elements or consensus sequences. But a possible promoter sequence was identified by comparing the intergene spacer regions of *lcf* and the peridinin chlorophyll

protein gene, *pcp;* a novel 13 nt sequence, CGTGAACGCAGTG, was found to be present in both. This is the first putative promoter sequence identified in dinoflagellates.

A most remarkable new feature of dinoflagellate genomes to emerge from these studies is the redundancy at the DNA level: individual genes have repeated regions, and multiple copies of genes occur in tandem. This, together with the absence of a TATA box or any known promoter elements, indicates that the organization, regulation and expression of dinoflagellate genes are very different from those of other eukaryotes.

SIGNIFICANCE: The genes now cloned and sequenced appear to have novel features,the further elucidation of which may lead to new basic knowledge of gene structure and expression. In marine ecosystems dinoflagellates are important, causing the brilliant "phosphorescence" in the wake of a ship at night. Dinoflagellates are photosynthetic and lie at the base of food chains; some produce potent neurotoxins poisonous to man and may cause massive blooms (red tides). An understanding of gene structures in these systems will contribute to knowledge of ecology and dynamics of the marine environment.

WORK PLAN (next 12 months) We will express both full length and shorter but active peptides of dinoflagellate luciferase in a pure form and in quantitites adequate to obtain crystals for structural studies by a lab specializing in this latter aspect. At the same time, additional clones will be tailored to minimize the size of a fully active peptide. Concurrently, work will be undertaken to clone the luciferase gene from other bioluminescent species, starting with *Pyrocystis lunula*. In work just completed, we have shown that these luciferases are similar, but can be distinguished by being about 10 kDa larger, shown by both Western and Northern blots. These species also lack the luciferin binding protein (LBP), so the occurrence and possible role of the N terminal part of luciferase, which in *Gonyaulax* is homologous with the corresponding region of LBP, will be investigated.

PUBLICATIONS (last 12 months)

Hastings, J.W. (1996) Chemistries and colors of bioluminescent reactions- a review. **Gene** 173: 5-11.

Mittag, M. & Hastings, J.W. (1996) Exploring the signaling pathway of circadian bioluminescence. **Physiol. Plantarum.**96: 727-732

Techel, D., Chuang, E., Lee, D.-H., Comolli, J. & Hastings, J.W. (1996) Changes in the ultra structural localization of mRNA for a circadian regulated protein. **Mol. Marine Biol. & Biotech.** 5:241-48

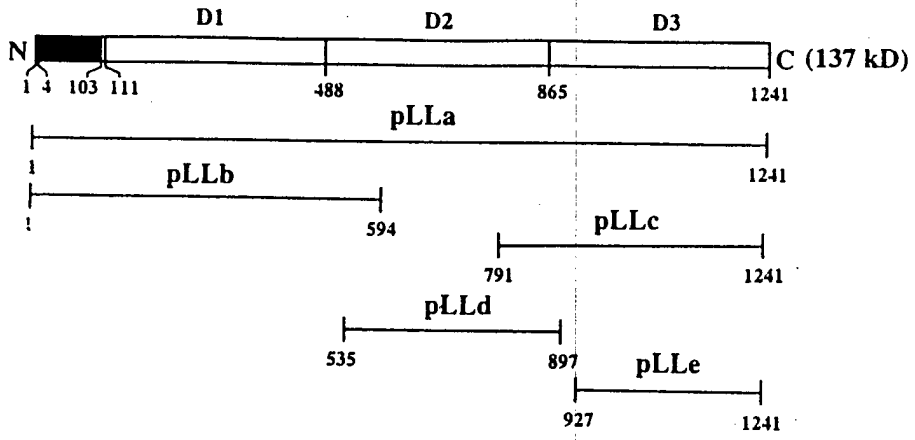
Li, Liming, Hong, R. & Hastings, J. W. (1997) Three functional luciferase domains in a single polypeptide chain. **PNAS** (In press).

Le, Q.H., Markovic, P., Hastings, J.W., Jovine, R.V.M. & Morse, D. (1997) Structure and organization of the Peridinin Chlorophyll Protein gene in *Gonyaulax polyedra*. **Mol. Gen. Genetics.** (In press)

Li, L. and Hastings, J.W. (1997) The structure and organization of the *Gonyaulax polyedra* luciferase gene. **Plant Mol. Biol.**(submitted)

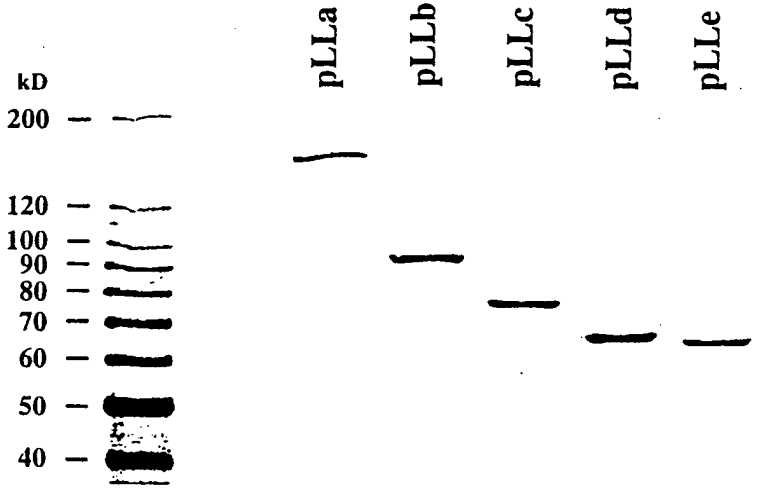
John W. Hastings, Harvard University

The *Gonyaulax* luciferase gene has three repeated units, D1, D2 & D3 cloned individually as pLLa. pLLb, pLLc, etc.



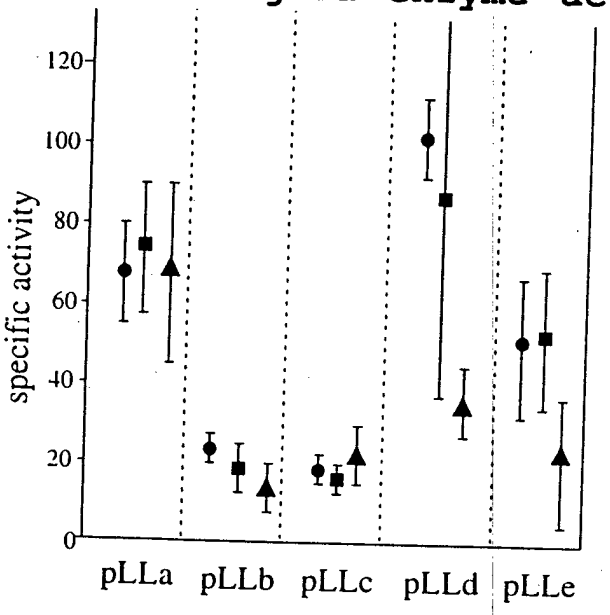
OBJECTIVES:

- To produce quantities of each peptide for crystals & structure determination



ACCOMPLISHMENTS:

- All three have good enzyme activity



SIGNIFICANCE

- Knowledge of gene structure and regulation will contribute to an understanding of the marine environment

- Organization and regulation of genes in the *Gonyaulax* nucleus, containing 50x as much DNA as the human cell and lacking histones

- An enhanced understanding of physiological and environmental aspects of "phosphorescence" of the ocean due to dinoflagellates.